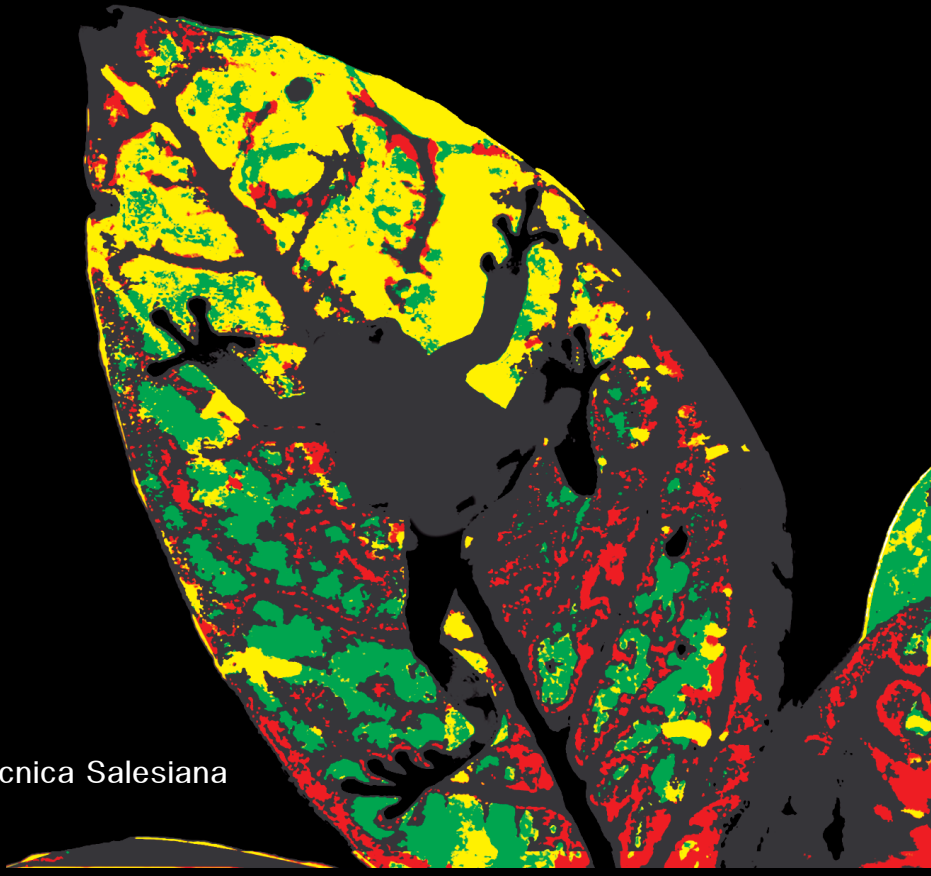


Leonardo Corral Dávalos

# ESTADÍSTICA y técnicas experimentales para la investigación BIOLÓGICA

Universidad Politécnica Salesiana



**Estadística y técnicas experimentales  
para la investigación biológica**

---



*Leonardo Corral Dávalos*

# **Estadística y técnicas experimentales para la investigación biológica**

---



**ABYA** | UNIVERSIDAD  
**YALA** | POLITÉCNICA  
SALESIANA

2019

**ESTADÍSTICA Y TÉCNICAS EXPERIMENTALES  
PARA LA INVESTIGACIÓN BIOLÓGICA**

© *Leonardo Corral Dávalos*

1ra edición: Universidad Politécnica Salesiana  
Av. Turuhuayco 3-69 y Calle Vieja  
Cuenca-Ecuador  
Casilla: 2074  
P.B.X. (+593 7) 2050000  
Fax: (+593 7) 4 088958  
e-mail: rpublicas@ups.edu.ec  
www.ups.edu.ec

CARRERAS DE BIOTECNOLOGÍA  
INGENIERÍA AMBIENTAL, INGENIERÍA AGROPECUARIA

Derechos de autor: 057790  
Depósito legal: 006530

ISBN: 978-9978-10-397-5

Edición, diseño,  
diagramación e impresión  
Editorial Universitaria Abya-Yala  
Quito-Ecuador

Tiraje: 300 ejemplares

Impreso en Quito-Ecuador, diciembre de 2019

Publicación arbitrada de la Universidad Politécnica Salesiana

<b>Dedicatoria</b> .....	11
<b>Prólogo</b> .....	13
<b>Reconocimientos</b> .....	15
<i>Capítulo 1</i>	
<b>Introducción</b> .....	17
1.1. ¿Qué es la Estadística?.....	18
1.2. Breve historia de la estadística .....	19
1.3. El razonamiento deductivo y el inferencial .....	22
1.4. El Método Científico.....	23
<i>Capítulo 2</i>	
<b>Datos y variables</b> .....	25
2.1. ¿Qué son datos y variables? .....	25
2.2. Tipos de variables.....	26
2.3. Representación de variables y datos.....	26
2.4. Símbolos y expresiones comunes en estadística.....	27
2.5. Reglas de la sumatoria .....	28
<i>Capítulo 3</i>	
<b>Poblaciones y muestras</b> .....	29
3.1. ¿Cuál es el concepto de población? .....	29
3.2. ¿Qué son parámetros?.....	29
3.3. Parámetros de Tendencia Central .....	30
3.4. Parámetros de dispersión .....	33
3.5. Propiedades de la media y varianza .....	36
3.6. Frecuencias, Histogramas y Curvas .....	38
3.7. La Distribución Normal de una Población .....	43
3.8. La distribución Z .....	44
3.9. ¿Qué es una muestra? .....	45
3.10. Estadísticas y sus características .....	46
3.11. El modelo lineal aditivo .....	50
<i>Capítulo 4</i>	
<b>Estadística y probabilidades</b> .....	53
4.1. Relación entre Estadística y Probabilidades .....	53

4.2. Reglas para el cálculo de probabilidades .....	54
4.3. La distribución binomial.....	58
4.4. La distribución Z y el cálculo de probabilidades .....	61

### Capítulo 5

<b>El teorema del límite central</b> .....	65
5.1. ¿Qué es el Teorema del Límite Central?.....	65
5.2. Los componentes del Teorema del Límite Central .....	67
5.3. El error estándar y el Teorema del Límite Central.....	70
5.4. La desviación estándar de las medias y las probabilidades .....	75

### Capítulo 6

<b>El planteamiento y las pruebas de hipótesis</b> .....	79
6.1. ¿Qué es una hipótesis? .....	79
6.2. ¿Cómo se plantean las hipótesis? .....	80
6.3. Procedimiento para las pruebas de hipótesis .....	81
6.4. Nivel de significación, valor crítico y prueba estadística .....	81
6.5. Hipótesis, decisiones y consecuencias .....	83
6.6. Error Tipo I y Error Tipo II .....	83
6.7. Pruebas de una cola y pruebas de dos colas .....	84

### Capítulo 7

<b>Pruebas de hipótesis para las medias</b> .....	87
7.1. Prueba Estadística Z.....	87
7.2. El Poder de la Prueba.....	90
7.3. Prueba Estadística $t$ de Student.....	98
7.4. La prueba estadística $t$ (P.E.t) y el control de calidad.....	101
7.5. Intervalo de Confianza para la media de la población .....	102

### Capítulo 8

<b>La distribución Ji-Cuadrado</b> .....	105
8.1. Origen de la distribución Ji – Cuadrado .....	105
8.2. Aplicaciones de la distribución Ji-Cuadrado.....	108
8.3. Intervalo de confianza para la varianza de la población .....	109
8.4. Ji – Cuadrado y la estadística no paramétrica.....	110
8.5. La prueba Ji – Cuadrado y la genética .....	115
8.6. Ji-Cuadrado y la prueba de contingencia .....	119
8.7. La prueba de Ji – Cuadrado en encuestas.....	123

### Capítulo 9

<b>Las distribuciones F y <math>t</math></b> .....	127
9.1. ¿Qué es la distribución F de Fisher? .....	127
9.2. Pruebas $t$ para la comparación de dos medias .....	131
9.3. Prueba de $t$ para observaciones NO pareadas.....	133

9.4. ¿Por qué la necesidad de dos pruebas: t y T?.....	138
9.5. La prueba de $t$ para observaciones pareadas .....	140
9.6. Intervalo de confianza para la diferencia de dos medias.....	144

*Capítulo 10*

<b>Regresión</b> .....	147
10.1. Las técnicas de regresión y sus aplicaciones.....	148
10.2. La relación de dependencia entre dos variables .....	148
10.3. Regresión lineal simple.....	149
10.4. El modelo de regresión lineal simple .....	150
10.5. Estimación de parámetros.....	152
10.6. Estimación del error experimental .....	157
10.7. Significado del coeficiente de regresión.....	160
10.8. Significado de la Intersección .....	160
10.9. Extrapolación de datos .....	161
10.10. Pruebas de hipótesis con relación a $\beta_1$ .....	161
10.11. Intervalo de confianza del coeficiente de regresión $\beta_1$ .....	167
10.12. Regresión polinómica.....	169

*Capítulo 11*

<b>Análisis de varianza de regresión</b> .....	183
11.1. Regresión: las fuentes de variación .....	183
11.2. El cuadro de análisis de varianza de la regresión .....	185
11.3. Relación entre $t$ y $F$ .....	190

*Capítulo 12*

<b>Análisis de correlación simple</b> .....	193
12.1. Relación entre regresión y correlación .....	194
12.2. Las unidades y límites del coeficiente de correlación .....	194
12.3. Evaluación de hipótesis en correlación.....	195
12.4. Correlación espuria.....	199
12.5. El coeficiente de determinación .....	200

*Capítulo 13*

<b>Introducción a diseño de experimentos</b> .....	203
13.1. El concepto de repetición.....	204
13.2. El concepto de error experimental.....	205
13.3. El concepto de Aleatorización.....	207
13.4. Procedimientos para la experimentación .....	208
13.5. Las pruebas t y el diseño de experimentos.....	210

*Capítulo 14*

<b>Diseño completamente al azar</b> .....	213
14.1. Ventajas del Diseño Completamente al Azar.....	213
14.2. Desventajas del Diseño Completamente al Azar.....	214



14.3. Componentes de la variación.....	214
14.4. Procedimiento para el análisis de datos.....	215
14.5. La aleatorización en el DCA.....	221
14.6. El Modelo Lineal Aditivo.....	222
14.7. La evaluación de hipótesis.....	226
14.8. El DCA con diferente número de repeticiones .....	232

### Capítulo 15

<b>Diseño de bloques completos al azar (bca)</b> .....	237
15.1. Ventajas del diseño de Bloques Completos al Azar.....	240
15.2. Desventajas del diseño de Bloques Completos al Azar.....	240
15.3. El modelo lineal aditivo para el diseño BCA.....	240
15.4. Componentes de la variación en el diseño BCA.....	241
15.5. Procedimiento para la evaluación de las hipótesis .....	242

### Capítulo 16

<b>Diseño cuadrado latino</b> .....	249
16.1. La aleatorización en el Cuadrado Latino.....	249
16.2. Ventajas y Desventajas del Diseño Cuadrado Latino.....	251
16.3. El modelo Lineal Aditivo para el Cuadrado Latino .....	251
16.4. El Cuadrado Latino en investigaciones pecuarias.....	256

### Capítulo 17

<b>Métodos de comparación de medias</b> .....	263
17.1. Todas las medias, igual número de repeticiones .....	264
17.2. Todas las medias, desigual número de repeticiones .....	281
17.3. Comparaciones con la media testigo .....	296
17.4. Comparaciones ortogonales entre grupos de medias.....	300
17.5. Cuando los tratamientos son cuantitativos.....	309

### Capítulo 18

<b>Diseños en látices</b> .....	317
18.1. Látices Balanceados.....	318
<i>Asignación de los tratamientos</i> .....	318
<i>Aleatorización de los tratamientos</i> .....	319
<i>El procedimiento para el análisis sigue los siguientes pasos:</i> .....	321
18.2. Látices parcialmente balanceados .....	327
<i>Pasos para el análisis</i> .....	331
<i>Comparación entre medias</i> .....	337

### Capítulo 19

<b>Experimentos factoriales</b> .....	339
La interacción.....	339
Designación de los factores y niveles.....	339
19.1. Experimentos Factoriales Completos .....	340

19.2. Experimentos factoriales con una repetición.....	362
19.3. Experimentos factoriales fraccionados .....	370
<i>Bloques incompletos</i> .....	370
<i>Los alias y la interacción generalizada</i> .....	378
<i>Fraccionamiento adicional de los tratamientos</i> .....	381
 <i>Capítulo 20</i>	
<b>Diseños anidados</b> .....	391
20.1. Diseño de Parcela Dividida .....	391
<i>Errores Estándar en el diseño de Parcela Dividida</i> .....	394
20.2. Diseño de Parcela Subdividida.....	400
<i>Errores Estándar en el diseño de Parcela Subdividida</i> .....	401
20.3. Diseño de Parcela Dividida en el Tiempo .....	412
20.4. Parcela Subdividida en el Espacio y Tiempo.....	415
20.5. Diseño de Bloque Dividido .....	420
 <i>Capítulo 21</i>	
<b>Experimentos combinados en el espacio y tiempo</b> .....	425
21.1. Homogeneidad de las varianzas.....	426
<i>La prueba de Bartlett</i> .....	427
<i>Modelo lineal aditivo</i> .....	427
21.2. Experimentos combinados, épocas de siembra.....	434
 <i>Capítulo 22</i>	
<b>Covarianza</b> .....	439
<i>Ventajas del análisis de covarianza</i> .....	440
<i>Requisitos para el uso de la covarianza</i> .....	440
<i>Pasos para el análisis de covarianza</i> .....	440
 <i>Capítulo 23</i>	
<b>Transformación de datos</b> .....	449
23.1. Independencia de los términos del error .....	449
23.2. Distribución normal de las poblaciones.....	450
23.3. Homogeneidad de las varianzas.....	461
23.4. Aditividad de los efectos de los tratamientos.....	471
 <i>Capítulo 24</i>	
<b>Estadística no paramétrica</b> .....	479
24.1. Análisis de Varianza de Kruskal – Wallis .....	480
<i>La asignación de rangos</i> .....	481
24.2. Análisis de Varianza de Friedman .....	486
24.3. La correlación de rangos de Spearman.....	493

<b>Referencias</b> .....	497
<b>Cuadro de fórmulas numeradas en el texto</b> .....	499
<b>Apéndices</b>	
Apéndice A. Tabla Z.....	509
Apéndice B. Tabla T .....	510
Apéndice C. Tabla JI-Cuadrado .....	511
Apéndice D1. Tabla F al 5% .....	512
Apéndice D2. Tabla F al 2.5% .....	513
Apéndice D3. Tabla F al 1% .....	514
Apéndice E1. Tabla Tukey al 5% .....	515
Apéndice E2. Tabla Tukey al 1% .....	516
Apéndice F1. Tabla Duncan al 5%.....	517
Apéndice F2. Tabla Duncan al 1%.....	518

Dedicado a mis nietas  
Andrea y María Emilia



Este libro resume las experiencias de 35 años de docencia e involucramiento en la investigación biológica, especialmente en el campo agronómico. Sin embargo, la estadística es una sola y las mismas técnicas que se presentan aquí pueden aplicarse en todas las áreas del conocimiento.

Los estudiantes, instructores y técnicos de las ciencias biológicas aplicadas, que requieren información para sus tesis y proyectos, tienen por lo general limitado acceso a especialistas a quienes consultar. Es cierto que hay varios textos de estadística y diseños experimentales, aunque relativamente pocos en español y con traducciones no siempre claras. Además, algunos son extremadamente complejos, con pocos ejemplos prácticos y presuponen un alto nivel de matemáticas para su comprensión. Estas restricciones fueron el principal motivo para escribir este libro, que se espera contribuya a un mejor y más eficiente uso de los métodos de análisis estadístico.

Este texto comienza desde los fundamentos de la estadística. Describe el significado de la variación y la importancia de su estudio. Establece tempranamente la relación entre población y las muestras que pueden extraerse de ella, un aspecto que se descuida y que es esencial para comprender esta ciencia. Enfatiza en el hecho de que la estadística sólo provee un nivel de probabilidad para las hipótesis planteadas, nunca certezas absolutas. Explica los procesos para la selección del diseño experimental que más se ajuste a las circunstancias, el análisis estadístico correspondiente al modelo y las pruebas posteriores para comparar las medias de los tratamientos. Incluye también algunos aspectos que no son comunes o suficientemente explicados, como la prueba de homogeneidad de las varianzas, la prueba de aditividad, los experimentos factoriales fraccionados, la transformación de datos que violan las presunciones del análisis de varianza y varias pruebas de la estadística no paramétrica.

Se ha tratado de mantener un adecuado equilibrio entre la teoría y las aplicaciones prácticas, con explicaciones fáciles de entender, si se dispone de conocimientos básicos de álgebra, en la mayoría de los casos. Varios ejercicios provienen de la experiencia propia del autor; otros, de su participación en decenas de tesis;

algunos, como resultado de consultorías para evaluación de proyectos y unos pocos, mediante simulaciones computarizadas. Por razones didácticas se presentan los análisis paso por paso, indicando las fórmulas necesarias para este objetivo. Los resultados se corroboran con análisis ejecutados con el programa *Statistix*<sup>1</sup>, pero puede usarse cualquier programa de preferencia del estudiante.

El material de este texto podría usarse en dos cursos universitarios. El primero, de estadística básica, del capítulo 1 al capítulo 10. El segundo, un curso de diseños y técnicas experimentales, del capítulo 11 al 24. En cursos de maestrías, el autor ha empleado todo el material, con un repaso rápido de los primeros capítulos y mayor énfasis en los capítulos posteriores.

Por último, se aspira a que investigadores y estudiantes, a través de estas páginas, puedan identificar mejores prácticas para la planificación de experimentos y análisis de datos.

---

1 Statistix 9. Analytical Software. Tallahassee, Florida. [www.statistix.com](http://www.statistix.com)

## Reconocimientos

---

Este libro se construyó sobre los conocimientos generados por científicos y autores que precedieron mi esfuerzo. Se reconoce esa enorme contribución al pie de las páginas y en la lista de referencias, que seguramente estarán incompletas.

El aporte derivado de mi interacción con cientos de estudiantes en clases, laboratorios y campos se manifiesta en todas las páginas del texto y especialmente en los ejercicios y comentarios. Sus inquietudes y preguntas aclararon el camino y aprendí de ellos.

La Escuela Agrícola Panamericana Zamorano me proveyó por algunos años de un ambiente favorable para el trabajo y la creación. De hecho, la semilla de este libro tuvo ese origen. Gracias a esa Institución y a su Rector, Jeffrey Lansdale, por el apoyo brindado para la publicación de este libro.

Rosita Espinoza, Catedrática e Investigadora de la Universidad Politécnica Salesiana merece un reconocimiento especial. Ella analizó el texto, lo envió a pares anónimos y lo puso a consideración de los coordinadores de las carreras de agroecología, biotecnología y ciencias agropecuarias de la UPS. Mi agradecimiento a Rosita y a los revisores del texto por los informes favorables.

Algunos colegas, profesores y exprofesores del Zamorano, analizaron el contenido de este texto. Entre ellos: Jorge Chang, Raúl Espinal, Jorge Román. Sus comentarios fueron de mucha ayuda. Merece mención especial el profesor Miguel Avedillo por la meticulosidad de su análisis. Miguel hizo una revisión en detalle de varias partes del texto. Sus observaciones contribuyeron a mejorarlo. Muchas gracias a todos ellos.

Varios profesionales de otras áreas también aportaron con valiosos comentarios. Mis reconocimientos a todos ellos. A Franklin López, economista, por su visión general del texto; a Paúl Corral, economista, por sus sugerencias sobre el tema de probabilidades; a Daniel Barrezueta y Luis Carrión, doctores en medicina, por las posibles aplicaciones en su campo.

Un reconocimiento muy especial merece el Contralmirante Luis Flores. Su campo es la ingeniería naval, pero su experiencia administrativa y sentido común sirvieron para lograr un mejor enfoque en función de los posibles usuarios de este libro.



La escritura de un libro con cientos de fórmulas, figuras y cuadros requiere de mucha acuciosidad y esfuerzo. Felizmente recibí un gran apoyo de varias personas. Gracias a Valeria Valenzuela por su clara visión y sutileza con las intrincadas fórmulas. Gracias a Isbela Álvarez y Patricia Obando por su precisión y diligencia en la transcripción de partes del texto.

Encontré que la escritura de un libro requiere de un gran soporte familiar. Gracias a todos, en especial a Eulalia, mi esposa, por el entendimiento y paciencia en todo este proceso.

Por último, pero no menos importante, quiero resaltar el extraordinario trabajo realizado por los funcionarios de Abya-Yala/UPS, Doctor Hernán Hermosa y Sra. Martha Vinueza. Su eficiencia hizo posible disponer de este libro en el tiempo requerido. Muchas gracias.

*“Un día llegará en que el pensamiento estadístico será tan necesario para ejercer la ciudadanía con eficiencia, como la capacidad de leer y escribir”.* Herbert G. Wells<sup>2</sup>.

La célebre frase de H.G. Wells demuestra la importancia que en la vida diaria tiene la estadística. Este conocimiento nos permite formar juicios críticos, con fundamentos lógicos, sobre aspectos tan diversos como el manejo de la economía, los avisos comerciales, el cambio climático, las tendencias en medicina, los diagnósticos médicos, las noticias, las encuestas comerciales y políticas, los postulados religiosos, etc.

La estadística como ciencia y técnica es una herramienta esencial de apoyo en todas las áreas del saber humano y un elemento imprescindible para la toma de decisiones. El avance de todas las ciencias ha sido en gran parte el resultado de la aplicación de métodos estadísticos para la evaluación de hipótesis. Este ha sido el caso de todas las ciencias experimentales puras y aplicadas, como biología, química, física, astronomía, agricultura, medicina, psicología, etc.

En el campo específico de las ciencias agrícolas, el conocimiento de la estadística es esencial para planificar experimentos; analizar el contenido de la literatura científica; presentar propuestas de investigación a organismos nacionales e internacionales; participar en proyectos de investigación para la generación, adaptación o validación de nuevas tecnologías; escribir artículos científicos; participar en programas de transferencia de tecnologías; realizar pruebas de evaluación de nuevos insumos agrícolas como fertilizantes, pesticidas, variedades, semillas, métodos de manejo, etc.

---

2 Escritor y filósofo inglés. [https://www.goodreads.com/author/quotes/880695.H\\_G\\_Wells](https://www.goodreads.com/author/quotes/880695.H_G_Wells)

## 1.1. ¿Qué es la Estadística?

Estadística es la parte de las matemáticas que estudia las características de las muestras y de las poblaciones de las que provienen. La estadística es una herramienta que nos señala los procedimientos para generar, organizar, analizar e interpretar datos con el propósito de solucionar los problemas cotidianos en todos los campos. Proporciona, además, un soporte de racionalidad a las decisiones. El lenguaje de la estadística es la matemática, en consecuencia, se requiere del conocimiento de los fundamentos matemáticos y de los modelos de representación para comprender los alcances de sus aplicaciones. Como generalmente las poblaciones son muy grandes o inaccesibles, para conocer sus características es necesario trabajar a partir de muestras.

La Estadística se divide en dos áreas: La Estadística Descriptiva y la Estadística Inferencial.

La **Estadística Descriptiva** tiene que ver con la recolección de datos, su organización en cuadros, la representación gráfica de los mismos y el cálculo de valores como frecuencias, mediana, media, desviación estándar y otros. Su papel esencial es proveer información básica inicial para la planificación y toma de decisiones.

La **Estadística Inferencial**, como su nombre lo indica, se emplea para realizar estimaciones o inferencias a partir de la información generada por la Estadística Descriptiva. Esto es, a partir de los resultados obtenidos con las muestras se llega a conclusiones relacionadas con las poblaciones. Para esto, la estadística inferencial se fundamenta en el cálculo de probabilidades, por lo que el nivel de certeza de sus conclusiones no es absoluto.

Podemos también decir que la Estadística Inferencial es la ciencia pura y aplicada que permite crear, desarrollar y aplicar técnicas para evaluar los resultados de experimentos. La Estadística Inferencial es una herramienta de la investigación, es decir un medio para arribar a conclusiones y no un fin en sí misma.

En todo caso es bueno tener en mente que la Estadística es sólo una herramienta. Cabe citar lo que indican Steel, Torrie y Dickey (1997), en el capítulo 1 con respecto a la interpretación de resultados: “Es el área de investigación la que debe proveer las respuestas. A veces los investigadores se olvidan de que son ellos los que deben pensar. La Estadística no puede pensar por ellos”.

## 1.2. Breve historia de la estadística<sup>3</sup>

La historia de la estadística es la historia del hombre sobre la tierra. Sin duda empezó con interrogantes sobre las distancias recorridas, los cambios de estaciones y de clima, la distribución de plantas y animales, la cantidad de alimento disponible, entre otras variables.

El nombre de “estadística” viene de Estado. Se conoce que las civilizaciones antiguas registraban la producción, el comercio, el número de habitantes y otros datos. El primer censo que registra la historia y que tuvo como finalidad reclutar soldados e imponer impuestos, se realizó en el imperio romano por orden del Emperador **César Augusto** (63 A.C.-14 D.C.)<sup>4</sup>. Estas “estadísticas” eran necesarias para el manejo del Estado y la planificación de la guerra y los territorios conquistados.

Pero la estadística como ciencia se inicia a mediados del siglo XVII, curiosamente por una discusión sobre las apuestas con dados y las experiencias del jugador **Chevalier de Meré**. Esto condujo a la formulación de la teoría de las probabilidades y la **distribución binomial** por parte de dos famosos matemáticos franceses: **Blaise Pascal** (1623-1662) y **Pierre Fermat** (1601-1665). Surgió entonces el concepto de la posibilidad de predecir eventos futuros, con cierto grado de certidumbre y sobre una base matemática. A partir de esto, muchos grandes pensadores han contribuido al desarrollo de la estadística. A continuación, se mencionan algunos personajes y sus aportes que han significado contribuciones importantes para esta ciencia.

El trabajo del matemático suizo **Jacob Bernoulli** (1655-1705) demostró que a medida que aumenta el tamaño de la muestra, la media de la muestra se aproxima más y más a la media de la población. En esto contribuyó también **Abraham de Moivre** (1667-1754), quien publicó la ecuación de la curva normal en 1733, aunque no la aplicó a datos experimentales. Estas observaciones condujeron a la formulación del **Teorema del Límite Central**, que es el fundamento en el que se basa la Estadística Inferencial.

3 Información sobre la historia e importancia de la estadística se encuentra en varios textos. Dos referencias recomendadas son: 1. Salsburg, D. 2001. *The Lady Tasting Tea. How statistics revolutionized science in the twentieth century*. W.H. Freeman and Company. New York. 2. Steel, R.G., J.H. Torrie, D.A. Dickey. 1997. *Principles and Procedures of Statistics: A biometrical Approach*. Third Edition. The McGraw-Hill Company, Inc. New York.

4 <http://www.imperivm.org/articulos/augusto.html>

Posteriormente, dos famosos matemáticos y astrónomos, el francés **Pierre Simon de Laplace** (1749-1827) y el alemán **Carl Friedrich Gauss** (1777-1855) desarrollaron en forma independiente la ecuación de la curva normal y aplicaron sus principios a su trabajo. Laplace demostró la importancia del uso de probabilidades para la interpretación de datos científicos, en su caso con relación a las órbitas de Júpiter y Saturno. Laplace también demostró matemáticamente el Teorema del Límite Central<sup>5</sup>. Por su parte Gauss, considerado uno de los más grandes matemáticos de todos los tiempos, desarrolló el método de “cuadrados mínimos”. Este método reduce a un mínimo la estimación del “error” en el análisis de datos y le sirvió a Gauss, en sus estudios de astronomía, para reencontrar el asteroide Ceres y calcular su órbita.

La contribución de **Simeón Denis Poisson** (1781-1840) tiene que ver con un caso especial de la distribución binomial, en la que la probabilidad de ocurrencia de un evento es sumamente baja. Una aplicación histórica de la llamada “**distribución de Poisson**” se dio en la Segunda Guerra Mundial, cuando el estadístico **R. D. Clarke** pudo determinar que las bombas V1 y V2 no iban dirigidas a blancos específicos en Londres, lo que hubiera cambiado la estrategia de defensa contra los ataques aéreos alemanes.

La investigación en estadística teórica y aplicada continuó en el siglo XIX con las contribuciones de los británicos **Francis Galton** (1822-1911) y **Karl Pearson** (1857-1935). Galton, primo de Charles Darwin, estudió la asociación de variables y desarrolló los conceptos de **regresión y correlación**, que fueron luego ampliados y perfeccionados por Pearson. El análisis del grado de asociación de variables cuantitativas independientes recibió el nombre de “correlación de Pearson”. El concepto de la distribución de **Ji-cuadrado** y las aplicaciones correspondientes se deben también a Pearson. Tanto Galton como Pearson, preconizadores de la eugenesia, hicieron múltiples observaciones biométricas, relacionadas especialmente con la herencia de características como la altura y la inteligencia humanas. Pearson creó la revista científica *Biometrika* en 1901 y que se publica hasta la fecha. Por sus grandes aportes a Pearson se lo considera como uno de los fundadores de la estadística.

Mientras Pearson trabajó con muestras grandes, uno de sus estudiantes, **William Sealy Gosset** (1876-1937) desarrolló empíricamente, extrayendo barajas al azar, la **distribución *t*** y la prueba del mismo nombre para muestras pequeñas.

---

5 <http://www.math.duke.edu/education/postcalc/probability/prob1.html>

El concepto de la **prueba  $t$  de Student**, seudónimo en alusión a su profesor Karl Pearson, se publicó en la revista *Biometrika* y se convirtió en un elemento de gran importancia para la experimentación e investigación. Un dato curioso, relacionado con lo anterior, es que el astrónomo alemán **Friedrich Helmert** había encontrado matemáticamente, ya en el año 1875, la ecuación de esta distribución. Sin embargo, posiblemente no vio sus aplicaciones prácticas.

Aunque ya se conocía que la variación de las observaciones experimentales podía separarse en varios componentes, fue el estadístico y genetista inglés **Ronald Fisher** (1890-1962) el creador de los métodos de análisis de varianza y de los diseños experimentales. Fisher trabajó como estadístico de la **Estación Experimental Rothamsted** e introdujo los conceptos de repetición, aleatorización, error experimental, bloqueo, covarianza, fusión de efectos, grados de libertad y otros en la investigación agropecuaria. Su libro “Métodos Estadísticos para Investigadores”, publicado en 1925, ha sido una guía clave para la experimentación en agricultura y otros campos. A Fisher se le debe el desarrollo de la distribución F (F por Fisher), de varias tablas estadísticas, de los experimentos factoriales y de los procedimientos conocidos con nombre propio como: Diseño Completamente al azar, Bloques Completos al Azar, Parcela Dividida, Cuadrado Latino, Bloques Incompletos, y otros.

Ronald Fisher hizo, además, aportes significativos al trabajo de Gosset, con quien desarrolló una cordial y fructífera relación. En cambio, con Pearson mantuvo una agria y larga disputa que se inició por una crítica de Pearson a un artículo suyo sobre la prueba de Ji-Cuadrado.

Otros aportes importantes más recientes ampliaron el campo de la aplicación de la estadística con nuevas teorías y aplicaciones. El trabajo del psicólogo **Abraham Wald** (1902-1950) se enfocó en los análisis secuenciales y en las funciones estadísticas para la toma de decisiones. Otro psicólogo, el inglés **Charles Spearman** (1863-1945) fue un pionero en el análisis factorial y desarrolló una prueba no paramétrica, el “coeficiente de correlación de escalas”, más conocido como la “Correlación de Spearman”, método que lo utilizó para medir la inteligencia humana.

A **Maurice Kendall** (1907-1983) se le reconocen contribuciones en el análisis de series en el tiempo, análisis multivariado y en una prueba de escalas que lleva su nombre. Las pruebas no paramétricas de “suma de escalas” y de “escalas signadas” fueron contribuciones de **Frank Wilcoxon** (1892-1965). Al ruso **Andrey Kolmogorov** (1903-1987) se le debe una nueva visión de los procesos estocásticos

y la prueba no paramétrica para la bondad de ajuste de distribuciones supuestamente normales (prueba de Kolmogorov-Smirnov). Otra prueba no paramétrica, que lleva el nombre de sus inventores, **William Kruskal** (1919-2005) y **Wilson Wallis** (1912-1998), substituye al análisis de varianza en una dirección, cuando las presunciones para el análisis de varianza normal (prueba paramétrica) no se cumplen. Igual cometido cumple la prueba de **Milton Friedman** (1912-2006), que reemplaza al análisis de varianza en dos direcciones.

En el campo de control estadístico de calidad en la producción industrial son importantes las contribuciones del japonés **Genichi Taguchi** (1924), quien propone novedosos análisis como parte de la ingeniería de la calidad. Además, incorpora nuevos conceptos de diseños de experimentos y análisis de medias, entre otros temas. Sin embargo, sus métodos han sido criticados por algunos investigadores occidentales.

En los últimos años los avances más importantes en la estadística han tenido que ver con nuevas teorías en el campo de las probabilidades, principalmente relacionadas con la física cuántica y el indeterminismo.

No podría concluirse este resumen sin mencionar que varios investigadores, mucho antes de que la estadística se consolidara como ciencia, hicieron contribuciones fundamentales al conocimiento científico. Realizaron observaciones, plantearon hipótesis, efectuaron experimentos, analizaron resultados y formularon sus teorías basándose en probabilidades y en la lógica que proporciona el sentido común. Ejemplos sobresalientes de esto fueron: **Leonardo de Vinci**, **Tycho Brahe**, **Galileo**, **Kepler**, **William Harvey**, **Isaac Newton**, **Johannes Mendel**, **Charles Darwin**.

### 1.3. El razonamiento deductivo y el inferencial

En el proceso de recoger información y llegar a conclusiones están involucrados dos formas de razonamiento lógico: 1) El *razonamiento deductivo*, que va de lo general a lo particular y 2) el *razonamiento inferencial*, que va de lo particular a lo general. Un ejemplo del primero es el siguiente: “Los genes para resistencia a piricularia (enfermedad del arroz causada por un hongo) se encuentran en especies silvestres del género *Oryza*. La planta *Oryza rufipogon* es una especie silvestre. Consecuentemente podría encontrarse genes para resistencia a piricularia en *Oryza rufipogon*”. Un ejemplo del segundo es el siguiente: “De diez plantas de maíz tomadas al azar, nueve tienen mazorcas sanas; consecuentemente se puede *inferir*

que alrededor del 90% de todas las plantas de maíz del sembrío probablemente tendrá mazorcas sanas”.

Los problemas en cualquier campo para su solución requieren del concurso del razonamiento inferencial. Sin embargo, las conclusiones derivadas del proceso inferencial deben ser siempre comprobadas mediante métodos deductivos. En otras palabras, las dos formas de razonamiento no son independientes, sino que se apoyan mutuamente.

#### 1.4. El Método Científico

El Método Científico es el procedimiento que se sigue para obtener respuestas en los trabajos de investigación. Aunque la metodología de las investigaciones varía de un trabajo a otro, el Método Científico tiene algunas características esenciales comunes. Estas son:

- Observación y revisión de hechos o fenómenos,
- formulación de hipótesis para explicar los hechos observados,
- sometimiento de la hipótesis a prueba por medio de métodos experimentales,
- evaluación de las hipótesis basándose en los resultados experimentales y
- aceptación, modificación o rechazo de las hipótesis formuladas.

En todo este proceso se emplea la herramienta llamada Estadística. Esta ciencia ayuda a planificar las pruebas experimentales, indica cómo y qué datos recoger, cómo analizarlos y qué nivel de confianza se puede tener en los resultados. Debido a que la Estadística se fundamenta en el cálculo de probabilidades, la evidencia que proporciona nunca es absoluta, por lo que siempre existe un margen de error.

Debe recordarse que la Estadística es una herramienta para la toma de decisiones. La evaluación de las hipótesis, su aceptación, modificación o rechazo se realiza a partir de la evidencia obtenida y bajo la luz del conocimiento del área de investigación.

Un ejemplo de la aplicación del método científico es el trabajo del botánico belga Juan Bautista van Helmont<sup>6</sup> en el siglo XVII. En términos de los pasos anteriores, el procedimiento fue el siguiente:

---

6 <https://sites.google.com/site/experimentsinphotosynthesis8b/jan-van-helmont>.



- El fenómeno observado era el crecimiento de las plantas.
- Van Helmont formuló la hipótesis de que las plantas crecían porque tomaban suelo por sus raíces y que, consecuentemente, el suelo era el principal constituyente de las plantas.
- Para someter la hipótesis a prueba, Van Helmont decidió cuantificar el crecimiento de un árbol a partir de una cantidad determinada de suelo. Sembró una pequeña planta de sauce que pesaba cinco libras en una maceta con doscientas libras de suelo. Durante cinco años únicamente añadió agua a la maceta. Al cabo de este período pesó cuidadosamente el suelo y el árbol. El suelo pesó 199 libras con 14 onzas y el árbol 169 libras con 3 onzas. El peso del suelo había variado únicamente en dos onzas, mientras que el peso de la planta había aumentado 164 libras y 3 onzas.
- Ante la evidencia, Van Helmont rechazó la hipótesis que había formulado de que el principal constituyente de las plantas era el suelo. Como sólo había añadido agua durante cinco años, concluyó lógicamente que el agua era el principal constituyente de las plantas. Hay que recordar que en ese tiempo no se conocía de la existencia del dióxido de carbono y menos aún de la fotosíntesis.

El procedimiento que siguió Van Helmont fue impecable y rigurosamente científico, a pesar de que la conclusión a la que arribó fue errónea. A la luz de los conocimientos actuales se le podría criticar que no tuvo repeticiones en diferentes ambientes. Una repetición del experimento en condiciones de oscuridad seguramente le hubiera llevado a otra conclusión. Pero no olvidemos que se trata de un experimento pionero, realizado en el siglo XVII (el experimento concluyó en 1648). La introducción de repeticiones en la experimentación fue una innovación de Ronald Fisher a principios del siglo XX.

Es importante notar el proceso del pensamiento inductivo en este trabajo. Van Helmont pudo generalizar que todos los sauces y probablemente todas las plantas responderían de la misma manera.

*“La estadística puede definirse como un conjunto de métodos para tomar decisiones frente a la incertidumbre”. W.A. Wallis<sup>7</sup>*

## **2.1 . ¿Qué son datos y variables?**

En Estadística se trabaja con datos. Los datos son observaciones que provienen de la población o poblaciones de interés para el investigador. Hay dos fuentes principales de datos: 1) los que se generan en experimentos controlados y 2) los que se obtienen directamente del entorno.

Un ejemplo del primer caso podría ser el de una investigación para evaluar el rendimiento de 10 nuevos híbridos de maíz. Para esto, el investigador prepara el terreno; siembra los híbridos generalmente en varias parcelas con repeticiones; los cuida durante todo el ciclo de vida y obtiene el rendimiento. Un ejemplo del segundo caso se tendría en una investigación sobre el rendimiento en sembríos ya establecidos por agricultores. El investigador únicamente toma los datos sin que haya intervenido para nada en el cultivo. Las dos fuentes de los ejemplos son válidas, obviamente, aunque los objetivos de las respectivas investigaciones sean distintos.

Una característica generalmente común de los datos es la variación existente entre ellos. Si se toma por ejemplo el rendimiento individual de 20 plantas de maíz es muy probable que los 20 datos sean diferentes, es decir, existe variación entre ellos.

Cuando una característica presenta variación recibe el nombre de *variable*. Ejemplos de variables comunes en un experimento con maíz son: porcentaje de germinación, días a floración, altura de planta, incidencia de enfermedades, porcentaje de acame, cobertura de la mazorca, peso de la mazorca, peso del grano, color del grano, entre otros.

---

7 Economista y estadístico estadounidense. <http://www.quotegarden.com/statistics.html>

## 2.2. Tipos de variables

Las variables pueden ser cualitativas y cuantitativas. Una *variable cualitativa* describe una característica con límites bien definidos; por ejemplo, la diferente coloración que presentan semillas de fréjol. Las semillas pueden ser blancas, rojas, negras, amarillas, bayas, cafés, o de otro color y combinaciones de colores. Las evaluaciones de experimentos que incluyen variables cualitativas se basan por lo general en la *Estadística No Paramétrica*.

Por razón de avances en la calidad de los instrumentos de laboratorio, reactivos y métodos de análisis, algunas variables tradicionalmente cualitativas pueden convertirse en variables cuantitativas. Por ejemplo, en ciertos trabajos, podría no bastar la indicación del color de las semillas. El investigador podría obtener resultados cuantitativos del contenido en las semillas de uno o más pigmentos y con estos datos apoyarse en forma más eficiente en sus trabajos de mejoramiento genético.

Una *variable cuantitativa* es aquella en la que las observaciones pueden expresarse numéricamente. Las variables cuantitativas pueden ser *continuas* y *discretas*. Los datos de las variables continuas no presentan límites definidos y pueden expresarse con fracciones o decimales. Ejemplos de este tipo de variables son: peso del grano, tamaño de la mazorca, contenido de grasa en la leche. Los datos de las variables discretas provienen de conteos y se expresan con números enteros por tener límites definidos. Como ejemplos se pueden citar los siguientes: número de semillas por vaina, número de larvas por fruto, número de bananos (dedos) por racimo.

Tanto las variables continuas como las discretas por lo general son evaluadas con métodos de la *Estadística Paramétrica*. Sin embargo, en varios casos los datos correspondientes a las variables discretas requieren de algún tipo de transformación, como se verá más adelante.

## 2.3. Representación de variables y datos

Las variables generalmente se representan con las últimas letras del alfabeto, como W, X, Y, Z. Las constantes se representan con las letras K o C. En este libro se empleará la letra Y para las *variables dependientes*, la letra X para las *variables independientes* y C para las constantes.

Cuando se tiene un conjunto o colección de datos numéricos se los puede designar así:  $Y_1, Y_2, \dots, Y_i, \dots, Y_n$ . Donde  $Y_i$  se refiere al dato "i" de la co-

lección;  $Y_n$  se refiere al último dato de la colección. Obviamente hay “n” datos en la colección.

## 2.4. Símbolos y expresiones comunes en estadística

En estadística la suma de los datos se expresa de la siguiente manera:  $\sum_{i=1}^n Y_i$

Donde  $\sum_{i=1}^n Y_i = Y_1 + Y_2 + \dots + Y_i + \dots + Y_n$

La letra griega sigma mayúscula  $\Sigma$  indica sumatoria y es uno de los símbolos más usados en Estadística. Los índices arriba y abajo de sigma indican los límites de la sumatoria. Por ejemplo:  $\sum_{i=1}^n Y_i$  significa la sumatoria de  $Y_1$  hasta  $Y_n$ , inclusive. Cuando se sobreentienden los límites, basta escribir:  $\sum Y_i$

Para los análisis estadísticos es necesario obtener la sumatoria de los cuadrados de todas las observaciones. Esto es:

$\sum Y_i^2 = Y_1^2 + Y_2^2 + Y_3^2 + \dots + Y_n^2$ . A esta expresión se la conoce como la “sumatoria de cuadrados sin corregir”.

La expresión anterior, “sumatoria de los cuadrados sin corregir”, no debe confundirse con la “sumatoria de las observaciones al cuadrado”:  $(\sum Y_i)^2 = (Y_1 + Y_2 + Y_3 + \dots + Y_n)^2$

Es decir que:  $\sum Y_i^2 \neq (\sum Y_i)^2$

Cuando se divide  $(\sum Y_i)^2$  para el número de observaciones ( $N$ , si se trata de una población o  $n$ , si se trata de una muestra, como se verá más adelante), se obtiene lo que se denomina el “factor de corrección”:  $FC = \frac{(\sum Y_i)^2}{n}$

Si restamos de la “sumatoria de cuadrados sin corregir” el “factor de corrección”, obtenemos la “sumatoria de cuadrados corregida”, o simplemente  $SC$ . Esto es:

$$SC = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$$

**Ejercicio 2.4.1.** Si  $Y_1 = 5$ ,  $Y_2 = 4$ ,  $Y_3 = 6$ ,  $Y_4 = 8$ , resuelva lo siguiente:

- a. la sumatoria de  $Y$ :

$$\sum Y_i = Y_1 + Y_2 + Y_3 + Y_4 = 5 + 4 + 6 + 8 = 23$$

- b. la sumatoria de los cuadrados de Y (se llama también: *Sumatoria de Cuadrados sin Corregir*)

$$\sum Y_i^2 = Y_1^2 + Y_2^2 + Y_3^2 + Y_4^2 = 5^2 + 4^2 + 6^2 + 8^2 = 141$$

- c. la sumatoria de Y al cuadrado:

$$(\sum Y_i)^2 = (5 + 4 + 6 + 8)^2 = 23^2 = 529$$

- d. La sumatoria de Y al cuadrado dividida para el número de observaciones (se llama también: *Factor de Corrección*)

$$\frac{(\sum Y_i)^2}{n} = \frac{529}{4} = 132.25$$

- e. La *Sumatoria de Cuadrados Corregida* (SCC o simplemente SC):

$$SC = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} = 141 - 132.25 = 8.75$$

## 2.5. Reglas de la sumatoria

Las reglas de la sumatoria son las siguientes:

- La sumatoria de una expresión que tiene dos o más variables es igual a la suma de las sumatorias de las variables individuales. Esto es:

$$\sum (Y_i + W_i + Z_i) = \sum Y_i + \sum W_i + \sum Z_i$$

- La sumatoria de una variable multiplicada por una constante  $K$ , es igual a la constante multiplicada por la sumatoria de la variable. Esto es:  $\sum KY_i = K \sum Y_i$
- La sumatoria de una constante  $K$  es igual a la constante multiplicada por el número de términos  $n$  en la sumatoria. Esto es:  $\sum_{i=1}^n K = nK$ . Es decir que si  $n=8$  y  $K=12$ , luego

$$\sum_{i=1}^{n=8} 12 = 8 * 12 = 96$$

*“Debemos admitir que cualquier inferencia de lo particular a lo general tiene algún grado de incertidumbre... pero el nivel de duda debe expresarse rigurosamente”.*  
Sir Ronald Fisher<sup>8</sup>

### 3.1. ¿Cuál es el concepto de población?

En estadística, una *población* o *universo* es el conjunto de todos los valores de una variable, cualitativa o cuantitativa, continua o discreta. Las poblaciones pueden ser finitas o infinitas, concretas o abstractas. Por ejemplo, los pesos de 50 estudiantes de un curso forman una población finita y concreta. Las alturas de todas las plantas de maíz cultivadas en el año 2014 forman una población prácticamente infinita, aunque concreta. El rendimiento de las plantas de maíz cultivadas en el Ecuador en el año 2020 es una población infinita y abstracta. En todo caso, es el investigador quien define la población de referencia en su trabajo.

### 3.2. ¿Qué son parámetros?

Si se conocieran todos los valores de una o más variables de una población sería posible describirla exactamente. Para describir una población se emplean sus *parámetros*, que son *características únicas y constantes*. En general, si pensamos en un gráfico que represente todos los valores de una población, nos interesa formarnos un criterio de la posible ubicación de la población en una escala determinada, su forma, inclinación y amplitud.

Hay dos tipos de parámetros: de *tendencia central* y de *dispersión*. Sin embargo, como rara vez se conocen los parámetros de una población, se pueden inferir sus valores mediante la obtención de muestras y el empleo de técnicas estadísticas. En todo caso es importante recordar que los parámetros están allí, pero que por lo general no los conocemos. La palabra “parámetro” aunque tiene algunas definiciones, etimológicamente quiere decir “medida desconocida”.

---

8 Estadístico y biólogo inglés. [https://www.brainyquote.com/authors/ronald\\_fisher](https://www.brainyquote.com/authors/ronald_fisher)

### 3.3. Parámetros de Tendencia Central

Se llaman parámetros de *tendencia central* a los que ubican a la población con referencia a puntos específicos del conjunto de datos (distribución de los datos). Los más usados son: el *modo*, la *mediana* y la *media aritmética*. Otros, como la *media geométrica*, la *media armónica*, los *cuartiles*, *quintiles* y *percentiles* son menos usados, aunque tienen aplicaciones específicas importantes en algunos campos, especialmente en las ciencias sociales.

- *Modo (Mo)*. Es la observación más frecuente de la población. Puede haber más de un modo.
- *Mediana (Me)*. Es el valor central en una serie ordenada de datos, de tal forma que el 50% de las observaciones queden a la izquierda de este valor y 50% a la derecha.
- *Media Aritmética* o simplemente *media* o *promedio* (Se representa con la letra griega Mu:  $(\mu)$ ). Es la suma de todas las observaciones dividida para el número de observaciones. La fórmula de la media de la población es:

$$\mu = \frac{\sum Y_i}{N} \quad (1)$$

- *Media Geométrica (G)*. Es la raíz “N” del producto de “N” datos. La media geométrica se emplea cuando se detecta en los datos la presencia de efectos multiplicativos. Ejemplos de esto son medidas relacionadas con el crecimiento de las plantas en el tiempo, el crecimiento poblacional, el interés compuesto, entre otros. La fórmula es:

$$G = \sqrt[N]{Y_1 * Y_2 * Y_3 \dots Y_N} \quad (2)$$

- *Media Armónica (H)*. La media armónica se emplea cuando los datos representan tasas o proporciones. La fórmula es:

$$H = \frac{N}{\sum_{i=1}^N \left(\frac{1}{Y_i}\right)} \quad (3)$$

- La media armónica es equivalente a la media aritmética ponderada. En la fórmula de la media ponderada se incluye la frecuencia ( $f$ ) o número de veces que se repiten las observaciones. La fórmula es:

$$\mu_P = \frac{\sum f_i Y_i}{\sum f_i} \quad (4)$$

- *Cuartiles, quintiles, percentiles.* Son puntos, en una serie de datos numéricos, que corresponden a observaciones que dividen la distribución en cuatro, cinco y cien segmentos, respectivamente. El segundo “cuartil” es equivalente a la mediana.

**Ejercicio 3.3.1.** Resolver lo siguiente:

- a. Encontrar la media o promedio de los pesos, en gramos, de una población de ocho mazorcas de maíz: 350, 285, 414, 342, 316, 287, 312, 294.

$$\text{Respuesta: } \mu = \frac{\sum Y_i}{N} = \frac{350+285+414+342+316+287+312+294}{8} = 325 \text{ gramos}$$

- b. Los requerimientos de agua de un cultivo en las etapas de elongación del tallo y de floración son de 20 metros cúbicos y 100 metros cúbicos por día por hectárea, respectivamente. ¿Cuál es una estimación del promedio diario de requerimiento de agua entre las dos etapas?

Respuesta: Como el requerimiento diario de agua se incrementa con el desarrollo del cultivo, la mejor estimación se obtiene con la media geométrica. Esto es:

$$G = \sqrt[N]{Y_1 * Y_2} = \sqrt[2]{Y_1 * Y_2} = \sqrt[2]{20 * 100} = 44.7 \text{ m}^3 \text{ de agua por hectárea.}$$

Este valor es una aproximación y puede variar dependiendo del cultivo, la densidad de siembra, textura y estructura del suelo, contenido de materia orgánica y otros factores.

Note que la media aritmética es 60 m<sup>3</sup>. Sin embargo, esto no hubiera sido una buena estimación, porque el requerimiento de agua con el desarrollo del cultivo no es lineal, sino más bien exponencial, con una fase lenta de incremento al principio.

- a. Se tiene dos reservorios vacíos con una capacidad de 5,000 metros cúbicos de agua cada uno. Para llenar los reservorios se dispone de dos bombas, A y B, con una capacidad de bombeo de 1,000 y de 500 metros cúbicos por hora, respectivamente. La bomba A llenará el reservorio uno en 5 horas, mientras que la bomba B llenará el reservorio dos en 10 horas. ¿Cuál es el promedio de bombeo de agua de las dos bombas para llenar los reservorios?

Respuesta: Para situaciones como la descrita, por tratarse de proporciones o tasas, lo correcto es encontrar la media armónica:

$$H = \frac{N}{\sum_{i=1}^N \left(\frac{1}{Y_i}\right)} = \frac{2}{\frac{1}{1000} + \frac{1}{500}} = \frac{2,000}{3} = 666.67 \text{ m}^3 / \text{hora}$$



(Nota: La media aritmética,  $750 \text{ m}^3/\text{hora}$ , no es una estimación correcta, porque no toma en cuenta el diferente tiempo de funcionamiento de cada bomba.)

También se puede emplear la fórmula de la *media aritmética ponderada* para resolver este problema. Las frecuencias, en este caso, son las horas de funcionamiento de cada bomba para llenar los reservorios. El resultado es, obviamente, el mismo:

$$\mu_p = \frac{\sum f_i Y_i}{\sum f_i} = \frac{5 \cdot 1000 + 10 \cdot 500}{15} = 666.67 \text{ m}^3/\text{hora}$$

- b. Si la bomba A llena un reservorio en 5 horas y la bomba B llena el mismo reservorio en 10 horas, ¿en qué tiempo llenarán el reservorio las dos bombas trabajando juntas, es decir en equipo?

Respuesta: para encontrar el tiempo de trabajo en equipo, en este caso para llenar un reservorio ( $N=1$ ), la solución se la encuentra con la misma fórmula de la media armónica:

$$H = \frac{N}{\sum \left(\frac{1}{Y_i}\right)} = \frac{1}{\sum \left(\frac{1}{Y_i}\right)} = \frac{1}{\frac{1}{5} + \frac{1}{10}} = \frac{1}{\frac{3}{10}} = 3.33 \text{ horas}$$

**Ejercicio 3.3.2.** En el Cuadro 1 se presenta las ventas de cacao de un productor durante tres períodos:

Cuadro 1. Ventas de cacao durante tres períodos

	I	II	III	Totales
Quintales de cacao	150	267	240	657 q
Precio de venta (US\$)	67	95	102	
Ingresos (US\$)	10050	25365	24480	59895 dólares

- a. ¿Cuál es el precio de venta promedio por quintal en los tres períodos? Esto es la media aritmética de los valores 67, 95 y 102, sin tomar en cuenta el volumen de ventas:

$$\mu = \frac{67+95+102}{3} = \$88.0$$

- b. ¿Cuál es el precio de venta promedio que el agricultor recibió por los 657 quintales?

Para esto, debe calcularse la media ponderada. Las frecuencias corresponden al número de quintales que se vendieron en cada período. Esto es:

$$\mu_p = \frac{\sum f_i Y_i}{\sum f_i} = \frac{150 \cdot 67 + 267 \cdot 95 + 240 \cdot 102}{657} = \$91.16$$

Es claro que el mismo resultado se obtiene al dividir el ingreso total para el número total de quintales de cacao vendidos.

El objetivo de este ejercicio fue demostrar el empleo de la media aritmética y de la media ponderada en respuesta a preguntas diferentes.

**Comentario:** Tanto la media geométrica como la media armónica tienen aplicaciones importantes. Sin embargo, en las pruebas estadísticas, la media aritmética es la que tiene mayor relevancia. En este texto nos referiremos desde este punto en adelante a la media aritmética, si es que no se especifica lo contrario.

### 3.4. Parámetros de dispersión

Los parámetros de dispersión son medidas que nos dan una idea de cuán diferentes o separados están los datos u observaciones de una población. Los parámetros de *dispersión* más utilizados son: *rango* (o *amplitud*), *varianza* (o *cuadrado medio*), *desviación estándar* y *coeficiente de variación*.

- *Rango* ( $Ra$ ) es la diferencia entre el valor máximo y el valor mínimo de un conjunto de datos:  $Ra = Y_{max} - Y_{min}$
- *Varianza*. Se representa con la letra griega sigma minúscula al cuadrado  $\sigma^2$ . Es una medida de la variación de los datos con relación a la media. Se calcula a partir del promedio de los cuadrados de las diferencias (o desviaciones) de cada uno de los datos que componen la población, con relación a la media. Por esta razón, un término sinónimo de *varianza* es *cuadrado medio*. Como esta fórmula demuestra el origen de la varianza, se la denomina “fórmula de origen” y es la siguiente:

$$\sigma^2 = \frac{\sum (Y_i - \mu)^2}{N} \quad (5)$$

Existe otra fórmula de la varianza que se deriva de la fórmula (5). Esta fórmula, denominada “fórmula computacional”, es la que más comúnmente se usa y está programada en las calculadoras y en los programas computarizados. La fórmula computacional es:

$$\sigma^2 = \frac{\sum Y_i^2 - \frac{(\sum Y_i)^2}{N}}{N} \quad (6)$$

Las dos fórmulas anteriores son obviamente equivalentes. Sin embargo, con la fórmula (5) se pierde precisión por efecto de redondeo.

La fórmula (6) se obtiene algebraicamente a partir de la fórmula (5) mediante los siguientes pasos:

1. Se despeja el denominador en la fórmula (5):

$$N\sigma^2 = \sum (Y_i - \mu)^2$$

2. Se expande el binomio:

$$\sum (Y_i - \mu)^2 = \sum (Y_i^2 + 2Y_i + \mu^2)$$

3. Se aplica las reglas de la sumatoria (Sección 2.5):

$$\sum (Y_i^2 + 2Y_i + \mu^2) = \sum Y_i^2 - 2\mu \sum Y_i + \sum \mu^2$$

4. De acuerdo con las reglas de la sumatoria,  $\sum \mu^2 = N\mu^2$ , por lo que se obtiene lo siguiente al reemplazar el valor de la media  $\mu$  por su igualdad  $\frac{\sum Y_i}{N}$ :

$$\sum Y_i^2 - 2 \frac{(\sum Y_i)^2}{N} + N \frac{(\sum Y_i)^2}{(N)^2}$$

5. Finalmente, después de simplificar se tiene:

$$\sum Y_i^2 - \frac{(\sum Y_i)^2}{N}$$

Con lo que se demuestra la igualdad de los términos:

$$N\sigma^2 = \sum (Y_i - \mu)^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{N}$$

Estas igualdades tienen un nombre común: “Sumatoria de Cuadrados Corregida” o sencillamente SC. Cuando se dividen para “N” se obtiene la Varianza que es, como se indicó, un parámetro de dispersión.

Continuamos con los parámetros de dispersión:

- *Desviación Estándar*. Se representa con la letra griega sigma minúscula  $\sigma$ . La desviación estándar es la raíz cuadrada de la varianza

$$\sigma = \sqrt{\sigma^2} \quad (7)$$

Las unidades de la varianza y de la desviación estándar están dadas por el tipo de medida. Por ejemplo, si la varianza de la producción de leche de cinco fincas es igual a  $\sigma^2 = 64 \text{ litros}^2$  la desviación estándar será igual a:  $\sigma = 8 \text{ litros}$ . La ventaja de la desviación estándar, con respecto a la varianza en la interpretación de los datos, es clara. Las unidades de la varianza son al cuadrado, mientras que las unidades de la desviación estándar son las mismas de las observaciones y de la media. Esto permite, por ejemplo, sumar desviaciones estándar a la media.

- *Coefficiente de Variación (CV)*. Es la medida que indica, en porcentaje, cuánto varían los datos alrededor de la media. La fórmula es la siguiente:

$$CV = \frac{\sigma}{\mu} * 100 \quad (8)$$

---

**Ejercicio 3.4.1** Encuentre el Modo, la Mediana, la Media, el Rango, la Varianza, la Desviación Estándar y el Coeficiente de Variación de la siguiente población: 9, 15, 9, 5, 6, 4, 8.

Para encontrar el Modo y la Mediana es conveniente ordenar los datos: 4, 5, 6, 8, 9, 9, 15.

- El Modo ( $Mo$ ), la observación que más se repite, es igual a 9
- La Mediana ( $Me$ ), la observación que se encuentra en el centro del conjunto de datos es igual a 8
- La Media  $\mu$  es igual a:  $\frac{\sum Y_i}{N} = \frac{56}{7} = 8$
- El Rango,  $Ra$ , es igual a:  $Y_{max} - Y_{min} = 15 - 4 = 11$
- La Varianza es igual a:

$$\sigma^2 = \frac{\sum (Y_i - \mu)^2}{N} = \frac{(4-8)^2 + (5-8)^2 + (6-8)^2 + (8-8)^2 + (9-8)^2 + (9-8)^2 + (15-8)^2}{7}$$

$$\sigma^2 = \frac{-4^2 - 3^2 - 2^2 + 0^2 + 1^2 + 1^2 + 7^2}{7} = \frac{80}{7} = 11.43$$

El cálculo de la varianza se fundamenta en el principio de cuadrados mínimos. Es decir, si en vez de la media se usaría un valor cualquiera para encontrar las desviaciones y luego se las elevaría al cuadrado, la “varianza” siempre sería, en este ejemplo, mayor que 11.43.

Note también que la varianza nunca puede ser un valor negativo. Los límites de la varianza son sólo valores positivos, de cero a infinito.

La varianza puede también encontrarse con la fórmula computacional, que es la más empleada:

$$\sigma^2 = \frac{\sum y_i^2 - \frac{(\sum y_i)^2}{N}}{N} = \frac{528 - \frac{(56)^2}{7}}{7} = 11.43$$

- La desviación estándar es igual a:  $\sigma = \sqrt{\sigma^2} = 3.38$
- El Coeficiente de Variación es igual a  $CV = \frac{\sigma}{\mu} * 100 = \frac{3.38}{8} * 100 = 42\%$

### 3.5. Propiedades de la media y varianza

La media, varianza y desviación estándar tienen las siguientes propiedades:

- Si a cada observación individual de un conjunto de datos se le suma o resta una cantidad constante C, la nueva media es igual a la media original sumada o restada la constante. La varianza y la desviación estándar permanecen inalteradas.
- Si a cada observación individual de un conjunto de datos se le multiplica o divide por una cantidad constante C, la nueva media es igual a la media original multiplicada o dividida por la constante. La nueva varianza es igual a la varianza original multiplicada o dividida por la constante al cuadrado. La nueva desviación estándar es igual a la original multiplicada o dividida por la constante.

Tomemos como ejemplo las siguientes cinco observaciones: 10, 12, 8, 6, 14. La media, la varianza y la desviación estándar de este grupo de datos son:  $\mu_o = 10$ ,  $\sigma_o^2 = 8$  y  $\sigma_o = 2.83$ , respectivamente. Llamemos a estos resultados, los parámetros originales (los subíndices con la letra “o” indican esto). Si a cada uno de esos datos le sumamos, como ejemplo, la constante

$C = 4$ , las nuevas media, varianza y desviación estándar (identificadas con el subíndice “n”) resultan ser:

$$\mu_n = \mu_o + C = 10 + 4 = 14$$

$$\sigma_n^2 = \sigma_o^2 = 8$$

$$\sigma_n = \sigma_o = 2.83$$

Si a las mismas cinco observaciones: 10, 12, 8, 6, 14 se las divide para una constante, digamos  $C = 2$ , los nuevos parámetros serán:

$$\mu_n = \mu_o / C = 10 / 2 = 5$$

$$\sigma_n^2 = \sigma_o^2 / C^2 = 8 / 2^2 = 2$$

$$\sigma_n = \sigma_o / C = 2.83 / 2 = 1.41$$

Note que a la nueva varianza se la encuentra dividiendo la varianza original para la *constante al cuadrado*.

Aunque no se ha presentado aún los estadígrafos *media, varianza y desviación estándar*, que son características de las muestras, las propiedades descritas para los parámetros de la población son iguales para los estadígrafos de las muestras.

---

**Ejercicio 3.5.1** Supongamos la siguiente población original: 15, 17, 11, 12, 13, 16. Los parámetros son:  $\mu_o = 14$ ,  $\sigma_o^2 = 5.60$ ,  $\sigma_o = 2.37$

- a. Si a cada observación se le suma la constante  $C = 10$ , tenemos una nueva población: 25, 27, 21, 22, 23, 26. De acuerdo con las propiedades de la media y de la varianza, ¿cuáles serán los valores de los parámetros de esta nueva población? Respuesta:

$$\mu_n = \mu_o + 10 = 24, \quad \sigma_n^2 = \sigma_o^2 = 5.60, \quad \sigma_n = \sigma_o = 2.37$$

- b. Si a cada observación de la población original (15, 17, 11, 12, 13, 16) se la multiplica por la constante  $C = 4$ , ¿cuáles serán los valores de la nueva población?

$$\mu_n = \mu_o * 4 = 56, \quad \sigma_n^2 = \sigma_o^2 * 4^2 = 89.60, \quad \sigma_n = \sigma_o * 4 = 9.4$$

---

**Ejercicio 3.5.2.** Si en una investigación se encuentra que la media es 60 libras, la varianza 400 libras cuadradas y la desviación estándar 20 libras, ¿cuáles son los

valores de la media, la varianza y la desviación estándar en kilogramos? (Recordar que 1lb es igual a  $1/2.2 \text{ kg} = 0.454 \text{ kg}$ .)

$$\mu_{kg} = \mu_{lb} * \frac{1}{2.2} = 60 * \frac{1}{2.2} = 27.3 \text{ kg}$$

$$\sigma_{kg}^2 = \sigma_{lb}^2 * \left(\frac{1}{2.2}\right)^2 = 400 * \frac{1}{4.84} = 82.6 \text{ kg}^2$$

$$\sigma_{kg} = \sigma_{lb} * \frac{1}{2.2} = 20 * \frac{1}{2.2} = 9.1 \text{ kg}$$

### 3.6. Frecuencias, Histogramas y Curvas

Como el número de datos que se recogen en los experimentos es generalmente elevado, por conveniencia se los agrupa en términos de la frecuencia  $f_i$  con que se presentan. De esta forma resumida, todos los datos o valores de una población pueden presentarse en un cuadro. Como ejemplo, si hubo 20 plantas que midieron 205 cm, no hace falta escribir 20 veces el dato 205 cm; basta indicar como centro de clase  $Y_i = 205 \text{ cm}$  y como frecuencia  $f_i = 20$ .

Todo un conjunto de observaciones puede presentarse en subconjuntos o *clases*, según el criterio del investigador. Los límites de cada *clase* son los valores extremos de los subconjuntos. El *intervalo* es la amplitud (o rango) de las clases. Por lo general y para mantener la simetría, los intervalos son iguales en todas las clases de un conjunto. Los *centros de las clases* son los promedios de los límites de cada clase; esto es, los valores centrales  $Y_i$  de los sub-conjuntos. Con este método, que simplifica los cálculos, los parámetros de la población pueden encontrarse con menor esfuerzo a partir del cuadro de frecuencias que se forme.

La agrupación de datos y el método relacionado para calcular la media, la varianza y otros parámetros y estadígrafos puede parecer obsoleto, por la velocidad actual de las computadoras y la facilidad de uso. Sin embargo, este procedimiento conduce a la generación de gráficos y curvas y, por lo tanto, a un mejor entendimiento del origen de las distribuciones.

---

**Ejercicio 3.6.1.** Las siguientes observaciones corresponden a un trabajo en el que se evaluó el número de nódulos en las raíces de 120 plantas de gandul o fréjol de palo (*Cajanus cajan*)

Cuadro 1. Número de nódulos por planta

11	21	26	27	29	31	32	35	36	38	40	42	44	46	50
14	21	26	27	30	31	32	35	37	38	40	43	44	47	50
15	23	27	27	30	31	33	35	37	39	40	43	44	47	50
15	23	27	27	30	31	33	35	37	39	41	43	45	47	54
17	24	27	28	30	31	33	35	37	39	42	43	46	48	57
17	24	27	28	30	31	33	36	38	39	42	43	46	49	57
18	24	27	28	30	31	34	36	38	39	42	43	46	49	58
20	25	27	29	31	32	35	36	38	40	42	43	46	49	58

- a. Presente un cuadro de frecuencias, mediante el empleo del programa *Statistix*, con un intervalo cuyo valor sea 5. Indique, además, las frecuencias de cada clase y la frecuencia acumulada para cada clase.

Cuadro 2. Distribución de frecuencias de la variable número de nódulos en la leguminosa gandul o fréjol de palo

Low	High	Freq	Cumulative		
			Percent	Freq	Percent
9.5	14.5	2	1,7	2	1,7
14.5	19.5	5	4,2	7	5,8
19.5	24.5	8	6,7	15	12,5
24.5	29.5	18	15,0	33	27,5
29.5	34.5	22	18,3	55	45,8
34.5	39.5	24	20,0	79	65,8
39.5	44.5	20	16,7	99	82,5
44.5	49.5	13	10,8	112	93,3
49.5	54.5	4	3,3	116	96,7
54.5	59.5	4	3,3	120	100,0
Total	120	100,0			

- b. A partir del cuadro de frecuencias anterior, presente un histograma para las 120 observaciones



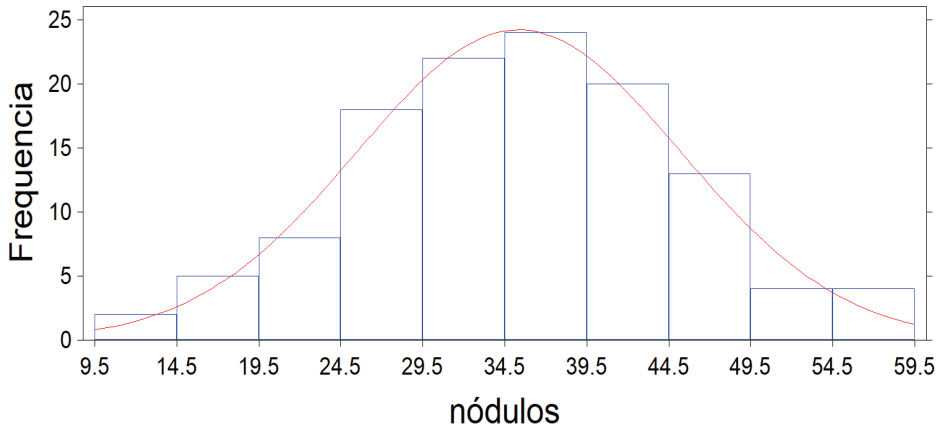


Gráfico 1. Histograma de las frecuencias de los nódulos de gándul

**Comentario:** El objetivo del ejercicio anterior fue demostrar cómo puede resumirse una gran cantidad de datos mediante un cuadro de frecuencias. La misma información puede resumirse también a través de un histograma. Las barras representan las clases y las alturas de las barras, las frecuencias. Puede verse, además, que esta distribución tiende a una curva “normal” (la línea curva en el gráfico), como se discutirá más adelante.

A partir de un cuadro de frecuencias, como se presenta en el ejercicio siguiente, pueden encontrarse las características de las poblaciones o muestras, esto es, los parámetros o estadígrafos, dependiendo de si se trata de una población o de una muestra.

**Ejercicio 3.6.2.** Supongamos que la siguiente colección de datos constituye una población: 20, 15, 5, 10, 20, 15, 15, 30, 10, 35, 15, 25, 30.

- a. Crear un cuadro de frecuencias.

Si los datos no están ordenados se procede a hacerlo:

5, 10, 10, 15, 15, 15, 15, 20, 20, 20, 25, 25, 30.

A continuación, se presenta un *Cuadro de Frecuencias*, en el que puede incluirse cálculos preliminares para encontrar los parámetros de la población:

Cuadro 1. Centros de clase y otras variables de un cuadro de frecuencias

$Y_i$ (Centros de clase)	$f_i$	$f_i Y_i$	$Y_i^2$	$f_i Y_i^2$
5	1	5	25	25
10	2	20	100	200
15	4	60	225	900
20	3	60	400	1200
25	2	50	625	1250
30	1	30	900	900
	$\sum f_i = 13$	$\sum f_i Y_i = 225$	$\sum Y_i^2 = 2275$	$\sum f_i Y_i^2 = 4475$

b. Encontrar el Modo, Mediana, Media, Rango, Varianza y Desviación estándar de esta población.

- El Modo (Mo) es igual a 15 (note que es el centro de clase con mayor frecuencia).
- La Mediana (Me) es también igual a 15 (ver datos ordenados).
- La Media ( $\mu$ ) se calcula partir del cuadro de frecuencias con la siguiente fórmula:

$$\mu = \frac{\sum f_i Y_i}{\sum f_i} = \frac{225}{13} = 17.31$$

- La varianza calculada a partir del cuadro de frecuencias es:

$$\sigma^2 = \frac{\sum f_i Y_i^2 - \frac{(\sum f_i Y_i)^2}{N}}{N = \sum f_i} = \frac{4475 - \frac{225^2}{13}}{13} = 44.67$$

- La desviación estándar es:  $\sigma = \sqrt{\sigma^2} = \sqrt{44.67} = 6.68$

**Comentario:** Con la disponibilidad de programas de computación el uso del método descrito en el ejercicio anterior para calcular parámetros parecería anticuado, sin embargo, ayuda a una mayor familiarización con los conceptos de parámetros, estadísticas y manejo de datos. Además, ayuda a comprender el origen de las diferentes distribuciones, incluyendo, en especial, la *distribución normal*.

Como se vio, un histograma es la representación gráfica mediante barras, de un conjunto de datos y tiene muchas aplicaciones. Como ejemplo, la producción mundial<sup>9</sup> en el año 2015 de algunos cultivos se presenta en el histograma siguiente:

9 Fuente: [www.faostat.org/site/339](http://www.faostat.org/site/339)

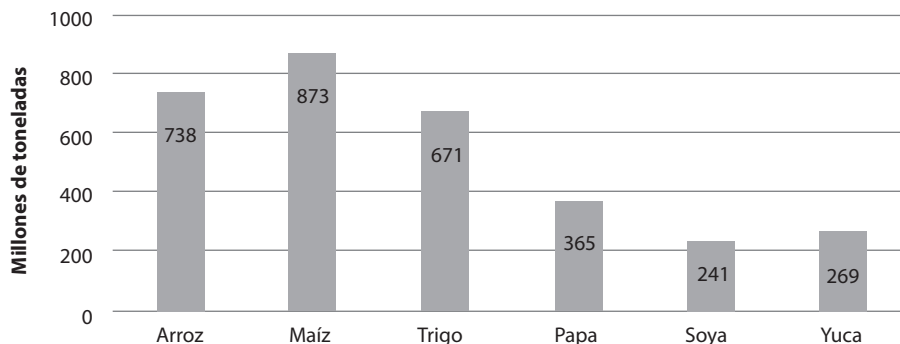
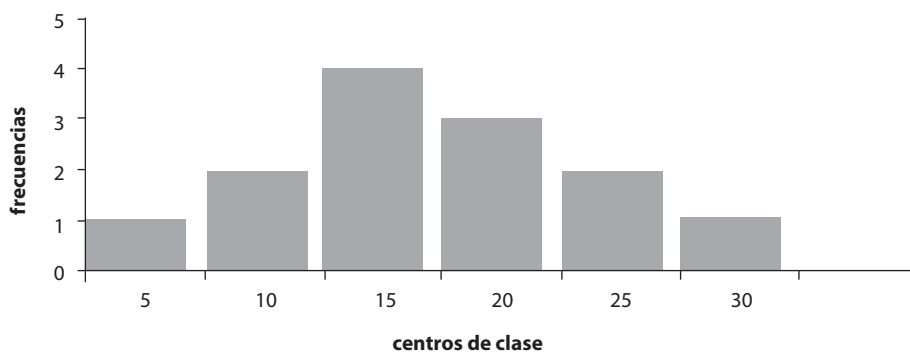


Gráfico 3.6. Los cultivos con el mayor volumen de producción a nivel mundial

En el siguiente ejercicio se pide presentar un histograma a partir del cuadro de frecuencias del Ejercicio 3.6.2. Esto sirve como introducción al concepto de *curva normal*.

**Ejercicio 3.6.3.** Presentar, a partir del Cuadro de frecuencias del Ejercicio 3.6.2, un histograma:



**Comentario:** En el gráfico del ejercicio 3.6.3 es importante observar que los centros de clase están en el punto medio de las barras. Los límites de las clases pueden estar fusionados o no. Si se unen los centros de las clases en la parte superior de las barras con una línea, se obtiene una *curva de frecuencias*. Mientras más barras haya, la curva será más suave (menos angulada). La distribución, esto es el tamaño relativo de las barras o la forma de la curva de frecuencias, es característica de los datos de cada población.

### 3.7. La Distribución Normal de una Población

Todas las curvas de frecuencias, construidas a partir de cuadros de frecuencias, tienen una distribución característica. Sin embargo, cuando la forma de la curva de frecuencias es simétrica y similar a una campana, decimos que la población presenta una *curva o distribución normal*.

Son características propias de la curva o distribución normal las siguientes:

- La fórmula para la ordenada o el alto de la curva normal es:

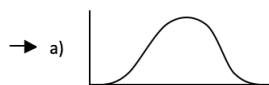
$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(Y-\mu)^2/2\sigma^2}$$

Note la presencia de los parámetros de la población en la fórmula.

- El modo, la mediana y la media son coincidentes en el centro de la distribución.

- Es simétrica. El parámetro de simetría es alfa-tres:  $\alpha_3 = \frac{\sum(Y_i-\mu)^3}{N\sigma^3}$

- a. Si  $\alpha_3$  es igual a cero, la curva es normal:



- b. Si  $\alpha_3$  es negativa, la curva no es normal. Tiene cola izquierda y sesgo negativo:



- c. Si  $\alpha_3$  es positiva, la curva tampoco es normal. Tiene cola derecha y sesgo positivo:



- La punta de la curva tiene una redondez característica, medida por el parámetro  $\alpha_4$ , llamado Curtosis:  $\alpha_4 = \frac{\sum(Y_i-\mu)^4}{N\sigma^4}$

- a. Si  $\alpha_4 = 3$ , la curva es normal (acúrtica).
- b. Si  $\alpha_4 < 3$ , la punta de la curva es muy aplanada y deja de ser normal (platicúrtica).
- c. Si  $\alpha_4 > 3$ , la punta es muy aguda y la distribución no es normal (leptocúrtica).

Mediante el uso de programas de computación se pueden encontrar rápidamente los valores de  $\alpha_3$  y  $\alpha_4$ . De esta forma es posible estimar si una población se acerca a una distribución normal o no.

- En teoría el rango de la distribución normal va de menos infinito a más infinito. Sin embargo, en la práctica se considera que el 100% de las observaciones están comprendidas dentro de la curva de la distribución normal.
- Una propiedad importante de la curva o distribución normal, como se indica en el gráfico 3.7.1, es que:
  - a. el 68.26% de las observaciones se encuentra entre  $\mu - \sigma$  y  $\mu + \sigma$
  - b. el 95.44% entre  $\mu - 2\sigma$  y  $\mu + 2\sigma$
  - c. el 99.74% entre  $\mu - 3\sigma$  y  $\mu + 3\sigma$

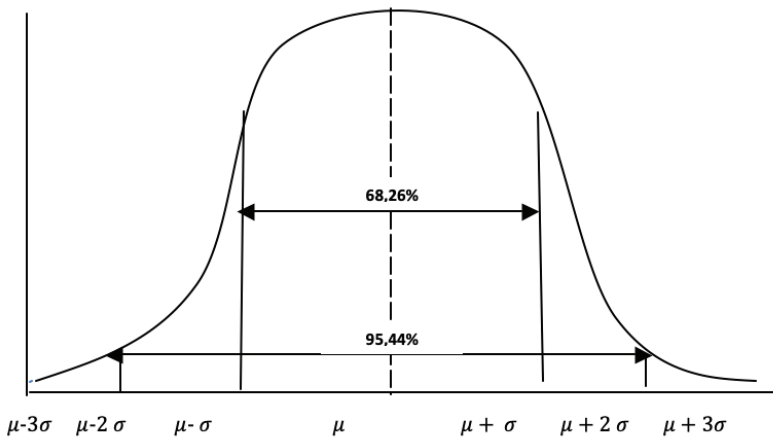


Gráfico 3.7.1. La curva de la distribución normal y sus características

Una notación conveniente para la distribución normal es la siguiente:  $Y_i \sim DN(\mu, \sigma^2)$ . Por ejemplo,  $Y_i \sim DN(20, 80)$  nos indica que se trata de una población en la que las observaciones  $Y_i$  están distribuidas normalmente ( $DN$ ) y cuya media y varianza son 20 y 80, respectivamente.

### 3.8. La distribución Z

La distribución  $Z$  es una distribución normal, estandarizada, en la que  $\mu = 0$  y  $\sigma^2 = 1$ . Cualquier población distribuida normalmente puede estandarizarse mediante la fórmula:

$$Z = \frac{(Y_i - \mu)}{\sigma} \quad (9)$$

La ventaja de la distribución  $Z$  es que podemos calcular fácilmente las probabilidades asociadas con esta curva. Si se tiene una estimación de la media y de la varianza de una población se puede calcular a partir de qué valor se encuentra, digamos, el 10 % más alto de los datos. Esto ayuda en varios procesos de selección de observaciones en diferentes campos. También podemos probar hipótesis relacionadas con la igualdad o no de medias de dos poblaciones. Para todo esto empleamos la tabla  $Z$  que se encuentra en el *Apéndice A*, o en su defecto, la opción de probabilidades del programa *Statistix*.

La importancia de la distribución normal estriba en el hecho de que muchas poblaciones de interés biológico, agronómico y de otros campos siguen esta distribución. Además, en el proceso de muestreo, como se verá más adelante, la distribución teórica de las medias de todas las muestras, independientemente de la forma de distribución de la población, se acerca o sigue la distribución normal. En otras palabras, la distribución normal es siempre el punto de referencia para cualquier población. Esta relación se estudia en el *Teorema del Límite Central*.

### 3.9. ¿Qué es una muestra?

Una muestra es una parte de la población. La Estadística nos proporciona los principios y métodos que permiten realizar inferencias válidas sobre las características de la población a partir de muestras. En general las poblaciones son muy grandes y muchas veces inasequibles. Aun si hubiera cómo llegar a conocer todas las características de la población con la que se trabaja, esto sería impráctico por el costo en tiempo y otros recursos. Por este motivo en investigación se trabaja con muestras.

Ejemplo de una muestra es un dato del peso de una mazorca de maíz del híbrido DK15. Esta es una muestra de una población muy grande, prácticamente infinita, de pesos de mazorca de ese híbrido. El tamaño de la muestra del ejemplo es  $n = 1$ ; sin embargo, las muestras por lo general son más grandes y pueden ser de cualquier tamaño.

Un requisito importante para que tengan validez las inferencias que se generen, es que la muestra sea representativa de la población. Para esto, la muestra debe ser una *muestra al azar*, que es aquella en la que *todos y cada uno de los componentes de la población tienen iguales probabilidades de ser incluidos en la muestra*.

La única forma posible de obtener una muestra al azar es mediante el empleo de algún proceso mecánico o electrónico, como la extracción de papeles numerados o fichas, el uso de tablas de números al azar, o un programa de computación. No es posible inventarse una muestra al azar, porque en general las personas tienden a favorecer algunos números más que otros.

### 3.10. Estadísticas y sus características

Las *estadísticas* o *estadígrafos* son características de las muestras. Así como las poblaciones se caracterizan por parámetros, las muestras se caracterizan por estadísticas. Los parámetros son constantes, las estadísticas son variables. Es muy poco probable que dos muestras al azar, aunque extraídas de la misma población, tengan las mismas estadísticas.

**Los estadígrafos más comunes de tendencia central son:**

- el modo de la muestra,  $mo$
- la mediana de la muestra,  $me$
- la media de la muestra,  $\bar{Y}$

La fórmula de la media de la muestra es:

$$\bar{Y} = \frac{\sum Y_i}{n} \quad (10)$$

Las definiciones de estos estadígrafos son similares a las dadas para los parámetros.

**Los estadígrafos de dispersión son:**

- El rango de la muestra,  $ra$
- La varianza de la muestra,  $S^2$ . La varianza de la muestra se puede encontrar con cualquiera de las dos fórmulas siguientes:

$$S^2 = \frac{\sum (Y_i - \bar{Y})^2}{n-1} \quad (11)$$

$$S^2 = \frac{\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}}{n-1} \quad (12)$$

Note el parecido de estas fórmulas de la varianza de la muestra con las fórmulas de la varianza de la población (sección 3.4), excepto por el denominador. Como la cantidad  $\frac{(\sum Y_i)^2}{n}$  es el Factor de Corrección (*FC*), la fórmula (12) puede expresarse más sucintamente de esta manera:  $S^2 = \frac{\sum Y_i^2 - FC}{n-1}$

El término  $(n-1)$  se denomina *grados de libertad* y es necesario en el cálculo de la varianza porque por el tamaño de la muestra generalmente se subestima la sumatoria de cuadrados corregida (los términos en el numerador de la fórmula).

Así como la media de la muestra  $\bar{Y}$  es el mejor estimador de la media de la población  $\mu$  así la varianza de la muestra  $S^2$  es el mejor estimador de la varianza de la población  $\sigma^2$ . Esto se fundamenta en el modelo lineal aditivo, como se verá más adelante.

- La desviación estándar de la muestra,  $S$

La fórmula es:

$$S = \sqrt{S^2} \quad (13)$$

*La desviación estándar de la muestra es el mejor estimador de la desviación estándar de la población.*

- La desviación estándar de la media de la muestra, más conocido como error estándar,  $S_{\bar{y}}$

La fórmula es:

$$S_{\bar{y}} = \sqrt{\frac{S^2}{n}} = \frac{S}{\sqrt{n}} \quad (14)$$

*El error estándar es el mejor estimador de la desviación estándar de la población de medias.* El Teorema del Límite Central, que se estudiará en el capítulo cinco, tiene que ver con el origen del error estándar, concepto en que se fundamenta la estadística inferencial.

- El coeficiente de variación de la muestra,  $cv$

La fórmula es:  $cv = \frac{S}{\bar{Y}} * 100 \quad (15)$

Por lo demás, los cálculos para encontrar los estadígrafos son similares a los empleados para los parámetros. Esto se notará en los ejercicios que se presentan a



continuación. Sin embargo, es de suma importancia no confundir los parámetros con los estadígrafos. Los parámetros son constantes, por lo general desconocidos y pertenecen a las poblaciones. Los estadígrafos son variables que se calculan a partir de muestras extraídas de poblaciones de interés para el investigador. De una misma población se pueden extraer muchas muestras. El fundamento de la Estadística Inferencial es estimar los valores de los parámetros a partir de los estadígrafos. En otras palabras, los estadígrafos sirven como estimadores de los parámetros.

**Ejercicio 3.10.1.** De una población de pesos de malezas por parcela, se extrajo la siguiente muestra de tamaño ocho, ( $n = 8$ ): 25, 32, 20, 18, 24, 30, 28, 23. El peso está en gramos. Encuentre las siguientes estadísticas: a) la sumatoria de las observaciones, b) la sumatoria de cuadrados no corregida, c) el factor de corrección, d) la media, e) la varianza, f) la desviación estándar, g) el error estándar, h) el coeficiente de variación. Indique las unidades en cada caso.

a.  $\sum Y_i = 25 + 32 + 20 + 18 + 24 + 30 + 28 + 23 = 200 \text{ gramos}$

b.  $\sum Y_i^2 = 25^2 + 32^2 + 20^2 + 18^2 + 24^2 + 30^2 + 28^2 + 23^2 = 5162 \text{ gramos}^2$

c. el factor de corrección:  $\frac{(\sum Y_i)^2}{n} = \frac{(200)^2}{8} = 5000 \text{ gramos}^2$

d. la media:  $\bar{Y} = \frac{\sum Y_i}{n} = \frac{200}{8} = 25 \text{ gramos}$

e. la varianza:  $S^2 = \frac{\sum Y_i^2 - FC}{n-1} = \frac{5162-5000}{7} = 23.14 \text{ gramos}^2$

f. la desviación estándar:  $S = \sqrt{S^2} = 4.81 \text{ gramos}$

g. el error estándar:  $S_{\bar{Y}} = \sqrt{\frac{S^2}{n}} = 1.70 \text{ gramos}$

h. el coeficiente de variación:  $cv = \frac{S}{\bar{Y}} * 100 = 19.24\%$

El siguiente ejercicio sirve como introducción al empleo del programa computarizado *Statistix*. Se escogió este programa para la realización de ejercicios y demostraciones por su facilidad de manejo, su rapidez y sus múltiples opciones para los análisis de datos. (*Statistix 9, User's Manual. 2008.*)

**Ejercicio 3.10.2.** Los siguientes datos corresponden a una muestra de 20 plantas de café robusta, tomadas al azar de una plantación. Note que los datos están organizados en Columnas (seis variables) y en hileras (20 casos que corresponden a las veinte plantas). Esta es la forma como los datos deben organizarse para los análisis en diferentes programas de computación. Las seis variables son: Var1=las

20 plantas, Var2=altura de planta, Var3=diámetro del tallo en cm, Var4=diámetro de la copa en cm, Var5=número de ramas, Var6=largo de las ramas en cm.

**Cuadro 1. Datos de 20 plantas de una plantación de café robusta en el primer año**

<b>Var1</b>	<b>Var2</b>	<b>Var3</b>	<b>Var4</b>	<b>Var5</b>	<b>Var6</b>
1	117	2.8	153	30	82
2	110	2.8	135	32	68
3	140	3	145	30	73
4	127	3.6	150	38	79
5	150	3	150	36	80
6	165	3.1	160	40	83
7	150	2.8	160	36	78
8	165	3.6	145	44	84
9	160	4	185	44	120
10	143	3.4	175	40	87
11	165	3.4	165	44	80
12	155	3.8	175	44	83
13	155	4	180	42	85
14	147	3.8	180	42	97
15	160	3.8	190	48	100
16	165	4.3	190	44	98
17	165	3	170	46	85
18	150	3	175	42	81
19	150	3.1	165	40	93
20	125	3.2	150	25	76

**Cuadro 2. Estadística descriptiva (Statistix Descriptive Statistics)**

	<b>Var2</b>	<b>Var3</b>	<b>Var4</b>	<b>Var5</b>	<b>Var6</b>
Media	148.20	3.3750	164.90	39.350	85.600
Desv. Est	16.719	0.4598	16.212	6.1067	11.486
Varianza	279.54	0.2114	262.83	37.292	131.94
Error Est	3.7386	0.1028	3.6251	1.3655	2.5684
V. Mínimo	110.00	2.8000	135.00	25.000	68.000
1er Cuartil	140.75	3.0000	150.00	36.000	79.250
Mediana	150.00	3.3000	165.00	41.000	83.000
3er Cuartil	163.75	3.8000	178.75	44.000	91.500
V. Máximo	165.00	4.3000	190.00	48.000	120.00

Con el programa *Statistix* se pueden encontrar varias estadísticas, como la media, la desviación estándar, la varianza, el error estándar, el valor mínimo, el 1er cuartil, la mediana o segundo cuartil, el 3er cuartil y el valor máximo de las variables. Puede verse en el Cuadro 2, como ejemplo, para la Var2 (altura de planta), los siguientes resultados: media, 148.20 cm; desviación estándar, 16.72 cm; varianza, 279.54 cm<sup>2</sup>; error estándar, 3.74 cm; valor mínimo, 110 cm; primer cuartil, 140.75 cm; mediana o segundo cuartil, 150 cm; tercer cuartil, 163.75 cm; y valor máximo, 165 cm.

### 3.11 . El modelo lineal aditivo

Es una práctica científica común el tratar de explicar fenómenos naturales mediante modelos. Como ejemplo, la producción de arroz en un área determinada puede explicarse a través de un modelo que tome en cuenta los componentes del rendimiento: número de panículas en un área determinada, número de granos por panícula y peso del grano. Influyen e interactúan en la expresión de los componentes del rendimiento, la constitución genética de las plantas y los factores ambientales.

En estadística cada observación, en teoría, está formada por la media, más un error por razón de muestreo (que puede ser de signo positivo o negativo). Este es el *modelo lineal aditivo*:

$Y_i = \mu + \varepsilon_i$  en donde  $\varepsilon_i$  es el error por razón de muestreo.

Las suposiciones necesarias para la validez de este modelo son:

- Los errores  $\varepsilon_i$  son aleatorios (al azar).
- Los errores  $\varepsilon_i$  no están correlacionados entre ellos, es decir son independientes.
- La sumatoria de los errores de las observaciones tiende a cero con el tamaño de la muestra ( $n$ ) y es igual a cero cuando  $n = N$ , es decir  $\sum \varepsilon_i = 0$ .

De acuerdo con el modelo la media de la muestra es igual a:

$$\bar{Y} = \mu + \frac{\sum \varepsilon_i}{n}$$

La cantidad  $\frac{\sum \varepsilon_i}{n}$  debe disminuir a medida que el tamaño de la muestra aumenta. Esto, porque valores positivos y negativos de  $\varepsilon_i$  tienden a cancelarse unos con otros y además, porque el valor del divisor  $n$  se incrementa. Es decir que  $\frac{\sum \varepsilon_i}{n}$  tiende a cero. En otras palabras  $\bar{Y}$  es un buen estimador de  $\mu$  y está más próximo

a la media  $\mu$  a medida que  $n$  aumenta. Los  $\varepsilon_i$  provienen de una muestra finita que nos permite obtener una estimación de la varianza de la población  $\sigma^2$ . Como no se conoce la media de la población  $\mu$ , los  $\varepsilon_i$  a su vez se estiman por las desviaciones  $(Y_i - \bar{Y})$ , de donde proviene la varianza de la muestra  $S^2$ . Esto es evidente en la fórmula número (11) presentada en páginas anteriores:

$$S^2 = \frac{\sum(Y_i - \bar{Y})^2}{n-1}$$



# Estadística y probabilidades

---

*“La idea de que las probabilidades aumentan o disminuyen dependiendo de cuán recientemente ocurrió un evento, se llama la falacia del jugador. Por ejemplo, si se obtienen 44 caras al lanzar 100 veces una moneda, ésta no desarrollará un sesgo hacia sellos para igualar lo que teóricamente se espera”.* Leonard Mlodinow<sup>10</sup>

## 4.1 . Relación entre Estadística y Probabilidades

Probabilidad se define como el estudio matemático de los eventos que suceden al azar. La estadística se fundamenta en el cálculo de probabilidades. Aunque a la probabilidad de ocurrencia de un evento se la puede expresar ambiguamente, como ejemplo: “es posible que mañana llueva”, en estadística se usa una escala numérica del cero al uno o porcentajes del 0 al 100%. Como ejemplo: “hay 50% de probabilidad que mañana llueva”. Cuando un evento es imposible que ocurra indicamos que tiene probabilidad igual a cero,  $P(E_i) = 0$ . En cambio cuando hay certeza sobre la ocurrencia de un evento decimos que su probabilidad es igual a uno,  $P(E_i) = 1$ .

En general, esto podemos expresarlo así:

$$0 \leq P(E_i) \leq 1, \text{ en donde } P(E_i) \text{ es la probabilidad del evento } E_i$$

Siempre es importante en el cálculo de probabilidades tener en cuenta el Espacio de Muestreo (EM). Como ejemplo, si lanzamos una moneda al aire se esperan dos posibles resultados o eventos: que caiga cara (C) o sello (S). Esto se representa de la siguiente forma:  $EM = [C, S]$

El Espacio de Muestreo cuando lanzamos simultáneamente dos monedas es:  $EM = [CC, CS, SC, SS]$

---

10 Físico estadounidense. [https://www.goodreads.com/author/quotes/1399.Leonard\\_Mlodinow](https://www.goodreads.com/author/quotes/1399.Leonard_Mlodinow)

La población de número de caras-caras, caras-sellos, sellos-caras y sellos-sellos, en el Espacio de Muestreo anterior, es infinita. Pero tanto si lanzamos dos monedas 100 veces como mil veces, la población de referencia es la misma e inalterable.

## 4.2. Reglas para el cálculo de probabilidades

Existen dos principios o reglas básicas para el cálculo de probabilidades: 1) la regla de la multiplicación y 2) la regla de la adición.

1. **La regla de la multiplicación** indica que: “Cuando dos eventos A y B son independientes, la probabilidad de la ocurrencia simultánea de los dos eventos es igual al producto de las probabilidades individuales de A y B”. Esto se representa con la siguiente fórmula en la que  $P$  denota “probabilidad”:  $P[A \cap B] = P(A) \cdot P(B)$  (16)

(El operador  $U$  invertida indica la independencia de los dos eventos.)

Al lanzar una moneda al aire puede caer “cara” o “sello”. Llamemos  $p$  a la probabilidad de que caiga cara, que por experiencia sabemos que es igual a  $\frac{1}{2}$ , si la moneda es “buena”. Igualmente, llamemos  $q$  a la probabilidad de que caiga sello y es también igual a  $\frac{1}{2}$ . Estos dos eventos son independientes, es decir, un resultado no afecta a otro resultado en una secuencia de eventos.

Con esta información podemos preguntar, ¿cuál es la probabilidad de que al lanzar dos monedas al aire las dos caigan cara? Si la probabilidad de que la primera moneda caiga cara es  $p = \frac{1}{2}$  y la probabilidad de que la segunda moneda también caiga cara es  $p = \frac{1}{2}$ , entonces la probabilidad de que las dos monedas caigan cara será:

$$P(E_i) = p^2 = 1/2 \cdot 1/2 = (1/2)^2 = 1/4$$

2. **La regla de la adición** plantea que “Cuando dos eventos A y B no pueden ocurrir simultáneamente, la probabilidad de ocurrencia de cualquiera de los dos eventos es igual a la suma de las probabilidades individuales de A y B”. Esto se representa en la siguiente fórmula, en la que el operador  $U$  indica la unión de dos eventos mutuamente excluyentes:

$$P[A \cup B] = P(A) + P(B) \quad (17)$$

En una baraja de 52 cartas (en la que hay cuatro cartas de cada denominación), la probabilidad de extraer un AS es  $4/52=1/13$ . La probabilidad de extraer una K también es  $4/52 = 1/13$ . Consecuentemente la probabilidad de extraer un AS o una K será:

$$1/13 + 1/13 = 2/13.$$

**Ejercicio 4.2.1** ¿Cuál es la probabilidad de que al lanzar tres monedas al aire obtengamos dos caras y un sello? Si se supone que cara =  $p = 1/2$  y sello =  $q = 1/2$ , entonces existen ocho posibles eventos al lanzar tres monedas al aire, como se demuestra en el cuadro a continuación:

Cuadro 1. Posibles eventos al lanzar tres monedas al aire

Eventos	Monedas			Probabilidades	Descripción
	1ra	2da	3ra		
1)	p	p	p	$(1/2)^3 = 1/8$	Tres caras y cero sellos
2)	p	p	q	$(1/2)^3 = 1/8$	Dos caras y un sello
3)	p	q	p	$(1/2)^3 = 1/8$	Dos caras y un sello
4)	p	q	q	$(1/2)^3 = 1/8$	Una cara y dos sellos
5)	q	p	p	$(1/2)^3 = 1/8$	Dos caras y un sello
6)	q	p	q	$(1/2)^3 = 1/8$	Una cara y dos sellos
7)	q	q	p	$(1/2)^3 = 1/8$	Una cara y dos sellos
8)	q	q	q	$(1/2)^3 = 1/8$	Cero caras y tres sellos

Los eventos 2), 3) y 5) tienen resultados idénticos, si no importa qué moneda en particular cayó cara o sello. Entonces, la probabilidad de obtener dos caras y un sello al lanzar tres monedas al aire es igual a  $1/8 + 1/8 + 1/8 = 3/8$

Una forma más sencilla para resolver problemas de este tipo es empleando la fórmula de la expansión del binomio,  $(p + q)^n$ . Note que si  $p + q = 1$ , entonces  $(p + q)^n = 1$ .

Como ejemplo en el ejercicio anterior se tiene:  $p = 1/2$ ,  $q = 1/2$  y  $n = 3$ .

$$\text{Entonces, } (p + q)^3 = p^3 + 3p^2q + 3pq^2 + q^3 = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$$

Los coeficientes de los términos pueden encontrarse directamente a partir de la expansión del binomio, pero es más fácil empleando el “Triángulo de Pascal”:



**Cuadro 4.2.1. Triángulo de Pascal.**  
**Coefficientes para la expansión del binomio  $(p + q)^n$**

				1						n
				1		1				1
			1		2		1			2
		1		3		3		1		3
	1		4		6		4		1	4
	1	5		10		10		5		5
	1	6	15		20		15	6	1	6
1	7	21	35		35		21	7	1	7
1	8	28	56	70	56	28	8	1	8	8

Note que en los extremos de las hileras van los coeficientes valor uno y que en la hilera, abajo de dos números contiguos, va en el medio la suma de esos dos números. El cuadro anterior se ha construido sólo hasta el valor de  $n = 8$ , pero  $n$  puede tomar cualquier valor. El número de términos en cada hilera es igual a  $n+1$ .

**Ejercicio 4.2.2.** Como resultado de un cruzamiento entre dos líneas puras de arveja, una de semillas amarillas y otra de semillas verdes, en la F2 se encuentra que  $\frac{3}{4}$  partes de las semillas son de color amarillo y  $\frac{1}{4}$  parte son de color verde. Si se toman al azar cinco semillas F2:

- ¿Cuál es la probabilidad de que tres sean amarillas y dos verdes?
- ¿Cuál es la probabilidad de que dos sean amarillas y tres verdes?
- ¿Cuál es la probabilidad de que por lo menos una sea amarilla?
- ¿Cuál es la probabilidad de que por lo menos una sea verde?

**Resolución:** El primer paso es expandir el binomio  $(p + q)^n$ , en el que la probabilidad de que la semilla sea amarilla es igual a:  $p = \frac{3}{4}$  y de que la semilla sea verde es igual a:  $q = \frac{1}{4}$ ; el número de individuos, en este caso semillas,  $n = 5$ . Es conveniente, para efectos de este ejercicio, el calcular la contribución de cada término en porcentaje. El resultado de la expansión del binomio es:

$$(p + q)^5 = p^5 + 5p^4q + 10p^3q^2 + 10p^2q^3 + 5pq^4 + q^5 = 1$$

$$(p + q)^5 = \frac{243}{1024} + \frac{405}{1024} + \frac{270}{1024} + \frac{90}{1024} + \frac{15}{1024} + \frac{1}{1024} = 1$$

$$(p + q)^5 = 23.7\% + 39.5\% + 26.4\% + 8.8\% + 1.5\% + 0.1\% = 100\%$$

Note que los exponentes de  $p$  representan el número de semillas amarillas, mientras que los exponentes de  $q$  representan el número de semillas verdes.

- a. Para responder a esta pregunta, la probabilidad de que la muestra esté compuesta por tres semillas amarillas y dos verdes, debemos reemplazar los valores correspondientes en el tercer término de la expansión del binomio. El tercer término ( $10p^3q^2$ ) representa la probabilidad, del evento al azar, de obtener una muestra formada por tres semillas amarillas y dos verdes. Esto es:

$$10p^3q^2 = 10 \left(\frac{3}{4}\right)^3 \left(\frac{1}{4}\right)^2 = 270/1024 = 0.2637, \text{ es decir, } 26.4\%.$$

- b. Es el cuarto término el que nos interesa ahora:

$$10p^2q^3 = 10 \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^3 = 90/1024 = 0.0879, \text{ es decir, } 8.8\%$$

- c. La respuesta se obtiene sumando los cinco primeros términos ya que cualquiera de ellos representa la presencia de por lo menos una semilla amarilla:

$$p^5 + 5p^4q + 10p^3q^2 + 10p^2q^3 + 5pq^4 = 1023/1024 = 0.9990, \text{ es decir, } 99.9\%$$

- d. Para esta pregunta debemos sumar los cinco últimos términos:

$$5p^4q + 10p^3q^2 + 10p^2q^3 + 5pq^4 + q^5 = 781/1024 = 0.7627, \text{ es decir, } 76.3\%.$$

### Otro procedimiento

Existe otro procedimiento cuando únicamente se trata de encontrar la probabilidad de ocurrencia de una determinada combinación. La fórmula que se emplea es la siguiente:

$$P = \frac{n!}{x!(n-x)!} p^x q^{(n-x)} \quad (18)$$

En la que  $P$  mayúscula es la probabilidad del evento,  $n$  es el tamaño del grupo y  $x$  el componente del grupo con probabilidad  $p$ . El signo de admiración (!) significa factorial.

Como ejemplo, para encontrar la probabilidad de dos semillas amarillas y tres verdes, cuando

$$n = 5, x = 2, p = 3/4 \text{ y } q = 1/4 \text{ (literal b del ejercicio anterior):}$$

$$P = \frac{5!}{(2!)(3!)} \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right)^3 = \frac{120 \cdot 9}{16 \cdot 64} = 0.088$$

La probabilidad de ocurrencia del evento descrito (dos semillas amarillas y tres verdes) es entonces 0.088 u 8.8% (como ya se determinó en el ejercicio anterior).

La fórmula anterior tiene mayor utilidad y es más práctica (ahorra tiempo) cuando se tienen valores altos de  $n$  y de  $x$ .

**Ejercicio 4.2.3** En una población de un cultivar de café robusta se conoce que el 10% de las plantas presentan resistencia a la roya.

a) Si se toman 15 plantas al azar, ¿cuál es la probabilidad de que todas tengan resistencia a la roya?

Según la fórmula (18) los datos son: número total de plantas,  $n = 15$ ; número esperado de plantas resistentes,  $x = 15$ ; proporción conocida de plantas resistentes,  $p = 1/10$ ; proporción plantas susceptibles,  $q = 9/10$ . Estos valores se reemplazan en la fórmula (18):

$$P = \frac{15!}{15!(15-15)!} (1/10)^{15} (9/10)^0 = \frac{1}{1 \cdot 10^{15}} = 1 \cdot 10^{-15}$$

Una probabilidad bastante baja. (Recuerde que factorial de cero es igual a uno. También recuerde que cualquier cantidad elevada a cero es igual a uno.)

b) Con la misma información del literal a), ¿cuál es la probabilidad de que 15 plantas tomadas al azar de esa población, 5 plantas presenten resistencia y 10 sean susceptibles?

El número total de plantas,  $n = 15$ ; número esperado de plantas resistentes,  $x = 5$ ; número esperado de plantas susceptibles,  $(n-x) = 10$ . Con estos datos se procede a reemplazar los valores en la fórmula:

$$P = \frac{15!}{5!(15-5)!} (1/10)^5 (9/10)^{10} = 3003 \cdot 0.00001 \cdot 0.3487 = 0.0105$$

### 4.3. La distribución binomial

Cuando las probabilidades de eventos excluyentes son dos,  $p$  y  $q$ , y la ocurrencia de un evento no tiene efecto en el siguiente, como en el lanzamiento de una moneda, se origina una *distribución binomial*. Los ejemplos de los tres ejercicios anteriores corresponden a distribuciones binomiales.

Si se lanzan seis monedas simultáneamente al aire varias veces y si se considera sólo el número de caras en cada lanzamiento, entonces los eventos posibles

son: 0 caras, 1 cara, 2 caras, 3 caras, 4 caras, 5 caras, y 6 caras. Las probabilidades de estos eventos son:

$$(p+q)^6 = p^6 + 6 p^5q + 15 p^4q^2 + 20 p^3q^3 + 15 p^2q^4 + 6 pq^5 + q^6$$

El denominador de cada término es igual a  $2^n$  cuando  $p = \frac{1}{2}$ . En este caso en el que  $n=6$ ,  $2^n$  es igual a 64.

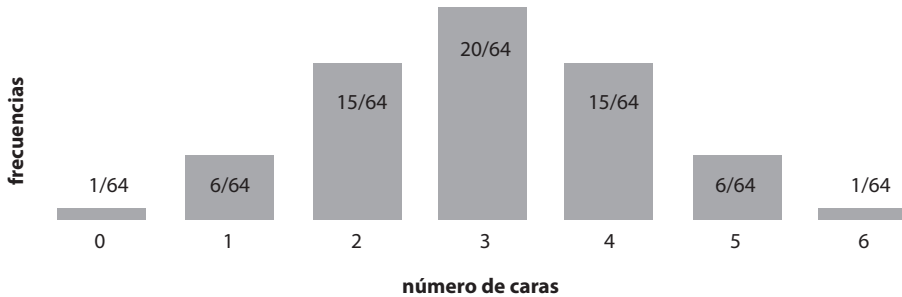


Gráfico 4.3.1. Distribución binomial del número de caras obtenido en una serie teórica de lanzamientos de seis monedas

En el Gráfico 4.3.1 se presenta un histograma de esta distribución binomial, cuando  $p = \frac{1}{2}$  (probabilidad del evento “cara”) en la que las frecuencias corresponden a los coeficientes, y el eje de la  $X$  al número de caras.

Al observar el gráfico se puede encontrar la probabilidad de cualquiera de los siete eventos posibles, al lanzar seis monedas ( $n = 6$  y  $p = \frac{1}{2}$ ). Así, la probabilidad de obtener tres caras es  $20/64 = 0.312$ . La de obtener seis caras es  $1/64 = 0.015$ .

A pesar de la distancia vertical entre las cúspides de las barras, se puede notar cierta semejanza con la distribución normal (Sección 3.6). Si en vez de  $n = 6$  tendríamos un valor más alto de  $n$ , habría más barras, la distancia entre las cúspides de las barras sería menor y consecuentemente el histograma resultante se asemejaría más y más a la distribución normal. Cuando  $n = 20$ , la distribución binomial es prácticamente idéntica a la distribución normal y consecuentemente pueden aplicarse las propiedades de la distribución normal (Sección 3.7).

La media y la varianza de una población con una distribución binomial pueden calcularse con el método empleado en el ejercicio 3.6.2, en el que se toma en cuenta las frecuencias. Pero primero se requiere estructurar un cuadro como el siguiente:

**Cuadro 4.3.1. Clases, frecuencias y cálculos complementarios para encontrar la media y varianza de una población binomial, cuando  $p = q = 1/2$  y  $n = 6$**

$Y_i$	$f_i$	$f_i Y_i$	$Y_i^2$	$f_i Y_i^2$
0	0.015625	0.00000	0	0.00000
1	0.093750	0.09375	1	0.09375
2	0.234375	0.46875	4	0.93750
3	0.312500	0.93750	9	2.81250
4	0.234375	0.93750	16	3.75000
5	0.093750	0.46875	25	2.34375
6	0.015625	0.09375	36	0.56250
$\Sigma Y_i = 21$	$\Sigma f_i = 1$	$\Sigma f_i Y_i = 3$	$\Sigma Y_i^2 = 91$	$\Sigma f_i Y_i^2 = 10.5$

En el Cuadro 4.3.1, el valor de  $Y_i$  corresponde al número de caras; las frecuencias resultan de la expansión del binomio  $(p + q)^6$  y se presentan sobre la base de la unidad (como ejemplo:  $1/64 = 0.015625$ ).

Con la información del Cuadro 4.3.1 se procede a encontrar la media y la varianza:

$$\mu = \frac{\Sigma f_i Y_i}{\Sigma f_i} = 3 \quad \sigma^2 = \frac{\Sigma f_i Y_i^2 - \frac{(\Sigma f_i Y_i)^2}{\Sigma f_i}}{\Sigma f_i} = \frac{10.5 - \frac{9}{1}}{1} = 1.5$$

Sin embargo, por las características propias de la distribución binomial, estos parámetros se encuentran más rápidamente con estas fórmulas:

$$\mu = n * p \quad (19)$$

$$\sigma^2 = n * p * q \quad (20)$$

En el ejemplo anterior estos parámetros son: la media  $\mu = 6 * \frac{1}{2} = 3$ , y la varianza,

$$\sigma^2 = 6 * \frac{1}{2} * \frac{1}{2} = 1.5$$

La media que hemos encontrado es la media del número de “caras”. Para encontrar la media del número de “sellos” la fórmula es:  $\mu = n * q$  que en este caso también es 3. Sin embargo, si las probabilidades de  $p$  y de  $q$  no fueran iguales, habría dos medias diferentes, una media para  $p$  y otra para  $q$ .

Cuando se trata de porcentajes, donde  $p + q = 100$ , las fórmulas para encontrar los parámetros de una población distribuida en forma binomial son:

Media:  $u = p$ , Varianza:  $\sigma^2 = (p*q)/n$  y Desviación Estándar:  $\sigma = \sqrt{(p*q)/n}$

#### 4.4. La distribución Z y el cálculo de probabilidades

Recuerde que la distribución Z es una distribución normal estandarizada, con  $\mu = 0$ ,  $\sigma^2 = 1$  y  $\sigma = 1$ . La fórmula, como se vio anteriormente es:  $Z = \frac{(Y_i - \mu)}{\sigma}$

Con el uso de la Tabla Z (Apéndice A) podemos calcular probabilidades acerca de la distribución normal. La primera columna y la primera fila de la Tabla Z son los valores en desviaciones estándar a partir de la media que es cero. Los valores en el cuerpo de la tabla corresponden a las áreas bajo la curva normal. Por lo general, estos valores se los multiplica por cien para expresarlos en porcentaje. Como ejemplo, para un valor de  $Z = 1.00$ , es decir una desviación estándar a partir de la media, el área correspondiente es 15.87%.

---

**Ejercicio 4.4.1.** Encuentre las siguientes probabilidades para la distribución Z:

a.  $P(Z \geq 0.58)$

Lo anterior equivale a preguntar: ¿cuál es la probabilidad de valores de Z iguales o mayores que 0.58? Para esto acudimos a la Tabla Z (Apéndice A). La columna de la extrema derecha equivale a las décimas, mientras que la hilera superior equivale a las centésimas. El área bajo la distribución Z es en la práctica igual a 1 o 100% (el cien por ciento de las observaciones está bajo la curva).

La tabla presenta únicamente valores de Z a partir de la media cero, hacia cualquiera de los dos extremos. Como la curva es simétrica, los valores de Z a la izquierda de la media son negativos y a la derecha son positivos. En este caso, por lo que se pide, el valor de Z es positivo y es igual a 0.58. Desde ese punto hacia el extremo derecho el área de la curva es 0.2810 y esa es la probabilidad (0.2810, o 28.10%). La respuesta obviamente hubiera sido la misma si se pedía la probabilidad de valores de Z iguales o menores que -0.58, aunque en este caso se refiere al extremo izquierdo de la curva.

b.  $P(Z \leq -1.35)$

Ubicamos el valor de Z igual a 1.35 en la Tabla Z. La probabilidad es igual a 0.0885, u 8.85%.

c.  $P(Z \leq 1.43)$

Ubicamos el valor de  $Z$  igual a 1.43 en la tabla. La probabilidad es igual a 0.0764. El problema pide la probabilidad de valores iguales o menores que 1.43. Por lo tanto, esta probabilidad está a la izquierda de 0.0764. Como el área total que cubre la distribución  $Z$  es igual a 1, la probabilidad es  $1 - 0.0764 = 0.9236$ , o 92.36%.

d.  $P(-1.26 \leq Z \leq 1.74)$

Aquí se pide la probabilidad de valores de  $Z$  comprendidos entre  $-1.26$  y  $1.74$ . Como son necesarios dos valores de  $Z$  se designan como  $Z_1$  y  $Z_2$ . Para  $Z_1 = -1.26$ , la probabilidad es 0.1038. Esto es, el área de la curva o distribución normal a la izquierda de  $-1.26$ . Para  $Z_2 = 1.74$ , la probabilidad es 0.0409. Entonces,

$$P(-1.26 \leq Z \leq 1.74) = 1 - 0.1038 - 0.0409 = 0.8553, \text{ o } 85.53\%.$$

**Ejercicio 4.4.2.** Para los siguientes problemas suponga que se trata de una población de pesos, en centigramos, de espigas de triticale (un híbrido entre trigo y centeno)  $Y_i \sim DN(20, 100)$ . Esto es, una población en la que las  $Y_i$  están distribuidas normalmente con  $\mu = 20 \text{ cg}$  y  $\sigma^2 = 100 \text{ cg}^2$

a.  $P(Y_i \geq 18)$

Esto equivale a preguntar: ¿qué porcentaje de las observaciones incluye valores iguales o mayores que 18? En otras palabras, ¿cuál es la probabilidad en esa población de encontrar espigas de triticale que tengan pesos iguales o mayores que 18 centigramos? Para contestar esto se debe transformar la población anterior a la distribución  $Z$ :

$$Z = \frac{Y_i - \mu}{\sigma} = \frac{18 - 20}{10} = -0.2. \text{ Entonces el problema se reduce a } P(Z \geq -0.2)$$

Por ser un valor negativo debemos restar de uno. La respuesta es  $1 - 0.4207 = 0.5793$ , o 57.93%.

b.  $P(15 \leq Y_i \leq 20)$

La pregunta ahora es: ¿cuál es la probabilidad de obtener al azar espigas con pesos comprendidos entre 15 y 20 centigramos? Para esto se requiere encontrar los valores de  $Z_1$  y  $Z_2$  equivalentes a los valores de 15 y 20:

$$Z_1 = \frac{15 - 20}{10} = -0.5 \quad Z_2 = \frac{20 - 20}{10} = 0$$

Consecuentemente  $P(15 \leq Y_i \leq 20)$  se reduce a  $P(-0.5 \leq Z \leq 0)$

Para  $Z_1 = -0.5$ , la probabilidad es 0.3085

Para  $Z_2 = 0.0$ , la probabilidad es 0.5000

Entonces, la respuesta es:  $1 - 0.3085 - 0.5000 = 0.1915$  ó 19.15%. En otras palabras, el 19.15% de la población de espigas de triticale tiene pesos comprendidos entre 15 y 20 centigramos.

- c. ¿Entre qué pesos de espiga está comprendido el 50% de las observaciones que se encuentra en el centro de la distribución?

Como ahora tenemos 25% en cada cola de la distribución, buscamos en el cuerpo de la tabla esa cantidad. Lo más cercano es 0.2514, es decir 25.14%. A este porcentaje le corresponden valores de  $Z$  de  $\pm 0.67$  desviaciones estándar.

Tenemos, entonces, dos valores de  $Z$  y conocemos la media ( $\mu = 20$ ) y la desviación estándar

( $\sigma = 10$ ) de la distribución normal de espigas de triticale. El problema se reduce ahora a encontrar los valores de  $Y_1$  y  $Y_2$  cuando los valores de  $Z$  son  $\pm 0.67$ . La solución se obtiene al despejar las siguientes ecuaciones:

$$Z_1 = \frac{Y_1 - \mu}{\sigma} = -0.67; Y_1 - \mu = -0.67\sigma; Y_1 - 20 = -0.67 * 10; Y_1 = 13.3 \text{ cg}$$

$$Z_2 = \frac{Y_2 - \mu}{\sigma} = 0.67; Y_2 - \mu = 0.67\sigma; Y_2 - 20 = 0.67 * 10; Y_2 = 26.7 \text{ cg}$$

La respuesta es entonces que el 50% de las observaciones del centro de la distribución se encuentra entre  $Y_1 = 13.3 \text{ cg}$  y  $Y_2 = 26.7 \text{ cg}$

**Ejercicio 4.4.3.** En una población normalmente distribuida se encuentra que el 33.0% de las plantas de café robusta tienen alturas inferiores a 140 cm, mientras que el 6.3% de las plantas tienen alturas superiores a 173 cm. Encuentre la desviación estándar y la media de esta población.

**Solución:** En el cuerpo de la Tabla  $Z$  (Apéndice A) se ubican los porcentajes (33.0% y 6.3%) que corresponden a los valores  $Z_1 = -0.44$  y  $Z_2 = +1.53$ , respectivamente. Luego, a partir de la fórmula  $Z = \frac{Y_i - \mu}{\sigma}$ , al reemplazar los valores de  $Z$  tenemos dos ecuaciones:

$$-0.44\sigma = 140 - \mu$$

$$+1.53\sigma = 173 - \mu$$



Al multiplicar una de las ecuaciones por -1 y al sumar las dos, se elimina  $\mu$  y se obtiene la ecuación  $1.97\sigma = 33$ . Entonces, la desviación estándar es:

$\sigma = \frac{33}{1.97=16.75} \text{ cm}$ . Este valor se reemplaza en cualquiera de las ecuaciones anteriores:

$$+1.53*16.75 = 173-\mu. \text{ Al despejar, se tiene } \mu = 147.4 \text{ cm}.$$

## El teorema del límite central

*“La persona no científica tiene una noción más clara de la física, química y biología que de la estadística. Muchos creen que los estadísticos son filatelistas numéricos, simples coleccionistas de números”.* Stephen Senn<sup>11</sup>

### 5.1. ¿Qué es el Teorema del Límite Central?

El *Teorema del Límite Central* establece la relación entre las muestras y la población de la que son extraídas. Esto tiene relación con el origen del error estándar, como se verá después.

Se ha señalado que  $\bar{Y}$  es un estimador de  $\mu$ , así como  $S^2$  lo es de  $\sigma^2$ . Esta relación entre las muestras y la población de la que provienen no es casual y más bien se fundamenta en el cálculo de probabilidades. Se pueden predecir los resultados de un muestreo en forma similar como se predicen los resultados al lanzar varias monedas al aire.

Cuadro 5.1.1. Todas las muestras posibles  $n = 2$ , extraídas de la Población 3, 5, 7, 9 y sus respectivas medias

Muestras, $n = 2$	$\sum Y_i$	$\sum Y_i/n$	$\bar{Y}_i$ (medias)
3 y 3	6	6/2	3
3 y 5	8	8/2	4
3 y 7	10	10/2	5
3 y 9	12	12/2	6
5 y 3	8	8/2	4
5 y 5	10	10/2	5
5 y 7	12	12/2	6

11 Bioestadístico suizo-británico. <http://www.vicbiostat.org.au/people/prof-stephen-senn>

Cuadro 5.1.1. Continuación

Muestras, n = 2	$\sum Y_i$	$\sum Y_i/n$	$\bar{Y}_i$ (medias)
5 y 9	14	14/2	7
7 y 3	10	10/2	5
7 y 5	12	12/2	6
7 y 7	14	14/2	7
7 y 9	16	16/2	8
9 y 3	12	12/2	6
9 y 5	14	14/2	7
9 y 7	16	16/2	8
9 y 9	18	18/2	9

Tomemos como ejemplo una población compuesta de las siguientes cuatro observaciones: 3, 5, 7, 9. Esta es una población intencionalmente pequeña por razón del ejemplo. La media, varianza y desviación estándar de esta población de cuatro observaciones son, respectivamente:  $\mu = 6$   $\sigma^2 = 5$   $\sigma = 2.24$

El valor de N, número de datos u observaciones en la población del ejemplo, es igual a 4. El número de muestras posibles, diferentes, con reemplazo, que se puede extraer de una población con N datos es igual a  $N^n$ , en donde n es el número de datos de la muestra. El caso es igual para todas y cada una de las muestras. Es necesario tener en cuenta que las muestras son tomadas al azar y que el muestreo es con reemplazo. Esto es, una vez que se extrae y se anota una observación, ésta vuelve a la población. De esta forma se garantiza que todas y cada una de las observaciones tengan igual probabilidad de ser incluidas en cualquier muestra.

En este ejemplo, si  $n = 1$ , todas las muestras posibles, sin repetir, serán cuatro:  $N^n = 4^1 = 4$ . Estas cuatro muestras son: 3, 5, 7, 9. Las medias de las cuatro muestras  $\bar{Y}$  serán también, obviamente, 3, 5, 7, 9. ¿Qué pasa si se aumenta el tamaño de la muestra? Tomemos, por facilidad, el caso de  $n = 2$ . Con este tamaño de muestra, el número de muestras posibles, diferentes, que pueden extraerse de la población 3, 5, 7, 9 será  $N^n = 4^2 = 16$ . Las 16 muestras y sus 16 medias se presentan en el Cuadro 5.1.1.

Se puede considerar a las 16 medias del Cuadro 5.1.1 como una nueva población, llamada “población de medias”. Las medias pueden agruparse, de acuerdo con su frecuencia, como se observa en el Cuadro 5.1.2 y formar un histograma, como se observa en el Gráfico 5.2.2.

Cuadro 5.1.2. Medias y frecuencias de la población de medias del cuadro anterior

Frecuencias ( $f_i$ )	1	2	3	4	3	2	1
Medias ( $\bar{Y}$ )	3	4	5	6	7	8	9

## 5.2. Los componentes del Teorema del Límite Central

El Teorema del Límite Central se compone básicamente de tres partes:

a) **La primera parte** señala que: “*A medida que el tamaño de la muestra aumenta la distribución de las medias de todas las muestras posibles del mismo tamaño, extraídas de la misma población original, se asemeja más y más a la **distribución normal***”. Esto se puede ver a continuación claramente, al comparar el histograma de las medias de la población 3, 5, 7, 9, cuando  $n = 1$ , con el histograma cuando  $n = 2$ . (Ver también el Gráfico 5.2.3.)

Si  $n = 3$ , entonces habrá  $4^3 = 64$  muestras posibles. El histograma con las frecuencias de estas 64 medias tendrá más barras y su distribución se aproximará aún más a la distribución normal que la del gráfico anterior.

Lo indicado tiene singular importancia ya que el investigador generalmente desconoce las características de la población con la que trabaja y si es que sigue o no la distribución normal. Sin embargo, si se extrae una muestra suficientemente grande, alrededor de  $n=30$ , las suposiciones basadas en una distribución normal tienen validez.

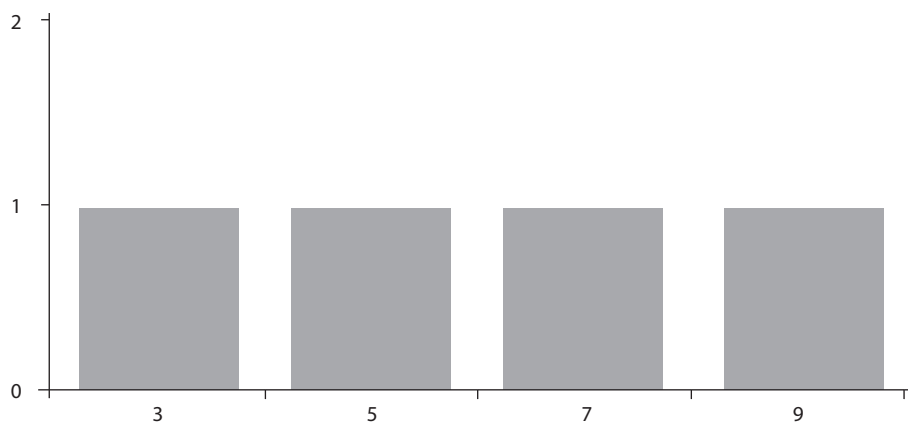


Gráfico 5.2.1. Histograma de las medias de las muestras obtenidas de la población 3, 5, 7, 9 cuando  $n=1$ .

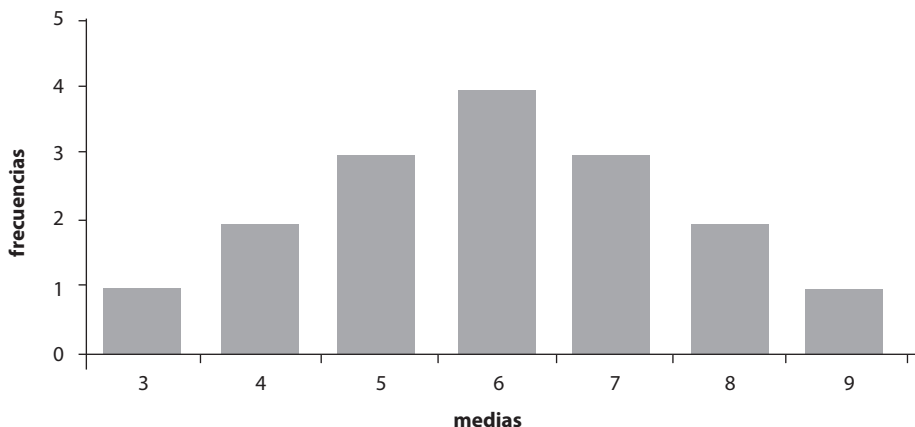


Gráfico 5.2.2. Histograma de las 16 medias de las muestras obtenidas de la población 3, 5, 7, 9 cuando  $n=2$

b) **La segunda parte** del Teorema del Límite Central dice: “Si la población inicial es normal, la distribución de las medias de las muestras seguirá exactamente la distribución normal independientemente del tamaño de la muestra”. En este caso, entonces, bastará con tomar una muestra pequeña, esto es de pocas observaciones, para tener como referencia las propiedades de la distribución normal. La relación entre diferentes poblaciones y la distribución de las medias de muestras de distintos tamaños se presenta en el Gráfico 5.2.3.

c) **La tercera parte** del teorema del límite central indica lo siguiente: “Si de una población  $Y$  con media  $\mu_Y$  y varianza  $\sigma_Y$  formada por  $N$  observaciones, se extraen todas las muestras posibles de tamaño  $n$ , se tendrá  $N^n$  muestras y  $N^n$  medias de muestras. Esta es una nueva población  $\bar{Y}$  formada por las medias de las muestras y que tiene sus propios parámetros, como media y varianza. La media  $\mu_{\bar{Y}}$  será igual a la media de la población  $Y$ , es decir  $\mu_Y$ .

La varianza de la nueva población formada por medias de las muestras,  $\sigma_{\bar{Y}}^2$  será igual a la varianza de la población  $Y$  dividida para  $n$ ”. Esta nueva población se denomina: “La Población de Medias”. Sus parámetros serán entonces:

$$\text{Media de la Población de Medias:} \quad \mu_{\bar{Y}} = \mu_Y \quad (21)$$

$$\text{Varianza de la Población de Medias:} \quad \sigma_{\bar{Y}}^2 = \frac{\sigma_Y^2}{n} \quad (22)$$

Desviación Estándar de la Población de Medias:  $\sigma_{\bar{Y}} = \frac{\sqrt{\sigma_Y^2}}{n}$  (23)

### ***Demostración***

Como demostración de lo señalado refirámonos a la población descrita anteriormente formada por las observaciones 3, 5, 7, 9. Los parámetros de esta población Y fueron:

Media  $\mu_Y = 6$ , Varianza  $\sigma_Y^2 = 5$ , Desviación Estándar  $\sigma_Y = \sqrt{5} = 2.24$

Cuando se extrajeron todas las muestras posibles  $n=2$  de esa población, las medias fueron (Cuadro 5.1.1): 3, 4, 4, 5, 5, 5, 6, 6, 6, 6, 7, 7, 7, 8, 8, 9

Tratemos a estas medias como lo que en verdad son: observaciones de una nueva población. Por ser esta una población formada por 16 medias, llamémosla “población de medias”. Encontramos ahora algunas características de esta población:

- La sumatoria de las observaciones. Como las observaciones son las medias, la fórmula es:

$$\sum \bar{Y}_i = 3 + 4 + 4 + 5 + \dots + 9 = 96$$

- La sumatoria de los cuadrados de las observaciones:

$$\sum \bar{Y}_i^2 = 3^2 + 4^2 + 4^2 + 5^2 + \dots + 9^2 = 616$$

- La sumatoria de las observaciones al cuadrado:

$$(\sum \bar{Y}_i)^2 = 96^2 = 9216$$

- La media de la población de medias (población  $\bar{Y}$ ) es igual a:

$$\mu_{\bar{Y}} = \frac{\sum \bar{Y}_i}{N} = \frac{96}{16} = 6$$

***Con lo que se comprueba que:***  $\mu_{\bar{Y}} = \mu_Y = 6$

- La varianza de la población de medias (población  $\bar{Y}$ ) es igual a:

$$\sigma_{\bar{Y}}^2 = \frac{\sum \bar{Y}_i^2 - \frac{(\sum \bar{Y}_i)^2}{N}}{N} = \frac{616 - \frac{9216}{16}}{16} = 2.5$$

**Con lo que se comprueba que:**  $\sigma_{\bar{Y}}^2 = \sigma_Y^2/n = 5/2 = 2.5$

- La desviación estándar de la población de medias (población  $\bar{Y}$ ) es igual a:

$$\sigma_{\bar{Y}} = \sqrt{\sigma_Y^2/n} = \sqrt{5/2} = 1.58$$

Lo anterior demuestra las relaciones entre población y muestras, de acuerdo con lo que postula el Teorema del Límite Central.

