



UNIVERSIDAD
DE MÁLAGA

TESIS DOCTORAL

**Aplicación del modelado estadístico y la detección de señales en memoria de
reconocimiento y potenciales evocados.**

Por

Agustín Wallace Ruiz

Dirigida por: Dr. Manuel Pelegrina del Rio

Programa de Doctorado en Psicología Cognitiva

Facultad de Psicología de Málaga

Departamento de Psicobiología y Metodología de las CC. Del Comportamiento

Universidad de Málaga

2015



Publicaciones y
Divulgación Científica

AUTOR: Agustín Wallace Ruiz

 <http://orcid.org/0000-0003-1437-0128>

EDITA: Publicaciones y Divulgación Científica. Universidad de Málaga



Esta obra está bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional:

<http://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>

Cualquier parte de esta obra se puede reproducir sin autorización pero con el reconocimiento y atribución de los autores.

No se puede hacer uso comercial de la obra y no se puede alterar, transformar o hacer obras derivadas.

Esta Tesis Doctoral está depositada en el Repositorio Institucional de la Universidad de Málaga (RIUMA): riuma.uma.es

EL DOCTOR MANUEL PELEGRINA DEL RÍO, PROFESOR TITULAR DEL
DEPARTAMENTO DE PSICOBIOLOGÍA Y METODOLOGÍA DE LAS CIENCIAS
DEL COMPORTAMIENTO DE LA UNIVERSIDAD DE MÁLAGA

Certifica:

Que la tesis doctoral realizada por D. Agustín Wallace Ruiz con el título:
**“Aplicación del modelado estadístico y la detección de señales en memoria de
reconocimiento y potenciales evocados”**, de la cual soy director, ha sido proyectada,
desarrollada y redactada bajo mi supervisión.

Que el mencionado trabajo de investigación reúne todas las características científicas y
técnicas para poder ser defendido públicamente. Asimismo merece nuestra valoración
en cuanto a actualidad y planteamiento metodológico.

De todo lo cual informo, como trámite preceptivo para su aceptación y posterior defensa
pública.

Málaga a 13 de febrero de 2015

Firmado: Manuel Pelegrina del Río

DEDICATORIA

A las personas más especiales de mi vida,
mi mujer Angelines y mi hija Alicia,
por su amor, compañía, paciencia y
apoyo incondicional durante estos años

AGRADECIMIENTOS

Son muchas las personas a las que quiero agradecer y dedicar esta tesis doctoral, tanto a las que han contribuido directamente con sus comentarios, consejos, enseñanzas o ayuda, así como las que me han apoyado, hecho pasar buenos momentos y dado su cariño a lo largo de los años de realización de esta tesis doctoral.

En primer lugar quiero agradecer al Dr. Manuel Pelegrina por su ayuda, paciencia, consejos y apoyo, y por haberme iniciado en mi vida académica, y su inestimable ayuda para finalizar este trabajo.

A mi familia: mis padres Pilar y Agustín, mis hermanos M^aIsabel, Laura y Jorge por su aportación emocional y apoyo incondicional.

A mis compañeros de la Facultad de Psicología de Málaga, y especialmente a los del departamento de Psicobiología y Metodología de la C.C. del Comportamiento por su apoyo, ánimo, y ser excelentes compañeros/as de trabajo.

A los alumnos de la Facultad de Psicología de Málaga que han contribuido con su “cerebro” y ayuda en el laboratorio a que esta tesis pudiese ser finalizada.

INDICE

PRIMERA PARTE: ENFOQUE TEÓRICO

1 – Introducción	1
1.1 Antecedentes históricos	2
1.2 Antecedentes históricos recientes	5
1.3 El presente trabajo	13
2 – El Modelado Estadístico	15
2.1 ¿Qué es el modelado estadístico?	16
2.2 ¿Cuál es el objetivo del modelado estadístico?	17
2.3 Aspectos Fundamentales del modelado estadístico	17
2.4 Etapas del modelado estadístico	19
2.5 Ventajas del modelado estadístico	22
2.6 Sistemas de codificación para variables categóricas	22
2.7 Selección del modelo	27
2.7.1 Selección de modelos conducida por el investigador	28
2.7.2 Selección de modelos mediante algoritmos automáticos	31
2.8 Concluyendo	35
3 – El Modelo de Regresión Lineal	36
3.1 Introducción	37
3.2 El modelo de regresión lineal simple	39
3.2.1 ¿Cómo se ajusta la recta de regresión?	41
3.2.2 Estimación de los parámetros del modelo	42
3.2.3 Ajuste y selección del modelo	44

3.3 Modelo de regresión lineal múltiple	48
3.3.1 Ajuste y selección del modelo con predictor polinómico	50
3.3.2 Ajuste y selección del modelo con dos predictores	53
3.3.3 Comparación de modelos	55
3.3.4 Modelos con interacción entre variables	56
3.3.5 Cálculo de los intervalos de confianza con interacciones	58
3.4 La Colinealidad	59
3.5 Problemas y soluciones a la colinealidad	61
3.6 Modelos con medidas repetidas	63
3.6.1 Diseños con medidas repetidas o apareadas	64
3.6.2 Diseños de muestras relacionadas	65
3.6.3 Diseños mixtos	67
3.7 Concluyendo	68
4 – Introducción a los Modelos Lineales Generalizados (MLG)	70
4.1 Introducción	71
4.2 Componentes del MLG	72
4.3 Clasificación de los modelos	75
4.4 Supuestos del MLG	77
4.5 Estimación de parámetros	77
4.6 Ajuste y selección del modelo	79
4.6.1 Modelos	79
4.6.2 Ajustando modelos	80
4.6.3 La selección del modelo	82
4.7 Interpretando el modelo	83

4.8 Concluyendo	84
5 – Modelos Loglineales	85
5.1 Introducción	86
5.2 Modelos loglineales con dos variables	90
5.3 Modelos loglineales con tres variables	93
5.4 Modelos loglineales con más de tres variables	99
5.5 Modelos loglineales con datos relacionados	100
5.5.1 Modelos de simetría	101
5.5.2 Modelos diagonales	103
5.5.3 Modelos de bloques	104
6 – Regresión de Poisson	106
6.1 Introducción	107
6.2 El modelo de regresión de Poisson simple	109
6.2.1 La variable de exposición	110
6.2.2 Estimación de los parámetros	110
6.2.3 Comparación de modelos	111
6.2.4 Modelo con variable explicativa binaria	112
6.3 El modelo de regresión de Poisson múltiple	112
7 – Regresión Logística	117
7.1 Introducción	118
7.2 Modelo de regresión logística simple	121
7.2.1 Estimación de los parámetros del modelo	122

7.2.2	Inferencias sobre los coeficientes de regresión	124
7.2.3	Comparación de modelos	125
7.2.4	Bondad de ajuste	125
7.2.5	Modelo con variable explicativa binaria	127
7.3	Modelo de regresión logística múltiple	128
7.3.1	Comparación de modelos	129
8	– Regresión Logit	130
8.1	Introducción	139
8.2	Modelo de regresión logit simple	132
8.2.1	Estimación de los parámetros del modelo	135
8.2.2	Comparación de modelos	136
8.3	Modelo de regresión logit múltiple	138
8.3.1	Modelo con una variable predictiva politómica	138
8.3.1.1	Estimación de los parámetros del modelo	139
8.3.1.2	Comparación de modelos	140
8.3.2	Modelo con varias variables predictoras	140
8.3.2.1	Introducción	140
8.3.2.2	Comparación de modelos	142
8.4	Equivalencia entre modelos logit y loglineales	146
9	– Evaluación de los Modelos Ajustados	148
9.1	Técnicas de evaluación de un modelo ajustado	147
9.2	Test diagnósticos	149
9.2.1	Leverages o valores de influencia	149

9.2.2	Distancia de Mahalanobis	153
9.2.3	Distancia de Cook	154
9.2.4	Índice DFFITS	154
9.2.5	Índice DFBETAS	155
9.3	Residuales	155
9.3.1	Residual ordinario	155
9.3.2	Residual de Pearson	156
9.3.3	Residual de Pearson estudentizado	156
9.3.4	Residual eliminado	156
9.3.5	Residual parcial	157
9.3.6	Residuales de discrepancia	157
9.3.7	Residuales de discrepancia estudentizados	157
9.3.8	Residuales de verosimilitud	157
9.3.9	Residuales “Jackknife”	158
9.4	Posibles causas de falta de validez de un modelo ajustado	158
9.5	La sobredispersión	159
9.6	Evaluación de la fiabilidad de un modelo predictivo	160
10	–Memoria de Reconocimiento	161
10.1	Introducción	162
10.2	Modelos teóricos de doble procesamiento	165
10.2.1	El modelo de Atkinson	166
10.2.2	Teoría de Detección de Señales (TDS)	166
10.2.3	Modelos Neuroanatómicos	167
10.2.4	El modelo de Mandler	168

10.2.5	El modelo de Jacoby	169
10.2.6	El Modelo de Tulving	170
10.2.7	El modelo de Yonelinas	171
10.3	Técnicas de medida	172
10.3.1	Técnicas de disociación de la tarea y de respuesta-velocidad	173
10.3.2	Técnicas de recuperación/reconocimiento	174
10.3.3	Técnicas de reconocimiento ítem/asociativo	175
10.3.4	Técnicas de estimación del proceso	176
10.3.5	El proceso recordar/conocer	177
10.3.6	El procedimiento ROC	181
10.4	Consideraciones sobre las teorías	182
10.5	La Teoría de Detección de Señales	185
10.5.1	Introducción	185
10.5.2	Detectando la señal	187
10.5.3	Diferenciando la señal del ruido	187
10.5.4	Emitiendo la respuesta	190
10.5.5	Curvas ROC	192
10.5.6	Concluyendo	194
10.6	Cronometría mental	194
10.7	Potenciales Evocados	195
10.7.1	Introducción	195
10.7.2	El registro de potenciales evocados	201

SEGUNDA PARTE: ESTUDIO EMPÍRICO

Experimento 1: Índices de detección de señales y modelado estadístico 205

en memoria de reconocimiento

1- Introducción	206
2- Método	212
Participantes	
Material	
Procedimiento	
3- Resultados	213
4- Discusión	217

Experimento 2: Modelado estadístico aplicado a memoria de 225

reconocimiento y potenciales evocados

1- Introducción	226
2- Método	228
Participantes	
Material	
Procedimiento	
3- Resultados	231
4- Discusión	233

Experimento 3: Modelado estadístico aplicado a memoria de 236

reconocimiento y tiempo de reacción

1- Introducción	237
2- Método	238
Participantes	
Material	

Procedimiento	
3- Resultados	239
4- Discusión	241
DISCUSIÓN GENERAL	244
CONCLUSIÓN	255
REFERENCIAS	259

PRIMERA PARTE

ENFOQUE TEÓRICO

CAPÍTULO 1 INTRODUCCIÓN

I am surely not alone in having suspected that some of Fisher's major views were adopted simply to avoid agreeing with his opponents. Leonard Savage.

1.1 ANTECEDENTES HISTÓRICOS

El intento de dar explicación a todo aquello que aparece en nuestro entorno, es algo inherente al ser humano, de este modo, la estadística ha proporcionado el marco conceptual que nos permite explicar¹ y predecir muchos de los fenómenos que nos rodean. Inicialmente egipcios, babilonios, árabes, griegos y romanos impulsaron esta ciencia como una forma de mejorar y controlar su extensa civilización, pero cayó en desuso tras la caída del imperio romano; no obstante, de nuevo a finales de la Edad Media vuelve a recobrar importancia.

Fue el matemático italiano, Leonardo de Pisa (conocido como Fibonacci) 1170-1250 la persona que revolucionó las matemáticas en Europa. Hasta entonces se seguía utilizando el sistema de numeración romana, muy sencillo de utilizar para sumas y restas; pero supone todo un reto en operaciones más complejas. La aportación principal de este matemático fue la introducción de la numeración árabe (de base 10) y el cero como dígito de valor nulo. Pero quizás, por lo que suele ser más conocido es por la secuencia de Fibonacci (1,1,2,3,5,8,13,21,34,55). En esta secuencia, cada número a partir del tercero, se obtiene sumando los dos que le preceden (ver figura 1.1).

¹ En este trabajo se ha utilizado la palabra explicar, o desde aquí explicativo, explicativa...de una manera amplia. Los diseños en el ámbito del modelado han “confundido” explicación con variable causal o predictiva. Mientras en filosofía de la ciencia el término explicar va siempre unido a una teoría, nunca a una variable. Porque, según la filosofía de la ciencia, ninguna variable, sin una teoría, puede explicar nada por sí sola.

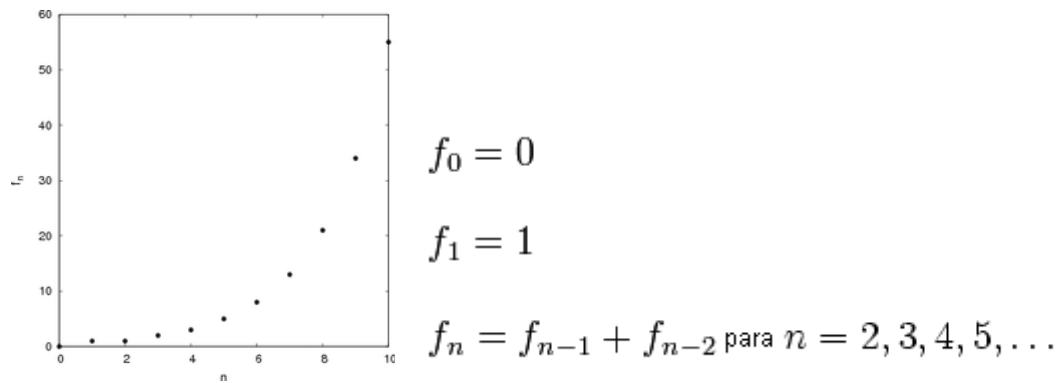


Figura 1.1 Gráfica de la secuencia de Fibonacci y sus ecuaciones

Donde una secuencia aparentemente aleatoria de números, sigue un patrón determinado por una función. Es interesante señalar también la relación de esta secuencia con el número $\phi = 1,61803\dots$, llamado número de oro y que representa la letra inicial en griego del nombre del escultor griego Fidias que lo utilizó en sus obras.

Este es el primer y más importante de los objetivos generales que pretendemos conseguir mediante esta tesis, buscar funciones que nos permitan encontrar patrones donde aparentemente solo existe el caos.

Más adelante, en el renacimiento se vive un gran furor por la estadística, y se producen avances muy importantes (notación, logaritmos, signos, etc.). Así, alrededor del año 1540, el alemán Muster realizó una compilación estadística de los recursos nacionales, que comprendía datos acerca de la organización política, instrucciones sociales, comercio y poderío militar. Durante el siglo XVII se aportaron indicaciones más concretas sobre los métodos de observación y análisis cuantitativo y se ampliaron los campos de la inferencia y la teoría estadística.

Por otra parte, la teoría de probabilidad tiene su origen en los juegos de azar. Hacia 1650, en Francia, un jugador llamado De Mére consultó al matemático Blaise Pascal sobre algunas cuestiones relacionadas con el juego de dados. Pascal mantuvo correspondencia con Fermat, Huygens y Bernoulli. Gracias a todos ellos, la teoría de la probabilidad pasó de ser una mera colección de problemas aislados, relativos a algunos juegos, a ser un sector importante de las matemáticas.

Pero fue Achenwall quien acuñó en 1760 la palabra estadística, que extrajo del término italiano *statista* (estadista). Creía, que los datos de la nueva ciencia serían el aliado más eficaz del gobernante consciente. La raíz etimológica de la palabra se halla, por otra parte, en el término latino *status*, que significa estado o situación. Esta etimología aumenta el valor intrínseco de la palabra, por cuanto la estadística revela el sentido cuantitativo de las más variadas situaciones.

Durante el siglo XVIII empieza el auge de la estadística descriptiva en asuntos sociales y económicos, y es a finales de ese siglo y comienzos del XIX cuando se comienzan a asentar verdaderamente las bases teóricas de la teoría de probabilidades con los trabajos de Lagrange (1770) y Laplace (1780), Gauss (1801), y de Poisson (1837). Previamente, cabe destacar el descubrimiento de la distribución normal por de Moivre (1738), distribución que será posteriormente “redescubierta” por Gauss (1801) y Poisson (1837).

No obstante, fue Quételet (1835) quien aplica la Estadística a las ciencias sociales, interpretando la teoría de la probabilidad para su uso en las ciencias sociales y así, resolver la aplicación del principio de promedios y de la variabilidad a los

fenómenos sociales. Quételect (1835) fue el primero en realizar la aplicación práctica del “método” Estadístico a las diversas ramas de la ciencia.

Entretanto, en el período del 1800 al 1820 se desarrollaron dos conceptos matemáticos fundamentales para la teoría Estadística; la teoría de los errores de observación, aportada por Laplace y Gauss; y la teoría de los mínimos cuadrados desarrollada por Laplace, Gauss y Legendre. A finales del siglo XIX, Galton (1888) ideó el procedimiento conocido por correlación, que tenía por objeto medir la influencia relativa de los factores sobre las variables. De aquí partió el desarrollo del coeficiente de correlación momento-producto creado por Pearson (1895), así como el importante concepto de la desviación estándar (Pearson, 1893). Por su parte, Yule (1907), en fechas muy próximas a las citadas realizaba amplios estudios sobre la medida de las correlaciones (Yule, 1897).

1.2 ANTECEDENTES HISTÓRICOS RECIENTES

El análisis de los datos categóricos hasta fechas recientes ha estado limitado al análisis de tablas de contingencia bidimensionales con el estadístico χ^2 de Pearson. Esto es debido a la controversia que surgió a principios de siglo entre Pearson y Yule.

Así, el año 1900 es el punto de comienzo del análisis de datos categóricos desde que Pearson (1900) introdujo el estadístico χ^2 y Yule (1911) presentó la razón de productos cruzados (*odds ratio*). En 1900 Pearson era muy conocido por la comunidad estadística, estaba a la cabeza del laboratorio estadístico del University College in London, y su trabajo en la década precedente incluía el desarrollo de una familia de

distribuciones de probabilidad sesgada (llamadas curvas de Pearson), obteniendo la estimación de los coeficientes de correlación, encontrando el error estándar y extendiendo el trabajo de Galton (1888) sobre la regresión lineal. El trabajo de Pearson (1900) tenía por objeto desarrollar la prueba ji-cuadrado e incluía tanto los resultados de la ruleta del casino de Monte Carlo (que variaba aleatoriamente) como las pruebas de ajuste de series de datos, las denominadas curvas de Pearson, y las pruebas estadísticas de independencia incluidas en las tablas de contingencia bidimensionales. En 1904, Pearson introdujo el término contingencia como una “*measure of the total deviation of the classification from independent probability*” e introdujo medidas para describir su magnitud. Defendía que las variables observadas eran transformaciones discretas de dos variables aleatorias continuas con distribución normal bivariada y pretendía extender la teoría de la correlación a las tablas de contingencia.

Yule era contemporáneo de Pearson, pero eligió una aproximación alternativa. Habiendo completado los trabajos pioneros sobre modelos de regresión múltiple y referidos a coeficientes de correlación parcial y múltiple, Yule (1911) desvió su atención a partir de 1900 hacia el estudio de la asociación en tablas de contingencia, ya que creía que muchas de las variables categóricas eran intrínsecamente discretas y definió índices en términos de conteo de celdillas, sin asumir un continuo subyacente. Sus medidas incluyen la razón de productos cruzados y la transformación de esta, llamada Q de Yule. Yule (1907) descubrió la discrepancia potencial entre asociaciones marginales y condicionales en tablas de contingencia, más tarde revisado por Simpson (1951) y llamada paradoja de Simpson. Yule (1907) por su parte argumentaba que las variables observadas debían ser el núcleo de interés independientemente de su nivel de medida. Pero Pearson (1906-1914) era el líder de la ciencia estadística en Inglaterra, se

impuso su mayor prestigio y personalidad, rechazando las ideas de Yule (1911) y desprestigiándolo argumentando, que sus coeficientes eran inadecuados.

Si esta controversia fue grande en su época, se quedó pequeña cuando entró en escena otro estadístico llamado Fisher. En 1922 introdujo el concepto de *grados de libertad*, usando una representación geométrica. Los grados de libertad caracterizan la familia de distribuciones ji-cuadrado, y Fisher (1922) afirmó que en las pruebas de independencia en tablas I x J, la χ^2 tenía (I-1)(J-1) grados de libertad. En contraste, Pearson (1900) argumentaba que para la aplicación de este estadístico, los grados de libertad igualaban el número de celdillas menos 1, o IJ-1 para tablas 2 x 2., Pearson (1900) sostenía que si utilizamos proporciones en filas y columnas para estimar las probabilidades esto tenía un efecto despreciable en distribuciones con muestras grandes.

La reputación de Fisher (1925) aumentó gracias a sus trabajos teóricos, en los que introdujo conceptos como suficiencia, información, y propiedades óptimas de los estimadores maximoverosímiles, pero además por sus contribuciones metodológicas en áreas como el diseño de experimentos y el análisis de varianza (Fisher 1935).

Por otra parte, Spearman (1927) que era psicólogo de profesión, estudió estadística y logro desarrollar notables aplicaciones de la estadística en el campo de la psicología, entre ellas, la correlación de datos que lleva su nombre. Además, desarrolló el Análisis Factorial, donde diferencia entre un factor general (factor G) y cierto número de factores específicos (factores E).

Sus aportaciones a la estadística son el coeficiente de correlación de rangos (Spearman 1904), así como la creación y desarrollo de los llamados *experimentos factoriales* en el ámbito de la estadística, que son aquellos experimentos en los que se estudian simultáneamente dos o más factores, y donde los tratamientos se forman por la combinación de los diferentes niveles de cada uno de los factores. Los experimentos factoriales se emplean en todos los campos de la investigación, son muy útiles en investigaciones exploratorias en las que se sabe poco sobre un número amplio de factores. En definitiva, Spearman (1927) logró establecer la relación entre psicología y estadística, y viceversa, ya que para él una se completaba con la otra.

Paralelamente, y por la misma época, mediados de los años 30 aparecieron una serie de trabajos sobre análisis de datos categóricos, en los que Bliss (1935) popularizó el modelo *probit* para respuestas binarias. En el apéndice de uno de los artículos de Bliss (1935), Fisher proporcionó un algoritmo para obtener estimaciones de parámetros maximoverosímiles en el modelo *probit*. Este algoritmo era el de Newton-Rapson, llamado también *scoring* de Fisher. Pero el prestigio de Pearson seguía imponiéndose, aunque Barlett (1935) había utilizado la razón de productos cruzados para definir la interacción de segundo orden en una tabla de contingencia tridimensional.

Por su parte, Fisher (1940) desarrolló también métodos de la correlación canónicos para las tablas de contingencia y muestra cómo asignar puntuaciones a las filas y columnas de una tabla de contingencia para aumentar al máximo la correlación. Su trabajo se relaciona más tarde con el desarrollo, particularmente en Francia, del análisis de correspondencias.

Además, Fisher y Yates (1938) propusieron el $\log[B / (1 - B)]$ como una de las posibles transformaciones de un parámetro binomial para analizar datos binarios. En este sentido, el físico y estadístico Berkson (1944) introdujo el término de 'logit' para esta transformación y demostró que el modelo de regresión logística se ajustaba, de manera similar, al modelo *probit*. Finalmente, Cornfield (1951), otro estadístico, mostró el uso de la razón de productos cruzados para aproximar los riesgos relativos en estudios de casos y controles (también denominados caso-control) con este modelo.

Por su parte, Cox (1969) también tenía influencia considerable popularizando el modelo de regresión logística sumamente utilizado en la actualidad. Además, es interesante destacar la importante aportación del estadístico danés Rash (1961) que propuso el ahora denominado modelo de Rash, que tuvo una gran influencia en el norte de Europa, estimuló muchas generalizaciones en la comunidad educativa en los Estados Unidos, hoy de uso generalizado. Es necesario señalar que la mayoría de sus trabajos son aplicados en psicología, muy específicamente en la corriente psicométrica.

Cabe también destacar a Cochran (1950) que publicó un trabajo sobre análisis de datos categóricos, así como, una generalización (la Q de Cochran) de la prueba de McNemar (1947) para comparar proporciones en muestras emparejadas. Mostró cómo dividir ji-cuadrado en componentes que describieran varios aspectos de asociación, como la tendencia lineal en proporciones binomiales en las filas cuantitativamente definidas de una tabla $I \times 2$. Cochran también propuso una prueba de independencia condicional para tablas $2 \times 2 \times K$, similar al propuesto después por Mantel y Haenszel (1959).

No obstante, durante los últimos cuarenta años ha sido Goodman (1962-2007) el contribuyente más prolífico al avance de metodología del análisis de datos categóricos. El campo debe una tremenda gratitud a su trabajo (*Contributions to the statistical analysis of contingency tables* 2002). Además, algunos de los estudiantes de Goodman en Chicago también han hecho contribuciones fundamentales como Haberman (1970-2009).

Otra nueva técnica importantísima, se refiere a la estimación por máxima verosimilitud (Birch 1963) procedimiento que permite estimar los parámetros de un modelo probabilístico, o los coeficientes de un modelo matemático, de tal manera que sean los más probables a partir de los datos obtenidos. Esta técnica forma parte fundamental de la corriente estadística bayesiana y de los modelos de detección de señales donde las probabilidades anteriores y posteriores a un suceso, así como las matrices de transición correspondientes, son fundamentales para entender los procesos. Además es una técnica fundamental para la unificación teórica de la estadística, como veremos seguidamente.

Por su parte, Birch formuló lo que llamó análisis loglineal. Este trabajo produjo un gran impacto, consolidando este enfoque de análisis dentro de los datos categóricos. La metodología usada por este enfoque es la estimación por máxima verosimilitud, llamándose por ello *enfoque MV*. Sus características principales eran: a) el empleo de métodos de estimación por máxima verosimilitud para estimar los parámetros; b) la carencia de una estructura analítica general, produciendo una dispersión de los procedimientos, cada uno de los cuales requiere su propia distribución y su modo específico de proceder; c) son técnicas analíticas muy versátiles que admiten variables

numéricas y son robustas ante la presencia de dificultades como celdillas con frecuencia cero; d) a causa de sus propiedades estadísticas hay restricciones en su aplicabilidad.

Otro trabajo de importancia fue el de de Grizzle, Starmer y Koch (1969) que formulaba un enfoque analítico alternativo dentro de la infraestructura del modelo lineal clásico, y que fue utilizado en sus comienzos por biólogos y psicólogos. Este es el *enfoque GSK* cuyas características principales son: a) el empleo de un método de estimación por mínimos cuadrados ponderados MCP para estimar los parámetros; b) un enfoque centrado en el modelo lineal clásico, utilizando la aproximación normal y produciendo homogeneidad en los procedimientos desarrollados; c) incluyen técnicas analíticas rígidas, que solo admiten variables categóricas y son poco robustas con datos atípicos como tablas de contingencia con celdillas nulas; d) permiten la aplicación a una amplia gama de diseños de investigación, similares al ANOVA, pero con variables categóricas.

Sin embargo, no sería hasta el trabajo de Nelder y Wedderburn (1972) cuando se inicia un enfoque que pretende unir todos los procedimientos conocidos, y se ponen las bases de lo que se conoce como *modelo lineal generalizado (MLG)*, que consiste en una extensión del modelo lineal clásico (MLC) y que abarca toda una familia de distribuciones, distinguiendo durante el proceso de estimación dos diferentes escalas vinculadas entre sí por una función de enlace. Utiliza un procedimiento común de estimación por máxima verosimilitud que tiene como cuerpo básico un algoritmo por mínimos cuadrados ponderados, articulado en un proceso iterativo. Este procedimiento de estimación se denomina por ello *mínimos cuadrados iterativamente reponderados MCIR*. Como resumen esquemático ver tabla 1.1.

Tabla 1.1: Hitos históricos en la estadística

Año	Suceso	Autor(es)
1202	Publicación de Liber Abaci	Leonardo de Pisa ("Fibonacci")
1654	Correspondencia relativa a los juegos de azar con los dados	P. de Fermat y B. Pascal
1656	Publicación de De ratiocinis in ludo aleae	Ch. Huygens
1713	Publicación de Ars Coniectandi	J. Bernuilli
1714	Publicación de Teoría de la Probabilidad	A. De Moivre
1860	Se acuña el nombre de estadística	G. Achenwall
1763	Publicación de essay towards solving a problema in the doctrine of chances	Rev. Bayes
1805	Desarrollo de método de mínimos cuadrados	A.M. Legendre
1809	Publicación de Theoria Motus Corporum Coelestium	C.F. Gauss
1812	Publicación de Théorie analytique des probabilités	P.S. Laplace
1834	Creación de la Statistical Society of London	
1837	Distribución de Poisson	D.S. Poisson
1839	Creación de la American Statistical Association	
1889	Publicación de Natural Inheritance	F. Galton
1895	Introducción a las distribuciones asimétricas	K. Pearson
1900	Desarrollo de la prueba ji-cuadrado	K. Pearson
1908	Publicación de The probable error of a mean	W.S. Gosset ("Student")
1910	Publicación de An introduction to the theory of statistics	G.U. Yule
1911	Publicación de The grammar of science	K. Pearson
1915	Distribución del coeficiente de correlación	R.A. Fisher
1921	Publicación de Treatise on probability	J.M. Keynes
1925	Publicación de Statistical Methods for Research Workers	R.A. Fisher
1927	Publicación de The Abilities of Man: Their Nature and Measurement	C. Spearman
1928	Pruebas para verificación de hipótesis	J. Neyman y E.S. Pearson
1933	Publicación de On the empirical determination of a distribution	A.N. Kolmogorov
1933	Análisis Probit	C. Bliss
1935	Publicación de The design of experiments	R.A. Fisher
1936	Publicación de relations between two sets of variables	H. Hotelling
1937	Publicación de Properties of sufficiency and statistical tests	M.S. Barlett
1938	Publicación de Statistical tables for biological, agricultural and medical research	F. Yates y R.A. Fisher
1940	Publicación de Statistical Methods	G.W. Snedecor
1945	Pruebas no paramétricas	F. Wilcoxon
1950	Publicación de Experimental design	W.G. Cochran y G.M. Cox
1961	Publicación de On general laws and the meaning of measurement in psychology	G. Rash
1963	Publicación de Maximum likelihood in three-way contingency tables	M.W. Birch
1969	Publicación de Analysis of binary data	D.R. Cox
1969	Publicación de Analysis of categorical data by linear models	Gizzle, Starmer y Koch
1972	Publicación de generalized linear models	J.A. Nelder y R.W. Wedderburn
1976	Publicación de Time series analysis forecasting and control	G. Box y G. Jenkins
1977	Publicación de Exploratory data analysis	J. Tukey
1979	Publicación de Measures of association for cross classifications	L.A. Goodman y W.H. Kruskal
1983	Publicación de Generalized linear models	P. McCullagh y J.A. Nelder

Por último, hemos de resaltar que la formulación definitiva del modelo lineal generalizado fue propuesta por McCullagh y Nelder (1983), los cuales desarrollaron en profundidad los fundamentos teórico-matemáticos del modelo. Desde entonces son varios los autores que han continuado estos trabajos, Agresti (1990), Dobson (1983), Lindsey (1997), Judd y McClelland (1998).

1.3 EL PRESENTE TRABAJO

El modelado estadístico como hemos visto, es un enfoque de análisis de datos que empezó a gestarse a principios del siglo XX y que se desarrolló a partir de la mitad de este siglo, gracias en gran parte, a los avances en informática, que han propiciado la potencia de cálculo necesaria, y a los avances en estadística de finales de los 80 y principios de los 90. Visto de este modo, se plantea como una alternativa integradora, no excluyente a los modelos estadísticos tradicionales.

El objetivo de esta investigación es comprobar si estas técnicas estadísticas aportan una mayor comprensión de los procesos implicados en la memoria de reconocimiento, para la que ya existen técnicas de medida desde diferentes enfoques teóricos. En este trabajo nos apoyamos en la Teoría de Detección de Señales (TDS) como base teórica y utilizaremos algunas de sus herramientas metodológicas, que como trataremos de demostrar tienen relación con el modelado estadístico (DeCarlo 1998).

La exposición del presente trabajo se ha estructurado en dos partes, la primera constituye el enfoque teórico de la investigación donde se exponen, en primer lugar, los fundamentos teóricos del modelado estadístico; en segundo lugar, una exposición de los

diferentes modelos del modelado estadístico que podemos aplicar al estudio de la memoria de reconocimiento; y en tercer lugar, una revisión de las diferentes teorías que tratan de dar explicación a los procesos implicados en la memoria de reconocimiento, una breve introducción a la TDS, y métodos de medida de la memoria de reconocimiento.

La segunda parte consiste en el estudio empírico llevado a cabo. En primer lugar se realiza un estudio para comprobar si el modelado estadístico aporta información adicional a las técnicas tradicionalmente utilizadas para el estudio de este fenómeno desde el punto de vista de la detección de señales. En segundo lugar se aplican de nuevo estas técnicas a un estudio de memoria de reconocimiento donde se incluyen medidas psicofísicas como son los potenciales evocados. En tercer lugar, aplicamos el modelado estadístico a las respuestas de los sujetos añadiendo la información que nos proporcionan los tiempos de reacción.

CAPÍTULO 2

EL MODELADO ESTADÍSTICO

2.1- ¿QUE ES EL MODELADO ESTADÍSTICO?

El modelado estadístico es un proceso en el que, mediante la aplicación de un conjunto de procedimientos estadísticos, se alcanzará una explicación sobre la variable respuesta a partir de una función ponderada de una o más variables explicativas (Dobson, 1983). Es una de las herramientas más útiles para estudiar y comprender la variabilidad de un conjunto de datos observados. Dentro de esta variabilidad podemos distinguir dos elementos:

- La variabilidad sistemática
- La variabilidad aleatoria

La variable de respuesta que estudiamos estaría compuesta de este modo por un componente sistemático y otro aleatorio:

El *componente sistemático* podemos representarlo mediante un modelo de regresión (donde la variabilidad de los datos está explicada por categorías o variables objeto de nuestro estudio).

El *componente aleatorio*, también denominado residual, incluye las diferentes posibilidades de respuesta. Representa una medida que describe hasta qué punto la variable de respuesta se desvía debido a fluctuaciones de las categorías o variables ajenas a nuestros objetivos de medida.

La relación entre los dos componentes de la variable de respuesta se puede especificar mediante una sencilla ecuación:

$$\text{DATOS} = \text{MODELO} + \text{ERROR}$$

donde MODELO, es la variabilidad de los datos explicada por el componente sistemático, y ERROR, es la variabilidad no explicada.

2.2 - ¿CUAL ES EL OBJETIVO DEL MODELADO ESTADÍSTICO?

Se trata de un proceso por el cual los modelos se construyen, evalúan y modifican en relación directa con los datos obtenidos. Por tanto, el objetivo es conseguir un modelo a partir de unos datos específicos en el cual la proporción de componente sistemático sea lo más grande posible con respecto al error. Asimismo, este modelo debe representar la relación entre la variable de respuesta y las variables explicativas.

2.3 – ASPECTOS FUNDAMENTALES DEL MODELADO ESTADÍSTICO

Antes de iniciar el proceso debemos dejar claras nuestras intenciones a la hora de utilizar los modelos resultantes del proceso de modelado. En efecto, mediante el modelado podemos conseguir dos objetivos:

- *Explicativo*: pretendemos descubrir la relación hipotética entre dos o más variables (como ocurre en los modelos de análisis clásicos).
- *Predictivo*: pretendemos predecir el valor de la variable de respuesta a partir de las variables explicativas (como en los modelos de regresión).

Así, durante el proceso de modelado estadístico podemos obtener diferentes modelos igualmente válidos, y es aquí donde aparece el problema de la selección del

modelo estadístico óptimo (Claeskens y Hjort, 2008). En general, el principio que debe guiar esta búsqueda es el de parsimonia: debemos elegir el modelo más simple siempre que nos encontremos ante igualdad de condiciones, o dicho de otro modo, no debemos incluir efectos sistemáticos en el modelo si no es absolutamente necesario.

La manera de reducir el error en modelado estadístico, consiste en comparar pares de modelos. Para ello utilizaremos procedimientos de naturaleza inferencial, centrándonos en el aumento o disminución del error, evaluado mediante técnicas de contraste de hipótesis.

Además, debemos evitar lo que se denomina “sobreajuste”, ya que el excesivo ajuste a los datos no nos proporcionará buenas predicciones en caso de que se presenten fluctuaciones al utilizar muestras diferentes (validez externa), siendo esta una gran aportación del modelado (Song y Lee, 2006).

En relación a los aspectos sustantivos del modelado estadístico, hay que realizar la distinción entre:

- *Modelos de confirmación*: provienen de la teoría sustantiva o de estudios previos, por lo tanto, los modelos quedarían especificados antes del análisis de los datos.
- *Modelos de hipótesis*: los modelos nos vendrán sugeridos durante el proceso de modelado estadístico. Aquí el análisis de los datos no confirmaría las hipótesis del investigador sino que el modelo se convertiría en una hipótesis a comprobar en relación a un conjunto nuevo de observaciones.

Como consecuencia de todo lo anterior, “El proceso de modelado genera nuevas hipótesis que modifican la teoría sustantiva inicial” (Losilla, *et al.* 2005 p. 54). Es aquí donde el investigador ha de considerar la teoría sustantiva en relación estrecha con la metodología.

Por ello, es importante indicar que el criterio de selección de los modelos, no podrá basarse únicamente en criterios estadísticos, sino que también tiene que basarse en la significación sustantiva, y es gracias a esto por lo que podemos llegar a teorías explicativas (Kadane y Lazar, 2004).

Una cita clásica ya en el modelado estadístico nos recuerda que “modelling in science remains partly, at least, an art” (McCullagh y Nelder 1989, p.8). Por lo que, gran parte de la responsabilidad del proceso de modelado recae sobre el investigador y su perspicacia a la hora de trabajar.

2.4 – ETAPAS DEL MODELADO ESTADÍSTICO

El proceso del modelado estadístico se estructura básicamente en 4 etapas, pero dado su carácter iterativo es probable que volvamos a etapas previas.

Para elaborar un modelado estadístico debemos seguir las siguientes etapas (Agresti, 1996):

- Etapa de *especificación*: en esta etapa seleccionaremos (de entre todos los modelos posibles) los modelos más relevantes y que mejor representen nuestra variable de respuesta. En esta etapa deberemos tomar decisiones

relacionadas con la formulación del componente sistemático, los supuestos del componente aleatorio y cómo se combinan ambos componentes. Esta etapa está fundamentada, tanto en la teoría sustantiva, como en la observación de los datos.

- Etapa de *selección*: representa un segundo paso en el proceso de modelado estadístico, y consiste en la estimación de los parámetros del componente sistemático, donde valoraremos la diferencia entre datos observados y ajustados. Cuando tenemos varios modelos posibles para explicar los datos, debemos utilizar además la comparación de modelos anidados para seleccionar el modelo más parsimonioso, y que a su vez, represente mejor los datos observados. Además, hemos de estimar los intervalos de confianza de los parámetros del modelo, para así poder determinar la eficacia predictiva y la interpretabilidad del modelo (Zucchini, 2000).
- Etapa de *evaluación*: en esta etapa debemos valorar si el modelo ajustado es válido, dejando relativamente al margen, si presenta un buen ajuste a los datos. Para ello, realizaremos un diagnóstico del modelo, evaluando los posibles errores de especificación del componente sistemático, de la distribución de probabilidad del componente aleatorio, y por último, de la relación de ambos componentes (etapa de especificación del modelo) (Wei y Shih, 1994). Finalmente, se evaluarán las observaciones extremas o influyentes, para evitar asumir estimaciones erróneas.
- Etapa de *interpretación*: la última etapa del proceso de modelado estadístico, consiste en elegir el modelo óptimo en función de los criterios de bondad de ajuste y parsimonia, pero integrado e interpretado en el marco teórico de referencia que hemos propuesto (ver figura 2.1).

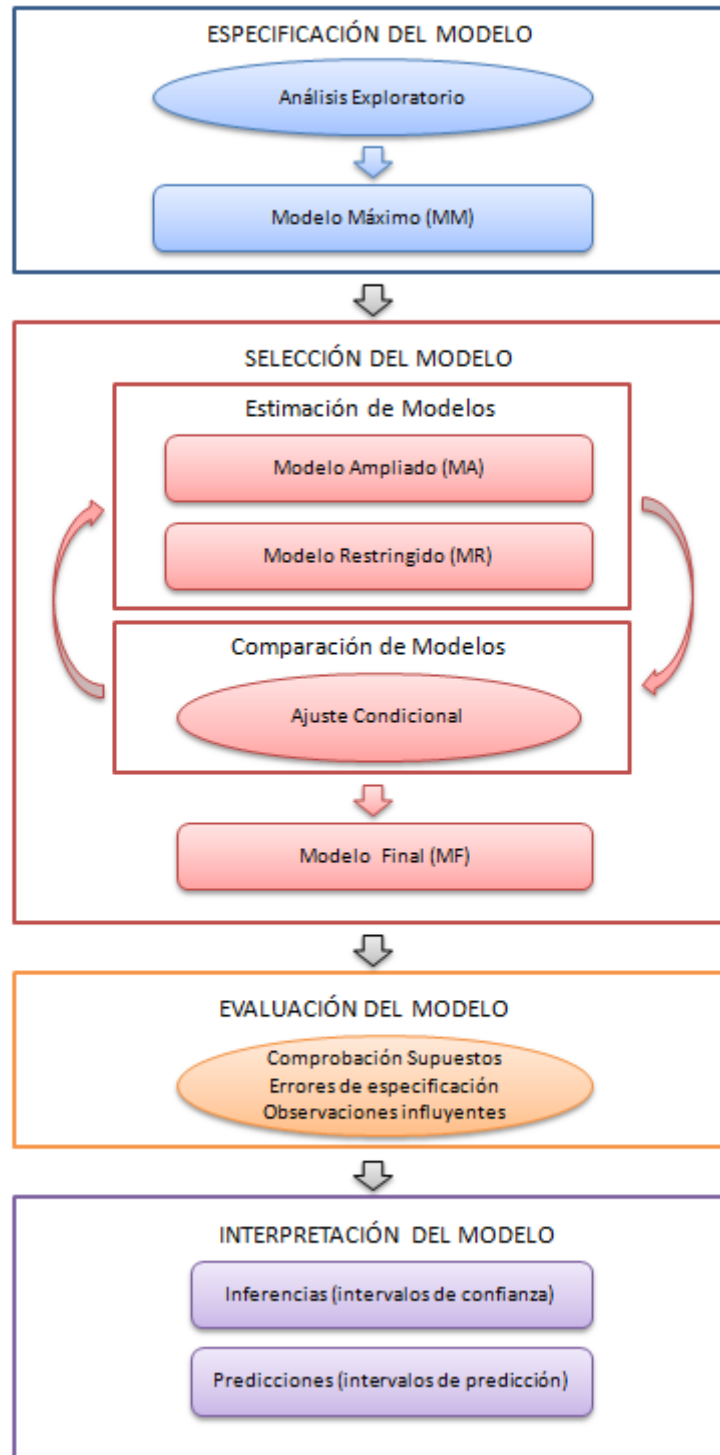


Figura 2.1: Etapas del Modelado estadístico (elaborado a partir de Ato, 1996)

2.5 – VENTAJAS DEL MODELADO ESTADÍSTICO

Las ventajas del modelado son tan evidentes que cualquier investigador desearía aplicar estas posibilidades a sus datos para así superar el simple contraste de hipótesis estadísticas, sin obviar estas, pero completándolas de forma muy eficiente de forma esquemática serían (Agresti, 2002):

- La aplicación de las técnicas de análisis estadístico adecuadas a la naturaleza de las variables. El análisis está al servicio de los datos y no al contrario.
- El desarrollo de conclusiones adecuadas a los datos que nos permitirá una correcta obtención de información sobre los datos.
- Garantizar y optimizar los objetivos de los resultados.
- Maximizar la información obtenida de los datos.
- Maximizar las posibilidades de tener éxito en el experimento (debido a la mayor potencia de la prueba estadística).

2.6 – SISTEMAS DE CODIFICACIÓN PARA VARIABLES CATEGÓRICAS

Antes de iniciar la exposición del proceso de selección de modelos, es importante recordar que las variables explicativas categóricas, sobre todo si tienen más de dos categorías, no debemos codificarlas con números consecutivos, ya que puede parecer que las estemos ordenando (López, *et al.*, 1997). Debemos codificarlas a partir de un conjunto de $c-1$ variables ficticias (donde c es el número de categorías), que contienen la misma información que la variable original. Además, las variables ficticias

pueden ser modelos creados por nosotros en base a una realidad que queremos contrastar.

Es importante considerar que la interpretación será diferente, en el caso de utilizar una variable ficticia únicamente que en el caso de $c-1$ variables ficticias. Por ello, en la fase de selección del modelo veremos si es conveniente la inclusión o exclusión de todas las variables ficticias.

