



Universidad de Valladolid

Facultad de Ciencias

TRABAJO FIN DE GRADO

Grado en Estadística

ANOVA con Medidas Repetidas

Autor: Ignacio Fernández Iniesta

Tutor/es: Lourdes Barba Escribá

CONTENIDOS

1. RESUMEN.....	9
2. ABSTRACT	9
3. ANÁLISIS CON MEDIDAS REPETIDAS.....	10
3.1. Introducción	10
3.2. Ventajas	10
3.3. Inconvenientes:	10
3.3.1. Diseño de contrapeso (<i>counterbalancing</i>):	11
3.4. Aplicaciones:.....	13
4. ANOVA CON MEDIDAS REPETIDAS EN UN FACTOR	14
4.1. Modelo básico	14
4.2. El estadístico de contraste.....	16
4.2.1. El estadístico F	17
4.3. Supuestos	18
4.3.1. Independencia	18
4.3.2. Normalidad	18
4.3.3. Aditividad o Esfericidad	18
4.4. Medidas del tamaño del efecto.....	24
4.5. Ejemplo práctico en SAS	24
A) PROC GLM	26
B) PROC MIXED.....	30
5. ANOVA CON MEDIDAS REPETIDAS DE DOS FACTORES	35
5.1. Medidas repetidas en un solo factor.....	35
5.1.1. Modelo	36
5.1.2. Supuestos	37
5.1.3. Estadísticos de contraste.....	38
5.1.4. Ejemplo práctico en SAS	42
5.2. Medidas repetidas en los dos factores.....	47
5.2.1. Modelo	47
5.2.2. Supuestos	49
5.2.3. Estadísticos de contraste.....	49
5.2.4. Ejemplo práctico en SAS	52
6. Otros ejemplos en SAS.....	58
Ejemplo 4:.....	58
A) PROC GLM	60

B) PROC MIXED	61
Ejemplo 5:	63
A) PROC GLM	64
B) PROC MIXED	66
7. Conclusiones	68
8. Posibles ampliaciones.....	68
9. Glosario de vocabulario inglés-español	70
10. Anexo	71
11. Bibliografía y enlaces	76

INDICE DE TABLAS

Tabla 1. Ejemplo diseño de contrapeso (R.Girden, Ellen [1])	12
Tabla 2. Resumen del ANOVA en medidas repetidas de un factor	16
Tabla 3. Respuesta afectiva de 20 sujetos hacia un perfume en cuatro grados de concentración.	25
Tabla 4. Medias cuadráticas y valores esperados en el modelo de dos factores con medidas repetidas en uno solo.	40
Tabla 5. Tabla resumen del ANOVA de dos factores mixto.....	41
Tabla 6. Tabla de datos para el ejemplo de dos factores con medidas repetidas en un solo factor.	42
Tabla 7. Medidas cuadráticas y valores esperados en el modelo no aditivo de dos factores con medidas repetidas en ambos. (Ximénez, Carmen y San Martín, Rafael [2]).....	51
Tabla 8. Valores esperados de las medias cuadráticas del modelo de dos factores con medidas repetidas en ambos (modelo no aditivo) (Ximénez, Carmen y San Martín, Rafael [2]).	52
Tabla 9. Respuesta afectiva de 10 sujetos hacia un perfume en dos grados de concentración y tres tipos de aroma.	53
Tabla 10. Datos para el Ejemplo 4.	58
Tabla 12. Datos para el ejemplo 5 (completos).....	63
Tabla 13. Datos para el ejemplo 5 (missing).....	64

INDICE DE FIGURAS

Figura 1. Procedimiento a seguir cuando usamos PROC GLM en medidas repetidas (Wolfinger y Chang [13])	27
Figura 2. Código para la lectura de datos de el ejemplo de medidas repetidas en un factor	28
Figura 3. Código para analizar el Ejemplo 1 utilizando PROC GLM.	28
Figura 4. Test de esfericidad usando PROC GLM para el Ejemplo 1.	29
Figura 5. Test multivariante para los efectos intrasujetos.	29
Figura 6. Test F univariantes para los efectos intrasujeto de DOSIS.	30
Figura 7. Procedimiento a seguir cuando usamos PROC MIXED en medidas repetidas (Wolfinger y Chang [13]).	31
Figura 8. Código para analizar el Ejemplo 1 utilizando PROC MIXED.	32
Figura 9. PROC MIXED con estructura de covarianza "simetría compuesta".	32
Figura 10. PROC MIXED con estructura de covarianza "Tipo H o Huynh-Feldt"	32
Figura 11. PROC MIXED con estructura de covarianza "sin estructura".	33
Figura 12. Estadísticos de ajuste para estructura de covarianza "simetría compuesta"	33
Figura 13. Estadísticos de ajuste para estructura de covarianza "Tipo H"	33
Figura 14. Estadísticos de ajuste para estructura de covarianza "No estructurado"	34
Figura 15. Test de tipo 3 para contrastar la hipótesis nula del modelo.	34
Figura 16. Lectura de datos para el Ejemplo 2.	43
Figura 17. Código para analizar el Ejemplo 2 utilizando PROC GLM.	43
Figura 18. Test de esfericidad para nuestro ejemplo usando PROC GLM.	44
Figura 19. Test F univariante para los efectos intersujeto.	44
Figura 20. Test F univariante para los efectos intrasujeto.	44
Figura 21. PROC MIXED con estructura de covarianza "simetría compuesta".	45
Figura 22. PROC MIXED con estructura de covarianza "Tipo H".	45
Figura 23. PROC MIXED con estructura de covarianza "sin estructura".	45
Figura 24. Estadísticos de ajuste para la estructura de covarianza "simetría compuesta".	45
Figura 25. Estadísticos de ajuste para la estructura de covarianza "Tipo H".	46
Figura 26. Estadísticos de ajuste para la estructura de covarianza "No estructurada".	46
Figura 27. Test de tipo III para contrastar las hipótesis nulas del modelo.	46
Figura 28. Lectura de datos para el Ejemplo 3.	53
Figura 29. Código para analizar el Ejemplo 3 con PROC GLM	53
Figura 30. Test de esfericidad para la variable FLOR.	54
Figura 31. Test de esfericidad para la interacción de FLOR x DOSIS	54
Figura 32. Test F univariantes para los efectos intrasujetos DOSIS y FLOR y su interacción	55
Figura 33. Lectura de datos para Ejemplo 3 con PROC MIXED.	56
Figura 34. PROC MIXED con estructura de covarianza "simetría compuesta"	56
Figura 35. PROC MIXED con estructura de covarianza "Tipo H"	56
Figura 36. PROC MIXED con estructura de covarianza "No estructurada"	56
Figura 37. Estadísticos de ajuste para la estructura de covarianza "simetría compuesta"	57
Figura 38. Estadísticos de ajuste para la estructura de covarianza "Tipo H"	57
Figura 39. Estadísticos de ajuste para la estructura de covarianza "No estructurada"	57
Figura 40. Test de tipo III para contrastar las hipótesis nulas del modelo	57
Figura 41. Lectura de datos para Ejemplo 4.	59
Figura 42. Código para analizar el Ejemplo 4 utilizando PROC GLM.	60
Figura 43. Test de esfericidad para el Ejemplo 4.	60
Figura 44. Test F univariante para los efectos intrasujetos.	60
Figura 45. Test F univariante para los efectos intersujetos.	61
Figura 46. PROC MIXED con estructura de covarianza "simetría compuesta"	61
Figura 47. PROC MIXED con estructura de covarianza "Tipo H".	61
Figura 48. PROC MIXED con estructura de covarianza "sin estructura".	61
Figura 49. Estadísticos de ajuste para la estructura de covarianza "simetría compuesta"	62
Figura 50. Estadísticos de ajuste para la estructura de covarianza "Tipo H".	62
Figura 51. Estadísticos de ajuste para la estructura de covarianza "sin estructura".	62
Figura 52. Test de tipo III para contrastar la hipótesis nula del modelo.	62
Figura 53. Test de esfericidad para el Ejemplo 5 (completos).	63

Figura 54. Test para contrastar la hipótesis nula del modelo.	64
Figura 55. Lectura de datos para el Ejemplo 5 (missing) usando PROC GLM.....	65
Figura 56. Código para analizar el Ejemplo 5 (missing) utilizando PROC GLM.....	65
Figura 57. Test de esfericidad para el Ejemplo 5 (missing).	65
Figura 58. Test F univariante para los efectos intrasujetos.....	65
Figura 59. Lectura de datos para el Ejemplo 5 (missing) usando PROC MIXED.	66
Figura 60. PROC MIXED con estructura de covarianza "simetría compuesta".....	66
Figura 61. PROC MIXED con estructura de covarianza "Tipo H".	66
Figura 62. PROC MIXED con estructura de covarianza "sin estructura".	67
Figura 63. Estadísticos de ajuste para la estructura de covarianza "simetría compuesta".....	67
Figura 64. Estadísticos de ajuste para la estructura de covarianza "Tipo H".	67
Figura 65. Estadísticos de ajuste para la estructura de covarianza "sin estructura".	67
Figura 66. Test de tipo III para contrastar la hipótesis nula del modelo.....	68

ANEXO

Anexo 1. Datos para Ejemplo 1. PROC GLM	71
Anexo 2. Datos para Ejemplo 1. PROC MIXED.....	72
Anexo 3. Datos para Ejemplo 2. PROC GLM	73
Anexo 4. Datos para Ejemplo 2. PROC MIXED.....	74
Anexo 5. Datos para Ejemplo 3. PROC GLM	75
Anexo 6. Datos para Ejemplo 3. PROC MIXED.....	75

1. RESUMEN

El análisis en medidas repetidas es una alternativa al modelo clásico de diseño cada vez más utilizada sobre todo en estudios relacionados con las ciencias sociales como por ejemplo en psicología o sociología. La principal ventaja se refiere a que un mismo individuo puede ser sometido a diferentes tratamientos o condiciones lo que hace que economizamos el número de individuos que participarán en el experimento. Pero no todo son ventajas, la fatiga, la posibilidad de tener datos no balanceados o el efecto residual son algunos de los inconvenientes que habrá que tener en cuenta al utilizar este tipo de diseño. En este trabajo se expondrán las ventajas, inconvenientes y supuestos que se deben cumplir al realizar un análisis con medidas repetidas.

Aparte se explicarán los diferentes modelos que podemos tener dependiendo de la estructura de nuestro experimento y para cada uno de estos se expondrá un ejemplo resuelto con SAS.

Palabras clave: medidas repetidas, ANOVA, SAS, PROC GLM, PROC MIXED, REPEATED

2. ABSTRACT

The repeated measures analysis is an alternative to the classic design model increasingly used specially in studies related to social sciences such as in psychology or sociology. The main advantage is referred to the fact that the same person can be subjected to different treatments or conditions what leads into an economitiation of the number of individuals that need to participate in the experiment. However not everything is an advantage, the fatigue, the possibility to have unbalanced data or the carry-over effect are some of the inconvenients that we will have to take into account when using this type of design. In this project the advantages, disadvantages and the assumptions that must be met when performing a repeated measures analysis will be adressed.

Also, we will explain the different types of models that we can have depending on the structure of our experiment and each of them will be presented with an example solved using SAS.

Key words: repeated measures, ANOVA, SAS, PROC GLM, PROC MIXED, REPEATED

3. ANÁLISIS CON MEDIDAS REPETIDAS

3.1. *Introducción*

El término “medidas repetidas” se refiere generalmente a los datos en los que la respuesta de cada unidad experimental o sujeto se observa en múltiples ocasiones o bajo múltiples condiciones. Otros términos utilizados para referirse a datos de medidas repetidas son: “datos longitudinales”, utilizado por algunos autores cuando el factor de las medidas repetidas es el tiempo (podríamos verlo como un caso especial de datos de medidas repetidas) o “diseño de bloques aleatorizados” designando cada individuo como “bloque”. La comparación de los tratamientos es intrasujeto.

En este trabajo utilizaremos el término “medidas repetidas” en su significado más amplio para referirnos a la situación en la que múltiples medidas de la variable respuesta son obtenidas de cada unidad experimental.

3.2. *Ventajas*

Una gran ventaja de los estudios resultantes de usar medidas repetidas es que es el único tipo de diseño en el que es posible obtener información sobre posibles patrones de cambio individuales. Este diseño también requiere un número menor de sujetos. Por ejemplo, cuando estamos estudiando el efecto de un tratamiento a lo largo del tiempo, se prefiere normalmente usar los mismos sujetos repetidamente antes que usar diferentes sujetos en cada periodo de tiempo.

Otra ventaja se refiere a que en este tipo de modelos la equivalencia previa de los sujetos sobre los que se toman las medidas no constituye un problema. El uso del procedimiento de medidas repetidas proporciona un control más efectivo de las fuentes de variación extrañas asociadas, por lo general, a las características individuales; es decir, se consigue una reducción de la varianza del error.

3.3. *Inconvenientes:*

Los diseños que usan medidas repetidas también presentan inconvenientes que han de considerarse para evitar su uso incorrecto. Uno de ellos se refiere al orden o secuencia de administración de los tratamientos, los cuales pueden sesgar los resultados. Otra desventaja es la posible aparición de fatiga si los participantes necesitan realizar múltiples tests.

Muchas veces, el investigador no puede controlar las circunstancias en las que se obtiene la medida por lo que los datos pueden estar no balanceados o parcialmente incompletos.

Otro problema es el llamado *carry-over effect* o efecto residual, se caracteriza por la persistencia de la acción de un tratamiento más allá del período o tiempo de aplicación. Representa la acumulación progresiva tanto de los efectos facilitadores de la respuesta (efecto de la práctica, aprendizaje, etc.) como de los efectos obstaculizadores (como la fatiga mental, cansancio físico, etc.). Por ejemplo, consideramos un experimento para probar el efecto que tiene la tasa de presentación en la memoria. A los sujetos se les presenta una lista de palabras y se les pide memorizar tantas como les sea posible. En la primera condición, las palabras se presentan una por segundo; en la otra condición, dos por segundo. La pregunta sería si el haber participado en la primera condición tiene efecto en los resultados bajo la segunda. A lo mejor aprenderse la primera lista de palabras interfiere en el aprendizaje de la segunda ya que será más complicado recordar qué palabras pertenecían a cada lista. O a lo mejor la práctica que se adquiere al aprender la primera lista hace más fácil aprender la segunda. En cualquiera de los casos, habría efecto *carry-over* ya que los resultados del experimento bajo la segunda condición se verían afectados por la experiencia adquirida bajo la primera condición. El efecto *carry-over* puede minimizarse prolongando los tiempos entre tratamientos.

Por último, los diseños intrasujetos presentan el inconveniente de la violación de algunos de los supuestos habituales de naturaleza estadística. Puesto que se emplean los mismos sujetos para todas las condiciones experimentales es probable que se produzca un efecto sistemático en las respuestas que de lugar a la aparición de correlación o dependencia entre los errores, lo que repercutiría gravemente en los resultados obtenidos ya que el estadístico F no es robusto ante observaciones correlacionadas (O'Brien y Kaiser, 1985).

Entre los métodos propuestos para neutralizar dichos efectos se encuentran "las técnicas de contrabalanceo" (Underwood y Shaughnessy, 1978). Se trata de equilibrar el orden de administración de los tratamientos o asignar aleatoriamente a cada sujeto la aplicación de los mismos. Con este procedimiento se controla otro de los inconvenientes de los diseños de medidas repetidas: los efectos de la práctica, del arrastre (o *carry-over*) y de la fatiga. Si solamente se sigue un único orden, la mejora en las puntuaciones puede atribuirse al tratamiento o al efecto de la práctica; y el empeoramiento al tratamiento o la fatiga.

Aunque se han estudiado muchos enfoques diferentes para el análisis de medidas repetidas, la mayoría están restringidos por el marco en el que la variable respuesta sigue una normal y los datos están balanceados y completos.

3.3.1. Diseño de contrapeso (*counterbalancing*):

En un experimento, el orden en que se realizan los tratamientos puede afectar de forma significativa en el comportamiento de los sujetos pudiendo provocar una respuesta falsa, por factores de fatiga, entrenamiento o externos que pueden

influir en el comportamiento de los sujetos. Por ello, los investigadores suelen utilizar un diseño de contrapeso que hace posible que el orden del tratamiento u otros factores no afecte negativamente a los resultados.

El *counterbalancing* es una manera de presentar los diferentes niveles del tratamiento de tal manera que cada uno (a) ocurre con la misma frecuencia en cada etapa de la práctica y (b) precede tantas veces como sigue a cada nivel. El cometido es contrarrestar la fatiga, entrenamiento y los efectos *carry-over*. Con un número par de niveles y un número de individuos múltiplo del número de niveles, se puede conseguir que estos dos requerimientos se cumplan aplicando las siguientes pautas:

1, 2, n, 3, n – 1, 4, n – 2, etc.

Cada número se refiere a un nivel del tratamiento. Si tenemos que testar cuatro niveles, entonces las pautas se reducen a que el primer orden de los niveles sea 1, 2, 4, 3. Cada orden que lo precede se deriva sumando “1” a los números del orden precedente. El orden dos sería 2 (1 + 1), 3 (2 + 1), 1 (4 + 1 no se aplica), 4 (3 + 1). Si los niveles corresponden a A(1), B(2), C(3), y D(4), entonces siguiendo las pautas, una secuencia de orden adecuada sería la presentada en la Tabla 1.

Tabla 1. Ejemplo diseño de contrapeso (R.Girden, Ellen [1])

	Etapas de la práctica			
Paloma	Primera	Segunda	Tercera	Cuarta
1	A(1)	B(2)	D(4)	C(3)
2	B(1+1 = 2)	C(2+1 = 3)	A(4+1 no se aplica)	D(3+1 = 4)
3	C(2 + 1 = 3)	D(3+1 = 4)	B(1+1 = 2)	A(4+1 no se aplica)
4	D(3+1 = 4)	A(4+1 no se aplica)	C(2+1 = 3)	B(1+1 = 2)

Cada nivel (A,B,C,D) ocurre una vez en cada etapa de la práctica, y cada nivel precede y sigue a cada otro nivel el mismo número de veces. Por ejemplo, A precede a B una vez (la primera paloma) y sigue a B una vez (para la tercera paloma). A precede a C una vez (para la cuarta paloma) y sigue a C una vez (para la segunda paloma). Finalmente, A precede a D una vez (para la segunda paloma) y sigue a D una vez (para la cuarta paloma). Cuando las respuestas bajo todos los niveles representados por A, B, C y D se juntan, las diferencias de medias no deberían de ser debido a la práctica, fatiga o *carry-over*. Con un número impar de

niveles, los requerimientos se pueden satisfacer usando en primer lugar el procedimiento que se acaba de describir para derivar un conjunto de órdenes y después construyendo otro conjunto dándole la vuelta a la primera fila (i.e. 3,4,2,1) y repitiendo el procedimiento.