(En el Gráfico 5.2.3 se observa como las distribuciones tienden a la curva normal a medida que aumenta el tamaño de la muestra.)

### 5.3. El error estándar y el Teorema del Límite Central

Conocemos, gracias al Teorema del Límite Central, la relación entre una población y las muestras que pueden extraerse de ella. Sin embargo, está claro que en la práctica no podemos extraer todas las muestras posibles de una población. Por lo general nos contentamos con obtener unas pocas muestras y muchas veces solo una muestra. Pero en todo caso, siempre debemos tener como referencia la población teórica formada por todas las medias de las muestras (Población de las Medias de las Muestras). De hecho, la media de la muestra que hayamos obtenido pertenece a esa población. Como se vio en la sección anterior, un parámetro muy importante de esa población es la *Desviación Estándar de la Población de Medias* ( $\sigma_{\bar{Y}}$ ). En los trabajos de investigación necesitamos contar con un estimador de ese parámetro. Puede deducirse que la estadística que resulta ser el mejor estimador de  $\sigma_{\bar{Y}}$  es  $S_{\bar{Y}}$ , que se lo conoce como *Desviación Estándar de la Media de la muestra* o simplemente *Error Estándar*. La fórmula del *Error Estándar* es igual a  $S_{\bar{Y}} = \sqrt{S^2/n}$ . Note la similitud con la fórmula de la *Desviación Estándar de la Población de Medias*:  $\sigma_{\bar{Y}} = \sqrt{\sigma_Y^2/n}$ .

Para tener mayor confianza en los resultados experimentales es importante que el error estándar sea lo más bajo posible. Hay dos formas cómo el investigador puede reducir el valor de  $S_{\bar{Y}}$ :

1. Refinando la técnica experimental. Esto es, controlando las fuentes de variación, lo que lleva a una reducción de la varianza  $S^2$  y consecuentemente del error estándar  $S_{\bar{Y}}$ .
2. Aumentando el tamaño de la muestra,  $n$ . En otras palabras, como la fórmula de  $S_{\bar{Y}}$  es igual a  $\sqrt{\frac{S^2}{n}} = \frac{s}{\sqrt{n}}$ , entonces el valor de  $S_{\bar{Y}}$  va a ser menor si el denominador se incrementa.

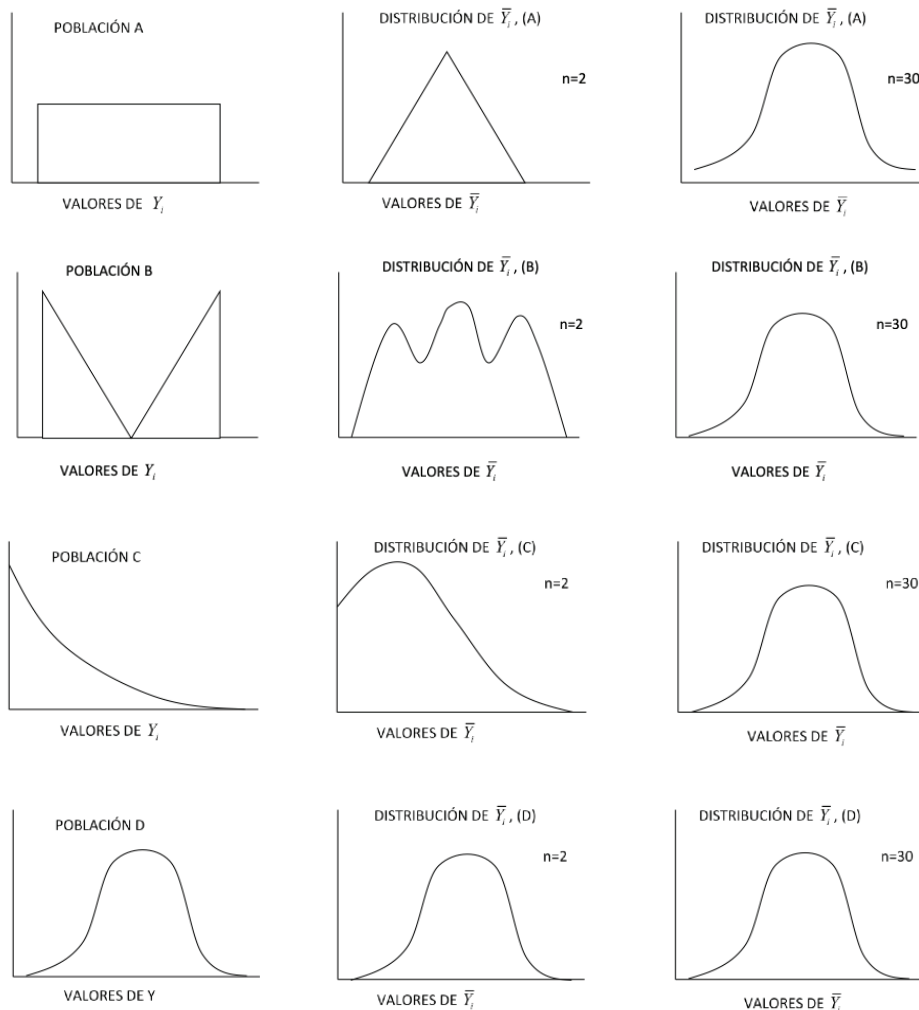


Gráfico 5.2.3. relación entre poblaciones diferentes y la distribución de medias de muestras de distintos tamaños

**Ejercicio 5.3.1** Al tomar una muestra al azar de tamaño 16 de una población normalmente distribuida, el valor de  $S_{\bar{Y}}$  fue igual a 5.

a) Encuentre la varianza de la muestra:

$$S_{\bar{Y}} = \frac{S}{\sqrt{n}}; 5 = \frac{S}{\sqrt{16}}; 25 = \frac{S^2}{16}; S^2 = 400$$



b) ¿Cuál debería ser el tamaño de una nueva muestra para reducir  $S_{\bar{y}}$  de 5 a 2.5, es decir, a la mitad?

Para encontrar el tamaño (n) de la nueva muestra, si queremos que  $S_{\bar{y}}$  sea igual a 2.5, se debe sustituir en la fórmula los valores conocidos:

$$S_{\bar{y}} = 2.5 = \sqrt{S^2/n}; 2.5^2 = 400/n; n = 400/2.5^2 = 64$$

c) Si los recursos solo alcanzaran para reducir de 5 a 4 el error estándar, ¿con qué tamaño de muestra se debería trabajar?

$$S_{\bar{y}} = 4 = \sqrt{S^2/n}; 4^2 = 400/n; n = 400/4^2 = 25$$

Note en el literal b) que al aumentar el tamaño de la muestra de 16 a 64, el error estándar sólo disminuyó a la mitad.

**Comentario:** Se presupone, con relación al ejercicio anterior, que la varianza de la nueva muestra no va a diferir significativamente del valor de 400, a no ser que haya cambios en la técnica de muestreo.

**Ejercicio 5.3.2.** Si se obtuvieran todas las muestras posibles  $n=20$  de una población en la que  $\mu = 100$  y  $\sigma^2 = 1280$ , ¿Cuáles serían los parámetros de la población de medias:  $\mu_{\bar{y}}$ ,  $\sigma_{\bar{y}}^2$ ,  $\sigma_{\bar{y}}$  ?

$$\mu_{\bar{y}} = \mu = 100$$

$$\sigma_{\bar{y}}^2 = \sigma^2/n = 1280/20 = 64$$

$$\sigma_{\bar{y}} = \sqrt{\sigma^2/n} = \sqrt{1280/20} = 8$$

**Ejercicio 5.3.3.** Se tiene una población normalmente distribuida, con media 40 y varianza 100

$$(\mu = 40, \sigma^2 = 100)$$

- a. Si se extrae una muestra  $n=4$ , ¿alrededor de qué valor probablemente estará el error estándar  $S_{\bar{y}}$ ?

Con los datos proporcionados no podemos encontrar  $S_{\bar{y}}$ . Sin embargo, sí podemos encontrar el valor del parámetro  $\sigma_{\bar{y}}$ , cuyo estimador es  $S_{\bar{y}}$ . Por lo tanto:

$\sigma_{\bar{y}} = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{100}{4}} = 5$ . Consecuentemente podemos inferir que el error estándar estará alrededor de 5.

- b. Con un programa de computación extraiga seis muestras  $n=4$  de una población con media igual a 40 y varianza igual a 100:  $\mu = 40$ ,  $\sigma^2 = 100$

Statistix 9

M1	M2	M3	M4	M5	M6
49.59	36.37	56.07	46.39	26.69	37.12
42.40	47.04	45.37	58.48	35.52	57.10
25.87	40.48	45.39	61.21	35.62	26.87
31.99	44.55	57.02	45.51	41.00	50.27

- c. Encuentre las medias, varianzas y errores estándar de las seis muestras

Statistix 9

Muestras	n	medias	Varianzas	Errores Estándar
M1	4	37.463	111.97	5.2908
M2	4	42.110	21.989	2.3446
M3	4	50.964	41.716	3.2294
M4	4	52.896	65.680	4.0521
M5	4	34.709	35.137	2.9638
M6	4	42.838	182.06	6.7464

Las medias, varianzas y errores estándar son estadísticas provenientes de muestras. Consecuentemente son variables. Estos valores son estimadores de los parámetros: media de la población ( $\mu = 40$ ), varianza de la población ( $\sigma^2 = 100$ ) y desviación estándar de la población de medias ( $\sigma_{\bar{y}} = 5$ ).

Note como los estimadores fluctúan alrededor de los parámetros. Sin embargo, las fluctuaciones son relativamente grandes por razón del pequeño tamaño de las muestras ( $n = 4$ ). El promedio de las seis medias es igual a 43.49, el promedio de las seis varianzas es igual a 76.43 y el promedio de los seis errores estándar es igual a 4.10.

**Comentario:** Las seis muestras al azar de tamaño cuatro del ejercicio anterior son parte de un total posible de  $N^n$  muestras. Si  $N$  fuera sólo 100, entonces  $N^n = 100^4 = 1 \cdot 10^8$ , es decir, habría 100 millones de muestras posibles. Por lo

tanto es muy improbable que haya dos muestras iguales o que al repetir el proceso se extraiga una muestra repetida.

Si quisiéramos reducir el error estándar a la mitad, es decir a un valor alrededor de 2.5, el tamaño de la muestra debería ser cuatro veces más grande, es decir  $n = 16$ . (Ver el ejercicio 5.3.1.) Para demostrar esto, saquemos de la misma población del ejercicio anterior, seis muestras de tamaño 16 (nada mágico con el número seis, pueden ser menos o más muestras).

**Ejercicio 5.3.4.** Se tiene una población normalmente distribuida, con media 40 y varianza 100

$$(\mu = 40, \sigma^2 = 100)$$

- a. Obtenga seis muestras de tamaño 16, ( $n = 16$ )

Statistix 9

<b>M1</b>	<b>M2</b>	<b>M3</b>	<b>M4</b>	<b>M5</b>	<b>M6</b>
50.96	31.3	40.59	29.8	47.73	27.2
32.73	38.09	34.32	35.54	32.26	42.97
41.75	39.99	43.05	34.35	38.34	32.75
29.18	44.95	40.82	37.39	63.52	41.03
43.01	56.58	39.72	31.67	11.39	44.42
36.09	36.91	53.54	36.88	49.11	47.06
27.23	50.29	52.14	32.04	45.38	43.98
34.56	45.52	22.07	36.79	42.93	30.22
34.18	52.11	43.52	53.45	56.95	31.33
40.69	40.3	34.84	30.71	38.58	42.66
25.36	35.19	35.64	46.8	32.63	37.27
45.94	45.39	36.66	39.97	28.64	51.43
38.18	58.32	45.3	57.79	36.79	41.28
47.61	52.54	31.37	33.21	44.85	24.04
39.78	58.09	59	41.65	38.26	40.72
28.29	51.38	40.49	42.21	51.33	43.13

b. Obtenga ahora las medias, la varianzas y errores estándar de las seis muestras

Statistix 9

Muestras	n	medias	Varianzas	Errores Estándar
M1	16	37.221	57.933	1.9028
M2	16	46.060	72.372	2.1268
M3	16	40.817	81.139	2.2519
M4	16	38.765	64.984	2.0153
M5	16	41.168	148.70	3.0486
M6	16	38.844	58.182	1.9069

Como puede verse las medias, las varianzas y los errores estándar fluctúan alrededor de la media de la población (40), la varianza de la población (100) y la nueva desviación estándar de la población de medias (2.5), respectivamente. Note también que la variación entre las muestras ha disminuido sensiblemente al incrementar el tamaño de la muestra de 4 a 16. El promedio de las seis medias es 40.47, el promedio de las seis varianzas es 80.55 y el promedio de los seis errores estándar es 2.21. Estos valores están ahora más cercanos a los parámetros de la población de medias:  $\mu_{\bar{y}} = 40.0$  y  $\sigma_{\bar{y}} = 2.50$ .

**Comentario:** A medida que aumenta el tamaño de la muestra, el error por muestreo disminuye. Esto se puede deducir del modelo lineal aditivo, que señala que:  $\bar{Y} = \mu + \sum \varepsilon_i/n$ . Es decir que la media de la muestra se acerca cada vez más a la media de la población a medida que se incrementa el tamaño de la muestra. Esto, bajo la presunción de que el muestreo sea al azar, por lo que las desviaciones o errores de signo positivo tienden a cancelarse con los errores de signo negativo.

## 5.4. La desviación estándar de las medias y las probabilidades

El error estándar interviene en la mayoría de las pruebas estadísticas que conllevan a decisiones sobre la base de probabilidades. Sin embargo puede ser que en algunos casos el investigador conozca o tenga evidencia suficiente acerca del valor de la varianza de la población de su interés.

Si fuera así, en las pruebas de hipótesis con relación a la media de la población debería emplear la *desviación estándar de la población de medias*  $\sigma_{\bar{y}}$  por ser una constante y no el *error estándar*  $S_{\bar{y}}$ , que es una variable. Para estos casos se emplea la fórmula de la distribución normal estandarizada  $Z$  para la población de medias. La fórmula es:

$$Z_{\bar{Y}} = (\bar{Y}_i - \mu_{\bar{Y}}) / \sigma_{\bar{Y}} \quad (24)$$

Veamos varios ejemplos:

**Ejercicio 5.4.1** Encuentre las siguientes probabilidades relacionadas con la Distribución  $Z$ : (comparar con los ejercicios de la Sección 4.4.)

- a. Si se extrae al azar una observación de una población normalmente distribuida con media igual a 30 y varianza igual a 40,  $Y_i \sim DN(\mu = 30, \sigma^2 = 40)$  ¿Cuál es la probabilidad de que esa **observación**  $Y_i$  se encuentre entre 28 y 34, es decir:  $P(28 \leq Y_i \leq 34)$ ?

Para contestar esta pregunta nos referimos a la distribución  $Z$ , cuya fórmula es:  $Z = \frac{(Y_i - \mu_Y)}{\sigma}$

Necesitamos dos valores que en la distribución  $Z$  sean equivalentes a 28 y a 34. Al reemplazar en la fórmula tenemos lo siguiente:

$$Z_1 = \frac{(28-30)}{6.32} = -0.32 \quad \text{y} \quad Z_2 = \frac{(34-30)}{6.32} = 0.63$$

En la Tabla  $Z$ , apéndice A, a la izquierda de -0.32 encontramos el área 37.45% y a la derecha de 0.63 el área 26.43%. Sumamos estas dos áreas y tenemos 63.88%. Consecuentemente, la probabilidad de que la observación se encuentre entre 28 y 34 será:  $100\% - 63.88\% = 36.12\%$ .

- b. Si de la misma población del literal a) se extrae una muestra al azar  $n=10$ , ¿Cuál es la probabilidad de que **la media de esa muestra**  $\bar{Y}_i$  se encuentre entre 28 y 34, es decir:  $P(28 \leq \bar{Y}_i \leq 34)$ ?

En este caso, como se trata de la media de una muestra  $\bar{Y}_i$  y no de una observación  $Y_i$ , debemos emplear la fórmula (24):

$$Z_{\bar{Y}} = (\bar{Y}_i - \mu_{\bar{Y}}) / \sigma_{\bar{Y}}$$

Con los datos de este ejercicio encontramos la desviación estándar de la población de medias,  $\sigma_{\bar{Y}} = \sqrt{\frac{40}{10}} = 2$ . Igualmente, como en el literal a), necesitamos dos valores que en la distribución  $Z$  sean equivalentes a 28 y 34:

$$Z_1 = \frac{(28-30)}{2} = -1 \quad \text{y} \quad Z_2 = \frac{(34-30)}{2} = 2$$

En la Tabla  $Z$ , apéndice A, a la izquierda de -1 encontramos el área 15.87% y a la derecha de 2 el área 2.28%. Sumamos estas dos áreas y tenemos 18.15%.

Consecuentemente la probabilidad de que la media se encuentre entre 28 y 34 será:  $100\% - 18.15\% = 81.85\%$ .

Note que en a) la probabilidad fue mucho menor que en b), a pesar de que las dos poblaciones tienen igual media. Esto resulta por cuanto el valor de la desviación estándar de la población de medias es siempre menor que el de la desviación estándar de la población original (excepto cuando  $n=1$ ).

Si una población sigue una distribución normal, el teorema del límite central nos indica cuál es la distribución de las medias de todas las muestras posibles extraídas al azar de esa población.

Si se tiene, como ejemplo, una población en la que los parámetros son  $\mu = 50$  y  $\sigma^2 = 100$

esperamos que  $\mu_{\bar{Y}} = 50$ , que  $\sigma_{\bar{Y}}^2 = 100/10 = 10$ , y que  $\sigma_{\bar{Y}} = \sqrt{10} = 3.16$ , cuando  $n=10$ .

Empleando la fórmula  $Z_{\bar{Y}} = (\bar{Y}_i - \mu_{\bar{Y}}) / \sigma_{\bar{Y}}$  podemos calcular la probabilidad de obtener valores determinados de  $\bar{Y}$ .

Si una población sigue una distribución normal, el teorema del límite central nos indica cuál es la distribución de las medias de todas las muestras posibles extraídas al azar de esa población.

Si se tiene, como ejemplo, una población en la que los parámetros son  $\mu = 50$  y  $\sigma^2 = 100$ , esperamos que  $\mu_{\bar{Y}} = 50$ , que  $\sigma_{\bar{Y}}^2 = \frac{100}{10} = 10$  y  $\sigma_{\bar{Y}} = \sqrt{10}$ , cuando  $n=10$ .

Mediante el uso de la fórmula  $Z_{\bar{Y}} = (\bar{Y}_i - \mu_{\bar{Y}}) / \sigma_{\bar{Y}}$  podemos calcular las probabilidades de obtener valores determinados de  $\bar{Y}$ .

**Ejercicio 5.4.2.** Si  $Y_i \sim DN(\mu = 50, \sigma^2 = 100)$ , calcule las siguientes probabilidades cuando  $n=10$ :

a.  $P(\bar{Y}_i \geq 51)$      $Z_{\bar{Y}} = (51 - 50) / 3.16 = 0.32$ ; R= 37.45%

b.  $P(\bar{Y}_i \geq 55)$      $Z_{\bar{Y}} = (55 - 50) / 3.16 = 1.58$ ; R= 5.71%

c.  $P(\bar{Y}_i \geq 60)$      $Z_{\bar{Y}} = (60 - 50) / 3.16 = 3.16$ ; R= 0.08%

d.  $P(\bar{Y}_i \leq 50)$      $Z_{\bar{Y}} = (50 - 50) / 3.16 = 0.00$ ; R= 50.00%

$$e. P(\bar{Y}_i \leq 46) \quad Z_{\bar{y}} = (46 - 50)/3.16 = -1.27; R = 10.20\%$$

$$f. P(\bar{Y}_i \leq 42) \quad Z_{\bar{y}} = (42 - 50)/3.16 = -2.53; R = 0.57\%$$

**Comentario:** Note en el ejercicio anterior que la probabilidad tiende a cero a medida que el valor de la media de la muestra se aleja del valor de la media de la población. Como ejemplo, en el literal f), la probabilidad de obtener una muestra con una media igual o menor que 42 es sólo 0.57%. Si desconociéramos que la media de la población es en verdad 50, no estaríamos tan seguros en afirmar que esa muestra con media  $\bar{Y} = 42$  provino de una población con una media  $\mu = 50$ , porque la probabilidad de que ocurra ese evento sólo por azar es muy baja. Más bien estaríamos tentados de pensar que esa estadística  $\bar{Y} = 42$  vino de una población con una media menor que 50.

## Capítulo 6

# El planteamiento y las pruebas de hipótesis

---

*“Hay solo dos posibilidades: si los resultados confirman la hipótesis, has realizado una medición; si los resultados son contrarios a la hipótesis, has hecho un descubrimiento”. Enrico Fermi<sup>12</sup>*

### 6.1. ¿Qué es una hipótesis?

Hipótesis es una proposición acerca de la forma de acción y posible relación de dos o más variables. Toda investigación requiere de hipótesis. Por lo general no es una idea sacada del aire, sino que tiene que sustentarse sobre el conocimiento actualizado del campo de investigación. La hipótesis es parte esencial del método científico, como se vio en la sección 1.4. La hipótesis debe ser sometida a prueba para la verificación de su postulado, su modificación o su rechazo. En todo caso, cualquier hipótesis es provisional. A Albert Einstein<sup>13</sup> se le atribuye la frase “ninguna cantidad de investigación podría probar que lo que propongo es cierto, pero un solo experimento podría demostrar que estoy errado”. De acuerdo con el epistemólogo Karl Popper<sup>14</sup>, considerado como uno de los grandes filósofos de la ciencia del siglo 20, una hipótesis no puede ser “confirmada”, porque siempre existe la posibilidad de que algún experimento futuro demuestre su falsedad.

La formulación de hipótesis es un hecho cotidiano en nuestras vidas. Expresiones como: “es posible que mañana llueva”, “probablemente el dólar se revalúe frente al euro”, o “tal equipo de fútbol puede quedar campeón”, son frases que contienen un componente de probabilidad y que las escuchamos frecuentemente. Sin embargo, en estos ejemplos poco o nada podemos hacer para provocar resul-

---

12 Físico italiano. [https://www.brainyquote.com/authors/enrico\\_fermi](https://www.brainyquote.com/authors/enrico_fermi)

13 Wynn, C.M; Wiggins, A.W; and Harris, S. (1997). *The Five Biggest Ideas in Science*. John Wiley and Sons. p. 107.

14 [http://es.wikipedia.org/wiki/Karl\\_Popper](http://es.wikipedia.org/wiki/Karl_Popper)



tados que nos lleven a confirmar estas hipótesis. Es decir, no podemos plantear experimentos controlados sobre la base de estas ideas.

En la aplicación del método científico es esencial la formulación de hipótesis, el sometimiento de la hipótesis a prueba y la evaluación final de la hipótesis. No pueden existir investigaciones científicas y experimentos sin una hipótesis previa.

## 6.2. ¿Cómo se plantean las hipótesis?

Para demostrar cómo se plantean las hipótesis analicemos el siguiente caso:

Al lanzar 10 monedas al aire, ¿cuál es la probabilidad de obtener 10 caras? El resultado de cada moneda es un evento independiente de lo que pueda ocurrir con las otras monedas. Recordemos que la probabilidad de obtener una cara al lanzar una moneda es  $p = 1/2$ . Consecuentemente para 10 monedas, la probabilidad es  $p^{10}$ . Esto es:  $p^{10} = (1/2)^{10} = 1/1024 = 0.00098$ , una cantidad un poco menor que 0.1%.

El punto es el siguiente. Supongamos que un tahúr le apuesta que al lanzar 10 monedas al aire obtendrá 10 caras. Como usted conoce acerca de probabilidades acepta la apuesta y ¡pierde! ¿Cuál sería su conclusión, que el señor que le propuso la apuesta era un timador o que sencillamente tuvo mucha suerte? La probabilidad de obtener 10 caras al lanzar 10 monedas al aire es tan baja que usted posiblemente concluya que lo estafaron.

En forma similar, si se desconoce la verdadera  $\mu$  y la probabilidad de obtener una media determinada  $\bar{Y}$  resulta muy baja con relación a una  $\mu$  hipotética, se podría concluir que la  $\bar{Y}$  encontrada proviene de una población con una  $\mu$  diferente de la que se tenía en mente. Sin embargo, tanto en el ejemplo de la apuesta como en este último, la decisión podría no ser la correcta.

En el ejemplo de la apuesta, antes de concluir si le estafaron o no, usted, consciente o inconscientemente, se planteó dos posibilidades: la primera, que las monedas eran de curso legal y que probablemente no todas caerían cara y la segunda, que las monedas habrían sido manipuladas y tenderían preferentemente a caer cara. En otras palabras, 1) el tahúr tuvo mucha suerte o, 2) el tahúr hizo trampa. En términos generales, estas dos posibilidades, mutuamente excluyentes, se llaman *hipótesis*.

La primera hipótesis se denomina “*Hipótesis Nula*” y se representa como  $H_0$ . Por lo general la  $H_0$  denota el “status quo”, es decir se basa en la información

o conocimiento previo que tenemos de una población. Así, si en un experimento se trata de evaluar el rendimiento de una nueva variedad, la  $H_0$  sería que el rendimiento de la nueva variedad va a ser igual al de la variedad tradicional de la zona. La segunda hipótesis se llama “*Hipótesis Alternativa*” y se representa como  $H_1$ . Como el nombre lo indica la  $H_1$  ofrece alternativas distintas a lo que plantea la  $H_0$ . En todo caso, las dos hipótesis deben abarcar todos los eventos posibles que puedan resultar.

En todo trabajo de investigación es esencial plantear las hipótesis antes de realizar el experimento y obtener los resultados. En el ejemplo de la apuesta, la hipótesis nula  $H_0$  se refiere a la información conocida de que las monedas pueden caer cara o sello indistintamente, pero con igual probabilidad, es decir  $p = q = \frac{1}{2}$ . Consecuentemente el tahúr no intenta hacer trampa; si gana es por puro azar. La hipótesis alterna ( $H_1$ ) es la probabilidad de que esas monedas tengan mayor probabilidad de caer cara, o tal vez que tengan cara en los dos lados, en cuyo caso  $p=1$  y  $q = 0$ . Si esto último es el caso, el tahúr hace trampa. Como la probabilidad de obtener 10 caras es tan baja, una vez conocido el resultado, usted posiblemente se decida por la  $H_1$ . Sin embargo, en cualquiera de los dos casos, tanto si se acepta la  $H_0$  como si la rechaza se puede cometer error. Esto determina que haya errores de dos tipos, Tipo I y Tipo II, como se verá más adelante.

### 6.3. Procedimiento para las pruebas de hipótesis

El procedimiento para las pruebas de hipótesis es el siguiente:

1. Planteamiento de la hipótesis nula ( $H_0$ ) y alternativa ( $H_1$ )
2. Determinación del *nivel de significación*  $\alpha$  y del *valor crítico* *V.C.*
3. Análisis de los resultados experimentales y obtención de la *prueba estadística*, *P.E.*
4. Comparación entre el valor de la prueba estadística y el valor crítico
5. Aceptación o rechazo de la hipótesis nula,  $H_0$

### 6.4. Nivel de significación, valor crítico y prueba estadística

Se llama en estadística *nivel de significación*  $\alpha$ , a la cantidad de evidencia necesaria para decidir cuán probable es que un evento haya ocurrido solo por azar. Generalmente y en forma convencional se ha fijado como nivel de signifi-

cación una probabilidad del 5%. Esto quiere decir que el investigador acepta que sus resultados puedan deberse al azar en ese porcentaje. En otras palabras, está dispuesto a rechazar la *hipótesis nula* si  $P \leq 0.05$  (probabilidad igual o menor que 5%). A veces, para dar mayor énfasis a los resultados experimentales se usa también el nivel de significación del 1%.

En las diferentes curvas de distribución de datos cuantitativos los niveles de significación que fije el investigador están en los extremos. Estos porcentajes de las distribuciones se denominan *áreas críticas* y separan las áreas de rechazo de las áreas de aceptación de la *hipótesis nula*. A cada área crítica le corresponde un *valor crítico*, que se lo encuentra con la fórmula de la distribución específica y que toma en cuenta la densidad de las áreas de la curva y los *grados de libertad* de la muestra. El *área crítica* se expresa en porcentaje, mientras que *el valor crítico* es una cantidad relacionada con el área crítica y que se ubica en la abscisa de los ejes cartesianos. Al valor crítico se lo encuentra en tablas específicas, como las de los apéndices de este texto. El *valor crítico* es un punto de referencia para compararlo con la *prueba estadística*.

La *prueba estadística* es el valor calculado con los datos experimentales obtenidos de la muestra y la aplicación de la fórmula de la distribución correspondiente, de acuerdo con la planificación del experimento. Si la *prueba estadística* excede el *valor crítico*, entonces la decisión es rechazar la *hipótesis nula*.

Para facilitar la evaluación de hipótesis, varios matemáticos e investigadores desarrollaron tablas estadísticas donde constan los *niveles de significación  $\alpha$*  y los correspondientes *valores críticos*. Por simplicidad, las tablas (ver apéndices) sólo especifican determinados niveles de significación. Los más usados tradicionalmente han sido los niveles de significación del 5% y del 1%. Actualmente, con los programas computarizados, el investigador puede determinar exactamente la probabilidad de un resultado y sobre esta base, decidir si acepta o rechaza la hipótesis nula planteada. Sin embargo, las tablas estadísticas más relevantes se incluyen en este texto por razones didácticas.

En investigación, una vez que se ha aceptado o rechazado la  $H_0$  el paso siguiente es la discusión de los resultados. Este es el punto crucial de una investigación; es el paso que explica y justifica los resultados obtenidos y en el que el aporte personal en términos de ideas tiene gran importancia.

## 6.5. Hipótesis, decisiones y consecuencias

El cálculo de probabilidades nos permite probar las hipótesis que nos hemos planteado. Una prueba de hipótesis es una evaluación probabilística de la evidencia a favor o en contra de la hipótesis nula y esto nos conduce a su aceptación o rechazo.

En el siguiente cuadro se presentan las alternativas y consecuencias al decidirse por determinada hipótesis:

Cuadro 6.5.1. Hipótesis, decisiones y consecuencias

		Condición de las hipótesis	
		$H_0$ verdadera	$H_1$ verdadera
Decisión 1	Aceptar $H_0$	Decisión correcta	Error Tipo II
Decisión 2	Rechazar $H_0$	Error Tipo I	Decisión correcta

En otras palabras, si se rechaza una hipótesis nula verdadera se comete *Error Tipo I*. Si se acepta una hipótesis nula falsa, se comete *Error Tipo II*.

## 6.6. Error Tipo I y Error Tipo II

En el proceso de evaluación de las hipótesis siempre se corre el riesgo de llegar a una conclusión errónea. Aunque el límite de las probabilidades es  $0 \leftrightarrow 1$ , las curvas de distribuciones en Estadística en teoría van de menos infinito a más infinito ( $-\infty \leftrightarrow +\infty$ ). Consecuentemente, *nunca* podemos tener certeza absoluta de la veracidad de nuestra decisión al aceptar o rechazar una hipótesis. En otras palabras, siempre existe la probabilidad de cometer error.

Como se señala en el Cuadro 6.5.1, existen dos tipos de error:

- Error Tipo I.** Se llama también “error de primera clase”. Se comete Error Tipo I cuando se *rechaza* una hipótesis nula *verdadera*.
- Error Tipo II.** Se llama también “error de segunda clase”. Se comete Error Tipo II cuando se *acepta* una hipótesis nula *falsa*.

El área crítica corresponde al nivel de significación de la prueba. A esto se le designa con la letra griega alfa  $\alpha$ . El nivel de significación es el riesgo de cometer *Error Tipo I*. En otras palabras, la probabilidad de cometer *Error Tipo I* es igual al nivel de significación  $\alpha$ .

La probabilidad de cometer *Error Tipo II* se designa con la letra griega beta  $\beta$ . El valor de  $\beta$  depende de  $\alpha$  y de  $\mu_0$  (la media hipotética) y tiene relación con el *poder de la prueba*. El poder de la prueba es igual a  $1 - \beta$ . Si el nivel de significación  $\alpha$  aumenta también aumenta la probabilidad de cometer Error Tipo I, pero disminuye la probabilidad de cometer Error Tipo II. Sin embargo, como se vio anteriormente, si el tamaño de la muestra aumenta el error estándar disminuye. Consecuentemente la prueba se vuelve más precisa. En otras palabras, tanto las probabilidades de cometer Error Tipo I como Error Tipo II pueden mantenerse bajas al incrementar el tamaño de la muestra,  $n$ . Más adelante se presentarán ejemplos de la estimación del poder de la prueba y de las relaciones con  $\alpha$ ,  $\beta$  y el tamaño de la muestra ( $n$ ).

En las investigaciones agropecuarias, el llegar a decisiones en forma apresurada y aceptar o rechazar las hipótesis sin mayor sustento o evidencia (Error Tipo I, Error Tipo II), puede ser sumamente costoso. Supongamos que sobre la base de una muestra pequeña se llega a la conclusión de que un producto X controla la sigatoka (enfermedad del banano) y nos apresuramos a aplicar el producto en 100 o más hectáreas. Bien podríamos estar botando el dinero, o peor, causando más mal que bien.

## 6.7. Pruebas de una cola y pruebas de dos colas

En las pruebas de hipótesis hay tres casos posibles, que son definidos por el investigador de acuerdo con sus intereses. Cuando las pruebas tienen relación con la media de la población los casos son:

**Caso No. 1.** Ignoramos la posibilidad de que  $\mu$  pueda ser menor que  $\mu_0$  ( $\mu_0$  es el símbolo que representa a la media que se plantea en la hipótesis nula. Se llama media hipotética):

$$H_0: \mu = \mu_0$$

$$H_1: \mu > \mu_0$$

Esta es una prueba de cola derecha y tiene signo positivo.

**Caso No. 2.** Ignoramos la posibilidad de que  $\mu$  pueda ser mayor que  $\mu_0$ :

$$H_0: \mu = \mu_0$$

$$H_1: \mu < \mu_0$$

Esta es una prueba de cola izquierda y tiene signo negativo.

**Caso No. 3.** Se considera la posibilidad de que  $\mu$  pueda ser mayor o menor que  $\mu_0$ :

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$

Esta es una prueba de dos colas y consecuentemente puede tener signo positivo o negativo.

En las figuras siguientes se presentan los tres casos posibles. Si la prueba estadística cae **dentro** del área crítica, o lo que es lo mismo excede al valor crítico absoluto, la decisión es rechazar la  $H_0$ . (La forma más aceptada de expresar lo mismo es indicar que no existe suficiente evidencia para rechazar la  $H_0$ ). En el caso contrario, se acepta la  $H_0$  con el nivel de significación especificado.

### Caso No. 1

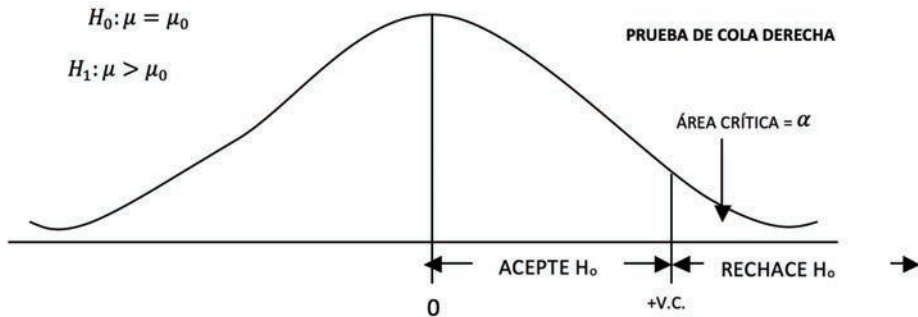


Gráfico 6.71. Prueba de cola derecha

## Caso No. 2

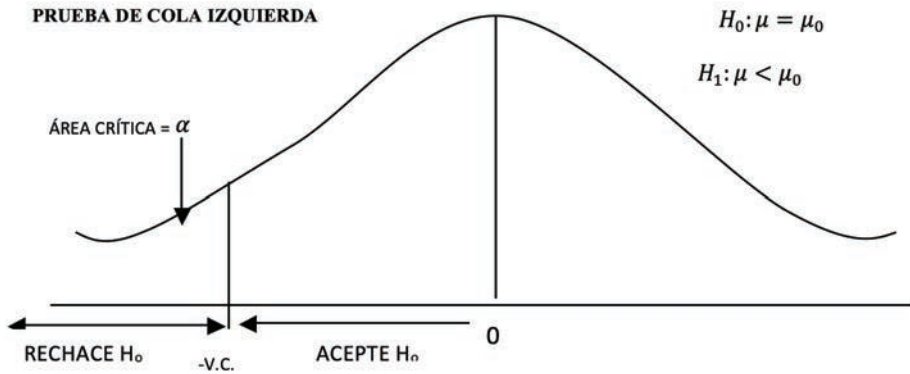


Gráfico 6.7.2. Prueba de cola izquierda

## Caso No. 3

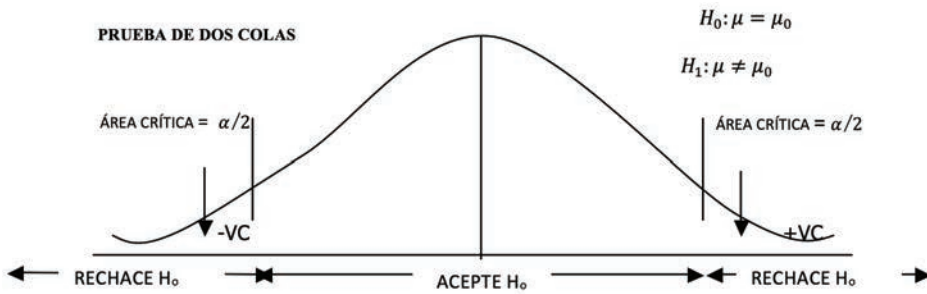


Gráfico 6.7.3. Prueba de dos colas

## Pruebas de hipótesis para las medias

*“La única prueba relevante de la validez de una hipótesis es comparar lo que se predijo con los resultados del experimento”. Milton Friedman<sup>15</sup>*

### 7.1 . Prueba Estadística Z

Si tenemos una población y conocemos su varianza  $\sigma^2$  pero desconocemos la media  $\mu$ , podemos probar la hipótesis nula que la media  $\mu$  es igual a una media hipotética  $\mu_0$  utilizando la distribución Z. La prueba estadística Z (*P.E.Z.*) es:

$$P.E.Z. = \frac{\bar{Y} - \mu_0}{\sigma_{\bar{Y}}} \quad (25)$$

Una vez planteadas las hipótesis y definidos el área crítica y el valor crítico, se procede a extraer una muestra al azar de la población, luego se encuentra la media de muestra  $\bar{Y}$  y la desviación estándar de la población de medias  $\sigma_{\bar{Y}}$ . A continuación se obtiene la prueba estadística *P.E.Z.* Luego evaluamos la prueba estadística comparándola con el valor crítico y decidimos si aceptar o rechazar la Hipótesis Nula. Los ejercicios a continuación son ejemplos de esto.

---

**Ejercicio 7.1.1.** Se tiene una población normalmente distribuida  $Y_i$  DN ( $\mu$ , 90). Es decir, una población con media desconocida y varianza igual a 90. Un investigador se plantea las siguientes hipótesis:

$$H_0: \mu = 20$$

$$H_1: \mu \neq 20$$

---

<sup>15</sup> Economista estadounidense. [https://www.brainyquote.com/authors/milton\\_friedman](https://www.brainyquote.com/authors/milton_friedman)



El nivel de significación que escoge es  $\alpha = 5\%$ .

Luego procede a la extracción de una muestra al azar y ésta es: 15, 14, 8, 10, 18, 20, 25, 10, 13, 12.

La media  $\bar{Y}$  es igual a:  $\bar{Y} = \frac{\sum Y_i}{n} = \frac{145}{10} = 14.5$

La desviación estándar de la población de medias es:  $\sigma_{\bar{Y}} = \sqrt{\frac{\sigma_y^2}{n}} = \sqrt{\frac{90}{10}} = 3$

La prueba estadística  $Z$  es igual a: P.E.  $Z = \frac{\bar{Y} - \mu_0}{\sigma_{\bar{Y}}} = \frac{14.5 - 20}{3} = -1.83$

La hipótesis alternativa plantea la posibilidad de que la media sea menor o mayor que 20, consecuentemente esta es una prueba de dos colas. Por lo tanto, la región o área crítica en cada cola es  $\alpha/2 = 2.5\%$ . En el cuerpo de la Tabla  $Z$  (Apéndice A) se busca 0.0250 (es decir, 2.5%). En la columna e hilera correspondientes se ubica el valor  $Z$ , que para este caso es 1.96.

Por ser prueba de dos colas los *valores críticos* son entonces -1.96 y +1.96. Como la *prueba estadística* -1.83 cae dentro de estos límites, en otras palabras, no excede a los valores críticos, la decisión consecuentemente es aceptar la  $H_0$  al 5% nivel de significación. Como en verdad nunca podemos estar seguros de que la verdadera media de la población sea igual a la media hipotética, una mejor manera de expresar lo anterior y que cubre la incertidumbre anotada, es señalar que “*no existe suficiente evidencia para rechazar la  $H_0$* ”.

Analicemos el significado de la conclusión a la que se llegó en el ejercicio anterior. La distribución  $Z$  representa la curva normal que seguirán las medias de todas las muestras posibles de tamaño 10, extraídas de la población  $Y_i$  DN ( $\mu$ , 90)

La probabilidad de extraer una muestra con una media como la obtenida,  $\bar{Y} = 14.5$ , es superior al 5%. Como ésta es una probabilidad no lo suficientemente baja, podemos concluir que no existe suficiente evidencia para rechazar la  $H_0$  que  $\mu = 20$ . Está claro que la decisión tomada sería correcta si el valor de  $\mu$  fuera en verdad 20.

Sin embargo, incluso si el valor de  $\mu$  fuera en verdad 20, es probable extraer al azar una muestra que tenga una media extrema y que la P.E. exceda al valor crítico absoluto de 1.96. Esta probabilidad es igual al 5%. Si este evento ocurriera, entonces se rechazaría una hipótesis nula verdadera y consecuentemente se cometería Error Tipo I.

En cualquier experimento podemos llegar a una conclusión equivocada y nunca conocer si se cometió Error Tipo I o II. Sólo la evidencia obtenida de otros experimentos similares al nuestro puede despejar estas dudas. En todo caso, en experimentos cuidadosamente conducidos, disminuye la probabilidad de arribar a una conclusión errada.

En el siguiente ejercicio se conoce que la población no está normalmente distribuida. Consecuentemente, el tamaño de la muestra debe ser alrededor de 30. Esto, de acuerdo con lo que postula el Teorema del Límite Central. Cuando  $n = 30$ , la distribución de la población de medias se convierte en una distribución normal y por lo tanto se pueden aplicar las propiedades de la distribución normal estandarizada  $Z$ .

---

**Ejercicio 7.1.2.** En un valle se ha sembrado fréjol en forma tradicional por muchos años. La media de la producción de fréjol en ese valle es  $\mu = 400$  kg/ha y la desviación estándar  $\sigma = 50$  kg/ha. (Es decir, varianza igual a 2500.) Se sabe además que la distribución de la producción no sigue la distribución normal. Un técnico agrícola trata de demostrar a los agricultores los beneficios del manejo integrado de plagas. Para esto realiza un experimento en 30 parcelas repartidas al azar en el valle y las cultiva de la misma forma que los agricultores, con la excepción del manejo de plagas. Las hipótesis que se plantea el técnico agrícola son:

$$H_0: \mu = 400$$

$$\alpha = 5\%$$

Note que la hipótesis nula es  $\mu = 400$ . Esto equivale a decir que el manejo integrado de plagas no afectará en forma alguna la producción de fréjol. Por lo contrario, la hipótesis alterna

$\mu > 400$  plantea la posibilidad de que la media sea mayor gracias a la aplicación del manejo de plagas. Además, descarta la posibilidad de que el manejo de plagas disminuya la producción. Por eso, se plantea la hipótesis de una sola cola. Una presunción importante es que la varianza de la producción se mantendrá igual al aplicar la tecnología que se pretende introducir.

Al analizar los resultados el técnico agrícola encuentra que la media de la muestra de 30 parcelas fue  $\bar{Y} = 450$ . Fue necesario tomar una muestra de tamaño 30 ya que la población original no presentaba una distribución normal. De esta forma la distribución de las  $\bar{Y}_i$  será normal, de acuerdo con el Teorema del Límite Central y podemos entonces aplicar las propiedades de esta distribución.

El valor  $\sigma_{\bar{y}}$  es igual a:  $\sigma_{\bar{y}} = \sqrt{\frac{2500}{30}} = 9.13$

La prueba estadística es igual a:  $P.E.Z = \frac{\bar{Y} - \mu_0}{\sigma_{\bar{y}}} = \frac{450 - 400}{9.13} = 5.48$

Note que, como se descarta la posibilidad de que el manejo integrado de plagas pueda disminuir la producción, la prueba estadística es una prueba de cola derecha.

La Tabla Z nos proporciona el *valor crítico*. El valor más próximo a  $\alpha = 5\%$  en el cuerpo de la tabla es 0.0495 (4.95%). Esto corresponde a un valor crítico igual a 1.65. Como la  $P.E. = 5.48$  excede claramente al valor crítico 1.65, la decisión es rechazar la  $H_0$  que  $\mu = 400$  al 5% nivel de significación. La prueba estadística incluso excede el valor crítico al nivel del 1% de significación, que es igual a 2.33 (ver Tabla Z). En otras palabras, el manejo integrado de plagas causó un aumento estadísticamente significativo en la producción de fréjol en las condiciones de ese experimento.

(Otra forma de indicar un valor significativo al 5% es colocando un asterisco al lado de la  $P.E.$  Si el valor es significativo al 1% se colocan dos asteriscos. La abreviación n.s. (no significativo) al lado de la prueba estadística indica que no hay suficiente evidencia para rechazar la  $H_0$ . En este caso tenemos  $P.E. = 5.48^{**}$ )

## 7.2. El Poder de la Prueba

Se define el poder de la prueba estadística como la probabilidad de rechazar la hipótesis nula cuando la hipótesis alternativa es verdadera. Como se mencionó anteriormente, el valor de alfa ( $\alpha$ ) es la probabilidad de cometer Error Tipo I. A partir del valor  $\alpha$  y de los valores de las hipótesis es posible estimar beta ( $\beta$ ), que es la probabilidad de cometer Error Tipo II. Una vez que se obtiene el valor de  $\beta$  se puede encontrar el *Poder de la Prueba* (PP). La fórmula del poder de la prueba es:

$$PP = 1 - \beta \quad (26)$$

Veamos un ejemplo:

Un investigador desea evaluar el rendimiento de una nueva variedad de arroz. Se conoce que el rendimiento promedio de la zona es 30 q/ha (q es el símbolo de quintal métrico y equivale a 100 kg). El investigador descarta la posibilidad de que la nueva variedad rinda menos que el promedio de la zona y más bien está predispuesto a rechazar la  $H_0$  si la media de la muestra es igual o mayor que 34 q/ha. Por lo tanto plantea las siguientes hipótesis:

$$H_0: \mu = 30$$

$$H_1: \mu > 30$$

También se conoce, a través de varios años de experiencia, que la población está normalmente distribuida y que la varianza es igual a  $100 \text{ q}^2/\text{ha}^2$ . **Antes de realizar el experimento** al investigador le interesa conocer cuál es la probabilidad de cometer Error Tipo I si toma una muestra de 10 observaciones ( $n=10$ ). Con la información que tiene, la condición  $n = 10$  y una supuesta media de la muestra igual a 34 ( $\bar{Y} = 34$ ), hace las siguientes estimaciones:

La desviación estándar de la población de medias será:

$$\sigma_{\bar{y}} = \sqrt{\sigma^2/n} = \sqrt{100/10} = 3.16$$

La prueba estadística Z tendrá el siguiente valor:

$$PEZ = (\bar{Y} - \mu_0) / \sigma_{\bar{y}} = (34 - 30) / 3.16 = 1.26$$

En la distribución Z, la media de la población de la hipótesis nula  $\mu = 30$  equivale a  $Z_0 = 0$ , mientras que el valor supuesto de la media de la muestra equivale a la prueba estadística, es decir  $Z_1 = 1.26$ . Este valor de  $Z_1$  corresponde a un área de 10.38 %. Esta probabilidad es demasiado alta para rechazar la hipótesis nula. El valor 10.38 % es el nivel alfa, la probabilidad de cometer Error Tipo I. Esto se resume en el siguiente gráfico:

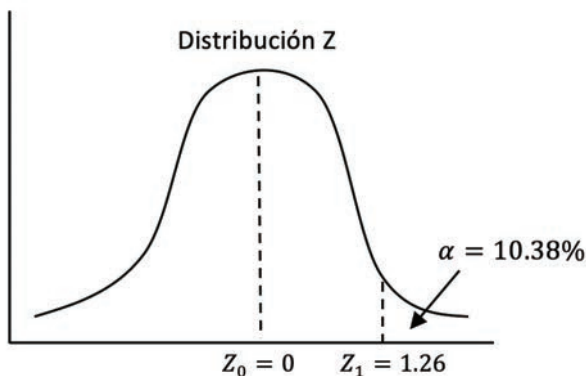


Gráfico 7.2.1. Ubicación de la media poblacional, media de la muestra y el nivel alfa

Sin embargo, parecería que una diferencia en el rendimiento de 4 q/ha, es decir 400 kg, sí es importante para el agricultor. El investigador podría pregun-

tarse: ¿qué tal si la media verdadera de la población de la nueva variedad de arroz no es 30 q/ha (que es el valor de la hipótesis nula), sino más bien 36 q/ha? Si fuera así, ¿cuál sería el valor de  $\beta$ , es decir, la probabilidad de cometer Error Tipo II? Este caso se presenta en el Gráfico 7.2.2. (Recuerde que todo esto es en la etapa de planificación del experimento.)

La prueba Z, con la misma información anterior, pero para el supuesto caso de que la media verdadera de la población sea 36 q/ha (la  $H_1$ ), es la siguiente:

$$PEZ = (\bar{Y} - \mu_1) / \sigma_{\bar{Y}} = (34 - 36) / 3.16 = -0.63$$

El valor de la prueba estadística  $Z = -0.63$  cae a la izquierda de la curva y equivale a un porcentaje de 26.43%. Este porcentaje es  $\beta$ , la probabilidad de Error Tipo II, es decir, la probabilidad de aceptar una hipótesis nula falsa (o rechazar una hipótesis alternativa verdadera).

Con esta información se puede calcular el Poder de la Prueba (PP) que, como se definió anteriormente, es la *probabilidad de rechazar la hipótesis nula cuando la hipótesis alternativa es verdadera*. El Poder de la Prueba es:

$PP = 1 - \beta = 1 - 0.2643 = 0.7357$ , que expresado en porcentaje es: 73.57%. En el Gráfico siguiente se presentan las curvas y las áreas respectivas:

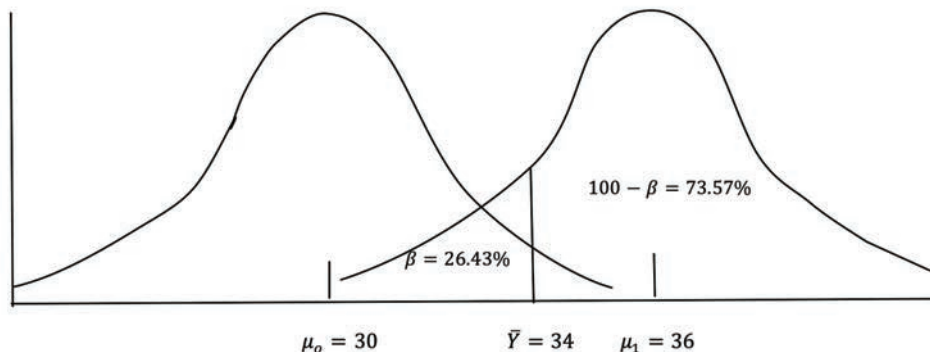


Gráfico 7.2.2. El Poder de la Prueba cuando la hipótesis alternativa es igual a 36 q/ha.

En resumen, en este ejemplo hemos calculado las siguientes probabilidades:

- Error Tipo I, 10.38 %
- Error Tipo II, 26.43 %
- Poder de la Prueba 73.57 %

Si la probabilidad del Poder de la Prueba no es suficientemente satisfactoria, el investigador siempre puede, en la planificación de su experimento, aumentar el tamaño de la muestra. Esto se demuestra en el siguiente ejercicio. (Siempre puede obtenerse un tamaño de muestra mayor, pero también deben considerarse los costos involucrados.)

---

**Ejercicio 7.2.1.** Con los mismos datos del ejemplo anterior, pero ahora con un tamaño de muestra  $n=18$ , encuentre las probabilidades de a) Error Tipo I, b) Error Tipo II y c) el Poder de la Prueba:

Las hipótesis son:

$$H_0: \mu = 30$$

$$H_1: \mu > 30$$

La varianza de la población es  $100 \text{ q}^2/\text{ha}^2$ , la supuesta media de la muestra es  $34 \text{ q/ha}$ , y el tamaño de la muestra  $n=18$ .

La desviación estándar de la población de medias será:

$$\sigma_{\bar{Y}} = \sqrt{\sigma^2/n} = \sqrt{100/18} = 2.36$$

### a) Error Tipo I

La prueba estadística  $Z$  tendrá el siguiente valor:

$$PEZ = (\bar{Y} - \mu_0) / \sigma_{\bar{Y}} = (34 - 30) / 2.36 = 1.70$$

El valor 1.70, de acuerdo con la Tabla  $Z$ , corresponde a un porcentaje de 4.46 %. Este es el valor  $\alpha$  encontrado, que es también igual a la probabilidad de cometer Error Tipo I.

### b) Error Tipo II

Ahora, bajo la suposición de una media poblacional verdadera de  $36 \text{ q/ha}$ , ¿cuál es el valor de  $\beta$ , la probabilidad de cometer Error Tipo II? Para esto debemos plantear una nueva prueba estadística  $Z$ :

$$PEZ = (\bar{Y} - \mu_1) / \sigma_{\bar{Y}} = (34 - 36) / 2.36 = -0.85$$

El valor  $-0.85$  en la distribución  $Z$  corresponde a un área, a la izquierda de la media de la población, equivalente a  $19.77\%$ . Este es el valor  $\beta$ , la probabilidad de cometer Error Tipo II.

### c) *El Poder de la Prueba*

El poder de la prueba, es decir, la probabilidad de rechazar una hipótesis nula falsa cuando se tiene una hipótesis alternativa verdadera es igual a:

$$PP = 1 - \beta = 1 - 0.1977 = 0.8023$$

En términos de porcentaje, la probabilidad del poder de la prueba es  $80.23\%$ .

**Comentario:** Observe como al cambiar el tamaño de la muestra de 10 a 18, disminuyó la probabilidad de cometer Error Tipo I, también disminuyó la probabilidad de cometer Error Tipo II y aumentó la probabilidad del Poder de la Prueba. Esto resalta la importancia que tiene el tamaño de la muestra en la investigación.

En el siguiente ejercicio se demuestra el procedimiento para encontrar el tamaño de la muestra cuando se desea un valor específico del poder de la prueba. Para esto es necesario fijar previamente los valores alfa y beta.

---

**Ejercicio 7.2.2.** La media de la población de contenido de proteína en 100 g de harina de trigo blanco de invierno es 8 g. Se conoce que la varianza de la población es  $4 \text{ g}^2$ . Un mejorador vegetal piensa que su nueva variedad tiene un contenido mayor. Las hipótesis que se plantea son:

$$H_0: \mu = 8$$

$$H_1: \mu > 8$$

El investigador espera que la media verdadera esté alrededor de 9.5 g. El nivel alfa (la probabilidad de cometer Error Tipo I) lo fija en  $5\%$  y el Poder de la Prueba (la probabilidad de rechazar una hipótesis nula falsa) en  $80\%$ . (Esto último indica que el valor de beta, la probabilidad de cometer Error Tipo II, es  $20\%$ .)

La pregunta que se hace es: ¿cuál debe ser el tamaño de muestra,  $n$ , para lograr el Poder de la Prueba deseado? Para responder esta pregunta se procede de la siguiente forma:

Primero se encuentra el valor  $Z$  correspondiente a alfa  $5\%$  en la tabla respectiva. Este valor es igual a  $1.645$ . Este valor se reemplaza en la fórmula de la Prueba Estadística  $Z$  (PEZ):

$$PEZ = (\bar{Y} - \mu_0) / \sigma_{\bar{Y}} = 1.645$$

$$1.645 = (\bar{Y} - 8) / (2 / \sqrt{n})$$

Ahora tenemos dos incógnitas:  $\bar{Y}$  y  $n$ . Despejamos  $\bar{Y}$ :

$$\bar{Y} = 8 + 1.645(2 / \sqrt{n}) = 8 + 3.29 / \sqrt{n} \quad (\text{Ecuación 1})$$

Luego, como se plantea que la hipótesis alternativa es la hipótesis verdadera ( $\mu_1 = 9.5 g$ ) se encuentra el valor beta al nivel del 20 %. Esto equivale a otro valor de  $PEZ$  y es igual a  $-0.842$ :

$$PEZ = (\bar{Y} - \mu_1) / \sigma_{\bar{Y}} = -0.842$$

$$-0.842 = (\bar{Y} - 9.5) / (2 / \sqrt{n})$$

Despejamos  $\bar{Y}$  y tenemos lo siguiente:

$$\bar{Y} = 9.5 - 0.842(2 / \sqrt{n}) = 9.5 - 1.684 / \sqrt{n} \quad (\text{Ecuación 2})$$

Ahora igualamos las dos ecuaciones y despejamos para  $n$ , el tamaño de la muestra:

$$8 + 3.29 / \sqrt{n} = 9.5 - 1.684 / \sqrt{n}$$

$$3.29 / \sqrt{n} + 1.684 / \sqrt{n} = 9.5 - 8$$

$$4.974 / \sqrt{n} = 1.5, \sqrt{n} = 4.974 / 1.5$$

$$(\sqrt{n})^2 = (3.32)^2, n = 11$$

Se concluye que con un tamaño de muestra igual a 11 se logra en este ejercicio obtener un Poder de la Prueba de 80 %. Esto es igual a la probabilidad de rechazar la hipótesis nula cuando la hipótesis alternativa es verdadera.

**Comentario:** La estimación de una media que corresponda a la hipótesis alternativa debe ser el resultado de la experiencia y conocimiento. No es un dato que se inventa el investigador.

En el caso de encuestas, el procedimiento para estimar el tamaño de la muestra tomando en cuenta el Poder de la Prueba, es similar al de los ejercicios anteriores. Las respuestas con dos opciones (como ejemplo: sí o no, cara o sello, alto o bajo) corresponden a una distribución binomial. Sin embargo, cuando el



tamaño de la muestra es alrededor de 30, la distribución binomial es prácticamente igual a la distribución normal o Z. Es decir, puede usarse el mismo procedimiento de los ejercicios anteriores para pruebas de hipótesis y para determinar el Poder de la Prueba.

En el caso de la distribución binomial, la desviación estándar de la población de medias de las proporciones es igual a:

$$\sigma_{\bar{p}} = \sqrt{pq/n} \quad (27)$$

Las letras p y q representan los dos eventos posibles, es decir que  $(p + q) = 1$ , mientras que n es el tamaño de la muestra.

En el siguiente ejercicio se aplica lo anterior para la prueba de hipótesis con relación a proporciones.

---

**Ejercicio 7.2.3.** Un extensionista agropecuario ha promocionado el uso de una nueva tecnología en varias comunidades rurales. Después de algún tiempo decide estimar qué porcentaje de los agricultores está utilizando la nueva tecnología en comparación con la tecnología anterior. Con esta finalidad planifica una encuesta en la que habrá una sola pregunta: ¿Utiliza usted la nueva tecnología? Sí o No. La proporción de las respuestas positivas la designa con la letra  $p$ . Las hipótesis que se plantea son las siguientes:

$$H_0: p_0 = 0.50$$

$$H_1: p_1 > 0.50$$

La media de la proporción de la hipótesis nula es  $p_0$ , mientras que la media de la proporción de la hipótesis alternativa es  $p_1$ . La media de la muestra del experimento será  $\bar{p}$ . Se supone que el valor de la estadística  $\bar{p}$  estará entre  $p_0$  y  $p_1$ .

El nivel de significación alfa que decide emplear es 0.01, pero espera que la proporción de agricultores que está aplicando la nueva tecnología sea igual a  $p_1 = 0.65$ . (Este valor es la *hipótesis alternativa esperada* por el investigador sobre la base de estudios previos y su propia experiencia.) Además, desea que el Poder de la Prueba sea 0.85 (beta igual a 0.15). Esto es la probabilidad de rechazar la hipótesis nula cuando la hipótesis alternativa es verdadera.

Con las condiciones indicadas, la pregunta es: ¿qué tamaño de muestra debe tomar el investigador?

Para encontrar la respuesta primero se calcula el valor de  $Z$  correspondiente a  $\alpha = 0.01$ . Este valor es  $Z = 2.32$ . El valor de  $Z$  correspondiente a  $\beta = 0.15$  es  $Z = -1.04$ . Estos valores se encuentran en la Tabla  $Z$  del apéndice A. En el siguiente gráfico compuesto por la población  $H_0$  y la población  $H_1$  se observa la ubicación relativa de los valores  $Z$  y los valores de las proporciones de cada hipótesis:

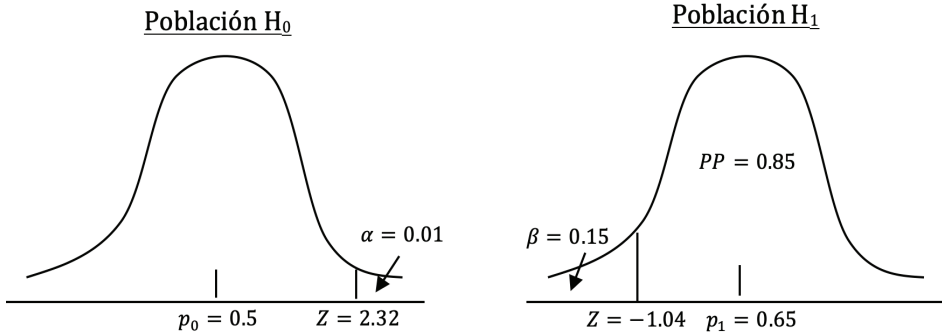


Gráfico 1. Valores de alfa, beta y del poder de la prueba en las poblaciones teóricas de las hipótesis nula y alternativa

Para encontrar el tamaño de la muestra con el que se cumplan las condiciones impuestas se realizan los siguientes cálculos:

La desviación estándar de la población de medias para la población de la hipótesis nula es:

$$\sigma_{\bar{0}} = \sqrt{pq/n} = \sqrt{0.5 * 0.5/n} = \sqrt{0.25/n}$$

La desviación estándar de la población de medias para la población de la hipótesis alternativa:

$$\sigma_{\bar{1}} = \sqrt{pq/n} = \sqrt{0.65 * 0.35/n} = \sqrt{0.23/n}$$

Luego se procede a reemplazar los valores en la fórmula de la Prueba Estadística  $Z$  y despejar para la media de la proporción de la muestra ( $\bar{p}$ ). (De acuerdo con la forma como están planteadas las hipótesis, el valor de la media de la muestra  $\bar{p}$  se encuentra entre el valor de la hipótesis nula:  $p_0 = 0.50$  y el valor de la hipótesis alternativa esperada:  $p_1 = 0.65$ .)

Para el caso de la hipótesis nula:

$$PEZ = (\bar{p} - p_0) / \sigma_{\bar{p}} = 2.32$$

$$2.32 = (\bar{p} - 0.50) / \sqrt{0.25/n}$$

$$\bar{p} = 0.50 + 2.32 * \sqrt{0.25/n}$$

$$\bar{p} = 0.50 + 1.16/\sqrt{n} \quad (\text{Ecuación 1})$$

Se plantea, a continuación, la fórmula de la Prueba Estadística Z en la que  $p_1$  es la media de la proporción de la hipótesis alternativa esperada:

$$PEZ = (\bar{p} - p_1) / \sigma_{\bar{p}} = -1.04$$

$$-1.04 = (\bar{p} - 0.65) / \sqrt{0.23/n}$$

$$\bar{p} = 0.65 - 1.04 * \sqrt{0.23/n}$$

$$\bar{p} = 0.65 - 0.50/\sqrt{n} \quad (\text{Ecuación 2})$$

Ahora, se igualan las Ecuaciones 1 y 2 y se despeja para el tamaño de la muestra n:

$$0.50 + 1.16/\sqrt{n} = 0.65 - 0.50/\sqrt{n}$$

$$1.66/\sqrt{n} = 0.15$$

$$(\sqrt{n})^2 = (1.66/0.15)^2$$

$$n \cong 122$$

En conclusión, si el extensionista agropecuario hace una encuesta a 122 agricultores seleccionados al azar, la probabilidad de aceptar una hipótesis alternativa verdadera igual a 0.65 es de 85%.

### 7.3. Prueba Estadística t de Student

Un inconveniente de la prueba estadística Z, para evaluar hipótesis concernientes a la media de la población, es que necesitamos conocer la varianza de la población  $\sigma^2$ . Con la prueba estadística t de Student no es necesario conocer  $\sigma^2$ .

La distribución  $t$  fue descubierta en forma empírica por W.S. Gosset. Él publicó sus resultados con el seudónimo de “Student” en 1908, en honor a su maestro Karl Pearson; de ahí el nombre de la distribución. Sin embargo, el astrónomo alemán F.R. Helmert ya había obtenido matemáticamente la distribución en 1875.

La fórmula desarrollada por Gosset es bastante similar a la de la distribución  $Z$ :

$$t = \frac{\bar{Y} - \mu_{\bar{Y}}}{S_{\bar{Y}}} \quad (28)$$

La diferencia es la inclusión de la estadística  $S_{\bar{Y}}$  a cambio de  $\sigma_{\bar{Y}}$ .

La distribución  $t$  también es simétrica, algo más aplanada que la distribución normal y con las colas más separadas de la abscisa. En el siguiente gráfico se comparan las dos distribuciones:

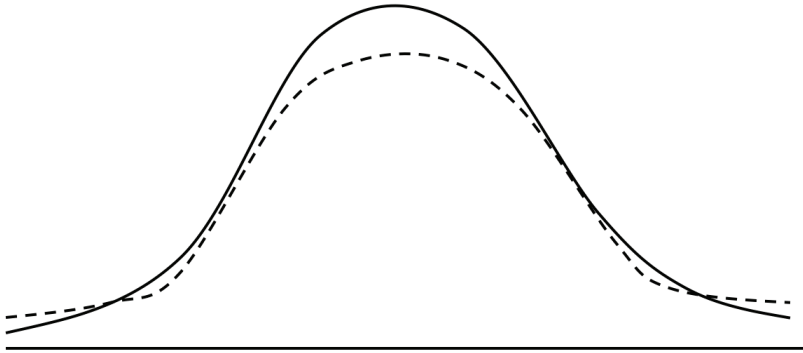


Gráfico 7.3.1. Comparación entre las distribuciones  $Z$  (línea sólida) y la  $t$  (línea punteada)

Sin embargo, la forma de la distribución  $t$  varía de acuerdo con el número de grados de libertad ( $n-1$ ), es decir, está relacionada con el tamaño de la muestra. Cuando  $n-1$  es un número alrededor de 30 las dos distribuciones son prácticamente idénticas. Note en la **Tabla  $t$**  (Apéndice B), que cuando  $(n-1) > 120$ ,  $t$  al 5% es igual a 1.6450. Este valor es el mismo que encontramos para  $\alpha = 5\%$  en la **Tabla  $Z$** .

El procedimiento para la prueba de hipótesis con relación a la media de la población es muy parecido al empleado anteriormente. Sin embargo, la ventaja de la prueba  $t$  es que no necesitamos conocer  $\sigma^2$ .

**Ejercicio 7.3.1.** Supongamos que es de interés probar las siguientes hipótesis referentes a una población de pesos de frutas, normalmente distribuida, con media y varianza desconocidas:

$$H_0: \mu = 10$$

$$\alpha = 1\%$$

Para el efecto, en un experimento se extrae la siguiente muestra: 6, 5, 10, 8, 8, 10, 8, 9.

Para probar las hipótesis planteadas se procede a encontrar las siguientes estadísticas: tamaño de la muestra, media, varianza y desviación estándar. Estas son:

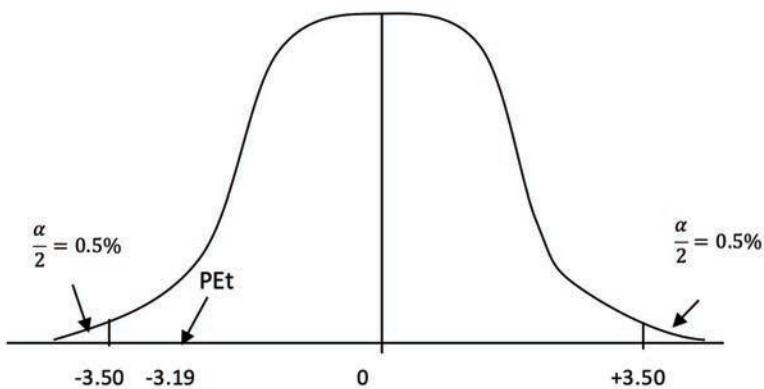
$$n = 8 \quad \bar{Y} = 8 \quad S^2 = 3.14 \quad S = 1.77$$

A continuación encontramos el *error estándar*:

$$S_{\bar{Y}} = \sqrt{\frac{S^2}{n}} = \sqrt{\frac{3.14}{8}} = 0.63$$

$$\text{La Prueba Estadística } t \text{ es: } P.E. t = \frac{\bar{Y} - \mu_0}{S_{\bar{Y}}} = \frac{8 - 10}{0.63} = -3.19$$

Para mayor claridad ubiquemos la *P.E. t* en un gráfico y los valores críticos encontrados en la Tabla *t* (Apéndice B) con 7 grados de libertad:



Como la prueba estadística *t* (*P.E. t*) no cae en las áreas críticas, la decisión es aceptar  $H_0$  que  $\mu = 10$  al 1% nivel de significación. En otras palabras, con el

nivel de significación del 1% no hay suficiente evidencia para creer que  $\mu$  sea distinta de 10.

Con el programa *Statistix* se pueden realizar fácilmente las pruebas  $t$  de una muestra. El resultado es el siguiente:

Null Hypothesis:  $\mu = 10$   
Alternative Hyp:  $\mu <> 10$

Variable	Mean	SE	T	DF	P
Pesos	8.0000	0.6268	-3.19	7	0.0153

Note que el programa no presenta el valor crítico, pero sí la probabilidad de haber obtenido una prueba estadística tan extrema como -3.19. Esta probabilidad, como podemos ver, es

$P = 0.0153$ . Es decir, 1.53%. Esto nos indica que la prueba estadística está entre el 1% y el 5%. En otras palabras, es estadísticamente significativa al 5% pero no al 1%

Los mismos conceptos, principios y discusión acerca de los tipos de error, se aplican también obviamente a estas pruebas, así como a todas las pruebas estadísticas.

## 7.4. La prueba estadística $t$ (P.E.t) y el control de calidad

La prueba estadística  $t$  es muy usada en el control de calidad de procesos industriales. A continuación se presenta un ejemplo:

**Ejercicio 7.4.1.** El departamento de control de calidad de una empresa productora de conservas estipula que la media del peso neto de palmito en lata no debe ser menor de 24 onzas. Pruebe las siguientes hipótesis al nivel alfa del 5% ( $\alpha = 5\%$ ):

$$H_0: \mu = 24$$

$$H_1: \mu < 24$$

Con este fin se extrajo una muestra al azar de 12 latas. Los pesos netos en onzas fueron los siguientes: 23.85 - 23.65 - 24.15 - 24.05 - 23.65 - 23.85 - 24.15 - 23.85 - 23.65 - 24.15 - 24.05 - 23.75

Primero encontramos  $\bar{Y}$ ,  $S^2$ ,  $S\bar{y}$ :

$$\bar{Y} = 23.90, \quad S^2 = 0.0409, \quad S_{\bar{Y}} = 0.0584$$

$$\text{Luego encontramos la P.E. } t = \frac{\bar{Y} - \mu_0}{S_{\bar{Y}}} = \frac{23.90 - 24.0}{0.0584} = -1.7120$$

Como el valor crítico encontrado en la Tabla t (Apéndice B), al 5% nivel de significación y con 11 grados de libertad, es igual a -1.7960 la decisión es aceptar la  $H_0$  que la media del peso es igual a 24 onzas al nivel de significación del 5%. ( $t = -1.7120$  n.s.)

Con el programa *Statistix* los resultados del análisis anterior son:

**Statistix: One-Sample T Test**

Null Hypothesis:  $\mu = 24$

Alternative Hyp:  $\mu < 24$

Variable	Mean	SE	T	DF	P
PesoLatas	23.900	0.0584	-1.71	11	0.0574

Observe que la probabilidad de haber obtenido un valor tan extremo como  $t = -1.71$  es  $P = 0.0574$ . Una probabilidad mayor que el 5%. Por lo tanto, se acepta la Hipótesis Nula. Otra forma de expresar lo mismo es indicar que no hay suficiente evidencia para rechazar la Hipótesis Nula que la media es igual a 24.

## 7.5. Intervalo de Confianza para la media de la población

De Acuerdo con el Teorema del Límite central, las medias de todas las muestras posibles que pueden extraerse de una población siguen una distribución normal. En el concepto de *Intervalo de Confianza* se presupone una distribución normal, en la que la media de la población es desconocida. Sin embargo, mediante la distribución  $t$  se puede estimar la probabilidad de que la media poblacional se encuentre entre dos valores de  $\bar{Y}_i$ . Estos dos valores son equidistantes de la media poblacional hipotética, por la simetría de la distribución  $t$  y se denominan “Límite Inferior” y “Límite Superior”.

El Intervalo de Confianza nos indica la probabilidad o grado de certeza de que la media verdadera esté comprendida entre los dos valores encontrados. El Intervalo de confianza es igual a  $1 - \alpha$ .

El desarrollo de la fórmula para el intervalo de confianza de la media poblacional desconocida o hipotética se basa en la naturaleza de la distribución  $t$ :

Si  $t = \frac{\bar{Y} - \mu}{S_{\bar{Y}}}$ , entonces se puede escribir:  $P\left(-t_{\alpha/2} \leq \frac{\bar{Y} - \mu}{S_{\bar{Y}}} \leq t_{\alpha/2}\right) = 1 - \alpha$

Despejando  $\mu$  obtenemos la fórmula para el intervalo de confianza de la media:

$$P\left(\bar{Y} - t_{\alpha/2} S_{\bar{Y}} \leq \mu \leq \bar{Y} + t_{\alpha/2} S_{\bar{Y}}\right) = 1 - \alpha \quad (29)$$

Fórmula en la que  $1 - \alpha$  es el intervalo de confianza;  $\bar{Y} - t_{\frac{\alpha}{2}} S_{\bar{Y}}$  es el límite inferior del intervalo, mientras que  $\bar{Y} + t_{\frac{\alpha}{2}} S_{\bar{Y}}$  es el límite superior. Note que  $\alpha / 2$  es el subíndice de  $t$  que nos señala el valor de  $t$  en la tabla y que se lo encuentra con los grados de libertad  $(n-1)$  de la muestra. El Intervalo de Confianza puede expresarse en decimales, en cuyo caso es  $1 - \alpha$ , o en porcentaje, esto es  $100 - \alpha$ .

**Ejercicio 7.5.1.** Encuentre los siguientes intervalos de confianza para la media de una población de la que se extrajo una muestra  $n = 16$  con estas estadísticas:  $\bar{Y} = 45.0$ ,  $S_{\bar{Y}} = 4.0$ :

a. **I.C. al 80%.**

Como el Intervalo de Confianza es del 80%,  $\alpha = 20\%$  y  $\alpha / 2 = 10\%$

El valor de  $t_{\alpha/2}$ , con  $(n - 1)$  grados de libertad, es decir 15 grados de libertad (g.l.) y a un nivel del 10% es:  $t_{(0.10, 15gl)} = 1.3410$

Entonces, el intervalo de confianza es: I.C. (80%) =  $45.0 \pm 1.3410 * 4 = 39.63$  y  $50.36$ .

En otras palabras, se puede indicar que la media de la población  $\mu$  se encuentra entre 39.63 y 50.36 con el 80% de probabilidad.

b. **I.C. al 90%,  $\alpha = 10\%$  y  $\alpha / 2 = 5\%$**

$t_{(0.05, 15gl)} = 1.7530$ , entonces el intervalo de confianza es:

I.C. (90%) =  $45.0 \pm 1.7530 * 4 = 37.98$  y  $52.01$ .

c. **I.C. al 95%,  $\alpha = 5\%$  y  $\alpha / 2 = 2.5\%$**

$t_{(0.025, 15gl)} = 2.1310$ , entonces el intervalo de confianza es:

I.C. (95%) =  $45.0 \pm 2.1310 * 4 = 36.47$  y  $53.52$ .



d. **I.C. al 99%**,  $\alpha = 1\%$  y  $\alpha / 2 = 0.5\%$

$t_{(0.005, 15gl)} = 2.947$ , entonces el intervalo de confianza es:

I.C. (99%) =  $45.0 \pm 2.9470 \cdot 4 = 33.21$  y  $56.78$ .

Un resumen de los resultados del ejercicio anterior se presenta en el Cuadro 7.5.1.

**Cuadro 7.5.1. Intervalos de Confianza, a diferentes probabilidades, para la media de una población cuando  $n = 16$ ,  $\bar{Y} = 45$ ,  $S_{\bar{y}} = 4$**

Intervalo de Confianza	Límite Inferior	Límite Superior	Amplitud del Intervalo
80 %	39.63	50.36	10.73
90 %	37.98	52.01	14.03
95 %	36.47	53.52	17.05
99 %	33.21	56.78	23.57

**Comentarios:** Note en el cuadro anterior que los intervalos se vuelven más amplios a medida que  $100 - \alpha / 2$  se aproxima al 100%. Sin embargo, nunca podemos tener certeza absoluta de que el intervalo contenga a la verdadera media de la población. Esto se explica por razón de que la distribución t, a semejanza de otras distribuciones de probabilidades, es una curva asintótica. Es decir que en sus extremos se aproxima, hasta el infinito, al eje de la abscisa.

En las pruebas de hipótesis si se acepta la hipótesis nula a un determinado nivel alfa, la media de la muestra estará dentro del intervalo de confianza. Si por el contrario, se rechaza la hipótesis nula, la media de la muestra quedará afuera del intervalo. Esto se cumple siempre que el nivel alfa de la prueba de hipótesis sea el mismo que para el intervalo de confianza.

## La distribución Ji-Cuadrado

*“No ponga su fe en lo que la estadística dice hasta que no haya analizado cuidadosamente qué es lo que no dice”. William W. Watt<sup>16</sup>*

### 8.1. Origen de la distribución Ji – Cuadrado

La distribución Ji-Cuadrado tiene que ver con la varianza y está relacionada con la distribución normal y con el teorema del límite central. Se representa con la letra griega *Ji* al cuadrado:  $\chi^2$ .

Si se tiene una población normalmente distribuida,  $Y_i$  DN  $(\mu, \sigma^2)$  y se extraen todas las muestras posibles de tamaño  $n$  y para cada muestra se calcula la estadística  $\frac{(n-1)S^2}{\sigma^2}$ , estas cantidades siguen la distribución Ji-Cuadrado con  $(n-1)$  grados de libertad.

La forma de la distribución  $\chi^2$  depende del número de grados de libertad, pero siempre es asimétrica. En la siguiente figura se presentan varias curvas de  $\chi^2$  cuyas formas específicas dependen de los grados de libertad.

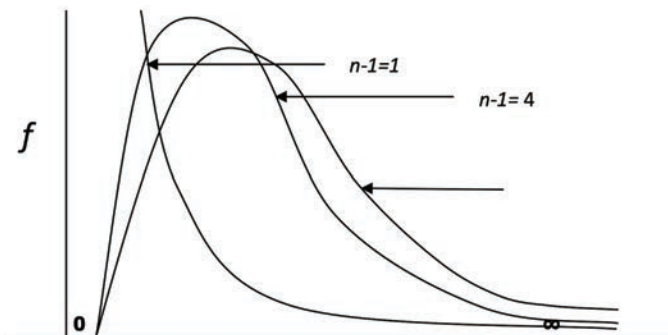


Gráfico 8.1.1. Formas de la distribución  $\chi^2$  en función de los grados de libertad

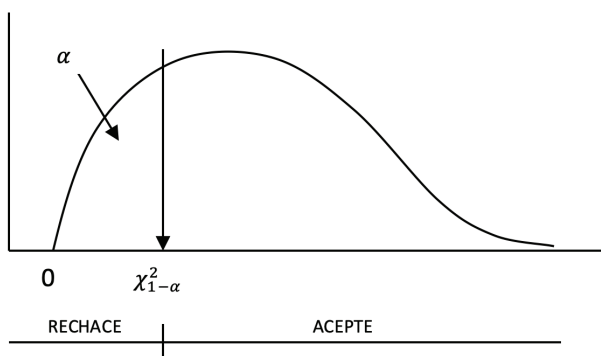
Así como las distribuciones  $Z$  y  $t$  nos permiten probar hipótesis con relación al parámetro  $\mu$ , así la distribución  $\chi^2$  nos permite probar hipótesis con relación al parámetro  $\sigma^2$ . El planteamiento de las pruebas es en general similar al usado para  $Z$  y  $t$ , sin embargo, como la distribución  $\chi^2$  no es simétrica, los valores críticos son distintos para las pruebas de cola izquierda y derecha.

La prueba estadística para evaluar una varianza poblacional hipotética es:

$$PE\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} \quad (30)$$

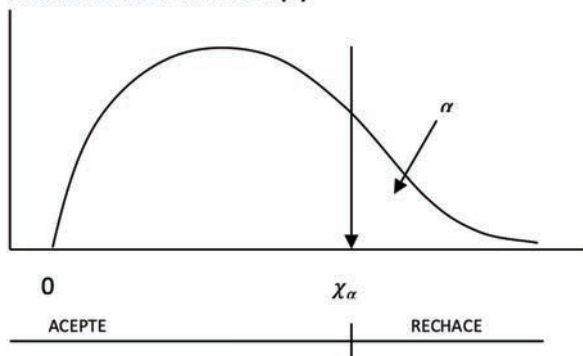
Note que el valor mínimo posible de  $\chi^2$  es cero. En otras palabras, no hay valores negativos en la distribución  $\chi^2$  ya que no puede haber varianzas negativas. En las siguientes figuras se presentan las áreas críticas para las pruebas de cola izquierda, cola derecha y dos colas.

#### PRUEBA DE COLA IZQUIERDA (A)



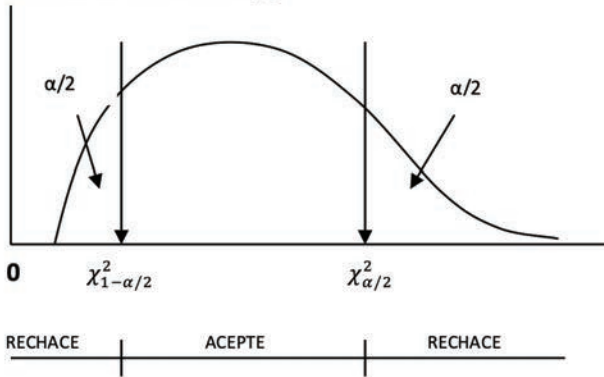
$$\begin{aligned} H_0: \sigma^2 &= \sigma_0^2 \\ H_1: \sigma^2 &< \sigma_0^2 \\ PE\chi^2 &= \frac{(n-1)S^2}{\sigma_0^2} \end{aligned}$$

#### PRUEBA DE COLA DERECHA (B)



$$\begin{aligned} H_0: \sigma^2 &= \sigma_0^2 \\ H_1: \sigma^2 &> \sigma_0^2 \\ PE\chi^2 &= \frac{(n-1)S^2}{\sigma_0^2} \end{aligned}$$

**PRUEBA DE DOS COLAS (C)**



$$H_0: \sigma^2 = \sigma_0^2$$

$$H_1: \sigma^2 \neq \sigma_0^2$$

$$PE_{\chi^2} = \frac{(n-1)S^2}{\sigma_0^2}$$

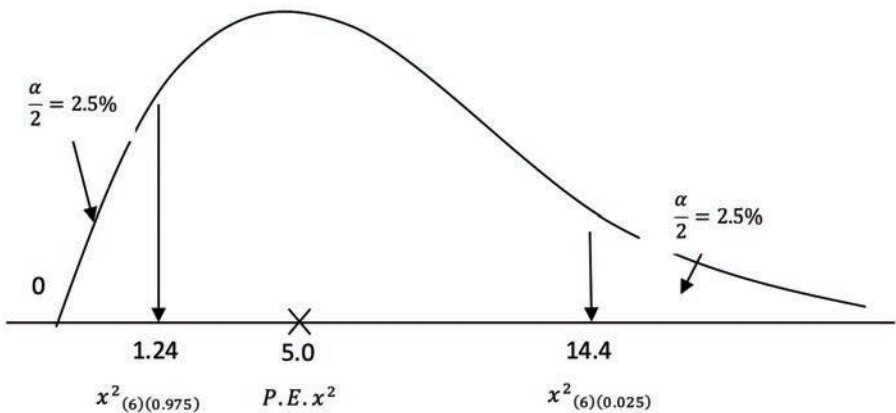
Gráfico 8.1.2. Ubicación de las zonas de aceptación y rechazo para la prueba de Ji Cuadrado. Pruebas de cola izquierda (A), cola derecha (B) y dos colas (C).

**Ejercicio 8.1.1.** En un experimento se desea probar la hipótesis nula de que la varianza de una población con distribución normal es igual a 20 versus la hipótesis alternativa de que no es igual a 20. La muestra  $n = 7$  que se extrajo tuvo varianza 16.67. El nivel de significación que se escogió fue  $\alpha = 5\%$ . Realice la prueba:

$$H_0: \sigma^2 = 20$$

$$H_1: \sigma^2 \neq 20 \quad P.E. \cdot \chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{6 \cdot 16.67}{20} = 5.0$$

Para mayor comprensión se ubica el valor de la P.E. en un gráfico:



Note, por el planteamiento de las hipótesis, que ésta es una prueba de dos colas. Es importante tener en cuenta que los valores críticos para las áreas de cola izquierda y cola derecha son distintos en la prueba de  $\chi^2$ . Esto es por la asimetría de la distribución. Como  $\frac{\alpha}{2} = 0.025$ , los valores críticos se los encuentra en la Tabla de Ji-Cuadrado (Apéndice C) con 6 grados de libertad, en las columnas de 0.975 y 0.025. Los valores críticos son, redondeando, 1.24 (cola izquierda) y 14.40 (cola derecha).

Como la Prueba Estadística *P.E.* no cae en las áreas críticas, se puede concluir que no hay suficiente evidencia para rechazar la  $H_0$ . En otras palabras, se acepta la  $H_0$  que  $\sigma^2=20$  al 5% nivel de significación.

## 8.2. Aplicaciones de la distribución Ji-Cuadrado

La distribución Ji-Cuadrado tiene aplicaciones y pruebas específicas tanto en la estadística **paramétrica** como en la estadística **no paramétrica**.

En el Ejercicio 7.4.1 se tuvo un ejemplo de una prueba de calidad con relación a la media de un producto. Sin embargo, en los procesos industriales no es solo importante que la media sea la estipulada (peso promedio de un producto, contenido de proteína, cantidad de ingrediente activo, etc.), sino también que la variación existente entre los ítems fabricados esté dentro de los estándares establecidos. Para esto último, la prueba paramétrica de  $\chi^2$  para evaluar la varianza de una población es lo recomendado.

---

**Ejercicio 8.2.1.** El departamento de control de calidad de una empresa productora de conservas estipula que la **varianza** del peso neto de palmito en lata no debe exceder el valor de 0.25 onzas<sup>2</sup>. Para realizar la prueba  $\chi^2$  se tomó una muestra de 12 latas. Los pesos netos en onzas fueron: 23.85 - 23.65 - 24.15 - 24.05 - 23.65 - 23.85 - 24.15 - 23.85 - 23.65 - 24.15 - 24.05 - 23.75. El valor de la varianza de la muestra fue  $S^2 = 0.0409$ . Pruebe las hipótesis correspondientes al nivel  $\alpha = 1\%$ :

$$H_0: \sigma^2 = 0.25$$

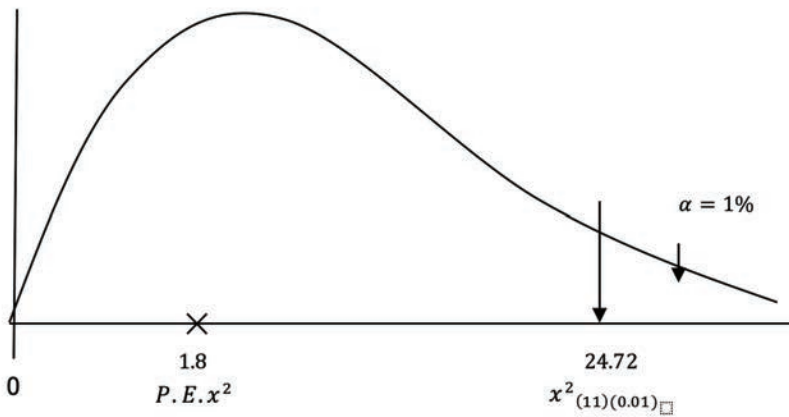
$$H_1: \sigma^2 > 0.25$$

Note que la hipótesis alternativa, en la forma que está planteada (“mayor que 0.25”), determina que sea prueba de cola derecha al indicar: “mayor que 0.25”. Las pruebas de una cola son más precisas y en este caso lo que interesa es probar que no se excede del valor estipulado.

$$\text{La prueba estadística es: } P.E. \chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{11 \cdot 0.0409}{0.25} = 1.8$$

El valor crítico correspondiente al área crítica  $\alpha = 1\%$  se lo encuentra en la Tabla de Ji-Cuadrado (Apéndice C), bajo la columna “0.010” y en la hilera de 11 grados de libertad. Este valor es 24.7. (Todos los valores de las probabilidades de las diferentes distribuciones pueden encontrarse también fácilmente con el programa *Statistix*.)

Para mayor claridad ubicamos la *P.E.* en un gráfico:



Se puede observar en el gráfico que la prueba estadística no cae en el área crítica. Consecuentemente se concluye que no hay suficiente evidencia para rechazar la  $H_0$  al 1% nivel de significación. Esto quiere decir que se cumple la estipulación de que la varianza no sea mayor, estadísticamente, que 0.25 onzas<sup>2</sup>. Esta prueba es importante en las industrias para garantizar la uniformidad del producto.

### 8.3. Intervalo de confianza para la varianza de la población

Con la distribución  $\chi^2$  podemos también determinar intervalos de confianza para la varianza de la población.

La fórmula puede deducirse de los ejemplos anteriores. Esta es:

$$P\left(\frac{(n-1)S^2}{\chi^2_{(\alpha/2)}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{(1-\alpha/2)}}\right) = 1-\alpha \quad (31)$$

La letra *P* indica “probabilidad”. Los conceptos y forma de aplicación son similares a los empleados para la media de la población.

**Ejercicio 8.3.1** Se extrajo una muestra al azar de una población  $Y_i \text{ DN } (\mu, \sigma^2)$ . El tamaño de la muestra fue  $n=20$  y el valor  $(n-1)S^2=483$ .

a) Se desea estimar un intervalo de confianza del 90% para la varianza de la población ( $\sigma^2$ ). Encuentre el intervalo:

$$P\left(\frac{(n-1)S^2}{\chi^2_{(19)(0.005)}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{(19)(0.95)}}\right) = 1-\alpha \quad \text{Reemplazando: } P\left(\frac{483}{30.1} \leq \sigma^2 \leq \frac{483}{10.1}\right) = 0.90$$

Intervalo de Confianza del 90% = 16.05 y 47.82

Es decir, se puede asegurar que la varianza  $\sigma^2$  se encuentra entre los valores 16.05 y 47.82 con un 90% de probabilidad.

b) Con los mismos datos anteriores, calcule ahora un I.C. del 99%:

$$P\left(\frac{(n-1)S^2}{\chi^2_{(19)(0.005)}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{(19)(0.995)}}\right) = 1-\alpha \quad \text{Reemplazando: } P\left(\frac{483}{38.6} \leq \sigma^2 \leq \frac{483}{6.8}\right) = 0.99$$

Intervalo de Confianza del 99% = 12.51 y 70.61

**Comentario:** Note como el intervalo se vuelve más amplio a medida que se incrementa el nivel de confianza. Un intervalo de confianza del 99% quiere decir que la probabilidad de que el intervalo contenga a la verdadera varianza de la población es 99%. En este caso hay 1% de probabilidad de error. Sin embargo, como se discutió en el caso de la media de la población, nunca podemos tener certeza absoluta de que la varianza esté incluida en el intervalo.

## 8.4. Ji – Cuadrado y la estadística no paramétrica

La distribución Ji- Cuadrado tiene aplicaciones muy útiles en el campo de la *estadística no paramétrica*. En este caso no nos interesan los parámetros de la población (media, varianza), sino más bien las formas de las distribuciones. En otras palabras, nos interesa probar la hipótesis de que los datos de una muestra provienen de una población que tiene una determinada distribución.

La prueba de  $\chi^2$  para aplicaciones no paramétricas se ha usado mucho en Genética y otros campos y es adecuada para datos que caen en categorías específicas, es decir para variables cualitativas, como por ejemplo: plantas enfermas, semillas de cotiledón amarillo, animales de pelo corto, número de machos y hembras, preferencias de los consumidores, etc. Cuando se trata de una sola característica o variable y nos interesa evaluar si los datos se encuentran en una proporción

determinada, la prueba de Ji – Cuadrado se denomina prueba para la “**bondad de ajuste**”. Si por lo contrario, lo que interesa es determinar si dos variables son independientes, la prueba de Ji – Cuadrado se denomina “**prueba de contingencia**”

La prueba estadística para la bondad de ajuste es:

$$\chi^2 = \sum \frac{(Ob-Es)^2}{Es} \quad (32)$$

En donde “**Ob**” es el número de ítems o eventos “observados” de una determinada característica y “**Es**” es el número teórico “esperado” de ítems o eventos, de acuerdo con la distribución teórica planteada en la hipótesis nula  $H_0$ .

La prueba se realiza con  $(n - 1)$  grados de libertad, en donde  $n$  es el número de categorías en las que se encuentran los eventos. Como ejemplo, una prueba relacionada con la proporción de una progenie en la que se tiene “semillas lisas” y “semillas rugosas” (dos categorías), habrá sólo un grado de libertad. Esto es así, por cuanto al conocer el número de semillas lisas de un total, la otra categoría, número de semillas rugosas, queda automáticamente determinado.

Sin embargo, cuando algunas de las categorías tienen un bajo número de observaciones (cinco o menos), en la fórmula se introduce un factor de corrección:

$$\chi^2 = \sum \frac{(|Ob-Es|-0.5)^2}{Es} \quad (33)$$

El restar 0.5 de los valores absolutos  $|Ob-Es|$  se denomina la **Corrección de Yates** y es necesaria para obtener una mejor aproximación a la distribución teórica planteada.

Es necesario recordar que, cuando se trata de dos categorías que se pueden expresar en proporciones,  $p$  y  $q$  representan las categorías (cara y sello, macho y hembra, sí y no, etc.) En todo caso, en la distribución binomial la suma de las probabilidades de los dos eventos contrastantes debe ser igual a uno. Esto es:  $(p + q) = 1$ .

---

**Ejercicio 8.4.1.** Al lanzar monedas al aire se espera igual número de caras y de sellos. Si en un experimento se contaron ocho caras y tres sellos, ¿Se puede creer que esta desviación sea solo producto de azar o que esté interviniendo algún otro factor? Para realizar la prueba estadística debe tomarse en cuenta el bajo número de observaciones en una categoría, por lo que debe aplicarse la corrección de Yates:



Cuadro 1. Prueba de Ji-cuadrado con el uso de la corrección de Yates

Categorías	Ob	Es	$ Ob-Es -0.5$	$( Ob-Es -0.5)^2$	$( Ob-Es -0.5)^2/Es$
Caras	8	5.5	+2	4	0.73
Sellos	3	5.5	- 2	4	0.73
TOTAL	11	11	0		$PE\chi^2 = 1.46$

El valor de Ji-Cuadrado en la tabla, con un grado de libertad y al 5% de probabilidad, es igual a 3.84. Como el valor de Ji-cuadrado calculado (1.46) no supera al valor de la tabla se puede concluir que no hay suficiente evidencia para rechazar la hipótesis nula. En otras palabras, la desviación existente puede atribuirse al azar, con una probabilidad de error del 5%.

**Ejercicio 8.4.2.** En un experimento, al lanzar 50 monedas al aire el resultado fue 30 caras y 20 sellos. En teoría, si la probabilidad de obtener cara es  $p=1/2$  y de obtener sello es  $q=1/2$  se esperarían 25 caras y 25 sellos. ¿Puede atribuirse la desviación observada únicamente al azar o, en su defecto, a que varias monedas de las 50 tienden a caer más de un lado que del otro? Para evaluar esta posibilidad pruebe la  $H_0$  de que el número de caras es igual al número de sellos, es decir, que están en la proporción uno a uno. Primero debemos plantear las hipótesis:

$H_0$ : la población de caras y sellos está en la proporción 1 a 1

$H_1$ : la población de caras y sellos NO está en la proporción 1 a 1

O, en forma más simplificada:

$H_0: p = 0.5$

$H_1: p \neq 0.5$

El nivel de significación que se emplea queda a criterio del investigador, pero por lo general se usa el 5%, ( $\alpha = 5\%$ ). Para encontrar la prueba estadística construimos un Cuadro de  $\chi^2$ . Es importante recordar, que de acuerdo con lo postulado en la  $H_0$ , las categorías de los datos esperados (Es) deben estar en la proporción 1 a 1:

Cuadro 1. Prueba estadística de Ji-cuadrado para evaluar la hipótesis nula  $p = 0.5$  en el lanzamiento al aire de 50 monedas

Categorías	Ob	Es	(Ob-Es)	(Ob-Es) <sup>2</sup>	$\frac{(Ob-Es)^2}{Es}$
Caras	30	25	+5	25	1.0
Sellos	20	25	-5	25	1.0
$\Sigma$	50	50	0		$P.E.\chi^2 = 2.0$

La prueba estadística  $\chi^2$ , o Ji-Cuadrado calculado, es  $PE\chi^2 = 2.0$ . Este valor debe compararse con el valor de  $\chi^2$  en la tabla correspondiente (Apéndice C) al nivel alfa del 5% y con un grado de libertad. Esto es,  $\chi_{0.05,1gl}^2 = 3.84$ .

Como el valor de la  $PE\chi^2$  no excede al valor de  $\chi^2$  en la tabla, la decisión es aceptar la  $H_0$  al 5% nivel de significación. En otras palabras, no hay la suficiente evidencia para rechazar la hipótesis nula. El número de caras y el número de sellos están en la proporción 1:1; la diferencia encontrada (30 vs. 20) puede atribuirse al azar, con una probabilidad de 5% de error.

El mismo ejercicio anterior puede resolverse fácil y rápidamente con el programa *Statistix 9.0*. El resultado es el siguiente:

#### Multinomial Test

Hypothesized Proportions Variable: Es  
Observed Frequencies Variable: Ob

Category	Hypothesized Proportion	Observed Frequency	Expected Frequency	Chi-Square Contribution
Caras	0,50000	30	25,00	1,00
Sellos	0,50000	20	25,00	1,00

Overall Chi-Square            2,00  
P-Value                            0,1573  
Degrees of Freedom            1

**Comentario:** Note que el programa *Statistix* presenta el porcentaje del área a la derecha del punto donde cae la prueba estadística. Este porcentaje es la probabilidad de un valor tan extremo como el obtenido. Como se esperaba, la prueba estadística fue igual a 2 y el área a la derecha, es decir la probabilidad, fue igual a 0.1573 (15.73%). Obviamente las conclusiones son similares que cuando se hizo el análisis manualmente: no hay suficiente evidencia para rechazar la hipótesis nula).

En el siguiente ejercicio también se usa la prueba de Ji – Cuadrado para la bondad de ajuste y se refiere a un experimento de biología aplicada:

**Ejercicio 8.4.3.** En un experimento se radiaron huevos de la mosca del vinagre *Drosophila melanogaster*. A los 10 días se contaron 300 hembras y 200 machos adultos. La proporción de los sexos en poblaciones de *Drosophila* y algunos otros organismos (incluyendo la especie humana) es  $\frac{1}{2}$  a  $\frac{1}{2}$ . Pruebe la hipótesis nula que la proporción es la esperada al nivel  $\alpha = 5\%$

$$H_0: p = 0.5$$

$$H_1: p \neq 0.5$$

Se procede en forma similar al ejercicio anterior. Estructuramos un cuadro de  $\chi^2$  con las categorías y los valores observados y esperados:

Cuadro 1. Prueba de Ji Cuadrado

Categorías	Ob	Es	(Ob-Es)	(Ob-Es) <sup>2</sup>	$\frac{(Ob-Es)^2}{Es}$
Hembras	300	250	+50	2500	10
Machos	200	250	-50	2500	10
$\Sigma$	500	500	0		$P.E. \chi^2 = 20$

La prueba estadística, o valor de Ji cuadrado calculado, es  $PE\chi^2 = 20$  con un grado de libertad. En la Tabla de  $\chi^2$  (Apéndice C), al 5% nivel de significación y con un grado de libertad, se encuentra el valor  $\chi_{0.05,1gl}^2 = 3.84$ . Como claramente el valor de la P.E. excede el valor de la tabla la decisión es rechazar la  $H_0$  al 5% nivel de significación. Es decir que la proporción de hembras y machos no es igual. Se puede válidamente pensar que las radiaciones afectan más a los huevos que llevan el cromosoma Y (huevos que producirán machos) que a los que no lo llevan.

Los programas de computación proporcionan directamente la probabilidad. En este caso, la probabilidad de haber obtenido únicamente por azar un valor tan extremo de  $\chi^2 = 20$  es igual o menor que 0.00009.

Si se realiza el ejercicio anterior con el programa Statistix el resultado es el siguiente:

Statistix 9.0

**Multinomial Test**

Hypothesized Proportions Variable: Es  
 Observed Frequencies Variable: Ob

Category	Hypothesized Proportion	Observed Frequency	Expected Frequency	Chi-Square Contribution
Hembras	0,50000	300	250,00	10,00
Machos	0,50000	200	250,00	10,00

Overall Chi-Square 20,00  
 P-Value 0,0000  
 Degrees of Freedom 1

**Comentario:** La prueba de  $\chi^2$  para la “bondad de ajuste” debe realizarse siempre a partir del número real de observaciones, esto es, números enteros y no sobre la base de porcentajes. Observe que los resultados de los ejercicios 8.4.2 y 8.4.3 son completamente diferentes, a pesar de que los valores observados tienen los mismos porcentajes: 30 caras y 20 sellos son 60% y 40%, así como 300 hembras y 200 machos son también 60% y 40%, respectivamente. La diferencia estriba en las cantidades. A mayor número de observaciones, menor error y consecuentemente, mayor precisión en las pruebas estadísticas.

## 8.5. La prueba Ji – Cuadrado y la genética

La prueba de  $\chi^2$  tiene varias aplicaciones en Genética. Se emplea en especial para determinar el posible modo de herencia y las relaciones entre genes. Como ejemplo, si se trata de una cruce di-híbrida (dos características), cuando hay segregación y recombinación independiente de los genes involucrados e igual viabilidad de los gametos, se espera en la  $F_2$  la proporción 9/16: 3/16: 3/16: 1/16. Las proporciones en la progenie pueden evaluarse con los procedimientos de los ejercicios anteriores.

**Ejercicio 8.5.1** En el maíz la presencia de lígula está determinada por un gene dominante  $L$ . Las plantas que tienen el genotipo  $ll$  carecen de lígula. De igual manera, el color del tallo morado es una característica determinada por el gene dominante  $V$ , mientras que las plantas  $vv$  presentan tallo de coloración verde. En la  $F_2$  de este cruzamiento  $LLVV * llvv$  se encontró lo siguiente:

Cuadro 1. Genotipos y fenotipos de las plantas en la  $F_2$  de la crucea  $LLVV * llvv$ 

Número de plantas	Genotipo	Fenotipo
780	$L-V-$	Con lígula, tallo morado
260	$L-vv$	Con lígula, tallo verde
245	$ll V-$	Sin lígula, tallo morado
85	$ll vv$	Sin lígula, tallo verde
Total: 1370 plantas		

Pruebe las hipótesis:

$$H_0: \text{proporción} = 9:3:3:1$$

$$H_1: \text{proporción} \neq 9:3:3:1$$

$$\alpha = 5\%$$

Los datos observados en el experimento (Ob), las cantidades teóricas esperadas de acuerdo con la hipótesis nula (Es) y demás componentes de la prueba de Ji-Cuadrado se presentan en el siguiente cuadro:

Cuadro 2. Valores observados, esperados y otras estadísticas de la prueba Ji-cuadrado para la  $F_2$  de la crucea  $LLVV * llvv$ 

Categorías	Ob	Es	$(Ob-Es)$	$\frac{(Ob-Es)^2}{Es}$
$L-V-$	780	$1370 \cdot 9/16 = 770.6$	+9.4	0.11
$L-vv$	260	$1370 \cdot 3/16 = 256.9$	+3.1	0.04
$ll V-$	245	$1370 \cdot 3/16 = 256.9$	-11.9	0.55
$ll vv$	85	$1370 \cdot 1/16 = 85.6$	-0.6	0.00
Suma	1370	1370	0	$PE\chi^2 = 0.70$

Note que los valores esperados resultan de multiplicar las proporciones 9/16, 3/16, 3/16 y 1/16 por 1370, que es el número total de plantas.

La  $P.E.\chi^2$  es igual a 0.70, con tres grados de libertad. El valor en la Tabla  $\chi^2$  (Apéndice C), al nivel alfa del 5% y con tres grados de libertad (hay cuatro categorías) es  $\chi^2_{0.05,3gl} = 7.81$ . La decisión es aceptar la  $H_0$  al 5% nivel de significación. Las desviaciones de los valores observados con los esperados son tan pequeñas que pueden ser atribuidas al azar. En términos genéticos puede señalarse que hay suficiente evidencia de que los dos pares de genes se separan y se recombinan independientemente.

Para obtener el análisis con el programa *Statistix* se crean dos variables: en la primera, los datos observados (Ob) y en la segunda, la proporción esperada (Es). en este caso es suficiente escribir los numeradores de la proporción, es decir 9: 3: 3: 1. Los resultados son:

#### Multinomial Test

Hypothesized Proportions Variable: Es  
Observed Frequencies Variable: Ob

Category	Hypothesized Proportion	Observed Frequency	Expected Frequency	Chi-Square Contribution
1	0,56250	780	770,63	0,11
2	0,18750	260	256,88	0,04
3	0,18750	245	256,88	0,55
4	0,06250	85	85,63	0,00

Overall Chi-Square 0,71  
P-Value 0,8719  
Degrees of Freedom 3

Observe el valor de P. La probabilidad de haber obtenido estas desviaciones, entre lo que se observó y lo que se esperaba, es igual a 0.8719 (87.19%). Esta es la probabilidad de que las desviaciones se deban al azar. Como el porcentaje es alto, la mejor decisión es aceptar la Hipótesis Nula.

Otro ejemplo en el que también se plantea como hipótesis nula la proporción 9:3:3:1, es decir, la segregación y recombinación independiente de dos pares de genes, se presenta a continuación. En este caso, la evidencia estadística lleva a rechazar la hipótesis nula y a buscar, consecuentemente, otra explicación para los resultados obtenidos.

**Ejercicio 8.5.2.** En la cebada, en la  $F_2$  de un cruzamiento entre líneas puras se encontró lo siguiente (Cuadro 1):

**Cuadro 1. Segregación en la  $F_2$  de dos características en la cebada: número de hileras y color del follaje**

Número de plantas	Genotipo	Fenotipo
1178	H-C-	Seis hileras, follaje verde
273	H-cc	Seis hileras, follaje clorótico
291	hhC-	Dos hileras, follaje verde
156	hhcc	Dos hileras, follaje clorótico
Total 1898		

Pruebe las hipótesis:

$$H_0: \text{proporción} = 9:3:3:1$$

$$H_1: \text{proporción} \neq 9:3:3:1$$

$$\alpha = 5\%$$

Cuadro 2. Valores observados y esperados y cálculo del valor de  $\chi^2$  para las características en la  $F_2$ : número de hileras y color del follaje en la cebada

Categorías	Ob	Es	(Ob-Es)	$\frac{(Ob-Es)^2}{Es}$
H- C-	1178	1898* 9/16=1067.6	+110.4	11.4
H- cc	273	1898* 3/16= 355.9	- 82.9	19.3
Hh C-	291	1898* 3/16= 355.9	- 64.9	11.8
hh cc	156	1898* 1/16= 118.6	+ 37.4	11.8
Totales	1898	1898	0	$P.E. \chi^2 = 54.3$

La  $P.E. \chi^2$  es igual a 54.3, con tres grados de libertad. El valor en la Tabla de  $\chi^2$  (Apéndice C) es  $\chi_{0.05,3gl}^2 = 7.81$ . La prueba estadística excede con un amplio margen al valor de la tabla. Consecuentemente la mejor decisión es rechazar la  $H_0$  al 5% nivel de significación. Las desviaciones entre los valores observados y los esperados son tan grandes que no podemos explicarlos únicamente como resultantes del azar. En términos genéticos, la evidencia apoya la idea de que para este par de características no hay segregación ni recombinación independientemente. Posiblemente los genes están ligados, es decir, se encuentran en el mismo cromosoma, en cuyo caso no se espera la proporción típica 9:3:3:1 de un cruzamiento di-híbrido.

Los resultados con el programa *Statistix* son:

#### Statistix Multinomial Test

Category	Hypothesized Proportion	Observed Frequency	Expected Frequency	Chi-Square Contribution
1	0.56250	1178	1067.63	11.41
2	0.18750	273	355.88	19.30
3	0.18750	291	355.88	11.83
4	0.06250	156	118.63	11.78

Overall Chi-Square: 54.31, P-Value: 0.0000, Degrees of Freedom: 3

Note que la probabilidad de obtener desviaciones tan grandes como las encontradas, sólo por azar, es bastante baja ( $P = 0.0000$ , es decir,  $\leq 1/10\ 000$ ). Consecuentemente la decisión es rechazar la  $H_0$ .

## 8.6. Ji-Cuadrado y la prueba de contingencia

La prueba de contingencia es un método de la estadística no paramétrica cuyo objetivo es determinar si existe o no interacción o dependencia entre dos variables. En los ejercicios anteriores se estudió el uso de la prueba de Ji-cuadrado cuando se tenía una variable nominal o cualitativa. En esta prueba se trata de dos variables cualitativas. Como los datos se organizan en un cuadro de dos o más hileras y dos o más columnas, el análisis correspondiente se denomina “prueba de Ji-Cuadrado en cuadros de contingencia”.

Una vez que se han obtenido los valores experimentales u observados (Ob), el siguiente paso es encontrar los valores teóricamente esperados (Es) en la ausencia de interacción. La fórmula para encontrar el valor esperado de una casilla cualquiera  $a_i b_j$  es la siguiente:

$$(Es) a_i b_j = \frac{C_i * H_j}{GT}$$

Fórmula en la que  $C_i$  indica una columna específica,  $H_j$  indica la hilera correspondiente y  $GT$  la suma total de los datos observados.

Para encontrar los grados de libertad para las pruebas de contingencia se aplica la siguiente fórmula:  $g. l. = (A-1)*(B-1)$ , en la que A y B representan el número de “niveles” de cada variable. A continuación, se presenta un ejemplo.

---

**Ejercicio 8.6.1** En un experimento se evaluó la posible interacción entre dos variables: A, híbridos de maíz y B, resistencia al marchitamiento causado por hongos del suelo. La variable A constó de dos híbridos: a1 y a2. La variable B presentó también dos niveles, b1 (plántulas muertas) y b2 (plántulas vivas). Para este fin, en una parcela inoculada con hongos del suelo, se sembraron 100 semillas de cada híbrido, en hileras alternas. Los resultados observados a los 15 días de la siembra se presentan en el siguiente cuadro:



Cuadro 1. Número de plántulas muertas y vivas de dos híbridos de maíz

## Variable B

Variable A	Plántulas Muertas, b1	Plántulas Vivas, b2	Totales, $H_j$
Híbrido a1	a1b1=26	a1b2=72	$H_1 = 98$
Híbrido a2	a2b1=33	a2b2=64	$H_2 = 97$
Totales, $C_i$	$C_1=59$	$C_2=136$	$GT=195$

Si estos híbridos de maíz tuvieran la misma resistencia al marchitamiento, la proporción de plantas muertas y vivas debería ser estadísticamente igual para los dos híbridos. Por lo tanto, las hipótesis que se plantean en términos de los híbridos y su resistencia al marchitamiento son:

$$H_0: a1 = a2$$

$$H_1: a1 \neq a2$$

$$\alpha = 5\%$$

La prueba estadística de Ji-cuadrado tiene la misma fórmula de los ejercicios anteriores. Sin embargo, para encontrar los valores esperados debe aplicarse la fórmula de la regla de

Proporciones:

$$(Es)a_i b_j = \frac{C_i * H_j}{GT}$$

Como ejemplo, el valor esperado para la primera casilla a1b1 del Cuadro 1 es:

$$(Es)a_1 b_1 = \frac{C_1 * H_1}{GT} = \frac{59 * 98}{195} = 29,65$$

Para la siguiente casilla a1b2, el valor esperado es:

$$(Es)a_1 b_2 = \frac{C_1 * H_2}{GT} = \frac{59 * 97}{195} = 29,35$$

Se puede continuar de forma similar para encontrar los valores esperados de las otras casillas. Sin embargo, en el caso de los cuadros de contingencia que tienen sólo dos niveles en cada una de las categorías o variables, sólo hace falta encontrar uno de los valores esperados. Los otros tres pueden encontrarse por diferencia. Esto tiene que ver con los grados de libertad.

En este ejercicio los grados de libertad de la prueba son:

$g.l. = (A-1)*(B-1) = (2-1)*(2-1) = 1$ . Note que la variable A tiene dos niveles: el híbrido a1 y el a2. La variable B también tiene dos niveles: plántulas muertas, b1 y plántulas vivas, b2.

Con esta información anterior se procede a formar un cuadro de Ji-cuadrado similar a los vistos anteriormente:

**Cuadro 2. Cálculo de la P.E.  $\chi^2$  para número de plántulas muertas y vivas de los dos híbridos de maíz**

Variable A*B	Ob	Es	(Ob-Es)	(Ob-Es) <sup>2</sup> /Es
a1b1	26	29.65	-3.65	0.45
a1b2	72	68.30	+3.70	0.20
a2b1	33	29.35	+3.65	0.45
a2b2	64	67.70	-3.70	0.20
Sumatorias	195	195	0	$P.E. \chi^2 = 1.30$

La prueba estadística es:  $\chi^2=1.30$ . Este valor se compara con el valor en la tabla (Apéndice C) con un grado de libertad y al 5% de probabilidad. Esto es:  
 $\chi^2_{(1gl)(0.05)} = 3.84$

Como el valor de la P.E.  $\chi^2$  no excede el valor de  $\chi^2$  en la Tabla, la decisión es aceptar la  $H_0$  y rechazar la  $H_1$  al nivel de probabilidad del 5%. En otras palabras, los híbridos de maíz a1 y a2 presentan niveles de resistencia estadísticamente iguales al marchitamiento. Es decir, las variables A y B son independientes.

Con el programa Statistix los resultados son:

**Two by Two Tables**

26	72	98
33	64	97
59	136	195

Fisher Exact Tests  
 Two-tail p 0,2779 (sum of probabilities for more extreme values)

Pearson's Chi-Square 1,30  
 P (Pearson's) 0,2549

**Comentario:** La Prueba Exacta de Fisher (Fisher Exact Test) es la mejor prueba, por cuanto proporciona la probabilidad real de los eventos observados. Para esto, el programa *Statistix* calcula y suma las probabilidades de los eventos más extremos que los observados. Sin embargo, los cálculos que requiere esta prueba toman mucho tiempo. El programa *Statistix* lo hace en experimentos con muestras en los que la suma total de observaciones, GT, sea igual o menor que 500 y sólo para cuadros  $2 \times 2$ . Para muestras más grandes la prueba de Ji-Cuadrado de Pearson provee una buena aproximación.

En el ejercicio anterior pudo verse que las probabilidades de las dos pruebas fueron bastante similares: Fisher, 27.79% y Pearson, 25.49%.

Otro ejemplo con muestras más grandes se presenta en el siguiente ejercicio:

**Ejercicio 8.6.2.** En una investigación (Luna, 1989<sup>17</sup>) se planteó la hipótesis nula de independencia entre las variables A “cobertura de mazorca” y B “método de emasculación”. Las hipótesis en términos de la cobertura de mazorca, en relación con el método de emasculación fueron:

$$H_0: a_1 = a_2$$

$$H_1: a_1 \neq a_2$$

$$\alpha = 1\%$$

En el siguiente cuadro se presentan los datos observados:

Cuadro 1. Número de mazorcas con buena y mala cobertura (variable A) en relación con dos métodos de emasculación (variable B)

Variable B

Variable A	Descogollado, b1	Despanojado, b2	Totales, $H_j$
Buena cobertura, a1	a1b1=1222	a1b2=1176	$H_1 = 2398$
Mala cobertura, a2	a2b1=51	a2b2=116	$H_2 = 167$
Totales, $C_i$	$C_1 = 1273$	$C_2 = 1292$	GT=2565

17 Luna, I. 1989. Efecto de densidad de siembra y métodos de emasculación en la producción de semilla de tres híbridos comerciales de maíz. Tesis. Escuela Agrícola Panamericana. Honduras, C.A.

Los valores esperados se encuentran con la fórmula utilizada en el ejercicio anterior. Estos valores son:  $a1b1=1190.1$ ,  $a1b2=1207.9$ ,  $a2b1=82.9$ ,  $a2b2=84.1$ . (El número de grados de libertad es uno, por lo que solo hubiera bastado calcular un valor esperado y por diferencia encontrar los otros valores.)

El valor de la prueba Ji-cuadrado se puede calcular directamente de esta manera:

$$P.E.\chi^2 = \frac{(1222-1190.1)^2}{1190.1} + \frac{(1176-1207.9)^2}{1207.9} + \frac{(51-82.9)^2}{82.9} + \frac{(116-84.1)^2}{84.1} = 26.04^{**}$$

El valor de Ji-cuadrado en la tabla, con un grado de libertad y alfa 0.01 es:  
 $\chi_{1gl,0.01}^2 = 6.63$

Este resultado nos indica que, estadísticamente, hay evidencias de que el método de emasculación influye sobre la cobertura de mazorca. Al descogollar se obtienen menos mazorcas con mala cobertura que al despanojar. Consecuentemente, se rechaza la hipótesis nula al nivel alfa del 1%.

El mismo análisis con *Statistix*:

**Two by Two Tables**

1222	1176	2398
51	116	167
1273	1292	2565

Pearson's Chi-Square 26.04

P (Pearson's) 0.0000

Note que la probabilidad de haber encontrado un valor tan extremo como 26.04 sólo por azar es igual o menor que 9/100000, es decir: ( $P \leq 0.00009$ )

## 8.7. La prueba de Ji – Cuadrado en encuestas

La prueba de Ji – Cuadrado puede aplicarse también para el análisis de datos resultantes de encuestas. Sin embargo, como en todos los casos en estadística

inferencial, la decisión de qué prueba usar debe hacerse en la planificación inicial del experimento. Veamos a continuación un ejemplo:

**Ejercicio 8.7.1.** En una investigación se quiso determinar si tres tipos de queso fabricados por una empresa tenían igual aceptación en mercados de Guayaquil que en mercados de Riobamba. Para esto se preguntó la preferencia a personas que se acercaron a comprar en puestos de venta de las dos ciudades. Los resultados fueron (número de personas que indicaron preferencia):

Cuadro 1. Preferencias de tres tipos de queso en dos ciudades del Ecuador

Variable B: tipos de queso

Variable A: ciudad	Queso 1=b1	Queso 2=b2	Queso 3=b3	Totales, $H_j$
Riobamba = a1	a1b1=56	a1b2=28	a1b3=46	$H_1 = 130$
Guayaquil = a2	a2b1=68	a2b2=62	a2b3=75	$H_2 = 205$
Totales, $C_i$	$C_1 = 124$	$C_2 = 90$	$C_3 = 121$	GT = 335

Si las preferencias en las dos ciudades fueran similares, entonces se podría decir que los tres tipos de queso tuvieron igual aceptación. Esta es la hipótesis nula. La hipótesis alterna indicaría que por lo menos uno de los quesos tuvo mayor aceptación en una ciudad en comparación con la otra. La prueba generalmente se la hace al nivel alfa del 5%.

$$H_0: a1 = a2$$

$$H_1: a1 \neq a2$$

$$\alpha = 5\%$$

$$\text{Los grados de libertad son: } g, l. = (\text{ciudades} - 1) * (\text{tipos de queso} - 1) = 2$$

Los valores esperados son:

$$Es(a1b1) = 124 * 130/335 = 48.12$$

$$Es(a1b2) = 90 * 130/335 = 34.93$$

Como los totales de los valores esperados deben ser iguales a los totales de los valores observados, basta con dos valores esperados (dos grados de libertad)

para, por diferencia, encontrar todos los valores. Con esta información procedemos a encontrar el valor de la  $PE\chi^2$

**Cuadro 2. Cálculo de la P.E.  $\chi^2$  para preferencias de tipos de queso (b1, b2, b3) en las ciudades de Riobamba (a1) y Guayaquil (a2)**

Combinaciones	Ob	Es	(Ob-Es)	$\frac{(Ob-Es)^2}{Es}$
a1b1	56	48.12	+7.88	1.29
a1b2	28	34.93	-6.93	1.37
a1b3	46	46.95	-0.95	0.02
a2b1	68	75.88	-7.88	0.82
a2b2	62	55.07	+6.93	0.87
a2b3	75	74.05	+0.95	0.01
Sumas	335	335	0	4.38

El valor de Ji-Cuadrado calculado es igual a 4.38. Este valor se compara con el valor en la Tabla de Ji-Cuadrado (Apéndice C) al 5% y con 2 grados de libertad que es 5.99. Como el valor calculado no supera al valor en la tabla, la conclusión es aceptar la  $H_0$ . Es decir, no hay suficiente evidencia de que las preferencias por esos tipos de queso sean distintas en Riobamba y Guayaquil.

Con el programa *Statistix*:

ciudad		tipoqueso			
		1	2	3	
1	Observed	56	28	46	130
	Expected	48.12	34.93	46.96	
	Cell Chi-Sq	1.29	1.37	0.02	
2	Observed	68	62	75	205
	Expected	75.88	55.07	74.04	
	Cell Chi-Sq	0.82	0.87	0.01	
		124	90	121	335

Overall Chi-Square 4.38

P-value 0.1116 (esto es igual a una probabilidad del 11.16%)

Degrees of Freedom 2

Como la probabilidad es superior al 5% se concluye que no hay suficiente evidencia para rechazar la hipótesis nula.

**Comentario:** La probabilidad de haber obtenido solo por azar un valor de Ji-Cuadrado igual a 4.38 es igual a 11.16%. Observe que en este caso, el programa NO efectúa la **Prueba Exacta de Fisher**. Solo lo hace para cuadros  $2 \times 2$  (el cuadro de este ejemplo es un  $3 \times 2$ ) y un número total de observaciones (GT) de 500 o menos.

*“El valor de  $P = 0.05$ , o 1 en 20, es conveniente tomarlo como un límite para juzgar si una desviación debe considerarse significativa o no. Efectos reducidos podrían no notarse si la muestra es pequeña, pero ninguna disminución del estándar de significancia corregirá esta dificultad”. Sir Ronald Fisher<sup>18</sup>*

## 9.1 . ¿Qué es la distribución F de Fisher?

La distribución F de Fisher, llamada así por el investigador que la desarrolló, es un tipo de distribución que se origina a partir de las varianzas. En este sentido, la distribución F tiene relación con la distribución Ji-Cuadrado. La distribución F se emplea para realizar pruebas que tienen que ver con las varianzas y también con las medias de dos o más poblaciones.

Si tenemos dos poblaciones A y B y extraemos una muestra de cada una, la proporción de las varianzas de las muestras:

$$F = \frac{s_A^2}{s_B^2} \quad (34)$$

sigue la distribución F si las poblaciones tienen una distribución normal y las varianzas  $\sigma_A^2$  y  $\sigma_B^2$  son iguales. La distribución F tiene una forma asimétrica similar a la de  $\chi^2$ . Su forma exacta varía de acuerdo con los grados de libertad de las dos muestras:  $\nu_A$ , que indica los grados de libertad de la muestra A y  $\nu_B$ , los grados de libertad de la muestra B. Las expresiones  $n_A$  y  $n_B$  son los tamaños respectivos de las muestras.

Como no puede haber varianzas negativas los límites de  $F = \frac{s_A^2}{s_B^2}$  son  $0 \leftrightarrow \infty$

---

18 Estadístico y biólogo inglés. [https://www.brainyquote.com/authors/ronald\\_fisher](https://www.brainyquote.com/authors/ronald_fisher)



Si  $S_A^2 = 0$  y  $S_B^2 \geq 0$  :  $F = 0$

Si  $S_A^2 > 0$  y  $S_B^2 = 0$  :  $F = \infty$

Si  $S_A^2 = S_B^2$ , entonces  $F = 1$ . Este es el punto central de la distribución, 50% de los valores F para cada lado, como se indica en el siguiente gráfico:

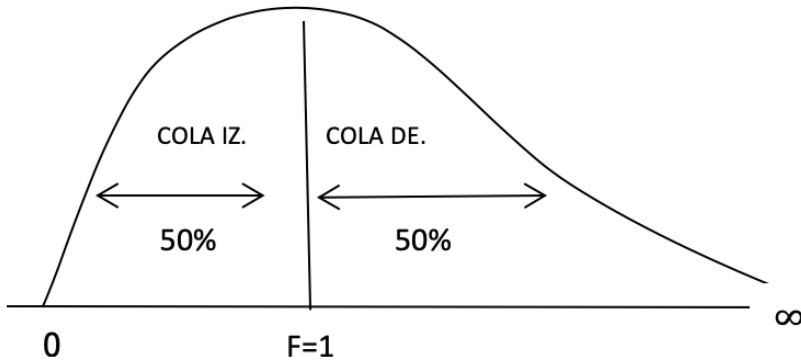


Gráfico 9.1.1. Los límites de la distribución F

La notación  $F_{(v_A, v_B)(\alpha)}$  indica un valor en la tabla F (Apéndice D) con  $v_A$  grados de libertad para la varianza del numerador  $S_A^2$  y  $v_B$  para la varianza del denominador;  $\alpha$  es el nivel de significación para la cola derecha.

Como la Tabla F presenta únicamente los valores críticos para la prueba de cola derecha, es conveniente colocar la varianza de la muestra numéricamente mayor en el numerador. Si se desea calcular el valor crítico para la cola izquierda se aplica esta relación:

$$F_{(v_B, v_A)(1-\alpha)} = \frac{1}{F_{(v_A, v_B)(\alpha)}} \quad (35)$$

Pero en la práctica no hace falta ya que la  $H_1: \sigma_A^2 < \sigma_B^2$  (prueba de cola izquierda) puede transformarse a prueba de cola derecha si  $H_1: \sigma_B^2 > \sigma_A^2$ .

El procedimiento para la prueba de las hipótesis es similar al visto en otros casos. La prueba estadística F es:

$$P.E.F = \frac{S_{Mayor}^2}{S_{menor}^2} \quad (36)$$

**Ejercicio 9.1.1.** En el mejoramiento vegetal se avanza más rápidamente cuando hay mayor variación en una característica deseada. Pruebe la hipótesis nula que la varianza del peso hectolítrico de la cruz A de trigo (Población A) es igual a la varianza de la cruz B (Población B):

$$H_0: \sigma_A^2 = \sigma_B^2$$

$$H_1: \sigma_A^2 \neq \sigma_B^2$$

$$\alpha = 10\%$$

Para esto, se extrajeron las siguientes muestras al azar:

Cuadro 1. Peso hectolítrico de muestras de trigo de dos cruza, A y B

Estadísticas	Muestra cruz A	Muestra cruz B
	83	73
	70	67
	80	75
	69	70
	70	74
	68	69
	82	74
	68	70
	69	-
$N$	9	8
$\bar{Y}$	73.22	71.50
$S^2$	41.19	8.29

$$P.E.F = \frac{S_A^2}{S_B^2} = \frac{41.19}{8.29} = 4.96$$

En el Apéndice D se encuentra el valor en la tabla F (con 8 y 7 grados de libertad y  $\alpha/2 = 0.05$ ). Este valor es:  $F_{(8y7g,l)(0.05)} = 3.73$

Como la prueba estadística F es igual a 4.96 y mayor que el valor F de la tabla, la decisión es rechazar la  $H_0$  que las dos varianzas de las poblaciones son iguales, al 10% nivel de significación.

Con el programa *Statistix* podemos encontrar la probabilidad de un valor de F tan alto y mayor que 4.96. Esta probabilidad es  $P = 0.0242$

En conclusión, se espera avanzar más rápidamente en el mejoramiento del peso hectolítrico con la población A, ya que presenta una varianza significativamente mayor que la varianza de la población B.

Otro ejemplo del uso de la prueba estadística F para la igualdad de dos varianzas se presenta en el siguiente ejercicio:

**Ejercicio 9.1.2** En un experimento se evaluó el efecto de una hormona en la sincronización del estro o celo en cerdas. Para esto se escogieron 20 cerdas de iguales características y al azar se asignaron 10 para ser inyectadas con la hormona (Tratamiento A). Las otras 10 fueron expuestas a la presencia de machos (Tratamiento B). La variable de respuesta fue el número de días transcurridos, a partir de los tratamientos, en que las cerdas entraron en celo. Las hipótesis fueron:

$$H_0: \sigma_A^2 = \sigma_B^2$$

$$H_1: \sigma_A^2 < \sigma_B^2$$

$$\alpha = 0.05$$

Cuadro 1. Número de días en entrar en celo de diez cerdas por tratamiento

Trat	Días al celo										$S^2$
A	6	3	4	7	5	4	3	6	4	5	1.79
B	4	8	10	6	12	9	7	8	5	13	8.40

$$PEF = \frac{S_B^2}{S_A^2} = \frac{8.40}{1.79} = 4.69$$

En la tabla F del Apéndice D se encuentran los valores críticos al 5% y al 1%, con nueve grados de libertad para el numerador y nueve para el denominador. Estos valores son respectivamente 3.18 y 5.35. Como el valor de la PEF es igual a 4.69, valor entre 3.18 y 5.35, la decisión es rechazar la hipótesis nula al nivel del 5% de significación. Con el programa *Statistix* se puede encontrar la probabilidad exacta de haber obtenido sólo por azar un valor tan extremo como 4.69. Este valor es  $P = 0.015$ , es decir 1.5%.