Uno de los aspectos más importantes a tener en cuenta, son las diferentes formas en que podemos codificar las $c-1$ variables ficticias. La elección de cada uno de los procedimientos o técnicas, dependerá de los objetivos y de las hipótesis del investigador, ya que, cada procedimiento estima diferentes valores, por lo tanto, supone diferentes interpretaciones, pero siempre sin variar las diferencias que existan en el modelo. Si nuestra intención es únicamente predictiva, será indiferente el método utilizado (la interpretación de los coeficientes de regresión no nos interesa, sino únicamente la predicción que nos proporciona la ecuación de regresión). Pero si nuestra intención es explicativa, son las hipótesis enmarcadas en una teoría sustantiva o modelo, las que nos deben guiar en la elección del método de codificación, debido a que, en función del método escogido las comparaciones entre categorías de la variable explicativa serán distintas (Azen y Walker 2011).

2.6.1. *Dummy Coding* o Codificación con respecto a una categoría de referencia

Se suele utilizar mucho por su facilidad a la hora de la interpretación. En este sistema cada variable ficticia representa una categoría, y la interpretación de su coeficiente de regresión se hace como la diferencia entre la media de la categoría que

representa (\bar{y}_j) y la media de la categoría de referencia (\bar{y}_r). En cuanto a la elección de la categoría de referencia, ésta viene dada por el tipo de variable explicativa, o dicho de otro modo, depende del investigador, aunque existen una serie de reglas no escritas: si por ejemplo, tratamos con variable nominales y queremos ver la influencia de una nueva técnica con respecto a otras ya existentes, esta debería ser nuestra categoría de referencia. Si la variable es nominal pero no hay un objetivo claro, la categoría de referencia puede ser cualquiera a excepción de la que tenga menor frecuencia. Si por el contrario tenemos una variable ordinal, y una categoría representa “ausencia”, esta debería ser nuestra categoría de referencia (Hardy, 1993).

Para generar las $c-1$ variables ficticias, primero debemos escoger la categoría de referencia, en segundo lugar, debemos asignar el valor 1 en la variable ficticia a los casos de las categorías que no sean la de referencia y asignar el valor 0 al resto de variables ficticias, y finalmente, a los casos que pertenecen a la categoría de referencia se les asigna el valor 0 en todas las variables ficticias.

De este modo el modelo de regresión queda así:

$$\hat{Y} = b_0 + b_1c_1 + b_2c_2 + \dots + b_jc_j \quad (2.1)$$

Con esta ecuación obtendremos los valores predichos para cada categoría. El coeficiente b_0 es el valor de la media para la categoría de referencia, y los demás b_j son iguales a la diferencia entre las medias de cada categoría y de la categoría de referencia.

2.6.2. *Effect Coding* o Codificación con respecto al promedio de categorías

Se suele utilizar cuando no tenemos claro cual puede ser la categoría de referencia. Cada variable ficticia representa una categoría y su coeficiente de regresión

es la diferencia entre la media de la categoría que representa (\bar{y}_j) y el promedio de las medias de cada categoría (\bar{y}^*). Así, el valor promedio de las medias de cada categoría es b_0 .

Utilizando este modelo de codificación, una de las categorías (c_p) no está reflejada en ninguna variable ficticia, pero su efecto puede ser obtenido a partir de los otros efectos, ya que la suma de todos los efectos es cero (Kuhn y Johnson, 2013).

$$\sum_{j=1}^c (\bar{y}_j - \bar{y}^*) = 0 \quad (2.2)$$

Donde el efecto de c_p es:

$$(\bar{y}_p - \bar{y}^*) = (-1) \sum_{j=1}^{c-1} b_j \quad (2.3)$$

Para generar las $c-1$ variables ficticias, primero elegiremos la categoría (j) que representa cada una de las $c-1$ variables ficticias y la categoría no representada (p), en segundo lugar, a la categoría j se le asigna el valor 1, y el valor 0 para el resto de variables, y por último, a la categoría p se le asigna el valor -1 .

El modelo resultante de este método de codificación es:

$$\hat{Y} = b_0 + b_1c_1 + b_2c_2 + \dots + b_jc_j \quad (2.4)$$

De este modo, el valor predicho de un sujeto se expresa como la combinación de dos componentes:

- El promedio de las medias (b_0)
- El efecto de la categoría a la que pertenece el sujeto (b_j)

2.6.3. Codificación en categorías anidadas o codificación de termómetro

Se utiliza cuando la variable categórica tiene categorías ordenadas y se quieren hacer comparaciones entre categorías adyacentes (López, *et al.*, 1997). En este método las variables ficticias representan la diferencia entre categorías adyacentes, y el coeficiente de regresión representaría la diferencia entre las dos medias de las categorías adyacentes. En este caso b_0 es igual a la media de la primera categoría, lo que nos permitirá calcular el resto de las categorías.

Para generar las $c-1$ variables ficticias, primero, asignaremos el valor 0 a todas las variables ficticias de la categoría de orden 1, en segundo lugar, asignaremos el valor 1 a la variable ficticia c_1 y el valor 0 a las variables ficticias de la categoría de orden 2, en tercer lugar, asignaremos el valor 1 a las variables ficticias c_1 y c_2 , y el valor 0 a las variables ficticias de la categoría de orden 3, y por último habrá que continuar esta asignación del valor 1 en la variable ficticias c_1 a c_{r-1} , y el valor 0 a las variables ficticias de la categoría de orden r .

2.6.4. Codificación de contrastes

Si nuestras hipótesis son específicas necesitaremos un tipo de codificación capaz de responder a esta necesidad, es la denominada codificación de contrastes (Andersen, 1997). Teniendo en cuenta que los contrastes son una combinación de c medias:

$$\varphi = q_1\bar{y}_1 + q_2\bar{y}_2 + \dots + q_c\bar{y}_c = \sum_{i=1}^c q_i\bar{y}_i \quad (2.5)$$

donde φ es el valor del contraste y q_c son los coeficientes del contraste cumpliendo la condición:

$$\sum_{i=1}^c q_i = 0 \quad (2.6)$$

De este modo, los coeficientes de cada contraste representan una variable ficticia, debiendo crear $c - 1$ contrastes, cuyos coeficientes definen las variables ficticias que introduciremos en el modelo de regresión. Pero existe otra condición que deben cumplir, y es que ningún contraste puede ser la combinación lineal perfecta de las demás, porque si fuera así, no nos permitiría estimar los parámetros. Una última condición que deben cumplir es que todos los contrastes sean ortogonales entre sí.

Así, cuando realizamos contrastes ortogonales se pueden interpretar directamente los resultados obtenidos del ajuste de un modelo (aunque siempre con precaución). Los contrastes ortogonales dependen de sus coeficientes (si los tamaños muestrales son iguales en cada categoría), que a su vez están determinados por las hipótesis de investigación.

2.6.5. Otros sistemas de codificación

Además de los sistemas de codificación presentados (que son los más comunes), existen otros métodos de codificación como, por ejemplo, la codificación Helmert donde cada vector compara un grupo con el efecto medio de los grupos que le siguen, a excepción del último grupo. O bien la codificación de diferencias, donde a partir del segundo grupo, se compara el grupo con el efecto medio de los grupos anteriores, por ello también se llama codificación Helmert inversa (Starmer, Yates y Moore, 2010).

Por último hemos de añadir que otro tipo de codificación es la polinómica donde se definen contrastes ortogonales polinómicos. Lo cual permite comprobar las tendencias lineales de los niveles de la variable explicativa.

2.7 - SELECCIÓN DEL MODELO

Cuando nuestro objetivo es predictivo (queremos obtener una ecuación con el menor error de predicción) los criterios estadísticos serán los determinantes, ya que buscamos las variables que nos proporcionen menor error de predicción. Sin embargo, cuando nuestro objetivo es explicativo, serán los criterios teóricos los determinantes, dependiendo, de este modo, la selección de modelo de los conocimientos y experiencia del investigador.

Existen dos formas de seleccionar los modelos (Claeskens y Hjort, 2008):

- La denominada “selección conducida por el investigador”, donde es el investigador el que decide que términos permanecen en el modelo y cuáles no. Normalmente se utiliza en investigaciones con finalidad explicativa.
- O bien, la selección mediante algoritmos automáticos en base a criterios fijados a priori. Se utiliza en investigaciones exploratorias o predictivas.

2.7.1 - SELECCIÓN DE MODELOS CONDUCTIDA POR EL INVESTIGADOR

En esta técnica hay que tener en cuenta el principio jerárquico por el cual si en un modelo se elimina un término cualquiera, todos los términos de mayor orden en los que esté incluido también se eliminarán, y a la inversa, si incluimos un término, todos los de orden inferior también deberían incluirlo (Bishop, Fienberg, y Holland, 1975).

Al trabajar con modelos jerárquicos, los términos de alto orden (interacciones) no dependen del método de codificación empleado (Kleimbaum 1994); sin embargo, si

el modelo no es jerárquico la significación estadística de los coeficientes de regresión varía dependiendo del método de codificación. Además, hemos de añadir que sólo pueden incluirse en el modelo variables resultantes del producto de varias variables, siempre y cuando cumplan el supuesto de que representen un concepto interpretable en la práctica.

Además, nuestra estrategia de selección la basaremos en el planteamiento de un modelo máximo (MMax), y sus sucesivos modelos restringidos (MR), a través de una estrategia de ajuste condicional, siempre teniendo en cuenta el principio jerárquico.

Los pasos a seguir serían los siguientes:

1. Especificar el modelo MMax: debe estar basado en la teoría sustantiva y la observación de los datos. Si trabajamos con diseños no experimentales es importante incluir términos de control para evitar sesgos en la estimación de los efectos. Este modelo máximo debe contener (Draper y Smith 1998; Kleinbaum, Kupper y Muller, 1998): potenciales variables explicativas de interés para el investigador, potenciales variables de control, potenciales variables modificadoras, y potencias y términos de interacción entre las variables anteriores.

También es importante la relación entre el número de términos del MMax y el número de casos. Sabemos que el número de casos deber ser igual o mayor que el numero de parámetros a estimar + 1. De hecho, se suele utilizar una matriz que contenga al menos 10 veces el número de parámetros del MMax ($n \geq 10p$) (McCullagh y Nelder 1989).

Los criterios que deberemos seguir para seleccionar los términos que incluiremos en el modelo deberían consistir en: primer lugar los que sean importantes desde el punto de vista práctico, en segundo lugar los que tengan una interpretación práctica clara, en tercer lugar los que se basen en medidas objetivas y fiables y por último, las que tengan un menor coste de obtención y registro.

Hemos de aclarar que un MMax grande minimiza la probabilidad de error tipo II o infraajuste, mientras que un MMax pequeño minimiza la probabilidad de error tipo I o sobreajuste.

2. Evaluar los términos de interacción: comenzaremos evaluando las interacciones de mayor orden mediante pruebas de significación estadística (Kleinbaum, Kupper y Muller 1998). Una opción es evaluar las interacciones de una en una y la otra es probar el grupo completo de interacciones de un determinado orden, “*Chunk test*”, y si resulta estadísticamente significativo, comprobar cada interacción por separado. Si el resultado es no significativo, eliminamos todos los términos, pasando a un nivel inferior de interacciones. Si este resultado es significativo, sólo eliminaríamos del modelo los términos que examinemos de forma individual y no sean estadísticamente significativos. Por último, evaluaremos los términos de interacción de nivel inferior que no estén incluidos en los términos de interacción de nivel superior y que sean estadísticamente significativos.

Como hemos comentado anteriormente, si el efecto de interacción es fácilmente interpretable, podemos incluirlo o descartarlo en base a la estimación por intervalo de sus coeficientes o de los índices ΔR^2 y PRE.

3. Evaluar las variables explicativas y de control: podemos mantener o no las variables de manera individual o por grupos, ya que si las eliminamos pueden producir una estimación sesgada del efecto de las variables explicativas (su eliminación produce un cambio relevante en el parámetro de al menos una variable explicativa). Este fenómeno se denomina “confusión” (Everitt y Howell, 2005). También la eliminación de esta variable puede un aumento del error estándar o pérdida de precisión en la estimación del parámetro de al menos una variable explicativa. Por último, puede resultar imposible su eliminación si la variable de control forma parte del Modelo Mínimo (MMin). Debemos señalar que “en el caso de modelos que incluyan términos de interacción, la evaluación de la confusión y la precisión es compleja, puesto que el coeficiente de las variables explicativas de interés no es único, sino que depende de los valores de las variables de interacción”. En estos casos, lo mejor es ajustar diferentes modelos que estimen los coeficientes de las variables explicativas y sus intervalos de confianza (IC) para cada combinación de las categorías de las variables modificadoras. Tomada esta decisión, nuestro siguiente paso consiste en evaluar los sucesivos MR. Hemos de añadir que para valorar la relevancia científico-práctica utilizaremos los IC95%, ΔR^2 y PRE, y para la significación estadística el grado de significación P_j .

2.7.2 - SELECCIÓN DE MODELOS MEDIANTE ALGORITMOS AUTOMÁTICOS

El uso de algoritmos de selección automática está basado únicamente en criterios estadísticos (Forster, 2000).

1. *La bondad de ajuste del modelo*: cuando trabajamos con modelos de regresión lineal se utilizan el coeficiente de determinación (R^2), el coeficiente de determinación ajustado (\bar{R}^2) y el estadístico C_p de Mallows (1973). Teniendo en cuenta el contexto en el que estamos trabajando. Para la inclusión y exclusión de términos habrá que utilizar la diferencia entre los grados de libertad de los modelos que se comparan:

$$\bar{R}^2 = 1 - \frac{D_T/n - p}{D_N/n - 1} \quad (2.7)$$

donde n es el número de observaciones y p el número de parámetros del modelo de trabajo. Se utiliza éste en vez del clásico R^2 , ya que no tiene el problema de que al incrementarse el número de términos tienda a 1. Por otra parte, podemos hacer el ajuste mediante el estadístico C_p de Mallows (1973).

$$C_p = \frac{D_{MT}}{D_{MMAX}/gl_{MMAX}} + 2p - n \quad (2.8)$$

Si el modelo de trabajo tiene un ajuste aceptable, C_p es casi igual a p , por tanto el criterio consistirá en seleccionar los modelos cuyo C_p sea menor o igual a p , y entre estos el que tenga el C_p menor (Draper y Smith, 1981). Hay que tener en cuenta que cuando el MR tiene distinto p , cada criterio puede llevar a seleccionar un modelo distinto. Este estadístico es de los más utilizados, ya que nos lleva a seleccionar los modelos más parsimoniosos. También se suele representar gráficamente mediante una gráfica bidimensional de los valores de C_p de diferentes modelos y los parámetros de cada uno de ellos. Los modelos óptimos estarán por debajo (o sobre la diagonal principal) y los menos adecuados por encima de esta (Neter, Wasserman y Kutner, 1983), ver figura 2.2.

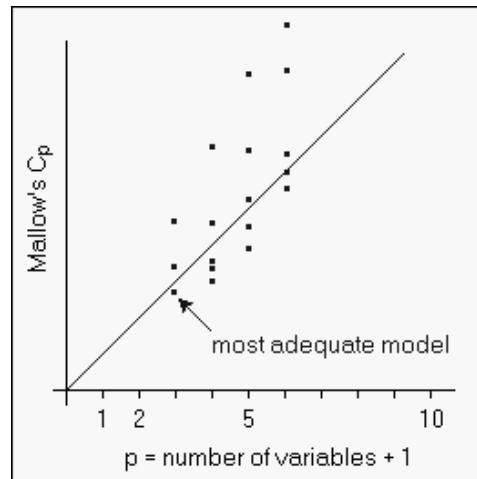


Figura 2.2 Grafico de los valores de C_p

2. Las técnicas automáticas de selección:

Hay dos tipos (Dobson 1983):

- “All posible regressions”, que es la estimación de todas la ecuaciones posibles.
- “Métodos por pasos”, que se construyen mediante la eliminación o adición de términos a los modelos (“*backward elimination*”, “*forward selection*”, y “*stepwise regresión*”) o definiendo conjuntos de términos para su evaluación global (“*chunkwise regresión*”).

La primera de las técnicas está basada en la estimación de todos los modelos posibles que se derivan del MMax. Hay que estimar $2^k - 1$ (donde k es el numero de términos del MMax). Después se agrupan estos modelos en función del número de términos que incluyen, y se ordenan en función de los criterios de ajuste calculados. La selección final dependerá de las ecuaciones que presenten mejores valores en estos criterios y también de cuestiones como, coste, tiempo, objetividad, etc. Tiene

el problema de que normalmente resulta inviable cuando el MMax tiene un número elevado de términos.

La segunda, consiste en la técnica por pasos (bien hacia atrás o bien hacia delante). En este caso, se construyen los modelos mediante la eliminación de términos a partir del MMax o su adición desde el MMin. Así, en cada uno de dichos pasos, se valorará un término o bloque en función de los criterios fundamentados en los índices de ajuste y tolerancia.

Finalizamos cuando ningún término cumpla con los criterios de exclusión o inclusión. Hemos de añadir que, el método “*stepwise regresión*” es una modificación del método “*forward selection*”. Éste nos permite reexaminar en cada paso las variables que se han incorporado al modelo y eliminarlas si no son relevantes. De este modo las variables pueden incluirse y excluirse del modelo más de una vez.

El denominado “*chunkwise regresión*” es una técnica muy utilizada ya que reduce el número de modelos a valorar, incorporando las variables necesarias siguiendo criterios sustantivos. (Kleinbaum, Kupper y Muller, 1998).

En el programa de análisis estadístico *R* (Ihaka y Gentleman, 1996) existe un paquete específico llamado *glmulti* (Calcagno, 2010) que realiza estos cálculos de manera automática.

Limitaciones: el principal problema de estas técnicas es que nos pueden llevar a seleccionar modelos inconsistentes o incoherentes desde un punto de vista estructural y teórico (Nuñez, *et al.*, 2011), podemos encontrarnos con dos problemas, uno es la sobreestimación de efectos cuando no tenemos claro qué modelo queremos, y tenemos un gran número de ellos. Este problema no debería ocurrirle a un investigador experto en el tema. El otro inconveniente está relacionado con los problemas de interpretación derivados de no mantener la jerarquía de términos. Así, al evaluar las interacciones puede ocurrir que seleccionemos un modelo que incluya un término de interacción, pero no los efectos principales. Este sería un problema de diseño (Myung, 2000).

2.8 CONCLUYENDO

En este capítulo se ha tratado de forma didáctica de aproximarnos a la “filosofía” del modelado estadístico, donde a diferencia de los métodos de análisis al uso, no existe un único resultado posible, sino que podemos obtener diferentes resultados “válidos” en función de los datos, la teoría sustantiva, el número de variables, los conocimientos del investigador, etc., lo cual enriquece y pone el valor los resultados obtenidos en nuestras investigaciones.

Una vez introducidas estas técnicas, pasaremos a repasar los modelos que podemos aplicar a nuestros datos de memoria de reconocimiento en función de la naturaleza de las variables con las que trabajamos. En el siguiente capítulo realizamos una exposición de los modelos a aplicar cuando la variable de respuesta es continua, el modelo de regresión lineal, desde la perspectiva del modelado estadístico.

CAPÍTULO 3

EL MODELO DE REGRESIÓN LINEAL

Las cosas complejas y estadísticamente improbables, son por naturaleza más difíciles de explicar que las cosas simples y estadísticamente probables, Richard Dawkins.

3.1 INTRODUCCIÓN

Es en los trabajos de Galton (1877-1889) donde aparece por primera vez el término de modelos de regresión. En sus trabajos sobre biología quería determinar la dependencia de la estatura de los hijos respecto a la de sus padres, y observó que se daba una “regresión a la media”. En la actualidad es un efecto incluido en la validez interna de los diseños. El efecto se produce cuando se maximiza la variable explicativa. Es decir, la consideración de valores extremos en la variable explicativa provoca en medidas sucesivas una regresión a la media.

No obstante, la regresión es mucho más general y actualmente se entiende por modelos de regresión una serie de procedimientos de modelado de datos con los que queremos estimar y realizar predicciones sobre una variable dependiente, a partir de una o más variables independientes, y mediante unos valores conocidos (Everitt y Howell, 2005).

$$Y' = \hat{\beta}_0 + \hat{\beta}_1 X_i + \varepsilon \quad (3.1)$$

Estos modelos deben cumplir una serie de supuestos:

- **Linealidad:** el valor esperado de la variable dependiente Y debe ser una función lineal de la variable explicativa X.
- **Homoscedasticidad:** la varianza de los errores debe ser constante.
- **Independencia:** los errores deben ser independientes entre sí, y además, independientes del predictor.
- **Normalidad:** la distribución de los errores debe seguir la distribución Normal.

La distribución de los datos deberá seguir una distribución normal cuyas características están resumidas en la tabla 3.1:

Tabla 3.1 Características de la distribución normal

Distribución	$y_i \sim N(\mu, \sigma^2 / w_i)$
Parámetro canónico	μ_i
Parámetro de escala	σ^2
Función de enlace	$\eta_i = \mu_i$
Función de varianza	$V(\mu) = 1$
Desviación	$D = \sum w_i (y_i - \mu)^2$

Es importante señalar su importancia por su relación con la estimación por mínimos cuadrados, por el teorema de Gauss-Markov. Siendo esta estimación para modelos lineales poco robusta frente a valores atípicos (*outliers*). Si la distribución de los valores atípicos es asimétrica, los estimadores pueden estar sesgados.

La gráfica de su función de densidad tiene una forma de campana y es simétrica respecto de un determinado parámetro. Esta curva se conoce como campana de Gauss y es el gráfico de una función gaussiana. Ver figura 3.1

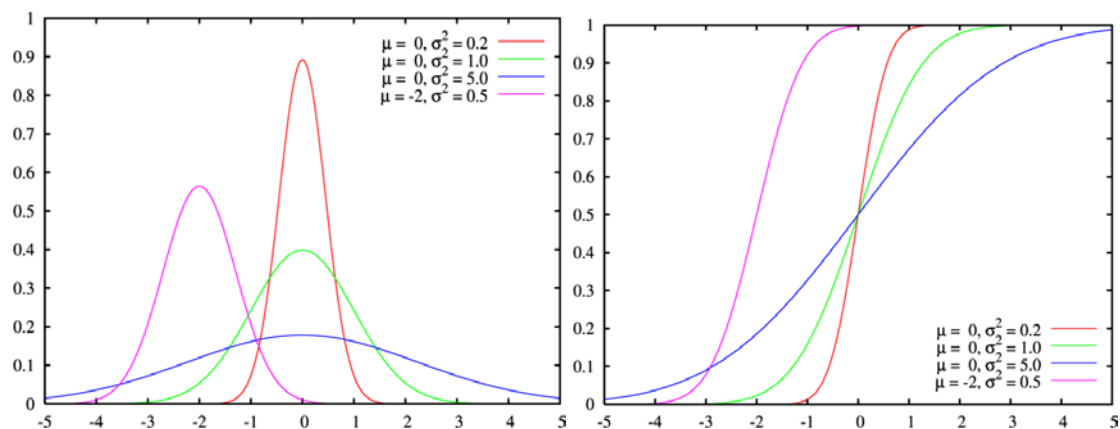


Figura 3.1 Función de densidad y función de distribución normal.

He incluido este modelo en este trabajo por ser el conocido por la mayoría de los investigadores y de esta manera constituye una buena introducción didáctica que sirva

para comprender mejor la filosofía del modelado estadístico y ayude la comprensión del resto de modelos que expondré. Además, si utilizamos medidas psicofísicas o proporcionadas por test psicológicos como variables de respuesta está plenamente justificado su uso. En el contexto de la memoria de reconocimiento podríamos aplicar estos modelos a medidas de la latencia de los potenciales evocados, su voltaje, los tiempos de reacción, etc.

3.2- EL MODELO DE REGRESIÓN LINEAL SIMPLE

En este capítulo veremos el modelo de regresión lineal a través de la comparación de modelos típica del modelado estadístico. En este tipo de modelos se trabaja con variables de respuesta continuas.

Para comenzar debemos realizar un “diagrama de dispersión” para poder comprobar las parejas de valores para cada observación (ver figura 3.2).

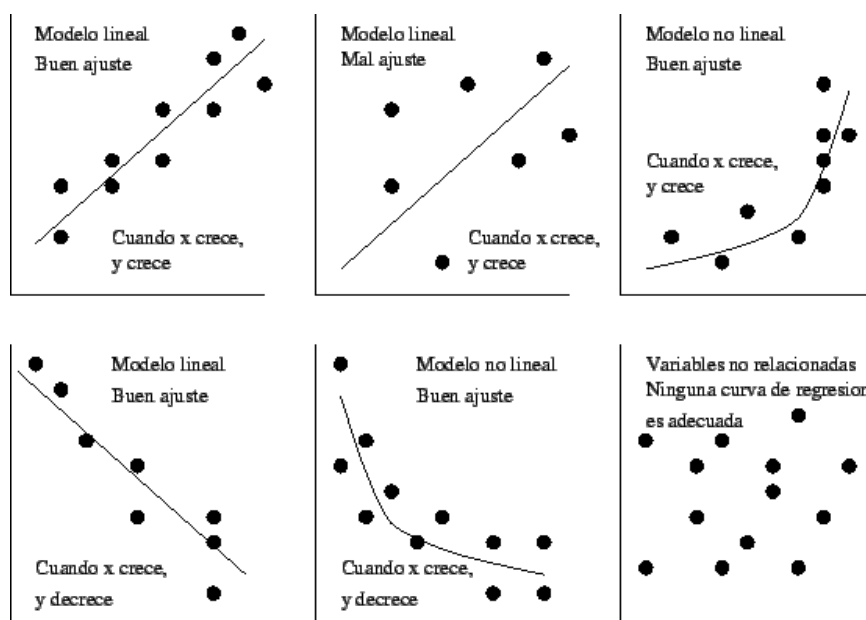


Figura 3.2 Clasificación de los “diagramas de dispersión” más comunes

Para valorar la asociación lineal entre dos variables continuas podemos utilizar:

La *covarianza*: que es una medida paramétrica que expresa el grado de variación conjunta de dos variables.

$$S_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1} \quad (3.2)$$

El *coeficiente de correlación producto-momento de Pearson*: este coeficiente nos permite determinar el grado y signo de covariación entre dos variables, y puede variar de -1 a +1.

$$r_{xy} = \frac{S_{xy}}{S_x S_y} \quad (3.3)$$

El valor que nos interesa es el del coeficiente de correlación lineal que varía desde +1 (relación lineal positiva perfecta) hasta -1 (relación lineal negativa perfecta). El valor 0 nos indica la ausencia de relación lineal (pero no descarta otros tipos de relación no lineal). Si elevamos al cuadrado el coeficiente de correlación lineal obtenemos el coeficiente de determinación (R^2), que nos indica la proporción de la variabilidad de la variable dependiente, explicada por la variable independiente (la bondad del ajuste del modelo), y su inversa nos proporciona el tanto por ciento de error. Entre las características más importantes del coeficiente de determinación consideraremos los siguientes (Montgomery, Peck y Vining, 2002):

- Nunca decrece al añadir nuevas variables, sino que aumenta hasta llegar al valor 1.
- Hay que comprobar la aportación de una nueva variable al modelo comparando el valor R^2 , con el que se produciría al incluir cualquier variable sin influencia en el modelo.

Así, ajustamos una recta para resumir la relación lineal, que podemos representar más formalmente mediante la siguiente fórmula:

$$\hat{Y} = b_0 + b_1X \quad (3.4)$$

donde b_0 es el valor de Y cuando $X = 0$, en el punto geométrico donde la recta cruza el eje de ordenadas, y b_1 es el denominado coeficiente de regresión. Da lugar a la pendiente de la recta de regresión. \hat{Y} es la variable dependiente estimada.

De este modo, dado un valor de la variable X , el modelo suministra una predicción \hat{Y} .

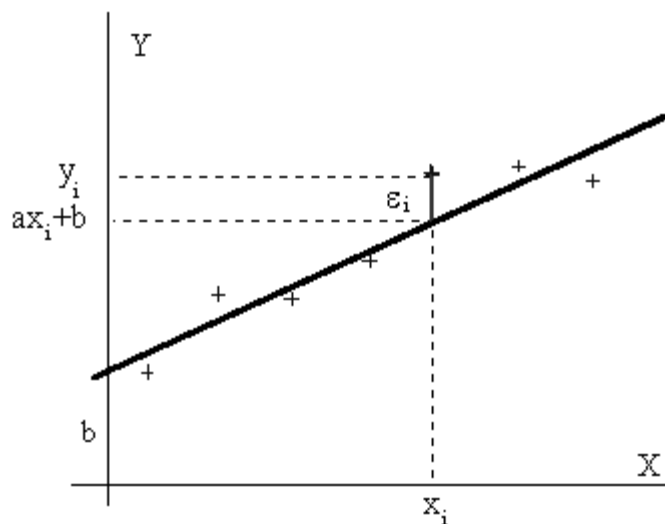


Figura 3.3 Recta de regresión

La diferencia entre el valor real (observado) y el valor predicho, es lo que se denomina “residual” o error en la predicción:

$$e_i = y_i - \hat{y}_i \quad (3.5)$$

3.2.1 - ¿COMO SE AJUSTA LA RECTA DE REGRESIÓN?

Los criterios para determinar los coeficientes de la recta de regresión que mejor ajusta a la nube de puntos son variados. El más utilizado es la suma del cuadrado de los

errores (SCE), ya que consigue que los errores para todas las observaciones sean lo más pequeños posibles, porque el cuadrado penaliza los errores grandes. Como dichos errores se sitúan en torno a una recta estimada, ello produciría errores que tienden a 0.

El método que nos asegura que esto se cumpla es el *ordinary least squares* (OLS) o en castellano mínimos cuadrados ordinarios (MCO), siguiendo los criterios de minimizar la SCE, que minimiza el error global que se comete al predecir los valores de Y a partir de los valores de X , y nos permite obtener b_0 y b_1 (Achen, 1982).

$$SCE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n (y_i - (b_0 + b_1 X))^2 \quad (3.6)$$

Los valores de b_0 y b_1 se obtienen mediante las formulas siguientes:

$$b_1 = r_{xy} \frac{S_y}{S_x} \quad b_0 = \bar{y} - b_1 \bar{x} \quad (3.7)$$

donde r_{xy} es el coeficiente de correlación de Pearson, S_y es la varianza de la variable dependiente y S_x la varianza del predictor.

3.2.2.- ESTIMACIÓN DE LOS PARÁMETROS DEL MODELO

Cuando deseamos utilizar el modelo de regresión lineal aplicado a una población, deberíamos conocer los valores de los parámetros β_0 y β_1 , pero al no conocer estos valores deberemos estimarlos a partir de datos obtenidos en nuestra investigación.

Para estimar los parámetros del modelo debemos comenzar por β_1 (primer parámetro a calcular), y nos plantearemos una serie de cuestiones (Berry, 1993):

En primer lugar debemos obtener la significación estadística del coeficiente b_1 . Evaluaremos la hipótesis nula $H_0 : \beta_1 = 0$, ya que deseamos comprobar en qué medida es b_1 diferente de 0 en la población. En el caso de rechazar esta hipótesis nos estaría indicando que el efecto es diferente de 0. (la variable explicativa tiene un efecto estadísticamente significativo sobre la variable de respuesta).

Asumimos que la distribución muestral de b_1 es normal con media β_1 . Para comprobar esta hipótesis utilizamos el estadístico t de Student:

$$t = \frac{b_1}{EE(b_1)} \quad (3.8)$$

donde $EE(b_1)$ es la estimación del error estándar de la distribución muestral de b_1 .

Para obtener $EE(b_1)$ tenemos que acudir a la siguiente fórmula:

$$EE(b_1) = \sqrt{\frac{SCE / (n - k - 1)}{SC_x}} \quad (3.9)$$

Partiendo del supuesto de que H_0 es cierta el estadístico t sigue una distribución t de Student con $n - k - 1$ grados de libertad, donde n es el número total de observaciones, y k el número de términos del modelo. Si el valor t obtenido es superior al valor crítico $t_{gl, \alpha/2}$ o $t_{gl, \alpha}$ (dependiendo de si la hipótesis es bilateral o unilateral) consideraremos que el parámetro β_1 es estadísticamente distinto de 0 (Allen, 1997).

En segundo lugar, debemos determinar la magnitud del efecto b_1 de la variable explicativa. Para ello necesitamos obtener el intervalo de confianza (IC) de β_1 . Los límites de este intervalo nos indican los valores entre los que puede variar el verdadero valor de b_1 en la población. Estos valores nos permitirán la interpretación del efecto de

X sobre Y en términos de magnitud del efecto. Para ello, debemos asociar b_1 a una medida de precisión de la estimación del valor del parámetro poblacional β_1 , es decir, el error estándar ($EE(b_1)$). Utilizando el valor de b_1 y su error estándar se puede construir el intervalo de confianza $1-\alpha$ de β_1 :

$$IC(1-\alpha)\beta_1 = b_1 \pm t_{n-k-1, \alpha/2} EE(b_1) \quad (3.10)$$

El intervalo de confianza es una alternativa a la prueba de significación t , (GroB, 2003) y además, nos “permite valorar la relevancia científica del efecto de la variable explicativa que refleja b_1 ”. Para ello es absolutamente imprescindible, establecer un valor previo de referencia.

3.2.3. – AJUSTE Y SELECCIÓN DEL MODELO

En modelado estadístico se utiliza el enfoque de la comparación de modelos, donde el modelo de regresión que incorpora la variable explicativa se compara con el modelo que no la incorpora. Este procedimiento se realiza comparando estos dos modelos alternativos, un modelo ampliado (MA) que incorpora la variable explicativa y el modelo nulo (MN) que no contiene la variable explicativa.

En el caso del modelo nulo (MN), sólo se estima el parámetro β_0 , prediciendo el mismo valor en la variable de respuesta para todos los sujetos, sin tener en cuenta el valor de los sujetos en la variable explicativa. Por tanto, la estimación de este parámetro, en este modelo es igual a la media de la variable de respuesta.

$$\begin{aligned} MA : \hat{Y} &= b_0 + b_1 X \\ MN : \hat{Y} &= b_0 \end{aligned} \quad (3.11)$$

Para comparar ambos modelos debemos valorar la diferencia de las discrepancias (D) entre los dos modelos. El procedimiento de comparación entre MN y MA es denominado como “estrategia de ajuste global” (Graybill e Iyer, 1994).

La discrepancia (D) está basada en la diferencia entre valores observados y los predichos por el modelo, lo que sería igual a la SCE de cada modelo. De este modo, la discrepancia del modelo ampliado será menor o igual a la discrepancia del modelo nulo. Esto es debido a que siempre que introduzcamos una variable explicativa al modelo, disminuirá el error de predicción, en el caso de que exista relación con la variable de respuesta, y por el contrario no cambiará su valor en el caso de que no exista relación con la variable de respuesta.

También tenemos que decidir si la disminución de la discrepancia (D) es suficiente para que sea estadísticamente significativa (Freund, *et al.*, 2006). Dicha comparación es equivalente a contrastar la hipótesis nula del parámetro β_1 , que vimos antes ($H_0: \beta_1 = 0$). Esta hipótesis indica la ausencia de relación lineal entre las variables X e Y , o en el contexto que estamos tratando, que la inclusión de la variable X no disminuye el error de predicción con respecto al MN. Para contrastar que no existen diferencias entre el ajuste de ambos modelos (MA y MN), utilizamos el estadístico F :

$$F = \frac{(D_{MN} - D_{MA}) / (p_{MA} - p_{MN})}{D_{MA} / (n - p_{MA})} \quad (3.12)$$

donde p_{MA} y p_{MN} representan el número de parámetros independientes estimados en cada modelo, n es el número total de observaciones y $n - p_{MA}$ los grados de libertad del MA.

Hay que señalar que el resultado de esta fórmula debe coincidir con el valor del estadístico t^2 , y refleja el error promedio que reduce cada variable del MA, respecto a la

inclusión de cualquier variable al azar. Si asumimos que los datos siguen una distribución normal de los errores, el estadístico F se distribuye según el modelo de probabilidad F de Fisher-Snedecor con grados de libertad iguales a los grados de libertad de numerador y denominador.

Su interpretación sería: la siguiente:

- Si $F \leq 1$ se acepta el MN.
- Si $F > 1$ se acepta el MA.

De este modo, si el valor F es menor que el valor crítico, aceptamos el MN (el error de predicción del MA no es menor que el del MN), por ser más parsimonioso. En el caso contrario, si el valor F es superior al valor crítico $F_{gl1,gl2,\alpha}$, el MN es rechazado, aceptando el MA (el efecto de variable explicativa sería estadísticamente significativo). Además, puede resultar muy útil calcular el valor de R^2 del MA, que nos dará la proporción de variabilidad de la respuesta Y explicada por la variable X (Lewis-Beck, 1980).

$$R_{MA}^2 = \frac{D_{MN} - D_{MA}}{D_{MN}} \quad (3.13)$$

Durante el proceso de comparación de modelos resulta bastante útil el uso de una tabla que resuma la información más relevante acerca de los modelos con los que estamos trabajando.

Al hacer un uso predictivo del modelo de regresión, podemos calcular el IC alrededor del cual estará cualquier valor \hat{y}_i predicho por la ecuación de regresión. Si consideramos el caso de la predicción de \hat{y}_i cuando $X = 0$, el IC para esta predicción se corresponde con el IC para β_0 :

$$\hat{y}_i = b_0 + b_1 0 = b_0 \quad (3.16)$$

Después deberíamos obtener el error estándar estimado $EE(b_0)$:

$$EE(b_0) = \sqrt{\frac{D}{n-p} \left[\frac{1}{n} + \frac{\bar{x}^2}{SC_x} \right]} \quad (3.14)$$

Calculando el IC de β_0 :

$$IC(1-\alpha)\beta_0 = b_0 \pm t_{n-k-1, \alpha/2} EE(b_0) \quad (3.15)$$

Como podemos comprobar este IC está construido de igual modo que el IC para β_1 , pero su interpretación es diferente, debido a que el valor de β_0 representa la media de Y cuando $X = 0$, por lo tanto su IC es el rango de valores entre los que se encuentra la media de Y en la población de sujetos cuyo valor $X = 0$. Para β_1 es el cambio en Y para cada incremento en una unidad en X , y su IC es el rango de valores de cambio en Y en la población.

También podemos calcular el error estándar del valor pronosticado \hat{y}_i , $EE(\hat{y}_i)$

$$EE(\hat{y}_i) = \sqrt{\frac{D}{n-p} \left[\frac{1}{n} + \frac{(x_i - \bar{x})^2}{SC_x} \right]} \quad (3.16)$$

y el IC de \hat{y}_i cuando $X = x_i$ tiene el valor

$$IC(1-\alpha)y_i = \hat{y}_i \pm t_{n-k-1, \alpha/2} EE(\hat{y}_i) \quad (3.17)$$

La distancia entre x_i y \bar{x} es la que determinará la amplitud del IC., así, a mayor diferencia entre x_i y \bar{x} , mayor será el error estándar de la puntuación predicha. De esta manera los IC para \hat{y}_i serán más estrechos a medida que la distancia entre x_i y \bar{x} sea menor.

Podemos calcular un intervalo alrededor de \hat{y}_i que nos permita saber el rango de valores entre los que se encontrará y_i para los sujetos de la población con un valor de $X = x_i$ (Cook y Weisberg, 1999). A esto se le denomina intervalo de predicción (IP), y su cálculo es igual que el del IC, pero con cambios en el error estándar, ya que la estimación de la respuesta para sujetos tiene más variabilidad que la estimación de la respuesta media. Para calcular el IP debemos calcular la desviación estándar de manera similar al error estándar:

$$DE(\hat{y}_i) = \sqrt{\frac{D}{n-p} \left[1 + \frac{1}{n} + \frac{(x_i - \bar{x})^2}{SC_X} \right]} \quad (3.18)$$

Así, nos encontramos que para un mismo valor de $X = x_i$ el IP para \hat{y}_i será más amplio que su IC.

Sin embargo, cuando utilizamos el modelo de regresión con fines predictivos debemos tener en cuenta que la recta de regresión se ha ajustado para el rango de valores de X observados, siendo incorrecto realizar predicciones para valores de X que se encuentren fuera del rango observado, debido a que la relación entre X e Y puede ser diferente a la estimada fuera de estos límites (Rawling, *et al.*, 1998).

3.3- MODELO DE REGRESIÓN LINEAL MÚLTIPLE DESDE EL MODELADO ESTADÍSTICO

Estamos trabajando con una generalización del modelo de regresión lineal simple con más de una variable explicativa o cuando nuestra variable explicativa es politómica. Esto implica que hay que determinar cuáles son las variables más relevantes para predecir o explicar la variable de respuesta. También nos facilita la predicción de la

variable de respuesta a partir de una serie de variables explicativas, esto nos va a ofrecer la ventaja de utilizar más información en la construcción del modelo, y consecuentemente, realizar estimaciones más precisas. Además podremos conocer cuáles son las variables explicativas que más influyen en la variable de respuesta. En este sentido es importante determinar si existe interacción entre las variables explicativas y controlar la influencia de variables de confundido.

Deberemos elegir cuidadosamente qué variables vamos a considerar como explicativas. Algunos criterios que deben de cumplir serán los siguientes (Losilla, *et al.* 2005):

- Tener sentido numérico.
- No deberá de haber variables repetidas o redundantes
- Las variables introducidas en el modelo deberán de tener una cierta justificación teórica.
- La relación entre variables explicativas en el modelo y casos debe de ser como mínimo de 1 a 10.
- La relación de las variables explicativas con la variable de respuesta debe de ser lineal, es decir, proporcional.

Cuando trabajamos con estos modelos incluiremos al menos un parámetro más a estimar con respecto al modelo de regresión simple.

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_j X_j + \varepsilon = \beta_0 + \sum \beta_j X_j + \varepsilon \quad (3.19)$$

Esta ecuación recibe el nombre de hiperplano, pues cuando tenemos al menos dos variables explicativas, en vez de recta de regresión tenemos un plano o superficie de regresión:

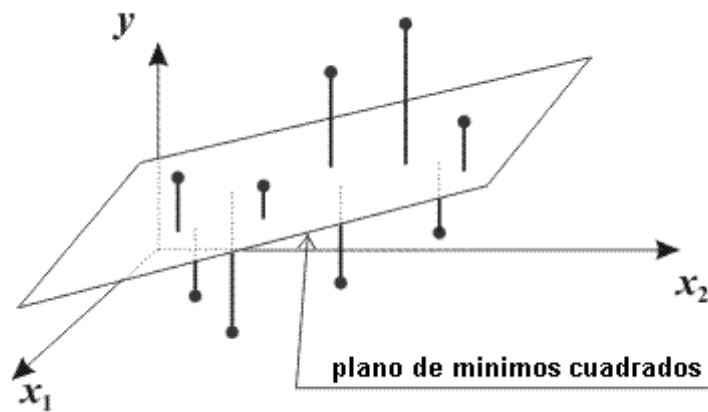


Figura 3.4 Superficie de regresión

Este es uno de los modelos más extendido en los estudios predictivos y explicativos, porque nos permite predecir más de una variable explicativa, y porque podemos observar la interacción entre variables explicativas e incluir variables de control.

3.3.1 – AJUSTE Y SELECCIÓN DEL MODELO CON UN PREDICTOR POLITÓMICO

Cuando tenemos una sola variable explicativa pero ésta es politómica, es decir, tiene más de dos categorías, estamos ante un modelo de regresión lineal múltiple, ya que es necesario codificar esta variable en $c-1$ variables ficticias, lo que da lugar a estimar $c-1$ parámetros más la constante (Peña, 2002).

Los modelos que se compararan mediante la estrategia de ajuste global para una variable explicativa con tres categorías serían:

$$\begin{aligned}
 MA : \hat{Y} &= b_0 + b_1 A_1 + b_2 A_2 \\
 MN : \hat{Y} &= b_0
 \end{aligned}
 \tag{3.20}$$

El ajuste global del modelo se hace comparando la discrepancia del MA, con la discrepancia del MN, utilizando la razón F :

$$F = \frac{(D_{MN} - D_{MA}) / (p_{MA} - p_{MN})}{D_{MA} / (n - p_{MA})} \quad (3.21)$$

De este modo podemos contrastar la hipótesis nula. Nos referimos al hecho de que el MA no reduzca el error de predicción respecto al MN.

El valor F lo comparamos con el valor crítico $F_{gl1, gl2, \alpha}$ donde $gl1 = p_{MA} - p_{MN}$ y $gl2 = n - p_{MA}$.

En el caso de que rechacemos H_0 , ello estaría indicando que al menos uno de los coeficientes b_i es estadísticamente distinto de 0, y no necesariamente que todos sean distintos de 0 (es decir, al menos hay diferencias entre una de las comparaciones).

No obstante, mediante esta prueba no podemos saber si el resultado significativo del modelo se debe a que hay diferencias entre dos categorías o entre todas. Para ello, deberemos acudir a la estrategia de ajuste parcial, que en este caso nos obliga a realizar dos ajustes parciales (Peña, 2002):

$$\begin{aligned} MA: \hat{Y} &= b_0 + b_1 A_1 + b_2 A_2 & MA: \hat{Y} &= b_0 + b_1 A_1 + b_2 A_2 \\ MR: \hat{Y} &= b_0 + b_2 A_2 & MR: \hat{Y} &= b_0 + b_1 A_1 \end{aligned} \quad (3.22)$$

En el primer caso, realizando la prueba de ajuste parcial, podemos comprobar si el efecto de la primera categoría es estadísticamente diferente de la categoría de referencia ($H_0: \beta_1 = 0$). Si este ajuste parcial es significativo, nos indicaría que $b_1 \neq 0$, y que hay diferencias estadísticamente significativas entre las dos categorías.