3.4. Aplicaciones:

Las medidas repetidas pueden emplearse con diferentes objetivos y funciones. Según Maxwell y Delaney (1990) se resumen en tres:

- 1. Evaluación longitudinal del cambio a lo largo del tiempo
- 2. Evaluación de la actuación de los sujetos bajo diferentes condiciones experimentales en estudios transversales (investigación observacional que analiza datos de variables recopiladas en un periodo de tiempo concreto sobre una muestra de una población) cuando se cuenta con muestras reducidas de sujetos.
- 3. Comparación de las puntuaciones obtenidas por los mismos sujetos en diferentes pruebas psicométricas o similares.

Los diseños de medidas repetidas se han aplicado en casi todas las áreas de investigación de las ciencias sociales y del comportamiento. Por ejemplo, en psicología, medicina, educación, sociología, ciencias políticas, economía y empresa. A continuación, se muestran varios ejemplos representativos:

1.- Cohen y Riel (1989) demostraron que la calidad de la escritura a mano de unos alumnos era mediocre cuando sabían que un profesor les iba a evaluar y asignar una nota. Sin embargo, mejoraba cuando el público al que iba dirigido era diferente. Una clase de 1º de ESO tuvo que escribir una redacción para la evaluación de su primer semestre y una semana más tarde tuvieron que escribir sobre el mismo tema pero a alumnos de otro país. Otra clase tuvo que realizar las mismas tareas pero en orden inverso. Los datos se analizaron mediante un ANOVA de un factor con medidas repetidas.

2.- Los estudiantes de la escuela secundaria deberían ser más capaces de aplicar lo que han estudiado si se les ha enseñado de manera aplicada que si se les ha mandado estudiar de memoria. A un grupo de estudiantes con problemas para aprender se les dio un curso durante 12 días consecutivos sobre cómo mejorar los factores de salud incrementa la longevidad. Para la primera parte de cada día, todos los estudiantes tuvieron clases estructuradas. Para la segunda parte, la mitad de los estudiantes siguieron revisando los contenidos de manera estructurada y utilizando técnicas de aplicación convencionales (ej. Mantener un seguimiento de su dieta). La otra mitad trabajó con ejercicios de simulación con ordenador analizando problemas de salud y cambiándolos para aliviar los problemas. A todos los estudiantes se les examinó un día después de las clases y otra vez 2 semanas

más tarde para ver cuánto habían retenido. Los datos fueron analizados con ANOVA de dos vías con medidas repetidas en un factor (Woodward, Carnine & Gersten, 1988)

3.- Cuando estudiantes discapacitados y no discapacitados van a la misma escuela, la interacción entre compañeros resulta complicada. Se compararon dos programas para facilitar una mayor integración. Alumnos no discapacitados de cuarto, quinto y sexto de primaria fueron emparejados con alumnos discapacitados y asistieron a tutoría de compañeros o sesiones de interacción de “amigos especiales”. Las interacciones entre parejas se calificaron y compararon durante el período de 8 semanas (Cole, Vandercook & Rynders, 1988)

4. ANOVA CON MEDIDAS REPETIDAS EN UN FACTOR

4.1. Modelo básico

Nos encontramos ante una situación experimental en la que se extrae una muestra aleatoria simple de n sujetos a los que se les aplica a cada uno J tratamientos establecidos por el investigador (efectos fijos). Es decir, todos y cada uno de los sujetos pasan por todos y cada uno de los niveles del factor y la muestra total (N) está compuesta por nJ observaciones.

El modelo de ANOVA de un factor con medidas repetidas en una versión del modelo lineal general adaptado al caso en el que solamente hay una variable independiente o un factor. Tal modelo describe la variable dependiente Y_{ij} a partir de las siguientes cuatro componentes:

Modelo:

$$Y_{ij} = \mu + \pi_i + \tau_j + e_{ij} \quad (4.1)$$

Para cada $i = 1, \dots, n$ y $j = 1, \dots, J$, en la ecuación (4.1), Y_{ij} es la respuesta del sujeto i en el j -ésimo nivel del factor o tiempo j , μ es la media general, π_i es el efecto aleatorio asociado al sujeto el cuál es constante bajo todas las repeticiones, τ_j es el efecto fijo en el j -ésimo nivel del factor, y e_{ij} es el error experimental asociado al sujeto i bajo el tratamiento o tiempo j

Se asume que los efectos aleatorios π_i son normales $N(0, \sigma\pi^2)$ independientes, los errores aleatorios e_{ij} son normales $N(0, \sigma e^2)$ independientes y los efectos aleatorios π_i y los errores aleatorios e_{ij} son independientes. Se asume que los efectos fijos τ_j satisfacen la restricción de que la suma de estos sea igual a 0.

$$\sum_{j=1}^t \tau_j = 0.$$

Las varianzas y covarianzas de las observaciones son:

$$\begin{aligned} \text{Var}(Y_{ij}) &= \text{Var}(\mu + \pi_i + \tau_j + e_{ij}) = \sigma_\pi^2 + \sigma_e^2 \\ \text{Cov}(Y_{ij}, Y_{i'j}) &= 0, \text{ para } i \neq i' \\ \text{Cov}(Y_{ij}, Y_{ij'}) &= \sigma_\pi^2, \text{ para } j \neq j' \end{aligned}$$

Por lo tanto, la matriz de covarianzas del vector $Y_i = (Y_{i1}, \dots, Y_{it})'$ viene dado por

$$\Sigma = \begin{pmatrix} \sigma_\pi^2 + \sigma_e^2 & & \sigma_\pi^2 \\ & \ddots & \\ \sigma_\pi^2 & & \sigma_\pi^2 + \sigma_e^2 \end{pmatrix} = (\sigma_\pi^2 + \sigma_e^2) \begin{pmatrix} 1 & & \rho \\ & \ddots & \\ \rho & & 1 \end{pmatrix} = \sigma_y^2 \begin{pmatrix} 1 & & \rho \\ & \ddots & \\ \rho & & 1 \end{pmatrix} \quad (4.2)$$

Donde

$$\rho = \frac{\sigma_\pi^2}{\sigma_\pi^2 + \sigma_e^2} = \text{Corr}(Y_{ij}, Y_{ij'}).$$

Aunque todas las variables aleatorias del modelo (4.1) son independientes, las medidas repetidas de un sujeto están correladas. Se dice que la matriz de covarianza resultante dada por la ecuación (4.2) con elementos diagonales iguales y elementos fuera de la diagonal iguales tiene simetría compuesta. Esta estructura de covarianza implica que la correlación entre cualquier par de observaciones repetidas es la misma, independientemente del espaciado entre las observaciones. Esta suposición es altamente restrictiva y, a menudo, poco realista, especialmente cuando el factor de mediciones repetidas es el tiempo.

Se deduce que el efecto del tratamiento (τ_j) es constante para cada sujeto, por lo tanto, se añade al efecto de los sujetos (π_i), de ahí que se denomine *modelo aditivo*.

Otro modelo alternativo plantea que ya que los individuos son expuestos en varios tratamientos, los efectos de estos no se pueden considerar constantes. Este modelo plantea lo mismo que el aditivo pero añadiendo un efecto adicional que consiste en la interacción entre el efecto fijo del factor τ_j y los efectos de los sujetos π_i . Es decir:

$$Y_{ij} = \mu + \tau_j + \pi_i + (\tau\pi)_{ji} + e_{ij}$$

Se denomina modelo no aditivo y todos los términos son idénticos al aditivo salvo $(\tau\pi)_{ji}$ que representa la variabilidad que queda después de eliminar los efectos separados τ_j y π_i , puesto que $(\tau\pi)_{ji}$ y e_{ij} no pueden calcularse separadamente. El modelo no aditivo suele utilizarse con menor frecuencia.

T_{+j} y T_{i+} son los totales correspondientes a cada tratamiento y a cada sujeto respectivamente. T_{+j} representa la suma de las puntuaciones obtenidas por todos los sujetos de la muestra en el tratamiento j . Del mismo modo T_{i+} representa la suma de las puntuaciones del sujeto i en todos los tratamientos o niveles de j . La forma de obtener estos totales es la siguiente

$$T_{+j} = \sum_i Y_{ij} \quad T_{i+} = \sum_j Y_{ij}$$

Finalmente, la media total ($\bar{Y}_{..}$) y la suma de puntuaciones (T o T_{++}) se obtienen mediante:

$$T = \sum_i \sum_j Y_{ij} \quad \bar{Y}_{..} = \frac{T}{N}$$

Tabla 2. Resumen del ANOVA en medidas repetidas de un factor

Fuentes de variación	SC	g.l.	MC	EMC	F
Intrasujetos (I)	$SCI = \sum_j \frac{T_{+j}^2}{n} - \frac{T^2}{N}$	$J-1$	$MCI = \frac{SCI}{J-1}$	$\sigma_e^2 + \frac{n \sum_j \tau_j^2}{J-1}$	$F = \frac{MCI}{MCE}$
Sujetos (P)	$SCP = \sum_i \frac{T_{i+}^2}{J} - \frac{T^2}{N}$	$n-1$	$MCP = \frac{SCP}{n-1}$	$\sigma_e^2 + J\sigma_\pi^2$	
Error (e)	$SCE = SCT - SCI - SCP$	$(J-1)(n-1)$	$MCE = \frac{SCE}{(J-1)(n-1)}$	σ_e^2	
Total (T)	$SCT = \sum_i \sum_j Y_{ij}^2 - \frac{T^2}{N}$	$N-1$			

4.2. El estadístico de contraste

El ANOVA de medidas repetidas en un solo factor pone a prueba la hipótesis sobre la igualdad de medias en los distintos J tratamientos o niveles de la variable independiente; lo que expresado formalmente es:

Siendo $\mu_{+j} = \mu + \tau_j$

$H_0: \mu_{+1} = \mu_{+2} = \dots = \mu_{+j}$; o bien $H_0: \sum \tau_j^2 = 0$

$H_1: \mu_{+j} \neq \mu_{+j'}$, para algún valor de j o de j' ($j \neq j'$); o bien $H_1: \sum \tau_j^2 \neq 0$

4.2.1. El estadístico F

El estadístico empleado para contrastar la hipótesis nula en el modelo de ANOVA de un factor con medidas repetidas es:

$$F = \frac{MCI}{MCE}$$

Mientras los supuestos del modelo se sostengan, este estadístico test sigue la distribución $F_{J-1,(n-1)(J-1)}$ si la hipótesis nula es verdad. Si H_0 es verdadera, dicho cociente se aproxima a 1. Sin embargo, si H_0 no es verdad, el numerador tenderá a ser más grande que el denominador. Ahora bien, incluso siendo H_0 verdadera, teniendo en cuenta que MCI y MCE son valores muestrales cabe esperar que calculadas en un conjunto de datos concretos, existan diferencias entre ellas. La cuestión radica en determinar cuán grandes han de ser dichas diferencias para pensar que H_0 no es verdadera. Para responder esta pregunta es necesario conocer la distribución muestral asociada al cociente MCI/MCE

Una variable del tipo $\frac{(n-1)S_{n-1}^2}{\sigma^2}$ se distribuye según una X^2 con $n-1$ grados de libertad.

Del mismo modo $\frac{(J-1)MCI}{\sigma^2}$ sigue una X^2 con $J-1$ grados de libertad bajo H_0 y $\frac{(J-1)(n-1)MCE}{\sigma^2}$ una X^2 con $(J-1)(n-1)$ grados de libertad.

El cociente entre dos X^2 independientes, cada una de ellas dividida por sus grados de libertad, es por definición una variable aleatoria distribuida según el modelo de probabilidad F con los grados de libertad del numerador y del denominador.

$$\frac{X^2/n_1}{X^2/n_2} \equiv F_{n_1, n_2}$$

Así pues:

$$\frac{\frac{(J-1)MCI}{\sigma^2} / (J-1)}{\frac{(J-1)(n-1)MCE}{\sigma^2} / (J-1)(n-1)} = \frac{MCI}{MCE} \equiv F_{J-1, (J-1)(n-1)}$$

F es un estadístico adecuado para contrastar la hipótesis nula del ANOVA con medidas repetidas ya que informa sobre el grado de discrepancia entre las μ_{+j} y posee una distribución muestral conocida.

4.3. Supuestos

4.3.1. Independencia

Este supuesto indica que los sujetos de la muestra han sido seleccionados de manera aleatoria de su población y asignado a todos los J tratamientos del factor, lo que implica que dentro de cada tratamiento, el valor de la respuesta observada de un sujeto es independiente al del resto de sujetos. En consecuencia, lo que una puntuación se desvía del promedio de su grupo (e_{ij}) es independiente de lo que otra puntuación cualquiera se desvía de ese mismo grupo. Es decir, $Cov(e_{ij}, e_{ij'}) = 0$ (donde $i \neq i'$).

Este supuesto suele comprobarse mediante la prueba de las rachas y su incumplimiento puede ser muy grave de modo que conviene cuidar los aspectos relacionados con la selección y asignación de sujetos.

4.3.2. Normalidad

Dentro de cada tratamiento, la muestra obtenida con los n sujetos sigue una distribución normal. Para comprobar esto podemos utilizar una prueba no paramétrica sobre normalidad como la propuesta por Kolmogorov-Smirnov u otras similares.

4.3.3. Aditividad o Esfericidad

En 1970, Huynh y Feldt y Rouanet y Lépine demostraron independientemente que una condición suficiente y necesaria para llevar a cabo un análisis de medidas repetidas, sin aplicar corrección a los grados de libertad, es que las diferencias entre pares de puntuaciones tengan igual varianza. Si nos fijáramos en la matriz de varianzas y covarianzas, a pesar de que no revelaría la simetría compuesta, podría mostrar igualdad de varianzas entre pares de valores. Esta matriz se llamó la matriz H por Huynh y Feldt y se dijo que tenía la propiedad de esfericidad; Rouanet y Lépine se refirieron a la propiedad como circularidad. Sin embargo, las matrices que demuestran simetría compuesta también tienen la propiedad de esfericidad. La simetría compuesta es un caso especial de la esfericidad, pero esta última es menos restrictiva que la simetría compuesta. Mientras que la simetría compuesta requiere $\{1/2[J(J+1)]\} - 2$ igualdades (de varianzas y covarianzas), la esfericidad necesita solo $\{1/2[J(J-1)]\} - 1$ igualdades (solo de las diferencias de las varianzas; Rouanet y Lépine, 1970)

Como $\sigma_{jk} = \rho\sigma_y^2$ y $\sigma_e^2 = \sigma_y^2(1 - \rho)$, la matriz de varianzas-covarianzas tiene que tener la siguiente forma, la cual desarrollamos anteriormente en la fórmula (4.2)

$$\Sigma = \sigma_y^2 \begin{pmatrix} 1 & & \rho \\ & \ddots & \\ \rho & & 1 \end{pmatrix}$$

Es decir, que las varianzas de los tratamientos (diagonal principal) son iguales y que las covarianzas entre tratamientos (fuera de la diagonal) también lo son. Las matrices que cumplen estas características se dice que tienen la propiedad de la simetría compuesta (o *compound symmetry*) y se conocen como matrices de tipo S (Box, 1954). Esto solo es posible bajo el modelo aditivo.

La simetría compuesta es una condición suficiente para comprobar que el estadístico F se distribuye mediante una F con $J-1$ y $(J-1)(n-1)$ grados de libertad bajo la hipótesis nula del ANOVA en medidas repetidas. Sin embargo no es una condición necesaria. Huynh y Feldt (1970), Rouanet y Lepine (1970) y más tarde Mendoza (1980) demostraron que el estadístico F también es válido si las varianzas de todos los pares posibles de diferencias entre tratamientos son iguales. Es decir:

$$\text{Si } \sigma_{Y_j - Y_k}^2 = \sigma_{Y_j}^2 + \sigma_{Y_k}^2 - 2\sigma_{Y_j Y_k} = \text{constante para todo } j \text{ y } k.$$

La esfericidad solamente requiere que la matriz Σ , una vez transformada mediante la utilización de polinomios ortonormales, sea una matriz diagonal que contenga varianzas iguales en la misma. La transformación polinómica ortonormal viene dada por:

$$C^* \Sigma C^* \quad (4.3)$$

Donde Σ es la matriz de varianzas-covarianzas poblacional de orden $J \times J$ y C^* la matriz de coeficientes ortonormales de orden $(J-1) \times J$.

La hipótesis de que la matriz resultante de la expresión (4.3) es una matriz diagonal que contiene varianzas iguales en la misma se plantea formalmente mediante

$$\begin{aligned} H_0: C^* \Sigma C^* &= \sigma^2 I_J \\ H_0: C^* \Sigma C^* &\neq \sigma^2 I_J \end{aligned} \quad (4.4)$$

Donde σ^2 es la varianza poblacional e I_J es la matriz identidad de orden J .

Para comprobar la hipótesis planteada en (4.4) utilizamos el estadístico W o criterio de razón de verosimilitud de Mauchly (1940), cuyo valor se calcula mediante:

$$W = \frac{|C^* \hat{\Sigma} C^*|}{\left| \text{traza} \frac{(C^* \hat{\Sigma} C^*)}{J-1} \right|^{J-1}} \quad (4.5)$$

Siendo $\hat{\Sigma}$ la matriz de varianzas y covarianzas muestral.