La conclusión es que con la hormona (tratamiento A) se logró una mayor sincronización del estro en comparación con el tratamiento B. Esta respuesta fue estadísticamente significativa ( $P = 0.015$ ).

**Comentario:** En el ejemplo del ejercicio anterior lo que se buscaba era reducir la variación del período en que las cerdas del experimento entraban en celo. En consecuencia, lo procedente es una prueba de varianzas y no una prueba de medias.

Hemos visto una de las aplicaciones de la distribución F. Esta distribución es, como se presentará más adelante, el fundamento de los *Análisis de Varianza* que se utilizan para comparar medias de diferentes tratamientos.

## 9.2. Pruebas $t$ para la comparación de dos medias

Se ha visto anteriormente las pruebas que se realizan para evaluar la igualdad o diferencia de una media poblacional con relación a una media hipotética. En esta sección se estudiarán las pruebas para evaluar la igualdad o diferencia de medias de dos poblaciones. Estas pruebas se basan en la distribución *t de Student*. Las hipótesis generalmente son:

$$H_0: \mu_A = \mu_B$$

$$H_1: \mu_A \neq \mu_B$$

$\alpha$

En donde  $\mu_A$  y  $\mu_B$  representan las medias de las poblaciones A y B respectivamente y  $\alpha$  es el nivel de significación.

Es un requerimiento importante para esta prueba el suponer, sobre la base de algún conocimiento previo, que las poblaciones A y B están normalmente distribuidas. Si hay evidencia de lo contrario el tamaño de las muestras debe ser alrededor de 30. Esto último, de acuerdo con el Teorema del Límite Central.

Las pruebas estadísticas  $t$  (*PEt*) para evaluar la igualdad o diferencia entre dos medias de poblaciones distintas, A y B, se utilizan en experimentos en todo campo de investigación.

Para continuar con esta sección es importante definir los siguientes conceptos:

- **Tratamiento.** Es la variable independiente cuyo efecto se busca conocer. En las pruebas  $t$  hay dos tratamientos, generalmente denominados 1 y 2 o A y B o cualquier otra designación. Cuando los tratamientos se aplican a las unidades experimentales se obtienen las muestras de las dos poblaciones en estudio.
- **Variable de Respuesta.** Es la variable dependiente, es decir la expresión del efecto de la variable independiente. Por ejemplo, si la variable independiente consiste de diferentes dosis de un fertilizante, la variable dependiente será el rendimiento de una variedad u otras características que pueden ser afectadas por el fertilizante (días a flor, altura de planta, resistencia a enfermedades, etc.)
- **Unidad Experimental.** Es la entidad o sujeto que recibe un determinado tratamiento y que da origen a una o varias observaciones. Dependiendo del tipo de experimento una unidad experimental puede ser una parcela en el campo, una planta en el invernadero, un plato Petri, un árbol, un grupo de árboles, una jaula de pollos, un estudiante en un experimento educativo, una persona voluntaria en un experimento médico, etc.
- **Par o Bloque.** El término “par” en las pruebas  $t$  se refiere a un grupo de dos unidades experimentales que se presume son homogéneas. Mediante sorteo, las unidades experimentales reciben el tratamiento 1 o el tratamiento 2. En experimentos con más de dos tratamientos, el conjunto de unidades experimentales que se suponen son homogéneas, se denomina “bloque”. El “par” es un “bloque” con dos unidades experimentales.

En las Pruebas Estadísticas  $t$  hay dos casos:

1) Cuando en la planificación del ensayo se presume que todas las unidades experimentales que dan origen a las muestras de A y de B son homogéneas. Como ejemplo, un campo experimental en que las parcelas sean parecidas en términos de textura del suelo y fertilidad. Así, en un sorteo totalmente al azar, un número de parcelas recibirán el tratamiento A y otro número de parcelas el tratamiento B. En este caso nos referimos como “**prueba de  $t$  para observaciones NO pareadas**” y

2) Cuando las unidades experimentales que dan origen a las muestras de A y de B presentan niveles de heterogeneidad, pero sin embargo podemos agruparlas en pares por alguna similitud entre ellas. Como ejemplo, en un experimento con 20 cerditos de diferente tamaño será necesario agruparlos en 10 pares por su similitud en el peso. Así, en un par tendremos cerditos con pesos entre 5 y 6 Kg

y en otro par a lo mejor cerditos con pesos entre 10 y 12 Kg, etc. Cuál cerdito de cada par recibe el tratamiento A o el B será al azar. En este caso el ensayo se llama “**prueba de  $t$  para observaciones pareadas**”.

### 9.3. Prueba de $t$ para observaciones NO pareadas

Esta es una prueba muy usada y es comparable a un **Análisis de Varianza de un Diseño Completamente al Azar** con dos tratamientos, como se verá más adelante. Los dos tratamientos representan dos poblaciones distintas, como por ejemplo: dos dosis de fertilizante, dos productos químicos, dos dietas de engorde, dos variedades de maíz, dos distancias de siembra, la aplicación o no de un insumo, etc.

Una presunción importante para esta prueba es que las varianzas de las dos poblaciones sean homogéneas. Sin embargo, cuando las observaciones NO son pareadas, existe el riesgo de que las dos poblaciones en estudio tengan diferentes varianzas. En este caso, la comparación de medias sería estadísticamente inválida si no se tomaran medidas al respecto.

En los Análisis de Varianza una medida correctiva cuando las varianzas no son homogéneas es la transformación de los datos a escalas que reduzcan la variación existente, pero esto se verá en capítulos posteriores.

En las pruebas  $t$  para observaciones no pareadas, si las varianzas poblacionales son estadísticamente iguales se puede promediar las varianzas de las dos muestras, tomando en cuenta el número de observaciones de cada una. Para decidir si las varianzas son estadísticamente iguales se requiere realizar una prueba  $F$  (ver sección 9.1). Si se acepta la hipótesis nula que las varianzas son iguales, entonces se puede incluir en la fórmula de la Prueba Estadística  $t$  ( $PEt$ ) el promedio de las dos varianzas. Este promedio se denomina **varianza ponderada**,  $S_p$ . (La ponderación se realiza en función del número de observaciones de cada varianza.)

Si las varianzas son diferentes no se deben promediar. En este caso, las varianzas intervienen en forma individual en la fórmula de la Prueba Estadística, que para diferenciar de la anterior se llama Prueba Estadística T ( $PET$ ). Esto tiene, sin embargo, un costo en términos de grados de libertad y consecuentemente una menor precisión en los resultados.

El procedimiento para efectuar la prueba de las hipótesis planteadas a un determinado nivel de significación es como sigue:

- Se toma una muestra al azar de cada una de las poblaciones en cuestión y se calcula la media y la varianza de cada muestra.
- Se decide si las varianzas de las dos poblaciones son estadísticamente iguales. Para esto, mediante la distribución F se prueba las siguientes hipótesis:

$$H_0: \sigma_A^2 = \sigma_B^2$$

$$H_1: \sigma_A^2 \neq \sigma_B^2$$

El nivel de significación que se usa generalmente es el mismo que se planteó inicialmente para la prueba de las medias. Esto conduce a dos casos:

**Caso 1.** Si se acepta la  $H_0: \sigma_A^2 = \sigma_B^2$ , entonces se emplea la siguiente prueba estadística:

$$t = \frac{\bar{Y}_A - \bar{Y}_B}{\sqrt{S_p^2 \left( \frac{1}{n_A} + \frac{1}{n_B} \right)}} \quad (37)$$

Los grados de libertad para esta prueba son:  $n_A + n_B - 2$

El término  $S_p^2$  se denomina **Varianza Ponderada** y es un estimador de la varianza común de las dos poblaciones, en la que se toma en cuenta el tamaño de cada muestra. Se calcula con la siguiente fórmula:

$$S_p^2 = \frac{(n_A - 1)S_A^2 + (n_B - 1)S_B^2}{n_A + n_B - 2} \quad (38)$$

Si  $n_A = n_B$ , entonces la fórmula se reduce a:  $S_p^2 = \frac{S_A^2 + S_B^2}{2}$  (39)

**Caso 2.** Si se rechaza la  $H_0: \sigma_A^2 = \sigma_B^2$ , entonces no hay una varianza común para las dos poblaciones y no se puede usar la fórmula anterior. En este caso se usa una aproximación a la distribución  $t$ , denominada por conveniencia distribución  $T$ . La fórmula de la prueba estadística  $T$  es:

$$T = \frac{\bar{Y}_A - \bar{Y}_B}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}} \quad (40)$$

La mayor dificultad para esta prueba es el encontrar el número de grados de libertad. La fórmula para esto es:

$$g.l.T = \frac{\left(\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}\right)^2}{\frac{\left(\frac{S_A^2}{n_A}\right)^2}{n_A-1} + \frac{\left(\frac{S_B^2}{n_B}\right)^2}{n_B-1}} \quad (41)$$

La expresión  $g.l.T$  indica el número de grados de libertad para la prueba. Como probablemente se obtendrá un número con una fracción decimal, el resultado debe aproximarse al entero más cercano posible. Por lo demás, el valor crítico se lo encuentra en la misma Tabla  $t$  al nivel de significación deseado y con  $g.l.T$  grados de libertad.

- Una vez que se encuentra el valor de la prueba estadística, en cualquiera de los dos casos,  $t$  o  $T$ , el siguiente paso es determinar el valor crítico.
- Compare la prueba estadística con el valor crítico. Si la prueba estadística cae dentro de las áreas críticas rechaza la  $H_0$ .

En el siguiente ejercicio se presenta un ejemplo de esta prueba:

---

**Ejercicio 9.3.1.** Un investigador desea evaluar el rendimiento de dos variedades híbridas de maíz, A y B. Para esto se plantea las siguientes hipótesis:

$$H_0: \mu_A = \mu_B$$

$$H_1: \mu_A \neq \mu_B$$

$$\alpha = 5\%$$

El terreno del que dispone para su experimento es bastante homogéneo. Así que decide marcar 20 parcelas en el campo (unidades experimentales). A continuación efectúa un sorteo totalmente al azar para sembrar 10 parcelas con la variedad A y 10 parcelas con la variedad B. Los resultados en kg por parcela fueron:

**Cuadro 1.** Estadísticas y rendimiento de la variedad A y B en kg por parcela

Estadísticas	Muestra Variedad A	Muestra Variedad B
	9.8	9.3
	8.1	13.2
	9.2	12.2
	10.1	11.6



	8.4	10.0
	8.8	9.2
	9.6	10.4
	7.4	13.6
	8.2	10.5
	10.4	11.2
$\sum Y_i$	90.0	111.2
$n$	10	10
$\bar{Y}$	9.00	11.12
$(n-1)S^2 = SC$	8.64	21.04
$S^2$	0.96	2.34

Una vez que se han calculado las estadísticas anteriores debe decidirse si se usa la prueba  $t$  (*minúscula*) o la prueba  $T$  (*mayúscula*). Para esto se plantea las siguientes hipótesis con relación a las varianzas de la población:

$$H_0: \sigma_A^2 = \sigma_B^2$$

$$H_1: \sigma_A^2 \neq \sigma_B^2$$

$$\alpha = 5\%$$

La prueba estadística  $F$  es:

$$P.E.F = \frac{S_B^2}{S_A^2} = \frac{2.34}{0.96} = 2.44$$

Note que la varianza numéricamente mayor va en el numerador. El valor crítico en la Tabla F es:  $F_{(9,9)(0.025)} = 4.03$

Como la  $P.E.F$  es igual a 2.99 y no excede al valor crítico de 4.03, la decisión es aceptar la  $H_0$  que las dos varianzas poblacionales son estadísticamente iguales. Consecuentemente, la prueba estadística que debe usarse es  $t$ .

Luego es necesario calcular la varianza ponderada o varianza común. Como  $n$  es igual para las dos muestras, la varianza común es el promedio de las dos varianzas:

$$S_p^2 = \frac{0.96 + 2.34}{2} = 1.65$$

Con este dato se procede a calcular La prueba estadística  $t$  con la fórmula (37), esto es:

$$PEt = \frac{11.12-9.00}{\sqrt{1.65\left(\frac{1}{10}+\frac{1}{10}\right)}} = \frac{2.12}{0.57} = 3.69$$

La prueba es de dos colas, de acuerdo con lo planteado en las hipótesis. Los Valores Críticos, al 2.5% y con 18 grados de libertad ( $g.l. = n_A + n_B - 2 = 18$ ), en la tabla t son:  $t_{(18)(0.025)} = \pm 2.10$

La  $PEt$  cae en la región crítica; por lo tanto, se rechaza la  $H_0$  al 5% nivel de significación. En otras palabras, la variedad B rinde significativamente más que la variedad A.

Con el programa *Statistix*, la hipótesis nula que se prueba es:

$H_0: \mu_D = 0$ , donde  $\mu_D$  es la media de la diferencia. Esto es igual a lo del ejercicio 9.3.1, ya que si  $\mu_A = \mu_B$ , entonces  $\mu_A - \mu_B = 0 = \mu_D$

Estos son los resultados del análisis con *Statistix*:

**Two-Sample T Tests for VarB vs VarA**

Variable	N	Mean	SD	SE
VarB	10	11.120	1.5288	0.4835
VarA	10	9.0000	0.9787	0.3095
Difference	2.1200	1.2836	0.5740	

**T-Tests for Mean Difference**

Null Hypothesis: difference = 0

Alternative Hyp: difference <> 0

95% CI for Difference						
Method	Variances	DF	T	P	Lower	Upper
Pooled	Equal	18	3.69	0.0017	0.9140	3.3260
Satterthwaite	Unequal	15.3	3.69	0.0021	0.8987	3.3413

Homogeneity of Variances			
	DF	F	P
Folded F Test	9,9	2.44	0.1000

Cases Included 20 Missing Cases 0

**Comentario:** Note que el programa *Statistix* presenta las dos alternativas: Cuando las varianzas son iguales (Pooled method) y cuando las varianzas no son

iguales (Satterhwaite method). Es decir, el investigador debe escoger la prueba más idónea. Como se decidió que las varianzas sí son iguales, este es el método que usamos en este caso (Pooled method). El valor de la prueba t es -3.69 y la probabilidad 0.0017. Como es inferior al 0.025, es decir 2.5%, se concluye que las medias de las dos variedades son estadísticamente diferentes.

Note también que el programa calcula el intervalo de confianza para la media de la población de diferencias ( $\mu_D$ ). Este intervalo, con el 95% de confianza, nos indica el límite inferior (0.91 kg) y el superior (3.32 kg). El intervalo no contiene al cero, que es el valor de la hipótesis nula. Esto obviamente se espera cuando la hipótesis nula es rechazada.

#### 9.4. ¿Por qué la necesidad de dos pruebas: t y T?

La pregunta tiene que ver con las varianzas. Si las varianzas de las dos poblaciones son significativamente diferentes, puede llegarse a conclusiones erradas con relación a las medias de esas poblaciones. En otras palabras, es un requisito para la prueba t la homogeneidad de las varianzas. Si las varianzas son distintas, con la prueba T y el menor número de grados de libertad se compensa por esa diferencia. Se verá más adelante en los análisis de varianza, que se usan principalmente para comparar tres o más medias de tratamientos, los procedimientos a seguir cuando hay evidencia de diferencias significativas entre las varianzas de las poblaciones.

A continuación se presenta un ejercicio en el que se emplea la prueba T.

**Ejercicio 9.4.1.** Con los datos del **ejercicio 9.1.1**, pruebe la hipótesis nula que las medias del peso hectolítrico de las dos poblaciones de trigo son iguales:

$$H_0: \mu_A = \mu_B$$

$$H_1: \mu_A \neq \mu_B$$

$$\alpha = 5\%$$

Cuadro 1. Datos del Ejercicio 9.1.1

Estadísticas	Muestra A Peso Hectolítrico	Muestra B Peso Hectolítrico
n	9	8
$\bar{Y}$	73.22	71.50
$S^2$	41.19	8.29

Como en el ejercicio 9.1.1 se concluyó que  $\sigma_A^2 \neq \sigma_B^2$ , entonces, para probar las hipótesis acerca de las medias poblacionales A y B, se debe usar la prueba estadística  $T$  (fórmula 40):

$$PET = \frac{73.22 - 71.50}{\sqrt{\frac{41.19}{9} + \frac{8.29}{8}}} = 0.73$$

Los grados de libertad para esta prueba son:

$$g.l.T = \frac{(41.19/9 + 8.29/8)^2}{\frac{(41.19/9)^2}{8} + \frac{(8.29/8)^2}{7}} = 11.4$$

Con esta fórmula rara vez se encontrará un número entero para los grados de libertad. Por lo tanto debemos tomar el valor más cercano, en este caso  $g.l.T = 11$

El valor crítico se lo encuentra en la Tabla  $t$  (Apéndice B). El valor, con 11 grados de libertad y al nivel alfa del 5% (0.025 a cada extremo, por ser prueba de dos colas), es:

$$t_{(11)(0.025)} = \pm 2.201$$

Como la  $PET$  es igual a 0.73 no excede al V.C. 2.201, la decisión es aceptar la  $H_0$  de que las medias del peso hectolítrico de las dos poblaciones son estadísticamente iguales al 5% nivel de significación.

Con *Statistix* los resultados, aunque se presentan en forma diferente, nos llevan a las mismas conclusiones:

#### Two-Sample T Tests for VarA vs VarB

Variable	N	Mean	SD	SE
V001	9	73.222	6.4183	2.1394
V002	8	71.500	2.8785	1.0177
Difference	1.7222	5.0830	2.4699	

#### T-Tests for Mean Difference

Null Hypothesis: difference = 0

Alternative Hyp: difference <> 0

Method	Variances	DF	T	P
Pooled	Equal	15	0.70	0.4963
Satterthwaite	Unequal	11.4	0.73	0.4820

Homogeneity of Variances	DF	F	P
Folded F Test	8, 7	4.97	0.0241

Observe en las últimas líneas la prueba para la homogeneidad de las varianzas. El valor de F es significativo al 5% ( $P=0.0241$ ). Por lo tanto las varianzas no son estadísticamente iguales (como ya se demostró en el ejercicio 9.1.1) y debe usarse la prueba T para evaluar las medias de las dos poblaciones. Esto corresponde al método “Satterthwaite” (tercera línea desde abajo). Los grados de libertad son 11.4 y  $T = 0.73$ . La probabilidad es igual a 48.2% ( $P=0.4820$ ), por lo que se acepta la hipótesis nula.

## 9.5. La prueba de $t$ para observaciones pareadas

Cuando las unidades experimentales son homogéneas, la prueba de  $t$  para observaciones NO pareadas es lo más indicado. Esto vimos en los casos anteriores. Sin embargo, algunas veces es difícil encontrar unidades experimentales homogéneas. En experimentos de campo las unidades experimentales, llamadas parcelas, difieren en varios aspectos como, por ejemplo: textura y estructura del suelo, inclinación del terreno, contenido de materia orgánica, nivel de fertilidad, cantidad de microorganismos beneficiosos y perjudiciales, manejo anterior, etc. Es decir, difieren en aspectos físicos, químicos y biológicos. En experimentos en que las unidades experimentales son animales, las diferencias son atribuibles a raza, sexo, edad, nutrición y en general a aspectos genéticos y ambientales.

Por lo anterior, para mejorar la precisión en experimentos en los que se conoce que hay variación entre las unidades experimentales, es conveniente agrupar éstas en pares. Las unidades experimentales que forman un par deben ser lo más similares que sea posible. Veremos posteriormente como el concepto de par se extiende al de “bloque”. Un bloque contiene dos o más unidades experimentales homogéneas. Un par es un bloque con dos unidades experimentales.

En la prueba  $t$  para observaciones pareadas las hipótesis que se plantean, con un nivel de significación  $\alpha$ , son:

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$

$\alpha$

El término  $\mu_d$  representa la media de una población formada por las diferencias entre las observaciones de las poblaciones A y B. Esto se explica mejor en el siguiente cuadro:

Cuadro 9.5.1. Organización de datos en el caso de observaciones pareadas

Par	Observaciones		Diferencias
	A	B	
1	$Y_{1A}$	$Y_{1B}$	$d_1 = Y_{1A} - Y_{1B}$
2	$Y_{2A}$	$Y_{2B}$	$d_2 = Y_{2A} - Y_{2B}$
3	$Y_{3A}$	$Y_{3B}$	$d_3 = Y_{3A} - Y_{3B}$
.	.	.	.
.	.	.	.
.	.	.	.
n	$Y_{nA}$	$Y_{nB}$	$d_n = Y_{nA} - Y_{nB}$

La prueba estadística es:

$$t = \frac{\bar{d} - \mu_d}{\sqrt{S_d^2/n}} \tag{42}$$

En la que:  $\mu_d = 0$ , de acuerdo con la hipótesis nula.

La media de las diferencias es:  $\bar{d} = \frac{\sum d_i}{n}$  (43)

La varianza de las diferencias es:  $S_d^2 = \frac{\sum d_i^2 - (\sum d_i)^2/n}{n-1}$  (44)

El error estándar de la diferencia de dos medias es:  $S_{\bar{d}} = \sqrt{S_d^2/n}$  (45)

La fórmula de la prueba estadística consecuentemente puede expresarse con mayor simplicidad de esta forma:  $t = \frac{\bar{d}}{S_{\bar{d}}}$  (46)

Para la validez de esta prueba  $t$  la única presunción necesaria es que las diferencias  $d_i$  sigan la distribución normal.

Para la asignación de los tratamientos A y B se realiza un sorteo individual dentro de cada par. Es decir, cualquiera de las dos unidades experimentales de cada par puede recibir, al azar, cualquiera de los dos tratamientos.

Por lo demás, el procedimiento para efectuar la prueba es similar al visto anteriormente.

El número de grados de libertad en este caso es igual al número de pares menos uno.

**Ejercicio 9.5.1.** Un investigador desea conocer si el rendimiento de una variedad de papa es afectado en forma diferente al emplear un nuevo sistema de manejo de malezas (A) en comparación con el sistema tradicional (B). Para el efecto dispone de un lote de terreno de 20 parcelas (unidades experimentales). Como sabe que el terreno que dispone no es homogéneo, marca cada par de parcelas en lotes contiguos (esto, porque por lo general áreas contiguas tienen suelos con características similares). Al final del ensayo de campo cosecha los 10 pares de parcelas y luego analiza las diferencias entre A y B. Las hipótesis son:

$$H_0: \mu_d = 0$$

$$H_1: \mu_d \neq 0$$

$$\alpha = 5\%$$

Note que  $\mu_d = 0$  es equivalente a  $\mu_A = \mu_B$ , ya que si las dos medias son iguales entonces la diferencia entre las dos es igual a cero. Los resultados experimentales se presentan en el cuadro siguiente:

Cuadro 1. Efecto de dos métodos de manejo de malezas, A y B, en el rendimiento de una variedad de papa, en kg por parcela

Par	Observaciones		Diferencias
	A	B	d
1	45	56	-11
2	54	61	-7
3	60	61	-1
4	58	62	-4
5	53	53	0
6	65	63	+2
7	50	56	-6
8	47	58	-11
9	38	44	-6
10	42	50	-8
n = 10	$\sum Y_A = 512$	$\sum Y_B = 564$	$\sum d = -52$

Ahora debemos encontrar  $\bar{d}$  y  $S_{\bar{d}}$  :

$$\bar{d} = \frac{\sum d_i}{n} = \frac{-52}{10} = -5.2$$

Para encontrar  $S_{\bar{d}}$  primero debemos encontrar  $S_d^2$  :

$$\sum d_i^2 = (-11)^2 + (-7)^2 \dots + (-8)^2 = 448$$

$$(\sum d_i)^2 = (-52)^2$$

$$S_d^2 = \frac{448 - (-52)^2 / 10}{9} = 19.73$$

(Note que, aunque la sumatoria de las diferencias  $\sum d_i$  puede ser negativa, la varianza siempre será positiva)

El valor  $S_{\bar{d}}$  es entonces:  $S_{\bar{d}} = \sqrt{\frac{19.73}{10}} = 1.4$

La prueba estadística  $t$  es:  $PEt = \frac{-5.2}{1.4} = -3.71$

El número de grados de libertad es:  $gl = 10 - 1 = 9$

Los valores críticos en la tabla  $t$  (prueba de dos colas) son:  $t_{(9)(0.025)} = \pm 2.26$

Como la prueba estadística excede al valor crítico absoluto, en otras palabras, cae dentro del área crítica, la decisión es rechazar la  $H_0$  que la media de la diferencia es igual a cero. Esto es, al 5% nivel de significación. En términos agronómicos y guiándonos por el signo de la sumatoria de las diferencias, se puede decir que el método de manejo de malezas B tiene un efecto significativamente mayor que el método A en el rendimiento de la papa, bajo las condiciones de este experimento.

Los resultados con *Statistix* son:

**Paired T Test for V001 - V002**

Null Hypothesis: difference = 0  
 Alternative Hyp: difference <> 0

Mean	-5.2000,	Std Error	1.4048
Mean - H0	-5.2000		
Lower 95% CI	-8.3778,	Upper 95% CI	-2.0222
T	-3.70,	DF	9
P	0.0049		

Cases Included 10    Missing Cases 0



Observe los valores de la prueba  $t$  (-3.70) y la probabilidad ( $P=0.0049$ ). Como la probabilidad es menor que el 0.05 se rechaza la Hipótesis nula ( $H_0$ ).

## 9.6. Intervalo de confianza para la diferencia de dos medias

Así como se estimaron intervalos de confianza para la media y varianza de una población, así también podemos estimar un intervalo de confianza para la media de la diferencia,  $\mu_d$ . Es importante recordar que  $\mu_d$  es un parámetro, que resulta de la diferencia de dos medias poblacionales:  $\mu_A - \mu_B = \mu_d$ . Para estimar un intervalo de confianza para  $\mu_d$  se emplea la siguiente fórmula que se deriva de las anteriores fórmulas:

$$P(d - t_{\alpha/2} * S_{\bar{d}} \leq \mu_d \leq d + t_{\alpha/2} * S_{\bar{d}}) = 1 - \alpha \quad (47)$$

Esta fórmula nos indica que la probabilidad de que  $\mu_d$  se encuentre entre los valores

$$d \pm t_{\alpha/2} * S_{\bar{d}} \text{ es igual a } 1 - \alpha.$$

---

**Ejercicio 9.6.1.** Encuentre el 90% intervalo de confianza para  $\mu_d$  con los datos del ejercicio anterior (ejercicio 9.5.1), en el que:

$$\bar{d} = -5.2 \quad S_{\bar{d}} = 1.4$$

Si se pide un intervalo de confianza al 90% el valor de alfa es  $\alpha = 10\%$  y  $\alpha/2 = 5\%$ . Consecuentemente el valor de  $t$  se lo encuentra en la Tabla  $t$  (Apéndice B) con 9 grados de libertad y al nivel de significación alfa del 5%. (También puede encontrarse el valor de  $t$ , más fácilmente, con la opción “probability functions” en el programa *Statistix*.) Este valor es igual a:

$$t_{(9)(0.05)} = \pm 1.83$$

Al reemplazar los valores en la fórmula tenemos:

$$P(-5.2 - 1.83 * 1.4 \leq \mu_d \leq -5.2 + 1.83 * 1.4) = 1 - 0.1 = 0.90$$

$$P(-7.76 \leq \mu_d \leq -2.64) = 90\%$$

Es decir que se puede indicar, con el 90% de confianza, que el valor del parámetro  $\mu_d$  se encuentra entre los límites -7.76 y -2.64. Note así mismo que la probabilidad de error es 10%.

**Ejercicio 9.6.2.** Con los datos del ejercicio 9.6.1 obtenga y compare entre ellos los intervalos de confianza al 80%, al 90%, al 95% y al 99%.

Debe tenerse en cuenta que  $\bar{d}$  es un estimador puntual de  $\mu_d$  y que ahora se trata de encontrar estimadores de intervalo, dentro de los cuales podamos afirmar, con una determinada probabilidad, que ahí se encuentra el parámetro  $\mu_d$

Las estadísticas que necesitamos son:

$$\bar{d} = -5.2 \quad S_{\bar{d}} = 1.4$$

Con el programa *Statistix* las respuestas son:

<b>Variable</b>	<b>Lo 80% CI</b>	<b>Up 80% CI</b>
difAB	-7.1428	-3.2572
<b>Variable</b>	<b>Lo 90% CI</b>	<b>Up 90% CI</b>
difAB	-7.7751	-2.6249
<b>Variable</b>	<b>Lo 95% CI</b>	<b>Up 95% CI</b>
difAB	-8.3778	-2.0222
<b>Variable</b>	<b>Lo 99% CI</b>	<b>Up 99% CI</b>
difAB	-9.7652	-0.6348

Como en verdad desconocemos el parámetro  $\mu_d$ , podemos observar que los intervalos de confianza se vuelven cada vez más amplios a medida que aumenta la probabilidad de que  $\mu_d$  se encuentre contenida dentro del intervalo.



*“Encontramos regresión a la media casi todos los días de nuestras vidas, por lo que deberíamos prepararnos, identificarla y no dejar que nos engañe”.* Gary Smith<sup>19</sup>

Regresión es el estudio de la posible relación de dependencia entre variables. Hasta el momento hemos visto aspectos relacionados con una sola variable Y, sea esta altura, peso, rendimiento, etc. En este capítulo se analizará la asociación que pueda existir entre una o más variables X, denominadas variables independientes y una variable Y, supuestamente la variable dependiente. Por lo general, tanto X como Y son variables cuantitativas. Ejemplos de X son: niveles de fertilización, distancias de siembra, horas luz en un invernadero, dosis de un producto, cambios de temperatura, etc. Ejemplos de Y son: rendimiento de un producto, altura de planta, peso del grano, contenido de azúcar, etc.

Los casos más comunes de estudio y análisis de regresión en las ciencias biológicas son tres:

1. *Regresión lineal*: Cuando los cambios en Y como respuesta a incrementos de X son relativamente constantes, en un determinado espacio de muestreo. El modelo de regresión solo incluye valores de X a la potencia 1.
2. *Regresión polinómica*: Cuando los cambios no son constantes, sino que se incrementan y/o disminuyen en un mismo espacio de muestreo. En este caso el modelo incluye valores de X a diferentes potencias:  $X^1, X^2, X^3 \dots X^n$ .
3. *Regresión múltiple polinómica*: En este caso el modelo de regresión incluye dos o más variables X (como ejemplo: niveles de nitrógeno, niveles de potasio, densidades de siembra, etc.). Este caso incluye los términos individuales, así como las interacciones entre variables:  $X_1, X_2, X_1^2, X_2^2, X_1X_2 \dots$  etc. Esto se verá en el capítulo de experimentos factoriales.

---

19 Economista estadounidense. [https://www.goodreads.com/author/quotes/86376.Gary\\_Smith](https://www.goodreads.com/author/quotes/86376.Gary_Smith)

## 10.1. Las técnicas de regresión y sus aplicaciones

Mediante las técnicas de regresión se puede:

- Determinar si la asociación de las variables en estudio es estadísticamente significativa o no. Es decir, ¿cambia Y en respuesta a cambios en la variable X? Si estos cambios son estadísticamente significativos se concluye que existe una respuesta lineal de Y cuando cambia X.
- Encontrar el punto en el que incrementos en X no causan cambios significativos en Y. Por ejemplo, si X corresponde a diferentes cantidades de un fertilizante llegará un punto en el que los rendimientos Y no se incrementan (o incluso disminuyen) en respuesta a mayores dosis de X.
- Conocer cuánto cambia Y en promedio cuando X cambia en una unidad. Por ejemplo, si se incrementa en una unidad la distancia de siembra X de un cultivo, cuánto en promedio sube o baja la producción Y.
- Predecir valores de Y en puntos de X no incluidos en el experimento. Supongamos que los valores de X en un experimento fueron: 2, 4, 6, 8 y 10. Si nos interesa podríamos predecir el valor de Y cuando  $X = 13$ , o cualquier otro valor de X, pero es importante que la interpretación del resultado tenga algún sentido biológico.

## 10.2. La relación de dependencia entre dos variables

En matemáticas la relación entre dos variables se expresa de la siguiente forma:  $Y = f(X)$

Lo anterior presupone que “*Y es una función de X*”, es decir que la variable Y depende de la variable X. Como ejemplo, suponga que el rendimiento de una variedad de arroz cultivado sin fertilización es igual a 2.0 t/ha y que por cada unidad de fertilizante (equivalente a, digamos, 50 kg de nitrógeno) el rendimiento sube 0.15 t/ha. Esta relación se expresa mediante la ecuación de la línea recta:

$$Y = b_0 + b_1X \quad (48)$$

El significado de  $b_0$  y  $b_1$  se explicará más adelante.

En el siguiente cuadro se presentan varios valores de X y los cambios correspondientes en Y:

Cuadro 10.2.1. Cambios de Y (producción de arroz por hectárea), en respuesta a X (unidades de fertilizante). Cada unidad equivale a 50 kg.

X en kg	X codificado	Y
0	0	$Y = 2.0 + 0.15 \cdot 0 = 2.00$
50	1	$Y = 2.0 + 0.15 \cdot 1 = 2.15$
100	2	$Y = 2.0 + 0.15 \cdot 2 = 2.30$
150	3	$Y = 2.0 + 0.15 \cdot 3 = 2.45$
200	4	$Y = 2.0 + 0.15 \cdot 4 = 2.60$

Puede notarse que la media  $\bar{X}$  es igual a 100 (2 en la escala codificada), y que la media  $\bar{Y}$  es 2.30. Esta relación funcional, perfecta, entre X y Y puede también presentarse en un gráfico, en el que la variable Y usualmente va en el eje de la ordenada y la variable X en el eje de la abscisa:

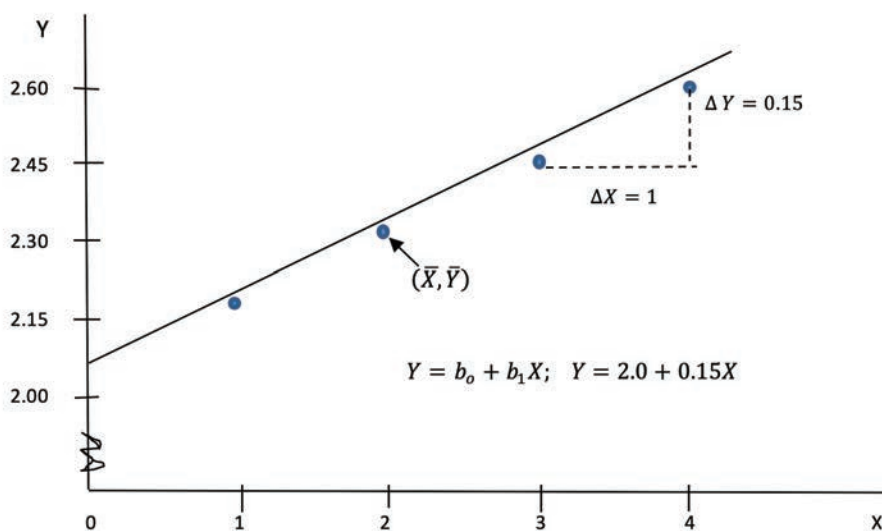


Gráfico 10.2.1. Representación gráfica teórica de la relación funcional entre Y (producción de arroz) y X (unidades de fertilizante).

### 10.3. Regresión lineal simple

La ecuación de la línea recta:  $Y = b_0 + b_1X$  describe la relación funcional que se observa en el Gráfico 10.2.1 en la que  $b_0$  es la intersección. Este es el punto en el que la línea o su prolongación corta el eje de las Y. En este ejemplo,  $b_0 = 2.0$ .

Note que en este punto  $X = 0$ . El otro valor de la ecuación es la pendiente o coeficiente de regresión  $b_1$ , que indica cuán inclinada está la línea de regresión con relación a la abscisa. En este ejemplo,  $b_1 = 0.15$ . La pendiente se define como el cambio en  $Y$  por unidad de cambio en  $X$ , como se puede ver en el gráfico anterior. Si  $b_1$  tiene un valor positivo la línea es ascendente, si  $b_1$  tiene un valor negativo la línea es descendente.

Observe en el gráfico que todos los puntos caen en la línea recta. Note también que la línea pasa por el punto de las medias ( $\bar{X}$ ,  $\bar{Y}$ ). Si se conoce la intersección y se conocen las medias puede entonces trazarse la línea recta descrita por la ecuación.

Sin embargo, una relación funcional perfecta como la descrita, por lo general, no se la encuentra en el campo de las ciencias biológicas. Por experiencia sabemos que el rendimiento del arroz o de otro cultivo depende de varios factores además del nivel de fertilización. Entre ellos: el clima de la zona, la constitución genética de la planta, características del suelo, disponibilidad de agua, densidad de siembra, incidencia de plagas, etc.

Por lo expuesto, en las ciencias biológicas estudiamos la relación estadística entre las variables  $X$  y  $Y$ . Al contrario de una relación funcional, una relación estadística no es perfecta y es variable, cualidad inherente a los sistemas biológicos y al proceso de muestreo. En una relación estadística las observaciones no caen necesariamente en la línea de la relación entre  $X$  y  $Y$ , que en lo sucesivo la llamaremos línea de regresión de  $Y$  sobre  $X$ .

## 10.4. El modelo de regresión lineal simple

Regresión es un campo bastante amplio en estadística. En esta sección nos concretaremos al modelo de regresión que se señala a continuación:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \quad (49)$$

Este modelo se llama simple y lineal. Es simple porque hay una sola variable independiente. Es lineal porque la variable independiente tiene únicamente exponente uno. Este modelo se llama también de primer orden. Los términos en el modelo son:

$Y_i$  Es el valor de la variable dependiente en la observación “i”

$\beta_0$  Es la intersección, es decir el punto en el que la línea de regresión corta el eje de las  $Y$  (eje de la ordenada). La intersección  $\beta_0$  es un parámetro.

$\beta_1$  Es el Coeficiente de Regresión o pendiente. Esto mide el cambio promedio en Y por cada unidad de cambio en X. La pendiente  $\beta_1$  es un parámetro.

$X_i$  Es el valor de la variable independiente en la observación "i".

$\varepsilon_i$  Es el error al azar proveniente de la medición de  $Y_i$  a un valor  $X_i$ . Por definición del modelo y como presunción importante para su validez, la media esperada de  $\varepsilon_i$  es igual a cero con varianza  $\sigma^2$ .

Las presunciones necesarias para la validez del modelo son las siguientes:

- Las varianzas  $\sigma^2$  de las poblaciones de  $Y_i$  son iguales, independientemente del nivel de  $X_i$
- Los términos del error  $\varepsilon_i$  no están correlacionados, es decir que la ocurrencia de un evento no influye en nada en otro evento.
- Las observaciones  $Y_i$  son independientes, es decir, no están correlacionadas entre ellas.

Aunque la variable independiente X puede medirse sin error, éste no es el caso para Y. Con cada nivel de X está asociada una población de observaciones  $Y_i$  que se presume siguen una distribución normal. Esto se representa en el gráfico siguiente:

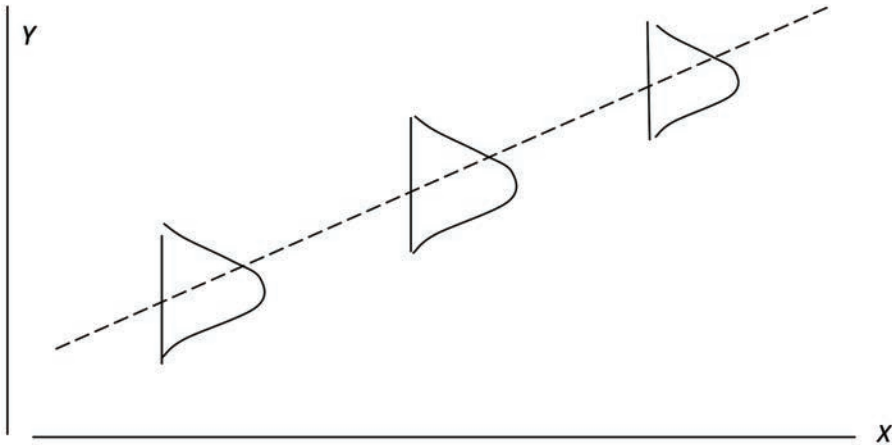


Gráfico 10.4.1. Ilustración gráfica del Modelo de Regresión Lineal Simple,  $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$  **bajo la suposición de un valor positivo de  $\beta_1$**



Como los parámetros son cantidades desconocidas, en estadística trabajamos con la siguiente ecuación que contiene los estimadores respectivos:

$$Y_i = b_0 + b_1X + e_i \quad (50)$$

Compare la fórmula (49) con la (50). Note la presencia de  $e_i$  en la fórmula (50). Este término es el error experimental. Esto indica que los puntos experimentales en un análisis de regresión NO caen necesariamente sobre la línea de regresión. El error experimental nos permite determinar, con cierto nivel de probabilidad, si la variable  $Y$  depende de la variable  $X$ .

En estadística nos interesa sobremanera el coeficiente de regresión  $b_1$ , también conocido como “la pendiente”. Es importante recordar que  $b_1$  proviene de una muestra y que es el estimador del parámetro  $\beta_1$ . A partir de los datos de la muestra se puede evaluar la hipótesis nula que  $\beta_1$  es igual a cero. Un valor positivo de  $b_1$  *estadísticamente significativo*, indica que  $Y$  aumenta a medida que  $X$  aumenta. Un valor negativo de  $b_1$  *estadísticamente significativo*, indica que  $Y$  disminuye a medida que  $X$  aumenta. En los dos casos se puede afirmar con un nivel de probabilidad previamente determinado que  $Y$  depende de  $X$ . Si el valor de  $b_1$  no es *estadísticamente significativo* al nivel de probabilidad de la prueba, se puede concluir que no hay suficiente evidencia para rechazar la hipótesis nula  $H_0: \beta_1 = 0$

Si estadísticamente se demuestra que  $\beta_1 = 0$ , entonces se puede señalar que el segmento de la línea recta es paralelo al eje de la abscisa, es decir que la variable  $Y$  aparentemente no sufre cambio alguno con el cambio de  $X$ . Lo indicado, siempre tomando en cuenta la probabilidad de cometer Error Tipo II (aceptar una hipótesis nula falsa).

## 10.5. Estimación de parámetros

Supongamos que nos interesa estudiar la relación entre las variables “edad” ( $X$ ) y “peso” ( $Y$ ) en una raza de cerdos. Como no conocemos los parámetros  $\beta_0$  y  $\beta_1$ , porque para eso tendríamos que trabajar con toda la población de cerdos de esa raza, tomamos una muestra al azar. Luego, con los pesos y edades de cada animal de la muestra buscamos el segmento de la línea recta que mejor se ajuste a las observaciones  $X$  y  $Y$ . Esta línea es la Línea de Regresión y tiene las siguientes características:

- Pasa por el punto  $(\bar{X}, \bar{Y})$ .
- La sumatoria de las desviaciones verticales de las observaciones  $(X_i, Y_i)$  con relación a la línea de regresión es igual a cero. Esto es:  $\sum e_i = 0$

- La sumatoria de los cuadrados de las desviaciones verticales es el valor mínimo posible. Esto es, a la línea de regresión se la encuentra basándose en el método de cuadrados mínimos, lo que nos permite estimar válidamente  $\beta_0$  y  $\beta_1$  a partir de una muestra.

La sumatoria de los cuadrados de las desviaciones se la conoce más comúnmente como la *Suma de Cuadrados del Error de Regresión* (SCE). La fórmula para encontrar este valor es la siguiente:

$$\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2 \quad (51)$$

Fórmula en la que:

$e_i$  representa el término del error, es decir, las desviaciones de los valores observados de  $Y_i$  con relación a los estimados  $\hat{Y}_i$

$\hat{Y}_i$  El valor de  $\hat{Y}_i$ , llamado “Y sombrero”, es el punto estimado de  $Y$  sobre la línea de regresión.

Los valores  $\hat{Y}_i$  se requieren para encontrar los errores  $e_i$  asociados con cada observación  $(Y, X)$ . Cada valor individual de  $\hat{Y}_i$  se lo encuentra con la siguiente fórmula:

$$\hat{Y}_i = b_0 + b_1 X \quad (52)$$

### **Las ecuaciones normales**

Las estadísticas  $b_0$  y  $b_1$  son los mejores estimadores de los parámetros  $\beta_0$  y  $\beta_1$ , respectivamente. Esto es así por cuanto  $b_0$  y  $b_1$  satisfacen los criterios de cuadrados mínimos. Puede demostrarse, mediante cálculo, que al estimar las estadísticas  $b_0$  y  $b_1$  se minimiza la Suma de Cuadrados del Error para cualquier conjunto de datos. La relación que se establece está dada por las siguientes ecuaciones simultáneas, denominadas *ecuaciones normales*:

$$\begin{aligned} \sum Y_i &= nb_0 + b_1 \sum X_i \\ \sum X_i Y_i &= b_0 \sum X_i + b_1 \sum X_i^2 \end{aligned} \quad (53)$$

A partir de las ecuaciones normales pueden encontrarse los estimadores  $b_0$  y  $b_1$ , como se demuestra en el siguiente ejercicio:

**Ejercicio 10.5.1.** Con los siguientes datos encuentre los estimadores  $b_0$  y  $b_1$

X	1	1	1	2	2	2	3	3	3	4	4	4	5	5	5
Y	10	13	12	14	15	14	18	16	15	20	22	16	17	20	18

Para resolver esto necesitamos las siguientes cantidades presentes en las ecuaciones normales:

$$n = 15$$

$$\sum Y_i = 10 + 13 + 12 + \dots + 18 = 240$$

$$\sum X_i = 1 + 1 + 1 + 2 + \dots + 5 = 45$$

$$\sum X_i Y_i = (1 * 10) + (1 * 13) + \dots + (5 * 18) = 775$$

$$\sum X_i^2 = 1^2 + 1^2 + 1^2 + 2^2 + \dots + 5^2 = 165$$

Ahora se procede a reemplazar las estadísticas encontradas en las ecuaciones normales:

$$15b_0 = 240 - 45b_1$$

$$45b_0 = 775 - 165b_1$$

Multiplicamos por -3 la primera ecuación para eliminar  $b_0$ :

$$-45b_0 = -720 + 135b_1$$

$$45b_0 = 775 - 165b_1$$

Al restar la una ecuación de la otra y al despejar, obtenemos el valor de  $b_1$ :

$$0 = 55 - 30b_1; b_1 = 1.83$$

Ahora podemos sustituir el valor de  $b_1$  en cualquiera de las ecuaciones para obtener el valor de  $b_0$ :

$$15b_0 = 240 - 82.5; b_0 = 10.50$$

### ***La Covarianza y el Coeficiente de Regresión***

Otra forma, quizás más simple, con la que podemos encontrar  $b_1$  (el coeficiente de regresión de la muestra o pendiente) es dividiendo la *covarianza* de  $XY$  para la *varianza* de  $X$ , es decir:

$$b_1 = \frac{Cov_{XY}}{S_X^2} \quad (54)$$

La covarianza es una medida de la variación conjunta de dos variables. La covarianza de la muestra  $Cov_{XY}$  se la encuentra con la siguiente fórmula:

$$Cov_{XY} = \frac{SP_{XY}}{n-1} = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{n-1} \quad (55)$$

La cantidad  $SP_{XY}$  es la Suma Corregida de Productos XY.

La varianza de la variable independiente  $S_X^2$  se la encuentra con la siguiente fórmula:

$$S_X^2 = \frac{SC_X}{n-1} = \frac{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}{n-1} \quad (56)$$

La cantidad  $SC_X$  es la Sumatoria Corregida de Cuadrados de X.

Por lo anterior, la estadística  $b_1$  puede encontrarse directamente con la siguiente fórmula:

$$b_1 = \frac{SP_{XY}}{SC_X} \quad (57)$$

El valor de La intersección  $b_0$  se lo encuentra a partir del punto  $(\bar{X}, \bar{Y})$  que, de acuerdo con el modelo, está necesariamente sobre la línea de regresión. Por lo tanto, se puede escribir la ecuación:

$$\bar{Y} = b_0 + b_1 \bar{X} \text{ Al despejar se obtiene la estadística } b_0:$$

$$b_0 = \bar{Y} - b_1 \bar{X} \quad (58)$$

El ejercicio a continuación hace uso de las fórmulas anteriores:

**Ejercicio 10.5.2.** Se desea encontrar las estadísticas  $b_0$  y  $b_1$  (estimadores de los parámetros  $\beta_0$  y  $\beta_1$ ) para la relación entre edad y peso de una raza criolla de cerdos. Para esto se tomaron muestras al azar de cerdos de diferentes edades en un criadero. Los resultados se indican a continuación:

semanas X	4	4	8	8	12	12	16	16	20	20	24	24
peso Y	4.3	5.8	17.4	15.6	30.2	36.1	53.3	60.0	75.9	68.4	84.5	88.7

Primero calculamos las estadísticas para la variable independiente X:

$$\begin{aligned} n=12 & & \sum X_i = 168 & & \bar{X} = 14 \\ \sum X_i^2 = 2912 & & \frac{(\sum X_i)^2}{n} = 2352 & & SC_X = 560 \end{aligned}$$

Luego calculamos las estadísticas para la variable dependiente Y:

$$\begin{aligned} n = 12 & & \sum Y_i = 540.20 & & \bar{Y} = 45.02 \\ \sum Y_i^2 = 34,701.7 & & \frac{(\sum Y_i)^2}{n} = 24,318.0 & & SC_Y = 10,383.7 \end{aligned}$$

A continuación, se calcula las estadísticas para las dos variables X y Y combinadas (note que

$n = 12$ , es decir, hay doce pares de observaciones XY):

$$\sum X_i Y_i = (4 \cdot 4.3) + (4 \cdot 5.8) + \dots + (24 \cdot 88.7) = 9,955.6$$

La expresión anterior se denomina: “La sumatoria de los productos XY”.

$$\sum X_i \sum Y_i = 168 \cdot 540.2 = 90,753.6$$

La expresión anterior se denomina: “La sumatoria de X por la sumatoria de Y”.

$$SP_{XY} = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} = 9,955.6 - 90,753.6/12 = 2,392.8$$

La expresión anterior se denomina: “La sumatoria corregida de productos XY”.

Tenemos ya todos los elementos necesarios para calcular el coeficiente de regresión  $b_1$ :

$$b_1 = \frac{SP_{XY}}{SC_X} = \frac{2,392.8}{560} = 4.2729$$

Luego calculamos la intersección  $b_0$ :

$$b_0 = \bar{Y} - b_1 \bar{X} = 45.02 - 4.2729 \cdot 14 = -14.8030$$

Las respuestas:  $b_1 = 4.27$  y  $b_0 = -14.80$ , son los estimadores de los parámetros  $\beta_1$  y  $\beta_0$ , respectivamente.

### La línea recta

Los datos del Ejercicio 10.5.2 pueden presentarse en un gráfico de coordenadas y a través de ellos trazar un segmento de línea recta. Podrían trazarse mu-

chos segmentos de líneas rectas, sin embargo, sólo uno es el que mejor representa los datos y éste es el que se fundamenta en el método de *cuadrados mínimos*. La *línea de Regresión* es el segmento de una línea recta que en promedio mejor representa los datos experimentales.

**Ejercicio 10.5.3.** Con los datos del Ejercicio 10.5.2 presente un gráfico con el segmento de línea recta (línea de regresión).

Para trazar una línea recta hacen falta únicamente dos puntos. Estos pueden ser: la intersección  $b_0$  y el punto correspondiente a las medias  $(\bar{X}, \bar{Y})$ . Los círculos del gráfico corresponden a las observaciones experimentales. El segmento de línea recta trazado es el que mejor representa en promedio la tendencia de las observaciones experimentales, de acuerdo con el modelo de regresión lineal. Con el programa *Statistix* el gráfico es el siguiente:

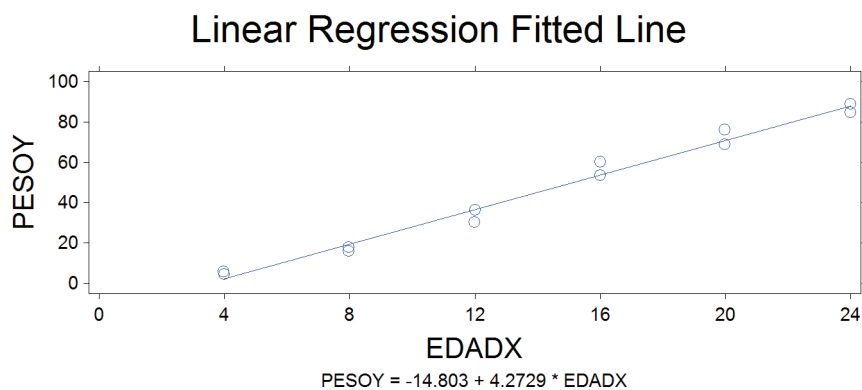


Gráfico 1. Segmento de la línea recta para los datos del ejercicio 10.5.2.

Observe la ecuación de regresión en la parte inferior del gráfico:  
 $Y = -14.803 + 4.272 * X$

## 10.6. Estimación del error experimental

Como todo lo relacionado con la estadística inferencial, el modelo es probabilístico y consecuentemente tiene un componente de error. Por lo tanto, se requiere estimar  $\varepsilon_i$ .

Si con los datos de un análisis de regresión, se traza un segmento de línea recta que pase por los puntos  $b_0$  y  $(\bar{X}, \bar{Y})$ , con seguridad muchas de las observacio-

nes experimentales no caerán sobre el segmento de la línea recta. Como ejemplo, varios de los círculos que representan a las observaciones experimentales no están exactamente sobre la línea de regresión (ver gráfico del Ejercicio 10.5.3). Cada una de las distancias o desviaciones verticales, de las observaciones experimentales con relación al segmento de línea, llamadas individualmente  $e_i$ , son parte del error y sirven en conjunto para estimar  $\varepsilon_i$ .

Las desviaciones verticales  $e_i$  son entonces las distancias entre las observaciones experimentales y el punto correspondiente estimado  $\hat{Y}_i$  sobre la línea de regresión. Para encontrar las  $\hat{Y}_i$  empleamos la ecuación de Regresión,  $\hat{Y} = b_0 + b_1X$ . Las  $e_i$  se encuentran con la fórmula:

$e_i = Y_i - \hat{Y}_i$ . La sumatoria de estas desviaciones de  $Y_i$  con relación a  $\hat{Y}_i$  es igual a cero, es decir:  $\sum e_i = \sum(Y_i - \hat{Y}_i) = 0$

La condición que  $\sum e_i = \sum(Y_i - \hat{Y}_i) = 0$  nos indica que la sumatoria de las desviaciones al cuadrado  $\sum e_i^2 = \sum(Y_i - \hat{Y}_i)^2$  es un *cuadrado mínimo*. En otras palabras, si trazáramos cualquier otra línea en que la sumatoria de las observaciones no fuera cero, es decir  $\sum e_i \neq 0$ , entonces el valor  $\sum e_i^2$  sería mayor que cuando  $\sum e_i = 0$ .

Si se elevan todos los valores de las desviaciones  $e_i$  al cuadrado se obtiene la estadística

$\sum e_i^2 = \sum(Y_i - \hat{Y}_i)^2$  que se denomina *Suma de Cuadrados del Error (SCE)*. Este es un valor mínimo, lo que implica que todas las desviaciones con relación a cualquier otra variable que no sea  $\hat{Y}_i$  darán un valor más alto. De lo anterior puede deducirse que todos los puntos estimados  $\hat{Y}_i$  están sobre la línea de regresión.

La SCE dividida para  $n - 2$  (los grados de libertad) se denomina “varianza del error” o más comúnmente conocido como *Cuadrado Medio del Error (CME)*, esto es:

$$CME = \frac{\sum(Y_i - \hat{Y}_i)^2}{n-2} = \frac{SCE}{n-2} \quad (59)$$

El CME es un estimador válido de  $\varepsilon_i$

---

**Ejercicio 10.6.1.** Con los datos del Ejercicio 10.5.2 encontrar los valores de  $\hat{Y}_i$ ,  $e_i$ ,  $e_i^2$ , de la SCE (Sumatoria de Cuadrados del Error) y del CME (Cuadrado Medio del Error).

Primero encontramos los valores de  $\hat{Y}$  para los doce valores de  $X$ , a partir de la ecuación

$$\hat{Y} = b_0 + b_1X, \text{ en la que } b_0 = -14.8030, \text{ y } b_1 = 4.2729.$$

Cuadro 1. Valores esperados de "Y sombrero",  $\hat{Y}$

$\hat{Y} = b_0 + b_1X$	$\hat{Y} = b_0 + b_1X$
$\hat{Y}_1 = -14.80 + 4.27 * 4 = 2.29$	$\hat{Y}_2 = -14.80 + 4.27 * 4 = 2.29$
$\hat{Y}_3 = -14.80 + 4.27 * 8 = 19.38$	$\hat{Y}_4 = -14.80 + 4.27 * 8 = 19.38$
<i>n y de x</i>	$\hat{Y}_6 = -14.80 + 4.27 * 12 = 36.47$
$\hat{Y}_7 = -14.80 + 4.27 * 16 = 53.56$	$\hat{Y}_8 = -14.80 + 4.27 * 16 = 53.56$
$\hat{Y}_9 = -14.80 + 4.27 * 20 = 70.66$	$\hat{Y}_{10} = -14.80 + 4.27 * 20 = 70.66$
$\hat{Y}_{11} = -14.80 + 4.27 * 24 = 87.75$	$\hat{Y}_{12} = -14.80 + 4.27 * 24 = 87.75$

Ahora encontramos las  $e_i = Y_i - \hat{Y}_i$  y las  $e_i^2 = (Y_i - \hat{Y}_i)^2$ , para las 12 observaciones y las respectivas sumatorias.

Cuadro 2. Datos para encontrar el Cuadrado Medio del Error

$Y_i$	$\hat{Y}_i$	$e_i = Y_i - \hat{Y}_i$	$e_i^2 = (Y_i - \hat{Y}_i)^2$
4.3	2.29	2.01	4.05
5.8	2.29	3.51	12.33
17.4	19.38	-1.98	3.92
15.6	19.38	-3.78	14.29
30.2	36.47	-6.27	39.34
36.1	36.47	-0.37	0.14
53.3	53.56	-0.26	0.07
60.0	53.56	6.44	41.43
75.9	70.66	5.25	27.51
68.4	70.66	-2.26	5.09
84.5	87.75	-3.25	10.54
88.7	87.75	0.95	0.91
$\sum Y_i = 540.2$	$\sum \hat{Y}_i = 540.2$	$\sum e_i = 0$	$\sum e_i^2 = 159.60$

Note que  $\sum Y_i = \sum \hat{Y}_i = 540.2$ , por lo que  $\sum e_i = \sum (Y_i - \hat{Y}_i) = 0$ .



La sumatoria de los cuadrados de las desviaciones, en otras palabras *Sumatoria de Cuadrados del Error*, es entonces:

$$SCE = \sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2 = 159.60$$

Por lo tanto, el *Cuadrado Medio del Error* es:

$$CME = \frac{SCE}{n-k} = \frac{159.60}{10} = 15.96$$

Como hubo 12 observaciones X Y, el número de grados de libertad,  $n - k$ , es 10 ( $n$  es el número de observaciones y  $k$  es el número de variables en el modelo, en este caso dos variables).

## 10.7. Significado del coeficiente de regresión

El coeficiente de regresión o *pendiente*  $b_1$  es la estadística que mejor estima el parámetro  $\beta_1$ . En el ejercicio 10.5.2 se encontró que el coeficiente de regresión fue:  $b_1 = 4.2729$ . Este valor indica que por cada semana de incremento en edad (unidad de X), el peso de los cerdos, en promedio, se incrementa en 4.2729 kilogramos (unidades de Y). Pero es importante recordar que  $b_1$  se obtuvo a partir de una muestra y por lo tanto la conclusión a la que se arrije está sujeta al cálculo de probabilidades y consecuentemente tiene un nivel de error.

Como puede verse en el gráfico del ejercicio 10.5.3,  $b_1$  determina la pendiente de la línea de Regresión. Si  $b_1 > 4.2729$ , entonces la inclinación de la línea o “pendiente” sería más pronunciada. Si  $b_1$  tuviera un valor negativo, entonces la línea iría de arriba hacia abajo, en otras palabras, Y disminuiría como respuesta a incrementos de X. Si  $b_1$  fuera cero, entonces la línea de regresión sería paralela a la abscisa o eje de las X, lo que significaría que Y no depende de X.

## 10.8. Significado de la Intersección

La intersección se estima por el valor  $b_0$ . Este es el punto en el que el segmento de la línea recta o su prolongación corta el eje de la ordenada Y. Se puede definir también como el valor de  $\hat{Y}_i$  cuando  $X = 0$ .

Se debe tener cuidado al interpretar  $b_0$ , en especial si datos experimentales de  $Y_i$  no fueron tomados al nivel  $X = 0$ . En el ejercicio 10.5.2 el valor de  $b_0$  fue igual a -14.8030. Como no se tomaron datos cuando  $X = 0$ , es decir al momento de la parición, este valor carece de sentido biológico. En todo caso  $b_0$  es un punto

que está sobre la línea de regresión (o en su proyección) y que define la muestra de la que proviene.

## 10.9. Extrapolación de datos

Una vez que se ha determinado la ecuación  $\hat{Y} = b_0 + b_1X$  se pueden extrapolar los resultados a valores de  $X$  diferentes a los empleados experimentalmente.

**Ejercicio 10.9.1.** Con los resultados obtenidos en el Ejercicio 10.6.1, ¿Cuál es el peso promedio esperado para cerdos de 14 semanas,  $\hat{Y}_{14}$ ?

$$\text{El valor es: } \hat{Y}_{14} = -14.8030 + 4.2729 \cdot 14 = 45.02 \text{ kg}$$

Recuerde que este valor de  $\hat{Y}_{14}$ , como cualquier otro valor esperado, está sobre la línea de regresión. Sin embargo, como la línea de Regresión es un segmento de una línea recta, la extrapolación tiene sentido únicamente para el intervalo que dicta el conocimiento del campo de estudio específico.

## 10.10. Pruebas de hipótesis con relación a $\beta_1$

**Cuando la hipótesis nula plantea que  $\beta_1$  es igual a un valor diferente de cero**

En estudios de regresión puede ser de interés probar la hipótesis nula que  $\beta_1$  es igual a algún valor hipotético  $\beta_{1h}$  versus la alternativa en la que se plantea que  $\beta_1$  no es igual a ese valor hipotético. Esto es:

$$H_0: \beta_1 = \beta_{1h}$$

$$H_1: \beta_1 \neq \beta_{1h}$$

La hipótesis alternativa  $H_1$  también puede tomar las formas  $\beta_1 > \beta_{1h}$  o  $\beta_1 < \beta_{1h}$  de acuerdo con la información previa o el interés que tenga el investigador.

Como ejemplo, en el ejercicio 10.52 se encontró que para esa muestra de cerdos de la raza A, el valor de  $b_1$  fue 4.2729. Como  $b_1$  es el estimador del parámetro  $\beta_1$  cuyo valor real desconocemos, un investigador puede estar interesado en probar, con la información disponible, si  $\beta_1$  es igual al parámetro  $\beta_{1h} = 6.0$ , valor ya establecido en otra raza de cerdos. La evaluación de estas hipótesis se realiza mediante la prueba  $t$ . Una presunción importante para esta prueba es que la muestra provenga de una población de  $Y_i$  normalmente distribuida. La fórmula que se utiliza es la siguiente:

$$t = \frac{b_1 - \beta_{1h}}{S_b} \quad (60)$$

Esta prueba estadística sigue la distribución  $t$  con  $n - 2$  grados de libertad.

La estadística  $S_b$  es el *error estándar del coeficiente de regresión* y es igual a la raíz cuadrada del Cuadrado Medio del Error (CME) dividido para la Suma de Cuadrados de X:

$$S_b = \sqrt{\frac{CME}{SC_X}} \quad (61)$$

En las secciones 10.5 y 10.6 se vieron los procedimientos para encontrar las estadísticas CME y  $SC_X$ .

---

**Ejercicio 10.10.1.** Con los datos del ejercicio 10.5.2 pruebe la hipótesis nula que el valor del parámetro  $\beta_1$  de esa población es estadísticamente igual a un incremento de 6.0 kg/semana.

$$H_0: \beta_1 = 6.0$$

$$H_1: \beta_1 \neq 6.0$$

$$\alpha = 5\%$$

El valor del coeficiente de Regresión hipotético  $\beta_{1h} = 6.0$  representa la ganancia en peso de una raza distinta de cerdos. El interés es entonces determinar si hay diferencias estadísticas en la ganancia de peso de las dos razas en función de la edad.

El primer paso es calcular el Cuadrado Medio del Error (CME). Esto lo hacemos a partir de la suma de cuadrados del Error (SCE) que es igual a  $\sum(Y_i - \hat{Y}_i)^2$  y ya fue encontrado en el ejercicio 10.6.1:

$$SCE = \sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2 = 159.60$$

Por lo tanto, el *Cuadrado Medio del Error* es:

$$CME = \frac{SCE}{n-k} = \frac{159.60}{10} = 15.96$$

Luego encontramos  $S_b$ . Para esto necesitamos conocer también la  $SC_X$  que fue igual a 560 (ejercicio 10.5.2):

$$S_b = \sqrt{\frac{CME}{SC_X}} = \sqrt{\frac{15.96}{560}} = 0.1688$$

El valor de la prueba estadística  $t$  es entonces:

$$t = \frac{4.2729-6.0}{0.1688} = -10.23$$

Los valores críticos en la Tabla  $t$  al nivel  $\alpha/2$  por ser prueba de dos colas, con 10 grados de libertad (12 pares de observaciones  $X Y$ ) son:

$$t_{(10)(0.025)} = \pm 2.2280$$

Como el valor de la prueba estadística  $t$  excede el valor crítico absoluto de la tabla, la decisión es rechazar la  $H_o$  al 5% nivel de significación. En otras palabras la ganancia de peso, con relación a la edad, en la raza de cerdos del experimento, no es igual a 6.0 kg/semana. En otras palabras, se ha encontrado evidencia estadística de que la ganancia de peso en la raza del experimento es menor que 6.0 kg/semana.

### Quando la hipótesis nula plantea que $\beta_1$ es igual a cero

En muchos casos, la prueba de hipótesis que interesa es si el valor de  $\beta_1$  es igual a cero, lo que indicaría que, al variar  $X$  en una unidad, no produce ningún efecto sobre  $Y$ . En otras palabras, que las variables  $X$  y  $Y$  son independientes.

Las hipótesis que se prueban son:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

La hipótesis alterna  $H_1$  también podría ser:  $H_1: B_1 < 0$  (menor que cero) o  $H_1: B_1 > 0$  (mayor que cero), pruebas de cola izquierda y derecha, respectivamente. La *prueba estadística* que se emplea es la misma del caso anterior:

$$t = \frac{b_1 - \beta_{1h}}{S_b}$$

Pero como  $\beta_{1h} = 0$ , la fórmula se reduce a:

$$t = \frac{b_1}{S_b}$$

**Ejercicio 10.10.2.** Con los datos del ejercicio 10.5.2 pruebe las siguientes hipótesis:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

$$\alpha = 1\%$$

La prueba estadística es:

$$t = \frac{b_1}{s_b} = \frac{4.2729}{0.1688} = 25.31$$

Como la prueba es al 1% (0.01), los valores críticos en la tabla  $t$ , por ser prueba de dos colas, se encuentran con los niveles de significación  $\alpha/2 = 0.005$  :

$$t_{(10)(0.005)} = \pm 3.1690$$

Como el valor 25.31 de la prueba estadística  $t$ , excede a los valores críticos de la tabla  $t$  ( $\pm 3.1690$ ), la decisión es rechazar la hipótesis nula  $H_0$  que  $\beta_1 = 0$  al nivel de significación del 1%. En otras palabras, la variable Y, peso de los cerdos, depende de la variable X, semanas de vida. Se puede señalar que hay evidencia estadística que el promedio de peso de los cerdos de esa raza se incrementa en 4.2729 kg por semana.

En este y otros casos podría ser que el modelo lineal no sea el más adecuado y que modelos curvilíneos proporcionen un mejor ajuste, algo que se verá posteriormente.

### **Respuestas con programas computarizados**

Con los programas computarizados se obtienen las respuestas con mayor celeridad, obviamente, pero también hay que interpretarlas correctamente. Como comparación se presentan los resultados de los ejercicios anteriores, sobre ganancia de peso de cerdos, con el programa *Statistix*:

#### **Least Squares Linear Regression of PESO EDAD CERDOS**

<b>Predictor Variables</b>	<b>Coefficient</b>	<b>Std Error</b>	<b>T</b>	<b>P</b>
Constant	-14.8033	2.62986	-5.63	0.0002
EDADX	4.2729	0.16882	25.31	0.0000

R-Squared = 0.9846    Resid. Mean Square (MSE) = 15.9604

Note que estos resultados:  $b_0 = -14.8033$ ,  $b_1 = 4.2729$ ,  $S_b = 0.16882$ ,  $t = 25.31$ ,  $P \leq 0.0000$ ,  $R^2 = 0.9846$ , MSE o CME = 15.96, fueron ya encontrados en este y otros ejercicios, con excepción de  $R^2$  y de la probabilidad  $P$ . Fijese, además, en el cuadro anterior de resultados del programa *Statistix*, la ubicación y el nombre o símbolo para cada uno de estos valores.

La estadística  $R^2$  se conoce como *coeficiente de determinación* y es una medida del grado de dependencia de Y con relación a X, que generalmente se expresa en porcentaje. En este ejemplo se puede indicar que la variación de Y se debe en un 98.46% a su asociación con X. Sin embargo, este valor es el *coeficiente de determinación simple*, que en este texto lo representaremos como  $r^2$ , para diferenciar del *coeficiente de determinación múltiple*,  $R^2$ .

Con relación a la probabilidad  $P$ , el programa *Statistix* en las pruebas que realiza no presenta los valores críticos, tampoco indica si se debe aceptar o rechazar la hipótesis nula. Lo que hace es presentar la probabilidad de que el valor encontrado para la prueba estadística se deba al azar. En este caso la  $P \leq 0.0000$  significa que la probabilidad de un resultado tan alto como  $t=25.31$ , sólo por azar, es muy baja (menor que 1/10000). Esto es, en todo caso, una probabilidad menor que la del nivel de aceptación planteado, que en este ejercicio fue del 1/100. Por lo tanto, hay fuerte evidencia estadística para rechazar la hipótesis nula y aceptar la hipótesis alterna.

**Ejercicio 10.10.3.** En los cultivos por lo general el ciclo de vida influye en el rendimiento. Con la siguiente información de un experimento con 16 líneas de arroz en que los datos fueron “días a floración y rendimiento en kg por parcela”:

Días a Flor (X):	50	82	56	70	72	82	64	60	84	71	64	54	80	58	65	68
Rendimiento (Y):	2.8	2.7	3.3	2.7	2.5	2.7	3.7	3.1	2.0	3.1	3.2	3.4	2.4	3.5	3.8	3.2

- encuentre los coeficientes  $b_0$  y  $b_1$
- pruebe las hipótesis:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

$$\alpha = 5\%$$

- presente un gráfico con la ecuación de regresión lineal

Los resultados, con el uso de un programa de estadística, son:

#### Least Squares Linear Regression of RENDkg

Predictor Variables	Coefficient	Std Error	T	P
Constant	5.05807	0.64270	7.87	0.0000
DIASflor	-0.03040	9.413E-03	-3.23	0.0061

R-Squared = 0.4269

Resid. Mean Square (MSE) = 0.14939

a) El coeficiente  $b_0$ , la intersección, es igual a 5.05 kg. En este caso es un valor teórico y sin sentido biológico. El coeficiente  $b_1$ , coeficiente de regresión o pendiente, es igual a -0.0304 kg.

b) Como el valor de la prueba estadística  $t$  es igual a -3.23 y tiene una probabilidad baja de ser un resultado del azar,  $P = 0.0061$ , la decisión es rechazar la hipótesis nula que  $\beta_1 = 0$ . Es decir, hay evidencia estadística que  $\beta_1 < 0$ . En otras palabras, el rendimiento en las poblaciones de estudio depende del ciclo de vida. Esta dependencia sin embargo no es muy fuerte. Si se observa el valor del coeficiente de determinación ( $r^2$ ) sólo se explica un 42.69% de la variación en Y por su dependencia con X.

c) el gráfico de la línea de regresión y la ecuación de regresión se presentan a continuación. Note la inclinación de la línea de regresión o pendiente negativa ( $b_1 = -0.0304$ )

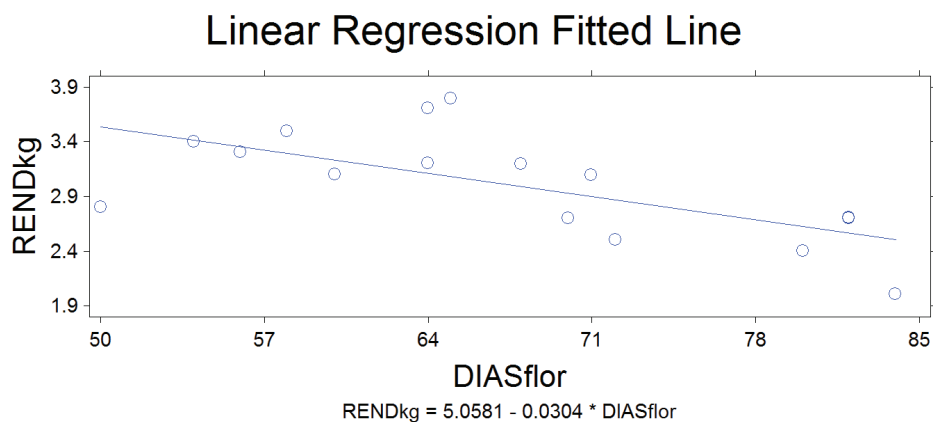


Gráfico 1. Regresión lineal de rendimiento sobre días a floración en 16 líneas de arroz

**Comentario:** Por lo general, a mayor ciclo de vida, mayor rendimiento. Resultados como los del ejercicio anterior pueden deberse a condiciones adversas en el desarrollo del cultivo (sequía, temperaturas altas, incidencia de enfermedades, daños por insectos, etc.).

### 10.11 . Intervalo de confianza del coeficiente de regresión $\beta_1$

En forma similar como se encontró intervalos de confianza para la media y varianza poblacional, igualmente se puede encontrar *intervalos de confianza para el coeficiente de regresión de la población*  $\beta_1$ . El procedimiento se fundamenta en la distribución t. La fórmula que se utiliza es la siguiente:

$$P(b_1 - t_{\alpha/2} * S_b \leq \beta_1 \leq b_1 + t_{\alpha/2} * S_b) = 1 - \alpha \quad (62)$$

Fórmula en la que:

- $P$  es la probabilidad
- $b_1$  es el coeficiente de regresión de las muestras
- $t_{\alpha/2}$  es el valor crítico en la tabla  $t$  con grados de libertad y al nivel especificados
- $S_b$  es el error estándar del coeficiente de regresión
- $\beta_1$  es el coeficiente de regresión de la población en estudio
- $1 - \alpha$  es el nivel de confianza del intervalo. Se expresa usualmente en porcentaje

Es importante recordar la importancia del concepto de *intervalo de confianza*. En el Ejercicio 10.5.2 encontramos que el coeficiente de regresión de la *muestra* fue  $b_1 = 4.2729$ . Aunque  $b_1$  es el mejor estimador del *coeficiente de regresión de la población*  $\beta_1$  no quiere decir que  $\beta_1$  necesariamente sea igual a 4.2729. Lo que nos dice es que la inclinación del segmento de línea recta fluctúa entre dos valores estipulados estadísticamente. Entre estos valores se encuentra el parámetro  $\beta_1$  con una determinada probabilidad o nivel de confianza. En el ejemplo, el valor 4.2729 es sólo un punto de referencia.

De lo anterior puede deducirse que mientras mayor sea el nivel de confianza más amplio será el intervalo con la probable ubicación del parámetro  $\beta_1$ . Es importante también recordar que el nivel de confianza, o probabilidad, puede aproximarse al 100% pero nunca llegar a ese valor. Esto último se debe a la propiedad asintótica de las distribuciones estadísticas.



Habrán casos en los que el intervalo de confianza incluya el cero. De ser así, el valor de

$\beta_1$ , de acuerdo con la prueba estadística, no será diferente de cero.

**Ejercicio 10.11.1.** Con los datos del ejercicio 10.5.2 (edad y ganancia de peso en cerdos) encuentre los Intervalos de Confianza (IC) para  $\beta_1$ : a) con el 90% de probabilidad y b) con el 99% de probabilidad. Observe las diferencias en la amplitud de los intervalos.

a) Si el IC es igual a  $100 - \alpha = 90\%$ , entonces  $\alpha = 0.1 = 10\%$  y  $\alpha/2 = 0.05 = 5\%$

Recordemos que  $b_1 = 4.2729$  y que  $S_b = 0.1688$

Encontramos en la tabla (Apéndice B) el valor de  $t$  con  $n - 2$  grados de libertad, al nivel alfa del 0.05 de probabilidad:

$$t_{\alpha/2} = t_{(10 \text{ gl}), 0.05} = 1.8120$$

Al reemplazar los valores en la fórmula tenemos:

$$P(4.2729 - 1.8120 * 0.1688 \leq \beta_i \leq 4.2729 + 1.8120 * 0.1688) = 1 - \alpha = 90\% IC$$

$$P(3.9670 \leq \beta_i \leq 4.5788) = 90\% IC$$

Se concluye que  $\beta_1$  se encuentra entre 3.9770 kg/semana y 4.5788 kg/semana con el 90% de probabilidad.

b) Si aumentamos el nivel de confianza, como ejemplo al 99%, podemos constatar que el intervalo se vuelve más amplio.

Ahora  $100 - \sigma = 99\%$ , entonces  $\alpha = 0.01 = 1\%$  y  $\alpha/2 = 0.005 = 0.5\%$

Encontramos en la tabla (Apéndice B) el valor de  $t$  con  $n - 2$  grados de libertad, al nivel alfa del 0.005 de probabilidad:

$$t_{\alpha/2} = t_{(10 \text{ gl}), 0.005} = 3.1690$$

Al reemplazar los valores en la fórmula tenemos:

$$P(4.2729 - 3.169 * 0.1688 \leq \beta_i \leq 4.2729 + 3.169 * 0.1688) = 1 - \alpha = 99\% IC$$

$$P(3.7380 \leq \beta_i \leq 4.8078) = 99\% IC$$

Se concluye que  $\beta_1$  se encuentra entre 3.7380 kg/semana y 4.8078 kg/semana con el 99% de probabilidad.

Note que, como se esperaba, al aumentar el nivel de confianza el intervalo se amplía.

## 10.12. Regresión polinómica

Cuando en los análisis de regresión lineal simple los resultados no son satisfactorios, el investigador puede pensar que el modelo que empleó no fue el adecuado. El coeficiente de determinación  $r^2$  es un buen indicador del grado de dependencia de Y con relación a X. Porcentajes bajos de esta estadística indican poca dependencia. El modelo de regresión polinómica se aplica cuando el investigador está interesado en encontrar respuestas que no siguen un solo patrón, es decir cuando los cambios no son constantes en la variable Y sino que dependen de los niveles de X. En otras palabras, los cambios en Y se incrementan o disminuyen en un mismo espacio de muestreo.

En este caso, el modelo incluye valores de X elevados a diferentes potencias. El modelo para la regresión polinómica es:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 X_i^3 \dots + \beta_n X_i^n + \varepsilon_i \quad (63)$$

El exponente de X define el grado del polinomio o respuesta:  $X_i^1$  corresponde a la respuesta lineal,  $X_i^2$  a la cuadrática,  $X_i^3$  a la cúbica,  $X_i^4$  a la cuártica y así sucesivamente. El exponente más alto de la ecuación es el que da el nombre al polinomio.

Una aplicación práctica de la regresión polinómica es, como ejemplo, el analizar el comportamiento de poblaciones de patógenos o de la misma enfermedad a través del tiempo. Supongamos que un investigador encuentra que el modelo cuártico es el que mejor se ajusta a un estudio de la población de un insecto en un cultivo. Esto le indica que inicialmente la población del insecto se incrementó (*respuesta lineal*), que luego la población decreció (*respuesta cuadrática*), que posteriormente volvió a crecer (*respuesta cúbica*) y que por último, la población volvió a decrecer (*respuesta cuártica*).

Lo anterior es solamente la parte de análisis de los datos. El porqué de ese determinado comportamiento debe venir del conocimiento del cultivo, del ciclo de vida del insecto, de la presencia de enemigos naturales, de cambios climáticos, etc.

Una aplicación, tal vez la más común en las ciencias agropecuarias, es la del modelo cuadrático en experimentos relacionados con niveles de fertilización, dosis de pesticidas, densidades de siembra, niveles de nutrimentos en dietas para animales, etc. El modelo de regresión polinómica es específicamente el cuadrático, porque por lo general se espera una respuesta inicial lineal y luego, una tendencia a rendimientos decrecientes. Como ejemplo, en experimentos con niveles de nitrógeno se espera que los rendimientos se incrementen linealmente hasta cierto nivel y que haya luego una tendencia decreciente en respuesta a niveles altos de nitrógeno. Con niveles aún más altos, la experiencia nos enseña que los rendimientos bajan sensiblemente por toxicidad de las plantas. En ningún caso se espera que vuelvan a incrementarse (lo que indicaría una respuesta cúbica). Lo particularmente interesante de la respuesta cuadrática, en experimentos con niveles de insumos, es el punto en el que se obtiene el *máximo técnico* y el *óptimo económico*. Esto se explicará posteriormente.

En todo caso, es el investigador el que decide qué modelo polinómico se ajusta mejor a sus datos. Un procedimiento que conduce a esto es incluir progresivamente diferentes variables  $X$  en el modelo, es decir  $X^1, X^2, X^3, etc$  y observar si el *coeficiente parcial de regresión* resultó significativo o no. Generalmente si no es significativo quiere decir que la inclusión de la nueva variable en el modelo no contribuyó a explicar mejor la variación de  $Y$ . En este caso también puede detectarse un cambio reducido en el *coeficiente de determinación*.

Para ilustrar lo anterior tomemos como ejemplo los siguientes datos de un experimento para evaluar el efecto de cinco niveles de fertilización (0, 1, 2, 3, 4) en el rendimiento de arroz en kilogramos por parcela:

Cuadro 10.12.1. Cinco niveles de fertilización (X)  
y su efecto en el rendimiento de arroz (Y) en kilogramos por parcela

X:	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Y:	3.2	5.9	7.5	8.3	7.1	2.6	4.8	6.3	6.9	6.5	3.0	6.2	7.1	7.5	6.8

**El modelo lineal.** Inicialmente el investigador podría plantearse un modelo lineal para explicar la tendencia de los datos:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

Como en ejercicios anteriores se presentó el procedimiento para el análisis del modelo lineal, en este ejemplo utilizamos directamente el programa *Statistix*:

## Least Squares Linear Regression of RendimY

Predictor Variables	Coefficient	Std Error	T	P
Constant	4.04667	0.49382	8.19	0.0000
fertNivel	0.96667	0.20160	4.79	0.0003
R-Squared	0.6388	Resid. Mean Square (MSE)	1.21928	

## Linear Regression Fitted Line

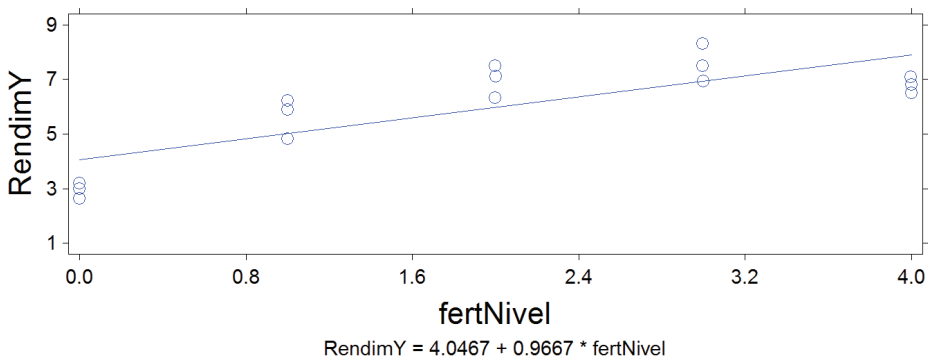


Gráfico 10.12.1. Respuesta lineal

El *coeficiente de regresión* de la muestra es igual a 0.9667. El valor de la prueba  $t$  igual a 4.79 nos sirve para evaluar la hipótesis nula de que el coeficiente de regresión es igual a cero. La probabilidad de haber obtenido por azar un valor tan extremo como 4.79 es  $P = 0.0003$ , por lo que debería rechazarse la hipótesis nula. En otras palabras, hay suficiente evidencia estadística de una respuesta lineal ascendente del rendimiento con relación a los niveles de fertilización. El valor del coeficiente de determinación simple,  $r^2 = 0.6388$  confirma lo señalado. Es decir, el 63.88% de la variación en el rendimiento  $Y$  puede atribuirse a su asociación lineal con los niveles de fertilización  $X$ .

**El modelo cuadrático.** Al observar el gráfico anterior puede verse que la línea recta representa bastante bien al conjunto de datos. El coeficiente de regresión,  $b_1$ , es significativamente diferente de cero y el coeficiente de determinación simple  $r^2$  es relativamente alto (63.88%). Sin embargo, podría ser que el modelo polinómico cuadrático provea un mejor ajuste:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i \quad (64)$$

Para encontrar las estadísticas  $b_0, b_1, b_2$ , estimadores de los parámetros  $\beta_0, \beta_1, \beta_2$ , es necesario encontrar varias sumatorias a partir de los datos de niveles de fertilización y rendimiento de arroz (Cuadro 10.12.1.) Los cálculos y sumatorias se presentan en el cuadro siguiente:

**Cuadro 10.12.2. Observaciones X, Y. Cálculo de sumatorias de cuadrados y sumatorias de productos para los datos del Cuadro 10.12.1.**

Datos	X	Y	X <sup>2</sup>	XY	X <sup>2</sup> Y	XX <sup>2</sup>	X <sup>2</sup> X <sup>2</sup>	Y <sup>2</sup>
1	0	3.2	0	0	0	0	0	10.24
2	1	5.9	1	5.9	5.9	1	1	34.81
3	2	7.5	4	15	30	8	16	56.25
4	3	8.3	9	24.9	74.7	27	81	68.89
5	4	7.1	16	28.4	113.6	64	256	50.41
6	0	2.6	0	0	0	0	0	6.76
7	1	4.8	1	4.8	4.8	1	1	23.04
8	2	6.3	4	12.6	25.2	8	16	39.69
9	3	6.9	9	20.7	62.1	27	81	47.61
10	4	6.5	16	26	104	64	256	42.25
11	0	3	0	0	0	0	0	9
12	1	6.2	1	6.2	6.2	1	1	38.44
13	2	7.1	4	14.2	28.4	8	16	50.41
14	3	7.5	9	22.5	67.5	27	81	56.25
15	4	6.8	16	27.2	108.8	64	256	46.24
<b>SUMATORIAS</b>	<b>30</b>	<b>89.7</b>	<b>90</b>	<b>208.4</b>	<b>631.2</b>	<b>300</b>	<b>1062</b>	<b>580.29</b>
<b>MEDIAS</b>	<b>2</b>	<b>5.98</b>	<b>6</b>	-	-	-	-	-

De acuerdo con el principio de cuadrados mínimos las ecuaciones normales para encontrar los coeficientes del modelo polinómico cuadrático son:

$$\sum Y_i = nb_0 + b_1 \sum X_i + b_2 \sum X_i^2$$

$$\sum X_i Y_i = b_0 \sum X_i + b_1 \sum X_i^2 + b_2 \sum X_i X_i^2 \quad (65)$$

$$\sum X_i^2 Y_i = b_0 \sum X_i^2 + b_1 \sum X_i X_i^2 + b_2 \sum X_i^2 X_i^2$$

De las ecuaciones anteriores, por procedimientos algebraicos, se derivan las siguientes fórmulas en términos de *Sumas de Cuadrados* y *Sumas de Productos*:

$$b_1 = \frac{(SCx^2)(SPxy) - (SPxx^2)(SPx^2y)}{(SCx)(SCx^2) - (SPxx^2)^2}$$

$$b_2 = \frac{(SCx)(SPx^2y) - (SPxx^2)(SPxy)}{(SCx)(SCx^2) - (SPxx^2)^2}$$

Resolvamos primero, con la información del Cuadro 10.12.2, cada una de las Sumas de Cuadrados y Sumas de Productos:

$$SCx^2 = \sum X_i^2 X_i^2 - \frac{\sum X_i^2 \sum X_i^2}{n} = 1062 - \frac{(90)^2}{15} = 522$$

$$SPxy = \sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n} = 208.4 - \frac{(30)(89.7)}{15} = 29$$

$$SPxx^2 = \sum X_i X_i^2 - \frac{\sum X_i \sum X_i^2}{n} = 300 - \frac{(30)(90)}{15} = 120$$

$$SPx^2y = \sum X_i^2 Y_i - \frac{\sum X_i^2 \sum Y_i}{n} = 631.2 - \frac{(90)(89.7)}{15} = 93$$

$$SCx = \sum X_i X_i - \frac{\sum X_i \sum X_i}{n} = 90 - \frac{(30)^2}{15} = 30$$

Note que tenemos ya todos los términos de los numeradores para encontrar  $b_1$  y  $b_2$

Observe que el denominador  $D$  es común para las dos fórmulas. Este valor es igual a:

$$D = (SCx)(SCx^2) - (SPxx^2)^2 = (30)(522) - (120)^2 = 1260$$

Por lo tanto, las estadísticas  $b_1$  y  $b_2$  son:

$$b_1 = \frac{(522)(29) - (120)(93)}{1260} = 3.1571$$

$$b_2 = \frac{(30)(93) - (120)(29)}{1260} = -0.5476$$

Para encontrar el coeficiente  $b_0$  o intersección se parte de la fórmula general del modelo, en la que se reemplaza con los valores de las medias de  $Y$ ,  $X$  y  $X^2$  que se encuentran en el Cuadro 10.12.2:

$$b_0 = \bar{Y} - b_1 \bar{X} - b_2 \bar{X}^2 = 5.98 - 3.1571 * 2 + 0.5476 * 6 = 2.9514$$

Consecuentemente, la ecuación de regresión polinómica de segundo grado o cuadrática, para los datos del Cuadro 10.12.1, es:

$$Y = 2.9514 + 3.1571X - 0.5476X^2$$

En la ecuación anterior, al reemplazar los 15 valores respectivos de  $X$  y  $X^2$ , se obtienen los 15 valores de  $\hat{Y}$  ( $Y$  sombrero). Estas cantidades se presentan en el Cuadro 10.12.3.

Podemos ahora probar hipótesis planteadas para determinar la significación estadística de los coeficientes parciales de regresión  $b_1$  y  $b_2$ . Para esto necesitamos primero encontrar el error estándar para cada uno de los coeficientes, es decir  $S_{b_1}$  y  $S_{b_2}$ .

Pero antes, en forma similar como se hizo en la sección 10.6, debemos calcular la *Sumatoria de Cuadrados del Error* (SCE) y el *Cuadrado Medio del Error* (CME). El CME es un estimador no sesgado de la varianza de regresión,  $\sigma^2$ . La *Sumatoria de Cuadrados del Error* se obtiene a partir de los cuadrados de las desviaciones  $(Y - \hat{Y})^2$  que se encuentran en la última columna del cuadro 10.12.3:

Cuadro 10.12.3. Valores de  $\hat{Y}$  ( $Y$  sombrero)  
y otras estadísticas para encontrar el Cuadrado Medio del Error

Datos	$X$	$Y$	$\hat{Y}$	$\hat{Y}^2$	$(Y - \hat{Y})$	$(Y - \hat{Y})^2$
1	0	3.2	2.95	8.71	0.25	0.0618
2	1	5.9	5.56	30.92	0.34	0.1150
3	2	7.5	7.08	50.06	0.42	0.1805
4	3	8.3	7.49	56.16	0.81	0.6492
5	4	7.1	6.82	46.49	0.28	0.0794
6	0	2.6	2.95	8.71	-0.35	0.1235
7	1	4.8	5.56	30.92	-0.76	0.5790
8	2	6.3	7.08	50.06	-0.78	0.6009
9	3	6.9	7.49	56.16	-0.59	0.3532
10	4	6.5	6.82	46.49	-0.32	0.1013
11	0	3	2.95	8.71	0.05	0.0024
12	1	6.2	5.56	30.92	0.64	0.4084
13	2	7.1	7.08	50.06	0.02	0.0006
14	3	7.5	7.49	56.16	0.01	0.0000
15	4	6.8	6.82	46.49	-0.02	0.0003
<b>SUMATORIAS</b>	<b>30</b>	<b>89.7</b>	<b>89.70</b>	<b>577.03</b>	<b>0.00</b>	<b>3.2554</b>
<b>MEDIAS</b>	<b>2</b>	<b>5.98</b>	-	-	-	-

La sumatoria de cuadrados del error (SCE) y la *varianza del error* o *cuadrado medio del error* (CME) son:

$$SCE = \sum(Y - \hat{Y})^2 = 3.2554$$

$$CME = \frac{SCE}{n-k} = \frac{3.2554}{15-3} = 0.2713$$

En la fórmula del CME, el valor de  $n$  es igual a 15 (número de observaciones de las variables) y  $k$  es igual a 3 (número de parámetros en el modelo:  $\beta_0, \beta_1, \beta_2$ )

Con esta información podemos pasar a encontrar los *errores estándar*  $S_{b_1}$  y  $S_{b_2}$  de la regresión. Las fórmulas son:

$$S_{b_1} = \sqrt{\frac{(CME)(SCx^2)}{(SCx)(SCx^2) - (SPxx)^2}} = \sqrt{\frac{(0.2713)(522)}{1260}} = 0.3353$$

$$S_{b_2} = \sqrt{\frac{(CME)(SCx)}{(SCx)(SCx^2) - (SPxx)^2}} = \sqrt{\frac{(0.2713)(30)}{1260}} = 0.0804$$

Disponemos ya de todo lo necesario para realizar las pruebas estadísticas de las siguientes hipótesis al nivel alfa del 5%:

$$H_0: \beta_1 = 0 \quad vs \quad H_1: \beta_1 \neq 0$$

$$H_0: \beta_2 = 0 \quad vs \quad H_1: \beta_2 \neq 0$$

Las pruebas estadísticas son pruebas  $t$ , similares a la usada en el ejercicio 10.10.2:

$$t_1 = \frac{b_1}{s_{b_1}} = \frac{3.1571}{0.3353} = 9.42 \quad t_2 = \frac{b_2}{s_{b_2}} = \frac{-0.5476}{0.0804} = -6.81$$

El valor crítico en la tabla  $t$  (Apéndice B) es  $\pm 2.179$ . Esto es al nivel del 2.5% de probabilidad, por ser prueba de dos colas, y con 12 grados de libertad. Como los valores de las dos pruebas estadísticas superan al valor crítico se puede concluir que, tanto  $\beta_1$  como  $\beta_2$  son estadísticamente diferentes de cero al nivel de significación especificado.

El significado de los coeficientes de regresión  $b_1$  y  $b_2$  es el siguiente:

- El coeficiente  $b_1$  indica que, en promedio, el rendimiento aumenta 3.1571 kg de arroz por cada unidad de incremento del fertilizante, si se mantiene  $b_2$  a un mismo nivel.



- El coeficiente  $b_2$  indica que, en promedio, el rendimiento disminuye 0.5476 kg de arroz por cada unidad de incremento del fertilizante, si se mantiene  $b_1$  a un mismo nivel.

**Intervalos de confianza para  $\beta_1$  y  $\beta_2$ .** Para encontrar intervalos de confianza para los parámetros de la regresión polinómica aplicamos fórmulas similares a las empleadas en la sección 10.11. Estas son:

$$P(b_1 - t_{\alpha/2} * S_{b1} \leq \beta_1 \leq b_1 + t_{\alpha/2} * S_{b1}) = 1 - \alpha$$

$$P(b_2 - t_{\alpha/2} * S_{b2} \leq \beta_2 \leq b_2 + t_{\alpha/2} * S_{b2}) = 1 - \alpha$$

Como ejemplo, podemos calcular intervalos de confianza para los parámetros  $\beta_1$  y  $\beta_2$ , digamos al 98%. En este caso, tenemos 1% a cada lado de la curva de  $t$ . En la tabla  $t$  para el 1% de área de la curva, con 12 grados de libertad encontramos el valor 2.6810. (Recuerde que  $n - k = 12$  en donde  $n$  es el número de observaciones y  $k$  el número de variables en el modelo.) Entonces  $t_{\alpha/2} = t_{(12gl),0.01} = 2.6810$

Al reemplazar los valores en la fórmula para  $\beta_1$ :

$$P(3.1571 - 2.6810 * 0.3353 \leq \beta_1 \leq 3.1571 + 2.6810 * 0.3353) = 98\%$$

Al resolver, se encuentra que  $\beta_1$  debe estar entre 2.2582 kg y 4.0560 kg de incremento en la producción de arroz, por unidad de fertilizante, con un 98% nivel de confianza (o probabilidad).

Ahora resolvemos para  $\beta_2$ :

$$P(-0.5476 - 2.6810 * 0.0804 \leq \beta_2 \leq -0.5476 + 2.6810 * 0.0804) = 98\%$$

Consecuentemente, se puede afirmar, con el 98% de confianza, que el parámetro  $\beta_2$  debe estar entre los valores - 0.7632 kg y - 0.3320 kg.

**Resultados con un programa para análisis estadísticos.** Pudo verse que las operaciones matemáticas, por la complejidad de las fórmulas, se vuelven bastante largas y requieren de considerable tiempo, atención y esfuerzo. Esto, aun para el modelo que contiene sólo tres variables. Con mayor número de variables, la cantidad de trabajo se incrementa exponencialmente. Felizmente, con las computadoras actuales y los programas estadísticos se ha simplificado grandemente el esfuerzo requerido. En verdad, ahora toma más tiempo introducir los datos y seleccionar las opciones correctas que obtener los resultados. A continuación, se

presentan los resultados del análisis de regresión polinómica (cuadrática) para el mismo ejemplo del Cuadro 10.12.1, con el programa *Statistix*:

#### Polynomial Regression of RendimY

Predictor Variables	Differenced Scale (i.e., $X_i - \bar{X}$ )				
	Coefficient	Std Error	T	P	VIF
Constant	7.07524	0.20958	33.76	0.0000	
fertNivel	0.96667	0.09509	10.17	0.0000	1.0
fertNivel^2	-0.54762	0.08037	-6.81	0.0000	1.0

#### Original Scale

Predictor Variables	Coefficient	Std Error
Constant	2.95143	0.28301
fertNivel	3.15714	0.33525
fertNivel^2	-0.54762	0.08037

R-Squared            0.9258            Resid. Mean Square (MSE)            0.27129

Cases Included 15            Missing Cases 0

Observe en los resultados que el programa presenta dos juegos de respuestas:

El primero, “differenced scale”, que presenta los resultados cuando se codifican los valores de  $X$  como  $X_i - \bar{X}$ . Esto normalmente se hace para simplificar la entrada de los datos, pero especialmente para obtener más fácilmente las respuestas mediante el método de polinomios ortogonales, como se verá más adelante.

El segundo, “Original Scale”, que emplea los valores reales de  $X$ , esto es, sin codificación. En este caso el programa no presenta los valores de  $t$  ni las probabilidades, pero no hace falta porque son equivalentes en las dos escalas. La intersección  $b_0$ , los coeficientes  $b_1$  y  $b_2$ , así como los valores  $S_{b1}$  y  $S_{b2}$  (los errores estándar -“Std Error”- de los coeficientes de regresión) son iguales en la “escala original” a los obtenidos manualmente en el ejemplo anterior.

Note en especial el valor de  $R^2$  igual a 0.9258. (En modelos con más de una  $X$ , se denomina  $R^2$  al *coeficiente de determinación múltiple* y se representa con mayúscula). Esto quiere decir que con el modelo cuadrático se explica en un 92.58% la variación en el rendimiento. Anteriormente, con el modelo lineal, se explicó el 63.88% de la variación de  $Y$  por su asociación con  $X$ . Este mejoramiento de casi

un 30% es significativo y demuestra que en este ejemplo el modelo cuadrático provee un mejor ajuste de los datos. Esto se puede visualizar al comparar el Gráfico 10.12.1 con el gráfico siguiente (Gráfico 10.12.2):

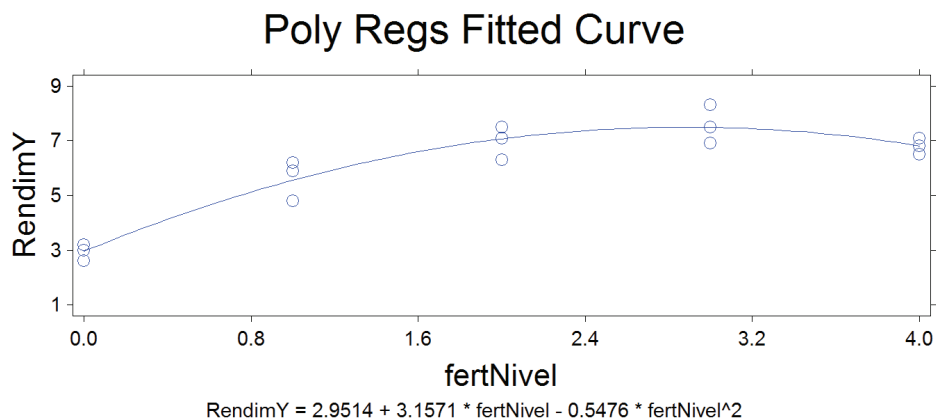


Gráfico 10.12.2. Respuesta cuadrática

Note que los datos se ajustan mejor (están más cerca) a la línea curva. Es decir que las desviaciones componentes del Cuadrado Medio del Error (*CME*) son menores que si se compara con el modelo lineal del Gráfico 10.12.1. El *CME* es una medida del *error experimental*.

El principio general en todos los casos es que mientras menor sea el *error experimental* mayor será la confianza en los resultados.

---

**Ejercicio 10.12.1.** Con los datos del ejercicio 10.10.3 (días a floración y rendimiento de arroz) y mediante un programa para análisis estadísticos encuentre el modelo polinómico que mejor se ajuste a los datos. Observe la significación de los coeficientes de regresión y los valores de los coeficientes de determinación

### Modelo Lineal

De acuerdo con los resultados del ejercicio 10.10.3 se obtuvo un valor de  $b_1 = -0.0304$ , estadísticamente significativo con una probabilidad  $P = 0.006$  (menor que el 1%). El valor del coeficiente de determinación  $r^2$  fue 42.69%.

## Modelo Cuadrático

Los coeficientes, su significación, el valor de  $R^2$ , y el gráfico con la ecuación de regresión se presentan a continuación:

Statistix, diasflor reg arroz,

### Polynomial Regression of Y

Differenced Scale (i.e., $X_i - \bar{X}$ )					
Predictor Variables	Coefficient	Std Error	T	P	VIF
Constant	3.22212	0.11865	27.16	0.0000	
X	-0.02800	8.061E-03	-3.47	0.0041	1.0
X <sup>2</sup>	-2.049E-03	8.123E-04	-2.52	0.0255	1.0

### Original Scale

Predictor Variables	Coefficient	Std Error
Constant	-4.22174	3.71997
X	0.24856	0.11090
X <sup>2</sup>	-2.049E-03	8.123E-04

R-Squared      0.6152      Resid. Mean Square (MSE)      0.10803

Note que los coeficientes de regresión son significativos. Al incluir una nueva variable en el modelo, esto es  $X^2$ , el valor del coeficiente de determinación,  $R^2$ , aumentó a 61.52%. Note también, en el gráfico que sigue, que la línea curva parece representar mejor al conjunto de datos, aunque existe una considerable dispersión de las observaciones. En todo caso, la significación de los coeficientes y la dirección de la curva parecen indicar que las variedades de arroz muy precoces y las muy tardías producen menos que las de ciclo intermedio.

## Poly Regs Fitted Curve

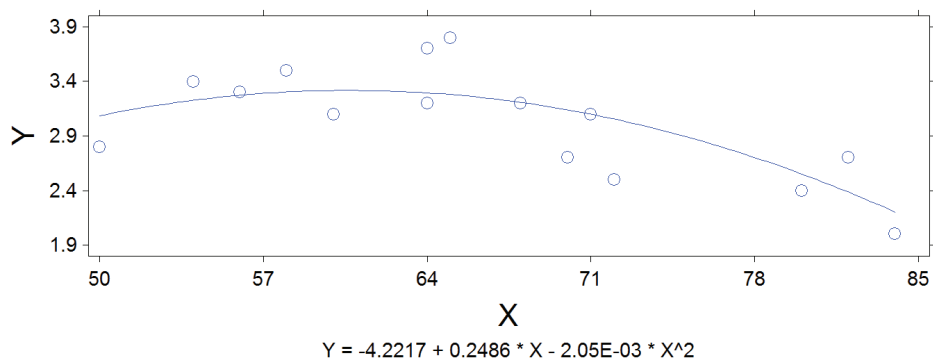


Gráfico 1. Respuesta cuadrática, rendimiento de arroz y días a flor

### Modelo Cúbico

En este modelo polinómico se incluyen cuatro variables: Y, X, X<sup>2</sup>, X<sup>3</sup>. De igual forma, se procede a encontrar los coeficientes de regresión, su significación, el coeficiente de determinación y se presenta el gráfico correspondiente:

diasflor reg arroz,

#### Polynomial Regression of Y

Predictor Variables	Differenced Scale (i.e., $X_i - \bar{X}$ )				
	Coefficient	Std Error	T	P	VIF
Constant	3.18275	0.11405	27.91	0.0000	
X	-0.06091	0.02136	-2.85	0.0146	8.1
X <sup>2</sup>	-1.862E-03	7.718E-04	-2.41	0.0327	1.0
X <sup>3</sup>	1.604E-04	9.732E-05	1.65	0.1253	8.0

#### Original Scale

Predictor Variables	Coefficient	Std Error
Constant	-50.5144	28.3089
X	2.38264	1.29923
X <sup>2</sup>	-0.03434	0.01961
X <sup>3</sup>	1.604E-04	9.732E-05

R-Squared      0.6862    Resid. Mean Square (MSE)    0.09544

Note que los coeficientes de regresión  $b_1$  y  $b_2$  son estadísticamente significativos al nivel del 5%. Sin embargo, el coeficiente de regresión  $b_3$  no es significativo ( $P = 12.53\%$ ). Esta información nos lleva a descartar el modelo cúbico y aceptar que para este juego de datos el mejor modelo es el cuadrático. El reducido incremento en el valor de  $R^2$  (de 61.52% en el modelo cuadrático a 68.62% en el modelo cúbico) confirma lo indicado. (Siempre que se incluye una nueva variable  $X$  en el modelo se incrementa el coeficiente de determinación  $R^2$ . Sin embargo, el incremento no es necesariamente significativo, como en este caso.)

Por lo señalado en el párrafo anterior y al observar el gráfico siguiente, la conclusión es que los datos no se ajustan bien al modelo cúbico. Consecuentemente, para este experimento el modelo cuadrático parece ser el más adecuado.

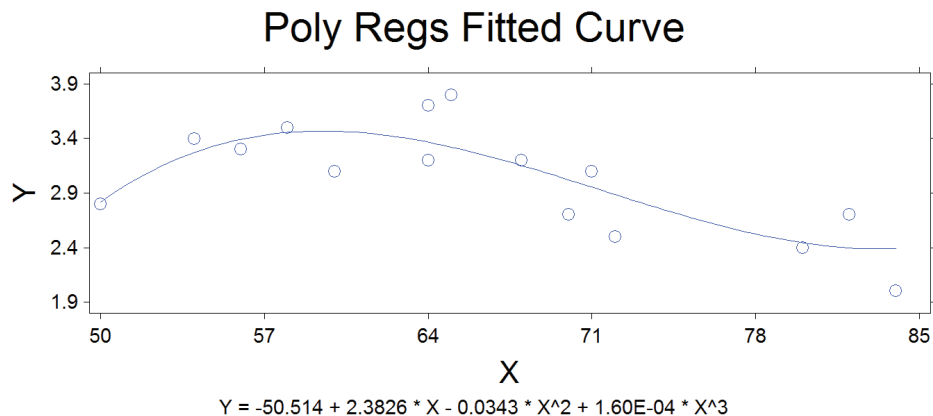


Gráfico 3. Respuesta cúbica, rendimiento de arroz y días a flor



# Análisis de varianza de regresión

*“Tengo mis resultados desde hace tiempo, pero no sé cómo llegar a ellos”.* Carl Gauss<sup>20</sup>

Hemos visto en las secciones precedentes las pruebas que nos permiten evaluar hipótesis como las siguientes, a un determinado nivel de significación alfa:

$$H_0: \beta_1 = 0 \text{ vs } H_1: \beta_1 \neq 0$$

$$H_0: \beta_2 = 0 \text{ vs } H_1: \beta_2 \neq 0$$

$$H_0: \beta_n = 0 \text{ vs } H_1: \beta_n \neq 0$$

Las hipótesis anteriores se evalúan mediante pruebas estadísticas *t* de Student, como se vio en los capítulos 9 y 10. En los análisis de varianza la prueba estadística es la prueba F de Fisher.

Un método alternativo para determinar el grado de dependencia de la variable Y con relación a una variable X, o con varias variables X, es la técnica que se conoce como *análisis de varianza de regresión*.

Con esta técnica podemos evaluar simultáneamente varias hipótesis concernientes a los coeficientes de regresión y determinar la pertinencia o no de un determinado modelo en forma más rápida.

Como el nombre lo indica, el método consiste en analizar la variación existente en el conjunto de datos y determinar las diferentes fuentes de esa variación y la importancia relativa de cada una de ellas.

## 11.1. Regresión: las fuentes de variación

En regresión, la variación proviene de tres fuentes:

20 Matemático y astrónomo alemán. [https://www.azquotes.com/author/5394-Carl\\_Friedrich\\_Gauss](https://www.azquotes.com/author/5394-Carl_Friedrich_Gauss)



**1. Variación total.** La variación total se estima por la *Suma de Cuadrados Total (SCT)*. Esta es la variación resultante de las desviaciones de las  $Y_i$  con respecto a  $\bar{Y}$ . La fórmula de origen es entonces:

$$SCT = \sum (Y_i - \bar{Y})^2 \quad (66)$$

Sin embargo, la fórmula de origen, por la presencia de decimales y consecuentes errores de redondeo, no provee un resultado exacto. Por lo tanto, mediante procedimientos algebraicos (ver sección 3.4) es conveniente obtener a partir de la fórmula (66), la *fórmula computacional*:

$$SCT = SC_Y = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} \quad (67)$$

**2. Variación por Regresión.** Esta cantidad se calcula por la *Suma de Cuadrados de la Regresión (SCReg)*. La SCReg resulta de las desviaciones de los valores esperados  $\hat{Y}_i$  con relación a la media  $\bar{Y}$ . La fórmula de origen es:

$$SC\ Re\ g = \sum (\hat{Y}_i - \bar{Y})^2 \quad (68)$$

En forma similar al caso anterior, por procedimientos algebraicos (ver sección 3.4) se obtiene, de la fórmula anterior, la fórmula computacional que nos sirve en todos los casos:

$$SC\ Re\ g = \sum \hat{Y}^2 - \frac{(\sum \hat{Y})^2}{n} \quad (69)$$

Sin embargo, en el caso de *regresión lineal simple* (una X y una Y) es más fácil emplear la siguiente fórmula en la que intervienen cantidades ya conocidas:

$$SC\ Re\ g = \frac{(SP_{xy})^2}{SC_x} \quad (70)$$

**3. Variación por Desviación o Error.** Esta variación es calculada por la *Suma de Cuadrados del Error (SCE)*. Es la variación resultante de las desviaciones de los valores observados  $Y_i$  con relación a los valores esperados  $\hat{Y}_i$ . La fórmula de origen, que se presentó en la sección 10.5, es:

$$SCE = \sum (Y_i - \hat{Y}_i)^2$$

Como la SCReg y la SCE son componentes de la SCT, en la práctica la SCE se la encuentra más directamente por la relación:

$$\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2 + \sum (Y_i - \hat{Y}_i)^2 \quad (71)$$

Fórmula de la cual resulta que:

$$SCT = SCReg + SCE \quad (72)$$

y consecuentemente, para encontrar directamente la Suma de Cuadrados del Error:

$$SCE = SCT - SCReg \quad (73)$$

**Ejercicio 11.1.1.** Con los datos del Cuadro 10.12.2 (efecto de cinco niveles de fertilización en el rendimiento de arroz) encuentre los componentes de la varianza, tanto para el modelo lineal como para el modelo cuadrático.

### Modelo Lineal

$$SCT = SCy = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} = 580.29 - \frac{(89.7)^2}{15} = 43.88$$

$$SC Reg = \frac{(SPxy)^2}{SCx} = \frac{(29)^2}{30} = 28.0333$$

$$SCE = SCT - SC Reg = 43.88 - 28.0333 = 15.8467$$

### Modelo Cuadrático:

$$SCT = 43.88$$

$$SC Reg = \sum \hat{Y}^2 - \frac{(\sum \hat{Y})^2}{n} = 577.03 - \frac{(89.70)^2}{15} = 40.62$$

$$SCE = SCT - SC Reg = 43.88 - 40.62 = 3.26$$

**Comentario:** Es importante notar que la Sumatoria de Cuadrados Total ( $SCT$ ) para un mismo grupo de datos, independientemente del modelo de que se trate, siempre será la misma. Lo que cambia, al incluir nuevas variables  $X$  en el modelo, es la Sumatoria de Cuadrados de Regresión ( $SCReg$ ) y la Sumatoria de Cuadrados del Error ( $SCE$ ).

## 11.2. El cuadro de análisis de varianza de la regresión

El análisis de las fuentes de variación se presenta en un cuadro que se denomina *Cuadro de Análisis de Varianza de Regresión* o *Cuadro ADEVA de Regresión*. El

Cuadro ADEVA de Regresión es un caso especial de análisis de varianza. Veremos posteriormente el caso más general de aplicación del Cuadro ADEVA, cuando se evalúan los efectos de tratamientos en experimentos diseñados con ese objetivo.

Un Cuadro ADEVA está constituido por varias columnas e hileras. Las columnas son constantes, mientras que el número y contenido de las hileras depende en gran parte del diseño del experimento, como se verá más adelante. Para los casos que hemos visto de regresión, lo que contiene el Cuadro ADEVA se presenta en el Cuadro 11.2.1:

Cuadro 11.2.1. Componentes del ADEVA de Regresión

fuelle de variación (FV)	suma de cuadrados (SC)	grados de libertad (gl)	cuadrado medio (CM)	Valor F	valor F (5%)	valor F (1%)
Regresión	$SC_{Reg}$	$k - 1$	$CM_{Reg}$	$CM_{Reg}/CME$	(tabla F)	(tabla F)
Error	$SCE$	$n - k$	$CME$			
Total	$SCT$	$n - 1$				

Puede notarse que las columnas de *sumas de cuadrados* y de *grados de libertad* tienen un total que resulta de las sumas de los otros componentes. Así,  $SCT = SC_{Reg} + SCE$ . Igual para los grados de libertad:  $n - 1 = (k - 1) + (n - k)$ . Esta propiedad nos sirve para comprobar si las cantidades encontradas son correctas. (El número de observaciones de la variable Y se denota con  $n$  y con  $k$  el número de parámetros en el modelo. En la regresión cuadrática, por ejemplo,  $k$  es igual a tres: Y, X,  $X^2$ .)

El nombre “análisis de varianza” indica que lo que principalmente nos interesa son los *cuadrados medios o varianzas* (términos sinónimos). Los cuadrados medios o varianzas se obtienen al dividir las sumatorias de cuadrados para los grados de libertad.

El Cuadrado Medio de Regresión:

$$CM_{Reg} = \frac{SC_{Reg}}{k-1} \quad (74)$$

El cuadrado Medio del Error:

$$CME = \frac{SCE}{n-k} \quad (75)$$

La columna del valor  $F$  se refiere a la distribución de Fisher. En los *análisis de varianza* (ADEVA) la prueba estadística es una *prueba F*, cuyo valor resulta del cociente entre dos varianzas o cuadrados medios, el *CMReg* y el *CME*:

$$F = \frac{CMReg}{CME} \quad (76)$$

Esta *prueba F* es similar a la que se usó anteriormente en la sección 9.1 para evaluar la hipótesis  $H_0: \sigma_A^2 = \sigma_B^2$ , en la que  $F = \frac{S_A^2}{S_B^2}$

Sin embargo, como teóricamente el *CMReg* debe ser por lo menos igual o mayor que el *CME*, el *CMReg* siempre va en el numerador. Consecuentemente, la prueba estadística es siempre una prueba de cola derecha.

Las dos últimas columnas del cuadro ADEVA corresponden a los valores críticos, que se los encuentra en la *tabla F* (Apéndice D) con  $k - 1$  grados de libertad para el numerador (*CMReg*) y con  $n - k$  grados de libertad para el denominador (*CME*). Estos son los grados de libertad asociados con cada varianza o cuadrado medio. (Recuerde que  $n$  es el número de observaciones y  $k$  el número de parámetros en el modelo.)

Convencionalmente se pone en el cuadro ADEVA junto al valor encontrado de la prueba estadística *F un asterisco* para representar significación estadística al 5% y *dos asteriscos* para el 1%. La abreviación *n.s.* al lado de la prueba *F* indica que el valor de *F* encontrado no fue estadísticamente significativo. Sin embargo, la evaluación de las hipótesis planteadas puede hacerse también a otros niveles de significación. De hecho, la mayoría de los programas estadísticos computarizados indican la probabilidad ( $P$ ) de cada prueba, permitiendo que el investigador tome la decisión que juzgue pertinente.

**Ejercicio 11.2.1.** Con los resultados obtenidos en el Ejercicio 11.1.1 (datos de niveles de fertilización y rendimiento de arroz del Cuadro 10.12.1) construya un cuadro ADEVA de regresión para el modelo lineal. Decida si acepta o rechaza la hipótesis nula que la variable  $Y$  no depende de  $X$  a los niveles de probabilidad del 5% y del 1%. Esto es igual a probar  $H_0: \beta_1 = 0$ .

Cuadro 1. Análisis de varianza (ADEVA) de regresión simple

FV	SC	gl	CM	F	F <sub>0.05</sub>	F <sub>0.01</sub>
Regresión	28.03	1	28.03	22.98**	4.67	9.07
Error	15.85	13	1.22			
Total	43.88	14				

Sobre la base de lo expuesto en el cuadro anterior, la decisión es rechazar la hipótesis nula al nivel del 1%. En otras palabras, la probabilidad de haber obtenido un valor de F tan extremo como 22.98, sólo por azar, es menor que el 1%.

El programa Statistix presenta simultáneamente los dos métodos de análisis, es decir la prueba t y la prueba F, como puede verse a continuación:

#### Least Squares Linear Regression of RendimY

Predictor Variables	Coefficient	Std Error	T	P
Constant	4.04667	0.49382	8.19	0.0000
fertNivel	0.96667	0.20160	4.79	0.0003

R-Squared 0.6388      Resid. Mean Square (MSE) 1.21928  
 Adjusted R-Squared 0.6110      Standard Deviation 1.10421  
 AICc 9.0092  
 PRESS 21.810

(Cuadro ADEVA DE REGRESION)

Source	DF	SS	MS	F	P
Regression	1	28.0333	28.0333	22.99	0.0003
Residual	13	15.8507	1.2193		
Total	14	43.8840			

Lack of Fit 3      12.6640      4.22133      13.25      0.0008  
 Pure Error 10      3.1867      0.31867

Cases Included 15      Missing Cases 0

### Comentarios:

Note que los dos análisis son equivalentes. En el análisis de los resultados que están sobre el ADEVA se emplea la prueba t. El análisis que se presenta en el cuadro ADEVA emplea, como se indicó, la prueba F. Esto quiere decir que debe haber una relación entre t y F. Esta relación es:  $F = t^2$ . Es decir, el valor de  $F = 22.99$  es igual a  $t^2 = 4.79^2$  (salvo error por redondeo.)

Observe que en el cuadro ADEVA que produce el programa no se presentan columnas con los valores críticos de F al 5% y al 1%. Más bien, con mayor exactitud, señala el área de la curva desde el punto en que cae la prueba estadística hacia el extremo. En este ejercicio, por ejemplo, la probabilidad P es igual a 0.0003. Es decir que, bajo la presunción de que la variable Y no depende de X, la probabili-

dad de un valor de la prueba F tan extremo como 22.99 es solo de 0.03%. Como la probabilidad es tan baja, la decisión debería ser que  $b_1$  no es igual a cero. En otras palabras, se puede concluir que la variable Y depende de la variable X, con una probabilidad de 0.03% de Error Tipo I.

Note también que el CME = 1.2193 está relacionado con el error estándar de regresión  $S_b = 0.2016$ . Esto se demuestra con la siguiente fórmula:

$$\sqrt{\frac{CME}{SCx}} = S_b = \sqrt{\frac{1.2193}{30}} = 0.2016$$

**Ejercicio 11.2.2.** Con la información obtenida en el Ejercicio 11.1.1 presente un cuadro ADEVA de regresión para el modelo cuadrático y pruebe la hipótesis nula de que no existe un efecto cuadrático a los niveles de significación del 5% y del 1%.

Cuadro 1. ADEVA de regresión cuadrática

FV	SC	gl	CM	F	$F_{0.05}$	$F_{0.01}$
Regresión	40.63	2	20.31	74.88**	3.49	6.93
Error	3.26	12	0.27			
Total	43.88	14				

Como el valor de la F calculada (74.86) excede el valor crítico al nivel del 1%, la decisión es rechazar la hipótesis nula. En otras palabras, hay evidencia estadística de la presencia de un efecto cuadrático.

Los resultados con el programa Statistix son:

Polynomial Regression of RendimY  
Diferenced Scale (i.e.,  $X_i - \bar{X}$ )

Predictor Variables	Coefficient	Std Error	T	P
Constant	7.07524	0.20958	33.76	0.0000
fertNivel	0.96667	0.09509	10.17	0.0000
fertNivel^2	-0.54762	0.08037	-6.81	0.0000

Original Scale

Predictor Variables	Coefficient	Std Error
Constant	2.95143	0.28301
fertNivel	3.15714	0.33525
fertNivel^2	-0.54762	0.08037

R-Squared	0.9258	Resid. Mean Square (MSE)	0.27129
Adjusted R-Squared	0.9135	Standard Deviation	0.52085
AICc	-10.916		
PRESS	4.5972		

**Anov table**

Source	DF	SS	MS	F	P
Regression	2	40.6286	20.3143	74.88	0.0000
Residual	12	3.2554	0.2713		
Total	14	43.8840			

Lack of Fit	2	0.06876	0.03438	0.11	0.8988
Pure Error	10	3.18667	0.31867		

Cases Included 15    Missing Cases 0

**11.3. Relación entre  $t$  y  $F$** 

Si es que tanto la prueba estadística  $t$  como la  $F$  pueden ser usadas para evaluar una misma hipótesis ( $H_0: B_1 = 0$  y otras que veremos posteriormente), debe entonces existir alguna relación entre  $t$  y  $F$ . Esta relación es la siguiente:  $t^2 = F$

Como la prueba  $t$  tiene sólo un grado de libertad en el numerador, la relación se cumple siempre que se trate de regresión lineal, que es cuando la Sumatoria de Cuadrados de la Regresión (SCReg) es igual al Cuadrado Medio de la Regresión (CMReg). Es decir que:

$$CMReg = \frac{SCReg}{1}$$

Para demostrar que  $t^2 = F$  partimos de las pruebas estadísticas correspondientes, donde la  $PEt$  está elevada al cuadrado. En la ecuación de la prueba  $F$  podemos poner la SCReg a cambio del CME por ser equivalentes en este caso:

$$t^2 = \frac{b_1^2}{s_b^2} \text{ y } F = \frac{SCReg}{CME}$$

$$\text{Entonces: } \frac{b_1^2}{s_b^2} = \frac{SCReg}{CME}$$

$$\text{Conocemos que: } SCReg = b_1 SP_{XY} \text{ y } CME = S_e^2 = S_b^2 SC_X$$

$$\text{Por lo tanto: } \frac{SCReg}{CME} = \frac{b_1 SP_{XY}}{S_b^2 SC_X}$$

Pero como:  $\frac{SP_{XY}}{SC_X} = b_1$ , entonces  $\frac{SC_{Reg}}{CME} = \frac{b_1^2}{S_b^2}$

Que es lo que se quería demostrar.

La relación matemática que existe entre  $t$  y  $F$  se mantiene también, obviamente, en las tablas respectivas (ver Apéndices B y D). La relación existente entre  $t$  y  $F$  se demuestra en el siguiente gráfico:

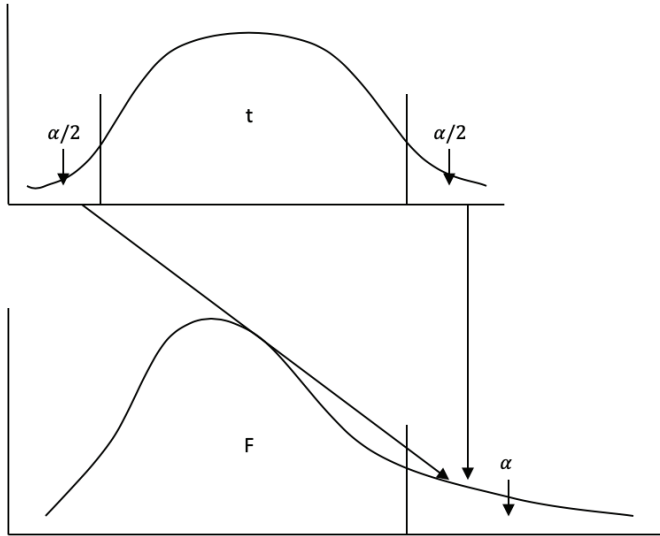


Gráfico 11.3.1. Relación entre las distribuciones  $t$  y  $F$  y las áreas críticas respectivas

Esta relación directa entre  $t$  y  $F$  ( $Y_i = b_0 + b_1$ ) se presenta cuando el número de grados de libertad para el numerador en la tabla  $F$  es igual a uno.

Veremos más adelante cómo mediante la prueba  $F$ , que es una prueba de varianzas, podemos realizar pruebas relacionadas con medias de poblaciones.





## **Análisis de correlación simple**

---

*“Una de las primeras cosas que enseñan los libros de estadística es que la correlación no implica causalidad. También es una de las primeras cosas que se olvidan”.*  
Thomas Sowell<sup>21</sup>

*Correlación* es una medida del grado de asociación de dos variables. La correlación determina la intensidad con la que las variables cambian en forma simultánea. La definición de correlación, como puede verse es, en parte, similar a la empleada para regresión, sin embargo, en correlación las dos variables son consideradas como independientes. En algunos casos la relación entre las dos puede deberse al efecto de una tercera variable.

La correlación entre variables cuantitativas se denomina “Correlación de Pearson” en homenaje a Karl Pearson, el científico que desarrolló este análisis. Cuando se trata de variables cualitativas, el procedimiento que se emplea es la “Correlación de Spearman”, que se verá en el Capítulo 24.

La correlación es un campo importante en las ciencias biológicas y en especial en el mejoramiento genético de plantas y animales. El conocimiento de la existencia de un grado de asociación entre dos variables puede ayudar en la selección de mejores materiales genéticos. Si se ha establecido, por ejemplo, que el nivel de resistencia a una enfermedad está positivamente correlacionada con la intensidad de la pigmentación en las hojas de un cultivo, podríamos obtener un cultivar resistente al seleccionar plantas que presenten la pigmentación adecuada.

Como en todo aspecto de la inferencia estadística debe tenerse cuidado en la interpretación de la correlación. Puede ser que la asociación de dos variables sea puramente casual y que no tenga ningún significado biológico (*correlación espuria*).

---

21 Economista y pensador estadounidense. <https://www.goodreads.com/quotes/3237859-one-of-the-first-things-taught-in-introductory-statistics-textbooks>

La correlación que estudiaremos es la *correlación lineal simple*. Se denomina correlación lineal porque se presume que en la población existe una relación lineal entre las variables y es simple porque se trata únicamente de dos variables.

## 12.1. Relación entre regresión y correlación

La correlación se mide a través del *coeficiente de correlación de la muestra*. Este coeficiente se lo representa con la letra  $r$ . La estadística  $r$  está relacionada con el coeficiente de regresión  $b_1$  ya que los dos se derivan de la covarianza.

Para demostrar lo anterior partamos de la ecuación que indica la regresión de Y sobre X. Esto es:

$$b_1 = b_{Y.X} = \frac{SP_{XY}}{SC_X}$$

Si se tratara de la regresión de X sobre Y la ecuación sería:

$$b_{X.Y} = \frac{SP_{XY}}{SC_Y}$$

La fórmula para el *coeficiente de correlación* proviene de las dos fórmulas anteriores. La relación es la siguiente:

$$r = \sqrt{b_{Y.X} * b_{X.Y}} \quad (77)$$

Sin embargo, es más fácil encontrar el valor de  $r$  a partir de las cantidades ya conocidas: Sumatoria de Productos XY ( $SP_{XY}$ ), Sumatoria de Cuadrados X ( $SC_X$ ), y la Sumatoria de Cuadrados de Y ( $SC_Y$ ). Consecuentemente, la fórmula más comúnmente usada para encontrar el coeficiente de correlación es:

$$r = \frac{SP_{XY}}{\sqrt{SC_X * SC_Y}} \quad (78)$$

Observe en las fórmulas anteriores que en el *análisis de correlación* también nos referimos a las variables como X y Y, pero sin que esto implique dependencia alguna. En otras palabras, en correlación tanto la variable X como la variable Y son consideradas independientes.

## 12.2. Las unidades y límites del coeficiente de correlación

Por la definición de correlación y al observar la fórmula (77) puede deducirse que la relación entre las dos variables es simétrica.

Al contrario de otras estadísticas como la media, la varianza, la desviación estándar o el coeficiente de regresión, el *coeficiente de correlación* no tiene unidad de medida. Esto también puede deducirse de la fórmula de  $r$ . Por ejemplo, si  $X$  = altura y  $Y$  = peso, las unidades en la fórmula serán las de altura y peso:

$$r = \frac{\text{altura} \cdot \text{peso}}{\sqrt{\text{altura}^2 \cdot \text{peso}^2}}$$

Al simplificar se demuestra que  $r$  no tiene unidad de medida alguna.