En el segundo caso, realizando la prueba de ajuste parcial, nos permite comprobar si el efecto de la segunda categoría es estadísticamente diferente de la

categoría de referencia ($H_0 : \beta_2 = 0$). Si este ajuste parcial es significativo, nos indicaría que $b_2 \neq 0$, y por lo tanto que hay diferencias estadísticamente significativas entre las dos categorías.

Para saber qué comparaciones son estadísticamente significativas debemos realizar una prueba F de ajuste parcial para cada una de las $c - 1$ variables ficticias. Para ello se compara el MA con un MR, que excluye una de las variables ficticias. Así, tendremos una $c - 1$ para cada coeficiente de regresión:

$$F = \frac{(D_{MR} - D_{MA}) / (p_{MA} - p_{MR})}{D_{MA} / (n - p_{MA})} \quad (3.23)$$

De este modo, el modelo seleccionado incluirá todas las variables ficticias o ninguna de ellas, y no se utilizan los modelos que sólo incluyen un subconjunto de variables ficticias. (Peña, 2002). Utilizando esta técnica obtendremos muchos datos como resultado de la comparación de modelos, por ello, es muy útil resumir toda esta información en un cuadro (ver tabla 3.3).

Tabla 3.2 Tabla resumen modelo lineal múltiple con un predictor polinómico

Modelo	Eliminado	ΔD	Δgl	F	P	R^2	b	$EE(b)$	$IC95\%$
MA (A_1, A_2)									
MN	A_1, A_2								
MR1(A_1)	A_2								
MR2(A_2)	A_1								
MF									

Además de toda la información incluida en esta tabla, podemos obtener más información calculando e interpretando el intervalo de confianza de los coeficientes β_j (que se obtiene a partir de la estimación del coeficiente de regresión y su error estándar).

También, deberemos tener en cuenta la interpretación de R^2 aplicable a la variable categórica.

3.3.2 – AJUSTE Y SELECCIÓN DEL MODELO CON DOS PREDICTORES

Los modelos que se compararan mediante la estrategia de ajuste global para dos variables explicativas serían:

$$\begin{aligned} MA : \hat{Y} &= b_0 + b_1A + b_2B \\ MN : \hat{Y} &= b_0 \end{aligned} \tag{3.24}$$

El ajuste global del modelo se hace comparando la discrepancia del MA, con la discrepancia del MN, utilizando la razón F :

$$F = \frac{(D_{MN} - D_{MA}) / (p_{MA} - p_{MN})}{D_{MA} / (n - p_{MA})} \tag{3.25}$$

De este modo podemos contrastar la hipótesis nula. Nos referimos al hecho de que el MA no reduzca el error de predicción respecto al MN. El valor F lo comparamos con el valor crítico $F_{gl1,gl2,\alpha}$ donde $gl1 = p_{MA} - p_{MN}$ y $gl2 = n - p_{MA}$.

En el caso de que rechacemos H_0 , ello estaría indicando que al menos uno de los coeficientes b_i es estadísticamente distinto de 0, y no necesariamente que todos sean distintos de 0.

No obstante, mediante esta prueba no podemos saber si el resultado significativo del modelo se debe a una o a las dos variables predictoras. Para ello, deberemos acudir a la estrategia de ajuste parcial, que en este caso nos obliga a realizar dos ajustes parciales (Peña, 2002).

$$\begin{aligned}
 MA : \hat{Y} &= b_0 + b_1A + b_2B & MA : \hat{Y} &= b_0 + b_1A + b_2B \\
 MR : \hat{Y} &= b_0 + b_2B & MR : \hat{Y} &= b_0 + b_1A
 \end{aligned}
 \tag{3.26}$$

En el primer caso, realizando la prueba de ajuste parcial, podemos comprobar si el efecto de la primera variable es estadísticamente significativo ($H_0 : \beta_1 = 0$). Si este ajuste parcial es significativo, nos indicaría que $b_1 \neq 0$, y que la variable A no puede excluirse del modelo.

En el segundo caso, realizando la prueba de ajuste parcial, nos permite comprobar si el efecto de la segunda variable es estadísticamente significativo ($H_0 : \beta_2 = 0$). Si este ajuste parcial es significativo, nos indicaría que $b_2 \neq 0$, que la variable A no puede excluirse del modelo.

Para saber qué comparaciones son estadísticamente significativas debemos realizar una prueba F de ajuste parcial para cada uno de los predictores. Para ello se compara el MA con un MR, que excluye uno de los predictores.

$$F = \frac{(D_{MR} - D_{MA}) / (p_{MA} - p_{MR})}{D_{MA} / (n - p_{MA})}
 \tag{3.27}$$

De este modo, el modelo seleccionado incluirá todas las variables o ninguna de ellas. De nuevo, es muy útil resumir toda esta información en la tabla 3.3.

Tabla 3.3 Tabla resumen modelo lineal múltiple con dos predictores.

Modelo	Eliminado	ΔD	Δgl	F	P	R^2	b	$EE(b)$	IC95%
MA (A,B)									
MR	A,B								
MR1(A)	B								
MR2(B)	A								
MF									

Además de toda la información incluida en esta tabla, podemos obtener más información calculando e interpretando el intervalo de confianza de los coeficientes β_j (que se obtiene a partir de la estimación del coeficiente de regresión y su error estándar). También, deberemos tener en cuenta la interpretación de R^2 aplicable a la variable categórica.

3.3.3 - COMPARACIÓN DE MODELOS

Ya hemos visto el ajuste global y el ajuste parcial que nos permiten seleccionar el modelo más parsimonioso. En el caso de la regresión múltiple además, es necesario determinar qué variables explicativas son relevantes. Por ello, debemos introducir un nuevo concepto, el de modelos “anidados”, que se da cuando un modelo es un caso particular de otro (los términos de un modelo menor deben estar incluidos en un modelo con mayor número de parámetros) (Dobson, 1983). Nuestro objetivo sigue siendo la elección del modelo más simple o MR siempre que su discrepancia no difiera significativamente de la discrepancia del modelo más complejo. El proceso es análogo al ajuste global, ya que se comparan ambas discrepancias mediante el estadístico F . La H_0 indicaría que no hay diferencias estadísticamente significativas (el MA no aporta una mayor explicación), si se rechazase seleccionaremos el MA, si se acepta nos quedamos con el MR.

Esta estrategia de comparación de modelos se denomina ajuste condicional, y nos permite evaluar, por una parte, la significación estadística de un parámetro mediante la formulación de dos modelos anidados (comparación focalizada), y por otro lado, la

significación estadística de un conjunto de parámetros, de manera que, el MA incluye el conjunto de parámetros y el MR los excluye (Dobson 1983).

Este es un proceso secuencial:

- Empezamos por el modelo máximo, que incluye todas las variables explicativas e iremos eliminando progresivamente términos del modelo cuyo efecto no sea significativo hasta que todos tengan significación o estén cercanos a tenerla.
- Si partimos del modelo nulo; comenzaremos por el modelo que no incluye variables explicativas e iremos añadiendo términos al modelo hasta que la adición de nuevos términos no mejore su ajuste.

3.3.4 - MODELOS CON INTERACCIÓN ENTRE VARIABLES

La formulación aditiva del modelo de regresión asume que el efecto de una variable explicativa es constante para el dominio de valores de las restantes variables incluidas en el modelo (Draper y Smith, 1981, 1998).

$$\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_j X_j + \varepsilon = \beta_0 + \sum \beta_j X_j + \varepsilon \quad (3.28)$$

Esta formulación tiene el inconveniente de que no nos permite contrastar que el efecto de una variable explicativa dependa de los valores que tomen otras variables incluidas en el modelo, o dicho de otro modo, la interacción entre variables. Para evitar estos inconvenientes podemos recurrir a la formulación indicada por Judd y McClelland (1989), denominado modelo de regresión “multiplicativo” o “interactivo”, también llamado “no lineal” o “polinómico”.

$$\hat{Y} = b_0 + b_1 X + b_2 Z + b_3 (XZ) \quad (3.29)$$

De este modo, el efecto principal de las variables está reflejado en b_1 y b_2 , y el efecto de interacción (o moderador) está reflejado en b_3 . Por lo tanto, la interpretación de estos coeficientes es más compleja que en un modelo aditivo, ya que incluye términos de interacción. Para facilitar esto, lo que se suele hacer es aislar la relación simple entre una de las variables explicativas y la respuesta, y analizar cómo cambia esta relación simple en función de los valores de una tercera variable. Para realizar esto, se agrupan los coeficientes de la ecuación de regresión que contienen y no contienen X , transformando la ecuación múltiple en una simple que es más fácil de interpretar (Judd y McClelland 1989)

$$\hat{Y} = (b_0 + b_2Z) + (b_1 + b_3Z)X \quad (3.30)$$

Así, b_0 es la constante de un modelo de regresión simple de Y sobre X cuando $Z = 0$, b_2 representa el cambio de la constante por cada incremento de una unidad en Z , b_1 es la pendiente de Y sobre X cuando $Z = 0$, y b_3 es el cambio que se produce en la pendiente por cada incremento de una unidad en Z .

Existe un procedimiento general para expresar la ecuación de regresión múltiple, y que se convierta en simple para la relación de X e Y propuesta por Judd y McClelland (1989). Esta propuesta considera el modelo como una función formada por un conjunto de k variables explicativas:

$$\hat{Y} = F(X_1, X_2, \dots, X_k) \quad (3.31)$$

3.3.5 - CÁLCULO DE LOS INTERVALOS DE CONFIANZA CON INTERACCIONES

En el caso de la regresión múltiple la interacción entre dos variables explicativas introduce algunas complejidades en el cálculo de los IC. Así por ejemplo, si tenemos tres variables, y las variables X y Z interaccionan, el efecto de X sobre Y , depende del valor de Z . Si $Z = 0$ el efecto de X es b_1 , pero si $Z \neq 0$ el cálculo se complica.

Existe un procedimiento propuesto mediante el cual se puede obtener el IC de X para los valores de $Z \neq 0$, desviando Z restando a sus valores el valor para el cual se desea obtener el IC de X , para después, regresar Y sobre X y la variable Z “desviada” (Judd y McClelland, 1989). Este cálculo es diferente en función de la naturaleza de la variable desviadora:

- Variable modificadora binaria: En este caso el coeficiente b_1 no da el efecto de X sobre Y para la categoría de Z codificada como 0 (o categoría de referencia). Para la categoría 1 de Z , esta pasara a ser la categoría de referencia recodificando los valores de Z cambiando 0 por 1. Para una mejor comprensión e interpretación de los coeficientes es aconsejable construir una tabla resumen.
- Variable modificadora politómica: en este caso el valor del coeficiente b_1 es el efecto de X para los sujetos de la categoría de referencia. De igual modo, recodificamos la variable de interacción cuando un conjunto de $k-1$ variables ficticias que reflejen esta nueva codificación.

- Variable modificadora continua: en este caso queremos obtener el efecto de X cuando Z toma el valor genérico $c(Z = c)$. Lo que haremos será centrar Z en el valor de referencia c, generando una nueva variable $Z_c = Z - c$, para después estimar el modelo de regresión de Y sobre X y Z. De este modo el IC de b_1 se corresponderá con el efecto de X para el valor c elegido.

3.4- LA COLINEALIDAD

La colinealidad ocurre cuando las variables explicativas están relacionadas entre si constituyendo una combinación lineal (se denomina multicolinealidad si son varias las variables explicativas relacionadas). La consecuencia es que no podremos distinguir la influencia de cada uno de ellas al quedar solapadas. Puede darse en diseños no experimentales y experimentales cuando los grupos están desequilibrados en tamaño. Si por el contrario están los grupos equilibrados debido a la aleatorización, no deberemos encontrarnos este problema (López, 1998).

Este es un aspecto muy importante, ya que la mayoría de las veces nos interesa conocer la aportación de cada una de las variables a la explicación de la respuesta. Para esto es muy práctico calcular el coeficiente de determinación R_{MA}^2 :

$$R_{MA}^2 = \frac{D_{MN} - D_{MA}}{D_{MN}} \quad (3.32)$$

Pero esto solo nos vale en el caso de una única variable explicativa. En regresión múltiple utilizaremos el incremento de $R^2 (\Delta R^2)$ que expresa la diferencia entre R_{MA}^2 y

R_{MR}^2 :

$$\Delta R_{MA}^2 = R_{MA}^2 - R_{MR}^2 = \frac{D_{MN} - D_{MA}}{D_{MN}} - \frac{D_{MN} - D_{MR}}{D_{MN}} = \frac{D_{MR} - D_{MA}}{D_{MN}} \quad (3.33)$$

A diferencia de R^2 que compara la explicación del MA con la del MN, ΔR_{MA}^2 compara la explicación del MA con la de un MR determinado, es decir, la explicación debida a ese término cuando el modelo ya incluye otros términos. De este modo, ΔR_{MA}^2 será menor al R_{MA}^2 que incluye ese término si existe colinealidad entre ese término y los que ya están incluidos en el modelo.

Existe otro índice que podemos utilizar denominado “reducción proporcional del error” o en inglés PRE. A diferencia de ΔR^2 , que mide la explicación de un término respecto al total de la variabilidad, este índice lo hace con respecto a lo que queda por explicar de dicha variabilidad. Expresa, por tanto, la reducción proporcional en la variabilidad que queda por explicar (o la reducción del error de predicción producida al incluir un término nuevo en el modelo).

$$PRE = \frac{R_{MA}^2 - R_{MR}^2}{1 - R_{MR}^2} = \frac{\Delta R_{MA}^2}{1 - R_{MR}^2} = \frac{(D_{MR} - D_{MA})/D_{MR}}{(D_{MN} - D_{MR})/D_{MN}} \quad (3.34)$$

Ambos índices son complementarios, y por lo tanto suelen utilizarse para evitar posibles conclusiones erróneas sobre la importancia de un término (López, 1998).

Si no existe colinealidad los valores de los efectos de las variables explicativas no difieren de los ajustados como modelos simples. Resumiendo, en ausencia de colinealidad los coeficientes de regresión b_j indican el efecto de cada variable explicativa. En el caso de que exista colinealidad los valores de b_j serán distintos ya que estiman los efectos de cada una de las variables explicativas.

Como ya ha quedado explicitado en esta tesis en el modelo de regresión lineal múltiple el coeficiente de una variable explicativa refleja el cambio (efecto) en la respuesta que se espera al incrementar en 1 unidad esa variable, y se conoce que dicho efecto se producirá para cualquier valor que tomen el resto de variable incluidas en el modelo; pero se desconoce si dicho cambio será el mismo para distintos valores de otras variables no incluidas en el modelo (McCullagh y Nelder, 1989).

3.5 - PROBLEMAS Y SOLUCIONES DE LA COLINEALIDAD

Podemos estudiar la colinealidad mediante (López 1998):

- El coeficiente de correlación múltiple al cuadrado (R_j^2): que se obtiene al regresar una variable explicativa sobre todas las demás variables explicativas del modelo.
- La tolerancia: que es la proporción de la varianza de esa variable no explicada por las otras variables que ya están en la ecuación ($1 - R_j^2$). Tendremos una tolerancia alta si obtenemos un valor cercano a 1 (la variable no está correlacionada con las demás), y tolerancia baja si obtenemos un valor cercano a 0 (existencia de relación de la variable con las demás), es decir, no aportaría nada al modelo.
- El factor de inflación de la varianza (FIV), que nos indica el impacto de la colinealidad sobre la precisión de los estimadores de los coeficientes de regresión: $\frac{1}{1 - R_j^2}$ (ver tabla 3.4).

Tabla 3.4: Resultados posibles del estudio de la colinealidad.

Tolerancia	FIV	Colinealidad
1	1	No hay colinealidad
1 - 0,3	1 - 3,33	Baja: el modelo no presenta efectos importantes
0,1-0,3	10-3,33	Alta: inconsistencia de los estimadores y pérdida de significación
0,1 - 0,01	10 - 100	Excesiva: graves problemas en la interpretación de los predictores
> 0,01	< 100	Colinealidad perfecta

Por otra parte, hay que tener en cuenta que la colinealidad no afecta a la capacidad predictiva del modelo, sino a la precisión e interpretación de los coeficientes de regresión de las variables colineales, que nos pueden llevar a conclusiones incorrectas. Si el modelo es predictivo, la colinealidad no nos afecta, pero en el caso de ser explicativo afectará a la valoración de los efectos de las variables explicativas. Esto ocurre en los casos de colinealidad elevada, pero también existe la colinealidad perfecta o singularidad, que se da cuando una variable explicativa representa una combinación lineal del resto de variables (Belsley, *et al.*, 2004)

Existen 3 situaciones en las que podemos encontrarnos en esta situación: en primer lugar, se da en el caso de incluir variables linealmente relacionadas, en segundo lugar, si en vez de introducir $k-1$ variables indicadora incluimos k variables, y en tercer lugar, si el tamaño de la muestra es pequeño, (cuando n es menor que el número de coeficientes a estimar) ver en Berry (1993).

Para solucionar estos problemas de colinealidad podemos: aumentar el tamaño de la muestra, combinar dos o más variables explicativas correlacionadas en una sola,

eliminar una de las variables explicativas si no es posible combinarla con otras (no es aconsejable), y centrar las dos variables explicativas (Cronbach 1987).

3.6 - MODELOS CON MEDIDAS REPETIDAS

Tendremos dependencia entre observaciones en estas situaciones (Oliver, *et al.*, 2000):

- *Medidas apareadas*: esto ocurre cuando se registran dos o más observaciones de una misma variable (o categoría).

- *Medidas repetidas*: ocurre cuando se registran dos o más observaciones de un mismo concepto provenientes de un mismo sujeto. Este tipo de medidas son menos variables que las medidas entre sujetos diferentes, debido a que en las mismas condiciones la variabilidad intra es más constante que la variabilidad entre. Lo importantes es medir el mismo concepto y en la misma escala de medida.

- *Datos secuenciales*: ocurre cuando se registra una misma información para cada sujeto del estudio varias veces a lo largo del tiempo, siendo la variable tiempo de interés para el estudio como variable explicativa de las diferencias entre medidas.

Para poder analizar estos datos, hay que eliminar esta dependencia, normalmente transformando los datos (en el caso de medidas apareadas y repetidas).

Los tipos de diseño que proporcionan este tipo de datos son:

- *Diseños entre sujetos*: donde la hipótesis que se suelen plantear son acerca de la diferencia entre distintos grupos de sujetos (cada sujeto solo pertenece a una categoría de la variable explicativa).
- *Diseños intra sujetos*: donde las hipótesis que se suelen plantear son acerca de la diferencia entre las medidas (repetidas o apareadas) sin la intervención de las variables que definen los grupos de sujetos (cada sujeto tiene un valor en cada categoría de la variable).
- *Diseños mixtos*: donde las hipótesis que se plantean son acerca de las diferencias entre las medidas y entre los grupos de sujetos.

3.6.1. DISEÑOS CON MEDIDAS REPETIDAS O APAREADAS

En este tipo de diseños las variables explicativas varían entre sujetos, teniendo más de una medida por sujeto, siendo éstas dependientes. Para el correcto análisis de este tipo de diseños deberemos transformar los datos combinando las medidas dependientes en una sola medida que elimine la variabilidad intra sujeto, y que podamos regresar sobre las variables explicativas entre sujetos de interés.

Tomaremos cada observación como y_{ih} , siendo i el número de sujeto y h el número de medida realizada. Obtendremos la medida combinada a partir de una agregación (P_0) de estas medidas originales, ponderando cada medida h por un factor δ_h que nos permite dar mayor importancia a una medida que a la otra, asignando el valor 1 a cada medida h .

$$P_0 = \frac{\sum_h \delta_h y_{ih}}{\sqrt{\sum_h \delta_h^2}} \quad (3.35)$$

El procedimiento a seguir será, generar la medida combinada P_0 para cada uno de los sujetos, utilizándose ésta como variable de respuesta en la comparación de los modelos para testar el efecto de la variable explicativa. Y para evaluar su significación estadística deberemos comparar un MA incluyendo esta variable con un MR que no la incluya:

$$\begin{aligned} MA: \hat{P}_0 &= b_0 + b_1 X \\ MR: \hat{P}_0 &= b_0 = \bar{P}_0 \end{aligned} \quad (3.36)$$

Para la interpretación de los coeficientes de regresión habrá que tener en cuenta que se hallan en la escala de medida de la variable combinada, pero se pueden convertir a los valores originales dividiéndolos por el denominador de P_0 . En el caso de los intervalos de confianza el procedimiento es el mismo.

En definitiva, mediante este procedimiento se colapsan las observaciones de cada sujeto en una única medida, eliminando de la discrepancia del modelo la variabilidad intra sujeto.

3.6.2. DISEÑOS DE MUESTRAS RELACIONADAS

En estos diseños la variable explicativa varía intra sujeto (cada sujeto proporciona más de una respuesta que será dependiente en cada categoría de la variable explicativa). Al igual que con los diseños anteriores habrá que transformar los datos para eliminar la dependencia, pero en este caso generaremos una variable que refleje la diferencia entre las medidas para cada sujeto (así contrastaremos la hipótesis sobre

diferencias entre los niveles del factor intrasujeto). Además, hemos de comprobar si la media de la nueva medida de diferencias es significativamente distinta de cero, a partir de una prueba de conformidad. Si es así, entonces asumimos que existen diferencias (la variable explicativa incide en la respuesta). Si es estadísticamente igual a cero asumimos que no existen diferencias.

Por otra parte, cuando realizamos el análisis de la diferencia entre dos medias, y partimos de que la variable explicativa intra sujeto tiene dos categorías, entonces debemos generar una variable que recoja la diferencia entre las medidas. La media de esta variable es la que contrastaremos para ver si es estadísticamente igual a cero. Esta nueva variable es:

$$d_1 = \frac{\delta_1 y_{i1} + \delta_2 y_{i2}}{\sqrt{\delta_1^2 + \delta_2^2}} \quad (3.37)$$

donde δ es un valor de ponderación que puede ir entre 1 y -1 , pero es mejor utilizar 0,5 y $-0,5$ para que la diferencia entre los dos valores sea la unidad.

Para realizar la prueba de conformidad, debemos utilizar la razón F , comparando un MA que predice la media observada y un MR que predice el valor cero:

$$\begin{aligned} MA : d_1 &= \bar{d}_1 \\ MR : \hat{d}_1 &= 0 \end{aligned} \quad (3.38)$$

El numerador de la razón F podemos calcularlo mediante la siguiente fórmula:

$$D_{MR} - D_{MA} = n\bar{d}_1^2 \quad (3.39)$$

En el caso de partir de una variable explicativa intrasujeto que tiene más de dos categorías, habrá que generar $c-1$ variables de diferencia. Para generar estas variables codificaremos el valor de interés mediante contrastes ortogonales. Así conseguimos que

la variabilidad total que atribuiremos a la variable explicativa se descomponga de forma independiente entre las $c - 1$ variables generadas:

$$d_i = \frac{\sum_h \delta_h y_{ih}}{\sqrt{\sum_h \delta_h^2}} \quad (3.40)$$

donde δ_h son los valores de ponderación correspondientes a los coeficientes de un conjunto de $c - 1$ contrastes ortogonales (definidos para codificar la variable explicativa). Al igual que en los diseños anteriores la media de cada variable d_i se debe comparar con cero:

$$\begin{aligned} MA : \hat{d}_1 &= \bar{d}_1 \\ MR : \hat{d}_1 &= 0 \end{aligned} \quad (3.41)$$

Finalmente habrá que obtener el valor de la prueba F global del análisis de la $H_0 : \mu_1 = \mu_2 = \dots = \mu_c$, que evalúa si existen diferencias en el conjunto de las c medidas repetidas

3.6.3. DISEÑOS MIXTOS

Como sabemos, en estos diseños al menos una de las variables explicativas cambia entre sujetos y otra intra sujeto. Por consiguiente, habrá que analizar las variables tanto de forma independiente como su interacción. Vamos a ver el caso de una variable entre y otra intra con dos categorías cada una.

Como en los casos anteriores habrá que generar una nueva variable d_1 que refleje la diferencia entre las categorías de la variable intra sujeto y utilizarla como variable de respuesta en un modelo que incluya la variable entre sujetos como variable

explicativa. Con ello pretendemos comprobar si las diferencias entre categorías de la variable intra sujeto se mantienen en las dos categorías de la variable entre sujetos.

$$d_1 = \frac{\sum_h \delta_h y_{ih}}{\sqrt{\sum_h \delta_h^2}} \quad (3.42)$$

Para comprobar el modelo se sigue la misma lógica que en el análisis de muestras independientes con medidas repetidas o apareadas (pero con d_1 en vez de P_0).

$$\begin{aligned} MA : \hat{d}_1 &= b_0 + b_1 X \\ MR : \hat{d}_1 &= b_0 \end{aligned} \quad (3.43)$$

En el caso de tener variables con más de dos categorías o mas variable entre o intra, habrá que generar un mayor número de variables de diferencia d_i , y variables explicativas. En la tabla ANOVA, que resume la comparación de modelos, debemos añadir las discrepancias y grados de libertad atribuibles a las fuentes de variabilidad que se excluyen del análisis, por lo que la tabla ANOVA debe incluir las fuentes de variabilidad con sus correspondientes sumas cuadráticas y grados de libertad.

3.7 CONCLUYENDO

En este capítulo se ha tratado el modelo de regresión lineal y sus extensiones desde la perspectiva del modelado estadístico. Este capítulo nos sirve de introducción didáctica a estas técnicas ya que la mayoría de investigadores están familiarizados con la regresión lineal y de este modo, podemos comprender mejor cómo funciona el modelado. En el contexto de este trabajo las medidas psicofísicas registradas en los experimentos pueden ser analizadas con los modelos lineales presentados en el capítulo permitiéndonos obtener un detallado análisis de los datos.

Una vez introducida esta técnica, pasaremos a repasar los modelos lineales generalizados ya que nuestros datos de memoria de reconocimiento no solo son psicofísicos, sino que además tenemos una serie de variables categóricas que pueden ser analizadas desde estos modelos.

CAPÍTULO 4

INTRODUCCIÓN A LOS MODELOS LINEALES GENERALIZADOS

Los fundamentos de la estadística están cambiando, no sólo en el sentido en que ellos fueron y continuarán evolucionando, sino también en el sentido idiomático de que ningún sistema es absolutamente estable. Leonard Savage.

4.1 – INTRODUCCIÓN

Los modelos lineales generalizados (MLG) son una extensión de los modelos lineales propuestos por Nelder y Wedderburn (1972) donde demostraron que la técnica de regresión lineal, podía ser utilizada para obtener estimadores de máxima verosimilitud de los parámetros, con observaciones distribuidas de acuerdo a la familia exponencial natural, y los efectos sistemáticos podían convertirse a una escala lineal mediante una transformación adecuada. Esta familia de modelos unifica los modelos con variables de respuesta cuantitativa y cualitativa (categóricos), utilizando para ellos además de la distribución normal, las distribuciones Binomial, de Poisson, etc.

Los MLG poseen una gran ventaja sobre los modelos lineales, ya que nos permiten la elección de la distribución de la variable dependiente dentro de una gama de distribuciones amplia, y además nos permiten un análisis estadístico del modelo válido para cualquier distribución que se haya escogido. Ello concuerda con la distribución de los datos en la realidad (McCullagh y Nelder 1989).

El MLG nos permite el modelado de cualquier variable de respuesta que se ajuste a una distribución de la familia exponencial, siempre que se cumplan unos supuestos básicos. El concepto de la familia exponencial fue introducido por Pitman, Darmois y Koopman en 1935 (ver tabla 4.1).

Tabla 4.1 Diferencias entre el modelo lineal y el MLG.

	Modelo Lineal	MLG
Parámetros	$\beta_0, \beta_1, \dots, \beta_k$	$\beta_0, \beta_1, \dots, \beta_k$
Estimación	Minimos cuadrados	Máxima verosimilitud
Bondad de ajuste	R^2	Desvianza
Comparación de modelos	AIC, BIC, F	AIC, BIC, Desvianza
Supuestos	Muy restrictivos	Menos restrictivos

Así, un modelo lineal generalizado es un modelo que vincula las respuestas (variables dependientes) con las variables de estímulo (independientes o explicativas).

Según lo anterior hemos de considerar tres componentes:

1. El componente sistemático, que indica la relación entre las variables independientes. Éste es un modelo lineal.
2. El componente aleatorio (la distribución de los y_i). En general, se supone que los y_i son independientes, con una distribución de la familia exponencial
3. La función de enlace, que es la que vincula la distribución de los y_i con el componente sistemático.

4.2 – COMPONENTES DEL MLG

Como ya hemos indicado, el modelado estadístico consiste en encontrar la función que mejor se ajuste a los valores de respuesta obtenidos. El objetivo de un investigador al realizar el modelado estadístico de una variable de respuesta, a partir de un conjunto de variables explicativas, consiste en encontrar la función que mejor se ajusta a los valores de respuesta (McCullagh, 2002).

Normalmente se utiliza la función lineal, que expresa la variable de respuesta a partir de una combinación lineal de los valores de las variables explicativas, como en los modelos de regresión lineal. A diferencia de los modelos lineales, el MLG tiene tres componentes:

1 – *Componente sistemático:*

Es el componente que reúne la variabilidad de Y explicada por X a partir de una combinación lineal de la variable X , incluyendo los parámetros β_0 y β_1 . En el MLG este componente se simboliza con η y representa el valor predicho, recibiendo el nombre de predictor lineal.

$$\eta_i = \beta_0 + \beta_1 X_i \quad (4.1)$$

Este componente es el que nos permite modelar la variación en la variable de respuesta por medio de un conjunto de variables explicativas, estimando sus coeficientes a partir de los datos observados de la muestra. Además, podemos incluir, las variables explicativas originales, potencias de variables, transformaciones de variables, así como las interacciones entre variables explicativas y variables de control (Agresti 1996, 2007).

Cuando las variables explicativas sean categóricas se codificaran por medio de un conjunto de $k-1$ variables ficticias. A diferencia del modelo de regresión lineal donde $\eta_i = \mu_i$, el modelo proporciona directamente los valores esperados. Sin embargo, en el MLG hay que incluir un tercer componente que relacione los dos anteriores (el valor esperado y el valor predicho).

2 - *Componente aleatorio:*

El valor observado y_i es una realización de la variable aleatoria Y , cuyo valor esperado viene dado por $E(y_i) = \mu_i$.

3 - *La función de enlace:*

Simbolizada como $g(\mu_i)$, transforma el valor esperado a la escala del predictor lineal.

$$g(\mu_i) = \eta_i \tag{4.2}$$

Así, la inversa de la función de enlace nos permite que al aplicar al resultado del predictor lineal η_i , obtener el valor esperado μ_i (que está en la escala de la variable de respuesta):

$$\mu_i = g^{-1}(\eta_i) = g^{-1}(\beta_0 + \beta_1 X_i) \tag{4.3}$$

Esta función debe asumir que la función sea conocida, monotónica y diferenciable. Es el investigador el que debe decidir cual es la función de enlace más adecuada en cada caso (Czado, 1997). Otra de las formas de seleccionar la función de enlace más adecuada, es la denominada “función de enlace canónica” que define el parámetro canónico de la distribución (ver tabla 4.2).

Tabla 4.2 Funciones de enlace canónicas (adaptado de Ato 1996)

Distribución	Función de enlace	
Normal	Identidad	$g(\mu_i) = \mu_i$
Poisson	Logaritmica	$g(\mu_i) = \log(\mu_i)$
Binomial	Logit	$g(\mu_i) = \log\left(\frac{\mu_i}{n_i - \mu_i}\right)$
Multinomial	Logit generalizado	$g(\mu_i) = \log\left(\frac{\mu_i}{\mu_i - \mu_j}\right)$

4.3 – CLASIFICACIÓN DE LOS MLG

La clasificación de los MLG se realiza considerando las siguientes características (Ato 1996):

- 1) Relación entre las variables: distinguiremos entre modelos simétricos y asimétricos. Los primeros tienen como objetivo el estudio de 2 o más variables con el mismo rol dentro del modelo, y los segundos tienen como objetivo la explicación o predicción de una respuesta a partir de un conjunto de variables explicativas.
- 2) Respuesta registrada: la variable de respuesta puede ser continua, categórica, binaria o politómica, de tiempo entre eventos y de recuento.
- 3) Respuesta modelada: podemos realizar el modelado sobre la variable de respuesta directamente o agrupando los datos. En este último caso, todas las variables explicativas deben ser categóricas, o continuas categorizadas mediante la agregación de los valores de respuesta de los sujetos que tienen un mismo patrón de valores en las variables explicativas. En el caso de trabajar con datos agrupados la respuesta modelada es diferente a la registrada. Si los datos individuales se han agrupado la respuesta modelada y registrada coinciden.
- 4) Componente sistemático: Distinguimos entre componente categórico como aquel componente organizado a partir de variables cualitativas, y mixto

como el componente que incluye términos continuos y categóricos mediante algún tipo de codificación. Pero si todos los términos son continuos también tendemos un componente sistemático mixto.

- 5) Componente aleatorio: en el MLG la distribución de la probabilidad del error corresponde a la familia exponencial. Los más utilizados son los de Poisson, Binomial y Multinomial en función de la naturaleza de los datos.

- 6) Función de enlace: se utilizará la función “logit” para distribuciones binomiales, “logit” para distribuciones de Poisson, y “logit” generalizado para las distribuciones multinomiales, para transformar el valor esperado a la escala del predictor lineal y así, poder realizar la correcta interpretación del modelo.

Tabla 4.3 Clasificación de modelos según los MLG (adaptado de Ato 1996).

Relación entre	Respuesta registrada	Respuesta modelada	Componente sistemático	Componente aleatorio	Funcion de enlace	Modelo
Asimétrica	Continua	Continua	Mixto	Normal	Identidad	Regresión Lineal
			Categórico	Normal	Identidad	ANCOVA
	Recuento	Recuento	Mixto	Poisson	Logarítmica	ANOVA
	Tiempo transcurrido	Tiempo transcurrido	Mixto	Poisson	Logarítmica	Regresión de Poisson
	Categórica binaria	Binaria	Mixto	Bernoulli	Logit	Regresión de Cox
		Frecuencias	Categórico	Binomial	Logit	Regresión Logística
					Probit	Regresion Logit
	Categórica polinómica	Polinómica	Mixto	Multinomial	Logit generalizado	Regresión Probit
		Frecuencias	Categórico	Multinomial	Logit generalizado	Regresión Logística Multinomial
	Simétrica	no hay VD	Frecuencias	Categórico	Poisson	Logarítmica
						Modelo loglineal

A esta clasificación podríamos añadir las variables de estímulo. Por ejemplo las probabilidades a priori de la frecuencia de las palabras y los procesos de transición entre los estímulos y las variables de respuesta, que utilizaremos en este trabajo.

4.4 – SUPUESTOS DEL MLG

Al igual que en los modelos lineales, es necesario cumplir una serie de supuestos, aunque en este caso sean menos restrictivos (McCullagh y Nelder 1989):

- Independencia de las observaciones
- Especificación correcta del modelo (será incorrecta si hay una ausencia de términos relevantes en el componente sistemático).
- Comprobar la presencia de observaciones problemáticas.
- Especificar correctamente la función de enlace
- Comprobar la presencia de “outliers” o valores mal justados por el modelo
- Comprobar la presencia de valores influyentes sobre los coeficientes estimados.

4.5 – ESTIMACIÓN DE PARAMETROS

En el caso de la distribución del componente aleatorio no normal se utiliza la Máxima Verosimilitud (ML). Cuando trabajamos con los OLS podemos aplicarlos sin realizar hipótesis sobre la distribución de la variable de respuesta, pero para la ML se hace necesario asumir dicha distribución. Ello es debido a la definición de estos modelos a partir de la probabilidad conjunta de las n observaciones (Lindsey, 1997).

En el MLG utilizaremos, para la estimación ML, el algoritmo de Mínimos Cuadrados Iterativamente ponderados (IWLS).

Los pasos a seguir en este algoritmo consisten en (Eliason, 1993):

- a) Estimar los parámetros, aplicando la función de enlace sobre los valores observados, obteniendo un vector G que se regresa sobre las variables explicativas X por medio de OLS. Es decir, $\hat{\beta}_j = (X'X)^{-1}(X'G)$.

- b) Una vez obtenidas estimados los parámetros calculamos los valores predichos $\hat{\eta}_i$, transformándolos por medio de la inversa de la función de enlace y obteniendo los valores esperados m_i .

- c) Obtenidos los valores predichos y esperados, construiremos una variable de respuesta de trabajo Z_i (la diferencia entre valores esperados y observados) y una matriz de ponderación W (la diagonal principal de esta matriz representa las varianzas de los valores esperados m_i).

- d) Posteriormente regresamos la variable Z_i sobre las variables explicativas X , teniendo en cuenta los valores de ponderación W , y obteniendo una estimación WLS mediante la siguiente ecuación: $\hat{\beta}_j = (X'WX)^{-1}(X'WZ)$

- e) Por último, hemos de repetir los pasos b y d de forma iterativa hasta cumplir el criterio de convergencia que se haya determinado.

Hemos de indicar que es muy importante obtener la estimación del error estándar de $\hat{\beta}_j$ a través de $V(\hat{\beta}_j) = (X'WX)^{-1}$, que representa es la matriz de varianzas-covarianzas de $\hat{\beta}_j$.

4.6 – AJUSTE Y SELECCIÓN DEL MODELO

Una vez que tenemos los parámetros del modelo debemos seleccionarlos y ajustarlos. Esta tarea consiste en proponer modelos con diferentes parámetros. Para ello nos basamos en la discrepancia entre los datos observados y los esperados.

4.6.1. MODELOS

Los modelos posibles son (Dobson, 1990, 2002):

En el modelo saturado (MS) el número de parámetros estimados es igual al número de observaciones. En el caso de datos individuales el número de observaciones es igual al tamaño de la muestra. En el caso de tener datos agrupados el número de observaciones es igual al número de casillas de la tabla de contingencia. No se utiliza con datos individuales ya que el número de parámetros a estimar sería igual al del tamaño muestral. Este modelo reproduce los datos observados de una manera exacta.

El modelo máximo (MMax) es el modelo más complejo. Éste incluye las variables explicativas, las variables de control y los términos de interacción. La elección del número de términos a incluir en el MMax (p) está en relación directa con el tamaño muestral. Hay que tener en cuenta que el número de grados de libertad del tamaño del error debe ser positivo $n > p - 1$. Con datos agrupados MMax = MS.

El modelo mínimo (MMin) incluye un número mínimo de parámetros que debe ser estimado. Está vinculado al muestreo utilizado y requiere la inclusión de un factor de bloqueo en el caso de diseños de bloques aleatorios.

El modelo nulo (MN) incluye un solo parámetro, proporcionando el mismo valor esperado para todas las observaciones. Normalmente sirve de referencia cuando se modelan datos individuales ya que no representa la estructura de los datos de una manera adecuada.

El modelo cero (MC) no estima ningún parámetro (es el investigador el que fija el valor de los parámetros). Se utiliza para las pruebas de conformidad sobre el valor del parámetro.

El modelo de trabajo (MT): es el objeto de comparación en cada paso del proceso de ajuste y selección del modelo final”. Estará entre el MMax y el MMin. (Lindsey 1997).

En resumen, como ya vimos anteriormente, en el modelado el objetivo es encontrar el modelo que reproduzca más fielmente los datos observados con un menor número de parámetros.

4.6.2- AJUSTANDO MODELOS

Una vez obtenido el modelo adecuado, hemos de seguir aplicando el ajuste de los modelos observados en relación con los esperados. Para obtener la bondad de ajuste

comparamos la diferencia entre valores observados y esperados. En MLG esto se denomina discrepancia (D) (Gill, 2001).

El cálculo se realiza de manera diferente según la distribución del componente aleatorio y su interpretación cambia en función de si los datos son individuales o agrupados.

Estos cálculos son necesarios para seleccionar el modelo óptimo. Para ello se comparan pares de modelos. En efecto, dos modelos están anidados si uno es un caso particular de otro (los términos del modelo menor están incluidos en el modelo mayor). Estas comparaciones se hacen utilizando el test de razón de verosimilitud ΔD .

Mediante el ajuste global trabajamos con datos agrupados el MT se compara con el MS (el ajuste global compara valores esperados y valores observados). En este caso la D del MS es igual a 0, y por lo tanto la diferencia de discrepancias es igual a la D del MT (la D del MT proporciona el ajuste global del modelo por sí mismo) (McCulloch y Searle, 2001).

Cuando partimos de datos individuales el MT se compara con el MN (se comparan los valores esperados por cada uno de ellos). En este caso la D no es interpretable por sí misma: $\Delta D = D_{MN} - D_{MT}$. Esta prueba tiene una distribución asintótica χ^2 , y su significación se obtiene situando el valor obtenido en la distribución χ^2 con los grados de libertad obtenidos de la diferencia de grados de libertad de los dos modelos (la diferencia de parámetros de los dos modelos.). En este caso la prueba ΔD evalúa si la adición de términos del MT reduce significativamente el valor D con

respecto al MN. En el caso de los datos agrupados evalúa si la eliminación de términos del MT con respecto al MS incrementa significativamente la D (un valor estadísticamente significativo nos lleva a la selección del MS).

Mediante el ajuste condicional comparamos las D de dos modelos anidados con el objetivo de escoger el modelo más simple (MR), cuya D no difiera significativamente de la D del modelo más complejo al que está anidado (MA). También utilizamos la prueba ΔD pero con los siguientes cambios: $\Delta D = D_{MR} - D_{MA}$. La significación se obtiene a través de la distribución χ^2 con los grados de libertad obtenidos de la diferencia de grados de libertad de los dos modelos. Esto nos permite evaluar la significación estadística de un parámetro mediante la formulación de dos modelos anidados siendo el MR el modelo que excluye parámetros. A su vez, también podemos evaluar la significación estadística de un subconjunto de parámetros formulando los correspondientes MA y MR.

4.6.3- LA SELECCIÓN DEL MODELO

Completado lo anterior, hemos de elegir el mejor modelo (Kadane 2004). Así, cuando el objetivo es predictivo hay que seleccionar las variables que expliquen el mayor porcentaje de variabilidad de respuesta, utilizando los criterios estadísticos en forma de algoritmos automáticos de selección.

Sin embargo, cuando el objetivo es explicativo: priman los criterios teóricos, siendo el investigador el que guía el proceso de selección, especificando un modelo inicial y un conjunto sucesivo de modelos restringidos, comprobándolos mediante una

estrategia de ajuste condicional; pero teniendo en cuenta una serie de aspectos. En efecto, en primer lugar, debemos respetar el principio de modelo jerárquico (el termino que introduzcamos deben incluir todos sus términos de orden inferior). En segundo lugar, deberíamos evaluar los términos de interacción a través de su significación estadística. En tercer lugar, analizaremos la posibilidad de incluir o excluir variables de control (confundidoras o de ajuste); y por ultimo decidiremos qué variables explicativas deben permanecer o no en el modelo.

4.7 – INTTERPRETANDO EL MODELO AJUSTADO

Para la interpretación del modelo debemos tener en cuenta que (Liao 1994):

En primer lugar, hemos de asumir que la significación estadística no puede ser el único criterio válido utilizado para la interpretación de los parámetros del modelo. Así, es conveniente el cálculo de los intervalos de confianza, además de la selección del modelo en base a criterios como: relevancia científica, tipo de diseño, etc.

Así, cuando trabajamos con MLG, al existir una función de enlace, e interpretar los coeficientes, hay que realizar la transformación a la escala de la variable de respuesta. De este modo la interpretación de los parámetros se hace en términos del factor de cambio en el valor esperado de respuesta asociado al incremento unitario de X .

Por otra parte, para la interpretación de los términos de interacción, debemos desviar relaciones simples. Además, es muy aconsejable utilizar tablas resumen para la interpretación de los coeficientes (Kam y Franzese, 2007) ver tabla 4.4.

Tabla 4.4 Tabla para la interpretación de los coeficientes de un modelo

Términos	b	$EE(b)$	gl	z	P	$exp(b)$	$IC95\%exp(\theta)$
Constante							
X							
Z							
X * Z							

Por último, en la predicción por intervalo de los valores esperados, utilizaremos el procedimiento de estimación correspondiente a los parámetros β_j . Ello en el caso en que haya interacción.

4.8 CONCLUYENDO

En este capítulo hemos expuesto como el MLG amplía el modelo lineal general, de manera que la variable dependiente está relacionada linealmente con los factores y las covariables mediante una determinada función de enlace. Esto permite que la variable de respuesta tenga una distribución no normal, ampliando de manera considerable los tipos de datos que podemos utilizar en nuestras investigaciones y que detallaremos en los próximos capítulos.