El estadístico W puede convertirse en una X^2 mediante:

$$X^2 = -\left(n - 1 - \frac{2J+5}{6}\right) \ln W$$

X^2 se distribuye aproximadamente según una chi-cuadrado con $J(J-1)/2$ grados de libertad. Para muestras suficientemente grandes, un valor $X^2 < X^2_{(1-\alpha)J(J-1)/2}$ indica que hay evidencia para apoyar el supuesto de esfericidad y por tanto, que las diferencias entre varianzas muestrales pueden atribuirse al azar.

Esta prueba, sin embargo, tiene poca potencia para tamaños de muestra pequeños. Además, para muestras de gran tamaño, es probable que la prueba muestre significación, aunque el efecto en la prueba F sea no significativo. La prueba de Mauchly para la esfericidad también ha demostrado ser sensible cuando usamos datos no-normales; en particular, es conservadora para distribuciones de colas ligeras y lo contrario para distribuciones de colas pesadas. También es muy sensible a los valores atípicos. Debido a estas propiedades, no es de gran utilidad práctica. Si el supuesto de esfericidad no se cumple, como ocurre frecuentemente, el ANOVA-MR de un factor proporciona una razón F sesgada positivamente, aumentando la probabilidad del error tipo I. En tal caso se tienen dos opciones:

- 1) Ajustar los grados de libertad del estadístico F.
- 2) Aplicar un análisis de varianza multivariado (o MANOVA) basado en J-1 contrastes independientes.

En cuanto a la primera opción, lo que se propone es que los grados de libertad se ajusten mediante una prueba más conservadora (Box, 1954). La prueba *épsilon* (ϵ) es una corrección que se aplica en los grados de libertad del numerador y denominador del estadístico F de forma que éste pasa a distribuirse según F con $\epsilon(J-1)$ y $\epsilon(J-1)(n-1)$ grados de libertad. El valor de *épsilon* depende del grado en que la matriz Σ se desvía de la esfericidad, es 1 cuando se cumple el supuesto de esfericidad y se acerca a $1/(J-1)$ cuando no se cumple.

En la práctica el valor de ϵ suele ser desconocido pero suele estimarse a partir de la matriz $\hat{\Sigma}$. Box (1954) fue el primero en proponer un estimador de ϵ para abordar el problema de la violación del supuesto de la esfericidad. Se trata de la prueba $\hat{\epsilon}$, que fue más tarde adaptada por Greenhouse y Geisser (1959), por lo que se conoce como la *épsilon* de Greenhouse y Geisser. $\hat{\epsilon}$ se calcula a partir de las sumas y medias de la matriz de varianzas-covarianzas muestral ($\hat{\Sigma}$) mediante la fórmula:

$$\hat{\epsilon} = \frac{J^2(\bar{s}_u - \bar{s})^2}{(J-1)(\sum \sum s_{jk}^2 - 2J\sum s_i^2 + J^2 s^2)} \quad (4.6)$$

Donde:

\bar{s} : media de todos los elementos de la matriz $\hat{\Sigma}$

$\overline{s_{ii}}$: media de las varianzas que aparecen en la diagonal

\bar{s}_i : media de los elementos de la fila i

s_{jk} : cualquiera de las covarianzas de la matriz $\hat{\Sigma}$

Los grados de libertad tanto del numerador como del denominador de F se multiplican por $\hat{\epsilon}$ y la significación del cociente se evalúa con los nuevos grados de libertad. Si el supuesto de esfericidad se cumple, el valor de $\hat{\epsilon}$ se acercará a 1. Por el contrario, la peor violación del supuesto ocurriría con $\hat{\epsilon} = 1/(J-1)$

La prueba $\hat{\epsilon}$ tiende a ser excesivamente conservadora para muestras de tamaño pequeño. Huynh y Feldt (1970) propusieron otra corrección epsilon para solventar este problema, la prueba $\tilde{\epsilon}$:

$$\tilde{\epsilon} = [N(J - 1)\hat{\epsilon}] - \frac{2}{(J-1)[N-g-(J-1)\hat{\epsilon}]} \quad (4.7)$$

Donde:

$\hat{\epsilon}$: epsilon de Greenhouse y Geisser

N : n° total de sujetos

J : n° de tratamientos

G : n° de grupos (en los diseños de un factor siempre es 1)

El valor de $\tilde{\epsilon}$ siempre es mayor que el de $\hat{\epsilon}$ y puede ser incluso mayor que 1, en cuyo caso se iguala a 1.

Teniendo en cuenta que el cálculo de $\tilde{\epsilon}$ es complicado Greenhouse y Geisser (1959) recomendaron que para decidir si es necesario obtenerlo pueden llevarse a cabo unos análisis previos basados en $\hat{\epsilon}$. Puesto que ϵ oscila entre 1 y $1/(J-1)$ los grados de libertad más pequeños en F son $(J-1) \times 1/(J-1) = 1$ para el numerador y $(J-1)(n-1) \times 1/(J-1) = n-1$ para el denominador.

Los análisis previos constan de los siguientes pasos:

1º. Comparar el valor crítico de la F obtenida en el ANOVA ($F_{(1-\alpha), (J-1), (J-1)(n-1)}$) con la obtenida con los grados de libertad ajustados ($F_{(1-\alpha), \hat{\epsilon}(J-1), \hat{\epsilon}(J-1)(n-1)}$). Si el valor de la última no aumenta, no es necesario continuar ya que F no será significativa. En caso contrario, seguir con el paso 2

2º. Comparar con el valor de la F ajustada mediante $\hat{\epsilon}$ con el de la F con 1 y $n-1$ grados de libertad. Si este último resulta menor entonces es significativo, por lo que no será necesario continuar. En caso contrario continuar con el paso 3.

3º. Si el valor de la F ajustada es mayor que el obtenido mediante el ANOVA-MR pero menor que el valor de las tablas correspondiente a una F con 1 y $n-1$ grados de libertad, utilizar la F ajustada mediante la prueba $\hat{\epsilon}$.

En cuanto a la decisión sobre si emplear $\hat{\epsilon}$ o $\tilde{\epsilon}$, algunos autores sugieren que $\hat{\epsilon}$ es demasiado conservadora para muestras pequeñas y que, una vez satisfechos los tres pasos, $\tilde{\epsilon}$ es la corrección más adecuada (Jaccard y Ackerman, 1985). Sin embargo, los estudios de tipo Montecarlo de Collier, Baker, Mandeville y Hayes (1967) encontraron que esto no era así incluso con muestras tan pequeñas como de 5 sujetos. Por su parte, Stevens (1986) recomienda realizar una media de ambos estimadores dada la tendencia de $\hat{\epsilon}$ y $\tilde{\epsilon}$ a infraestimar y sobrestimar (respectivamente) el valor de ϵ . Sobre la base de estos criterios Barcikowski y Robey (1983) establecieron las siguientes recomendaciones generales:

1) Si, dado que se conoce ϵ y $\epsilon \geq 0,75$ emplear la corrección $\tilde{\epsilon}$

2) Si $\epsilon < 0,75$, emplear la corrección más conservadora $\hat{\epsilon}$

3) Si no se sabe nada sobre ϵ , como suele ocurrir, emplear la corrección $\hat{\epsilon}$ ya que es menos sensible al aumento de error de tipo I.

4.3.3.1. Enfoque Multivariante

Otra alternativa al análisis univariante cuando el supuesto de esfericidad no se cumple consiste en utilizar un análisis multivariante (MANOVA). En este caso el supuesto de esfericidad no es un requisito para el análisis de datos. Para transformar los datos de ANOVA en formato MANOVA las observaciones originales se reconvirtan en nuevas variables. Esta transformación puede hacerse de diversas maneras pero siempre tendremos $J-1$ nuevas variables. Una de las formas es convertir los J grupos de observaciones en $J-1$ diferencias. El análisis se realizará sobre estas nuevas $J-1$ variables (Girden, 1992). Otra posible transformación consiste en primero crear una matriz de coeficientes ortonormales (como la matriz C^* definida anteriormente), ponderando cada observación por su coeficiente ortonormal llevando a cabo el análisis sobre las nuevas $J-1$ variables resultantes. Ambos procedimientos producen los mismos resultados.

Con las $J-1$ variables nuevas resultantes se realiza el contraste multivariante T^2 de Hotelling cuyo estadístico se calcula mediante:

$$T^2 = (n)(\bar{y})(S_d^{-1})(\bar{y}')$$

Donde S_d^{-1} es la inversa de la matriz de varianzas-covarianzas de las nuevas variables e \bar{y} el vector fila de medias.

La T^2 puede convertirse en una F mediante:

$$F = \frac{n - J + 1}{(n - 1)(J - 1)} T^2$$

El estadístico F se distribuye según F con $(J-1)$ y $(n-J+1)$ grados de libertad. Si su valor es significativo, indica que hay diferencias entre las medias de las variables.

Estas operaciones son relativamente complicadas por lo que suelen realizarse con la ayuda del ordenador.

Ya que en el MANOVA los contrastes sobre las variables son independientes, las correlaciones entre las nuevas variables son muy bajas o nulas. Por lo tanto, el análisis puede realizarse sin tener que asumir esfericidad en la matriz de varianzas-covarianzas, aunque se asume que los datos siguen una distribución multivariada normal y el número de sujetos tiene que ser mayor que el número de tratamientos (Bray y Maxwell, 1985)

4.3.3.2. ANOVA Vs. MANOVA

Diversos estudios sugieren que no hay una clara ventaja entre usar el enfoque univariado o multivariado. De hecho, para algún conjunto de datos el enfoque univariado es más potente, mientras que para otros casos, con la violación del supuesto de esfericidad, el enfoque multivariante puede ser más adecuado (Romaniuk 1977).

Existen tests preliminares para determinar si el supuesto de esfericidad es plausible o no. (véase Kirk, 1982). Mientras que algunos recomiendan realizar estos tests antes del ANOVA (e.g. Huynh y Mandeville, 1979), la mayoría no. Estos tests son poco sensibles a las desviaciones de la normalidad y producen mayores errores de tipo I. De hecho, Keselman, Mendoza, Rogan y Breen (1980) encontraron evidencias de que la esfericidad (probada por Mauchley) casi siempre se viola y recomienda seguir el procedimiento de los tres pasos de Greenhouse y Geisser.

No obstante, el ANOVA es más potente que el MANOVA cuando existe homogeneidad en las varianzas ya que los grados de libertad asociados al término error son mayores que en la T^2 de Hotelling. Sin embargo, si existe heterogeneidad, los efectos pueden quedar enmascarados en el ANOVA y el MANOVA resultar más robusto. Asimismo, el MANOVA no es apropiado para situaciones donde $(n-1) < J$, la matriz de varianza-covarianza es singular (no tiene inversa) y no tiene solución (Greenhouse y Geisser, 1959). (Ver *pág.* 26 R. Girden, Ellen [1])

4.4. Medidas del tamaño del efecto

El análisis de la varianza permite averiguar si la variable independiente influye sobre la variable dependiente pero no informa sobre el grado de asociación entre ambas (es decir, ρ). Las medidas del tamaño del efecto son estimaciones de la proporción de varianza de la variable dependiente que es explicada por la variable independiente (es decir ρ^2) y permiten completar la conclusión alcanzada con la F del ANOVA-MR.

De entre las medidas del tamaño del efecto propuestas la más simple de todas es la de Pearson (1905):

$$\eta^2 = \frac{SCI}{SCT}$$

η^2 (eta al cuadrado) es el cociente entre la variabilidad debida al factor y la variabilidad total. Pese a que esta medida es un estimador sesgado de ρ^2 , sigue siendo una de las más utilizadas.

Otra de las medidas del tamaño del efecto más utilizadas es la estimación del índice omega al cuadrado ω^2 de Hayes (1988) cuyo valor depende del tipo de modelo planteado y sus supuestos. Si se asume el modelo aditivo la variabilidad total depende de la suma de la variabilidad debida al factor, a los sujetos y a los errores aleatorios y se calcula mediante:

$$\hat{\omega}^2 = \frac{SCI - (J - 1)MCE}{SCT + MCP}$$

Por otra parte, si se asume el modelo no aditivo, hay que añadir la variabilidad debida a la interacción intrasujetos y tratamientos. Esto tiene el problema de que el denominador del cociente suma más que la variabilidad total con la que el valor de ω^2 queda ligeramente infraestimado. Una alternativa para solucionar este problema consiste en aplicar la siguiente fórmula:

$$\tilde{\omega}^2 = \frac{SCI - (J - 1)MCE}{SCT + MCP + (n_j)MCE}$$

De las medidas propuestas la más recomendable es ω^2 ya que puede aplicarse en diferentes diseños experimentales (Maxwell, Camp y Arvey, 1980)

4.5. Ejemplo práctico en SAS

Para mostrar cómo se lleva a cabo el análisis con medidas repetidas en un factor en SAS extraemos el siguiente ejemplo de "Análisis de varianza con medidas repetidas" de Carmen Ximénez y Rafael San Martín [2].

Ejemplo 1:

“Un psicólogo investigador de mercados pretende estudiar la respuesta afectiva hacia un perfume en función del grado de concentración en sus componentes aromáticas con el fin de orientar una campaña publicitaria para el lanzamiento del mismo. El perfume está elaborado con cuatro grados de concentración diferentes: el primero, de menor concentración tiene 500 partes de aroma por millón (ppm), el segundo 1000 ppm, el tercero 1500 ppm y el cuarto, la máxima concentración, 2000 ppm. La respuesta afectiva se mide en una escala de tipo Likert de 7 puntos. En un primer estudio piloto el psicólogo seleccionó una muestra aleatoria de 20 sujetos. Para evitar los efectos de arrastre y fatiga, evaluó la respuesta afectiva de los 20 sujetos hacia el perfume dejando transcurrir 5 días tras la exposición a cada grado de concentración aromática del mismo. Los resultados obtenidos aparecen en la tabla “ (Ximénez, Carmen y San Martín, Rafael [2])

Tabla 3. Respuesta afectiva de 20 sujetos hacia un perfume en cuatro grados de concentración.

Sujetos:	Grado de concentración del perfume			
	1	2	3	4
1	1	3	4	4
2	2	2	5	6
3	2	5	4	6
4	2	3	5	4
5	1	3	4	7
6	1	4	4	5
7	2	3	5	4
8	4	3	4	5
9	2	4	4	6
10	2	2	5	6
11	2	3	3	3
12	1	4	1	7
13	4	2	5	4
14	1	3	6	5
15	2	2	5	5
16	1	3	5	3
17	1	2	3	6
18	3	4	4	4
19	3	4	7	4
20	3	2	4	3

Observando la tabla vemos que estamos ante un diseño con un factor (grado de concentración del perfume) con cuatro niveles que son las concentraciones fijadas por el investigador, por lo tanto de efectos fijos.

En SAS podemos utilizar dos procedimientos para llevar a cabo el análisis con medidas repetidas: PROC GLM que involucra los enfoques univariante y multivariante tradicionales y PROC MIXED que utiliza un enfoque más general de la estructura de la covarianza.

A) PROC GLM

La primera cosa a destacar de PROC GLM es que es necesario que los datos estén balanceados dentro de los sujetos; esto es, no utiliza los datos de ningún sujeto que tenga valores que faltan. Después de identificar los sujetos con datos completos, tenemos que seleccionar un modelo útil en términos de los efectos fijos intersujetos (*between-subject*) e intrasujetos (*within-subject*). Los efectos intersujetos son aquellos cuyos niveles permanecen constantes dentro de los sujetos, mientras que los efectos intrasujetos cambian dentro de los sujetos.

En nuestro ejemplo univariante solo tenemos efectos intrasujetos que corresponden a DOSIS. Usando PROC GLM tenemos que seleccionar una transformación. Esto se debe a que PROC GLM lleva a cabo sus cálculos en un conjunto de variables de contraste enumerando una menos que el número de variables de medidas repetidas. La transformación por defecto es el contraste de los niveles.

PROC GLM realiza un test de significación estándar para los efectos intersujetos. Por otra parte, hay dos tipos de test diferentes disponibles para los efectos intrasujetos: univariante y multivariante. Como ya se ha dicho anteriormente, los test univariantes son adecuados cuando la matriz de varianzas-covarianzas intrasujetos tiene una cierta estructura conocida como de Tipo H (Huynh y Feldt 1970). PROC GLM realiza un test de esfericidad para esta estructura. Cuando el test de esfericidad no tiene un p-valor significativo, debemos usar el test univariante para los efectos dentro de los sujetos ya que bajo la suposición de Tipo H serán normalmente más potentes que los tests multivariantes.

Cuando el test de esfericidad resulta significativo, PROC GLM nos ofrece dos maneras de probar la significación de los efectos dentro de los sujetos. La primera manera es ajustando las pruebas univariantes, GLM muestra dos de estos ajustes: G-G (Greenhouse y Geisser 1959) y H-F (Huynh y Feldt 1976). La segunda manera se refiere a cuatro diferentes pruebas multivariantes: Lambda de Wilk, Traza de Pillai, Traza de Hotelling-Lawley y la raíz mayor de Roy. Estos tests están basados en una matriz de varianzas-covarianzas intrasujetos completamente general (sin una estructura concreta).