Los límites del valor de  $r$  son:  $-1 \leftrightarrow +1$ . Un valor de  $r = -1$  demuestra una relación inversa perfecta, mientras que  $r = +1$  demuestra una relación directa perfecta.

### 12.3. Evaluación de hipótesis en correlación

A semejanza de lo estudiado en capítulos anteriores,  $r$  es el coeficiente de correlación de la muestra y es el mejor estimador del *coeficiente de correlación de la población*, parámetro  $\rho$  que se representa con la letra griega “rho”. Por lo tanto, en correlación se pueden evaluar las siguientes hipótesis:

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

La prueba estadística que se emplea es una prueba  $t$ , que como se puede haber supuesto, está relacionada con la prueba estadística empleada para el coeficiente de regresión.

La prueba estadística es:

$$PEt = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \quad (79)$$

Los valores críticos para esta prueba se los encuentra en la tabla  $t$  con  $n-2$  grados de libertad.

La prueba puede ser de una cola o de dos colas de acuerdo con el conocimiento y experiencia que tenga el investigador sobre las variables en estudio.

Matemáticamente se puede demostrar que:

$$\frac{b_{Y.X}}{\sqrt{S_b^2}} = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$$

Por lo tanto, esta prueba estadística  $t$  sirve para discernir sobre la presencia de una relación o no de las dos variables, a un determinado nivel de significación.

Es decir que las hipótesis nulas,  $H_0: \beta_1 = 0$  vista anteriormente, y  $H_0: \rho = 0$ , son matemáticamente equivalentes, aunque tienen diferente interpretación biológica.

Como se indicó,  $r$  es el coeficiente de correlación de la muestra. Se parte de la presunción de que la muestra fue tomada al azar de dos poblaciones normalmente distribuidas. Esto permite que  $r$  sea el mejor estimador del coeficiente de correlación de la población  $\rho$  (letra griega “rho”).

**Ejercicio 12.3.1.** En un experimento los rendimientos de maíz provenientes de cruza simples medio hermanas fueron los siguientes:

Cuadro 1. Rendimientos en t/ha de dos cruza de maíz con un progenitor común

Sitios	1	2	3	4	5	6	7	8
Cruza AB, X	4.2	5.3	2.4	2.5	3.2	3.1	4.2	3.0
Cruza AC, Y	3.5	3.8	3.0	2.9	3.2	4.0	4.1	3.6

La denominación X y Y es únicamente para identificación y no indica grado de dependencia alguno.

Pruebe que el rendimiento de las cruza simples está positivamente correlacionado. Es decir, pruebe las siguientes hipótesis al nivel de significación  $\alpha = 0.05$ :

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

Para esto debemos primero encontrar las siguientes estadísticas:

$$n, \sum X_i, \sum X_i^2, \sum Y_i, \sum Y_i^2, \sum X_i Y_i,$$

$$n = 8$$

$$\sum X_i = 27.9 \quad \sum X_i^2 = 104.23 \quad SCX = 6.93$$

$$\sum Y_i = 28.10 \quad \sum Y_i^2 = 100.11 \quad SCY = 1.41$$

$$SPXY = 1.95$$

El valor de  $r$  es:

$$r = \frac{SP_{XY}}{\sqrt{SC_X * SC_Y}} = \frac{1.95}{\sqrt{6.39 * 1.41}} = 0.62$$

La prueba estadística  $t$  es:

$$PEt = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.62\sqrt{8-2}}{\sqrt{1-0.62^2}} = 1.94$$

En el Apéndice B, el valor crítico de  $t$  al nivel  $\alpha/2 = 0.025$  con seis grados de libertad es 2.45

Como el valor de la prueba estadística no excede al valor crítico, la decisión es aceptar la  $H_o$  que  $\rho = 0$ . En otras palabras, no existe suficiente evidencia estadística para rechazar la hipótesis nula.

Se esperaba que las dos variables estuvieran positivamente correlacionadas por razón de que los híbridos en cada caso tenían un progenitor común. Posiblemente el reducido tamaño de la muestra impidió que se detectara una correlación significativa, a pesar del relativamente elevado valor de  $r$ .

El programa Statistix, en el sub-menú correspondiente, presenta el coeficiente de correlación  $r$  y la probabilidad de encontrar un valor tan alto como el obtenido:

```
CRUZAS
Statistix 9.0
```

**correlations (Pearson)**

```

                X
Y                0.6246
P-VALUE         0.0979
```

```
Cases Included 8      Missing Cases 0
```

Note el valor de  $r = 0.62$  y la probabilidad de 9.79%

Otro ejemplo de correlación se presenta en el siguiente ejercicio:

**Ejercicio 12.3.2.** Los siguientes datos corresponden a una investigación en la que se trató de determinar si los rendimientos de maíz y soya, sembrados en secuencia en una zona ecológicamente similar, estaban relacionados. Para esto se seleccio-

naron al azar fincas que cultivan maíz en el invierno y luego soya en el verano, en condiciones de secano. Los resultados se presentan en el siguiente cuadro:

**Cuadro 1. Producción de maíz (invierno) y soya (verano), en t/ha.**

Finca No.	1	2	3	4	5	6	7	8	9	10
Maíz, X	4.49	4.03	4.52	1.55	3.72	2.85	2.45	4.19	3.04	5.85
Soya, Y	1.85	1.52	2.09	1.02	1.46	2.11	0.80	1.82	1.48	1.97

Evaluar las siguientes hipótesis al nivel de significación  $\alpha = 0.05$  :

$$H_0: \rho = 0$$

$$H_1: \rho \neq 0$$

Para encontrar el coeficiente de correlación de la muestra se requieren las siguientes cantidades de acuerdo con la fórmula (77):

$$SC_X = 13.61 \quad SC_Y = 1.76 \quad SP_{XY} = 3.35$$

Consecuentemente, el valor del coeficiente de correlación de esta muestra es:

$$r = \frac{SP_{XY}}{\sqrt{SC_X * SC_Y}} = 0.68$$

La prueba estadística t es igual a:

$$PEt = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.68\sqrt{8}}{\sqrt{1-0.46}} = 2.63$$

En el Apéndice B, el valor de t al nivel  $\alpha/2 = 0.025$  con ocho grados de libertad es 2.31. Por lo tanto, existe suficiente evidencia para rechazar la hipótesis nula. En este caso se puede señalar que los rendimientos de maíz en invierno y soya en verano están positivamente correlacionados. Desde el punto de vista agronómico este resultado podría explicarse por razón del manejo del cultivo de maíz en el invierno, lo que incide en el rendimiento del siguiente cultivo.

El programa Statistix, en el sub-menú correspondiente, presenta el coeficiente de correlación r y la probabilidad de un valor tan alto como el obtenido:

FINCAS

Statistix 9.0

**Correlations (Pearson)**

MAIZtha

SOYAtha 0.6847

P-VALUE 0.0289

Cases Included 10 Missing Cases 0

## 12.4. Correlación espuria

Como se señaló anteriormente, en correlación se presume que las dos variables son independientes. Cuando se encuentra una correlación significativa, el investigador tiene que buscar una *explicación lógica y de sentido común* para ese resultado. Lo más probable es que la correlación se deba al efecto de una tercera variable. Es decir, se puede determinar una posible causa de ese resultado. Los dos ejercicios anteriores (Sección 12.3) son ejemplos de correlaciones que tienen una explicación lógica.

Una frase favorita de los estadísticos es que la correlación no implica causa. Sin embargo, en la literatura se pueden encontrar cientos de ejemplos del mal uso de esta técnica, en los que se presentan como causales relaciones que no guardan lógica.

Como ejemplo, en una investigación publicada en el *New England Journal of Medicine*<sup>22</sup> se reporta una correlación positiva,  $r = 0.79$  con una probabilidad de error  $P < 0.0001$ , entre el consumo de chocolate en 14 países del primer mundo y el número de premios Nobel. La conclusión del autor es que el chocolate aumenta la capacidad intelectual de las personas, lo que conduce a ganar premios Nobel. Bien por el chocolate, pero en las pequeñas fincas de América Latina miles de familias campesinas han consumido casi diariamente por siglos chocolate artesanal de su propia producción y hasta ahora no ha asomado ningún premio Nobel. Parece más lógico pensar que los premios Nobel son el resultado de buenos sistemas educativos y de un gran apoyo a la investigación científica. Una correlación entre los recursos asignados para la investigación y el número de premios Nobel seguramente hubiera también resultado significativa, aunque el resultado no fuera tan novedoso como en el caso del chocolate.

22 Messerli, Franz. Annual per capita chocolate consumption and the number of Nobel laureates. *New England Journal of Medicine*, 2012; 367:1562-1564



Que la cantidad de cigüeñas en Europa está correlacionada positivamente con el número de nacimientos, o que el consumo de helados aumenta el número de personas ahogadas en el verano, son otros ejemplos de correlación espuria que pueden encontrarse en la web.

## 12.5. El coeficiente de determinación

Como se vio en el capítulo anterior, *el coeficiente de determinación* se define como la proporción de la variabilidad de  $Y$  que puede ser explicada por su asociación con una  $X$  o varias  $X$ .

Si el valor del coeficiente de correlación se eleva al cuadrado, entonces se obtiene el *coeficiente de determinación*. El coeficiente de determinación se expresa generalmente en porcentaje. Es una estadística que nos ayuda a interpretar los resultados tanto en regresión como en correlación. La fórmula más comúnmente usada para el *coeficiente de determinación simple* (sólo una variable  $X$ ) es:

$$r^2 = \frac{(SP_{XY})^2}{SC_X * SC_Y} \quad (80)$$

En el Ejercicio 11.2.1 el valor del coeficiente de determinación  $r^2$  fue 0.6388. Esto significa, en el ejemplo en mención, que la variabilidad en el rendimiento del arroz ( $Y$ ) depende en un 63.88% de la variable independiente “niveles de fertilización” ( $X$ ).

Cuando se trata de regresión polinómica o múltiple, es decir una variable  $Y$  y dos o más variables  $X$ , se emplea la notación  $R^2$ . Esta estadística se denomina *coeficiente de Determinación Múltiple*. Para este caso la fórmula, derivada de las anteriores, es:

$$R^2 = \frac{SC_{Reg}}{SCT} \quad (81)$$

En el Cuadro de Análisis de Varianza de Regresión (Cuadro ADEVA de Regresión) se encuentran, entre otra información, la Sumatoria de Cuadrados de Regresión ( $SC_{Reg}$ ) y la Sumatoria de Cuadrados Total ( $SCT$ ). Debe recordarse que esta última es igual a la Sumatoria de Cuadrados de  $Y$ , es decir  $SCT = SC_Y$ . (La fórmula (81) puede también emplearse para encontrar el valor de  $r^2$  a partir del Cuadro de Análisis de Varianza de Regresión, como se demuestra en el Ejercicio 12.5.1.)

**Ejercicio 12.5.1.** a) Con los datos del Ejercicio 12.3.1 encuentre el coeficiente de determinación empleando la fórmula (80):

$$r^2 = \frac{(SP_{XY})^2}{SC_X * SC_Y} = \frac{(1.95)^2}{6.93 * 1.41} = 0.39, \text{ lo que es igual a } 39\%$$

b) con los mismos datos del Ejercicio 12.3.1, pero ahora empleando la fórmula (81), encuentre el valor del coeficiente de determinación.

Para esto necesitamos encontrar primero la *SCReg* y la *SCT*:

$$SCReg = \frac{(SP_{XY})^2}{SC_X} = \frac{(1.95)^2}{6.93} = 0.55$$

$$SCT = SC_Y = 1.41$$

$$r^2 = \frac{SCReg}{SCT} = \frac{0.55}{1.41} = 0.39, \text{ es decir } 39\%$$

Se puede señalar entonces que 39% de la variación de *Y* puede ser explicada por su relación con *X* y viceversa. En otras palabras, 39% de la variación en producción de las cruza simples es atribuible al hecho de tener un progenitor común. El 61% de la variación restante se debe a otros factores, como posiblemente: tipo de suelo, humedad, fertilización, métodos de cultivo, etc.

El siguiente ejercicio trata del coeficiente de determinación múltiple:

**Ejercicio 12.5.2.** En el Ejercicio 11.2.2 se presentó el siguiente Cuadro ADEVA de Regresión de *Y* sobre *X* y  $X^2$  (regresión polinómica):

Cuadro 1. Análisis de varianza de regresión

FV	SC	gl	CM	F	F <sub>0.05</sub>	F <sub>0.01</sub>
Regresión	40.63	2	20.31	74.88**	3.49	6.93
Error	3.26	12	0.27			
Total	43.88	14				

Encontrar el coeficiente de determinación múltiple:

$$R^2 = \frac{CMReg}{SCT} = \frac{40.63}{43.88} = 0.9258, \text{ es decir } 92.58\%$$

El resultado anterior indica que la variación de *Y* se explica en un 93% por la dependencia con *X* y  $X^2$ .

### Coefficiente de determinación ajustado

Con la inclusión de varias variables independientes en el modelo de regresión puede presentarse un sobre dimensionamiento de la estadística  $R^2$ . En estos casos es conveniente encontrar el *coeficiente de determinación ajustado*  $R_{aj}^2$ . La fórmula es:

$$R_{aj}^2 = 1 - \frac{(n-1)*SCE}{(n-k)*SCT} \quad (82)$$

Fórmula en la que  $(n - 1)$  son los grados de libertad de la SCT y  $(n - k)$  son los grados de libertad de la SCE.

**Ejercicio 12.5.3.** Resultados parciales obtenidos con el programa Statistix con los datos del ejercicio 11.2.2, se presentan a continuación:

R-Squared                      0.9258  
Adjusted R-Squared        0.9135

Source	DF	SS	MS	F	P
Regression	2	40.6286	20.3143	74.88	0.0000
Residual	12	3.2554	0.2713		
Total	14	43.8840			

Compruebe los resultados, empleando la fórmula (81), que el coeficiente de determinación ajustado es igual a 0.9135

$$R_{aj}^2 = 1 - \frac{(n-1)*SCE}{(n-k)*SCT} = 1 - \frac{14*3.2554}{12*43.8840} = 0.9135, \text{ es decir } 91.35\%$$

Nota: El programa Statistix siempre presenta el valor  $R^2$  y el valor  $R_{aj}^2$  en los análisis de regresión múltiple.

# Introducción a diseño de experimentos

*“El estadístico que cree que su mayor contribución al planificar un experimento tiene que ver con la teoría estadística, encuentra que su mayor aporte lo hace simplemente al persuadir al investigador que explique por qué desea llevar a cabo ese experimento”. Gertrude M. Cox<sup>23</sup>*

## **El origen de la investigación**

El espectacular desarrollo de nuestra civilización se debe a la acumulación de conocimiento adquirido a través de los sentidos y almacenado en las células cerebrales. Pero el simple ver, oler, sentir, gustar y escuchar no hubiera conducido al desarrollo actual de la civilización a no ser por el innato deseo de saber el porqué de las cosas. Sin que seamos únicos en muchos aspectos, esta característica es posiblemente la que más nos diferencia de otros miembros del reino animal.

El procedimiento de observar hechos y tratar de explicarlos es la forma de acción de los “genes de la curiosidad” de nuestra especie. A este procedimiento se lo conoce como el *método científico*.

## **El método científico**

El método científico<sup>24</sup> se define como el conjunto de pasos que siguen los científicos para encontrar posibles explicaciones de la ocurrencia de un fenómeno. Un ejemplo clásico de la aplicación del método científico es el trabajo del botánico holandés Van Helmont en el siglo XVII (ver la Sección 1.4).

---

23 Profesora de Estadística estadounidense. <https://www.stat.wisc.edu/quotes>

24 En la Sección 1.4 se presentó el método científico y el experimento del Van Helmont, pero es importante mencionarlo de nuevo en esta Sección.

El procedimiento que siguió Van Helmont fue impecable y rigurosamente científico, a pesar de que la conclusión a la que arribó fue errónea. Obviamente, no se conocía el fenómeno de la fotosíntesis ni la existencia del dióxido de carbono. Sobre la base de los conocimientos actuales se le podría criticar que no tuvo repeticiones en diferentes ambientes. Una repetición del experimento en condiciones de oscuridad seguramente le hubiera llevado a otra conclusión. Pero no olvidemos que se trató de un experimento pionero realizado en el siglo XVII (el experimento concluyó en 1648). La introducción de repeticiones en los experimentos fue una innovación creada por Ronald Fisher a inicios del siglo XX.

## El diseño de experimentos

Por Diseño de Experimentos se entiende el proceso de planificación de los experimentos, la selección del material experimental, la asignación al azar de los tratamientos a las unidades experimentales, la forma de conducción de los experimentos y el análisis de los datos.

El científico inglés Ronald Fisher<sup>25</sup> inventó los diseños experimentales, el análisis de varianza y la prueba F (por Fisher) entre los años 20 y 30 del siglo XX. Su libro, “The Design of Experiments” se publicó en 1935 y ha tenido hasta ahora 21 ediciones. En su trabajo en la Estación Experimental Agrícola de Rothamsted, R. Fisher desarrolló los conceptos de *repetición*, *error experimental* y *aleatorización* que son esenciales en todo diseño experimental. Estos tres conceptos están íntimamente relacionados. Es importante resaltar que las innovaciones de Ronald Fisher aportaron enormemente al desarrollo de las ciencias, no sólo en el campo agrícola sino también en otros campos del conocimiento.

### 13.1 . El concepto de repetición

Se entiende por repetición la aplicación de un mismo tratamiento a más de una unidad experimental. Si un tratamiento se aplicara a una sola unidad experimental, entonces no habría forma de separar el efecto del tratamiento del efecto inherente a la unidad experimental. Por ejemplo, si se dispone únicamente de dos unidades experimentales (parcelas, en este caso) y se observa que la variedad A de maíz produce 10 kg en una de ellas, mientras que la variedad B produce 15 kg en la otra parcela, no habría forma de saber si la diferencia se debe a que la variedad

---

25 <https://www.famousscientists.org/ronald-fisher/>

B es mejor que la A o que por azar le tocó una unidad experimental con un nivel superior de fertilidad.

En los experimentos la repetición tiene varias funciones:

- Proporciona una estimación del error experimental.
- Reduce el error estándar de la media de los tratamientos, en cuya fórmula consta  $r$  que es el número de repeticiones:  $S_{\bar{y}} = \sqrt{CME/r}$ . A mayor número de repeticiones mayor es la precisión del experimento. En este sentido por precisión se entiende la mayor probabilidad de tomar una decisión correcta con relación a la hipótesis nula.
- Aumenta la amplitud de las inferencias que puedan hacerse al emplear adecuadamente unidades experimentales más variables.

En experimentos en el campo, una repetición es un área con varias unidades experimentales o parcelas que recibieron cada una un tratamiento diferente (en algunos diseños esto recibe el nombre de “bloque”). Por lo general, todos los tratamientos están representados una vez en cada repetición. En los experimentos se requiere que por lo menos haya dos repeticiones de cada tratamiento (dos parcelas que reciban el mismo tratamiento) para poder estimar el error experimental. No es necesario que las repeticiones estén físicamente juntas. Pueden estar ubicadas en diferentes localidades y también en diferentes épocas.

### 13.2. El concepto de error experimental

El error experimental es el resultado de la variación que se genera en las unidades experimentales que recibieron el mismo tratamiento.

Suponga un experimento en el que se desea comparar la producción de dos variedades de maíz, sembradas cada una en una parcela. En este experimento la producción proviene de dos componentes que se suman (son aditivos): el efecto de la variedad y el efecto de la parcela. Estos efectos están fusionados y no hay forma de separarlos. En otras palabras, no se puede saber si la producción observada se debe a diferencias entre las variedades, a diferencias entre las parcelas, o a una combinación de los dos efectos. Es decir, no se puede estimar el error experimental si se tiene una sola parcela (o repetición).

En los experimentos el objetivo es medir el efecto de cada uno de los tratamientos. Para esto se suman por separado las observaciones de cada tratamiento en todas las repeticiones. Pero luego se necesita una medida, un patrón de comparación. Esa medida es un valor que se obtiene a partir de las parcelas que recibie-

ron un mismo tratamiento. Ese valor se denomina *error experimental*. Pero sólo es posible estimarlo si se tiene por lo menos dos repeticiones para cada tratamiento.

El error experimental es un componente inevitable en los experimentos. Lo que se busca con una buena planificación y ejecución es que sea lo más pequeño posible. Un error experimental muy inflado tapa el efecto de los tratamientos y conduce a conclusiones erróneas.

## Las fuentes del error experimental

El error experimental proviene de dos fuentes: a) de la variación inherente (propia) de las unidades experimentales y b) de la variación introducida por técnicas defectuosas en el manejo del experimento (errores humanos y mecánicos).

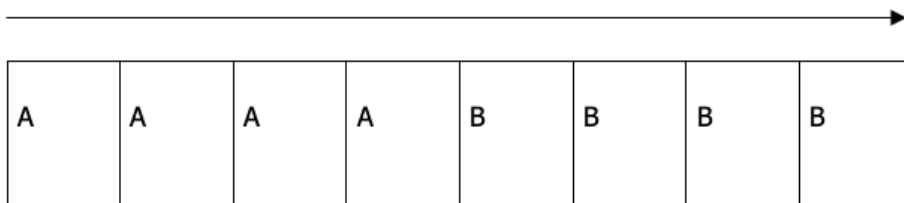
- a. **Variación inherente a las unidades experimentales.** No hay parcelas en el campo, o plantas de un mismo híbrido, o animales de una misma raza exactamente iguales. En experimentos bajo condiciones controladas, como en el caso de invernaderos y cámaras de crecimiento, siempre habrá variación. En organismos homocigóticos e incluso en clones la variación está presente. La norma de la naturaleza es la variación. Los organismos presentan variación genética, variación ambiental y variación que resulta de la interacción genotipo-ambiente. Esta es la variación inherente que en los experimentos es parte del error experimental. Sin embargo, el éxito de una investigación depende en gran medida de la selección de unidades experimentales lo más homogéneas que sea posible.
- b. Para disminuir la variación inherente se puede seleccionar unidades experimentales homogéneas, agrupar unidades experimentales similares en bloques, emplear la técnica de covarianza. Estas alternativas pueden usarse solas o en conjunto.
- c. **Variación por errores humanos y mecánicos.** Esta es la variación que se produce por poca uniformidad en la conducción del experimento. Cuando la técnica con la que se conduce un experimento es deficiente, el error experimental se incrementa. Consecuentemente la probabilidad de encontrar diferencias reales entre los tratamientos disminuye. Si por ejemplo se trata de un experimento que tiene como objetivo comparar varias dosis de un fertilizante, la variación puede originarse por fluctuaciones en la aplicación de la dosis programada, por incorporación desigual del fertilizante al suelo, por desigual calidad del producto en diferentes bolsas, etc.

Para disminuir la variación causada por errores humanos y mecánicos se debe tener cuidado en la ejecución del experimento. es necesario evitar el agotamiento mental y físico y supervisar adecuadamente el trabajo de colaboradores. Cualquier actividad como aplicación de productos, toma de datos, etcétera, debe efectuarse por bloques y no por tratamientos. Es también necesario revisar el equipo disponible. Una balanza defectuosa puede echar a perder toda una investigación.

Es importante recordar que mientras más bajo sea el error experimental mayor será la precisión del experimento.

### 13.3. El concepto de Aleatorización

Para realizar una estimación válida del error experimental no debe introducirse ningún sesgo que pueda favorecer a uno o varios tratamientos en detrimento de otros. Suponga un experimento en el que se siembran sistemáticamente dos variedades A y B (o tratamientos) en un grupo de parcelas. Un mapa del campo sería como sigue:



Si existiera una gradiente de fertilidad de más a menos en el sentido que indica la flecha, el tratamiento A resultaría sistemáticamente favorecido porque ocupa un mejor terreno que el tratamiento B. Esto invalida los resultados porque parte de la diferencia entre A y B va a deberse a que la variedad A creció en un suelo con mayor fertilidad. Uno de los supuestos en que se basa el análisis de varianza es precisamente que el error siga una distribución normal y que sea independiente del efecto de las parcelas.

Para que los resultados y el análisis sean válidos, se deben eliminar los posibles sesgos. Esto se logra mediante la *aleatorización o sorteo*. La asignación al azar de los tratamientos a las unidades experimentales busca que todos los tratamientos tengan igual probabilidad de ser ubicados en cualquier unidad experimental, dentro de las restricciones que imponga el diseño que se use, como se verá más adelante.



### 13.4. Procedimientos para la experimentación

Para el sometimiento de la hipótesis a prueba se requiere de procedimientos experimentales adecuados, que nos permitan obtener la mayor información posible con relación a la inversión de recursos. Es importante tener en cuenta los siguientes aspectos:

1. **Definición del problema.** Este es el paso más importante. Si no se ha definido con claridad el problema que se va a investigar, muy difícilmente se podrá conducir una investigación exitosa.
2. **Determinación de los objetivos.** Los objetivos deben guardar estrecha relación con la hipótesis planteada. Si la hipótesis es que el maíz responde a aplicaciones de roca fosfórica, un objetivo puede ser evaluar el rendimiento de varios híbridos de maíz en respuesta a diferentes cantidades de roca fosfórica. Los objetivos deben ser claros y específicos. Cuando los objetivos de una investigación son ambiguos o demasiado ambiciosos, el trabajo no aportará mayor cosa hacia la resolución del problema.
3. **Selección de los tratamientos.** Los tratamientos constituyen las variables independientes cuyos efectos sobre las unidades experimentales deseamos medir. La selección de los tratamientos deberá hacerse teniendo en cuenta los objetivos de la investigación. Siempre debe incluirse un tratamiento “testigo” que sirve como patrón de comparación del efecto de los otros tratamientos.
4. **Selección de las unidades experimentales.** Las unidades experimentales son las entidades que reciben los tratamientos cuyos efectos deseamos medir. Dependiendo del experimento, la unidad experimental puede ser una parcela en el campo, una maceta en el invernadero, un tubo de ensayo, un plato Petri, un animal, un grupo de animales, un árbol, una trampa para insectos, etc. La unidad experimental debe ser representativa de la población de referencia, es decir de aquella a la que se aplicarán las conclusiones y recomendaciones.
5. **Diseño experimental.** Las características de las unidades experimentales determinan el diseño que deba usarse. El mejor diseño no es el más grande ni el más complicado, sino aquel que conduce a resultados confiables. En todo caso, hay que buscar el diseño más simple posible. Es necesario recordar siempre que los diseños experimentales y los análisis estadísticos no son un fin sino un medio. Son herramientas de la investigación.
6. **Determinación del número de repeticiones.** Basándose en conocimientos anteriores del campo de investigación y procedimientos matemáticos se

puede determinar el número adecuado de repeticiones, en función de la cantidad de información y precisión deseada. Mucho de esto depende de la variación existente en el material experimental. Algunos textos indican que la combinación entre el número de tratamientos y repeticiones debe proveer por lo menos de 20 grados de libertad para el error experimental, pero esto no es una regla fija y depende de la variabilidad del material experimental.

7. **Mediciones por hacerse.** Esto se refiere a las variables dependientes o de respuesta que se obtendrán como resultado de la aplicación de los tratamientos. Deben ser parte de la planificación del experimento. Son ejemplos de variables dependientes: altura de planta, ganancia de peso en animales, peso del grano, incidencia de enfermedades, daño por insectos, etc. Las mediciones deben ser relevantes y estar relacionadas con los objetivos.
8. **Esquema del análisis de varianza.** Es necesario elaborar un esquema del análisis de varianza en el que consten, principalmente, las fuentes de variación y los grados de libertad. Este esquema ayuda a identificar posibles errores en la planificación del experimento. En este punto es importante también señalar el método de separación de medias se empleará.
9. **Conducción del experimento.** Los pasos anteriores corresponden a la planificación del experimento. La conducción del experimento se refiere a la ejecución del trabajo. Lo que se intenta es llevar a cabo el experimento con el menor número posible de errores. Los errores mecánicos y humanos aumentan el error experimental y consecuentemente disminuye la probabilidad de detectar diferencias reales que podrían existir entre tratamientos. En otras palabras, el experimento pierde validez y no hay confianza en los resultados obtenidos. Se debe tener cuidado especialmente en la medición de las parcelas, el pesaje de los insumos, la aplicación de los tratamientos, el manejo agronómico del ensayo, la toma de datos y la transcripción de los datos a algún programa de computación.
10. **Efectos de borde.** En experimentos de campo, las filas (o surcos) de los extremos deben eliminarse. El nivel de competencia por luz, agua y nutrientes es diferente para esas filas con relación a las filas del centro. También deben eliminarse las plantas de los extremos de las filas. El área que queda se denomina *parcela útil*.
11. **Análisis de los datos e interpretación.** Una vez concluido el experimento se procede al análisis de los datos en la forma planificada de antemano. La interpretación de los resultados debe hacerse a la luz del conocimiento actual en ese campo o campos afines, pero siempre atentos ante la posibilidad de hechos que apoyen la hipótesis o que proporcionen argumen-

tos para modificarla o rechazarla. No importa cuán bien se conduzca un experimento si al momento de la interpretación de los resultados se cometen errores. En todo caso, y sobre todas las cosas, el investigador debe ser honesto y objetivo. No es ético el descartar o modificar resultados que no nos agradan. Además, podría ser que algún resultado inesperado sea la pista necesaria para el descubrimiento de algo nuevo.

12. **Conclusiones y recomendaciones.** las conclusiones y recomendaciones deben basarse en los resultados obtenidos y tomando en cuenta la población original de referencia. Un experimento es un modelo en pequeño que trata de determinar cuáles serían los resultados en una situación real. Es decir que a partir de la muestra se infiere características de la población. Si en el modelo (o experimento) se determina que la variedad A rinde mejor que la B en suelos arcillosos y siembras de invierno, no quiere decir que A necesariamente rendirá mejor que B en otras condiciones.

### 13.5. Las pruebas $t$ y el diseño de experimentos

En el Capítulo IX se empleó las pruebas de  $t$  (observaciones no pareadas y pareadas) para comparar el efecto de dos tratamientos; procedimiento que conduce a la aceptación o rechazo de la hipótesis nula planteada.

En el caso de observaciones no pareadas se asignan los dos tratamientos a las unidades experimentales aleatoriamente, es decir al azar. Esto es equivalente a un Diseño Completamente al Azar (DCA), como se verá más adelante.

Cuando puede agruparse a las unidades experimentales en pares homogéneos, se tiene el caso de observaciones pareadas. La aleatorización se restringe para que las unidades experimentales de cada par reciban un tratamiento distinto. Esto es equivalente a un Diseño de Bloques Completos al Azar (BCA), en el que el bloque consta de dos unidades experimentales.

Para las pruebas de hipótesis con los procedimientos de observaciones no pareadas y observaciones pareadas se puede emplear indistintamente la Prueba Estadística  $t$  o la Prueba  $F$  (recordar que  $t^2=F$ ). En los diseños con tres o más tratamientos la prueba que se emplea es la Prueba Estadística  $F$ .

En la investigación agropecuaria y también en otros campos, los diseños DCA y BCA son los más utilizados. Sin embargo, existen varias decenas de diseños para experimentos, algunos relativamente sencillos como los señalados y otros bastante complejos.

El investigador, de acuerdo con los objetivos de su investigación, las características de las unidades experimentales y otros recursos a su disposición debe decidir qué diseño emplear en cada caso. Sin embargo, es necesario tener en mente que la calidad de la investigación no está relacionada con la complejidad del diseño experimental. El principio básico es emplear el diseño experimental más simple que sea posible y que mejor se ajuste a las circunstancias.



## **Diseño completamente al azar**

---

*“El modelo que se emplee en el diseño de un experimento debe escogerse cuidadosamente. De hecho, el modelo dicta el marco conceptual estadístico para la generación de los datos”. Morris Hamburg<sup>26</sup>*

Cuando en un experimento los tratamientos son el único criterio de clasificación, el análisis que se realiza se denomina *análisis de varianza en una dirección*. Los resultados de un experimento conducido con un Diseño Completamente al Azar (DCA) dan origen a un análisis de varianza en una dirección.

En un DCA, la asignación de los tratamientos a las unidades experimentales (UE) es completamente al azar, lo que indica que no existe ninguna restricción en la aleatorización. Cualquier unidad experimental puede recibir cualquier tratamiento.

El DCA es el mejor diseño cuando las unidades experimentales son homogéneas. Es decir, cuando la variación entre las unidades experimentales es baja. Homogeneidad no significa igualdad. Homogeneidad entre parcelas en el campo es difícil de encontrar. Lo más usual es detectar la presencia de alguna gradiente. Mientras más grande es el terreno mayor será la variación entre parcelas. Sin embargo, en cierto tipo de experimentos es posible contar con unidades experimentales (UE) homogéneas. En experimentos en el laboratorio, en el invernadero, en cámaras de crecimiento, donde las UE son tubos de ensayo, platos Petri, macetas, etcétera, es relativamente fácil controlar la variación. En estos casos, el diseño experimental recomendado es el Diseño Completamente al Azar (DCA).

### **14.1 . Ventajas del Diseño Completamente al Azar**

1. Es flexible en cuanto al número de tratamientos y de repeticiones.
2. El número de repeticiones de cada tratamiento puede ser diferente, bien porque se planificó así por no contar con suficientes UE o porque se per-

---

26 Hamburg, M. 1974. *Basic Statistics: A Modern Approach*. Harcourt Brace Jovanovich, Inc. New York. Pg 17.

dieron en la ejecución del experimento una o varias UE. Esto complica en algo el análisis, pero se lo puede realizar sin necesidad de emplear fórmulas especiales para estimar valores perdidos.

3. El análisis estadístico es fácil, en especial si el número de repeticiones de cada tratamiento es igual.
4. Es el diseño con el que se obtiene el mayor número de grados de libertad para el error, en comparación con otros diseños de igual tamaño. Esta es una ventaja importante, en especial en experimentos con pocos tratamientos y pocas repeticiones, porque la precisión de un experimento aumenta con el número de grados de libertad.

## 14.2. Desventajas del Diseño Completamente al Azar

En sí mismo, el DCA no tiene ninguna desventaja. Sin embargo, si se conduce un experimento con un DCA, cuando por la variación entre UE debió escogerse otro diseño, el error experimental resulta muy inflado. Esto causa una disminución en la precisión del ensayo, que por lo general es mayor que lo que se gana al tener unos cuantos grados de libertad más. Un error experimental muy inflado puede impedir que se detecten diferencias reales entre los tratamientos, lo que resulta ser una grave desventaja por el mal uso del Diseño Completamente al Azar.

## 14.3. Componentes de la variación

El concepto de componentes de la variación se introdujo en el caso del Análisis de Varianza de Regresión (Sección 11.1). El concepto y el procedimiento es similar para los diseños experimentales. La variación se analiza a través un Cuadro de Análisis de Varianza, que se abrevia como Cuadro ADEVA. Todo experimento debidamente planificado tiene diferentes fuentes de variación. En el caso más simple, que es el de un DCA, los componentes de variación son:

*Variación total = Variación entre los tratamientos + Variación dentro de los tratamientos.*

El último componente, la variación dentro de los tratamientos, es en verdad el error experimental. Note el carácter aditivo de los componentes de variación.

Las fuentes de variación se expresan en términos de Sumas de Cuadrados. Por lo tanto, la igualdad anterior puede escribirse así:

$$SCT = SCTrat + SCE$$

expresión en la que  $SCT$  es la Suma de Cuadrados Total,  $SCTrat$  es la Suma de Cuadrados de los Tratamientos y  $SCE$  es la Suma de Cuadrados del Error.

En diseños diferentes al DCA hay adicionalmente otras fuentes de variación; sin embargo, el carácter aditivo de las mismas se mantiene.

#### 14.4. Procedimiento para el análisis de datos

En un experimento conducido con un DCA en el que hubo  $k$  tratamientos y  $r$  repeticiones los datos pueden presentarse como en el Cuadro 14.1. Los símbolos del Cuadro 14.1 son:

$Y_{ij}$  = cualquier observación de  $Y_{11}$  a  $Y_{kr}$

$k$  = número de tratamientos

$r_j$  = número de repeticiones en cada tratamiento

$n$  = número total de observaciones. El valor de  $n$  es igual a  $\sum r_j$ , o a  $r \cdot k$ , si  $r$  es igual para todos los tratamientos.

$T_i$  = Sumatoria de las observaciones  $Y_{ij}$  en el tratamiento "i"

$\bar{Y}_i$  = Media del tratamiento "i";  $\bar{Y}_i = \frac{T_i}{r_j}$

$\sum Y_{ij}$  = La sumatoria de todas las observaciones en el experimento.

$\bar{\bar{Y}}$  La gran media experimental. Es la media de todas las observaciones en el experimento,

$$\bar{\bar{Y}} = \frac{\sum Y_{ij}}{n} \quad (83)$$

Cuadro 14.4.1. Presentación de los datos de un DCA con  $k$  tratamientos y  $r$  repeticiones

Tratamientos	1	2	3		$k$	Total
	$Y_{11}$	$Y_{21}$	$Y_{31}$	-----	$Y_{k1}$	
	$Y_{12}$	$Y_{22}$	$Y_{32}$	-----	$Y_{k2}$	
	$Y_{13}$	$Y_{23}$	$Y_{33}$	-----	$Y_{k3}$	
	.	.	.		.	
	.	.	.		.	
	.	.	.		.	
	.	.	.		.	
	$Y_{1j}$	$Y_{2j}$	$Y_{3j}$		$Y_{kj}$	
$r_j$	$r_1$	$r_2$	$r_3$		$r_k$	$n$
$T_i$	$T_1$	$T_2$	$T_3$		$T_k$	$\sum T_i = \sum Y_{ij}$
$\bar{Y}_i$	$\bar{Y}_1$	$\bar{Y}_2$	$\bar{Y}_3$		$\bar{Y}_k$	$\bar{\bar{Y}}$



Note que cada observación está asociada a un tratamiento y a una repetición.

En el Análisis de Varianza, dependiendo del modelo que se planifique en la investigación, *fijo o al azar* como se verá más adelante, se tiene varios parámetros de referencia. Estos son:

- La media de la población de todos los tratamientos,  $\mu$
- Las medias de las poblaciones de cada uno de los tratamientos,  $\mu_i$
- Los efectos de los tratamientos ( $\tau_i$ ) medidos como desviaciones de la media; de esto resulta que  $\mu_i = \mu + \tau_i$
- La varianza del error experimental,  $\sigma^2$
- La varianza de los efectos de los tratamientos,  $\sigma_\tau^2$  (parámetro de importancia en los modelos al azar)

Como en los experimentos con diseños experimentales se toman muestras de poblaciones de tratamientos, se requiere de los estimadores respectivos de cada uno de los parámetros señalados. Estos son:

- El estimador de la media poblacional,  $\mu$ , es la gran media del experimento  $\bar{Y}$
- Los estimadores de las medias de los tratamientos,  $\mu_i$ , son los valores  $\bar{Y}_i$
- Los estimadores de los efectos de los tratamientos  $\tau_i$  son los valores  $\bar{Y}_i - \bar{Y}$
- El estimador de la varianza del error  $\sigma^2$  es el Cuadrado Medio del Error,  $CME$
- El estimador de la varianza entre los efectos de los tratamientos  $\sigma_\tau^2$  es el valor  $S_\tau^2$

Lo indicado con relación a los parámetros y sus estimadores obviamente no solo se refieren al DCA sino también a todos los otros diseños que se verán más adelante (aunque tendrán, además de los señalados, otros parámetros y estimadores).

Una vez organizados los datos como en el Cuadro 14.4.1, se pueden calcular los componentes de la variación en un experimento conducido con un Diseño Completamente al Azar (DCA):

1. La Suma de Cuadrados Total,  $SCT$ . Esto es una medida de la variación total en el experimento. Las fórmulas son:

- a. Fórmula de Origen o de Definición

$SCT = \sum (Y_{ij} - \bar{Y})^2$ , a partir de esta fórmula, por álgebra y aplicando las propiedades de la sumatoria se deriva la fórmula computacional:

- b. Fórmula Computacional<sup>27</sup>

$$SCT = \sum Y_{ij}^2 - FC \quad (84)$$

Fórmula en la que el Factor de Corrección  $FC$  es:

$$FC = \frac{(\sum Y_{ij})^2}{n} \quad (85)$$

2. La Suma de Cuadrados de los Tratamientos ( $SCTrat$ ). Esto es una medida de la variación *entre* los tratamientos. Las fórmulas son:

- c. Fórmula de Origen o Definición

$$SCTrat = r \sum (\bar{Y}_i - \bar{Y})^2$$

- d. Fórmula Computacional

$$SCTrat = \sum \frac{T_i^2}{r_j} - FC \quad (86)$$

3. La Suma de Cuadrados del Error ( $SCE$ ). Esto es una medida de la variación *dentro* de los tratamientos. Las fórmulas son:

- e. Fórmula de Origen o Definición

$$SCE = \sum (Y_i - \bar{Y}_i)^2$$

- f. Fórmula general de la suma de cuadrados del error

$$SCE = SCT - SCTrat$$

<sup>27</sup> Las fórmulas computacionales son más precisas porque evitan los errores por redondeo

Una vez obtenidas las fuentes de variación se colocan en un Cuadro ADEVA, como el que se presenta en el siguiente cuadro:

Cuadro 14.4.2. Ejemplo de un Cuadro ADEVA para un DCA

Fuente de Variación (FV)	Suma de Cuadrados (SC)	Grados de Libertad (gl)	Cuadrado Medio (CM)	Prueba F (F)
Tratamientos	$SCTrat$	$k-1$	$CMTrat$	$CMTrat/CME$
Error	$SCE$	$n-k$	$CME$	
Total	$SCT$	$n-1$		

El Cuadrado Medio o varianza se obtiene al dividir la Suma de Cuadrados para los grados de libertad en cada hilera. Por lo tanto:

- $CMTrat = SCTrat / (k - 1)$
- $CME = SCE / (n - k)$

Es importante notar las características aditivas del modelo:

- $SCTrat + SCE = ACT$
- $(k + 1) + (n - k) = n - 1$

**Ejercicio 14.4.1.** Mediante las fórmulas de origen y las computacionales encuentre los componentes de variación para un diseño completamente al azar (DCA), con cuatro tratamientos (A, B, C, D) y tres repeticiones. Los datos son los siguientes:

Cuadro 1. Resultados de un experimento ficticio con un DCA

Totales y medias	A	B	C	D	Gran media
	6	2	9	10	
	8	7	5	11	
	4	3	7	12	
$T_i$	18	12	21	33	
$\bar{Y}_i$	6	4	7	11	$\bar{Y} = 7$

Cuadro 2a. Sumatoria desviaciones: <i>SCT</i>			
A	B	C	D
-1	-5	+2	+3
+1	0	-2	+4
-3	-4	0	+5
$\sum(Y_{ij}-\bar{Y}) = 0$			

Cuadro 3a. Sumatoria desviaciones: <i>SCTrat</i>			
A	B	C	D
-1	-3	0	+4
-1	-3	0	+4
-1	-3	0	+4
$r \sum(\bar{Y}_i-\bar{Y}) = 0$			

Cuadro 4a. Sumatoria de desviaciones <i>SCE</i>			
A	B	C	D
0	-2	+2	-1
-2	+3	-2	0
+2	-1	0	+1
$\sum(Y_i-\bar{Y}_i) = 0$			

Cuadro 2b. <i>SCT</i>			
A	B	C	D
1	25	4	9
1	0	4	16
9	16	0	25
$\sum(Y_{ij}-\bar{Y})^2 = 110$			

Cuadro 3b. <i>SCTrat</i>			
A	B	C	D
1	9	0	16
1	9	0	16
1	9	0	16
$r \sum(\bar{Y}_i-\bar{Y})^2 = 78$			

Cuadro 4b. <i>SCE</i>			
A	B	C	D
0	4	4	1
4	9	4	0
4	1	0	1
$\sum(Y_i-\bar{Y}_i)^2 = 32$			

En los cuadros anteriores, a partir del Cuadro 1 de resultados de un DCA, se observa el procedimiento para encontrar los componentes de variación mediante las fórmulas de origen. En el Cuadro 2a pueden verse las desviaciones de las observaciones con relación a la gran media. (Según el principio de cuadrados mínimos, toda desviación de un valor con relación a una media es igual a cero.) En el Cuadro 2b, la sumatoria de los cuadrados de las desviaciones del cuadro 2a dan como resultado la  $SCT=110$ . En forma similar, las desviaciones de las medias de los tratamientos con relación a la gran media que se presentan en el Cuadro 3a, al elevarse al cuadrado y sumarse dan como resultado la  $SCTrat=78$  (cuadro 3b). El origen de la *SCE*, a partir de las desviaciones de las observaciones de cada tratamiento con relación a su propia media, se observa en los Cuadros 4a y 4b: la *SCE* es igual a 32.

Con las fórmulas denominadas computacionales los resultados son:

$$FC = \frac{(\sum Y_{ij})^2}{n} = \frac{(84)^2}{12} = 588$$

$$SCT = \sum Y_{ij}^2 - FC = 698 - 588 = 110$$

$$SCTrat = \sum \frac{T_j^2}{r_j} - FC = \frac{18^2}{3} + \frac{12^2}{3} + \frac{21^2}{3} + \frac{33^2}{3} - 588 = 78$$

$$SCE = SCT - SCTrat = 110 - 78 = 32$$

Como se esperaba, se obtienen resultados iguales con los dos tipos de fórmulas.

Nota: En el cálculo de la SCTrat, como el número de repeticiones fue igual para todos los tratamientos, pudo usarse el valor de tres como denominador común y simplificar la fórmula.

Para comprender en mejor forma las relaciones entre los componentes de la variación se plantea el siguiente ejercicio:

**Ejercicio 14.4.2.** Suponga un DCA con cuatro tratamientos (A, B, C, D) y dos repeticiones. Construya tres cuadros que satisfagan lo siguiente:

1. SCT igual a 0. Gran media experimental igual a ocho.
2. SCT mayor que cero. Medias de los tratamientos iguales a cinco. SCTrat igual a cero.
3. SCT mayor que cero. SCE igual a cero.

Cuadro 1			
A	B	C	D
8	8	8	8
8	8	8	8

Cuadro 2			
A	B	C	D
5	8	1	4
5	2	9	6

Cuadro 3			
A	B	C	D
3	6	4	1
3	6	4	7

### Comentarios:

- En el Cuadro 1 no existe variación, todos los datos son iguales. Consecuentemente la SCT, así como la SCTrat y la SCE son iguales a cero. La gran media experimental es  $\bar{Y} = 8$ .
- En el Cuadro 2, las medias de los cuatro tratamientos son iguales a cinco. No hay variación entre los tratamientos, consecuentemente la SCTrat es igual a cero. La SCT es mayor que cero e igual a la SCE. (Nota: cualquier combinación de datos que en cada tratamiento sume diez también satisface la condición impuesta.)

- En el Cuadro 3 no hay variación entre las observaciones dentro de un mismo tratamiento. Consecuentemente, la SCE es igual a cero. La SCT es mayor que cero e igual a la SCTrat. Un número infinito de valores satisface también la condición impuesta.

## 14.5. La aleatorización en el DCA

El sorteo o aleatorización en el diseño de experimentos se refiere a la asignación al azar de los tratamientos a las unidades experimentales. En el caso del Diseño Completamente al Azar, cualquier unidad experimental puede recibir cualquier tratamiento.

Para el sorteo se han empleado tradicionalmente varios procedimientos, como papeles numerados, fichas y tablas de números al azar. Las tablas de números al azar pueden encontrarse en varios textos de Estadística. Sin embargo, con los programas computarizados se puede obtener con mayor facilidad el conjunto de números al azar que se desee. Esta opción la presentan casi todos los programas, incluyendo *Excel* y *Statistix*. Pero además, en internet se encuentran varios sitios que permiten la generación de números al azar. En el siguiente ejercicio se presenta un ejemplo con el empleo del programa *Excel*.

**Ejercicio 14.5.1.** Se planifica realizar un experimento con un diseño completamente al azar (DCA) con cinco tratamientos (A, B, C, D, E) y cuatro repeticiones. Genere números al azar con el programa *Excel* para asignar en forma aleatorizada los tratamientos a las unidades experimentales.

Lo primero que debe hacerse es numerar las unidades experimentales de 1 al 20. Luego, mediante el empleo de la opción ALEATORIO.ENTRE(...), se procede a generar números enteros al azar entre 1 y 20. Los números repetidos que se produzcan deben descartarse. Generalmente puede decidirse que los cuatro números al azar más pequeños se asignen al tratamiento A; los siguientes cuatro números, al tratamiento B y así sucesivamente. En el siguiente cuadro se resumen los resultados:

Cuadro 1. Asignación al azar de los tratamientos a las unidades experimentales

No. U.E.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
No. Azar	10	5	18	9	19	8	3	20	16	12	1	11	17	15	13	4	7	14	2	6
Trat. Asign.	C	B	E	C	E	B	A	E	D	C	A	C	E	D	D	A	B	D	A	B

En el cuadro anterior, como ejemplo, las unidades experimentales 7, 11, 16 y 19 correspondieron a los números al azar 3, 1, 4 y 2. Estas unidades experimentales recibirán el tratamiento A.

La ubicación física de las unidades experimentales va a depender de las condiciones y características del experimento. En el caso de experimentos de campo, un mapa con las unidades experimentales o parcelas, numeradas en orden y el tratamiento correspondiente, se presenta en el siguiente cuadro:

**Cuadro 1. Mapa de campo para un Diseño Completamente al Azar, con cinco tratamientos (A, B, C, D, E) asignados al azar a cuatro repeticiones**

1 C	2 B	3 E	4 C	5 E
10 C	9 D	8 E	7 A	6 B
11 A	12 C	13 E	14 D	15 D
20 B	19 A	18 D	17 B	16 A

Nota: La marcación de las parcelas en zigzag es optativa, pero tiene como propósito facilitar la toma de datos en el terreno y el registro en el libro de campo.

## 14.6. El Modelo Lineal Aditivo

El *modelo lineal aditivo* identifica las fuentes de variación y es la base para las suposiciones necesarias del análisis de varianza para todos los diseños experimentales. Para comprender el origen del análisis de varianza (ADEVA) en un diseño DCA es necesario referirse a este modelo:

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad (87)$$

en donde:  $i = 1, \dots, k$  y  $j = 1, \dots, r$

( $k$  = número de tratamientos;  $r$  = número de repeticiones)

En el modelo se tiene que una observación cualquiera  $Y_{ij}$  es el resultado aditivo de varios componentes:

$\mu$ , es la media general de la población de referencia en el experimento

$\tau_i$ , es el efecto del tratamiento  $i$  medido como desviación de la media

$\varepsilon_{ij}$ , es el componente del error que se supone proviene del azar y está asociado con cada observación

Lo que interesa es estimar el valor de los parámetros  $m$  y  $\tau_i$  en el modelo y evaluar las hipótesis que acerca de ellos se plantean. Como el modelo descrito es lineal podríamos, en forma similar a lo que hicimos en regresión, usar el procedimiento de *cuadrados mínimos* para obtener estimadores válidos de los parámetros. Las suposiciones necesarias para esto son:

1. Las poblaciones tienen igual varianza, es decir  $\sigma_1^2 = \sigma_2^2 \dots = \sigma_k^2 = \sigma^2$ .
2. Existe independencia entre las observaciones de las diferentes poblaciones, en otras palabras, no están correlacionadas.
3. Las poblaciones están normalmente distribuidas.
4. Los  $\varepsilon_{ij}$  son independientes y tienen media igual a cero y varianza  $\sigma^2$ . Esto es, no hay un componente sistemático del error que se acarre de observación a observación.

El modelo lineal aditivo puede ser fijo (modelo I) o al azar (modelo II).

### ***Modelo Fijo o Modelo I***

En el modelo fijo los efectos de los tratamientos están establecidos (son fijos) y cumplen la siguiente condición:  $\sum \tau_i = 0$ . Estos efectos forman una población finita y son los parámetros que nos interesan. Otro parámetro de interés es la varianza del error,  $\sigma_\varepsilon^2 = CME = \sigma^2$

### ***Modelo al Azar o Modelo II***

En el modelo al azar los efectos de los tratamientos  $\tau_i$  representan una muestra al azar de una población de efectos de tratamientos que tienen media igual a cero y una varianza  $\sigma_\tau^2$  (varianza de los efectos de los tratamientos). Estos son los parámetros que nos interesan, además de la varianza del error experimental,  $\sigma^2$ .

### ***Diferencias entre el modelo fijo y el modelo al azar***

Tanto en el modelo fijo como en el modelo al azar el planteamiento del experimento y el análisis de los datos son iguales. La diferencia entre los dos mo-



delos se refiere a las poblaciones de referencia en cada una, a las hipótesis que se prueban y a las conclusiones a las que se puede arribar.

En el **modelo fijo** la población de referencia es la formada por los tratamientos seleccionados. Si en un experimento nos interesa evaluar el comportamiento de cuatro variedades de maíz HB-104, HPB, H-29, y D-833, estas cuatro variedades forman la población de referencia. Las conclusiones a las que se lleguen tendrán validez únicamente para esas variedades. Si el experimento fuera a repetirse el próximo año, en él se incluirían las cuatro variedades indicadas. En otras palabras, las inferencias que se realizan en el modelo fijo se refieren al grupo específico de tratamientos del experimento.

En el modelo fijo, la hipótesis nula que se prueba es que los efectos de los tratamientos son iguales a cero (y por consiguiente iguales entre sí):

Las hipótesis son:

$$H_0: \tau_i = 0$$

$$H_1: \tau_i \neq 0$$

Como puede verse, la hipótesis nula plantea que todos los efectos de los tratamientos son iguales versus la hipótesis alternativa que por lo menos dos efectos no son iguales.

La restricción impuesta  $\sum \tau_i = 0$  determina que los efectos de los tratamientos puedan medirse como desviaciones de la media total. Consecuentemente las hipótesis pueden plantearse también de la siguiente forma:

$$H_0: \mu_{Di} = 0$$

$$H_1: \mu_{Di} \neq 0$$

Consecuentemente, la hipótesis nula plantea que todas las diferencias entre las medias de los tratamientos son iguales a cero, mientras que la hipótesis alternativa indica que por lo menos una diferencia entre las medias de los tratamientos no es igual a cero.

Puede afirmarse que la mayoría de los experimentos se realizan con el modelo fijo. Esto por cuanto el investigador, por lo general, está interesado específicamente en un grupo de tratamientos.

En el **modelo al azar**, los tratamientos son una muestra aleatoria (al azar) de una población de referencia. Si se realiza un experimento para evaluar la calidad de los insecticidas que se expenden en un país, se tomará una muestra al azar de todos los insecticidas que se encuentran en el mercado. Los resultados y conclusiones a las que se lleguen se referirán a toda la población de insecticidas y no únicamente a los que se incluyeron en la muestra. Si el experimento fuera a repetirse el próximo año, se tomaría otra muestra al azar de la población de insecticidas. Esta muestra seguramente incluirá diferentes insecticidas que los de la muestra del año anterior. Por lo tanto, las inferencias que se realizan en el modelo al azar se refieren a la población de tratamientos.

Las empresas productoras y expendedoras de semillas utilizan el modelo al azar porque constantemente están evaluando nuevos materiales y les interesa conocer el comportamiento a través de los ciclos agrícolas de sus nuevos materiales, en comparación con otros del mercado.

En el modelo al azar la hipótesis nula que se prueba es que la varianza de los efectos de los tratamientos es igual a cero. Las hipótesis son:

$$H_0: \sigma_\tau^2 = 0$$

$$H_1: \sigma_\tau^2 \neq 0$$

El interés principal es estimar el valor de los parámetros (media y efecto de los tratamientos) en el modelo y evaluar las hipótesis que acerca de ellos se planteen.

### Valores esperados de las Varianzas o Cuadrados medios

Así como la media de la muestra es un estimador de la media de la población, los cuadrados medios son estimadores de los cuadrados medios esperados. Si los experimentos planeados tanto con el modelo fijo o con el modelo al azar fueran repetidos indefinidamente y se calculara el promedio de los cuadrados medios, se obtendrían los valores esperados. Esto se presenta en el siguiente cuadro:

Cuadro 14.6.1. Cuadrados medios esperados para el modelo fijo y el modelo al azar

Fuente de Variación	Grados de libertad	Cuadrado Medio Modelo Fijo	Cuadrado Medio Modelo al Azar
Tratamientos	$k - 1$	$\sigma^2 + r \sum \tau_i^2 / (k - 1)$	$\sigma^2 + r \sigma_\tau^2$
Error	$n - k$	$\sigma^2$	$\sigma^2$
Total	$n - 1$		

En el Capítulo 9 se estudió acerca de la distribución F y las pruebas estadísticas correspondientes. De igual manera, la prueba estadística esperada F resulta de dividir el Cuadrado Medio esperado de los tratamientos, para el Cuadrado Medio esperado del Error. Esto es, para el modelo fijo:

$$F = \frac{\sigma^2 + r \sum \tau_i^2 / (k-1)}{\sigma^2} \quad (88)$$

Si el valor  $r \sum \tau_i^2 / (k-1)$  fuera igual a cero, F sería igual a uno. Un valor de F igual a uno indica que no hubo efecto de los tratamientos. Teóricamente no puede darse un valor de F menor de uno, sin embargo, en algunos experimentos con poco control sobre el error experimental, se presenta este caso.

Los cuadrados medios esperados del cuadro anterior corresponden al caso más común en el DCA, que es cuando el número de repeticiones es igual en todos los tratamientos. Cuando los tratamientos tienen diferentes repeticiones, hay varios valores de  $r$  y consecuentemente difieren las fórmulas de los cuadrados medios esperados. Para este caso y para otros diseños con fuentes adicionales de variación, puede consultarse el texto de Steel et al; (1997). De igual manera se puede proceder para el Modelo Mixto, que combina efectos fijos y al azar en el mismo diseño.

## 14.7. La evaluación de hipótesis

Para una mejor comprensión del proceso de evaluación de hipótesis, analicemos los resultados de un experimento:

---

**Ejercicio 14. 7.1.** Un investigador desea conocer si seis tratamientos tienen igual efecto en el control de la roya del café. Para esto selecciona 24 arbustos iguales y asigna al azar cuatro para cada uno de los seis tratamientos. Se decide que el efecto se medirá basándose en la producción de cada arbusto.

Las hipótesis que se plantea, con las connotaciones señaladas en la sección anterior, son:

$$H_0: \mu_{Di} = 0$$

$$H_1: \mu_{Di} \neq 0$$

El nivel de significación alfa generalmente se fija en 5% ( $\alpha = 5\%$ ). Esto es la probabilidad de rechazar una hipótesis nula verdadera o Error Tipo I.

**Cuadro 1. Rendimientos de cerezas de café, en hectogramos por arbusto, en respuesta a seis tratamientos para el manejo de la roya.**

A	B	C	D	E	F
16.8	18.5	24.3	15.8	14.4	8.4
20.6	21.2	27.1	11.8	12.0	5.2
18.3	20.0	19.4	13.9	10.9	9.3
22.2	18.1	23.8	12.4	13.2	4.5

**Cuadro 2. Totales de los tratamientos, número de repeticiones y medias**

Tratamientos	A	B	C	D	E	F
Totales de los Tratamientos, $T_i$	77.90	77.80	94.60	53.90	50.50	27.40
Repeticiones, $r_j$	4	4	4	4	4	4
Medias, $\bar{Y}_i$	19.47	19.45	23.65	13.47	12.62	6.85

La gran media experimental es:  $\bar{\bar{Y}} = 15.92$ . En el caso de las repeticiones, como todos los tratamientos tienen igual número de repeticiones, se tiene que:  $r_j = r = 4$ .

Para encontrar los componentes de las fuentes de variación en primer lugar, mediante la fórmula computacional, se calcula el Factor de Corrección:

$$FC = \frac{(\sum Y_i)^2}{\sum r_j = n} = \frac{(16.8 + 18.5 + \dots + 4.5)^2}{24} = 6083.35$$

Luego se procede a obtener la Suma de Cuadrados Total:

$$SCT = \sum Y_i^2 - FC = (16.8^2 + 18.5^2 + \dots + 4.5^2) - 6083.35 = 822.58$$

A continuación, se encuentra la Suma de Cuadrados de los Tratamientos:

$$SCT_{\text{trat}} = \frac{\sum T_i^2}{r_j} - FC = \frac{(77.9^2 + 77.8^2 + \dots + 27.4^2)}{4} - 6083.35 = 735.81$$

Finalmente, por diferencia, se encuentra la Suma de Cuadrados del Error:

$$SCE = SCT - SCT_{\text{trat}} = 822.58 - 735.81 = 86.77$$

Se dispone ya de todos los componentes de Sumas de Cuadrados para preparar el cuadro ADEVA:

Cuadro 3. Análisis de Varianza (ADEVA)

E.V.	S.C.	G.L.	C.M.	F	$F_{0.05}$	$F_{0.01}$
Tratamientos	735.81	5	147.16	30.53**	2.77	4.25
Error	86.77	18	4.82			
Total	822.58	23				

Cuando la prueba F calculada excede el valor crítico F de la tabla, esto es,  $F_{cal} \geq F_{tabla}$ , se declara que el valor de F es significativo. Es decir, se rechaza la hipótesis nula y se acepta la hipótesis alternativa al nivel de significación alfa ( $\alpha$ ). Esto indica que la probabilidad de haber encontrado un valor de F tan alto como 30.53, sólo por azar, es baja. Por lo tanto, puede decidirse que hay diferencias reales entre los tratamientos.

Los valores F de la tabla se los encuentra en el Apéndice D1 al 5%, Apéndice D2 al 2.5% y Apéndice D3 al 1%, con los grados de libertad correspondientes a tratamientos (numerador) y al error (denominador). Esto es 5 y 16 grados de libertad en este ejercicio. Los valores encontrados en la tabla, que constan en las dos últimas columnas del Cuadro ADEVA son:

$$F_{(3,16)(0.05)} = 2.77$$

$$F_{(3,16)(0.01)} = 4.25$$

En forma convencional se pone junto al valor de F calculado un asterisco para indicar significación al 5% y dos asteriscos para significación al 1%. En este caso (Cuadro 3) el valor F lleva dos asteriscos, lo que indica que el valor calculado de F excede el valor  $F_{(0.05)}$  de la tabla y el valor  $F_{(0.01)}$ .

La conclusión es, entonces, rechazar la hipótesis nula de que las diferencias entre medias son iguales a cero. Esto conlleva a aceptar la hipótesis alterna de que por lo menos una diferencia entre medias no es igual a cero, al nivel de significación alfa previamente fijado.

Con el programa Statistix se obtienen los siguientes resultados:

Statistix 9.0 ROYA CAFÉ 2,

**Completely Randomized AOV for REND**

Source	DF	SS	MS	F	P
FUNG	5	735.807	147.161	30.53	0.0000
Error	18	86.773	4.821		
Total	23	822.580			

Grand Mean 15.921 CV 13.79

<b>Homogeneity of Variances</b>	<b>F</b>	<b>P</b>
Levene's Test	1.22	0.3420
O'Brien's Test	0.78	0.5784
Brown and Forsythe Test	0.59	0.7080

<b>FUNG</b>	<b>Mean</b>
A	19.475
B	19.450
C	23.650
D	13.475
E	12.625
F	6.850

Observations per Mean 4  
 Standard Error of a Mean 1.0978  
 Std Error (Diff of 2 Means) 1.5525

### Comentarios:

- Note en el ejercicio anterior que el programa Statistix presenta, en la última columna del Cuadro ADEVA, la probabilidad de un valor F tan extremo como el calculado, esto es
- $P = 0.0000$ . Se puede entonces señalar que la probabilidad de un resultado así, sólo por azar, es  $P < 0.00009$  (menor que 1/10000). Este resultado es obviamente mucho más preciso que el encontrado mecánicamente, porque no está limitado sólo a las tablas del 1 y 5%.
- Observe también que Statistix incluye en el análisis tres pruebas diferentes de homogeneidad de las varianzas: Levene, O'Brien y Brown y Forsythe. La homogeneidad de las varianzas de los tratamientos es un requisito para la validez del análisis estadístico. Si las varianzas no son homogéneas se puede arribar a conclusiones erradas. En este ejemplo, a juzgar por las altas probabilidades de las tres pruebas ( $P > 0.05$ ), se concluye que hay evidencia suficiente para señalar que las varianzas sí son homogéneas.
- En otros casos en que haya discrepancias entre las pruebas de homogeneidad de varianzas, el investigador puede guiarse por el promedio de las tres pruebas. Si la probabilidad se mantiene en valores menores que el 5%, entonces lo aconsejado es emplear alguno de los procedimientos de transformación de datos. La transformación de datos no siempre corrige la falta

de homogeneidad. Si este fuera el caso, podría realizarse el análisis con métodos de la estadística no paramétrica. Estos métodos no tienen como requisito la homogeneidad de las varianzas. Tanto lo relacionado con la transformación de datos, así como el empleo de pruebas no paramétricas, se verá en capítulos posteriores.

## Separación de medias

En todo caso, el análisis de varianza corresponde a la primera parte del manejo de datos. El paso siguiente es determinar qué medias o grupos de medias son estadísticamente diferentes. Esto último, cuando los tratamientos son cualitativos. Cuando los tratamientos son cuantitativos, el procedimiento es establecer tendencias mediante regresión. Estos métodos, que complementan el análisis de varianza, se verán en el Capítulo 17.

## El Coeficiente de Variación

El Coeficiente de Variación (C.V.) es una estadística importante que debe estar presente en todo informe de investigación. Por lo general se lo incluye junto al Cuadro ADEVA. El C.V. se calcula con la siguiente fórmula:

$$C.V. = \frac{\sqrt{S_e^2}}{\bar{Y}} * 100 = \frac{\sqrt{CME}}{\bar{Y}} * 100$$

Por observación de la fórmula se puede indicar que el Coeficiente de Variación es una medida de la variación *dentro* de los tratamientos y alrededor de la gran media experimental. Se lo expresa usualmente en porcentaje.

Mientras mayor sea el CME mayor será el C.V. El valor del C.V. en algunos experimentos con poco control de la variación dentro de los tratamientos puede exceder el 100%.

Los resultados experimentales son más confiables cuando el C.V. es bajo. Sin embargo, el C.V. por sí solo no nos indica mayor cosa, si es que no se lo compara con otros C.V. encontrados en experimentos similares.

Supongamos que en un experimento encontramos un C.V. = 22%. ¿Es esto bueno o malo? No se sabe, pero si al examinar los resultados de otros experimentos similares vemos que el C.V. oscila entre 5% y 10% entonces se puede pensar que nuestro C.V. está inflado. Un C.V. alto puede deberse a que no se seleccionó

cuidadosamente las unidades experimentales, o que el diseño escogido no era el adecuado, o que hubo fallas en la conducción del experimento, o una combinación de estas falencias. El número de repeticiones también afecta el C.V. Si las otras limitaciones permanecen constantes, entonces a mayor número de repeticiones, menor C.V. En todo caso, con C.V. elevados, el grado de confiabilidad en los resultados con relación a las hipótesis planteadas se reduce.

**Ejercicio 14.7.2.** Encuentre el coeficiente de variación (C.V.) del experimento del ejercicio 14.7.1.

Para encontrar el C.V. se requiere conocer la gran media experimental y el Cuadrado Medio del Error CME:

$$\bar{Y} = \frac{\sum Y_{ij}}{n} = \frac{382.10}{24} = 15.92$$

$$CME = \frac{SCE}{n-k} = \frac{86.77}{18} = 4.82$$

Consecuentemente:

$$C.V. = \frac{\sqrt{CME}}{\bar{Y}} * 100 = \frac{\sqrt{4.82}}{15.92} * 100 = 13.79\%$$

El siguiente ejercicio tiene relación con el Modelo Lineal Aditivo:

**Ejercicio 14.7.3.** Encuentre la estimación de la variabilidad de los efectos fijos del experimento presentado en el ejercicio 14.7.1.

En el ejercicio 14.7.1, se encontró lo siguiente:  $CMTrat = 147.16$  y  $CME = \sigma^2 = 4.82$

Según el modelo fijo, para encontrar el valor estimado de los efectos,  $\frac{\sum \tau_i^2}{(k-1)}$ , se parte del Cuadrado Medio de los Tratamientos,  $CMTrat$ . Esto es, según el Cuadro 14.6.1:

$$CMTrat = \sigma^2 + r \frac{\sum \tau_i^2}{(k-1)}, \text{ consecuentemente:}$$

$$r \sum \frac{\tau_i^2}{(k-1)} = CMTrat - \sigma^2 = 147.16 - 4.82 = 142.34$$

$$\sum \frac{\tau_i^2}{(k-1)} = \frac{142.34}{4} = 35.58$$



Se concluye que el valor estimado de los efectos de los tratamientos es 35.58. Como se trata del Modelo Fijo, si se repitiera el experimento un sinnúmero de veces, la cantidad  $\frac{\sum \tau_i^2}{(k-1)}$  sólo variará, teóricamente, por fluctuaciones en el error experimental, es decir, por las estimaciones de  $\sigma^2$ .

## 14.8. El DCA con diferente número de repeticiones

El único diseño que puede planificarse con diferente número de repeticiones es el DCA. La conducción del ensayo y el análisis de datos no presentan mayor dificultad. Sin embargo, en el empleo de métodos de separación de medias (que se verán en el Capítulo 17) debe tenerse en cuenta el diferente número de repeticiones de los tratamientos.

En experimentos con diseños completamente al azar es siempre preferible contar con igual número de repeticiones, pero esto no siempre es posible. Como ejemplo, en experimentos con animales es muchas veces difícil contar con suficientes unidades experimentales homogéneas. En otros casos, la cantidad de insumos disponibles para cada tratamiento puede determinar más repeticiones para unos tratamientos y menos repeticiones para otros. Además, siempre existe la posibilidad de que se dañen parcelas en el campo o que, por error, se omitan observaciones. En otros diseños, a diferencia del DCA, deben estimarse los datos perdidos para proceder al análisis.

El siguiente ejercicio se refiere a un DCA con diferente número de repeticiones. En este experimento los datos están expresados en porcentajes de germinación. Como se recordará, una presunción para la validez del modelo lineal aditivo es que la población de la que provienen las muestras siga la distribución normal. Sin embargo, algunas mediciones en porcentaje provienen de poblaciones con distribuciones binomiales, es decir que corresponden a dos categorías. En este ejemplo las categorías son: las semillas germinaron o no germinaron. Para resolver este problema se requiere de una transformación de datos. (Vale anotar que no todos los datos medidos en porcentaje corresponden a distribuciones binomiales. Como ejemplo, la humedad se mide en porcentaje y también los contenidos de proteína, grasa y minerales en diferentes productos. En estos casos, no es sí o no, o sí o no. Estas variables siguen, por lo general, la distribución normal.)

Lo concerniente a transformación de datos se tratará en un capítulo posterior, pero por el momento basta indicar que los datos  $Y_i$  en porcentaje, se convierten a una escala  $Y_i'$  de grados (del 0 al 90) mediante la fórmula

$Y'_i = \arcsen(\sqrt{Y_i/100}) * \frac{180}{\pi}$ . La escala en grados está más cerca de la distribución normal que la escala en porcentaje.

No siempre se requiere realizar la transformación indicada. Varios textos recomiendan efectuarla sólo cuando se observa excesiva variación entre los porcentajes. Sin embargo, lo más conveniente es realizar una prueba de normalidad de los datos, lo que se facilita con los programas computarizados. Con *Statistix* se puede realizar la Prueba W desarrollada por Shapiro y Wilk (1965). La hipótesis nula que se plantea es la normalidad de la población. Si la probabilidad es igual o menor que 5% ( $P \leq 0.05$ ) hay evidencia suficiente para rechazar la hipótesis nula y concluir que la distribución no es normal. En este caso se requiere transformar los datos. Veamos el siguiente ejercicio:

**Ejercicio 14.8.1.** En un experimento conducido con un DCA se evaluó el porcentaje de germinación de siete grupos experimentales de la leguminosa andina chocho *Lupinus mutabilis*. Un alto porcentaje de germinación puede indicar algún nivel de resistencia a patógenos del suelo. En cada parcela se sembraron 100 semillas, pero los datos de algunas parcelas se perdieron. Los resultados se presentan en el siguiente cuadro:

**Cuadro 1. Porcentaje de germinación de siete grupos experimentales de chocho con diferentes niveles de resistencia a hongos del suelo**

A	B	C	D	E	F	G
42	45	54	36	38	48	66
33	74	35	55	42	73	57
56	65	43	60	25	34	38
35	48	30	64	30	24	42
58	82	45	53	X	29	54
X	62	X	X	X	41	74

\*Las X en el cuadro indican que los datos de esas parcelas se perdieron.

Al examinar los datos se nota que existe considerable variación entre los porcentajes, de un mínimo de 24 a un máximo de 82. Esto podría indicar que hace falta transformar los datos. Pero es mejor asegurarse con la prueba W y obtener un gráfico que compare los porcentajes con una distribución normal (“Normal Probability Plot” o “Rankit”). Esto se obtiene con *Statistix* en la opción “Randomness/Normality tests”. El Gráfico, la prueba W y la probabilidad se presentan a continuación:

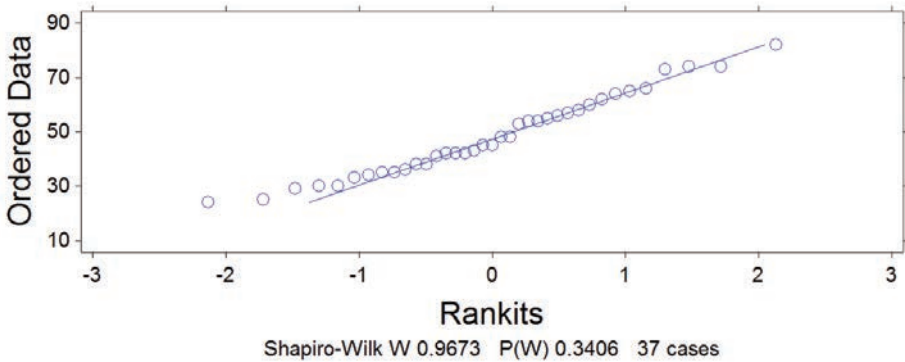


Gráfico 1. Prueba de normalidad de los datos de porcentaje de germinación de chocho

Como puede observarse, los datos de porcentaje (los círculos en el gráfico) están sobre la línea recta o cerca de ella, lo que indica que el conjunto de datos se aproxima a una distribución normal. Esto se confirma con el valor de la probabilidad de la prueba  $W$  que es  $P = 0,3406$ . Por lo tanto, se puede concluir que hay suficiente evidencia para aceptar la hipótesis nula de la normalidad de los porcentajes. En consecuencia, en este caso, no hace falta transformar los datos y se puede continuar con el análisis, como se observa en el Cuadro 2.

Cuadro 2. Totales de los tratamientos, número de repeticiones y medias para los datos de porcentaje de germinación de siete líneas de chocho

Tratamientos	A	B	C	D	E	F	G
Totales de los Tratamientos, $T_i$	224	376	207	268	135	249	331
Repeticiones, $r_j$	5	6	5	5	4	6	6
Medias, $\bar{Y}_i$	44.80	62.66	41.40	53.60	33.75	41.50	55.16

Para encontrar los componentes de las fuentes de variación se procede en forma similar al del ejercicio anterior. En primer lugar, mediante la fórmula computacional, se calcula el Factor de Corrección:

$$FC = \frac{(\sum Y_i)^2}{\sum r_j = n} = \frac{(42+45+\dots+74)^2}{37} = 86597.29$$

Luego se obtiene la Suma de Cuadrados Total:

$$SCT = \sum Y_i^2 - FC = (42^2 + 45^2 + \dots + 74^2) - 86597.29 = 8144.70$$

A continuación, se encuentra la Suma de Cuadrados de los Tratamientos:

$$SCTrat = \frac{\sum T_i^2}{r_j} - FC = \left( \frac{224^2}{5} + \frac{376^2}{6} \dots + \frac{331^2}{6} \right) - FC = 3085.09$$

(note que en la fórmula de la  $SCTrat$  el número de repeticiones es diferente y que por lo tanto no tenemos un denominador común, como fue el caso en el ejercicio anterior.)

Finalmente, por diferencia, se encuentra la Suma de Cuadrados del Error:

$$SCE = SCT - SCTrat = 8144.70 - 3085.09 = 5059.62$$

Se dispone ya de todos los componentes de Sumas de Cuadrados para preparar el cuadro ADEVA:

Cuadro 3. Análisis de Varianza

F.V.	S.C.	G.L.	C.M.	F	F <sub>0.05</sub>	F <sub>0.01</sub>
Líneas	3085.09	6	514.18	3.05*	2.42	3.47
Error	5059.62	30	168.65			
Total	8144.71	36				

La gran media experimental es:  $\bar{Y} = \frac{\sum Y_i}{n} = \frac{42+45+\dots+74}{37} = \frac{1790}{37} = 48.38$

El Coeficiente de Variación (CV) es:  $CV = \frac{\sqrt{CME}}{\bar{Y}} * 100 = \frac{\sqrt{168.65}}{48.38} * 100 = 26.84\%$

La prueba F igual a 3.05\* señala que por lo menos la diferencia de dos medias es estadísticamente significativa al nivel del 5% de probabilidad.

Con Statistix el análisis es el siguiente:

Statistix 9.0      DCA germ chocho,

One-Way AOV for: A B C D E F G

Source	DF	SS	MS	F	P
Lineas	6	3085.09	514.181	3.05	0.0189
Error	30	5059.62	168.654		
Total	36	8144.70			

Grand Mean 48.378      CV 26.84

<b>Homogeneity of Variances</b>	<b>F</b>	<b>P</b>
Levene's Test	0.82	0.5624
O'Brien's Test	0.61	0.7234
Brown and Forsythe Test	0.50	0.8068

<b>Variable</b>	<b>N</b>	<b>Mean</b>	<b>SE</b>
A	5	44.800	5.8078
B	6	62.667	5.3018
C	5	41.400	5.8078
D	5	53.600	5.8078
E	4	33.750	6.4933
F	6	41.500	5.3018
G	6	55.167	5.3018

### **Comentarios:**

- A pesar de que las observaciones eran porcentajes de germinación, se demostró, mediante la Prueba W, que en este caso no hacía falta transformar los datos. En otras palabras, la distribución de los datos no se apartaba sensiblemente de una distribución normal, requisito importante para la validez del modelo lineal aditivo.
- Además de la normalidad de los datos, se puede observar en los resultados que provee Statistix, que también se cumple el requisito de homogeneidad de las varianzas. Esto se deduce por los altos valores de las probabilidades de las tres pruebas: Levene, O'Brien y Brown.
- En el Cuadro ADEVA el valor de F igual a 3.05\* indica que por lo menos una diferencia entre dos medias es estadísticamente significativa. Cómo determinar qué diferencias son significativas se analizará en el Capítulo 17.

## Diseño de bloques completos al azar (bca)

*“Consultar a un estadístico al final del experimento equivale a pedirle que haga un examen ‘post mortem’. Quizá pueda decir la causa de su muerte”. Ronald Fisher<sup>28</sup>*

Cuando las unidades experimentales son heterogéneas, no es conveniente emplear un Diseño Completamente al Azar (DCA). El procedimiento más idóneo, en este caso, es formar grupos con las unidades experimentales en función de posibles semejanzas que puedan tener. Estos grupos reciben el nombre de “bloques”.

Suponga un experimento en el que se quiere evaluar el efecto de cuatro tratamientos. El investigador dispone de varias unidades experimentales. Como observa que existe variación entre ellas, decide agruparlas por semejanzas que logra identificar. Por lo tanto, forma tres bloques con cuatro unidades experimentales cada una (ver Gráfico 15.1.) Luego asignará al azar, los cuatro tratamientos, a las cuatro unidades experimentales dentro de cada bloque. Este diseño experimental se denomina *Diseño de Bloques Completos al Azar*, nombre que se abrevia como BCA.

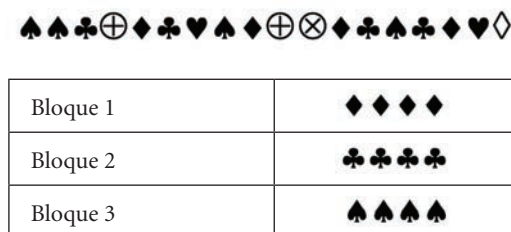


Gráfico 15.1. Unidades experimentales heterogéneas y su arreglo en bloques completos. Diseño de Bloques Completos al Azar.

La situación que se presenta en el Gráfico 15.1 sería la situación ideal. Parecería que las unidades experimentales en cada bloque son exactamente iguales.

28 Estadístico y biólogo inglés. [https://www.brainyquote.com/authors/ronald\\_fisher](https://www.brainyquote.com/authors/ronald_fisher)

Pero esto no se da en la vida real. Siempre habrá algunas diferencias entre las unidades experimentales *dentro* de cada bloque. En todo caso, lo que se busca es minimizar las diferencias dentro de los bloques. Las diferencias *entre* bloques disminuyen el error experimental, como se verá más adelante.

Un experimento conducido con un diseño de Bloques Completos al Azar (BCA) genera un análisis de varianza en dos direcciones. Es decir, hay dos criterios de clasificación de las observaciones: *Tratamientos* y *Bloques*.

Un bloque completo contiene todos los tratamientos, repetidos una sola vez. El término “bloque” en este caso es sinónimo de “repetición”. En capítulos posteriores se presentarán experimentos con bloques incompletos, en los que varios bloques incompletos forman una repetición.

El término “bloques” se aplicó por primera vez en experimentos en el campo y son secciones homogéneas del terreno experimental. Por extensión, el término “bloques” se emplea también en experimentos de producción animal o de otra índole. Por ejemplo, en un experimento en el que se desea probar el efecto de varias dietas en el engorde de cerdos, los bloques podrían ser varios corrales diferentes entre ellos, pero con condiciones similares dentro de cada corral. Los bloques también podrían ser diferentes razas de cerdos.

En experimentos en el laboratorio, en cámaras de crecimiento, en invernaderos, en experimentos con animales, el investigador puede físicamente mover sus unidades experimentales. En experimentos agronómicos esto no es posible, obviamente. Las unidades experimentales son, por lo general, pequeñas parcelas de terreno. Sin embargo, se sabe que parcelas que se encuentran cercanas entre ellas son bastante homogéneas. Por el contrario, parcelas distantes o alejadas unas de otras presentan variación considerable o gradientes en sus propiedades físicas, químicas y biológicas. Por esto se recomienda que los bloques sean lo más cuadrados que sea posible, con parcelas rectangulares dentro de los bloques.

En investigación agrícola el diseño más utilizado es el diseño BCA, pero muchas veces se decide a priori el diseño sin conocer el área experimental, lo cual seguramente llevará a resultados poco confiables. Para conducir un experimento se debe recabar información de posibles experimentos anteriores en el sitio, reconocer el área y en el caso de detectar una o más gradientes, planificar el diseño acorde con esto. En áreas nuevas que vayan a dedicarse a la investigación, pruebas de uniformidad permiten identificar la dirección de las gradientes.

Una vez que se conoce o se sospecha de la existencia de una gradiente, los bloques en el campo se organizan en forma perpendicular a la gradiente. Veamos un ejemplo:

**Ejercicio 15.1.** Suponga que se desea evaluar seis diferentes híbridos (A, B, C, D, E, F) de un cultivo. También suponga que en el terreno experimental hay una gradiente de fertilidad, de secciones más fértiles del terreno a secciones menos fértiles. El diseño más recomendado para este caso es el de BCA. Presente un cuadro con la dirección de la pendiente, cinco bloques completos y la aleatorización correspondiente:

Bloques							Dirección de la pendiente
1	F	A	D	C	E	B	++
2	C	D	B	A	E	F	+
3	D	C	F	E	B	A	0
4	E	A	B	F	C	D	-
5	B	A	F	C	D	E	--

Note que la pendiente corre en dirección perpendicular a los bloques. Con esto se logra que todos y cada uno de los tratamientos estén representados tanto en las áreas fértiles como en las áreas de baja fertilidad. De no haberse planificado de esta forma, uno o más tratamientos podrían estar sobre representados en las áreas más fértiles en desmedro de otros tratamientos. La introducción de un sesgo como el descrito, puede llevar a conclusiones erradas.

La aleatorización de los seis tratamientos dentro de cada bloque se realizó con el programa Statistix, en la forma descrita en el capítulo anterior.

En general, con el diseño BCA los tratamientos están presentes sólo una vez en cada bloque completo. El número de tratamientos determina el número de parcelas (unidades experimentales) en cada bloque. El número de bloques o repeticiones depende de algunos factores. Entre ellos se pueden señalar:

1. El grado de variación entre las unidades experimentales. A mayor variación puede considerarse aumentar el número de bloques.
2. El número de tratamientos. A mayor número de tratamientos, el número de bloques puede ser menor. La idea es lograr entre 14 y 20 grados de libertad para el error experimental.
3. Los recursos disponibles.
4. El grado de precisión que se desea tener en el experimento.



En términos generales, la precisión de un experimento se incrementa con un número mayor de bloques o repeticiones. Pero en la práctica, experimentos muy grandes no necesariamente proveen mejor información. Además, son más costosos, más difíciles de conducir y consecuentemente, más propensos a errores humanos.

### 15.1. Ventajas del diseño de Bloques Completos al Azar

Las ventajas del diseño BCA son:

1. Puede usarse con cualquier número de tratamientos y de bloques.
2. Los bloques disminuyen la variación del error experimental.
3. El análisis estadístico no presenta mayores complicaciones.
4. La pérdida de observaciones no es un problema insalvable. Hay métodos para estimar datos perdidos y continuar con el análisis.
5. Es el diseño que, después del DCA, presenta el mayor número de grados de libertad para el error experimental.

### 15.2. Desventajas del diseño de Bloques Completos al Azar

Si el diseño BCA se aplica correctamente, no existe ninguna desventaja en su uso. Sin embargo, cuando por las características del terreno, se usa este diseño cuando lo correcto hubiera sido usar algún otro, disminuye la precisión del experimento.

### 15.3. El modelo lineal aditivo para el diseño BCA

El modelo estadístico en el que se fundamenta el análisis del diseño BCA es el siguiente:

$$Y_{ij} = \mu + \tau_i + r_j + \varepsilon_{ij}$$

el modelo indica que una observación cualquiera  $Y_{ij}$  es el resultado de la acción aditiva de varios factores:

$\mu$ : la media general de todas las observaciones en las poblaciones de referencia en el experimento.

$\tau_i$ : el efecto de los tratamientos medidos como desviación de  $\mu$

$r_j$ : el efecto de los bloques medidos como desviación de  $\mu$

$\varepsilon_{ij}$  = El error al azar asociado con cada observación.

**Presunciones para la validez del modelo:**

1. Las poblaciones de las que provienen los tratamientos tienen igual varianza:

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$$

2. Las poblaciones de las que provienen los bloques tienen igual varianza:

$$\sigma_1^2 = \sigma_2^2 = \dots = \sigma_r^2$$

3. Las sumatorias de los efectos son iguales a cero:

$$\sum \tau_i = \sum r_j = 0$$

4. Los  $\varepsilon_{ij}$  son variables independientes al azar, con distribución normal, media igual a cero y varianza  $\sigma^2$ . Los  $\varepsilon_{ij}$  representan las desviaciones del modelo aditivo debidas al error experimental.

**15.4. Componentes de la variación en el diseño BCA**

Por lo señalado anteriormente se puede deducir que la variación en un diseño BCA proviene de las siguientes fuentes:

1. *Variación Total*. Mide la variación total en el experimento. Se estima por medio de la Suma de Cuadrados Total (SCT).
2. *Variación por tratamientos*. Mide la variación entre tratamientos. Se estima por medio de la Suma de Cuadrados de los Tratamientos (SCTrat).
3. *Variación por Bloques*. Mide la variación entre bloques o repeticiones. Se estima por medio de la Suma de Cuadrados de Bloques (SCR).
4. *Variación por el Error Experimental*. Mide las desviaciones de las observaciones con relación a los valores esperados de acuerdo con el modelo lineal aditivo. Se estima a través de la Suma de Cuadrados del Error (SCE), que resulta de la interacción entre bloques y tratamientos.

La relación entre los componentes de la variación es la siguiente, de acuerdo con el modelo lineal aditivo:

$$SCT = SCTrat + SCR + SCE$$

## 15.5. Procedimiento para la evaluación de las hipótesis

Las hipótesis que principalmente nos interesa evaluar son:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_k$$

$H_1$ : “Por lo menos dos medias de los **tratamientos** no son iguales”

Puede ser de interés, en algunos casos, también evaluar las siguientes hipótesis con relación a los bloques:

$$H_0: \mu_1 = \mu_2 = \dots = \mu_r$$

$H_1$ : “Por lo menos dos medias de los **bloques** no son iguales”

Así como en el DCA, en el diseño de BCA el modelo puede ser Rígido o fijo (Modelo I) y al Azar (Modelo II). El modelo puede también ser Mixto, en el que por ejemplo los tratamientos sean al azar y los bloques fijos. En el caso de un Modelo al Azar (Modelo II) las hipótesis que se evalúan son:

$$H_0: \sigma_\tau^2 = 0$$

$$H_1: \sigma_\tau^2 \neq 0$$

En un experimento conducido con un diseño BCA el número de tratamientos se representa con la letra  $k$ , mientras que el número de repeticiones o bloques con la letra  $r$ . En este diseño los términos “repeticiones” y “bloques” son sinónimos. Los resultados pueden presentarse como se indica en el Cuadro 15.5.1:

Cuadro 15.5.1. Presentación de los datos de un diseño BCA con  $k$  tratamientos y  $r$  bloques

Bloques	Tratamientos						
	1	2	3			$k$	
							$R_j$
1	$Y_{11}$	$Y_{12}$	$Y_{13}$	-	-	$Y_{1k}$	$R_1$
2	$Y_{21}$	$Y_{22}$	$Y_{23}$	-	-	$Y_{2k}$	$R_2$
3	$Y_{31}$	$Y_{32}$	$Y_{33}$	-	-	$Y_{3k}$	$R_3$
-	-	-	-	-	-	-	-
-	-	-	-	-	-	-	-
R	$Y_{r1}$	$Y_{r2}$	$Y_{r3}$	-	-	$Y_{rk}$	$R_r$
$T_i$	$T_1$	$T_2$	$T_3$	-	-	$T_k$	$\sum Y_{ij}$
$\bar{Y}_i$	$\bar{Y}_1$	$\bar{Y}_2$	$\bar{Y}_3$	-	-	$\bar{Y}_k$	$\bar{Y}$
							$n=rk$

Los símbolos en el cuadro anterior representan lo siguiente:  $Y_{ij}$  es la observación  $Y$  en el tratamiento  $i$ , bloque  $j$ ;  $T_i$  es la suma total de las observaciones en el tratamiento  $i$ ;  $R_j$  es la suma total de las observaciones en el bloque  $j$ . Los otros símbolos son conocidos o se explican por sí solos.

Una vez que se ha estructurado un cuadro similar al 15.1, pero con los datos reales, se puede proceder a calcular los componentes de la variación en un diseño de BCA:

1. Suma de Cuadrados Total (SCT):

$$SCT = \sum Y_{ij}^2 - FC, \text{ en la que el Factor de Corrección (FC) es: } FC = \frac{(\sum Y_{ij})^2}{rk=n}$$

2. Suma de Cuadrados de los Bloques o Repeticiones (SCR):

$$SCR = \frac{\sum R_j^2}{k} - FC \quad (89)$$

3. Suma de Cuadrados de los Tratamientos (SCTrat):

$$SCTrat = \frac{\sum T_i^2}{r} - FC \quad (90)$$

4. La Suma de Cuadrados del Error (SCE) se encuentra por diferencia:

$$SCE = SCT - SCR - SCTrat$$

Una vez obtenidos los valores de las fuentes de variación se estructura un cuadro ADEVA, cuyo esquema se presenta en el cuadro 15.5.2:

Cuadro 15.5.2. Esquema de un CUADRO ADEVA, en un Diseño de Bloques Completos al Azar

Fuente de Variación (FV)	Suma de Cuadrados (SC)	Grados de Libertad (gl)	Cuadrado Medio (CM)	Valores Fisher (F)
Bloques	$SCR$	$r-1$	$\frac{SCR}{r-1} = CMR$	$\frac{CMR}{CME}$
Tratamientos	$SCTrat$	$k-1$	$\frac{SCTrat}{k-1} = CMTrat$	$\frac{CMTrat}{CME}$
Error	$SCE$	$(r-1)(k-1)$	$\frac{SCE}{(r-1)(k-1)} = CME$	
TOTAL	$SCT$	$rk-1 = n-1$		

Puede ser de interés calcular la Eficiencia Relativa (ER) del diseño de Bloques Completos Al Azar (BCA) en comparación con el Diseño Completamente al Azar (DCA). Es decir, cuánto se ganó o se perdió al realizar el experimento con un BCA y no con un DCA. Esto puede ayudar en la planificación de futuros experimentos en el mismo sitio. La fórmula para este cálculo es:

$$ER = \frac{(r-1)*CMR+r(k-1)*CME}{(rk-1)*CME} \quad (91)$$

Fórmula en la que CMR es el Cuadrado Medio de las Repeticiones, CME es el Cuadrado Medio del Error, r es el número de repeticiones o bloques y k es el número de tratamientos.

Un ejercicio similar a un diseño BCA se vio ya en la sección 9.5. Se trata de la prueba *t* para observaciones pareadas. La igualdad se basa en la relación  $t^2 = F$  cuando se tiene únicamente dos tratamientos. Las unidades experimentales se agrupan en pares o bloques con dos unidades experimentales homogéneas. Sin embargo, cuando se requiere evaluar más de dos tratamientos, el procedimiento es mediante un análisis de varianza.

En el siguiente ejercicio se presenta el análisis de un experimento conducido con un diseño de bloques completos al azar (BCA):

---

**Ejercicio 15.5.1.** En un experimento se generaron progenies de cruzas entre cinco materiales de maíz reventón (canguil), con un híbrido de maíz semiduro dentado (Pardo, 1992)<sup>29</sup> En este ensayo se evaluó el volumen de expansión de las primeras cruzas (Filial 1 o F1). Como se detectó una gradiente en el terreno, el diseño seleccionado fue un BCA, con seis tratamientos (se añadió un testigo comercial) y cinco bloques. Realizar el análisis correspondiente con los resultados que se presentan en el siguiente cuadro. Calcular también, la eficiencia relativa del diseño empleado en comparación con un DCA. (Datos simplificados para facilidad del ejemplo):

---

29 Pardo, R.R. 1992. Estimación de heterosis en la primera generación de cruzas entre maíces reventones y un híbrido de maíz dentado blanco. Tesis. Escuela Agrícola Panamericana. Honduras, C.A.

Cuadro 1. Volumen de expansión de cruzas de maíz tipo canguil

Bloques	Tratamientos						$R_j$
	Te	A1	A2	Z1	Z2	Z3	
1	23.5	15.9	19.5	13.4	17.6	13.1	103.0
2	19.6	17.3	12.7	12.5	16.6	9.0	87.7
3	16.3	14.6	10.3	11.8	18.8	7.2	71.0
4	23.9	20.8	18.2	15.0	13.5	10.9	102.3
5	20.2	15.4	16.3	6.8	12.0	15.3	86.0
$T_i$	103.5	84.0	77.0	59.5	70.5	55.5	$\sum Y_{ij} = 450.0$
$\bar{Y}_i$	20.7	16.8	15.4	11.9	14.1	11.1	$\bar{Y} = 15$

Los maíces reventones de las cruzas fueron: A1 y A2, americanos introducidos de Nebraska e Indiana; Z1, Z2 y Z3 provenientes del banco de germoplasma, los dos primeros de origen ecuatoriano y el tercero de origen colombiano. El testigo, tratamiento Te, fue un maíz reventón comercial.

Para encontrar los componentes de variación debe calcularse primero el Factor de Corrección. Esto es:

$$FC = \frac{(\sum Y_{ij})^2}{rk = n} = \frac{(450)^2}{5*6} = 6750$$

La Suma de Cuadrados Total es:

$$SCT = \sum Y_{ij}^2 - FC = (23.5^2 + 15.9^2 + \dots + 15.3^2) - 6750 = 542.8$$

La Suma de Cuadrados de las Repeticiones es:

$$SCR = \frac{\sum R_j^2}{k} - FC = \frac{(103.0^2 + 87.7^2 + \dots + 86.0^2)}{5} - 6750 = 117.1$$

La Suma de Cuadrados de los Tratamientos es:

$$SCTrat = \frac{\sum T_i^2}{r} - FC = \frac{(103.5^2 + 87.7^2 + \dots + 55.5^2)}{6} - 6750 = 307.6$$

La Suma de Cuadrados del Error se encuentra por diferencia:

$$SCE = SCT - SCR - SCTrat = 542.8 - 117.1 - 307.6 = 118.1$$

Con esta información se estructura el cuadro ADEVA:

Cuadro 2. Análisis de Varianza de los datos de volumen de expansión en maíz canguil

FV	SC	gl	CM	F	F <sub>0.05</sub>	F <sub>0.01</sub>
Bloques	117.1	4	29.27	4.96**	2.87	4.43
Tratamientos	307.6	5	61.52	10.43**	2.71	4.10
Error	118.1	20	5.90			
Total	542.8	29				

Nota: dos asteriscos (\*\*) denotan valores de F significativos al nivel alfa de 0.01

Para los tratamientos, el valor de F igual a 10.43, significativo a una probabilidad  $P \leq 0.01$ , señala que hay diferencias entre las medias. En el Capítulo 17 se verá lo concerniente a los métodos de identificación de las medias estadísticamente diferentes de otras medias.

El valor de F significativo para bloques, a un nivel alfa también menor de 0.01, indica que se logró remover variación del error experimental gracias al empleo del diseño BCA.

La eficiencia relativa (fórmula 103) es:

$$ER = \frac{4*29.27+5*5*5.90}{29*5.90} = 1.55$$

El valor 1.55 indica que se ganó un 55% en eficiencia al emplear un diseño BCA y no un DCA. Generalmente, cuando se detecta una diferencia significativa entre bloques, la eficiencia relativa señalará que la precisión del experimento se incrementó al emplear un BCA.

Los resultados obtenidos con Statistix se presentan a continuación:

Randomized Complete Block AOV Table for Volumen

Source	DF	SS	MS	F	P
Bloques	4	117.097	29.2742		
Trat	5	307.600	61.5200	10.42	0.0000
Error	20	118.123	5.9062		
Total	29	542.820			

Grand Mean 15.000 CV 16.20

**Tukey's 1 Degree of Freedom Test for Nonadditivity**

Source	DF	SS	MS	F	P
Nonadditivity	1	5.534	5.53415	0.93	0.3460
Remainder	19	112.589	5.92575		

Relative Efficiency, RCB 1.55

**Means of Volumen for Trat**

Trat	Mean
Te	20.700
A1	16.800
A2	15.400
Z1	11.900
Z2	14.100
Z3	11.100

Observations per Mean                    5  
 Standard Error of a Mean                1.0868  
 Std Error (Diff of 2 Means)            1.5370

**Comentarios:**

- El Coeficiente de Variación de 16.20% en el Ejercicio 15.1 es relativamente alto, pero se justifica por tratarse de materiales heterocigóticos de diferentes procedencias.
- De acuerdo con el modelo lineal, los datos deben ser aditivos. El programa Statistix, en esta opción, realiza la prueba de Tukey para la falta de aditividad ( "*nonadditivity*", en los resultados que presenta). En el Ejercicio anterior, la probabilidad encontrada para la prueba de falta de aditividad fue de 0.3460. Como la probabilidad es mayor que 0.05 se acepta que los datos de este ejemplo sí son aditivos. En general, cuando los datos no son aditivos, las varianzas de los tratamientos no son homogéneas. Una posible solución cuando se presentan estos casos es el empleo de procedimientos de la estadística no paramétrica (Capítulo 24).





*Los medios principales para adquirir conocimiento son tres: observación, reflexión y experimentación. La observación recoge hechos, la reflexión los combina y la experimentación verifica el resultado”. Denis Diderot<sup>30</sup>*

El Cuadrado Latino (CL) es un diseño que tiene un número igual de tratamientos, columnas e hileras. Se designan como columnas e hileras a los equivalentes de los bloques de un diseño BCA. El nombre se debe al famoso matemático suizo Euler<sup>31</sup> que identificaba las casillas de un cuadrado con letras griegas y latinas para plantear conjeturas combinatorias. Sin embargo, las casillas o unidad experimentales pueden identificarse con cualquier letra, símbolo o número.

La característica del Cuadrado Latino es que cada tratamiento debe estar presente solo una vez en cada columna y solo una vez en cada hilera. Este diseño es de gran utilidad cuando en el terreno experimental se detectan dos gradientes, perpendiculares una de otra. Se lo utiliza también en ganadería, en experimentos en que la misma unidad experimental rota por diferentes períodos. En este caso, “las gradientes” son las unidades experimentales y los períodos. Con el mismo concepto, el diseño de Cuadrado Latino se utiliza en la industria y en otros campos.

## 16.1. La aleatorización en el Cuadrado Latino

La aleatorización de los tratamientos en un diseño Cuadrado Latino está restringida dentro de las hileras y columnas, en tal forma que todas las hileras y todas las columnas tengan todos los tratamientos. Esta restricción determina que el número de tratamientos sea igual al número de columnas y de hileras. Los diseños CL se identifican con el número de tratamientos  $k$ , elevado al cuadrado.

---

30 Filósofo y enciclopedista francés. [https://www.brainyquote.com/authors/denis\\_diderot](https://www.brainyquote.com/authors/denis_diderot)

31 <http://www.maa.org/press/periodicals/convergence/euler-squares-euler-squares>

Como ejemplo,  $k^2 = 4^2 = 16$  es un diseño CL con cuatro tratamientos y 16 unidades experimentales.

Desde el punto de vista matemático, el número de permutaciones<sup>32</sup> posibles que puede darse, para formar los cuadrados latinos se incrementa aceleradamente con el número  $k$ . Como ejemplo, para un cuadrado latino  $4^2$  hay 576 permutaciones. Para un  $5^2$  hay 161,280 permutaciones. Para un cuadrado latino  $9^2$ , el número de permutaciones posibles es la astronómica cantidad de  $5.5247 \times 10^{27}$ . El popular juego Sudoku tiene como base un cuadrado latino  $9^2$ , aunque por las restricciones propias del juego el número de permutaciones es menor que el señalado.

Para la aleatorización se disponen los tratamientos en forma sistemática, como se presenta en el Cuadro 16.1.1 para un Cuadrado Latino  $5^2$ . Note en el cuadro que la primera hilera tiene las letras que designan a los cinco tratamientos en orden alfabético. Luego, la segunda hilera empieza con la segunda letra y termina con la primera. Se sigue este procedimiento hasta completar el cuadro. Con este arreglo sistemático se cumple el requisito clave que ningún tratamiento puede repetirse en ninguna hilera o columna.

A continuación, se sortea el orden de las columnas e hileras del Cuadro 16.1.1. Los números al azar pueden obtenerse con cualquier método de los descritos. Para las columnas, el sorteo obtenido con Statistix fue: 1, 3, 4, 5, 2. Para las hileras fue: 2, 5, 3, 4, 1. El resultado, siguiendo el orden del sorteo, se presenta en el Cuadro 16.1.2.

Cuadro 16.1.1. Distribución sistemática de los tratamientos A, B, C, D, E en un Cuadrado Latino  $5^2$ , previa a la aleatorización

Hileras	Columnas				
	1	2	3	4	5
1	A	B	C	D	E
2	B	C	D	E	A
3	C	D	E	A	B
4	D	E	A	B	C
5	E	A	B	C	D

32 [https://en.wikipedia.org/wiki/Latin\\_square](https://en.wikipedia.org/wiki/Latin_square)

**Cuadro 16.1.2. Cuadrado Latino. Tratamientos aleatorizados en columnas e hileras a partir del Cuadro 16.1**

Hileras	Columnas				
	1	3	4	5	2
2	B	D	E	A	C
5	E	B	C	D	A
3	C	E	A	B	D
4	D	A	B	C	E
1	A	C	D	E	B

Note que el arreglo obtenido en el cuadro anterior es uno de los 161,280 posibles.

Como el número de columnas e hileras no tiene significado alguno, se puede, por conveniencia para evitar errores, enumerar nuevamente en forma ordinal las columnas e hileras.

## 16.2. Ventajas y Desventajas del Diseño Cuadrado Latino

El CL tiene la ventaja que remueve variación en dos direcciones, es decir, por hileras y columnas. Esta variación, de no usarse un CL, inflaría el Error Experimental. Otra ventaja es que el análisis y conducción del experimento son relativamente fáciles.

Las desventajas del CL son:

1. No se recomienda para experimentos con menos de cuatro tratamientos ni más de ocho. En el primer caso, por haber pocos grados de libertad para el error y en el segundo por ser experimentos demasiado grandes.
2. Si no se detectan claramente las fuentes de variación se pierden, innecesariamente, grados de libertad del error.
3. Si se pierden los datos de una o más unidades experimentales se deben estimar estos valores mediante el uso de fórmulas algo complicadas. (Sin embargo, los programas estadísticos por defecto estiman los valores perdidos.)

## 16.3. El modelo Lineal Aditivo para el Cuadrado Latino

El modelo lineal aditivo para el análisis del diseño Cuadrado Latino es:

$$Y_{ijk} = \mu + h_i + c_j + \tau_k + \varepsilon_{ijk}$$

Modelo en el que una observación cualquiera  $Y_{ijk}$  es el resultado aditivo de una media, del efecto de la hilera, del efecto de la columna, del efecto del tratamiento y del error experimental. Los componentes del error deben seguir una distribución normal, con media igual a cero y varianza desconocida. Una presunción importante en este modelo es la ausencia de interacción real entre tratamientos, hileras y columnas. La interacción resultante, se supone, proviene únicamente del azar y constituye el error experimental.

Las fuentes de variación en el Cuadrado Latino incluyen, además de las sumas de cuadrados ya identificadas en los diseños anteriores, las sumas de cuadrados de hileras y de columnas. La relación, que puede intuirse por el modelo lineal, es la siguiente:

$SCT = SCH + SCC + SCTrat + SCE$ . Las fórmulas para encontrar cada componente se presentan a continuación:

1. Suma de Cuadrados Total  $SCT = \sum Y_{ijk}^2 - FC$
2. Suma de Cuadrados Hileras  $SCH = \sum \frac{H_i^2}{k} - FC$  (92)
3. Suma de Cuadrados Columnas  $SCC = \sum \frac{C_j^2}{k} - FC$  (93)
4. Suma de Cuadrados Tratamientos  $SCTrat = \sum \frac{T_k^2}{k} - FC$  (94)
5. Suma de Cuadrados Error  $SCE = SCT - SCH - SCC - SCTrat$

$$\text{El Factor de Corrección es: } FC = \frac{(\sum Y_{ijk})^2}{k \cdot k}$$

Con la información anterior se puede estructurar el modelo del Análisis de Varianza (ADEVA):

Cuadro 16.3.1. Modelo del Cuadro ADEVA para el Diseño Cuadrado Latino

Fuente de Variación (FV)	Suma de Cuadrados (SC)	Grados de Libertad (gl)	Cuadrado Medio (CM)	Valores Fisher (F)
Hileras	$SCH$	$k-1$	$\frac{SCH}{k-1}$	$\frac{CMH}{CME}$
Columnas	$SCC$	$k-1$	$\frac{SCC}{k-1}$	$\frac{CMC}{CME}$
Tratamientos	$SCTrat$	$k-1$	$\frac{SCTrat}{k-1}$	$\frac{CMTrat}{CME}$
Error	$SCE$	$(k-1)(k-2)$	$\frac{SCE}{(k-1)(k-2)}$	
TOTAL	$SCT$	$k^2-1 = n-1$		

Veamos un ejemplo:

**Ejercicio 16.3.1.** En un experimento de cultivos asociados, conducido con un diseño Cuadrado Latino, se evaluó el efecto de cinco distancias de siembra de sorgo sobre el rendimiento de una variedad de soya (Romero, 1993)<sup>33</sup>. La aleatorización empleada fue la del Cuadro 16.2. Las distancias de siembra del sorgo fueron: A = 5 cm, B = 10 cm, C = 15 cm, D = 20 cm, E = 25cm. (Algunas características del ensayo se modificaron por razones del ejemplo.) Los rendimientos de la variedad de soya, en quintales métricos por hectárea (q/ha), se presentan en el siguiente cuadro:

**Cuadro 1. Resultados de rendimiento de soya en q/ha, bajo el efecto de cinco distancias de sorgo en cultivos asociados. Diseño Cuadrado Latino**

Hileras	Columnas					$H_i$
	1	2	3	4	5	
1	B=23.68	D=20.80	E=20.50	A=16.82	C=18.45	100.25
2	E=20.86	B=19.15	C=14.44	D=26.21	A=16.06	96.72
3	C=28.25	E=32.17	A=29.50	B=17.73	D=29.59	137.24
4	D=25.32	A=14.54	B=18.53	C=19.44	E=25.47	103.30
5	A=14.85	C=17.57	D=16.31	E=14.09	B=22.02	84.84
$C_j$	112.96	104.23	99.28	94.29	111.59	$\sum Y_{ijk} = 522.35$

	Distancias de siembra (tratamientos)				
	A=5 cm	B=10 cm	C=15 cm	D=20 cm	E=25 cm
$T_k$	91.77	101.11	98.15	118.23	113.09
$\bar{Y}_k$	18.35	20.22	19.63	23.65	22.62

Con los resultados del Cuadro 1 se realiza el Análisis de Varianza pertinente. Primero es necesario encontrar el Factor de Corrección:

$$FC = \frac{(\sum Y_{ijk})^2}{k*k} = \frac{(522.35)^2}{25} = 10,913.98$$

Debe, a continuación, calcularse las diferentes Sumas de Cuadrados:

33 Romero, A.R. 1993. Niveles de competencia de maíz y sorgo sobre el cultivo de soya en asocio, en el valle del Zamorano y en fincas de pequeños productores. Tesis. Escuela Agrícola Panamericana. Honduras, C.A.

$$SCT = \sum Y_{ijk}^2 - FC = (23.68^2 + 20.80^2 + \dots + 22.02^2) - 10,913.98 = 672.53$$

$$SCH = \sum \frac{H_i^2}{k} - FC = \frac{100.25^2 + 96.72^2 + \dots + 84.84^2}{5} - 10,913.98 = 307.69$$

$$SCC = \sum \frac{C_j^2}{k} - FC = \frac{112.96^2 + 104.23^2 + \dots + 111.59^2}{5} - 10,913.98 = 50.68$$

$$SCTrat = \frac{\sum T_k^2}{k} - FC = \frac{91.77^2 + 101.11^2 + \dots + 113.09^2}{5} - 10,913.98 = 95.23$$

$$SCE = SCT - SCH - SCC - SCTrat = 218.93$$

Los resultados del análisis de varianza (Cuadro 2) indican que las distancias de siembra de sorgo aparentemente no influyeron en el cultivo asociado de soya. El valor F de 1.31, con probabilidad  $P > 0.05$ , así lo sugiere. Se esperaba que las poblaciones más altas de sorgo ejercieran mayor competencia sobre las plantas de soya, por luz y nutrientes. Posiblemente la distancia entre surcos de soya y sorgo pudo influir en la falta de efecto.

Cuadro 2. Cuadro ADEVA para el rendimiento de soya. Diseño Cuadrado Latino

EV.	S.C.	g. l.	C.M	F	F <sub>05</sub>	F <sub>01</sub>
Hileras	307.69	4	76.92	4.22*	3.26	5.41
Columnas	50.68	4	12.67	0.69 n.s.		
Distancias	95.23	4	23.81	1.31 n.s.		
Error	218.93	12	18.24			
TOTAL	672.53	24				

$$CV = \frac{\sqrt{18.24}}{20.89} * 100 = 20.44\%$$

Las hileras removieron variación que hubiera inflado el error experimental. Este efecto de las hileras fue significativo ( $P < 0.05$ ). La remoción de variabilidad por parte de las columnas no fue significativa. De realizarse otro experimento en el mismo terreno se recomendaría emplear un BCA en el que las hileras serían los bloques.

Los resultados con el programa Statistix son los siguientes:

Statistix 9.0                   soya dist sorgo CL,

**Latin Square AOV Table for Rendimien**

Source	DF	SS	MS	F	P
Hileras	4	307.690	76.9225		
Columnas	4	50.680	12.6700		
Distancia	4	95.233	23.8082	1.30	0.3228
Error	12	218.931	18.2442		
Total	24	672.534			

Grand Mean 20.894       CV 20.44

**Tukey's 1 Degree of Freedom Test for Nonadditivity**

Source	DF	SS	MS	F	P
Nonadditivity	1	0.632	0.6322	0.03	0.8616
Remainder	11	218.299	19.8453		

	Relative Efficiency
Completely Randomized Design	1.49
Randomized Complete Block, Hileras	0.91
Randomized Complete Block, Columnas	1.59

**Means of Rendimien for Distancia**

Distancia	Mean
5	18.354
10	20.222
15	19.630
20	23.646
25	22.618

Observations per Mean           5  
 Standard Error of a Mean       1.9102  
 Std Error (Diff of 2 Means)   2.7014



## Comentarios:

- El coeficiente de variación del ejercicio anterior es muy alto para un cultivo de auto-polinización como la soya. Posiblemente se presentaron problemas en la conducción del ensayo, como la presencia de plagas y períodos de sequía, que pudieron afectar diferencialmente los rendimientos de las parcelas.
- De acuerdo con la información de la eficiencia relativa, que provee Statistix, hubiera sido mejor realizar el experimento con un diseño de Bloques Completos al Azar. Si los bloques hubieran sido las hileras, se hubiera ganado un 59% más en precisión. Esto, principalmente porque se tendría 16 grados de libertad para el error. Con un Diseño Completamente al Azar también se hubiera ganado en precisión, en comparación con el Cuadrado Latino. Esta información es importante para futuros experimentos en el mismo sitio.
- Como los tratamientos fueron cuantitativos, se podrían evaluar posibles tendencias mediante análisis de regresión o el método simplificado de regresión que se conoce como *comparaciones ortogonales*. En el Capítulo 17 se verá un ejemplo con los datos de este ejercicio.

## 16.4. El Cuadrado Latino en investigaciones pecuarias

En investigaciones pecuarias puede emplearse cualquier diseño experimental. Sin embargo, el principal problema en algunas investigaciones es contar con suficientes unidades experimentales. En este caso es importante el uso de alguno de los diseños reversibles<sup>34</sup>, en los que un mismo animal o grupo de animales recibe dos o más tratamientos en secuencia. El Cuadrado Latino, en experimentos con animales, es parte de los diseños reversibles. En este caso, cada animal recibe en secuencia, a través del tiempo, todos los tratamientos en estudio. Las secuencias o períodos constituyen las hileras, mientras que los animales se constituyen en las columnas.

Como se mencionó, una limitación importante en los cuadrados latinos es el tamaño del experimento, lo que determina pocos grados de libertad para el error. Con tres tratamientos se tiene solo dos grados de libertad para el error. Con cuatro tratamientos, los grados de libertad del error son seis. Con pocos grados de libertad se pierde precisión en el experimento. A este inconveniente se une el

34 Morris T.R. 1999. *Experimental Design and Analysis in Animal Sciences*. CABI Publishing, New York. Chapter 5, Change-Over Designs, p. 42-52

hecho de que por lo general se dispone de pocos animales que pueden asignarse a la investigación.

La solución a los dos problemas anteriores es el empleo de un experimento con dos o más cuadrados latinos de igual tamaño. Obviamente, los tratamientos son los mismos en los dos cuadrados, pero también los períodos (hileras) o los animales (columnas) podrían ser los mismos.

Una vez que se ha ejecutado el experimento, se realiza el análisis de cada cuadrado en forma individual y si las varianzas son homogéneas, se pueden combinar los resultados en un solo cuadro ADEVA. El esquema del análisis de varianza para el caso de dos cuadrados latinos  $4^2$ , en los que las hileras son “períodos” y las columnas “animales”, se presenta en el Cuadro 16.4.1.

El Cuadro 16.4.1 se explica por sí solo. En la situación que consta en el cuadro, se supone que hubo ocho períodos y ocho animales o grupos de animales. De ahí los seis grados de libertad de cada uno. Esto implica que los ensayos se realizaron en diferentes épocas. Si los dos ensayos se realizaran simultáneamente, los períodos serían los mismos y habría solo tres grados de libertad. Los tres restantes pasarían al error experimental.

**Cuadro 16.4.1. Esquema del Análisis de Varianza combinado para dos cuadrados latinos  $4^2$ , ejecutados en períodos diferentes y con diferentes animales**

Fuente de Variación	Grados de libertad
Cuadrados	$c-1 = 1$
Períodos dentro de los cuadrados	$c(k-1) = 6$
Animales dentro de los cuadrados	$c(k-1) = 6$
Tratamientos	$k-1 = 3$
Error	$c(k-1)(k-2) + (c-1)(k-1) = 15$
TOTAL	$ck^2-1 = 31$

Veamos un ejemplo:

**Ejercicio 16.4.1.** En una empresa avícola se planificó evaluar el efecto de cuatro dietas en el peso de los huevos. La empresa produce su propio alimento (el tratamiento testigo T), mientras que las otras dietas (tratamientos A, B y C) provienen de la competencia. Se decidió emplear dos cuadrados latinos  $4^2$ . Para el efecto, se designaron al azar ocho jaulas con 10 ponedoras en cada jaula. Las ponedoras

eran de la misma edad y procedencia genética. Los animales de cada jaula fueron alimentados durante cuatro períodos con cada una de las dietas. Los períodos de alimentación duraron 24 días cada uno, pero sólo se registraron los datos de peso de huevos en los últimos 12 días de cada período. El objetivo de esto era eliminar cualquier efecto remanente de la dieta anterior. Los dos cuadrados latinos se ejecutaron simultáneamente. La aleatorización y resultados preliminares obtenidos se presentan en los Cuadros 1 y 2. Realizar los análisis estadísticos pertinentes:

Cuadro 1. Cuadrado Latino 1. Peso promedio de huevos, gramos

Períodos	Jaulas: 1 a 4				$P_i$
	1	2	3	4	
1	T=47.7	C=41.9	B=58.3	A=50.5	198.4
2	C=44.6	B=60.2	A=54.1	T=52.2	211.1
3	B=57.3	A=59.7	T=51.1	C=47.2	215.3
4	A=55.2	T=49.8	C=45.3	B=61.4	211.7
$J_j$	204.8	211.6	208.8	211.3	$\sum Y_{ijk} = C_1 = 836.5$

	Dietas			
	A	B	C	T
$T_k$	219.5	237.2	179.0	200.8
$\bar{Y}_k$	54.9	59.3	44.8	50.2

Cuadro 2. Cuadrado Latino 2. Peso promedio de huevos, gramos

Períodos	Jaulas: 5 a 8				$P_i$
	5	6	7	8	
1	B=55.7	T=51.6	C=46.5	A=54.6	208.4
2	A=56.4	C=43.9	B=52.3	T=48.8	201.4
3	C=42.9	A=53.1	T=54.3	B=59.0	209.3
4	T=52.4	B=62.8	A=57.4	C=46.7	219.3
$J_j$	207.4	211.4	210.5	209.1	$\sum Y_{ijk} = C_2 = 838.4$

	Dietas			
	A	B	C	T
$T_k$	221.5	229.8	180.0	207.1
$\bar{Y}_k$	55.4	57.5	45.0	51.8

Observe que en los Cuadros 1 y 2, los períodos  $P_i$  son las hileras, mientras que las jaulas  $J_j$  son las columnas. Los resultados de los análisis con Statistix se presentan a continuación:

CUADRADO LATINO 1

Statistix 9.0

**Latin Square 1 AOV Table for PESOHUEVO**

Source	DF	SS	MS	F	P
PERIODO	3	40.922	13.641		
JAULAS	3	7.417	2.472		
DIETAS	3	468.167	156.056	30.52	0.0005
Error	6	30.679	5.113		
Total	15	547.184			

Grand Mean 52.281 CV 4.33

**Means of PESOHUEVO for DIETAS**

DIETAS	Mean
A	54.875
B	59.300
C	44.750
T	50.200

CUADRADO LATINO 2

Statistix 9.0

**Latin Square 2 AOV Table for PESOHUEVO**

Source	DF	SS	MS	F	P
PERIODO	3	40.715	13.572		
JAULAS	3	2.285	0.762		
DIETAS	3	358.015	119.338	13.03	0.0049
Error	6	54.945	9.157		
Total	15	455.960			

Grand Mean 52.400    CV 5.78

**Means of PESOHUEVO for DIETAS**

DIETAS	Mean
A	55.375
B	57.450
C	45.000
T	51.775

Para combinar los dos cuadrados latinos, en un solo análisis similar (pero no igual) al del esquema del Cuadro 16.4.1, es importante que las varianzas del error experimental de los dos cuadros ADEVA presenten homogeneidad. De lo contrario, los resultados se discutirían en forma individual para cada caso.

La homogeneidad de las varianzas se evalúa mediante la prueba F (ver sección 9.1). Generalmente se escoge el nivel alfa de probabilidad de 0.05. En la tabla F (Apéndice D) el valor correspondiente a seis grados de libertad para el numerador y seis para el denominador es igual a 4.28. La prueba F es:

$$F = \frac{CME_{(mayor)}}{CME_{(menor)}} = \frac{9.16}{5.11} = 1.79$$

El valor de la prueba F no supera al valor de la tabla. Se concluye que las varianzas son homogéneas (el nivel de error es el 5%). Por lo tanto, pueden combinarse los dos cuadrados latinos en un solo cuadro ADEVA. Si las varianzas no fueran homogéneas se violaría una presunción necesaria para el análisis y se correría el riesgo de arribar a conclusiones erróneas.

Para combinar los dos cuadrados latinos en un solo cuadro es importante tener en cuenta que los períodos y los tratamientos (dietas) fueron los mismos, no así las jaulas. Por lo tanto, deben sumarse las cantidades correspondientes. Esto se presenta en el Cuadro 3:

**Cuadro 3. Totales de períodos y dietas de los dos cuadrados latinos combinados**

	Períodos				Total Cuadrados
	1	2	3	4	
Cuadrado 1	198.4	211.1	215.3	211.7	$C_1 = 836.5$
Cuadrado 2	208.4	201.4	209.3	219.3	$C_2 = 838.4$
Total $P_{ii}$	406.8	412.5	424.6	431.0	$C = C_1 + C_2 = 1674.9$

	Tratamientos				Total Cuadrados
	A	B	C	T	
Cuadrado 1	219.5	237.2	179.0	200.8	$C_1 = 836.5$
Cuadrado 2	221.5	229.8	180.0	207.1	$C_2 = 838.4$
Total $T_{kk}$	441.0	467.0	359.0	407.9	$C = C_1 + C_2 = 1674.9$
Medias $\bar{Y}_{kk}$	55.13	58.38	44.88	50.99	

El procedimiento para construir el cuadro ADEVA del análisis combinado es el siguiente:

1. Primero se obtiene la Suma de Cuadrados Total de todas las observaciones (32 en total). Esto es:

$$SCT_{CL} = \sum \sum Y_{ijk} - \frac{(\sum \sum Y_{ijk})^2}{2k^2} = (47.7^2 + 41.9^2 + \dots + 46.7^2) - \frac{(1674.9)^2}{32} = 1003.26$$

Note que, para el análisis combinado de los dos cuadrados latinos, el Factor de Corrección es:

$$FC = \frac{(\sum \sum Y_{ijk})^2}{2k^2} = \frac{C^2}{2k^2} = \frac{1674.9^2}{32} = 87665.3$$

2. Luego se encuentra la Suma de los Cuadrados de los cuadrados latinos:

$$SC(CL) = \frac{C_1^2 + C_2^2}{k^2} - FC = \frac{836.5^2 + 838.4^2}{16} - \frac{1674.9^2}{32} = 0.11$$

3. La Suma de Cuadrados de los Períodos se encuentra a partir de los totales  $P_{ii}$  (Cuadro 3):

$$SCP_{ii} = \frac{\sum P_{ii}^2}{2k} - FC = \frac{406.8^2 + 412.5^2 + 424.6^2 + 431.0^2}{8} - \frac{1674.9^2}{32} = 45.77$$

4. Como las jaulas del cuadrado 1 eran distintas de las jaulas del cuadrado 2, tan solo se añaden las Sumas de Cuadrados de Jaulas de los dos ADEVA. Esta cantidad es igual a 9.7 y recibe el nombre de "Suma de Cuadrados de Jaulas dentro de los cuadrados". Tiene, en este caso, seis grados de libertad (un total de ocho jaulas, menos un grado por jaulas y otro por los cuadrados).

5. La Suma de Cuadrados de los Tratamientos se encuentra a partir de los totales  $T_{kk}$  (Cuadro 3):

$$SCTrat_{kk} = \frac{\sum T_{kk}^2}{2k} - FC = \frac{441.0^2 + 467.0^2 + 359.0^2 + 407.9^2}{8} - \frac{1674.9^2}{32} = 813.86$$

6. La Suma de Cuadrados del Error se puede encontrar por diferencias a partir de la Suma de Cuadrados Total. Pero también, a partir de la suma de cuatro componentes ortogonales (independientes) del error. Los dos primeros son las  $SCE$  de los dos cuadrados latinos: 30.67 y 54.94. El tercer valor es la diferencia entre la suma de las  $SCP$  de los dos cuadrados latinos y el valor de la  $SCP_{ii}$ :  $(40.92 + 40.71) - 45.77 = 35.86$ . El cuarto valor es la diferencia entre la suma de las  $SCTrat$  de los dos cuadrados latinos y el valor de la  $SCTrat_{kk}$ :

$(468.16 + 358.01) - 813.86 = 12.31$ . Al sumar los valores de los cuatro componentes se obtiene la Suma de Cuadrados del Error del análisis combinado:

$SCE = 30.67 + 54.94 + 35.86 + 12.31 = 133.78$  (hay una pequeña diferencia por redondeo).

Al sumar los grados de libertad de los cuatro componentes, el aporte total es de 18 grados.

Con todos estos elementos se puede construir el cuadro ADEVA:

**Cuadro 4. Análisis de Varianza de dos Cuadrados Latinos  $4^2$  combinados**

Fuente de Variación	<i>S. C.</i>	<i>g. l.</i>	<i>C. M.</i>	<i>F</i>	$F_{0.05}$	$F_{0.01}$
Cuadrados Latinos	0.11	1				
Jaulas, dentro de los cuadrados	9.70	6				
Períodos	45.77	3	15.26	2.05 n.s.	3.16	5.09
Tratamientos	813.86	3	271.29	36.51**		
Error	133.82	18	7.43			
TOTAL	1003.26	31				

Puede observarse que no hubo diferencias entre los períodos. Las diferencias entre tratamientos o dietas fueron estadísticamente significativas ( $P < 0.01$ ).

## Métodos de comparación de medias

*“Si crees que tu experimento necesita un estadístico, lo que quizá en verdad necesitas es un mejor experimento”.* Ernest Rutherford<sup>35</sup>

En la planificación de los experimentos hay varios puntos importantes que considerar, entre ellos y en este orden, la selección de los tratamientos que conduzcan a las respuestas esperadas, la selección de las unidades experimentales en relación con sus características y la decisión de qué diseño experimental emplear. Aspectos de gran importancia en la planificación de experimentos y que muchas veces no se toman en cuenta son las pruebas posteriores que deben realizarse después del Análisis de Varianza (ADEVA).

En el análisis de varianza la hipótesis nula se refiere a todas las medias de los tratamientos. En las pruebas posteriores, que reciben el nombre de “métodos de separación o comparación de medias”, hay varias hipótesis nulas cuyo número depende del número de comparaciones que se hagan. Estos métodos no pueden usarse indiscriminadamente y dependen de varias condiciones, como se demostrará posteriormente. En todo caso, si el método escogido no es el adecuado no se logrará extraer toda la información posible de los resultados obtenidos y lo que es peor, se puede llegar a conclusiones erradas.

Es importante señalar que los métodos de separación de medias pueden emplearse con cualquier diseño experimental.

La primera condición necesaria para decidir el método de separación de medias tiene que ver con los tratamientos o variable independiente. Hay dos tipos de tratamientos: los cualitativos y los cuantitativos.

---

35 Ernest Rutherford. Físico neozelandés. [https://todayinsci.com/R/Rutherford\\_Ernest/RutherfordErnest-Quotations.htm](https://todayinsci.com/R/Rutherford_Ernest/RutherfordErnest-Quotations.htm)



**Tratamientos cualitativos** son aquellos que pueden ser descritos con palabras y que tienen un grupo de características propias, limitadas a un espacio muy corto. Un grupo de variedades de un cultivo, o los fungicidas que pueden comprarse en un mercado, son ejemplos de tratamientos cualitativos. Cada variedad tiene un nombre y un conjunto de características genéticas específicas. Así mismo, cada fungicida tiene un nombre y un ingrediente activo específico. En un experimento se puede determinar qué variedad fue igual o mejor que otras o qué fungicida tuvo mayor efecto. No hay puntos intermedios.

**Tratamientos cuantitativos** son aquellos que se describen con números y que se refieren a medidas de peso, distancia, tiempo, temperatura, etc. Sus límites sólo dependen de la precisión de los instrumentos de medida. Diferentes niveles de un fertilizante o diferentes dosis de un pesticida son tratamientos cuantitativos. Si en un experimento se evalúan cuatro dosis de un fertilizante: 0, 50, 100, 200 kg, y el resultado indica que 100 kg causó el mayor efecto, esa cantidad es sólo un punto de referencia de una escala continua. Bien pudo cualquier valor, menor o mayor que 100 kg, dar igual o mejor resultado que el obtenido.

## Comparación de medias con tratamientos cualitativos

Cuando los tratamientos son cualitativos se presentan tres opciones para la identificación de diferencias estadísticamente significativas:

1. comparaciones entre todas las medias, igual número de repeticiones
2. comparaciones entre todas las medias, desigual número de repeticiones
3. comparaciones de las medias con la media testigo y
4. comparaciones ortogonales entre grupos de medias.

### 17.1 . Todas las medias, igual número de repeticiones

La comparación de todas las medias de par en par es pertinente cuando los tratamientos no pueden agruparse con algún criterio específico. Como ejemplo, en un experimento para evaluar diferentes variedades de un cultivo, este tipo de comparaciones es el procedimiento más indicado. Cuando se realizan todas las comparaciones posibles entre medias, de par en par, el número de comparaciones está dado por la siguiente fórmula:

$$w = \frac{k*(k-1)}{2} \quad (95)$$

en la que  $w$  es el número de comparaciones y  $k$  es el número de tratamientos o medias. Como ejemplo, si se tiene cuatro medias A, B, C, D, las seis comparaciones posibles serán: A-B, A-C, A-D, B-C, B-D, C-D. A medida que se incrementa el número de medias el número de comparaciones sube exponencialmente. Esto se demuestra en el siguiente cuadro:

**Cuadro 171.1. Relación entre  $k$  número de medias  
y  $w$  número de comparaciones posibles entre todas las medias**

Valor de $k$	3	4	5	6	7	8	9	10	11	12	13	14	$k$
Valor de $w$	3	6	10	15	21	28	36	45	55	66	78	91	$w$

en forma convencional, las pruebas se realizan al nivel de significación alfa del 5% o del 1%. Esto, por la restricción que imponen las tablas respectivas. Sin embargo, mediante el empleo de programas computarizados, el investigador puede escoger el nivel de significación que desee, aunque esto determine mayor o menor riesgo de cometer algún tipo de error. (Error Tipo I – rechazar una hipótesis nula verdadera – si el valor alfa es muy alto. Error Tipo II – aceptar una hipótesis nula falsa – si el valor alfa es muy bajo.)