Así, el MLG cubre los tipos de datos más utilizados, cuantitativos (regresión lineal), frecuencias (modelos loglineales), recuento (regresión de Poisson), datos binarios (modelos logísticos), etc., además de muchos otros modelos estadísticos a través de la propia formulación general del modelo. Esto nos abre un abanico de posibilidades analíticas muy amplio permitiéndonos utilizar variables categóricas aumentando el valor de los resultados que podemos obtener en nuestras investigaciones.

CAPÍTULO 5

MODELOS LOGLINEALES

5.1 – INTRODUCCIÓN

Una vez expuestos los modelos, y por problemas de espacio, desarrollaremos los modelos de loglineales, de Poisson, logístico y logit, debido a que son los más frecuentes en nuestro ámbito, dejando para posteriores trabajos los restantes modelos.

Los orígenes de esta prueba se remontan a Birch (1963) que formuló el modelo loglineal, posteriormente Goodman (1970) presentó el análisis multivariante de datos cualitativos para tablas de contingencia. Fienberg (1976) presentó las formas de análisis multivariante para clasificación cruzada de datos categóricos. Así, los modelos loglineales nos muestran las asociaciones entre las variables de igual modo que los modelos de regresión y el ANOVA usando modelos jerárquicos. Bishop (1975) presentó la estimación máximo verosímil a tablas completas, aplicándola a tablas de contingencia 2x2 y 2x3, luego generalizó la metodología para el caso multidimensional e interpretó los modelos loglineales en función de la razón de productos cruzados y al principio de jerarquía. También describe el estadístico de verosimilitud G^2 el cual ayuda a determinar el modelo que se ajusta mejor a los datos categóricos.

Finalmente, Agresti (1990) recopiló los trabajos realizados en modelos log-lineales utilizando una nueva notación, pero además, las generalizó utilizando las condiciones de colapsibilidad, mostrando su aplicación a datos nominales y ordinales.

Los modelos loglineales, permiten analizar la interacción entre todas las variables de una tabla de contingencia simultáneamente, incluyendo las interacciones entre grupos de variables. Se trata de ajustar un modelo que busca predecir la frecuencia esperada en

cada celda de la tabla como un producto de los efectos principales y sus interacciones (Christensen, 1997). Cuando no hay ningún efecto significativo, el valor esperado en cada celda es igual al total de observaciones dividido por el número de celdas; si los totales marginales son diferentes, se debe a la presencia de algún efecto. Si un modelo se ajusta a los datos, es decir, explica de manera satisfactoria las relaciones existentes entre las variables, el valor esperado en cada celda no se diferencia significativamente de la frecuencia observada en la misma; si son diferentes, debe probarse otro modelo. Además, dependiendo de los parámetros que resulten significativos en el modelo ajustado, este tipo de análisis permite describir las características de las variables y sus interacciones.

En general, se trata de ajustar modelos donde la variable dependiente es la frecuencia esperada en la celda, bajo la hipótesis nula de independencia de los efectos.

En los modelos loglineales (a diferencia de la mayoría de los modelos que expondremos) no distinguimos entre variables explicativas y de respuesta, y sólo incluyen variables categóricas. Sus principales aplicaciones son las tablas de contingencia y el análisis de la asociación entre variables. El modelo especifica la forma en que las frecuencias esperadas dependen de los niveles de las variables categóricas para cada celda, así como de sus interacciones.

Los modelos loglineales se caracterizan porque (Fienberg y Rinaldo, 2007):

- No existe distinción entre variables explicativas y variables de respuesta (son modelos simétricos).
- El componente sistemático está formado por variables categóricas.

- El componente aleatorio sigue la distribución de Poisson (ver figura y tabla 5.1)
- La función de enlace es la transformación logarítmica $\eta_i = \log(\mu_i)$

Superan las principales limitaciones de la aplicación tradicional de la distribución ji-cuadrado en una, tabla de contingencia multidimensional:

- Confusión en la determinación de la asociación marginal entre un par de variables categóricas con respecto a otras variables presentes en el estudio.
- Incapacidad para probar los casos de independencia ya sea conjunta o parcial entre las variables categóricas en el experimento.
- Incapacidad para realizar el análisis simultáneo de la asociación de pares de variables.
- Desconocer la posibilidad de interacciones de tercer orden y de orden más alto entre las variables.

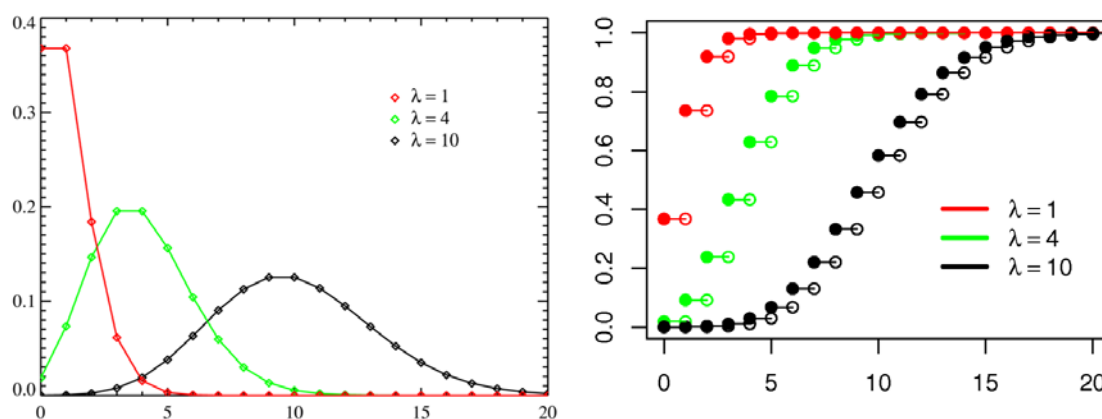


Figura 5.1 Función de densidad y función de la distribución de Poisson (Wikipedia)

Tabla 5.1 Características de la distribución de Poisson.

Distribución	$y_i \sim P(\mu_i)$
Parámetro canónico	$\log(\mu_i)$
Parámetro de escala	1
Función de enlace	$\eta_i = \log(\mu_i)$
Función de varianza	$V(\mu_i) = \mu_i$
Desviación	$D = \sum w [y_i - \log(y_i / \hat{\mu}_i) - (y_i - \hat{\mu}_i)]$

Inconvenientes de los modelos loglineales (Von Eye y Mun 2013):

- Colapsabilidad: se produce cuando se omiten variables relevantes de una tabla de contingencia.
- La Paradoja de Simpson: es el cambio en el sentido de una asociación entre dos variables cuando se controla el efecto de una tercera variable, normalmente se produce cuando se colapsa de manera inadecuada una tabla.
- Tablas dispersas: cuando tenemos frecuencias esperadas pequeñas o valores cero.

Inicialmente explicaremos los modelos más sencillos (con dos variables), para posteriormente ir aumentando el número de variables, lo que evidentemente aumenta su complejidad. En el ámbito de la memoria de reconocimiento pueden ser muy útiles estos modelos, ya que es habitual la presencia de variables categóricas como el tipo de respuesta del sujeto, frecuencias *a priori*, condición experimental, escalas de estimación, estímulos, respuestas, sesión, etc.

5.2 MODELOS LOGLINEALES CON DOS VARIABLES

Estos modelos podemos formularlos de dos maneras diferentes, con la primera nos referimos a la escala de las frecuencias esperadas (modelo multiplicativo), con la segunda manera, a la escala del logaritmo de las frecuencias esperadas (modelo aditivo).

El objetivo del modelado loglineal, consiste en encontrar un modelo que represente óptimamente los resultados empíricos, es decir, que no difiera del modelo saturado (ajuste global), y que sea además el más simple de todos los modelos posibles con ajuste óptimo. La evaluación de los modelos puede hacerse con ayuda de algún criterio estadístico (que veremos más adelante). Una vez hallado el modelo óptimo se interpretan sus parámetros desde una perspectiva sustantiva.

Para el caso de tablas 2 x 2 los modelos posibles son:

- El modelo de equiprobabilidad [1]
- Los dos modelos con efecto marginal, el de fila [A], y el de columna [B].
- El modelo de independencia [A,B]
- El modelo saturado [AB]

La notación de los modelos es la sugerida por Goodman (1971) y Fienberg (1980). Para la presentación y explicación de estos modelos utilizaremos la codificación con respecto a una categoría de referencia (codificación ficticia o “Dummy coding”).

- *Modelo de Equiprobabilidad* [1]

El modelo de equiprobabilidad o nulo supone que todas las frecuencias esperadas m_{ij} son iguales (exponemos el modelo multiplicativo y aditivo):

$$m_{ij} = e^{\lambda} \quad (5.1)$$

$$\log(m_{ij}) = \lambda \quad (5.2)$$

La codificación de este modelo es un vector de unos, y es un componente necesario para todos los modelos lineales y loglineales. Nos permite representar la constante, y requiere 1 grado de libertad. Es un modelo cuya utilidad reside en que es un punto de referencia con el que comparar otros modelos, aunque puede tener también un sentido sustantivo. Es el modelo con la discrepancia máxima, con un sólo parámetro (la constante), y en el que sus valores ajustados son todos iguales.

- *Modelo con Efecto Marginal de Fila* [A]

En este modelo se supone que todas las frecuencias esperadas presentan valores marginales diferentes para los valores de A pero iguales para B.

$$m_{ij} = e^{\lambda} e^{\lambda_i^A} \quad (5.3)$$

$$\log(m_{ij}) = \lambda + \lambda_i^A \quad (5.4)$$

Dado que el efecto marginal produce dos valores diferentes necesita 2 grados de libertad. Debido a que la constante necesita un grado de libertad debemos introducir una restricción en el modelo (restricción de nulidad) referida a un nivel de la variable codificada (usualmente el primer nivel). De este modo la restricción $\lambda_1^A = 0$ define como objetivo el segundo nivel de A y como referencia el primer nivel. Por lo que sólo se utiliza un grado de libertad.

- *Modelo con Efecto Marginal de Columna* [B]

Desde este modelo se supone que las frecuencias esperadas presentan valores marginales diferentes en la variable B, pero iguales para A.

$$m_{ij} = e^{\lambda} e^{\lambda_j^B} \quad (5.5)$$

$$\log(m_{ij}) = \lambda + \lambda_j^B \quad (5.6)$$

Dado que le efecto marginal produce dos valores diferentes, necesita 2 grados de libertad. Debido a que la constante necesita un grado de libertad debemos introducir una restricción en el modelo (restricción de nulidad) referida a un nivel de la variable codificada (usualmente el primer nivel). De este modo la restricción $\lambda_1^B = 0$ define como objetivo el segundo nivel de B y como referencia el primer nivel. Por ello se utiliza un grado de libertad.

- *Modelo de Independencia* [A,B]

El modelo de independencia supone que las frecuencias esperadas no presentan asociación entre las variables A y B, es decir, solo supone efectos marginales tanto de fila como de columna.

$$m_{ij} = e^{\lambda} e^{\lambda_i^A} e^{\lambda_j^B} \quad (5.7)$$

$$\log(m_{ij}) = \lambda + \lambda_i^A + \lambda_j^B \quad (5.8)$$

Dado que los efectos marginales producen 4 valores diferentes, necesita 4 grados de libertad. Debido a que la constante necesita un grado de libertad debemos introducir una restricción en el modelo (restricción de nulidad) referida al primer nivel de las variables codificadas (usualmente el primer nivel). De este modo las restricciones $\lambda_1^A = \lambda_1^B = 0$ definen como objetivo el segundo nivel de A y B y como referencia el primer nivel. Debido a esto solo se utilizan dos grados de libertad.

Este modelo tiene la propiedad de que la razón de odds (OR) es igual a la unidad y su logaritmo es 0, de aquí el nombre del modelo (Vermunt, 1997).

- *Modelo Saturado* [AB]

El modelo saturado supone que las frecuencias esperadas son una mezcla de efectos marginales de las variables A y B, y también de su asociación mutua.

$$m_{ij} = e^{\lambda} e^{\lambda_i^A} e^{\lambda_j^B} e^{\lambda_{ij}^{AB}} \quad (5.9)$$

$$\log(m_{ij}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} \quad (5.10)$$

Este modelo necesita 4 grados de libertad, tantos como casillas de la tabla de contingencia, de aquí viene el nombre del modelo. El ajuste global del modelo es perfecto y el ajuste parcial relativo al MN es máximo. Este modelo no es interpretable, está utilizando toda la información de los datos (deja 0 grados de libertad).

5.3. MODELOS LOGLINEALES CON TRES VARIABLES

Aquí, consideraremos los modelos jerárquicos, donde los componentes interactivos incorporan todos los parámetros de orden inferior, a diferencia de los modelos no jerárquicos. Debido a ello, el aumento de modelos con respecto al de dos dimensiones es importante, ya que pasamos de 5 a 19 modelos posibles:

- Un modelo de equiprobabilidad [1]
- Tres modelos de efectos marginales [A], [B] y [C]
- Tres modelos de independencia marginal [A,B], [A,C] y [B,C]
- Un modelo de independencia mutua [A,B,C]
- Tres modelos de asociación marginal [AB], [AC] y [BC]
- Tres modelos de asociación condicional [AB,C], [AC,B] y [BC,A]

- Tres modelos de doble asociación [AB,AC], [AB,BC] y [AC,BC]
- Un modelo de triple asociación [AB,AC,BC]
- Un modelo saturado [ABC]

La interpretación de estos modelos se suele realizar describiendo la estructura de su asociación marginal y condicional utilizando las OR debido a que los parámetros son funciones de la OR. Al igual que en los modelos para dos dimensiones habrá que introducir las restricciones de nulidad para los parámetros (Fienberg y Rinaldo, 2012).

Ahora pasamos a exponer los posibles modelos a ajustar para tres variables, obviando los ya expuestos para dos variables.

- *Modelo de Independencia* [A,B,C]

El modelo de independencia supone que las tres variables son mutuamente independientes:

$$m_{ijk} = e^{\lambda} e^{\lambda_i^A} e^{\lambda_j^B} e^{\lambda_k^C} \quad (5.11)$$

$$\log(m_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C \quad (5.12)$$

Dado que los efectos marginales producen 6 valores diferentes necesita 6 grados de libertad. Debido a que la constante necesita un grado de libertad, debemos introducir una restricción en el modelo (restricción de nulidad) referida al primer nivel de las variables codificadas (usualmente el primer nivel). De este modo las restricciones $\lambda_1^A = \lambda_1^B = \lambda_1^C = 0$ definen como objetivo el segundo nivel de A, B y C, y como referencia el primer nivel. Por lo tanto solo se utilizan tres grados de libertad (más el de

la constante). De este modo, los parámetros de este modelo se interpretan con respecto a este nivel de referencia.

En este modelo, la asociación marginal es idéntica a la condicional entre dos variables, dada una tercera e igual a 1 (son asociaciones nulas que revelan la independencia entre variables). En este modelo las $OR = 1$.

- *Modelos de asociación condicional* [AB,C], [AC,B] y [BC,A]

En estos modelos se parte del supuesto de que hay una asociación entre dos variables, pero tal asociación es independiente de una tercera variable:

$$m_{ijk} = e^{\lambda} e^{\lambda_i^A} e^{\lambda_j^B} e^{\lambda_k^C} e^{\lambda_{ij}^{AB}} \quad (5.13)$$

$$\log(m_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} \quad (5.14)$$

$$m_{ijk} = e^{\lambda} e^{\lambda_i^A} e^{\lambda_j^B} e^{\lambda_k^C} e^{\lambda_{ik}^{AC}} \quad (5.15)$$

$$\log(m_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC} \quad (5.16)$$

$$m_{ijk} = e^{\lambda} e^{\lambda_i^A} e^{\lambda_j^B} e^{\lambda_k^C} e^{\lambda_{jk}^{BC}} \quad (5.17)$$

$$\log(m_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{jk}^{BC} \quad (5.18)$$

En estos modelos la asociación marginal entre dos variables es igual que la asociación parcial, dada la tercera variable.

- *Modelos de doble asociación* [AB,AC], [AB,BC] y [AC,BC]

Estos modelos suponen que dos variables son condicionalmente independientes de otro par de variables.

$$m_{ijk} = e^{\lambda} e^{\lambda_i^A} e^{\lambda_j^B} e^{\lambda_k^C} e^{\lambda_{ij}^{AB}} e^{\lambda_{ik}^{AC}} \quad (5.19)$$

$$\log(m_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} \quad (5.20)$$

$$m_{ijk} = e^\lambda e^{\lambda_i^A} e^{\lambda_j^B} e^{\lambda_k^C} e^{\lambda_{ij}^{AB}} e^{\lambda_{ik}^{AC}} \quad (5.21)$$

$$\log(m_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{jk}^{BC} \quad (5.22)$$

$$m_{ijk} = e^\lambda e^{\lambda_i^A} e^{\lambda_j^B} e^{\lambda_k^C} e^{\lambda_{ij}^{AB}} e^{\lambda_{jk}^{BC}} \quad (5.23)$$

$$\log(m_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} \quad (5.24)$$

Las tablas marginales de las variables asociadas tienen las mismas asociaciones que sus tablas parciales. Este tipo de modelos nos permiten revelar la existencia de relaciones espurias, es decir, de relaciones no lógicas. Una asociación será espuria si encontramos una tercera variable que haga que esta asociación desaparezca si controlamos la tercera variable.

En estos modelos las OR marginales son iguales a la OR parciales en el caso de las asociaciones marginales que se incluyen en el modelo, pero no en las asociaciones no incluidas en el modelo.

- *Modelo de triple asociación [AB,AC,BC]*

Este modelo parte del supuesto de que ningún par de variables es condicionalmente independiente.

$$m_{ijk} = e^\lambda e^{\lambda_i^A} e^{\lambda_j^B} e^{\lambda_k^C} e^{\lambda_{ij}^{AB}} e^{\lambda_{ik}^{AC}} e^{\lambda_{jk}^{BC}} \quad (5.25)$$

$$\log(m_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} \quad (5.26)$$

En este modelo la asociación parcial entre dos variables es la misma para cada uno de los niveles de la tercera, pero diferente de su asociación marginal, al contrario que los anteriores modelos expuestos.

- *Modelo saturado* [ABC]

Desde este modelo suponemos que cada par de variables puede ser condicionalmente dependiente, y además se parte del supuesto de que la asociación entre pares de variables puede depender del nivel de una tercera.

$$m_{ijk} = e^{\lambda} e^{\lambda_i^A} e^{\lambda_j^B} e^{\lambda_k^C} e^{\lambda_{ij}^{AB}} e^{\lambda_{ik}^{AC}} e^{\lambda_{jk}^{BC}} e^{\lambda_{ijk}^{ABC}} \quad (5.27)$$

$$\log(m_{ijk}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC} \quad (5.28)$$

- *Selección del modelo* (Agresti, 1990)

Para la selección del modelo más adecuado debemos seguir tres criterios básicos:

- El criterio lógico: que sigue el principio de parsimonia
- El criterio estadístico: que descansa en la bondad de ajuste del modelo
- El criterio sustantivo: que se basa en la interpretabilidad del modelo.

En resumen, buscamos el modelo más parsimonioso, con una bondad de ajuste aceptable y que pueda ser interpretado. Suele ocurrir que hay más de un modelo que cumple estas características. Un modelo parsimonioso es aquel que con el mayor número de grados de libertad tiene todos sus parámetros significativos. Además, un modelo bien ajustado es aquel cuya razón entre discrepancia y grados de libertad se aproxima a la unidad.

Debemos considerar ajustado un modelo cuando $P \geq 0.10$.

Para la selección del modelo más adecuado podemos ayudarnos del índice Akaike Information Criterion (AIC). Este índice se calcula mediante la siguiente fórmula (Akaike 1981):

$$AIC = D - 2gl \quad (5.29)$$

donde D es la desviación menos el doble de los grados de libertad (entendemos el mínimo que puede haber)

O bien podemos utilizar el Bayesian Information Criterion BIC desarrollado por Gideon Schwarz (1978) que se calcula:

$$BIC = D - gl \log(N) \quad (5.30)$$

donde D es la desviación menos los grados de libertad por el logaritmo de N .

Por eso cuando hay un BIC o un AIC muy pequeño, es porque la discrepancia es muy pequeña y tenemos por el contrario muchos grados de libertad y/o muchos datos.

Como sabemos, estos índices no son interpretables en si mismos, sino en comparación con los valores obtenidos para cada modelo, por ello, si tenemos dos modelos, escogeremos aquel que tenga el valor AIC o BIC más bajo.

Hay que recordar que el MS se descarta al no ser posible su interpretación. También recordar que no debemos interpretar un modelo sin haber evaluado su ajuste. Especialmente recomendable es realizar un análisis de los residuales. Como norma general, consideraremos un ajuste normal si los residuales estandarizados no presentan una proporción significativa de valores muy altos. Consideraremos un ajuste inapropiado cuando hay una proporción mayor de 0,5 valores residuales sobre 1 en términos absolutos, o bien la presencia de un valor residual mayor que 2 entendido este valor en términos absolutos.

5.4. MODELOS LOGLINEALES CON MÁS DE TRES VARIABLES

No es necesario indicar aquí (ya lo veremos en la parte aplicada de forma más clara) la gran utilidad de este arsenal estadístico cuando diseñamos experimentos con múltiples categorías y variables que se relaciona de forma progresivamente compleja y cuyas relaciones (relaciones entre variables y categorías) son relevantes para entender los modelos sustantivos. Al aumentar el número de variables aumenta de manera importante la complejidad de los modelos, ya que supone un incremento del número de interacciones y casillas de la tabla de contingencia. El tamaño muestral necesario para rellenar una tabla de contingencia con muchas casillas es considerable. Además, la selección del modelo más adecuado se complica notablemente. Se ha propuesto diferentes estrategias, como la de Goodman (1971) y la de Brown y Benedetti (1977) basados en la G^2 pero la que se suele utilizar es la propuesta por Bishop, Fienberg y Holland (1975) Aitkin (1980), y Christensen (1990) donde la decisión está basada en un procedimiento estadístico de selección de modelos:

- Selección proactiva o *forward selection*: el proceso se inicia con un modelo muy simple, y después se van introduciendo parámetros y mejorando el ajuste hasta encontrar un modelo con un ajuste aceptable, con el menor número de parámetros posible.
- Selección retroactiva o *backward selection*: el proceso se inicia con el MS y después se van eliminando parámetros hasta encontrar un modelo que no difiera del MS pero que incluye menos parámetros.
- Selección por pasos o *stepwise selection*: se combinan las técnicas de selección proactiva y retroactiva.

En los modelos loglineales se suele utilizar la estrategia retroactiva, ya que es muy similar a la estrategia de ajuste parcial. El número máximo de etapas en la estrategia de búsqueda de un modelo depende de la dimensión del MS. En cada etapa debemos comprobar el ajuste y los parámetros de las interacciones. Se descartaran los parámetros con un nivel de probabilidad $P > 0.05$ o $P > 0.01$, en función del rigor que demanda la naturaleza de las variables. A continuación se prueba el modelo resultado de eliminar estos parámetros. Si el ajuste no difiere continuaremos el proceso, pero si hay diferencias debemos detener el proceso y comprobar el modelo anterior. Debemos recordar que al trabajar con modelos jerárquicos, si se descarta una interacción, deben mantenerse en el modelo las interacciones de orden inferior.

5.5. MODELOS LOGLINEALES CON DATOS RELACIONADOS

Este tipo de modelos se aplica cuando la muestra de individuos se evalúa en dos momentos en el tiempo o se les ha aplicado una técnica de emparejamiento. Para estos casos las frecuencias de las casillas de la tabla de contingencia pueden considerarse estadísticamente dependientes, no siendo correcto aplicar los modelos loglineales anteriormente expuestos (Upton 1980, Frees 2004).

Los resultados de estas situaciones se resumen en una tabla de contingencia de forma cuadrada, donde la variable de fila tiene el mismo número de categorías que la variable de columna. Es por esto que suelen ser llamados modelos para tablas cuadradas.

Podemos distinguir en estas tablas tres cuerpos diferenciados (ver figura 5.2):

- El cuerpo de la diagonal principal
- El cuerpo triangular superior
- El cuerpo triangular inferior

Triangular superior

$$\begin{bmatrix} 1 & 3 & -4 \\ 0 & 6 & 2 \\ 0 & 0 & -5 \end{bmatrix}$$

Triangular inferior

$$\begin{bmatrix} 4 & 0 & 0 \\ -2 & -1 & 0 \\ 7 & 5 & 3 \end{bmatrix}$$

Diagonal

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

Unitaria

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Figura 5.2 Representaciones matriciales de modelos para tablas cuadradas.

5.5.1 MODELOS DE SIMETRÍA

El primer modelo que se debe plantear como hemos visto es el modelo [I], que plantea la independencia de las variables.

1) *Modelo de Cuasi-Independencia* [QI]

Uno de los efectos más comunes de las tablas cuadradas es la alta proporción de frecuencias concentradas en la diagonal principal, lo que suele provocar que el modelo [I] no presente un ajuste adecuado (Hsiao 2003). Sin embargo, la hipótesis de cuasi-independencia se centra en las frecuencias de la tabla que ocupan los triángulos superior e inferior, excluyendo las que ocupan la diagonal principal. Así, una tabla cuadrada

presentará cuasi-independencia si A y B son independientes en las casillas que no pertenecen a la diagonal.

$$\log(m_{ij}) = \lambda + \lambda_i^A + \lambda_j^B + \delta_i I(i = j) \quad (5.31)$$

donde $I(i = j)$ es una función indicador, que es igual a 1 cuando $(i = j)$ e igual a 0 cuando $(i \neq j)$. Este modelo aporta I parámetros al modelo [I], para poder representar los elementos de la diagonal principal. Si los parámetros son mayores a 0, quiere decir que existe una proporción mayor de frecuencias que con el modelo [I].

El modelo [QI] supone que las OR para todas las subtablas rectangulares sacadas de los triángulos superior e inferior son iguales a 1 y su LOR es igual a 0 (Markus 1979).

2) El modelo de simetría [SI]

Una tabla cuadrada será simétrica cuando los elementos a ambos lados de la diagonal principal son iguales.

$$\log(m_{ij}) = \lambda + \lambda_i + \lambda_j + \lambda_{ij} \quad (5.32)$$

Hay que señalar que éste es muy restrictivo y es muy raro que se obtenga un ajuste aceptable del modelo. Tiene la propiedad de que cuando se somete a prueba la hipótesis de simetría esto implica la hipótesis de homogeneidad marginal, pero esto no pasa al contrario, de aquí la dificultad para obtener un ajuste aceptable (Bassi, *et al.*, 2000).

3) Modelo de cuasi-simetría [QSI]

Para evitar el supuesto de homogeneidad marginal, debemos estimar de forma independiente los efectos principales de fila y columna. La hipótesis que trata de contrastar es que los valores esperados simétricos son iguales.

$$\log(m_{ij}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} \quad (5.33)$$

Una de las propiedades de este modelo es que las OR locales tienen la propiedad de la simetría.

4) *Modelo de homogeneidad marginal [HM]*

Hay que señalar que no es un modelo loglineal convencional, en consecuencia, para poder realizar el análisis podemos; en primer lugar, utilizar el enfoque condicional, realizándose el cálculo a partir de los modelos [SI] y [QSI] (Causinus 1966). Aquí partimos del supuesto de que el modelo [SI] implica simetría más homogeneidad marginal, y el modelo [QSOI] sólo simetría. Si el modelo [QSI] obtiene un ajuste aceptable, la diferencia de discrepancias entre ambos modelos nos permitirá probar las hipótesis de homogeneidad marginal.

En segundo lugar, utilizar el enfoque no condicional, realizándose el cálculo a partir de la metodología de los modelos marginales (Lang y Agresti, 1994; Bergsma, 1997).

5.5.2 MODELOS DIAGONALES

Estos modelos atienden a diferencia de los modelos de simetría en la diagonal principal de la tabla cuadrada. Se centran en la diferencia entre las casillas de la diagonal principal y las casillas externas a la diagonal principal. En estos modelos el MR es el modelo [I] (Christensen 1997). Veamos algunos a continuación:

1) Modelo diagonal principal [Dp]

Este modelo supone que los elementos de la diagonal principal son homogéneos y diferentes a los elementos externos a la diagonal principal.

$$\log(m_{ij}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^D \quad (5.34)$$

donde D es una nueva variable de clasificación, que se comporta así: si $i \neq j$ entonces $D = 1$, y si $i = j$ entonces $D = 2$

2) Modelo diagonal menor simétrico [Dms]

Su formulación es igual que la del modelo [Dp], pero la nueva variable D se comporta de forma diferente, si $i = j$ entonces $D = 1$, el resto de valores depende de las posiciones de i o j . Una de sus características principales es que considera siempre la misma probabilidad de manera independiente a la ocasión de medida que se considere.

3) Modelo diagonal menor asimétrico [Dma]

Se formula igual que el modelo [Dp], pero la nueva variable D se comporta de forma diferente, si $i = j$ entonces $D = 1$, el resto de valores depende de las posiciones de i y j . Una de sus características principales es que considera probabilidades diferentes en cada ocasión de medida que se considere.

5.5.3 MODELOS DE BLOQUES

Estos modelos comparan pautas específicas guiadas por el diseño de investigación. Siempre parten del modelo [QI], pero introduciendo restricciones en los parámetros de interacción, lo que conlleva una cierta dependencia de filas y columnas. Ello genera nuevas hipótesis específicas (Knoke y Burke, 1980).

1) Modelo triangular de bloques [T]

La hipótesis a contrastar es si existe un incremento de las frecuencias en el triangulo superior (postest) respecto a las frecuencias del triangulo inferior (pretest). Si los resultados son diferentes en cada bloque entonces hay dos bloques diferenciados, en caso contrario aceptaríamos la hipótesis nula.

Las variables categóricas deben ser ordinales.

$$\log(m_{ij}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^T \quad (5.35)$$

donde λ_{ij}^T es un parámetro que toma los siguientes valores, 2 si $i < j$ o 1 si $i > j$. Se somete a prueba la hipótesis de nulidad de bloques triangulares, por lo tanto si $\lambda_{ij}^T = 0$ se acepta H_0 , no existen bloques triangulares pero si se rechaza no podemos eliminar del modelo este parámetro.

2) Modelo de bloques mixtos [M]

En este modelo, las variables categóricas deben ser nominales y además, es más complejo, ya que se supone que en los datos podemos distinguir conglomerados de respuesta con pautas definidas.

$$\log(m_{ij}) = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^M \quad (5.36)$$

CAPÍTULO 6

REGRESIÓN DE POISSON

6.1 – INTRODUCCIÓN

Este es el modelo más adecuado para analizar variables de respuesta de recuento (Long, 1997). Inicialmente se aplicó en el estudio de conductas criminales, para posteriormente expandirse a otras áreas. Actualmente es una distribución muy utilizada en la investigación médica y, en particular, en epidemiología.

La distribución de Poisson es la que suelen seguir los datos de recuento y nos permite obtener la probabilidad de que se produzca un determinado k de ocurrencias de un evento, siendo su función de probabilidad:

$$\pi_i = P(y_i = k) = \frac{\exp(-\mu)\mu^{y_i}}{y_i!} \quad (6.1)$$

Si $\mu > 0$ el parámetro media de la distribución coincide con la varianza, denominándose esta característica como “equidispersión”.

Esta distribución debe su nombre, como hemos indicado anteriormente en este trabajo, al estadístico Poisson que la describió en 1837, aunque ya había sido introducida en 1718 por De Moivre como una forma límite de la distribución binomial que surge cuando se observa un evento raro después de un número grande de repeticiones. El concepto de evento “raro” o poco frecuente debe ser entendido en el sentido de que la probabilidad de observar k eventos decrece rápidamente a medida que k aumenta.

Sus principales características son:

- Tiene un único parámetro: la media.
- La varianza es igual a la media

- Si el cociente entre la varianza y la media es mayor de 1 tendremos sobredispersión.

Una característica de la distribución de Poisson es que tiene la propiedad de que la suma de las variables aleatorias de Poisson independientes es también la variable de Poisson, cuya medida es igual a la suma de las medias de las variables aleatorias independientes (Haight 1967).

Ventajas sobre el modelo lineal:

- Los modelos lineales podrían predecir valores negativos.
- La varianza de la variable de respuesta probablemente aumente con la media.
- Los errores no se distribuyen según la distribución normal.
- Los ceros no se pueden transformar satisfactoriamente.

¿Que diferencias hay entre modelar respuestas de frecuencias y respuestas de recuento?

En la regresión de Poisson nos interesa saber de qué manera una o más variables pueden explicar o predecir el número de ocurrencias (recuento) que se darán en un determinado evento. Tradicionalmente se han analizado este tipo de variables con un modelo de regresión lineal (Gürtler y Henze 2000), no siendo esto del todo correcto, aunque sí es cierto que con valores de recuento grandes sus resultados no difieren mucho de los de la regresión de Poisson.

De este modo, aplicaremos este modelo cuando se dispone de datos individuales, para el caso de datos agrupados deberemos acudir a los modelos loglineales. Pero si

todas las variables son categóricas se pueden utilizar indistintamente aunque es recomendable utilizar el modelo loglineal (Cameron y Trivedi, 1998, 2013).

En el contexto de este trabajo pueden ser muy útiles para modelar el recuento de respuestas o conductas de los sujetos a los estímulos utilizando una gran variedad de variables explicativas incluyendo las medidas psicofísicas.

6.2 – EL MODELO DE REGRESIÓN DE POISSON SIMPLE

Se utiliza cuando la variable de respuesta es de recuento y la variable explicativa es continua o categórica binaria, y se estudia si el valor esperado m_i del recuento está influenciado por la variable explicativa. Como hemos comentado antes, si la variable explicativa es binaria es mejor utilizar el modelo loglineal.

La ecuación del modelo de regresión de Poisson simple es:

$$\eta_i = \log(\mu_i) = \beta_0 + \beta_1 X_i \quad (6.2)$$

Este es un MLG cuyo componente aleatorio sigue la distribución de Poisson, utiliza una función de enlace logarítmica, con una variable de respuesta cuantitativa discreta no negativa. La relación entre el parámetro μ y la variable explicativa viene dada por:

$$\log(m_i) = b_0 + b_1 X_i \quad (6.3)$$

y en su forma multiplicativa es:

$$m_i = \exp(b_0 + b_1 X_i) = \exp(b_0) \exp(b_1 X_i) \quad (6.4)$$

Como vemos no hay relación lineal entre el valor de la variable de respuesta y el valor de la variable explicativa. Además, la expresión multiplicativa nos permitirá interpretar los efectos de b_1 directamente en la escala de la variable de respuesta,

mientras que la interpretación de dichos efectos en la expresión aditiva se realiza sobre la escala del predictor lineal (logarítmica).

6.2.1 VARIABLE DE EXPOSICIÓN

En el modelo de regresión de Poisson, debemos incluir la variable de exposición o también llamado “multiplicador de tasa” simbolizado por t . De este modo se tiene en cuenta que el recuento de cada una de las observaciones se puede basar en periodos de tiempo, espacio o tamaño de población diferentes (Lunneborg 1994, Lindsey 1995). Si el valor t es igual para todas las observaciones no hace falta incluirla en el modelo (sería como una constante). Por otra parte, la variable t debe ser como mínimo de 10 a 100 veces mayor que el dato de recuento para que los supuestos de la distribución de Poisson se cumplan, en caso contrario habría que utilizar un modelo loglineal. (Liao 1994).

6.2.2 ESTIMACIÓN DE LOS PARÁMETROS

En este modelo, y siguiendo la expresión aditiva, el valor del coeficiente b_1 representa el cambio en el logaritmo de recuento esperado por cada unidad de cambio en la variable explicativa X , siendo este cambio constante entre los valores X que estén separados equidistantemente (es un cambio lineal). Pero si tomamos la expresión multiplicativa, la interpretación del coeficiente b_1 será más sencilla, debido a que el valor exponentado de b_1 representa el factor de cambio del recuento esperado para el incremento de una unidad en la variable explicativa X , siendo este cambio no constante entre los valores de X , ya que de hecho es un cambio no lineal.

Si deseamos saber el factor de cambio en el recuento esperado cuando la puntuación en X se incrementa en δ unidades, solo tendremos que calcular $\exp(\delta b_1)$, que a su vez se puede transformar a porcentaje (Lindsey, 1995). Para obtener la significación estadística de los coeficientes del modelo, podemos contrastar la hipótesis de que $\beta = 0$ con la prueba de Wald (1945):

$$z = \frac{b_1}{EE(b_1)} \quad (6.5)$$

Siguiendo su valor una distribución normal $N(0,1)$, además podemos calcular su intervalo de confianza: $IC(1-\alpha)\beta_1 : b_1 \pm z_{\alpha/2}EE(b_1)$. De este modo, el coeficiente será significativo si el intervalo de confianza no contiene el valor 0, aunque es el investigador el que puede decidir el valor criterio. Para la expresión multiplicativa del modelo exponenciaremos la fórmula anterior: $IC(1-\alpha)\exp(\beta_1) : \exp[b_1 \pm z_{\alpha/2}EE(b_1)]$

Finalmente es necesario obtener el IC para los valores esperados, para ello, debemos obtener el valor del error estándar, su cálculo en este contexto conlleva cierta dificultad (Liao 1994).

6.2.3 COMPARACIÓN DE MODELOS

Para la comparación de los modelos debemos utilizar la prueba de diferencias que en este modelo se obtiene de:

$$D = -2 \sum_{i=1}^n \left(y_i \log \left(\frac{m_i}{y_i} \right) + (y_i - m_i) \right) \quad (6.6)$$

De este modo comparamos el MA con el MR en modelos simples, donde en este último el valor esperado m_i es el valor medio de los recuentos registrados para las diferentes observaciones. Así: $\Delta D = D_{MN} - D_{MA}$, distribuyéndose ΔD como una

distribución χ^2 con grados de libertad $\Delta gl = gl_{MN} - gl_{MA}$ (la diferencia en el número de parámetros del modelo).

En los MLG no tenemos un R^2 (índice de bondad de ajuste), pero si “pseudo- R^2 ” (Cameron y Trivedi, 1998):

$$pseudo-R^2 = \frac{D_{MN} - D_{MA}}{D_{MN}} \quad (6.7)$$

El valor obtenido varía entre 0 y 1, y aumentará según se vayan introduciendo variables al MA. Se interpreta como la proporción de reducción en la discrepancia del MN debida a la inclusión de la variable explicativa en el MA.

6.2.4 MODELO CON VARIABLE EXPLICATIVA BINARIA

Si la variable explicativa es categórica binaria se puede incluir directamente al modelo con la adecuada codificación. Si codificamos con respecto a una categoría de referencia el valor $\exp(b_1)$ es el factor de cambio entre el número promedio de recuentos para el grupo codificado con el valor 1 respecto al grupo de referencia codificado como 0.

Si la variable explicativa es politómica hay que convertirla en $k-1$ variables, quedándonos un modelo de Regresión de Poisson Múltiple.

6.3 – EL MODELO DE REGRESIÓN DE POISSON MÚLTIPLE

Se utiliza cuando la variable de respuesta es de recuento y las variables explicativas son continuas o categóricas, y se estudia si el valor esperado m_i del recuento está

influenciado por las variables explicativas. A diferencia de el modelo simple, tenemos que distinguir entre modelos que incluyen interacciones entre variable explicativas y los que no las incluyen (Cameron y Trivedi 1998).

La ecuación aditiva del modelo sin interacción es:

$$\log(\mu_i) = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} = \beta_0 + \sum_{j=1}^k \beta_j X_{ji} \quad (6.8)$$

donde X_j es la representación genérica de las variables explicativas y de control. El ajuste se realiza de igual manera que con el modelo simple comparando MA y MR, y la interpretación de los coeficientes de una variable explicativa proporciona el efecto de esta variable manteniendo constantes las restantes variables del modelo, de manera independiente al valor que tomen estas variables. En este caso se asume que el efecto de cada una de las variables explicativas es constante para el dominio de valores de las restantes variables del modelo.

La ecuación aditiva del modelo con dos variables en interacción es:

$$\log(\mu_i) = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i \quad (6.9)$$

donde X y Z serían las variables explicativas y de control. Los efectos principales del modelo serían β_1 y β_2 , y el efecto de interacción entre las variables sería β_3 . Es recomendable comenzar el proceso de selección con las interacciones de mayor orden, eliminando las que no resulten estadísticamente significativas (Claeskens y Hjort 2008). Podemos realizar esto con las interacciones, una a una, o mediante un *chunk test* (con interacciones de un orden similar). El proceso se sigue realizando sobre las interacciones de orden inmediatamente inferior, realizando sucesivamente este proceso según el principio jerárquico.

Para la interpretación de los coeficientes de regresión, como hemos visto, la formulación aditiva no permite una interpretación directa de los coeficientes, de este modo, será mucho más útil su expresión multiplicativa

$$m_i = \exp(b_0) \exp(b_1 X_i) \exp(b_2 Z_i) \exp(b_3 X_i Z_i) \quad (6.10)$$

La interpretación se hará de la siguiente manera:

- La constante $\exp(b_0)$ es el valor esperado cuando X e $Z = 0$
- El coeficiente $\exp(b_1)$ es el factor de cambio en el valor esperado por cada cambio de 1 unidad en X cuando $Z = 0$
- El coeficiente $\exp(b_2)$ es el factor de cambio en el valor esperado por cada cambio de 1 unidad en Z cuando $X = 0$
- El coeficiente $\exp(b_3)$ es el factor de cambio en el coeficiente $\exp(b_1)$ cuando X cambia en 1 unidad y además depende del valor que tome Z . También, es el factor de cambio en el coeficiente $\exp(b_2)$ cuando Z cambia en 1 unidad y además depende del valor que tome X .

Para una mejor interpretación de este último coeficiente, se recomienda aislar la relación entre una de las variables explicativas y la variable de respuesta, para después analizar como cambia esta relación en función de los valores de una tercera variable (Lindsey 1995).

Si queremos calcular los IC cuando tenemos términos de interacción en un modelo de regresión de Poisson, el efecto de una de las variables explicativas sobre la variable de respuesta dependerá del valor que tome la otra variable explicativa. Si $Z = 0$ el

efecto de X es b_1 . Para los restantes valores de Z el efecto de X sobre Y viene dado por $b_1 + b_3 Z$. De este modo el cálculo del error estándar incluirá las varianzas y covarianzas de los coeficientes b_1 y b_3 necesarios para la estimación por intervalo del efecto de X cuando $Z \neq 0$. (es aconsejable que se realice con un programa informático debido a su complejidad) (Dobson 2002).

Si la variable modificadora Z es categórica binaria, entonces el coeficiente b_1 nos proporciona el efecto de X sobre Y para la categoría de Z codificada como 0 (categoría de referencia). Además, la suma de los coeficientes b_1 y b_3 nos proporciona el efecto de X para la categoría 1 de Z . Pero para simplificar esta operación es mejor recodificar la categoría 1 de Z como categoría de referencia y es conveniente realizar una tabla resumen con toda la información.

No obstante, si la variable Z tiene más de dos categorías (politómica), la complejidad es aún mayor, ya que a pesar de que es sencillo el cálculo del valor del coeficiente, tanto el error estándar, como su estimación del IC se vuelven muy complejos. En efecto, la solución pasa por recodificar la variable creando un conjunto de $k-1$ variables ficticias.

Finalmente, si la variable modificadora Z es continua, lo que deseamos conocer es el efecto de la variable X cuando Z toma el valor genérico c . Lo que se hace es centrar Z en el valor de referencia c generando una variable $Z_c = Z - c$, estimando después el modelo de regresión Y sobre X y Z_c . De este modo el valor del coeficiente b_1 de X nos

indica el efecto de esta variable cuando Z toma el valor c y el IC de este coeficiente se corresponderá con el efecto de X para el valor c .

CAPÍTULO 7

REGRESIÓN LOGÍSTICA

7.1 – INTRODUCCIÓN

Cuando tenemos una variable de respuesta categórica binaria y un conjunto de variables explicativas continuas y/o categóricas se utiliza el modelo de regresión logística. Es un modelo que encaja muy bien y se integra perfectamente en los modelos de análisis propuestos desde las teorías propias de la detección de señales. Todos ellos utilizados conjuntamente, cuando sea necesario, aumentan la precisión y validez de los datos en los diseños correspondientes.

Fue Bartlett (1937) el que utilizó la transformación $\log[y/(1-y)]$ para analizar proporciones. Más tarde, Fisher y Yates (1938) propusieron el uso de esa transformación para analizar datos binarios. El término logit fue introducido por Berkson (1944) para designar esta transformación y sus trabajos popularizaron su utilización, posteriormente Cornfield (1966) utilizó la regresión logística para el cálculo de *odds ratio* como valores aproximados del riesgo relativo en estudios de casos y controles. Pero quizás el principal difusor de la regresión logística fue Cox (1970).

Es una de las técnicas de regresión más utilizada en medicina, debido a que permite el uso de múltiples variables con relativamente pocos casos. Se ha sugerido que el número de sujetos para poder usar esta técnica estadística sin problemas debe ser superior a $10(k+1)$ donde k es el número de variables explicativas (Agresti 1999).

En el modelo de regresión logística la variable de respuesta es la probabilidad de que se dé el atributo de estudio (probabilidad de éxito), y estará comprendida entre 0 y 1. La función logística nos permite que el valor predicho se encuentre entre estos

límites. Es una función monótona creciente o decreciente, no lineal y con forma de S (ver figura 7.1).