Figura 1. Procedimiento a seguir cuando usamos PROC GLM en medidas repetidas (Wolfinger y Chang [13])

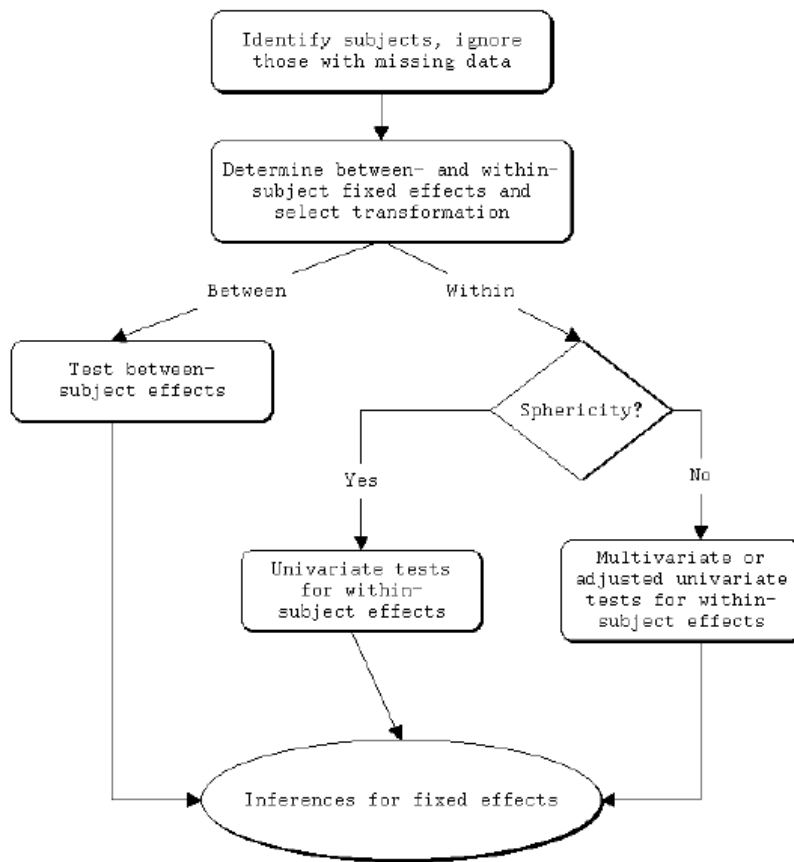


Figura 2. Código para la lectura de datos de el ejemplo de medidas repetidas en un factor

```
data perfume(keep=persona y1-y4);
input persona y1-y4;
output perfume;
datalines;
1 1 3 4 4
2 2 2 5 6
3 2 5 4 6
4 2 3 5 4
5 1 3 4 7
6 1 4 4 5
7 2 3 5 4
8 4 3 4 5
9 2 4 4 6
10 2 2 5 6
11 2 3 3 3
12 1 4 1 7
13 4 2 5 4
14 1 3 6 5
15 2 2 5 5
16 1 3 5 3
17 1 2 3 6
18 3 4 4 4
19 3 4 7 4
20 3 2 4 3
;
```

Hemos llamado al dataset perfume y hemos indicado que tiene 5 variables, la primera es el indicador del sujeto y las otras cuatro corresponden a las diferentes concentraciones del perfume.

El siguiente código es el utilizado para realizar el PROC GLM con medidas repetidas.

Figura 3. Código para analizar el Ejemplo 1 utilizando PROC GLM.

```
proc glm data = perfume;
model y1-y4=/ nouni;
repeated dosis 4 / printe;
run;
```

Como el ejemplo que estamos realizando es de medidas repetidas en un factor, no tendremos efecto intersujetos. Las cuatro variables de medidas repetidas están indicadas en el lado izquierdo de la ecuación indicada en la sentencia MODEL, en el lado derecho iría la variable de efectos fijos intrasujetos. La opción NOUNI hace que no salgan por pantalla los ANOVA unidireccionales para cada una de las 4 variables. La sentencia REPEATED contiene el efecto de medidas repetidas intrasujeto DOSIS el cual tiene 4 niveles. La opción PRINTE en la sentencia REPEATED hace que el PROC GLM realice el test para contrastar la esfericidad.

Figura 4. Test de esfericidad usando PROC GLM para el Ejemplo 1.

Tests de esfericidad				
Variables	DF	Criterio de Mauchly	Chi-cuadrado	Pr > ChiSq
Variables transformadas	5	0.1968396	28.805095	<.0001
Componentes ortogonales	5	0.681172	6.8042779	0.2356

La Figura 4 muestra los resultados de la opción PRINTE. Este test de esfericidad lo que hace es comprobar si la matriz de varianzas-covarianzas para los efectos de dentro de los sujetos tiene o no una estructura de covarianza de Tipo H.

Para nuestros datos el p-valor es 0.2356 no significativo, por lo tanto no podemos rechazar la hipótesis nula de que la matriz es de Tipo H. Por consiguiente, como hemos visto antes en la Figura 1, los tests estándar univariantes para DOSIS son apropiados.

Figura 5. Test multivariante para los efectos intrasujetos.

Criterio de test MANOVA y estadísticos F exactos para la hipótesis de efecto no dosis H = Tipo III Matriz SSCP para T. indedosis Matriz SSCP de error E = S=1 M=0.5 N=7.5					
Estadístico	Valor	F-Valor	Num DF	Den DF	Pr > F
Lambda de Wilks	0.16293172	29.11	3	17	<.0001
Traza de Pillai	0.83706828	29.11	3	17	<.0001
Traza Hotelling-Lawley	5.13754045	29.11	3	17	<.0001
Raíz más grande de Roy	5.13754045	29.11	3	17	<.0001

En la Figura 5 vemos los resultados de los test multivariantes para los efectos intrasujeto de DOSIS, aunque, teniendo en cuenta los resultados de Salidas 1, no es necesario suponer que estos datos son multivariantes. Aun así estos test son válidos pero menos potentes que los tests univariantes.

Figura 6. Test F univariantes para los efectos intrasujeto de DOSIS.

Procedimiento GLM							
Análisis de medidas repetidas de la varianza							
Test de hipótesis univariante para efectos Within Subject							
Fuente	DF	Tipo III SS	Cuadrado de la media	F-Valor	Pr > F	Adj Pr > F	
						G - G	H - F
dosis	3	99.63750000	33.21250000	24.87	<.0001	<.0001	<.0001
Error(dosis)	57	76.11250000	1.33530702				

Greenhouse-Geisser Epsilon	0.7821
Huynh-Feldt Epsilon	0.8992

Los dos ajustes mostrados se basan en ajustar los grados de libertad y ya hablamos de ellos en el apartado 4.3.3. Viendo los resultados de los tests vemos que no es necesario realizar ningún ajuste a los grados de libertad.

El p-valor del test es significativo por lo que rechazamos la hipótesis nula de igualdad de medias entre las diferentes concentraciones. Por tanto, podemos decir que el grado de concentración influye en la respuesta afectiva que el sujeto experimenta hacia el perfume

B) PROC MIXED

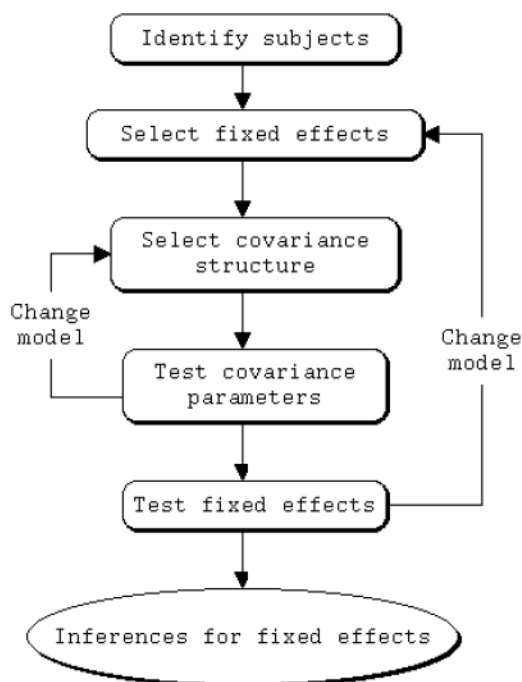
La primera diferencia con el PROC GLM es que con PROC MIXED podemos usar todos los datos en vez de ignorar aquellos no completos. Esto se debe a que PROC MIXED usa una estimación basada en la verosimilitud mientras que PROC GLM el método de los momentos el cual requiere que los datos estén completos.

La segunda es que no hay que hacer una distinción entre los efectos fijos intrasujetos e intersujetos. Simplemente hay que determinar el modelo de la media y colocar todos los efectos fijos en la sentencia MODEL.

La especificación de la media en PROC MIXED es más general que la de PROC GLM en dos maneras:

1. Puedes omitir los efectos de interacción inter e intrasujetos del modelo de media PROC MIXED mientras que no puedes en PROC GLM
2. Se pueden usar variables continuas en los efectos inter e intra sujetos en PROC MIXED, mientras que todos los efectos inter e intra sujetos deben de consistir en variables de clasificación en PROC GLM.

Figura 7. Procedimiento a seguir cuando usamos PROC MIXED en medidas repetidas (Wolfinger y Chang [13]).



Otra diferencia es que en PROC MIXED es necesario especificar la estructura de la covarianza. Esto es importante ya que los estadísticos test en PROC MIXED son funciones de este. Por lo tanto, se deberán comparar diferentes estructuras y seleccionar aquella que sea más razonable. Una estrategia para el proceso de selección de la estructura de la covarianza ha sido proporcionada por Wolfinger (1993) y está indicada en la Figura 7 mediante el bucle de retorno después de probar los parámetros de covarianza.

Después de seleccionar una estructura de covarianza, tenemos que comprobar la significación. Si algún factor no es significativo entonces lo podremos sacar del modelo y otra vez tendremos que asegurarnos que la estructura de la covarianza es apropiada.

Con un modelo final de media-varianza somos capaces de realizar inferencias sobre los efectos fijos llevando a cabo contrastes de hipótesis y construyendo intervalos de confianza.

Para poder utilizar el procedimiento PROC MIXED creamos una variable categórica DOSIS que indicará a qué concentración del perfume se referirá el valor de la nueva variable Y, la cual sustituye a la antigua y_1, \dots, y_4 .

Figura 8. Código para analizar el Ejemplo 1 utilizando PROC MIXED.

```
data perfume2(keep=persona dosis y);
input persona y1-y4;
y=y1; dosis =500; output perfume2;
y=y2; dosis =1000; output perfume2;
y=y3; dosis =1500; output perfume2;
y=y4; dosis =2000; output perfume2;
datalines;
1 1 3 4 4
2 2 2 5 6
3 2 5 4 6
4 2 3 5 4
5 1 3 4 7
6 1 4 4 5
7 2 3 5 4
8 4 3 4 5
9 2 4 4 6
10 2 2 5 6
11 2 3 3 3
12 1 4 1 7
13 4 2 5 4
14 1 3 6 5
15 2 2 5 5
16 1 3 5 3
17 1 2 3 6
18 3 4 4 4
19 3 4 7 4
20 3 2 4 3
;
```

Primero necesitamos seleccionar la estructura de las covarianzas que resulte más apropiada. Para ello vamos cambiando en el código la opción type con los diferentes tipos de estructuras de covarianzas

Figura 9. PROC MIXED con estructura de covarianza “simetría compuesta”.

```
proc mixed data = perfume2;
class dosis persona;
model y = dosis;
repeated / type = cs sub = persona;
run;
```

Figura 10. PROC MIXED con estructura de covarianza “Tipo H o Huynh-Feldt”

```
proc mixed data = perfume2;
class dosis persona;
model y = dosis;
repeated / type = hf sub = persona;
run;
```


Figura 11. PROC MIXED con estructura de covarianza "sin estructura".

```
proc mixed data = perfume2;
class dosis persona;
model y = dosis;
repeated / type = un sub = persona;
run;
```

Los tres PROC MIXED realizados anteriormente tienen las mismas declaraciones CLASS y MODEL. La declaración CLASS especifica los efectos de clasificación que en nuestro ejemplo serán DOSIS y PERSONA y la declaración MODEL indica el modelo de nuestros datos. La única diferencia entre los tres códigos anteriores está en la opción TYPE de la declaración REPEATED, aquí es donde seleccionamos qué tipo de estructura de covarianza queremos utilizar. La opción SUB indica cuáles son los sujetos del modelo, en nuestro ejemplo indicados por la variable PERSONA.

En nuestro ejemplo solo tenemos efectos intrasujeto pero si tuviéramos efectos intersujetos se colocarían también en la declaración MODEL. Esto contrasta con el PROC GLM ya que en este teníamos que indicar cuáles eran los principales efectos intrasujetos en la declaración REPEATED y los efectos intersujetos en la declaración MODEL.

Para ver cuál de los tres tipos de estructura es el más adecuado, nos fijamos en el índice de Akaike (AIC) y en el índice de Schwarz (BIC).

Figura 12. Estadísticos de ajuste para estructura de covarianza "simetría compuesta"

Estadísticos de ajuste	
Verosimilitud -2 Res Log	240.8
AIC (Smaller is Better)	244.8
AICC (Smaller is Better)	245.0
BIC (Smaller is Better)	246.8

Figura 13. Estadísticos de ajuste para estructura de covarianza "Tipo H".

Estadísticos de ajuste	
Verosimilitud -2 Res Log	240.2
AIC (Smaller is Better)	250.2
AICC (Smaller is Better)	251.0
BIC (Smaller is Better)	255.2

Figura 14. Estadísticos de ajuste para estructura de covarianza “No estructurado”.

Estadísticos de ajuste	
Verosimilitud -2 Res Log	232.9
AIC (Smaller is Better)	252.9
AICC (Smaller is Better)	256.3
BIC (Smaller is Better)	262.8

Fijándonos en los índices AIC y BIC de las salidas vemos que la estructura de covarianza más apropiada es “simetría compuesta” ya que sus índices son los más bajos. Por lo tanto, nos quedamos con el modelo de la *Figura 9*.

Figura 15. Test de tipo 3 para contrastar la hipótesis nula del modelo.

Tests de tipo 3 de efectos fijos				
Efecto	Num DF	Den DF	F-Valor	Pr > F
dosis	3	57	24.87	<.0001

En la *Figura 15* observamos que el p-valor para el contraste de la hipótesis nula es significativo por lo que podemos decir que rechazamos la hipótesis nula y por lo tanto no hay igualdad de medias como pasaba utilizando PROC GLM.

PROC MIXED calcula el estadístico F usando la forma cuadrática general de tipo Wald

$$F = \frac{\hat{\beta}L[L'(X'\hat{V}^{-1}X) - L]^{-1}L'\hat{\beta}}{\text{rango}(L)}$$

Donde $\hat{\beta}$ es una estimación del vector de parámetros de efectos fijos, L es una matriz de coeficientes de Tipo III, X es la matriz diseño de los efectos fijos y \hat{V} es la estimación de la matriz de varianzas y covarianzas de los datos, PROC GLM usa las razones de las medias cuadráticas para construir los tests F.

La ventaja que tenemos al usar el PROC MIXED es que puede utilizarse para cualquier tipo de estructura de la covarianza pero una desventaja es que los grados de libertad tienen que asignarse usando otros criterios. Para los análisis con medidas repetidas, PROC MIXED obtiene los grados de libertad dividiendo los grados de libertad residuales en intra e intersujetos y asignándoles a los efectos intra y intersujetos respectivamente. Para nuestro ejemplo tenemos solo efectos intrasujetos por lo que sus grados de libertad residuales equivalen a $80 - 4 = 76$

5. ANOVA CON MEDIDAS REPETIDAS DE DOS FACTORES

Otra situación que podemos tener con medidas repetidas es cuando tenemos dos factores. Mientras que en el anterior diseño teníamos un sólo factor con J niveles por los que pasan n sujetos de la muestra, en este tendremos J niveles en el factor A y K niveles en el factor B. A parte de poder determinar si cada uno de los factores afecta a la variable dependiente por separado, en los modelos ANOVA de dos factores podemos ver si la interacción de estos factores afecta a la variable dependiente.

En estos modelos el orden de administración de las JK combinaciones de tratamientos es aleatorizado independientemente para cada uno de los sujetos de la muestra.

En los modelos de ANOVA-MR de dos factores tenemos que considerar si las medidas repetidas se dan en uno o en ambos factores. Ambas opciones serán estudiadas en este apartado.

5.1. Medidas repetidas en un solo factor

La extensión del modelo ANOVA de un factor de efectos fijos completamente aleatorizado a un modelo que incluya un segundo factor con medidas repetidas dentro de los niveles del primero es uno de los más frecuentes en las ciencias sociales. Por ejemplo, diferentes tipos de psicosis pueden ser tratados por el mismo medicamento, con diferentes dosis efectivas. En vez de tener JK grupos independientes, se pueden incluir los diferentes tipos de psicosis y dentro de cada grupo probar las diferentes dosis en todos los individuos. Por consiguiente, el primer factor es de grupos independientes, y el segundo factor es medidas repetidas en todos los individuos de cada grupo. Los diseños que incluyen un factor de tipo intersujetos y otro de tipo intrasujetos se conocen como diseños mixtos.

En otro ejemplo, podríamos considerar una preocupación entre sociólogos sobre cómo los hombres y las mujeres adolescentes atribuyen la pobreza o riqueza a diferentes causas. El género (factor B) sería manipulado encuestando a hombres y mujeres adolescentes formando dos grupos independientes, pero a todos los adolescentes se les mediría la importancia que le dan a las cuatro explicaciones (factor A).

En general, si uno de los factores es una variable de la personalidad o del organismo, entonces el factor (B) sólo puede variar seleccionando grupos independientes. Si a cada individuo de la segunda variable (A) no le afectaría participar en cada condición (i.e. no hay efecto *carry-over*) y esta variable puede introducirse como cuantitativa (e.g. diferentes edades) o cualitativa (e.g. diferentes explicaciones de la pobreza), entonces se podrían utilizar las mismas personas en todas las condiciones.