Es importante considerar que, en algunas pruebas, a medida que aumenta el número de comparaciones, también aumenta la probabilidad de declarar diferentes medias que en realidad no lo son. Es decir, se incrementa la probabilidad de cometer Error Tipo I (rechazar una hipótesis nula verdadera).

Con el objetivo de disminuir la probabilidad de cometer Error Tipo I, el valor crítico de algunas pruebas varía de acuerdo con el número de medias en comparación. Sin embargo, en este último caso, con valores críticos más altos, puede cometerse Error Tipo II, que es la probabilidad de aceptar una hipótesis nula falsa. De acuerdo con el tipo de experimento, el investigador deberá decidir qué implica mayor riesgo: rechazar una hipótesis nula verdadera o aceptar una hipótesis nula falsa. Deberá consecuentemente decidir qué prueba emplear. Estas son las incertidumbres que debe enfrentar el investigador.

Existen varias pruebas para efectuar todas las comparaciones posibles entre medias. Se analizarán las siguientes:

1. la prueba de la Diferencia Mínima Significativa (DMS)
2. la prueba de Tukey (TUK)
3. la prueba de Scheffé (SCH)
4. la prueba de Bonferroni (BON)

5. la prueba de amplitud múltiple de Duncan (AMD)
6. la prueba de Student-Newman-Keuls (SNK)

Las pruebas mencionadas, con excepción de la DMS, pueden realizarse incluso si se obtiene, en el cuadro ADEVA, un valor F no significativo. El valor F se refiere a una sola hipótesis nula y en la práctica representa un promedio de todas las medias. Este promedio podría encubrir algunas diferencias reales entre medias.

La prueba de la DMS para comparar más de tres o cuatro medias no es adecuada. Esto, por cuanto la probabilidad de cometer Error Tipo I sube a medida que se incrementa el número de medias en comparación. Cochran y Cox (1968)<sup>36</sup> señalan que, si las pruebas son independientes y se realizan cinco comparaciones supuestamente al nivel alfa 5%, la probabilidad de declarar diferentes dos medias que no lo son sube a 23%; si son 10 comparaciones, sube a 40%; si son 20 comparaciones, sube a 64%.

Solo para efectos del ejemplo y para compararla con otras pruebas que ofrecen algún nivel de protección de cometer Error Tipo I, se incluye en esta sección la prueba de la DMS.

En todos los casos de las pruebas mencionadas se evalúan las diferencias entre pares de medias. La fórmula para decidir si una diferencia es estadísticamente significativa o no, a un nivel de probabilidad prefijado, es la siguiente desigualdad:  $|\bar{Y}_i - \bar{Y}_j| \geq V.C. Prueba$ . Es decir que, si el valor absoluto<sup>37</sup> de la diferencia es igual o mayor que el valor crítico de la prueba, las dos medias son estadísticamente diferentes. Note que en forma explícita se acepta o se rechaza la hipótesis de igualdad de las dos medias.

### 1. La Diferencia Mínima Significativa (DMS)

Como la mayoría de las pruebas, La Diferencia Mínima Significativa (DMS) se basa en la distribución t. La fórmula para calcular el valor crítico de la DMS es la siguiente:

$$DMS = t_{\frac{\alpha}{2}} * S_{\bar{d}} \quad (96)$$

36 Cochran, W.G. and G.M. Cox. 1968. *Experimental Designs*. Second Edition, first corrected printing, John Wiley & Sons, New York. p. 74.

37 Valor absoluto es el valor numérico sin tomar en cuenta el signo.

En la fórmula anterior,  $t_{\alpha/2}$  es el valor en la distribución t con los grados de libertad del Cuadrado Medio del Error (CME) y al nivel de probabilidad  $\frac{\alpha}{2}$ . El valor alfa sobre dos es la mitad del nivel de significación al que se evaluó la hipótesis nula en el análisis de varianza. Se divide para dos porque sólo interesa evaluar la diferencia entre dos medias, sin importar si están a la izquierda o a la derecha del centro de la distribución t. Se recordará que el error estándar de la diferencia de dos medias,  $S_{\bar{a}}$ , se usó en el Capítulo 7 en la prueba t para observaciones pareadas. La fórmula es:  $S_{\bar{a}} = \sqrt{2 * CME / r}$

La Diferencia Mínima Significativa (DMS), como el nombre lo indica, es el valor mínimo que la diferencia absoluta (independiente del signo) entre dos medias debe superar para que las medias se declaren estadísticamente diferentes. (En otras pruebas de separación de medias, el equivalente al concepto de la DMS, se denomina sencillamente *valor crítico*.) Este concepto, expresado en una fórmula, es el siguiente:

$$|\bar{Y}_a - \bar{Y}_b| \geq DMS$$

En otras palabras, si la diferencia absoluta entre dos medias es igual o mayor que la DMS, la diferencia es estadísticamente significativa al nivel que se haya realizado la prueba. Veamos un ejemplo:

**Ejercicio 17.1.1.** En el Ejercicio 14.7.1, en un experimento conducido con un DCA se estudió el efecto de seis tratamientos (fungicidas) en el rendimiento de café. Separe las medias usando la prueba DMS. El cuadro ADEVA y el cuadro de medias se presentan a continuación:

Cuadro 1. Análisis de Varianza (Cuadro ADEVA)

F.V.	S.C.	G.L.	C.M.	F	F <sub>0.05</sub>	F <sub>0.01</sub>
Fungicidas	735.81	5	147.16	30.53**	2.77	4.25
Error	86.77	18	4.82			
Total	822.58	23				

Cuadro 2. Medias de rendimientos de café

Fungicidas	A	B	C	D	E	F
Medias, $\bar{Y}_i$	19.47	19.45	23.65	13.47	12.62	6.85

El valor de la prueba F fue significativo al nivel alfa del 1%, es decir: ( $P \leq 0.01$ ). Consecuentemente se rechazó la hipótesis nula de igualdad de todas las diferencias entre medias. El problema es ahora determinar qué medias son diferentes.

Para encontrar el valor de la DMS a un nivel alfa del 0.01 debe definirse primero las cantidades que intervienen en la fórmula de la DMS:

El valor de  $t_{\alpha/2}$  con 18 grados de libertad y al nivel de 0.005 se lo encuentra en la Tabla t del Apéndice B. El valor es igual a 2.88.

El error estándar de la diferencia de dos medias es igual a:

$$S_{\bar{d}} = \sqrt{\frac{2CME}{r}} = \sqrt{\frac{2(4.82)}{4}} = 1.55$$

Por lo tanto, el valor de la Diferencia Mínima Significativa es:

$$DMS = t_{\alpha/2} S_{\bar{d}} = 2.88 * 1.55 = 4.46$$

Como en el experimento hubo seis medias, el valor de  $w$ , número total de comparaciones, es igual a 15, de acuerdo con la fórmula (95). Según la expresión  $|\bar{Y}_a - \bar{Y}_b| \geq DMS$ , si la diferencia entre las medias supera el valor de 4.46, entonces la diferencia se declara estadísticamente significativa al nivel alfa del 1%. Tomemos como ejemplo las primeras cuatro comparaciones:

**Cuadro 3. Una muestra de comparaciones de medias con la DMS**

Comparación	A menos B	A menos C	A menos D	A menos E
Diferencia de Medias	19.47-19.45	19.47-23.65	19.47-13.47	19.47-12.62
Diferencias absolutas	0.02	4.18	6.00	6.85
Significación	n.s.	n.s.	**	**

Nota: n.s. indica que la diferencia no fue significativa, dos asteriscos (\*\*) indica que la diferencia fue significativa al 1%.

Todas las 15 comparaciones se pueden obtener con el programa *Statistix*:

Statistix 9.0 ROYA CAFÉ 2,

**LSD All-Pairwise Comparisons Test of REND by FUNG**

FUNG	Mean	A	B	C	D	E
A	19.475					
B	19.450	0.025				
C	23.650	4.175	4.200			
D	13.475	6.000*	5.975*	10.175*		
E	12.625	6.850*	6.825*	11.025*	0.850	
F	6.8500	12.625*	12.600*	16.800*	6.625*	5.775*

Alpha 0.01 Standard Error for Comparison 1.5525  
 Critical T Value 2.878 Critical Value for Comparison 4.4689

Las 15 comparaciones se presentan en una matriz triangular. De las 15 comparaciones, 11 resultaron significativas. Las diferencias significativas tienen un asterisco, aunque el nivel alfa fue el 1%. Al pie de la matriz observe el valor t, el valor del error estándar de la diferencia de dos medias y la Diferencia Mínima Significativa (Critical Value for Comparison) igual a 4.46. (Es importante recordar que esto sólo sirve de ejemplo. En la práctica, el empleo de la prueba DMS para todas las comparaciones posibles, como el de este ejemplo, no es adecuado.)

## 2. La prueba de Tukey (TUK)

La prueba de Tukey (TUK) puede usarse en experimentos con cualquier número de tratamientos cualitativos para evaluar estadísticamente todas las diferencias entre pares de medias.

El valor crítico para la diferencia entre dos medias de tratamientos es un solo valor que toma en cuenta el número de tratamientos. Esto es, protege en buena medida contra la probabilidad de cometer Error Tipo I (rechazar una hipótesis nula verdadera), pero en cambio aumenta la probabilidad de cometer Error Tipo II (aceptar una hipótesis nula falsa). De todas maneras, la prueba de Tukey es una de las más usadas en la investigación, por el nivel de protección que ofrece y por la facilidad de ejecución.

El valor para la prueba de Tukey (TUK), o valor crítico, se encuentra mediante la siguiente fórmula:

$$TUK = q_{\alpha} * S_{\overline{Y}} \quad (97)$$

El valor  $q_{\alpha}$  se encuentra en la tabla de Tukey, en el apéndice E, al nivel  $\alpha$  que se vaya a realizar la prueba (5% o 1%), con los grados de libertad del CME y en la columna que corresponda al número de medias ( $p$ ) en comparación.

La estadística  $S_{\overline{Y}}$  es el error estándar de la media. La fórmula es:  $S_{\overline{Y}} = \sqrt{\frac{CME}{r}}$ .

El valor crítico de la prueba TUK es igual al de la DMS cuando se comparan únicamente dos medias. Si son más de dos medias, el valor crítico de TUK es más alto que el de DMS. Por esta razón, en la prueba TUK se reduce la probabilidad de cometer Error Tipo I. Veamos un ejemplo:

**Ejercicio 17.1.2.** Con los datos del Ejercicio 14.7.1 realice todas las 15 comparaciones posibles entre pares de medias con la prueba de Tukey, al nivel alfa del 1%.

El cuadro del análisis de varianza (Cuadro ADEVA) y las medias de los seis tratamientos (fungicidas) se presentaron en los Cuadros 1 y 2 del ejercicio anterior (Ejercicio 17.1.1).

Del cuadro ADEVA se requiere la siguiente información para calcular el error estándar:

$$CME = 4.82, \quad gl \text{ Error} = 18, \quad r = 4$$

$$\text{El error estándar de la media es: } S_{\overline{Y}} = \sqrt{\frac{CME}{r}} = \sqrt{\frac{4.82}{4}} = 1.10$$

Para encontrar el valor de  $q_{\alpha}$  debemos acudir a la Tabla de Tukey en el Apéndice E. Allí, al nivel de alfa al 1%, con  $p = 6$  y grados de libertad del error = 18, se encuentra el valor  $q_{\alpha} = 5.60$

Por lo tanto, el valor crítico de la prueba de Tukey es:

$$TUK = q_{\alpha} * S_{\overline{Y}} = 5.60 * 1.10 = 6.16$$

Las diferencias en valor absoluto (sin tomar en cuenta los signos), entre cualquier par de medias que supere el valor crítico de la prueba TUK, se declara significativa al nivel alfa que se haya escogido. Esto es:  $|\overline{Y}_i - \overline{Y}_j| \geq TUK$ . Como ejemplo, al comparar la media A con la C, se tiene lo siguiente:

El programa Statistix presenta los siguientes resultados:

Statistix 9.0 ROYA CAFÉ 2,

**Tukey HSD All-Pairwise Comparisons Test of REND by FUNG**

FUNG	Mean	A	B	C	D	E
A	19.475					
B	19.450	0.025				
C	23.650	4.175	4.200			
D	13.475	6.000	5.975	10.175*		
E	12.625	6.850*	6.825*	11.025*	0.850	
F	6.8500	12.625*	12.600*	16.800*	6.625*	5.775

Alpha 0.01 Standard Error for Comparison 1.5525  
 Critical Q Value 5.602 Critical Value for Comparison 6.1503

Puede verse que con la prueba de Tukey son ocho las diferencias estadísticamente significativas al nivel alfa del 1%. Note que el programa encuentra el valor de Tukey con la siguiente fórmula:

$$TUK = q_{\alpha} * \frac{S_{\bar{d}}}{\sqrt{2}} = 5.60 * \frac{1.55}{1.41=6.15}$$

### 3. La prueba de Scheffé (SCH)

La prueba de Scheffé se basa en la distribución F. Se emplea para realizar todas las comparaciones posibles entre pares de medias. También se utiliza para realizar todas las comparaciones posibles o contrastes entre grupos de medias, aunque no hayan sido planificadas de antemano. Para el primer caso, el valor crítico de la prueba SCH se encuentra con la siguiente fórmula:

$$SCH = S_{\bar{d}} * \sqrt{gk * F_{\alpha}} \quad (98)$$

La expresión  $gk$  se refiere a los grados de libertad de los tratamientos. La expresión  $F_{\alpha}$  se refiere al valor de F, con el nivel alfa escogido para la prueba y con los grados de libertad de tratamientos en el numerador y del error en el denominador. El valor de F se lo encuentra en la tabla F del Apéndice D. Veamos un ejemplo:

---

**Ejercicio 17.1.3.** Con la información del Ejercicio 14.7.1 realice con la prueba de Scheffé todas las comparaciones posibles entre pares de medias e identifique las diferencias significativas al nivel alfa del 1%.



Se procede a encontrar los componentes de la fórmula de la prueba SCH. El valor de los grados de libertad de los tratamientos  $gk$  es 5. El valor de  $F_\alpha$  al 1%, con 5 grados de libertad de los tratamientos en el numerador y 18 grados de libertad del error en el denominador es igual a 4.25 (tablas F en los Apéndices D).

El valor de la prueba SCH es:

$$SCH = S_{\bar{a}} * \sqrt{gk * F_\alpha} = 1.5525 * \sqrt{5 * 4.25} = 7.15$$

Statistix 9.0

ROYA CAFÉ 2,

**Scheffe All-Pairwise Comparisons Test of REND by FUNG**

FUNG	Mean	A	B	C	D	E
A	19.475					
B	19.450	0.025				
C	23.650	4.175	4.200			
D	13.475	6.000	5.975	10.175*		
E	12.625	6.850	6.825	11.025*	0.850	
F	6.8500	12.625*	12.600*	16.800*	6.625	5.775

Alpha	0.01	Standard Error for Comparison	1.5525
Critical F Value	4.248	Critical Value for Comparison	7.1550

El valor crítico de Scheffé para las comparaciones fue 7.16. Observe en este ejemplo que, con la prueba SCH, cinco diferencias fueron estadísticamente significativas al nivel alfa del 1%.

**Comentario:** la prueba de Scheffé, en comparación con otras pruebas, disminuye sensiblemente la probabilidad de cometer Error Tipo I (rechazar una hipótesis nula verdadera). En este sentido, la prueba es excesivamente conservadora.

#### 4. La prueba de Bonferroni (BON)

La prueba de Bonferroni se basa en la distribución t. El valor crítico de la prueba de Bonferroni se encuentra con la siguiente fórmula:

$$BON = t_{\alpha'} * S_{\bar{a}} \quad (99)$$

Note en el subíndice de t la presencia de alfa prima,  $\alpha'$ . Este valor, conocido como la “corrección de Bonferroni”, tiene la finalidad de disminuir la probabilidad de cometer Error Tipo I. Alfa prima se encuentra con la siguiente fórmula:

$$\alpha' = \frac{\alpha}{2*w} \quad (100)$$

En esta corrección,  $\alpha$  es el nivel de significación de la prueba y  $w$  es el número de todas las comparaciones posibles entre pares de medias. Se divide para dos porque los signos de las diferencias pueden ser positivos o negativos, es decir, pueden ubicarse a la izquierda o a la derecha de la curva de la distribución t.

En el Apéndice B se encuentran los valores de t correspondientes a diferentes probabilidades alfa, prefijadas de antemano. Sin embargo, las combinaciones de valores de alfa y  $m$  son infinitas y consecuentemente no hay una tabla para valores alfa prima. Sin embargo, mediante programas computarizados como Statistix, Excel u otros se pueden generar valores de alfa prima.

Veamos un ejemplo:

---

**Ejercicio 17.1.4.** Con los datos del Ejercicio 14.7.1, empleando la prueba de Bonferroni, identifique las diferencias estadísticamente significativas entre pares de medias al nivel alfa del 0.01, es decir 1%.

El primer paso es encontrar el valor de alfa prima,  $\alpha'$ . Esto es:

$$\alpha' = \frac{\alpha}{2*m} = \frac{0.01}{2*15} = 0.00033$$

Esta cantidad representa una fracción de la distribución t a cada extremo de la curva. Mediante la función t inversa, en la opción “probability functions” de Statistix se encuentra el valor de t para alfa prima igual a 0.00033 y 18 grados de libertad para el error:

$$t_{(0.00033,18gl)} = 4.11$$

Con esta información se puede proceder a determinar el valor crítico de la prueba de Bonferroni:

$$BON = t_{\alpha'} * S_{\bar{d}} = 4.11 * 1.55 = 6.37$$

Toda diferencia entre dos medias, que sea igual o superior al valor crítico de BON (6.37), es declarada estadísticamente significativa al nivel alfa del 1%, en este ejemplo.

Con Statistix los resultados son:

Statistix 9.0 ROYA CAFÉ 2,

**Bonferroni All-Pairwise Comparisons Test of REND by FUNG**

FUNG	Mean	A	B	C	D	E
A	19.475					
B	19.450	0.025				
C	23.650	4.175	4.200			
D	13.475	6.000	5.975	10.175*		
E	12.625	6.850*	6.825*	11.025*	0.850	
F	6.8500	12.625*	12.600*	16.800*	6.625*	5.775

Alpha 0.01 Standard Error for Comparison 1.5525  
 Critical T Value 4.104 Critical Value for Comparison 6.3712

Note que con esta prueba, ocho de las 15 diferencias resultaron estadísticamente significativas.

**Comentario:** a semejanza de otras pruebas que presentan valores críticos altos, la probabilidad de cometer Error Tipo II se incrementa en la prueba de Bonferroni. Como ejemplo, la diferencia entre A y D es 6.00. Con la prueba de Bonferroni se concluye que la diferencia no es significativa. Es decir, se acepta la hipótesis nula de que las dos medias son iguales. Pero suponga que esa diferencia fuera realmente significativa. Si este fuera el caso se estaría aceptando una hipótesis nula falsa. Es decir, se cometería Error Tipo II.

### 5. La prueba de la Amplitud Múltiple de Duncan (AMD)

La prueba de la Amplitud Múltiple de Duncan (AMD) es una de las pruebas más utilizadas en el análisis de experimentos. Es una prueba diferente a las anteriores, porque no compara todas las medias de par en par. Es algo más complicada que las otras pruebas porque emplea diferentes amplitudes (“rangos”) de acuerdo con el número de medias en comparación. Es decir, no hay un solo valor crítico como en las otras pruebas. Si  $k$  es el número de medias, la prueba AMD tiene  $k-1$  valores críticos.

Para su uso, deben ordenarse las medias en forma ascendente o descendente. Una particularidad de esta prueba es que se pueden comparar dos medias adyacentes entre sí, pero también tres o más medias contiguas, simultáneamente.

Un principio importante es que medias que compartan una misma amplitud no pueden ser declaradas diferentes.

Para marcar las diferencias estadísticas entre medias no debe emplearse una matriz de diferencias, como se hizo en las pruebas anteriores. Como son amplitudes que incluyen varias medias, lo procedente es presentar un cuadro con las medias ordenadas y segmentos de líneas rectas que marquen las amplitudes no significativas. También se puede señalar con una misma letra, generalmente minúscula, la amplitud no significativa.

La prueba AMD ofrece niveles de protección contra los Errores Tipo I, pero esto se logra a expensas de aumentar el riesgo de Errores Tipo II.

El valor crítico AMD se lo encuentra con la siguiente fórmula:

$$AMD = q_{\alpha'} * S_{\bar{Y}} \quad (101)$$

El error estándar de la media es igual a:  $S_{\bar{Y}} = \sqrt{\frac{CME}{r}}$ .

Los valores  $q$  prima de la amplitud múltiple ( $q_{\alpha'}$ ) son variables que dependen del número de medias ( $p$ ) en comparación, los grados de libertad del Cuadrado Medio del Error (CME) y el nivel alfa al que se realizarán las comparaciones. Los valores  $q_{\alpha'}$  se encuentran en la Tabla de Duncan, Apéndice F, al nivel alfa indicado, en las columnas marcadas con una  $p$  (número de medias en comparación) y en la hilera que corresponda a los grados de libertad del error.

Veamos un ejemplo:

---

**Ejercicio 17.1.5.** Con los datos del Ejercicio 14.7.1 compare las medias empleando la prueba de la amplitud múltiple de Duncan (AMD)

Las medias ordenadas se presentan en el siguiente Cuadro:

**Cuadro 1. Medias ordenadas para la prueba AMD**

Fungicidas	C	A	B	D	E	F
Medias, $\bar{Y}_i$	23.65	19.48	19.45	13.48	12.63	6.85

En el Cuadro 2 se presentan el número de medias ( $p$ ) en comparación, los valores  $q$  alfa prima de la tabla de Duncan con 18 grados de libertad (los del error experimental) y al nivel alfa del 1%, el valor del error estándar y los valores críticos de la prueba AMD.

Cuadro 2. Cálculo de los valores críticos para la prueba AMD

Medias en Comparación	p = 2	p = 3	p = 4	p = 5	p = 6
Valor $q_{\alpha}$	4.07	4.27	4.38	4.46	4.53
Valor $S_{\bar{y}}$	1.10	1.10	1.10	1.10	1.10
$AMD = q_{\alpha} * S_{\bar{y}}$	4.47	4.70	4.82	4.91	4.98

Note los valores críticos AMD en la última fila del Cuadro 2.

En el Cuadro 3 se presentan las diferencias entre las medias previamente ordenadas. (Son, obviamente, las mismas 15 diferencias de otras pruebas, pero presentadas en forma diferente.)

Cuadro 3. Diferencias entre medias

	A	B	D	E	F
C	4.17	4.20	10.17	11.02	16.80
A	-	0.03	6.00	6.85	12.63
B	-	-	5.97	6.82	12.60
D	-	-	-	0.85	6.63
E	-	-	-	-	5.78

Para definir qué grupo de medias es igual o estadísticamente diferente, se procede en forma similar que en las otras pruebas. Es decir que, si la diferencia absoluta entre medias supera el valor crítico, las medias son significativamente diferentes:  $|\bar{Y}_i - \bar{Y}_j| \geq AMD$ .

Generalmente se comparan primero las dos medias de los extremos, en este caso la media C y la media F. Si la diferencia entre las medias no fuera significativa, entonces ninguna otra diferencia entre las medias de la amplitud será estadísticamente diferente. Si la diferencia fuera significativa, se procede a comparar la misma media C con la siguiente del extremo, es decir con la media E, en este ejemplo. Se continúa de esta forma hasta que se haya comparado todas las medias. Veamos todo el proceso de separación de medias:

1. La diferencia absoluta entre C y F es 16.80 (Cuadro 3). El valor crítico AMD, para toda la amplitud de seis medias en comparación, es 4.98 (Cuadro 2). Como la diferencia supera el valor crítico, las dos medias C y F son estadísticamente diferentes al nivel alfa de 0.01.

2. La diferencia entre C y la siguiente media E es 11.02. El valor crítico para la amplitud de cinco medias es 4.91. Las medias C y E son estadísticamente diferentes.

3. La diferencia entre C y D es 10.17. El valor crítico para la amplitud de cuatro medias es 4.82. Las dos medias C y D son estadísticamente diferentes.

4. La diferencia entre C y B es 4.20. El valor crítico para la amplitud de tres medias es 4.70. Las dos medias C y B y la que se encuentra entre ellas (media A) son estadísticamente iguales al nivel alfa del 0.01. La media C ya no interviene en ninguna otra comparación. La amplitud no significativa (medias C, A, y B) se identificará con la letra minúscula a.

5. La comparación que sigue es entre A y F. La diferencia es significativa, así como las diferencias entre A y E y entre A y D.

6. La diferencia entre B y F es significativa, así como las diferencias entre B y E y entre B y D.

7. La diferencia entre D y F es igual a 6.63 y es estadísticamente significativa. La diferencia entre D y E es 0.85 y no es significativa. Las medias D y E se marcarán con una letra común.

8. La última comparación es entre las medias E y F. La diferencia es igual a 5.78. El valor crítico es igual a 4.47. La diferencia es estadísticamente significativa.

El resultado final se presenta en el Cuadro 4, en el que la significación se indica con letras diferentes y la no significación con letras iguales, debajo de cada media. (Se puede utilizar también rayas que unan las medias iguales, pero se prefiere el uso de letras.)

**Cuadro 4. Separación de medias mediante la prueba de Duncan al 1%**

Fungicidas	C	A	B	D	E	F
Medias, $\bar{Y}_i$	23.65	19.48	19.45	13.48	12.63	6.85
Significación	a	a	a	b	b	c

Se puede concluir, entonces, que en el experimento se identificaron tres grupos de medias o tratamientos. Los fungicidas C, A y B son estadísticamente iguales entre ellos y diferentes de los otros dos grupos. El tratamiento F (testigo) resultó diferente de los otros tratamientos.

**Comentario:** habrá casos en los que las igualdades y diferencias no estén tan claramente definidas como en este ejercicio. Es decir, casos en que algunas medias tengan más de una letra en la fila de “significación”.

## 6. La prueba de Student-Newman-Keuls (SNK)

La prueba de Student-Newman-Keuls es una prueba de separación de medias de amplitud múltiple, similar a la prueba de Duncan. Es decir, se pueden comparar simultáneamente varias medias. La fórmula que se emplea para encontrar los diferentes valores críticos es la siguiente:

$$SNK = Q_{\alpha} * S_{\bar{Y}} \quad (102)$$

Los valores  $Q_{\alpha}$  se encuentran en la tabla de amplitudes estudentizadas del Apéndice E (la misma tabla que se emplea para la prueba de Tukey), al nivel de significación que se haya planificado (1% o 5%) y con los grados de libertad del Cuadrado Medio del Error. Dependiendo del número de medias que se vaya a comparar, se seleccionan los valores correspondientes de la variable  $Q_{\alpha}$ . El estadígrafo  $S_{\bar{Y}}$  es el ya conocido error estándar de la media. Veamos un ejemplo:

**Ejercicio 17.1.6.** Con los datos del Ejercicio 14.7.1 compare las medias empleando la prueba de Student-Newman-Keuls (SNK) al nivel alfa del 1w% de significación.

En primer lugar, se requiere ordenar las medias como se hizo en la prueba Duncan. En este caso, como se trata del mismo ejemplo, se puede usar el cuadro con las medias ordenadas del ejercicio anterior (Cuadro 1, Ejercicio 17.1.5).

Luego se procede a encontrar los valores Q alfa prima en la tabla del Apéndice E. Los grados de libertad y el nivel alfa se encuentran en la primera columna de la tabla. En el caso de este ejercicio, en la hilera con 18 grados de libertad (los del error experimental) y al nivel alfa del 1%. Como son seis medias, los valores que se seleccionan están en las columnas del 2 al 6 en el cuerpo de la tabla. Esto se presenta en el Cuadro 1.

Cuadro 1. Número de medias, Q alfa prima, error estándar y cálculo de los valores críticos para la prueba SNK

Número de medias en Comparación	p = 2	p = 3	p = 4	p = 5	p = 6
Valores $Q_{\alpha}$	4.07	4.70	5.09	5.38	5.60
Valor $S_{\bar{Y}}$	1.10	1.10	1.10	1.10	1.10
$SNK = Q_{\alpha} * S_{\bar{Y}}$	4.47	5.17	5.60	5.92	6.15

Los valores críticos de la prueba SNK resultan de multiplicar los datos de la variable Q alfa prima por el error estándar de la media. Como ejemplo, cuando la

amplitud incluye cuatro medias se tiene:  $SNK = Q_{\alpha'} * S_{\overline{Y}} = 5.09 * 1.10 = 5.60$ . Esta información se encuentra en el Cuadro 1.

Para definir qué medias son iguales o estadísticamente diferentes, se procede en forma similar a la de otras pruebas. Es decir que, si la diferencia absoluta entre medias supera el valor crítico SNK, las medias son diferentes:  $|\overline{Y}_i - \overline{Y}_j| \geq SNK$ .

Igual que con la prueba de Duncan, en esta prueba también se comparan simultáneamente grupos de medias. Observe, por ejemplo, que para la diferencia de 16.80 entre las medias C y F, el valor crítico SNK es igual a 6.16. Esta comparación incluye a las seis medias. Si resultara no significativa, no habría nada más que hacer. Igualmente, para evaluar, por ejemplo, la diferencia de 6.85 entre A y E, el valor crítico de SNK es 5.60. Esta comparación incluye la amplitud de cuatro medias (A, B, D, E).

El resultado final se presenta en el cuadro siguiente:

**Cuadro 3. Separación de medias mediante la prueba de Student-Newman-Keuls, al nivel alfa del 1%**

Fungicidas	C	A	B	D	E	F
Medias, $\overline{Y}_i$	23.65	19.48	19.45	13.48	12.63	6.85
Significación, al nivel alfa del 1%	a	a	a	b	b	c

La significación se representa con letras minúsculas diferentes y la no significación con letras iguales, debajo de cada media, como se hizo en el caso de la prueba AMD (ejercicio anterior). Se concluye, al nivel de significación del 1%, que los fungicidas C, A y B son estadísticamente iguales entre ellos y diferentes de los otros dos grupos. El tratamiento F fue el testigo. La letra c indica que resultó diferente de los demás tratamientos.

**Comentario:** Vale el mismo comentario que se hizo para la prueba Duncan. Esto es, que habrá casos en los que las igualdades y diferencias de medias con la prueba SNK no estén tan claramente definidas como en este ejercicio. Es decir, casos en que algunas medias tengan más de una letra en la fila de “significación”.

## Resumen comparativo entre los métodos de separación de medias

En el cuadro siguiente se presenta un breve análisis comparativo de los métodos de separación de medias:



Cuadro 171.2. Comparación entre métodos de separación de medias, sobre la base de los resultados del Ejercicio 14.71

Prueba	Valor crítico Alfa 0.01	Diferencias Significativas
DMS	4.47	11
TUK	6.15	8
SCH	7.16	5
BON	6.37	8
AMD, p=2	4.47	-
AMD, p=6	4.98	-
SNK, p=2	4.47	-
SNK, p=6	6.15	-

Hay varias observaciones que pueden hacerse sobre el cuadro anterior:

1. Con la prueba que se obtuvo el mayor número de diferencias significativas fue con la prueba de la Diferencia Mínima Significativa (DMS). Esto se debe a que tiene el valor crítico más bajo y consecuentemente la probabilidad de cometer Error Tipo I es alta. Esta prueba es excesivamente liberal y no se recomienda para comparar todos los pares de medias.
2. Las pruebas de Tukey (TUK) y Bonferroni (BON) tuvieron resultados similares: ocho diferencias entre medias resultaron significativas. Las dos protegen en buen grado contra el Error Tipo I, por lo que se las considera conservadoras. Esto tiene un precio y es una mayor probabilidad de cometer Error Tipo II. En general con la prueba BON se obtendrá un valor crítico algo mayor que con la prueba TUK. Podría considerarse un inconveniente de la prueba BON el que no haya tablas específicas y haya que recurrir a programas computarizados para determinar el valor crítico.
3. La prueba de Scheffé (SCH) es una prueba conservadora. Note que dio como resultado cinco diferencias significativas en el ejemplo analizado. Esta prueba se basa en la distribución F, a diferencia de las otras del cuadro anterior que tienen como fundamento la distribución t. Una característica de esta prueba es que, si el valor de F en el cuadro ADEVA no es significativo, ninguna diferencia entre medias será significativa.
4. Tanto la prueba de Amplitud Múltiple de Duncan (AMD) como la prueba de Student-Newman-Keuls (SNK) son diferentes de las otras. Las comparaciones que se realizan no son de par en par de medias, sino que se comparan todas las medias simultáneamente dentro de una amplitud determinada. La amplitud de los valores críticos, en el ejemplo para seis

medias, varió entre 4.47 y 4.98 para la prueba AMD y entre 4.47 y 6.15 para la prueba SNK.

5. El resultado similar de las dos pruebas AMD y SNK en los ejemplos respectivos se debió a que el valor F del ADEVA tuvo una probabilidad bastante baja ( $P < 0.0000$ ), mientras que la separación de medias se hizo con el nivel alfa 1%, es decir, ( $P=0.01$ ). Las dos pruebas, AMD y SNK, tienen el inconveniente de que su cálculo y ejecución son más complicados por el hecho de requerir de varios valores críticos.
6. Observe que, en todas las pruebas cuando se trata de dos medias, el valor crítico es igual al de la DMS, esto es, 4.47 en el ejemplo. En el caso de la prueba SNK, el valor crítico más alto es igual al valor crítico de la prueba de Tukey, es decir, en este ejemplo, 6.15.

Como se indicó en secciones anteriores, el investigador debe tomar la decisión de qué prueba usar, conociendo los riesgos que se presentan. Sin embargo, se recomienda el uso de la prueba SNK. No es tan liberal como la prueba AMD ni tan conservadora como las pruebas de Tukey, Scheffé o Bonferroni.

## 17.2. Todas las medias, desigual número de repeticiones

Cuando el número de repeticiones  $r$  no es igual para todos los tratamientos, las fórmulas para encontrar los valores críticos de las diferentes pruebas se modifican para compensar por esta variable.

El proceso para comparar todos los pares de medias se vuelve más laborioso cuando el número de repeticiones de las medias es diferente. Mientras que algunas pruebas sólo presentan un valor crítico cuando las repeticiones son iguales, las mismas pruebas presentan varios valores críticos cuando el número de repeticiones no es igual.

### 1. La Diferencia Mínima Significativa (DMS), $r$ no iguales

La fórmula para calcular la DMS es la misma que se presentó en la sección anterior, esto es:  $DMS = t_{\alpha/2} * S_{\bar{d}}$ . Sin embargo, el error estándar de la diferencia de dos medias  $S_{\bar{d}}$  debe tomar en cuenta el diferente número de repeticiones. La fórmula es la siguiente:

$$S_{\bar{d}} = \sqrt{CME * \frac{r_m + r_n}{r_m * r_n}} \quad (103)$$

Las variables  $r_m$  y  $r_n$  corresponden al número de repeticiones de cualquier par de medias en comparación.

Como se indicó en la sección anterior, no se recomienda el uso de la DMS para realizar todas las posibles comparaciones entre pares de medias. Sin embargo, por razones didácticas, veamos el siguiente ejemplo del experimento en que se evaluaron siete líneas de chocho:

**Ejercicio 17.2.1.** Con los datos del Ejercicio 14.8.1 y empleando la prueba DMS, compare la media de la línea A con la media de la línea B, al nivel alfa de 0.05, es decir, 5%. El cuadro ADEVA se presenta a continuación:

**Cuadro 1. Análisis de Varianza (ADEVA) de los datos de germinación de siete líneas de chocho**

F.V.	S.C.	G.L.	C.M.	F	F <sub>0.05</sub>	F <sub>0.01</sub>
Líneas	3085.09	6	514.18	3.05*	2.42	3.47
Error	5059.62	30	168.65			
Total	8144.71	36				

Las medias y el número de repeticiones de cada tratamiento se presentan en el siguiente cuadro:

**Cuadro 2. Número de repeticiones y medias de los datos de porcentaje de germinación de siete líneas de chocho**

Líneas de chocho	A	B	C	D	E	F	G
Repeticiones, $r$	5	6	5	5	4	6	6
Medias, $\bar{Y}_i$	44.80	62.66	41.40	53.60	33.75	41.50	55.16

La fórmula para la DMS es:  $DMS = t_{\alpha/2} * S_{\bar{a}}$ . El valor  $t_{\alpha/2}$  se lo encuentra en el Apéndice B, al 0.025 de probabilidad y 30 grados de libertad (los del Cuadrado Medio (C.M)). Esto es:

$$t_{(0.025)(30gl)} = 2.0420$$

El error estándar, como ejemplo, para la diferencia de las medias A (cinco repeticiones) y B (seis repeticiones) es:

$$S_{\bar{a}} = \sqrt{CME * \frac{r_m + r_n}{r_m * r_n}} = \sqrt{168.65 * \frac{5+6}{30}} = 7.86$$

Consecuentemente, el valor de la DMS es:  $DMS = t_{\alpha/2} * S_{\bar{d}} = 2.0420 * 7.86 = 16.05$

La diferencia entre las dos medias A y B es, en valor absoluto, igual a 17.86. Dos medias son estadísticamente diferentes, al nivel de significación especificado, cuando la diferencia absoluta entre las medias es igual o mayor que la DMS. Es decir que:

$|\bar{Y}_A - \bar{Y}_B| \geq DMS$ . En este caso, la diferencia es 17.86 y supera a la DMS que es 16.05. Por lo tanto, las medias A y B son estadísticamente diferentes al nivel del 0.05 de significación.

Todas las comparaciones entre medias de este ejemplo se pueden realizar con el programa Statistix. Los resultados se presentan a continuación:

Statistix 9.0      DCA germ chocho,

**LSD All-Pairwise Comparisons Test**

Variable	Mean	A	B	C	D	E	F
A	44.800						
B	62.667	17.867*					
C	41.400	3.400	21.267*				
D	53.600	8.800	9.067	12.200			
E	33.750	11.050	28.917*	7.650	19.850*		
F	41.500	3.300	21.167*	0.100	12.100	7.750	
G	55.167	10.367	7.500	13.767	1.567	21.417*	13.667

Alpha      0.05      Standard Error for Comparison      7.4979 TO 8.7117

Critical T Value 2.042      Critical Value for Comparison      15.313 TO 17.792

Observe, en los resultados que presenta Statistix, que los errores estándar de las diferencias y los valores críticos para la comparación son variables. Esto es lo que se espera por razón de los diferentes números de repeticiones en los tratamientos.

Note que, de las 21 comparaciones posibles, seis diferencias resultaron estadísticamente significativas.

## 2. La prueba de Tukey (TUK), $r$ no iguales

Cuando se tiene diferente número de repeticiones en las medias, la fórmula de la prueba TUK es la siguiente:

$$TUK' = \frac{q_{\alpha} * S_{\bar{d}}}{\sqrt{2}} \quad (104)$$

Como se vio en la sección anterior, el valor de  $q_{\alpha}$  se lo encuentra en la tabla de Tukey, Apéndice E. Debe buscarse al nivel de significación del 1% o 5%, con los grados de libertad del Cuadrado Medio del Error (CME) y en la columna correspondiente al número de medias.

La estadística  $S_{\bar{d}}$  es el error estándar de la diferencia de dos medias, que se lo encuentra con la fórmula (103) vista anteriormente:  $S_{\bar{d}} = \sqrt{CME * \frac{r_m + r_n}{r_m * r_n}}$ . (Recuerde que el término Cuadrado Medio del Error (CME) es sinónimo de Varianza del Error.) Las variables  $r_m$  y  $r_n$  corresponden al número de repeticiones de cualquier par de medias en comparación. Veamos un ejemplo:

**Ejercicio 17.2.2.** Con los datos del ejercicio 14.8.1, compare las medias mediante la prueba de Tukey, al nivel alfa de 0.05.

El valor de  $q_{\alpha}$  cuando alfa es 0.05, con 30 grados de libertad del error y para siete medias, es igual a 4.46 (Apéndice E). El Cuadrado Medio del Error (CME) que se encontró en el análisis de varianza fue igual a 168.65. Los tratamientos, medias y número de repeticiones se presentan en el siguiente cuadro:

Cuadro 1. Tratamientos, medias y número de repeticiones del Ejercicio 14.8.1

Tratamientos	A	B	C	D	E	F	G
Repeticiones, $r_j$	5	6	5	5	4	6	6
Medias, $\bar{Y}_i$	44.80	62.66	41.40	53.60	33.75	41.50	55.16

Los errores estándar de la diferencia de dos medias van a variar de acuerdo con el número de repeticiones. Como ejemplo, para calcular el error estándar de la diferencia de dos medias cuando las repeticiones respectivas son 4 y 5, para las comparaciones entre los tratamientos E y A, E y C, E y D, se aplica la fórmula (97):

$$S_{\bar{d}} = \sqrt{CME * \frac{r_m + r_n}{r_m * r_n}} = \sqrt{168.65 * \frac{9}{20}} = 8.71$$

En el siguiente cuadro se presentan las combinaciones de las repeticiones (valores  $r_m$  y  $r_n$ ), los errores estándar de las diferencias y los valores críticos de Tukey:

Cuadro 2. Errores estándar de las diferencias entre dos medias y valores críticos de TUK', cuando el valor de  $q_\alpha$  al 0.05 y con 18 grados de libertad es igual a 4.46

Combinación de Repeticiones, $r_m$ y $r_n$	6 y 6	5 y 6	5 y 5	4 y 6	4 y 5
$S_{\bar{d}}$	7.50	7.86	8.21	8.38	8.71
$q_\alpha * S_{\bar{d}}$	33.44	35.06	36.62	37.37	38.85
$TUK' = \frac{q_\alpha * S_{\bar{d}}}{\sqrt{2}}$	23.65	24.79	25.89	26.43	27.47

Todas las diferencias entre pares de medias que tengan 4 y 5 repeticiones y que superen el valor crítico de 27.47, son estadísticamente diferentes al nivel alfa de 0.05. Como ejemplo, la diferencia entre las medias E y D es 19.85. Se concluye, entonces, que E y D no son estadísticamente diferentes. De igual manera se procede para evaluar las otras 20 comparaciones entre medias.

De hecho, la única diferencia significativa al nivel de probabilidad de 0.05, es la diferencia de 28.91 entre las medias E (con 4 repeticiones) y B (con 6 repeticiones). La diferencia supera al valor crítico de Tukey que es 26.43. Este resultado también demuestra que la prueba de Tukey es conservadora, como lo señala Hayter (1984)<sup>38</sup>

Los resultados con Statistix son:

Statistix 9.0      DCA germ chocho,

#### Tukey HSD All-Pairwise Comparisons Test

Variable	Mean	A	B	C	D	E	F
A	44.800						
B	62.667	17.867					
C	41.400	3.400	21.267				
D	53.600	8.800	9.067	12.200			
E	33.750	11.050	28.917*	7.650	19.850		
F	41.500	3.300	21.167	0.100	12.100	7.750	
G	55.167	10.367	7.500	13.767	1.567	21.417	13.667

Alpha 0.05      Standard Error for Comparison 7.4979 TO 8.7117

Critical Q Value 4.464      Critical Value for Comparison 23.668 TO 27.500

Puede observarse que la única diferencia significativa fue entre las medias B y E. Observe también, en los resultados que provee Statistix, lo siguiente:

38 Hayter, A.J. 1984. A proof of the conjecture that the Tukey-Kramer multiple comparisons procedure is conservative. J. American Statist. Ass., 12:61-75.

1. El valor crítico  $q$  en la tabla Tukey es 4.46

2. Los errores estándar para las comparaciones van de 7.50, cuando se comparan dos medias, ambas con 6 repeticiones; a 8.71, cuando se comparan dos medias con 4 y 5 repeticiones.

3. Los valores TUK' se obtienen al multiplicar el valor de  $q$  por los errores estándar, cantidad que se divide para la raíz cuadrada de dos.

### 3. La prueba de Scheffé (SCH), $r$ no iguales

En el caso de la prueba de Scheffé, la fórmula (98) presentada en la sección anterior, sirve también para estos casos. Sin embargo, debe calcularse más de un error estándar para compensar por el diferente número de repeticiones.

Veamos un ejemplo:

**Ejercicio 17.2.3.** Con los datos del Ejercicio 14.8.1 compare las medias mediante la prueba de Scheffé (SCH), al nivel alfa de 0.05.

La fórmula de la prueba de Scheffé es la siguiente:  $SCH = S_{\bar{a}} * \sqrt{gk * F_{\alpha}}$

Los errores estándar de la diferencia de dos medias se los encuentra con las fórmulas señaladas en ejercicios previos. El valor de los grados de libertad de los tratamientos  $gk$  es 6. El valor de  $F_{\alpha}$  al 0.05, con 6 grados de libertad de los tratamientos en el numerador y 30 grados de libertad del error en el denominador, es igual a 2.42 (tabla F en el Apéndice D).

El valor de  $\sqrt{gk * F_{\alpha}}$  es por lo tanto igual a  $\sqrt{6 * 2.42} = 3.81$

Los errores estándar de la diferencia de dos medias son iguales a los que se encontraron en el ejercicio previo. El cuadro siguiente resume los resultados:

**Cuadro 1. Repeticiones, errores estándar de las diferencias y el valor crítico de Scheffé**

Combinación de Repeticiones, $r_m$ y $r_n$	6 y 6	5 y 6	5 y 5	4 y 6	4 y 5
$S_{\bar{a}}$	7.50	7.86	8.21	8.38	8.71
$\sqrt{gk * F_{\alpha}}$	3.81	3.81	3.81	3.81	3.81
Valores críticos SCH	28.58	29.95	31.28	31.93	33.19

Si se comparan los valores SCH con las diferencias de las medias, se concluye que ninguna diferencia es significativa. Este resultado contradice lo encontrado en el Cuadro ADEVA, en el que se determinó que por lo menos una diferencia era estadísticamente significativa, al nivel alfa 0.05. Por ser excesivamente conservadora, no se recomienda el uso de esta prueba.

Contribuyó a este resultado un error experimental muy inflado ( $CME = 168.65$ ), que también determinó un coeficiente de variación muy alto ( $CV = 26.84\%$ ).

Los resultados con Statistix son:

Statistix 9.0      DCA germ chocho,

**Scheffe All-Pairwise Comparisons Test**

Variable	Mean	A	B	C	D	E	F
A	44.800						
B	62.667	17.867					
C	41.400	3.400	21.267				
D	53.600	8.800	9.067	12.200			
E	33.750	11.050	28.917	7.650	19.850		
F	41.500	3.300	21.167	0.100	12.100	7.750	
G	55.167	10.367	7.500	13.767	1.567	21.417	13.667

Alpha 0.05      Standard Error for Comparison 7.4979 TO 8.7117

Critical F Value 2.421      Critical Value for Comparison 28.574 TO 33.200

#### 4. La prueba de Bonferroni (BON), $r$ no iguales

En la prueba de Bonferroni, cuando se tienen medias con diferentes repeticiones, la fórmula para encontrar el valor crítico BON es la misma que se presentó en la sección anterior, esto es:

$BON = t_{\alpha'} * S_{\bar{a}}$ . El valor  $t_{\alpha'}$  es uno solo, sin embargo, igual que para las otras pruebas, el valor de  $S_{\bar{a}}$  variará de acuerdo con los números de repeticiones de las medias en comparación. Veamos un ejemplo:

**Ejercicio 17.2.4.** Con los mismos datos del Ejercicio 14.8.1 encuentre los valores críticos para la prueba de separación de medias de Bonferroni, al nivel alfa de 0.05.

Primero se debe encontrar el valor de  $\alpha'$  al nivel alfa de 0.05.



La fórmula es:  $\alpha' = \frac{\alpha}{2*w} = \frac{0.05}{2*21} = 0.00119$ , en la que  $w$  es el número de combinaciones de  $k$  medias tomadas de dos en dos. (Hubo siete tratamientos = siete medias, por lo que  $w = 21$ )

Como se indicó, no hay tablas para encontrar el valor de  $t_{\alpha'}$ , por lo que se recomienda emplear algún programa estadístico. Con Statistix el valor de  $t_{\alpha'}$ , cuando  $\alpha'$  es igual a 0.00119 y con 30 grados de libertad de la varianza del error es:  $t_{(p=0.00119;30gl,error)} = 3.32$

Los errores estándar de las diferencias se determinaron con la fórmula (97) en los ejercicios anteriores. Con la información disponible se prepara el Cuadro 1.

En el Cuadro 1 se observa que los valores críticos son algo más altos que los de la prueba de Tukey. Es decir que, en términos generales, la prueba de Bonferroni es más conservadora que la de Tukey. Por esta razón y por la mayor dificultad para encontrar los valores  $t_{\alpha'}$ , se recomienda, de entre las dos, escoger la prueba de Tukey.

Cuadro 1. Combinación de repeticiones, valor de  $t_{\alpha'}$ , errores estándar de las diferencias y los valores críticos de Bonferroni

Combinación de Repeticiones, $r_m$ y $r_n$	6 y 6	5 y 6	5 y 5	4 y 6	4 y 5
$t_{\alpha'}$	3.32	3.32	3.32	3.32	3.32
$S_{\bar{d}}$	7.50	7.86	8.21	8.38	8.71
$BON = t_{\alpha'} * S_{\bar{d}}$	24.90	26.10	27.26	27.82	28.92

En los resultados que presenta Statistix, la única diferencia significativa al nivel alfa de 0.05 es la de la comparación entre las medias B y E. Igual resultado se obtuvo con la prueba de Tukey.

Statistix 9.0                      DCA germ chocho,

**Bonferroni All-Pairwise Comparisons Test**

Variable	Mean	A	B	C	D	E	F
A	44.800						
B	62.667	17.867					
C	41.400	3.400	21.267				
D	53.600	8.800	9.067	12.200			

E	33.750	11.050	28.917*	7.650	19.850		
F	41.500	3.300	21.167	0.100	12.100	7.750	
G	55.167	10.367	7.500	13.767	1.567	21.417	13.667

Alpha 0.05 Standard Error for Comparison 7.4979 TO 8.7117

Critical T Value 3.319 Critical Value for Comparison 24.882 TO 28.911

### 5. La prueba de amplitud múltiple de Duncan (AMD), $r$ no iguales

La fórmula (101) que se utilizó para calcular los valores críticos de la prueba AMD, se modifica para tomar en cuenta el diferente número de repeticiones de las medias. Esto es:

$$AMD = \frac{q_{\alpha} * S_{\bar{d}}}{\sqrt{2}} \quad (105)$$

Esta modificación fue propuesta por Kramer<sup>39</sup> (1956), sin embargo, de acuerdo con Steel et al<sup>40</sup> (1997), la validez estadística de este procedimiento no ha sido verificada.

Los valores  $q_{\alpha}$  se los encuentra en la tabla de Duncan (Apéndice F) en las columnas de los números de medias, al nivel alfa del 0.05 o 0.01 y con los grados de libertad del Cuadrado Medio del Error. Los errores estándar de las diferencias de dos medias  $S_{\bar{d}}$  se calculan con la fórmula (103), tomando en cuenta el número de repeticiones que originó cada media.

Veamos un ejemplo:

---

**Ejercicio 17.2.5.** Con los datos del Ejercicio 14.8.1 realice la separación de medias, al nivel alfa de 0.05, mediante el procedimiento de la amplitud múltiple de Duncan, para diferente número de repeticiones.

En el Ejercicio 14.8.1 se compararon siete tratamientos, con diferente número de repeticiones. El Cuadrado Medio del Error (CME) fue 168.65 y tuvo 30 grados de libertad.

39 Kramer, C.Y. 1956. Extension of multiple range tests to group means with unequal numbers of replication. *Biometrika* 12:307-310.

40 Steel, R.G.D., J.H. Torrie and D.A. Dickey. 1997. *Principles and Procedures of Statistics, a Biometrical Approach*. Third Edition. McGraw-Hill, Inc. New York. p.199

En la tabla de Duncan se encuentran los valores de  $q_{\alpha'}$  al nivel de 0.05, con 30 grados de libertad y en las columnas del 2 al 7, que corresponden a los números de medias de la amplitud. Los valores  $q_{\alpha}$  se presentan en el Cuadro 1:

Cuadro 1. Valores de  $q_{\alpha'}$  al nivel de 0.05 con 30 grados de libertad, para medias de la amplitud de la 2 a la 7. Tabla de Duncan

Número de medias en comparación, $p$	$p = 2$	$p = 3$	$p = 4$	$p = 5$	$p = 6$	$p = 7$
Valores $q_{\alpha'}$	2.89	3.04	3.12	3.20	3.25	3.29

En el siguiente cuadro se presentan las medias ordenadas ascendentemente y el número de repeticiones de cada media:

Cuadro 2. Medias ordenadas ascendentemente y número de repeticiones

Tratamientos	E	C	F	A	D	G	B
Repeticiones, $r$	4	5	6	5	5	6	6
Medias, $\bar{Y}_i$	33.75	41.40	41.50	44.80	53.60	55.16	62.66

Cuando se realiza comparación de medias con diferente número de repeticiones debe calcularse el error estándar de la diferencia de medias para cada caso. Como ejemplo, si se trata de comparar la media de E con la media de C, los valores de  $m$  y  $n$  en la fórmula (97) serán 4 y 5 (ver Cuadro 2). En este ejercicio hay cinco posibles combinaciones diferentes de  $m$  y  $n$ . Por lo tanto, habrá cinco errores estándar de la diferencia de medias. Las combinaciones  $m$  y  $n$  y los errores estándar se presentan en el siguiente cuadro:

Cuadro 3. Combinaciones de repeticiones y errores estándar de la diferencia de medias

Combinación de Repeticiones, $r_m$ y $r_n$	6 y 6	5 y 6	5 y 5	4 y 6	4 y 5
$S_{\bar{a}}$	7.50	7.86	8.21	8.38	8.71

Con la información contenida en los cuadros anteriores, se puede proceder a comparar las medias incluidas en las diferentes amplitudes:

1. Es preferible comparar primero las medias de los extremos, es decir la E y la B (con 4 y 6 repeticiones, respectivamente). La diferencia de las dos medias es 28.91. Note que esta comparación involucra a las siete medias. Este es el

concepto de amplitud en la prueba. Por esta razón, el valor  $q$  alfa es 3.29 (Cuadro 1). El error estándar de la diferencia es 8.38. Consecuentemente el valor crítico de la prueba AMD es:

$AMD = \frac{q_{\alpha} * S_{\bar{d}}}{\sqrt{2}} = \frac{3.29 * 8.38}{1.4142} = 19.50$ . Como la diferencia entre las medias supera al valor crítico AMD, se declaran diferentes las dos medias, E y B, al nivel alfa de 0.005.

2. La siguiente comparación es entre la E y la G, con 4 y 6 repeticiones. La diferencia es 21.41. La comparación involucra a seis medias, por lo que el valor de  $q$  alfa es 3.25. El valor crítico es:

$AMD = \frac{3.25 * 8.38}{1.4142} = 19.26$ . La diferencia de las medias, E y G, es significativa.

3. Se continúa con la comparación entre E y D, con 4 y 5 repeticiones. La diferencia es 19.85. Como la comparación involucra a 5 medias, el valor de  $q$  alfa es 3.20. El error estándar, cuando  $m$  igual a 4 y  $n$  igual a 5 es 8.71. El valor crítico consecuentemente es:

$AMD = \frac{3.20 * 8.71}{1.4142} = 19.70$ . Como la diferencia de 19.85 supera al valor crítico de 19.70, las medias E y D son diferentes estadísticamente al nivel alfa del 0.05.

4. La siguiente comparación es entre E y A. La diferencia es 11.50. El valor crítico es:

$AMD = \frac{3.12 * 8.71}{1.4142} = 19.22$ . La diferencia entre E y A no es significativa. Esto determina que el grupo de las cuatro medias, de la E a la A, compartan una misma letra (ver Cuadro 4), lo que representa la igualdad estadística de ese grupo. La media E ya no interviene en ninguna otra comparación.

5. La comparación es ahora entre la media que sigue, la media C, con la más distante, la media B. La diferencia es 21.26. El valor crítico es:

$AMD = \frac{3.25 * 7.86}{1.4142} = 18.06$ . La diferencia entre C y B es significativa.

6. La comparación que sigue es entre C y G. La diferencia es 13.76. El valor crítico es:

$AMD = \frac{3.20 * 7.86}{1.4142} = 17.78$ . La diferencia no es significativa. Esto indica que todo el grupo de medias, de la C a la G compartan una misma letra. La media C ya no interviene en ninguna otra comparación.

7. Se comparan ahora las medias F y B. La diferencia es 21.16. El valor crítico es:

$$AMD = \frac{3.20 \cdot 7.50}{1.4142} = 16.97. \text{ Las medias F y B son diferentes.}$$

8. Sigue la comparación entre las medias F y G. La diferencia es 13.66. El valor crítico es:

$$AMD = \frac{3.12 \cdot 7.50}{1.4142} = 16.55. \text{ La diferencia entre F y G no es significativa. Todo el grupo, de la F a la G, deben llevar una misma letra. La media F ya no interviene en ninguna otra comparación.}$$

9. Se compara ahora la media A con la media B. La diferencia es 17.86. El valor crítico para la comparación es:  $AMD = \frac{3.12 \cdot 7.86}{1.4142} = 17.34$ . La diferencia entre A y B es significativa.

10. La siguiente comparación es entre A y G. La diferencia es 10.36. El valor crítico es:

$$AMD = \frac{3.04 \cdot 7.86}{1.4142} = 16.90. \text{ La diferencia entre A y G no es significativa. Todo el grupo, de la A a la G, deben llevar una misma letra. (Ver el Cuadro 4.) La media A ya no interviene en ninguna otra comparación.}$$

11. La siguiente comparación es entre D y B. La diferencia es 9.06. El valor crítico para comparar esta amplitud de tres medias es:

$$AMD = \frac{3.04 \cdot 7.86}{1.4142} = 16.90. \text{ La diferencia no es significativa. El grupo de tres medias de la D a la B deben llevar una misma letra, que indica que no hay diferencias significativas entre ellas.}$$

Estas son todas las comparaciones que se pueden hacer con este juego de datos, empleando la prueba AMD.

Para resumir la información es conveniente marcar con una misma letra los grupos de medias que resultaron no significativos. Empecemos con la primera comparación que fue la número 4. Las medias de la E a la A se marcan con la letra a. El siguiente grupo no significativo fue el de la comparación número 6, de las medias de la C a la G. Este grupo se marca con la letra b. Luego sigue la comparación número 8, de la F a la G, que se marca con la letra c. La comparación número 10, del grupo de la A a la G, se marca con la letra d. Finalmente, la última comparación no significativa fue la número 11 del grupo de la D a la B, que se marca con la letra e. Los resultados se resumen en el siguiente cuadro:

**Cuadro 4. Prueba de la Amplitud Múltiple de Duncan (AMD) al nivel alfa de 0.05. Evaluación de porcentajes de germinación de siete líneas de chocho**

Líneas de chocho	E	C	F	A	D	G	B
Repeticiones, $r$	4	5	6	5	5	6	6
Medias, $\bar{Y}_i$	33.75	41.40	41.50	44.80	53.60	55.16	62.66
Significación estadística, 5%	a	ab	abc	abcd	bcde	bcde	de

Nota: Letras iguales minúsculas, bajo las medias denotan igualdad estadística al 0.05. Letras distintas señalan diferencias.

Cuando las medias comparten letras iguales y distintas a la vez, quiere decir que las diferencias o igualdades estadísticas no están claramente definidas. En este ejercicio, resalta la diferencia entre E y las medias D, G, y B. Se puede señalar también que las medias C y F son estadísticamente diferentes de la media B, al nivel alfa de 0.05.

**Comentario:** como la prueba de Duncan se refiere a amplitudes, no es correcto presentar los resultados de comparaciones entre medias a través de una matriz de diferencias, como se puede hacer con otras pruebas. Lo correcto es representar las amplitudes mediante letras o también mediante segmentos de líneas rectas que unan las medias estadísticamente iguales.

## 6. La prueba de Student-Newman-Keuls (SNK), $r$ no iguales

Como se indicó en la sección anterior, la prueba de Student-Newman-Keuls es una prueba de separación de medias de amplitud múltiple, similar a la prueba de Duncan. Es decir, se pueden comparar simultáneamente varias medias.

Cuando el número de repeticiones no es el mismo para todas las medias, es necesario calcular errores estándar de la diferencia de dos medias para cada caso. Consecuentemente, la fórmula (102) debe modificarse de acuerdo con esto. A semejanza de la prueba de Duncan, en esta prueba habrá tantos valores críticos como errores estándar de la diferencia. La fórmula para encontrar los valores críticos es la siguiente:

$$SNK = \frac{Q_{\alpha^*} S_{\bar{d}}}{\sqrt{2}} \quad (106)$$

Los valores  $Q_{\alpha^*}$  se encuentran en la tabla de amplitudes estudentizadas del Apéndice E. El nivel alfa de significación de la prueba puede ser de 0.05 o de 0.01,

según decida el investigador, con los grados de libertad del Cuadrado Medio del Error (CME) del Cuadro ADEVA. Los valores  $Q_{\alpha'}$  constan en las columnas marcadas con la letra p, que indica el número de medias de tratamientos.

En todo caso, es importante señalar que la validez estadística de la prueba SNK, cuando se tiene diferente número de repeticiones entre medias, no ha sido estadísticamente verificada (Steel, et al. 1997). Una acotación similar se hizo para el caso de la prueba AMD.

A pesar de lo anterior, veamos un ejemplo:

**Ejercicio 17.2.6.** Con los datos del Ejercicio 14.8.1 realice la separación de medias, al nivel alfa de 0.05, mediante la prueba de Student-Newman-Keuls (SNK) para diferente número de repeticiones.

En el Ejercicio 14.8.1 se compararon siete tratamientos, con diferente número de repeticiones. El Cuadrado Medio del Error (CME) fue 168.65 y tuvo 30 grados de libertad.

En la tabla del Apéndice E (la misma que se emplea para la prueba de Tukey) constan los valores de  $Q_{\alpha'}$  al nivel de 0.05, con 30 grados de libertad y en las columnas del 2 al 7, que corresponden a los números de medias de tratamientos. Los valores  $Q_{\alpha'}$  se presentan en el Cuadro 1:

Cuadro 1. Valores de  $Q_{\alpha'}$  al nivel de 0.05 con 30 grados de libertad, para comparar amplitudes de dos a siete medias

Número de medias en comparación, p	p = 2	p = 3	p = 4	p = 5	p = 6	p = 7
Valores $Q_{\alpha'}$	2.89	3.49	3.84	4.10	4.30	4.46

En el Cuadro 2, a continuación, se presentan las medias ordenadas ascendentemente y el número de repeticiones que dio origen a cada media. (Este cuadro es igual al Cuadro 2 del ejemplo anterior, así como el Cuadro 3, pero se los presenta por facilidad de los cálculos.)

Cuadro 2. Medias ordenadas ascendentemente y numero de repeticiones

Tratamientos	E	C	F	A	D	G	B
Repeticiones, $r$	4	5	6	5	5	6	6
Medias, $\bar{Y}_i$	33.75	41.40	41.50	44.80	53.60	55.16	62.66

Los errores estándar de las diferencias  $S_{\bar{d}}$  se encuentran con la fórmula (103). Como ejemplo, para encontrar el error de la diferencia entre las medias de los tratamientos E y B, con 4 y 6 repeticiones respectivamente (los valores m y n), se tiene:

$$S_{\bar{d}} = \sqrt{CME * \frac{r_m + r_n}{r_m * r_n}} = \sqrt{168.65 * \frac{10}{24}} = 8.38$$

En el análisis de estos resultados hay cinco posibles combinaciones diferentes de m y n, los números de repeticiones. Por lo tanto, habrá cinco errores estándar de la diferencia de medias, como se presenta en el siguiente cuadro:

**Cuadro 3. Combinaciones de repeticiones y errores estándar de la diferencia de medias**

Combinación de Repeticiones, $r_m$ y $r_n$	6 y 6	5 y 6	5 y 5	4 y 6	4 y 5
$S_{\bar{d}}$	7.50	7.86	8.21	8.38	8.71

Con la información anterior, se procede a efectuar la separación de las medias:

1. Para comparar la media E y la media B, las dos de los extremos, con 4 y 6 repeticiones, respectivamente, se debe tener en cuenta que son siete medias en la amplitud. Por lo tanto, el valor  $Q_{\alpha'}$  es igual a 4.46 (Cuadro 1). El valor  $S_{\bar{d}}$  para esta comparación es igual a 8.38. El valor crítico SNK es entonces:

$$SNK = \frac{Q_{\alpha'} * S_{\bar{d}}}{\sqrt{2}} = \frac{4.46 * 8.38}{1.4142} = 26.43$$

La diferencia absoluta entre la media E y B es igual a 28.91. Si la diferencia es igual o mayor que el valor crítico, es decir:  $|\bar{Y}_i - \bar{Y}_j| \geq SNK$ , se declara estadísticamente diferentes a las medias, al nivel de 0.05 de probabilidad. En este caso:  $|28.91| \geq 26.43$ , las dos medias son diferentes.

2. La siguiente comparación es entre las medias E y G, también con 4 y 6 repeticiones, respectivamente. La diferencia absoluta entre las dos medias es igual a 21.41. El valor crítico es igual a:

$$SNK = \frac{Q_{\alpha'} * S_{\bar{d}}}{\sqrt{2}} = \frac{4.30 * 8.38}{1.4142} = 25.48$$



La diferencia entre las dos medias es menor que el valor crítico SNK ( $|21.41| < 25.48$ ), por lo tanto, todo el grupo de seis medias, de la E a la G, son estadísticamente iguales al nivel de probabilidad alfa 0.05.

Los resultados se resumen en el siguiente cuadro:

**Cuadro 4. Separación de medias, prueba SNK al nivel alfa de 0.05. Evaluación de porcentajes de germinación de siete líneas de chocho**

Líneas de chocho	E	C	F	A	D	G	B
Repeticiones, $r$	4	5	6	5	5	6	6
Medias, $\bar{Y}_i$	33.75	41.40	41.50	44.80	53.60	55.16	62.66
Significación estadística	a	ab	ab	ab	ab	ab	b

\*Letras minúsculas iguales denotan igualdad estadística al 0.05. Letras distintas señalan diferencias.

**Comentario:** puede notarse que la prueba SNK tiene mayor poder de discriminación que la prueba de la Amplitud Múltiple de Duncan (AMD). Sin embargo, la probabilidad de cometer Error Tipo II (aceptar una hipótesis nula falsa) es mayor con la prueba SNK que con la prueba AMD.

### 17.3. Comparaciones con la media testigo

Cuando en un experimento se incluye un testigo, pueden compararse todas las medias de los tratamientos en forma individual con la media del tratamiento testigo. En este caso, las pruebas recomendadas son la DMS y la prueba de Dunnett (DNT). El empleo de estas pruebas, para su validez, requiere que hayan sido planificadas de antemano. Por lo tanto, es correcto llevarlas a cabo incluso si el valor de F no fuera significativo.

#### 1. La Diferencia Mínima Significativa (DMS)

El empleo de la prueba DMS es adecuado cuando se trata de comparar varios tratamientos, de uno a uno, con el tratamiento testigo. Si el número de repeticiones es igual para todos los tratamientos debe usarse la fórmula (96). De lo contrario, la fórmula de la DMS debe incluir el error estándar de la diferencia de dos medias con desigual número de repeticiones (fórmula 103). Veamos un ejemplo del primer caso:

**Ejercicio 17.3.1.** En el Ejercicio 16.4.1, se presentó un experimento en el que una empresa avícola comparó su propio alimento (dieta testigo T), con tres alimentos de la competencia (dietas A, B, C). La variable de respuesta fue el peso de los huevos, en gramos. Las medias de los datos combinados de dos cuadrados latinos fueron: A = 55.13, B = 58.38, C = 44.88, T = 50.99. El número de repeticiones de cada media fue ocho. El CME fue igual a 7.43, con 18 grados de libertad. Compare una a una todas las medias con la media testigo, mediante la prueba de la Diferencia Mínima Significativa (DMS) al nivel alfa de 0.05.

Primero debe encontrarse el valor de  $t$  al nivel alfa de 0.05 y con 18 grados de libertad. Este valor en la tabla respectiva es igual a 2.101. La DMS es:

$$DMS = t_{\frac{\alpha}{2}}^* \sqrt{\frac{2 * CME}{r}} = 2.101 * \sqrt{\frac{2 * 7.43}{8}} = 2.86 \text{ gramos}$$

El resultado se presenta en el siguiente cuadro:

Cuadro 1. Diferencias, en gramos, entre las dietas A, B, C vs la dieta T.

Dieta A	Dieta B	Dieta C	Dieta T
4.14	7.39	-6.11	0
*	*	*	Significación

Todas las diferencias sobrepasaron el valor absoluto de la DMS. En conclusión, la dieta T, al nivel alfa de 0.05, fue estadísticamente superior a la dieta C, pero inferior a las dietas A y B.

## 2. La prueba de Dunnett (DNT)

La prueba de Dunnett (DNT) tiene, a semejanza de la prueba de la DMS, un solo valor. Pero, en forma parecida a la prueba de Tukey, toma en cuenta el número de medias del experimento.

La prueba DNT tiene la ventaja de que se pueden calcular intervalos de confianza para las diferencias y pueden ser para un solo extremo o los dos extremos (límites inferior y superior). De hecho, hay tablas para cada caso.

La fórmula para encontrar el valor crítico es:

$$DNT = t_{\alpha}(Dunnett)^* \sqrt{\frac{2 * CME}{r}} \quad (107)$$

El valor de  $t_{\alpha}$  (*Dunnett*) se lo encuentra en tablas en varios textos, pero también en internet<sup>41</sup> (igual que las otras tablas). Para su correcto uso, debe tomarse en cuenta el valor alfa, el número de grados de libertad del error y el número de medias, valor que encabeza las columnas de la tabla. Este número, en la tabla de la referencia, incluye también al testigo.

La fórmula para el Intervalo de Confianza de la diferencia entre cualquiera de las medias y la media del testigo es:

$$IC = (\bar{Y}_i - \bar{Y}_T) \pm t_{\alpha} * \sqrt{\frac{2 * CME}{r}} \quad (108)$$

La fórmula anterior, con el signo  $\pm$ , indica que se obtienen los dos límites, el inferior y el superior. Por lo tanto, el valor  $t$  de Dunnett debe encontrarse en la tabla específica para los dos límites. Si se requiere sólo el límite inferior o el superior, la fórmula (108) llevará sólo el signo menos o más. En este caso, el valor  $t$  de Dunnett también será diferente. Veamos un ejemplo:

**Ejercicio 17.3.2.** Con la misma información empleada en el Ejercicio 17.3.1 compare individualmente las medias de las dietas A, B y C, con la media de la dieta T (testigo). Use la prueba de Dunnett, para los dos límites, al nivel alfa de 0.05. Presente, además, los intervalos de confianza:

Primero se debe encontrar el valor  $t_{\alpha}$  (*Dunnett*) en la tabla para los dos límites, al nivel alfa de 0.05 y con 18 grados de libertad. Las columnas en la tabla corresponden al número de medias para la comparación. En este caso, el número de medias, incluyendo la media del testigo, es igual a cuatro. El valor es igual a 2.56.

El valor crítico para las comparaciones es:  $DNT = 2.56 * \sqrt{\frac{2 * 7.43}{8}} = 3.49$

Las diferencias entre las medias en comparación con la media del testigo se presentan en el siguiente cuadro:

Cuadro 1. Diferencias, en gramos, entre las dietas A, B, C y la dieta T.

Dieta A	Dieta B	Dieta C	Dieta T
4.14	7.39	-6.11	0
*	*	*	Significación

41 <http://www.stat.ufl.edu/~winner/tables/dunnett-2side.pdf>

Todas las diferencias sobrepasaron el valor absoluto de DNT. En conclusión, la dieta T, al nivel alfa de 0.05, fue estadísticamente superior a la dieta C, pero inferior a las dietas A y B.

Los intervalos de confianza al 95% (alfa igual a 0.05) para las diferencias entre las medias y el testigo, se los encuentra con la fórmula (108). Estos son:

$$IC.(AvsT) = 4.14 \pm 3.49 = (+0.65).y.(+7.63)$$

$$IC.(BvsT) = 7.39 \pm 3.49 = (+3.90).y.(+10.88)$$

$$IC.(CvsT) = -6.11 \pm 3.49 = (-9.60).y.(-2.62)$$

Con el programa Statistix los resultados son:

Statistix 9.0 dos Latin sq huevos,

**Two-sided Dunnett's Multiple Comparisons with a Control of PESOHUEVO**

Control: DIETAS = T

Simultaneous 95% confidence intervals of treatment mean - control mean

DIETASlit	Mean	Lower Bound	Difference	Upper Bound
A	55.125	0.643	4.138*	7.632
B	58.375	3.893	7.388*	10.882
C	44.875	-9.607	-6.112*	-2.618
T	50.987			

Alpha 0.05 Standard Error for Comparison 1.3633

Critical D Value 2.563 Critical Value for Comparison 3.4947

Error term used: CUADRADO\*JAULAS\*PERIODO\*DIETASlit, 18 DF

## Comentarios:

- La prueba de la Diferencia Mínima Significativa es más liberal que la prueba de Dunnett. Esta última, sin embargo, ofrece más protección contra la posibilidad de cometer Error Tipo I.
- Las dos pruebas son válidas si fueron planificadas de antemano. En otras palabras, no es legítimo su uso cuando se decide emplearlas una vez que se conocen los resultados. Tampoco es correcto emplear más de una prueba para un mismo grupo de medias. El investigador, de acuerdo con su experiencia y tipo de experimento, resolverá cuál prueba emplear.

- En el caso de que el número de repeticiones no sea igual para las medias, deberá emplearse el error estándar de la fórmula (103) con cualquiera de los dos métodos.
- Observe que en ninguno de los intervalos calculados con la fórmula de Dunnett está incluido el cero. Cuando el intervalo incluye al cero, las diferencias no son estadísticamente significativas al nivel que se haya planteado.

## 17.4. Comparaciones ortogonales entre grupos de medias

### *Cuando el número de repeticiones es igual*

Si en la selección de tratamientos cualitativos se incluyen algunos que tengan características comunes y a la vez sean diferentes de otros, se puede formar varios grupos de tratamientos. Todas las comparaciones son entre dos grupos de tratamientos y se denominan *contrastos lineales*. Sin embargo, es preferible para evitar sesgos, que los contrastes lineales sean también *contrastos ortogonales*. Este es el mejor método de separación de medias cuando los tratamientos son cualitativos y debe preferirse sobre los otros, siempre que sea posible.

Los contrastes o comparaciones ortogonales son pruebas  $F$  con un solo grado de libertad (o pruebas  $t$ , por la relación  $F = t^2$ ). Por ortogonalidad se entiende que las comparaciones son independientes entre ellas. En otras palabras, con este método, la variación entre tratamientos se descompone en cantidades que contribuyen por separado a la Suma de Cuadrados de Tratamientos ( $SCTrat$ ). Pueden hacerse tantos contrastes como haya grados de libertad para tratamientos. Este método no requiere de una  $F$  significativa, porque se supone que las comparaciones fueron planificadas de antemano. Esta característica determina que la probabilidad de cometer Error Tipo I o Tipo II se mantenga constante.

No hay reglas fijas para organizar un juego de comparaciones ortogonales. Con un mismo grupo de tratamientos cualitativos pueden generarse varios juegos de comparaciones ortogonales diferentes. Sin embargo, en la planificación del experimento y en especial, en la selección de los tratamientos, el investigador deberá identificar de antemano el juego de comparaciones ortogonales que mejor responda a sus objetivos.

### *Linealidad*

Una comparación o contraste lineal da como resultado la diferencia entre los totales de dos grupos de tratamientos. El procedimiento puede también

realizarse con medias, pero es más fácil con los totales de los tratamientos T. Si llamamos Q a un contraste lineal, entonces:

$$Q_i = \sum c_i T_i - \sum c_j T_j \quad (109)$$

En la comparación, cada total de tratamiento lleva un coeficiente  $c_i$ . Sin embargo, la condición para que el contraste sea lineal y la comparación válida es que:

$$\sum c_i = 0$$

El término  $c_i$  representa los coeficientes que van delante de los totales de los tratamientos. Si un tratamiento no interviene en la comparación, el coeficiente de ese tratamiento es cero.

La suma de cuadrados de una comparación o contraste lineal ( $SCQ_i$ ) es parte de la Suma de Cuadrados de los Tratamientos ( $SCTrat$ ). La fórmula de un contraste lineal es:

$$SCQ_i = \frac{Q_i^2}{D_i} \quad (110)$$

El divisor  $D_i$  es igual a:

$$D_i = r \sum c_i^2 \quad (111), \text{ en donde } r \text{ es el número de repeticiones.}$$

### Ortogonalidad

Dos comparaciones son ortogonales si la suma de los productos de los respectivos coeficientes es igual a cero:  $\sum c_{1i} c_{2i}$

Si todas las comparaciones posibles son ortogonales entre sí, entonces:

$$\sum SCQ_i = SCTrat \quad (112)$$

Veamos un ejemplo:

**Ejercicio 17.4.1.** En el Ejercicio 15.5.1 se evaluó el volumen de expansión de diferentes cruza de maíz reventón: A1 y A2, americanos introducidos de Nebraska e Indiana; Z1, Z2 y Z3 provenientes del banco de germoplasma, los dos primeros de origen ecuatoriano y el tercero de origen colombiano. El testigo T<sub>e</sub> fue un maíz reventón comercial. El diseño fue un BCA con cinco repeticiones. La Suma de

Cuadrados de los Tratamientos resultó igual a 307.6, con cinco grados de libertad. La Suma de Cuadrados del Error fue igual a 118.1, con 20 grados de libertad. Los totales de los tratamientos y las medias se presentan en el siguiente cuadro:

**Cuadro 1. Totales y medias del volumen de expansión de cinco cruza de maíz reventón y un testigo comercial**

	Tratamientos					
	Te	A1	A2	Z1	Z2	Z3
$T_i$	103.5	84.0	77.0	59.5	70.5	55.5
$\bar{Y}_i$	20.7	16.8	15.4	11.9	14.1	11.1

Identifique los grupos de tratamientos que sean estadísticamente iguales o diferentes.

Respuesta: Este es un ejemplo en el que no es adecuado el empleo de ninguno de los métodos de separación de medias analizados anteriormente. Por la estructura del conjunto de tratamientos, el método más idóneo para este caso es el de comparaciones ortogonales. Como hubo cinco grados de libertad para los tratamientos, son cinco las comparaciones que pueden realizarse. Las cinco que parecen más obvias se presentan en el Cuadro 2. Es importante tomar en cuenta que la suma de los coeficientes debe ser igual a cero.

**Cuadro 2. Comparaciones ortogonales, valores del contraste lineal  $Q_i$  y divisor  $D_i$**

Tratamientos	Te	A1	A2	Z1	Z2	Z3			
Totales $T_i$	103.5	84.0	77.0	59.5	70.5	55.5			
Contrastes	Coeficientes						$Q_i$	$D_i$	$SCQ_i$
	$c_1$	$c_2$	$c_3$	$c_4$	$c_5$	$c_6$			
Te - (A1+A2+Z1+Z2+Z3)	-5	+1	+1	+1	+1	+1	-171.0	150	194.9
(A1+A2) - (Z1+Z2+Z3)	0	-3	-3	+2	+2	+2	-112.0	150	83.6
A1 - A2	0	-1	+1	0	0	0	-7.0	10	4.9
(Z1+Z2) - Z3	0	0	0	-1	-1	+2	-19.0	30	12.0
Z1 - Z2	0	0	0	-1	+1	0	-11.0	10	12.1

Note en el cuadro anterior que la suma de los coeficientes, de  $c_1$  a  $c_6$ , en todas las comparaciones, es igual a cero.

La prueba de la ortogonalidad se realiza mediante la fórmula  $\sum c_{1i}c_{2i}$ . Puede demostrarse que, para cualquier par de contrastes, la suma de los productos de los

coeficientes es igual a 0. Tomemos como ejemplo los dos primeros contrastes. Al multiplicar en orden, la una fila de coeficientes con la otra fila se tiene lo siguiente:

$$\sum c_1 c_2 = -5*(0) + 1*(-3) + 1*(-3) + 1*(2) + 1*(2) + 1*(2) = 0$$

De igual manera, para cualquier otro par de contrastes del Cuadro 2, la suma de los productos de los coeficientes es igual a cero.

Aunque no importa el orden de las comparaciones, generalmente cuando hay un tratamiento testigo, la primera comparación es la del testigo versus los demás tratamientos. En este caso, en el Cuadro 2, el valor  $Q_i$  para el primer contraste es, según la fórmula (110):

$$Q_i = \sum c_i T_i - \sum c_j T_j = (-5*103.5) + (84.0 + 77.0 + 59.5 + 70.5 + 55.5) = -171.0$$

Con el valor del contraste  $Q_i$  se puede determinar la Suma de Cuadrados del mismo contraste, pero primero hace falta encontrar el divisor:

$$D_i = r \sum c_i^2 = 5*(-5^2 + 1^2 + 1^2 + 1^2 + 1^2 + 1^2) = 150$$

La Suma de Cuadrados del Contraste lineal es:  $SCQ_i = \frac{Q_i^2}{D_i} = \frac{-171.0^2}{150} = 194.9$

De la misma forma se procede con las otras comparaciones. Note que la suma de todos los valores  $SCQ_i$  es igual a 307.6 (salvo error por redondeo). Esta cantidad es la Suma de Cuadrados de los Tratamientos ( $SCTrat$ ) del Ejercicio 15.1. Es decir que:

$$\sum SCQ_i = SCTrat = 194.9 + 83.6 + 4.9 + 12.0 + 12.1 = 307.6$$

Esto demuestra la ortogonalidad de los contrastes y su aporte independiente a la  $SCTrat$ . Como cada contraste tiene un grado de libertad, las sumas de cuadrados de los contrastes son iguales a los cuadrados medios.

Con la información anterior se puede presentar un Cuadro ADEVA completo, en el que se integran los resultados de este ejercicio con los del ejercicio 15.5.1.



Cuadro 3. Análisis de Varianza del experimento sobre volumen de expansión de maíz reventón

Fuente de Variación	Suma de Cuadrados	Grados de Libertad	Cuadrado Medio	F	F <sub>0.05</sub>	F <sub>0.01</sub>
Bloques	117.1	4	29.27	4.96**	2.87	4.43
Tratamientos	307.6	5	61.52	10.43**	2.71	4.10
Te vs otros	194.9	1	194.9	33.03**	4.35	8.10
A <sub>1</sub> A <sub>2</sub> vs Z <sub>1</sub> Z <sub>2</sub> Z <sub>3</sub>	83.6	1	83.6	14.17**		
A <sub>1</sub> vs A <sub>2</sub>	4.9	1	4.9	0.83 n.s		
Z <sub>1</sub> Z <sub>2</sub> vs Z <sub>3</sub>	12.0	1	12.0	2.03 n.s		
Z <sub>1</sub> vs Z <sub>2</sub>	12.1	1	12.1	2.05 n.s		
Error	118.1	20	5.90			
Total	542.8	29				

Como puede verse (por el signo de  $Q_i$  en el Cuadro 2), el testigo fue superior a los demás tratamientos en volumen de expansión. Las cruza A<sub>1</sub> y A<sub>2</sub> superaron a las tres cruza Z. Las diferencias fueron estadísticamente significativas al nivel alfa 0.01. Las otras comparaciones entre cruza no fueron significativas. (Para estas conclusiones deben tomarse en cuenta los signos de los contrastes, o lo que es igual, los totales de los tratamientos.)

Los resultados con Statistix son los siguientes:

Statistix 9.0 maiz canguil BCA

**AOV Contrasts of VOLUMEN by TRATAMIEN**

Contrast Number 1

Contrast Coefficients: -5 1 1 1 1 1

<b>Contrast</b>	<b>-34.200</b>	<b>SS (Contrast)</b>	<b>194.94</b>
Scheffe's F	6.60	P (Scheffe's F)	0.0009
T-Statistic	-5.75	P (T-Statistic)	0.0000
SE (Contrast)	5.9529		

Contrast Number 2

Contrast Coefficients: 0 -3 -3 2 2 2

<b>Contrast</b>	<b>-22.400</b>	<b>SS (Contrast)</b>	<b>83.627</b>
Scheffe's F	2.83	P (Scheffe's F)	0.0431
T-Statistic	-3.76	P (T-Statistic)	0.0012
SE (Contrast)	5.9529		

Contrast Number 3

Contrast Coefficients: 0 -1 1 0 0 0

<b>Contrast</b>	<b>-1.4000</b>	<b>SS (Contrast)</b>	<b>4.9000</b>
Scheffe's F	0.17	P (Scheffe's F)	0.9722
T-Statistic	-0.91	P (T-Statistic)	0.3732
SE (Contrast)	1.5370		

Contrast Number 4

Contrast Coefficients: 0 0 0 -1 -1 2

<b>Contrast</b>	<b>-3.8000</b>	<b>SS (Contrast)</b>	<b>12.033</b>
Scheffe's F	0.41	P (Scheffe's F)	0.8379
T-Statistic	-1.43	P (T-Statistic)	0.1689
SE (Contrast)	2.6622		

Contrast Number 5

Contrast Coefficients: 0 0 0 -1 1 0

<b>Contrast</b>	<b>2.2000</b>	<b>SS (Contrast)</b>	<b>12.100</b>
Scheffe's F	0.41	P (Scheffe's F)	0.8363
T-Statistic	1.43	P (T-Statistic)	0.1678
SE (Contrast)	1.5370		

Error term used: BLOQUES\*TRATAMIEN, 20 DF

## Comentario:

Observe en el ejercicio anterior que, además de las Sumas de Cuadrados de los contrastes, Statistix presenta pruebas t. Estas pruebas, que se realizan con las medias, tienen un solo grado de libertad y son equivalentes a las pruebas F por la relación:  $F = t^2$ .

### Cuando el número de repeticiones no es igual

En las comparaciones ortogonales, cuando los tratamientos tienen diferente número de repeticiones, la principal complicación se presenta al momento de definir los coeficientes. Las fórmulas anteriores se modifican para tener en cuenta esta variable. En este caso, el contraste lineal es válido si  $\sum r_i c_i = 0$

La otra modificación tiene que ver con el divisor de la cantidad  $Q^2$ . Esto es:  

$$D_{id} = \sum r_i c_i^2$$

Veamos un ejemplo:

**Ejercicio 17.4.2.** En un experimento conducido con un DCA se obtuvo los siguientes resultados:

Cuadro 1. Totales de cuatro tratamientos y número de repeticiones

	Tratamientos			
	H1	H2	H3	H4
	10	14	20	16
	12	15	22	19
	15	18	24	11
	8	20	-	-
$T_i$	45	67	66	46
$r_i$	4	4	3	3

Cuadro 2. Análisis de Varianza con Statistix

Source	DF	SS	MS	F	P
Tratam	3	201.833	67.2778	7.46	0.0065
Error	10	90.167	9.0167		
Total	13	292.000			

Con la información provista realice las siguientes comparaciones ortogonales:

1.  $(H1+H2) - (H3 + H4)$ ,
2.  $(H1 - H2)$ ,
3.  $(H3 - H4)$

Primero se deben definir los coeficientes. En el caso de la comparación 1, si los tratamientos tuvieran el mismo número de repeticiones, los coeficientes serían:

+1 +1 -1 -1. Sin embargo, estos coeficientes no satisfacen la condición  $\sum r_i c_i = 0$  :

$$\sum r_i c_i = 4*(+1) + 4*(+1) + 3*(-1) + 3*(-1) = 8-6 = 2$$

Es decir:  $\sum r_i c_i \neq 0$

Obviamente, se requiere que los coeficientes de los tratamientos con tres repeticiones sean numéricamente mayores que uno. Si en vez de -1 y -1 se escogiera -1.3333 y -1.3333, se lograría el objetivo deseado, pero se prefiere, por conveniencia y exactitud de los cálculos, emplear números enteros o fracciones exactas. Esto se logra, si los coeficientes de los tratamientos con

tres repeticiones, en este caso, fueran -2 y -2, y los de cuatro repeticiones +1.5 y +1.5. Con esto se tendría:

$$\sum r_i c_i = 4*(+1.5) + 4*(+1.5) + 3*(-2) + 3*(-2) = 0$$

A continuación, se realizan los demás cálculos:

Cuadro 3. Comparaciones ortogonales, valores del contraste lineal  $D_{id}$

Tratamientos	H1	H2	H3	H4			
Totales $T_i$	45	67	66	46			
Contrastes	Coeficientes				$Q_{id}$	$D_{id}$	$SCQ_{id}$
$(H1+H2) - (H3+H4)$	+1.5	+1.5	-2	-2	-56	42	74.67
$H1 - H2$	+1	-1	0	0	-22	8	60.50
$H3 - H4$	0	0	+1	-1	20	6	66.67

El divisor para el primer contraste es:

$$D_{id} = \sum r_i c_i^2 = 4*(1.5)^2 + 4*(1.5)^2 + 3*(-2)^2 + 3*(-2)^2 = 42$$

La suma de cuadrados del primer contraste es:  $SCQ_{id} = \frac{Q_{id}^2}{D_{id}} = \frac{-56^2}{42} = 74.67$

En forma similar se calculan las otras cantidades que se presentan en el Cuadro 3.

Se puede verificar que los contrastes lineales son ortogonales. La fórmula para esto es:  $\sum r_i c_{1i} c_{2i} = 0$ . Si la sumatoria da un valor distinto de cero, los contrastes no son ortogonales. Como ejemplo, veamos si los contrastes 1 y 2 son en verdad ortogonales:

$$\sum r_i c_{1i} c_{2i} = (4 \cdot 1.5 \cdot 1) + (4 \cdot 1.5 \cdot -1) + (3 \cdot -2 \cdot 0) + (3 \cdot -2 \cdot 0) = 0$$

Se concluye, entonces, que los dos contrastes son ortogonales. La misma prueba puede hacerse para los contrastes 1 y 3 y para los contrastes 2 y 3.

Es también una prueba de ortogonalidad el que la sumatoria de todos los valores  $SCQ_{id}$  sea igual a la  $SCTrat$ . De acuerdo con la información de los Cuadros 2 y 3 esto se cumple. Es decir,

$$\sum SCQ_{id} = SCTrat = 201.83$$

Con Statistix los resultados son:

Statistix 9.0 contrastes dif r,

**AOV Contrasts for: T1 T2 T3 T4**

Contrast Number 1

Contrast Coefficients: 1 1 -1 -1

<b>Contrast</b>	<b>-9.3333</b>	<b>SS (Contrast)</b>	<b>74.667</b>
Scheffe's F	2.76	P (Scheffe's F)	0.0976
T-Statistic	-2.88	P (T-Statistic)	0.0165
SE (Contrast)	3.2434		

Contrast Number 2

Contrast Coefficients: 1 -1 0 0

<b>Contrast</b>	<b>-5.5000</b>	<b>SS (Contrast)</b>	<b>60.500</b>
Scheffe's F	2.24	P (Scheffe's F)	0.1466
T-Statistic	-2.59	P (T-Statistic)	0.0269
SE (Contrast)	2.1233		

Contrast Number 3

Contrast Coefficients: 0 0 1 -1

<b>Contrast</b>	<b>6.6667</b>	<b>SS (Contrast)</b>	<b>66.667</b>
Scheffe's F	2.46	P (Scheffe's F)	0.1224
T-Statistic	2.72	P (T-Statistic)	0.0216
SE (Contrast)	2.4518		

**Comentario:** *Statistix* toma en cuenta el diferente número de repeticiones. Sin embargo, los coeficientes que acepta el programa *Statistix* son los mismos que se usarían si las repeticiones fueran iguales. Esto, obviamente, facilita la obtención de los resultados.

## 17.5. Cuando los tratamientos son cuantitativos

Cuando los tratamientos son cuantitativos no es adecuado el uso de ninguno de los métodos de separación de medias estudiados en la primera parte de este capítulo. El método correcto es el empleo de *polinomios ortogonales*, que no es más que una forma simplificada de identificación de tendencias (lineal, cuadrática, cúbica, etcétera).

En los contrastes ortogonales, de la sección 17.4, el investigador decide qué coeficientes emplear para sus comparaciones. En cambio, en el caso de los polinomios ortogonales, los coeficientes se determinan algebraicamente y solo dependen del grado del polinomio y del número de tratamientos. La condición sigue siendo, en todo caso, que la sumatoria de los coeficientes sea igual a cero:  $\sum c_i = 0$

Para facilitar los cálculos se recomienda que los tratamientos (niveles o dosis), es decir la variable independiente  $X$ , estén igualmente espaciados y que el número de repeticiones sea igual. Si se tratara de un experimento con cinco niveles de un fertilizante, los valores de  $X_i$  podrían ser, como ejemplo: 0, 15, 30, 45, 60. Note que la diferencia, o intervalo, entre todos los niveles contiguos es 15. Esto es lo que se denomina “igualmente espaciados”. Cuando los niveles no están igualmente espaciados, la definición de los coeficientes en forma manual es una tarea bastante laboriosa, más aún si el número de repeticiones no es igual para todos los tratamientos. Sin embargo, los programas estadísticos computarizados resuelven estos problemas con facilidad. *Statistix*, por ejemplo, genera los coeficientes de acuerdo con los datos disponibles.

En el Cuadro 17.5.1 se encuentran los coeficientes  $c_i$ , los que se multiplican por los totales de los tratamientos y luego se suman los productos, teniendo en cuenta los signos. Las sumatorias dan como resultado los valores de  $Q_i$ . Puede notarse que esto es similar a lo que se hizo en la sección anterior. Pero en este caso, a diferencia de la sección anterior, los coeficientes del Cuadro 17.5.1 se basan en una codificación de los valores de los tratamientos o niveles  $X_i$ . En todo caso lo que se busca, para facilitar los cálculos, es que los coeficientes sean números enteros mínimos posibles y que su suma, para cualquier comparación, sea igual a cero.

Veamos, como ejemplo, el origen de los coeficientes para las respuestas lineales y cuadráticas para tres, cuatro y cinco niveles o tratamientos cuantitativos.

### 1. Con tres niveles igualmente espaciados

**Respuesta lineal:**  $X$ . Con tres niveles, los coeficientes para la respuesta lineal son:

-1, 0, +1. Estos son los valores más bajos, en números enteros, para que la sumatoria de los coeficientes sea igual a cero.

**Respuesta cuadrática:**  $X^2$ . Para encontrar los coeficientes de la respuesta cuadrática es necesario elevar al cuadrado los coeficientes de la respuesta lineal. El resultado es: +1, 0, +1. Sin embargo, la sumatoria de los coeficientes no es igual a cero. Es decir:  $\sum c_i \neq 0$ . Para obviar esto se calcula la media de los coeficientes y se la resta de cada uno de ellos. La media es  $2/3$ . Los nuevos valores son, entonces: +1/3, -2/3, +1/3. Como es preferible trabajar con números enteros, se multiplica cada uno de los términos por 3. El resultado es: +1, -2, +1. Estos son los coeficientes de la respuesta cuadrática.

Esta variable, doblemente codificada podemos denominarla  $X_2$ , que es igual a:

$$X_2 = \left(X_i^2 - \frac{2}{3}\right) * 3 = 3X^2 - 2$$

La ecuación de la respuesta cuadrática, para los valores esperados,  $\hat{Y}$ , es:

$$\hat{Y} = \bar{Y} + b_1X + b_2X_2, \text{ o lo que igual: } \hat{Y} = \bar{Y} + b_1X + b_2(3X^2 - 2)$$

### 2. Con cuatro niveles igualmente espaciados

**Respuesta lineal:**  $X$ . Con cuatro niveles, los coeficientes para la respuesta lineal son:

$$-3, -1, +1, +3.$$

**Respuesta cuadrática:**  $X^2$ . Al elevar al cuadrado los coeficientes de la respuesta lineal, los valores son: +9, +1, +1, +9. Como estos valores no suman cero, se resta de cada uno la media, que es igual a 5. Los nuevos valores son: +4, -4, -4, +4. Ahora, para simplificar, se dividen para 4 y se obtienen los coeficientes: +1, -1, -1, +1. Esta doble codificación se expresa de esta manera:  $X_2 = \frac{(X^2 - 5)}{4}$

La ecuación de la respuesta cuadrática, para los valores esperados,  $\hat{Y}$ , es:

$$\hat{Y} = \bar{Y} + b_1X + b_2X_2, \text{ o lo que igual: } \hat{Y} = \bar{Y} + b_1X + b_2 \frac{(X^2 - 5)}{4}$$

### 3. Con cinco niveles igualmente espaciados

**Respuesta lineal:**  $X$ . Con cinco niveles, los coeficientes para la respuesta lineal son:

$$-2, -1, 0, +1, +2.$$

**Respuesta cuadrática:**  $X^2$ . Al elevar al cuadrado los coeficientes de la respuesta lineal, los valores son: +4, +1, 0, +1, +4. Como estos valores no suman cero, se resta de cada uno la media, que es igual a 2. Los nuevos valores son: +2, -1, -2, -1, +2. Esta codificación da como resultado:  $X_2 = (X^2-2)$

La ecuación de la respuesta cuadrática, para los valores esperados,  $\hat{y}$ , es:

$$\hat{Y} = \bar{\bar{Y}} + b_1X + b_2X_2, \text{ o lo que igual: } \hat{Y} = \bar{\bar{Y}} + b_1X + b_2(X^2-2)$$

Los coeficientes para las respuestas cúbica, cuártica y más altas pueden obtenerse en la forma descrita en los numerales anteriores. En todo caso, cada tendencia tiene tan solo un grado de libertad. Es decir, por ejemplo, no se puede encontrar una tendencia cúbica si únicamente se tiene tres tratamientos (sólo dos grados de libertad). En el Cuadro 17.5.1 se presentan los coeficientes para las distintas tendencias hasta para cinco tratamientos.

Debe tomarse en cuenta que el procedimiento descrito en párrafos anteriores no es más que un método simplificado de regresión polinómica. Estos métodos se desarrollaron para evitar largos y tediosos cálculos. Actualmente, con el empleo de programas computarizados, pueden realizarse los análisis de regresión polinómica rápidamente y sin necesidad de codificar los niveles.

**Cuadro 17.5.1. Número de tratamientos cuantitativos, coeficientes y divisores para realizar comparaciones ortogonales polinómicas**

Número de Tratamientos	Tendencias (grado del polinomio)	Totales de los Tratamientos					Divisor $D_i$ $r \sum c_i^2$
		T1	T2	T3	T4	T5	
2	Lineal	-1	+1				$r*2$
3	Lineal	-1	0	+1			$r*2$
	cuadrática	+1	-2	+1			$r*6$
4	Lineal	-3	-1	+1	+3		$r*20$
	cuadrática	+1	-1	-1	+1		$r*4$
	Cúbica	-1	+3	-3	+1		$r*20$



5	Lineal	-2	-1	0	+1	+2	<b>r*10</b>
	cuadrática	+2	-1	-2	-1	+2	<b>r*14</b>
	Cúbica	-1	+2	0	-2	+1	<b>r*10</b>
	Cuártica	+1	-4	+6	-4	+1	<b>r*70</b>

Veamos un ejemplo:

**Ejercicio 17.5.1.** En el Ejercicio 16.3.1 se presentó un experimento conducido con un diseño de cuadrado latino  $5^2$ , en el que se evaluó, en cultivos asociados, el efecto de distancias de siembra en sorgo sobre el rendimiento de una variedad de soya. Los totales de los tratamientos se muestran en el Cuadro 1. Mediante el método de polinomios ortogonales encontrar las tendencias y analizar los resultados.

**Cuadro 1.** Distancias de siembra de sorgo y totales del rendimiento del cultivo asociado soya, en quintales métricos

	Distancias de siembra (tratamientos)				
	A=5 cm	B=10 cm	C=15 cm	D=20 cm	E=25 cm
$T_k$	91.77	101.11	98.15	118.23	113.09
$\bar{Y}_k$	18.35	20.22	19.63	23.65	22.62

El valor de la media del experimento es:  $\bar{Y} = 20.89$

En el Cuadro 2 se presentan los coeficientes de los polinomios ortogonales para cinco tratamientos, según el Cuadro 17.2.

**Cuadro 2.** Respuestas ortogonales, coeficientes y sumas de cuadrados de los contrastes

Distancias de siembra	A=5 cm	B=10 cm	C=15 cm	D=20 cm	E=25 cm			
Totales $T_k$	91.77	101.11	98.15	118.23	113.09			
<b>Respuesta</b>	<b>Coefficientes</b>					$Q_i$	$D_i$	$SCQ_i$
Lineal	-2	-1	0	+1	+2	59.76	5*10	71.42
Cuadrática	+2	-1	-2	-1	+2	-5.92	5*14	0.50
Cúbica	-1	+2	0	-2	+1	-12.92	5*10	3.33
Cuártica	+1	-4	+6	-4	+1	-83.6	5*70	19.96

Con el programa Statistix no hace falta que el investigador introduzca los coeficientes. Sólo debe indicar el grado del polinomio. El programa calcula las tendencias y su significación estadística:

Statistix 9.0                   soya dist sorgo CL,

**Polynomial Contrasts of Rendimien by Distancia**

Degree = 1, Linear Trend

Contrast	3.7796	SS (Contrast)	71.425
T-Statistic	1.98	P (T-Statistic)	0.0713
SE (Contrast)	1.9102		

Degree = 2, Quadratic Trend

Contrast	-0.3164	SS (Contrast)	0.5007
T-Statistic	-0.17	P (T-Statistic)	0.8712
SE (Contrast)	1.9102		

Degree = 3, Cubic Trend

Contrast	-0.8171	SS (Contrast)	3.3385
T-Statistic	-0.43	P (T-Statistic)	0.6764
SE (Contrast)	1.9102		

Degree = 4

<b>Contrast</b>	<b>-1.9984</b>	<b>SS (Contrast)</b>	<b>19.968</b>
T-Statistic	-1.05	P (T-Statistic)	0.3161
SE (Contrast)	1.9102		

Error term used: Hileras\*Columnas\*Distancia, 12 DF

En el Cuadro 3 se presenta el ADEVA del ejercicio 16.3.1, en el que se incluye las respuestas polinómicas ortogonales.

Cuadro 3. Análisis de Varianza con los componentes polinómicos para el cuadrado latino del ejercicio 16.3.1

F.V.	S.C.	g.l.	C.M	F	F <sub>05</sub>	F <sub>01</sub>
Hileras	307.69	4	76.92	4.22*	3.26	5.41
Columnas	50.68	4	12.67	0.69 n.s.		
Distancias	95.23	4	23.81	1.31 n.s.		
Respuesta Lineal	71.42	1	71.42	3.92 n.s.	4.75	9.33
Cuadrática	0.50	1	0.50	0.03 n.s.		
Cúbica	3.33	1	3.33	0.18 n.s.		
Cuártica	19.96	1	19.96	1.09 n.s.		
Error	218.93	12	18.24			
TOTAL	672.53	24				

Puede probarse la ortogonalidad de los polinomios con esta fórmula:  
 $\sum SCQ_i = SCTrat$

Se observa, en el cuadro anterior, que ninguna de las respuestas fue significativa. Es decir, la curva de regresión es estadísticamente paralela al eje de la abscisa.

El gráfico de la curva del polinomio cuártico, generado con el programa Statistix, se presenta a continuación:

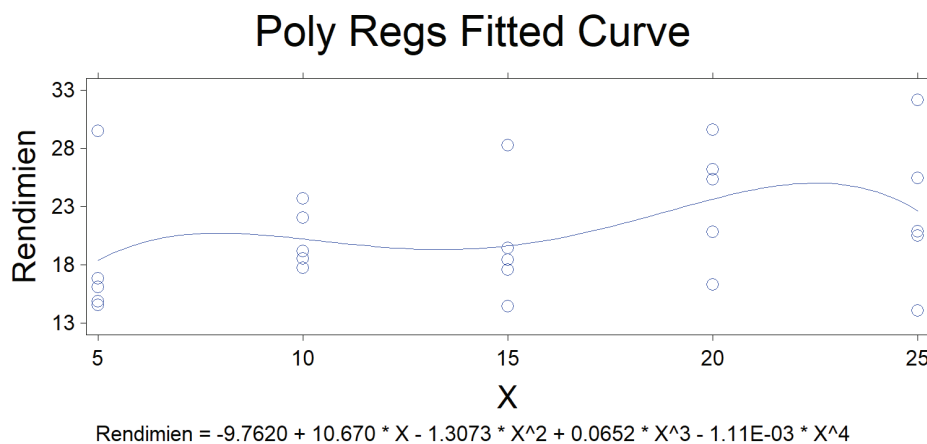


Gráfico 1. Curva y ecuación del polinomio cuártico

En términos del experimento y con una alta probabilidad, se demuestra que las distancias de siembra en el cultivo asociado no afectaron en modo alguno el rendimiento de la soya. El comportamiento errático de las respuestas denota que hubo problemas en la conducción del ensayo. Los círculos en el gráfico corresponden a las observaciones. Puede notarse la excesiva variación dentro de cada distancia. Esto determinó un error experimental inflado y consecuentemente, un alto coeficiente de variación (20.44%)

Por lo general, cuando se busca respuestas a diferentes niveles de un insumo, se busca el punto en el que se alcanza el máximo rendimiento. Esto se logra mediante la respuesta cuadrática. Solo por razones del ejemplo, se presenta el correspondiente polinomio en el gráfico 2.

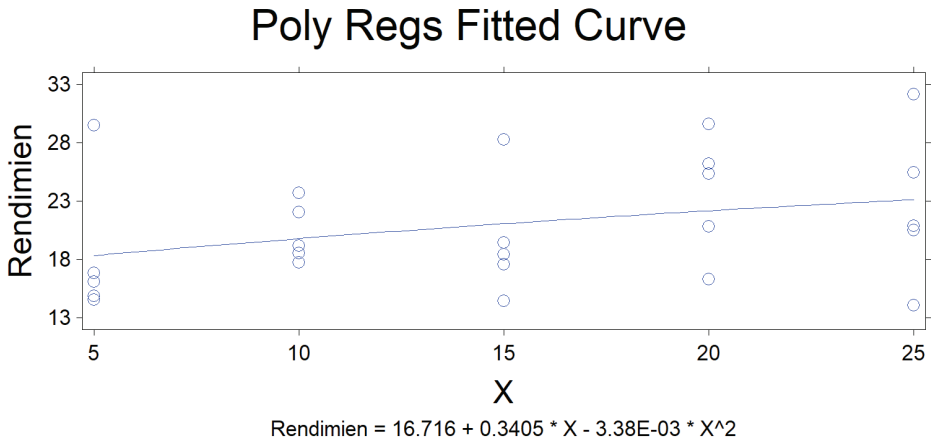


Gráfico 2. Curva del polinomio cuadrático

Puede verse al pie del gráfico 2, la ecuación de regresión cuadrática. Esta ecuación es igual a  $\hat{Y} = b_0 + b_1X + b_2X^2$ . Esta es la ecuación que resulta cuando se emplean los niveles sin codificar; en este ejemplo, los valores 5, 10, 15, 20, 25 cm. Con el método de polinomios ortogonales la fórmula, cuando se tiene cinco tratamientos, es:  $\hat{Y} = \bar{Y} + b_1X + b_2(X^2-2)$ . Cuando se busca el punto de máximo rendimiento los resultados con las dos fórmulas son obviamente iguales, aunque es necesario decodificar el valor de  $X$  de la última ecuación. Ejemplo de esto se verá más adelante.

En todo caso, con la información del Cuadro 2 pueden encontrarse los coeficientes de regresión  $b_1$  y  $b_2$ :

$$b_1 = \frac{Q_{lineal}}{r \sum c_i^2} = \frac{59.76}{50} = 1.1952 \quad b_2 = \frac{Q_{cuad}}{r \sum c_i^2} = \frac{-5.92}{70} = -0.08457$$

La media del experimento fue igual a:  $\bar{Y} = 20.89$ . Al reemplazar los coeficientes  $b_1$  y  $b_2$  en la fórmula, la media se convierte en la intersección  $b_0$ :

$$\hat{Y} = 20.89 + 1.1952X - 0.0846(X^2 - 2), \quad \hat{Y} = 21.06 + 1.1952X - 0.0846X^2$$

*“Más datos significa más información, pero también significa más falsa información.” Nassim Nicholas Taleb<sup>42</sup>*

Los *látices* son diseños experimentales que se usan cuando el número de tratamientos es muy elevado. Como el tamaño de la repetición necesariamente debe ser muy grande, no se logra homogeneidad entre las unidades experimentales. Al dividir las repeticiones en áreas más pequeñas, denominadas *bloques incompletos*, se aspira a reducir la variabilidad entre las unidades experimentales. Los látices tienen especial importancia en programas de mejoramiento vegetal, en los que se requiere evaluar un considerable número de líneas y variedades. Fue Yates (1939) quien creó estos diseños.

El término “látice” no existe en español. Se lo ha tomado del inglés “lattice” que se refiere a una estructura entrelazada de tiras de diferentes materiales. Se verá que, por la distribución armónica de los tratamientos en los bloques y los bloques en las repeticiones, el término parece adecuado.

En los látices, una repetición contiene todos los tratamientos distribuidos en bloques incompletos; por lo tanto, en estos diseños, el término “repetición” no es sinónimo de “bloque”, como sí lo fue en el diseño de Bloques Completos al Azar. Es decir, los látices son diseños de bloques incompletos. Una restricción importante en estos diseños es que el número de tratamientos tiene que ser un cuadrado perfecto.

Existen varios tipos de látices, como lo señalan Cochran y Cox (1968), Kempthorne (1979) y otros varios autores. Sin embargo, los más utilizados son los látices balanceados y los parcialmente balanceados. En este texto se verán ejemplos de los dos tipos. En cualquier caso, si con cualquiera de los dos látices no se

---

42 Escritor y epistemólogo libanés. <https://www.wired.com/2013/02/big-data-means-big-errors-people/>

ganó en eficiencia, el experimento puede analizarse como un diseño de Bloques Completos al Azar.

### 18.1. Látices Balanceados

En los látices balanceados, si  $k^2$  es el número de tratamientos,  $k$  es el número de bloques incompletos en cada repetición. El número de repeticiones es igual a  $r = k + 1$ . Como ejemplo, si  $k^2 = 64$ ,  $k = 8$  y  $r = 9$ .

Se deduce, del párrafo anterior, que el número de tratamientos en los Látices debe ser un cuadrado perfecto  $k^2$ . Esta condición es generalmente fácil de lograr. Si por ejemplo se dispone de 23 líneas experimentales que se desea evaluar, se puede añadir dos otras líneas o variedades para un total de  $k^2=25$ .

El término balanceado se refiere al requisito de que todos y cada uno de los tratamientos deben estar en combinación, una sola vez en un mismo bloque, con otros diferentes tratamientos. Sin embargo, esta restricción también determina que no sea posible estructurar Látices completamente balanceados para 36, 100 y 144 tratamientos, a pesar de que estos números son cuadrados perfectos.

#### *Asignación de los tratamientos*

La asignación de los tratamientos a los bloques incompletos se vuelve una tarea tediosa, en especial con los látices más grandes. Felizmente, en textos como el de Cochran y Cox (1968), Gomez y Gomez (1984) y también en internet se encuentran los planos de varios tipos de látices. Una vez escogido el plano se debe sortear el orden de los bloques en cada repetición y el orden de los tratamientos en cada bloque, en forma similar a lo que se hace en el Diseño Cuadrado Latino. Como ejemplo, para un Látice  $k^2 = 9$  se tiene el siguiente plano:

Cuadro 18.1.1. Látice Completamente Balanceado con nueve tratamientos ( $k^2 = 9$ ), cuatro repeticiones ( $r = 4$ ) y tres bloques en cada repetición ( $k = 3$ )

Repetición 1		Repetición 2		Repetición 3		Repetición 4	
Bloq*	Tratam.	Bloq.	Tratam.	Bloq.	Tratam.	Bloq.	Tratam.
1	(1)(2)(3)	4	(1)(4)(7)	7	(1)(5)(9)	10	(1)(6)(8)
2	(4)(5)(6)	5	(2)(5)(8)	8	(2)(6)(7)	11	(2)(4)(9)
3	(7)(8)(9)	6	(3)(6)(9)	9	(3)(4)(8)	12	(3)(5)(7)

Note el arreglo sistemático del plano del cuadro anterior. Como ejemplo, el tratamiento 1 está en la primera unidad experimental de cada repetición. El arreglo sistemático se corrige con la aleatorización.

En el cuadro anterior, los números en paréntesis indican los tratamientos. Note que cada bloque es diferente de los otros. Por eso los bloques se numeran del 1 al 12.

Puede observarse también, que todos y cada uno de los tratamientos está en combinación con dos distintos tratamientos en cada repetición. Este es el significado del término “balanceado”. Tomemos como ejemplo el tratamiento 5. Este tratamiento está distribuido así: en la repetición 1, con los tratamientos 6 y 4; en la repetición 2, con los tratamientos 2 y 8; en la repetición 3, con los tratamientos 1 y 9; en la repetición 4, con los tratamientos 3 y 7. Es decir, el tratamiento 5 está en las cuatro repeticiones, en bloques incompletos de tres tratamientos, en combinación con todos los otros tratamientos tomados de dos en dos. Gracias a esto, las diferencias entre medias de tratamientos se calculan con igual precisión en los látices balanceados.

### ***Aleatorización de los tratamientos***

Una vez que se ha seleccionado el plano correspondiente para el experimento, debe procederse a la aleatorización de la siguiente forma: 1. Asignar a los tratamientos un número al azar, 2. Sortear el orden de los bloques en cada repetición, 3. Sortear los tratamientos dentro de cada bloque.

El esquema del ADEVA de un látice completamente balanceado, con  $k^2$  número de tratamientos, se presenta en el Cuadro 18.1.2:

**Cuadro 18.1.2. Esquema del ADEVA de un Látice Balanceado  $k^2$**

F.V.	S.C.	g.l.
Repeticiones	$SCR_{BCA}$	$k$
Tratamientos (sin ajustar)	$SCTrat_{sin\ ajust}$	$k^2-1$
Bloques (ajustados)	$SCB_{ajust}$	$k^2-1$
Error Intra-bloque	$SCE_{intrbl}$	$(k-1)(k^2-1)$
Tratamientos (ajustados)	$SCTrat_{ajust}$	$k^2-1$
Error efectivo	$SCE_{efectivo}$	$(k-1)(k^2-1)$
Total	$SCT$	$k^3 + k^2-1$



En este esquema se combinan fuentes de variación del BCA con otras propias del Látice. La Suma de Cuadrados de Repeticiones, Suma de Cuadrados de Tratamientos (sin ajustar) y Suma de Cuadrados Total son las mismas del diseño BCA y se calculan de forma igual que en un BCA. El procedimiento para calcular las otras fuentes de variación se presenta mediante un ejemplo:

**Ejercicio 18.1.1.** Con los datos de los siguientes dos cuadros, realice el análisis correspondiente a un látice balanceado  $k_2 = 9$ . (Experimento simulado con la ayuda de un programa estadístico.)

Cuadro 1. Repeticiones, bloques, tratamientos y totales de los 12 bloques ( $B_k$ ) para la variable dependiente rendimiento

REPS	BLOQ	TREAT	REND	$B_k$	REPS	BLOQ	TREAT	REND	$B_k$
1	1	1	12.0		3	7	1	11.2	
1	1	2	21.4		3	7	5	11.6	
1	1	3	19.8	53.2	3	7	9	15.6	38.4
1	2	4	17.9		3	8	2	14.4	
1	2	5	13.7		3	8	6	14.0	
1	2	6	16.7	48.3	3	8	7	16.8	45.2
1	3	7	12.8		3	9	3	17.2	
1	3	8	18.9		3	9	4	19.5	
1	3	9	19.1	50.8	3	9	8	15.7	52.4
2	4	1	14.5		4	10	1	10.1	
2	4	4	22.9		4	10	6	19.1	
2	4	7	16.6	54.0	4	10	8	16.4	45.6
2	5	2	18.4		4	11	2	15.4	
2	5	5	12.6		4	11	4	19.8	
2	5	8	12.5	43.5	4	11	9	16.9	52.1
2	6	3	15.7		4	12	3	14.6	
2	6	6	18.2		4	12	5	11.7	
2	6	9	20.6	54.5	4	12	7	18.0	44.3

Cuadro 2. Totales de los tratamientos y repeticiones

Repeti- ciones	Tratamientos									Totales $R_j$
	1	2	3	4	5	6	7	8	9	
1	12.0	21.4	19.8	17.9	13.7	16.7	12.8	18.9	19.1	152.3
2	14.5	18.4	15.7	22.9	12.6	18.2	16.6	12.5	20.6	152.0
3	11.2	14.4	17.2	19.5	11.6	14.0	16.8	15.7	15.6	136.0
4	10.1	15.4	14.6	19.8	11.7	19.1	18.0	16.4	16.9	142.0
Totales $T_i$	47.8	69.6	67.3	80.1	49.6	68.0	64.2	63.5	72.2	Suma Total $\sum T_i = 582.3$

*El procedimiento para el análisis sigue los siguientes pasos:*

#### PRIMER PASO

Con los datos del Cuadro 2 se realiza el análisis de varianza como si se tratara de un diseño BCA, con cuatro repeticiones y nueve tratamientos (en este análisis no se toma en cuenta los bloques):

$$\text{El Factor de Corrección es: } FC = \frac{582.3^2}{36} = 9418.7$$

$$\text{Suma Cuadrados Total: } SCT = (12.0^2 + 21.4^2 + \dots + 16.9^2) - FC = 350.21$$

$$\text{Suma Cuadrados Repeticiones: } SCR = \frac{152.3^2 + \dots + 142.0^2}{9} - FC = 21.22$$

$$\text{Suma Cuadrados Tratamientos: } SCTrat = \frac{47.8^2 + 69.6^2 + \dots + 72.2^2}{4} - FC = 212.60$$

$$\text{Suma Cuadrados Error: } SCE = SCT - SCR - SCTrat = 116.39$$

Luego, se forma el Cuadro ADEVA para el diseño BCA:

Cuadro 3. Análisis de Varianza como si se tratara de un diseño BCA

FV	SC	GL	CM	F	P	$F_{0.05}$	$F_{0.01}$
REP	21.22	3	7.07				
TRAT	212.60	8	26.57	5.48**	0.054%	2.36	3.36
Error	116.39	24	4.85				
Total	350.21	35					

## SEGUNDO PASO

Se procede ahora a encontrar los valores  $B_t$ . Un valor  $B_t$  es igual a la sumatoria de los totales de los bloques en los que aparece un tratamiento. Como ejemplo, el tratamiento 1 se encuentra en los bloques 1, 4, 7 y 10 (Cuadro 1). El valor es:  $B_1 = 53.2 + 54.0 + 38.4 + 45.6 = 191.2$ . Los valores  $B_t$  para los nueve tratamientos, constan en el Cuadro 4.

La sumatoria de los valores  $B_t$  es  $k$  veces la sumatoria de los totales de los tratamientos

$T_i$ . En los látices,  $k$  es el número de bloques en cada repetición. Por lo tanto, la relación  $\sum B_t = k \sum T_i$  sirve para comprobar los cálculos.

## TERCER PASO

Se deben encontrar también Los valores  $D_i$ , que son desviaciones de los tratamientos con relación a los bloques incompletos. Se encuentran, para cada tratamiento, con la siguiente fórmula:

$$D_i = (kT_i - rB_t + \sum T_i) \quad (113)$$

Como ejemplo, para el tratamiento 1:

$$D_1 = (3 \cdot 47.8 - 4 \cdot 191.2 + 582.3) = -39.1$$

Con esta información, para todos los tratamientos, se estructura el siguiente cuadro:

Cuadro 4. Totales de los tratamientos ( $T_i$ ), valores  $B_t$  y valores  $D_i$

Tratamientos	$T_i$	$B_t$	$D_i$	$D_i^2$
1	47.8	191.2	-39.1	1528.81
2	69.6	194.0	15.1	228.01
3	67.3	204.4	-33.4	1115.56
4	80.1	206.8	-4.6	21.16
5	49.6	174.5	33.1	1095.61
6	68.0	193.6	11.9	141.61
7	64.2	194.3	-2.3	5.29
8	63.5	192.3	3.6	12.96
9	72.2	195.8	15.7	246.49
Sumas Totales	$\sum T_i = 582.3$	$\sum B_t = 1746.9$	$\sum D_i = 0$	$\sum D_i^2 = 4395.50$

Note que  $\sum D_i = 0$ . Esta ecuación siempre debe ser igual a cero (como otras sumatorias de desviaciones vistas anteriormente), de lo contrario, hay errores en los cálculos.

#### CUARTO PASO

Se dispone ya de la información necesaria para partir la Suma de Cuadrados del Error del BCA, en dos componentes: la Suma de Cuadrados de Bloques ajustados por repeticiones:  $SCB_{ajust}$  y la Suma de Cuadrados del Error intra-bloque:  $SCE_{int rb}$ . La fórmula para encontrar la  $SCB_{ajust}$  es:

$$SCB_{ajust} = \frac{\sum D_i^2}{k^3 * (k+1)} = SCB_{ajust} = \frac{4395.50}{108} = 40.70 \quad (114)$$

Luego se encuentra la Suma de Cuadrados del Error Intra-bloque,  $SCE_{int rb}$  que es la diferencia entre la Suma de Cuadrados del Error del BCA (Cuadro 18.1.5) y la estadística  $SCB_{ajust}$ . Es decir:

$$SCE_{int rb} = SCE_{BCA} - SCB_{ajust} = 116.39 - 40.70 = 75.69$$

#### QUINTO PASO

A continuación, debe determinarse el Factor de Ajuste,  $q$ . Para esto se necesita estimar los cuadrados medios de las dos sumas de cuadrados encontradas en el paso anterior. Los grados de libertad son, respectivamente,  $(k^2-1)$  y  $(k-1)(k^2-1)$ , como se presentó en el esquema del ADEVA (Cuadro 18.1.2):

$$CMB_{ajust} = \frac{SCB_{ajust}}{(k^2-1)} = \frac{40.70}{8} = 5.09$$

$$CME_{int rb} = \frac{SCE_{int rb}}{(k-1)(k^2-1)} = \frac{75.69}{16} = 4.73$$

El factor de ajuste se encuentra con la siguiente fórmula:

$$q = \frac{CMB_{ajust} - CME_{int rb}}{k^2 * CMB_{ajust}} = q = \frac{5.09 - 4.73}{9 * 5.09} = 0.0079 \quad (115)$$

Si el valor de ajuste  $q$  resulta igual a cero, o es negativo, no hace falta proseguir con otros cálculos. De darse este caso, se realiza el análisis de los datos como si fuera un diseño de Bloques Completos al Azar.

En este ejercicio el factor de ajuste es positivo, aunque cercano a cero, por lo que seguramente se ganará muy poco en precisión al analizar los resultados con el diseño Láctice Balanceado. Sin embargo, por razones del ejemplo, se continua con el análisis.

#### SEXTO PASO

Como  $q > 0$ , se procede a encontrar los totales de los tratamientos ajustados. La fórmula es:

$$T_{i(ajust)} = T_i + q * D_i \quad (116)$$

El ajuste es necesario para compensar por errores de muestreo en los valores  $D_i$ . Los valores de los tratamientos ajustados y las medias ajustadas se presentan en el siguiente cuadro:

Cuadro 5. Tratamientos, medias, tratamientos ajustados y medias ajustadas

No. Tratamientos	Tratamientos $T_i$	Medias $\bar{Y}_i$	Trat (ajust) $T_i + q * D_i$	Medias (ajust) $\bar{Y}_{i(ajust)}$
1	47.80	11.95	47.49	11.87
2	69.60	17.40	69.72	17.43
3	67.30	16.83	67.04	16.76
4	80.10	20.03	80.06	20.02
5	49.60	12.40	49.86	12.46
6	68.00	17.00	68.09	17.02
7	64.20	16.05	64.18	16.05
8	63.50	15.88	63.53	15.88
9	72.20	18.05	72.32	18.08
Totales	582.30		582.30	

Luego, debe encontrarse la Suma de Cuadrados de Tratamientos ajustados. El número de repeticiones es  $(k + 1) = 4$ . La fórmula (117) es:

$$SCTrat_{ajust} = \frac{\sum T_{i(ajust)}^2}{k+1} - FCSCTrat_{ajust} = \frac{47.49^2 + 69.72^2 + \dots + 72.32^2}{4} - \frac{582.30^2}{36} = 213.56$$

#### SÉPTIMO PASO

Por último, se requiere encontrar el Cuadrado Medio del Error Efectivo. La fórmula es:

$$CME_{effect} = CME_{intrb} * (1 + qk) \quad (118)$$

$$CME_{effect} = 4.73 * (1 + 0.0077 * 3) = 4.84$$

Con la información obtenida se estructura el Cuadro ADEVA:

**Cuadro 6. Análisis de Varianza. Látice Completamente Balanceado**

F.V.	S.C.	g.l.	C.M	F	P	F <sub>05</sub>	F <sub>01</sub>
Repeticiones	21.21	3					
Tratamientos (sin ajustar)	212.59	8	26.57				
Bloques * Repeticiones	40.70	8	5.09				
Error Intra-bloque	75.69	16	4.73				
Tratamientos (ajustados)	(213.56)	(8*)	26.70	5.51**	0.19%	2.59	3.89
Error efectivo		(16)	4.84				
Total	350.20	35					

\*los paréntesis en las sumas de cuadrados y en los grados de libertad indican que cantidades equivalentes ya fueron tomadas en cuenta para los totales respectivos.

Se puede concluir que por lo menos dos medias de los tratamientos son estadísticamente diferentes, con una probabilidad incluso inferior al nivel alfa del 1%.

$$CV = \frac{\sqrt{CME_{effect}}}{\bar{y}} = \frac{\sqrt{4.84}}{16.18} * 100 = 13.60\%$$

El error estándar para la diferencia de dos medias es igual a:

$$S_{\bar{d}} = \sqrt{\frac{2 * CME_{effect}}{r}} = \sqrt{\frac{2 * 0.019}{4}} = 0.098$$

La eficiencia relativa (E) del látice balanceado, en comparación con el diseño de bloques completos al azar (BCA), se encuentra con la siguiente fórmula:

$$E = \frac{CME_{BCA}}{CME_{effect}} = \frac{4.85}{4.84} = 1.00$$

El análisis con Statistix es el siguiente:

Statistix 9.0

LATICE COMPL BALANC

**Balanced Lattice AOV for REND**

Source	DF	SS	MS	F	P
REP	3	21.219			
TRAT(unadj)	8	212.595			
BLOQ*REP	8	40.699	5.0874		
Intrablock error	16	75.695	4.7309		
TRAT(adj)	(8)	26.6952	5.51	0.0019	
Effective error	(16)	4.8414			
Total	35	350.207			

Grand Mean 16.175 CV 13.60

**Relative efficiency, RCB 1,00**

**Means of REND for TRAT**

TRAT	Mean
1	11.874
2	17.429
3	16.760
4	20.016
5	12.464
6	17.023
7	16.046
8	15.882
9	18.081

Observations per Mean 4  
 Standard Error of a Mean 1.1002  
 Std Error (Diff of 2 Means) 1.5559

**Comentarios:**

- En la práctica, tan solo con nueve tratamientos posiblemente no se justificaría usar este diseño, a no ser que hubiera una buena razón, como la que presentan Cochran y Cox (1968), página 396. De lo contrario, bien se

podría emplear un diseño completamente al azar o, de haber una pendiente, un diseño de bloques completos al azar.

- Note que, en las fuentes de variación, Statistix presenta el componente BLOQ\*REP. Esta interacción tiene el mismo origen y es igual a la Suma de Cuadrados de Bloques ajustados,  $SCB_{ajust}$ .
- Observe, en los resultados de Statistix, que la *Eficiencia Relativa* fue igual a 1.00. Esta estadística compara el Láti-ce con el BCA. Un valor de 1.00 indica que se habría obtenido un resultado similar si el análisis se conducía como el de un diseño BCA. En otras palabras, no hubo heterogeneidad significativa dentro de las repeticiones, como para justificar dividir las en bloques incompletos. Este resultado por lo general está relacionado con el número de tratamientos. En látices más grandes se incrementa la probabilidad de encontrar una mayor variación dentro de las repeticiones.
- Statistix presenta también las medias ajustadas por el efecto de los bloques. (Ver también el Cuadro 5.) Los procedimientos de separación de medias se realizan con las medias ajustadas si se ganó en eficiencia con el diseño láti-ce. De lo contrario, como en este ejemplo, no hace diferencia si se emplean las medias obtenidas con el diseño BCA o las obtenidas con el láti-ce.

## 18.2. Látices parcialmente balanceados

Por el número de tratamientos y el número de repeticiones, los látices completamente balanceados son experimentos muy grandes. Como ejemplo, si  $k^2 = 25$  habría seis repeticiones ( $r = k + 1 = 6$ ). Para un experimento así, se requeriría de  $25 * 6 = 150$  unidades experimentales (U.E). Obviamente, con látices completamente balanceados más grandes que el señalado, el número de unidades experimentales requeridas será muy alto. Los experimentos se encarecen y se vuelven más difíciles de manejar.

Una manera de reducir el tamaño del láti-ce es tener menos repeticiones. Si el investigador decidiera tener únicamente dos repeticiones, para el mismo caso de 25 tratamientos, necesitaría 50 unidades experimentales. Este tipo de diseño se conoce como *Láti-ce Parcialmente Balanceado*.

Para llevar a cabo un experimento con el diseño Láti-ce Parcialmente Balanceado, el investigador puede generar su propio arreglo o, para ganar tiempo, escoger el número de repeticiones que desee en internet. Otra opción es seleccionar, en las referencias de la sección anterior, el arreglo y número de repeticiones que mejor se ajusten a los objetivos de la investigación, tomando también en cuenta los recursos disponibles.



Para el análisis de los datos debe tomarse en cuenta que cualquier tratamiento está en combinación únicamente con algunos otros tratamientos en un mismo bloque. Esto determina que las comparaciones entre medias no tengan la misma precisión.

Para el caso de un látice parcialmente balanceado con 25 tratamientos ( $k_2$ ), dos repeticiones y cinco bloques por repetición, se tiene como ejemplo lo siguiente:

**Cuadro 18.2.1. Arreglo para un Látice Parcialmente Balanceado:  
25 tratamientos, dos repeticiones y cinco bloques incompletos por repetición**

Bloques	repetición I					bloques	repetición II				
1	1	2	3	4	5	6	24	14	9	4	19
2	6	7	8	9	10	7	22	12	17	7	2
3	11	12	13	14	15	8	5	25	10	20	15
4	16	17	18	19	20	9	6	21	1	11	16
5	21	22	23	24	25	10	8	23	3	13	18

Se puede ver, como ejemplo, que el tratamiento 3 está en la primera repetición en el bloque uno, junto con los tratamientos 1, 2, 4 y 5. En la segunda repetición está en el bloque 10 con los tratamientos 8, 23, 13, 18. Con ninguno de los otros 16 tratamientos aparece en un mismo bloque el tratamiento 3. Este es un requisito de este diseño.