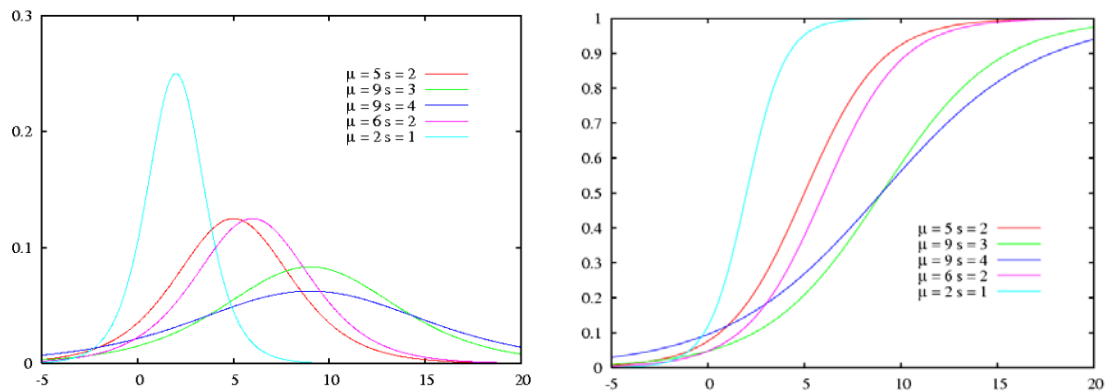


Figura 7.1 Función de densidad y función de distribución logística.

Se representa por $p(z)$, lo que refleja que se interpreta en términos de probabilidad:

$$p(z) = \frac{\exp(z)}{1 + \exp(z)} \quad (7.1)$$

Una variable dicotómica se describe como la proporción de casos que pertenecen a una categoría ($p_1 = n_1/n$) y su valor estará comprendido entre 0 y 1. Esto se interpreta como la probabilidad de que un sujeto presente dicha característica, donde π es la proporción poblacional y p la proporción muestral. Como ya hemos visto, la proporción muestral de casos que pertenecen a la primera categoría de una variable binaria es $p_1 = n_1/n$, para calcular la proporción de la segunda categoría se realiza mediante $1 - p_1$. Normalmente 0 significa ausencia de la característica de estudio y 1 presencia.

También se puede utilizar un índice complementario denominado *odds* (Hosmer y Lemeshow, 1989, 2000), que se calcula como el cociente del número de sujetos de una categoría dividido entre el número de sujetos de la otra categoría o dicho de otro

modo, el cociente entre la probabilidad de que ocurra frente a la probabilidad de que no

ocurra : $O_1 = \frac{p}{1-p}$. Suele aparecer sin el subíndice, siendo el cociente entre la categoría

1 y la categoría 0. Sus valores van de 0 a ∞ y se interpreta:

- Si $O < 1$ hay más sujetos en la categoría 0 que en la categoría 1.
- Si $O = 1$ habría el mismo número de sujetos en las dos categorías.
- Si $O > 1$ hay menos sujetos en la categoría 0 que en la categoría 1.

También podemos utilizar la *odds ratio* (OR), que no es más que comparar las *odds* obtenidas en dos grupos de sujetos

$$O_{A,B} = \frac{O_A}{O_B} = \frac{p_A/1-p_A}{p_B/1-p_B} \quad (7.2)$$

Sus valores también van de 0 a ∞ y se interpreta:

- Si $OR < 1$ nuestro objeto de estudio es menor en el grupo A que en el Grupo B (hay un incremento de *odds*).
- Si $OR = 1$ nuestro objeto de estudio es igual en los dos grupos (no hay variación de *odds*).
- Si $OR > 1$ nuestro objeto de estudio es mayor el grupo A que en el grupo B (hay una disminución de *odds*).

Es importante indicar que habrá que hacer distinciones en el caso de trabajar con datos agrupados o datos individuales: si las variables explicativas son continuas no se pueden agrupar datos y se realizaría un modelo de regresión logística, pero si las variables explicativas son categóricas se pueden agrupar los datos y se realizaría un modelo de regresión logit (Kleinbaum 1994).

Los inconvenientes que podemos encontrarnos con esta prueba son (Long 1997):

- Implica conocer el nivel de separación entre los valores de la variable de respuesta.
- Problemas numéricos derivados de la frecuencia observada en las combinaciones entre los niveles de la/s variable/s predictor/a/s y la variable de respuesta.
- El tamaño de muestra: en el caso de que el tamaño muestral sea reducido es conveniente utilizar el test de la razón de verosimilitud.
- Es sensible a los casos perdidos.
- La categorización de las variables debe estar bien fundamentada.
- Sesgo en presencia de colinealidad.

7.2. MODELO DE REGRESIÓN LOGÍSTICA SIMPLE

Este modelo nos “permite predecir la probabilidad de presentar una determinada característica de interés en función del valor en una variable explicativa de tipo continuo”. Su ecuación es:

$$\text{logit}(\mu_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 X_i \quad (7.3)$$

Como en todos los MLG tenemos el componente sistemático (el predictor lineal que nos proporciona el valor predicho), la función de enlace (*logit*) que es el logaritmo de todas las *odds* esperadas por el modelo, y por último el componente aleatorio que supone la variabilidad de la variable de respuesta *Y* no explicada por el predictor lineal (η) siguiendo una distribución de probabilidad de Bernoulli (distribución binomial con $n = 1$). Debido a que los datos no se agrupan, cada predicción del modelo condicionada

a un valor determinado de la variable explicativa incluye un único sujeto, por eso el error es binomial ya que la respuesta solo puede tomar valores 0 y 1.

La función de enlace *logit* se utiliza cuando el componente aleatorio sigue una distribución Binomial $B(n_i, \pi_i)$, ya sea con n_i sujetos o con 1 sujeto. Esta función de enlace nos permite linealizar la relación entre la probabilidad modelada y la combinación lineal de variables explicativas, y también convertir los valores esperados a una misma escala de medida. El valor transformado se presenta en la escala de medida del predictor lineal.

7.2.1. ESTIMACIÓN DE LOS PARÁMETROS DEL MODELO

La ecuación aditiva del modelo de regresión logística simple para una muestra concreta es:

$$\text{logit}(m_i) = \log(o_i) = \log\left(\frac{m_i}{1-m_i}\right) = b_0 + b_1 X_i \quad (7.4)$$

donde m_i es la probabilidad esperada de éxito para la observación i . El *logit* de m_i es igual al logaritmo del cociente entre la proporción esperada de éxito y su complementario. Esta misma ecuación expresada en términos multiplicativos permite que la respuesta sea ahora la *odds* esperada.

$$o_i = \exp(b_0 + b_1 X_i) = \exp(b_0) \exp(b_1 X_i) \quad (7.5)$$

Para la estimación de los parámetros podemos utilizar dos métodos diferentes, el condicional y el incondicional la elección de alguno de estos métodos depende del número sujetos de la muestra (Menard, 1995). Si el número de parámetros es grande con respecto a la muestra utilizaremos el método condicional, y si el número de parámetros es pequeño con respecto a la muestra se utiliza el método incondicional.

Ésto es debido a que la estimación por máxima verosimilitud es óptima con muestras grandes en relación al número de parámetros. Es importante señalar que la interpretación de los parámetros es exactamente la misma con los dos métodos. Los programas de análisis de datos estadísticos habitualmente utilizan el método incondicional si no se les indica lo contrario.

La interpretación de los coeficientes del modelo de regresión logística se suelen hacer en términos de la ecuación multiplicativa, ya que se expresa la razón entre la probabilidad de ocurrencia esperada m_i y la no ocurrencia $1-m_i$ de la característica modelada (Pampel, 2000). De este modo:

- Si $b_1 > 0$ entonces $\exp(b_1) > 1$, es decir, un incremento de X produce un incremento de o_i (que aumente la probabilidad de que la característica estudiada aparezca).
- Si $b_1 = 0$ entonces $\exp(b_1) = 1$, es decir un incremento de X no cambia o_i
- Si $b_1 < 0$ entonces $\exp(b_1) < 1$, es decir, un incremento de X produce una disminución de o_i (que disminuya la probabilidad de que la característica estudiada aparezca).

Como hemos visto, el valor de $\exp(b_1)$ es el factor por el que se multiplica o_i cuando la variable explicativa X se incrementa en una unidad, pero también puede ser interpretado directamente como una OR. Si la variable explicativa es continua es preferible considerar un incremento en X de u unidades, así el factor por el que se multiplica o_i cuando X se incrementa en u unidades viene dado por $\exp(ub_1)$ (Hosmer y Lemeshow, 1989).

En el caso de $\exp(b_0)$ este es el valor de o_i para un sujeto con valor 0 en la variable explicativa X , y como constante se considera un valor de referencia respecto al cual interpretar las *odds* de sujetos con valores diferentes de 0 en la variable explicativa (siempre que 0 sea un valor posible de X).

7.2.2. INFERENCIAS SOBRE LOS COEFICIENTES DE REGRESIÓN

Comenzaremos con la comprobación de la significación estadística del modelo mediante una prueba de ajuste global. Existen dos procedimientos alternativos (Menard 2002):

En primer lugar, aplicar una prueba de significación para comprobar si el parámetro que se estima es estadísticamente igual o diferente de 0. Después se realiza la prueba de significación de Wald (que se fundamenta en que la distribución muestral de una estimación máximo verosímil sigue la normal cuando n es grande):

$$W = \left(\frac{b_1}{EE(b_1)} \right)^2 \quad (7.6)$$

En segundo lugar se puede aplicar el procedimiento de comparar la discrepancia de dos modelos (un MN y un MA).

También se puede calcular el IC del parámetro β_1 , pero en este caso resulta mejor calcular el IC del OR, ya que su distribución muestral tiende a ser simétrica en muestras de tamaño pequeño:

$$IC(1-\alpha)\beta_1 = b_1 \pm z_{\alpha/2}EE(b_1) \quad (7.7)$$

Por su parte, los límites del IC para el OR vienen dados por:

$$IC(1-\alpha)OR = \exp(b_1 \pm z_{\alpha/2}EE(b_1)) \quad (7.8)$$

7.2.3. COMPARACIÓN DE MODELOS

La discrepancia de un modelo de regresión logística ($-2\log L$) se calcula a partir de la diferencia entre los valores observados y_i y los valores esperados m_i del modelo:

$$D = -2 \log L = -2 \sum_{i=1}^n (y_i \log(m_i(1 - m_i)) + \log(1 - m_i)) \quad (7.9)$$

Pudiéndose expresar así:

$$D = -2 \sum_{i=1}^n (m_i \text{logit}(m_i) + \log(1 - m_i)) \quad (7.10)$$

Para conocer la bondad de ajuste debemos recurrir a la prueba de ajuste global, donde realizaremos el test de la diferencia de discrepancias (ΔD):

$$\Delta D = D_{MR} - D_{MA} = -2 \log L_{MR} - (-2 \log L_{MA}) = -2 \log \left(\frac{L_{MR}}{L_{MA}} \right) \quad (7.11)$$

Como sabemos, este índice sigue una distribución χ^2 con grados de libertad iguales a la diferencia de parámetros entre los dos modelos comparados. En el caso de el MR = MN y el MA es el que incluye la variable explicativa. Por lo tanto el ΔD que evalúa el ajuste global de un MA es:

$$\Delta D = D_{MN} - D_{MA} = -2 \log L_{MN} - (-2 \log L_{MA}) = -2 \log \left(\frac{L_{MN}}{L_{MA}} \right) \quad (7.12)$$

Que sigue una distribución χ^2 con 1 grado de libertad.

7.2.4. BONDAD DE AJUSTE DEL MODELO

Para determinar la bondad de ajuste del modelo podemos utilizar los siguientes índices (Hosmer y Hjort 2002):

- χ_{HM}^2 compara el número de sujetos observados en cada categoría de respuesta con el número de sujetos predicho por el modelo. El procedimiento debe comenzar dividiendo a los sujetos en g grupos (lo normal es $g = 10$), a partir de la probabilidad esperada de respuesta m_i . De este modo en el primer grupo estarán los sujetos con $m_i \leq 0,1$, en el segundo los sujetos con $0,1 < m_i \leq 0,2$ y así hasta 1. Después para cada grupo se recuenta el número de sujetos observados en la muestra en cada categoría de respuesta, y también se calcula el número de sujetos predicho por el modelo en cada categoría de respuesta. Es muy recomendable representar estos datos en una tabla (ver tabla 7.1).

Tabla 7.1 Valores observados y esperados para el cálculo del estadístico χ_{HM}^2

Grupo	Respuesta r = 0		Respuesta r = 1	
	Observado (o_i)	Esperado (e_i)	Observado (o_i)	Esperado (e_i)
g = 1	$o_{0,1}$	$e_{0,1}$	$o_{1,1}$	$e_{1,1}$
g = 2	$o_{0,2}$	$e_{0,2}$	$o_{1,2}$	$e_{1,2}$
⋮	⋮	⋮	⋮	⋮
g	$o_{0,g}$	$e_{0,g}$	$o_{1,g}$	$e_{1,g}$

El estadístico se calcula a partir de la tabla:

$$\chi_{HM}^2 = \sum_{g=1}^g \frac{(o_{0,g} - e_{0,g})^2}{e_{0,g}} + \frac{(o_{1,g} - e_{1,g})^2}{e_{1,g}} \quad (7.13)$$

Sigue una distribución χ^2 con $g - 2$ grados de libertad, pudiéndose obtener su significación estadística. El modelo estará ajustado si no hay significación estadística.

Hay que indicar que no puede calcularse para un MN, ya que predeciría un único m_i para todos los sujetos, siendo imposible establecer g grupos.

- R_L^2 de la familia pseudo- R^2 , y se calcula a partir de la discrepancia del MN y

MA:

$$R_L^2 = \frac{D_{MN} - D_{MA}}{D_{MN}} \quad (7.14)$$

Un inconveniente importante consiste en su difícil interpretación, ya que no refleja el porcentaje de varianza de Y explicada por X , porque las D no son sumas de cuadrados.

7.2.5. MODELO CON VARIABLE EXPLICATIVA BINARIA

Si nuestro modelo incluye un predictor binario X con 2 categorías, la ecuación del modelo será:

$$\text{logit}(m_i) = b_0 + b_1 X \quad (7.15)$$

El proceso es similar a cuando tenemos un predictor continuo, lo que debemos tener en cuenta es la interpretación de los coeficientes de regresión.

Si la codificación es respecto a la categoría de referencia 0, el coeficiente b_0 es el valor del $\text{logit}(m_i)$ de los sujetos del grupo codificado con el valor 0, y el coeficiente b_1 es el cambio que se produce en el $\text{logit}(m_i)$ debido a un incremento en una unidad en X . Pero si lo que deseamos es interpretar la ecuación multiplicativa entonces, el término $\exp(b_1)$ es la OR del grupo codificado con el valor 1 con respecto al codificado con el valor 0.

Finalmente, si la codificación es respecto al promedio de categorías, el paso del grupo codificado con valor -1 al codificado con el valor $+1$ implica un incremento de 2 unidades en la variable explicativa.

7.3. MODELO DE REGRESIÓN LOGÍSTICA MÚLTIPLE

En este modelo al menos uno de los predictores debe ser cuantitativo, ya que si todos fuesen categóricos, deberíamos utilizar el modelo de regresión logit (por la agrupación de los datos) (Glonck y McCullagh 1995).

Este modelo puede incluir una combinación lineal de variables explicativas categóricas o cuantitativas, así como sus términos de interacción. Además, nos permite ajustar el efecto de una variable mediante variables de control. Su ecuación es:

$$\text{logit}(m_i) = b_0 + b_1 X_{i1} + b_2 X_{i2} + \dots + b_k X_{ik} = b_0 + \sum_{j=1}^k b_j X_{ij} \quad (7.16)$$

Si el modelo no presenta términos de interacción se interpreta igual que el modelo de regresión logística simple, teniendo en cuenta que al interpretar un parámetro se consideran fijados los demás. Así, $\exp(b_1)$ se interpreta como la OR asociada a un incremento de una unidad en X (Skrondal y Rabe-Hesketh, 2003).

En el caso de que el modelo presente términos de interacción, éste se incluye en el modelo como un nuevo término. Con dos variables explicativas la ecuación sería:

$$\text{logit}(m_i) = b_0 + b_1 X_i + b_2 Z_2 + b_3 XZ_i \quad (7.17)$$

donde b_3 es el coeficiente de regresión del término de interacción entre la variable explicativa X y la variable modificadora Z .

No se recomienda incluir términos de interacción con más de dos variables ya que, por una parte, su interpretación es bastante compleja, y por otra, sucede que suele

aparecer colinealidad, lo que nos impedirá una correcta estimación de los parámetros (Jaccard 2001).

7.3.1 COMPARACIÓN DE MODELOS

En los modelos de predictores múltiples debemos acudir al test de la diferencia de discrepancias para el ajuste y selección del modelo más adecuado.

De este modo se compara un MR y un MA:

$$\Delta D = D_{MR} - D_{MA} = -2 \log L_{MR} - (-2 \log L_{MA}) = -2 \log \left(\frac{L_{MR}}{L_{MA}} \right) \quad (7.18)$$

El valor resultante sigue una distribución χ^2 con grados de libertad igual a la diferencia de parámetros ($\Delta gl = p_{MR} - p_{MA}$).

En el caso de que incluyamos términos de interacción en el modelo la interpretación se complica debido a que la estimación de la OR de la variable explicativa depende del valor de otra variable. De este modo, lo mejor para interpretar es derivar relaciones simples (Jaccard 2001).

CAPÍTULO 8

REGRESIÓN LOGIT

8.1 – INTRODUCCIÓN

A diferencia de la regresión logística, estos modelos nos permiten modelar la relación entre una variable de respuesta categórica (binaria o politómica) y una o varias variables explicativas categóricas. Su nombre proviene del logaritmo de las *odds* que se conoce como *logit* (DeMaris 1992).

El modelo logit fue introducido por Berk (1944), y su origen es una analogía del modelo probit desarrollado por Bliss (1934). Más tarde Barnard (1949) definió el logaritmo de las *odds* como el *logit* de la probabilidad de un evento.

De este modo, se pueden organizar los datos en una tabla de contingencia. En estos modelos, una de las variables es de respuesta y el resto son variables explicativas, por lo tanto, tenemos una relación de asimetría entre las variables, los cuales podemos modelar mediante un modelo de regresión *logit*. Sin embargo hay que señalar que algunos modelos *logit* pueden ser expresados bajo la forma de un modelo loglineal.

Los modelos de *logit* pueden seguir una distribución binomial (ver tabla y figura 8.1).

Tabla 8.1 Características de la distribución de Binomial

Distribución	$y_i \sim B(N, \pi_i)$
Parámetro canónico	$\log(\pi_i / 1 - \pi_i)$
Parámetro de escala	1
Función de enlace	$\eta_i = \log(\mu_i / N - \mu_i)$
Función de varianza	$V(\mu) = \frac{\mu(N - \mu)}{N}$
Desvianza	$D = \sum w [y_i - \log(y_i / \hat{\mu}_i) + (N - y_i) \log(N - y_i / (N - \hat{\mu}_i))]$

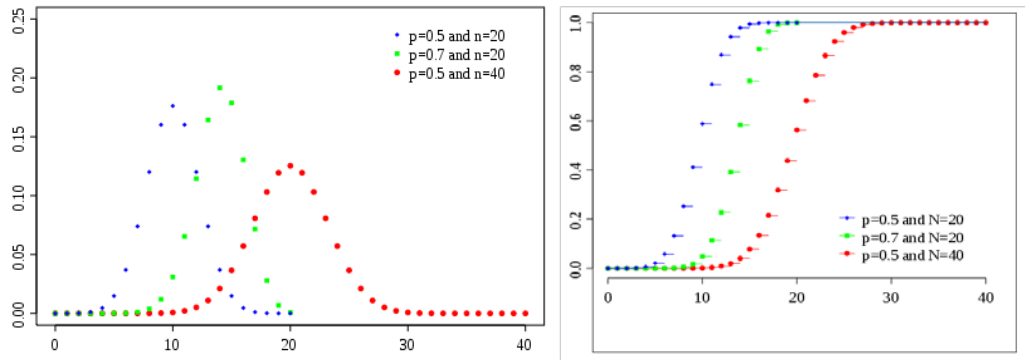


Figura 8.1 Función de densidad y función de la distribución Binomial (Wikipedia)

En el ámbito de la memoria de reconocimiento estos modelos tienen sentido ya que es posible trabajar con datos exclusivamente categóricos si trabajamos con la Teoría de Detección de Señales (TDS) y además como expone DeCarlo (1998) existe relación entre la TDS y los modelos logit.

8.2. MODELO DE REGRESIÓN LOGIT SIMPLE

Este es el caso en el que tenemos dos variables categóricas (para ejemplificar utilizaremos una variable dicotómica) que podemos agrupar en una tabla 2 x 2 (ver tabla 8.2):

Tabla 8.2 Tabla de contingencia 2x2

A	B		Total	O_{il}^{AB}
	1	2		
1	n_{11} (P_{11})	n_{12} (P_{21})	n_{1+}	$O_{11}^{AB} = \frac{n_{11}}{n_{12}}$
2	n_{21} (P_{12})	n_{22} (P_{22})	n_{2+}	$O_{21}^{AB} = \frac{n_{21}}{n_{22}}$
Total	n_{+2}	n_{+1}	n_{++}	$O_{+1}^{AB} = \frac{n_{+1}}{n_{+2}}$

En la tabla A es la variable explicativa y B la variable de respuesta, n_{ij} son las frecuencias observadas, n_{i+} y n_{+j} son los totales marginales de fila y columna, y n_{++} es la suma de los totales marginales (Agresti, 1990).

Dada esta tabla podemos obtener la probabilidad condicional $p_{ij} = n_{ij}/n_{i+}$, y la *odds* observada que nos indicará la probabilidad relativa de pertenecer a la categoría uno de la variable de respuesta:

$$O_{i1}^{AB} = \frac{P_{(1i)}}{P_{(2i)}} = \frac{n_{i1}/n_{i+}}{n_{i2}/n_{i+}} = \frac{n_{i1}}{n_{i2}} \quad (8.1)$$

Como podemos observar esta *odds* se obtiene a partir de las frecuencias observadas, y sus valores oscilan entre 0 y $+\infty$. Su interpretación sería la que proponemos a continuación (Liao 1994):

- Si $O_{i1}^{AB} = 1$, la probabilidad de una respuesta favorable es igual a la de una respuesta desfavorable.
- Si O_{i1}^{AB} está entre 0 y 1, la probabilidad de una respuesta favorable es menor que la de una respuesta desfavorable.
- Si $O_{i1}^{AB} > 1$, la probabilidad de una respuesta favorable es mayor que la de una respuesta desfavorable.

El modelo se formula de la misma manera que el de regresión logística, pero definiéndolo en términos de frecuencias. Su ecuación general sería:

$$\eta = \text{logit}(\mu) = \beta_0 + \beta_1 A \quad (8.2)$$

Y para una muestra:

$$\text{logit}(m_i) = \log(o_{i1}^{AB}) = \log\left(\frac{m_{i1}}{m_{i2}}\right) = b_0 + b_1 A \quad (8.3)$$

donde o_{i1}^{AB} es el valor esperado de las *odds* de la variable de respuesta en la categoría i de la variable explicativa A, m_{i1} y m_{i2} son las frecuencias esperadas en las categorías 1 y 2 de la variable de respuesta B para la categoría i de A.

Los componentes básicos del modelo son (Ato y López 1996):

- El componente aleatorio: la distribución del error sigue una ley binomial n .
- El componente sistemático: el predictor lineal que proporciona el valor predicho es $\eta = \beta_0 + \beta_1 A$.
- La función de enlace: es la función *logit* $\eta = \text{logit}(\mu)$.

En el MS para una tabla bidimensional 2 x 2, las frecuencias y las *odds* esperadas coinciden con las observadas y tiene 0 grados de libertad. En los modelos *logit* los grados de libertad se calculan como la diferencia entre el número de observaciones y el número de parámetros estimados por el modelo. De este modo, el número de observaciones es realmente el número de *odds* o *logits* observados, es decir, el número de categorías de la variable explicativa (Aldrich y Nelson 1984).

Al codificar la variable explicativa A mediante la codificación respecto a una categoría de referencia, el coeficiente b_0 será el valor *logit* cuando $A = 0$. El coeficiente b_1 es el cambio que se produce en el *logit* por un incremento de 1 unidad en A. Será más fácil la interpretación si utilizamos la forma multiplicativa del modelo:

$$o_{i1}^{AB} = \left(\frac{m_{i1}}{m_{i2}} \right) = \exp(b_0) \exp(b_1 A) \quad (8.4)$$

De esta manera podemos obtener el valor de la *odds* esperado por el modelo para cada categoría de la variable explicativa. Así $\exp(b_0)$ es el valor de la *odds* cuando $A = 0$ y $\exp(b_1)$ es el valor de la OR (la razón entre la *odds* cuando $A = 1$ y la *odds* cuando $A = 0$) (Ato, *et al.*, 2005):

$$\exp(b_0) = o_{11}^{AB} \quad \exp(b_1) = \frac{o_{21}^{AB}}{o_{11}^{AB}} \quad (8.5)$$

El valor de OR estará entre 0 y $+\infty$:

- si $or = 1$: la probabilidad de una respuesta favorable es igual en ambas categorías de respuesta.
- Si or está entre 0 y 1: la probabilidad de respuesta favorable es menor en el grupo de comparación que en el grupo de referencia.
- Si $or > 1$: la probabilidad de respuesta favorable es mayor en el grupo de comparación con respecto al grupo de referencia.

8.2.1 ESTIMACIÓN DE LOS PARÁMETROS DEL MODELO

Utilizaremos el estadístico z como prueba de significación estadística del coeficiente de regresión b_1 (Lindsey, 1997):

$$z = \frac{b_1}{EE(b_1)} \quad (8.6)$$

Se distribuye asintóticamente según $N(0,1)$. También podemos utilizar el estadístico de Wald (1945) (z^2) que se distribuye según χ^2 con un grado de libertad.

Para calcular el IC utilizamos la siguiente fórmula:

$$IC95\%(\beta_1) = b_1 \pm z_{0,025} EE(b_1) \quad (8.7)$$

Los límites superior e inferior del IC representan los valores máximo y mínimo del cambio esperado del *logit* por un incremento en una unidad de la variable explicativa de la población. También se puede interpretar en términos de OR.

8.2.2 COMPARACIÓN DE MODELOS

En el modelo simple solo tenemos dos modelos que ajustar, por una parte, el MA, que incluirá el efecto de la única variable explicativa:

$$\log\left(\frac{m_{i1}}{m_{i2}}\right) = b_0 + b_1 A \quad (8.8)$$

En este caso MA=MS, por lo tanto $m_{ij} = n_{ij}$, $D = 0$, y $gl = 0$

Por otra parte tenemos el MR, que no incluye la variable explicativa A:

$$\log\left(\frac{m_{i1}}{m_{i2}}\right) = b_0 \quad (8.9)$$

En este caso MR = MN. Este modelo estima solo un odds para la variable de respuesta. Se denomina también este modelo como modelo de independencia, ya que su hipótesis es que A y B son estadísticamente independientes.

Utilizaremos una estrategia de ajuste global para comparar el MR y MS, de esta manera evaluaremos la diferencia entre las frecuencias esperadas por el MR y las frecuencias observadas. En los modelos logit si es interesante evaluar el ajuste global con respecto a un MS, ya que nos proporciona un criterio de ajuste muy bueno para comparar otros modelos más parsimoniosos (Ato y López 1996). Lo que

haremos será comparar la verosimilitud del MR con la del MS mediante la razón verosimilitud:

$$LR = -2(\log L_{MR} - \log L_{MS}) = -2 \log(L_{MR}/L_{MS}) \quad (8.10)$$

También podemos obtener el valor LR mediante el estadístico G^2 , que para el caso de tablas 2x2 se calcula como:

$$G^2 = 2 \sum n_{ij} \log \left(\frac{n_{ij}}{m_{ij}} \right) \quad (8.11)$$

Al igual que LR evalúa la diferencia de discrepancias de los dos modelos, y dado que para el MS, $G^2 = 0$, entonces $\Delta D = \Delta G^2 = G_{MR}^2 - G_{MS}^2 = G_{MR}^2$. Este estadístico sigue una distribución χ^2 con grados de libertad igual al número de *odds* menos el número de parámetros independientes estimados. Hemos de indicar que cuando G^2 es estadísticamente significativo, entonces MR y MS difieren significativamente (el MR no explica la variabilidad), pero si no es estadísticamente significativo indica que los modelos no difieren eligiendo por tanto el MR.

Como podemos observar en los modelos de regresión logit, el valor de la discrepancia es interpretable en si mismo, ya que representa la comparación entre los valores esperados y los observados (Agresti 1990). En este sentido es especialmente recomendable aquí aplicar la estrategia de comparación de modelos, ya que el estadístico de Wald es poco fiable cuando lo aplicamos con muestras pequeñas.

8.3 MODELO DE REGRESIÓN LOGIT MÚLTIPLE

En los modelos de regresión logit múltiples, tenemos que distinguir entre modelos que incluyen una variable predictora politómica y los que incluyen varias variables, además, también habrá que diferenciar entre aquellos con interacciones entre variables explicativas y los que no las incluyen

8.3.1. MODELO CON VARIABLE EXPLICATIVA POLITÓMICA

Siendo A la variable explicativa con número de categorías I, y B la variable de respuesta con número de categorías $J = 2$, la expresión aditiva del modelo es:

$$\log\left(\frac{m_{i1}}{m_{i2}}\right) = b_0 + b_i^A \quad (8.12)$$

donde b_i^A es la representación de cada uno de los efectos de la categoría i de la variable explicativa A. al tener la variable explicativa más de dos categorías hay que introducir $I - 1$ variables ficticias. Si codificamos con respecto a una categoría de referencia, y siendo ésta la primera categoría de la variable explicativa, cada una de las variables ficticias compara la categoría i con la primera categoría. De este modo $\beta_1^A = 0$, y el modelo quedaría de la siguiente manera:

$$\log\left(\frac{m_{i1}}{m_{i2}}\right) = b_0 + b_2 A_2 + b_3 A_3 + \dots + b_I A_I \quad (8.13)$$

Hemos de indicar que este modelo tiene cero grados de libertad (numero de *odds* = número de parámetros independientes estimados). La expresión multiplicativa de este modelo sería:

$$o_{i1}^{AB} = \frac{m_{i1}}{m_{i2}} = \exp(b_0) \exp(b_i^A) \quad (8.14)$$

Los coeficientes del modelo multiplicativo se pueden interpretar en términos de odds y OR (Liao 1994):

- El coeficiente $\exp(b_0)$: es el valor de *odds* predicho cuando $A_i = 0$
- El coeficiente $\exp(b_i^A)$: es el valor de la OR entre el grupo i de A y el grupo de referencia A_1 .

$$o_{11}^{AB} = \exp(b_0) \quad \frac{o_{i1}^{AB}}{o_{11}^{AB}} = \exp(b_i^A) \quad (8.15)$$

8.3.1.1 ESTIMACIÓN DE LOS PARÁMETROS DEL MODELO

Al igual que con el modelo simple, el estadístico z se utiliza como prueba de significación estadística del coeficiente de regresión b_i^A . Cuando un valor de z no es estadísticamente significativo, significa que los *logits* predichos por el modelo son iguales para las categorías de la variable explicativa (Liao 1994). Si utilizamos el modelo multiplicativo, esto significaría que la *odds* de ocurrencia del fenómeno estudiado en la categoría i de A es igual a la *odds* de la categoría de referencia.

El cálculo del IC para β_i^A se realiza siguiendo la siguiente ecuación:

$$IC95\%(\beta_i^A) = b_i^A \pm z_{0,025} EE(b_i^A) \quad (8.17)$$

donde los límites superior e inferior del IC representan los valores máximo y mínimo del cambio esperado en el *logit* al comparar la categoría i de A con la categoría de referencia en la población. El IC correspondiente para la OR se calcula:

$$IC95\%(\exp(\beta_i^A)) = \exp(b_i^A) \pm z_{0,025} EE(b_i^A) \quad (8.18)$$

8.3.1.2 COMPARACIÓN DE MODELOS

Se realiza de manera similar al modelo de regresión *logit* simple, ya que lo que hacemos es evaluar la discrepancia del MR con respecto al MS mediante el estadístico G^2 .

El MR no incluye el efecto de la variable explicativa A:

$$\log\left(\frac{m_{i1}}{m_{i2}}\right) = b_0 \quad (8.19)$$

Este modelo predice un solo *odds* para la variable de respuesta de manera independiente a la variable explicativa: $b_0 = n_{+1}/n_{+2}$, $o_{i1}^{AB} = \exp(b_0)$.

El ajuste global del MR se evalúa calculando la discrepancia del modelo con respecto al MS, de este modo comparamos frecuencias observadas y esperadas.

Se calcula mediante el estadístico G^2 (Agresti, 1990):

- Si G^2 es estadísticamente significativo entonces los ajustes de MR y MS van a diferir significativamente.
- Cuando G^2 no es estadísticamente significativo los ajustes de MR y MS no van a diferir, escogiéndose por tanto el MR.

8.3.2 MODELO CON VARIAS VARIABLES PREDICTORAS

8.3.2.1 INTRODUCCIÓN

Aquí lo que nos interesa es estudiar el efecto de una variable explicativa sobre una variable de respuesta, y comprobar si esta relación puede verse modificada con

la introducción de una nueva variable en el modelo (Agresti, 1990). Para comprobarlo, como hemos expuesto anteriormente, la asociación marginal no tiene en cuenta la segunda variable explicativa y, sin embargo, la asociación parcial si la tiene en cuenta.

Si la asociación marginal y parcial difieren, se da un fenómeno denominado “Paradoja de Simpson” como vimos en los modelos loglineales. Este fenómeno se debe a que la variable a través de la cual ha sido colapsada la tabla está relacionada tanto con la variable explicativa como con la variable de respuesta, lo cual, está relacionado con el fenómeno de la confusión que se da en los modelos de regresión lineal. El hecho de que aparezca este fenómeno no es suficiente para que esta variable sea considerada confundidora, ya que si el estudio de esta variable era uno de los objetivos de la investigación no se puede considerar como tal. Pero, sin embargo, si no hay efecto entre esta variable y la variable explicativa objetivo, entonces estaremos ante una variable confundidora.

Pongamos como ejemplo que tuviésemos un modelo logit con dos variables explicativas, en este caso serían 5 los modelos posibles a ajustar (ver tabla 8.3).

Tabla 8.3 Modelos logit posibles para una tabla con dos variables explicativas.

Modelo	$\log (m_{ij1}/m_{ij2})$	Nº parámetros
[1]	$= b_0$	1
[A]	$= b_0 + b_i^A$	$1 + (I - 1)$
[B]	$= b_0 + b_i^B$	$1 + (J - 1)$
[A,B]	$= b_0 + b_i^A + b_i^B$	$1 + (I - 1) + (J - 1)$
[AB]	$= b_0 + b_i^A + b_i^B + b_{ij}^{AB}$	$1 + (I - 1) + (J - 1) + (I - 1) * (J - 1)$

En la tabla 8.3 b_i^A es el efecto de la categoría i de A sobre el *logit* de la variable de respuesta, b_j^B es el efecto de la categoría j de B sobre el *logit* de la variable de respuesta y b_{ij}^{AB} es el efecto de interacción de las variables explicativas sobre la variable de respuesta.

Aquí el MS es [A,B], ya que incorpora los efectos marginales y la interacción entre las variables explicativas, y el MN es [1] que estima un solo parámetro y predice solo un *logit* independientemente de las variables explicativas.

8.3.2.2 COMPARACIÓN DE MODELOS

Para ello podremos utilizar la estrategia de ajuste global, que consiste en evaluar la significación estadística de la discrepancia de cada modelo con respecto al MS. Si el valor de la discrepancia no es significativo para un modelo, las frecuencias esperadas no difieren de las observadas (por lo tanto el modelo será adecuado para explicar la relación entre las variables). Para obtener el valor de la discrepancia utilizaremos nuevamente el estadístico G^2 que para el caso de dos variables explicativas se calcula:

$$G^2 = 2 \sum n_{ijk} \log \left(\frac{n_{ijk}}{m_{ijk}} \right) \quad (8.20)$$

Si después de aplicar la estrategia de ajuste global tenemos varios modelos con un ajuste adecuado, debemos utilizar la estrategia de ajuste condicional, para seleccionar el modelo que represente mejor a los datos, pero con el menor número de parámetros. Es un procedimiento similar a la Razón F parcial para comparar modelos de regresión lineal (Agresti 1990).

Entonces evaluaremos la significación estadística de la diferencia de discrepancias entre un MR y un MA en el que está anidado. De este modo, podremos evaluar la significación estadística de un parámetro y la significación estadística de un conjunto de parámetros. Esta diferencia de discrepancias es la diferencia entre los valores G^2 de los dos modelos anidados. Dicha diferencia se aproxima asintóticamente a χ^2 con grados de libertad iguales a la diferencia de grados de libertad de los modelos comparados, suponiendo que el MR es correcto.

Si la diferencia es estadísticamente no significativa la eliminación de términos del MA no disminuye el ajuste significativamente, por lo tanto elegiríamos el MR. Si por el contrario, la diferencia es estadísticamente significativa, la eliminación de términos del MA reduce significativamente el ajuste del modelo, por lo tanto elegiríamos el MA.

Podemos iniciar el proceso desde el MN o el MS. Si lo iniciamos desde el MN iremos incorporando parámetros al modelo, finalizaremos cuando la adición de parámetros no disminuya la discrepancia significativamente, pero si comenzamos por el MS, lo que haremos será ir eliminando términos del modelo, y finalizaremos cuando esta eliminación de términos provoque que disminuya el ajuste del modelo.

Hemos de añadir que debemos respetar siempre el principio de jerarquía de los modelos, ya sea por adición o eliminación (DeMaris 2002).

Por otra parte, si tomamos los modelos posibles para dos variables explicativas los pasos a seguir para la selección del modelo más adecuado serían (Agresti 1990):

- Evaluar el ajuste global del modelo [A,B]: si la discrepancia es estadísticamente significativo habría que elegir el MS, y no seguiríamos con el proceso.
- Si la discrepancia es estadísticamente no significativa: seleccionaremos este modelo y comenzaremos a realizar los ajustes condicionales.
- Se compararían los modelos anidados [A,B]-[A] y [A,B]-[B]: si resultasen estadísticamente significativos no pueden ser eliminados del modelo, por lo cual nos quedaríamos con el modelo [A,B].
- Si es significativo solo uno de los modelos marginales elegiríamos este, ya sea [A] o [B].
- Si no son estadísticamente significativos deberíamos elegir el MN.

Es necesario señalar que son varios los índices para medir la bondad de ajuste de los modelos, sobre todo si se trabaja con modelos no anidados. Pero uno de los criterios más utilizados, sobre todo cuando todas las variables son categóricas, es el criterio de información bayesiano: $BIC = D - (\log n)gl$ o $AIC = D - 2gl$ (Shibata, 2002). Es importante señalar que los valores del BIC o AIC no son interpretables en sí mismos, sino que hay que hacerlo en comparación con los valores de este índice de los otros modelos comparados, ya que lo que favorece es la parsimonia del modelo, penalizando la inclusión de parámetros a este. Por lo tanto, como se vio anteriormente, se seleccionará el modelo con el valor del índice BIC o AIC menor, ya que sabremos que este modelo nos está proporcionando el máximo de información con el menor número de parámetros.

En el caso de que nuestro modelo no incluyese términos de interacción y siguiendo con nuestro ejemplo con dos variables explicativas, el modelo [A,B] sería:

$$\log \frac{m_{ij1}}{m_{ij2}} = b_0 + b_i^A + b_j^B \quad (8.21)$$

Si expresamos el modelo en su forma multiplicativa nos será más fácil realizar la interpretación de los coeficientes:

$$O_{ij1}^{ABC} \frac{m_{ij1}}{m_{ij2}} = \exp(b_0) \exp(b_i^A) \exp(b_j^B) \quad (8.22)$$

donde $\exp(b_0)$ es la *odds* esperada de la variable de respuesta en las categorías de referencia de las variables explicativas: $\exp(b_0) = O_{111}^{ABC} = \frac{m_{111}}{m_{222}}$

donde $\exp(b_i^A)$ es la cantidad por la que se multiplica la *odds* esperada en la categoría i de A con respecto a la categoría 1 de A para cada nivel de B, o dicho de otro modo, la OR de la variable de respuesta entre la categoría i de A y la categoría

uno de A para cualquier categoría de B: $\exp(b_i^A) = \frac{O_{ij1}^{ABC}}{O_{1j1}^{ABC}} = \frac{m_{ij1}}{m_{1j1}} / \frac{m_{ij2}}{m_{1j2}}$

donde $\exp(b_j^B)$ es la cantidad por la que se multiplica la *odds* esperada en la categoría i de A con respecto a la categoría j de B para cada nivel de A, o dicho de otro modo, la OR de la variable de respuesta entre la categoría j de B y la categoría

1 de B para cualquier categoría de A: $\exp(b_j^B) = \frac{O_{ij1}^{ABC}}{O_{i11}^{ABC}} = \frac{m_{ij1}}{m_{i11}} / \frac{m_{ij2}}{m_{i12}}$

En el caso de que nuestro modelo incluya términos de interacción y siguiendo con nuestro ejemplo de dos variables explicativas, el modelo [AB] sería:

$$\log \frac{m_{ij1}}{m_{ij2}} = b_0 + b_i^A + b_j^B + b_{ij}^{AB} \quad (8.23)$$

Si expresamos el modelo en su forma multiplicativa nos será más fácil realizar la interpretación de los coeficientes:

$$o_{ij1}^{ABC} \frac{m_{ij1}}{m_{ij2}} = \exp(b_0) \exp(b_i^A) \exp(b_j^B) \exp(b_{ij}^{AB}) \quad (8.24)$$

donde $\exp(b_{ij}^{AB})$ es el efecto sobre el *logit* de la variable de respuesta de la interacción entre A y B, es decir, como el efecto de una de las variables cambia en función del nivel de la otra variable, o también el valor por el que se multiplica la OR entre la categorías j de B al pasar de la categoría 1 a la categoría i de A, o a la inversa (ver fórmula 8.25).

$$\exp(b_{ij}^{AB}) = \frac{o_{ij1}^{ABC}}{o_{i11}^{ABC}} \bigg/ \frac{o_{1j1}^{ABC}}{o_{111}^{ABC}} = \frac{\frac{m_{ij1}}{m_{i11}}}{\frac{m_{1j1}}{m_{111}}} \bigg/ \frac{\frac{m_{ij2}}{m_{i12}}}{\frac{m_{1j2}}{m_{112}}} \quad (8.25)$$

8.4 EQUIVALENCIA ENTRE MODELOS LOGIT Y LOGLINEALES

Los modelos logit pueden verse como un caso especial de los modelos log-lineales, cuando la variable respuesta se considera dependiente de las otras variables, por lo que en realidad es una simplificación del modelo loglineal (Christensen 1997).

Los principios a seguir para esta equivalencia son: (Ato y López 1996):

- Un modelo loglineal contendrá como componente específico la estructura asociativa de todas las variables explicativas. Es decir, la interacción entre todas las variables explicativas.
- Un modelo loglineal contendrá los mismos componentes del logit más la variable de respuesta con todas las interacciones con los componentes del modelo logit.

Al realizar el modelado correspondiente obtendríamos los mismos ajustes, pero diferente número de parámetros. También hay que señalar que no todos los modelos loglineales tienen un equivalente logit.

CAPÍTULO 9

EVALUACIÓN DE MODELOS AJUSTADOS

If you torture the data enough, nature will always confess. Ronald Coase.

9.1. TÉCNICAS DE EVALUACIÓN DE UN MODELO AJUSTADO

Un modelo se dice que presenta un buen ajuste a los datos si los valores predichos reflejan de forma adecuada a los valores observados. Si el modelo presenta un mal ajuste, este no puede ser utilizado para extraer conclusiones ni efectuar predicciones.

Los modelos ajustados mediante modelado estadístico pueden resultar inadecuados por diferentes razones. Una de ellas es la invalidez del componente lineal del modelo, esta situación es frecuente cuando ciertas variables predictoras o términos de interacción no son incluidos cuando deberían estar incluidos, o también cuando no se realizan transformaciones en ellas que permiten mejorar el ajuste a los datos. La existencia de observaciones influyentes u *outliers* puede ser también determinante en un mal ajuste. Otra de las situaciones habituales en un mal ajuste se tiene cuando la función de enlace no es la apropiada.

El mal ajuste es posible también cuando no se cumple la hipótesis de distribución de la variable de respuesta, que puede ser debido a que las puntuaciones observadas no sean independientes, o a la existencia de sobredispersión o infradispersión (Collett 1991) provocando una estimación de la varianza mayor o inferior que la que debería ser.

Los llamados diagnósticos del modelo se pueden realizar a partir de:

- Test estadísticos diagnósticos
- Análisis de los residuales

- Representaciones gráficas

En este trabajo con centraremos únicamente en los dos primeros.