“Cuando los sujetos se asignan a los niveles del factor A de forma aleatoria, reciben el nombre de diseños split-plot. Una desventaja de este diseño es que el efecto del factor intersujetos puede confundirse con el del factor intrasujetos ya que cada grupo de sujetos es asignado solamente a uno de esos A_j niveles. Por tanto, en este modelo se plantean también tres hipótesis: una referida al factor intersujetos (A), otra al del factor intrasujetos (B) y una última referida a los efectos de la interacción.” (Ximénez, Carmen y San Martín, Rafael [2])

5.1.1. Modelo

El modelo de ANOVA de dos factores con medidas repetidas en uno de ellos es una versión del modelo lineal general adaptado al caso en el que hay dos variables independientes; una que divide la muestra en dos o más grupos y otra referida a la muestra total. El modelo que describe la variable dependiente Y_{ijk} , es el siguiente:

$$Y_{ijk} = \mu + \alpha_j + \beta_k + \pi_{i(k)} + \alpha\beta_{jk} + \pi\alpha_{i(k)j} + e_{ijk} \quad (5.1)$$

En este modelo:

1. Y_{ijk} es la puntuación obtenida en la variable Y por el sujeto i del grupo j en el tratamiento k del factor B.
2. μ representa la parte de media común a todos los sujetos, se calcula mediante:

$$\mu = \frac{\sum \mu_{+jk}}{JK}$$

Siendo su estimador: $\hat{\mu} = \bar{Y}$

3. El parámetro α_j representa el efecto propio del j -ésimo nivel del factor A o el efecto atribuible al grupo j bajo el que se obtuvo la puntuación Y_{ijk} :

$$\begin{aligned} \alpha_j &= \mu_{+j+} - \mu \\ \hat{\alpha}_j &= \bar{Y}_{+j+} - \bar{Y} \end{aligned}$$

4. El parámetro β_k representa el efecto propio del k -ésimo nivel del factor B o el efecto atribuible al tratamiento k bajo el que se obtuvo la puntuación Y_{ijk} :

$$\begin{aligned} \beta_k &= \mu_{++k} - \mu \\ \hat{\beta}_k &= \bar{Y}_{++k} - \bar{Y} \end{aligned}$$

5. $(\alpha\beta)_{jk}$ representa el efecto de la interacción entre el grupo j del factor A y el nivel k del factor B:

$$(\alpha\beta)_{jk} = \mu_{+jk} - \mu_{+j+} - \mu_{++k} + \mu$$

$$(\widehat{\alpha\beta})_{jk} = \bar{Y}_{+jk} - \bar{Y}_{+j+} - \bar{Y}_{++k} + \bar{Y}$$

6. π_i se refiere a los efectos debidos a las diferencias entre los sujetos:

$$\pi_i = \mu_{i++} - \mu_{+j+}$$

$$\hat{\pi}_i = \bar{Y}_{i++} - \bar{Y}_{+j+}$$

7. $(\pi\beta)_{ijk}$ representa el término de errores aleatorios o las desviaciones de las puntuaciones del sujeto i respecto al promedio de su grupo y de los tratamientos:

$$(\pi\beta)_{ijk} = Y_{ijk} - \mu_{+jk} - \mu_{i++} + \mu_{+j+}$$

$$(\widehat{\pi\beta})_{ijk} = Y_{ijk} - \bar{Y}_{+jk} - \bar{Y}_{i++} + \bar{Y}_{+j+}$$

8. e_{ijk} representa el término de errores aleatorios o las desviaciones de las puntuaciones del sujeto i respecto al promedio de su grupo y de los tratamientos:

$$e_{ijk} = Y_{ijk} - \mu_{+j+} - \mu_{++k} - \mu_{i++} + \mu$$

$$\hat{e}_{ijk} = Y_{ijk} - \bar{Y}_{+j+} - \bar{Y}_{++k} - \bar{Y}_{i++} + \bar{Y}$$

Los términos π_i , $(\pi\beta)_{ijk}$ y e_{ijk} son independientes y normales con media 0 y varianza σ_π^2 , $\sigma_{\pi\beta}^2$ y σ_e^2 respectivamente.

5.1.2. Supuestos

Las suposiciones subyacentes al estudio de dos factores con medidas repetidas en un factor son una mezcla de aquellas de los grupos independientes y las de un diseño de medidas repetidas de un factor. La variabilidad dentro del grupo es la misma para cada grupo y las observaciones se distribuyen normal e independientemente entre los grupos.

Sin embargo, en este caso el factor intrasujetos es el único que asume el supuesto de esfericidad para sus K niveles dentro de cada nivel A_j del factor intersujetos. Esto es, la matriz Σ de los K niveles del factor B ha de ser esférica e igual para todos los grupos A_j . Huynh (1978) se refirió a estas dos suposiciones como esfericidad multimuestra. Si los supuestos del modelo se sostienen, el tratamiento y los test F de interacción son válidos. Si se viola cualquiera de los supuestos, entonces la esfericidad no se sostiene y el contraste de la hipótesis sobre los efectos de la interacción es más sensible al error tipo I (Huynh y Feldt, 1980). El supuesto de esfericidad se viola la mayoría de las veces por lo que se han

investigado diferentes técnicas alternativas de corrección. Estas incluyen la corrección Greenhouse-Geisser que es la menos robusta, el ajuste $\hat{\epsilon}$, el ajuste $\tilde{\epsilon}$ y análisis multivariante. La robustez varía también en función del grado de la desviación de la esfericidad, del tipo de distribución de puntuaciones, del tamaño de la muestra y del número de niveles del factor intrasujetos en comparación con el tamaño de la muestra.

Cuando la desviación de la esfericidad es baja ($\epsilon \geq 0.75$), podemos utilizar $\tilde{\epsilon}$ para realizar el ajuste de los grados de libertad, pero solo cuando la muestra es mayor que el número de niveles del factor intrasujetos (Rogan et al., 1979) y el factor intersujetos tiene sólo dos niveles (Maxwell y Arvey, 1982). Por el contrario, cuando la desviación hacia la esfericidad es mayor ($\epsilon < 0.75$) la corrección $\hat{\epsilon}$ es más robusta, sobre todo si el factor intersujetos tiene más de dos niveles, la muestra es amplia y los datos siguen el modelo de distribución normal (Maxwell y Arvey, 1982).

5.1.3. Estadísticos de contraste

El modelo ANOVA de dos factores mixto es un contraste en el que se ponen a prueba tres hipótesis. Una sobre los efectos del factor A, otra sobre los efectos del factor B y una última sobre la interacción AB.

Planteamiento de las hipótesis:

$H_{0(A)}: \mu_{+1+} = \mu_{+2+} = \dots \mu_{+j+} = \mu_{+j'+}$; o bien $\alpha_j = 0$ para todo valor de j o j' ($j \neq j'$)

$H_{1(A)}: \mu_{+j+} \neq \mu_{+j'+}$ para algún valor de j o de j' ($j \neq j'$)

$H_{0(B)}: \mu_{++1} = \mu_{++2} = \dots \mu_{++k} = \mu_{++k'}$; o bien $\beta_k = 0$ para todo valor de k o k' ($k \neq k'$)

$H_{1(B)}: \mu_{++k} \neq \mu_{++k'}$ para algún valor de k o k' ($k \neq k'$)

$H_{0(AB)}: \mu_{+jk} - \mu_{+j'k} = \mu_{+j+} - \mu_{+j'+}$; o bien $(\alpha\beta)_{jk} = 0$ para todo valor de j o k

$H_{1(AB)}: \mu_{+jk} - \mu_{+j'k} \neq \mu_{+j+} - \mu_{+j'+}$ para algún valor de j, j' o k ($j \neq j'$)

Partiendo del modelo planteado y sus supuestos podemos derivar los estadísticos de contraste con distribución muestral conocida que permitan poner a prueba las hipótesis planteadas. Si sustituimos los valores poblacionales del modelo planteado en (5.1) por sus correspondientes estimadores y trasladamos a la izquierda el término \bar{Y} se obtienen las siete fuentes de variabilidad del modelo:

$$Y_{ijk} - \bar{Y} = (\bar{Y}_{+j+} - \bar{Y}) + (\bar{Y}_{++k} - \bar{Y}) + (\bar{Y}_{+jk} - \bar{Y}_{+j+} - \bar{Y}_{++k} + \bar{Y}) + (\bar{Y}_{i++} - \bar{Y}_{+j+}) + (Y_{ijk} - \bar{Y}_{+jk} - \bar{Y}_{i++} + \bar{Y}_{+j+}) + (Y_{ijk} - \bar{Y}_{+j+} - \bar{Y}_{++k} + \bar{Y}) \quad (5.2)$$

Calculando la suma de cuadrados de cada uno de los términos de (5.2) en el conjunto de las N puntuaciones y transformándolo, obtenemos las sumas de cuadrados y sus grados de libertad.

Sumas de cuadrados

g.l.

$$SCT = \sum_i \sum_j \sum_k Y_{ijk}^2 - \frac{T^2}{N} \quad N - 1$$

$$SCInter = \sum_i \sum_j \frac{T_{ij+}^2}{K} - \frac{T^2}{N} \quad nJ - 1$$

$$SCA = \sum_j \frac{T_{+j+}^2}{nK} - \frac{T^2}{N} \quad J - 1$$

$$SCP = \sum_i \sum_j \frac{T_{ij+}^2}{K} - \sum_j \frac{T_{+j+}^2}{nK} \quad J(n - 1)$$

$$SCIntra = \sum_i \sum_j \sum_k Y_{ijk}^2 - \sum_i \sum_j \frac{T_{ij+}^2}{K} \quad nJ(K - 1)$$

$$SCB = \sum_k \frac{T_{++k}^2}{nJ} - \frac{T^2}{N} \quad K - 1$$

$$SCAB = \sum_j \sum_k \frac{T_{+jk}^2}{n} - \sum_j \frac{T_{+j+}^2}{nK} - \sum_k \frac{T_{++k}^2}{nJ} + \frac{T^2}{N} \quad (J - 1)(K - 1)$$

$$SC(B \times Suj.) = \sum_i \sum_j \sum_k Y_{ijk}^2 - \sum_i \sum_j \frac{T_{ij+}^2}{K} - \sum_j \sum_k \frac{T_{jk+}^2}{n} + \sum_j \frac{T_{+j+}^2}{nK} \quad J(K - 1)(n - 1)$$

Las medias cuadráticas, su formulación y sus valores esperados aparecen en la siguiente tabla:

Tabla 4. Medias cuadráticas y valores esperados en el modelo de dos factores con medidas repetidas en uno solo.

Fuente de variación	Media cuadrática	Valor esperado
Factor A	$MCA = \frac{SCA}{J-1}$	$\sigma^2 + K\sigma_{\pi}^2 + nK \sum \alpha^2 / (J-1)$
Sujetos (P)	$MCP = \frac{SCP}{J(n-1)}$	$\sigma^2 + K\sigma_{\pi}^2$
Factor B	$MCB = \frac{SCB}{K-1}$	$\sigma^2 + \sigma_{\beta\pi}^2 + nJ \sum \beta^2 / (K-1)$
Interacción AB	$MCAB = \frac{SCAB}{(J-1)(K-1)}$	$\sigma^2 + \sigma_{\beta\pi}^2 + n_{jk} \sum \sum (\alpha\beta)^2 / (J-1)(K-1)$
B x Sujetos	$MC(B \times Suj.) = \frac{SC(B \times Suj.)}{(J-1)(K-1)}$	$\sigma^2 + \sigma_{\beta\pi}^2$

Cada estadístico de contraste resulta del cociente entre la media cuadrática de la fuente de variabilidad y su correspondiente media cuadrática error (la intrasujetos, MC(B x Suj.) y la intersujetos, MCP). En la siguiente tabla tenemos un resumen sobre los estadísticos de contraste y sus grados de libertad correspondientes (la F se distribuye con los grados de libertad del numerador y del denominador.)

Tabla 5. Tabla resumen del ANOVA de dos factores mixto.

FV	SC	g.l.	MC	F
Intersujetos	SCInter	$nJ - 1$		
Factor A (A)	SCA	$J - 1$	MCA	$F_A = \frac{MCA}{MCP}$
Sujetos (P)	SCP	$J(n - 1)$	MCP	
Intrasujetos	SCIntra	$nJ(K - 1)$		
Factor B	SCB	$K - 1$	MCB	$F_B = \frac{MCB}{MC(B \times Suj.)}$
Interacción AB	SCAB	$(J - 1)(K - 1)$	MCAB	$F_{AB} = \frac{MCAB}{MC(B \times Suj.)}$
B x Sujetos	SC(B x Suj.)	$J(K - 1)(n - 1)$		
Total (T)	SCT	$N - 1$		

5.1.4. Ejemplo práctico en SAS

Volvemos al ejemplo del ANOVA de un factor con medidas repetidas en el que se estudiaba la respuesta efectiva hacia un perfume en cuatro concentraciones diferentes (Apartado 4.5).

Ejemplo 2:

“Se amplía este ejemplo incluyendo un factor de tipo intersujetos que divide la muestra en dos sexos, se construiría el factor A con dos niveles (A1: hombres, A2: mujeres). El factor B, grado de concentración del perfume es de tipo intrasujetos y tiene cuatro niveles por los que pasan todos los sujetos (B1:500 ppm, B2: 1000 ppm, B3: 1500 ppm, B4: 2000 ppm). En este caso el objetivo es estudiar el efecto de sexo, el grado de concentración del perfume y el de ambos en interacción sobre la respuesta afectiva hacia el perfume y el de ambos en interacción sobre la respuesta afectiva hacia el perfume, por lo que se plantearán tres hipótesis. Los resultados obtenidos con datos de una muestra de 20 sujetos aparecen en la tabla.”

Tabla 6. Tabla de datos para el ejemplo de dos factores con medidas repetidas en un solo factor.

Sujetos:	Hombres				Mujeres			
	Concentración del perfume				Concentración del perfume			
	1	2	3	4	1	2	3	4
1	1	3	4	4	2	5	4	6
2	2	2	5	6	4	2	5	4
3	1	4	1	7	1	3	4	7
4	1	4	4	5	2	3	5	4
5	3	2	4	3	1	3	6	5
6	1	2	3	6	2	3	5	4
7	2	2	5	5	3	4	4	4
8	1	3	5	3	3	4	7	4
9	2	3	3	3	4	3	4	5
10	2	2	5	6	2	4	4	6

A) PROC GLM

En este ejemplo se ha unificado el código para la lectura de datos y el dataset perfume2f será el utilizado cuando usamos PROC GLM y perfume2f2 cuando usamos PROC MIXED.

Figura 16. Lectura de datos para el Ejemplo 2.

```
data perfume2f(keep=persona genero y1-y4)
perfume2f2 (keep=persona genero dosis y);
input persona genero$ y1-y4;
output perfume2f;
y=y1; dosis =500; output perfume2f2;
y=y2; dosis =1000; output perfume2f2;
y=y3; dosis =1500; output perfume2f2;
y=y4; dosis =2000; output perfume2f2;
datalines;
```

```
1 H 1 3 4 4
2 H 2 2 5 6
3 H 1 4 1 7
4 H 1 4 4 5
5 H 3 2 4 3
6 H 1 2 3 6
7 H 2 2 5 5
8 H 1 3 5 3
9 H 2 3 3 3
10 H 2 2 5 6
11 M 2 5 4 6
12 M 4 2 5 4
13 M 1 3 4 7
14 M 2 3 5 4
15 M 1 3 6 5
16 M 2 3 5 4
17 M 3 4 4 4
18 M 3 4 7 4
19 M 4 3 4 5
20 M 2 4 4 6
;
```

Figura 17. Código para analizar el Ejemplo 2 utilizando PROC GLM..

```
proc glm data=perfume2f;
class genero;
model y1-y4=genero / nouni;
repeated dosis 4 / printe;
run;
```

En este ejemplo tenemos un factor de medidas repetidas que sigue siendo DOSIS y un factor independiente que es GENERO. En el modelo se indican las cuatro variables de medidas repetidas como en el ejemplo de un factor y el factor independiente se añade a la derecha de la igualdad indicada en MODEL.

Figura 18. Test de esfericidad para nuestro ejemplo usando PROC GLM.

Tests de esfericidad				
Variables	DF	Criterio de Mauchly	Chi-cuadrado	Pr > ChiSq
Variables transformadas	5	0.2109196	26.024434	<.0001
Componentes ortogonales	5	0.6960332	6.0594296	0.3005

Para nuestros datos el p-valor del test de esfericidad es 0.3005, no significativo. Esto significa que no rechazamos la hipótesis de esfericidad. En consecuencia, los tests estándar univariantes son apropiados.

Figura 19. Test F univariante para los efectos intersujeto.

Procedimiento GLM Análisis de medidas repetidas de la varianza Tests de hipótesis para efectos Between Subjects					
Fuente	DF	Tipo III SS	Cuadrado de la media	F-Valor	Pr > F
genero	1	7.81250000	7.81250000	17.31	0.0006
Error	18	8.12500000	0.45138889		

En este ejemplo tenemos como único factor intersujeto GENERO, el test es significativo a nivel 1%. Esto significa que rechazamos la hipótesis nula y por lo tanto podemos decir que hay una diferencia de percepción del olor del perfume dependiendo de nuestro género.

Figura 20. Test F univariante para los efectos intrasujeto.

Procedimiento GLM Análisis de medidas repetidas de la varianza Test de hipótesis univariante para efectos Within Subject							
Fuente	DF	Tipo III SS	Cuadrado de la media	F-Valor	Pr > F	Adj Pr > F	
						G - G	H-F-L
dosis	3	99.63750000	33.21250000	24.18	<.0001	<.0001	<.0001
dosis*genero	3	1.93750000	0.64583333	0.47	0.7043	0.6608	0.6889
Error(dosis)	54	74.17500000	1.37361111				

Greenhouse-Geisser Epsilon	0.7919
Huynh-Feldt-Lecoutre Epsilon	0.9203

En la salida anterior podemos ver que el factor de medidas repetidas DOSIS tiene un p-valor significativo y que la interacción de este con el factor intersujeto GENERO es no significativa. Esto significa que la concentración del perfume sí que afecta a la respuesta afectiva que el sujeto experimenta hacia el perfume, sin embargo, la interacción entre el género y la concentración no.

B) PROC MIXED

Figura 21. PROC MIXED con estructura de covarianza “simetría compuesta”.

```
proc mixed data=perfume2f2;
class genero dosis persona;
model y = genero|dosis;
repeated / type=cs sub=persona;
run;
```

Figura 22. PROC MIXED con estructura de covarianza “Tipo H”.

```
proc mixed data=perfume2f2;
class genero dosis persona;
model y = genero|dosis;
repeated / type=hf sub=persona;
run;
```

Figura 23. PROC MIXED con estructura de covarianza “sin estructura”.

```
proc mixed data=perfume2f2;
class genero dosis persona;
model y = genero|dosis;
repeated / type=un sub=persona;
run;
```

Como ya vimos en el ejemplo anterior, los tres PROC MIXED tienen las mismas declaraciones exceptuando la opción TYPE que va cambiando para cada tipo de estructura de covarianza.

Para ver cuál de los tres tipos de estructura es la más adecuada, nos fijamos en el índice de Akaike (AIC) y el índice de Schwarz (BIC).

Figura 24. Estadísticos de ajuste para la estructura de covarianza “simetría compuesta”.

Estadísticos de ajuste	
Verosimilitud -2 Res Log	225.6
AIC (Smaller is Better)	229.6
AICC (Smaller is Better)	229.7
BIC (Smaller is Better)	231.6

Figura 25. Estadísticos de ajuste para la estructura de covarianza "Tipo H".