Del párrafo anterior puede deducirse que la precisión es mayor cuando se comparan medias de tratamientos que compartan un mismo bloque, que cuando no es éste el caso. Debe entonces calcularse más de un error estándar.

El esquema del ADEVA es el siguiente:

**Cuadro 18.2.2. Esquema del ADEVA para un Látice Parcialmente Balanceado,  
con  $k_2$  tratamientos y  $r$  repeticiones**

F.V.	S.C.	G.L.
Repeticiones	$SCR$	$r-1$
Tratamientos (sin ajustar)	$SCTrat_{sinajust}$	$k^2-1$
Bloques (ajustados)	$SCB_{ajust}$	$r(k-1)$
Tratamientos (ajustados)	$SCTrat_{ajust}$	$k^2-1$
Error intra-bloque	$SCE_{intra}$	$(k-1)(rk-k-1)$
Total	$SCT$	$rk^2-1$

A continuación, veamos un ejemplo:

**Ejercicio 18.2.1.** En el Cuadro 1 se presenta un látice parcialmente balanceado, con 25 tratamientos y dos repeticiones. Los datos provienen de un ensayo de evaluación de varios genotipos de triticale (cruza entre trigo y centeno) provistos por el Centro Internacional de Mejoramiento del Maíz y Trigo (CIMMYT) (Romero y Díaz, 1992)<sup>43</sup>. Algunas características del ensayo se modificaron para facilidad del ejemplo.

Para el arreglo del experimento se tomaron las dos repeticiones presentadas en el Cuadro 18.2.1. Este arreglo o plano proviene del libro de Cochran y Cox (1968), página 428. Para la aleatorización, se asignaron al azar los números del 1 al 25 a los diferentes genotipos de triticale. De esta forma, la primera repetición queda aleatorizada. En la segunda repetición, se sorteó el orden de los bloques incompletos y de los tratamientos en cada bloque.

**Cuadro 1. Rendimiento de 25 genotipos de triticale, en toneladas por hectárea y totales de las repeticiones y bloques  $B$**

Bloque	Trat	Rep I	Total, $B$	Bloque	Trat	Rep II	Total, $B$
1	1	2.73		6	24	1.78	
1	2	2.86		6	9	1.65	
1	3	2.80		6	19	2.24	
1	4	2.98	bloque 1 = 14.88	6	4	1.94	bloque 6= 9.96
1	5	3.51		6	14	2.35	
2	6	3.02		7	7	2.58	
2	7	2.85		7	22	2.74	
2	8	2.79		7	12	2.85	
2	9	2.34	bloque 2= 14.25	7	17	3.08	bloque 7= 14.47
2	10	3.25		7	2	3.22	
3	11	3.46		8	20	2.73	
3	12	3.21		8	25	2.46	

43 Romero, A.R. y O. Díaz. 1992. Comportamiento agronómico de triticale. Departamento de Agronomía, Informe Anual de Investigación, volumen 5, pg. 85-86. Escuela Agrícola Panamericana, El Zamorano. Honduras, C.A.

3	13	2.54		8	15	3.18	
3	14	3.45	bloque 3= 15.50	8	5	3.21	bloque 8= 14.54
3	15	2.84		8	10	2.96	
4	16	2.28		9	1	2.64	
4	17	2.38		9	21	3.15	
4	18	2.71		9	6	3.22	
4	19	2.95	bloque 4= 13.22	9	11	2.84	bloque 9= 14.81
4	20	2.90		9	16	2.96	
5	21	2.70		10	23	2.28	
5	22	3.20		10	18	2.81	
5	23	2.34		10	13	2.66	
5	24	2.55	bloque 5= 13.41	10	3	2.35	bloque 10= 13.34
5	25	2.62		10	8	3.24	
Total, reps		71.26				67.12	138.38

En el siguiente cuadro se presentan los rendimientos de los 25 genotipos de triticale:

**Cuadro 2. Rendimientos totales ( $T_i$ ) de los 25 tratamientos, suma de dos observaciones, en toneladas métricas.**

Trat No.	$T_i$	Trat No.	$T_i$	Trat No.	$T_i$	Trat No.	$T_i$	Trat No.	$T_i$
1	5.37	6	6.24	11	6.30	16	5.24	21	5.85
2	6.08	7	5.43	12	6.06	17	5.46	22	5.94
3	5.15	8	6.03	13	5.20	18	5.52	23	4.62
4	4.92	9	3.99	14	5.80	19	5.19	24	4.33
5	6.72	10	6.21	15	6.02	20	5.63	25	5.08

Note en el Cuadro 1 que los totales de las repeticiones son  $R_1 = 71.26$  y  $R_2 = 67.12$  toneladas. La suma total del rendimiento es  $\sum Y_{ij} = 138.38$  toneladas. (Estas cantidades provienen de la transformación de los rendimientos originales, de kilogramos por parcela, a toneladas métricas por hectárea.)

## Pasos para el análisis

El análisis del látice parcialmente balanceado procede en varios pasos:

### PRIMER PASO

Los datos se analizan como que se tratara de un diseño de Bloques Completos al Azar. Las estadísticas que se calculan son: Factor de Corrección (FC), Suma de Cuadrados Total (SCT), Suma de Cuadrados de Repeticiones (SCR), Suma de Cuadrados de Tratamientos (SCTrat) y Suma de Cuadrados del Error (SCE).

- $FC = \frac{138.38^2}{50} = 382.98$
- $SCT = (2.73^2 + 2.86^2 + \dots + 3.24^2) - FC = 8.44$
- $SCR = \frac{(71.26^2 + 67.12^2)}{25} - FC = 0.3428$
- $SCTrat_{sinajust} = \frac{5.37^2 + 6.08^2 \dots + 5.08^2}{2} - FC = 5.0548$
- $SCE = SCT - SCR - SCTrat = 3.0437$

Luego, a partir de los totales de los bloques del Cuadro 1 se calcula la Suma de Cuadrados de los Bloques dentro de las repeticiones. A esta cantidad se la conoce como  $SCB_{sinajust}$ . (Esta es una cantidad que se la necesitará posteriormente para encontrar la Suma de Cuadrados de los Tratamientos ajustados.) La fórmula es:

$$SCB_{sinajust} = \frac{\sum B_j^2}{k} - \frac{\sum R_h^2}{k^2} \quad (119)$$

$$SCB_{sinajust} = \frac{(14.88^2 + 14.25^2 + \dots + 13.34^2)}{5} - \frac{(71.26^2 + 67.12^2)}{25} = 4.00$$

### SEGUNDO PASO

Ahora es necesario calcular los valores de las desviaciones  $D_i = P - rB$ , en que  $P$  es la suma de los totales de los tratamientos que están en cada bloque y  $rB$  es el producto del número de repeticiones por el total de cada bloque. Pero primero se debe calcular los valores  $P$ . Con la información del Cuadro 2 se construye el Cuadro 3:

Cuadro 3. Los valores P para cada bloque

Bloques	Tratamientos					totales P
1	(1) 5.37	(2) 6.08	(3) 5.15	(4) 4.92	(5) 6.72	28.24
2	(6) 6.24	(7) 5.43	(8) 6.03	(9) 3.99	(10) 6.21	27.90
3	(11) 6.30	(12) 6.06	(13) 5.20	(14) 5.80	(15) 6.02	29.38
4	(16) 5.24	(17) 5.46	(18) 5.52	(19) 5.19	(20) 5.63	27.04
5	(21) 5.85	(22) 5.94	(23) 4.62	(24) 4.33	(25) 5.08	25.82
6	(24) 4.33	(9) 3.99	(19) 5.19	(4) 4.92	(14) 5.80	24.23
7	(7) 5.43	(22) 5.94	(12) 6.06	(17) 5.46	(2) 6.08	28.97
8	(20) 5.63	(25) 5.08	(15) 6.02	(5) 6.72	(10) 6.21	29.66
9	(1) 5.37	(21) 5.85	(6) 6.24	(11) 6.30	(16) 5.24	29.00
10	(23) 4.62	(18) 5.52	(13) 5.20	(3) 5.15	(8) 6.03	26.52

Note, como ejemplo, que el total del tratamiento 11 aparece en los bloques 3 y 9. Este valor es igual a 6.30 toneladas. En forma similar, todos los totales de los tratamientos constan en dos bloques, uno por repetición.

En el Cuadro 4 se presentan los valores  $P$ ,  $B$ ,  $rB$ ,  $D_i$ ,  $R_{nj}$  y  $q^*D_i$ . Como se trata de un látice con dos repeticiones, hay dos valores  $R_n$ . Los valores  $B$  de los totales de los 10 bloques se encuentran en el Cuadro 1. El cálculo para encontrar el factor de ajuste  $q$  se presenta más adelante, en el tercer paso.

Cuadro 4. Cálculo de los valores  $D_i = P-rB$ ,  $R_{nj}$  y  $qD_i$ .  
Ensayo de evaluación de genotipos de triticale

No.Trat	$T_i$	$qD_1$ Repetición I	$qD_2$ Repetición II	$T_{i(ajust)} = T_i + qD_1 +$	$\bar{Y}_{i(ajust)} = T_{i(ajust)}/2$
1	5.37	-0.2140	0.6068	5.76	2.88
2	6.08	-0.2140	0.6068	6.47	3.24
3	5.15	-0.2140	0.6068	5.54	2.77
4	4.92	-0.2140	0.6068	5.31	2.66
5	6.72	-0.2140	0.6068	7.11	3.56
6	6.24	-0.0845	0.0042	6.16	3.08
7	5.43	-0.0845	0.0042	5.35	2.67
8	6.03	-0.0845	0.0042	5.95	2.97

9	3.99	-0.0845	0.0042	3.91	1.95
10	6.21	-0.0845	0.0042	6.13	3.06
11	6.30	-0.2281	0.0817	6.15	3.08
12	6.06	-0.2281	0.0817	5.91	2.96
13	5.20	-0.2281	0.0817	5.05	2.53
14	5.80	-0.2281	0.0817	5.65	2.83
15	6.02	-0.2281	0.0817	5.87	2.94
16	5.24	0.0845	-0.0873	5.24	2.62
17	5.46	0.0845	-0.0873	5.46	2.73
18	5.52	0.0845	-0.0873	5.52	2.76
19	5.19	0.0845	-0.0873	5.19	2.59
20	5.63	0.0845	-0.0873	5.63	2.81
21	5.85	-0.1408	-0.0225	5.69	2.84
22	5.94	-0.1408	-0.0225	5.78	2.89
23	4.62	-0.1408	-0.0225	4.46	2.23
24	4.33	-0.1408	-0.0225	4.17	2.08
25	5.08	-0.1408	-0.0225	4.92	2.46

Las cantidades  $R_{nj}$  corresponden a las sumatorias de los valores  $D_i$  en cada repetición. Note que  $\sum D_i$  es siempre igual a cero; de lo contrario, hay error en los cálculos.

### TERCER PASO

Con la información obtenida se puede calcular la Suma de Cuadrados de Bloques ajustados ( $SCB_{aj}$ ) con la siguiente fórmula:

$$SCB_{aj} = \frac{\sum D_i^2}{kr(r-1)} - \frac{\sum R_{ni}^2}{k^2r(r-1)}$$

$$SCB_{aj} = \frac{(-1.52^2) + (-0.60)^2 + \dots + (-0.16)^2}{5 \cdot 2 \cdot 1} - \frac{(-4.14)^2 + (4.14)^2}{25 \cdot 2 \cdot 1} = \frac{25.98}{10} - \frac{34.28}{50} = 1.9122$$

Luego debe calcularse la Suma de Cuadrados del Error Intra-bloque ( $SCE_{intrbl}$ ):

$$SCE_{intrbl} = SCT - SCR - SCTrat_{s\text{in}aj} - SCB_{aj} \quad (119)$$

$$SCE_{intrbl} = 8.4413 - 0.3428 - 5.0548 - 1.9122 = 1.1315$$

Este valor, ( $SCE_{intrbl}$ ), es la Suma de Cuadrados del Error cuando se efectúa el análisis de varianza como que fuera un BCA sólo con bloques, sin repeticiones.

Como se requiere calcular el factor de ajuste  $q$ , deben encontrarse los cuadrados medios del error intra bloque y de los bloques ajustados. Los grados de libertad son los que se indican en el Cuadro 18.2.2, es decir 16 y 8, respectivamente:

$$CME_{intrbl} = \frac{SCE_{intrbl}}{(k-1)*(rk-k-1)} = \frac{1.1315}{16} = 0.0707 \quad (120)$$

$$CMB_{aj} = \frac{SCB_{aj}}{r*(k-1)} = \frac{1.9122}{8} = 0.2390 \quad (121)$$

Cuando el valor del  $CME_{intrbl}$  es mayor que el del  $CMB_{aj}$ , el valor  $q$  se lo toma como cero y no hace falta realizar ningún ajuste. Si fuera este el caso, se realiza el análisis como si se tratara de un diseño BCA. Sin embargo, en este caso,  $CME_{intrbl} < CMB_{aj}$ , por lo que debe encontrarse el factor de ajuste  $q$ . La fórmula es:

$$q = \frac{CMB_{aj} - CME_{intrbl}}{k(r-1)*CMB_{aj}} \quad (122)$$

$$q = \frac{0.2390 - 0.0707}{5*1*0.2390} = 0.1408$$

Como  $q$  resultó mayor que cero, se computan los productos  $qD_i$  para todos los bloques. Los resultados se exhiben en el Cuadro 4. Observe, en el mencionado cuadro, que los valores  $qD_1$  y  $qD_2$  corresponden a la primera y segunda repetición, respectivamente.

#### CUARTO PASO

A continuación, se procede a calcular los totales ajustados de los tratamientos, como se demuestra en el siguiente cuadro:

Cuadro 5. Cálculo de los totales de los tratamientos ajustados y medias ajustadas

No.Trat	$T_i$	$qD_1$ Repetición I	$qD_2$ Repetición II	$T_{i(ajust)} = T_i + qD_1 +$	$\bar{Y}_{i(ajust)} = T_{i(ajust)}/2$
1	5.37	-0.2140	0.6068	5.76	2.88
2	6.08	-0.2140	0.6068	6.47	3.24
3	5.15	-0.2140	0.6068	5.54	2.77
4	4.92	-0.2140	0.6068	5.31	2.66
5	6.72	-0.2140	0.6068	7.11	3.56
6	6.24	-0.0845	0.0042	6.16	3.08
7	5.43	-0.0845	0.0042	5.35	2.67
8	6.03	-0.0845	0.0042	5.95	2.97
9	3.99	-0.0845	0.0042	3.91	1.95
10	6.21	-0.0845	0.0042	6.13	3.06
11	6.30	-0.2281	0.0817	6.15	3.08
12	6.06	-0.2281	0.0817	5.91	2.96
13	5.20	-0.2281	0.0817	5.05	2.53
14	5.80	-0.2281	0.0817	5.65	2.83
15	6.02	-0.2281	0.0817	5.87	2.94
16	5.24	0.0845	-0.0873	5.24	2.62
17	5.46	0.0845	-0.0873	5.46	2.73
18	5.52	0.0845	-0.0873	5.52	2.76
19	5.19	0.0845	-0.0873	5.19	2.59
20	5.63	0.0845	-0.0873	5.63	2.81
21	5.85	-0.1408	-0.0225	5.69	2.84
22	5.94	-0.1408	-0.0225	5.78	2.89
23	4.62	-0.1408	-0.0225	4.46	2.23
24	4.33	-0.1408	-0.0225	4.17	2.08
25	5.08	-0.1408	-0.0225	4.92	2.46

La Gran Media ajustada es igual a 2.77 toneladas por hectárea. Este valor nos sirve para calcular el Coeficiente de Variación.

Una vez que se han obtenido los totales para los tratamientos ajustados se procede a encontrar la Suma de Cuadrados de los Tratamientos ajustados. La fórmula es la siguiente:

$$SCTrat_{aj} = SCTrat_{sin aj} - Q \quad (123)$$



fórmula en la que:

$$Q = qk(r-1) * \left( \frac{r * SCB_{sinaj}}{(r-1)(1+qk)} - SCB_{aj} \right) \quad (124)$$

En el *primer paso* se encontró que:  $SCTrat_{sinaj} = 5.05$  y  $SCB_{sinaj} = 4.00$ . Los valores de

$SCB_{aj} = 1.9122$  y  $q = 0.1408$  se calcularon en el *tercer paso*. El valor  $k = 5$  es la raíz cuadrada del número de tratamientos y  $r = 2$  el número de repeticiones. Por lo tanto, los valores de  $Q$  y de la  $SCTrat_{aj}$  son:

$$Q = 0.7040 * \left( \frac{8}{1.7040} - 1.9122 \right) = 1.9590$$

#### QUINTO PASO

Se tienen ya todos los elementos necesarios para construir el cuadro ADEVA.

**Cuadro 18.2.8. Análisis de varianza (ADEVA) para el Látice Parcialmente Balanceado; datos de rendimiento de 25 genotipos de triticale**

Fuente de Variación	Suma de Cuadrados	grados de libertad	Cuadrado Medio	F	P	F <sub>.05</sub>	F <sub>.01</sub>
Repeticiones	0.3428	1	0.3428				
Tratamientos (sinajust)	5.0548	24	0.2106	1.66 n.s	11.1%	1.98	2.66
Error (BCA)	3.0437	24	0.1268				
Bloques (ajust)	1.9122	(8)					
Tratamientos (ajust)	3.0959	24	0.1290	1.82 n.s	10.8%	2.24	3.18
Error Intrabloque	1.1315	(16)	0.0707				
TOTAL	8.4413	49					

Puede verse que no se detectaron diferencias significativas, al nivel alfa del 5%, entre las medias de los tratamientos ajustados. En todo caso, a juzgar por los valores de F y las probabilidades respectivas, se ganó en eficiencia con el látice en comparación con el diseño BCA. Esto se puede cuantificar con la siguiente fórmula, donde E representa la eficiencia:

$$E = \frac{CME_{BCA}}{CME_{int rbl} \left( 1 + \frac{r k q}{k+1} \right)} \quad (125)$$

$$E = \frac{0.1268}{0.0707 \left(1 + \frac{1.4080}{6}\right)} = 1.45$$

Es decir que, al conducir el experimento con un látice parcialmente balanceado, la ganancia en eficiencia en comparación con un BCA fue de 45%. Con esta información se puede calcular el coeficiente de variación del látice:

$$CV = \frac{\sqrt{CME_{intrbl}}}{\bar{Y}_{aj}} * 100 = \frac{\sqrt{0.0707}}{2.77} * 100 = 9.60\%$$

### Comparación entre medias

Las medias de los tratamientos ajustados constan en el Cuadro 5. Como los látices fueron planificados para la evaluación de germoplasma en programas de mejoramiento vegetal, los tratamientos son generalmente cualitativos. En los látices parcialmente balanceados para la comparación entre medias se presentan dos casos:

1. Cuando los tratamientos comparten un mismo bloque. En este caso, la fórmula del error estándar para comparar medias ajustadas es:

$$S_{\bar{Y}} = \sqrt{\frac{CME_{intrbl} * [1 + q * (r-1)]}{r}} = \sqrt{\frac{0.0707 * [1 + 0.1408 * (2-1)]}{2}} = 0.2008$$

2. Cuando los tratamientos no comparten bloques. En este caso, el error estándar es mayor que en el caso anterior (lo que implica menor precisión) y se encuentra con la siguiente fórmula:

$$S_{\bar{Y}} = \sqrt{\frac{CME_{intrbl} * (1 + q * r)}{r}} = \sqrt{\frac{0.0707 * (1 + 0.1408 * 2)}{2}} = 0.2128$$

Estas fórmulas para encontrar el error estándar son las adecuadas para todos los látices parcialmente balanceados. Sin embargo, su uso se vuelve complicado y laborioso en la prueba de Tukey, comparaciones entre grupos de medias y otras pruebas. Por esta razón, Cochran y Cox (1968) sugieren el empleo de un promedio ponderado de los errores estándar, excepto en el caso de látices pequeños. La fórmula es:

$$S_{\bar{Y}_P} = \sqrt{\frac{CME_{intrbl} * \left(1 + \frac{qkr}{k+1}\right)}{r}} = \sqrt{\frac{0.0707}{2} * \left(1 + \frac{0.1408 * 5 * 2}{6}\right)} = 0.2089$$

En el experimento de evaluación de 25 genotipos de triticale, el valor crítico de la prueba de Tukey para comparar las medias ajustadas al nivel alfa del 5% es:

$$TUK = q_{\alpha} * S_{\bar{Y}_P} = 6.13 * 0.2089 = 1.28 \text{ toneladas por hectárea.}$$

El valor de  $q_{\alpha}$  para experimentos con un número alto de medias se puede encontrar en varias páginas de internet<sup>44</sup>. En este caso, con 25 medias y 16 grados de libertad del error intra-bloque, se hizo una interpolación entre los valores 6.08 y 6.24, correspondientes a 24 y 28 medias, respectivamente.

---

44 <http://www.real-statistics.com/statistics-tables/studentized-range-q-table/>

## Capítulo 19

# Experimentos factoriales

---

*“Si crees que las estadísticas son aburridas, debes tener los números equivocados”.* Edward R. Tufte<sup>45</sup>

En capítulos anteriores se han visto ejemplos de experimentos con una sola variable independiente. Los diferentes niveles de la variable independiente se denominan “tratamientos”. En los experimentos factoriales hay dos o más variables independientes. Estas variables independientes reciben el nombre de “factores”.

Los experimentos factoriales no son otro tipo de diseño experimental. De hecho, los experimentos factoriales pueden realizarse con diseños Completamente al Azar, Bloques Completos al Azar, Cuadrados Latinos e incluso, con Látices.

### La interacción

Suponga que a un investigador le interesa conocer el efecto de dos densidades de siembra en el rendimiento de cuatro variedades de un cultivo. Podría planificar dos experimentos, en los que evaluaría el rendimiento de las cuatro variedades sembradas con cada una de las densidades. Sin embargo, este plan no le permitiría cuantificar, estadísticamente, si alguna de las variedades respondió en forma diferente al efecto de las densidades. Es decir, no conocería si hubo o no *interacción* entre las dos variables: densidades y variedades.

*Interacción* se define como la respuesta diferente, multiplicativa, de los niveles de un factor con relación a los niveles de otro factor. Puede afirmarse que el objetivo principal de los experimentos factoriales es el estudio de la interacción.

### Designación de los factores y niveles

Por lo general, los factores se designan con letras mayúsculas (A, B, C, etcétera). Los niveles de cada factor se identifican con las letras minúsculas corres-

---

45 Estadístico estadounidense. [https://www.goodreads.com/author/quotes/10775.Edward\\_R\\_Tufte](https://www.goodreads.com/author/quotes/10775.Edward_R_Tufte)

pondientes y números, o sólo números. En un experimento con tres factores, si el factor A tiene tres niveles, estos serán  $a_1$ ,  $a_2$  y  $a_3$ . El término “niveles” se aplica tanto a los factores cuantitativos (por ejemplo, dosis de un insumo) como a los cualitativos (por ejemplo, variedades de un cultivo).

El producto de los números de niveles de los factores es el número de tratamientos. Si se planifica un experimento factorial  $3 \times 2 \times 4$ , el número de tratamientos será 24. Sin embargo, en los experimentos factoriales no interesa el efecto de los tratamientos. De hecho, el procedimiento para el análisis estadístico consiste en partir la suma de cuadrados de los tratamientos en diferentes componentes: sumas de cuadrados de los factores y sumas de cuadrados de las interacciones entre factores.

## Aleatorización

La aleatorización depende del diseño experimental que se empleará. A su vez, el diseño depende de la homogeneidad o no de las unidades experimentales, como se ha visto en capítulos anteriores. En todo caso, la aleatorización se realiza con los tratamientos, cuyo número es el resultado de la combinación de los niveles de los factores.

## Tipos de experimentos factoriales

Si el experimento tiene dos o más repeticiones, el factorial se denomina completo. En experimentos muy grandes, por el número de factores, se puede tener una sola repetición e incluso, fracciones de repetición. La técnica de factoriales fraccionados se emplea generalmente en experimentos exploratorios. Esto es posible porque los experimentos factoriales tienen lo que se conoce como “repetición escondida”. Esto se verá posteriormente en este capítulo.

### 19.1. Experimentos Factoriales Completos

Como se indicó, un factorial completo tiene dos o más repeticiones. El siguiente ejemplo de un factorial completo ilustra lo que es la interacción y las precauciones que deben tomarse en la interpretación de los resultados.

---

**Ejercicio 19.1.1.** Se condujo un experimento con un diseño completamente al azar (DCA), en un arreglo factorial  $2 \times 2$  con tres repeticiones. Los dos factores en estudio fueron:

Factor A: Variedades de arroz, con dos niveles a1 y a2 (una variedad alta y otra semi-enana).

Factor B: Fertilizante nitrogenado, con dos niveles b1 y b2 (los equivalentes de 50 kg/ha y 100 kg/ha de nitrógeno).

El esquema para el análisis de varianza se presenta en el siguiente cuadro:

**Cuadro 1. Esquema del ADEVA de un DCA con tres repeticiones, dos factores y dos niveles por factor**

FV	SC	GL
TRAT	<i>SCTrat</i>	$(k-1) = 3$
A	<i>SCA</i>	$(a-1) = 1$
B	<i>SCB</i>	$(b-1) = 1$
A*B	<i>SC(A*B)</i>	$(a-1)*(b-1) = 1$
Error	<i>SCE</i>	$(n-k) = 8$
TOTAL	<i>SCT</i>	$(n-1) = 11$

Nota: k es el número de tratamientos y n es el número total de observaciones en el experimento

Los datos de rendimiento, en quintales métricos por hectárea, se presentan en el Cuadro 2. Los tratamientos, en número de cuatro, resultan de la combinación de los niveles de los dos factores. Cada tratamiento se repitió tres veces en el ensayo.

Se pide encontrar los efectos principales de los factores y la interacción entre ellos. Discutir los resultados que se obtengan.

**Cuadro 2. Experimento factorial: dos variedades de arroz, a1 y a2, y dos dosis de nitrógeno, b1 y b2. Resultados en q/ha.**

Trat	Factores		Repeticiones			$T_i$
	A	B				
1	a1	b1	45	50	40	135
2	a1	b2	20	15	25	60
3	a2	b1	18	20	18	56
4	a2	b2	50	52	40	142
$\sum y_i$						393

En todo experimento factorial debe partirse la Suma de Cuadrados de los Tratamientos en sus componentes. En el caso de dos factores hay tres componentes: la Suma de Cuadrados del Factor A, la Suma de Cuadrados del Factor B y la Suma de Cuadrados de la Interacción A\*B.

### Análisis

1. Lo primero que debe hacerse es encontrar las fuentes de variación del diseño empleado, en este caso, un DCA. Las sumas de cuadrados son:

$$SCT = 2356.25 \quad SCTrat = 2170.92 \quad SCE = 185.33$$

2. A continuación, para partir la Suma de Cuadrados de Tratamientos en sus componentes, se forma un cuadro de doble entrada. A este cuadro, para simplificar, se lo llama Cuadro A B:

Cuadro 3. Cuadro A B para dos variedades y dos dosis de fertilizante nitrogenado

	Variedades=A		
Nitrógeno=B	Alta, a1	Enana, a2	Totales $B_j$
50 kg/ha, b1	135	56	191
100 kg/ha, b2	60	142	202
Totales $A_i$	195	198	393

Los cuatro valores en el centro del cuadro A B (*en itálicas*) son los totales de los cuatro tratamientos. Con los totales de los factores se procede a encontrar las sumas de cuadrados respectivas:

$$SCA = \frac{\sum A_i^2}{r*b} - FC = \frac{195^2 + 198^2}{3*2} - \frac{393^2}{12} = 0.75$$

$$SCB = \frac{\sum B_j^2}{r*a} - FC = \frac{191^2 + 202^2}{3*2} - \frac{393^2}{12} = 10.08$$

La Suma de Cuadrados de la Interacción A\*B es igual a:

$$SC(A*B) = SCTrat - SCA - SCB = 2170.92 - 0.75 - 10.08 = 2160.09$$

Con los resultados obtenidos se procede a estructurar el Cuadro ADEVA:

Cuadro 4. Análisis de varianza del experimento factorial  
(diseño completamente al azar)

F.V.	S.C.	g. l.	C.M	F	$F_{05}$	$F_{01}$
Tratamientos	2,170.92	3				
Factor A	0.75	1	0.75	< 1 ns	5.32	11.26
Factor B	10.08	1	10.08	< 1 ns		
interacción A*B	2,160.09	1	2,160.09	93.24 **		
Error	185.33	8	23.16			
Total	2,356.25	11				

Los resultados del Cuadro 4 parecen indicar que no hay diferencia significativa entre los niveles del factor A. Tampoco hay, aparentemente, diferencia significativa entre los niveles del factor B. Sin embargo, se observa una interacción altamente significativa.

No se podría concluir el análisis con lo señalado en el párrafo anterior. Se requiere de un examen algo más detenido para extraer toda la información posible del experimento. La regla básica es que, antes de decir algo con relación a los factores individuales, se debe tratar de explicar el porqué de una interacción estadísticamente significativa. Para esto es importante un cuadro similar al Cuadro 3, pero ahora con las medias:

Cuadro 5. Medias de los efectos de los factores y de la interacción

Nitrógeno=B	Variedades=A		<b>Medias <math>\bar{Y}_b</math></b>
	Alta, a1	Enana, a2	
50 kg/ha, b1	$a1b1 = 45.00$	$a2b1 = 18.67$	$b1 = 31.83$
100 kg/ha, b2	$a1b2 = 20.00$	$a2b2 = 47.33$	$b2 = 33.67$
<b>Medias <math>\bar{Y}_a</math></b>	$a1 = 32.50$	$a2 = 33.00$	$\bar{Y} = 32.75$

Puede verse en el Cuadro 5 que la media de a1, igual a 32.50, difiere muy poco de la media de a2, igual a 33.00. Eso es lo que indica el resultado estadísticamente no significativo que consta en el cuadro del ADEVA (Cuadro 3). Igual comentario puede hacerse para las medias de b1 y b2. Sin embargo, cuando se consideran las medias de la interacción, es decir, cuando se combinan los niveles de un factor con el otro, los resultados son completamente diferentes. Como ejemplo, la media de la variedad alta a1, en combinación con el nivel b1 de nitrógeno, fue



45 quintales por hectárea. La misma variedad a 1, con el nivel 2 de nitrógeno, tuvo una media de 20 quintales por hectárea. Una diferencia que contribuye a la interacción altamente significativa, detectada en el Cuadro 4.

La mejor manera para explicar una interacción es a través de un gráfico, como el que se presenta a continuación:

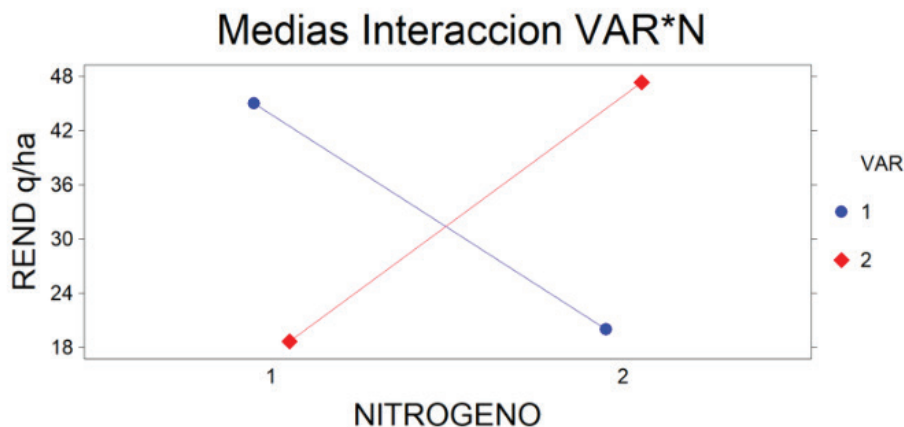


Gráfico 1. Interacción entre variedades de arroz y niveles de Nitrógeno

Este ejemplo ilustra lo que es la interacción. Apresuradamente se podría concluir que las dos dosis de fertilizante tienen igual efecto y que no hay diferencia entre las dos variedades. Pero esto no es lo que se puede ver en la Figura 1. Las respuestas de las dos variedades a las dosis del fertilizante son totalmente diferentes. Desde el punto de vista agronómico se puede indicar que la variedad alta (1) no tolera dosis altas de fertilizante, las plantas se acaman y eso determina una baja producción al nivel alto de nitrógeno. La variedad semi-enana (2), por el contrario, hace buen uso del nitrógeno disponible, gracias a sus tallos más cortos y gruesos.

**Comentario:** Cuando no hay interacción, las líneas en el gráfico son estadísticamente paralelas. En una interacción tan alta como la del ejemplo anterior, las líneas que representan a las variables incluso llegan a cruzarse. Sin embargo, puede haber interacción estadísticamente significativa sin que las líneas necesariamente se crucen. De estos casos se verá en ejemplos posteriores.

## Un experimento con tres factores

En el caso de experimentos con tres factores A, B y C hay cuatro interacciones: AB, AC, BC, y ABC. Las interacciones entre dos factores se llaman también “interacciones de primer grado” y la interacción entre tres factores, “interacción de segundo grado”.

Un ejemplo de un experimento factorial con tres factores se presenta en el siguiente ejercicio:

**Ejercicio 19.1.2.** En un experimento conducido con un diseño BCA se evaluó el comportamiento de dos variedades de arroz en respuesta a cuatro diferentes dosis de nitrógeno y cuatro de fósforo (Suárez, 1990)<sup>46</sup>. Las variedades fueron Cica 8 y Cuyamel 3820 (Factor VAR). Las dosis o niveles de nitrógeno fueron las cantidades equivalentes a 0, 50, 100 y 150 kilogramos por hectárea (Factor N). Los niveles de fósforo fueron 0, 40, 80 y 120 kilogramos por hectárea (Factor P). Esto da como resultado un experimento factorial  $2 \times 4 \times 4$ . Es decir, 32 tratamientos. Los resultados, en toneladas por hectárea, se presentan en el siguiente cuadro:

**Cuadro 1. Tratamientos, niveles, repeticiones y rendimientos del experimento factorial  $2 \times 4 \times 4$  conducido con un diseño BCA**

Trat	VAR	N	P	R1, t/ha	R2, t/ha	Trat	VAR	N	P	R1, t/ha	R2, t/ha
1	1	1	1	5.80	3.80	17	2	1	1	3.46	4.69
2	1	1	2	9.18	4.03	18	2	1	2	3.85	3.76
3	1	1	3	4.56	6.19	19	2	1	3	2.83	3.52
4	1	1	4	5.53	3.65	20	2	1	4	3.74	4.48
5	1	2	1	8.43	5.92	21	2	2	1	5.54	5.13
6	1	2	2	7.33	8.38	22	2	2	2	5.78	5.49
7	1	2	3	7.96	6.35	23	2	2	3	4.64	3.99
8	1	2	4	6.86	6.98	24	2	2	4	4.74	4.52
9	1	3	1	8.47	7.98	25	2	3	1	5.64	4.14
10	1	3	2	8.94	8.97	26	2	3	2	4.88	5.53

46 Suárez, G.P. 1990. Evaluación de niveles de fertilización, densidades de siembra y uso de herbicidas en dos variedades de arroz. Tesis. Escuela Agrícola Panamericana. Honduras, C.A.

11	1	3	3	8.94	6.53	27	2	3	3	7.25	5.56
12	1	3	4	8.59	6.97	28	2	3	4	5.23	5.86
13	1	4	1	7.69	7.79	29	2	4	1	5.80	5.61
14	1	4	2	7.43	8.76	30	2	4	2	4.76	5.03
15	1	4	3	7.08	9.37	31	2	4	3	5.21	4.91
16	1	4	4	9.46	7.20	32	2	4	4	3.74	5.54

Las variedades (VAR) de arroz fueron Cica 8 y Cuyamel 3820 (valores 1 y 2 en el Cuadro 1). Los niveles de nitrógeno (N) seleccionados fueron las cantidades equivalentes por parcela a 0, 50, 100 y 150 kilogramos por hectárea. En el Cuadro 1 estos valores están codificados como 1, 2, 3 y 4. El mismo caso con el tercer factor niveles de fósforo (P). Los valores reales de 0, 40, 80 y 120 kilogramos de  $P_2O_5$  están codificados como 1, 2, 3 y 4 en el Cuadro 1. Las dos repeticiones constan como R1 y R2.

Cuadro 2. Esquema del ADEVA

FV	SC	GL
REP	<i>SCR</i>	$(r-1) = 1$
TRAT	<i>SCTrat</i>	$(k-1) = 31$
V	<i>SCV</i>	$(v-1) = 1$
N	<i>SCN</i>	$(n-1) = 3$
P	<i>SCP</i>	$(p-1) = 3$
V*N	<i>SC(V*N)</i>	$(v-1)*(n-1) = 3$
V*P	<i>SC(V*P)</i>	$(v-1)*(p-1) = 3$
N*P	<i>SC(N*P)</i>	$(n-1)*(p-1) = 9$
V*N*P	<i>SC(V*N*P)</i>	$(v-1)*(n-1)*(p-1) = 9$
Error	<i>SCE</i>	$(r-1)*(k-1) = 31$
TOTAL	<i>SCT</i>	$(k*r)-1 = 63$

Nota: **r** es el número de repeticiones, **k** es el número de tratamientos, **v** el número de variedades, **n** el número de niveles de nitrógeno y **p**, el número de niveles de fósforo.

Con la información contenida en el Cuadro 1, se pueden encontrar las sumas de cuadrados de todas las fuentes de variación del Cuadro 2. Por conveniencia, primero se calcula el Factor de Corrección (FC) y luego, la Suma de Cuadra-

dos Total (SCT), la Suma de Cuadrados de las Repeticiones (SCR), la Suma de Cuadrados de los Tratamientos (SCTrat) y la Suma de Cuadrados del Error (SCE):

$$FC = \frac{(\sum Y_{ij})^2}{r*k} = \frac{(385.97)^2}{2*32} = 2327.70$$

$$SCT = \sum Y_{ij}^2 - FC = (5.81^2 + 9.18^2 + \dots + 5.54^2) - 2327.70 = 201.00$$

$$SCR = \frac{\sum R_j^2}{k} - FC = \frac{199.34^2 + 186.63^2}{32} - 2327.70 = 2.52$$

$$SCTrat = \frac{\sum T_i^2}{r} - FC = \frac{(5.80 + 3.80)^2 + (9.18 + 4.03)^2 + \dots + (3.74 + 5.54)^2}{2} - FC = 160.92$$

$$SCE = SCT - SCR - SCTrat = 201.00 - 2.52 - 160.92 = 37.56$$

Los siguientes cálculos conducen a la determinación de las sumas de cuadrados de los factores y de las interacciones entre factores. Para esto es necesario construir tres cuadros de doble entrada, los cuadros V\*N, V\*P y N\*P.

Cuadro 3. Cuadro de doble entrada V\*N

Variedad	Nitrógeno				Totales $V_v$
	n1	n2	n3	n4	
v1	42.74	58.21	65.39	64.78	231.12
v2	30.33	39.83	44.09	40.60	154.85
Totales $N_n$	73.07	98.04	109.48	105.38	$\sum Y_{ij} = 385.97$

Con los datos del centro del Cuadro 3 (*datos en itálicas*) se calcula la Suma de Cuadrados del Cuadro V\*N, las sumas de cuadrados de los factores en el cuadro y de la interacción V\*N:

$$SC_{cuadro}(V*N) = \frac{\sum (v_v n_n)^2}{r*p} - FC = \frac{42.74^2 + 58.21^2 + \dots + 40.60^2}{2*4} - FC = 145.56$$

$$SCV = \frac{\sum V_v^2}{r*n*p} - FC = \frac{231.12^2 + 154.85^2}{2*4*4} - FC = 90.89$$

$$SCN = \frac{\sum N_n^2}{r*v*p} - FC = \frac{73.7^2 + 98.04^2 + 109.48^2 + 105.38^2}{2*2*4} - FC = 49.92$$

Con la información anterior se procede a encontrar la Suma de Cuadrados de la Interacción V\*N. Esta estadística es igual a:

$$SC(V*N) = SC_{cuadro}(V*N) - SCV - SCN = 145.56 - 90.89 - 49.92 = 4.75$$

Luego, se forma el cuadro V\*P:

Cuadro 4. Cuadro de doble entrada V\*P

Variedad	Fósforo (P)				Totales $V_v$
	p1	p2	p3	p4	
v1	55.88	63.02	56.98	55.24	231.12
v2	40.01	39.08	37.91	37.85	154.85
Totales $P_p$	95.89	102.10	94.89	93.09	$\sum Y_{ij} = 385.97$

Con los datos del centro del Cuadro 4 (*datos en itálicas*) se calcula la Suma de Cuadrados del Cuadro V\*P, la suma de cuadrados del factor P (fósforo) y la suma de cuadrados de la interacción V\*P:

$$SC_{cuadro}(V*P) = \frac{\sum (v_p p_p)^2}{r*n} - FC = \frac{55.88^2 + 63.02^2 + \dots + 37.85^2}{2*4} - FC = 96.06$$

$$SCP = \frac{\sum P_p^2}{r*v*n} - FC = \frac{95.89^2 + 102.10^2 + 94.89^2 + 93.09^2}{2*2*4} - FC = 2.87$$

$$SC(V*P) = SC_{cuadro}(V*P) - SCV - SCP = 96.06 - 90.89 - 2.87 = 2.30$$

Para encontrar la suma de cuadrados de la interacción N\*P se requiere armar el cuadro N\*P:

Cuadro 5. Cuadro de doble entrada N\*P

Nitrógeno	P				Totales $N_n$
	p1	p2	p3	p4	
n1	17.75	20.82	17.10	17.40	73.07
n2	25.02	26.98	22.94	23.10	98.04
n3	26.23	28.32	28.28	26.65	109.48
n4	26.89	25.98	26.57	25.94	105.38
Totales $P_p$	95.89	102.10	94.89	93.09	$\sum Y_{ij} = 385.97$

$$SC_{\text{cuadro}}(N^*P) = \frac{\sum(n_n * p_p)^2}{r * \eta} - FC = \frac{17.75^2 + 20.82^2 + \dots + 25.94^2}{2 * 2} - FC = 55.90$$

$$SC(N^*P) = SC_{\text{cuadro}}(N^*P) - SC_N - SC_P = 55.90 - 49.92 - 2.87 = 3.11$$

Solo falta calcular la suma de cuadrados de la interacción entre los tres factores. Esta estadística se la encuentra al restar de la suma de cuadrados de los tratamientos todas las sumas de cuadrados de los factores y todas las sumas de cuadrados de las interacciones entre dos factores:

$$SC(V^*N^*P) = SC_{\text{Trat}} - SC_V - SC_N - SC_P - SC(V^*N) - SC(V^*P) - SC(N^*P) = 7.09$$

Se tiene ya toda la información necesaria para construir el cuadro ADEVA:

**Cuadro 6. Análisis de Varianza. Experimento factorial con un diseño BCA**

FV	SC	GL	CM	F	P	F <sub>.05</sub>	F <sub>.01</sub>
REP	2.52	1	2.52				
TRAT	(160.92)	(31)	-				
V	90.89	1	90.89	75.11**	0*10 <sup>-5</sup>	4.16	7.53
N	49.92	3	16.64	13.75**	1*10 <sup>-5</sup>	2.82	4.49
P	2.87	3	0.96	0.79 ns	0.50	2.82	4.49
V*N	4.75	3	1.58	1.31 ns	0.28	2.82	4.49
V*P	2.30	3	0.77	0.64 ns	0.59	2.82	4.49
N*P	3.11	9	0.35	0.29 ns	0.97	2.20	3.05
V*N*P	7.09	9	0.79	0.65 ns	0.74	2.20	3.05
Error	37.56	31	1.21				
TOTAL	201.00	63					

Los valores F en el cuadro ADEVA se obtienen al dividir los cuadrados medios de cada uno de los factores e interacciones para el cuadrado medio del error experimental. Al comparar los valores F calculados con los valores críticos de la Tabla F, puede verse que ninguna interacción resultó estadísticamente significativa, es decir, la probabilidad es por lo menos mayor que 0.05 ( $P \geq 0.05$ ). Las probabilidades para cada caso, obtenidas con el programa *Statistix*, también constan en el cuadro ADEVA.

Los factores, variedades y fertilizante nitrogenado, resultaron significativos ( $P < 0.01$ ). Los niveles de fósforo, por el contrario, no tuvieron efecto alguno sobre el rendimiento; resultado que puede atribuirse a que el terreno experimental tuvo suficiente cantidad de ese nutrimento.

El análisis de varianza generado con el programa *Statistix* presenta resultados similares. Se incluye también, como referencia, las medias de las variedades y del fertilizante nitrogenado:

**Statistix 9.0**

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
REP	1	2.524	2.5241		
VAR	1	90.892	90.8924	75.03	0.0000
N	3	49.917	16.6390	13.73	0.0000
P	3	2.872	0.9573	0.79	0.5086
VAR*N	3	4.745	1.5816	1.31	0.2902
VAR*P	3	2.299	0.7662	0.63	0.5996
N*P	9	3.105	0.3449	0.28	0.9740
VAR*N*P	9	7.087	0.7875	0.65	0.7460
Error	31	37.556	1.2115		
Total	63	200.996			

Grand Mean 6.0308 CV 18.25

**Means of REND for N**

N	Mean
0	4.5669
50	6.1275
100	6.8425
150	6.5862

**Means of REND for VAR**

VAR	Mean
1	7.2225
2	4.8391

A pesar de que la interacción VAR\*N resultó no significativa, un gráfico ayuda a interpretar en mejor forma las respuestas de las variedades con relación a los niveles de nitrógeno:

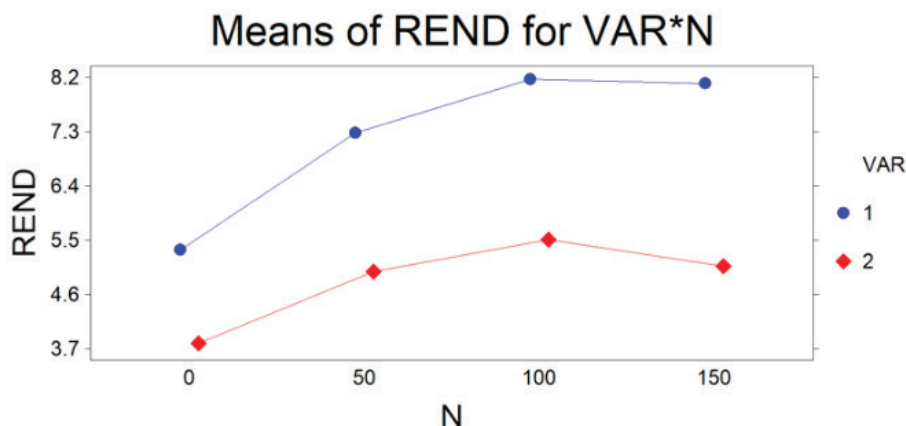


Gráfico 1. Efecto de nitrógeno en el rendimiento de dos variedades de arroz

Del gráfico anterior se concluye que las curvas de respuesta de las dos variedades de arroz fueron curvas cuadráticas, o parábolas, similares. Es decir, no hubo interacción significativa. Sin

embargo, el rendimiento de las dos variedades fue claramente diferente. Las medias de Cica 8 y de Cuyamel 3820 fueron: 7.2 y 4.8 toneladas métricas por hectárea, respectivamente.

En el caso de las dosis de nitrógeno, como se vio en el capítulo 17 para los tratamientos o factores cuantitativos, el método adecuado para estudiar su efecto es el de polinomios ortogonales. El cálculo de las comparaciones ortogonales se presenta en el Cuadro 7:

Cuadro 7. Comparaciones ortogonales y sumas de cuadrados

Niveles de Nitrógeno, kg/ha	0	50	100	150			
Totales $T_i$ , t/ha	73.07	98.04	109.48	105.38			
Respuesta	Coeficientes, $C_i$				$Q_i$	$D_i$	$SCQ_i$
Lineal	-3	-1	+1	+3	108.37	320	36.70 **
Cuadrática	+1	-1	-1	+1	-29.07	64	13.20**
Cúbica	-1	+3	-3	+1	-2.01	320	0.01n.s.

Nota: Los totales de los tratamientos se encuentran en el Cuadro 3 de este ejercicio. El número de repeticiones ( $r$ ) de cada nivel de nitrógeno es 16. Los coeficientes vienen del Cuadro 17.5.1 (capítulo 17).



Es bueno recordar que:  $Q_i = \sum c_i * T_i$ ,  $D_i = r \sum c_i^2$ ,  $SCQ_i = \frac{Q_i^2}{D_i}$ .

Como cada comparación tiene un solo grado de libertad, las sumas de cuadrados son iguales a los cuadrados medios. El valor F en la tabla, con 1 y 31 grados de libertad, es 7.53 al 0.01 nivel de probabilidad. Por lo tanto, la respuesta lineal y la cuadrática son altamente significativas.

Con el programa *Statistix* los resultados de las comparaciones ortogonales son obviamente iguales:

**Polynomial Contrasts of REND by N**

Degree = 1, Linear Trend			
Contrast	1.5145	SS (Contrast)	36.700
T-Statistic	5.50	P (T-Statistic)	0.0000
SE (Contrast)	0.2752		

Degree = 2, Quadratic Trend			
Contrast	-0.9084	SS (Contrast)	13.204
T-Statistic	-3.30	P (T-Statistic)	0.0024
SE (Contrast)	0.2752		

Degree = 3, Cubic Trend			
Contrast	-0.0281	SS (Contrast)	0.0126
T-Statistic	-0.10	P (T-Statistic)	0.9193
SE (Contrast)	0.2752		

Error term used: REP\*VAR\*N\*P, 31 DF

Al observar las probabilidades de la prueba t, se ratifica que sólo las respuestas lineal y cuadrática fueron significativas.

### **Estimación del máximo rendimiento**

Cuando se realizan experimentos para evaluar respuestas a diferentes dosis de insumos lo adecuado es determinar el punto y sector de la parábola en los que se maximiza la producción. Sin embargo, desde el punto de vista económico, es posible que, por la diferencia entre el ingreso por la venta del producto y el costo del insumo, el ingreso obtenido al nivel del máximo rendimiento sea menor que algún otro punto en la curva. A esto se llama “el óptimo económico”.

Para estimar el máximo rendimiento y el óptimo económico se requiere, primero, encontrar la ecuación de regresión cuadrática. Luego, por cálculo diferencial, se encuentra el valor de  $X_{max}$ , es decir, la cantidad del insumo con el que se maximiza el rendimiento. En el siguiente ejercicio se presenta un ejemplo de esto:

**Ejercicio 19.1.3.** Con los datos del Ejercicio 19.1.2, encuentre el valor de  $X_{max}$  con el que se obtiene el rendimiento máximo estimado  $\hat{Y}_{max}$ .

La ecuación cuadrática codificada, cuando se tiene cuatro niveles de un insumo, se vio en la Sección 17.2. Esta ecuación es:  $\hat{Y} = \bar{Y} + b_1X + b_2 \frac{(X^2-5)}{4}$ . Con la información obtenida en el ejercicio anterior se procede a reemplazar en la ecuación los valores correspondientes.

La media del experimento fue 6.0308. Los coeficientes de regresión lineal y cuadrática se encuentran a partir del Cuadro 7 del ejercicio anterior:

$$b_1 = \frac{Q_L}{D_L} = \frac{108.37}{320} = 0.3387, \quad b_2 = \frac{Q_C}{D_C} = \frac{-29.07}{64} = -0.4542$$

Al reemplazar los valores en la ecuación de regresión y luego, al simplificar, se tiene:

$$\hat{Y} = 6.0308 + 0.3387X - 0.4542 \frac{(X^2-5)}{4}, \quad \hat{Y} = 6.5986 + 0.3387X - 0.1136X^2.$$

Note que, en la última ecuación, el valor de la intersección sustituye al valor de la media. La ecuación es entonces:  $\hat{Y} = b_0 + b_1X + b_2X^2$ .

### *Estimación del Máximo Rendimiento*

Mediante cálculo diferencial se puede encontrar, a partir de la ecuación  $\hat{Y} = b_0 + b_1X + b_2X^2$  un valor estimado de  $X$  con el que se obtiene el máximo rendimiento  $\hat{Y}$ . Para esto, se obtiene la derivada de  $\hat{Y}$  sobre  $X$  y se iguala a cero:

$d\hat{Y}/dX = b_1 + 2b_2X = 0$ . Al despejar  $X$  se tiene:

$$X = -b_1/2b_2.$$

Luego, se sustituyen en la ecuación los coeficientes  $b_1$  y  $b_2$ . El valor de  $X_{max}$  es:

$$X_{max} = -0.3387/2 * (-0.1136) = 1.4907$$

Al reemplazar el valor de  $X_{\max}$  en la ecuación cuadrática, se obtiene el valor de  $\hat{Y}_{\max}$  en toneladas de arroz por hectárea:

$$\hat{Y} = 6.5986 + 0.3387X - 0.1136X^2,$$

$$\hat{Y} = 6.5986 + 0.3387(1.4907) - 0.1136(1.4907^2) = 6.85 \text{ t/ha}$$

El máximo rendimiento estimado es, entonces, 6.85 t/ha. A continuación, es importante conocer con qué cantidad de fertilizante nitrogenado  $X$  se obtiene este rendimiento.

La variable  $X$  está en la escala codificada de los coeficientes lineales  $C_i$ . Estos son: -3, -1, +1, +3 (ver Cuadro 7 del ejercicio anterior); por lo que deben transformarse a los valores originales del experimento: 0, 50, 100 y 150 kilogramos de nitrógeno por hectárea. La fórmula para decodificar toma en cuenta el intervalo real (50 kg/ha), el intervalo codificado (esto es: 2) y la media de los niveles de nitrógeno originales (75 kg):

$$X_{real} = \frac{I_{real} * X_{cod}}{I_{cod}} + \bar{X}_{real} = \frac{50 * 1.49}{2} + 75 = 112.25 \text{ kg/ha}$$

Es decir que, con alrededor de 112.25 kg/ha de fertilizante nitrogenado, se obtiene un rendimiento de 6.85 t/ha de arroz. Este es un valor estimado de la máxima producción posible, en las condiciones con las que se llevó a efecto el experimento.

**Comentario:** en el ejercicio anterior, la información viene de un experimento factorial con un diseño BCA. Sin embargo, vale aclarar que la estimación del máximo rendimiento puede realizarse con cualquier tipo de diseño y con datos de experimentos con un factor o varios factores. El único requisito es, obviamente, que la variable independiente sea cuantitativa.

El procedimiento del ejercicio 19.1.3 ha sido y es generalmente empleado en muchos trabajos por razón de que simplifica los cálculos. Sin embargo, con varios programas estadísticos puede trabajarse directamente con los valores reales. Como ejemplo, el programa *Statistix* genera el análisis completo y presenta un gráfico con la ecuación de regresión. Esto se vio en la sección 10.12 sobre regresión polinómica. Veamos el siguiente ejercicio:

**Ejercicio 19.1.4.** Con la misma información del Ejercicio 19.1.2, pero con los datos reales de  $X$ , encuentre mediante el programa Statistix el valor de  $X$  con el que se maximiza  $\hat{Y}$ .

La ecuación cuadrática es:  $\hat{Y} = b_0 + b_1X + b_2X^2$ . El primer paso es encontrar los coeficientes  $b_0$ ,  $b_1$ ,  $b_2$ :

Con Statistix 9

Polynomial Regression of REND

Predictor Variables	Original Scale Coefficient	Std Error
Constant	4.56059	0.38349
N	0.04080	0.01232
N^2	-1.817E-04	7.869E-05

Note que estos coeficientes se generan a partir de los datos reales de  $X$ , es decir, sin codificación:  $X_1 = 0$ ,  $X_2 = 50$ ,  $X_3 = 100$ ,  $X_4 = 150$ . Esto es, kilogramos de fertilizante nitrogenado por hectárea. Con la información de los coeficientes podemos escribir la ecuación de regresión cuadrática:

$$\hat{Y} = 4.5606 + 0.0408X - 1.8170 \cdot 10^{-4}X^2$$

El programa Statistix también genera un gráfico de la respuesta cuadrática que incluye la ecuación:

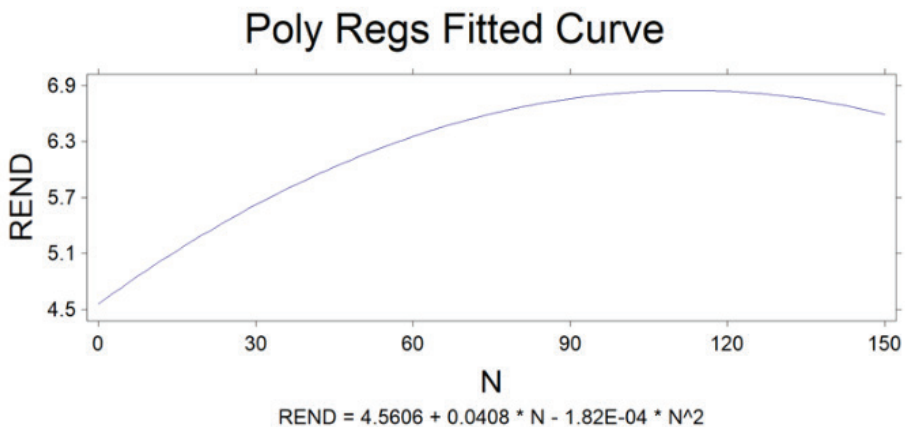


Gráfico 1. Respuesta cuadrática y ecuación de la regresión polinómica

A partir de la ecuación del Gráfico 1, obtenida con los datos reales, se puede estimar el máximo rendimiento, como se hizo con la ecuación codificada. Pero primero se requiere determinar el valor de  $X$  con el que se maximiza  $\hat{Y}$ . Mediante cálculo diferencial se obtiene el valor de  $X$ :

$$X = -b_1/2b_2$$

$$X = -0.0408/2*(-1.8170*10^{-4}) = 112.27 \text{ kg/ha}$$

Al sustituir el valor de  $X$  en la ecuación de regresión se tiene:

$$\hat{Y} = 4.5605 + 0.0408X - 1.8170*10^{-4}X^2$$

$$\hat{Y} = 4.5605 + 0.0408*112.27 - 1.8170*10^{-4}*112.27^2 = 6.85 \text{ t/ha}$$

Observe que, tanto el valor de  $X$  como el valor de  $\hat{Y}$  son idénticos a los obtenidos en el ejercicio 19.1.3, salvo por errores de redondeo.

### **Comentario:**

Los resultados obtenidos en los ejercicios anteriores corresponden al promedio del rendimiento de las dos variedades. Sin embargo, en el análisis de varianza del ejercicio 19.2 se observó que la variedad Cica 8 tuvo un rendimiento significativamente superior que el de la otra variedad. Por lo tanto, más importante para los productores sería conocer la cantidad de fertilizante nitrogenado con la que se obtiene el máximo rendimiento con la variedad Cica 8.

---

**Ejercicio 19.1.5.** Con la información del Ejercicio 19.1.2 determine la cantidad de fertilizante nitrogenado con el que se obtiene una estimación del máximo rendimiento de la variedad Cica 8.

Para lograr lo planteado debe realizarse el análisis tan solo con la mitad de los datos del Cuadro 1 del Ejercicio 19.1.2. Esto es, con 32 observaciones que corresponden a los tratamientos del 1 al 16 (los de la variedad Cica 8) y a las dos repeticiones. El Cuadro ADEVA generado con Statistix se presenta a continuación:

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
REP	1	5.5945	5.5945		
N	3	41.6623	13.8874	7.40	0.0029
P	3	4.7699	1.5900	0.85	0.4896
N*P	9	3.4321	0.3813	0.20	0.9897
Error	15	28.1690	1.8779		
Total	31	83.6278			

Grand Mean 7.2225    CV 18.97

Como no hubo interacción N\*P y el efecto del fósforo (P) no fue significativo es válido incluirlos en el error experimental. El nuevo análisis se presenta a continuación:

**Randomized Complete Block AOV Table for REND**

Source	DF	SS	MS	F	P
REP	1	5.5945	5.5945		
N	3	41.6623	13.8874	10.31	0.0001
Error	27	36.3710	1.3471		
Total	31	83.6278			

Grand Mean 7.2225    CV 16.07

**Means of REND for N**

N	Mean
0	5.3425
50	7.2762
100	8.1738
150	8.0975

Observations per Mean            8  
 Standard Error of a Mean        0.4103

Observe que el efecto del nitrógeno (N) fue altamente significativo. Los tres grados de libertad del N pueden descomponerse mediante el procedimiento de contrastes polinómicos ortogonales, como se presenta a continuación con el programa Statistix:

**Polynomial Contrasts of REND by N**

Degree = 1, Linear Trend			
Contrast	2.0488	SS (Contrast)	33.581
T-Statistic	4.99	P (T-Statistic)	0.0000
SE (Contrast)	0.4103		

Degree = 2, Quadratic Trend			
Contrast	-1.0050	SS (Contrast)	8.0802
T-Statistic	-2.45	P (T-Statistic)	0.0211
SE (Contrast)	0.4103		

Degree = 3, Cubic Trend			
Contrast	0.0140	SS (Contrast)	1.56E-03
T-Statistic	0.03	P (T-Statistic)	0.9731
SE (Contrast)	0.4103		

Error term used: Error, 27 DF

Lo anterior demuestra que únicamente las tendencias lineal y cuadrática fueron significativas. Mediante los métodos de regresión, empleando los valores reales de  $X$  (0, 50, 100, 150), se puede calcular los coeficientes, encontrar el ADE-VA de Regresión, generar un gráfico de la tendencia y determinar la ecuación cuadrática. Con *Statistix* se obtienen los siguientes resultados:

**Polynomial Regression of REND**

Predictor Variables	Original Scale Coefficient	Std Error
Constant	5.34563	0.41455
N	0.04848	0.01331
N <sup>2</sup>	-2.010E-04	8.506E-05

R-Squared 0,4982    Resid. Mean Square (MSE) 1.44714  
 Adjusted R-Squared 0,4636    Standard Deviation 1.20297

Source	DF	SS	MS	F	P
Regression	2	41.6608	20.8304	14,39	0,0000
Residual	29	41.9670	1.4471		
Total	31	83.6278			

Lack of Fit 1    0.0016    0.00156    0,00    0,9745  
 Pure Error 28    41.9655    1.49877  
 Cases Included 32    Missing Cases 0

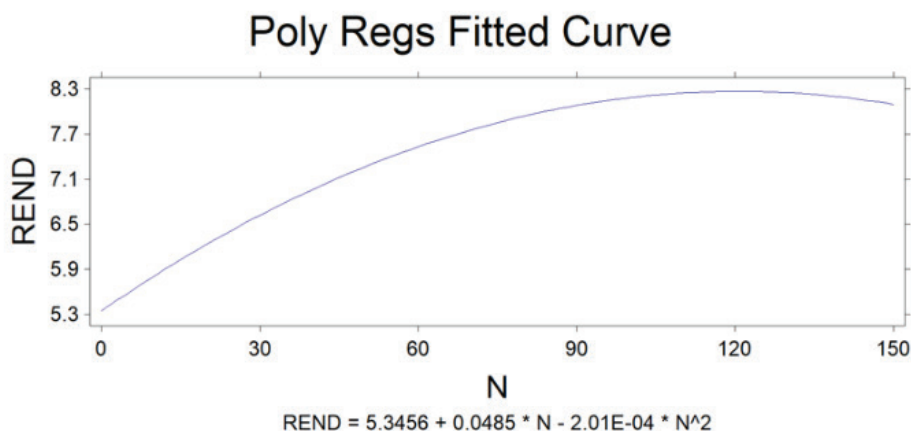


Gráfico 1. Respuesta de la variedad de arroz Cica 8 a diferentes niveles de nitrógeno

Ahora se puede estimar el máximo rendimiento a partir de la ecuación:

$$\hat{Y} = b_0 + b_1X + b_2X^2 = 5.3456 + 0.0485X - 2.01 \cdot 10^{-4}X^2$$

Mediante cálculo diferencial se tiene que:  $X = b_1 / -2b_2 = 120.60$  kg/ha

Al reemplazar el valor de  $X$  en la ecuación:

$$\hat{Y} = 5.3456 + 0.0485 \cdot 120.60 - 2.01 \cdot 10^{-4} \cdot 120.60^2 = 8.27 \text{ t/ha}$$

Es decir que, con 120.60 kg/ha de fertilizante nitrogenado se espera, con la variedad de arroz Cica 8, en las condiciones del experimento, un rendimiento máximo estimado de alrededor de 8.27 t/ha.

### *Intervalo de Confianza*

El valor de 8.27 t/ha es una estimación puntual de la media poblacional. En todo caso, como se trata de una estadística, este valor es una variable. En capítulos anteriores se presentó el concepto y aplicaciones del intervalo de confianza. Un intervalo de confianza proporciona un límite inferior y un límite superior, dentro de los cuales se puede señalar, con una determinada probabilidad, que se encuentra la media poblacional de  $\hat{Y}$ , es decir  $\mu_{\hat{Y}}$ . Con el programa Statistix se calculan los siguientes límites con el intervalo de confianza que se desee, en este caso con el 95% de probabilidad:



**Predicted/Fitted Values of REND** Statistix 9.0

Lower Fitted Bound	7.6783	LÍMITE INFERIOR, 95%
Fitted Value	8.2683	
Upper Fitted Bound	8.8583	LÍMITE SUPERIOR, 95%

SE (Fitted Value) 0.2885

Percent Coverage 95,0

Corresponding T 2,05

Predictor Values: N = 120.60

Este resultado indica que, si se mantienen las mismas condiciones del experimento, se puede afirmar con un 95% de confianza, que futuros rendimientos de la variedad Cica 8 estarán entre 7.67 y 8.85 t/ha.

En el gráfico siguiente, Statistix presenta los límites superior e inferior dentro de los cuales se encuentran los rendimientos, para cualquier valor de  $X$ , con un 95% de confianza.

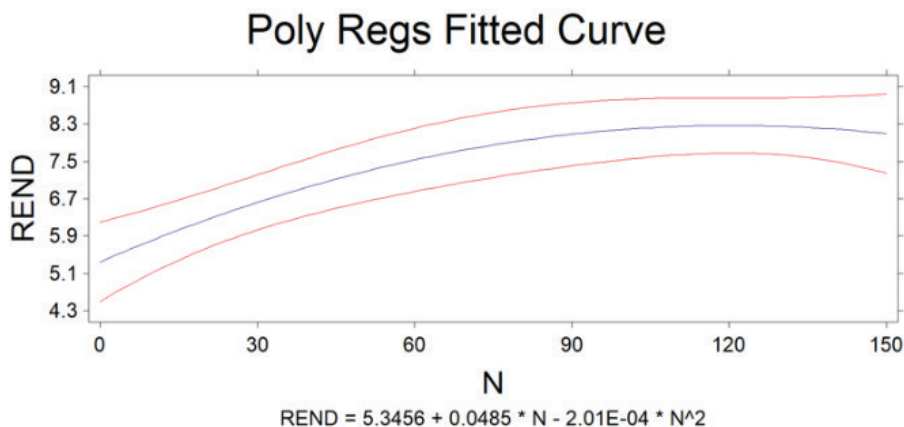


Gráfico 2. Intervalo de confianza al 95% del rendimiento de la variedad Cica 8

**Estimación del Óptimo Económico**

En el ejercicio anterior, el fertilizante nitrogenado representa el único costo variable. Todos los demás se consideran costos fijos, para facilidad del ejemplo. El mayor retorno económico no necesariamente ocurre cuando el nivel de

$X$  genera la mayor producción posible. Esto es, en el ejemplo anterior, cuando  $X = 120.60 \text{ kg/ha}$  de fertilizante nitrogenado y  $\hat{Y} = 8.27 \text{ t/ha}$  de arroz. Todo depende del precio de venta del producto y del costo del insumo.

Para determinar la cantidad del insumo con la que se logra el mayor ingreso económico, la derivada de  $Y$  sobre  $X$  debe igualarse a la tasa del costo del insumo sobre el precio de venta del producto. En el siguiente ejercicio se presenta un ejemplo:

**Ejercicio 19.1.6.** Con la información del ejercicio anterior, determine la cantidad de fertilizante nitrogenado con la que se obtiene el rendimiento más rentable (“óptimo económico”). Suponga un costo por kilogramo de nitrógeno de \$2.00 y un precio de venta del arroz de \$250.00 por tonelada:

Se parte de la misma ecuación:

$$\hat{Y} = b_0 + b_1X + b_2X^2 = 5.3456 + 0.0485X - 2.01 \cdot 10^{-4}X^2$$

La derivada de  $Y$  sobre  $X$  debe igualarse a la tasa del costo del insumo sobre el costo del producto:

$$dY/dX = b_1 + 2b_2X = 2.00/250.00 = 0.008, \text{ de donde } X \text{ es igual a:}$$

$$X = (0.008 - b_1)/2b_2 = \frac{(-0.0405)}{2 \cdot (-0.000201)} = 100.75 \text{ kg/ha}$$

Al reemplazar en la ecuación el valor de  $X$ , se espera un rendimiento de:

$$\hat{Y}_{opt} = b_0 + b_1X + b_2X^2 = 5.3456 + 0.0485 \cdot (100.75) - 0.000201 \cdot (100.75)^2 = 8.19 \text{ t/ha.}$$

También se puede, como en el caso anterior, encontrar intervalos de confianza para el valor  $\hat{Y}_{opt}$  mediante el programa Statistix 9.0:

**Predicted/Fitted Values of REND**

Lower Fitted Bound	7.5468	LIMITE INFERIOR, 95%
Fitted Value	8.1892	
Upper Fitted Bound	8.8317	LIMITE SUPERIOR, 95%
SE (Fitted Value)	0.3141	

Percent Coverage 95,0

Corresponding T 2,05

Predictor Values: N = 100.75

## 19.2. Experimentos factoriales con una repetición

### *El concepto de “Repetición Escondida”*

El concepto de “repetición escondida” se refiere al hecho de que, en un experimento factorial, un nivel de cualquier factor se presenta varias veces en una misma repetición. Esto posibilita que se pueda realizar un ensayo con una sola repetición, lo que equivale a un Diseño Completamente al Azar. Como no se tiene un error experimental que provenga de la interacción entre factores y repeticiones, el investigador puede decidir que algunas interacciones, usualmente las de mayor grado, se conviertan en el error experimental. Esto se justifica porque las interacciones de alto grado son generalmente no significativas y producto del azar.

Veamos el siguiente ejercicio:

---

**Ejercicio 19.2.1.** Se realizó un experimento factorial para estudiar el efecto de tres factores en el rendimiento de maíz. Los factores fueron: A, dos densidades de siembra; B, cuatro variedades de maíz y C, cinco niveles de fertilización.

Por el elevado número de tratamientos se planificó llevar a cabo el experimento con una sola repetición. En este experimento factorial  $2 \times 4 \times 5$ , con 40 tratamientos, se marcó un número igual de unidades experimentales en un terreno homogéneo y se asignaron los tratamientos a las unidades experimentales mediante un sorteo totalmente al azar. Como hubo una sola repetición, el investigador decidió que la interacción  $A \times B \times C$  sea el error experimental. Es decir, le interesan todas las interacciones entre dos factores y los efectos principales. Los resultados del rendimiento de maíz en kilogramos por parcela se presentan en el Cuadro 1.

Como se indicó, este experimento factorial tiene una sola repetición de los 40 tratamientos. Sin embargo, puede verse en el Cuadro 1, que cada nivel de A se repite 20 veces, cada nivel de B se repite 10 veces y cada nivel de C se repite ocho veces. Esto es lo que se conoce como “repetición escondida”.

Para el análisis estadístico debe seguirse el mismo esquema que el empleado en el Ejercicio 19.1.2. Es decir, encontrar el Factor de Corrección, la Suma de Cuadrados de los tratamientos y construir los cuadros de doble entrada  $A \times B$ ,  $A \times C$  y  $B \times C$ . Luego, se calculan las sumas de cuadrados de los factores y de las interacciones. La interacción  $A \times B \times C$  se constituye en el error experimental. Para los cálculos, en este ejercicio, se toma en cuenta lo siguiente:

$k = 40$ ,  $r = 1$ ,  $a = 2$ ,  $b = 4$ ,  $c = 5$ . Estos son, respectivamente, el número de tratamientos, el número de repeticiones y los números de niveles de cada factor.

Cuadro 1. Rendimiento de maíz en kg por parcela en respuesta a tres factores

Trat #	A	B	C	REND
1	1	1	1	10.4
2	1	1	2	10.6
3	1	1	3	15.3
4	1	1	4	14.5
5	1	1	5	12.1
6	1	2	1	8.3
7	1	2	2	7.5
8	1	2	3	9.6
9	1	2	4	13.2
10	1	2	5	12.8
11	1	3	1	12.3
12	1	3	2	14.6
13	1	3	3	15.1
14	1	3	4	10.4
15	1	3	5	10.2
16	1	4	1	13.5
17	1	4	2	12.4
18	1	4	3	16.6
19	1	4	4	13.2
20	1	4	5	10.8
Trat #	A	B	C	REND
21	2	1	1	8.2
22	2	1	2	9.3
23	2	1	3	15.5
24	2	1	4	10.1
25	2	1	5	7.6
26	2	2	1	15.4
27	2	2	2	16.2
28	2	2	3	13.7
29	2	2	4	9.4
30	2	2	5	7.9
31	2	3	1	16.7
32	2	3	2	18.4
33	2	3	3	19.8
34	2	3	4	16.4
35	2	3	5	14.7
36	2	4	1	7.5
37	2	4	2	12.3
38	2	4	3	16.5
39	2	4	4	12.8
40	2	4	5	10.6

El análisis obtenido con el programa Statistix es el siguiente:

Analysis of Variance Table for REND

Source	DF	SS	MS	F	P
A	1	6.084	6.0840	0.95	0.3478
B	3	80.792	26.9307	4.23	0.0296
C	4	90.641	22.6603	3.56	0.0391
A*B	3	80.724	26.9080	4.22	0.0296
A*C	4	24.676	6.1690	0.97	0.4601
B*C	12	46.573	3.8811	0.61	0.7987

Error	12	76.466	6.3722
Total	39	405.956	
Grand Mean	12.560	CV 20.10	

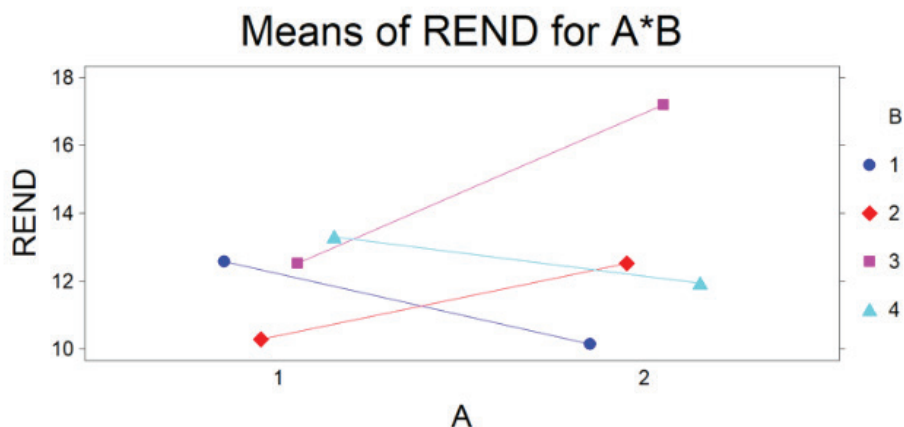


Gráfico 1. Interacción entre densidades (factor A) y variedades (factor B)

En los resultados del análisis de varianza se observa que hubo una interacción significativa entre densidades (factor A) y variedades (factor B). Puede verse en el Gráfico 1 que el rendimiento de las variedades 2 y 3 se incrementó con la densidad 2. En cambio, el rendimiento de las variedades 1 y 4 disminuyó por efecto de la densidad 2. Esto puede atribuirse a diferentes características morfológicas y genéticas de las variedades. En todo caso, la variedad 3 fue la que mejor rindió con la densidad 2, mientras que, con la densidad 1, el rendimiento fue similar para las variedades 1, 3 y 4. Esto se puede corroborar con las medias respectivas.

La fertilización (factor C) no interactuó con los otros factores, pero tuvo un efecto significativo en el rendimiento, como puede verse en el cuadro ADEVA ( $P = 0.0391$ ). Por tratarse de una variable cuantitativa, lo adecuado es determinar las tendencias mediante el método de polinomios ortogonales. Como los grados de libertad del factor C son cuatro se podría estimar hasta una tendencia cuártica, pero lo que generalmente interesa, en ensayos de rendimiento de variedades, es la respuesta cuadrática. Con Statistix se obtiene lo siguiente:

Statistix 9.0 Factorial una rep,

**Polynomial Contrasts of REND by C**

Degree = 1, Linear Trend

Contrast	-0.4941	SS (Contrast)	1.9531
T-Statistic	-0.55	P (T-Statistic)	0.5900
SE (Contrast)	0.8925		

Degree = 2, Quadratic Trend

Contrast	-2.9232	SS (Contrast)	68.359
T-Statistic	-3.28	P (T-Statistic)	0.0066
SE (Contrast)	0.8925		

Error term used: A\*B\*C, 12 DF

Como puede observarse, la tendencia cuadrática es significativa. El siguiente gráfico corrobora estos resultados:

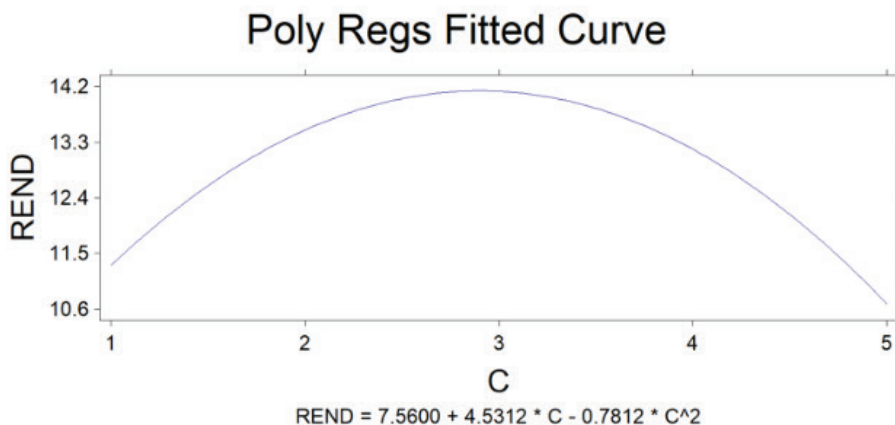


Gráfico 2. Respuesta cuadrática a diferentes niveles de fertilización

A partir de esta información se podrían obtener los valores esperados de máximo rendimiento y del punto de mayor rentabilidad.

### *Quando varias interacciones forman el error*

En otros casos de experimentos factoriales con una sola repetición, el investigador puede decidir que varias interacciones se constituyan en el error experimental.

Como ejemplo, En un experimento factorial con cuatro factores ( $2^2 \times 2^2 = 2^4$ ) y una sola repetición, el investigador puede decidir que las interacciones

entre tres y cuatro factores no son de interés y más bien constituirse en el error experimental. (Por experiencia se conoce que por lo general estas interacciones no son estadísticamente significativas y su inclusión en el error mejora el análisis.) El esquema del ADEVA se presenta en el cuadro siguiente:

Cuadro 19.2.1. Esquema del ADEVA de un factorial 2<sup>4</sup> con una sola repetición en el que el error experimental lo forman las interacciones de 2do y 3er grado

Fuente de variación	Grados de libertad
A	1
B	1
A*B	1
C	1
A*C	1
B*C	1
D	1
A*D	1
B*D	1
C*D	1
Error: (A*B*C + A*B*D + A*C*D + B*C*D + A*B*C*D)	5
TOTAL	15

En el ejercicio a continuación se presenta este caso:

**Ejercicio 19.2.2.** Suponga un experimento factorial con cuatro factores A, B, C, D y dos niveles de cada factor. El experimento consta de una sola repetición, por lo que el investigador decide que las interacciones entre tres y cuatro factores se constituyan en el Error Experimental. Realice el Análisis de Varianza con los siguientes datos:

Cuadro 1. Datos de un experimento factorial 2<sup>4</sup>

A	B	C	D	REND
1	1	1	1	12
2	1	1	1	15
1	2	1	1	22
2	2	1	1	16
1	1	2	1	21
2	1	2	1	15
1	2	2	1	19

2	2	2	1	20
1	1	1	2	22
2	1	1	2	24
1	2	1	2	25
2	2	1	2	15
1	1	2	2	18
2	1	2	2	16
1	2	2	2	23
2	2	2	2	24

Statistix 9.0

#### Analysis of Variance Table for REND

Source	DF	SS	MS	F	P
A	1	18.062	18.0625	0.79	0.4138
B	1	27.562	27.5625	1.21	0.3213
C	1	1.563	1.5625	0.07	0.8038
D	1	45.562	45.5625	2.00	0.2163
A*B	1	7.562	7.5625	0.33	0.5893
A*C	1	1.562	1.5625	0.07	0.8038
B*C	1	7.562	7.5625	0.33	0.5893
A*D	1	0.062	0.0625	0.00	0.9602
B*D	1	3.062	3.0625	0.13	0.7288
C*D	1	14.062	14.0625	0.62	0.4675
Error	5	113.812	22.7625		
Total	15	240.437			

Grand Mean 19.188    CV 24.87

Nota: Este análisis sirve sólo como ejemplo. Un experimento, con tan pocos grados de libertad para el error, por lo general carece de precisión.

#### Comentarios:

- Si se dispone de un programa como Statistix, los experimentos con una repetición o fracciones de repetición (que se verán más adelante) se analizan con facilidad. Hay dos opciones en Statistix con las que se pueden realizar estos análisis: 1. En el menú “modelos lineales”, “análisis de varianza”,



“experimentos factoriales” y 2. En el menú “modelos lineales”, “análisis de varianza”, “AOV/AOCV”.

- Con la primera opción sólo se puede incluir en el error todas las interacciones de un mismo tamaño. En el caso de la segunda opción (Análisis de Varianza/Análisis de Covarianza) puede incluirse en el error cualquier interacción. Para esto se requiere escribir el comando en el que constan los nombres de los factores y las interacciones cuyos efectos se desean. Las interacciones que no se incluyen, por defecto, se constituyen en el error experimental.

### Número de interacciones entre factores

En la planificación previa de experimentos factoriales con una repetición o fracciones de repetición es conveniente conocer el número de interacciones que se generará entre factores. Esta información es importante para decidir qué interacciones formarán el error experimental y cuántos grados de libertad tendrá el error. Esto además facilita la preparación de un esquema del análisis de varianza y, si se emplea un programa computarizado, la definición del modelo del análisis.

La fórmula general para el número de interacciones (NI) es:

$$NI = \frac{n!}{r!(n-r)!} \quad (126)$$

Fórmula en la que:

- El número de factores se representa con la letra n,
- El número de interacciones entre un determinado número de factores es r,
- El signo de factorial es (!). Es importante recordar que factorial de cero es igual a uno:  $0! = 1$ .

Ejemplos:

1. Si se tienen cuatro factores, como en el caso del Ejercicio 19.3.1, ¿Cuántas interacciones se generan entre dos, tres y cuatro factores?:

a. entre dos factores:  $NI = \frac{n!}{r!(n-r)!} = \frac{4!}{2!*2!} = \frac{4*3*2*1}{2*1*2*1} = 6$

b. entre tres factores:  $NI = \frac{n!}{r!(n-r)!} = \frac{4!}{3!*1!} = \frac{4*3*2*1}{3*2*1*1} = 4$

c. entre cuatro factores:  $NI = \frac{n!}{r!(n-r)!} = \frac{4!}{4!*0!} = \frac{4*3*2*1}{4*3*2*1*1} = 1$

2. En un experimento con seis factores, todos con dos niveles (esto es  $2^6$ ) y una sola repetición, en el que las interacciones entre cuatro, cinco y seis factores formarán el error experimental, ¿Cuál es el número de interacciones entre dos y tres factores y cuántos grados de libertad tendrá el error? Presente un esquema del ADEVA para este caso:

a. entre dos factores: 
$$NI = \frac{n!}{r!(n-r)!} = \frac{6!}{2!*4!} = \frac{6*5*4*3*2*1}{2*1*4*3*2*1} = 15$$

b. entre tres factores: 
$$NI = \frac{n!}{r!(n-r)!} = \frac{6!}{3!*3!} = \frac{6*5*4*3*2*1}{3*2*1*3*2*1} = 20$$

- c. El total de grados de libertad es  $2^n - 1 = 63$ . Efectos principales = 6, interacciones entre dos factores = 15, interacciones entre tres factores = 20. Es decir,  $6 + 15 + 20 = 41$ . Por lo tanto, el número de grados de libertad para el error será  $63 - 41 = 22$ . (Debe notarse que, como todos los factores tienen dos niveles, el número de grados de libertad de los efectos principales es uno en todos los casos. Consecuentemente, las interacciones, sin importar el grado, también tendrán únicamente un grado de libertad.)

Lo indicado se resume en el Cuadro 19.2.2:

**Cuadro 19.2.2. Esquema del ADEVA  
de un experimento factorial  $2^6$  con una repetición**

FUENTE DE VARIACIÓN	GL
Efectos principales (A, B, C, D, E, F)	6
Interacciones de 2 factores (AB, AC, etc.)	15
Interacciones de 3 factores ABC, ABD, etc.)	20
Interacciones de 4 factores (15), de 5 factores (6), de 6 factores (1)	22 (error)
TOTAL	63

Para facilitar el análisis, en una etapa inicial de una investigación, podría presumirse que incluso las interacciones entre tres factores no son significativas, por lo que pasarían a formar parte del error experimental.

Otra forma para encontrar los números de efectos principales e interacciones es a partir de los coeficientes de la expansión del binomio de Newton, esto es:  $(a + b)^n$ , en la que  $n$  es igual al número de factores. Por ejemplo, si se tienen seis factores, al resolver el binomio  $(a + b)^6$  los coeficientes serán, a partir del segundo término: 6, 15, 20, 15, 6, 1. Es decir, los mismos que constan en el Cuadro 19.2.2.

### 19.3. Experimentos factoriales fraccionados

Por razón de la presencia de repeticiones escondidas en los experimentos factoriales, no sólo que se puede tener experimentos con una sola repetición, sino que también experimentos con una fracción de repetición. Esto es lo que se conoce como “experimentos factoriales fraccionados”. Los experimentos factoriales fraccionados son especialmente usados en *experimentos exploratorios*, tanto en el laboratorio como en el campo. El emplear sólo una fracción de repetición se vuelve necesario en los experimentos exploratorios porque generalmente el número de tratamientos es muy elevado. En estos experimentos el error experimental lo constituyen una o varias interacciones, usualmente del orden más alto, según decida el investigador.

#### *Experimentos exploratorios 2<sup>n</sup>*

Cuando se introducen nuevos métodos, nuevas tecnologías, nuevas variedades o especies, de las que se tiene poca información, se recomienda el empleo de *experimentos exploratorios* en los que se evalúan varios factores a la vez, cada factor con dos niveles. Esto se representa con  $2^n$ , en que el 2 indica el número de niveles de todos los factores y la  $n$  indica el número de factores. En estos casos el investigador está interesado en obtener información rápida sobre qué factores e interacciones de primer orden son más importantes. De esta forma, en próximos experimentos, podrá descartar algunos factores y aumentar el número de niveles de los que tuvieron respuestas significativas. La teoría y análisis de los experimentos factoriales exploratorios están explicados en varios textos (Kempthorne, 1979; Cochran y Cox, 1968; Petersen, 1985).

#### *Bloques incompletos*

Se conoce que el Cuadrado Medio, o Varianza del Error experimental (CME), está directamente correlacionado con el tamaño de la repetición o bloque completo. Por lo general, a mayor tamaño de la repetición, mayor heterogeneidad entre unidades experimentales y, consecuentemente, un CME más alto. Por esta razón es conveniente trabajar con *bloques incompletos*, que son áreas más pequeñas e idealmente homogéneas. El concepto de bloques incompletos se vio en el Capítulo 18, que versó sobre diseños en látices.

Es importante recordar que en el Diseño de Bloques Completos al Azar (BCA), los términos “bloque” y “repetición” son sinónimos. En los diseños de bloques incompletos, “bloque” y “repetición” no significan lo mismo. Una repetición

contiene todos los tratamientos. Un bloque incompleto contiene sólo parte de los tratamientos. Varios bloques incompletos pueden formar una repetición.

### *Número de tratamientos*

Si se planificara un experimento factorial exploratorio con ocho factores, todos con dos niveles, el número de tratamientos sería igual a  $2^8 = 256$ . Incluso con una sola repetición, el experimento sería demasiado grande y muy difícil de manejar. La solución a este problema es trabajar tan solo con una fracción de repetición. El investigador tendrá que decidir, de acuerdo con los recursos a su disposición, el tamaño de la fracción que empleará.

En el caso anterior, como ejemplo, podría pensarse en un experimento tan solo con la octava parte de los tratamientos, es decir:  $\frac{1}{8} * 2^8 = 32$  tratamientos. Si se tuviera un área homogénea, los tratamientos se asignarían al azar a 32 unidades experimentales. Sin embargo, en experimentos de campo es difícil encontrar áreas homogéneas grandes. En este evento, el investigador podría conducir el experimento en dos bloques incompletos, cada uno con 16 tratamientos, o incluso, en cuatro bloques incompletos con ocho tratamientos cada uno. Este procedimiento es también práctico en experimentos de laboratorio, en los que por limitaciones de equipo y materiales podría llevarse a cabo el experimento en diferentes períodos: un bloque incompleto por período.

En los experimentos factoriales fraccionados, qué tratamientos van en un determinado bloque incompleto no es una selección al azar. Debe existir ortogonalidad entre los grupos de tratamientos en los diferentes bloques y esto se logra mediante un proceso que se describirá más adelante.

### *Representación de los tratamientos*

Como hemos visto, los factores en todo experimento factorial se representan con letras mayúsculas, usualmente A, B, C, etc. Para el análisis computarizado de los datos es necesario crear variables que representen a los factores y etiquetar, en la base de datos, los niveles de cada factor con números. La combinación de los números, en cada caso, indica un tratamiento específico. Como ejemplo, en un experimento factorial con cuatro factores, todos con dos niveles ( $2^4$ ), la combinación 1, 1, 1, 1 identifica al tratamiento uno. Este es el tratamiento en el que, por lo general, todos los factores están en el nivel más bajo. La combinación 2, 2, 2, 2 identifica al tratamiento 16, aquel en que todos los factores están en el nivel más alto.

En los experimentos factoriales fraccionados  $2^n$ , para determinar los tratamientos que se asignarán a los bloques incompletos se emplea otro sistema de designación de los tratamientos. Como todos los factores están en dos niveles, basta indicar con una letra minúscula la presencia del nivel alto de ese factor. La ausencia de la letra indica que ese factor está en el nivel bajo. El tratamiento en que todos los factores están en el nivel bajo se representa con el número 1 entre paréntesis: (1). Esta es la forma tradicional como se designan los tratamientos en los experimentos factoriales  $2^n$ . Como ejemplo, en un experimento factorial  $2^4$ , es decir 16 tratamientos, se designan los tratamientos como se presenta en el Cuadro 19.3.1.

### *Asignación de los tratamientos a los bloques incompletos*

Para decidir qué tratamientos irán en los bloques incompletos, el “método tradicional” (Cuadro 19.3.1) de designación de los tratamientos es el más indicado. La selección de los tratamientos que se asignan a los bloques incompletos se basa en un proceso similar al de los contrastes ortogonales (Sección 17.1.3), en que los signos del contraste ortogonal que escoja el investigador (generalmente una interacción alta) definen qué tratamientos van en cada bloque. Esto quiere decir que los efectos de los bloques estarán fusionados con el efecto de la interacción escogida. En otras palabras, estos efectos son indistinguibles. Otro término que se usa es el de “efectos confundidos”, con la acepción en español de “confundido” como sinónimo de “mezclado” o “fusionado”. A la interacción que se emplea para dividir en dos grupos los tratamientos se la llama *contraste definidor*.

Cuadro 19.3.1. Designación de los tratamientos en un experimento factorial  $2^4$

Trat No.	Método para Ordenadores				Método Tradicional	Trat No.	Método para Ordenadores				Método Tradicional
	A	B	C	D			A	B	C	D	
1	1	1	1	1	(1)	9	1	1	1	2	D
2	2	1	1	1	a	10	2	1	1	2	Ad
3	1	2	1	1	b	11	1	2	1	2	Bd
4	2	2	1	1	ab	12	2	2	1	2	Abd
5	1	1	2	1	c	13	1	1	2	2	Cd
6	2	1	2	1	ac	14	2	1	2	2	Acd
7	1	2	2	1	bc	15	1	2	2	2	Bcd
8	2	2	2	1	abc	16	2	2	2	2	Abcd

Debe tenerse en cuenta que en el caso de factores cualitativos (como variedades, prácticas culturales, etc.) el investigador decide arbitrariamente cuál “nivel” es el alto y cuál es el bajo.

Suponga, por facilidad de la explicación, un experimento factorial  $2^3$ , es decir con  $2 \times 2 \times 2$  tratamientos. Si se decide tener cuatro bloques incompletos, con cuatro tratamientos cada uno, se tendrá un experimento factorial fraccionado  $\frac{1}{2} \times 2^3 = 4$ , con dos repeticiones. Sin embargo, los tratamientos de cada bloque deben estar balanceados con relación a los efectos principales de los factores y de las interacciones.

Para comprender cómo se seleccionan los tratamientos que irán en cada bloque es necesario estructurar un cuadro, en el que consten los factores e interacciones en una columna y los tratamientos en la primera hilera. Luego, en forma similar a lo que se hace con los contrastes ortogonales, escribimos los coeficientes que nos permiten encontrar los efectos principales y las interacciones del experimento. Como ejemplo, en el caso del factor A se designa con -1 los tratamientos en los que no consta letra “a” (minúscula) y con +1 los tratamientos en los que “a” se encuentra presente. Note que la sumatoria de los coeficientes de cada hilera debe ser igual a cero. Para encontrar la interacción AB, por ejemplo, multiplicamos, en orden, los coeficientes de la hilera de A por los de la hilera de B. Se procede en forma similar hasta llenar todas las casillas. Un ejemplo de esto se presenta en el Cuadro 19.3.2 para el caso de un experimento factorial  $2^3$ :

**Cuadro 19.3.2. Método para dividir los tratamientos de un experimento factorial  $2^3$  en dos bloques incompletos, de cuatro tratamientos cada uno**

Factores e interacciones	Tratamientos y coeficientes								Sumatoria de Coeficientes
	(1)	a	b	ab	c	ac	bc	abc	
A	-1	+1	-1	+1	-1	+1	-1	+1	0
B	-1	-1	+1	+1	-1	-1	+1	+1	0
AB	+1	-1	-1	+1	+1	-1	-1	+1	0
C	-1	-1	-1	-1	+1	+1	+1	+1	0
AC	+1	-1	+1	-1	-1	+1	-1	+1	0
BC	+1	+1	-1	-1	-1	-1	+1	+1	0
ABC	-1	+1	+1	-1	+1	-1	-1	+1	0

Como se decidió que el efecto de la interacción de más alto grado, esto es ABC, sea el contraste definidor, separamos los tratamientos con los signos po-

sitivos y negativos en la hilera de la interacción ABC. Se forman así dos grupos ortogonales de tratamientos, Grupo 1: (1), ab, ac, bc y Grupo 2: a, b, c, abc. Es decir, cada bloque tendrá 4 parcelas (unidades experimentales) y los tratamientos seleccionados se asignarán por sorteo a cada parcela dentro del bloque. Como la interacción ABC, llamada en este caso el contraste definidor, está confundida con bloques, no será posible tener una estimación de esa interacción. Sin embargo, cualquier factor u otra interacción pudo ser el contraste definidor. Se pudo, como ejemplo, seleccionar la interacción BC. Si hubiera sido así, al observar los signos del cuadro anterior para BC, los dos grupos de tratamientos seleccionados hubieran sido, Grupo uno: (1), a, bc, abc y Grupo dos: b, ab, c, ac.

### *Análisis de varianza de los experimentos factoriales fraccionados*

Para el análisis de varianza de los experimentos factoriales fraccionados hay varias opciones:

1. Se pueden formar cuadros de doble entrada, como se hizo en el Ejercicio 19.1.2. A partir de cada cuadro se calculan las sumas de cuadrados de los efectos principales de dos factores y de su interacción. Sin embargo, si se tiene un experimento factorial  $2^5$  se requiere de 10 cuadros de doble entrada, lo que demanda considerable esfuerzo. Para un experimento factorial  $2^6$  habrá que estructurar 15 cuadros.
2. Mediante el empleo de un procedimiento similar al descrito en la sección 17.1.3, en el que se calculan las sumas de cuadrados de los contrastes ortogonales. En este procedimiento se multiplican, para cada caso, los totales de los tratamientos por los coeficientes y se suman los productos. Este método se empleará en el Ejercicio 19.3.1.
3. Otra opción es utilizar el método desarrollado por Yates y que se conoce como el “Algoritmo de Yates”. Este método está descrito en varios libros (Cochran y Cox, 1968; Petersen, 1985). Es un método de cálculo manual poco usado actualmente.

En el siguiente ejercicio se presenta un ejemplo del análisis de varianza de un experimento factorial fraccionado con la segunda opción, en el que la interacción confundida o contraste definidor es ABC:

---

**Ejercicio 19.3.1.** Datos de rendimiento y análisis de varianza para el supuesto caso de un experimento factorial fraccionado  $\frac{1}{2} * 2^3 = 4$ , con dos repeticiones y cuatro bloques incompletos, en el que el efecto de la interacción ABC está confundido con las repeticiones.

Cuadro 1. Datos de rendimiento. Experimento factorial fraccionado  $\frac{1}{2} * 2^3$  con dos repeticiones y cuatro bloques.

Rep	Blq	A	B	C	Rend	Trat	Rep	Blq	A	B	C	Rend	$T_i$
1	1	0	0	0	12	(1)	2	3	0	0	0	8	20
1	1	1	1	0	13	ab	2	3	1	1	0	10	23
1	1	1	0	1	5	ac	2	3	1	0	1	7	12
1	1	0	1	1	14	bc	2	3	0	1	1	11	25
1	2	1	0	0	15	a	2	4	1	0	0	17	32
1	2	0	1	0	6	b	2	4	0	1	0	8	14
1	2	0	0	1	11	c	2	4	0	0	1	13	24
1	2	1	1	1	9	abc	2	4	1	1	1	9	18

En el Cuadro 2, mediante el procedimiento de contrastes ortogonales, se pueden calcular las sumas de cuadrados.

Cuadro 2. Cálculo de las sumas de cuadrados de los factores e interacciones

F.V.	Totales de los Tratamientos								$Q_i$	$\frac{Q_i^2}{r \sum c_i^2}$
	(1)	a	b	ab	c	ac	bc	abc		
A	-20	+32	-14	+23	-24	+12	-25	+18	2	0.25
B	-20	-32	+14	+23	-24	-12	+25	+18	-8	4.00
AB	+20	-32	-14	+23	+24	-12	-25	+18	2	0.25
C	-20	-32	-14	-23	+24	+12	+25	+18	-10	6.25
AC	+20	-32	+14	-23	-24	+12	-25	+18	-40	100.00
BC	+20	+32	-14	-23	-24	-12	+25	+18	22	30.25
ABC	-20	+32	+14	-23	+24	-12	-25	+18	8	4.00

Note en el Cuadro 2 que, los signos negativos o positivos de los coeficientes ortogonales, todos de valor 1, se asignan de acuerdo con el procedimiento indicado para el Cuadro 19.3.2. Los productos de los coeficientes por los totales de los tratamientos constan en las hileras del Cuadro 2 de este ejercicio. El contraste lineal  $Q_i$  es el resultado de la sumatoria de los productos. Como ejemplo, en el caso del factor A:

$$Q_i = \sum c_i T_i = -20 + 32 - 14 + \dots + 18 = 2$$

$$\text{La suma de cuadrados del factor A es igual a: } SCA = \frac{Q_i^2}{r \sum c_i^2} = \frac{4}{16} = 0.25$$



La sumatoria de los cuadrados de los coeficientes  $\sum c_i^2$  es igual a 8 para todos los contrastes. Cada total de los tratamientos fue el resultado de la suma de dos cantidades, es decir que  $r = 2$

El Cuadro ADEVA completo, obtenido con Statistix, se presenta a continuación:

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
BLQ	3	16.500	5.500		
A	1	0.250	0.250	0.12	0.7409
B	1	4.000	4.000	1.92	0.2152
C	1	6.250	6.250	3.00	0.1340
A*B	1	0.250	0.250	0.12	0.7409
A*C	1	100.000	100.000	48.00	0.0004
B*C	1	30.250	30.250	14.52	0.0089

Error 6 12.500 2.083  
 Total 15 170.000  
 Grand Mean 10.500 CV 13.75

En este ejercicio, los cuatro bloques forman dos repeticiones (REPS), aunque no constan explícitamente en el ADEVA. El efecto de los bloques es indistinguible de los efectos de las REPS y del contraste definidor o interacción ABC. En otras palabras, los efectos están confundidos, es decir, fusionados. Los tres grados de libertad de bloques provienen de lo siguiente:

- un grado de libertad por REPS,
- un grado de libertad por el contraste definidor ABC y
- un grado de libertad por la interacción REPS\*ABC.

En cambio, los grados de libertad del error experimental resultan de las interacciones de REPS (un grado de libertad) con los factores A, B, C y las interacciones AB, AC, BC. En total, seis grados de libertad.

Este ejercicio sirve solo de ejemplo. En la práctica, por lo general, no sería conveniente tener un experimento con tan pocos grados de libertad del error.

### *El arreglo ortogonal de los tratamientos*

En el caso de un experimento diseñado como un experimento factorial fraccionado  $\frac{1}{2} * 2^n$ , lo más fácil es decidir que la interacción más alta sea el contras-

te definidor. Para constituir los dos grupos basta separar los tratamientos que tienen un número de letras impares de los que tienen un número de letras pares (al tratamiento (1) se lo toma como par). Los dos grupos son ortogonales y el experimento queda perfectamente balanceado. Con este procedimiento la interacción de grado más alto se constituye automáticamente en la interacción confundida con el efecto de bloques. Veamos el siguiente ejercicio:

**Ejercicio 19.3.2.** En un experimento se estudió el efecto de cuatro factores en el rendimiento de alfalfa: A, nitrógeno (0 y 120 kg/ha); B, fósforo (0 y 35 kg/ha); C, potasio (0 y 30 kg/ha); D, cal (0 y 4 t/ha). Por las condiciones de campo se planificó tener cuatro bloques incompletos en dos repeticiones. Se tomó la interacción ABCD como el contraste definidor para formar dos grupos: un grupo de tratamientos con número de letras pares y el otro grupo con número de letras impares. Los tratamientos de cada grupo se asignaron al azar a los bloques. Este es un experimento factorial  $\frac{1}{2} * 2^4$ . Los resultados se presentan a continuación:

**Cuadro 1. Rendimiento de materia seca de alfalfa en kilogramos por parcela. Experimento factorial fraccionado  $\frac{1}{2} * 2^4$  en cuatro bloques incompletos\***

BLQ	A	B	C	D	REND	TRAT	BLQ	A	B	C	D	REND
1	1	1	1	1	10	(1)	3	1	1	1	1	12
1	2	2	1	1	22	ab	3	2	2	1	1	24
1	2	1	2	1	24	ac	3	2	1	2	1	25
1	2	1	1	2	25	ad	3	2	1	1	2	19
1	1	2	2	1	12	bc	3	1	2	2	1	15
1	1	2	1	2	13	bd	3	1	2	1	2	13
1	1	1	2	2	9	cd	3	1	1	2	2	16
1	2	2	2	2	15	abcd	3	2	2	2	2	26
2	2	1	1	1	15	a	4	2	1	1	1	18
2	1	2	1	1	12	b	4	1	2	1	1	13
2	1	1	2	1	10	c	4	1	1	2	1	12
2	1	1	1	2	13	d	4	1	1	1	2	14
2	2	2	2	1	18	abc	4	2	2	2	1	26
2	2	2	1	2	16	abd	4	2	2	1	2	24
2	2	1	2	2	14	abd	4	2	1	2	2	20
2	1	2	2	2	13	acd	4	1	2	2	2	15

\*Note que los bloques 1 y 2 forman una repetición, así como lo hacen los bloques 3 y 4.

El análisis de varianza se presenta a continuación:

Statistix 9.0

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
BLOQ	3	107.844	35.948		
A	1	520.031	520.031	61.47	0.0000
B	1	13.781	13.781	1.63	0.2226
C	1	1.531	1.531	0.18	0.6770
D	1	0.281	0.281	0.03	0.8579
A*B	1	0.031	0.031	0.00	0.9524
A*C	1	0.281	0.281	0.03	0.8579
A*D	1	16.531	16.531	1.95	0.1839
B*C	1	0.031	0.031	0.00	0.9524
B*D	1	3.781	3.781	0.45	0.5147
C*D	1	19.531	19.531	2.31	0.1509
A*B*C	1	5.281	5.281	0.62	0.4426
A*B*D	1	0.031	0.031	0.00	0.9524
A*C*D	1	13.781	13.781	1.63	0.2226
B*C*D	1	30.031	30.031	3.55	0.0805

Error 14 118.438 8.460  
 Total 31 851.219  
 Grand Mean 16.656 CV 17.46

Similar al caso del ejercicio anterior, los 14 grados de libertad del error resultan de la interacción entre repeticiones (un grado de libertad, aunque confundido con bloques) y los componentes del análisis. La interacción ABCD, seleccionada como el contraste definidor, no se puede calcular. Está confundida con el efecto de bloques.

### ***Los alias y la interacción generalizada***

Otro concepto importante en los experimentos factoriales fraccionados, es el de “alias”, término con el que se identifica a los efectos factoriales idénticos. En los experimentos con repeticiones fraccionadas, todo contraste tiene uno o más contrastes como alias.

El procedimiento para reconocer los alias consiste en multiplicar los efectos por el contraste definidor. Si resulta un término al cuadrado, se lo elimina del producto. Al resultado se lo denomina la “interacción generalizada”. El alias de un efecto factorial es su interacción generalizada.

Como ejemplo, en el arreglo de un experimento factorial  $2^3$ , en el que el contraste definidor es ABC, las interacciones generalizadas para los efectos principales A, B y C son:

- $A*ABC = A^2BC = BC$ ; por lo tanto, A y BC son alias
- $B*ABC = AB^2C = AC$ ; por lo tanto, B y AC son alias
- $C*ABC = ABC^2 = AB$ ; por lo tanto, C y AB son alias

Esto se puede ver más claramente en el Cuadro 19.3.3.

Cuadro 19.3.3. Experimento factorial fraccionado  $\frac{1}{2}$  de  $2^3$ . Coeficientes ortogonales de los tratamientos a, b, c, abc, cuando la interacción ABC es el contraste definidor

Efectos	Tratamientos				$\sum c_i$
	a	b	c	abc	
A	+1	-1	-1	+1	0
B	-1	+1	-1	+1	0
AB	-1	-1	+1	+1	0
C	-1	-1	+1	+1	0
AC	-1	+1	-1	+1	0
BC	+1	-1	-1	+1	0
ABC	+1	+1	+1	+1	4

Note en el cuadro anterior que los signos de los coeficientes de A son los mismos que los de BC. Consecuentemente los dos efectos son idénticos o alias. Lo mismo para B y su alias AC y para C y su alias AB. Consecuentemente, en un experimento con varios bloques incompletos que contengan los tratamientos a, b, c, abc, la estimación del valor de A será la misma que la de BC. Igual para los casos de B y C. Esto implica que en el cuadro ADEVA sólo podrá constar uno de los alias. Veamos el siguiente ejemplo:

**Ejercicio 19.3.3.** En un experimento factorial fraccionado  $\frac{1}{2} * 2^3$  en que la interacción ABC es el contraste definidor, los tratamientos a, b, c, abc, se asignaron a cuatro bloques incompletos. Se pide demostrar, mediante el análisis de varianza, la presencia de alias. Los datos se presentan en el siguiente cuadro:

**Cuadro 1. Rendimiento simulado del efecto de los factores A, B, C en cuatro bloques incompletos**

Blq	A	B	C	Rend	Trat	Blq	A	B	C	Rend
1	1	0	0	12	a	3	1	0	0	6
1	0	1	0	4	b	3	0	1	0	12
1	0	0	1	5	c	3	0	0	1	14
1	1	1	1	10	abc	3	1	1	1	8
2	1	0	0	11	a	4	1	0	0	13
2	0	1	0	6	b	4	0	1	0	15
2	0	0	1	14	c	4	0	0	1	8
2	1	1	1	9	abc	4	1	1	1	16

El análisis de varianza con Statistix se presenta a continuación:

**Statistix: Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
BLQ	3	55.688	18.5625		
A	1	3.063	3.0625	0.18	0.6798
B	1	0.562	0.5625	0.03	0.8590
C	1	1.562	1.5625	0.09	0.7676
Error	9	151.563	16.8403		
Total	15	212.438			

Se recordará que el alias de A es BC, el de B es AC y el de C es AB. Con el método de contrastes ortogonales, procedimiento que consta en el Cuadro 2 del ejercicio 19.3.1, se puede demostrar que las sumas de cuadrados de las interacciones son:  $BC = 3.0630$ ,  $AC = 0.5620$  y  $AB = 1.5620$ . Estos valores son exactamente iguales a las sumas de cuadrados de A, B y C, respectivamente.

Por defecto, el programa presenta el factor principal y no la interacción que es su alias. Como ejemplo, con el programa Statistix, si se deja fuera del análisis al

factor A, se obtendrá su alias, es decir la interacción BC, cuya suma de cuadrados es 3.0630. Esto puede verse en el siguiente análisis:

**Statistix: Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
BLQ	3	55.688	18.5625		
B	1	0.563	0.5625	0.03	0.8590
C	1	1.562	1.5625	0.09	0.7676
B*C	1	3.063	3.0625	0.18	0.6798
Error	9	151.563	16.8403		
Total	15	212.438			

Los nueve grados de libertad del error, en este último análisis, provienen de las interacciones entre bloques con B, C y BC.

### *Fraccionamiento adicional de los tratamientos*

Hasta el momento se han analizado experimentos factoriales con una sola repetición y otros con la mitad de una repetición. En esta sección se verá el caso de experimentos con la cuarta parte de una repetición. A partir de este punto se podrá generalizar para otros casos de experimentos factoriales fraccionados  $2^{n-2}$ , en los que  $n > 4$ . Esto implica que no es posible construir un diseño  $\frac{1}{4} * 2^4$  en el que los mismos cuatro tratamientos se repitan en diferentes bloques incompletos.

Cuando se planifica un experimento factorial fraccionado  $\frac{1}{4} * 2^n$ , se requiere de dos efectos factoriales como contrastes definidores. Con el primer contraste definidor, guiándose por los signos de los efectos factoriales, se obtienen dos grupos, cada uno con la mitad de los tratamientos. Con el segundo contraste definidor, una de las mitades anteriores (la que se escoja), se divide nuevamente en dos. De esta forma se generan dos nuevos grupos con la cuarta parte de los tratamientos. Los dos grupos son ortogonales y puede seleccionarse solo uno de ellos, o los dos grupos de tratamientos  $\frac{1}{4} * 2^n$ , para asignarlos a los bloques incompletos.

Para este doble fraccionamiento de los tratamientos hay cuatro reglas:

1. El investigador escoge los dos efectos factoriales que servirán como contrastes definidores
2. La interacción generalizada de los dos contrastes definidores también actúa como un contraste definidor.

3. Todos los efectos factoriales (los que no son contrastes definidores) tienen, cada uno, tres alias.
4. Por la presencia de alias, no se podrá estimar todas las interacciones entre dos factores, excepto a partir de experimentos factoriales  $\frac{1}{4} * 2^8$ .

Cuadro 19.3.4. Contraste definidor ABE para un experimento factorial  $2^5$ .  
Formación de dos grupos de 16 tratamientos

Trat	Signos del contraste				Trat	Signos del contraste			
	A	B	E	ABE		A	B	E	ABE
(1)	-	-	-	-	e	-	-	+	+
a	+	-	-	+	ae	+	-	+	-
b	-	+	-	+	be	-	+	+	-
ab	+	+	-	-	abe	+	+	+	+
c	-	-	-	-	ce	-	-	+	+
ac	+	-	-	+	ace	+	-	+	-
bc	-	+	-	+	bce	-	+	+	-
abc	+	+	-	-	abce	+	+	+	+
d	-	-	-	-	de	-	-	+	+
ad	+	-	-	+	ade	+	-	+	-
bd	-	+	-	+	bde	-	+	+	-
abd	+	+	-	-	abde	+	+	+	+
cd	-	-	-	-	cde	-	-	+	+
acd	+	-	-	+	acde	+	-	+	-
bcd	-	+	-	+	bcde	-	+	+	-
abcd	+	+	-	-	abcde	+	+	+	+

Suponga el caso de un experimento factorial  $\frac{1}{4} * 2^5$ . Si se escoge, como el primer contraste definidor a la interacción ABE, se formarán dos grupos de 16 tratamientos. (El procedimiento para la asignación de los tratamientos se presentó en la sección correspondiente y Cuadro 19.3.2.) Los signos denotan la presencia o ausencia de los niveles de los factores. Los productos de los signos de A, B y E generan los signos del contraste definidor ABE, signos que sirven para separar los tratamientos en dos grupos. Esto consta en el Cuadro 19.3.4.

De los signos negativos y positivos, bajo la columna del contraste definidor (Cuadro 19.3.4), se forman dos grupos de 16 tratamientos:

Grupo 1: (1), ab, c, abc, d, abd, cd, abcd, ae, be, ace, bce, ade, bde, acde, bcde  
Grupo 2: a, b, ac, bc, ad, bd, acd, bcd, e, abe, ce, abce, de, abde, cde, abcde

Si se selecciona como segundo contraste definidor a la interacción BCD y se decide tomar el grupo 2 de 16 tratamientos, se obtiene el resultado que se presenta en el Cuadro 19.3.5.

Los dos grupos de ocho tratamientos que se forman al tomar como segundo contraste definidor a la interacción BCD son (Cuadro 19.3.5):

Grupo 1: b, ac, ad, bcd, abe, ce, de, abcde

Grupo 2: a, bc, bd, acd, e, abce, abde, cde

**Cuadro 19.3.5. Segundo contraste definidor BCD.  
Formación de dos grupos de ocho tratamientos**

Trat	Signos del contraste				Trat	Signos del contraste			
	B	C	D	BCD		B	C	D	BCD
a	-	-	-	-	e	-	-	-	-
b	+	-	-	+	abe	+	-	-	+
ac	-	+	-	+	ce	-	+	-	+
bc	+	+	-	-	abce	+	+	-	-
ad	-	-	+	+	de	-	-	+	+
bd	+	-	+	-	abde	+	-	+	-
acd	-	+	+	-	cde	-	+	+	-
bcd	+	+	+	+	abcde	+	+	+	+

Para determinar la interacción generalizada se multiplican los dos contrastes definidores y se eliminan los factores que resulten elevados al cuadrado. Esto es:

$ABE*BCD = AB^2CDE = ACDE$ . Por lo tanto, los contrastes definidores son tres.

Los alias se identifican al multiplicar los efectos factoriales por los tres contrastes definidores ABE, BCD y ACDE. Consecuentemente, cada efecto factorial tiene tres alias. Como ejemplo, para el efecto factorial A, sus tres alias son:

$$A*ABE = A^2BE = BE$$

$$A*BCD = ABCD$$

$$A*ACDE = A^2CDE = CDE$$

En el siguiente cuadro se presentan los alias de los demás efectos factoriales:



Cuadro 19.3.6. Los alias en un experimento factorial fraccionado  $\frac{1}{4} * 2^5$  cuando los contrastes definidores son ABE, BCD y ACDE

$$\begin{aligned}
 I &= ABE = BCD = ACDE \\
 A &= BE = ABCD = CDE \\
 B &= AE = CD = ABCDE \\
 C &= ABCE = BD = ADE \\
 D &= ABDE = BC = ACE \\
 E &= AB = BCDE = ACD \\
 AC &= BCE = ABD = DE \\
 AD &= BDE = ABC = CE
 \end{aligned}$$

Se observa en el Cuadro 19.3.6 que ninguno de los efectos principales, o factores, (A, B, C, D, E) tiene como alias a otro efecto principal. Esto determina que en el análisis de varianza pueda estimarse el efecto de cada uno. El caso de las interacciones es distinto. Todas las interacciones entre dos factores tienen como alias a un efecto principal, con excepción de AC y AD. Consecuentemente, de las 10 interacciones entre dos factores, solo puede estudiarse el efecto de las interacciones AC y AD. Todas las otras interacciones entre tres, cuatro y cinco factores tienen como alias a efectos factoriales de menor grado. Veamos el siguiente ejercicio:

**Ejercicio 19.3.4.** Se planificó un experimento factorial fraccionado  $\frac{1}{4} * 2^5$  con cuatro bloques incompletos y un grupo de ocho tratamientos. Los contrastes definidores fueron las interacciones ABE, BCD y la interacción generalizada ACDE. Efectuar el análisis de varianza con los datos que se presentan a continuación:

Cuadro 1. Datos de un experimento factorial fraccionado  $\frac{1}{4} * 2^5$ , en cuatro bloques incompletos y con un grupo de ocho tratamientos

BLQ	A	B	C	D	E	REND	Trat	BLQ	A	B	C	D	E	REND
1	1	2	1	1	1	15	b	3	1	2	1	1	1	19
1	2	1	2	1	1	18	ac	3	2	1	2	1	1	21
1	2	1	1	2	1	25	ad	3	2	1	1	2	1	11
1	1	2	2	2	1	32	bcd	3	1	2	2	2	1	22
1	2	2	1	1	2	23	abe	3	2	2	1	1	2	16
1	1	1	2	1	2	12	ce	3	1	1	2	1	2	14
1	1	1	1	2	2	21	de	3	1	1	1	2	2	19

1	2	2	2	2	2	16	abcde	3	2	2	2	2	2	17
2	1	2	1	1	1	10	b	4	1	2	1	1	1	23
2	2	1	2	1	1	11	ac	4	2	1	2	1	1	26
2	2	1	1	2	1	14	ad	4	2	1	1	2	1	13
2	1	2	2	2	1	23	bcd	4	1	2	2	2	1	17
2	2	2	1	1	2	15	abe	4	2	2	1	1	2	21
2	1	1	2	1	2	9	ce	4	1	1	2	1	2	18
2	1	1	1	2	2	13	de	4	1	1	1	2	2	14
2	2	2	2	2	2	17	abcde	4	2	2	2	2	2	26

El análisis de varianza es el siguiente:

Statistix 9.0

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
BLQ	3	195.344	65.1146		
A	1	2.531	2.5313	0.12	0.7343
B	1	87.781	87.7812	4.10	0.0557
C	1	22.781	22.7813	1.06	0.3139
D	1	26.281	26.2812	1.23	0.2803
E	1	26.281	26.2812	1.23	0.2803
A*C	1	0.031	0.0312	0.00	0.9699
A*D	1	87.781	87.7812	4.10	0.0557
Error	21	449.406	21.4003		
Total	31	898.219			

Observe que en el análisis de varianza solo constan los efectos factoriales de los cinco factores y de las interacciones AC y AD. Note, en el Cuadro 19.3.6, que las interacciones entre dos factores AC y AD son las únicas que no son alias de algún factor. De las 24 otras interacciones, 21 son alias con los efectos del análisis de varianza (8 entre dos factores, 8 entre tres factores, 4 entre cuatro factores y 1 entre cinco factores). Las tres restantes interacciones son los contrastes definidores ABE, BCD y ACDE, que están confundidas con el efecto de bloques.

En el caso de los alias, el programa Statistix calcula primero, por defecto, la suma de cuadrados de los efectos principales. Luego calcula las de las interaccio-

nes entre dos factores y así sucesivamente. Como ejemplo, si se excluye del análisis al factor A, el programa presentará la interacción BE que es alias de A. El ADEVA en este caso es el siguiente:

Statistix 9.0

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
BLQ	3	195.344	65.1146		
B	1	87.781	87.7813	4.10	0.0557
C	1	22.781	22.7813	1.06	0.3139
D	1	26.281	26.2812	1.23	0.2803
E	1	26.281	26.2812	1.23	0.2803
B*E	1	2.531	2.5312	0.12	0.7343
C*E	1	87.781	87.7813	4.10	0.0557
D*E	1	0.031	0.0312	0.00	0.9699
Error	21	449.406	21.4003		
Total	31	898.219			

Si se comparan los dos cuadros ADEVA se demuestra que A y BE son alias. Las sumas de cuadrados son idénticas, así como el término del error. Esto es lo que se esperaba.

**Comentario:** En los experimentos factoriales, el investigador está generalmente interesado no sólo en los efectos principales sino también en la presencia de posibles interacciones. En la práctica, un experimento factorial fraccionado  $\frac{1}{4} \cdot 2^5$  con un solo grupo de ocho tratamientos aporta muy poca información, como se demostró en el ejercicio anterior. Únicamente a partir de experimentos  $\frac{1}{4} \cdot 2^8$  se pueden obtener los efectos de todos los factores y de todas las interacciones entre dos factores. Pero esto también depende de la selección adecuada de los contrastes definidores.

La idea principal atrás del concepto de bloques incompletos y experimentos factoriales fraccionados es evaluar varios factores a la vez, analizar algunas interacciones de interés, y controlar en mejor forma la variación propia de las unidades experimentales. Estos objetivos pueden incluso alcanzarse con experimentos  $\frac{1}{4} \cdot 2^5$  si se utilizan de dos a cuatro grupos ortogonales de ocho tratamientos. En el siguiente ejercicio se presenta este caso:

**Ejercicio 19.3.5.** En una investigación se estudió el efecto de cinco factores, cada factor con dos niveles, en el rendimiento de soya (Vargas, 1990)<sup>47</sup> Los cinco factores fueron: factor A, dos variedades de soya, Siatsa 194 (nivel 1) y Clark (nivel 2); B, fertilizante nitrogenado, 0 kg/ha y 100 kg/ha; C, fertilizante fosfatado, 0 kg/ha y 35 kg/ha; D, inoculación de la semilla, sin y con inóculo; E, tratamiento de la semilla, sin y con fungicida. De lo anterior resulta un experimento con 32 tratamientos.

Como el terreno en el que se conduciría el ensayo no era homogéneo, se procedió a dividir los 32 tratamientos en cuatro grupos de ocho tratamientos cada grupo. Los grupos se asignarían al azar a cuatro bloques incompletos, que son áreas más pequeñas y con menor variabilidad. El resultado es un experimento factorial fraccionado  $\frac{1}{4} \cdot 2^5$ .

Para dividir los 32 tratamientos en dos grupos de 16 se escogió el contraste definidor ABE. El siguiente contraste definidor, para generar cuatro grupos ortogonales de ocho tratamientos, fue BCD. Con esta selección de contrastes definidores y la decisión de usar los cuatro grupos en el ensayo, el investigador se aseguró que obtendrá los efectos de los cinco factores y de las 10 interacciones entre dos factores. Los resultados, con modificaciones para facilidad del ejemplo, se presentan a continuación:

**Cuadro 1. Rendimiento de soya, en toneladas por hectárea, en respuesta al efecto de cinco factores**

BLOQ	A	B	C	D	E	REND	trat	BLOQ	A	B	C	D	E	REND	Trat
1	1	2	1	1	1	3.34	b	3	1	1	1	1	1	3.81	(1)
1	2	1	2	1	1	1.97	ac	3	2	2	2	1	1	1.91	abc
1	2	1	1	2	1	1.50	ad	3	2	2	1	2	1	2.32	abd
1	1	2	2	2	1	4.37	bcd	3	1	1	2	2	1	3.64	cd
1	2	2	1	1	2	1.78	abe	3	2	1	1	1	2	1.60	ae
1	1	1	2	1	2	3.62	ce	3	1	2	2	1	2	3.84	bce
1	1	1	1	2	2	4.63	de	3	1	2	1	2	2	4.70	bde
1	2	2	2	2	2	2.41	abcde	3	2	1	2	2	2	1.82	acde

47 Vargas, G.D. 1990. Efecto de la época de siembra, fertilización, inoculación y tratamiento de semilla en el comportamiento de dos líneas de soya. Tesis. Escuela Agrícola Panamericana. Honduras, C.A.

2	2	1	1	1	1	1.38	a	4	2	2	1	1	1	1.86	ab
2	1	2	2	1	1	3.39	bc	4	1	1	2	1	1	3.71	c
2	1	2	1	2	1	3.52	bd	4	1	1	1	2	1	3.57	d
2	2	1	2	2	1	2.97	acd	4	2	2	2	2	1	2.93	abcd
2	1	1	1	1	2	4.66	e	4	1	2	1	1	2	2.90	be
2	2	2	2	1	2	2.16	abce	4	2	1	2	1	2	1.65	ace
2	2	2	1	2	2	1.78	abde	4	2	1	1	2	2	1.54	ade
2	1	1	2	2	2	3.40	cde	4	1	2	2	2	2	2.99	bcde

El análisis de varianza, con el programa Statistix, es el siguiente:

Statistix 9.0

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
BLOQ	3	0.5321	0.1774		
A	1	25.4006	25.4006	112.73	0.0000
B	1	0.0167	0.0167	0.07	0.7900
C	1	0.1116	0.1116	0.50	0.4939
D	1	0.6356	0.6356	2.82	0.1169
E	1	0.0158	0.0158	0.07	0.7956
A*B	1	0.6933	0.6933	3.08	0.1029
A*C	1	1.2129	1.2129	5.38	0.0372
A*D	1	0.0621	0.0621	0.28	0.6083
A*E	1	0.3806	0.3806	1.69	0.2163
B*C	1	0.0914	0.0914	0.41	0.5353
B*D	1	0.3140	0.3140	1.39	0.2589
B*E	1	0.0657	0.0657	0.29	0.5983
C*D	1	0.0001	0.0001	0.00	0.9854
C*E	1	0.8745	0.8745	3.88	0.0705
D*E	1	0.1785	0.1785	0.79	0.3896
Error	13	2.9291	0.2253		
Total	31	33.5146			

Grand Mean 2.8647 CV 16.57

Los 13 grados de libertad del error provienen de ocho interacciones de segundo grado (entre tres factores), más cuatro interacciones de tercer grado, más la

interacción de cuarto grado ABCDE. Las interacciones que faltan son ABE, BCD y ACDE. Estos son los contrastes definidores que no forman parte del error, sino que están confundidos con el efecto de bloques.

Con relación a los resultados, únicamente la interacción entre variedades y la fertilización fosfatada ( $A \times C$ ) fue significativa, con una probabilidad  $P = 3.72\%$ . El siguiente gráfico ayuda a interpretar la interacción:

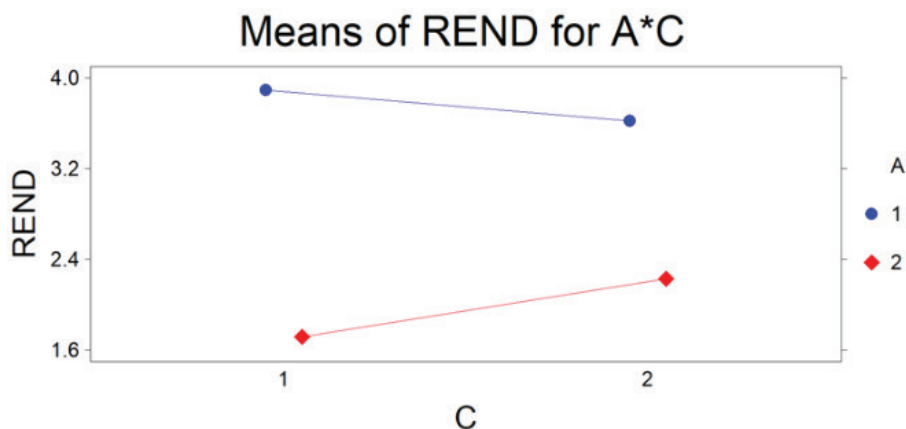


Gráfico 1. Interacción entre variedades (A) y fertilización fosfatada (C)

Se pudo esperar que las dos variedades tuvieran respuestas similares con relación a la fertilización fosfatada. Podría pensarse que las condiciones químicas dentro de los bloques fueron diferentes. El coeficiente de variación, demasiado alto para un cultivo de autopolinización, abona a esta idea.

En todo caso, sin necesidad de ver las medias, la diferencia entre variedades resalta en el gráfico anterior. La variedad Siatsa 194 fue estadísticamente superior en rendimiento que la variedad Clark. En el ADEVA, la probabilidad de que este resultado sea producto del azar es sumamente baja ( $P \leq 0.00009$ ), a pesar del relativamente alto coeficiente de variación.



*“Un análisis estadístico propiamente conducido es una delicada disección de incertidumbres, una cirugía de suposiciones”.* Michael J. Moroney<sup>48</sup>

Los diseños anidados son experimentos factoriales en los que, por características del ensayo, conveniencia del investigador y material disponible, unidades experimentales pequeñas están dentro de unidades experimentales de mayor tamaño. Estos experimentos son muy utilizados en diferentes campos de investigación. En este capítulo se verán varios ejemplos relacionados con la agricultura.

### **20.1 . Diseño de Parcela Dividida**

Parcela Dividida es un diseño en el que se analizan los efectos principales y la interacción de dos factores en parcelas de diferente tamaño, llamadas respectivamente, *parcelas grandes o principales* y *parcelas pequeñas o subparcelas*. Todo experimento conducido con un diseño de Parcela Dividida es un experimento factorial. Su uso se recomienda cuando por la naturaleza de los factores o de las unidades experimentales es conveniente instalar el experimento en parcelas de diferentes tamaños. Los casos más comunes en agricultura se presentan cuando hay interés en los efectos, por ejemplo, de diferentes tipos de labranza mecanizada, sistemas de riego, fumigación aérea, etcétera, o cuando, por conveniencia en el manejo del experimento, se usa implementos grandes como sembradoras, aspersores de insumos, cosechadoras y otros. En estos ejemplos, el empleo de parcelas grandes en los experimentos se hace necesario. Las parcelas grandes, o principales, convencionalmente reciben el factor A. Las parcelas pequeñas, o subparcelas, reciben el factor B.

Por razón del diferente número de unidades experimentales a las que se asignan los niveles de los factores, las varianzas del error o cuadrados medios del

---

48 Estadístico inglés. [https://todayinsci.com/M/Moroney\\_Michael/MoroneyMichael-Quotations.htm](https://todayinsci.com/M/Moroney_Michael/MoroneyMichael-Quotations.htm)



error son diferentes. El efecto del factor A se mide con relación al Cuadrado Medio del Error (a), mientras que los efectos del factor B y de la interacción A\*B se miden con relación al Cuadrado Medio del Error (b).