9.2 TEST DIAGNÓSTICOS

Una forma de medir la adecuación de un modelo es proporcionando medidas globales de bondad de ajuste a través de test estadísticos construidos a tal fin.

9.2.1 Leverages o valores de influencia

Se considera que una observación es influyente a priori si su inclusión en el análisis modifica sustancialmente el sentido del mismo. Una observación puede ser influyente si es un *outlier* respecto a alguna de las variables explicativas (Stevens 1984).

El estimador insesgado de la varianza de los errores es:

$$s^2 = \frac{R'R}{n-p} \quad (9.1)$$

donde R es el vector de los residuales, n el número de observaciones y p el número de parámetros estimados. Los elementos diagonales de la matriz H simbolizadas como h_{ij} son los valores influyentes (*leverages*). De este modo las observaciones de las variables explicativas con un valor de influencia alto pueden estar condicionando el ajuste del modelo. Este valor queda acotado por:

$$\frac{1}{n} \leq h_{ij} \leq \frac{1}{r} \quad (9.2)$$

donde n es el número de observaciones, y r el numero de repeticiones con el mismo patrón de valores de las variables explicativas. El valor de influencia tenderá a ser menor a medida que el número de observaciones aumente (Atkinson 2000).

También, podemos disminuir la influencia de una observación incluyendo otras con su mismo patrón de respuesta. De este modo el valor predicho \hat{Y}_i se puede expresar:

$$\hat{Y}_i = h_{ij}Y_i + \sum_{i \neq j} h_{ij}Y_j \quad (9.3)$$

Con esta formulación para los valores influyentes altos los valores \hat{Y}_i estarán determinados por el valor de observado para este sujeto, y dado que $Var(r_i) = (1 - h_{ii})\sigma^2$, estos valores influyentes altos tendrán una varianza en su residual pequeña. También podemos obtener la covarianza entre dos residuales (Kraznowski, 1998) $Covar(r_i, r_j) = -h_{ij}\sigma^2$, no siendo independientes los residuales entre si, la covarianza será baja siempre que tengamos una muestra grande.

Hoaglin y Welsch (1978) proponen que la siguiente fórmula para determinar si una observación será influyente:

$$h_{ii} > \frac{2p}{n} \quad (9.4)$$

Mediante la Regresión Lineal Simple podemos calcular el valor de influencia de cada observación mediante:

$$h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum (x_j - \bar{x})^2} \quad (9.5)$$

donde h_i nos proporciona la magnitud la magnitud en la que x_i se diferencia del valor promedio del conjunto de las observaciones. Cuanto más lejano esté un valor x_i de su promedio, mayor será su valor de influencia. Podemos obtener los valores de influencia centrados o *centered leverage*, restando del valor de influencia de una observación el valor $1/n$. La media de la variable de los valores de influencia centrados será $(p-1)/n$ y la media del valor de influencia será p/n (donde p es el numero de parámetros del

modelo). Así, cuanto mayor sea el valor h_i de una observación, mayor impacto tendrá sobre la constante, la pendiente y los valores predichos.

Pero cuando eliminamos esa observación entonces podemos evaluar su influencia mediante la obtención del cambio que se produce en los coeficientes. (Fox 1991). Para evaluar el cambio producido en el estimador b_1 al no incluir la observación i utilizamos la siguiente fórmula:

$$b_1 - b_{1(i)} = B_i \frac{r_i}{1 - h_i} \quad (9.6)$$

donde $b_{1(i)}$ es el valor de b_1 resultante de eliminar la observación i , y r_i es el residual

$$\text{ordinario para la observación } i (y_i - \hat{y}_i) \text{ y } B_i = \frac{x_i - \bar{x}}{\sum (x_j - \bar{x})^2}. \quad (9.7)$$

Sin embargo, si lo que deseamos es evaluar el cambio producido en la constante b_0 al no incluir la observación i , utilizamos la siguiente fórmula:

$$b_0 - b_{0(i)} = A_i \frac{r_i}{1 - h_i} \quad (9.8)$$

donde $b_{0(i)}$ es el valor de b_0 resultante de eliminar la observación i , y r_i es el residual

$$\text{ordinario para la observación } i (y_i - \hat{y}_i) \text{ y } A_i = \frac{\sum x_j^2}{n \sum (x_j - \bar{x})^2} - \frac{x_i \bar{x}}{\sum (x_j - \bar{x})^2} \quad (9.9)$$

Por otra parte, si necesitamos evaluar el cambio producido en el valor predicho para la observación i utilizamos la fórmula:

$$\hat{y}_i - \hat{y}_{i(i)} = \frac{h_i}{1 - h_i} r_i \quad (9.10)$$

No obstante, mediante la Regresión Lineal Múltiple podemos calcular el cambio en el vector b de estimación del modelo cuando se excluye la observación i con la siguiente fórmula:

$$b_i - b_{(i)} = (X'X)^{-1} x_i' \frac{r_i}{1 - h_i} \quad (9.11)$$

donde $x_i' = (1, x_{i1}, \dots, x_{ik})'$ y h_i es el elemento diagonal de la matriz H .

Finalmente, para evaluar el cambio en el vector de los valores predichos lo hacemos mediante:

$$\hat{y} - \hat{y}_{(i)} = X(X'X)^{-1} x_i' \frac{r_i}{1 - h_i} \quad (9.12)$$

Es decir, utilizamos la misma fórmula que en la regresión simple.

9.2.2 Distancia de Mahalanobis

Es un índice alternativo al de los valores de influencia introducido por Mahalanobis (1936). Mediante éste podemos medir la distancia de una observación i hasta el centroide (o punto de intersección de las medias de las variables explicativas):

$$MD = (v_i - \bar{v})C^{-1}(v_i - \bar{v})' \quad (9.13)$$

donde v es el vector de fila de los valores de las k variables explicativas,

$v_i = (x_{i1}, \dots, x_{ik})$, \bar{v} es el centroide y $C = \frac{1}{n-1} \sum_{i=1}^n (v_i - \bar{v})'(v_i - \bar{v})$ es la matriz de covarianzas.

El valor resultante de la fórmula podemos compararlo con los valores de la distribución χ^2 con k grados de libertad, siendo k el número de variables explicativas del modelo de regresión sin incluir la constante.

Para calcular la relación entre la MD y el valor de h_{ii} se utiliza:

$$MD_i = (n-1) \left[h_{ii} - \frac{1}{n} \right] \text{ o } h_{ii} = \frac{1}{n} + \frac{MD_i}{n-1} \quad (9.14)$$

La interpretación de este índice es que a mayor valor de la MD mayor será el valor de influencia.

9.2.3 Distancia de Cook

Introducida por Cook (1977), este índice nos permite evaluar la influencia de una observación sobre la estimación de los coeficientes de regresión. Lo hace comparando los coeficientes de regresión estimados al incluir y eliminar la observación i del modelo. La influencia de una observación se refleja no solo en las variables explicativas sino también en la variable de respuesta.

$$D_i = \frac{(\theta - \theta_{(i)})' X' W X (\theta - \theta_{(i)})}{p \hat{\phi}} \quad (9.15)$$

donde θ es el vector de estimaciones de los parámetros, $\theta_{(i)}$ es el vector cuando la observación i ha sido eliminada y $\hat{\phi}$ es el parámetro de escala estimado. Los valores altos de D son los que tienen influencia en el modelo (Cook y Weisberg 1982). Un valor D igual o cercano a uno puede considerarse grande, aunque Hutchenson y Sofroniu (1999) lo consideran demasiado conservador y proponen como punto de corte $4/n - p$.

9.2.4 Índice DFFITS

Mediante este índice se mide la influencia sobre el valor predicho de una observación i cuando esta es eliminada (Belsley, Kuh y Welsch, 1980):

$$DFFITS(i) = \frac{r_i}{s_{(i)}} \frac{\sqrt{h_{ii}}}{1 - h_{ii}} \quad (9.16)$$

donde r_i es el residual ordinario de la observación i , y $s_{(i)}$ es la desviación estándar de la variable de respuesta del modelo sin la observación i .

El criterio para saber si un valor DFFITS es elevado nos lo proporciona en valores absolutos la siguiente fórmula $2\sqrt{p/n}$.

9.2.5 Índice DFBETAS

Aplicando este índice podemos evaluar el cambio que se produce en la estimación del parámetro j al eliminar la observación i (Belsley, Kuh y Welsch, 1980).

$$DFBETAS_j(i) = \frac{c_{ji}}{\sqrt{\sum_{i=1}^n c_{ji}^2}} \frac{r_i}{s_{(i)}(1-h_{ii})} \quad (9.17)$$

donde c_{ji} es el elemento diagonal de la matriz $C = (X'X)^{-1}$. Un criterio para considerar si un valor de este índice es alto o no es: $2/\sqrt{n}$.

9.3 RESIDUALES

Como es sabido, los residuales son la diferencia entre el valor observado y el valor ajustado por el modelo. Si obtenemos residuales con valores altos nos estarían indicando la presencia de observaciones que no han sido bien ajustadas (Cook y Weisberg 1982).

9.3.1 Residual ordinario

Nos permite comparar el valor observado con el valor predicho.

$$r_i = y_i - \hat{y}_i = y_i - \hat{\mu}_i \quad (9.18)$$

9.3.2 Residual de Pearson

Resulta de dividir r_i por su desviación estándar propuesto por Pearson (1900).

$$r_i^P = \frac{y_i - \hat{\mu}_i}{\sqrt{\text{Var}(y_i)}} \quad (8.19)$$

A partir de este residual podemos obtener el estadístico χ^2 de bondad de ajuste:

$$\chi^2 = \sum (r_i^P)^2 \quad (9.20)$$

9.3.3 Residual de Pearson estudentizado

Se calcula dividiendo r_i por su desviación estándar teniendo en cuenta los valores de influencia (Pierce y Schafer, 1986). Se utiliza para detectar *outliers*. Señalar que estos residuales tienen un problema de asimetría, por lo que se aconseja transformarlos a los residuales de Anscombe (1963).

$$r_{i,s}^P = \frac{y - \hat{\mu}}{\sqrt{\hat{\phi} \text{Var}(\mu)(1-h_{ii})}} \quad \text{o} \quad r_{i,s}^P = \frac{y_i - \hat{\mu}_i}{\sqrt{(1-h_i)\sigma^2}} \quad (9.21)$$

9.3.4 Residual eliminado

Nos permite comparar el valor observado con el valor esperado cuando eliminamos la observación asociada a un residual alto.

$$r_{(i)} = \frac{Y_i - \hat{\mu}_{(i)}}{s_{(i)}\sqrt{1+h_{(i)}}} \quad (9.22)$$

9.3.5 Residual parcial

Resulta de eliminar la influencia de las variables explicativas de la variable de respuesta:

$$e(i) = r_i + b_j x_{ij} \quad (9.23)$$

9.3.6 Residuales de discrepancia

Nos informan de la contribución de cada observación a la discrepancia global. Hay expresiones diferentes para modelos normales, de Poisson, logit y loglineales (Cordeiro 2004).

$$r_i^D = \text{sign}(y_i - \hat{\mu}_i) \sqrt{(y_i - \hat{\mu}_i)^2} \quad (9.24)$$

$$r_i^D = \text{sign}(y_i - \hat{\mu}_i) \sqrt{2y_i \log \frac{y_i}{\hat{\mu}_i} - 2 \frac{y_i}{\hat{\mu}_i}} \quad (9.25)$$

$$r_i^D = \text{sign}(y_i - \hat{\mu}_i) \sqrt{2y_i \log \frac{y_i}{\hat{\mu}_i} + 2(n_i - y_i) \log \left(\frac{n_i - y_i}{n_i - \hat{\mu}_i} \right)} \quad (9.26)$$

9.3.7 Residuales de discrepancia estudentizados

Estos residuales consiguen una mayor aproximación a la distribución normal que los residuales de discrepancia.

$$r_i^{AD} = \frac{\text{sign}(y_i - \hat{\mu}_i) \sqrt{D_i}}{\sqrt{1 - h_{ii}}} \quad (9.27)$$

9.3.8 Residuales de verosimilitud

Este procedimiento nos permite comparar la discrepancia de las observaciones con la discrepancia del modelo en el caso concreto en el que se elimina una observación.

$$r_i^L = \text{sign}(y_i - \hat{\mu}_i) \sqrt{h_{ii} (r_{i,S}^P)^2 + (1 - h_{ii}) (r_i^{AD})^2} \quad (9.28)$$

Sin despreciar las demás pruebas expuestas, es necesario resaltar, como ejemplo, el sentido intuitivo para la representación de modelos de algunas de las pruebas. Nos referimos al concepto de verosimilitud y memoria en detección de señales previamente observadas.

9.3.9 Residuales “jackknife”

Son similares al residual de Pearson estudentizado pero utiliza la estimación $s_{(i)}^2$ cuando se elimina una observación (Efron, 1982, Shao, 1992).

$$t_i = \frac{y_i - \hat{\mu}_i}{s_{(i)} \sqrt{1 - h_{ii}}} \quad (9.29)$$

Se suelen utilizar para la detección de observaciones atípicas.

9.4 – POSIBLES CAUSAS DE FALTA DE VALIDEZ DE UN MODELO AJUSTADO

Hay que recordar que a pesar de que un modelo pueda presentar un ajuste adecuado a los datos, este puede no ser del todo correcto o, dicho de otra manera, aceptable. Ello puede ser debido a la existencia de observaciones influyentes, incumplimiento de supuestos, o una especificación errónea del modelo, entre otras.

De este modo, podemos decir que el ajuste adecuado no es criterio suficiente para decir que el modelo es correcto. En efecto, será la evaluación de ese modelo la que nos permita definir mejor el modelo más adecuado.

Así, un modelo puede ver afectada su validez por (Ato, *et al.*, 2005):

- *Outliers* y valores influyentes.
- Especificación errónea del componente sistemático.
- Variables explicativas omitidas.
- Observaciones correlacionadas.
- Supuestos incorrectos.
- Homogeneidad de varianzas (solo en los modelos normales).
- Sobre-dispersión (en modelos logit, log-lineales y Poisson).

9.5. LA SOBREDISPERSIÓN

En este contexto del modelado estadístico y sobre todo en los MLG se produce un fenómeno llamado sobredispersión que aparece cuando tenemos grupos de sujetos con valores homogéneos en las variables explicativas dentro de la población, y también puede ocurrir que exista dependencia entre las observaciones, o que se haya omitido alguna variable importante del modelo. Todo esto produce una incorrecta estimación de los errores estándar de los parámetros del predictor lineal (Dey 1997).

Lo que se suele hacer es evaluar la relación entre la discrepancia y los grados de libertad:

$$s = \frac{D}{gl} \quad (9.30)$$

Su interpretación es la siguiente:

- Si $s = 1$ no existe sobredispersión, el modelo es correcto.
- Si $s > 1$ indica sobredispersión.

- Si $s < 1$ indica infradispersión

9.6. EVALUACIÓN DE LA FIABILIDAD DE UN MODELO

Podemos utilizar dos estrategias diferentes para evaluar la eficacia predictiva del modelo ajustado (Zheng y Agresti, 2000):

- 1) Los índices de eficacia predictiva (solo para variables de respuesta categóricas).
 - Tasas de falsos positivos y negativos.
 - Índices de sensibilidad y especificidad.
 - Índices basados en la proporcionalidad del error.
 - R^2 ajustada para recuentos propuesta por Long (1997).
- 2) Los *split sample* o técnicas de división en mitades (validos para todo tipo de variables de respuesta). Estas técnicas dividen de forma aleatoria la muestra en dos grupos para después hacer dos análisis diferentes:
 - Ajustar y seleccionar un modelo de cada grupo y comparar sus resultados.
 - Utilizar la estrategia de “validación cruzada” que consiste en seleccionar el mejor modelo del primer grupo, y utilizarlo para predecir la respuesta del segundo grupo.

CAPÍTULO 10

MEMORIA DE RECONOCIMIENTO

10.1 INTRODUCCIÓN

El estudio de la memoria de comienzo en 1885 con Ebbinghaus, que realizó una serie de pruebas para evaluar su propia memoria utilizando pseudopalabras que memorizaba, para posteriormente comprobar su grado de recuerdo a través de diversos periodos de tiempo. De esta manera, descubrió que la pérdida de memoria tenía lugar rápidamente a lo largo de las primeras horas o días, pero este declive se hacía más gradual durante los siguientes días, semanas y meses.

Estudiar el modelado estadístico en memoria nos retrotrae a su origen y a la gran profusión de estudios que se realizaron en este sentido en la década de los años cincuenta y sesenta. Por ello entendemos que la memoria (especialmente la memoria de reconocimiento) tal y como la presentamos a continuación puede representar un contenido sustantivo especialmente privilegiado para incluir las matemáticas en general y el modelado en particular. El interés de algunos autores en este sentido no ha dejado de crecer nunca como veremos a continuación.

En este capítulo, trataré de definir la memoria de reconocimiento y los diversos modelos que pretenden explicarla, así como describir los hallazgos científicos que sostienen la existencia de dos procesos involucrados en ese tipo de memoria (el recuerdo y la familiaridad) y los métodos utilizados a lo largo de las últimas décadas para la medición y estimación de esos procesos.

Los resultados de los estudios de velocidad de procesamiento nos muestran que la familiaridad es más rápida que el recuerdo, lo que sugiere que se trata de dos procesos distintos. Del mismo modo, los estudios de curvas ROC (Receiver Operating

Characteristic) indican que por lo menos dos componentes son necesarios para la memoria de reconocimiento, mientras que los estudios de potenciales evocados por ejemplo muestran la memoria como recuerdo o como familiaridad, presenta distintos correlatos electrofisiológicos. En los estudios con pacientes amnésicos muestran un déficit del recuerdo mucho más severo que de la familiaridad.

El aprendizaje y la memoria permiten la acumulación y evocación de diferentes tipos de informaciones que caracterizan a cada ser humano. Squire (1987), definió el aprendizaje como el proceso por el cual se adquieren nuevas informaciones, y la memoria como la persistencia del aprendizaje en un estado al que se pueda acceder posteriormente. La capacidad para identificar un estímulo, objeto, persona o situación como algo ya vivido o visto anteriormente se debe a la memoria de reconocimiento.

Sin embargo, hasta mediados del siglo XX muchos psicólogos dudaban que la memoria fuese una función diferenciada e independiente, sobre todo, debido al hecho de que el almacenamiento de la memoria afecta a muchas regiones diferentes del cerebro. Se ha confirmado que extensas lesiones del lóbulo temporal medial (áreas del hipocampo, la corteza entorrinal, el subículo y la corteza parahipocámpica) y regiones diencefálicas (cuerpos mamilares y ciertos núcleos talámicos) producen ese tipo de déficit cognitivo (Sandi, Venero y Cordero, 2001).

Estos estudios junto a otros estudios cognitivos y de neuroimagen, han demostrado que la memoria de reconocimiento refleja dos procesos distintos de memoria, el recuerdo y la familiaridad. Esa distinción está apoyada por cuatro hallazgos empíricos, que indican que los test de reconocimiento centrados en la familiaridad son

funcionalmente distintos de aquellos basados en el recuerdo y que cada uno de los procesos depende de substratos cerebrales parcialmente separados (Yonelinas, 2002). Además, los test basados en la familiaridad se refieren a test de reconocimiento de ítems, abarcando tareas como la discriminación entre ítems estudiados previamente e ítems que no hayan sido estudiados, mientras que los test basados en el recuerdo se suelen referir a test de reconocimiento asociativo, en los cuales los sujetos deben ser capaces de determinar cuándo o dónde fue estudiado un ítem previamente.

Una línea de investigación que sostiene la existencia de ambos procesos es la que estudia la velocidad de procesamiento, cuyos datos indican que la familiaridad es más rápida que el recuerdo. Según estos investigadores, los sujetos son capaces de ejecutar con precisión y de forma más rápida las tareas que involucran la familiaridad que las tareas que requieren el recuerdo de informaciones específicas (Hintzman, Caulton y Levitin, 1998).

Además, los estudios sobre la confianza en las respuestas de reconocimiento indican la existencia de estos dos procesos, ya que los resultados muestran que el recuerdo y la familiaridad producen distintas curvas ROC. Cuando se introduce en una función de confianza de respuesta de los sujetos tanto la tasa de éxito como la de falsa alarma, las funciones empíricas observadas son curvilíneas, y por eso requieren por lo menos dos parámetros de memoria funcionalmente independientes, lo que indica que por lo menos dos componentes son necesarios para la memoria de reconocimiento.

Hay que añadir que en los test de reconocimiento de ítems, las funciones ROC serán curvilíneas, mientras que en los test de reconocimiento asociativo las curvas ROC

presentan una forma más lineal (Kelly y Wixted, 2001; Rotello, Macmillan y Van Tassel, 2000). Por lo tanto, los modelos que proporcionan un buen ajuste de la curva ROC del reconocimiento de ítems tienden a propiciar un pobre ajuste de la curva ROC del reconocimiento asociativo, lo que indica que diferentes procesos contribuyen en la ejecución de esos dos tipos de test.

El tercer hallazgo se refiere al hecho de que el recuerdo y la familiaridad presentan distintos correlatos electrofisiológicos. Los estudios en los que se utilizan los potenciales evocados de eventos (ERP) durante tareas de reconocimiento, indican que sus distribuciones varían según el proceso requerido por la tarea, es decir, es diferente cuando el sujeto se acuerda de un ítem o lo asocia a algún detalle del evento estudiado (proceso de recuerdo), o cuando el sujeto reconoce al ítem basándose en el proceso de familiaridad (Curran, 2000; Klimesch, *et al.*, 2001).

Finalmente, el cuarto hallazgo que apoya la existencia de ambos procesos es el hecho de que determinados casos de daño cerebral producen un déficit mucho más grave en el recuerdo que en la familiaridad. Por ejemplo, pacientes amnésicos obtienen resultados significativamente inferiores en tareas de reconocimiento asociativo (recuerdo) si los comparamos con las tareas de reconocimiento de ítems (familiaridad), lo que sugiere que las regiones dañadas en pacientes amnésicos son más importantes para el recuerdo que para la familiaridad (Aggleton, *et al.*, 2000; Huppert y Piercy, 1978; Hurst y Volpe, 1982).

Han sido propuestos muchos modelos que asumen la existencia del doble procesamiento, pero que difieren en puntos importantes. Esos modelos son descritos a continuación.

10.2 MODELOS TEÓRICOS DE DOBLE PROCESAMIENTO DE LA MEMORIA DE RECONOCIMIENTO

10.2.1 EL MODELO DE ATKINSON

Atkinson y colaboradores (Atkinson y Juola, 1974; Atkinson, Hertmann, y Wescourt, 1974) proponen un modelo de búsqueda condicional en el cual los sujetos contestan rápidamente al ítem, basándose en la familiaridad, a no ser que ese proceso produzca una respuesta ambigua, produciendo esto una extensa búsqueda en memoria. Este modelo considera la familiaridad como la activación de nódulos en un almacén léxico, donde cada nódulo representa una palabra u objeto individual. Cuando se accede a un nódulo, se produce su activación, que disminuye gradualmente con el paso del tiempo. Por lo tanto, en un test de memoria de reconocimiento, los nódulos correspondientes a los ítems estudiados están, de media, más activos que aquellos correspondientes a los ítems no estudiados, y por esto, medir la activación de los nódulos produce una discriminación entre esas dos clases de ítems.

10.2.2 TEORÍA DE DETECCIÓN DE SEÑALES (TDS)

El proceso de familiaridad es explicado por la TDS en el sentido de que los niveles de activación de ítems nuevos y viejos se solapan cuando forman una

distribución Gaussiana. Según esa teoría, los sujetos establecen un criterio y consideran los ítems que lo excedan como ya estudiados. Además, el sujeto establece un criterio adicional inferior y clasifica a los ítems que estén por debajo del criterio como nuevos (no estudiados). Sin embargo, la medición de la activación de los ítems localizados entre ambos criterios produce errores, y para esos ítems el sujeto utiliza el proceso de recuerdo, con el cual busca la información necesaria en otro almacén que contiene los ítems estudiados y codificados anteriormente. El modelo sostiene también que la familiaridad apoya la memoria para informaciones perceptivas, mientras que el recuerdo apoya las informaciones semánticas. En este capítulo extenderemos esta teoría ya que forma parte de la fundamentación teórica de este trabajo.

10.2.3 MODELOS NEUROANATÓMICOS

Muchos modelos de doble procesamiento fueron desarrollados para explicar el déficit de memoria observado en pacientes con lesión cerebral focal. Los modelos proponen que el daño en el lóbulo temporal medial (hipocampo y alrededores) produce un deterioro del proceso de recuerdo, pero no afecta a la familiaridad. Por ejemplo, Mayes (1988) argumenta que los pacientes amnésicos presentan un déficit de la memoria para el contexto y mantienen intacta la memoria para el ítem. Otros autores presentan argumentos similares (Verfaellie y Treadwell, 1993). Del mismo modo, Huppert y Piercy (1978) argumentan que sujetos sanos basan sus decisiones relacionadas con la memoria de reconocimiento en las evaluaciones de la familiaridad de los ítems o en la recuperación de la información contextual (o de una lista) de los eventos estudiados, mientras que los pacientes amnésicos están limitados a las evaluaciones de la familiaridad.

Sin embargo, modelos más recientes rechazan la idea de que haya un déficit selectivo del recuerdo en pacientes que presentan lesiones en el lóbulo temporal medial y, en cambio, sugieren que la región del hipocampo es esencial para el recuerdo y las estructuras periféricas son importantes para la familiaridad (Aggleton y Brown, 1999). Por eso, pacientes que presentan daño específico en el hipocampo manifiestan déficit selectivo del recuerdo, mientras que pacientes con daños más extensivos y que afectan el giro hipocampal manifiestan deterioro tanto del recuerdo como de la familiaridad.

Aunque la mayoría de esos modelos se centra en el lóbulo temporal medial, algunos modelos neuroanatómicos indican la importancia del tálamo y del lóbulo frontal en estos dos procesos. Por ejemplo, Aggleton y Brown (1999) argumentan que un circuito que conecta el hipocampo y el tálamo anterior por medio del fórnix soporta el proceso de recuerdo, mientras que un circuito que conecta la corteza perirrinal al tálamo medial dorsal soporta el proceso de familiaridad. Además, argumentan que las proyecciones que salen del tálamo hacia el lóbulo frontal indican que éste participa en ambos procesos y que puede ser importante para las operaciones tanto de codificación como de recuperación.

10.2.4 EL MODELO DE MANDLER

Mandler y colaboradores (Mandler, 1980, 1991; Graf y Mandler, 1984) argumentan que los procesos involucrados en la memoria de reconocimiento pueden basarse tanto en la evaluación de la familiaridad del ítem como en el proceso de búsqueda en la memoria. Cuando un evento es estudiado, hay un incremento en la activación o familiaridad, que es la integración intraítem de los aspectos perceptivos del

ítem estudiado. Mientras la familiaridad apoya los procesos implicados en la memoria de reconocimiento, el recuerdo refleja el proceso de búsqueda que apoya la ejecución del reconocimiento y recuperación, con los cuales se recuperan informaciones interítems (informaciones que relacionen el evento a su contexto o a otros eventos). Los autores consideran que los dos procesos son independientes y operan paralelamente, pero la familiaridad es más rápida que el recuerdo. También asocian el bajo rendimiento en tareas de recuerdo a lesiones del lóbulo temporal medial, y asumen que la familiaridad refleja la activación de las representaciones de memorias existentes en el neocortex.

10.2.5 EL MODELO DE JACOBY

Jacoby y colaboradores (Jacoby, 1991; Jacoby y Kelley, 1992; Jacoby, Kelley y Dywan, 1989) proponen que los procesos implicados en la memoria de reconocimiento pueden apoyarse tanto en la evaluación de la fluidez de procesamiento (familiaridad) como en la recuperación del contexto de un ítem estudiado (recuerdo).

Una diferencia fundamental entre los dos procesos es que el recuerdo es un proceso analítico y conscientemente controlado, mientras que la familiaridad es un proceso relativamente automático.

A diferencia de otros modelos, la familiaridad no es considerada una característica inherente al ítem o evento, sino que surge cuando el procesamiento fluente de un ítem es atribuido a una experiencia pasada con el mismo ítem. Por ejemplo, en un test de memoria de reconocimiento, si algunos ítems son procesados con mayor fluencia

que otros, los sujetos lo atribuyen a experiencias pasadas con ellos y los consideran familiares.

Este modelo también propone que la familiaridad y el recuerdo son procesos independientes y paralelos, de modo, que la familiaridad es más rápida que el recuerdo. Además, la familiaridad no refleja la intervención de un sistema de memoria distinto al del recuerdo, más bien, ambos procesos dependen de una memoria detallada de episodios anteriores, y por eso soportan nuevos aprendizajes, tales como pseudo palabras y nuevas asociaciones.

10.2.6 EL MODELO DE TULVING

Tulving y colaboradores (Tulving y Markowitsch, 1998; Nyberg, Cabeza y Tulving, 1996; Wheeler, Stuss, y Tulving, 1997) argumentaron que existen diversos sistemas de memoria funcionalmente distintos, como por ejemplo, la memoria episódica que origina la experiencia consciente de acordarse (recuerdo) y, la memoria semántica, que origina la experiencia consciente de “saber” (el sentimiento de familiaridad en la ausencia del recuerdo). Aunque el modelo de Tulving no es considerado como de doble procesamiento, se asemeja al modelo de Atkinson al asumir que la ejecución de tareas de memoria de reconocimiento refleja la intervención de dos sistemas diferentes de memoria. El sistema episódico almacena experiencias personales y sus relaciones temporales, mientras que el sistema semántico almacena los conocimientos generales acerca del mundo. Sin embargo, las tareas de reconocimiento dependen de ambos sistemas. También se supone que los dos sistemas son flexibles y apoyan el aprendizaje

de nuevas informaciones, aunque el sistema semántico es un poco más lento para ese tipo de aprendizaje.

Este modelo sostiene que la información pasa por el sistema semántico antes de ser codificada en el sistema episódico. No obstante, durante la recuperación de la información los dos sistemas operan paralela e independientemente, en el sentido de que la información puede ser recuperada a partir de uno de los dos sistemas (Tulving y Markowitsch, 1998).

Estos autores argumentan que el hipocampo es esencial para la memoria episódica. En contraposición, las zonas cercanas a esta región son esenciales para la memoria semántica (Tulving y Markowitsch, 1998). Además, las regiones prefrontales derechas son responsables por la recuperación de la información perteneciente a la memoria episódica, mientras que las regiones frontales izquierdas son importantes para la recuperación de información de la memoria semántica y la codificación de la memoria episódica (Nyberg, Cabeza y Tulving, 1996).

10.2.7 EL MODELO DE YONELINAS

Este modelo (Yonelinas 1999, 2001a, 2001b; Yonelinas, Kroll, Dobbins, y Soltani, 1999) propone que el recuerdo y la familiaridad se diferencian en dos aspectos: el tipo de información que producen y cómo influyen en la confianza del reconocimiento. La familiaridad refleja la evaluación de la información acerca de la fuerza de la memoria “cuantitativa”. A su vez, el recuerdo se refiere a un proceso de recuperación de información que tiene como base un umbral, por el cual se recompone

la información “cualitativa” de un determinado evento previo. Los sujetos pueden recuperar muchos aspectos de un evento estudiado (por ejemplo, el contexto temporal y espacial o las asociaciones entre distintos componentes del ítem), pero cuando no son capaces de acordarse de las informaciones cualitativas precisas de algunos ítems (los que están bajo del umbral), utilizan la evaluación por familiaridad. Así que el recuerdo propicia decisiones de reconocimiento relativamente más fiables que la familiaridad, ya que soporta un amplio rango de respuestas de reconocimiento fiables. En el momento de la recuperación de la información el modelo propone que ambos procesos se inician paralela e independientemente, pero la familiaridad suele ser más rápida que el recuerdo (Yonelinas y Jacoby, 1994, 1996). Respecto a las características neuroanatómicas del modelo, las lesiones en el hipocampo están relacionadas al bajo rendimiento en las tareas de recuerdo. Por el contrario, el rendimiento en tareas que requieren el proceso de familiaridad es inferior cuando estructuras adicionales del lóbulo temporal están lesionadas (Yonelinas, *et al.*, 1998).

Según este modelo, tanto el recuerdo como la familiaridad sostienen el aprendizaje de nuevas asociaciones, pero la familiaridad sólo bajo condiciones muy limitadas (Yonelinas, 1999; Yonelinas, Kroll, Dobbins y Soltani, 1999). Pero no explica la memoria asociativa de dos ítems distintos, a no ser que puedan ser considerados un único ítem más grande (por ejemplo, la boca, los ojos y la nariz forman un rostro).

10.3 TÉCNICAS DE MEDIDA

Para evaluar los argumentos teóricos de los distintos modelos descritos anteriormente se han desarrollado algunas técnicas capaces de medir los procesos

involucrados y todos ellos tienen en cuenta los supuestos básicos de cada modelo. Dichas técnicas son utilizadas para examinar el recuerdo y la familiaridad y pertenecen a dos clases generales, denominadas técnicas de disociación de la tarea y de estimación del proceso.

Las primeras tienen como objetivo identificar las condiciones de las tareas y los test que aíslan uno de los dos procesos. Si el rendimiento en una tarea propuesta por esa técnica se disocia del rendimiento de una tarea de un test estándar de reconocimiento, en el cual ambos procesos son requeridos, entonces se puede inferir sobre los efectos de diferentes variables en tales procesos. Sin embargo, algunos patrones de resultados producidos por esa técnica son ambiguos, por lo cual se ha desarrollado otra clase de técnicas.

10.3.1 TÉCNICAS DE DISOCIACIÓN DE LA TAREA Y TÉCNICAS DE RESPUESTA-VELOCIDAD

Como algunos modelos suponen que la familiaridad es más rápida que el recuerdo (Atkinson y Juola, 1974; Mandler, 1991; Jacoby, 1991; Yonelinas; 2001b), la velocidad de respuesta es útil para separar la contribución de los dos procesos. Son usadas tres técnicas distintas basadas en esto. En la técnica respuesta-tiempo el rendimiento en tareas de reconocimiento analiza separadamente las respuestas de reconocimiento rápidas y lentas, bajo condiciones estándares de reconocimiento que no involucran la velocidad. Las respuestas rápidas reflejan familiaridad, mientras que las lentas reflejan la contribución del recuerdo.

En la técnica, respuesta-límite, los sujetos deben contestar dentro del límite de tiempo determinado tras la presentación del estímulo. El rendimiento en esta tarea es comparado con el de otra tarea de reconocimiento que no establece tiempo de respuesta.

Las respuestas con tiempo determinado se basan principalmente en la familiaridad y las de tiempo libre reflejan ambos procesos. La técnica de intercambio velocidad-precisión, es una variación del anterior. Los sujetos deben contestar según intervalos de tiempo variados, tras el comienzo de cada ítem del test. Esto permite establecer una función de tiempo que mide el incremento de la precisión de la recuperación. El punto en el que el rendimiento deja de ser casual puede ser usado para inferir la velocidad de recuperación en diferentes condiciones del test (Yonelinas, 2002).

Una limitación potencial de la técnica de respuesta-límite se refiere a la variación de las instrucciones dadas a los sujetos en las dos condiciones del test. En una de ellas se les pide que contesten al ítem, teniendo en cuenta el tiempo, mientras que en la otra no se hace referencia al tiempo. Las diferencias en las instrucciones pueden influenciar en la manifestación de los procesos, lo que complica la comparación entre distintas condiciones. Sin embargo, en las otras dos técnicas, las instrucciones se mantienen constantes

10.3.2 TÉCNICAS DE RECUPERACIÓN/RECONOCIMIENTO

Si el proceso de recuerdo que está implicado en el reconocimiento es similar al proceso de búsqueda usado en los test de recuperación de información (Mandler, 1991;

Tulving y Markowitsch, 1998), entonces el rendimiento en los test de recuperación de información puede ser usado como indicador del recuerdo. Al comparar el rendimiento en tareas de recuperación de información con el rendimiento en tareas de reconocimiento (que refleja recuerdo y familiaridad) es posible inferir acerca del efecto de diferentes variables en los procesos de recuerdo y familiaridad. Por ejemplo, si una variable produce un efecto más grande en la recuperación de información que en el reconocimiento, se supone que se debe al hecho de que la variable produce un efecto más grande en el recuerdo que en la familiaridad (Yonelinas, 2002). Una limitación potencial asociada a esa técnica se debe a que las dos condiciones del test difieren en la pista para la recuperación y en el tipo de respuestas requeridas, lo que puede generar medidas sesgadas en los dos procesos. Otra limitación se refiere a las escalas, ya que las escalas del reconocimiento y de la recuperación puede que no sean lineales dentro del rango de las puntuaciones observadas (Loftus, 1978), es decir, las puntuaciones de cada escala no son equivalentes.

10.3.3 TÉCNICAS DE RECONOCIMIENTO ÍTEM/ASOCIATIVO

Esta técnica asume que el recuerdo refleja la recuperación de información cualitativa del evento estudiado (Atkinson y Juola, 1974; Jacoby, 1991; Yonelinas, 2001b; Mandler, 1991). Así que el recuerdo es útil en los test de memoria de reconocimiento asociativo, en los cuales los sujetos deben tomar decisiones acerca de la ocurrencia concomitante de diferentes ítems o de distintos aspectos del evento estudiado, como asociaciones ítem-ítem (por ejemplo, si dos palabras estaban emparejadas en el estudio), asociaciones ítem-contexto (si el ítem fue presentado al lado

derecho o izquierdo) o asociaciones ítem-características (si determinada palabra fue producida por una voz masculina o femenina).

Respecto a la familiaridad, la técnica asume que refleja la fuerza de la memoria para ítems simples y por eso debe discriminar los ítems estudiados de los que no fueron estudiados previamente, como en un test de reconocimiento de ítems. Sin embargo, la familiaridad no tiene la misma utilidad en los test de reconocimiento asociativo, por lo tanto, pueden ser usados como indicadores del recuerdo, mientras que los test de reconocimiento de ítems requieren respuestas tanto de recuerdo como de familiaridad.

La ventaja es que mantiene constantes el estímulo de recuerdo y el tipo de respuesta requerida en las diferentes condiciones de reconocimiento de ítem y reconocimiento asociativo, aunque algunos test asociativos pueden tener cierto grado de influencia de la familiaridad, como por ejemplo cuando los sujetos deben decidir sobre el orden temporal, pues ítems más recientes pueden ser percibidos como más familiares.

Además, algunos modelos proponen que la familiaridad puede soportar el aprendizaje de algunos tipos de asociación y por eso bajo algunas condiciones puede contribuir, aunque en menor grado en la memoria de reconocimiento.

10.3.4 TÉCNICAS DE ESTIMACIÓN DEL PROCESO

El procedimiento de estimación del proceso se basa en la modificación de tareas de reconocimiento asociativo y de ítems. El recuerdo es medido como la capacidad para recordar dónde o cuándo un ítem fue estudiado previamente. Si un sujeto es capaz de

recordar un ítem, debería ser capaz de determinar cuándo o dónde fue inicialmente estudiado, mientras que la familiaridad no soporta dicha discriminación.

Esta técnica fue inicialmente desarrollada por Jacoby (1991) y consiste en que los sujetos inicialmente estudian una lista de ítems presentada visualmente, bajo condiciones de codificación accidental. Luego, estudian otra lista de ítems presentados oralmente, bajo condiciones de codificación intencional. En una de las condiciones del test (test de inclusión) los sujetos deben contestar “sí”, si el ítem fue estudiado anteriormente (incluye ítems de las dos listas estudiadas). En la otra condición del test (test de exclusión), los sujetos tienen que contestar “sí” solamente para los ítems de la lista oral (excluye los ítems de la lista visual).

Como ambos procesos pueden ser usados en la situación de inclusión, la probabilidad de aceptar correctamente que un ítem pertenezca a la lista visual es equivalente a la probabilidad de que el ítem sea recordado, más la probabilidad de que el ítem no sea recordado, sino aceptado basándose en el proceso de familiaridad [$P(\text{inclusión}) = R + (1 - R)F$]. En cambio, la probabilidad de que un ítem leído sea aceptado incorrectamente en la situación de exclusión es igual que la probabilidad de que actúe la familiaridad en la ausencia del recuerdo [$P(\text{exclusión}) = (1 - R)F$], ya que la familiaridad es capaz de distinguir entre ítems estudiados y no estudiados anteriormente, pero no de determinar a qué lista pertenece el ítem. Los parámetros R y F pueden ser estimados al contrastar el rendimiento en las condiciones de inclusión y exclusión.

Esta técnica coincide con la mayoría de modelos en el argumento de que el recuerdo y la familiaridad son independientes. La principal técnica para medir la independencia de los dos procesos es buscar la disociación entre ellos. Si de hecho son completamente independientes, es posible encontrar alteraciones que influyan en el recuerdo, pero no en la familiaridad y otras que produzcan el efecto opuesto. En cambio, si los dos procesos no son independientes, dichas disociaciones no son posibles.

Se han llevado a cabo estudios con el objetivo de comprobar la relación entre esos procesos, y en general, indicaron que estos actúan de manera independiente el uno del otro (Caldwell y Masson, 2001; Yonelinas, 2001a, 2001b). Sin embargo, Yonelinas (2002) señala que los dos procesos operan de manera independiente en la mayoría de las condiciones experimentales. Pese a eso, bajo otras condiciones la suposición de la independencia no se sostiene. Así que, si existe o no esta independencia, es algo que todavía está estudiándose.

Una limitación potencial es el uso de una medida muy estricta del recuerdo, la habilidad para determinar a qué lista de las estudiadas pertenece el ítem. La medida del recuerdo es la capacidad del sujeto para rescatar alguna de las características del evento estudiado (codificación incidental o intencional, a cuál lista pertenece el ítem y la modalidad auditiva o visual) y usar la información como base para excluir el ítem. Sin embargo, cuando el sujeto se acuerda de otros aspectos del evento estudiado que no apoyan la discriminación requerida, estos no son considerados medidas de recuerdo.

Otra limitación se refiere a la diferencia de las instrucciones dadas en las condiciones de exclusión e inclusión. El modelo considera que la probabilidad de que el

recuerdo actúe en ambas condiciones es la misma, pero, de hecho, los sujetos usan con más frecuencia el recuerdo en las condiciones de exclusión. Del mismo modo, la técnica también asume que la contribución de la familiaridad es constante en las dos condiciones. Una alternativa para garantizar el uso similar del recuerdo y de la familiaridad en las dos condiciones es mezclar los test de inclusión y exclusión (Jacoby, Toth y Yonelinas, 1993).

10.3.5 EL PROCESO RECORDAR/CONOCER

Tulving (1985) desarrolló un modelo con el objetivo de medir la contribución de diferentes tipos de procesos en el rendimiento total de la memoria. Los sujetos deben reflexionar sobre la base de sus decisiones y relatar si reconocen a los ítems apoyándose en el proceso de “recordar” (recuerdo de informaciones episódicas del evento estudiado) o en el proceso de “conocer” (el ítem es familiar en la ausencia del recuerdo). Como las respuestas de “recordar” y “conocer” se excluyen mutuamente (los sujetos pueden contestar solamente una de las dos posibilidades para cada ítem), el método asume que los procesos subyacentes son también mutuamente excluyentes, lo que es consistente con todos los modelos de doble procesamiento.

Pese a esto, Tulving (1985) no especificó exactamente cómo se relacionan las respuestas con los sistemas subyacentes de la memoria. Ese hecho condujo a una variabilidad en el relato y en la interpretación de los resultados del método porque la medida de la familiaridad no es tan sencilla como propone. En estudios anteriores, las respuestas “conocer” eran usadas como la estimación de la familiaridad (Gardiner, 1988), pero los resultados han demostrado que las respuestas producen una medida

sesgada de la familiaridad (Gardiner y Richardson-Klavehn, 2000; Wagner, Gabrieli y Verfaellie, 1997), debido a que los sujetos no son instruidos a contestar “conocer” cuando un ítem es familiar, sino que son instruidos a dar esa respuesta siempre que un ítem sea familiar y no recordado. Por ello, la proporción de esa respuesta tiende a subestimar la probabilidad de que un ítem sea familiar.

La medición de la familiaridad y del recuerdo, basada en relatos subjetivos, es lo que hace único este procedimiento. Por una parte, una ventaja es que los sujetos pueden añadir cualquier información acerca del evento estudiado. Por otra parte, confiar en relatos subjetivos puede ser problemático si los sujetos no tienen acceso directo a los procesos que apoyan el recuerdo o si sus relatos no son precisos. Sin embargo, los estudios indican que las estimaciones del recuerdo derivadas de las respuestas recordar/conocer son coincidentes con las derivadas de otras medidas de recuerdo que no dependen de relatos subjetivos, lo que sugiere que los sujetos, generalmente, tienen acceso a esos procesos de la memoria (Yonelinas, 2001b).

Una crítica común a estos aspectos de la memoria se refiere a que los relatos de recordar y conocer no reflejan las contribuciones del recuerdo y la familiaridad, sino que reflejan estados subjetivos de conciencia que son completamente independientes a los procesos de recuerdo y familiaridad (Gardiner y Richardson-Klavehn, 2000). Otra crítica es que conocer y recordar puede que no reflejen dos formas distintas de memoria, sino más bien diferencias en la familiaridad o confianza (Hirshman y Master, 1997; Inoue y Bellezza, 1998).