Estadísticos de ajuste	
Verosimilitud -2 Res Log	223.2
AIC (Smaller is Better)	233.2
AICC (Smaller is Better)	234.1
BIC (Smaller is Better)	238.2

Figura 26. Estadísticos de ajuste para la estructura de covarianza "No estructurada".

Estadísticos de ajuste	
Verosimilitud -2 Res Log	216.7
AIC (Smaller is Better)	236.7
AICC (Smaller is Better)	240.3
BIC (Smaller is Better)	246.6

Como podemos observar el índice AIC y BIC en simetría compuesta es el menor por lo tanto nos quedaremos con el modelo cuya estructura de covarianza es "simetría compuesta".

Figura 27. Test de tipo III para contrastar las hipótesis nulas del modelo.

Tests de tipo 3 de efectos fijos				
Efecto	Num DF	Den DF	F-Valor	Pr > F
genero	1	18	17.31	0.0006
dosis	3	54	24.18	<.0001
genero*dosis	3	54	0.47	0.7043

Fijándonos en los p-valores observamos que tanto para el factor GENERO como para el factor DOSIS su p-valor es muy significativo (< 0.01) mientras que para la interacción no. Esto nos indica que tanto el género como la concentración del perfume influyen en la respuesta afectiva que el sujeto experimenta hacia el perfume mientras que la interacción no produce ningún cambio.

5.2. Medidas repetidas en los dos factores

Otra situación con la que podemos encontrarnos es tener dos factores con medidas repetidas en ambos. Este diseño no se utiliza mucho, pero se ha utilizado para estudios sobre aprendizaje y atención. En este modelo asumimos que los n sujetos de la muestra pasan por todas las JK combinaciones posibles de tratamientos.

Tenemos un grupo de n sujetos del que se toman J medidas del factor A y K del factor B fijadas por el investigador (efectos fijos). En cada nivel de A tendremos K niveles del factor B. Los n sujetos pasan por todos los niveles de ambos factores y la muestra total estará constituida por nJK observaciones.

5.2.1. Modelo

El modelo aditivo asume que no existe interacción entre los i sujetos y los jk tratamientos y se representa con:

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \pi_i + \varepsilon_{ijk} \quad (5.2.1)$$

Puesto que este modelo sólo se puede justificar si la ausencia de interacción puede demostrarse estadísticamente (Kirk, 1982), el foco de atención será el modelo no aditivo ya que asume esta interacción.

El modelo estadístico que guiará el análisis será:

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \pi_i + (\pi\alpha)_{ij} + (\pi\beta)_{ik} + (\pi\alpha\beta)_{ijk} + e_{ijk} \quad (5.2.2)$$

En este modelo:

1. Y_{ijk} es la puntuación obtenida en la variable Y por el sujeto i del grupo j en el tratamiento k del factor B.
2. μ representa la parte de media común a todos los sujetos, se calcula mediante:

$$\mu = \frac{\sum \mu_{+jk}}{JK}$$

Siendo su estimador: $\hat{\mu} = \bar{Y}$

3. El parámetro α_j representa el efecto propio del j -ésimo nivel del factor A o el efecto atribuible al grupo j bajo el que se obtuvo la puntuación Y_{ijk} :

$$\begin{aligned} \alpha_j &= \mu_{+j+} - \mu \\ \hat{\alpha}_j &= \bar{Y}_{+j+} - \bar{Y} \end{aligned}$$

4. El parámetro β_k representa el efecto propio del k-ésimo nivel del factor B o el efecto atribuible al tratamiento k bajo el que se obtuvo la puntuación Y_{ijk} :

$$\beta_k = \mu_{++k} - \mu$$

$$\hat{\beta}_k = \bar{Y}_{++k} - \bar{Y}$$

5. $(\alpha\beta)_{jk}$ representa el efecto de la interacción entre el grupo j del factor A y el nivel k del factor B:

$$(\alpha\beta)_{jk} = \mu_{+jk} - \mu_{+j+} - \mu_{++k} + \mu$$

$$(\widehat{\alpha\beta})_{jk} = \bar{Y}_{+jk} - \bar{Y}_{+j+} - \bar{Y}_{++k} + \bar{Y}$$

6. π_i se refiere a los efectos debidos a las diferencias entre los sujetos:

$$\pi_i = \mu_{i++} - \mu$$

$$\hat{\pi}_i = \bar{Y}_{i++} - \bar{Y}_{+j+}$$

7. $(\pi\alpha)_{ij}$ es la interacción entre el sujeto i y el tratamiento j del factor A:

$$(\pi\alpha)_{ij} = \mu_{ij+} - \mu_{+j+} - \mu_{i++} + \mu$$

$$(\widehat{\pi\alpha})_{ij} = \bar{Y}_{ij+} - \bar{Y}_{+j+} - \bar{Y}_{i++} + \bar{Y}$$

8. $(\pi\beta)_{ik}$ es la interacción entre el sujeto i y el tratamiento k del factor B:

$$(\pi\beta)_{ik} = \mu_{i+k} - \mu_{++k} - \mu_{i++} + \mu$$

$$(\widehat{\pi\beta})_{ik} = \bar{Y}_{i+k} - \bar{Y}_{++k} - \bar{Y}_{i++} + \bar{Y}$$

9. $(\pi\alpha\beta)_{ijk}$ es la interacción entre el sujeto i y el efecto de la combinación jk de los JK tratamientos.

$$(\pi\alpha\beta)_{ijk} = Y_{ijk} - \mu_{+jk} - \mu_{ij+} - \mu_{i+k} + \mu_{+j+} + \mu_{++k} + \mu_{i++} - \mu$$

$$(\widehat{\pi\alpha\beta})_{ijk} = Y_{ijk} - \bar{Y}_{+jk} - \bar{Y}_{ij+} - \bar{Y}_{i+k} + \bar{Y}_{+j+} + \bar{Y}_{++k} + \bar{Y}_{i++} - \bar{Y}$$

10. e_{ijk} representa el término de errores aleatorios o las desviaciones de las puntuaciones del sujeto i no sólo respecto al promedio de su grupo sino también respecto a su promedio en todos los tratamientos:

$$e_{ijk} = Y_{ijk} - \mu_{+j+} - \mu_{++k} - \mu_{i++} + \mu$$

$$\hat{e}_{ij} = Y_{ijk} - \bar{Y}_{+j+} - \bar{Y}_{++k} - \bar{Y}_{i++} + \bar{Y}$$

5.2.2. Supuestos

En los dos modelos anteriores asumimos que las observaciones Y_{ijk} son normales e independientes y que existe homogeneidad en las varianzas de las variables y en las diferencias entre variables. Y_{ijk} es una variable aleatoria independiente y normalmente distribuida con media μ_{+jk} y varianza $\sigma_{\pi}^2 + \sigma_e^2$ para el modelo aditivo y $\sigma_{\pi}^2 + \sigma_e^2 + \sigma_{\pi\alpha}^2 + \sigma_{\pi\beta}^2 + \sigma_{\pi\alpha\beta}^2$ para el no aditivo.

Como en el modelo de un factor, en el de dos factores se asumen los supuestos de independencia, normalidad y esfericidad. En este último hay dos casos generales que considerar. El primero se refiere al modelo aditivo. Asumimos igualdad de varianzas $\sigma_{\pi\alpha}^2 = \sigma_{\pi\beta}^2 = \sigma_{\pi\alpha\beta}^2$. Cada varianza de la interacción parcial constituye un estimador de la varianza de la interacción total y el promedio ponderado de sus medidas cuadráticas, la media cuadrática del error para la prueba F del ANOVA. Para que esto se cumpla es necesario que la matriz de varianzas-covarianzas de las K combinaciones de tratamiento satisfaga el supuesto de esfericidad. El segundo caso es cuando el modelo no aditivo se considera más adecuado o cuando la matriz de varianzas-covarianzas total no cumple el supuesto de esfericidad. Si esto ocurre, aún es posible que cada una de las tres submatrices cumpla este supuesto y que la prueba F sea válida. En este caso es necesario utilizar la media cuadrática error derivada de las distintas medias cuadráticas de las interacciones parciales y los grados de libertad del denominador de F resultan más pequeños que en el primer caso. Si no se cumple el supuesto de esfericidad en cada una de las submatrices se procede como en el modelo de un factor; empleando las correcciones en los grados de libertad o los contrastes multivariados.

5.2.3. Estadísticos de contraste

$H_{0(A)}: \mu_{+1+} = \mu_{+2+} = \dots \mu_{+j+} = \mu_{+j'+}$; o bien $\alpha_j = 0$ para todo valor de j o j' ($j \neq j'$)

$H_{1(A)}: \mu_{+j+} \neq \mu_{+j'+}$ para algún valor de j o de j' ($j \neq j'$)

$H_{0(B)}: \mu_{++1} = \mu_{++2} = \dots \mu_{++k} = \mu_{++k'}$; o bien $\beta_k = 0$ para todo valor de k o k' ($k \neq k'$)

$H_{1(B)}: \mu_{++k} \neq \mu_{++k'}$ para algún valor de k o k' ($k \neq k'$)

$H_{0(AB)}: \mu_{+jk} - \mu_{+j'k} = \mu_{+j+} - \mu_{+j'+}$; o bien $(\alpha\beta)_{jk} = 0$ para todo valor de j o k

$H_{1(AB)}: \mu_{+jk} - \mu_{+j'k} \neq \mu_{+j+} - \mu_{+j'+}$ para algún valor de j, j' o k ($j \neq j'$)

Partiendo del modelo planteado y sus supuestos pueden derivarse los estadísticos de contraste con distribución muestral conocida que permitan poner a prueba las tres hipótesis planteadas tanto desde el modelo aditivo como desde el no aditivo. En primer lugar nos referimos al modelo aditivo.

Si sustituimos los valores poblacionales del modelo aditivo planteado en (5.2.2) por sus correspondientes estimadores muestrales obtenemos lo siguiente

$$Y_{ijk} - \bar{Y} = (\bar{Y}_{+j+} - \bar{Y}) + (\bar{Y}_{++k} - \bar{Y}) + (\bar{Y}_{+jk} - \bar{Y}_{+j+} - \bar{Y}_{++k} + \bar{Y}) + (\bar{Y}_{i++} - \bar{Y}) + (Y_{ijk} - \bar{Y}_{+j+} - \bar{Y}_{++k} - \bar{Y}_{i++} + \bar{Y})$$

Sumas de cuadrados

g.l.

$$SCT = \sum_i \sum_j \sum_k Y_{ijk}^2 - \frac{T^2}{N} \quad N - 1$$

$$SCInter = \sum_j \sum_k \frac{T_{+jk}^2}{n} - \frac{T^2}{N} \quad JK - 1$$

$$SCA = \sum_j \frac{T_{+j+}^2}{nK} - \frac{T^2}{N} \quad J - 1$$

$$SCB = \sum_k \frac{T_{++k}^2}{nJ} - \frac{T^2}{N} \quad K - 1$$

$$SCAB = \sum_j \sum_k \frac{T_{+jk}^2}{n} - \sum_j \frac{T_{+j+}^2}{nK} - \sum_k \frac{T_{++k}^2}{nJ} + \frac{T^2}{N} \quad (J - 1)(K - 1)$$

$$SCP = \sum_i \frac{T_{i++}^2}{JK} - \frac{T^2}{N} \quad (n - 1)$$

$$SCE = \sum_i \sum_j \sum_k Y_{ijk}^2 - \sum_j \frac{T_{+j+}^2}{nK} - \sum_i \frac{T_{i++}^2}{JK} + \frac{T^2}{N} \quad (JK-1)(n-1)$$

Tabla 7. Medidas cuadráticas y valores esperados en el modelo no aditivo de dos factores con medidas repetidas en ambos. (Ximénez, Carmen y San Martín, Rafael [2]).

Fuente de variación	Media cuadrática	Valor esperado
Factor A	$MCA = \frac{SCA}{J-1}$	$\sigma^2 + n_j \sum \alpha^2 / (J-1)$
Factor B	$MCB = \frac{SCB}{K-1}$	$\sigma^2 + n_k \sum \beta^2 / (K-1)$
Interacción AB	$MCAB = \frac{SCAB}{(J-1)(K-1)}$	$\sigma^2 + n_{jk} \sum \sum (\alpha\beta)^2 / (J-1)(K-1)$
Sujetos (P)	$MCP = \frac{SCP}{n-1}$	$\sigma^2 + JK\sigma_\pi^2$
Error (E)	$MCE = \frac{SCE}{(JK-1)(n-1)}$	σ^2

En el modelo no aditivo los efectos de la interacción entre cada factor (A, B y AB) y los sujetos están incluidos en el término MCE cuyo valor podemos descomponer en los siguientes tres componentes:

$$MC(A \times Suj.) = \frac{SC(A \times Suj.)}{(J-1)(n-1)}$$

$$MC(B \times Suj.) = \frac{SC(B \times Suj.)}{(K-1)(n-1)}$$

$$MC(AB \times Suj.) = \frac{SC(AB \times Suj.)}{(J-1)(K-1)(n-1)}$$

Donde las sumas de cuadrados se calculan:

$$SC(A \times Suj.) = \sum_i \sum_j \frac{T_{ij+}^2}{K} - \sum_j \frac{T_{+j+}^2}{nK} - \sum_i \frac{T_{i++}^2}{JK} + \frac{T^2}{N}$$

$$SC(B \times Suj.) = \sum_i \sum_k \frac{T_{i+k}^2}{J} - \sum_k \frac{T_{++k}^2}{nJ} - \sum_i \frac{T_{i++}^2}{JK} + \frac{T^2}{N}$$

$$SC(AB \times Suj.) = \sum_i \sum_j \sum_k Y_{ijk}^2 - \sum_i \sum_j \frac{T_{ij+}^2}{K} - \sum_i \sum_k \frac{T_{i+k}^2}{J} - \sum_j \sum_k \frac{T_{+jk}^2}{n} + \sum_i \frac{T_{i++}^2}{JK} + \sum_j \frac{T_{+j+}^2}{nK} + \sum_k \frac{T_{++k}^2}{nJ} + \frac{T^2}{N}$$

En la siguiente tabla aparecen indicados los valores esperados de las nuevas medias cuadráticas (ya que hemos introducido nuevas fuentes de variación).

Tabla 8. Valores esperados de las medias cuadráticas del modelo de dos factores con medidas repetidas en ambos (modelo no aditivo) (Ximénez, Carmen y San Martín, Rafael [2]).

Fuente de variación	Valor esperado
Factor A	$\sigma^2 + K\sigma_{\pi\alpha}^2 + n_j \sum \alpha^2 / (J - 1)$
Factor B	$\sigma^2 + J\sigma_{\pi\beta}^2 + n_k \sum \beta^2 / (K - 1)$
Interacción AB	$\sigma^2 + \sigma_{\alpha\beta}^2 + n_{jk} \sum \sum (\alpha\beta)^2 / (J - 1)(K - 1)$
Sujetos (P)	$\sigma^2 + JK\sigma_{\pi}^2$
A x Sujetos	$\sigma^2 + K\sigma_{\alpha p}^2$
B x Sujetos	$\sigma^2 + J\sigma_{\beta p}^2$
AB x Sujetos	$\sigma^2 + \sigma_{\alpha\beta p}^2$

5.2.4. Ejemplo práctico en SAS

Ejemplo 3:

“En el ejemplo del apartado sobre ANOVA-MR de un factor se estudiaba la respuesta afectiva hacia un perfume en cuatro grados diferentes de concentración fijados por el investigador (500 ppm, 1000 ppm, 1500 ppm, 2000 ppm). Ahora dicho investigador está interesado en estudiar un nuevo factor (factor B): el tipo de aroma del que está compuesto el perfume. Ha considerado tres tipos de aroma: B1: lavanda (el más suave); B2: espliego; y B3: rosas (el más fuerte). En cuanto al factor A, en este caso solamente decide considerar dos tipos de concentración de perfume: A1: la mínima (500 ppm) y A2: la máxima (2000 ppm). Ambas variables se miden en una nueva muestra que consta de 10 sujetos, los cuales pasan por todos los niveles de los factores considerados. Por tanto, el objetivo ahora es estudiar el efecto del grado de concentración del perfume y el tipo de aroma, y el de ambos factores en interacción sobre la respuesta afectiva hacia el perfume. La respuesta afectiva de los sujetos se mide en una escala de 1 a 7. Los resultados obtenidos aparecen en la Tabla 9.” (Ximénez, Carmen y San Martín, Rafael, [2])

Tabla 9. Respuesta afectiva de 10 sujetos hacia un perfume en dos grados de concentración y tres tipos de aroma.

Sujetos:	Concentración de 500 ppm			Concentración de 2000 ppm		
	Lavanda	Espliego	Rosas	Lavanda	Espliego	Rosas
1	3	4	5	5	6	5
2	4	4	4	6	7	4
3	4	6	7	4	6	7
4	4	4	5	5	7	7
5	2	4	5	5	4	7
6	2	3	7	5	6	7
7	4	3	5	4	7	6
8	6	6	5	6	6	6
9	4	6	6	7	3	7
10	4	5	4	4	6	6

A) PROC GLM

Figura 28. Lectura de datos para el Ejemplo 3.

```
data perfume2fMRglm;
input persona y1-y6;
datalines;
1 3 4 5 5 6 5
2 4 4 4 6 7 4
3 4 6 7 4 6 7
4 4 4 5 5 7 7
5 2 4 5 5 4 7
6 2 3 7 5 6 7
7 4 3 5 4 7 6
8 6 6 5 6 6 6
9 4 6 6 7 3 7
10 4 5 4 4 6 6
;
```

Figura 29. Código para analizar el Ejemplo 3 con PROC GLM

```
proc glm data=perfume2fMRglm;
model y1-y6= / nouni;
repeated dosis 2, flor 3 /printe;
run;
```

Estamos ante un estudio de dos factores con medidas repetidas en ambos por lo que a la hora de su lectura tendremos que indicarlo de forma jerárquica en la sentencia REPEATED siendo DOSIS el primer factor con dos niveles y dentro de cada dosis tenemos FLOR con tres niveles.