Más adelante se verá que, de haber un factor C, las subparcelas pueden dividirse y formar *sub-subparcelas*. El diseño se llama en este caso *Parcela Subdividida*. En teoría, se podría seguir con más subdivisiones, pero los experimentos se vuelven muy grandes y complejos. Si es de interés del investigador incluir un factor C, bien podría combinar los factores B y C en las subparcelas, como un experimento factorial dentro de la Parcela Dividida. Puede inferirse de lo señalado que las opciones son muchas, sin embargo, la recomendación es mantener los experimentos en la forma más simple posible.

Parcela Dividida es un diseño compuesto. Si se conoce que el terreno es homogéneo, los niveles del Factor A se asignarán totalmente al azar a las parcelas principales. Esto da origen a un experimento de Parcela Dividida en un Diseño Completamente al Azar. Si el terreno tiene una gradiente, el experimento será de Parcela Dividida en un Diseño de Bloques Completos al Azar. Si son dos gradientes perpendiculares, lo más conveniente será un experimento de Parcela Dividida en un diseño de Cuadrado Latino.

Independientemente del diseño, Los niveles del Factor B se asignan completamente al azar a las subparcelas. Esto implica que se realizan dos aleatorizaciones en el diseño de Parcela Dividida, una para las parcelas principales y otra para las subparcelas.

Suponga un diseño de Parcela Dividida en Bloques Completos al Azar, en el que el factor A tenga tres niveles (a1, a2, a3) y el factor B cuatro niveles (b1, b2, b3, b4). Una vez realizadas las aleatorizaciones, un mapa de un bloque podría ser el siguiente:

Cuadro 20.1.1. Mapa de un bloque de un diseño de Parcela Dividida en Bloques Completos al Azar: tres parcelas principales y 12 subparcelas

PP1		PP2		PP3	
a1b1	a1b3	a3b2	a3b1	a2b4	a2b1
a1b4	a1b2	a3b3	a3b4	a2b2	a2b3

Note en el mapa anterior que en el Bloque hay tres parcelas principales (PP). Por sorteo, el nivel a1 se asignó a la PP1 (por eso las cuatro subparcelas tienen a1), el nivel a3 a la PP2 y el nivel a2 a la PP3. Tanto los niveles de A como los

niveles de B deben sortearse de nuevo en cada uno de los otros bloques que tenga el experimento.

### *El Modelo Lineal Aditivo*

El modelo lineal aditivo depende del tipo de diseño. Como ejemplo, en el caso más simple, cuando se tiene Parcela Dividida en un Diseño Completamente al Azar el modelo es el siguiente:

$$Y_{ij} = \mu + A_i + \varepsilon_a + B_j + (A*B)_{ij} + \varepsilon_b$$

Observe que existen dos errores: el error  $\varepsilon_a$  asociado con la parcela principal y el error  $\varepsilon_b$  asociado con la subparcela. Al igual que con los diseños anteriores, el modelo puede ser fijo, al azar o mixto. Esto, como se vio anteriormente, tiene implicaciones en el alcance de las recomendaciones que se puedan hacer a partir de los resultados obtenidos.

Los esquemas del ADEVA para los tres diseños básicos en Parcela Dividida se presentan en el Cuadro 20.1.2.

**Cuadro 20.1.2. Esquemas del ADEVA para los tres diseños básicos en Parcela Dividida (dos factores A y B)**

Diseño Completamente al Azar		Diseño Bloques Completos al Azar		Diseño Cuadrado Latino	
FV	Gl	FV	Gl	FV	Gl
		Bloques	r-1	Hileras	a-1
				Columnas	a-1
A	a-1	A	a-1	A	a-1
error (a)	a(r-1)	error (a)	(a-1)(r-1)	error (a)	(a-1)(a-2)
B	b-1	B	b-1	B	b-1
A*B	(a-1)(b-1)	A*B	(a-1)(b-1)	A*B	(a-1)(b-1)
error (b)	a(r-1)(b-1)	error (b)	a(r-1)(b-1)	error (b)	a(a-1)(b-1)
<b>Total</b>	<b>rab-1</b>	<b>Total</b>	<b>rab-1</b>	<b>Total</b>	<b>a2b-1</b>

## Errores Estándar en el diseño de Parcela Dividida

El procedimiento para la separación de medias es más complicado que para los otros diseños. En el diseño de parcela dividida se debe considerar qué medias se van a separar y de acuerdo con esto se calcula el error estándar. Sin embargo, esto debe hacerse a partir del Cuadrado Medio del Error más idóneo. En algunos casos es necesario incluso combinar los cuadrados medios y obtener un error estándar sopesado. En el Cuadro 20.1.3 se presentan los errores estándar para la comparación de dos medias en un Diseño de Parcela Dividida.

Cuadro 20.1.3. Errores estándar para la diferencia de dos medias en un diseño de Parcela Dividida (dos factores: A en la parcela principal y B en la subparcela)

Medias en comparación	ejemplos de diferencias $\bar{Y}_{ai} - \bar{Y}_{aj}$	errores estándar de las diferencias entre dos medias
dos medias de A	$a_1 - a_2$	$\sqrt{\frac{2 * CME_a}{rb}}$
dos medias de B	$b_1 - b_2$	$\sqrt{\frac{2 * CME_b}{ra}}$
dos medias de B al mismo nivel de A	$a_1 b_1 - a_1 b_2$	$\sqrt{\frac{2 * CME_b}{r}}$
dos medias de A al mismo o diferente nivel de B	$a_1 b_1 - a_2 b_1, a_1 b_1 - a_2 b_2$	$\sqrt{\frac{2[(b-1) * CME_b + CME_a]}{rb}}$

En el Cuadro 20.1.3,  $CME$  es el Cuadrado Medio del Error para cada uno de los factores. Las letras  $r$ ,  $a$ ,  $b$ , son, respectivamente, el número de repeticiones o bloques, el número de niveles del factor A y el número de niveles del factor B.

A continuación, se presenta un ejemplo del análisis de varianza de un experimento conducido con un diseño de parcela dividida:

**Ejercicio 20.1.1.** Se realizó un experimento con un diseño de Parcela Dividida en Bloques Completos al Azar para evaluar el efecto de diferentes fungicidas y cuatro fertilizantes foliares en el peso del racimo de banano. El factor A, fungicidas, tuvo tres niveles:  $a_1$ ,  $a_2$ ,  $a_3$  y el factor B, fertilizantes foliares, cuatro niveles:  $b_1$ ,  $b_2$ ,  $b_3$ ,  $b_4$ . Los fungicidas se esparcieron, con bombas nebulizadoras, en las parcelas grandes. Los fertilizantes foliares se aplicaron a las plantas de las sub-parcelas. Hubo tres bloques o repeticiones. Se pide efectuar el análisis correspondiente. Los resultados se presentan en el siguiente cuadro:

Cuadro 1. Peso de racimos de banano, promedio de cinco plantas, en kilogramos

Factores		R1	R2	R3	
Fungicidas	F. Foliares				
a1	b1	35.0	28.0	29.1	
a1	b2	37.7	29.6	30.7	
a1	b3	40.9	35.9	33.8	
a1	b4	27.5	29.8	26.0	
Parcelas grandes		141.1	123.3	119.6	
a2	b1	28.9	25.7	24.8	
a2	b2	27.3	25.4	29.6	
a2	b3	32.4	32.7	33.5	
a2	b4	34.1	37.3	31.6	
Parcelas grandes		122.7	121.1	119.5	
a3	b1	29.5	31.8	24.4	
a3	b2	30.9	33.2	36.1	
a3	b3	37.3	35.2	29.1	
a3	b4	31.8	33.4	27.2	
Parcelas grandes		129.5	133.6	116.8	
Bloques Total, $R_j$		393.3	378.0	355.9	$\sum Y_{ij} = 1127.2$

Note en el Cuadro 1 la presencia de los totales de las parcelas grandes o principales. Debe calcularse la Suma de Cuadrados de las Parcelas Grandes (SCPG) para encontrar el Error(a) y el Error(b). Igual que en los experimentos factoriales, en el diseño de Parcela Dividida también se requiere formar una tabla de doble entrada A\*B, para calcular los efectos principales y la interacción. (Si se tratara de un diseño de Parcela Sub-dividida, esto es tres factores, entonces habría que formar tres tablas de doble entrada: A\*B, A\*C y B\*C).

Cuadro 2. Cuadro A\*B para el cálculo de los efectos principales de A y de B y de la interacción A\*B

Factor A	Factor B				Totales $A_a$
	b1	b2	b3	b4	
a1	92.1	98.0	110.6	83.1	384.0
a2	79.4	82.3	98.6	103.0	363.3
a3	85.7	100.2	101.6	92.4	379.9
Totales $B_b$	257.2	280.5	310.8	278.7	$\sum Y_{ij} = 1127.2$

Sobre la base de los Cuadros 1 y 2 se pueden estimar los componentes del análisis de varianza:

$$FC = \frac{(\sum Y_{ij})^2}{r * a * b} = \frac{(1127.2)^2}{3 * 3 * 4} = 35293.88$$

$$SCT = (35.0^2 + 28.0^2 + \dots + 27.2^2) - FC = 578.42$$

$$SCR = \frac{(393.3^2 + 378.0^2 + 355.9^2)}{a * b} - FC = 58.92$$

$$SC_{cuadro}(A * B) = \frac{(92.1^2 + 98.0^2 + \dots + 92.4^2)}{r} - FC = 343.16$$

$$SCA = \frac{(384.0^2 + 363.3^2 + 379.9^2)}{r * b} - FC = 20.02$$

$$SCB = \frac{(257.2^2 + 280.5^2 + 310.8^2 + 278.7^2)}{r * a} - FC = 161.94$$

$$SC(A * B) = SC_{cuadro}(A * B) - SCA - SCB = 161.20$$

Hasta aquí el análisis de los resultados es similar al de Bloques Completos al Azar con un arreglo factorial, con la diferencia que en Parcela Dividida hay dos errores: Error (a) y Error (b). Para encontrar el Error (a) se calcula primero la Suma de Cuadrados de las Parcelas Grandes (SCPG) (ver el Cuadro 1) y luego, de esta cantidad, se resta la SCR y la SCA:

$$SCPG = \frac{(141.1^2 + 123.3^2 + \dots + 116.8^2)}{b} - FC = 125.73$$

$$SCE(a) = SCPG - SCR - SCA = 46.79$$

El Error (b) se encuentra con la siguiente fórmula:

$$SCE(b) = SCT - SCPG - SCB - SC(A * B) = 129.55$$

Con las cantidades anteriores se procede al análisis de varianza del diseño de Parcela Dividida en Bloques Completos al Azar:

Cuadro 3. Análisis de Varianza

E.V.	S.C.	g.l.	C.M.	F	F <sub>05</sub>	F <sub>01</sub>	P
Repeticiones	58.92	2	29.46				
Factor A	20.02	2	10.01	0.86 n.s.	6.94	18.00	49.0%
Error (a)	46.78	4	11.70				
Factor B	161.94	3	53.98	7.50**	3.16	5.09	0.18%
A*B	161.19	6	26.87	3.73*	2.66	4.01	1.37%
Error (b)	129.55	18	7.20				
Total	578.42	35					

Gran Media:  $\bar{Y} = 31.3$

Los coeficientes de variación son:

$$CV(a) = \frac{\sqrt{11.7}}{31.3} * 100 = 10.9\% \quad CV(b) = \frac{\sqrt{7.20}}{31.3} * 100 = 8.57\%$$

El análisis de varianza con Statistix es el siguiente:

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
BLQ	2	58.924	29.4619		
A	2	20.024	10.0119	0.86	0.4904
Error BLQ*A	4	46.783	11.6957		
B	3	161.940	53.9800	7.50	0.0018
A*B	6	161.192	26.8653	3.73	0.0137
Error BLQ*A*B	18	129.553	7.1974		
Total	35	578.416			

Grand Mean 31.311

CV (BLQ\*A) 10.92

CV (BLQ\*A\*B) 8.57

De acuerdo con los resultados del análisis de varianza, se detectó una interacción significativa entre fungicidas (factor A) y fertilizantes foliares (factor B). La probabilidad de que este efecto, sobre la variable “peso de racimo”, se deba al azar es bastante baja ( $P = 1.37\%$ ). Esta probabilidad menor que 5% genera confianza en el resultado. Para interpretar la interacción se puede hacer un grá-

fico con las medias u obtenerlo directamente con el programa Statistix. Con esta segunda opción se tiene:

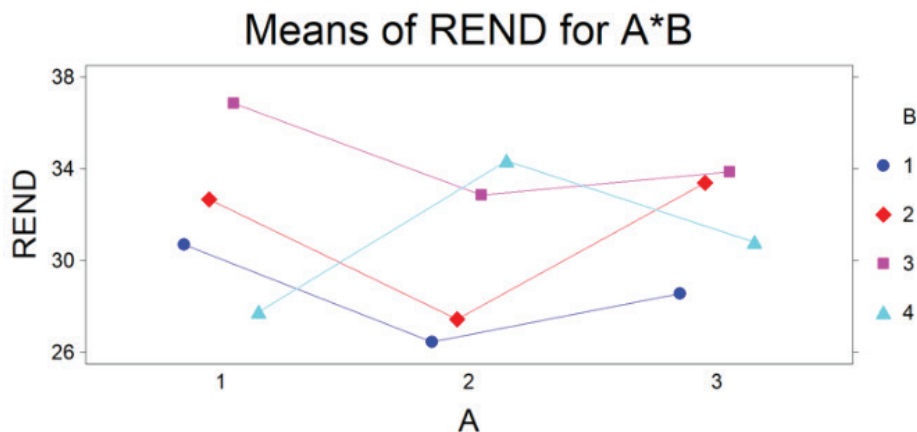


Gráfico 1. Interacción significativa entre fungicidas (factor A) y fertilizantes foliares (factor B) para la variable “peso de racimo”.

Se observa en el Gráfico 1 que la interacción significativa proviene principalmente del comportamiento del fertilizante foliar 4. Mientras el peso de racimo para los fertilizantes 1, 2 y 3 es menor en presencia del fungicida 2, con el fertilizante 4 ocurre lo contrario. (El investigador deberá buscar la explicación de este resultado, posiblemente en términos de los elementos de los fertilizantes foliares y de los ingredientes activos de los fungicidas.)

A continuación, se presenta un cuadro con las medias:

Cuadro 4. Medias de peso de racimo para los factores A, B y la interacción A\*B

Factor A	Factor B				$\bar{Y}_a$
	b1	b2	b3	b4	
a1	30.70	32.67	36.87	27.77	32.00
a2	26.47	27.43	32.87	34.33	30.28
a3	28.57	33.40	33.87	30.80	31.66
$\bar{Y}_b$	28.58	31.17	34.53	30.97	$\bar{Y} = 31.31$

En el Cuadro 4, las medias de la interacción A\*B están en *itálicas* en el cuerpo de la tabla. Las medias de A se encuentran en la última columna y las de B en la última hilera.

## Separación de medias

Si uno o dos de los factores representaran variables cuantitativas, lo adecuado sería el empleo del método de polinomios ortogonales para determinar la existencia o no de tendencias. El procedimiento no difiere del aplicado en capítulos anteriores.

En este ejemplo, las variables son cualitativas, por lo que debería usarse alguno de los métodos de separación de medias descritos en el Capítulo 17. En este caso, se requiere conocer el error estándar específico para una determinada comparación.

**Medias del factor A.** Como se indica en el Cuadro 20.1.3, para comparar dos medias del Factor A, el error estándar de la diferencia de dos medias de A es:

$$S_{\bar{a}} = \sqrt{\frac{2 * CME_a}{rb}} = \sqrt{\frac{2 * 11.70}{3 * 4}} = 1.40$$

El error estándar de la diferencia de dos medias se usa para calcular la Diferencia Mínima Significativa (DMS). Sin embargo, para emplear el método de Tukey o el de Duncan, se requiere el error estándar, cuya fórmula es similar a la anterior, pero sin el 2 que multiplica al Cuadrado Medio del Error. Es decir:

$S_{\bar{y}} = \sqrt{\frac{CME_a}{rb}} = 0.99$ , o lo que es igual:  $S_{\bar{y}} = \frac{S_{\bar{a}}}{\sqrt{2}} = 0.99$ . En todo caso es importante recordar que, por los pocos grados de libertad del Error(a), la precisión de las estimaciones relacionadas con el factor A es baja.

**Medias del factor B.** Cuando se trata de comparar dos medias del Factor B, con el método DMS, el error estándar de la diferencia de dos medias de B es:

$$S_{\bar{a}} = \sqrt{\frac{2 * CME_b}{ra}} = \sqrt{\frac{2 * 7.20}{3 * 3}} = 1.26$$

El error estándar para comparar medias con el método de Tukey o el de Duncan es:

$$S_{\bar{y}} = \frac{S_{\bar{a}}}{\sqrt{2}} = 0.89$$



**Comparación de medias de la interacción.** Aquí se presentan dos casos. El primer caso, cuando la comparación es entre medias de B, al mismo nivel de A. El error estándar de la diferencia de dos medias es:

$$S_{\bar{a}} = \sqrt{\frac{2 * CME_b}{r}} = \sqrt{\frac{2 * 7.20}{3}} = 2.19 \text{ mientras que el error estándar es:}$$

$$S_{\bar{y}} = \frac{S_{\bar{a}}}{\sqrt{2}} = 1.55$$

El segundo caso se da cuando la comparación es entre medias de A, al mismo o diferente nivel de B. El error estándar de la diferencia de dos medias, para este ejemplo, es:

$$S_{\bar{a}} = \sqrt{\frac{2[(b-1) * CME_b + CME_a]}{rb}} = \sqrt{\frac{2[(4-1) * 7.20 + 11.70]}{3 * 4}} = 2.36$$

El error estándar es:

$$S_{\bar{y}} = \frac{S_{\bar{a}}}{\sqrt{2}} = 1.66$$

**Comentario:** por razón del diferente número de unidades experimentales a las que se asignan los niveles de los factores, las estimaciones de los efectos no tienen igual precisión. En otras palabras, en el diseño de Parcela Dividida se sacrifica precisión en la estimación del efecto del factor A, para ganar precisión en la estimación del efecto del factor B y de la interacción A\*B.

## 20.2. Diseño de Parcela Subdividida

Cuando se estudia el efecto de tres factores y estos, por conveniencia o necesidad, se asignan a parcelas de diferentes tamaños, se tiene un Diseño de Parcela Subdividida. En este diseño hay parcelas de tres tamaños: parcelas grandes, subparcelas y sub-subparcelas.

En forma convencional, el factor A se asigna a las parcelas principales, el factor B a las subparcelas y el factor C a las sub-subparcelas. En forma similar al caso de Parcela Dividida, la asignación de los niveles del factor A depende del diseño básico, sea este DCA, BCA o CL. La asignación de los niveles de B a las subparcelas y los de C a las sub-subparcelas es completamente al azar. Esto, por cuanto se supone homogeneidad dentro de cada parcela principal.

### El Modelo Lineal Aditivo

El modelo lineal aditivo depende del tipo de diseño básico. Como ejemplo, en el caso más simple, cuando se tiene Parcela Sub-dividida en un Diseño Completamente al Azar, el modelo es el siguiente:

$$Y_{ijk} = \mu + A_i + \varepsilon_a + B_j + (AB)_{ij} + \varepsilon_b + C_k + (AC)_{ik} + (BC)_{jk} + (ABC)_{ijk} + \varepsilon_c$$

Los esquemas de los análisis de varianza de los diseños de Parcela Sub-dividida, en los tres diseños básicos, son idénticos a los de Parcela Dividida hasta el error (b) (ver Cuadro 20.1.2). Por esta razón, en el Cuadro 20.2.1 solo se presentan los complementos de los esquemas a partir de los efectos del factor C.

**Cuadro 20.2.1. Complementos de los esquemas del ADEVA para los tres diseños básicos en Parcela Subdividida (tres factores A, B y C)**

Diseño Completamente al Azar		Diseño Bloques Completos al Azar		diseño Cuadrado Latino	
FV	Gl	FV	Gl	FV	Gl
C	c-1	C	c-1	C	c-1
A*C	(a-1)(c-1)	A*C	(a-1)(c-1)	A*C	(a-1)(c-1)
B*C	(b-1)(c-1)	B*C	(b-1)(c-1)	B*C	(b-1)(c-1)
A*B*C	(a-1)(b-1)(c-1)	A*B*C	(a-1)(b-1)(c-1)	A*B*C	(a-1)(b-1)(c-1)
error (c)	ab(r-1)(c-1)	error (c)	ab(r-1)(c-1)	error (c)	ab(r-1)(c-1)
<b>Total</b>	<b>abcr-1</b>	<b>Total</b>	<b>abcr-1</b>	<b>Total</b>	<b>a<sup>2</sup>bc-1</b>

En el poco frecuente caso de experimentos en Parcela Sub-subdividida (cuatro factores: A, B, C, D), los efectos principales, interacciones y errores pueden deducirse del contenido del cuadro anterior.

### Errores Estándar en el diseño de Parcela Subdividida

Los errores estándar para las diferencias entre dos medias se encuentran en el Cuadro 20.2.2.

Cuadro 20.2.2. Errores estándar para la diferencia de dos medias en un diseño de Parcela Subdividida (tres factores: A en la parcela principal, B en la subparcela y C en la sub-subparcela)

Medias en comparación	ejemplos de diferencias $\bar{Y}_{ai} - \bar{Y}_{aj}$	errores estándar de las diferencias entre dos medias
dos medias de A	$a1 - a2$	$\sqrt{\frac{2 * CME_a}{rbc}}$
dos medias de B	$b1 - b2$	$\sqrt{\frac{2 * CME_b}{rac}}$
dos medias de C	$c1 - c2$	$\sqrt{\frac{2 * CME_c}{rab}}$
dos medias de B al mismo o diferente nivel de A	$a1b1 - a1b2, a1b1 - a2b2$	$\sqrt{\frac{2 * CME_b}{rc}}$
dos medias de C al mismo nivel de A	$a1c1 - a1c2$	$\sqrt{\frac{2 * CME_c}{rb}}$
dos medias de C al mismo nivel de B	$b1c1 - b1c2$	$\sqrt{\frac{2 * CME_c}{ra}}$
dos medias de C en igual combinación de A y B	$a1b2c1 - a1b2c2$	$\sqrt{\frac{2 * CME_c}{r}}$
dos medias de A al mismo o diferente nivel de B	$a1b1 - a2b1, a1b1 - a2b2$	$\sqrt{\frac{2[(b-1) * CME_b + CME_a]}{rbc}}$
dos medias de B al mismo o diferente nivel de C	$b1c1 - b2c1, b1c1 - b2c2$	$\sqrt{\frac{2[(c-1) * CME_c + CME_b]}{rac}}$
dos medias de B en igual combinación de A y C	$a2b1c1 - a2b2c1$	$\sqrt{\frac{2[(c-1) * CME_c + CME_b]}{rc}}$
dos medias de A al mismo o diferente nivel de C	$a1c1 - a2c1, a1c1 - a2c2$	$\sqrt{\frac{2[(c-1) * CME_c + CME_a]}{rbc}}$

El siguiente ejemplo es de un diseño de Parcela Subdividida en Bloques Completos al Azar:

**Ejercicio 20.2.1.** En este experimento se estudió el efecto de dos herbicidas (Factor A, en las parcelas grandes), dos tiempos de aplicación del herbicida (Factor B, en las subparcelas) y cuatro fertilizantes foliares (Factor C, en las sub-subparcelas). La variable de respuesta fue el rendimiento en t/ha de una variedad de soya “round-up ready”. (El experimento se realizó en Bolivia.) Como se identificó un gradiente de fertilidad, la asignación de los herbicidas a las parcelas grandes fue en Bloques Completos al Azar. Hubo cuatro bloques o repeticiones. Los resultados se presentan en el siguiente cuadro:

Cuadro 1. Rendimiento de soya en t/ha

Trat No.	Herbicidas Factor A	TiempoApl Factor B	FertFoliar Factor C	R1	R2	R3	R4	Totales $T_i$
1	1	1	1	1.27	1.14	1.37	1.21	4.99
2	1	1	2	1.56	1.57	1.53	1.79	6.45
3	1	1	3	1.68	1.14	1.81	1.59	6.22
4	1	1	4	1.53	1.81	1.88	2.04	7.26
5	1	2	1	1.43	1.01	1.87	1.40	5.71
6	1	2	2	1.55	1.31	1.39	1.23	5.48
7	1	2	3	1.46	1.90	1.75	1.90	7.01
8	1	2	4	1.49	1.92	1.86	2.41	7.68
9	2	1	1	2.53	1.76	2.46	1.02	7.77
10	2	1	2	2.29	2.47	2.34	2.18	9.28
11	2	1	3	1.95	1.96	2.71	1.79	8.41
12	2	1	4	1.92	2.05	2.07	2.11	8.15
13	2	2	1	2.26	2.04	2.58	2.13	9.01
14	2	2	2	1.47	2.32	1.29	1.11	6.19
15	2	2	3	1.70	1.89	2.04	1.62	7.25
16	2	2	4	1.48	1.38	1.43	2.05	6.34
	Totales de los bloques, $R_j$			27.57	27.67	30.38	27.58	$\sum Y_{ij} = 113.2$

En el caso de Parcela Subdividida, los cálculos para el análisis se facilitan si se obtienen primero los componentes de los diseños básicos, aunque varios de estos no se incluyan en el ADEVA correspondiente. En este ejemplo, el diseño básico es Bloques Completos al Azar (BCA).

Con los datos del Cuadro 1 se puede proceder a calcular las estadísticas del diseño BCA, que conducirán posteriormente al análisis de varianza de la Parcela Subdividida. Estas son: Factor de Corrección (FC), Suma de Cuadrados Total (SCT), Suma de Cuadrados de las Repeticiones o Bloques (SCR), Suma de Cuadrados de los Tratamientos (SCTrat) y Suma de Cuadrados del Error ( $SCE_{BCA}$ ). El valor  $SCE_{BCA}$  es igual a la sumatoria de los tres errores de la Parcela Subdividida:

$$FC = \frac{(\sum Y_{ij})^2}{r*a*b*c} = \frac{(1.27 + 1.14 + \dots + 2.05)^2}{4*2*2*4} = \frac{(113.20)^2}{64} = 200.22$$

$$SCT = (1.27^2 + 1.14^2 + \dots + 2.05^2) - FC = 10.86$$

$$SCR = \frac{(27.57^2 + 27.67^2 + 30.38^2 + 27.58^2)}{a*b*c = t = 16} - FC = 0.36$$

$$SCTrat = \frac{(4.99^2 + 6.45^2 + \dots + 6.34^2)}{r = 4} - FC = 10.86$$

$$SCE_{(BCA)} = SCT - SCR - SCTrat = 4.58$$

Los anteriores son los componentes del Diseño de Bloques al Azar. Para continuar con el análisis de los componentes de la Parcela Subdividida se debe construir tres cuadros de doble entrada: Cuadro A\*B, Cuadro A\*C y Cuadro B\*C. A partir del Cuadro 1 se obtienen los totales respectivos para formar los cuadros de doble entrada:

Cuadro 2. Cuadro A\*B para el cálculo de los efectos principales de A y de B y de la interacción A\*B

Factor A	Factor B		Totales $A_a$
	b1	b2	
a1	24.92	25.88	50.80
a2	33.61	28.79	62.40
Totales $B_b$	58.53	54.67	$\sum Y_{ij} = 113.20$

$$SC_{\text{cuadro}}(A*B) = \frac{(24.92^2 + 25.88^2 + 33.61^2 + 28.79^2)}{r*b*c = 16} - FC = 2.86$$

$$SCA = \frac{(50.80^2 + 62.40^2)}{r*b*c = 32} - FC = 2.10$$

$$SCB = \frac{(58.53^2 + 54.67^2)}{r*a*c = 32} - FC = 0.23$$

$$SC(A*B) = SC_{\text{cuadro}}(A*B) - SCA - SCB = 0.52$$

Cuadro 3. Cuadro A\*C para el cálculo del efecto principal de C y de la interacción A\*C

Factor A	Factor C				Totales $A_a$
	c1	c2	c3	c4	
a1	10.70	11.93	13.23	14.94	50.80
a2	16.78	15.47	15.66	14.49	62.40
Totales $C_c$	27.48	27.40	28.89	29.43	$\sum Y_{ij} = 113.20$

$$SC_{\text{cuadro}}(A*C) = \frac{(10.70^2 + 11.93^2 + \dots + 14.49^2)}{r*b = 8} - FC = 3.67$$

$$SCC = \frac{(27.48^2 + 27.40^2 + 28.88^2 + 29.46^2)}{r*a*b = 16} - FC = 0.19$$

$$SC(A*C) = SC_{\text{cuadro}}(A*C) - SCA - SCC = 1.37$$

Cuadro 4. Cuadro de doble entrada B\*C para el cálculo de la interacción B\*C

Factor B	Factor C				Totales $B_b$
	c1	c2	c3	c4	
b1	12.76	15.73	14.63	15.41	58.53
b2	14.72	11.67	14.26	14.02	54.67
Totales $C_c$	27.48	27.40	28.89	29.43	$\sum Y_{ij} = 113.20$

$$SC_{\text{cuadro}}(B*C) = \frac{(12.76^2 + 15.73^2 + \dots + 14.02^2)}{r*a = 8} - FC = 1.60$$

$$SC(B*C) = SC_{\text{cuadro}}(B*C) - SCB - SCC = 1.17$$

La Suma de Cuadrados de la interacción entre los tres factores se calcula por diferencia a partir de la Suma de Cuadrados de los Tratamientos ( $SCTrat$ ):

$$SC(A*B*C) = SCTrat - SCA - SCB - SC(A*B) - SCC - SC(A*C) - SC(B*C) = 0.32$$

Hasta aquí, el análisis de los resultados es similar al de Bloques Completos al Azar con un arreglo factorial A – B - C, con la diferencia que en Parcela Subdividida hay tres errores: Error (a), Error (b) y Error (c). El Error (a) se encuentra a partir de la Suma de Cuadrados de las Parcelas Grandes (SCPG) y el Error (b) a partir de la Suma de Cuadrados de las Subparcelas (SCSP).

Puede notarse en el Cuadro 1 que cada repetición tiene dos parcelas grandes (o principales). Las parcelas grandes corresponden al factor A. Los totales de estas parcelas se presentan en el Cuadro 5.

**Cuadro 5. Totales de las parcelas grandes (factor A) por repeticiones**

	Repeticiones			
	R1	R2	R3	R4
Parcelas principales, a1	11.97	11.80	13.46	13.57
Parcelas principales, a2	15.60	15.87	16.92	14.01
Repeticiones, $R_j$	27.57	27.67	30.38	27.58

Con la información del Cuadro 5 se procede a calcular la Suma de Cuadrados de las Parcelas principales o grandes (SCPG) y luego, la Suma de Cuadrados del Error (a):

$$SCPG = \frac{(11.97^2 + 11.80^2 + \dots + 30.38^2)}{b*c = 8} - FC = 2.98$$

$$SCE(a) = SCPG - SCR - SCA = 0.52$$

En el Cuadro 6 se presentan los totales de las subparcelas (esto es, la suma de cuatro valores de las sub-subparcelas, es decir, valores  $Y_{ij}$ ). Las subparcelas corresponden al factor B. Cada repetición contiene cuatro subparcelas, según consta en el Cuadro 1.

Cuadro 6. Totales de las subparcelas (factor B) por repeticiones

	Repeticiones			
	R1	R2	R3	R4
Sub-parcelas, (a1) b1	6.04	5.66	6.59	6.63
Sub-parcelas, (a1) b2	5.93	6.14	6.87	6.94
Sub-parcelas (a2) b1	8.69	8.24	9.58	7.10
Sub-parcelas (a2) b2	6.91	7.63	7.34	6.91
Repeticiones, $R_j$	27.57	27.67	30.38	27.58

Con la información del Cuadro 6 se obtiene:

$$SCSP = \frac{(6.04^2 + 5.66^2 + \dots + 6.91^2)}{c = 4} - FC = 4.11$$

$$SCE(b) = SCSP - SCPG - SCB - SC(A*B) = 0.37$$

El Error (c) puede calcularse a través de más de un procedimiento, pero la forma más rápida es por diferencia con relación a la Suma de Cuadrados del Error del diseño BCA:

$$SCE(c) = SCE_{(BCA)} - SCE(a) - SCE(b) = 3.70$$

Se dispone ya de todos los componentes para formar el cuadro del análisis de varianza:

Cuadro 7. Análisis de varianza de la variable rendimiento de soja.  
Diseño de Parcela Subdividida en Bloques Completos al Azar

Fuente de Variación	Suma de Cuadrados	Grados de libertad	Cuadrado medio	Valor F	F al 5%	F al 1%	Probabilidad
Repeticiones	0.36	3	0.12				
Factor A	2.10	1	2.10	12.35 *	10.13	34.12	3.9 %
Error (a)	0.52	3	0.17				
Factor B	0.23	1	0.23	3.83 n.s	5.99	13.75	10.0 %
A*B	0.52	1	0.52	8.67 *			2.7 %
Error (b)	0.37	6	0.06				
Factor C	0.19	3	0.06	0.60 n.s	2.88	4.42	60.9 %
A*C	1.37	3	0.46	4.60 **			0.9 %
B*C	1.17	3	0.39	3.90 *			1.8 %
A*B*C	0.32	3	0.11	1.10 n.s			38.1 %
Error (c)	3.70	36	0.10				
TOTAL	10.86	63					



El coeficiente de variación de A es:

$$CV_{(A)} = \frac{\sqrt{CME_a}}{\bar{Y}} * 100 = \frac{\sqrt{0.1722}}{1.7688} * 100 = 23.46\%$$

De forma similar se calculan los coeficientes de variación de los otros factores.

Las interacciones A\*B y B\*C resultaron estadísticamente significativas al nivel del 5%, mientras que la interacción A\*C fue significativa al 1%. Sólo el efecto del factor A presentó diferencias estadísticas, a pesar del reducido número de grados de libertad del Error (a). Gráficos de las interacciones ayudarán a interpretar en mejor forma estos resultados.

Con el programa Statistix los resultados son iguales, salvo pequeñas diferencias por error de redondeo:

**Analysis of Variance Table for REND. Statistix 9.0**

Source	DF	SS	MS	F	P
REPS	3	0.3609	0.12030		
A	1	2.1025	2.10250	12.21	0.0397
Error REPS*A	3	0.5167	0.17223		
B	1	0.2328	0.23281	3.76	0.1006
A*B	1	0.5220	0.52201	8.43	0.0272
Error REPS*A*B	6	0.3716	0.06193		
C	3	0.1942	0.06474	0.63	0.6002
A*C	3	1.3728	0.45761	4.46	0.0092
B*C	3	1.1668	0.38894	3.79	0.0185
A*B*C	3	0.3243	0.10809	1.05	0.3813
Error REPS*A*B*C	36	3.6969	0.10269		
Total	63	10.8615			

Grand Mean 1.7688, CV(REPS\*A) 23.46, CV(REPS\*A\*B) 14.07,  
CV(REPS\*A\*B\*C) 18.12

Las medias de las respuestas significativas se las puede obtener a partir de los cuadros de doble entrada. Las medias con Statistix son:

**Means of REND for A**

A	Mean
1	1.5875
2	1.9500

Std Error (Diff of 2 Means)      0.1038

**Means of REND for A\*B**

A	B	Mean
1	1	1.5575
1	2	1.6175
2	1	2.1006
2	2	1.7994

Standard Error of a Mean      0.0622

**Means of REND for A\*C**

A	C	Mean
1	1	1.3375
1	2	1.4913
1	3	1.6538
1	4	1.8675
2	1	2.0975
2	2	1.9337
2	3	1.9575
2	4	1.8112

Standard Error of a Mean      0.1133

**Means of REND for B\*C**

B	C	Mean
1	1	1.5950
1	2	1.9663
1	3	1.8288
1	4	1.9262
2	1	1.8400
2	2	1.4588
2	3	1.7825
2	4	1.7525

Observations per Mean      8  
 Standard Error of a Mean      0.1133  
 Error term used: REPS\*A\*B\*C, 36 DF.

Con los valores de las medias y con los errores estándar respectivos se puede, utilizando la información del Cuadro 20.2.2, realizar las comparaciones que se consideren pertinentes. En todo caso, los gráficos de las interacciones significativas ayudan a interpretar los resultados:

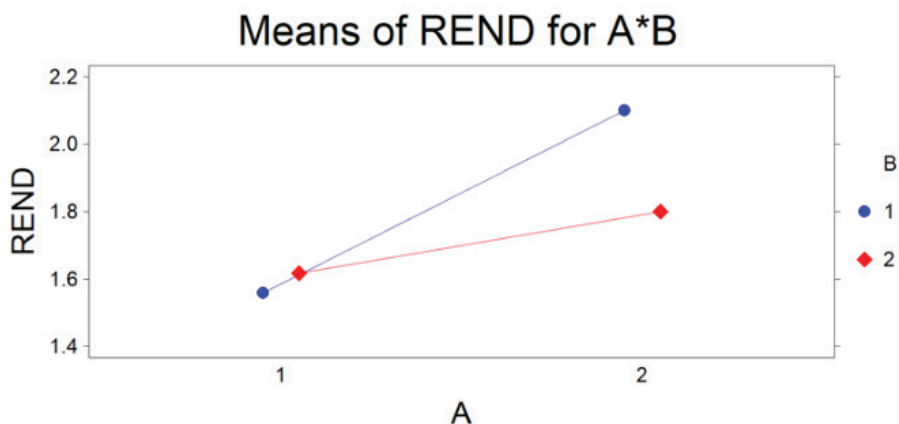


Gráfico 1. Presencia de interacción entre los factores A y B

La interacción A\*B fue significativa ( $P=2.7\%$ ). El Gráfico 1 indica que hubo una respuesta diferente de los herbicidas a1 y a2 (factor A) con relación al tiempo de aplicación (factor B). La combinación a2b1 determinó un mayor rendimiento que las otras combinaciones.

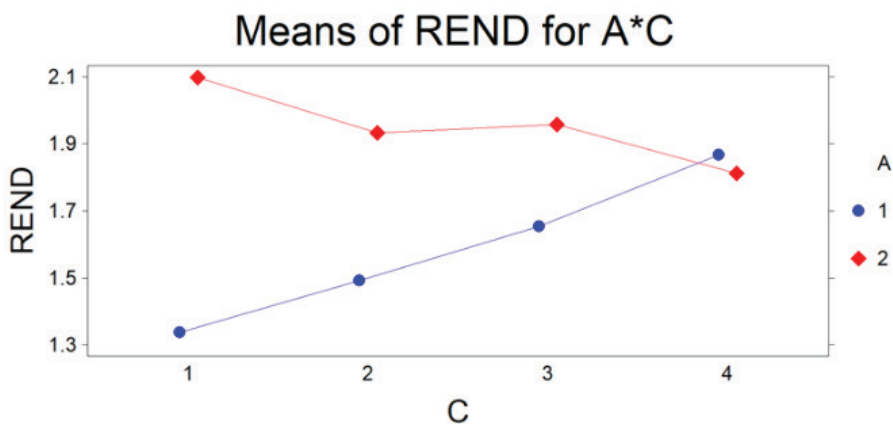


Gráfico 2. Presencia de interacción entre los factores A y C

La interacción A\*C resultó altamente significativa ( $P=0.92\%$ ), como se demuestra en el Gráfico 2. La mejor combinación fue la del herbicida a2 con el fertilizante foliar c1, con un rendimiento promedio de 2.09 t/ha.

La interacción B\*C también resultó significativa ( $P=1.85\%$ ). Las tendencias pueden observarse en el Gráfico 3. La respuesta es muy diferente, en términos de rendimiento de soja, entre b1c2 (promedio 1.96 t/ha) y b2c2 (promedio 1.45 t/ha). Esta respuesta variable del fertilizante c2 en relación con los tiempos de aplicación del herbicida (b1 y b2) fue la causa principal de la interacción significativa detectada.

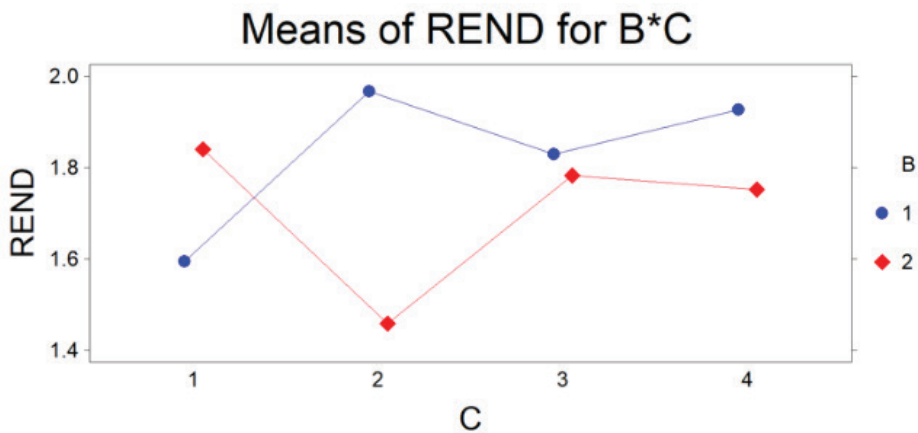


Gráfico 3. Presencia de interacción entre los factores B y C

### Comentarios:

- Los coeficientes de variación, excesivamente altos para un cultivo endogámico como la soja, indican que hubo deficiencias en la conducción del experimento.
- En los diseños de parcela dividida se espera, teóricamente, que los coeficientes de variación disminuyan progresivamente con el menor tamaño y mayor número de parcelas. Esto no ocurrió en este experimento, lo que corrobora lo expresado en el comentario anterior.
- En igual forma que en otros ejemplos, las posibles explicaciones agronómicas y biológicas de los resultados obtenidos, no han sido objeto de una discusión detallada en este texto.

### 20.3. Diseño de Parcela Dividida en el Tiempo

En los ejemplos anteriores, las parcelas grandes se dividen físicamente para formar parcelas más pequeñas. Por esta característica se los conoce también como *diseños de parcelas divididas en el espacio*. Sin embargo, cuando en los experimentos se toman datos a través del tiempo, como es el caso de cosechas periódicas de cultivos perennes, se origina un nuevo factor, el factor tiempo. Aquí no hay una división física del terreno, por lo que las parcelas del factor A, como las del factor B (tiempo) son exactamente iguales en tamaño.

Este arreglo se denomina *Parcela Dividida en el Tiempo*. El concepto es algo similar al de parcela dividida y frecuentemente se analizan los datos con ese modelo. Sin embargo, debe considerarse el hecho de que el factor B no se asigna al azar dentro de las parcelas del factor A. Esto determina que A, B y la interacción A\*B tengan su propio error, por lo que mayor similitud existe con el diseño de Bloque Dividido, que se verá en la sección 20.5 de este capítulo.

Este diseño también puede utilizarse con cultivos anuales. Como ejemplo, si se estudia el efecto de sistemas de manejo de plagas; evaluaciones periódicas durante el ciclo del cultivo pueden arrojar luz sobre su eficiencia.

Si el diseño base fuera un BCA, el modelo lineal aditivo sería el siguiente:

$$Y_{ij} = \mu + R_{i,j} + A_i + \varepsilon_a + B_j + \varepsilon_b + (A*B)_{ij} + \varepsilon_{ab}$$

Note que cada término: A, B, A\*B tienen su propio error, que proviene de la interacción con las repeticiones. De acuerdo con el modelo, el esquema del análisis de varianza se presenta en el Cuadro 20.3.1.

Cuadro 20.3.1. Esquema del ADEVA para un diseño de Parcela Dividida en el Tiempo en un BCA

Fuente de variación	grados de libertad
Bloques, R	r-1
Factor A	a-1
Error (a) = R*A	(r-1)*(a-1)
Factor B	b-1
Error (b) = R*B	(r-1)*(b-1)
A*B	(a-1)*(b-1)
Error (ab) = R*A*B*	(r-1)*(a-1)*(b-1)
TOTAL	r*a*b*-1

Veamos el siguiente ejercicio:

**Ejercicio 20.3.1.** Suponga un experimento con un cultivo perenne, conducido con un diseño BCA, ( $r=3$ ). El objetivo es evaluar el rendimiento de dos cultivares (factor A: a1, a2) a través de cuatro cosechas en períodos de igual intervalo (factor B: b1, b2, b3, b4). Los datos se presentan en el Cuadro 1. Se pide realizar el análisis de varianza.

Cuadro 1. Resultados para un ejemplo de Parcela Dividida en el Tiempo en un BCA

Repeticiones	Factor A	Factor B	Rendimiento
1	1	1	12
1	1	2	18
1	1	3	15
1	1	4	14
1	2	1	16
1	2	2	25
1	2	3	23
1	2	4	20
2	1	1	7
2	1	2	13
2	1	3	13
2	1	4	11
2	2	1	14
2	2	2	20
2	2	3	17
2	2	4	16
3	1	1	13
3	1	2	19
3	1	3	20
3	1	4	16
3	2	1	18
3	2	2	28
3	2	3	26
3	2	4	21

El procedimiento para encontrar los componentes del análisis de varianza es muy similar al del Ejercicio 20.1.1, por lo que se presenta directamente el resultado con el programa Statistix 9.0:

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
REP	2	160.333	80.167		
A	1	222.042	222.042	1332.25	0.0007
Error REP*A	2	0.333	0.167		
B	3	178.792	59.597	42.91	0.0002
Error REP*B	6	8.333	1.389		
A*B	3	5.458	1.819	1.42	0.3253
Error REP*A*B	6	7.667	1.278		
Total	23	582.958			

Grand Mean 17.292

CV(REP\*A) 2.36

CV(REP\*B) 6.82

CV(REP\*A\*B) 6.54

En el diseño de Parcela Dividida en el Tiempo, una vez ejecutado el análisis de varianza, lo más adecuado es el análisis de tendencias a través de los diferentes períodos (polinomios ortogonales). Si la interacción A\*B fuera significativa, habría que efectuar este análisis en forma individual para cada nivel del factor A. En este ejemplo la interacción no fue estadísticamente significativa ( $P = 32.5\%$ ), por lo que el método de polinomios ortogonales puede aplicarse a todos los datos:

**Polynomial Contrasts of REND by B. Statistix 9.0**

Degree = 1, Linear Trend

Contrast	1.6771	SS (Contrast)	16.875
T-Statistic	3.49	P (T-Statistic)	0.0131
SE (Contrast)	0.4811		

Degree = 2, Quadratic Trend

Contrast	-4.9167	SS (Contrast)	145.04
T-Statistic	-10.22	P (T-Statistic)	0.0001
SE (Contrast)	0.4811		

Error term used: REP\*B, 6 DF

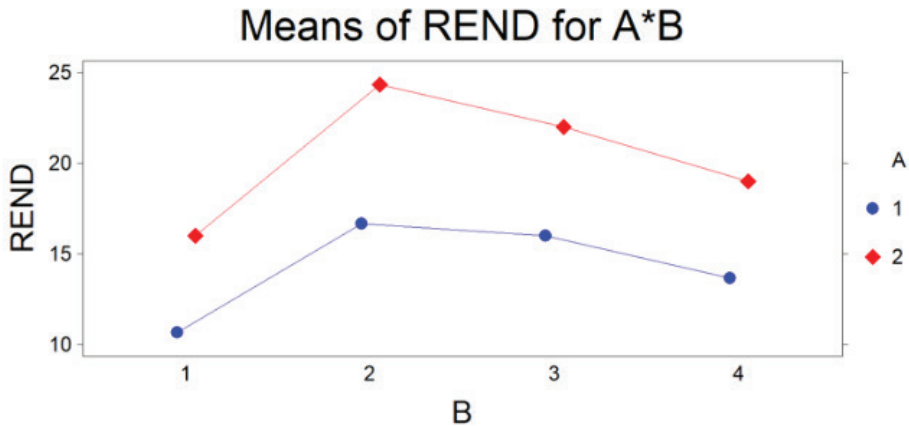


Gráfico 1. Rendimiento de dos cultivares en cuatro cosechas.

La ausencia de una interacción significativa indica que los dos cultivares tuvieron tendencias similares, aunque cuantitativamente distintas (Gráfico 1). Los rendimientos se incrementaron hasta el segundo período y luego decrecieron. Esta información puede servir para planificar prácticas agronómicas de manejo. El cultivar a2 fue significativamente superior al a1.

**Comentario:** No es un requisito que los intervalos estén igualmente espaciados. En períodos de rápido crecimiento pueden acortarse los tiempos. En estos casos habría que codificar adecuadamente los niveles del factor tiempo o introducir directamente, en la base de datos, los valores reales de los períodos.

#### 20.4. Parcela Subdividida en el Espacio y Tiempo

En este diseño se evalúan tres factores. Los niveles del factor A se asignan al azar a las parcelas principales, de las cuales se obtendrán a través del tiempo (factor B) las evaluaciones que se deseen. Esto es exactamente igual, hasta aquí, a lo presentado en el subcapítulo anterior (Parcela Dividida en el Tiempo). Una vez estructuradas las parcelas principales, su área física se divide para formar las subparcelas en las que se sortean los niveles del factor C.



### Modelo lineal aditivo

El modelo lineal aditivo de este diseño, en el caso de Bloques Completos al Azar, es el siguiente:

$$Y_{ij} = \mu + R_{i,j} + A_i + \varepsilon_a + B_j + \varepsilon_b + (AB)_{ij} + \varepsilon_{ab} + C_k + (AC)_{ik} + (BC)_{jk} + (ABC)_{ij}$$

Se sobreentiende, en el modelo, que los errores provienen de la interacción de los términos respectivos con las repeticiones.

De acuerdo con el modelo, el esquema del análisis de varianza para este diseño es:

**Cuadro 20.4.1.** Esquema del análisis de varianza de un diseño de parcela subdividida en el espacio y en el tiempo, en bloques al azar, en el que el factor B corresponde a diferentes períodos de evaluación de la variable de respuesta

Fuentes de Variación	Grados de Libertad
Bloques, R	r-1
Factor A	a-1
Error (a) = R*A	(r-1)(a-1)
Factor B	b-1
Error (b) = R*B	(r-1)(b-1)
A*B	(a-1)(b-1)
Error (ab) = R*A*B	(r-1)(a-1)(b-1)
Factor C	c-1
A*C	(a-1)(c-1)
B*C	(b-1)(c-1)
A*B*C	(a-1)(b-1)(c-1)
Error (abc) = R*A*B*C	(ab)(r-1)(c-1)
TOTAL	r*a*b*c-1

Veamos el siguiente ejemplo:

**Ejercicio 20.4.1.** En un experimento, conducido con un Diseño de Parcela Subdividida, el objetivo fue evaluar el rendimiento, a través del tiempo, de dos variedades de alfalfa en suelos enmendados con hidróxido de calcio o cal apagada. La cal se aplicó a las parcelas principales (factor A: a1 = 2 t/ha, a2 = 4 t/ha). Los períodos de corte de la alfalfa correspondieron al factor B. Los cortes se realizaron cada 45 días: b1, b2, b3, b4. Las variedades de alfalfa se asignaron a las sub-parcelas (factor C: c1, c2). El diseño base fue un BCA con tres repeticiones. Los resultados en hectogramos de materia seca de cada sub-parcela se presentan en el Cuadro 1:

Cuadro 1. Rendimientos de alfalfa en hectogramos de materia seca por sub-parcela

Cal Factor A	Cortes Factor B	Alfalfa Factor C	R1	R2	R3
1	1	1	12.54	11.45	12.38
1	2	1	13.15	12.43	14.42
1	3	1	13.22	13.16	13.97
1	4	1	14.08	13.20	15.64
1	1	2	11.56	10.86	13.23
1	2	2	11.85	10.42	13.34
1	3	2	12.48	11.60	14.08
1	4	2	13.23	12.58	15.76
2	1	1	14.43	13.41	15.04
2	2	1	14.89	13.52	16.34
2	3	1	15.35	14.80	16.85
2	4	1	16.20	15.45	15.43
2	1	2	12.38	12.90	14.68
2	2	2	14.23	13.05	14.95
2	3	2	15.12	13.67	15.12
2	4	2	15.72	14.02	16.04

El análisis de varianza con el programa Statistix se presenta a continuación:

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
R	2	29.638	14.8191		
A	1	38.449	38.4492	107.47	0.0092
Error R*A	2	0.716	0.3578		
B	3	23.020	7.6732	55.24	0.0001
Error R*B	6	0.833	0.1389		
A*B	3	0.731	0.2438	0.55	0.6678
Error R*A*B	6	2.671	0.4451		
C	1	7.115	7.1148	23.51	0.0002
A*C	1	0.029	0.0290	0.10	0.7609
B*C	3	0.877	0.2923	0.97	0.4331
A*B*C	3	0.733	0.2445	0.81	0.5079
Error R*A*B*C	16	4.842	0.3027		
Total	47	109.654			

Grand Mean 13.838, CV(R\*A) 4.32, CV(R\*B) 2.69, CV(R\*A\*B) 4.82  
CV(R\*A\*B\*C) 3.98

Si no se cuenta con un programa computarizado, para encontrar los componentes de variación debe procederse de forma similar a la señalada en el ejercicio 20.2.1. Esto es:

- Calcular el factor de corrección de las observaciones y las sumas de cuadrados del diseño base BCA:  $FC, SCT, SCR, SCTrat, SCE$
- Formar tres cuadros de doble entrada: AB, AC, BC y hallar las sumas de cuadrados de los factores y de las interacciones.
- Encontrar los totales de las parcelas A (seis parcelas) y de las parcelas B (24 parcelas) por repeticiones. Con esta información se determinan las sumas de cuadrados de las parcelas A y B, lo que conduce a calcular los errores del caso.

No se detectaron interacciones en el análisis de varianza. Sin embargo, hubo diferencias estadísticas significativas para los tres factores. Aunque la interacción entre los niveles de cal y las variedades de alfalfa no fue significativa, un gráfico ayuda a visualizar el comportamiento de estos factores:

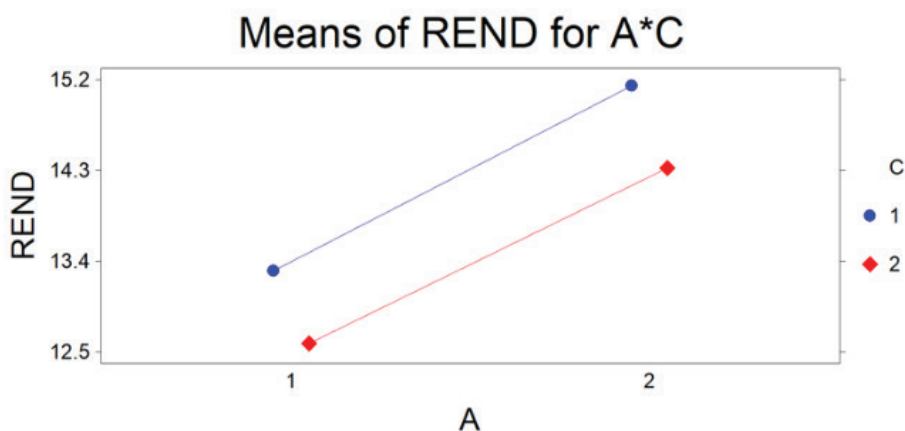


Gráfico 1. Comportamiento de dos variedades de alfalfa (c1 y c2) con relación a la aplicación de dos (a1) y cuatro (a2) toneladas de cal por hectárea

Al observar el gráfico 1 se concluye que la variedad c1 fue superior a la c2 en rendimiento y que el nivel a2 de cal determinó una mayor respuesta, independientemente de las variedades.

Con relación al factor B, rendimiento de materia seca a través del tiempo, lo más adecuado es realizar un análisis de tendencias mediante polinomios ortogonales. Los resultados con Statistix son:

**Polynomial Contrasts of REND by B**

Degree = 1, Linear Trend (tendencia lineal)

Contrast	1.3845	SS (Contrast)	23.002
T-Statistic	12.87	P (T-Statistic)	0.0000
SE (Contrast)	0.1076		

Degree = 2, Quadratic Trend (tendencia cuadrática)

Contrast	8.33E-03	SS (Contrast)	8.33E-04
T-Statistic	0.08	P (T-Statistic)	0.9408
SE (Contrast)	0.1076		

Error term used: R\*B, 6 DF

La suma de cuadrados de la tendencia lineal fue 23.002, con una probabilidad ( $P \leq 0.0000$ ). La tendencia cuadrática no fue estadísticamente significativa. Esto puede verse en el siguiente gráfico:

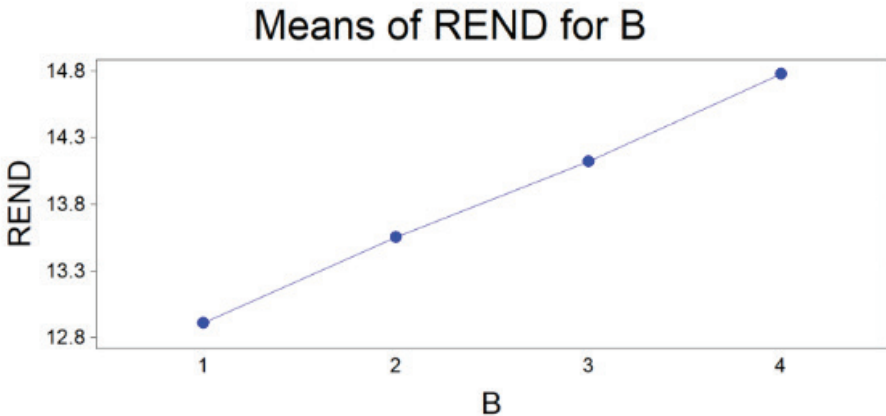


Gráfico 2. Rendimiento de materia seca de alfalfa. Tendencia lineal ascendente

La respuesta lineal señala que la producción no ha llegado a un máximo posible, por lo que sería conveniente continuar la toma de datos a través del tiempo. Esto permitiría, si fuera de interés, determinar el máximo rendimiento y el óptimo económico de este cultivo perenne.

## 20.5. Diseño de Bloque Dividido

El diseño de Bloque Dividido es un experimento factorial con dos factores, en el que los niveles de un factor se asignan al azar a franjas horizontales y los del otro factor, a franjas verticales. Se lo conoce también como *diseño de parcelas en franjas*.

A diferencia del diseño de Parcela Dividida, en este diseño no hay parcelas principales ni subparcelas, por el hecho de que la aleatorización se realiza entre franjas y no entre parcelas individuales. Sin embargo, como en todos los casos, es importante recordar que la aleatorización de las franjas debe realizarse para cada bloque. Un ejemplo de *un bloque* en el que se sortearon las franjas se presenta en el Cuadro 20.5.1:

Cuadro 20.5.1. Diseño de Bloque Dividido: factor A, tres niveles asignados a las franjas horizontales; factor B, cuatro niveles asignados a las franjas verticales

Factor A	Factor B			
	Nivel b2	b4	b1	b3
Nivel a1	$a1b2$	$a1b4$	$a1b1$	$a1b3$
Nivel a3	$a3b2$	$a3b4$	$a3b1$	$a3b3$
Nivel a2	$a2b2$	$a2b4$	$a2b1$	$a2b3$

### ¿Cuándo usar un diseño de Bloque Dividido?

Este diseño se aplica por lo general cuando se requiere disponer de parcelas grandes para los dos factores. Este es el caso de experimentos de fumigación aérea, sistemas de riego, preparación del suelo, siembra mecanizada, uso de maquinaria para la aplicación de insumos, etcétera. En estos experimentos el mayor interés está relacionado con el estudio de la posible interacción entre los factores, ya que los términos del error para los efectos principales tienen por lo general pocos grados de libertad.

El modelo lineal aditivo para este diseño es:

$$Y_{ij} = \mu + R_{i,j} + A_i + \varepsilon_a + B_j + \varepsilon_b + (A*B)_{ij} + \varepsilon_{ab}$$

Puede verse que los efectos principales y la interacción tienen su propio error. De acuerdo con el modelo, el esquema del análisis de varianza se presenta en el Cuadro 20.5.2.

Cuadro 20.5.2. Esquema del ADEVA para un diseño de Bloque Dividido

Fuente de variación	grados de libertad
Bloques, R	$r-1$
Factor A	$a-1$
Error (a) = $R^*A$	$(r-1)*(a-1)$
Factor B	$b-1$
Error (b) = $R^*B$	$(r-1)*(b-1)$
$A*B$	$(a-1)*(b-1)$
Error (ab) = $R^*A^*B^*$	$(r-1)*(a-1)*(b-1)$
TOTAL	$r*a*b-1$

Puede notarse que, tanto el modelo lineal aditivo como el esquema del análisis de varianza del diseño de Bloque dividido son iguales a los presentados para el diseño de Parcela Dividida en el Tiempo (Sección 20.3). Sin embargo, la estructura del experimento, los conceptos y aplicaciones son diferentes en los dos diseños.

Para el análisis estadístico se procede en forma similar que para cualquier experimento factorial. Mediante un cuadro de doble entrada  $A*B$  se obtienen las sumas de cuadrados de A, de B y de la interacción  $A*B$ . Sin embargo, de acuerdo con el modelo, debe tomarse en cuenta que cada uno de estos componentes tiene su propio error experimental.

Veamos un ejemplo:

**Ejercicio 20.5.1.** En un experimento se trató de evaluar el efecto de tres tipos de fertilizantes orgánicos ( $a_1, a_2, a_3$ ) en el rendimiento de cuatro híbridos de maíz ( $b_1, b_2, b_3, b_4$ ). El experimento constó de cuatro bloques. Se planificó que tanto la distribución del abono como la siembra se realizaran con implementos agrícolas. Por esta razón, el diseño que se seleccionó fue el de Bloque Dividido. En este diseño, las franjas horizontales de uno de los factores se cruzan con las franjas verticales del otro factor. El sorteo o aleatorización del orden de los niveles de cada factor debe realizarse en cada bloque. Los resultados se presentan en el siguiente cuadro:

Cuadro 1. Efecto de tres fertilizantes orgánicos, factor A, en el rendimiento (t/ha) de cuatro híbridos de maíz, factor B. Arreglo experimental en Bloque Dividido.

Bloques	Tratamientos											
	a1b1	a1b2	a1b3	a1b4	a2b1	a2b2	a2b3	a2b4	a3b1	a3b2	a3b3	a3b4
Blq1	2.95	3.22	3.45	2.84	3.39	4.54	3.74	3.25	3.95	4.41	5.23	4.63
Blq2	3.31	3.46	4.09	3.48	3.07	3.04	3.77	4.23	3.68	3.82	4.02	3.84
Blq3	3.41	3.57	3.88	3.12	2.98	4.35	3.85	3.66	4.42	5.15	4.83	3.76
Blq4	2.98	3.71	3.66	3.24	3.54	2.83	4.25	4.18	3.56	4.35	3.82	2.48

Para el análisis estadístico se procede en forma similar al caso del diseño de parcela dividida en el tiempo. Sin embargo, por el usualmente reducido número de grados de libertad de los errores A y B, debe procederse con precaución al llegar a conclusiones sobre los efectos de los factores A y B. En todo caso, como en todo experimento factorial, es importante determinar primero la presencia o ausencia de una interacción significativa.

Los resultados con el programa Statistix se presentan a continuación:

Statistix 9.0

**Analysis of Variance Table for REND**

Source	DF	SS	MS	F	P
BLOQUES	3	0.9335	0.31115		
A	2	4.2830	2.14151	4.70	0.0592
Error BLOQUES*A	6	2.7342	0.45569		
B	3	2.8431	0.94770	6.29	0.0137
Error BLOQUES*B	9	1.3566	0.15074		
A*B	6	1.0838	0.18063	0.79	0.5863
Error BLOQUES*A*B	18	4.0924	0.22735		
Total	47	17.3264			

Grand Mean 3.7290

CV(BLOQUES\*A) 18.10

CV(BLOQUES\*B) 10.41

CV(BLOQUES\*A\*B) 12.79

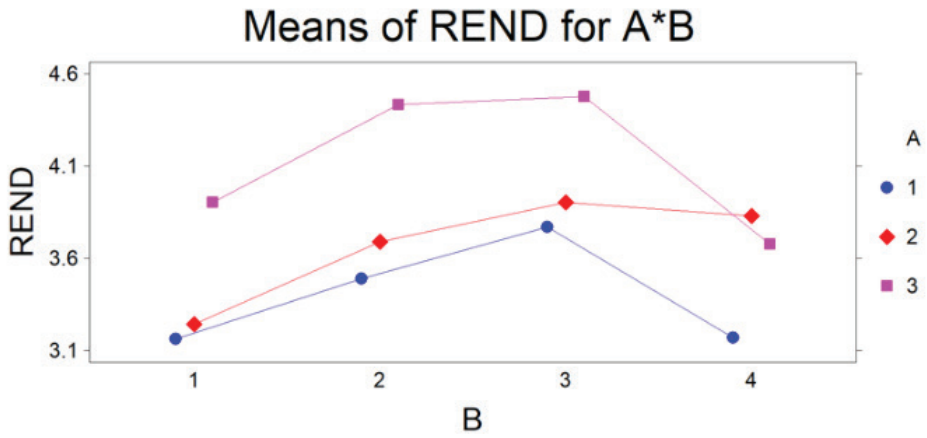


Gráfico 1. Rendimiento de cuatro híbridos de maíz en respuesta a tres tipos de fertilizantes orgánicos

Puede concluirse que la respuesta de los cuatro híbridos fue similar con relación a los fertilizantes orgánicos. Es decir, no hubo interacción.





## Capítulo 21

# Experimentos combinados en el espacio y tiempo

---

*“Para hacer producir hay que salir de las oficinas, internarse en el campo y ensuciarse las manos. Es el único lenguaje que entienden las plantas y los animales.”*  
Norman Borlaug<sup>49</sup>

Los experimentos que se realizan en varias zonas de cultivo o en diferentes épocas de siembra pueden combinarse en un solo análisis si tienen la misma estructura y el mismo tamaño.

En las ciencias agrícolas, las recomendaciones que pueden hacerse a partir de un solo experimento son limitadas. Los ensayos preliminares de evaluación de nuevos cultivos, por ejemplo, se realizan generalmente en granjas experimentales, pero luego es esencial que los materiales más promisorios sean evaluados en diferentes zonas de producción y durante varios ciclos del cultivo.

En los experimentos conducidos en una sola localidad, los resultados que se miden (por ejemplo: rendimiento de maíz, ganancia de peso en cerdos, contenido de aceite en soya) son la expresión fenotípica de los organismos. El fenotipo se define como el conjunto de características físicas, morfológicas y fisiológicas<sup>50</sup>. El fenotipo de un organismo se debe a la acción de los genes (genotipo), a los efectos del ambiente y a la interacción entre genotipo y ambiente. Sin embargo, estos componentes están fusionados y no hay forma de conocer cuál es la contribución individual de cada uno.

---

49 Agrónomo estadounidense. Premio Nobel de la Paz. <http://www.agbioworld.org/biotech-info/topics/borlaug/quotes.html>

50 Algunas características fenotípicas se deben a la acción de un solo gene. Estas características monogénicas como color del tallo, forma de las hojas, marcas en la piel de animales, son poco afectadas por el ambiente. Las características fenotípicas de interés económico, como rendimiento, producción de leche, contenido de aceite deben su expresión al efecto combinado de varios genes. El ambiente influye grandemente en la manifestación de estas características poligénicas.

Cuando experimentos iguales se llevan a cabo en diferentes localidades o diferentes épocas de siembra es posible combinar los análisis de varianzas de las localidades en un solo análisis. Muchos de estos trabajos se llevan a cabo para identificar las mejores variedades y prácticas agronómicas para diferentes zonas y épocas de siembra. Pero también, dependiendo del germoplasma y del modelo que se emplee, se pueden separar los efectos genéticos de los ambientales. Esto último es un campo de estudio bastante amplio que no se tratará en este texto. Una buena referencia al respecto es el libro de genética cuantitativa de Hallauer y Miranda (1981).

### **21.1. Homogeneidad de las varianzas**

Para combinar los resultados de varios experimentos en un solo análisis, las varianzas de los errores experimentales (cuadrados medios) deben ser estadísticamente iguales. Si no se cumple con este requisito los resultados que se obtengan no serán confiables.

Cuando se trata de combinar dos experimentos puede emplearse la prueba estadística F para la homogeneidad de las varianzas de los errores experimentales. En los ejercicios 9.1.1 y 16.4.1 hay ejemplos de su uso. Cuando se trata de combinar tres o más experimentos, el procedimiento adecuado es la prueba de Bartlett para la homogeneidad de las varianzas. Si se concluye que las varianzas de los errores experimentales son estadísticamente iguales pueden combinarse los experimentos en un solo análisis. Si por lo contrario se decide que las varianzas son heterogéneas, se excluyen del análisis los experimentos que presenten las varianzas más altas. Otra opción es transformar los datos de los experimentos que se desee combinar. Sin embargo, no siempre la transformación de datos corrige el problema de la heterogeneidad de las varianzas. Los temas relacionados con la transformación de datos se verán en el Capítulo 23.

Cuando el número de experimentos que se pretende combinar es muy elevado se vuelve engorroso el uso de la prueba de Bartlett. En este caso puede fijarse un límite para el coeficiente de variación de los experimentos individuales y descartar aquellos que lo sobrepasen. Como ejemplo, en coordinación con el Centro Internacional de Mejoramiento del Maíz y Trigo (CIMMYT) se evaluaban en Centroamérica y el Caribe 49 híbridos de maíz en 20 o más localidades<sup>51</sup>. El diseño era un látice parcialmente balanceado con dos repeticiones. Los experimentos que tenían coeficientes de variación de 20% o más se excluían del análisis combinado.

---

51 Programa Colaborativo Centroamericano para el Mejoramiento de Cultivos Alimenticios (PCCMCA). [http://www.mag.go.cr/rev\\_meso/v01n01\\_093.pdf](http://www.mag.go.cr/rev_meso/v01n01_093.pdf)

### La prueba de Bartlett

La prueba de Bartlett se basa en la distribución Ji-cuadrado ( $\chi^2$ ) y puede usarse en el análisis de cualquier experimento cuando se sospecha que las varianzas no son homogéneas. La fórmula (127) para encontrar el valor de la Prueba de Bartlett  $\chi^2$  es la siguiente:

$$PB\chi^2 = \frac{f(k \cdot \ln \bar{S}^2 - \sum \ln S_i^2)}{C}$$

En la fórmula:

- $f$  es el número de grados de libertad de cada varianza
- $k$  es el número de tratamientos. Para los análisis combinados  $k$  es el número de experimentos
- $\ln \bar{S}^2$  es el logaritmo natural del promedio de las varianzas:  $\bar{S}^2 = \sum S_i^2 / k$
- $\sum \ln S_i^2$  es la sumatoria de los logaritmos naturales de las varianzas
- $C$  es un factor de ajuste:  $C = 1 + \frac{k+1}{3kf}$

La fórmula (127) se aplica cuando las varianzas tienen igual número de grados de libertad. De no ser este el caso, la fórmula se modifica para compensar por esta variación (Snedecor y Cochran, 1967). Sin embargo, en el caso de experimentos combinados, las varianzas del error experimental tienen el mismo número de grados de libertad.

### Modelo lineal aditivo

El modelo lineal aditivo para los experimentos combinados es:

$\bar{Y}_{ij} = \mu + \pi_i + \tau_j + \pi_i\tau_j + \bar{\varepsilon}_{ij}$  fórmula en la que:

- $\bar{Y}_{ij}$  es la media observada de un tratamiento en una localidad o época de siembra
- $\mu$  es la media poblacional de todo el experimento combinado
- $\pi_i$  es el efecto de la localidad o época de siembra
- $\tau_j$  es el efecto del tratamiento
- $\pi_i\tau_j$  es el efecto de la interacción entre localidades (o épocas de siembra) y tratamientos
- $\bar{\varepsilon}_{ij}$  es el error promedio de todas las parcelas del experimento.

El anterior modelo se refiere al caso de experimentos iguales, con un solo factor en estudio. Cuando se trata de experimentos factoriales deben incluirse en el modelo los términos correspondientes.

### *Cuadrados medios esperados*

De acuerdo con el modelo, los valores esperados de las sumas de cuadrados se presentan en el Cuadro 21.1 (Cochran y Cox, 1968).

En el Cuadro 21.1 no consta el cuadrado medio esperado de las localidades, pero éstas pueden ser una muestra al azar de sitios de toda una región o una muestra de sitios seleccionados por el investigador por razón de su potencial de producción. Este último caso es el más común, por lo que se trataría de un modelo fijo, tanto para localidades como para tratamientos.

**Cuadro 21.1. Cuadrados medios esperados de experimentos en diferentes localidades**

Fuente de variación	Cuadrados medio esperados
Tratamientos	$\sigma_{\varepsilon}^2 + r\sigma_{\pi\tau}^2 + \frac{r\ell}{(k-1)} \sum (\tau_j - \bar{\tau})^2$
Tratamientos*localidades	$\sigma_{\varepsilon}^2 + r\sigma_{\pi\tau}^2$
Error experimental combinado	$\sigma_{\varepsilon}^2$

Según los cuadrados medios esperados, el efecto de los tratamientos se estima mediante  $\frac{r\ell}{(k-1)} \sum (\tau_j - \bar{\tau})^2$ , donde  $r$ ,  $\ell$  y  $k$  son los números de repeticiones, localidades y tratamientos, respectivamente.

El Cuadro 21.1 sirve para identificar el correcto denominador para la prueba F de tratamientos. Si la interacción es estadísticamente significativa, debe emplearse el cuadrado medio de la interacción como denominador de la prueba F. Si la interacción no es significativa, es decir que se supone que  $\sigma_{\pi\tau}^2$  es igual a cero, puede sumarse la interacción al error experimental combinado y emplearlo como denominador para la prueba F de tratamientos.

### *Esquema del análisis de varianza*

En términos generales, el esquema para el análisis de varianza es el siguiente:

Cuadro 21.2. Esquema del ADEVA para experimentos combinados

F.V.	S.C.	G.L.	C.M.	Prueba F
Localidades, Loc	$SCL$	$l-1$	$CM_{loc}$	$CM_{loc}/CMEa$
Reps dentro de Loc (Error a, Loc*Rep)	$SCR/L$	$l(r-1)$	$CMEa$	
Variedades, Var	$SCV$	$v-1$	$CM_{var}$	$CM_{var}/CMEb$
Loc*Var	$SC(L*V)$	$(l-1)(v-1)$	$CM(l*v)$	
Error combinado (Error b, Loc*Rep*Var)	$SCE$	$l(r-1)(v-1)$	$CMEb$	
TOTAL	$SCT$	$lrv-1$		

Veamos un ejemplo:

**Ejercicio 21.1.** Se realizaron cuatro experimentos en zonas productoras de fréjol para evaluar la capacidad de producción de cinco variedades de distinta procedencia. Todos los experimentos se condujeron con un diseño de Bloques Completos al Azar (BCA) con tres repeticiones. La variable dependiente estudiada fue el rendimiento en decagramos por parcela. Los resultados de las cuatro localidades se presentan en el Cuadro 1.

Los análisis de varianza de los cuatro experimentos, cada uno en una localidad diferente, se presentan en el Cuadro 2.

Cuadro 1. Rendimiento, en hg/parcela, de cinco variedades de fréjol (A, B, C, D, E) en cuatro localidades (L) y tres bloques o repeticiones (R).

Localidades y Repeticiones	A	B	C	D	E	Localidades y repeticiones	A	B	C	D	E
L1 R1	72.5	53.8	72.7	62.4	26.0	L3 R1	48.2	52.4	62.7	46.4	35.8
L1 R2	71.9	35.9	82.0	49.7	36.5	L3 R2	55.2	44.1	65.2	51.4	51.0
L1 R3	67.3	38.5	70.1	57.7	41.3	L3 R3	62.4	43.2	70.1	49.4	39.0
L2 R1	59.5	50.3	73.0	62.7	45.3	L4 R1	65.9	43.1	84.6	56.1	41.8
L2 R2	70.6	47.3	71.6	51.8	32.5	L4 R2	60.6	47.9	78.2	42.4	36.0
L2 R3	61.3	42.0	63.5	50.8	54.6	L4 R3	64.1	40.8	67.6	47.2	51.6

**Cuadro 2. Análisis de varianza de experimentos  
conducidos en cuatro diferentes localidades**

Localidad 1					Localidad 2				
E.V.	S.C.	g.l.	C.M.	F	E.V.	S.C.	g.l.	C.M.	F
Rep	19.16	2	9.58		Rep	42.50	2	21.25	
Var	3614.76	4	903.69	15.5**	Var	1411.58	4	352.89	6.3*
Error	467.32	8	58.42		Error	449.80	8	56.23	
Total	4101.24	14			Total	1903.88	14		
$\bar{Y} = 55.9$ $CV = 13.7\%$					$\bar{Y} = 55.7$ $CV = 13.4\%$				

Localidad 3					Localidad 4				
E.V.	S.C.	g.l.	C.M.	F	E.V.	S.C.	g.l.	C.M.	F
Rep	54.12	2	27.06		Rep	76.23	2	38.12	
Var	1037.59	4	259.40		Var	2557.64	4	639.41	15.4**
Error	267.59	8	33.45	7.8**	Error	332.98	8	41.62	
Total	1359.29	14			Total	2966.85	14		
$\bar{Y} = 51.8$ $CV = 11.2\%$					$\bar{Y} = 55.2$ $CV = 11.7\%$				

El siguiente paso es efectuar la prueba de Bartlett de homogeneidad de las *varianzas o cuadrados medios del error*. Sobre esta base, si las varianzas no son iguales se procede a transformar los datos o, en su defecto, a descartar las localidades con los CM más altos. Si hay evidencia de que las varianzas son homogéneas, entonces se emplean los datos originales (Cuadro 1) para combinar los resultados de las cuatro localidades en un solo análisis de varianza.

**Cuadro 3. Cálculos preliminares para la prueba de Bartlett**

	Localidades				Sumatoria	$\bar{S}^2$	$\ln \bar{S}^2$
	1	2	3	4			
$S_i^2$	58.42	56.23	33.45	41.62	189.72	47.43	3.8593
$\ln S_i^2$	4.0677	4.0295	3.5101	3.7286	15.3357		

En la fórmula (127),  $f = 8$  y  $k = 4$ . Estos son los grados de libertad del error experimental y el número de experimentos para el análisis combinado. Con esta información, la prueba estadística de Bartlett es:

$$PB\chi^2 = \frac{8 * (4 * 3.8593 - 15.3357)}{1.0521} = 0.78$$

El valor de  $\chi^2$  al nivel de 5% y con  $k-1 = 3$  grados de libertad es igual a 7.81, de acuerdo con la tabla respectiva (Apéndice C). Por lo tanto, como el valor de Ji-cuadrado de la prueba estadística de Bartlett no supera el valor de Ji-cuadrado de la tabla, se puede afirmar que hay suficiente evidencia estadística de que las varianzas son homogéneas. El análisis combinado, con los datos originales, puede efectuarse.

Mucha de la información para el análisis de varianza combinado está contenida en los análisis de varianza individuales presentados en el Cuadro 2. Sin embargo, de acuerdo con el esquema del ADEVA para el análisis combinado (Cuadro 21.2) deben encontrarse la suma de cuadrados de las variedades (*SCV*), la suma de cuadrados de las localidades (*SCL*) y la suma de cuadrados de la interacción variedades\*localidades (*SC(V\*L)*). Para esto, en forma similar a lo que se hizo en el caso de los experimentos factoriales, a partir de los datos del Cuadro 1 se estructura una tabla de doble entrada como el que se presenta en el Cuadro 4:

Cuadro 4. Variable rendimiento, tabla de doble entrada: Variedades\*Localidades.

Localidades	Variedades					$\sum Loc_i$
	1	2	3	4	5	
1	211.7	128.2	224.8	169.8	103.8	838.3
2	191.4	139.6	208.1	165.3	132.4	836.8
3	165.8	139.7	198.0	147.2	125.8	776.5
4	190.6	131.8	230.4	145.7	129.4	827.9
$\sum Var_j$	759.5	539.3	861.3	628.0	491.4	$\sum Y_{ij} = 3279.5$

Primero debe encontrarse el Factor de Corrección:

$$FC = \frac{3279.5^2}{60} = 179252$$

luego se requiere calcular la suma de cuadrados del cuadro  $V*L$  (datos en *italica*) y los demás componentes:

$$SCTrat_{kk} = \frac{\sum T_{kk}^2}{2k} - FC = \frac{441.0^2 + 467.0^2 + 359.0^2 + 407.0^2}{8} - 179252$$



$$SCV = \frac{759.5^2 + 539.3^2 + \dots + 491.4^2}{12} - FC = 7863.03$$

$$SCL = \frac{833.3^2 + 836.8^2 + 776.5^2 + 827.9^2}{15} - FC = 171.45$$

$$SC(V*L) = SCC(V*L) - SCV - SCL = 758.53$$

Con esta información y la del Cuadro 2 se puede organizar el Cuadro ADE-VA combinado que se presenta en el Cuadro 5:

**Cuadro 5. Análisis de varianza combinado para evaluar la respuesta de cinco variedades de fréjol en cuatro localidades**

E.V.	S.C.	G.L.	C.M.	Prueba F
Localidades, Loc	171.45	3	57.15	2.38 ns
Reps dentro de Loc (Error Loc*Rep)	192.00	8	24.00	
Variedades, Var	7863.03	4	1965.76	41.45**
Var*Loc	758.53	12	63.21	1.33 ns
Error combinado (Error Loc*Rep*Var)	1517.7	32	47.43	
TOTAL	10502.7	59		

Algunas de las sumas de cuadrados que se presentan en el Cuadro 5 provienen de los análisis de varianza individuales que se presentaron en el Cuadro 2. Estas son:

- La Suma de Cuadrados Total (10502.7) se obtiene al añadir las cuatro sumas de cuadrados totales de los análisis individuales (10331.3), más las sumas de cuadrados de las localidades (171.4).
- La suma de cuadrados del error combinado (1517.7) es el resultado de la suma de los errores (SCE) de las cuatro localidades.
- El error Localidades\*Repeticiones (192.0) se obtiene al añadir las SCR de las cuatro localidades.

El error Localidades\*Repeticiones es la medida para evaluar diferencias entre localidades. La prueba F correspondiente (Cuadro 5) no fue significativa. En todo caso, Cochran y Cox (1968) y Snedecor y Cochran (1967) advierten de posibles sesgos en la interpretación de estos resultados si la muestra de localidades no fue tomada al azar.

La interacción Variedades\*Localidades tampoco fue significativa, por lo que los efectos de los tratamientos (variedades) se miden directamente, de acuerdo con los cuadrados medios esperados, con el error combinado. El valor de la prueba F indica que hubo diferencias significativas entre las variedades al nivel de 1%. El método de separación de medias de Student-Newman-Keuls o algún otro (Capítulo 17) podría usarse para identificar las variedades más promisorias.

Con el programa *Statistix* los resultados del análisis combinado son los siguientes:

Statistix 9.0 LOCS4 REPS3 VAR5 Frejol,

#### Analysis of Variance Table for Rend

Source	DF	SS	MS	F	P
LOC	3	171.4	57.15	2.38	0.1453
Error LOC*REP	8	192.0	24.00		
VAR	4	7863.0	1965.76	41.45	0.0000
LOC*VAR	12	758.5	63.21	1.33	0.2489
Error LOC*REP*VAR	32	1517.7	47.43		
Total	59	10502.7			

Grand Mean 54.658

CV(LOC\*REP) 8.96

CV(LOC\*REP\*VAR) 12.60

#### Comentarios:

- En los experimentos combinados se espera que la interacción Variedades\*Localidades sea significativa. Si esto ocurre, pueden identificarse las mejores variedades o genotipos para zonas específicas. Que el comportamiento de diferentes variedades sea similar a través de distintas localidades, como en el caso del ejercicio anterior, puede indicar que las condiciones del suelo, clima y manejo del ensayo fueron similares en las cuatro localidades seleccionadas o, tal vez lo más probable, que las variedades eran homoestáticas, es decir, que sus características y propiedades no variaron significativamente en respuesta a diferencias ambientales.
- En la producción agrícola es importante identificar qué genotipos, en promedio, producen más que otros en diferentes condiciones ambientales. Esta característica se conoce como *estabilidad*. Con la información de experimentos en diferentes localidades se pueden identificar estos genotipos.

Una metodología muy usada para este proceso la desarrollaron Eberhart y Russell (1966). Sobre esta base se pueden hacer recomendaciones que cubran zonas más amplias de cultivo.

## 21 .2. Experimentos combinados, épocas de siembra

En experimentos para evaluar el rendimiento de líneas promisorias de cultivos o para comparar diferentes prácticas de manejo, se recomienda llevar a cabo ensayos en diferentes épocas de siembra. Esto permite estudiar la posible interacción entre los tratamientos (líneas promisorias, prácticas de manejo, etcétera) y las épocas. El procedimiento es similar al empleado para localidades.

Como las épocas de siembra tienen por lo general patrones climáticos más o menos constantes, se considera que los efectos son fijos. Por lo tanto, lo señalado en el Cuadro 21.1 y 21.2, con relación a localidades, tiene también aplicación para este caso. Sin embargo, si se trata de combinar experimentos de diferentes años, el caso es distinto. Los patrones climáticos varían de año a año y por lo tanto el modelo es al azar. Esto tiene aplicación para experimentos con cultivos anuales y perennes. Las características del modelo al azar pueden encontrarse en las referencias citadas.

En el siguiente ejemplo se analiza un experimento factorial conducido con un diseño de Bloques Completos al Azar durante dos épocas de siembra. En Centroamérica se denomina “primera” la época que se inicia con las lluvias de abril o mayo. La segunda época de siembra, que generalmente se inicia en septiembre, se denomina “postrera”.

---

**Ejercicio 21.2.** En un trabajo de investigación, Andrade, 1992<sup>52</sup> estudió, durante las épocas de siembra de primera y postrera, el efecto de varios factores sobre la Razón de Equivalencia de la Tierra (RET). La RET es un índice de la eficiencia de los cultivos asociados en comparación con los monocultivos. Como ejemplo, un valor de RET igual a 1.50 indica que para obtener la misma producción de los cultivos asociados se requiere un 50% más de tierra en los monocultivos.

Por facilidad del ejemplo se analizarán únicamente los factores: soya (variedad Siatsa-194) asociada con maíz y sorgo, y dos dosis de fertilización nitrogenada. Para calcular la RET se sembraron, junto a las parcelas de cultivos asociados,

---

52 Andrade, J.C. 1991. Evaluación agronómica de cultivos asociados soya-maíz, soya-sorgo en el valle del Zamorano y en fincas de agricultores. Tesis. Escuela Agrícola Panamericana. Honduras, C.A.

parcelas con monocultivos de soya, maíz y sorgo. Los experimentos se condujeron con diseños de bloques completos al azar. Los resultados se presentan en el siguiente cuadro:

Cuadro 1. Valores de RET\* para la asociación soya – maíz y soya – sorgo

G	N	E1R1	E1R2	E1R3	E1R4	E2R1	E2R2	E2R3	E2R4
1	1	1.42	1.61	1.22	1.32	1.36	1.12	1.22	1.24
1	2	1.74	1.79	1.56	1.86	1.45	1.05	1.34	1.38
2	1	1.56	1.35	1.75	1.64	1.39	1.34	1.18	1.28
2	2	1.84	1.74	1.62	1.70	1.64	1.28	1.14	1.15
3	1	1.15	1.18	1.34	1.12	1.86	1.88	1.56	1.65
3	2	1.38	1.44	1.50	1.56	1.74	1.94	1.45	1.73
4	1	1.12	1.24	1.23	1.32	1.68	1.77	1.60	1.88
4	2	1.36	1.45	1.56	1.46	1.45	1.48	1.65	1.75

En el Cuadro 1, las letras y números significan lo siguiente:

- La letra E seguida de un número indica la época del ensayo: 1=Primera, 2=Postrera
- La letra R marca el número de la repetición o bloque
- La letra G señala las gramíneas en asociación con soya: 1) maíz H-27, 2) maíz HPB, 3) sorgo sureño y 4) sorgo Isiap-dorado.
- La letra N representa las dosis de nitrógeno: 1) 0 kg N/ha y 2) 100 kg N/ha.

En el Cuadro 2 constan los análisis de varianza correspondientes a cada época de siembra:

Cuadro 2. Análisis de varianza de la variable RET en las dos épocas de siembra

Primera					Postrera				
F.V.	S.C.	g.l.	C.M.	F	F.V.	S.C.	g.l.	C.M.	F
Rep	0.0106	3	0.0035		Rep	0.1316	3	0.0438	
G	0.6097	3	0.2032	13.35**	G	1.3468	3	0.4489	19.68**
N	0.4975	1	0.4975	32.67**	N	0.0047	1	0.0047	0.21 n.s
G*N	0.0398	3	0.0132	0.87 n.s	G*N	0.0511	3	0.0170	0.75 n.s
Error	0.3198	21	0.0152		Error	0.4791	21	0.0228	
Total	1.4774	31			Total	2.0136	31		
$\bar{Y} = 1.47$ CV = 8.38					$\bar{Y} = 1.49$ CV = 10.15				

Para decidir si los dos experimentos pueden combinarse en uno solo se requiere de una prueba F para la homogeneidad de las varianzas (cuadrados medios del error). Si las varianzas son estadísticamente homogéneas, los dos experimentos pueden combinarse en un solo análisis:

La prueba es:  $PEF = 0.0228/0.0152 = 1.50$ . El valor de F en la tabla al nivel del 1%, con 21 gl para el numerador y 21 gl para el denominador, es igual a 2.86. Como el valor de F calculado (1.50) no supera el valor de la tabla, se concluye que las dos varianzas son homogéneas.

El procedimiento para el análisis de varianza combinado para épocas de siembra es similar al que se presentó para localidades. Sin embargo, en este ejemplo, debe tenerse en cuenta que se trata de un experimento factorial llevado a cabo en dos épocas de siembra. El análisis combinado es el siguiente:

Statistix 9.0            AAJCfinalcomb

**Analysis of Variance Table for RET**

Source	DF	SS	MS	F	P
Epoca	1	0.00391	0.00391	0.16	0.6989
Error Epoca*REP	6	0.14224	0.02371		
G	3	0.10928	0.03643	1.91	0.1419
N	1	0.20250	0.20250	10.64	0.0022
Epoca*G	3	1.84732	0.61577	32.37	0.0000
Epoca*N	1	0.29976	0.29976	15.76	0.0003
G*N	3	0.06265	0.02088	1.10	0.3607
Epoca*G*N	3	0.02832	0.00944	0.50	0.6869
Error Epoca*REP*G*N	42	0.79901	0.01902		
Total	63	3.49498			

Grand Mean 1.4806

CV(Epoca\*REP) 10.40

CV(Epoca\*REP\*G\*N) 9.32

En el Cuadro ADEVA se observa que hubo dos interacciones altamente significativas: época por gramíneas y época por nitrógeno. Además, se observa que el efecto del nitrógeno fue también significativo. Para interpretar las interacciones se presentan los siguientes gráficos:

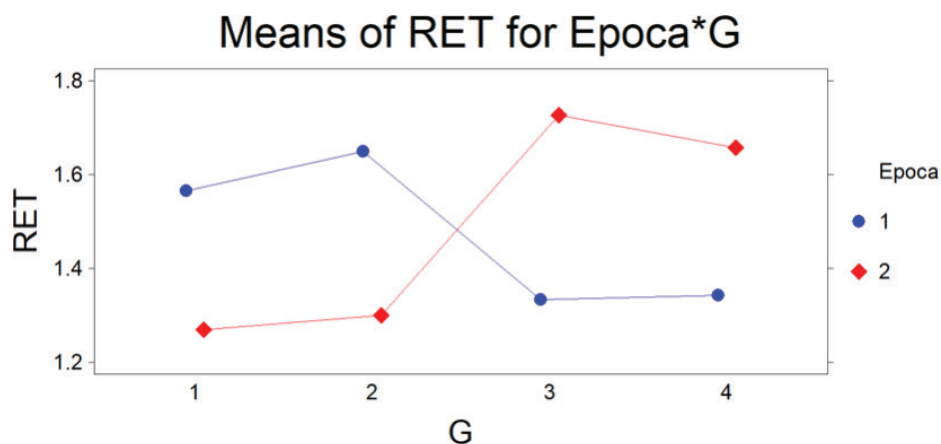


Gráfico 1. Interacción estadísticamente significativa entre épocas y gramíneas

Puede notarse que el valor de RET fue mayor para las dos variedades de maíz (G1 y G2) en la época de siembra primera. Esta es la época de lluvias y de días más largos en Honduras. En cambio, en la época postrera, el RET de las variedades de sorgo fue mayor que en el caso del maíz. Este resultado no sorprende ya que el sorgo tiene estructuras que le permiten soportar condiciones de sequía.

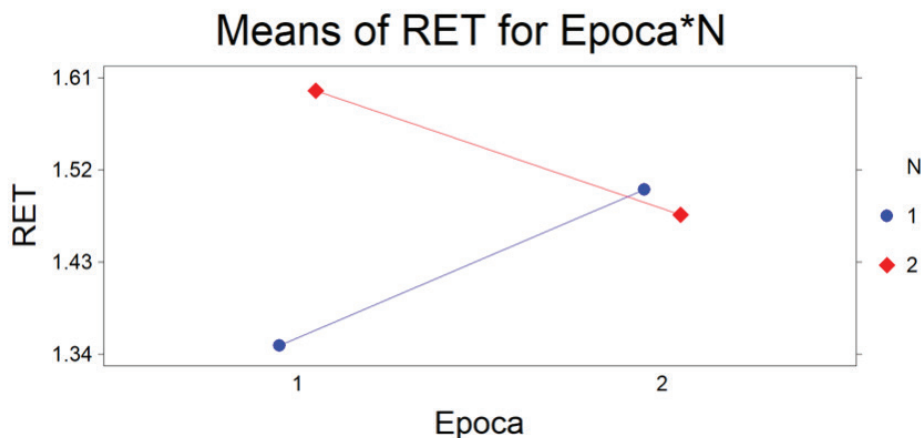


Gráfico 2. Interacción estadísticamente significativa entre épocas y nitrógeno

En el Gráfico 2 se observa que en la época primera el valor RET fue afectado en forma diferente por los niveles de nitrógeno. Sin embargo, en la época seca (postrera) no hubo ma-

yor diferencia. Esto se explica por un mayor desarrollo de las plantas en la época lluviosa y en la presencia de una dosis alta de nitrógeno. Sin embargo, en la época seca (postrera) no hubo efecto del nitrógeno sobre el valor RET. Este resultado pudo esperarse, ya que en condiciones de sequía las plantas no aprovechan el fertilizante que está en el suelo.

*“La formulación de un problema es a menudo más importante que su solución, que simplemente puede ser cuestión de destreza matemática o experimental. Plantear nuevas preguntas, nuevas posibilidades, mirar viejos problemas desde un nuevo ángulo requiere imaginación creativa y marca avances reales en la ciencia”.*  
Albert Einstein<sup>53</sup>

En el análisis de covarianza se examinan simultáneamente las varianzas y covarianzas. No es un diseño experimental sino una técnica que puede ser usada con cualquier diseño. El análisis de covarianza se lo usa especialmente para incrementar la precisión de los experimentos.

El procedimiento en sí es simple: se miden las variables dependientes y una o más variables asociadas que reciben el nombre de *covariables*. Si la variable dependiente es rendimiento de grano, las covariables podrían ser número de plantas por parcela, porcentaje de daño por insectos, cantidad inicial de algún elemento esencial en cada parcela, etc. Sin embargo, en la mayoría de los casos se trabaja con una sola covariable.

Para explicar lo que se logra con el análisis de covarianza se presenta el siguiente ejemplo. Supongamos que la variable que nos interesa es el rendimiento de varios cultivares de fréjol. Los resultados experimentales se deberán a dos variables: 1) diferencias genéticas entre los cultivares y 2) efectos ambientales diferentes, bióticos y abióticos, en cada parcela. Por ejemplo, si los niveles de competencia no fueron los mismos en las parcelas, un análisis de covarianza, en el que se use como covariable el número de plantas por parcela, permitirá remover la variación en rendimiento por causa del diferente número de plantas. Es decir, el análisis se realiza en tal forma como si todas las parcelas tuvieran el mismo número de plantas.



### *Ventajas del análisis de covarianza*

El empleo del análisis de covarianza tiene las siguientes ventajas:

- Aumenta la precisión de los experimentos al disminuir el error experimental
- En los casos en que la variación entre las unidades experimentales no siga un patrón fijo y no se pueda agrupar las unidades experimentales en bloques o hileras y columnas, el análisis de covarianza remueve parte de la variación entre unidades experimentales que recibieron el mismo tratamiento
- El análisis de covarianza puede planificarse de antemano, pero no es un requisito indispensable
- El análisis de covarianza reduce la variación que puede resultar de daños impredecibles por plagas y patógenos, pero es indispensable cuantificar esos daños lo más objetivamente que sea posible.

### *Requisitos para el uso de la covarianza*

Para el uso adecuado del análisis de covarianza hay dos condiciones básicas:

- se requiere que la variable y la covariable estén relacionadas biológica o físicamente. De lo contrario, los resultados son inválidos y carecen de sentido. En el ejemplo de rendimiento de cultivos de fréjol y el número de plantas, las dos variables están asociadas ya que el número de plantas por parcela influye en el rendimiento. Por lo general, se supone una asociación lineal entre las dos variables, pero es posible y a veces común la existencia de una asociación cuadrática o de otro grado más alto. En este caso la covarianza debe ser también del grado más alto.
- Es importante que la covariable no sea afectada por los tratamientos. Si los tratamientos son diferentes herbicidas y algunos de estos reducen el número de plantas, no es correcto usar el número de plantas como covariable para corregir el rendimiento.

### *Pasos para el análisis de covarianza*

El análisis de Covarianza es una técnica que puede usarse con cualquiera de los diseños, sean estos experimentos de uno o varios factores. Una vez que se han obtenido los datos de la variable principal ( $Y_i$ ) y de la covariable ( $X_i$ ), se procede de la siguiente forma:

**Primer paso:** se obtienen las sumas de cuadrados y se generan los análisis de varianza en la forma usual, de acuerdo con el diseño experimental empleado, tanto para la variable Y como para la covariable X.

**Segundo paso:** para efectuar el análisis de covarianza se estructura un cuadro que contenga las sumas de cuadrados (encontrados en el primer paso) y las sumas de productos.

Para el diseño completamente al azar (DCA) las sumas de productos se encuentran con las siguientes fórmulas:

1. Factor de Corrección:

$$FC_{xy} = \frac{\sum X_i \sum Y_i}{rk} \quad (128)$$

2. Suma de Productos Total

$$SPT = \sum X_i Y_i - FC_{xy} \quad (129)$$

3. Suma de Productos de los Tratamientos

$$SPT_{rat} = \frac{\sum T_x T_y}{r} - FC_{xy} \quad (130)$$

4. Suma de Productos del Error

$$SPE = SPT - SPT_{rat} \quad (131)$$

A partir de la información anterior pueden deducirse las fórmulas necesarias para otros diseños.

**Tercer paso:** es necesario, a continuación, encontrar las sumas de cuadrados corregidas en función de la covariable. Las fórmulas son las siguientes:

1. Suma de Cuadrados Corregida Total (Fórmula 132):

$$SCCT_y = SCT_y - \frac{(SPT_{xy})^2}{SCT_x}$$

2. Suma de Cuadrados Corregida del Error (Fórmula 133):

$$SCCE_y = SCE_y - \frac{(SPE_{xy})^2}{SCE_x}$$

## 3. Suma de Cuadrados Corregida de Tratamientos (Fórmula 134)

$$SCC_{Trat_y} = SCCT_y - SCCE_y$$

Veamos un ejemplo:

**Ejercicio 22.1.** En un experimento se evaluó el rendimiento de seis variedades de soya con un diseño completamente al azar (DCA). Cada cultivar se asignó a cuatro parcelas, es decir, hubo cuatro repeticiones ( $r=4$ ). La germinación de la semilla no fue uniforme, por lo que se decidió emplear el número de plantas por parcela como covariable. Se pide realizar los análisis pertinentes con los resultados que se presentan en el Cuadro 1.

Cuadro 1. Rendimiento en kilogramos por parcela (variable Y) y número de plantas (covariable X) en la evaluación de seis cultivares de soya

Variedades	Valores de X y Y								Totales Tratamientos (Variedades)	
	X1	Y1	X2	Y2	X3	Y3	X4	Y4	$T_x$	$T_y$
A	24	1.25	20	1.1	27	1.43	22	1.38	93	5.16
B	34	1.75	32	1.54	28	1.38	26	1.46	120	6.13
C	33	1.65	18	1.38	23	1.22	32	1.43	106	5.68
D	35	1.68	25	1.24	30	1.45	23	1.04	113	5.41
E	34	1.83	25	1.74	22	1.48	28	1.6	109	6.65
F	31	1.58	33	1.64	26	1.37	29	1.47	119	6.06
Sumatorias de X y de Y									$\sum X_i = 660$	$\sum Y_i = 35.09$

A partir de las observaciones del Cuadro 1, mediante los procedimientos estudiados en los capítulos 14 y otros capítulos se realizan los análisis de varianza tanto para la variable principal (Cuadro 2), como para la covariable (Cuadro 3). Estos son:

Cuadro 2. Análisis de varianza del rendimiento de seis variedades de soya

F.V.	S.C.	G.L.	C.M.	F	$F_{0.05}$	$F_{0.01}$
Variedades	0.3653	5	0.0731	2.28 n.s.	2.77	4.25
Error	0.5771	18	0.0321			
Total	0.9424	23				

$$\bar{Y} = 1.46 \text{ kg/parcela } CV = 12.25\%$$

Cuadro 3. Análisis de varianza del número de plantas de soya por parcela

F.V.	S.C.	G.L.	C.M.	F	F <sub>0.05</sub>	F <sub>0.01</sub>
Variedades	124.00	5	24.80	1.07 n.s.	2.77	4.25
Error	416.00	18	23.11			
Total	540.00	23				

$$\bar{X} = 27.50 \text{ plantas/parcela } CV = 17.48\%$$

Sobre la base del Cuadro 1 se forma una tabla con los productos de X por Y, como se presenta en el Cuadro 4:

Cuadro 4. Tabla de productos X por Y

Variedades (Tratamientos)	X1Y1	X2Y2	X3Y3	X4Y4	Tratamientos $T_{xy}$
A	30.00	22.00	38.61	30.36	120.97
B	59.50	49.28	38.64	37.96	185.38
C	54.45	24.84	28.06	45.76	153.11
D	58.80	31.00	43.50	23.92	157.22
E	62.22	43.50	32.56	44.80	183.08
F	48.98	54.12	35.62	42.63	181.35
Sumatoria de X*Y					$\sum X_i Y_i = 981.11$

Con la información de los Cuadros 1 y 4 se procede a calcular las sumas de productos. El factor de corrección es lo primero que debe calcularse. Luego se calcula la suma de productos total, la suma de productos de los tratamientos y finalmente, la suma de productos del error experimental:

- Factor de Corrección:

$$FC_{xy} = \frac{660 \cdot 35.09}{24} = 964.98$$

- Suma de Productos Total

$$SPT = (30.00 + 59.50 + \dots + 42.63) - 964.98 = 16.1350$$

- Suma de Productos de los Tratamientos

$$SPT_{rat} = \frac{(120.97 + 185.38 + \dots + 181.35)}{r = 4} - 964.98 = 3.7450$$

## 4. Suma de Productos del Error

$$SPE = SPT - SPT_{rat} = 16.1350 - 3.7450 = 12.3900$$

Con la información generada se estructura un cuadro con las sumas de cuadrados y de productos. Esto se presenta en el Cuadro 5:

Cuadro 5. Sumas de Cuadrados y Productos

F.V.	G.L.	Sumas de Cuadrados y Productos		
		X*X	X*Y	Y*Y
Tratamientos	5	124	3.7450	0.3653
Error	18	416	12.3900	0.5771
Total	23	540	16.1350	0.9424

Los datos del cuadro anterior sirven para calcular las sumas de cuadrados corregidas, o ajustadas, de Y en función de X. Estas son:

## 1. Suma de Cuadrados Corregida Total

$$SCCT_y = 0.9424 - \frac{(16.1350)^2}{540} = 0.4603$$

## 2. Suma de Cuadrados Corregida del Error

$$SCCE_y = 0.5771 - \frac{(12.3900)^2}{416} = 0.2081$$

## 3. Suma de Cuadrados Corregida de Tratamientos

$$SCCTrat_y = 0.4603 - 0.2081 = 0.2522$$

Se dispone ya de la información completa para formar el cuadro de análisis de covarianza. Esto se presenta en el Cuadro 6:

Cuadro 6. Análisis de Covarianza (ANCOVA) de la variable Y, rendimiento de seis variedades de soja, corregido en función de la covariable X, número de plantas

F.V.	S.C.C.	G.L.	C.M.C.	F	F <sub>0.05</sub>	F <sub>0.01</sub>
Variedades	0.2522	5	0.0504	4.13 *	2.81	4.34
Error	0.2081	17	0.0122			
Total	0.4603	22				

$$\bar{Y} = 1.46 \text{ kg/parcela } CV = \frac{\sqrt{CMCE}}{\bar{Y}} * 100 = \frac{\sqrt{0.0122}}{1.46} * 100 = 7.57\%$$

Note que se pierde un grado de libertad del error por el efecto de la covariable.

Para conocer si la covarianza contribuyó a mejorar la precisión del análisis estadístico puede compararse el Cuadro 2 con el Cuadro 6. En el Cuadro 2, el valor de la prueba F fue igual a 2.28, una cantidad no significativa por ser inferior al nivel alfa del 0.05 (Apéndice D). Por el contrario, en el Cuadro 6, el valor de la prueba F fue igual a 4.13. Como este valor supera el nivel alfa del 0.05 puede concluirse que, una vez que se removió el efecto del diferente número de plantas, afloraron diferencias estadísticas entre las medias de rendimiento de las variedades de soya. También puede observarse, en los dos cuadros referidos, la diferencia entre los coeficientes de variación.

El incremento en precisión puede cuantificarse mediante el cálculo de la eficiencia relativa entre el análisis de covarianza y el análisis de varianza. La fórmula para esto es (135):

$$RE = \frac{CME_y}{CMCE_y * \left(1 + \frac{CMTrat_x}{SCE_x}\right)} = \frac{0.0321}{0.0122 * \left(1 + \frac{24.80}{416.00}\right)} = 2.47$$

Este valor indica que la precisión que se ganó está alrededor de un 147% (2.47-1.00). Una cantidad que demuestra lo eficiente que puede ser el análisis de covarianza.

Como de acuerdo con el análisis de covarianza hubo diferencias significativas entre las medias, el siguiente paso es encontrar las medias corregidas de las variedades. La fórmula para calcular las medias corregidas es:  $\bar{Y}_{ic} = \bar{Y}_i - b_\epsilon(\bar{X}_i - \bar{\bar{X}})$ , fórmula en la que:

- $\bar{Y}_{ic}$  es la media corregida de la variedad (o tratamiento) en cuestión
- $\bar{Y}_i$  es la media sin corregir de la variedad (o tratamiento) en cuestión
- $b_\epsilon$  es el coeficiente de regresión del error
- $\bar{X}_i$  es la media de la covariable correspondiente al tratamiento en cuestión
- $\bar{\bar{X}}$  es la gran media de la covariable (el promedio de todas las observaciones X)

El valor del coeficiente de regresión del error se calcula a partir de la información del Cuadro 5 (Sumas de Cuadrados y Productos). La ecuación es la siguiente:

$$b_{\varepsilon} = \frac{SPE}{SCE_x} = \frac{12.3900}{416} = 0.0298$$

Las medias de rendimiento (variable Y) y de número de plantas (X) se encuentran a partir de los datos del Cuadro 1. En el Cuadro 7 se presentan las estadísticas y cálculos para encontrar las medias corregidas:

**Cuadro 7. Procedimiento para encontrar las medias corregidas de rendimiento en función de la covariable**

Variedades	Medias $\bar{Y}_i$	Medias $\bar{X}_i$	Desviación $\bar{X}_i - \bar{\bar{X}}$	$b_{\varepsilon}(\bar{X}_i - \bar{\bar{X}})$	$\bar{Y}_{ic} = \bar{Y}_i - b_{\varepsilon}(\bar{X}_i - \bar{\bar{X}})$
A	1.29	23.25	-4.25	-0.1267	1.42
B	1.53	30.00	2.50	0.0745	1.46
C	1.42	26.50	-1.00	-0.0298	1.45
D	1.35	28.25	0.75	0.0224	1.33
E	1.66	27.25	-0.25	-0.0075	1.67
F	1.52	29.75	2.25	0.0671	1.45
Gran Media	$\bar{\bar{Y}} = 1.46$	$\bar{\bar{X}} = 27.5$			$\bar{\bar{Y}}_c = 1.46$

Debe notarse que los procedimientos de comparación de medias se realizan con las medias corregidas.

Con el programa Statistix 9 los resultados son:

**Cuadro 8. Análisis de covarianza, coeficiente de regresión del error y medias corregidas.**

**Analysis of Variance Table for Rend**

Source	DF	SS	MS	F	P
VAR	5	0.25218	0.05044	4.12	0.0124
NPlant	1	0.36902	0.36902	30.14	0.0000
Error	17	0.20811	0.01224		
Total	23				

Grand Mean 1.4621 CV 7.57

Covariate Coefficient	Std Error	T	P
NPlant	0.02978	5.425E-03	5.49 0.0000

Como los tratamientos, las variedades de soya, corresponden a una variable cualitativa puede emplearse alguno de los métodos de separación de medias recomendados para este caso. Como ejemplo, la prueba de Tukey con el programa Statistix presenta el siguiente resultado de la comparación entre las medias corregidas:

Statistix 9.0 Covarianza soya

**Tukey HSD All-Pairwise Comparisons Test of Rend for VAR**

<b>VAR</b>	<b>Mean</b>	<b>Homogeneous Groups</b>
5	1.6699	A
2	1.4580	AB
3	1.4498	AB
6	1.4480	AB
1	1.4166	AB
4	1.3302	B

Alpha 0.05      Standard Error for Comparison      0.0782

Critical Q Value 4,524      Critical Value for Comparison 0.2503

Error term used: Error, 17 DF

There are 2 groups (A and B) in which the means are not significantly different from one another.

**Comentario:** El resultado de la prueba de Tukey indica que la única diferencia clara, estadísticamente significativa al nivel del 5%, es entre las medias de las variedades 4 y 5.





*“No importa cuán bella sea tu teoría, no importa cuán inteligente seas. Si no está de acuerdo con el experimento, tu teoría está errada.” Richard Feynman<sup>54</sup>*

Es importante conocer las presunciones en las que se sustentan los análisis de varianza. Si los datos departen sensiblemente de las bases teóricas, el investigador podría llegar a conclusiones erradas. La solución en varios casos, pero no siempre, es una transformación de datos. Cuando la transformación no corrige las desviaciones del modelo, la alternativa es el empleo de los métodos de análisis de la estadística no paramétrica. Estos métodos se verán en el capítulo 24.

*Las presunciones para la validez de los análisis de varianza son básicamente las siguientes:*

1. Independencia de los términos del error.
2. Distribución normal de las poblaciones.
3. Homogeneidad de las varianzas.
4. Aditividad de los efectos de los tratamientos.

## 23.1. Independencia de los términos del error

Para la validez de los análisis es importante la presunción de la independencia de los términos del error. Esto se logra principalmente a través de la aleatorización. Esto es, la asignación por sorteo de los tratamientos a las unidades experimentales. Mediante la aleatorización se espera que las desviaciones que generan el componente del error tengan signos positivos y negativos que se cancelen mutuamente. En otras palabras, se espera que la sumatoria de los términos del error tienda a cero, idealmente:  $\sum \varepsilon_i = 0$ . Si los errores no son independientes

---

54 Físico estadounidense. [https://www.brainyquote.com/quotes/richard\\_p\\_feynman\\_160383?src=t\\_exp](https://www.brainyquote.com/quotes/richard_p_feynman_160383?src=t_exp)

y más bien existe correlación positiva entre ellos, las pruebas F se distorsionan y pueden generar resultados falsamente significativos.

Sin embargo, si a pesar de la aleatorización inicial, la conducción del experimento es deficiente, puede perderse la independencia del error. Como ejemplo, si se tiene ocho tratamientos y se toma los datos del tratamiento A el día 1, del tratamiento B el día 2 y así sucesivamente hasta el tratamiento H el día 8, los efectos de los tratamientos van a fusionarse con las variaciones del clima (temperatura, humedad, luminosidad) de esos días y con las particularidades del desarrollo evolutivo de plagas y enfermedades. Algo similar ocurre cuando a trabajadores o laboratoristas se les pide tomar datos de tratamientos específicos. Estas fallas en la ejecución de los trabajos generan correlaciones positivas entre las repeticiones de un mismo tratamiento y consecuentemente se producen errores que no son independientes. La solución para esto es, obviamente, trabajar por repeticiones y no por tratamientos. Con la aleatorización y una buena conducción del experimento se evita en lo posible la aparición de resultados falsamente positivos.

## 23.2. Distribución normal de las poblaciones

Si las poblaciones de las que provienen las muestras no siguen la distribución normal pueden producirse sesgos en los errores de los tratamientos. Sin embargo, en la mayoría de los casos, las desviaciones de la normalidad tienen relativamente poco efecto negativo en los resultados. Por lo general, cuando se trata de poblaciones de variables cuantitativas continuas puede suponerse que siguen la distribución normal.

Una característica de las poblaciones que no siguen la distribución normal es que las medias están frecuentemente relacionadas con las varianzas.

### *Distribución binomial*

Cuando en experimentos las observaciones se registran en porcentajes, la distribución es binomial:  $p + q = 100$ . La media y la varianza de poblaciones que tienen una distribución binomial están relacionadas. Sus parámetros son:  $\mu = p$  y  $\sigma^2 = pq/n$ . Puede notarse que la media y la varianza tienen elementos comunes en sus fórmulas.

Con la finalidad de romper la relación entre la media y la varianza y cambiar una distribución binomial a una distribución más cercana a la normal, se emplea la transformación angular *arcoseno*. El efecto que tiene es convertir los

porcentajes a grados. Esto reduce las varianzas sin afectar mayormente a las medias. Existen tablas para encontrar los valores de la transformación. Sin embargo, en Excel y en otros programas puede realizarse mediante esta ecuación:

$$Y'_i = \arcsen(\sqrt{Y_i/100}) * \frac{180}{\pi} \quad (136)$$

Fórmula en la que:

- $Y'_i$  es el valor del dato transformado a grados.
- $\arcsen$  es la función inversa del seno de la raíz cuadrada de un número en el rango de 0 a 1. La respuesta está en radianes, por lo que debe multiplicarse por  $\frac{180}{\pi}$  para la conversión a grados. La constante matemática Pi ( $\pi$ ) es aproximadamente igual a 3.1416.
- $Y_i$  es el dato u observación en porcentaje.

La transformación no es siempre necesaria. Se recomienda realizarla cuando en un experimento los datos fluctúan entre porcentajes bajos (menores de 30%) y porcentajes altos (mayores de 70%). En todo caso, antes de proceder con el análisis de varianza se recomienda efectuar la prueba de la normalidad desarrollada por Shapiro y Wilk (1965). Un ejemplo del uso de esta prueba se presentó en la sección 14.8. Otro ejemplo se presenta a continuación:

**Ejercicio 23.2.1.** En un experimento en bloques completos al azar se evaluó el efecto de ocho fungicidas en el control del hongo *Diplodia maydis* (Steel et al. 1997, página 216). La variable de respuesta fue el número de semillas germinadas de un total de 25 semillas sembradas por parcela. Estos datos corresponden a una distribución binomial (semillas germinadas vs semillas sin germinar). Sin embargo, los autores procedieron como si se tratara de una distribución normal. En este ejemplo, se analizan los resultados en porcentaje y grados. Las observaciones en porcentaje se presentan en el Cuadro 1.

Cuadro 1. Efecto de ocho fungicidas en el porcentaje de semillas germinadas

BLQ	TRATAMIENTOS: FUNGICIDAS								$R_j$
	A	B	C	D	E	F	G	H	
1	32	64	56	40	32	32	28	48	332
2	32	76	64	44	28	32	24	76	376
3	36	96	56	48	4	12	24	36	312
4	28	88	52	32	4	12	24	44	284

5	28	76	56	28	12	12	16	36	264
6	20	76	52	12	8	28	16	20	232
$T_i$	176	476	336	204	88	128	132	260	$\sum Y_{ij} = 1800$

El análisis de varianza con el programa *Statistix* es el siguiente:

**Cuadro 2. Análisis de varianza para la variable porcentaje de germinación de semillas. (STX)**

Source	DF	SS	MS	F	P
BLOQUES	5	1640.0	328.00		
FUNGI	7	19369.3	2767.05	29.94	0.0000
Error	35	3234.7	92.42		
Total	47	24244.0			
Grand Mean 37.500 CV 25.64					

En el siguiente cuadro se presentan las medias y varianzas de los tratamientos:

**Cuadro 3. Medias y varianzas de los porcentajes de germinación de semillas, por tratamientos (fungicidas). (STX)**

Variable	Level	N	$T_i$	Mean	Variance
FUNGI	A	6	176	29.333	29.867
FUNGI	B	6	476	79.333	124.27
FUNGI	C	6	336	56.000	19.200
FUNGI	D	6	204	34.000	171.20
FUNGI	E	6	88	14.667	151.47
FUNGI	F	6	128	21.333	106.67
FUNGI	G	6	132	22.000	24.000
FUNGI	H	6	260	43.333	348.27
Total	48	1800	37.500	515.83	

Sin embargo, como se hizo en el Ejercicio 14.8.1, por tratarse de muestras provenientes de poblaciones binomiales, debió primero realizarse la prueba de normalidad (Prueba W). Esto es especialmente necesario cuando la variación entre los datos es alta. Los datos en el Cuadro 1 van de un 4% a un 96%. Con el programa *Statistix* se obtiene el valor de la prueba y el gráfico correspondiente:

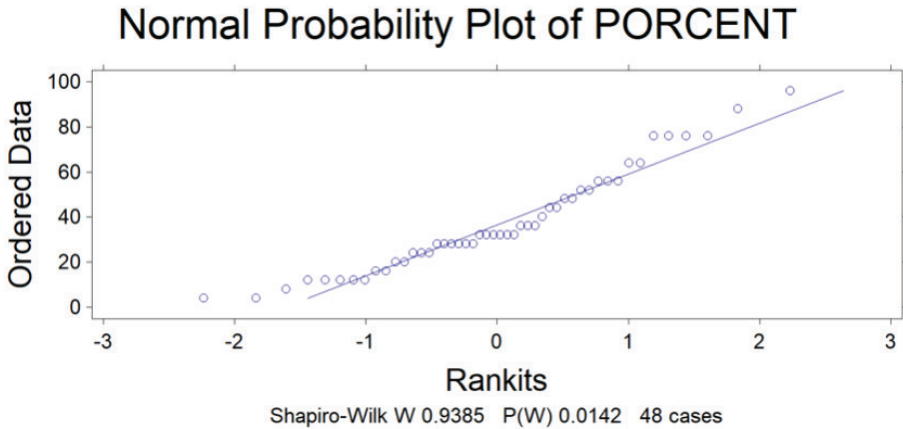


Gráfico 1. Datos en porcentaje de semillas germinadas y su dispersión alrededor de la línea de referencia normal

Puede notarse que muchos datos no están sobre la línea de referencia de la normalidad. La probabilidad de que los datos correspondan a una distribución normal es, según la Prueba W,

$P = 1,42\%$ . Como la probabilidad es baja, se concluye que los datos no siguen una distribución normal sino que mantienen las características propias de una distribución binomial. Esto determina la necesidad de una transformación arcoseno, para luego realizar el análisis de varianza definitivo. Los datos transformados se presentan en el Cuadro 4.

Cuadro 4. Efecto de ocho fungicidas. Datos transformados de porcentajes a grados mediante la función arcoseno

BLQ	TRATAMIENTOS: FUNGICIDAS								$R_j$
	A	B	C	D	E	F	G	H	
1	34.45	53.13	48.45	39.23	34.45	34.45	31.95	43.85	319.96
2	34.45	60.67	53.13	41.55	31.95	34.45	29.33	60.67	346.2
3	36.87	78.46	48.45	43.85	11.54	20.27	29.33	36.87	305.64
4	31.95	69.73	46.15	34.45	11.54	20.27	29.33	41.55	284.97
5	31.95	60.67	48.45	31.95	20.27	20.27	23.58	36.87	274.01
6	26.57	60.67	46.15	20.27	16.43	31.95	23.58	26.57	252.19
$T_i$	196.24	383.33	290.78	211.3	126.18	161.66	167.1	246.38	1782.97

Con la finalidad de constatar el efecto de la transformación arcoseno puede obtenerse la prueba W y observar la dispersión de los datos alrededor de la línea de normalidad. Esto se presenta en el Gráfico 2.

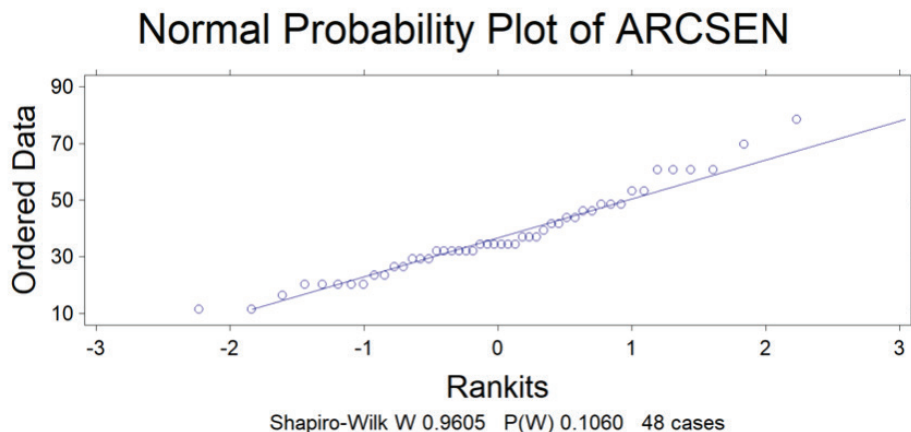


Gráfico 2. Datos transformados mediante la función arcoseno y su dispersión alrededor de la línea de referencia normal

En el Gráfico 2, la prueba W tiene una probabilidad de  $P = 10.60\%$ . Como la probabilidad supera el límite del 5% puede concluirse que los datos transformados se aproximan a la distribución normal. El análisis de varianza con los datos transformados se presenta en el Cuadro 5.

Cuadro 5. Análisis de varianza de la variable ARCSEN en grados (STX)

Source	DF	SS	MS	F	P
BLOQUES	5	713.3	142.66		
FUNGI	7	7993.1	1141.88	24.81	0.0000
Error	35	1610.9	46.02		
Total	47	10317.3			

Grand Mean 37.144 CV 18.26

Si se compara el Cuadro 2 del ADEVA en porcentaje, con el Cuadro 5 del ADEVA en grados, se observa una reducción considerable del cuadrado medio del error, de 92.42 a 46.02. Similar efecto tiene la transformación en el coeficiente de variación. En este ejemplo bajó de 25.64% a 18.26%. Estos cambios determinan, en términos generales, una disminución en las probabilidades de arribar a conclusiones y recomendaciones erradas en los experimentos.

En el Cuadro 6 se presentan los totales de los tratamientos, las medias y las varianzas de los datos transformados a grados.

De igual manera, si se comparan los Cuadros 3 y 6, puede notarse que las medias de los tratamientos cambian poco con la transformación, al contrario de las varianzas que se reducen sensiblemente. Compare en especial las medias y varianzas totales. En el caso del análisis con porcentajes, la media es 37.5 y la varianza total 515.8. Con los datos transformados a grados, la media es 37.1 y la varianza total 219.5. De esta forma, la transformación arcoseno rompe la relación entre medias y varianzas.

**Cuadro 6. Medias y varianzas expresadas en grados para la variable germinación de semillas, por tratamientos. (STX)**

Variable	Level	N		Mean	Variance
FUNGI	A	6	196.23	32.705	12.456
FUNGI	B	6	383.32	63.887	78.692
FUNGI	C	6	290.76	48.460	6.5035
FUNGI	D	6	211.31	35.218	73.123
FUNGI	E	6	126.17	21.028	100.26
FUNGI	F	6	161.65	26.942	54.286
FUNGI	G	6	167.11	27.851	11.979
FUNGI	H	6	246.38	41.063	127.53
Total	48	1782.9	37.144	219.52	

Lo anterior tiene efecto sobre los resultados. En el experimento original se realizaron comparaciones ortogonales entre grupos de medias con la metodología que se describió en la sección 17.1.3. Una de esas comparaciones fue entre el tratamiento A (el testigo sin fungicida) versus los otros siete tratamientos con fungicidas. El valor de F para esta comparación con los datos en porcentajes fue:

$$F_p = \frac{CMQ_1}{CME} = \frac{457.33}{92.42} = 4.95$$

Este valor de F, con 1 y 35 grados de libertad, tiene una probabilidad  $P = 3.3\%$ , un valor que no supera el nivel del 5%. De acuerdo con esto, el efecto del testigo sobre el porcentaje de germinación de las semillas de maíz fue estadísticamente diferente al de los otros siete tratamientos combinados. Sin embargo, cuando la misma comparación ortogonal se realiza con los datos transformados a grados, el resultado es el siguiente:



$$F_g = \frac{CMQ_1}{CME} = \frac{135.13}{46.02} = 2.93$$

La probabilidad de este nuevo valor de F, con 1 y 35 grados de libertad, es  $P = 9.6\%$ . Esta probabilidad supera el nivel del 5% e indica que las diferencias entre el tratamiento testigo y los tratamientos con fungicidas no son estadísticamente significativas.

La transformación también tuvo efectos en las otras comparaciones, pero no causó cambios en la interpretación de la significancia estadística.

### Comentarios:

- Idealmente, el tamaño de la muestra por parcela debe ser igual. En el experimento del ejercicio anterior se sembraron en todas las parcelas 25 semillas. Sin embargo, habrá casos en que el tamaño de la muestra, por condiciones propias o extrañas al experimento, podría fluctuar. Si la variación es mínima, no alterará grandemente los resultados. En todo caso, debe primar el buen criterio del investigador.
- Los métodos de separación de medias se llevan a cabo con los valores de las medias transformadas. Sin embargo, en los informes, como referencia, deben señalarse los valores de las medias en la escala original.

### Distribución de Poisson

La distribución de Poisson es muy similar a la binomial, ya que también se refiere a la ocurrencia de dos eventos excluyentes:  $(p + q) = 1$ . En la distribución binomial, al aumentar el tamaño de la muestra la distribución se aproxima rápidamente a la distribución normal, en especial cuando  $p$  y  $q$  tienen valores cercanos a 0.5. Sin embargo, cuando la diferencia entre las proporciones  $p$  y  $q$  es bastante amplia, como ejemplo  $p = 0.05$  y  $q = 0.95$ , la distribución permanece sesgada, incluso en el caso de muestras grandes. Esto da origen a la distribución de Poisson, en honor al científico que la estudió. Como  $n$  tiende al infinito y  $p$  hacia cero en forma simultánea,

$np$  se convierte en una constante igual a la media. No sólo que  $\mu = np$ , sino que también

$$\sigma^2 = \mu = np.$$

La distribución de Poisson tiene aplicaciones para el caso de eventos raros. El más conocido ejemplo de su uso se dio en la Segunda Guerra Mundial, cuando el estadístico inglés R. D. Clarke pudo determinar que las bombas alemanas V1 y V2 no iban dirigidas a blancos específicos (Salsburg, 2001). En las ciencias biológicas se utiliza en experimentos en los que la ocurrencia de un evento tiene una baja probabilidad. Como ejemplo, el número de una especie de insectos atrapados al pasar  $n$  veces una red sobre un cultivo, el número de plantas albinas mutantes vs plantas normales en  $n$  parcelas de soya, el número de frutos con larvas vs frutos sin daño en  $n$  cajas para exportación, etc.

En experimentos, cuando se supone que los datos se aproximan a una distribución de Poisson, en la que como se indicó la varianza y la media están relacionadas de la siguiente manera:

$\sigma^2 = \mu$ , la transformación raíz cuadrada es frecuentemente efectiva para romper o disminuir la relación entre las varianzas y las medias. La transformación también aproxima la distribución de los datos a la normalidad. Si se presentan ceros, lo adecuado es sumar 1 a todos los datos antes de obtener la raíz cuadrada. A continuación, se presenta un ejemplo:

**Ejercicio 23.2.2.** En un experimento se evaluó el número de mazorcas podridas en cinco híbridos dobles experimentales de maíz (H1 a H5). Se incluyó un híbrido comercial tratado con insecticida, como testigo (TC). El diseño fue un DCA con cinco repeticiones. Los resultados, número de mazorcas podridas en un total de 90 mazorcas por parcela, se presentan en el Cuadro 1:

Cuadro 1. Número de mazorcas podridas

TRATAMIENTOS					
TC	H1	H2	H3	H4	H5
2	5	12	3	12	5
0	10	4	1	10	0
1	6	8	0	7	2
4	0	3	5	2	4
0	3	9	3	8	6

Cuadro 2. Análisis de varianza de la variable “número de mazorcas podridas” (STX)

Source	DF	SS	MS	F	P
TRAT	5	167.500	33.5000	3.72	0.0123
Error	24	216.000	9.0000		
Total	29	383.500			

Grand Mean 4.5000 CV 66.67

En el Cuadro 3 se presentan los totales de los tratamientos, las medias y varianzas de la variable “mazorcas podridas”.