10.3.6 EL PROCEDIMIENTO RECEIVER OPERATING CHARACTERISTIC

La principal característica de esta propuesta es que posibilita estimar la contribución de la familiaridad y el recuerdo examinando el efecto de la variación del criterio de respuesta de las tasas de aciertos y falsas alarmas (Yonelinas, 1994). El modelo deriva una ecuación que describe como esas tasas están relacionadas, es decir, si el rendimiento refleja una combinación del recuerdo y la de familiaridad. Después, la ecuación es ajustada a los datos observados, y las tasas de acierto y falsa alarma son introducidas en una función confianza de los sujetos respecto a sus respuestas sí/no.

Una ventaja es el hecho de requerir solo una condición de test de reconocimiento, por ello, el recuerdo y la familiaridad pueden ser analizados a partir de una única tarea en vez de confiar en la comparación de dos tareas que involucran diferentes estímulos de recuperación o tipos de respuestas. Sin embargo, el modelo asume otros supuestos, como por ejemplo que la familiaridad refleja un proceso de detección de señal, con el cual viejos ítems son más familiares que los nuevos y todos los ítems que excedan el criterio de respuesta son considerados como ya estudiados. La distribución de la familiaridad es gaussiana y las distribuciones nuevas y viejas poseen varianzas iguales. Además, cuando actúa el recuerdo, la respuesta de reconocimiento es relativamente fiable, mientras que las respuestas de familiaridad tienen un rango más amplio de respuestas fiables. Los estudios confirman estos supuestos así como las comparaciones directas entre las estimaciones de los parámetros del modelo y aquellas basadas en los métodos de proceso-disociación y recordarse/saber (Yonelinas, 2001a, 2001b).

10.4 CONSIDERACIONES SOBRE LAS TEORÍAS

Los modelos descritos coinciden en algunos argumentos fundamentales. Por ejemplo, hay un acuerdo general de que la familiaridad es más rápida que el recuerdo, y la mayoría de los modelos asumen que ambos procesos se inician paralelamente, salvo el modelo de Atkinson (Atkinson Y Juola, 1974; Atkinson, Hertmann, Y Wescourt, 1974) que propone que la familiaridad se completa antes que el recuerdo.

Otro punto fundamental de estos modelos es la relación entre los dos procesos. Los modelos de Mandler (1980, 1991), Jacoby (1991), Tulving (Tulving Y Markowitsch, 1998) y Yonelinas (2001b) concuerdan que los dos procesos funcionan independientemente en el momento de recuperar la información. Pocos modelos neuroanatómicos explican la relación entre ellos, pero asumen que es posible interrumpir el proceso de recuerdo, y suponen que sea parcialmente independiente de la familiaridad. A su vez, Atkinson (Atkinson y Juola, 1974) propone dos sistemas de almacenamiento distintos, que involucran mecanismos de codificación independientes. En cambio, Tulving (Tulving y Markowitsch, 1998) argumenta que los procesos actúan según una secuencia durante la codificación, aunque puedan operar independientemente durante la recuperación.

Algunos modelos también coinciden en que la familiaridad suele ser descrita como un indicador de la fuerza de la memoria, reflejando el recuerdo la recuperación de informaciones específicas del evento estudiado (Atkinson y Juola, 1974; Jacoby, 1991; Yonelinas, 2001b). Además, se asemejan en la consideración de que el recuerdo es un proceso conceptual y la familiaridad, perceptivo. Sin embargo, para Jacoby (1991) la familiaridad refleja tanto un proceso conceptual, como perceptivo, y para Tulving

(Tulving y Markowitsch, 1998) está relacionada con la memoria semántica y el conocimiento abstracto.

Muchos modelos son consistentes con la idea de que el recuerdo y la familiaridad son descritos como procesos controlados y automáticos, respectivamente (Jacoby, 1991), de hecho, el asumir que la familiaridad es un proceso rápido, puede ser considerado relativamente automático (Yonelinas, 2001b). Además, el proceso de activación que subyace la familiaridad en algunos modelos es relativamente automático, mientras que el proceso de búsqueda requiere control y esfuerzo (Atkinson, Hertmann y Wescourt, 1974; Mandler 1991).

Respecto a los desacuerdos entre los modelos, uno de los más importantes se refiere a cuál de los dos procesos puede soportar el aprendizaje de nueva información. Algunos consideran que el recuerdo es necesario para el aprendizaje de nuevas asociaciones y representaciones. Por otra parte, la familiaridad refleja la activación de representaciones ya existentes (Atkinson, Hertmann, y Wescourt, 1974; Mandler, 1980, Mandler, 1991). A su vez, Jacoby (1991) argumenta que ambos procesos pueden soportar nuevos aprendizajes, pero Yonelinas (2001b) propone que la familiaridad no soporta la memoria asociativa para diferentes aspectos de un evento, a no ser que estos aspectos sean considerados un todo unificado.

Hay también desacuerdos referentes a la correspondencia cerebral de la familiaridad y del recuerdo. Aquí la mayoría de los modelos proponen que el recuerdo depende del lóbulo temporal medial y por eso asumen que pacientes amnésicos presentan déficit en tareas que requieren ese proceso (Jacoby, 1991; Mandler, 1991;

Yonelinas, 2001b). Sin embargo, los modelos no coinciden respecto a la familiaridad en esos pacientes, Mandler (1980, 1991) y otros autores de algunos de los modelos neuroanatómicos proponen que la familiaridad se conserva en la amnesia. Otros modelos proponen que la familiaridad depende de la localización precisa del daño en el lóbulo temporal medial. Por ejemplo, Tulving (Tulving y Markowitsch, 1998), Yonelinas (2001b) y algunos de los modelos neuroanatómicos sugieren que el hipocampo es fundamental para el proceso de recuerdo, mientras que regiones cercanas a esa estructura (como el giro hipocampal) son importantes para la familiaridad.

Estos desacuerdos entre los modelos se deben a las diferencias de enfoque y procedimientos propuestos para la investigación de la memoria de reconocimiento.

A pesar de ello, los modelos de doble procesamiento han sido fundamentales en los estudios de la memoria en los últimos 30 años y han sido de extrema importancia tanto para la elaboración de una base teórica de la memoria de reconocimiento como para la diferenciación entre los procesos involucrados

En el presente trabajo, nos decantamos por la TDS como sustrato teórico sobre el que plantear los experimentos de memoria de reconocimiento que deberán realizar los sujetos, y por ello, a continuación realizaremos una exposición más detallada de esta teoría.

10.5 LA TEORÍA DE DETECCIÓN DE SEÑALES (TDS)

10.5.1 INTRODUCCIÓN

Ya nadie duda de que exista una estrecha relación entre percepción y cognición, cuando estamos hablando de procesamiento de la información. Desde este punto de vista podemos distinguir los procesos relacionados con la sensibilidad o discriminación, y los procesos de toma de decisiones. También podemos controlar la distribución y probabilidad de la señal, así como un registro de todas las respuestas del sujeto (curva ROC).

Por lo que tenemos 4 niveles de análisis:

- Ordenar, clasificar y categorizar la señal.
- Analizar la sensibilidad o discriminación de los sujetos.
- Medir y controlar el criterio de decisión de los sujetos.
- Clasificar los resultados, aplicar pruebas de significación estadística y obtener modelos matemáticos descriptivos o predictivos de los datos.

En sus inicios este modelo se aplicó en experimentos de visión (Tanner y Swets 1953), pero poco tiempo después ya empiezan a realizarse con palabras (Egan 1958). En el primero no resulta complicado obtener las distribuciones a priori de las variables, pero con palabras necesitamos establecer formas de control de las variables de este tipo de señales.

De este modo, el control de una sola variable no nos permite obtener la distribución completa de la señal para este tipo de variables, que pueden clasificarse de

manera diferente. Si tenemos un conjunto de palabras y consideramos una única variable, la distribución obtenida la podemos dividir en dos conjuntos equivalentes y de igual probabilidad a partir de la media. Siendo la probabilidad a priori de la señal $P(S)$ y la del ruido $P(R)$ de modo que $P(S) + P(R) = 1$.

Pero como esta probabilidad a priori depende del experimentador esta probabilidad no tiene porque ser siempre equiprobable.

En la matriz de probabilidades condicionadas tenemos las claves para el cálculo de la probabilidad de la sensación.

La TDS pretende describir formalmente (matemáticamente) la capacidad de un organismo, o de una máquina artificial, para reconocer un estímulo determinado y decidir una respuesta en base a ese reconocimiento. Se ha utilizado profusamente en la psicología (recientemente, por ejemplo, se ha propuesto como paradigma en el estudio de percepción de contingencias: Allan, Siegel y Tangen, (2005)), pero también se ha aplicado en muchos otros campos, como la inteligencia artificial y la medicina, y algunos de sus aspectos todavía están presentes en la ingeniería aplicada (las curvas ROC que se usan en el desarrollo y evaluación de los sistemas de control de calidad). Aunque se trata de una teoría con evidente despliegue matemático es un tema razonablemente accesible incluso para los profanos en la materia.

10.5.2 DETECTANDO LA SEÑAL

Tenemos un sistema cuya misión es detectar un determinado estímulo, que llamaremos "señal", distinguiéndolo de cualquier otro, y la tarea consiste en reconocer una señal determinada en presencia de ruido. Cada vez que sea presentada, el sistema tiene que emitir inmediatamente una respuesta, indicando que ha captado y reconocido la señal (respuesta positiva). En cualquier otro caso, es decir, cuando no ha sido presentada la señal especificada, el sistema ha de emitir una respuesta negativa. Este comportamiento que llevan a cabo las personas y los animales en los experimentos típicos de psicología no es tan diferente del que realizan aparatos tan sencillos como los detectores de movimiento o los termostatos.

Sin embargo, la detección de la señal siempre se lleva a cabo en un entorno con "ruido", es decir, repleto de otros estímulos (auditivos o de cualquier otro tipo), que no son la señal pero en todo caso la acompañan. De forma que, cuando se presenta la señal, realmente el sistema está captando "señal + ruido", y debe distinguir esta estimulación de la situación alternativa "sólo ruido". Esto es lo que nos ocurre a todos cuando escuchamos la televisión o la radio: creemos escuchar sólo la voz del locutor, pero realmente nuestros oídos están captando muchos otros sonidos a la vez, seamos conscientes de ellos o no (el tráfico de la calle, las voces del público, las imperfecciones de la señal de audio de la TV o radio, y las imperfecciones de nuestro propio sistema auditivo).

Evidentemente, el ruido hace más difícil la tarea de detección de la señal, provocando falsas alarmas (el sistema cree detectar la señal cuando ésta no ha sido

presentada) y rechazos incorrectos u omisiones (el sistema pasa por alto la presentación de la señal). En general, con respuestas binarias como ésta (es decir, el sistema sólo responde "Sí" o "No"), sólo pueden darse las siguientes cuatro situaciones:

- a) Acierto (A): La señal estaba ahí (señal + ruido) y el sistema la detecta.
- b) Falsa alarma (FA): El sistema cree detectar la señal, pero realmente sólo había ruido.
- c) Rechazo Incorrecto (RI): La señal estaba ahí (señal + ruido) pero el sistema no la detecta (esto es, la confunde con el ruido).
- d) Rechazo correcto (RC): El sistema no detecta la señal, y efectivamente sólo había ruido.

	Ruido	Ruido + Señal
SI	Falsa Alarma (FA) <i>FALLO</i>	Acierto (A) <i>ACIERTO</i>
NO	Rechazo correcto (RC) <i>ACIERTO</i>	Omisión (O) <i>FALLO</i>

Figura 10.1: Tabla de TDS.

Un sistema sensible y específico minimizará el número de rechazos incorrectos y falsas alarmas, a la vez que mostrará un número grande de aciertos y rechazos correctos.

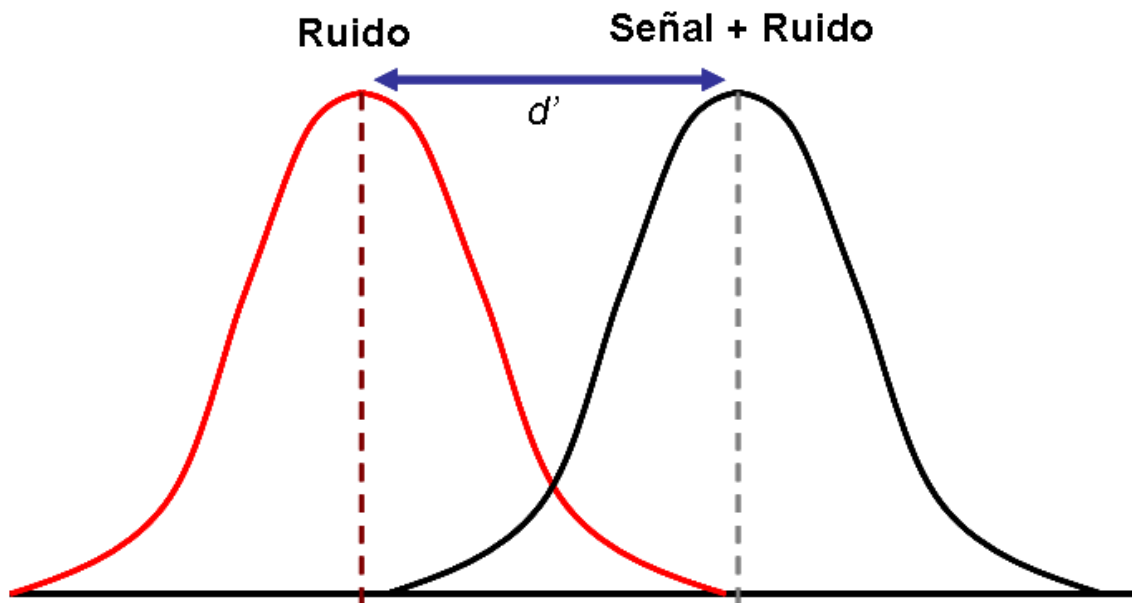


Figura 10.2: Sensibilidad

Si la frecuencia del ruido es aleatoria, sus valores se distribuyen según una curva de frecuencias que en este caso sigue una ley normal. En la imagen, he representado las distribuciones normales del ruido (curva roja) y de la "señal + ruido" (curva negra), suponiendo que en el eje horizontal tenemos los distintos valores de frecuencia sonora (en Hz).

10.5.3 DIFERENCIANDO LA SEÑAL DEL RUIDO: ÍNDICE DE SENSIBILIDAD

Las medias de las frecuencias del ruido y de la "señal + ruido" son diferentes (cada una está en el centro de su distribución correspondiente) y están representadas con dos líneas verticales discontinuas. Cuanto más diferentes sean esas medias, más sencillo será el problema de discriminar entre los dos estímulos, más fácil reconocer cuándo hemos captado "ruido" y cuándo hemos captado "señal + ruido". A la distancia que

separa ambas medias la representamos mediante el parámetro d' , que es el índice de discriminabilidad o de sensibilidad.

La diferencia entre la media de las frecuencias del ruido y la media de las frecuencias de la "señal + ruido" es grande, visualmente identificaremos esta situación porque las curvas "ruido" y "señal + ruido" estarán más separadas. Por lo tanto, no tendremos excesivos problemas en reconocer cuándo se ha presentado la señal. Sin embargo, esto se vuelve más peliagudo si la frecuencia media del ruido es similar a la frecuencia media de la señal. En este último caso las curvas se solaparían en mayor medida que en el ejemplo anterior.

10.5.4 EMITIENDO UNA RESPUESTA: ¿ESTABA LA SEÑAL AHÍ?

El segundo componente importante en la TDS es el criterio de respuesta, identificado con el parámetro β . Este parámetro pretende representar la tendencia del sistema a sesgar su respuesta en una dirección u otra (hacia el "Sí, he captado la señal" o hacia el "No, no la he captado"). En muchos casos, la respuesta no es una función directa del índice de sensibilidad d' (es decir, de la dificultad de diferenciar ambos estímulos). Hay veces en las que estamos casi seguros de que hemos captado la señal, pero por si acaso preferimos no arriesgarnos, así que damos una respuesta negativa. Y a veces sucede justo lo contrario. En el siguiente gráfico se representa el criterio de respuesta como una línea verde vertical. Si el sistema capta un sonido cuya frecuencia está a la derecha de la línea vertical (β es positiva), tenderá a responder positivamente, y lo inverso ocurrirá si la frecuencia captada está a la izquierda del criterio.

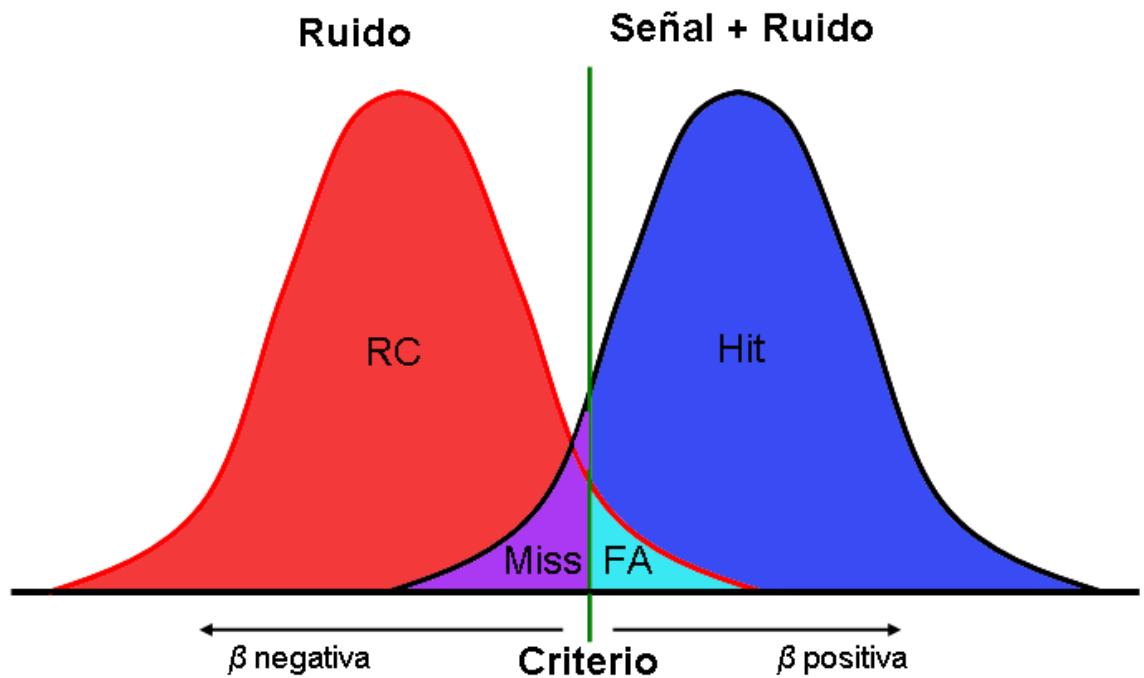


Figura 10.3: Grafico de señales y ruido

La inclusión del criterio de respuesta en esta imagen nos deja ver 4 áreas bajo las curvas (en 4 colores):

- a) Acierto (A): Se ha presentado la señal (curva negra), y el sonido captado por el sistema (señal + ruido) está a la derecha del criterio (la línea vertical), por lo que la respuesta es positiva. El sistema detecta correctamente la señal.
- b) Falsa alarma (FA): Se ha presentado un estímulo que no es la señal (ruido, curva roja), pero está a la derecha del criterio, por lo que la respuesta es positiva, y errónea.
- c) Rechazo Incorrecto (RI): Se ha presentado la señal (curva negra), pero el sonido captado por el sistema (señal + ruido) está a la izquierda del criterio, por lo que la respuesta es negativa. Por lo tanto, el sistema no ha detectado la señal.

d) Rechazo correcto (RC): Se ha presentado un estímulo que no es la señal (ruido, curva roja), y que además está a la izquierda del criterio, por lo que la respuesta es negativa y correcta.

Lo interesante del gráfico es que, podemos calcular las áreas bajo las curvas, de forma que es posible traducirlas a probabilidades. La probabilidad bajo la curva completa es igual a 1.

Esto significa, en la práctica, que cuando conocemos d' y el criterio de respuesta, podemos calcular las probabilidades (representadas como áreas) de que un sistema produzca aciertos, falsas alarmas, rechazos incorrectos y rechazos correctos, y de este modo predecirlos. Tenemos por tanto, dos conceptos clave, la sensibilidad y la especificidad del sistema detector, y en clave probabilística: la sensibilidad es la probabilidad de acierto sobre el total de veces que se presenta la señal (es el cociente entre los aciertos y la suma de los aciertos y falsas alarmas); la especificidad es la probabilidad de cometer una falsa alarma sobre el total de veces que no se presenta la señal (es el cociente entre los rechazos correctos y la suma de los rechazos correctos y los rechazos incorrectos). El sistema perfecto es sensible (detecta la señal siempre o casi siempre que aparece) y específico (es capaz de distinguir la señal del ruido y no da una respuesta positiva cuando no está la señal).

Ya en un plano psicológico, hay ciertas manipulaciones que podemos planear para alterar las probabilidades de cada tipo de situación. Una de las más obvias es el coste de respuesta. Dependiendo del peso que se le dé a los aciertos y los fallos los

sujetos se arriesgaran más o menos a la hora de emitir la respuesta.

10.5.5 CURVAS ROC

Por último, debemos mencionar uno de los instrumentos más útiles en la TDS y en sus aplicaciones a distintos campos: las curvas ROC (Receiver Operating Characteristic). En estas curvas se confronta gráficamente la sensibilidad del sistema con su especificidad. La curva ROC se construye mediante 2 ejes: el eje vertical representa la probabilidad de acierto; el eje horizontal la probabilidad de una falsa alarma.

Un sistema detector de señales que no ha sido entrenado correctamente mostrará un patrón de respuestas aleatorio, que se representa mediante una línea recta de 45°. Es decir, para ese sistema ineficiente es igual de probable producir un acierto que una falsa alarma. No es sensible ni específico. Esta línea representa la aleatoriedad de la respuesta.

Sin embargo, los sistemas suelen tener un rendimiento mucho mejor que éste, y sus respuestas se representan por medio de una curva que nos da una indicación de su eficiencia: si el número de aciertos es alto y el de falsas alarmas reducido (esto es, si el sistema es muy sensible y muy específico), las respuestas van a representarse mediante una curva que va a alejarse de la línea de aleatoriedad, y por el contrario va a acercarse mucho a la esquina superior izquierda del plano. Cuanto más se acerque a este punto la curva, mayor será el área bajo dicha curva, y más eficaz es el sistema en su tarea de

reconocer la señal (muchos aciertos, pocas falsas alarmas). Y podremos cuantificar esa eficacia, que es lo interesante (ver figura 10.4).

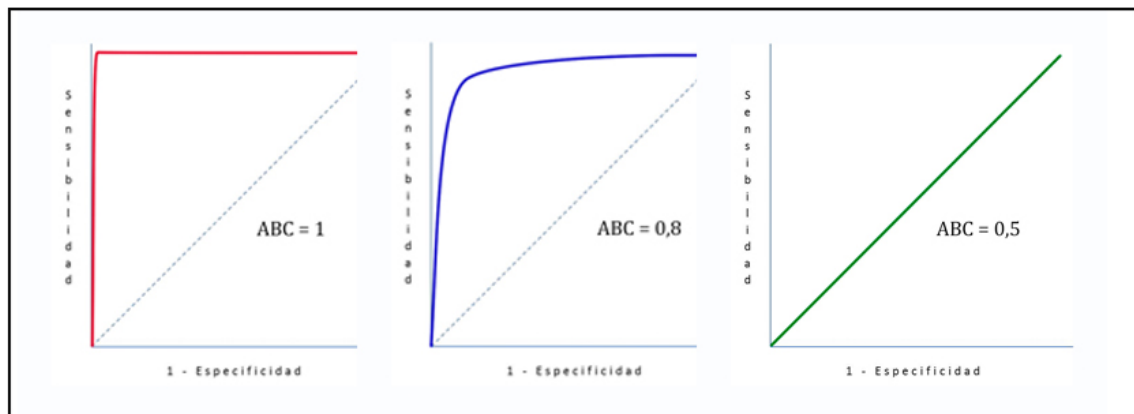


Figura 10.4: Diferentes curvas ROC.

10.5.6 CONCLUYENDO

La TDS nos permite analizar los factores que pueden estar influyendo en las respuestas de los sujetos y que actúan independientemente de la habilidad de los sujetos para detectar o discriminar entre estímulos.

La TDS ofrece un procedimiento de análisis que va a permitir no sólo calcular el nivel de detectabilidad del estímulo sino también el criterio o criterios que utiliza el observador para emitir sus respuestas. Por lo que intervienen variables asociadas a los procesos cognitivos y la personalidad junto con los estímulos y las respuestas.

10.6 CRONOMETRÍA MENTAL

Para comprobar nuestra hipótesis sobre los procesos cognitivos una de las técnicas más empleada es el registro de los tiempos de reacción (TR), definido como el intervalo transcurrido entre la presentación del estímulo y el inicio de la respuesta. Es

importante no confundirlo con el tiempo de ejecución (que sería el tiempo que tarda el sujeto en realizar la tarea). Además, nuestro sujeto debe responder de forma rápida y precisa (sin cometer errores) a la tarea demandada.

El estudio de la “cronometría mental” proviene del fisiólogo holandés Donders (1868), que pretendía establecer la duración de los eventos mentales. Estas técnicas tuvieron un gran auge a finales del XIX y principios del XX, pero con los asociacionistas y conductistas estas técnicas decayeron, volviéndose a recuperar el interés con el ascenso de la psicología cognitiva.

Aunque el estudio de los TR puede tener inconvenientes (inestabilidad, sensibilidad a los estados del sujeto, duración, práctica, etc.), su gran ventaja es su estrecha relación con la complejidad del proceso mental requerido para ejecutar la respuesta. Es importante indicar que el TR no tiene significado en si mismo, sino que el experimentador asuma unos supuestos teóricos y utilice artificios en el diseño de la tarea experimental.

10.7 POTENCIALES EVOCADOS

10.7.1 INTRODUCCIÓN

Cuando un tren de estímulos sensoriales de cualquier tipo llega al cerebro, provoca secuencias características de ondas en el trazado electroencefalográfico (EEG), que denominamos potenciales evocados. Estos son diferentes para cada modalidad sensorial y su variabilidad también depende de la intensidad del estímulo. Además, presentan una relación estable en el tiempo respecto al estímulo.

Coles y Rugg (1995) definen los potenciales evocados como fluctuaciones en el voltaje del electroencefalograma (EEG) provocadas por sucesos sensoriales, motores o cognitivos. Se observa que tras la presentación de un determinado tipo de estímulo o la realización de alguna tarea se producen cambios constantes en el EEG. Estos cambios adoptan la forma de picos o valles y se supone que nos informan acerca de los procesos cognitivos que subyacen.

En referencia al sustrato fisiológico de los potenciales evocados, se sabe que la actividad eléctrica registrada en el cuero cabelludo es la suma de los potenciales post-sinápticos generados por la despolarización e hiperpolarización de las células cerebrales.

Se utiliza esta técnica de registro para investigar estos procesos psicológicos ya que los procesos cognitivos se suponen relacionados con la actividad eléctrica cerebral, y puesto que los potenciales evocados son fluctuaciones de los potenciales eléctricos del cerebro provocadas por la ocurrencia de un suceso o por la presentación de un estímulo, pueden ser buenos candidatos para ayudarnos a comprender mejor el sistema cognitivo humano. Los investigadores consideran que los potenciales evocados relacionados con eventos (ERP en inglés) son indicadores de procesos o subprocesos cognitivos o perceptivos.

Para obtener los ERP se colocan electrodos sobre varias posiciones en la cabeza, registrándose la actividad eléctrica siendo la amplitud de entre 1 y 20 microvoltios en los ERP, y de entre 50 y 100 microvoltios para el EEG (Kutas y Van Petten, 1994).

Las características definatorias de los ERP según Donchin, Ritter y McCallum (1978) son una combinación de su polaridad, latencia, topografía y sensibilidad a las características de la manipulación experimental, donde las características primera y tercera recogen información sobre la fuente fisiológica, mientras que las otras dos se refieren a la función psicológica:

- a) La polaridad: los componentes pueden ser de dos tipos: positivos o negativos. En este sentido, denominaremos P a los componentes con polaridad positiva, y N a los que tengan polaridad negativa.

- b) La latencia: se mide tomando el tiempo en milisegundos desde la presentación del estímulo hasta la aparición del pico o del valle; esto es, hasta el punto de máxima o mínima amplitud dentro de una ventana de latencia concreta.

- c) La topografía: es la distribución en el cuero cabelludo, ya que el registro de la actividad eléctrica se realiza en varias localizaciones, lo que nos permite, entre otras cosas, detectar si existe un lugar en concreto donde aparece un determinado componente y si hay diferencias hemisféricas. Pero, debemos tener muy en cuenta lo siguiente: las fluctuaciones en el voltaje recogidas con un electrodo sobre el cuero cabelludo no han de tomarse como actividad originada por el tejido cerebral directamente subyacente a esa localización, es decir, que la topografía sobre el cuero cabelludo no suministra un mapa de la localización neuronal. De hecho, la actividad generada en un área concreta del cerebro puede haberse registrado en una localización situada a una considerable distancia de su generador.

- d) La sensibilidad a la manipulación experimental: es el criterio más difícil, puesto que existen componentes bastante inespecíficos. Podemos designar un componente en función de las características físicas de la onda medida en un grupo de datos, o por otra parte, fundamentado en el proceso psicofisiológico que supuestamente la origina.

Existe una clasificación que diferencia entre ERP exógenos y endógenos:

- Los potenciales exógenos son de latencia más corta que los potenciales endógenos (los primeros aparecen en los primeros 100 o 150 ms tras la presentación del estímulo), mientras que los segundos son más tardíos
- Los exógenos se ven afectados por las características físicas del estímulo (p.ej., la modalidad de presentación -auditiva, visual, somestésica, etc.- o la intensidad), mientras que los endógenos no. Estos últimos se ven influidos por factores tales como la atención, la relevancia de la tarea o el estado del sujeto durante el procesamiento del estímulo, además, algún potencial endógeno aparece incluso en ausencia de un suceso externo, como, por ejemplo, cuando un estímulo esperado no ocurre.

Dentro de los componentes sensoriales se recogen todos aquellos denominados exógenos. Este grupo de potenciales se caracteriza por ser respuestas obligatorias del sistema nervioso a estímulos externos. De hecho, se considera que recogen la actividad de las vías sensoriales cuando transmiten la señal generada por los receptores periféricos al sistema de procesamiento central. Esto se apoya en el hecho de que se

trata de componentes específicos de la modalidad sensorial, de manera que difieren en forma y en distribución cortical dependiendo del tipo de modalidad sensorial en que se presente el estímulo (Hillyard, Picton y Regan, 1978).

- N100 o Processing Negativity: se refiere a una onda negativa que aparece aproximadamente a los 100 milisegundos. Este potencial surge al comparar la actividad evocada por estímulos atendidos con aquella evocada por estímulos no atendidos. Concretamente esta negatividad puede estar asociada a la atención.
- P300: se trata de una onda positiva que suele aparecer ante estímulos inesperados que son relevantes para la tarea que ha de realizar el sujeto, y su latencia varía dependiendo de la dificultad de la tarea. Además, su amplitud se ve modulada por el grado de expectación que el sujeto tiene de un determinado estímulo (Callaway 1978 y Duncan-Johnson y Donchin 1979).

El componente P300 se ha utilizado en una gran variedad de aplicaciones teóricas, empíricas y clínicas. En particular, se considera como un índice de la memoria de trabajo (Donchin, Kans, Bashore, Coles y Gratton, 1986), de forma que su amplitud informa de la asignación de los recursos atencionales y su latencia es un indicador del tiempo de evaluación del estímulo (Katayama y Polich, 1998).

- N400. Se trata de una onda negativa, con topografía centroparietal y latencia en torno a los 400 ms, que se considera un indicador del procesamiento semántico

En consecuencia, parece que este componente era un indicador de dos procesos: el primero, la expectativa que tenemos de una palabra a partir del contexto y, el segundo, la asociación semántica que una palabra guarda con la mejor terminación. En definitiva, es un índice de la cantidad de preparación semántica o activación que una palabra recibe del contexto que la precede (ver figura 10.5).

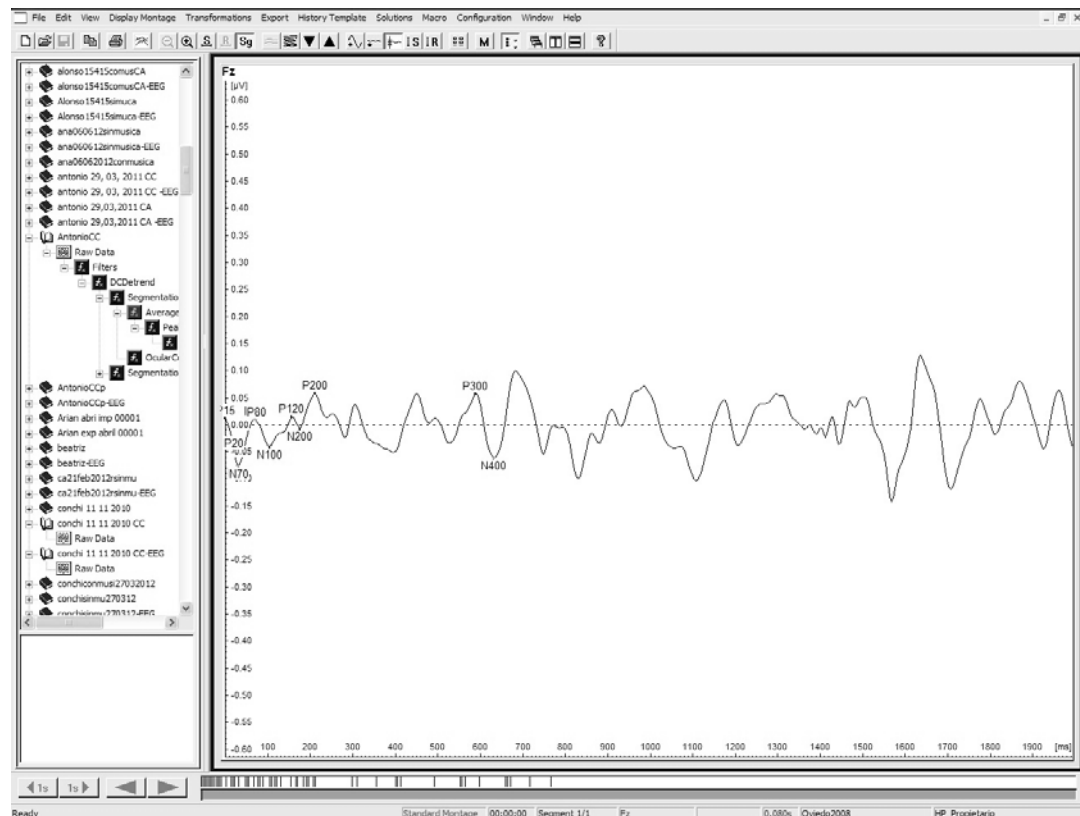


Figura 10.5: Detalle de onda cerebral con los potenciales indicados.

Además de los componentes que se han detallado en párrafos anteriores, existen otros, de descubrimiento más reciente, con los que se está trabajando en la actualidad; pero cuya naturaleza específica todavía está por delimitar, aunque la mayoría entra en la categoría de muy larga latencia.

10.7.3 EL REGISTRO DE POTENCIALES EVOCADOS

La utilidad fundamental del registro de ERP nos permite obtener información sobre el funcionamiento de la cognición humana. En este sentido vamos a destacar las principales ventajas que presenta esta técnica para el estudio de los procesos cognitivos:

- La primera radica en el hecho de que el registro de potenciales evocados es una técnica no invasiva. Consiste en la utilización de electrodos ubicados en la superficie de la cabeza, siendo el registro totalmente inocuo para la persona.
- En segundo lugar, permite obtener una medida de la actividad cerebral en contextos en los que la obtención de una respuesta conductual puede ser problemática o imposible.
- La tercera ventaja consiste en que el registro de ERP nos aporta múltiples datos. A diferencia de las medidas conductuales, como por ejemplo el tiempo de reacción, que nos aporta una única fuente de información para explicar el proceso psicológico bajo estudio, los potenciales evocados nos ofrecen una perspectiva multidimensional del proceso. Nos referimos a que se puede obtener para cada condición experimental información sobre la latencia, la amplitud, la polaridad y la topografía del componente de interés. Todas estas fuentes de información pondrán a nuestro alcance un conocimiento más preciso del proceso psicológico subyacente.
- Cuarta ventaja: la alta resolución temporal del registro de ERP -en orden de milisegundos- nos permite estudiar la evolución o secuencia de actividad neuronal subyacente a funciones cerebrales superiores. De este

modo se podrán investigar distintos procesos cognitivos en tiempo casi real.

Todo lo expuesto anteriormente, nos plantea el registro de ERP como una herramienta que puede ayudarnos y apoyar nuestra investigación en el estudio de la memoria de reconocimiento.

SEGUNDA PARTE

ESTUDIO EMPÍRICO

Aquellos que dicen que no se puede hacer, no deberían interrumpir a quienes lo estamos haciendo.
Thomas Edison

EXPERIMENTO 1

Cómo es posible que la matemática, un producto del pensamiento humano independiente de la experiencia, se adapte tan admirablemente a los objetos de la realidad. Albert Einstein.

ÍNDICES DE DETECCIÓN DE SEÑALES Y MODELADO ESTADÍSTICO EN MEMORIA DE RECONOCIMIENTO²

1. INTRODUCCIÓN

El primer problema que debemos enfrentar en la aplicación del modelado estadístico a la memoria de reconocimiento es: ¿existen ya técnicas fiables para detectar la señal diferenciada del ruido y realizar los análisis correspondientes?, porque no vamos a aplicar una técnica relativamente compleja a este campo, si ya existe alguna más simple que resuelve el problema de una forma más parsimoniosa (Pelegina, Wallace, *et al.*, 2011).

Revisando la bibliografía sobre la TDS nos encontramos con los índices de discriminación (Green y Swets, 1998; Swets, 1986) y curvas ROC (Egan, 1975; Metz, Herman y Shen, 1998) aplicados en psicofísica, pero también en atención, percepción, memoria, etc. Una de estas aplicaciones se ha venido desarrollando en memoria de reconocimiento (Egan; 1958)

Una tarea típica de memoria de reconocimiento consiste en utilizar un tipo de información y a partir de ésta, extraer aquellas acciones (eventos o frases) que ejercerán de señal y ruido. Básicamente el modelo es el mismo para cualquier tipo de estímulos: señales luminosas, sonidos, variaciones de gris en una imagen, respuestas de un juez sobre la presencia o no de unos hechos, diagnóstico médico, opinión a favor o en contra de algo, etc. Para conseguirlo hay que utilizar indicadores o índices que permitan medir de forma adecuada los procesos citados (discriminación, decisión...).

² Los datos de este experimento provienen del laboratorio de psicofísica de la Facultad de Psicología de Málaga y han sido publicados en artículos pero nunca con el enfoque del modelado utilizado en este trabajo.

Una vez obtenida una clasificación adecuada de la señal y el ruido, el problema es metodológico: discriminación, representación de las repuestas y modelado estadístico.

Entre los índices de sensibilidad destacamos d' , índice paramétrico, que asume el supuesto de distribución normal de la señal y el ruido. Es una medida de la sensibilidad del sujeto en el proceso sensorial y se define como la diferencia estimada de medias entre dos distribuciones normales respecto a la señal y al ruido, dada en las desviaciones típicas del ruido:

$$d' = \frac{\mu_{sr} - \mu_r}{\sigma_r}$$

Donde se evalúa la discriminación entre las repuestas a la señal (sr) y al ruido(r), dada la desviación típica del ruido.

El criterio de decisión β es un índice paramétrico bajo el supuesto de distribuciones normales, al igual que d' . Es el cociente estimado entre los valores de la ordenada de la señal con respecto al ruido.

$$\beta = \frac{\hat{y}_{sr}}{\hat{y}_r}$$

donde \hat{y}_{sr} representa el valor de la ordenada de la señal y \hat{y}_r el valor de la ordenada del ruido (McNicol, 1972).

Por su parte Luce (1963a) propone el índice η que no sólo incorpora aciertos y errores (o alternativamente rechazos correctos y rechazos incorrectos), como ocurre con d' , sino las cuatro posibilidades de respuesta consideradas conjuntamente e incluidas en la matriz 2x2.

$$\eta = \left| \frac{P(NO/s) * P(SI/r)}{P(SI/s) * P(NO/r)} \right| = (b * c / a * d)^{1/2}$$

donde el numerador corresponde a los errores y el denominador a los aciertos.

Este modelo es el cociente entre los productos cruzados de errores y aciertos. Se caracteriza porque en el numerador de la ecuación se encuentran los errores en lugar de los aciertos. El índice *LOR* se obtiene mediante la fórmula:

$$LOR = \ln(ad / bc)$$

En esta fórmula se aplica una transformación logarítmica a las proporciones la cual permite "suavizar" (en caso necesario) la excesiva variabilidad de los resultados e incorpora todos los resultados de la tabla de contingencia. Es un índice que ha sido utilizado en el ámbito de los modelos categóricos y propuesto como índice de sensibilidad por Swets (1986).

La *Q* de Yule (ver Bishop, Fienberg y Holland, 1975) se ha desarrollado en el contexto de los modelos categóricos, por lo que se presenta como la mejor alternativa en nuestro caso. Ha sido propuesta como índice de precisión por Nelson (1984). Su cálculo viene dado por:

$$Q = (ad - bc) / (ad + bc)$$

permite obtener valores entre -1 y 1. Esta fórmula penaliza los errores.

Hay que destacar que η , *LOR* y *Q* corresponden a distribuciones logísticas (Swets, 1986). En este sentido, DeCarlo (1998) deduce cómo es posible establecer una relación entre los MLG y la TDS a partir de la regresión logística como modelo de distribución de los datos. En el caso de respuestas SI-NO, y suponiendo una distribución logística, la medida de discriminabilidad adecuada es *d* (McMillan y Creelman, 1991):

$$d = \frac{\psi_s - \psi_r}{\tau}$$

donde ψ_s es la moda de la distribución de la señal y ψ_r , es la moda de la distribución del ruido, mientras que τ está relacionado con la desviación típica σ (DeCarlo, 1998 y Hasselblad y Hedges, 1995). En una distribución normal d es equivalente a d' y τ es igual a σ .

Además, es posible obtener escalas ordinales que pueden ser representadas mediante un modelo *logit* acumulativo (Agresti, 1989a). Esto, en el contexto de la TDS ha tenido una versión propia (DeCarlo, 1998). En efecto, aplicando la transformación *logit* a los aciertos y falsas alarmas obtenemos:

$$\text{logit } p(Y=1|S) = \frac{\psi_s - c}{\tau} \quad \text{y} \quad \text{logit } p(Y=1|R) = \frac{\psi_r - c}{\tau}$$

donde Y es la probabilidad de repuestas SI, s es la señal y R el ruido. Además, en la fórmula, se ha sustituido por c uno de los elementos de la fórmula anterior: En la primera parte de la expresión se sustituye por ψ_r y en el segundo caso por ψ_s . Si consideramos estas últimas fórmulas:

$$\text{logit } p(Y=1|X) = \frac{\psi_r - c}{\tau} + \frac{\psi_s - \psi_r}{\tau}$$

donde $X = 1$ indica los aciertos y $X = 0$ las falsas alarmas (codificación dummy). Considerando estas fórmulas e igualando $\psi_r = 0$ y $\tau = 1$, entonces $\psi_s = 1$ podemos escribir conjuntamente ambas fórmulas de una forma lineal mediante la siguiente expresión:

$$\text{logit } p(Y=1|X) = -c + dX$$

donde X es el valor de la pendiente, $-c$ es el intercepto y d la distancia. Esta fórmula corresponde al procedimiento SI-NO. Además, se puede reescribir para el procedimiento de escalas de estimación de la forma siguiente:

$$\text{logit } p(Y > j | X) = -c_j + dX$$

donde se han añadido una serie de valores j ordenados de mayor a menor y un criterio de decisión múltiple c_j que corresponde a los criterios de cada uno de los niveles de j .

Como podemos observar en los modelos representados en estas fórmulas se pueden relacionar fácilmente con los MLG. En efecto, consideremos el modelo para una respuesta binaria (Agresti, 1989a):

$$\pi_1 = \alpha + \beta x$$

que corresponde a un modelo de regresión logit con codificación dummy. Es decir, es el modelo de regresión logística. En detección de señales el predictor lineal básico viene determinado por:

$$\eta = c_j - dX$$

donde X es la señal o el ruido, c_j representa un criterio simple o múltiple y d es la medida escalada de la distancia (DeCarlo, 1998). Para el caso de la escala ordinal el modelo logit viene representado por:

$$\text{logit } p(Y \leq j) = \alpha_j + \beta x$$

Esta fórmula representa un modelo lineal logit acumulado (Agresti, 1989a).

Estas fórmulas muestran un evidente paralelismo entre TDS y MLG que proporciona un punto de vista complementario entre los modelos de regresión y detección de señales.