El test de esfericidad se realiza para aquellos factores con más de dos niveles por lo que en nuestro ejemplo tendremos un test para FLOR y otro para la interacción FLOR*DOSIS.

FLOR:

Figura 30. Test de esfericidad para la variable FLOR.

Tests de esfericidad				
Variables	DF	Criterio de Mauchly	Chi-cuadrado	Pr > ChiSq
Variables transformadas	2	0.3284426	8.9071466	0.0116
Componentes ortogonales	2	0.700471	2.8480183	0.2407

El p-valor es alto (0.2407) por lo que no rechazamos la hipótesis nula y podemos considerar que la matriz de varianzas y covarianzas es de Tipo H.

FLOR*DOSIS:

Figura 31. Test de esfericidad para la interacción de FLOR x DOSIS

Tests de esfericidad				
Variables	DF	Criterio de Mauchly	Chi-cuadrado	Pr > ChiSq
Variables transformadas	2	0.8691112	1.1222732	0.5706
Componentes ortogonales	2	0.7635467	2.1582484	0.3399

El p-valor es alto (0.3399) por lo que no rechazamos la hipótesis nula y podemos considerar que la matriz de varianzas y covarianzas de FLOR*DOSIS es de Tipo H.

Figura 32. Test F univariantes para los efectos intrasujetos DOSIS y FLOR y su interacción

Procedimiento GLM
Análisis de medidas repetidas de la varianza
Test de hipótesis univariante para efectos Within Subject

Fuente	DF	Tipo III SS	Cuadrado de la media	F-Valor	Pr > F
dosis	1	21.60000000	21.60000000	26.27	0.0006
Error(dosis)	9	7.40000000	0.82222222		

Fuente	DF	Tipo III SS	Cuadrado de la media	F-Valor	Pr > F	Adj Pr > F	
						G - G	H - F
flor	2	18.30000000	9.15000000	6.33	0.0083	0.0156	0.0110
Error(flora)	18	26.03333333	1.44629630				

Greenhouse-Geisser Epsilon	0.7695
Huynh-Feldt Epsilon	0.8973

Fuente	DF	Tipo III SS	Cuadrado de la media	F-Valor	Pr > F	Adj Pr > F	
						G - G	H - F
dosis*flor	2	0.70000000	0.35000000	0.26	0.7745	0.7292	0.7659
Error(dosis*flor)	18	24.30000000	1.35000000				

Greenhouse-Geisser Epsilon	0.8088
Huynh-Feldt Epsilon	0.9601

En nuestro ejemplo solo tenemos efectos intrasujetos y como observamos en Figura 32 tanto el factor DOSIS como FLOR influyen en la variable respuesta pero su interacción no.

B) PROC MIXED

Figura 33. Lectura de datos para Ejemplo 3 con PROC MIXED.

```
data perfume2fMRmixed;
set perfume2fMRglm;
y = y1; dosis=1;flor=1; output;
y = y2; dosis=1;flor=2; output;
y = y3; dosis=1;flor=3; output;

y = y4; dosis=2;flor=1; output;
y = y5; dosis=2;flor=2; output;
y = y6; dosis=2;flor=3; output;
drop y1-y6;
run;
```

Para usar PROC MIXED tenemos que cambiar los datos de forma. En este caso al tener dos factores con medidas repetidas en ambos tendremos que crear dos nuevas columnas, una referida a la concentración del perfume (DOSIS) y otra a la variedad de flor (FLOR). La variable respuesta se tendrá en una sola columna Y. Podemos observar cómo hay que presentar los datos en Anexo 6.

Como siempre que utilizamos el procedimiento PROC MIXED, pasamos a evaluar qué estructura de la matriz de varianzas-covarianzas parece ser la más adecuada.

Figura 34. PROC MIXED con estructura de covarianza "simetría compuesta"

```
proc mixed data = perfume2fMRmixed;
class persona dosis flor;
model y = dosis flor dosis*flor;
repeated / subject=persona type=cs;
run;
```

Figura 35. PROC MIXED con estructura de covarianza "Tipo H"

```
proc mixed data = perfume2fMRmixed;
class persona dosis flor;
model y = dosis flor dosis*flor;
repeated / subject=persona type=hf;
run;
```

Figura 36. PROC MIXED con estructura de covarianza "No estructurada"

```
proc mixed data = perfume2fMRmixed;
class persona dosis flor;
model y = dosis flor dosis*flor;
repeated / subject=persona type=un;
run;
```

Para ello tenemos que fijarnos en los índices AIC y BIC que aparecen a continuación para cada tipo de estructura.

Figura 37. Estadísticos de ajuste para la estructura de covarianza "simetría compuesta"

Estadísticos de ajuste	
Verosimilitud -2 Res Log	180.1
AIC (Smaller is Better)	184.1
AICC (Smaller is Better)	184.4
BIC (Smaller is Better)	184.7

Figura 38. Estadísticos de ajuste para la estructura de covarianza "Tipo H"

Estadísticos de ajuste	
Verosimilitud -2 Res Log	170.8
AIC (Smaller is Better)	184.8
AICC (Smaller is Better)	187.3
BIC (Smaller is Better)	187.0

Figura 39. Estadísticos de ajuste para la estructura de covarianza "No estructurada"

Estadísticos de ajuste	
Verosimilitud -2 Res Log	149.4
AIC (Smaller is Better)	191.4
AICC (Smaller is Better)	220.3
BIC (Smaller is Better)	197.7

Como podemos observar, la estructura de covarianza "simetría compuesta" es la más adecuada ya que sus índices AIC y BIC son los más bajos de las 3 estructuras que hemos estudiado. Seguiremos el estudio con esta estructura, es decir, ejecutando el código de Figura 39.

Figura 40. Test de tipo III para contrastar las hipótesis nulas del modelo

Tests de tipo 3 de efectos fijos				
Efecto	Num DF	Den DF	F-Valor	Pr > F
dosis	1	9	16.84	0.0027
flor	2	18	7.13	0.0052
dosis*flor	2	18	0.27	0.7643

En la Figura 40 podemos observar los resultados de realizar el contraste de hipótesis para nuestro ejemplo. Teniendo 3 contrastes y siendo la hipótesis la igualdad de las medias.

Para los factores DOSIS y FLOR por separado vemos que su p-valor es significativo por lo que rechazamos la hipótesis nula y podemos decir que tanto la concentración como el tipo de flor afecta a la respuesta afectiva mientras que para la interacción esto no sucede ya que su p-valor es no significativo.

6. Otros ejemplos en SAS

En este apartado trataremos situaciones que no se han dado en los ejemplos vistos anteriormente como que se rechace la hipótesis de esfericidad (Ejemplo 4) o cómo PROC MIXED aprovecha los datos incompletos (Ejemplo 5).

Ejemplo 4:

“La Tabla 10 muestra las medidas repetidas de la cantidad total de colesterol medida cada 4 semanas durante 24 semanas para 12 sujetos tratados con el fármaco y 11 sujetos tratados con el placebo.” (S.Davis, Charles [3])

Tabla 10. Datos para el Ejemplo 4.

Grupo:	Persona	Semana					
		4	8	12	16	20	24
Droga	1	317	280	275	270	274	266
	2	186	189	190	135	197	205
	3	377	395	368	334	338	334
	4	229	258	282	272	264	265
	5	276	310	306	309	300	264
	6	272	250	250	255	228	250
	7	219	210	236	239	242	221
	8	260	245	264	268	317	314
	9	284	256	241	242	243	241
	10	365	304	294	287	311	302
	11	298	321	341	342	357	335
	12	274	245	262	263	235	246
Placebo	1	232	205	244	197	218	233
	2	367	354	358	333	338	355
	3	253	256	247	228	237	235
	4	230	218	245	215	230	207
	5	190	188	212	201	169	179
	6	290	263	291	312	299	279
	7	337	337	383	318	361	341
	8	283	279	277	264	269	271
	9	325	257	288	326	293	275
	10	266	258	253	284	245	263
	11	338	343	307	274	262	309

Como podemos observar, estamos ante un problema de análisis de medidas repetidas con dos factores (GRUPO y SEMANA) con medidas repetidas en uno de ellos (SEMANA).

Figura 41. Lectura de datos para Ejemplo 4.

```
data colesterol (keep = persona grupo y1-y6)
colesterol2 (keep=persona grupo semana y);
input persona grupo y1-y6;
output colesterol;
y=y1; semana =4; output colesterol2;
y=y2; semana =8; output colesterol2;
y=y3; semana =12; output colesterol2;
y=y4; semana =16; output colesterol2;
y=y5; semana =20; output colesterol2;
y=y6; semana =24; output colesterol2;
datalines;
1 1 317 280 275 270 274 266
2 1 186 189 190 135 197 205
3 1 377 395 368 334 338 334
4 1 229 258 282 272 264 265
5 1 276 310 306 309 300 264
6 1 272 250 250 255 228 250
7 1 219 210 236 239 242 221
8 1 260 245 264 268 317 314
9 1 284 256 241 242 243 241
10 1 365 304 294 287 311 302
11 1 298 321 341 342 357 335
12 1 274 245 262 263 235 246
13 2 232 205 244 197 218 233
14 2 367 354 358 333 338 355
15 2 253 256 247 228 237 235
16 2 230 218 245 215 230 207
17 2 190 188 212 201 169 179
18 2 290 263 291 312 299 279
19 2 337 337 383 318 361 341
20 2 283 279 277 264 269 271
21 2 325 257 288 326 293 275
22 2 266 258 253 284 245 263
23 2 338 343 307 274 262 309
;
```

A) PROC GLM

La lectura de datos y el código para analizar este ejemplo es parecido al del Ejemplo 2.

Figura 42. Código para analizar el Ejemplo 4 utilizando PROC GLM.

```
proc glm data=colesterol;
class grupo;
model y1-y6=grupo / nouni;
repeated semana 6 / printe;
run;
```

Figura 43. Test de esfericidad para el Ejemplo 4.

Tests de esfericidad				
Variables	DF	Criterio de Mauchly	Chi-cuadrado	Pr > ChiSq
Variables transformadas	14	0.102917	43.430191	<.0001
Componentes ortogonales	14	0.2091833	29.882799	0.0079

Nos fijamos en el p-valor de las componentes ortogonales y observamos que es significativo, por lo que rechazamos que la matriz de varianzas y covarianzas del factor con medidas repetidas SEMANA sea de Tipo H por consiguiente no podemos interpretar correctamente el valor de F.

Podemos fijarnos en los tests univariantes ajustados F (G-G o H-F) o los tests MANOVA (Wilks Lambda, Pillai's Trace, Hotelling-Lawley Trace, and Roy's Greatest Root).

Figura 44. Test F univariante para los efectos intrasujetos.

Procedimiento GLM							
Análisis de medidas repetidas de la varianza							
Test de hipótesis univariante para efectos Within Subject							
Fuente	DF	Tipo III SS	Cuadrado de la media	F-Valor	Pr > F	Adj Pr > F	
						G - G	H-F-L
semana	5	3584.36605	716.87321	1.81	0.1165	0.1471	0.1334
semana*grupo	5	923.67040	184.73408	0.47	0.7998	0.7252	0.7606
Error(semana)	105	41510.69192	395.33992				

Nos fijamos en el p-valor de los tests ajustados G-G y H-F-L del efecto intrasujeto. Este es no significativo tanto para el factor SEMANA como para su interacción con GRUPO lo que significa que no rechazamos la hipótesis nula de que estos no influyan en la variable respuesta.

Figura 45. Test F univariante para los efectos intersujetos.

Procedimiento GLM					
Análisis de medidas repetidas de la varianza					
Tests de hipótesis para efectos Between Subjects					
Fuente	DF	Tipo III SS	Cuadrado de la media	F-Valor	Pr > F
grupo	1	27.3625	27.3625	0.00	0.9656
Error	21	301560.8838	14360.0421		

Como podemos observar, el p-valor para el efecto intersujeto del factor GRUPO no es significativo por lo que no rechazamos la hipótesis nula y podemos deducir que no hay diferencia en cuanto al índice de colesterol dependiendo de si nos han suministrado el fármaco o el placebo.

B) PROC MIXED

Primero elegimos la estructura más adecuada para la matriz de varianzas y covarianzas.

Figura 46. PROC MIXED con estructura de covarianza "simetría compuesta".

```
proc mixed data=colesterol2;
class grupo semana persona;
model y = grupo|semana;
repeated / type=cs sub=persona;
run;
```

Figura 47. PROC MIXED con estructura de covarianza "Tipo H".

```
proc mixed data=colesterol2;
class grupo semana persona;
model y = grupo|semana;
repeated / type=hf sub=persona;
run;
```

Figura 48. PROC MIXED con estructura de covarianza "sin estructura".

```
proc mixed data=colesterol2;
class grupo semana persona;
model y = grupo|semana;
repeated / type=un sub=persona;
run;
```

Figura 49. Estadísticos de ajuste para la estructura de covarianza "simetría compuesta".

Estadísticos de ajuste	
Verosimilitud -2 Res Log	1215.8
AIC (Smaller is Better)	1219.8
AICC (Smaller is Better)	1219.9
BIC (Smaller is Better)	1222.0

Figura 50. Estadísticos de ajuste para la estructura de covarianza "Tipo H".

Estadísticos de ajuste	
Verosimilitud -2 Res Log	1212.8
AIC (Smaller is Better)	1226.8
AICC (Smaller is Better)	1227.8
BIC (Smaller is Better)	1234.7

Figura 51. Estadísticos de ajuste para la estructura de covarianza "sin estructura".

Estadísticos de ajuste	
Verosimilitud -2 Res Log	1179.9
AIC (Smaller is Better)	1221.9
AICC (Smaller is Better)	1230.8
BIC (Smaller is Better)	1245.8

Fijándonos en los índices AIC y BIC observamos que la estructura más adecuada es la de simetría compuesta

Figura 52. Test de tipo III para contrastar la hipótesis nula del modelo.

Tests de tipo 3 de efectos fijos				
Efecto	Num DF	Den DF	F-Valor	Pr > F
grupo	1	21	0.00	0.9656
semana	5	105	1.81	0.1165
grupo*semana	5	105	0.47	0.7998

Observando los tests para contrastar la hipótesis nula de no efecto observamos que tanto para los factores GRUPO y SEMANA como para la interacción el p-valor es no significativo y por tanto no rechazamos la hipótesis nula.

Ejemplo 5:

“ Soy famoso, sácame de aquí” es un programa de TV en el que personajes celebres en un intento lamentable de salvar sus carreras van a una selva y se someten a pruebas humillantes. Una de ellas es “ensayos de Bushtucker” donde los famosos comen cosas como insectos palo, larvas de coso, ojos de pez, testículos de canguro y muchas otras cosas.

Siempre me he preguntado cuál de estas comidas es la más repugnante. Así que cogí a 8 famosos y les hice comer 4 animales diferentes (insectos palo, larvas de coso, ojos de pez y testículos de canguro) en orden de equilibrio (counterbalanced). En cada ocasión medí el tiempo que tardaba el famoso en tener arcadas en segundos. (Discovering statistics: repeated measures ANOVA [19])

Para este ejemplo vamos a probar cómo los resultados cambian utilizando PROC GLM o PROC MIXED cuando tenemos datos missing.

Estamos ante un ejemplo de medidas repetidas en un factor. (Como en el ejemplo 1)

Tabla 11. Datos para el ejemplo 5 (completos).

Famoso	Insecto palo	Testículos de canguro	Ojo de pez	Larvas de coso
1	8	7	1	6
2	9	5	2	5
3	6	2	3	8
4	5	3	1	9
5	8	4	5	8
6	7	5	6	7
7	10	2	7	2
8	12	6	8	1

Para el conjunto de datos completo vamos a exponer solo el contraste de hipótesis relacionado con el efecto de los factores en la variable respuesta usando PROC GLM, ya que como vimos anteriormente para datos balanceados tanto con PROC GLM como con PROC MIXED se sacan las mismas conclusiones.

Figura 53. Test de esfericidad para el Ejemplo 5 (completos).

Tests de esfericidad				
Variables	DF	Criterio de Mauchly	Chi-cuadrado	Pr > ChiSq
Variables transformadas	5	0.0194624	22.541397	0.0004
Componentes ortogonales	5	0.136248	11.405981	0.0439

Para ello nos vamos a fijar en los test MANOVA ya que no se cumple la hipótesis del test de esfericidad

Figura 54. Test para contrastar la hipótesis nula del modelo.

Criterio de test MANOVA y estadísticos F exactos para la hipótesis de efecto no comida H = Tipo III Matriz SSCP para T. indecomida Matriz SSCP de error E = S=1 M=0.5 N=1.5					
Estadístico	Valor	F-Valor	Num DF	Den DF	Pr > F
Lambda de Wilks	0.05823014	26.96	3	5	0.0016
Traza de Pillai	0.94176986	26.96	3	5	0.0016
Traza Hotelling-Lawley	16.17323850	26.96	3	5	0.0016
Raíz más grande de Roy	16.17323850	26.96	3	5	0.0016

Como vemos el p-valor es significativo por lo que podemos decir que el tipo de “comida” que el concursante ingiere sí que influye en el tiempo que tarda este en tener una arcada.

Tabla 12. Datos para el ejemplo 5 (missing)

Famoso	Insecto palo	Testículos de canguro	Ojo de pez	Larvas de coso
1	8	7	1	6
2	9	5	2	5
3	6	2	3	8
4	5	3	1	9
5	8	4	5	-
6	7	5	6	7
7	10	2	7	2
8	12	6	8	1

A) PROC GLM

Como vimos en el punto 4.5 PROC GLM no utiliza los datos de ningún sujeto que tenga valores missing. Esto significa que para nuestro conjunto de datos PROC GLM tratará al sujeto 5 como si no apareciera.

Figura 55. Lectura de datos para el Ejemplo 5 (missing) usando PROC GLM.

```
data concurso;
input participante y1-y4;
datalines;
1 8 7 1 6
2 9 5 2 5
3 6 2 3 8
4 5 3 1 9
5 8 4 5 -
6 7 5 6 7
7 10 2 7 2
8 12 6 8 1
;
```

Figura 56. Código para analizar el Ejemplo 5 (missing) utilizando PROC GLM.

```
proc glm data = concurso;
model y1-y4= / nouni;
repeated comida 4 / printe;
run;
```

Figura 57. Test de esfericidad para el Ejemplo 5 (missing).