Cuadro 3. Medias y varianzas de la variable “mazorcas podridas” (STX)

Variable	Level	N	$T_i$	Mean	Variance
TRAT	TC	5	7	1.4000	2.8000
TRAT	H1	5	24	4.8000	13.700
TRAT	H2	5	36	7.2000	13.700
TRAT	H3	5	12	2.4000	3.8000
TRAT	H4	5	39	7.8000	14.200
TRAT	H5	5	17	3.4000	5.8000
Overall	30	135	4.5000	13.224	

Sin embargo, es importante probar si los datos se aproximan a la distribución normal. La prueba W y el gráfico correspondiente son:

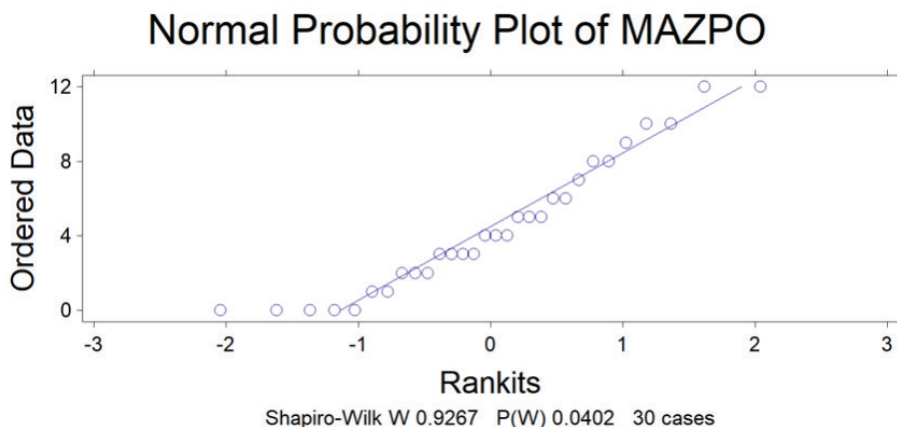


Gráfico 1. Dispersión de los datos de número de mazorcas podridas alrededor de la línea de referencia normal

Puede verse que la prueba  $W$  es significativa, con una probabilidad menor que 5%. Esto indica la necesidad de transformar los datos para buscar aproximarse a la distribución normal. A continuación, se presentan los datos más 1, transformados mediante raíz cuadrada. Se suma 1 a los datos para eliminar los ceros:  $Y'_i = \sqrt{Y_i + 1}$

**Cuadro 4. Mazorcas podridas. Datos transformados mediante raíz cuadrada**

TRATAMIENTOS					
TC	H1	H2	H3	H4	H5
1.73	2.45	3.61	2.00	3.61	2.45
1.00	3.32	2.24	1.41	3.32	1.00
1.41	2.65	3.00	1.00	2.83	1.73
2.24	1.00	2.00	2.45	1.73	2.24
1.00	2.00	3.16	2.00	3.00	2.65

**Cuadro 5. Medias y varianzas de los datos transformados**

Variable	Level	N	$T'_i$	Mean	Variance
TRAT	TC	5	7.3800	1.4760	0.2764
TRAT	H1	5	11.420	2.2840	0.7410
TRAT	H2	5	14.010	2.8020	0.4448
TRAT	H3	5	8.8600	1.7720	0.3227
TRAT	H4	5	14.490	2.8980	0.5161
TRAT	H5	5	10.070	2.0140	0.4386
Overall	30	66.230	2.2077	0.6534	

Si se comparan los Cuadros 3 y 5, se observa cierta relación de las medias con las varianzas en los datos originales. En los datos transformados se nota que las medias y las varianzas son independientes.

En todo caso, es importante obtener la prueba  $W$  y generar el gráfico correspondiente con *Statistix*, para comprobar si la transformación logró que los datos se aproximen a la distribución normal.

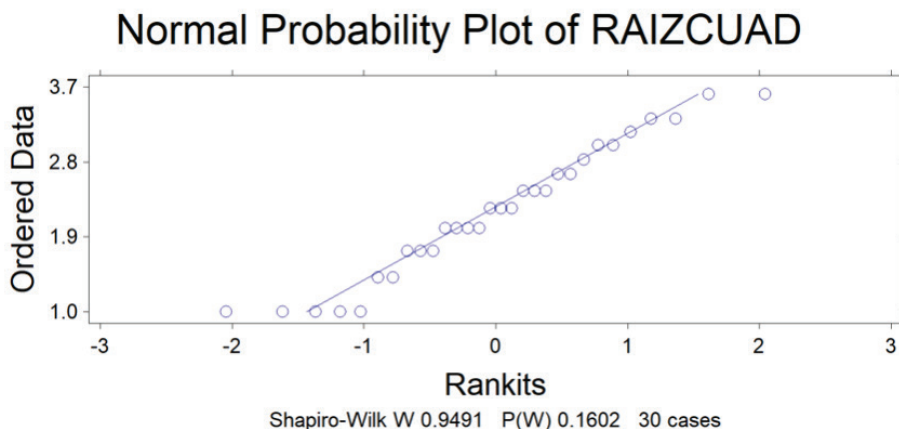


Gráfico 2. Distribución de los datos de raíz cuadrada alrededor de la línea de referencia de la normalidad

La prueba W, en el Gráfico 2, con una probabilidad igual a 16 %, sugiere que la transformación realizada aproximó los datos a una distribución normal. Con esta información se puede realizar en forma válida el análisis de varianza, como se observa en el Cuadro 6.

**Cuadro 6. Análisis de varianza de los datos transformados mediante "raíz cuadrada del número de mazorcas podridas" (STX)**

Source	DF	SS	MS	F	P
TRAT	5	7.9913	1.59827	3.50	0.0162
Error	24	10.9586	0.45661		
Total	29	18.9499			

Grand Mean 2.2077 CV 30.61

Si se comparan los Cuadros 2 y 6, los valores de F y de las probabilidades son similares en este ejercicio. Sin embargo, puede notarse que el coeficiente de variación se redujo a la mitad con los datos transformados. Esto contribuye a una mayor precisión en las conclusiones a las que se llegue. En todo caso, la finalidad de la transformación fue convertir datos que seguían la distribución de Poisson a una distribución normal, lo que validó el análisis de varianza.

### 23.3. Homogeneidad de las varianzas

A través del texto se ha mencionado varias veces la importancia de la homogeneidad de varianzas. Si las varianzas de los tratamientos no son estadísticamente homogéneas puede arribarse a conclusiones erradas. Cuando sólo se trata de dos varianzas se realiza una prueba F. Esto se vio en el Ejercicio 9.1.1 y también en el Ejercicio 16.4.1. En este último caso, las varianzas del error fueron homogéneas y pudieron combinarse los resultados de dos cuadrados latinos en uno solo.

En el caso de experimentos con tres o más tratamientos existen varias pruebas para determinar la homogeneidad de las varianzas. entre ellas la desarrollada por Bartlett (1937) y la de Brown y Forsythe (1974).

Para la prueba de Bartlett se aplica la fórmula (127) presentada en el Capítulo 21. Esta es:

$$PB\chi^2 = \frac{f(k^* \ln \bar{S}^2 - \sum \ln S_i^2)}{C}$$

La prueba de Brown y Forsythe es una prueba F que se basa en las desviaciones de los datos de cada tratamiento con relación a su mediana. (Se recordará que la mediana es el dato que queda en el centro una vez que se ordenan los datos.) Las desviaciones se elevan al cuadrado y se procede, con los nuevos datos, a realizar el análisis de varianza. La hipótesis nula que se evalúa, como en todas las pruebas, es que las varianzas de las poblaciones son iguales.

Sobre la base de los resultados de la prueba que se use, si se decide que las varianzas no son homogéneas y que por lo tanto se viola una de las presunciones básicas del análisis de varianza, puede efectuarse una transformación logarítmica de los datos. Si la transformación no surte efecto, una opción es realizar el análisis con métodos de la estadística no paramétrica (Capítulo 24).

Sin embargo, existe una prueba paramétrica más potente que las no paramétricas y que puede usarse sin importar si las varianzas son homogéneas o heterogéneas. Esta opción es el análisis de varianza en una dirección desarrollado por Welch (1951). Sólo tiene aplicación para los tratamientos y no considera bloques u otros factores. La fórmula para encontrar el valor de F en el análisis de varianza de Welch es la siguiente (Fórmula 137):

$$F = \frac{\frac{1}{k-1} * \sum w_i (\bar{Y}_i - \bar{Y}')^2}{1 + \frac{2(k-2)}{k^2-1} * \sum \left( \frac{1}{n_i-1} \right) * \left( 1 - \frac{w_i}{w} \right)^2}$$

Fórmula en la que:

- $k$ . Es el número de tratamientos
- $w_i = \frac{n_i}{S_i^2}$ . Este valor es el “peso” de cada tratamiento.
- $w = \sum w_i$ . Es la sumatoria de los pesos de todos los tratamientos.
- $\bar{Y}_i$ . Es la media del tratamiento “i”.
- $\bar{Y}' = \frac{\sum w_i \bar{Y}_i}{w}$ . Esta cantidad equivale al cuadrado medio.

El número de grados de libertad para el numerador de la prueba F es igual a  $k-1$ . Para el denominador, el número de grados de libertad se encuentra con la fórmula (138). En muchos casos esta fórmula generará un número con decimales, por lo que debe redondearse al entero más cercano (Fórmula 138):

$$g.l. denom. = \frac{k^2-1}{3 * \sum \left( \frac{1}{n_i-1} \right) * \left( 1 - \frac{w_i}{w} \right)^2}$$

Veamos un ejemplo de estos procedimientos:

**Ejercicio 23.3.1.** En un experimento se evaluaron los rendimientos de tres híbridos comerciales de maíz y dos variedades de polinización libre. El diseño experimental fue completamente al azar (DCA) con siete repeticiones. El rendimiento, en quintales métricos por hectárea, se presenta en el Cuadro 1. Realice los análisis que sean pertinentes.

Cuadro 1. Rendimiento de maíz, en q/ha, de cinco tratamientos, sus medias y varianzas

TRAT	REPETICIONES							$\bar{Y}_i$	$S_i^2$
H1	62	58	52	47	54	55	44	53.14	38.14
H2	54	45	46	40	56	55	50	49.43	35.95
H3	63	62	54	62	58	57	60	59.42	10.61
PL1	35	44	20	53	24	42	17	33.57	184.95
PL2	18	25	38	45	16	50	40	33.14	180.81

Al inspeccionar las varianzas de los tratamientos en el Cuadro 1 se nota amplias diferencias entre algunas de ellas. Las varianzas aparentemente no son

iguales; pero en todo caso esto debe comprobarse mediante el empleo de una prueba estadística.

El programa *Statistix*, junto al análisis de varianza, presenta tres pruebas de homogeneidad de varianzas, como puede verse en el siguiente cuadro:

**Cuadro 2. Análisis de varianza del rendimiento de maíz de tres híbridos y dos variedades de polinización abierta.STX)**

Source	DF	SS	MS	F	P
TRAT	4	3937.83	984.457	10.93	0.0000
Error	30	2702.86	90.095		
Total	34	6640.69			

Grand Mean 45.743 CV 20.75

Homogeneity of Variances	F	P
Levene's Test	5.83	0.0014
O'Brien's Test	4.82	0.0040
Brown and Forsythe Test	3.87	0.0119

Observe en el Cuadro 2 que las tres pruebas de homogeneidad de varianzas tienen probabilidades menores que el nivel de 5% ( $P < 0.05$ ). Esto indica que debe rechazarse la hipótesis nula de homogeneidad de varianzas de los tratamientos a los niveles alfa indicados. En todo caso, parece que la prueba de Levene, que se basa en diferencias entre las observaciones y sus medias, es excesivamente conservadora.

A continuación, se presenta la prueba de Bartlett, que es la prueba más indicada de homogeneidad de las varianzas, por diferentes autores (Steel et al. 1997; Snedecor y Cochran, 1967; Morris, 1999). Detalles de la fórmula se encuentran en el Capítulo 21.

**Cuadro 3. Varianzas de los tratamientos, logaritmos naturales y cálculos previos para la prueba de Bartlett**

	TRATAMIENTOS					Sumatoria	$\bar{S}^2$	$\ln \bar{S}^2$
	H1	H2	H3	PL1	LP2			
$S_i^2$	38.14	35.95	10.61	184.95	180.81	450.46	90.09	4.50
$\ln S_i^2$	3.6413	3.5821	2.3618	5.2201	5.1974	20.00		



Al aplicar la fórmula (127) se tiene lo siguiente:

$$PB\chi^2 = \frac{6(5 \cdot 4.50 - 20.00)}{1.07} = 14.06$$

En la fórmula, el valor de  $f = 6$  es el número de grados de libertad de cada varianza ( $7-1$ ) y  $k = 5$ , el número de tratamientos. El denominador en la fórmula es un factor de ajuste

$$C = 1 + \frac{k + 1}{3kf}$$

En la Tabla de Ji-cuadrado, el valor crítico al nivel alfa del 5%, con  $k-1$  grados de libertad es igual a 9.49. Como la prueba estadística supera al valor crítico se concluye que se rechaza la hipótesis nula de homogeneidad de las varianzas. Con el programa *Statistix* se puede encontrar la probabilidad exacta. Esta es  $P = 0.0071$ . Es decir que existe amplia evidencia de que las varianzas de las poblaciones de las que se extrajeron las muestras no eran iguales.

Mediante una transformación logarítmica de los datos originales puede tratarse de homogeneizar las varianzas. El análisis de varianza con los datos transformados se presenta en el Cuadro 4 (*Statistix* 9).

**Cuadro 4. Análisis de varianza de los datos de rendimiento de maíz transformados a logaritmos de base 10\*. (STX)**

Source	DF	SS	MS	F	P
TRAT	4	0.51104	0.12776	7.87	0.0002
Error	30	0.48714	0.01624		
Total	34	0.99818			

Grand Mean 1.6329    CV 7.80

Homogeneity of Variances	F	P
Levene's Test	7.63	0.0002
O'Brien's Test	6.30	0.0008
Brown and Forsythe Test	4.63	0.0050

\*Nota: el mismo resultado se obtiene con los logaritmos naturales o neperianos.

Con la información del Cuadro 4, al observar las tres pruebas F para la homogeneidad de las varianzas, se puede afirmar que la transformación logarítmica

no solucionó el problema de la heterogeneidad de las varianzas. Las bajas probabilidades de las tres pruebas hacen innecesario efectuar la prueba de Bartlett con los datos transformados.

Como no se puede honestamente efectuar el análisis de varianza convencional, porque además se corre el riesgo de arribar a conclusiones erradas, la solución y mejor alternativa es llevar a cabo el análisis de varianza de Welch. Para llegar a la prueba F de la fórmula (137) es conveniente, primero, encontrar los componentes intermedios. La información se genera a partir de los datos del Cuadro 1 y se presenta en el Cuadro 5 a continuación:

**Cuadro 5. Algunos cálculos previos para encontrar el valor de F en el análisis de varianza de Welch**

TRAT	$\bar{Y}_i$	$S_i^2$	$w_i = n_i/S_i^2$	$w_i(\bar{Y}_i - \bar{Y}')^2$	$(1-w_i/w)^2$
H1	53.14	38.14	0.18	0.54	0.70
H2	49.43	35.95	0.19	5.73	0.68
H3	59.42	10.61	0.66	13.80	0.17
PL1	33.57	184.95	0.04	17.14	0.93
PL2	33.14	180.81	0.04	18.25	0.93
Sumatoria			$w = 1.11$	55.46	3.41

En este ejercicio,  $k = 5$  y  $n_i = n = 7$

Las cantidades  $w_i = \frac{n_i}{S_i^2}$  son los pesos de cada tratamiento. La sumatoria de los pesos individuales se representa como el valor  $w$ :

$$w = \sum \frac{n_i}{S_i^2} = \frac{7}{38.14} + \frac{7}{35.95} + \dots + \frac{7}{180.81} = 0.18 + 0.19 + \dots + 0.04 = 1.11$$

El valor equivalente al cuadrado medio  $\bar{Y}'$  se encuentra con esta ecuación:

$$\bar{Y}' = \frac{\sum w_i \bar{Y}_i}{w} = \frac{(0.18 \cdot 53.14) + (0.19 \cdot 49.43) + \dots + (0.04 \cdot 33.14)}{1.11} = 54.85$$

Se dispone ya de todas las cantidades necesarias para encontrar el valor  $F$  del análisis de varianza de Welch, de acuerdo con la fórmula (137):

$$F = \frac{\frac{1}{4} \cdot 55.46}{1 + \frac{2(5-2)}{24} \cdot \frac{1}{6} \cdot 3.41} = \frac{13.86}{1.14} = 12.13$$

El número de grados de libertad del numerador es igual a  $k-1 = 4$ . Para encontrar los grados de libertad del denominador se emplea la fórmula (138):

$$g.l. denom. = \frac{24}{3 * \frac{1}{6} * 3.41} = 14.1$$

El valor crítico de F en la tabla, al nivel del 5% con 4 y 14 grados de libertad, es igual a 3.11. Como el valor de la prueba estadística F supera al valor crítico, se acepta la hipótesis alternativa de que por lo menos dos medias de los tratamientos no son estadísticamente iguales.

El análisis de varianza de Welch con el programa *Statistix* (Cuadro 6) presenta resultados iguales, salvo por errores de redondeo y por la probabilidad exacta de haber obtenido solo por azar un valor F tan alto ( $P = 0.0002$ ).

**Cuadro 6. Análisis de varianza de Welch con los datos de rendimiento de maíz de tres híbridos y dos variedades de polinización libre. (STX)**

Welch's Test for Mean Differences			
Source	DF	F	P
TRAT	4.0	12.14	0.0002
Error	14.1		

Para identificar las medias que no son estadísticamente iguales puede usarse la prueba de Games-Howell. Pero también, dependiendo del diseño de los tratamientos, puede llevarse a efecto un juego de comparaciones ortogonales.

**Comentario:** El análisis de varianza de Welch es la mejor alternativa cuando las varianzas son heterogéneas en los diseños completamente al azar y con un solo factor. Aparentemente, no se han desarrollado procedimientos para otros diseños.

### **La prueba de Games-Howell**

La prueba de Games-Howell (GH) es una modificación de la prueba de Tukey, que se usa cuando en el análisis de varianza de Welch el valor de F tiene una probabilidad menor que el 5%. Con esta prueba pueden identificarse las medias de tratamientos que difieren estadísticamente de otras medias. El procedimiento es similar al de la prueba de Tukey. Sin embargo, el error estándar y los grados de libertad se encuentran con diferentes fórmulas.

La Diferencia entre medias tiene significación estadística cuando se cumple la siguiente desigualdad:

$$|\bar{Y}_i - \bar{Y}_j| > q_\alpha * S'_{\bar{Y}} \quad (139)$$

Expresión en la que  $|\bar{Y}_i - \bar{Y}_j|$  es el valor absoluto de la diferencia de dos medias y  $q_\alpha * S'_{\bar{Y}}$  es el valor de la prueba GH. El error estándar común para las dos varianzas ( $S'_{\bar{Y}}$ ) se calcula con la siguiente fórmula:

$$S'_{\bar{Y}} = \sqrt{\frac{1}{2} * \left( \frac{S_A^2}{n_A} + \frac{S_B^2}{n_B} \right)} \quad (140)$$

Los grados de libertad para la prueba GH tiene la misma fórmula (41) presentada en la sección 9.3 para la prueba T; en este caso:

$$g.l.T = \frac{\left( \frac{S_A^2}{n_A} + \frac{S_B^2}{n_B} \right)^2}{\frac{\left( \frac{S_A^2}{n_A} \right)^2}{n_A - 1} + \frac{\left( \frac{S_B^2}{n_B} \right)^2}{n_B - 1}}$$

Debe tomarse en cuenta que el empleo de esta fórmula compleja se debe a que las varianzas no son iguales.

El valor crítico para la prueba GH se encuentra en la Tabla de Tukey (Apéndice E1 y E2), al nivel de significación alfa 0.05 o 0.01, con los grados de libertad obtenidos a través de la fórmula anterior. Las columnas de la Tabla de Tukey se refieren a k, que es el número de tratamientos (p en la tabla) del experimento. Para las comparaciones, las medias deben haber sido previamente ordenadas en forma ascendente o descendente. Veamos un ejemplo:

---

**Ejercicio 23.3.2.** Con los datos del Ejercicio 23.3.1, mediante la prueba de Games-Howell, compare las medias de los tratamientos previamente ordenadas. Utilice el nivel de significación alfa del 5%. En el Cuadro 1 se presentan las estadísticas necesarias.

Cuadro 1. Medias ordenadas, varianzas y número n de observaciones. Datos de rendimiento de maíz de tres híbridos y dos variedades de polinización abierta

Estadísticas	Tratamientos				
	PL2	PL1	H2	H1	H3
Medias $\bar{Y}_i$	33.14	33.57	49.43	53.14	59.42
Varianzas $S_i^2$	180.81	184.95	35.95	38.14	10.61
Valores $n_i$	7	7	7	7	7

La primera comparación se realiza, generalmente, entre la media más baja y la segunda más alta. Esto por cuanto el valor altamente significativo de F en el análisis de varianza de Welch, asegura que por lo menos las dos medias extremas, la menor y la mayor, son estadísticamente diferentes.

La diferencia absoluta entre las dos medias de PL2 y H1 es:

$$|33.14 - 53.14| = 20.00$$

Luego se encuentra el error estándar común de las dos medias:

$$S'_{\bar{Y}} = \sqrt{\frac{1}{2} * \left( \frac{180.81}{7} + \frac{38.14}{7} \right)} = 3.95$$

El número de grados de libertad para la prueba son:

$$g.l.GH = \frac{\left( \frac{180.81}{7} + \frac{38.14}{7} \right)^2}{\frac{\left( \frac{180.81}{7} \right)^2}{7-1} + \frac{\left( \frac{38.14}{7} \right)^2}{7-1}} = 8.42$$

(Este número se redondea al entero más cercano, en este caso se redondea a ocho.)

En la Tabla de Tukey al 5%, con 8 grados de libertad y en la columna del número de medias de tratamientos (cinco) se encuentra el valor  $q_\alpha = 4.89$ . El valor crítico, consecuentemente es:  $q_\alpha * S'_{\bar{Y}} = 19.31$

Como la diferencia absoluta entre las medias PL2 y H1 (igual a 20.00), excede el valor crítico (igual a 18.72), se concluye que las dos medias son estadísticamente diferentes al nivel del 5%.

De igual manera, se puede proseguir con comparaciones entre otras medias. En todos los casos habrá que calcular las diferencias entre dos medias, los errores estándar comunes y los grados de libertad para la prueba.

### Comparaciones ortogonales

En la Sección 17.1.3 se presentó el método de comparaciones ortogonales. Se usa para identificar medias o grupos de medias estadísticamente diferentes, después de que se haya efectuado el análisis de varianza. Este es el mejor método de comparación de medias cuando los tratamientos son cualitativos. Sin embargo, se requiere que las comparaciones o contrastes se hayan planificado de antemano, para evitar comparaciones guiadas por los resultados. El método de comparaciones o contrastes ortogonales no debe confundirse con el método de *polinomios ortogonales*, procedimiento que se realiza cuando los tratamientos son *cuantitativos*.

Cada comparación ortogonal tiene un solo grado de libertad. Esto hace que pueda evaluarse el resultado mediante una prueba F o una prueba t, gracias a la relación  $F = t^2$ .

En experimentos con diseños completamente al azar, que presenten varianzas heterogéneas y diferente número de repeticiones, la mejor forma para realizar comparaciones ortogonales es mediante la prueba T para observaciones no pareadas. Esto, obviamente, si el arreglo o diseño de los tratamientos lo permite. En la Sección 9.3 se presentó la fórmula (40) para la prueba T:

$$T = \frac{\bar{Y}_A - \bar{Y}_B}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}}$$

Esta prueba T es la indicada cuando las varianzas no son iguales. Los grados de libertad se encuentran con la fórmula (41), también empleada en el ejercicio anterior. Veamos un ejemplo:

---

**Ejercicio 23.3.3.** Con los datos del rendimiento de maíz de tres híbridos y dos variedades de polinización libre (Ejercicio 23.3.1) realice las siguientes comparaciones ortogonales: 1) los tres híbridos vs las dos variedades; 2) los híbridos H1 y H2 vs el híbrido H3; 3) el híbrido H1 vs el híbrido H2; y 4) la variedad PL1 vs la variedad PL2. Como antecedentes, se conoce que los híbridos H1 y H2 fueron desarrollados por una entidad local, mientras que el híbrido H3 proviene de una

empresa comercial. Las dos variedades de polinización libre tienen también diferente procedencia.

Para realizar las comparaciones, de acuerdo con la fórmula (40), deben encontrarse las diferencias entre las medias de los grupos y las varianzas de los grupos, tomando en cuenta el número de observaciones en cada caso. Con los datos del Cuadro 1 del Ejercicio 23.3.1 se puede formar el siguiente Cuadro:

Cuadro 1. Estadísticas de los grupos para las comparaciones ortogonales

Estadísticas	Grupos para las comparaciones							
	H1+H2+H3	PL1+PL2	H1+H2	H3	H1	H2	PL1	PL2
$\bar{Y}$	54.00	33.36	51.29	59.42	53.14	49.43	33.57	33.14
$S^2$	43.30	168.86	37.91	10.61	38.14	35.95	184.95	180.81
$n$	21	14	14	7	7	7	7	7

La prueba estadística T para la primera comparación ortogonal y los grados de libertad son:

$$T = \frac{54.00 - 33.36}{\sqrt{\frac{43.30}{21} + \frac{168.86}{14}}} = \frac{20.64}{3.76} = 5.49$$

$$g.l.T = \frac{\left(\frac{43.30}{21} + \frac{168.86}{14}\right)^2}{\frac{\left(\frac{43.30}{21}\right)^2}{21-1} + \frac{\left(\frac{168.86}{14}\right)^2}{14-1}} = \frac{199.47}{11.40} \approx 18$$

Ahora debe buscarse en la Tabla t el valor crítico al nivel alfa al que se haya planteado la prueba, con 18 grados de libertad. Si se toma alfa al 5%, el valor crítico es igual a 2.10. Como el valor de la prueba estadística T (5.49) supera el valor crítico (2.10), la decisión es aceptar la hipótesis alterna de que la diferencia entre las medias de los dos grupos es estadísticamente significativa. Es decir que los híbridos en conjunto rindieron significativamente más que las variedades.

A continuación, se presentan los resultados obtenidos con el programa *Statistix* y resumidos en el siguiente cuadro:

Cuadro 2. Resultados de las cuatro comparaciones ortogonales realizadas

Comparación	Diferencia De medias	Homogeneidad De varianzas	Valor de T	Grados de Libertad	Probabilidad
H1+H2+H3 vs PL1+PL2	20.64	no	5.49	17.5	0.0000
H1+H2 vs H3	-8.14	sí	-3.25	19	0.0042
H1 vs H2	3.71	Sí	1.14	12	0.2759
PL1+PL2	0.43	Sí	0.06	12	0.9537

En el cuadro anterior se verifica que el rendimiento de los híbridos fue claramente superior al rendimiento de las variedades de polinización libre. Al comparar los híbridos entre ellos, el H3 tuvo un rendimiento significativamente mayor que el de los dos otros híbridos. No hubo diferencia significativa entre los híbridos H1 y H2, ambos desarrollados por una entidad local. Tampoco hubo diferencia significativa entre las dos variedades de polinización libre.

Puede notarse que tres de las comparaciones tuvieron varianzas homogéneas. Aunque el resultado sería el mismo, para estos casos pudo emplearse la fórmula (37) que es más simple: las varianzas homogéneas se promedian y los grados de libertad se encuentran con la fórmula  $gl = n_i + n_j - 2$ , como se señala en la Sección 9.3.

### 23.4. Aditividad de los efectos de los tratamientos

La aditividad de los efectos experimentales es otra de las presunciones necesarias para la validez de los análisis estadísticos. El modelo lineal aditivo, como ejemplo en el caso de los diseños de bloques al azar, es el siguiente:  $Y_{ij} = \mu + \tau_i + r_j + \varepsilon_{ij}$ .

En el modelo lineal está implícita la aditividad de los efectos de los tratamientos ( $\tau_i$ ) con el efecto de las repeticiones o bloques ( $r_j$ ). Pero, además, la suposición básica del modelo es que el error experimental ( $\varepsilon_{ij}$ ) resulta de la interacción *al azar* entre tratamientos y bloques. Sin embargo, si la interacción no es al azar, el error experimental resulta inflado por causa de efectos multiplicativos entre tratamientos y bloques.

En un modelo lineal multiplicativo,  $Y_{ij} = \mu + \tau_i + r_j + (\tau_i r_j + \varepsilon_{ij})$ , el efecto multiplicativo  $\tau_i r_j$  está íntimamente fusionado con el error experimental, lo que invalida el análisis de varianza. Cuando los efectos son multiplicativos, las varianzas o las desviaciones estándar son proporcionales a las medias.



Sin embargo, Tukey (1949) desarrolló un procedimiento para estimar cuánto del error experimental se debe a la No Aditividad. El procedimiento consiste en partir la suma de cuadrados del error experimental de los diseños de bloques completos al azar. Esta prueba tiene un solo grado de libertad.

Si mediante la prueba de Tukey se demuestra que los datos no son aditivos debe tratarse de corregir este problema. De lo contrario, los resultados de los análisis de varianza pueden llevar a conclusiones erradas. La corrección para este caso, que no siempre produce los resultados esperados, consiste en una transformación logarítmica de los datos. Esta transformación tiene dos posibles efectos: elimina la No Aditividad y homogeniza las varianzas.

Para realizar la prueba de Tukey se sigue los siguientes pasos:

- Primer paso. Encontrar las desviaciones entre las medias de los tratamientos y la gran media experimental:  $d_i = \bar{Y}_i - \bar{\bar{Y}}$ . Encontrar también las desviaciones entre las medias de los bloques y la gran media experimental:  $d_j = \bar{Y}_j - \bar{\bar{Y}}$
- Segundo paso. Calcular los valores  $w_i$ , que resultan de la sumatoria de los productos de las observaciones  $Y_{ij}$  por las desviaciones  $d_i$ . Esto es:  $w_i = \sum Y_{ij} * d_i$
- Tercer paso. Hallar el valor  $Q$ , que es igual a  $\sum w_i * d_j$
- Cuarto paso. Encontrar las cantidades  $\sum d_i^2$  y  $\sum d_j^2$
- Quinto paso. Calcular la cantidad  $\frac{Q^2}{\sum d_i^2 * \sum d_j^2}$ . Este valor es la Suma de Cuadrados de la No Aditividad.
- Sexto paso. Determinar la significación o no del componente de la No Aditividad.
- Séptimo paso. Si la No Aditividad resultara significativa debe procederse a una transformación de datos para tratar de corregir la desviación del modelo.

Veamos un ejemplo:

---

**Ejercicio 23.4.1.** En un experimento conducido con un diseño de bloques al azar se evaluó el efecto de cuatro prácticas de control de malezas (A, B, C, D) en cinco

repeticiones o bloques. Los resultados, en peso seco de malezas por parcela, se presentan en el Cuadro 1.

Cuadro 1. Peso seco de malezas en decagramos por parcela

BLOQ	TRATAMIENTOS				$R_j$
	A	B	C	D	
1	1.54	22.36	5.05	32.57	61.52
2	3.68	14.05	1.58	53.83	73.14
3	1.02	8.73	1.26	23.34	34.35
4	0.58	35.49	3.29	42.59	81.95
5	6.31	24.08	9.05	19.86	59.30
$T_i$	13.13	104.71	20.23	172.19	$\sum Y_{ij} = 310.26$
$\bar{Y}_i$	2.63	20.94	4.05	34.44	$\bar{Y} = 15.51$

En el Cuadro 2 se presenta el análisis de varianza:

Cuadro 2. Análisis de varianza de la variable peso de maleza

F.V.	S.C.	G.L.	C.M.	F	$F_{0.05}$	$F_{0.01}$	P
BLOQ	323.5	4	80.9				
TRAT	3426.0	3	1142.0	14.6	3.49	6.93	0.0003
Error	941.4	12	78.5				
Total	4690.9	19					

Sin embargo, antes de llegar a conclusiones, es conveniente inspeccionar las medias y desviaciones estándar de los tratamientos. Si se observa alguna relación entre ellas puede deberse a la presencia de efectos multiplicativos. En el Cuadro 3 se presentan las medias, varianzas y desviaciones estándar de los tratamientos.

Cuadro 3. Medias ( $\bar{Y}_i$ ), desviaciones estándar ( $S_i$ ) y varianzas ( $S_i^2$ ) de la variable peso de malezas de los tratamientos A, B, C y D

Tratamientos	$\bar{Y}_i$	$S_i$	$S_i^2$
A	2.63	2.38	5.66
B	20.94	10.25	105.03
C	4.05	3.18	10.12
D	34.44	13.98	195.42

Puede notarse que las medias de los tratamientos parecen estar correlacionadas con las desviaciones estándar respectivas. A medias pequeñas corresponden desviaciones estándar pequeñas y a medias altas, desviaciones estándar altas. Esto podría ameritar una transformación de datos para romper esa dependencia.

La prueba de No Aditividad, desarrollada por Tukey, permite determinar estadísticamente si se requiere transformar los datos. En el Cuadro 4 se presentan los cálculos preliminares para estimar la No Aditividad.

**Cuadro 4. Cálculo de las estadísticas que se requieren para la prueba de No Aditividad de Tukey. Datos de peso de malezas del Cuadro 1**

BLOQ	TRATAMIENTOS				ESTADÍSTICAS			
	A	B	C	D	$\bar{Y}_j$	$d_j$	$w_i = \sum Y_{ij} * d_i$	$w_i * d_j$
1	1.54	22.36	5.05	32.57	15.38	-0.13	660.03	-87.78
2	3.68	14.05	1.58	53.83	18.29	2.77	1029.47	2853.69
3	1.02	8.73	1.26	23.34	8.59	-6.93	461.51	-3196.20
4	0.58	35.49	3.29	42.59	20.49	4.97	953.49	4743.14
5	6.31	24.08	9.05	19.86	14.83	-0.69	321.49	-221.18
$\bar{Y}_i$	2.63	20.94	4.05	34.44				$Q = 4091.66$
$d_i$	-12.89	5.43	-11.47	18.93				

Para encontrar el divisor se requieren estos dos valores:

$$\sum d_i^2 = (-12.89^2) + \dots + (18.93^2) = 685.20 \quad \text{y} \quad \sum d_j^2 = (-0.13^2) + \dots + (-0.69^2) = 80.88$$

La contribución a la Suma de Cuadrados del Error Experimental por razón de la No Aditividad es:

$$\frac{Q^2}{\sum d_i^2 * \sum d_j^2} = \frac{4091.66^2}{685.20 * 80.88} = 302.1$$

Esta prueba tiene un solo grado de libertad. La prueba se resume en el Cuadro 5.

**Cuadro 5. Prueba de aditividad de Tukey**

E.V.	S.C.	G.L.	C.M.	F	$F_{0.05}$	$F_{0.01}$	P
No Aditividad	302.1	1	302.1	5.2	4.84	9.65	0.0435
Residuo	639.3	11	58.1				
Error	941.4	12	78.5				

De acuerdo con el contenido del Cuadro 5, los datos no siguen el modelo aditivo. Puede verse que el valor de F igual a 5.2 tiene una probabilidad menor que el 5% ( $P = 0.04$ ). Por lo tanto, los resultados obtenidos en el Análisis de Varianza del Cuadro 2 pueden llevar a conclusiones erradas.

Con el programa Statistix, a continuación de los análisis de varianza de los diseños de bloques completos al azar (BCA) se presenta la prueba de No Aditividad. Con los datos del Cuadro 1 los resultados son:

**Cuadro 6. Análisis de varianza de la variable peso de malezas, en un diseño de bloques completos al zar. (STX)**

Source	DF	SS	MS	F	P
BLOQ	4	323.53	80.88		
TRAT	3	3425.98	1141.99	14.56	0.0003
Error	12	941.37	78.45		
Total	19	4690.89			

Grand Mean 15.513    CV 57.09

**Tukey's 1 Degree of Freedom Test for Nonadditivity**

Source	DF	SS	MS	F	P
Nonadditivity	1	302.082	302.082	5.20	0.0435
Remainder	11	639.288	58.117		

Como puede constatarse, los resultados son iguales a los presentados en los Cuadros 2 y 5.

Para la corrección de la No Aditividad se recomienda una transformación logarítmica de los datos. Puede usarse cualquier base de logaritmos. En este ejercicio se empleará los logaritmos de base 10. Cuando se presentan datos menores de uno, para evitar logaritmos negativos, como en este ejercicio, la transformación que debe emplearse es:  $\log_{10}(Y_{ij} + 1)$ .

Los datos originales del Cuadro 1 se presentan, transformados a la escala logarítmica, en el Cuadro 7:

Cuadro 7. Transformación de la variable “peso de malezas” en decagramos a la escala  $\log_{10}(Y_{ij} + 1)$

BLOQ	TRATAMIENTOS				$R_j$
	A	B	C	D	
1	0.4048	1.3685	0.7818	1.5260	4.0810
2	0.6702	1.1775	0.4116	1.7390	3.9984
3	0.3054	0.9881	0.3541	1.3863	3.0339
4	0.1987	1.5622	0.6325	1.6394	4.0327
5	0.8639	1.3993	1.0022	1.3193	4.5847
$T_i$	2.4430	6.4956	3.1821	7.6100	$\sum Y_{ij} = 19.7307$
$\bar{Y}_i$	0.4886	1.2991	0.6364	1.5220	$\bar{Y} = 3.9461$

Por lo demás, el análisis de varianza procede en forma normal con los datos transformados. En el programa Statistix se incluye, a continuación del ADEVA, la prueba de No Aditividad. Los resultados del análisis se presentan a continuación:

Cuadro 8. Análisis de varianza de la variable peso de malezas transformada a una escala logarítmica

Source	DF	SS	MS	F	P
BLOQ	4	0.31710	0.07927		
TRAT	3	3.77475	1.25825	25.92	0.0000
Error	12	0.58251	0.04854		
Total	19	4.67436			

Grand Mean 0.9865 CV 22.33

**Tukey's 1 Degree of Freedom Test for Nonadditivity**

Source	DF	SS	MS	F	P
Nonadditivity	1	0.05928	0.05928	1.25	0.2881
Remainder	11	0.52323	0.04757		

Relative Efficiency, RCB 1,10

Observe que el valor de F para la prueba de No Aditividad es igual a 1.25, con una probabilidad P igual a 28.81%. Es decir que las desviaciones de la aditividad no fueron significativas. En otras palabras, la transformación logarítmica surtió efecto. Los datos ahora no se apartan del modelo lineal aditivo. Consecuentemente las conclusiones a las que se llegue a partir de estos resultados tienen validez estadística.

Observe también que gracias a la transformación logarítmica de los datos aumentó sensiblemente la precisión del experimento. El coeficiente de variación disminuyó de 57.09% a 22.23%. El valor de F es ahora igual a 25.92. La probabilidad de que estos resultados se deban exclusivamente al azar es  $P \leq 0.00009$ .

Para reportar los resultados, como ejemplo, después de realizar pruebas de Tukey u otras, las medias deben expresarse en la escala original. Para este caso en que se sumó 1 al valor original del peso de las malezas, debe encontrarse el antilogaritmo de la media transformada menos uno. Así, como ejemplo para el tratamiento A, que tiene media transformada igual a 0.4886 (Cuadro 7), se tiene:  $\text{anti log}_{10}(\bar{Y}_A) = (10^{\bar{Y}_A} - 1) = (3.08 - 1) = 2.08$ . Note que este valor, 2.08 decagramos, es diferente al de la media de los datos originales.

**Comentario:** John Tukey desarrolló también la prueba de falta de aditividad para el diseño Cuadrado Latino. Este procedimiento puede encontrarse en el texto de Snedecor y Cochran (1967), página 334.



## Capítulo 24

# Estadística no paramétrica

---

*“Una respuesta aproximada al problema correcto vale mucho más que una respuesta exacta a un problema indefinido”.* John Tukey<sup>55</sup>

Hasta aquí se han visto principalmente varios tipos de análisis para variables cuantitativas. Los métodos empleados fueron los de la estadística paramétrica, es decir, la relacionada con poblaciones que tienen media y varianza. Con cierto tipo de datos, aquellos que no pueden ser analizados con los métodos normales del análisis de varianza, porque violan las presunciones requeridas, puede procederse al análisis empleando técnicas distintas. Estas técnicas pertenecen al campo de la *estadística no paramétrica*. Su uso depende de las características de los datos, de cómo pueden agruparse y sobre todo, del objetivo de la investigación.

En el Capítulo 8 se presentó la distribución Ji-Cuadrado y se vieron varias aplicaciones, incluyendo dos relacionadas con la estadística no paramétrica: la prueba para la “bondad de ajuste” y la “prueba de contingencia”. En este capítulo se estudiarán tres nuevos procedimientos de análisis con métodos no paramétricos. Estos métodos se emplean cuando hay evidencia que los datos no tienen una distribución normal o que las varianzas de las poblaciones no son iguales. Estas condiciones se presentan generalmente en los siguientes casos:

- Las observaciones provienen de escalas preestablecidas por el investigador, como ejemplo, una escala de 1 a 10 para calificar el grado de susceptibilidad de a una enfermedad vegetal. Otro ejemplo puede ser las respuestas de un panel con relación a un nuevo producto. Este es el caso de variables cualitativas transformadas a escalas numéricas.
- Las observaciones provienen de una variable cuantitativa discreta, en la que puede suponerse una distribución diferente a la normal o heterogénei-

---

55 Estadístico estadounidense. [https://www.azquotes.com/author/14847-John\\_Tukey](https://www.azquotes.com/author/14847-John_Tukey)



dad de varianzas. Ejemplos: número de larvas en tallos de maíz, número de semillas germinadas, número de nódulos en raíces de leguminosas, etc.

- Las observaciones provienen de una variable cuantitativa continua, pero las varianzas de las poblaciones no son estadísticamente iguales. En estos casos podrían transformarse los datos y usar los métodos paramétricos, como se vio en el capítulo anterior. Sin embargo, las transformaciones, cuando las diferencias son muy grandes, no siempre producen el resultado deseado.

En todo caso, los análisis paramétricos son más eficientes y precisos que los no paramétricos y generan más información acerca de las poblaciones en estudio. Esto implica que debe preferirse su uso si se cumplen los requisitos necesarios.

Los requerimientos que las muestras deben ser tomadas al azar, que haya independencia de las observaciones dentro de cada muestra e independencia entre muestras son también necesarios para la validez de los experimentos no paramétricos.

## 24.1. Análisis de Varianza de Kruskal – Wallis

El análisis de varianza de Kruskal – Wallis (KW) es un método no paramétrico con el que se comparan varios tratamientos, que pueden tener igual o diferente número de observaciones. En este sentido es similar al diseño completamente al azar (DCA). Sin embargo, en la prueba KW el análisis se basa en las ubicaciones ordinales de los datos experimentales. Esto es, las observaciones experimentales se transforman a rangos, que se designan como valores  $r_i$ .

El valor de la prueba estadística KW se encuentra con la siguiente fórmula:

$$KW = \frac{1}{V^2} \left[ \sum \frac{R_i^2}{n_i} - n^* \frac{(n+1)^2}{4} \right] \quad (141)$$

Ecuación en la que:

- $R_i$  son las sumatorias de los rangos de cada tratamiento
- $n_i$  son los números de rangos de cada tratamiento
- $n$  es el número total de rangos
- $V^2$  es una cantidad similar a la varianza (“cuasi varianza”) que se calcula con la siguiente fórmula:

$$V^2 = \frac{1}{n-1} \left[ \sum r_i^2 - n^* \frac{(n+1)^2}{4} \right] \quad (142)$$

Ecuación en la que  $r_i$  son todos y cada uno de los rangos del experimento.

## Las hipótesis

La hipótesis nula que se somete a prueba es que las poblaciones de las que se obtuvieron las muestras tienen distribuciones similares. La hipótesis alternativa plantea que por lo menos la distribución de una población es diferente a la de las otras. Note que no se especifica la forma o tipo de distribución. El valor de la Prueba KW se compara con valores de Ji-cuadrado al nivel de significación deseado y con  $k - 1$  grados de libertad, donde  $k$  es el número de tratamientos.

## La asignación de rangos

Para realizar la prueba KW todos los datos u observaciones experimentales, independientemente del tratamiento del que provengan, se ordenan del menor al mayor, se numeran ordinalmente y luego, se transforman a una escala de rangos. *En el caso de datos repetidos, o empates, se asigna el rango correspondiente a la media de los números ordinales.* Por facilidad del ejemplo supongamos solo dos tratamientos A y B con estos resultados: tratamiento A: 10,15, 17, 20 y tratamiento B: 13, 15, 17, 17, 18, 21. (Note que los tratamientos pueden tener diferente número de observaciones.) En el siguiente cuadro se observa el procedimiento:

**Cuadro 24.1. Transformación de datos a una escala ordinal y asignación de rangos.**

Datos ordenados	10	13	15	15	17	17	17	18	20	21
Número ordinal	1	2	3	4	5	6	7	8	9	10
Rangos asignados	1.0	2.0	3.5	3.5	6.0	6.0	6.0	8.0	9.0	10.0

El análisis consecuentemente procedería con los rangos de los dos tratamientos:

Tratamiento A: rangos 1.0, 3.5, 6.0, 9.0

Tratamiento B: rangos 2.0, 3.5, 6.0, 6.0, 10.0

## El ajuste de la prueba estadística KW

En teoría, los valores KW siguen aproximadamente la distribución Ji-Cuadrado. Sin embargo, por causa especialmente de la presencia de rangos repetidos (empates), pueden subestimarse los valores KW. Para compensar esta tendencia, algunos investigadores, entre ellos Lehmann (2006), sugieren el empleo de un factor de ajuste A como denominador del valor KW. La fórmula del factor de ajuste es:

$$A = 1 - \frac{\sum(e_i^3 - e_i)}{n^3 - n} \quad (143)$$

Fórmula en la que  $e_i$  es el número total de rangos empatados en cada tratamiento.

Si se rechaza la hipótesis nula puede identificarse, a un nivel de probabilidad determinado, qué distribuciones son estadísticamente similares y cuáles son diferentes. El procedimiento es similar a la prueba de la Diferencia Mínima Significativa (Capítulo 17). La fórmula desarrollada por Conover e Iman (1981) es la siguiente desigualdad:

$$|\bar{r}_a - \bar{r}_b| > t_{\alpha/2} * \sqrt{V_{2*} * \left(\frac{n-1-k}{n-k}\right) * \sqrt{\frac{1}{n_a} + \frac{1}{n_b}}} \quad (144)$$

El valor  $|\bar{r}_a - \bar{r}_b|$  es la diferencia absoluta (signo positivo) entre cualesquiera de dos promedios de rangos que se quiera comparar. El valor de  $t_{\alpha/2}$  se encuentra en el Apéndice B, con  $n-k$  grados de libertad.

Veamos un ejemplo:

**Ejercicio 24.1.1.** En un experimento se evaluó el efecto de varios fungicidas en la germinación de semillas de maíz. Los tratamientos fueron: A, B, C, D y el testigo T. Cada tratamiento se asignó al azar a ocho parcelas en las que se sembraron 20 semillas. El suelo fue previamente inoculado con hongos patógenos. Los resultados se presentan en el siguiente cuadro:

Cuadro 1. Número de plántulas obtenidas en los cinco tratamientos

Tratamientos	Número de plántulas							
A	12	13	15	11	14	16	10	12
B	6	9	8	11	8	10	7	10
C	11	15	9	12	8	14	13	12
D	16	18	15	14	17	16	13	17
T	8	5	7	4	3	5	2	7

Note que los datos del Cuadro 1 corresponden a una variable cuantitativa discreta.

El siguiente paso es ordenar, en forma ascendente, todos los datos y asignarles un rango. Esto se presenta en el Cuadro 2.

Cuadro 2. Datos ordenados ascendentemente y rango asignado

Datos	Rango	Datos	Rango	Datos	Rango	Datos	Rango
2	1	8	11.5	11	20	14	30
3	2	8	11.5	12	23.5	15	33
4	3	8	11.5	12	23.5	15	33
5	4.5	9	14.5	12	23.5	15	33
5	4.5	9	14.5	12	23.5	16	36
6	6	10	17	13	27	16	36
7	8	10	17	13	27	16	36
7	8	10	17	13	27	17	38.5
7	8	11	20	14	30	17	38.5
8	11.5	11	20	14	30	18	40

En el Cuadro 3 se presentan los tratamientos con los datos convertidos a rangos:

Cuadro 3. Sumatorias  $R_i$ , número de rangos  $n_i$ , promedio de los rangos  $\bar{r}_i$  y número de empates  $e_i$ .

Tratamientos	Observaciones convertidas a rangos								$R_i$	$n_i$	$\bar{r}_i$	$e_i$
A	23.5	27	33	20	30	36	17	23.5	210	8	26.3	8
B	6	14.5	11.5	20	11.5	17	8	17	105.5	8	13.2	7
C	20	33	14.5	23.5	11.5	30	27	23.5	183	8	22.9	8
D	36	40	33	30	38.5	36	27	38.5	279	8	34.9	7
T	11.5	4.5	8	3	2	4.5	1	8	42.5	8	5.3	5

Puede observarse en el Cuadro 2 que hubo 35 rangos empatados. Los rangos que no tuvieron empates fueron: 2, 3, 4, 6 y 40. Con esta información se crea la columna  $e_i$  en el Cuadro 3.

Antes de calcular la prueba estadística KW es necesario encontrar la sumatoria de los cuadrados de todos los rangos (Cuadro 3). Esto es:

$$\sum r_i^2 = 23.5^2 + 27^2 + 33^2 + \dots + 8^2 = 22115$$

Ahora se puede encontrar el valor de  $V^2$ :

$$V^2 = \frac{1}{39} \left[ 22115 - 40 * \frac{(40 + 1)^2}{4} \right] = 136.03$$

Con la información del Cuadro 3 se procede a encontrar el valor de la sumatoria de los cuadrados de los rangos:

$$\sum \frac{R_i^2}{n_i} = \frac{210^2}{8} + \frac{105^2}{8} + \frac{183^2}{8} + \frac{279^2}{8} + \frac{42.5^2}{8} = 21045.81$$

Con los resultados anteriores se encuentra la prueba estadística  $KW$ :

$$KW = \frac{1}{136.03} \left[ 21045.81 - 40 * \frac{(40 + 1)^2}{4} \right] = 31.14$$

Por el alto valor de la prueba estadística  $KW$  no hace falta aplicar el factor de ajuste  $A$ . Sin embargo, como ejemplo, el valor  $A$  se encuentra con la fórmula (143) y a partir del número de rangos empatados de cada tratamiento (última columna del Cuadro 3):

$$A = 1 - \frac{(8^3-8) + (7^3-7) + (8^3-8) + (7^3-7) + (5^3-5)}{40^3-40} = 0.9719$$

El valor ajustado de la prueba estadística sería entonces  $KW_A$ :

$$KW_A = \frac{KW}{A} = \frac{31.14}{0.97} = 32.04$$

El valor de Ji-cuadrado con  $k - 1 = 4$  grados de libertad y al nivel alfa del 0.01 es igual a 13.27 (Apéndice C). Por lo tanto, la prueba estadística es altamente significativa. Esto quiere decir que debe rechazarse la hipótesis nula y aceptar la hipótesis alternativa de que por lo menos una distribución poblacional es diferente de las otras.

Mediante la fórmula (144) se puede determinar, estadísticamente, qué distribuciones fueron diferentes. Si se toma, como ejemplo, el nivel de significación alfa 0.01 con 35 grados de libertad ( $n-k$ ), entonces el lado derecho de la desigualdad es:

$$2.72 * \sqrt{136.03 * \left( \frac{40-1-31.14}{40-5} \right)} * \sqrt{\frac{1}{8} + \frac{1}{8}} = 7.5$$

Es decir que, cualquier diferencia entre los promedios de dos rangos, mayor que 7.5, indica que las poblaciones de las que se extrajeron los tratamientos (muestras) tienen diferentes distribuciones. En el Cuadro 3 constan los prome-

dios de los rangos. Los promedios ordenados y la significación estadística de las diferencias se presentan en el Cuadro 4:

**Cuadro 4. Promedio de rangos y diferencias significativas al nivel alfa del 0.01**

Tratamientos	D	A	C	D	E
Rangos promedio	34.9	26.3	22.9	13.2	5.3
Significación al 1%*	a	b	b	c	d

\*Letras minúsculas distintas señalan diferencias entre rangos.

Con el programa *Statistix* los resultados de la prueba de Kruskal - Wallis son los siguientes:

Statistix 9.0

**Kruskal-Wallis One-Way Nonparametric AOV**

Variable	Mean Rank	Sample Size
A	26,3	8
B	13,2	8
C	22,9	8
D	34,9	8
T	5,3	8
Total	20,5	40

Kruskal-Wallis Statistic 31.1427  
 P-Value, Using Chi-Squared Approximation 0.0000

**Parametric AOV Applied to Ranks**

Source	DF	SS	MS	F	P
Between	4	4235.81	1058.95	34.68	0.0000
Within	35	1068.69	30.53		
Total	39	5304.50			

Total number of values that were tied 35  
 Max. diff. allowed between ties 0,00001

Cases Included 40 Missing Cases 0

## Comentarios:

- Note que Statistix presenta también el análisis de varianza paramétrico. Este análisis se realiza con los rangos. Es una alternativa más rápida que la prueba KW, sin embargo la prueba es poco conservadora. Esto determina que los valores de F sean generalmente más altos que los valores de la prueba KW, como en este caso ( $KW = 31.14$  vs  $F = 34.68$ ).
- En experimentos pequeños, con tres tratamientos y hasta 15 observaciones, la aproximación de Ji-Cuadrado se desvía excesivamente de las probabilidades reales. Para estos casos se han desarrollado tablas con las probabilidades exactas (Conover, 1999). Estos valores también pueden encontrarse en internet, como ejemplo, en la referencia de pie de página<sup>56</sup>.

## 24.2. Análisis de Varianza de Friedman

El Análisis de Varianza de Friedman es otra prueba muy importante de la estadística no paramétrica. Este análisis es equivalente al de un Diseño de Bloques al Azar (BCA), pero tiene como ventaja, igual que otras pruebas no paramétricas, que no requiere el cumplimiento de las presunciones de las pruebas paramétricas. Idealmente se emplea para estudios que involucren variables cualitativas y cuantitativas discretas. Sin embargo, también se usa para casos de variables cuantitativas continuas que no provengan de poblaciones de distribución normal o que tengan varianzas heterogéneas.

Como se trata de un diseño equivalente a un BCA, las unidades experimentales están agrupadas en bloques, dentro de los cuales se sortean los tratamientos. En algunos experimentos los resultados relacionados a los bloques no interesan mayormente. Pero debe tomarse en cuenta que el BCA es un experimento con dos factores (tratamientos y bloques) y habrá casos en que el investigador tenga interés en las respuestas de los dos.

### *La prueba estadística*

La prueba estadística MF (por Milton Friedman) genera, a semejanza de la prueba de KW, un valor que se aproxima a la distribución Ji-Cuadrado. La fórmula es:

---

56 National Academy of Sciences (US) Statistical Tables. <https://onlinepubs.trb.org/onlinepubs/nchrp/cd-22/manual/v2appendixc.pdf>

$$MF_T = \frac{12}{(k^2+k)} * \frac{\sum R_T}{b} - 3b*(k + 1) \quad (145)$$

Fórmula en la que:

- $MF_T$  es la prueba de Friedman para los tratamientos
- $k$  es el número de tratamientos
- $R_T$  es la suma de los rangos en el tratamiento  $k$
- $b$  es el número de bloques
- Los valores 12 y 3 son constantes en la fórmula

Si es que hay interés en probar la hipótesis nula de que los bloques provienen de poblaciones con distribuciones similares, la fórmula de la prueba estadística es:

$$MF_B = \frac{12}{(b^2+b)} * \frac{\sum R_B}{k} - 3k*(b + 1) \quad (146)$$

Fórmula en la que  $R_B$  es la suma de los rangos en el bloque  $j$

### ***El factor de ajuste***

A semejanza de la prueba KW, por la presencia de rangos empatados ( $e_i$ ) se presenta la tendencia a subestimar el valor de la prueba de Friedman, por lo que se emplea un factor de ajuste para el efecto de los tratamientos:

$$A_T = 1 - \frac{\sum(e_i^3 - e_i)}{bk*(k^2-1)} \quad (147)$$

La expresión  $e_i$  de la fórmula (147) representa el número de rangos empatados, que puede ser dos o más. Esto se toma en cuenta dentro de cada bloque o hilera. Suponga que en una hilera los valores de los tratamientos tienen estos rangos a, a, a, b, b. Hay por lo tanto tres y dos rangos empatados. El aporte de esta hilera al numerador de la fórmula será:  $(3^3-3) + (2^3-2) = 30$ . Esta cantidad se sumará a las cantidades de los otros bloques o hileras.

De igual manera, si hay interés en los bloques, la fórmula para el factor de ajuste tiene en el denominador la cantidad  $bk*(b^2-1)$ . El numerador  $\sum(e_i^3 - e)$  se obtiene a partir de los empates dentro de cada uno de los tratamientos o columnas.



Si se rechaza la hipótesis nula es posible identificar cuál o cuáles poblaciones son estadísticamente diferentes. Con esta finalidad se han desarrollado varios procedimientos. Como ejemplo, en este texto se empleará un método no paramétrico que se basa en los valores de la distribución t de Student para la prueba de Tukey. Esta prueba no es excesivamente conservadora, como su contraparte paramétrica, ni excesivamente liberal como otras pruebas (Siegel y Castellan, 1988). La fórmula es:

$$|R_i - R_j| > q_\alpha * \sqrt{\frac{bk^*(k+1)}{12}} \quad (148)$$

En esta prueba  $R_i$  y  $R_j$  representan las sumatorias de los rangos de cualesquiera de dos tratamientos. Si se desea probar diferencias entre los bloques,  $R_i$  y  $R_j$  serán las sumatorias de los rangos de cualesquiera de dos bloques. En el caso para bloques, el numerador dentro del signo de raíz cuadrada cambia a  $bk^*(b + 1)$ .

Veamos un ejemplo:

**Ejercicio 24.2.** En un experimento se realizó una prueba de sabor de seis productos por parte de cinco catadores. Se les pidió calificar los productos en una escala de 1 (poco agradable) a 5 (muy agradable). En las calificaciones se permitió que los catadores expresaran algunas de sus decisiones con valores intermedios. Los resultados se presentan en el Cuadro 1.

La asignación de rangos para los tratamientos (productos) se realiza por hileras. Por ejemplo, para la primera hilera las calificaciones puestas en orden son: 3, 3, 3, 3.5, 4, 5. Los puestos 1,2 y 3 tienen la misma calificación, por lo tanto, el rango que se asigna a la calificación 3 es el promedio de la suma de los números ordinales, esto es  $6/3=2$ . Los rangos se presentan en el Cuadro 2.

**Cuadro 1. Calificaciones originales de cinco catadores para la prueba de sabor de seis productos**

Catadores	Productos					
	A	B	C	D	E	F
Cat1	3	3	4	3.5	3	5
Cat2	3.5	2.5	4	3	1.5	4
Cat3	2.5	2.5	3.5	4	3.5	3.5
Cat4	2.5	4	3.5	3.5	2	4
Cat5	3.5	2.5	4	3	2	4.5

Cuadro 2. Calificaciones convertidas a rangos para la prueba de sabor de seis productos

Hileras	Productos					
	A	B	C	D	E	F
H1	2	2	5	4	2	6
H2	4	2	5.5	3	1	5.5
H3	1.5	1.5	4	6	4	4
H4	2	5.5	3.5	3.5	1	5.5
H5	4	2	5	3	1	6
$\sum r_i = R_T$	13.5	13	23	19.5	9	27
$\bar{r}_i$	2.7	2.6	4.6	3.9	1.8	5.4

Al sumar los rangos de cada tratamiento se obtienen los valores  $R_T$  (Cuadro 2). Mediante la fórmula (145) se encuentra la prueba estadística MF para los tratamientos (o productos):

$$MF_T = \frac{12}{(36 + 6)} * \frac{13.5^2 + 13.0^2 + \dots + 27.0^2}{5} - 15 * 7 = 13.31$$

El factor de ajuste (fórmula 144) se calcula a partir de las hileras. Esto es igual a:

$$A_T = 1 - \frac{(3^3 - 3) + (2^3 - 2) + (2^3 - 2) + (3^3 - 3) + (2^3 - 2) + (2^3 - 2)}{30 * 35} = 0.9314$$

Note, como ejemplo, que en la primera hilera (H1) hay tres datos empata- dos, por lo que  $(e^3 - e) = (3^3 - 3)$ . En la tercera hilera hay dos y tres datos empata- dos, por lo que la contribución de esta hilera al ajuste es:  $(2^3 - 2) + (3^3 - 3)$ .

El valor de la prueba MF para tratamientos, ajustada por razón de los em- pates, es igual a:

$$MF_T(A) = \frac{13.31}{0.9314} = 14.29$$

En la tabla de Ji-Cuadrado (Apéndice C), al nivel del 5% y con  $k-1 = 5$  grados de libertad, el valor crítico es igual a 11.1. Como el valor de la prueba es- tadística 14.29 es mayor que el valor de Ji-Cuadrado en la tabla, se concluye que hay suficiente evidencia para rechazar la hipótesis nula. Es decir que, por lo menos una de las poblaciones tiene una distribución diferente.

Para identificar la o las poblaciones con distribuciones diferentes se puede emplear la fórmula (148). El procedimiento es semejante al de la prueba de Tukey. El valor  $q_{\alpha}$  se encuentra en la tabla respectiva (Apéndice E), al nivel alfa del 5% o 1% y en la columna para el número de tratamientos, en este caso, seis. El número de grados de libertad para la prueba es  $bk-k = 24$ . El lado derecho de la desigualdad es el valor crítico para comparar las diferencias entre los valores  $R$ :

$$q_{0.05}^* \sqrt{\frac{bk^*(k+1)}{12}} = 4.37 * 4.18 = 18.26$$

Si la diferencia absoluta de dos rangos excede el valor crítico de 18.26 se concluye que las dos poblaciones tienen estadísticamente diferentes distribuciones. Los resultados ordenados se presentan en el Cuadro 3.

**Cuadro 3. Sumatorias de rangos de los productos y diferencias significativas al 5%**

Productos	F	C	D	A	B	E
Sumatoria de rangos	27.0	23.0	19.5	13.5	13.0	9.0
Significación al 5%	a	ab	ab	ab	ab	b

\*Letras minúsculas distintas marcan diferencias estadísticas significativas

Note que la única diferencia significativa es entre los productos F y E.

En este ejemplo, el análisis de las respuestas de los catadores (bloques) podría ser de interés. En algunos casos, esta información ayuda a interpretar en mejor forma los resultados y a refinar la técnica para futuros ensayos. El procedimiento para identificar posibles diferencias entre catadores (bloques) es similar al utilizado para los productos (tratamientos).

A partir del Cuadro 1 de los datos originales, la asignación de rangos para los catadores (bloques) se realiza por columnas. Por ejemplo, para la primera columna (C1) las calificaciones originales (Cuadro 1) fueron 3, 3.5, 2.5, 2.5, 3.5. Puestas en orden del 1 al 5 son: 2.5, 2.5, 3, 3.5, y 3.5. Los puestos 1 y 2 están empatados por lo que se asignan los rangos 1.5 y 1.5. Esto es:  $(1+2)/2$  para cada rango. También están empatados los puestos 4 y 5, por lo que se asignan los rangos 4.5 y 4.5. Esto es:  $(4+5)/2$  para cada rango. A la calificación 3 le corresponde en este caso el rango 3. Por lo tanto, la columna C1 queda así: 3, 4.5, 1.5, 1.5 y 4.5. Se continúa de manera similar con las otras columnas. El resultado se presenta en el Cuadro 4.

Cuadro 4. Calificaciones convertidas a rangos para la variable “Catadores”

Catadores	Columnas						$R_B$	$\bar{r}_j$
	C1	C2	C3	C4	C5	C6		
Cat1	3	4	4	3.5	4	5	23.5	3.92
Cat2	4.5	2	4	1.5	1	2.5	15.5	2.58
Cat3	1.5	2	1.5	5	5	1	16	2.67
Cat4	1.5	5	1.5	3.5	25	2.5	16.5	2.75
Cat5	4.5	2	4	1.5	2.5	4	18.5	3.08

A continuación, mediante la fórmula (146), se calcula la prueba estadística MF para los bloques, en este caso, para los catadores:

$$MF_B = \frac{12}{(25 + 5)} * \frac{23.5^2 + 15.5^2 + \dots + 18.5^2}{6} - 18 * 6 = 2.87$$

El valor del ajuste para los catadores, por razón de los empates en las columnas o tratamientos, tiene el siguiente numerador  $\sum(e_i^3 - e)$ , esto es:

$$(2^3 - 2) + (2^3 - 2) + (3^3 - 3) + (3^3 - 3) + (2^3 - 2) + (2^3 - 2) + (2^3 - 2) + (2^3 - 2) + (2^3 - 2) = 90$$

Note, como ejemplo, que en la C1 hay dos empates de dos rangos cada uno. En la C2 hay tres rangos empatados. Observe los rangos empatados en las otras columnas.

Por lo anterior, el valor del ajuste para la variable “catadores” es igual a:

$$A_B = 1 - \frac{\sum(e_j^3 - e_j)}{bk*(b^2 - 1)} = 1 - \frac{90}{30 * 24} = 0.875$$

El valor de la prueba estadística MF ajustada en función de los empates es:

$$FM_B(A) = \frac{FM_B}{A_B} = \frac{2.87}{0.87} = 3.28$$

El valor de la prueba estadística 3.28 es menor que el valor crítico de Ji-Cuadrado, al nivel de probabilidad alfa 5% y con b – 1 grados de libertad. Esto es:  $\chi^2_{(0.05)(4gl)} = 9.49$ . (Apéndice C.) Consecuentemente se concluye que no hay suficiente evidencia para rechazar la hipótesis nula.

El resultado anterior indica que no hubo diferencias estadísticas significativas entre los catadores. Es decir, en términos generales, no existió mayor discre-

pancia entre los catadores al calificar el sabor de los productos. Esta conclusión genera más confianza en los resultados obtenidos con la prueba de Friedman para los productos.

Con el programa Statistix los resultados fueron:

**Friedman Two-Way Nonparametric AOV for Calificac = Productos  
Catadores**

<b>Productos</b>	<b>Mean Rank</b>	<b>Sample Size</b>
1	2,70	5
2	2,60	5
3	4,60	5
4	3,90	5
5	1,80	5
6	5,40	5

Friedman Statistic, Corrected for Ties 14.294  
P-value, Chi-Squared Approximation 0,0138  
Degrees of Freedom 5

<b>Catadores</b>	<b>Mean Rank</b>	<b>Sample Size</b>
1	3,92	6
2	2,58	6
3	2,67	6
4	2,75	6
5	3,08	6

Friedman Statistic, Corrected for Ties 3.2762  
P-value, Chi-Squared Approximation 0,5127  
Degrees of Freedom 4

### **Comentarios:**

- Observe que Statistix presenta los promedios de los rangos. Las fórmulas que emplea para calcular las pruebas estadísticas son diferentes a las del ejemplo, pero obviamente los resultados son iguales, salvo errores por redondeo.
- En la prueba de Friedman, la aproximación de Ji-Cuadrado, cuando el número de tratamientos es tres o cuatro o el número total de observacio-

nes es menor que 13, se desvía sensiblemente de los valores críticos reales. Para estos casos, igual que en la prueba de Kruskal-Wallis, debe acudirse a tablas, en internet u otros textos, que presentan las probabilidades exactas.

### 24.3. La correlación de rangos de Spearman

La correlación simple es una medida del grado de asociación de dos variables. Cuando se puede presumir que por lo menos una de las dos variables tiene una distribución normal, se usa el análisis de correlación de Pearson, método descrito en el Capítulo 12. Si las variables no siguen la distribución normal, como es generalmente el caso de las variables discretas, una alternativa es transformar los datos. Si esto no surte efecto, entonces el camino a seguir es emplear el procedimiento no paramétrico conocido como *correlación de rangos de Spearman*.

Para el análisis los datos pueden colectarse como rangos o puede asignarse los rangos después. La correlación de Spearman mide la correspondencia entre los rangos de observaciones pareadas que provienen de dos variables,  $X$  y  $Y$ . Sin embargo, la designación  $X$  y  $Y$  no presupone la dependencia de una de las variables con relación a la otra.

Los rangos deben asignarse en forma independiente a cada par de observaciones de las variables  $X$  y  $Y$ . Los rangos se asignan mediante un procedimiento similar al empleado para la prueba de Kruskal y Wallis o la prueba de Friedman.

#### *Las hipótesis*

La hipótesis nula plantea que el parámetro de la correlación de Spearman, identificado con la letra griega “rho” con el subíndice “s” ( $\rho_s$ ), es igual a cero. La hipótesis alternativa por el contrario señala que  $\rho_s$  no es igual a cero:

$$H_0: \rho_s = 0 \quad \text{vs} \quad H_1: \rho_s \neq 0$$

#### *El coeficiente de correlación de Spearman*

El coeficiente de correlación de Spearman  $r_s$  es la estadística que mejor estima  $\rho_s$ . Igual que en el caso de la correlación de Pearson, los límites de  $r_s$  son  $-1$  y  $+1$ . El valor  $-1$  indica una correlación inversa perfecta, mientras que el valor  $+1$  indica una correlación directa perfecta (sin error). El coeficiente de correlación de Spearman es una medida de correspondencia entre los rangos, sin que esto implique necesariamente la presencia de una correlación lineal.

Para encontrar el valor de  $r_s$  se utiliza la siguiente fórmula:

$$r_s = \frac{SP_{XY}}{\sqrt{SC_X \cdot SC_Y}} \quad (149)$$

Puede notarse que el lado derecho de la fórmula (149) es igual al de la fórmula (78) del Capítulo 12, la que se usa para encontrar el coeficiente de correlación de Pearson ( $r$ ). Sin embargo, para calcular el coeficiente de correlación de Spearman ( $r_s$ ) las sumas de productos y las sumas de cuadrados se calculan a partir de los rangos de  $X$  y  $Y$ .

Cuando no se presentan empates en los rangos, tanto de  $X$  como de  $Y$ , puede utilizarse un procedimiento más rápido para encontrar el valor de  $r_s$ . La fórmula, derivada algebraicamente a partir de la fórmula anterior, es:

$$r_s = 1 - \frac{6 \sum d_i^2}{(n^3 - n)} \quad (150)$$

En esta última ecuación se tiene que:

- $d_i$  es la diferencia entre los rangos de un mismo par.
- $n$  es el número de pares.
- El valor 6 es una constante.

Incluso si hubiera pocos empates y el número de pares fuera relativamente elevado (alrededor de 20), los resultados con las dos fórmulas tenderían a diferir muy poco. Las conclusiones por lo general serían las mismas.

La prueba estadística que se utiliza es una prueba  $t$ , similar al caso de la correlación de Pearson:

$$PEt = \frac{r_s \sqrt{n-2}}{\sqrt{1-r_s^2}} \quad (151)$$

Veamos un ejemplo:

**Ejercicio 24.3.** En un experimento se estudió la posible relación entre el número de mazorcas de maíz buenas (sin daño por hongos) y el número de pupas de lepidópteros en el suelo, antes de la siembra. Los resultados obtenidos en parcelas de

20 m<sup>2</sup> se presentan en el Cuadro 1. Para el análisis se decidió emplear la prueba t para la correlación de Spearman.

**Cuadro 1. Observaciones pareadas X y Y para número de mazorcas buenas y número de pupas, respectivamente. Rangos y diferencias X-Y entre rangos**

X	63	57	48	53	74	42	54	49	66	45	68	60	51	57
Y	13	16	23	20	14	22	19	24	21	20	18	12	16	18
Rango X	11	8.5	3	6	14	1	7	4	12	2	13	10	5	8.5
Rango Y	2	4.5	13	9.5	3	12	8	14	11	9.5	6.5	1	4.5	6.5
d <sub>i</sub>	9	4	-10	-3.5	11	-11	-1	-10	1	-7.5	6.5	9	0.5	2
d <sub>i</sub> <sup>2</sup>	81	16	100	12.25	121	121	1	100	1	56.25	42.25	81	0.25	4

Debe recordarse que las hipótesis que normalmente se evalúan a un nivel alfa de 5% y 1% de probabilidad son:

$$H_0: \rho_s = 0$$

$$H_1: \rho_s \neq 0$$

A partir de la información del Cuadro 1 se puede calcular el coeficiente de correlación de Spearman.

Por razón del ejercicio se compararán los resultados que se obtengan con las dos fórmulas.

Para el caso de la fórmula (149) se requiere encontrar las sumas de productos y de cuadrados de los rangos:

- $SP_{XY} = (11*8) + (8.5*4.5) + \dots + (8.5*6.5) - \frac{105*105}{14} = -142$
- $SC_X = (11^2 + 8.5^2 + \dots + 8.5^2) - \frac{105^2}{14} = 227$
- $SC_Y = (2^2 + 4.5^2 + \dots + 6.5^2) - \frac{105^2}{14} = 226$

Con la información anterior se calcula el coeficiente de correlación de Spearman:

$$r_s = \frac{-142}{\sqrt{227*226}} = -0.6269$$



El valor crítico de la prueba estadística t, según la fórmula (151), es:

$$PEt = \frac{-0,6269 \cdot \sqrt{14-2}}{\sqrt{1-0,6269^2}} = 2,7870$$

El valor de t en la tabla B del apéndice, con  $n - 2 = 12$  grados de libertad y al nivel alfa del 5% (prueba de dos colas), es igual a 2.1790. Como el valor de la prueba estadística excede al valor de la tabla se puede señalar que existe una correlación estadísticamente significativa entre las dos variables. Puede señalarse que un menor número de pupas en el suelo incide en un mayor número de mazorcas no atacadas por hongos. La probabilidad de error es de 5%.

Con la fórmula (150) el resultado, con la información del Cuadro 1, es:

$$r_s = 1 - \frac{6 \cdot (81 + 16 + 100 + \dots + 4)}{14^3 - 14} = -0,6198$$

Como puede verse, los resultados al aplicar las dos fórmulas son bastante similares. La pequeña diferencia en este caso se debe a la presencia de dos rangos empatados en la variable X y seis rangos empatados en la variable Y.

Con el programa *Statistix* los resultados son:

Statistix 9.0 Spearman r ej cap 24

**Spearman Rank Correlations, Corrected for Ties**

MzMaizR	
PupasR	-0,6269
P-Value	0.0192

Maximum Difference Allowed Between Ties 0,00001

Cases Included 14 Missing Cases 0

## Comentario final

¡Espero que se hayan divertido!

## Referencias

---

- Bartlett, M.S. 1937. Some examples of statistical methods of research in agriculture and applied biology. *Journal of the Royal Statistical Society*, Supplement 4: 137-183.
- Brown, M.B. and A.B. Forsythe. 1974. Robust test for equality of variances. *Journal of the American Statistical Association*. 69:364-367.
- Campbell, S. K. 1981. *Equívocos y Falacias en la Interpretación de Estadísticas*. Editorial Limusa, México, D.F.
- Cochran, W.G. and G.M. Cox. 1968. *Experimental Designs*. Second Edition (first corrected printing). John Wiley & Sons, Inc. New York.
- Conover, W.J. 1999. *Practical nonparametric statistics*. (3th edition). John Wiley & Sons, Inc. New York.
- Conover, W.J. and R.L. Iman. 1981. Rank transformations as a bridge between parametric and nonparametric statistics. *The American Statistician*, 35:124-1129
- Eberhart, S.A; and W.A. Russell. 1966. Stability parameters for comparing varieties. *Crop Science*, 6: 36-40.
- Fisher, R.A. 1947. *The Design of Experiments*. Fourth Edition. Oliver and Boyd. Edinburgh, United Kingdom.
- Gomez, K.A. and A.A. Gomez. 1984. *Statistical procedures for agricultural research*. Second Edition. John Wiley & Sons, Inc. New York.
- Hallauer A.R; and J.B. Miranda, Filho. 1981. *Quantitative Genetics in Maize Breeding*. First Edition. The Iowa State University Press. Ames, Iowa. USA.
- Hamburg, M. 1974. *Basic Statistics: A Modern Approach*. Harcourt Brace Jovanovich, Inc. New York.
- Hayter, A.J. 1984. A proof of the conjecture that the Tukey-Kramer multiple comparisons procedure is conservative. *J. American Statist. Ass.*, 12, 61-75.
- Kramer, C.Y. 1956. Extension of multiple range tests to group means with unequal number of replications. *Biometrika*, 12, 307-310.
- Hogg, R.V., and E.A. Tanis. 2001. *Probability and Statistical Inference*, Prentice Hall, 6th edition. New Jersey, USA.
- Kempthorne, O. 1979. *Design and Analysis of Experiments*. First edition (with corrections). Krieger Publishing Co. Inc. New York
- Lehmann, E.L. 2006. *Nonparametric Statistical Methods*. Springer-Verlag. New York.
- Little, T. M. and F. Jackson Hills. 1972. *Statistical Methods in Agricultural Research*. Agricultural Extension, University of California. USA.
- Mendenhall, W. 1990. *Estadística para Administradores*. Segunda edición. Grupo Editorial Iberoamérica. México, D.F.

- Morris, T.R. 1999. *Experimental Design and Analysis in Animal Sciences*. CABI Publishing. Wallingford, United Kingdom.
- Neter, J. and W. Wasserman. 1974. *Applied Linear Statistical Models*. Richard D. Irwin, Inc. Homewood, Illinois. USA.
- Petersen, R.G. 1985. *Design and Analysis of Experiments*. Marcel Dekker, INC. New York
- Ruxton, G.D. and N. Colegrave. 2011. *Experimental Design for the Life Sciences*. Third Edition. Oxford University Press. United Kingdom.
- Salsburg, D. 2001. *The Lady Tasting Tea. How statistics revolutionized science in the twentieth century*. W.H. Freeman and Company. New York
- Shapiro, S. and M. Wilk. 1965. An analysis of variance test for normality. *Biometrika*, 52: 591-611.
- Siegel, S. and N. J. Castellan. 1988. *Nonparametric statistics for the behavioral sciences*. 2nd ed. Mc. Graw-Hill. New York.
- Snedecor, G.W. and W.G. Cochran. 1967. *Statistical Methods*. Sixth Edition. The Iowa State University Press. Ames, Iowa, USA.
- Statistix 9. 2008. *User's Manual*. Analytical Software. Tallahassee, Florida.
- Steel, R.G., J.H. Torrie, D.A. Dickey. 1997. *Principles and Procedures of Statistics: A biometrical Approach*. Third Edition. The McGraw-Hill Company, Inc. New York.
- Tukey, J.W. 1949. One degree of freedom for non-additivity. *Biometrika*, 5: 232-242.
- Welch, B.L. 1951. On the comparison of several mean values: an alternative approach. *Biometrika*, 38:330-336.
- Wynn, C.M; Wiggins, A.W; and Harris, S. 1997. *The Five Biggest Ideas in Science*. New York: John Wiley & Sons, Inc.
- Yates, F. 1936. A new method of arranging variety trails involving a large number of varieties. *Journal of Agricultural Science*: 26, 424-455.

## Cuadro de fórmulas numeradas en el texto

---

1.	Media de la población	$\mu = \frac{\sum Y_i}{N}$
2.	Media geométrica	$G = \sqrt[N]{Y_1 * Y_2 * Y_3 \dots Y_N}$
3.	Media armónica	$H = N / \sum_{i=1}^N \left( \frac{1}{Y_i} \right)$
4.	Media ponderada	$\mu_P = \frac{\sum f_i Y_i}{\sum f_i}$
5.	Varianza de la población, FO <sup>57</sup>	$\sigma^2 = \frac{\sum (Y_i - \mu)^2}{N}$
6.	Varianza de la población, FC <sup>58</sup>	$\sigma^2 = \frac{\sum Y_i^2 - \frac{(\sum Y_i)^2}{N}}{N}$
7.	Desviación Estándar (población)	$\sigma = \sqrt{\sigma^2}$
8.	Coficiente Variación (población)	$CV = \frac{\sigma}{\mu} * 100$
9.	Distribución Z población normal	$Z = (Y_i - \mu) / \sigma$
10.	Media de la muestra	$\bar{Y} = \sum Y_i / n$

57 FO, fórmula de origen

58 FC, Fórmula computacional

11.	Varianza de la muestra, FO	$S^2 = \frac{\sum(Y_i - \bar{Y})^2}{n-1}$
12.	Varianza de la muestra, FC	$S^2 = \frac{\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}}{n-1}$
13.	Desviación Estándar (muestra)	$S = \sqrt{S^2}$
14.	Error Estándar	$S_{\bar{Y}} = \sqrt{\frac{S^2}{n}}$
15.	Coefficiente de variación (muestra)	$cv = \frac{S}{\bar{Y}} * 100$
16.	Probabilidad, eventos independientes	$P[A \cap B] = P(A) * P(B)$
17.	Probabilidad, eventos excluyentes	$P[A \cup B] = P(A) + P(B)$
18.	Probabilidad, eventos combinados	$P = \frac{n!}{x!(n-x)!} p^x q^{(n-x)}$
19.	Media de la distribución binomial	$\mu = n * p$
20.	Varianza de la distribución binomial	$\sigma^2 = n * p * q$
21.	Media de la población de medias	$\mu_{\bar{Y}} = \mu_Y$
22.	Varianza de la población de medias	$\sigma_{\bar{Y}}^2 = \sigma_Y^2 / n$
23.	Desviación estándar población medias	$\sigma_{\bar{Y}} = \sqrt{\sigma_Y^2 / n}$
24.	Distribución Z población de medias	$Z_{\bar{Y}} = (\bar{Y}_i - \mu_{\bar{Y}}) / \sigma_{\bar{Y}}$
25.	Prueba estadística Z	$P.E.Z. = \frac{\bar{Y} - \mu_0}{\sigma_{\bar{Y}}}$

26.	Poder de la prueba	$PP = 1 - \beta$
27.	DE <sup>59</sup> población de medias, distribución binomial	$\sigma_{\bar{p}} = \sqrt{pq/n}$
28.	Distribución t de Student	$t = \frac{\bar{Y} - \mu_{\bar{Y}}}{S_{\bar{Y}}}$
29.	Intervalo de confianza, media	$P(\bar{Y} - t_{\alpha/2} S_{\bar{y}} \leq \mu \leq \bar{Y} + t_{\alpha/2} S_{\bar{y}}) = 1 - \alpha$
30.	Prueba Estadística Ji-Cuadrado	$PE\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$
31.	Intervalo de confianza, varianza	$P\left(\frac{(n-1)S^2}{\chi^2_{(\alpha/2)}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{(1-\alpha/2)}}\right) = 1 - \alpha$
32.	Prueba Ji-Cuadrado, proporciones	$\chi^2 = \sum \frac{(Ob-Es)^2}{Es}$
33.	Ji-Cuadrado, corrección de Yates	$\chi^2 = \sum \frac{( Ob-Es -0.5)^2}{Es}$
34.	Distribución F de Fisher	$F = \frac{S_A^2}{S_B^2}$
35.	Distribución F, alfa cola izquierda	$F_{(V_B, V_A)(1-\alpha)} = \frac{1}{F_{(V_A, V_B)(\alpha)}}$
36.	Prueba F, homogeneidad de varianzas	$P.E.F = \frac{S_{Mayor}^2}{S_{menor}^2}$
37.	Prueba t, observaciones no pareadas	$t = \frac{\bar{Y}_A - \bar{Y}_B}{\sqrt{S_p^2 \left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$
38.	Varianza ponderada	$S_p^2 = \frac{(n_A-1)S_A^2 + (n_B-1)S_B^2}{n_A + n_B - 2}$

39. Varianza ponderada, $n_A = n_B$	$S_p^2 = \frac{S_A^2 + S_B^2}{2}$
40. Prueba T, varianzas heterogéneas	$T = \frac{\bar{Y}_A - \bar{Y}_B}{\sqrt{\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}}}$
41. Prueba T, grados de libertad	$g. l. T = \frac{\left(\frac{S_A^2}{n_A} + \frac{S_B^2}{n_B}\right)^2}{\frac{\left(\frac{S_A^2}{n_A}\right)^2}{n_A - 1} + \frac{\left(\frac{S_B^2}{n_B}\right)^2}{n_B - 1}}$
42. PE t, observaciones pareadas	$t = \frac{\bar{d} - \mu_d}{\sqrt{S_d^2/n}}$
43. Media de las diferencias	$\bar{d} = \frac{\sum d_i}{n}$
44. Varianza de las diferencias	$S_d^2 = \frac{\sum d_i^2 - (\sum d_i)^2/n}{n - 1}$
45. Error estándar de las diferencias	$S_{\bar{d}} = \sqrt{S_d^2/n}$
46. Prueba t, diferencia entre dos medias	$t = \frac{\bar{d}}{S_{\bar{d}}}$
47. Intervalo, media de la diferencia	$P(\bar{d} - t_{\alpha/2} * S_{\bar{d}} \leq \mu_d \leq \bar{d} + t_{\alpha/2} * S_{\bar{d}}) = 1 - \alpha$
48. Ecuación de la línea recta	$Y = b_0 + b_1 X$
49. Modelo regresión lineal simple	$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$
50. Estimadores, regresión lineal simple	$Y_i = b_0 + b_1 X + e_i$

51.	Suma cuadrados del error	$\sum e_i^2 = \sum (Y_i - \hat{Y}_i)^2$
52.	Valores estimados de Y	$\hat{Y}_i = b_0 + b_1 X$
53.	Ecuaciones normales, regresión	$\sum Y_i = nb_0 + b_1 \sum X_i$ $\sum X_i Y_i = b_0 \sum X_i + b_1 \sum X_i^2$
54.	Coefficiente de regresión lineal <sup>60</sup>	$b_1 = \frac{cov_{XY}}{S_X^2}$
55.	Covarianza, muestra	$cov_{XY} = \frac{SP_{XY}}{n-1} = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{n-1}$
56.	Suma de cuadrados, X	$S_X^2 = \frac{SC_X}{n-1} = \frac{\sum X_i^2 - \frac{(\sum X_i)^2}{n}}{n-1}$
57.	Coefficiente de regresión lineal	$b_1 = \frac{SP_{XY}}{SC_X}$
58.	La intersección	$b_0 = \bar{Y} - b_1 \bar{X}$
59.	Cuadrado Medio del Error	$CME = \frac{\sum (Y_i - \hat{Y}_i)^2}{n-2} = \frac{SCE}{n-2}$
60.	Prueba t, coeficiente de regresión	$t = \frac{b_1 - \beta_1 h}{S_b}$
61.	Error estándar, coeficiente de regresión	$S_b = \sqrt{\frac{CME}{SC_X}}$
62.	Pendiente, intervalo	$P(b_1 - t_{\alpha/2} * S_b \leq \beta_1 \leq b_1 + t_{\alpha/2} * S_b) = 1 - \alpha$

60 Fórmulas 54 y 57 son equivalentes



63.	Modelo, regresión múltiple	$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 X_i^3 \dots + \beta_n X_i^n + \varepsilon_i$
64.	Modelo, regresión cuadrática	$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i$
65.	Ecuaciones normales, regresión cuadrática	$\sum Y_i = nb_0 + b_1 \sum X_i + b_2 \sum X_i^2$ $\sum X_i Y_i = b_0 \sum X_i + b_1 \sum X_i^2 + b_2 \sum X_i X_i^2$ $\sum X_i^2 Y_i = b_0 \sum X_i^2 + b_1 \sum X_i X_i^2 + b_2 \sum X_i^2 X_i^2$
66.	Suma de cuadrados total <sup>61</sup>	$SCT = \sum (Y_i - \bar{Y})^2$
67.	Suma de cuadrados total	$SCT = SC_Y = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$
68.	Regresión lineal <sup>62</sup> , fórmula original	$SC\ Reg = \sum (\hat{Y}_i - \bar{Y})^2$
69.	Regresión lineal	$SC\ Reg = \sum \hat{Y}^2 - \frac{(\sum \hat{Y})^2}{n}$
70.	Regresión lineal	$SC\ Reg = \frac{(SP_{xy})^2}{SC_x}$
71.	Componentes de Regresión	$\sum (Y_i - \bar{Y})^2 = \sum (\hat{Y}_i - \bar{Y})^2 + \sum (Y_i - \hat{Y}_i)^2$
72.	Componentes de Regresión	$SCT = SC\ Reg + SCE$
73.	Suma de Cuadrados del Error	$SCE = SCT - SC\ Reg$
74.	Cuadrado Medio de Regresión	$CM\ Reg = \frac{SC\ Reg}{k - 1}$

61 Suma de Cuadrados Total, fórmula original

62 Suma de Cuadrados de regresión, fórmula original

75.	Cuadrado Medio del Error	$CME = \frac{SCE}{n-k}$
76.	Prueba F en regresión	$F = \frac{CM\text{ Reg}}{CME}$
77.	Coefficiente de correlación simple	$r = \sqrt{b_{Y.X} * b_{X.Y}}$
78.	Coefficiente de correlación simple <sup>63</sup>	$r = \frac{SP_{XY}}{\sqrt{SC_X * SC_Y}}$
79.	Prueba t, correlación	$PEt = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$
80.	Coefficiente de determinación simple	$r^2 = \frac{(SP_{XY})^2}{SC_X * SC_Y}$
81.	Coefficiente de determinación <sup>64</sup> múltiple	$R^2 = \frac{SCReg}{SCT}$
82.	Coefficiente R ajustado	$R_{aj}^2 = 1 - \frac{(n-1)*SCE}{(n-k)*SCT}$
83.	La gran media experimental	$\bar{Y} = \sum Y_{ij}/n$
84.	Suma de Cuadrados Total	$SCT = \sum Y_{ij}^2 - FC$
85.	Factor de Corrección	$FC = \frac{(\sum Y_{ij})^2}{n}$
86.	Suma de Cuadrados de Tratamientos	$SCTrat = \sum \frac{T_i^2}{r_j} - FC$
87.	El modelo lineal aditivo del DCA <sup>65</sup>	$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}$

63 Fórmulas 77 y 78 son equivalentes

64 Esta fórmula también sirve para encontrar r2 a partir del Análisis de Varianza de Regresión

65 DCA son las siglas de Diseño Completamente al Azar

88.	Modelo I, F de Cuadrados Medios esperados	$F = \frac{\sigma^2 + \frac{r \sum \tau_i^2}{(k-1)}}{\sigma^2}$
89.	SC Repeticiones, diseño BCA	$SCR = \frac{\sum R_j^2}{k} - FC$
90.	SC Tratamientos, diseño BCA	$SCTrat = \frac{\sum T_i^2}{r} - FC$
91.	Eficiencia Relativa, BCA sobre DCA	$RE = \frac{(r-1)*CMR + r(k-1)*CME}{(rk-1)*CME}$
92.	Cuadrado Latino, SC hileras	$SCH = \sum \frac{H_i^2}{k} - FC$
93.	Cuadrado Latino, SC columnas	$SCC = \sum \frac{C_j^2}{k} - FC$
94.	Cuadrado Latino, SC tratamientos	$SCTrat = \sum \frac{T_k^2}{k} - FC$
95.	Número de comparaciones entre pares de medias	$w = \frac{k*(k-1)}{2}$
96.	La Diferencia Mínima Significativa	$DMS = t_{\alpha} * S_{\bar{d}}$
97.	La Prueba de Tukey	$TUK = q_{\alpha} * S_{\bar{y}}$
98.	La Prueba de Scheffé	$SCH = S_{\bar{d}} * \sqrt{gk * F_{\alpha}}$
99.	La Prueba de Bonferroni	$BON = t_{\alpha'} * S_{\bar{d}}$
100.	La Corrección de Bonferroni	$\alpha' = \frac{\alpha}{2 * w}$
101.	La Prueba de Duncan	$AMD = q_{\alpha} * S_{\bar{y}}$
102.	La Prueba de Student-Newman-Keuls	$SNK = Q_{\alpha'} * S_{\bar{y}}$

103. Error estándar, diferentes r	$S_{\bar{d}} = \sqrt{CME * \frac{r_m + r_n}{r_m * r_n}}$
104. Prueba de Tukey, diferentes r	$TUK' = \frac{q_{\alpha} * S_{\bar{d}}}{\sqrt{2}}$
105. Prueba de Duncan, diferentes r	$AMD = \frac{q_{\alpha'} * S_{\bar{d}}}{\sqrt{2}}$
106. Prueba SNK, diferentes r	$SNK = \frac{Q_{\alpha'} * S_{\bar{d}}}{\sqrt{2}}$
107. Prueba de Dunnett	$DNT = t_{\alpha}(Dunnett) * \sqrt{\frac{2 * CME}{r}}$
108. Intervalo de Confianza, Dunnett	$IC = (\bar{Y}_i - \bar{Y}_T) \pm t_{\alpha} * \sqrt{\frac{2 * CME}{r}}$
109. Contraste lineal	$Q_i = \sum c_i T_i - \sum c_j T_j$
110. Suma de Cuadrados del contraste lineal	$SCQ_i = \frac{Q_i^2}{D_i}$
111. Divisor de la Suma de Cuadrados	$D_i = r \sum c_i^2$
112. Sumatoria de SC de contrastes ortogonales	$\sum SCQ_i = SCTrat$
113. Látices, desviación tratamientos bloques	$D_i = (kT_i - rB_t + \sum T_i)$
114. Error (B*R), Látices balanceados	$SC(B^*R) = \frac{\sum D_i^2}{k^3 * (k + 1)}$
115. Látice balanceado, factor de ajuste	$q = \frac{CME(B^*R) - CME_{int\ rb}}{k^2 * CME(B^*R)}$
116. Tratamientos ajustados, látice balanceado	$T_{i(ajust)} = T_i + q * D_i$

117. STrat ajustados, látice balanceado	$SCTrat_{aj} = \frac{\sum T_{i(aj)}^2}{k+1} - FC SCTrat_{aj}$
118. Cuadrado Medio Error efectivo	$CME_{effect} = CME_{intrbl} * (1 + qk)$
119. Látice parcial, bloques sin ajuste	$SCB_{sinaj} = \frac{\sum B_j^2}{k} - \frac{\sum R_n^2}{k^2}$
120. CM Error Intra bloque	$CME_{intrbl} = \frac{SCE_{intrbl}}{(k-1) * (rk-k-1)}$
121. CM Bloques ajustados	$CMB_{ajust} = \frac{SCB_{ajust}}{r*(k-1)}$
122. Látice parcial, factor de ajuste	$q = \frac{CMB_{ajust} - CME_{intrbl}}{k(r-1) * CMB_{ajust}}$
123. Látice parcial, tratamientos ajustados	$SCTrat_{ajust} = SCTrat_{sinajust} - Q$
124. Látice parcial, valor de Q	$Q = qk(r-1) * \left( \frac{r*SCB_{sinajust}}{(r-1)(1+qk)} - SCB_{ajust} \right)$
125. Látice parcial, eficiencia relativa al BCA	$E = \frac{CME_{BCA}}{CME_{intrbl} \left( 1 + \frac{rkq}{k+1} \right)}$
126. Número de interacciones	$NI = \frac{n!}{r!(n-r)!}$
127. Prueba de Bartlett	$PB\chi^2 = \frac{f(k * \ln \bar{S}^2 - \sum \ln S_i^2)}{C}$
128. Covarianza, factor de corrección	$FC_{xy} = \frac{\sum X_i \sum Y_i}{rk}$
129. Covarianza, suma de productos total	$SPT = \sum X_i Y_i - FC_{xy}$
130. Suma de productos, tratamientos	$SPTrat = \frac{\sum T_x T_y}{r} - FC_{xy}$

131. Covarianza, suma productos del error	$SPE = SPT - SPT_{rat}$
132. Suma de cuadrados corregida Total	$SCCT_y = SCT_y - \frac{(SPT_{xy})^2}{SCT_x}$
133. Suma de cuadrados corregida, Error	$SCCE_y = SCE_y - \frac{(SPE_{xy})^2}{SCE_x}$
134. Suma de cuadrados corregida Tratamientos	$SCCTrat_y = SCCT_y - SCCE_y$
135. Covarianza, eficiencia relativa	$RE = \frac{CME_y}{CMCE_y * \left(1 + \frac{CMTrat_x}{SCE_x}\right)}$
136. Transformación arco seno	$Y'_i = \arcsen(\sqrt{Y_i/100}) * \frac{180}{\pi}$
137. Prueba F de Welch	$F = \frac{\frac{1}{k-1} * \sum w_i (\bar{Y}_i - \bar{Y}')^2}{1 + \frac{2(k-2)}{k^2-1} * \sum \left(\frac{1}{n_i-1}\right) * \left(1 - \frac{w_i}{w}\right)^2}$
138. Grados de libertad, prueba de Welch	$g.l. denom. = \frac{k^2 - 1}{3 * \sum \left(\frac{1}{n_i-1}\right) * \left(1 - \frac{w_i}{w}\right)^2}$
139. Prueba de medias, Games-Howell	$ \bar{Y}_i - \bar{Y}_j  > q_\alpha * S'_{\bar{Y}}$
140. Error estándar, prueba Games-Howell	$S'_{\bar{Y}} = \sqrt{\frac{1}{2} * \left(\frac{S_i^2}{n_i} + \frac{S_j^2}{n_j}\right)}$
141. Kruskal-Wallis (KW), ADEVA	$KW = \frac{1}{V^2} \left[ \sum \frac{R_i^2}{n_i} - n * \frac{(n+1)^2}{4} \right]$
142. Cuasi varianza, KW	$V^2 = \frac{1}{n-1} \left[ \sum r_i^2 - n * \frac{(n+1)^2}{4} \right]$
143. Factor de ajuste, KW	$A = 1 - \frac{\sum(e_i^3 - e_i)}{n^3 - n}$

144. Comparación entre rangos, KW	$ \bar{r}_a - \bar{r}_b  > t_{\alpha/2} * \sqrt{V^2 * \left(\frac{n-1-k}{n-k}\right) * \sqrt{\frac{1}{n_a} + \frac{1}{n_b}}}$
145. Milton Friedman (MF), ADEVA	$MF_T = \frac{12}{(k^2 + k)} * \frac{\sum R_T}{b} - 3b*(k + 1)$
146. MF, ADEVA bloques	$MF_B = \frac{12}{(b^2 + b)} * \frac{\sum R_B}{k} - 3k*(b + 1)$
147. Factor de ajuste, MF	$A_T = 1 - \frac{\sum(e_i^3 - e_i)}{bk*(k^2 - 1)}$
148. Comparación entre rangos (trat), MF	$ R_i - R_j  > q_{\alpha} * \sqrt{\frac{bk*(k + 1)}{12}}$
149. Coeficiente de correlación (CC) de Spearman	$r_s = \frac{SP_{XY}}{\sqrt{SC_X * SC_Y}}$
150. CC Spearman, sin empates	$r_s = 1 - \frac{6 \sum d_i^2}{(n^3 - n)}$
151. Prueba t, correlación de Spearman	$PEt = \frac{r_s \sqrt{n-2}}{\sqrt{1-r_s^2}}$

## Apéndice A. Tabla Z<sup>66</sup>

La columna bajo Z indica el número entero y el primer decimal de los valores críticos. Los valores que encabezan las columnas son el segundo decimal de los valores críticos.

Los valores en el cuerpo de la tabla son las áreas desde el centro al extremo izquierdo o derecho, dependiendo del signo.

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.496	0.492	0.488	0.484	0.48	0.476	0.472	0.468	0.464
0.1	0.4602	0.456	0.452	0.448	0.444	0.44	0.436	0.433	0.429	0.425
0.2	0.4207	0.417	0.413	0.409	0.405	0.401	0.397	0.394	0.39	0.386
0.3	0.3821	0.378	0.375	0.371	0.367	0.363	0.359	0.356	0.352	0.348
0.4	0.3446	0.341	0.337	0.334	0.33	0.326	0.323	0.319	0.316	0.312
0.5	0.3085	0.309	0.302	0.298	0.295	0.291	0.288	0.284	0.281	0.278
0.6	0.2743	0.271	0.268	0.264	0.261	0.258	0.255	0.251	0.248	0.245
0.7	0.2420	0.239	0.236	0.233	0.23	0.227	0.224	0.221	0.218	0.215
0.8	0.2119	0.209	0.206	0.203	0.201	0.198	0.195	0.192	0.189	0.187
0.9	0.1841	0.181	0.179	0.176	0.174	0.171	0.169	0.166	0.164	0.161
1.0	0.1587	0.156	0.154	0.152	0.149	0.147	0.145	0.142	0.14	0.138
1.1	0.1357	0.134	0.131	0.129	0.127	0.125	0.123	0.121	0.119	0.117
1.2	0.1151	0.113	0.111	0.109	0.108	0.106	0.104	0.102	0.1	0.099
1.3	0.0968	0.095	0.093	0.092	0.098	0.089	0.087	0.085	0.084	0.082
1.4	0.0808	0.079	0.078	0.076	0.075	0.074	0.072	0.071	0.069	0.068
1.5	0.0668	0.066	0.064	0.063	0.062	0.061	0.059	0.058	0.057	0.056
1.6	0.0548	0.054	0.053	0.052	0.051	0.05	0.049	0.048	0.047	0.046
1.7	0.0446	0.044	0.043	0.042	0.041	0.04	0.039	0.038	0.038	0.037
1.8	0.0359	0.035	0.034	0.034	0.033	0.032	0.031	0.031	0.03	0.029
1.9	0.0287	0.028	0.026	0.027	0.026	0.026	0.025	0.024	0.024	0.023
2.0	0.0228	0.022	0.022	0.021	0.021	0.02	0.02	0.019	0.019	0.018
2.1	0.0179	0.017	0.017	0.017	0.016	0.016	0.015	0.015	0.015	0.014
2.2	0.0139	0.014	0.013	0.013	0.013	0.012	0.012	0.012	0.011	0.011
2.3	0.0107	0.01	0.01	0.01	0.01	0.009	0.009	0.009	0.009	0.008
2.4	0.0082	0.008	0.008	0.008	0.007	0.007	0.007	0.007	0.007	0.006
2.5	0.0062	0.006	0.006	0.006	0.006	0.005	0.005	0.005	0.005	0.005
2.6	0.0047	0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
2.7	0.0035	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
2.8	0.0026	0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
2.9	0.0019	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.001	0.001
3.0	0.0013	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
3.1	0.0010	9E-04	9E-04	9E-04	8E-04	8E-04	8E-04	8E-04	7E-04	7E-04
3.2	0.0007	7E-04	6E-04	6E-04	6E-04	6E-04	6E-04	5E-04	5E-04	5E-04
3.3	0.0005	5E-04	5E-04	4E-04	4E-04	4E-04	4E-04	4E-04	4E-04	3E-04
3.4	0.0003	3E-04	3E-04	3E-04	3E-04	3E-04	3E-04	3E-04	3E-04	2E-04



## Apéndice B. Tabla T<sup>67</sup>

Las probabilidades, de cola derecha (o izquierda por la simetría), se encuentran en la hilera que encabeza la tabla. En el cuerpo de la tabla se encuentran los valores críticos correspondientes a los grados de libertad, que se comparan con la prueba estadística.

g.l/prob	0.4	0.25	0.1	0.05	0.025	0.01	0.005	0.0005
1	0.325	1.000	3.078	6.314	12.706	31.821	63.657	636.619
2	0.289	0.816	1.886	2.920	4.303	6.965	9.925	31.599
3	0.277	0.765	1.638	2.353	3.182	4.541	5.841	12.924
4	0.271	0.741	1.533	2.132	2.776	3.747	4.604	8.610
5	0.267	0.727	1.476	2.015	2.571	3.365	4.032	6.869
6	0.265	0.718	1.440	1.943	2.447	3.143	3.707	5.959
7	0.263	0.711	1.415	1.895	2.365	2.998	3.499	5.408
8	0.262	0.706	1.397	1.860	2.306	2.896	3.355	5.041
9	0.261	0.703	1.383	1.833	2.262	2.821	3.250	4.781
10	0.260	0.700	1.372	1.812	2.228	2.764	3.169	4.587
11	0.260	0.697	1.363	1.796	2.201	2.718	3.106	4.437
12	0.259	0.695	1.356	1.782	2.179	2.681	3.055	4.318
13	0.259	0.694	1.350	1.771	2.160	2.650	3.012	4.221
14	0.258	0.692	1.345	1.761	2.145	2.624	2.977	4.141
15	0.258	0.691	1.341	1.753	2.131	2.602	2.947	4.073
16	0.258	0.690	1.337	1.746	2.120	2.583	2.921	4.015
17	0.257	0.689	1.333	1.740	2.110	2.567	2.898	3.965
18	0.257	0.688	1.330	1.734	2.101	2.552	2.878	3.922
19	0.257	0.688	1.328	1.729	2.093	2.539	2.861	3.883
20	0.257	0.687	1.325	1.725	2.086	2.528	2.845	3.850
21	0.257	0.686	1.323	1.721	2.080	2.518	2.831	3.819
22	0.256	0.686	1.321	1.717	2.074	2.508	2.819	3.792
23	0.256	0.685	1.319	1.714	2.069	2.500	2.807	3.768
24	0.256	0.685	1.318	1.711	2.064	2.492	2.797	3.745
25	0.256	0.684	1.316	1.708	2.060	2.485	2.787	3.725
26	0.256	0.684	1.315	1.706	2.056	2.479	2.779	3.707
27	0.256	0.684	1.314	1.703	2.052	2.473	2.771	3.690
28	0.256	0.683	1.313	1.701	2.048	2.467	2.763	3.674
29	0.256	0.683	1.311	1.699	2.045	2.462	2.756	3.659
30	0.256	0.683	1.310	1.697	2.042	2.457	2.750	3.646
40	0.255	0.681	1.303	1.684	2.021	2.423	2.704	3.551
60	0.255	0.679	1.296	1.671	2.000	2.390	2.660	3.460
120	0.254	0.677	1.289	1.658	1.980	2.358	2.617	3.373
inf	0.253	0.674	1.282	1.645	1.960	2.326	2.576	3.291

67 <https://www.stat.tamu.edu/~lzhou/stat302/T-Table>.

## Apéndice C. Tabla JI-Cuadrado<sup>68</sup>

Los valores que encabezan las columnas son las áreas a la derecha del valor crítico. Las áreas a la izquierda del valor crítico se encuentran restando de uno. Como ejemplo, 0.05 a la izquierda es 0.95 a la derecha.

g.l.	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
1	0	0	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.01	0.02	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952

**Apéndice D1 . Tabla F<sup>69</sup> al 5%**

Los grados de libertad del numerador corresponden a las columnas y los del denominador a las hileras

g.l.	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	120
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	253
2	18.5	19.0	19.2	19.3	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.55
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.66
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.4
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.7
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.27
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	2.97
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.75
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.58
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.45
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.34
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.25
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.18
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.11
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.06
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.01
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	1.97
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.93
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.9
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.87
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.84
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.81
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.79
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.77
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.75
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.73
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.71
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.7
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.68
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.58
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.47
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.35

## Apéndice D2. Tabla F<sup>70</sup> al 2.5%

Los grados de libertad del numerador corresponden a las columnas y los del denominador, a las hileras

g.l.	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	120
1	648	800	864	900	922	937	948	957	963	969	977	985	993	997	1001	1006	1014
2	38.5	39	39.2	39.3	39.3	39.3	39.4	39.4	39.4	39.4	39.4	39.4	39.5	39.5	39.5	39.5	39.5
3	17.4	16	15.4	15.1	14.9	14.7	14.6	14.5	14.5	14.4	14.3	14.3	14.2	14.1	14.1	14	14
4	12.2	10.6	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84	8.75	8.66	8.56	8.51	8.46	8.41	8.31
5	10.0	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62	6.52	6.43	6.33	6.28	6.23	6.18	6.07
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46	5.37	5.27	5.17	5.12	5.07	5.01	4.9
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76	4.67	4.57	4.47	4.42	4.36	4.31	4.2
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30	4.20	4.1	4	3.95	3.89	3.84	3.73
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96	3.87	3.77	3.67	3.61	3.56	3.51	3.39
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72	3.62	3.52	3.42	3.37	3.31	3.26	3.14
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53	3.43	3.33	3.23	3.17	3.12	3.06	2.94
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37	3.28	3.18	3.07	3.02	2.96	2.91	2.79
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25	3.15	3.05	2.95	2.89	2.84	2.78	2.66
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15	3.05	2.95	2.84	2.79	2.73	2.67	2.55
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06	2.96	2.86	2.76	2.7	2.64	2.59	2.46
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99	2.89	2.79	2.68	2.63	2.57	2.51	2.38
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92	2.82	2.72	2.62	2.56	2.5	2.44	2.32
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87	2.77	2.67	2.56	2.5	2.44	2.38	2.26
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82	2.72	2.62	2.51	2.45	2.39	2.33	2.2
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77	2.68	2.57	2.46	2.41	2.35	2.29	2.16
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73	2.64	2.53	2.42	2.37	2.31	2.25	2.11
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70	2.60	2.5	2.39	2.33	2.27	2.21	2.08
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67	2.57	2.47	2.36	2.3	2.24	2.18	2.04
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64	2.54	2.44	2.33	2.27	2.21	2.15	2.01
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61	2.51	2.41	2.3	2.24	2.18	2.12	1.98
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59	2.49	2.39	2.28	2.22	2.16	2.09	1.95
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57	2.47	2.36	2.25	2.19	2.13	2.07	1.93
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55	2.45	2.34	2.23	2.17	2.11	2.05	1.91
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53	2.43	2.32	2.21	2.15	2.09	2.03	1.89
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51	2.41	2.31	2.2	2.14	2.07	2.01	1.87
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39	2.29	2.18	2.07	2.01	1.94	1.88	1.72
60	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33	2.27	2.17	2.06	1.94	1.88	1.82	1.74	1.58
120	5.15	3.8	3.23	2.89	2.67	2.52	2.39	2.3	2.22	2.16	2.05	1.94	1.82	1.76	1.69	1.61	1.43

## Apéndice D3. Tabla F<sup>71</sup> al 1%

Los grados de libertad del numerador corresponden a las columnas y los del denominador a las hileras

g.l.	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	120
2	98.5	99.0	99.2	99.3	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5
3	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.4	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.2
4	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.9	13.8	13.7	13.6
5	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.11
6	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	6.97
7	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.74
8	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	4.95
9	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.40
10	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.00
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.69
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.45
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.67	3.59	3.51	3.43	3.26
14	8.86	6.52	5.56	5.04	4.70	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.09
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.90	3.81	3.67	3.52	3.37	3.29	3.21	3.13	2.96
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.85
17	8.40	6.11	5.19	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.75
18	8.29	6.01	5.09	4.58	4.25	4.02	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.66
19	8.19	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.93	2.84	2.76	2.58
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.70	2.52
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.46
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.40
23	7.88	5.66	4.77	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.35
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.31
25	7.77	5.57	4.68	4.18	3.86	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.27
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.82	2.66	2.59	2.50	2.42	2.23
27	7.68	5.49	4.60	4.11	3.79	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.20
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.17
29	7.60	5.42	4.54	4.05	3.73	3.50	3.33	3.20	3.09	3.01	2.87	2.73	2.57	2.50	2.41	2.33	2.14
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.11
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.67	2.52	2.37	2.29	2.20	2.11	1.92
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.73
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.04	1.95	1.86	1.76	1.53

## Apéndice E1 . Tabla Tukey<sup>72</sup> al 5%

Valores q alfa

Número de medias de los tratamientos, p

g.l.	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18	20
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.8	6.99	7.17	7.32	7.47	7.60	7.72	7.83	8.03	8.21
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65	6.79	6.92	7.03	7.14	7.24	7.43	7.59
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43	6.55	6.66	6.76	6.85	7.02	7.17
8	3.26	4.04	4.53	4.89	5.17	5.40	5.6	5.77	5.92	6.05	6.18	6.29	6.39	6.48	6.57	6.73	6.87
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98	6.09	6.19	6.28	6.36	6.51	6.64
10	3.15	3.88	4.33	4.65	4.91	5.12	5.3	5.46	5.6	5.72	5.83	5.93	6.03	6.11	6.19	6.34	6.47
11	3.11	3.82	4.26	4.57	4.82	5.03	5.2	5.35	5.49	5.61	5.71	5.81	5.90	5.98	6.06	6.20	6.33
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.26	5.39	5.51	5.61	5.71	5.80	5.88	5.95	6.09	6.21
13	3.06	3.73	4.15	4.45	4.69	4.88	5.05	5.19	5.32	5.43	5.53	5.63	5.71	5.79	5.86	5.99	6.11
14	3.03	3.70	4.11	4.41	4.64	4.83	4.99	5.13	5.25	5.36	5.46	5.55	5.64	5.71	5.79	5.91	6.03
15	3.01	3.67	4.08	4.37	4.59	4.78	4.94	5.08	5.2	5.31	5.4	5.49	5.57	5.65	5.72	5.85	5.96
16	3.00	3.65	4.05	4.33	4.56	4.74	4.90	5.03	5.15	5.26	5.35	5.44	5.52	5.59	5.66	5.79	5.9
17	2.98	3.63	4.02	4.30	4.52	4.70	4.86	4.99	5.11	5.21	5.31	5.39	5.47	5.54	5.61	5.73	5.84
18	2.97	3.61	4.00	4.28	4.49	4.67	4.82	4.96	5.07	5.17	5.27	5.35	5.43	5.50	5.57	5.69	5.79
19	2.96	3.59	3.98	4.25	4.47	4.65	4.79	4.92	5.04	5.14	5.23	5.31	5.39	5.46	5.53	5.65	5.75
20	2.95	3.58	3.96	4.23	4.45	4.62	4.77	4.9	5.01	5.11	5.20	5.28	5.36	5.43	5.49	5.61	5.71
21	2.94	3.56	3.94	4.21	4.42	4.60	4.74	4.87	4.98	5.08	5.17	5.25	5.33	5.40	5.46	5.58	5.68
22	2.93	3.55	3.93	4.20	4.41	4.58	4.72	4.85	4.96	5.06	5.14	5.23	5.30	5.37	5.43	5.55	5.65
23	2.93	3.54	3.91	4.18	4.39	4.56	4.7	4.83	4.94	5.03	5.12	5.20	5.27	5.34	5.41	5.52	5.62
24	2.92	3.53	3.90	4.17	4.37	4.54	4.68	4.81	4.92	5.01	5.10	5.18	5.25	5.32	5.38	5.49	5.59
25	2.91	3.52	3.89	4.15	4.36	4.53	4.67	4.79	4.9	4.99	5.08	5.16	5.23	5.3	5.36	5.47	5.57
26	2.91	3.51	3.88	4.14	4.35	4.51	4.65	4.77	4.88	4.98	5.06	5.14	5.21	5.28	5.34	5.45	5.55
27	2.90	3.51	3.87	4.13	4.33	4.50	4.64	4.76	4.86	4.96	5.04	5.12	5.19	5.26	5.32	5.43	5.53
28	2.90	3.50	3.86	4.12	4.32	4.49	4.62	4.74	4.85	4.94	5.03	5.11	5.18	5.24	5.3	5.41	5.51
29	2.89	3.49	3.85	4.11	4.31	4.47	4.61	4.73	4.84	4.93	5.01	5.09	5.16	5.23	5.29	5.4	5.49
30	2.89	3.49	3.85	4.10	4.3	4.46	4.6	4.72	4.82	4.92	5.00	5.08	5.15	5.21	5.27	5.38	5.48
40	2.86	3.44	3.79	4.04	4.23	4.39	4.52	4.63	4.73	4.82	4.90	4.98	5.04	5.11	5.16	5.27	5.36
60	2.83	3.40	3.74	3.98	4.16	4.31	4.44	4.55	4.65	4.73	4.81	4.88	4.94	5.00	5.06	5.15	5.24
120	2.80	2.80	3.36	3.68	3.92	4.10	4.24	4.36	4.47	4.56	4.64	4.71	4.78	4.84	4.90	5.00	5.09

## Apéndice E2. Tabla Tukey<sup>73</sup> al 1%

Valores q alfa

Número de medias de los tratamientos, p

g.l.	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18	20
5	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.2	10.5	10.7	10.9	11.1	11.2	11.4	11.7	11.9
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.1	9.3	9.49	9.65	9.81	9.95	10.1	10.3	10.5
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37	8.55	8.71	8.86	9.00	9.12	9.24	9.46	9.65
8	4.75	5.64	6.20	6.63	6.96	7.24	7.47	7.68	7.86	8.03	8.18	8.31	8.44	8.55	8.66	8.85	9.03
9	4.60	5.43	5.96	6.35	6.66	6.92	7.13	7.33	7.50	7.65	7.78	7.91	8.03	8.13	8.23	8.41	8.57
10	4.48	5.27	5.77	6.14	6.43	6.70	6.86	7.06	7.21	7.36	7.49	7.60	7.71	7.81	7.91	8.08	8.23
11	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99	7.13	7.25	7.36	7.47	7.56	7.65	7.81	7.95
12	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81	6.94	7.06	7.17	7.26	7.36	7.44	7.59	7.73
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67	6.79	6.90	7.01	7.10	7.19	7.27	7.42	7.55
14	4.21	4.90	5.32	5.63	5.88	6.09	6.26	6.41	6.54	6.66	6.77	6.87	6.92	7.05	7.13	7.27	7.40
15	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44	6.56	6.66	6.76	6.85	6.93	7.00	7.14	7.26
16	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35	6.46	6.56	6.66	6.74	6.82	6.90	7.03	7.15
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27	6.38	6.48	6.57	6.66	6.73	6.81	6.94	7.05
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20	6.31	6.41	6.50	6.58	6.66	6.73	6.85	6.97
19	4.05	4.67	5.05	5.33	5.55	5.74	5.89	6.02	6.14	6.25	6.34	6.43	6.51	6.59	6.65	6.78	6.89
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09	6.19	6.29	6.37	6.45	6.53	6.59	6.71	6.82
21	4.00	4.61	4.99	5.26	5.47	5.65	5.79	5.92	6.04	6.14	6.23	6.32	6.40	6.47	6.53	6.66	6.76
22	3.99	4.59	4.96	5.23	5.44	5.61	5.75	5.88	5.99	6.09	6.19	6.27	6.35	6.42	6.48	6.60	6.71
23	3.97	4.57	4.93	5.20	5.40	5.57	5.71	5.84	5.96	6.05	6.14	6.23	6.30	6.37	6.44	6.55	6.66
24	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92	6.02	6.11	6.19	6.26	6.33	6.39	6.51	6.61
25	3.94	4.53	4.88	5.14	5.35	5.51	5.66	5.78	5.89	5.98	6.07	6.15	6.22	6.29	6.36	6.47	6.57
26	3.93	4.51	4.87	5.12	5.32	5.49	5.63	5.75	5.86	5.95	6.04	6.12	6.19	6.26	6.32	6.43	6.53
27	3.92	4.50	4.85	5.10	5.30	5.46	5.60	5.72	5.83	5.92	6.01	6.09	6.16	6.23	6.29	6.40	6.50
28	3.91	4.48	4.83	5.08	5.28	5.44	5.58	5.70	5.80	5.90	5.98	6.06	6.13	6.20	6.26	6.37	6.47
29	3.90	4.47	4.81	5.06	5.26	5.42	5.56	5.67	5.78	5.87	5.96	6.03	6.08	6.17	6.23	6.34	6.44
30	3.89	4.46	4.80	5.05	5.24	5.40	5.54	5.65	5.76	5.85	5.93	6.01	6.08	6.14	6.20	6.31	6.41
40	3.83	4.37	4.70	4.93	5.11	5.27	5.40	5.50	5.60	5.69	5.76	5.84	5.90	5.96	6.02	6.12	6.21
60	3.76	4.28	4.60	4.82	4.99	5.13	5.25	5.36	5.45	5.53	5.60	5.67	5.73	5.79	5.84	5.93	6.02
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30	5.38	5.44	5.51	5.56	5.61	5.66	5.75	5.83

## Apéndice F1 . Tabla Duncan<sup>74</sup> al 5%

La primera columna corresponde a los grados de libertad del error; las demás columnas, al número de medias en comparación.

### Número de Medias, p

g.l.	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18	20
1	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0	18.0
2	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09	6.09
3	4.50	4.52	4.52	4.52	4.52	4.52	4.52	4.52	4.52	4.52	4.52	4.52	4.52	4.52	4.52	4.52	4.52
4	3.93	4.01	4.03	4.03	4.03	4.03	4.03	4.03	4.03	4.03	4.03	4.03	4.03	4.03	4.03	4.03	4.03
5	3.64	3.75	3.80	3.81	3.81	3.81	3.81	3.81	3.81	3.81	3.81	3.81	3.81	3.81	3.81	3.81	3.81
6	3.46	3.59	3.65	3.68	3.69	3.70	3.70	3.70	3.70	3.70	3.70	3.70	3.70	3.70	3.70	3.70	3.70
7	3.34	3.48	3.55	3.59	3.61	3.62	3.63	3.63	3.63	3.63	3.63	3.63	3.63	3.63	3.63	3.63	3.63
8	3.26	3.40	3.48	3.52	3.55	3.57	3.58	3.58	3.58	3.58	3.58	3.58	3.58	3.58	3.58	3.58	3.58
9	3.20	3.34	3.42	3.47	3.50	3.52	3.54	3.54	3.55	3.55	3.55	3.55	3.55	3.55	3.55	3.55	3.55
10	3.15	3.29	3.38	3.43	3.47	3.49	3.51	3.52	3.52	3.53	3.53	3.53	3.53	3.53	3.53	3.53	3.53
11	3.11	3.26	3.34	3.40	3.44	3.46	3.48	3.49	3.50	3.51	3.51	3.51	3.51	3.51	3.51	3.51	3.51
12	3.08	3.23	3.31	3.37	3.41	3.44	3.46	3.47	3.48	3.49	3.50	3.50	3.50	3.50	3.50	3.50	3.50
13	3.06	3.20	3.29	3.35	3.39	3.42	3.44	3.46	3.47	3.48	3.48	3.49	3.49	3.49	3.49	3.49	3.49
14	3.03	3.18	3.27	3.33	3.37	3.40	3.43	3.44	3.46	3.47	3.47	3.48	3.48	3.48	3.48	3.48	3.48
15	3.01	3.16	3.25	3.31	3.36	3.39	3.41	3.43	3.45	3.46	3.47	3.47	3.48	3.48	3.48	3.48	3.48
16	3.00	3.14	3.24	3.30	3.34	3.38	3.40	3.42	3.44	3.45	3.46	3.47	3.47	3.47	3.48	3.48	3.48
17	2.98	3.13	3.22	3.29	3.33	3.37	3.39	3.41	3.43	3.44	3.45	3.46	3.47	3.47	3.47	3.48	3.48
18	2.97	3.12	3.21	3.27	3.32	3.36	3.38	3.40	3.42	3.44	3.45	3.45	3.46	3.47	3.47	3.47	3.47
19	2.96	3.11	3.20	3.26	3.31	3.35	3.38	3.40	3.42	3.43	3.44	3.45	3.46	3.46	3.47	3.47	3.47
20	2.95	3.10	3.19	3.26	3.30	3.34	3.37	3.39	3.41	3.42	3.44	3.45	3.45	3.46	3.46	3.47	3.47
21	2.94	3.09	3.18	3.25	3.30	3.33	3.36	3.39	3.40	3.42	3.43	3.44	3.45	3.46	3.46	3.47	3.47
22	2.93	3.08	3.17	3.24	3.29	3.33	3.36	3.38	3.40	3.41	3.43	3.44	3.45	3.45	3.46	3.47	3.47
23	2.93	3.07	3.17	3.23	3.28	3.32	3.35	3.37	3.39	3.41	3.42	3.43	3.44	3.45	3.46	3.47	3.47
24	2.92	3.07	3.16	3.23	3.28	3.32	3.35	3.37	3.39	3.41	3.42	3.43	3.44	3.45	3.46	3.47	3.47
25	2.91	3.06	3.15	3.22	3.27	3.31	3.34	3.37	3.39	3.40	3.42	3.43	3.44	3.45	3.45	3.46	3.47
26	2.91	3.05	3.15	3.22	3.27	3.31	3.34	3.36	3.38	3.40	3.41	3.43	3.44	3.45	3.45	3.46	3.47
27	2.90	3.05	3.14	3.21	3.26	3.30	3.33	3.36	3.38	3.40	3.41	3.42	3.43	3.44	3.45	3.46	3.47
28	2.90	3.04	3.14	3.21	3.26	3.30	3.33	3.36	3.38	3.39	3.41	3.42	3.43	3.44	3.45	3.46	3.47
29	2.89	3.04	3.14	3.20	3.25	3.29	3.33	3.35	3.37	3.39	3.41	3.42	3.43	3.44	3.45	3.46	3.47
30	2.89	3.04	3.13	3.20	3.25	3.29	3.32	3.35	3.37	3.39	3.41	3.42	3.43	3.44	3.45	3.46	3.47
40	2.86	3.01	3.10	3.17	3.22	3.27	3.30	3.33	3.35	3.37	3.39	3.40	3.42	3.43	3.44	3.46	3.47
60	2.83	2.98	3.07	3.14	3.20	3.24	3.28	3.31	3.33	3.36	3.37	3.39	3.41	3.42	3.43	3.45	3.47
100	2.80	2.95	3.05	3.12	3.17	3.22	3.25	3.29	3.31	3.34	3.36	3.38	3.39	3.41	3.42	3.45	3.47

74 <https://www.york.ac.uk/depts/maths/tables/duncan>



## Apéndice F2. Tabla Duncan<sup>75</sup> al 1%

La primera columna corresponde a los grados de libertad del error; las demás columnas, al número de medias en comparación.

### Número de Medias, p

g.l.	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	18	20
1	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0	90.0
2	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0	14.0
3	8.26	8.32	8.32	8.32	8.32	8.32	8.32	8.32	8.32	8.32	8.32	8.32	8.32	8.32	8.32	8.32	8.32
4	6.51	6.68	6.74	6.76	6.76	6.76	6.76	6.76	6.76	6.76	6.76	6.76	6.76	6.76	6.76	6.76	6.76
5	5.70	5.89	5.99	6.04	6.07	6.07	6.07	6.07	6.07	6.07	6.07	6.07	6.07	6.07	6.07	6.07	6.07
6	5.24	5.44	5.55	5.61	5.66	5.68	5.69	5.70	5.70	5.70	5.70	5.70	5.70	5.70	5.70	5.70	5.70
7	4.95	5.15	5.26	5.33	5.38	5.42	5.44	5.45	5.46	5.47	5.47	5.47	5.47	5.47	5.47	5.47	5.47
8	4.75	4.94	5.06	5.13	5.19	5.23	5.26	5.28	5.29	5.30	5.31	5.31	5.32	5.32	5.32	5.32	5.32
9	4.60	4.79	4.91	4.99	5.04	5.09	5.12	5.14	5.16	5.17	5.19	5.19	5.20	5.20	5.21	5.21	5.21
10	4.48	4.67	4.79	4.87	4.93	4.98	5.01	5.04	5.06	5.07	5.09	5.10	5.11	5.11	5.12	5.12	5.12
11	4.39	4.58	4.70	4.78	4.84	4.89	4.92	4.95	4.98	4.99	5.01	5.02	5.03	5.04	5.05	5.05	5.06
12	4.32	4.50	4.62	4.71	4.77	4.82	4.85	4.88	4.91	4.93	4.94	4.96	4.97	4.98	4.99	5.00	5.01
13	4.26	4.44	4.56	4.64	4.71	4.75	4.79	4.82	4.85	4.87	4.89	4.90	4.92	4.93	4.94	4.95	4.96
14	4.21	4.39	4.51	4.59	4.65	4.70	4.74	4.78	4.80	4.82	4.84	4.86	4.87	4.88	4.89	4.91	4.92
15	4.17	4.35	4.46	4.55	4.61	4.66	4.70	4.73	4.76	4.78	4.80	4.82	4.83	4.85	4.86	4.87	4.89
16	4.13	4.31	4.43	4.51	4.57	4.62	4.66	4.70	4.72	4.75	4.77	4.79	4.80	4.81	4.83	4.84	4.86
17	4.10	4.28	4.39	4.47	4.54	4.59	4.63	4.66	4.69	4.72	4.74	4.76	4.77	4.79	4.80	4.82	4.83
18	4.07	4.25	4.36	4.45	4.51	4.56	4.60	4.64	4.66	4.69	4.71	4.73	4.75	4.76	4.77	4.79	4.81
19	4.05	4.22	4.34	4.42	4.48	4.53	4.58	4.61	4.64	4.66	4.69	4.71	4.72	4.74	4.75	4.77	4.79
20	4.02	4.20	4.31	4.40	4.46	4.51	4.55	4.59	4.62	4.64	4.66	4.68	4.70	4.72	4.73	4.75	4.77
21	4.00	4.18	4.29	4.37	4.44	4.49	4.53	4.57	4.60	4.62	4.65	4.66	4.68	4.70	4.71	4.73	4.75
22	3.99	4.16	4.27	4.36	4.42	4.47	4.51	4.55	4.58	4.60	4.63	4.65	4.66	4.68	4.69	4.72	4.74
23	3.97	4.14	4.25	4.34	4.40	4.45	4.50	4.53	4.56	4.59	4.61	4.63	4.65	4.67	4.68	4.70	4.72
24	3.96	4.13	4.24	4.32	4.39	4.44	4.48	4.52	4.55	4.57	4.60	4.62	4.63	4.65	4.67	4.69	4.71
25	3.94	4.11	4.22	4.31	4.37	4.42	4.47	4.50	4.53	4.56	4.58	4.60	4.62	4.64	4.65	4.68	4.70
26	3.93	4.10	4.21	4.29	4.36	4.41	4.45	4.49	4.52	4.55	4.57	4.59	4.61	4.63	4.64	4.67	4.69
27	3.92	4.09	4.20	4.28	4.35	4.40	4.44	4.48	4.51	4.54	4.56	4.58	4.60	4.62	4.63	4.66	4.68
28	3.91	4.08	4.19	4.27	4.33	4.39	4.43	4.47	4.50	4.52	4.55	4.57	4.59	4.60	4.62	4.65	4.67
29	3.90	4.07	4.18	4.26	4.32	4.38	4.42	4.46	4.49	4.51	4.54	4.56	4.58	4.60	4.61	4.64	4.66
30	3.89	4.06	4.17	4.25	4.31	4.37	4.41	4.45	4.48	4.50	4.53	4.55	4.57	4.59	4.60	4.63	4.65
40	3.83	3.99	4.10	4.18	4.24	4.30	4.34	4.38	4.41	4.44	4.46	4.48	4.50	4.52	4.54	4.57	4.59
60	3.76	3.92	4.03	4.11	4.17	4.23	4.27	4.31	4.34	4.37	4.39	4.42	4.44	4.46	4.47	4.50	4.53
100	3.70	3.86	3.96	4.04	4.11	4.16	4.20	4.24	4.27	4.30	4.33	4.35	4.37	4.39	4.41	4.44	4.47

Un componente esencial para el desarrollo de los países ha sido la investigación científica. Con este libro, que resume las experiencias de 35 años, el autor pretende contribuir al mejoramiento de los procesos de investigación en las ciencias biológicas. Se escribió con el propósito de que sirva como un texto de enseñanza y una herramienta para la planificación y ejecución de experimentos. Se caracteriza por mantener un balance adecuado entre la teoría y la práctica, lo que consigue el autor a través de ejercicios que explican detalladamente las razones del experimento, los pasos del análisis y la interpretación de los resultados. Además de las técnicas más conocidas, el autor incluye otros análisis que no son comunes en otros textos. Entre ellos, lo relacionado con los experimentos factoriales fraccionados, varias pruebas no paramétricas y los procedimientos a seguir cuando se violan las presunciones del análisis de varianza. El autor aspira a que este libro, por su contenido, se convierta en un material de frecuente consulta para estudiantes, investigadores y técnicos de campo involucrados en la investigación.

#### Leonardo Corral Dávalos

Nació en Riobamba, Ecuador. Tiene un B.Sc en Ciencias Agrícolas y un M.Sc en Agronomía de la Universidad Estatal de Ohio. La Universidad Estatal de Oregon le otorgó un PhD en Genética Vegetal y Estadística. Ha sido profesor de varias universidades en cursos de pregrado y postgrado. Ha dirigido decenas de tesis y proyectos de investigación. Ha sido consultor de programas nacionales e internacionales. Actualmente brinda asesoramiento a empresas agrícolas y ofrece cursos cortos de capacitación. Reside en Guayaquil, Ecuador, y está dedicado a la lectura de la historia de la biología, especialmente en relación con las consecuencias de mezclar la ciencia con la ideología. Espera publicar un libro sobre el tema. [leocorral2001@yahoo.com](mailto:leocorral2001@yahoo.com)