Tests de esfericidad				
Variables	DF	Criterio de Mauchly	Chi-cuadrado	Pr > ChiSq
Variables transformadas	5	0.0196048	18.567681	0.0023
Componentes ortogonales	5	0.1355991	9.4352479	0.0929

P-valor no significativo por lo tanto no rechazamos la hipótesis de que la matriz de varianzas y covarianzas sea de tipo H. Podremos utilizar el estadístico F.

Figura 58. Test F univariante para los efectos intrasujetos.

Procedimiento GLM							
Análisis de medidas repetidas de la varianza							
Test de hipótesis univariante para efectos Within Subject							
Fuente	DF	Tipo III SS	Cuadrado de la media	F-Valor	Pr > F	Adj Pr > F	
						G - G	H - F
comida	3	74.9642857	24.9880952	3.02	0.0566	0.1037	0.0848
Error(comida)	18	148.7857143	8.2658730				

Greenhouse-Geisser Epsilon	0.5261
Huynh-Feldt Epsilon	0.6821

Como observamos el p-valor del factor COMIDA es no significativo a nivel 0,05 por lo que no rechazamos la hipótesis nula y podemos decir que el tipo de “comida” que se le da al concursante no afecta al tiempo que tarda este en que le dé una arcada.

B) PROC MIXED

Fijándonos otra vez en el punto 4.5 vimos que con un conjunto de datos donde hay observaciones missing PROC MIXED utiliza el conjunto de datos completo ya que usa una estimación basada en la verosimilitud.

Figura 59. Lectura de datos para el Ejemplo 5 (missing) usando PROC MIXED.

```
data concurso2 (keep=participante comida y);
input participante y1-y4;
y=y1; comida =1; output concurso2;
y=y2; comida =2; output concurso2;
y=y3; comida =3; output concurso2;
y=y4; comida =4; output concurso2;
datalines;
1 8 7 1 6
2 9 5 2 5
3 6 2 3 8
4 5 3 1 9
5 8 4 5 -
6 7 5 6 7
7 10 2 7 2
8 12 6 8 1
;
```

Primero seleccionamos la estructura de la matriz de varianzas y covarianzas.

Figura 60. PROC MIXED con estructura de covarianza "simetría compuesta".

```
proc mixed data = concurso2;
class comida participante;
model y = comida;
repeated / type = cs sub = participante;
run;
```

Figura 61. PROC MIXED con estructura de covarianza "Tipo H".

```
proc mixed data = concurso2;
class comida participante;
model y = comida;
repeated / type = hf sub = participante;
run;
```

Figura 62. PROC MIXED con estructura de covarianza "sin estructura".

```
proc mixed data = concurso2;  
class comida participante;  
model y = comida;  
repeated / type = un sub = participante;  
run;
```

Nos fijamos en los índices AIC y BIC y la estructura bajo la cual tome los valores más bajos será la elegida.

Figura 63. Estadísticos de ajuste para la estructura de covarianza "simetría compuesta".

Estadísticos de ajuste	
Verosimilitud -2 Res Log	131.0
AIC (Smaller is Better)	135.0
AICC (Smaller is Better)	135.5
BIC (Smaller is Better)	135.1

Figura 64. Estadísticos de ajuste para la estructura de covarianza "Tipo H".

Estadísticos de ajuste	
Verosimilitud -2 Res Log	112.7
AIC (Smaller is Better)	122.7
AICC (Smaller is Better)	125.5
BIC (Smaller is Better)	123.0

Figura 65. Estadísticos de ajuste para la estructura de covarianza "sin estructura".

Estadísticos de ajuste	
Verosimilitud -2 Res Log	98.8
AIC (Smaller is Better)	118.8
AICC (Smaller is Better)	132.6
BIC (Smaller is Better)	119.6

Seleccionamos la estructura "sin estructura" para la matriz de varianzas y covarianzas y nos fijamos en el contraste de hipótesis.

Figura 66. Test de tipo III para contrastar la hipótesis nula del modelo.

Tests de tipo 3 de efectos fijos				
Efecto	Num DF	Den DF	F-Valor	Pr > F
comida	3	7	35.74	0.0001

Como observamos en la Figura 66 el p-valor es significativo por lo que, al contrario que cuando usamos PROC GLM, podemos decir que el tipo de “comida” que se le da al concursante para ingerir influye en el tiempo que este tarda en tener una arcada. Esto se debe a que mientras que PROC MIXED no necesita que sus datos estén balanceados, PROC GLM elimina aquellos sujetos donde se encuentra alguna observación missing lo que supone una pérdida de información. Las conclusiones que sacamos con PROC MIXED para los datos con observaciones missing son las mismas que cuando usamos datos balanceados.

7. Conclusiones

El análisis de medidas repetidas es un diseño muy utilizado en ciencias sociales. Resulta muy útil cuando queremos economizar el número de sujetos al que preguntar, pero para poder utilizarlo debemos tener en cuenta los posibles problemas como la fatiga, los datos no balanceados o el efecto residual. Además, tenemos que comprobar que nuestros datos cumplan los supuestos de independencia, normalidad y esfericidad.

Hemos visto que estos datos pueden ser analizados con SAS y que deberemos tener en cuenta que PROC GLM utiliza solo a los individuos cuyos datos están completos y sin embargo PROC MIXED aprovecha toda la información ya que usa una estimación basada en la verosimilitud.

8. Posibles ampliaciones

A lo largo del desarrollo de este trabajo han ido surgiendo ciertos posibles temas de ampliación que por la limitación de tiempo no se han podido desarrollar.

Comparaciones múltiples:

Cuando rechazamos la hipótesis nula del modelo, parece razonable estudiar entre qué medias existen diferencias.

Comparaciones de tendencia:

“Si los J niveles de la variable independiente analizada son cuantitativos y se ha rechazado la hipótesis nula del ANOVA, puede resultar interesante conocer el tipo de relación que existe entre la variable dependiente e independiente.” (Ximénez, Carmen y San Martín, Rafael [2])

R:

Tanto SAS como R son dos de los programas más utilizados en los estudios estadísticos y ambos se han tratado en diferentes asignaturas a lo largo de la carrera. La ventaja de R es que este es libre y por lo tanto cualquier persona puede acceder a ello. Mi primera intención era realizar todos los ejemplos en ambos para compararlos y sin duda me parece interesante la aplicación de estos métodos en R.

9. Glosario de vocabulario inglés-español

Unbalanced design = diseño no balanceado

Counterbalancing = contrapeso

Attrition = desgaste

Carry-over = arrastre

Carry-over effect = efecto residual

Compound symmetry = simetría compuesta

Score = valor/observación

Working mean model = modelo útil de la media

Powerful = potente

Between-subject = intersujeto

Within-subject = intrasujeto

10. Anexo

Anexo 1. Datos para Ejemplo 1. PROC GLM

	persona	y1	y2	y3	y4
1	1	1	3	4	4
2	2	2	2	5	6
3	3	2	5	4	6
4	4	2	3	5	4
5	5	1	3	4	7
6	6	1	4	4	5
7	7	2	3	5	4
8	8	4	3	4	5
9	9	2	4	4	6
10	10	2	2	5	6
11	11	2	3	3	3
12	12	1	4	1	7
13	13	4	2	5	4
14	14	1	3	6	5
15	15	2	2	5	5
16	16	1	3	5	3
17	17	1	2	3	6
18	18	3	4	4	4
19	19	3	4	7	4
20	20	3	2	4	3

Anexo 2. Datos para Ejemplo 1. PROC MIXED

	persona	y	dosis		persona	y	dosis
2	1	3	1000	34	9	4	1000
3	1	4	1500	35	9	4	1500
4	1	4	2000	36	9	6	2000
5	2	2	500	37	10	2	500
6	2	2	1000	38	10	2	1000
7	2	5	1500	39	10	5	1500
8	2	6	2000	40	10	6	2000
9	3	2	500	41	11	2	500
10	3	5	1000	42	11	3	1000
11	3	4	1500	43	11	3	1500
12	3	6	2000	44	11	3	2000
13	4	2	500	45	12	1	500
14	4	3	1000	46	12	4	1000
15	4	5	1500	47	12	1	1500
16	4	4	2000	48	12	7	2000
17	5	1	500	49	13	4	500
18	5	3	1000	50	13	2	1000
19	5	4	1500	51	13	5	1500
20	5	7	2000	52	13	4	2000
21	6	1	500	53	14	1	500
22	6	4	1000	54	14	3	1000
23	6	4	1500	55	14	6	1500
24	6	5	2000	56	14	5	2000
25	7	2	500	57	15	2	500
26	7	3	1000	58	15	2	1000
27	7	5	1500	59	15	5	1500
28	7	4	2000	60	15	5	2000
29	8	4	500	61	16	1	500
30	8	3	1000	62	16	3	1000
31	8	4	1500	63	16	5	1500
32	8	5	2000	64	16	3	2000
33	9	2	500	65	17	1	500
66	17	2	1000				
67	17	3	1500				
68	17	6	2000				
69	18	3	500				
70	18	4	1000				
71	18	4	1500				
72	18	4	2000				
73	19	3	500				
74	19	4	1000				
75	19	7	1500				
76	19	4	2000				
77	20	3	500				
78	20	2	1000				
79	20	4	1500				
80	20	3	2000				

Anexo 3. Datos para Ejemplo 2. PROC GLM

	persona	genero	y1	y2	y3	y4
1	1	H	1	3	4	4
2	2	H	2	2	5	6
3	3	H	1	4	1	7
4	4	H	1	4	4	5
5	5	H	3	2	4	3
6	6	H	1	2	3	6
7	7	H	2	2	5	5
8	8	H	1	3	5	3
9	9	H	2	3	3	3
10	10	H	2	2	5	6
11	11	M	2	5	4	6
12	12	M	4	2	5	4
13	13	M	1	3	4	7
14	14	M	2	3	5	4
15	15	M	1	3	6	5
16	16	M	2	3	5	4
17	17	M	3	4	4	4
18	18	M	3	4	7	4
19	19	M	4	3	4	5
20	20	M	2	4	4	6

Anexo 4. Datos para Ejemplo 2. PROC MIXED

	persona	genero	y	dosis		persona	genero	y	dosis
1	1	H	1	500	33	9	H	2	500
2	1	H	3	1000	34	9	H	3	1000
3	1	H	4	1500	35	9	H	3	1500
4	1	H	4	2000	36	9	H	3	2000
5	2	H	2	500	37	10	H	2	500
6	2	H	2	1000	38	10	H	2	1000
7	2	H	5	1500	39	10	H	5	1500
8	2	H	6	2000	40	10	H	6	2000
9	3	H	1	500	41	11	M	2	500
10	3	H	4	1000	42	11	M	5	1000
11	3	H	1	1500	43	11	M	4	1500
12	3	H	7	2000	44	11	M	6	2000
13	4	H	1	500	45	12	M	4	500
14	4	H	4	1000	46	12	M	2	1000
15	4	H	4	1500	47	12	M	5	1500
16	4	H	5	2000	48	12	M	4	2000
17	5	H	3	500	49	13	M	1	500
18	5	H	2	1000	50	13	M	3	1000
19	5	H	4	1500	51	13	M	4	1500
20	5	H	3	2000	52	13	M	7	2000
21	6	H	1	500	53	14	M	2	500
22	6	H	2	1000	54	14	M	3	1000
23	6	H	3	1500	55	14	M	5	1500
24	6	H	6	2000	56	14	M	4	2000
25	7	H	2	500	57	15	M	1	500
26	7	H	2	1000	58	15	M	3	1000
27	7	H	5	1500	59	15	M	6	1500
28	7	H	5	2000	60	15	M	5	2000
29	8	H	1	500	61	16	M	2	500
30	8	H	3	1000	62	16	M	3	1000
31	8	H	5	1500	63	16	M	5	1500
32	8	H	3	2000	64	16	M	4	2000

65	17	M	3	500
66	17	M	4	1000
67	17	M	4	1500
68	17	M	4	2000
69	18	M	3	500
70	18	M	4	1000
71	18	M	7	1500
72	18	M	4	2000
73	19	M	4	500
74	19	M	3	1000
75	19	M	4	1500
76	19	M	5	2000
77	20	M	2	500
78	20	M	4	1000
79	20	M	4	1500
80	20	M	6	2000

Anexo 5. Datos para Ejemplo 3. PROC GLM

	persona	y1	y2	y3	y4	y5	y6
1	1	3	4	5	5	6	5
2	2	4	4	4	6	7	4
3	3	4	6	7	4	6	7
4	4	4	4	5	5	7	7
5	5	2	4	5	5	4	7
6	6	2	3	7	5	6	7
7	7	4	3	5	4	7	6
8	8	6	6	5	6	6	6
9	9	4	6	6	7	3	7
10	10	4	5	4	4	6	6

Anexo 6. Datos para Ejemplo 3. PROC MIXED

	persona	y	dosis	flor
1	1	3	1	1
2	1	4	1	2
3	1	5	1	3
4	1	5	2	1
5	1	6	2	2
6	1	5	2	3
7	2	4	1	1
8	2	4	1	2
9	2	4	1	3
10	2	6	2	1
11	2	7	2	2
12	2	4	2	3
13	3	4	1	1
14	3	6	1	2
15	3	7	1	3
16	3	4	2	1
17	3	6	2	2
18	3	7	2	3
19	4	4	1	1
20	4	4	1	2
21	4	5	1	3
22	4	5	2	1
23	4	7	2	2
24	4	7	2	3
25	5	2	1	1
26	5	4	1	2
27	5	5	1	3
28	5	5	2	1
29	5	4	2	2
30	5	7	2	3
31	6	2	1	1
32	6	3	1	2
33	6	7	1	3
34	6	5	2	1
35	6	6	2	2
36	6	7	2	3
37	7	4	1	1
38	7	3	1	2
39	7	5	1	3
40	7	4	2	1
41	7	7	2	2
42	7	6	2	3
43	8	6	1	1
44	8	6	1	2
45	8	5	1	3
46	8	6	2	1
47	8	6	2	2
48	8	6	2	3
49	9	4	1	1
50	9	6	1	2
51	9	6	1	3
52	9	7	2	1
53	9	3	2	2
54	9	7	2	3
55	10	4	1	1
56	10	5	1	2
57	10	4	1	3
58	10	4	2	1
59	10	6	2	2
60	10	6	2	3

11. Bibliografía y enlaces

- [1]. R. Girden, Ellen: *ANOVA Repeated Measures*, EEUU, SAGE, 1992
- [2]. Ximénez, Carmen y San Martín, Rafael: *Análisis de varianza con medidas repetidas*, La Muralla, Madrid, 2000
- [3]. S. Davis, Charles: *Statistical Methods for the Analysis of Repeated Measurements*, San Diego, Springer, 2002
- [4]. Nombre de los cursos académicos en USA <http://tic-y-docencia.blogspot.com/2012/03/nombre-de-los-cursos-academicos-en-usa.html>
- [5]. Universitat de Barcelona: Medidas Repetidas
http://www.ub.edu/deia/Temari_archivos/8_aplicats_classe.ppt
- [6]. HyperStat Online Contents: Carryover Effects <http://davidmlane.com/hyperstat/A134100.html>
- [7]. Explorable: Diseño de medidas con contrapeso <https://explorable.com/es/disenno-de-medidas-con-contrapeso>
- [8]. Laerd Statistics: Sphericity <https://statistics.laerd.com/statistical-guides/sphericity-statistical-guide.php>
- [9]. Biostat: test for sphericity <https://biostats.w.uib.no/test-for-sphericity-mauchly-test/>
- [10]. R Tutorial Series: One-Way Repeated Measures ANOVA
<https://www.r-bloggers.com/r-tutorial-series-one-way-repeated-measures-anova/>
- [11]. R - One Way Repeated Measures ANOVA Lecture 1
<https://www.youtube.com/watch?v=vl4nPUxhQTK>
- [12]. One Way Within-participant or Repeated-measures Analysis of Variance, Balanced Designs
http://psych.wisc.edu/moore/Rpdf/610-R8_OneWayWithin.pdf
- [13]. Wolfinger, Ross y Chang, Ming: *Comparing the SAS GLM and MIXED Procedures for Repeated Measures* <https://stats.idre.ucla.edu/wp-content/uploads/2016/02/mixedglm.pdf>
- [14]. Abebe, Asheber: *Introduction to Design and Analysis of Experiments with the SAS System*
<http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/Disenno/CursoDisExpSAS.pdf>

[15]. Wikipedia: Repeated measures design

https://en.wikipedia.org/wiki/Repeated_measures_design

[16]. Mixed-Up Mixed Models: Things That Look Like They Should Work But Don't, and Things That Look Like They Shouldn't Work But Do

<https://support.sas.com/resources/papers/proceedings/proceedings/sugi25/25/aa/25p020.pdf>

[17]. Comparison of PROC MIXED and PROC GLM for Analysis of Repeated Measures Data

https://www.researchgate.net/profile/Mohamed_Mourad_Lafifi/post/GLM_or_ANCOVA_in_multivariate_repeated_within_and_between_analysis/attachment/5acc9ba2b53d2f63c3c4edcd/AS%3A613824084340736%401523358474555/download/Comparison+of+PROC+MIXED+and+PROC+GLM+for+Analysis+of+Repeated+Measures+Data+SD06-Yarandi.pdf

[18]. Testing for sphericity in PROC GLM <http://support.sas.com/kb/22/589.html>

[19]. Discovering statistics: repeated measures ANOVA

<https://www.discoveringstatistics.com/docs/repeatedmeasures.pdf>

[20]. Example 39.7 Repeated Measures Analysis of Variance

https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_lm_sect052.htm

[21]. The Effect of Missing Data on Repeated Measures Models

<https://support.sas.com/resources/papers/proceedings/proceedings/sugi24/Stats/p262-24.pdf>

[22]. PROC GLM Contrasted with Other SAS Procedures

https://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_lm_sect003.htm

[23]. Wikipedia: Mauchly's test https://en.wikipedia.org/wiki/Mauchly%27s_sphericity_test