Los índices d' , η , LOR y Q son consistentes con la existencia de un criterio de decisión independiente (Pelegrina, Ruiz-Soler, *et al.*, 2000). Ello está de acuerdo con la propuesta de Wald (1950) que partiendo de la teoría de la decisión establece en términos estadísticos que la precisión en la toma de decisiones tiene que ser independiente del resto de probabilidades.

Los índices anteriores representan una serie de posibilidades analíticas que utilizan técnicas diferentes: operaciones que incluyen solo una parte de las celdillas, las que incluyen relaciones cruzadas, las que incluyen marginales de fila o de columna, las que incluyen solo una columna y las que incluyen todos los valores a , b , c y d . Corresponde al investigador una vez obtenidos los datos empíricos, decidir el índice más adecuado a la hipótesis.

Finalmente, y en orden a mostrar la aplicabilidad de lo anterior proponemos una investigación que se desarrolla en un ámbito sustantivo. Todo el proceso se resume en tres fases sucesivas que presentamos en los apartados restantes de esta investigación: categorización de información organizada de forma esquemática, medida mediante índices y modelado estadístico.

La categorización viene determinada por la elaboración de un *script* mediante categorías altamente significativas (muy típicas) y mediante dos intervalos de retención, un intervalo de retención largo (IRL) y un intervalo de retención inmediato (IRI) (Pelegrina, Ruiz-soler, *et al.*, 2000). El IRL permite suponer que la discriminación será mínima debido a dos razones: un efecto de tipicidad que produce una contratransferencia en el aprendizaje y el IRL que producirá un olvido relacionado con

el tamaño largo del intervalo. Por el contrario, cuando el mismo tipo de información se aplica con un IRI, la discriminación será mejor, ya que disminuirá el olvido, debido a un intervalo corto, pero también baja, debido a que seguirá "activo" el efecto de tipicidad en un contexto esquemático. Dicho efecto produce un nivel alto de interferencia en memoria de reconocimiento.

Ahora bien, de acuerdo con la metodología propuesta en este trabajo, no sólo obtendremos índices determinados para comparar hipótesis puntuales (elección del índice más adecuado); sino que, además, obtendremos los modelos mejor ajustados mediante el modelado estadístico. Todo ello, considerado conjuntamente, nos permitirá elaborar un análisis bastante completo de los datos obtenidos en la esta investigación: adquisición de conocimientos, y dentro de ésta adquisición conocer mejor cómo se adquieren conceptos muy significativos, y muy relacionados entre sí.

2 MÉTODO

Participantes

Los datos proceden de dos muestras de 32 sujetos que cursaban asignaturas del área de Metodología de las Ciencias del Comportamiento (Facultad de Psicología de la Universidad de Málaga). Ambas muestras estaban formadas por hombres y mujeres de edad comprendida entre 18 y 22 años. En este trabajo se reinterpretan resultados ya obtenidos en el laboratorio desde la perspectiva del modelado estadístico, con objeto de analizar las supuestas ventajas de dicho enfoque.

Material

Estaba formado por las 22 frases representadas en el anexo 1. Estas fueron seleccionadas mediante un análisis factorial en el que inicialmente habrá 75 categorías. Se eligieron aquellos cuya carga fue igual o mayor a 0.70. Este criterio define la alta tipicidad de las acciones del anexo 1.

Procedimiento

Consistió en una tarea típica de memoria de reconocimiento. La tarea que debían realizar los sujetos era leer lo más rápidamente posible una lista de frases escrita en un folio DIN A4, en letra "france". Las frases estaban formadas por contenidos y actividades (o acciones) realizadas en asignaturas del área de Metodología de las Ciencias del Comportamiento (anexo 1). El tiempo de lectura, fijo para todos los sujetos, fue de 2 segundos por frase. Esta prueba la realizaron 32 sujetos.

Pasada una semana, a la misma hora, se realizó la prueba de reconocimiento (anexo 2). Cada alumno/a contestó "SI" o "NO". Esta prueba se realizó sin límite de tiempo. El segundo grupo (grupo de IRI) era de 32 sujetos. El procedimiento fue equivalente al grupo de IRL a diferencia de que el intervalo de retención era inmediato. Es decir, los sujetos pasaron por la prueba de reconocimiento (anexo 2) inmediatamente después de leer las frases del anexo 1.

3 RESULTADOS

Los datos obtenidos se representan en las tablas 1 y 2. En la tabla 1 los resultados corresponden a 32 sujetos por 22 estímulos (11 señales y 11 ruidos). Ello dio lugar a 704 estímulos, 352 señales y 352 ruidos. Estos se han organizado en una escala unificada de 10 puntos que incluyen desde la respuesta "NO 5" hasta la respuesta "SI 5".

Tabla 1: Condición IRL

Condición IRL										
	Respuesta No					Respuesta SI				
Esc.	5	4	3	2	1	1	2	3	4	5
S	83	37	35	13	2	2	6	30	39	105
	$\Sigma = 170$					$\Sigma = 182$				
R	68	45	34	20	2	2	3	33	35	96
	$\Sigma = 169$					$\Sigma = 169$				
Q de Yule =0,090										

Podemos observar que los valores referidos a las falsas alarmas son similares a los valores correspondientes a los aciertos. Ésta es una de las razones para aplicar el índice Q. La otra razón es que los valores de la a, c y b, d no son simétricos. Es decir, 435, 339 y 210, 239 respectivamente (tabla 1). En efecto, este índice incorpora todas las repuestas de los sujetos a, b, c y d y no sólo a y c (aciertos y falsas alarmas): el valor de Q es 0.090, lo cual indica una sensibilidad muy pequeña.

Por otra parte, observamos que en la condición IRI (tabla 2) se obtuvieron los resultados para un total de 32 sujetos por 22 estímulos. Ello permitió registrar las respuestas a 352 señales y 352 ruidos. En los demás extremos la tabla 2 es similar a la 1, salvo la puntualización de que el valor de Q fue 0.023, aún más bajo que en la condición de intervalo largo de retención, lo cual resulta ser contrario a una de nuestras hipótesis. Obtener índices adecuados a los datos es indispensable para proceder seguidamente al

modelado. En efecto, el modelado es fiel a la realidad a la que se adapta y los indicadores que se utiliza deben medir adecuadamente la realidad. Sin los índices adecuados el modelado no es correcto.

Tabla 2: Condición IRI

		Condición IRI									
		Respuesta No					Respuesta SI				
Esc.		5	4	3	2	1	1	2	3	4	5
S		108	35	33	14	2	2	3	15	23	114
		$\Sigma = 192$					$\Sigma = 157$				
R		85	41	35	17	2	2	3	33	20	103
		$\Sigma = 180$					$\Sigma = 161$				
Q de Yule =0,023											

El modelado estadístico es un ejercicio analítico complementario y necesario, con la elección de índices, ya que es una forma de análisis que establece comparaciones cualitativas y cuantitativas entre las variables y condiciones experimentales, de tal forma que nos permite extraer información específica y global sobre todas las variables y categorías del experimento, de una forma conjunta, eligiendo el modelo mejor ajustado y más parsimonioso.

Mediante el modelado obtenemos una panorámica completa de los datos empíricos (tablas 1 y 2), y mediante el ejemplo que hemos propuesto para ilustrar nuestro trabajo en este sentido cabe señalar que en la condición IRL han resultado ajustados cinco modelos mediante el programa *lem* (Vermunt, 1997). Donde *A* es la escala de estimación, *B* es la señal-ruido y *C* la respuesta si-no (ver tabla 3).

Tabla 3: Modelos de la condición IRL

Modelo	χ^2	D	g.l.	P	BIC
Nulo	553,30	568,19	19	0,000	443,99
A	27,18	27,96	15	0,021	-70,09
B	567,91	537,64	18	0,000	450,25
C	550,07	567,98	18	0,000	450,32
A,B	27,19	27,67	14	0,016	-63,84
A,C	27,03	27,74	14	0,015	-63,77
B,C	548,59	567,70	17	0,000	456,58
A,B,C	27,06	27,46	13	0,011	-57,51
AB	24,20	25,80	10	0,004	-39,56
AC	5,53	5,57	10	0,850	-59,80
BC	548,59	567,50	16	0,000	462,91
AB,C	24,00	25,59	9	0,002	-33,23
AC,B	5,25	5,28	9	0,809	-53,55
BC,A	26,65	27,26	12	0,007	-51,18
AB,AC	3,40	3,41	5	0,636	-29,27
AB,BC	23,80	25,39	8	0,001	-26,90
AC,BC	5,05	5,08	8	0,749	-47,21
AB,AC,BC	3,34	3,36	4	0,499	-22,79
ABC	0,00	0,00	0	1,000	0,00

En la condición IRI ha resultado ser también cinco los modelos ajustados (ver tabla 4). IRL1. En la discusión consideramos de nuevo y de una forma integrada los dos niveles de análisis: índices, y modelado estadístico (ver tabla 4).

Tabla 4: Modelos de la condición IRI

Modelo	χ^2	D	g.l.	P	BIC
Nulo	773,07	718,27	19	0,000	594,08
A	40,96	43,89	15	0,000	-54,16
B	771,45	718,18	18	0,000	600,52
C	784,99	714,04	18	0,000	596,38
A,B	40,93	43,80	14	0,000	-47,71
A,C	37,17	39,66	14	0,000	-51,85
B,C	783,42	713,95	17	0,000	602,83
A,B,C	37,15	39,57	13	0,000	-45,41
AB	35,32	37,28	10	0,000	-28,08
AC	11,08	11,26	10	0,338	-54,11
BC	783,20	713,61	16	0,000	609,02
AB,C	31,29	33,05	9	0,000	-25,77
AC,B	10,99	11,16	9	0,265	-47,67
BC,A	37,20	39,22	12	0,000	-39,21
AB,AC	4,60	4,65	5	0,460	-28,03
AB,BC	30,94	32,71	8	0,000	-19,58
AC,BC	10,65	10,82	8	0,212	-41,48
AB,AC,BC	3,72	3,74	4	0,442	-22,40
ABC	0,00	0,00	0	1,000	0,00

4 DISCUSIÓN

Para obtener un índice de sensibilidad hemos optado por la Q de Yule, ya que d' no es adecuada por ser un índice paramétrico bajo el supuesto de distribución normal de los datos (Pelegriña, Ruiz-Soler, *et al.*, 2000). Tampoco es adecuada d' si las celdillas de los aciertos y falsas alarmas no son complementarias a los rechazos correctos y rechazos incorrectos (ello se debió a la presencia de ceros empíricos en los datos de algunos sujetos).

Si consideramos los índices restantes: η , LOR y Q , hemos de tener en cuenta otros extremos. En efecto, si observamos los datos vemos que la discriminación es muy baja. Es decir, el parecido entre frases organizadas en torno a un tema central produce el

denominado "efecto de tipicidad en un contexto esquemático". Esto ha dado lugar a un bajo nivel de reconocimiento, lo cual produce una cantidad alta de falsas alarmas en ambas condiciones experimentales (IRL e IRI).

En tal caso, hay que valorar las falsas alarmas y los rechazos incorrectos (los errores). Si considerarnos lo anterior entenderemos que la fórmula más sensible a los errores es la Q de Yule, ya que en este índice los errores son sustractivos en el dividendo y aditivos en el divisor. Realizado el cálculo, observamos que el índice de sensibilidad ha resultado ser muy bajo en ambas condiciones experimentales ($Q = 0.090$ en la condición IRL y $Q=0,023$ en la condición IRI), ello podría indicar que el "efecto de tipicidad en un contexto esquemático" ha sido tan potente que no se ha permitido observar el "efecto del intervalo de retención".

Realizamos un ajuste de los modelos mediante lem (Vermunt 1997). Este análisis aporta otras posibilidades analíticas, dados los datos que hemos obtenido. Así, hemos visto en los resultados que en la condición IRL aparecen dos modelos a los que se pueden ajustar los datos. El modelo mejor ajuste de la condición IRL ha resultado ser el AC (ver tabla 5). El modelo indica que la escala de estimación y la respuesta están relacionadas; aunque en otro de los modelos ajustados aparece también la relación entre la señal-ruido y la respuesta (modelo AC,BC). Estas asociaciones indican que las respuestas se relacionan tanto con el tipo de escala como la señal-ruido. Recordemos: A es la escala de estimación, B es la señal-ruido y C la respuesta si-no. La respuesta C está asociada tanto a la escala como a la señal- ruido (al tipo de estímulo), asociada dos a dos y como grupo, tal y como indica el modelo, sin proponer, en principio, ningún

tipo de direccionalidad entre ellos. En este modelo predomina C. Estaríamos ante un modelo de estímulo- respuesta (más bien asociativo) (ver tabla 5 y 6).

Tabla 5: Comparación de modelos condición IRL

Modelo	Eliminado	ΔD	Δgl	P	BIC	
MA	ABC	0	0	1,000	0	
MR	AB,AC,BC	ABC	3,36	4	0,499	-22,79
MA	AB,AC,BC		3,36	4	0,499	-22,79
MR	AB,AC	BC	0,05	1	0,636	-29,27
MR	AB,BC	AC	22,03	4	0,001	-26,90
MR	AC,BC	AB	1,72	4	0,749	-22,79
MA	AC,BC		5,08	8	0,749	-22,79
MR	AC,B	BC	0,20	1	0,809	-53,55
MR	BC,A	AC	22,18	4	0,007	-51,18
MA	AC,B		5,28	9	0,809	-53,55
MR	AC	B	0,29	1	0,850	-59,80
MR	B	AC	532,36	9	0,000	450,25
MF	AC	B	5,57	10	0,850	-59,80

Tabla 6: Valores observados, estimados y residuos del modelo ajustado

*** FREQUENCIES ***

A B C	observed	estimated	std. res.
1 1 1	2.000	2.000	0.000
1 1 2	2.000	2.000	0.000
1 2 1	2.000	2.000	0.000
1 2 2	2.000	2.000	0.000
2 1 1	13.000	16.500	-0.862
2 1 2	6.000	4.500	0.707
2 2 1	20.000	16.500	0.862
2 2 2	3.000	4.500	-0.707
3 1 1	35.000	34.500	0.085
3 1 2	30.000	31.500	-0.267
3 2 1	34.000	34.500	-0.085
3 2 2	33.000	31.500	0.267
4 1 1	37.000	41.000	-0.625
4 1 2	39.000	37.000	0.329
4 2 1	45.000	41.000	0.625
4 2 2	35.000	37.000	-0.329
5 1 1	83.000	75.500	0.863
5 1 2	105.000	100.500	0.449
5 2 1	68.000	75.500	-0.863
5 2 2	96.000	100.500	-0.449

Por otra parte, mediante la condición IRI se ha obtenido, al igual que en el caso anterior el ajuste del modelo AC , en el que A , C , y AC han resultado significativos (ver tabla 6). Se ha seleccionado por ser el modelo más parsimonioso, del que destacamos la ausencia de B , es decir, que influye menos el estímulo. Esto coincide una vez más con nuestra hipótesis de partida, así como con el modelo citado anteriormente. Aunque, atendiendo a criterios teóricos, el modelo AB,AC en este caso nos parecería más adecuado (ver tabla 7).

En definitiva, el efecto de tipicidad llega a ser tan fuerte que anula el intervalo de retención (recordemos que en esta condición el intervalo de retención es inmediato). Sin embargo, hay una asociación entre escala (A) y respuestas (C). Ello, presente en todos los modelos indicaría un "efecto general de escala" asociado a la respuesta. Así, se acepta la hipótesis propuesta sobre el efecto de tipicidad en un contexto esquemático. Esto nos podría estar indicando que en memoria los esquemas mentales son "más fuertes" que la señal-ruido. Lo que nos lleva a pensar que estamos ante un modelo de memoria cognitivo (que va de arriba a abajo) y no ante un modelo asociacionista (estímulo-respuesta). La aplicación del modelado nos permite delimitar qué modelo psicológico está implicado en estos procesos de reconocimiento, ya que el modelo formal es el mismo, coincide con el modelo sustantivo, se identifica con él, tal y como ha venido ocurriendo de forma general en la ciencia. En este contexto la significación estadística habitualmente prescriptiva por científicos y editores deviene en una técnica necesaria pero no suficiente para elaborar los modelos matemáticos.

Tabla 7: Comparación de modelos condición IRI

Modelo		Eliminado	ΔD	Δgl	P	BIC
MA	ABC		0	0	1,000	0
MR	AB,AC,BC	ABC	3,74	4	0,442	-22,40
MA	AB,AC,BC		3,74	4	0,442	-22,40
MR	AB,AC	BC	0,91	1	0,460	-28,03
MR	AB,BC	AC	28,97	4	0,000	-19,58
MR	AC,BC	AB	7,08	4	0,212	-41,48
MA	AB,AC		4,65	8	0,460	-28,03
MR	AB,C	AC	28,40	1	0,000	-25,77
MR	AC,B	AB	6,51	1	0,265	-47,67
MA	AC,B		11,60	9	0,265	-47,67
MR	AC	B	0,34	1	0,338	-54,11
MR	B	AC	706,58	9	0,000	682,18
MF	AC	B	11,26	10	0,338	-54,11

Tabla 8: Valores observados, estimados y residuos del modelo ajustado

*** FREQUENCIES ***

A B C	observed	estimated	std. res.
1 1 1	2.000	2.000	0.000
1 1 2	2.000	2.000	0.000
1 2 1	2.000	2.000	0.000
1 2 2	2.000	2.000	0.000
2 1 1	14.000	15.500	-0.381
2 1 2	3.000	3.000	0.000
2 2 1	17.000	15.500	0.381
2 2 2	3.000	3.000	0.000
3 1 1	33.000	34.000	-0.171
3 1 2	15.000	24.000	-1.837
3 2 1	35.000	34.000	0.171
3 2 2	33.000	24.000	1.837
4 1 1	35.000	38.000	-0.487
4 1 2	23.000	21.500	0.323
4 2 1	41.000	38.000	0.487
4 2 2	20.000	21.500	-0.323
5 1 1	108.000	96.500	1.171
5 1 2	114.000	108.500	0.528
5 2 1	85.000	96.500	-1.171
5 2 2	103.000	108.500	-0.528

Además, hemos de destacar un resultado post hoc, es decir un efecto de escala asociado a la respuesta en todos los modelos y a un efecto de escala asociado a la señal-

ruido en las dos condiciones. En efecto, esto no es nuevo en ciencia, el hecho de que la medida “transforma” siempre por definición el objeto medido. Un problema fundamental de las ciencias físicas y conductuales o psicológicas. Pero de nuevo mediante el enfoque aquí presentado evaluamos el alcance de este problema. Es decir, la necesidad de trabajar el tema relacionado con la naturaleza de la medida con más ahínco si cabe que en ciencias físicas, debido a la mayor complejidad del objeto medido.

Resumiendo lo anterior, hemos visto que el modelado ofrece posibilidades diferentes de análisis. Posibilidades que al ser compatibles y complementarias enriquecen las posibilidades del análisis en líneas diferentes y fundamentales para la ciencia psicológica. Entre todo lo anterior cabe destacar que el modelado estadístico ofrece amplias posibilidades para analizar no sólo datos cuantitativamente diferentes, sino también modelos cualitativamente distintos. Por ejemplo, en esta aplicación las condiciones IRL e IRI ofrecen posibles teorías explicativas en diferentes direcciones y en el mismo modelo estructural, cuestión menos plausible de obtener mediante los modelos experimentales clásicos o la estadística que habitualmente utilizamos (solo mediante pruebas de significación). Hemos de insistir en el hecho de que las aplicaciones desarrolladas aquí han resultado ser complementarias entre sí (datos cualitativos, datos cuantitativos, significación estadística, y modelado sustantivo y formal de las variables y categorías). No obstante, constituyen hipótesis de trabajo que debe ser aún mejor desarrollado en el futuro. Todo lo cual permite un trato adecuado, completo y preciso de los resultados empíricos (ver anexos 1 y 2).

Anexo 1. Frases utilizadas en la memorización. (Experimento 1)

(Pelegrina, Ruiz-Soler, *et al.*, 2000)

Frases señal

Interés del alumno en las tareas de investigación

Participación del alumno en las tareas de investigación

Aprendizaje de las tareas de investigación

Definición de variables en la investigación por parte del alumno

Formular problemas de investigación por parte del alumno

Interés en la exposición por parte del profesor

Participación durante la exposición del profesor

Exposición por parte del profesor de los paquetes estadísticos

Investigación del alumno con ayuda de Internet

Elección del tema por el alumno y conocimiento bibliográfico

Elaboración de hipótesis por parte del alumno

Anexo 2 (Experimento 1)

Frases señal más ruido (al 50%) para el reconocimiento posterior

Interés del alumno en las tareas de investigación

Motivación del alumno en las tareas de investigación

Participación del alumno en las tareas de investigación

Asistencia del alumno a tareas de investigación

Aprendizaje de las tareas de investigación

Elaboración del plan de investigación por el alumno

Definición de variables en la investigación por parte del alumno

Diseño específico de la investigación por parte del alumno

Formular problemas de investigación por parte del alumno

Toma de decisiones del alumno en la investigación

Interés en la exposición por parte del profesor

Motivación en la exposición por parte del profesor

Participación durante la exposición del profesor

Investigación del alumno mediante CD-ROM

Elección del tema por el alumno interés del alumno

Elección del tema por el alumno y motivación del alumno

Elección del tema por el alumno y conocimiento bibliográfico

Elaboración de hipótesis por parte del alumno

EXPERIMENTO 2

El estudio profundo de la naturaleza es la fuente más fértil
de descubrimientos matemáticos. Joseph Fourier.

MODELADO ESTADÍSTICO APLICADO EN MEMORIA DE RECONOCIMIENTO Y POTENCIALES EVOCADOS³

1 INTRODUCCIÓN

Para el segundo de los experimentos de memoria de reconocimiento, vamos a evaluar tareas de memoria “explícita” e “implícita”, manipulando los estímulos y las condiciones experimentales para contrastar las diferencias entre ambos tipos de memoria (Baddeley, 1999). Hemos de añadir que los constructos a los que nos referimos son elaboraciones para este experimento sin pretender teorizar sobre el tema de la memoria.

La tarea planteada consistió en el estudio del resultado obtenido en el rendimiento de un observador al valorar una señal. El procedimiento más sencillo consiste en establecer juicios de valor, basados en respuestas SI- NO. Desde la TDS partimos del supuesto de que el sujeto elige un valor como criterio de respuesta basado en la máxima verosimilitud. En nuestra investigación, los valores de sensación de la señal sobre los valores de sensación del ruido en el ámbito de la memoria de palabras, (Egan, 1958).

Los datos de este estudio proceden de una serie de estímulos sensoriales que proponemos al participante y que elicitán respuestas electrofisiológicas en el cerebro, provocando secuencias características de ondas en el trazado electroencefalográfico (EEG) (Hillyard y Kutas, 1983). Estas ondas son diferentes para cada modalidad

³ Los datos de este experimento provienen del laboratorio de psicofísica de la Facultad de Psicología de Málaga y no han sido publicados.

sensorial y su variabilidad depende del tipo de estímulo. En nuestra aplicación utilizamos los potenciales relacionados con eventos (ERP en inglés).

Consideramos el tiempo que tarda en activarse un P300 como variable de respuesta ante los estímulos, la señal y el ruido durante diferentes sesiones de registro. Mediante estas técnicas podemos aislar y registrar áreas del cerebro específicas. En la práctica, la forma de onda de la P300 se debe evocar utilizando un estímulo provisto por alguna de las modalidades sensoriales. El procedimiento típico es el paradigma *odd-ball*, en el que un estímulo diana se presenta entre estímulos de fondo más frecuentes. También puede usarse un estímulo distractor para asegurarse de que la respuesta se debe a la diana en lugar de a un cambio en el patrón de fondo o para manipular la atención. El paradigma *odd-ball* clásico ha sufrido muchas variaciones, pero al final la mayor parte de los protocolos utilizados para evocar la P300 implica alguna forma de hacer consciente una situación o una toma de decisiones. Es en este contexto en el que se introduce el paradigma de la TDS.

La investigación científica a menudo descansa en la medición de la P300 para examinar potenciales relativos a eventos, especialmente cuando se trata de la toma de decisiones, ya que existe una relación bastante directa entre la reducción de la incertidumbre del sujeto y la amplitud del P300. Por otra parte, el deterioro cognitivo a veces está relacionado con modificaciones en el P300, de manera que se puede utilizar como medida para detectar la eficiencia de la respuesta en la evaluación de diferentes procesos cognitivos. Así es que hay un amplio rango de usos del P300 que va, por ejemplo, desde el estudio de la depresión y la drogadicción, a los trastornos de ansiedad, o los procesos cognitivos, como la atención, percepción y memoria.

Las hipótesis que queremos contrastar aquí se refiere a la constatación de las diferencias entre memoria “implícita” y “explícita”, añadiendo a ello la obtención del modelo mejor ajustado, y determinando además en qué medida contribuyen a ello las diferentes condiciones experimentales y variables o categorías del experimento.

2 MÉTODO

Participantes

Participaron 48 Estudiantes de psicología con edades comprendidas entre los 19 y 23 años, que cursaban asignaturas del área de Metodología de las Ciencias del Comportamiento (Facultad de Psicología de la Universidad de Málaga). La muestra estaba formada por hombres y mujeres.

Material

En primer lugar obtuvimos una muestra de 40 palabras a partir de su frecuencia de uso y clase. Las palabras fueron presentadas mediante el programa E-prime v1.1 (Schneider, Eshman y Zucolotto, 2002). El registro se realizó con un equipo de potenciales evocados QuickAmp de 72 canales y el programa de registro Brainvision Recorder v2. Se utilizaron 11 canales: 3 frontales (F3, FZ, F4), 3 centrales (C3, CZ, C4), 3 parietales (P3, PZ, P4), y 2 occipitales (O1, O2).

Procedimiento

Se aplicó el paradigma de la TDS en memoria de reconocimiento de ítems. La mitad de las palabras, 20 para cada condición, fueron consideradas señal y otras 20 ruido. En la condición de memoria “implícita” el participante realizaba una tarea

distractora al observar la primera lista de palabras. Esta tarea consistía en evaluar la simetría de dos números colocados encima y debajo de la palabra a recordar. Además, el participante no era informado de que estaba ante una prueba de memoria. En la condición de memoria “explícita” aparecía la palabra a memorizar y el participante era informado de que debía someterse posteriormente a una prueba de memoria. La aparición de cada palabra estaba precedida por un punto negro en la pantalla durante 500 milisegundos. El participante respondía mediante respuestas SI-NO en caso de que la palabra perteneciera o no a la lista original. Las sesiones tenían una duración de 20 minutos, registrando los datos del EEG en el programa *Brainvision Recorder*. Posteriormente fueron tratados y segmentados mediante el programa *Brainvision Analyzer*.

En este experimento nos decidimos por los Potenciales evocados cognitivos P300. La onda P300 (EP300, P3) es un (ERP) que puede ser registrado mediante Electroencefalografía como una deflexión positiva de voltaje con una latencia de unos 300ms en el EEG. La presencia, magnitud, topografía y duración de esta señal se utiliza a veces en la medición de la función cognitiva de los procesos de toma de decisiones.

Esta metodología permite aislar y registrar localmente un área sin el ruido de otras señales adquiridas a través de electrodos sobre la piel del cráneo. La investigación científica a menudo descansa en la medición del P300 para examinar potenciales relativos a eventos, especialmente cuando se trata de la toma de decisiones.

La obtención del P300 se realizó mediante el programa *Brainvisión Analyzer*. Ambos programas son comercializados para tareas de investigación por *BrainProducts* (ver figura 1).

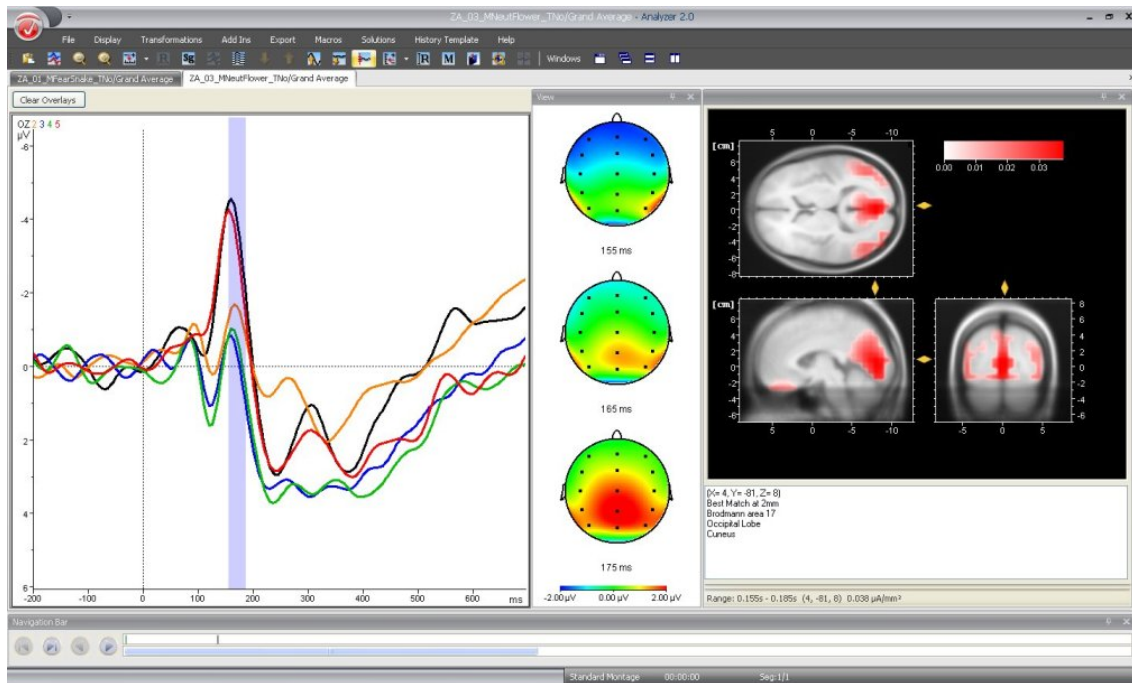


Figura 1: Ejemplo de procesamiento y análisis de señal con *Brainvision Analyzer* v2.

En la figura 1 tenemos un ejemplo de la información que nos puede proporcionar de manera visual el programa *Brainvisión Recorder* tras el procesamiento de la actividad cerebral, que incluye, la forma de la onda del potencial, su latencia, su voltaje, la activación eléctrica por zonas de la corteza cerebral y tiempo, el tipo de onda cerebral, y gracias al módulo *Low Resolution Electromagnetic Tomography (LORETA)* imágenes de resonancia magnética funcional de baja resolución. Paralelamente, el programa registra mucha más información que va recogiendo durante el proceso de registro y que podemos recuperar en forma de archivo de datos. El programa además de realizar el

filtrado de las ondas cerebrales, también permite realizar algunos análisis estadísticos con ellos, comparación de ondas, cálculo de la *grand average*, etc.

3 RESULTADOS

De los resultados obtenidos en el registro de potenciales evocados se seleccionaron para el análisis los pertenecientes a los electrodos frontales (F3, FZ, F4) que son los habitualmente utilizados en el estudio de los ERP (tiempos de latencia en milisegundos). Para el ajuste de estos datos utilizamos el programa *R* (Ihaka y Gentleman, 1996). Se ha ajustado 1 modelo, el modelo lineal de interacción entre las variables: *A* (tipo de memoria) y *B* (señal-ruido) (ver figura 2 y 3).

```
Residuals:
  Min      1Q  Median      3Q      Max
-36.83 -14.73   3.75  11.85  44.17

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      269.250     5.554  48.482 < 2e-16 ***
Memoria[T.Imp]    37.417     7.854   4.764 2.09e-05 ***
SR[T.Ruido]      35.500     7.854   4.520 4.62e-05 ***
Memoria[T.Imp]:SR[T.Ruido] -61.333    11.107  -5.522 1.69e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 19.24 on 44 degrees of freedom
Multiple R-squared:  0.4265, Adjusted R-squared:  0.3874
F-statistic: 10.91 on 3 and 44 DF,  p-value: 1.759e-05
```

Figura 2: Salida del programa R con el modelo ajustado

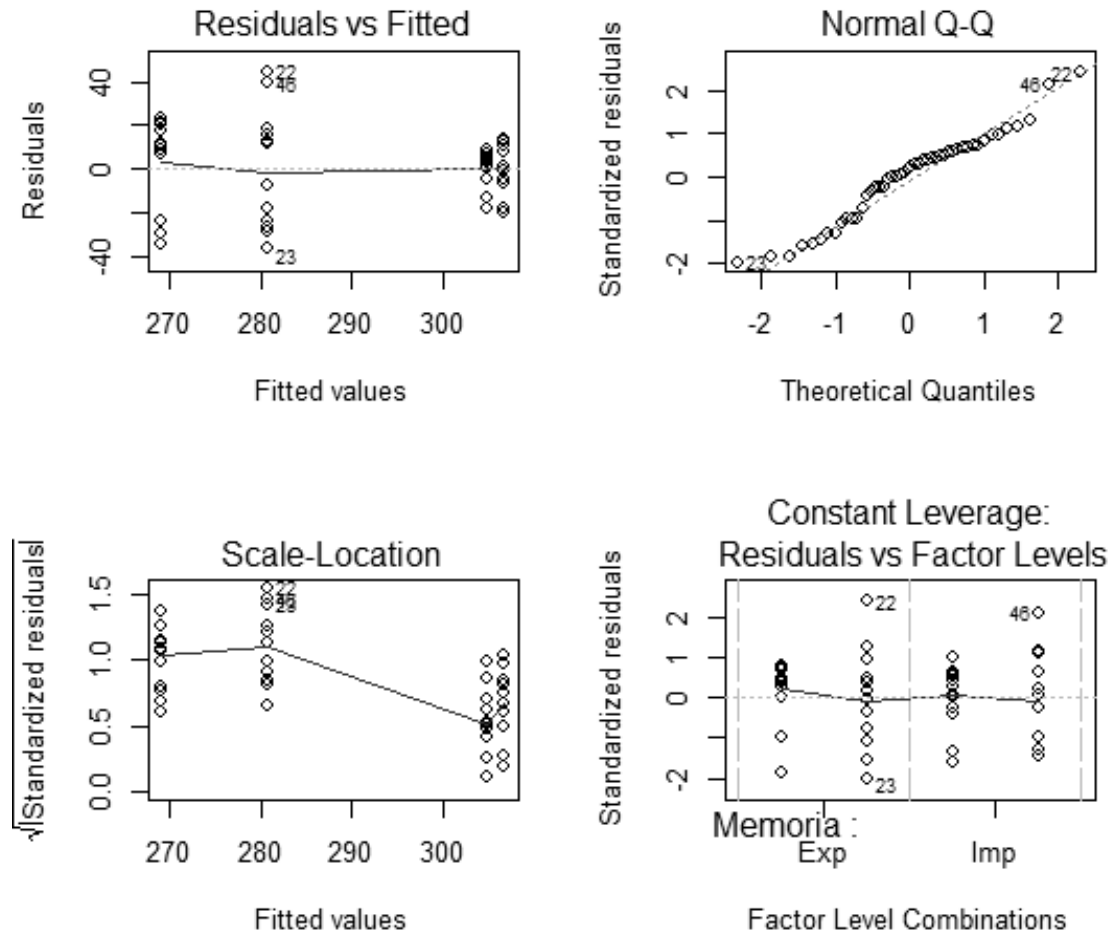


Figura 3: Gráficos de diagnóstico

El modelo que mejor hemos ajustado [$F=10,91$, $p<0,001$] es un modelo lineal de interacción que incluye las variables A (tipo de memoria) y la variable B (señal-ruido), es decir el modelo $A:B$. Este modelo incluye A que resulta de la comparación de los dos tipos de memoria y que es el resultado más significativo desde el punto de vista estadístico (ver tabla 1).

Tabla 1: Resumen del modelo ajustado

Términos	b	EE(b)	gl	t	p	IC95%
Constante	269,250	5,554	1	48,482	<0,001	258,057 a 280,422
(M)Tipo Memoria	37,417	7,854	1	4,764	<0,001	21,588 a 53,245
(SR)Señal-Ruido	35,500	7,854	1	4,52	<0,001	19,671 a 51,328
M*SR	-61,333	11,107	1	-5,522	<0,001	-83,718 a -38,948

Pero también, observamos otros datos relevantes, por ejemplo, la significación entre la señal y el ruido como dos elementos distintos del experimento. Es decir, que los tipos de memoria, además de ser estadísticamente significativos entre sí, se comportan de forma diferente y varían de forma estadísticamente significativa en función de la señal-ruido. Considerando los resultados obtenidos tras el ajuste de los modelos, y teniendo en cuenta que estamos trabajando con un modelo exploratorio (y que hemos obtenido un R^2 ajustado moderado), podríamos decir en primer lugar que ambos tipos de memoria son diferentes, y en segundo lugar que los sujetos harían uso de un tipo u otro de memoria en función de los estímulos presentados y de las sesiones. Por lo que concluimos que la memoria, al ser la más significativa de todas las comparaciones incluye una variabilidad propia y sólo relativamente independiente de la señal-ruido y de las sesiones. En efecto, éstas últimas variables incorporan una variabilidad propia asociada a la memoria, pero con entidad propia y diferenciada, como vemos, ya que la significación estadística, a pesar de que existe, es menor a la comparación específica de tipos de memoria.

4 DISCUSIÓN

El modelo mejor ajustado ha resultado ser de interacción, que incluye las variables que aquí hemos llamado A (tipo de memoria) y la variable B (señal-ruido), es decir el modelo $A:B$ siendo la señal y el ruido diferentes en el modelo. Por consiguiente, podemos asumir que hay interacción entre ambas memorias, Que la memoria denominada “implícita” no es independiente de la memoria que hemos denominado “explícita”. Según estos modelos no existiría una memoria pura, sino una interacción entre memorias que dependería selectivamente del tipo de estímulo utilizado. Con ello

se abre una oportunidad para estudiar la memoria humana mediante modelos de conjuntos difusos donde se deberían trabajar las evidencias a favor de un tipo u otro de memoria y no tanto los modelos o teorías puras de la memoria.

Todo ello aporta un modelo estructural general (saturado y articulado por partes), de acuerdo con unas variables fundamentales: memoria “explícita” e “implícita”; así como unos modelos parciales que permiten evaluar todos los efectos relevantes que se detectan en la respuesta de los participantes a partir de las variables estímulo y de las sesiones de la investigación, en la línea seguida por DeCarlo (1998) y Wallace, Pelegrina, Ruiz-Soler, (2002) en el ámbito de la memoria, la detección de señales y el modelado estadístico. Además, hemos incorporado los potenciales evocados en el mismo contexto, añadiendo una medida psicofísica que tiene una gran estabilidad. Estos resultados nos llevan, a diferencia del experimento anterior, a un modelo asociativo (de abajo a arriba) más que cognitivo, donde no se aprecian los efectos del esquema, ya que, la interacción obtenida entre ambas memorias como modelo general se ve afectada por el tipo de estímulo de forma diferenciada (tanto de la señal como del ruido). Por consiguiente, “lo implícito y lo explícito “, tal y como fue operacionalizado en este experimento, es un constructo unido a variables de estímulo-respuesta más cercano a modelos asociativos que cognitivos. Es decir, es algo unido a lo corpóreo o físico y no tanto a los esquemas de conocimiento (ver en anexo 1 un ejemplo de palabras asociadas a las condiciones experimentales).

Anexo 1 (Experimento 2)

(Material habitual de nuestro laboratorio utilizado en diferentes experimentos)

Lista de palabras	
tuya	según
casa	cara
allí	cerca
duda	tarde
mía	quizá
año	hijo
otra	cuya
fondo	hora
acaso	allá
caso	lado

EXPERIMENTO 3

La difusión de los métodos estadísticos puede contribuir a una sociedad más abierta, más liberal y más tolerante. Daniel Peña.

MODELADO ESTADÍSTICO APLICADO EN MEMORIA DE RECONOCIMIENTO Y TIEMPOS DE REACCIÓN

1 INTRODUCCIÓN

La atención al procesamiento de palabras organizadas según su clase se inicia en 1970. Los estudios se centran en el agramatismo asociado con la afasia de Broca. A partir de estos estudios preliminares se asumió un modelo de almacenamiento y acceso diferencial en palabras de clase abierta (CA) y clase cerrada (CC) (Bradley, 1978; Biassou, Obler, Nespoulous, Dordain y Harris, 1997) (ver anexo de este experimento). Típicamente las palabras de clase abierta son nombres, adjetivos, verbos y adverbios, y las cerradas son artículos, preposiciones y conjunciones. Estos estudios, además, asumen que hay diferentes conjuntos de vocabulario en el léxico mental, y que su procesamiento es computacionalmente distinto (Biassou, *et al.*, 1997).

Partimos de la hipótesis de la aceptación de un modelo de dos procesos en el reconocimiento de palabras de CA y CC (Biassou, *et al.*, 1997), asumiendo que las palabras de CA se discriminarán mejor que las palabras de CC. (hipótesis de la doble ruta). Esperamos que los sujetos sean capaces de discriminar mejor las palabras de CA que las palabras de CC, y además encontrar diferencias en los tiempos de reacción (TR). En otras palabras, en las tareas de memoria de palabras de CA se discrimina de forma cualitativamente distinta a como ocurre en memoria de palabras de CC.

2 MÉTODO

Participantes

36 estudiantes de psicología de entre 19 y 25 años, que cursaban asignaturas del área de Metodología de las Ciencias del Comportamiento (Facultad de Psicología de la Universidad de Málaga). La muestra estaba formada por hombres y mujeres.

Material

Las palabras fueron presentadas mediante el programa E-prime v1.1 (Schneider, Eshman y Zucolotto, 2002). El registro se realizó con un equipo de potenciales evocados *QuickAmp* de 72 canales y el programa de registro *Brainvision Recorder* v1. Se utilizaron 11 canales: 3 frontales (F3, FZ, F4), 3 centrales (C3, CZ, C4), 3 parietales (P3, PZ, P4), y 2 occipitales (O1, O2).

Procedimiento

Se realizó un registro de 20 señales y 20 situaciones de “ruido” para 20 palabras de CC y 20 palabras de CA. Las palabras aparecían en pantalla cada 2 segundos. El participante respondía mediante respuestas SI-NO en caso de que la palabra perteneciera o no a la lista original. Se midieron los TRs de todas las respuestas para cada una de las palabras. Para su registro utilizamos el programa *Brainvision Recorder*. Posteriormente fueron tratados y segmentados mediante el programa *Brainvision Analyzer*.

3 RESULTADOS

Antes de analizar los datos se realizó una tabla resumen (ver tabla 1) para comprobar el recuento de respuestas de los sujetos y los tiempos de reacción medio para cada tipo de respuesta.

Tabla 1: Recuento de respuestas obtenidas

		A	RI	FA	RC
CA	Rec.	261	83	53	258
	TR	1058	1214	915	1146
CC	Rec.	259	56	101	199
	TR	1092	1214	1220	1289

Como podemos comprobar, el número de aciertos y rechazos correctos es similar en ambas condiciones, el número de rechazos incorrectos es algo menor en las palabras de clase cerrada, pero encontramos que hay el doble de falsas alarmas en las palabras de clase cerrada, y además los tiempos de reacción de estas es mayor que en las de clase abierta.

En vista de estos datos, nos proponemos a ajustar un modelo de regresión de Poisson utilizando el programa “R” (Robert Gentleman Ihaka y Gentleman 1996) de libre distribución.

Tras el proceso de modelado, en el que resultaban también ajustados los modelos con los efectos principales de las variables explicativas, se llegó al ajuste de un modelo de Regresión de Poisson que incluye los efectos principales de Clase de Palabra (*C*) y Tiempo de Reacción (*TR*), y la interacción entre *C* y *TR*. ($\Delta D=15,94$ y $p<0,001$) (ver figura 1).

```

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.9992  -1.1329  -0.3085   0.4247   3.0138

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  -0.5564585  0.4844075  -1.149 0.250663
CA.CC[T.CC]   3.1237656  0.7076816   4.414 1.01e-05 ***
TR            0.0014839  0.0003733   3.975 7.03e-05 ***
CA.CC[T.CC]:TR -0.0021852  0.0005680  -3.847 0.000119 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

Null deviance: 114.57  on 35  degrees of freedom
Residual deviance: 74.39  on 32  degrees of freedom
    
```

Figura 1: Resultados del modelo ajustado con R

En la salida del programa R, el modelo que mejor hemos ajustado es un modelo que incluye los efectos principales de las variables explicativas, y la interacción entre dos variables explicativas (ver tabla 2 y figura 2). Obteniéndose un *pseudo-R*² de 0,54.

Tabla 2: Resumen del modelo ajustado

Términos	b	EE(b)	gl	z	P	exp(b)	IC95% exp(β)
Constante	-0,556	0,484	1	-1,149	0,250	0,573	0,199 a 1,344
(M)Memoria	3,123	0,707	1	4,414	<0,001	22,714	5,859 a 95,297
TR	0,001	0,000	1	3,975	<0,002	1,001	1,000 a 1,002
M*TR	-0,002	0,000	1	-3,847	<0,003	0,998	0,997 a 0,999

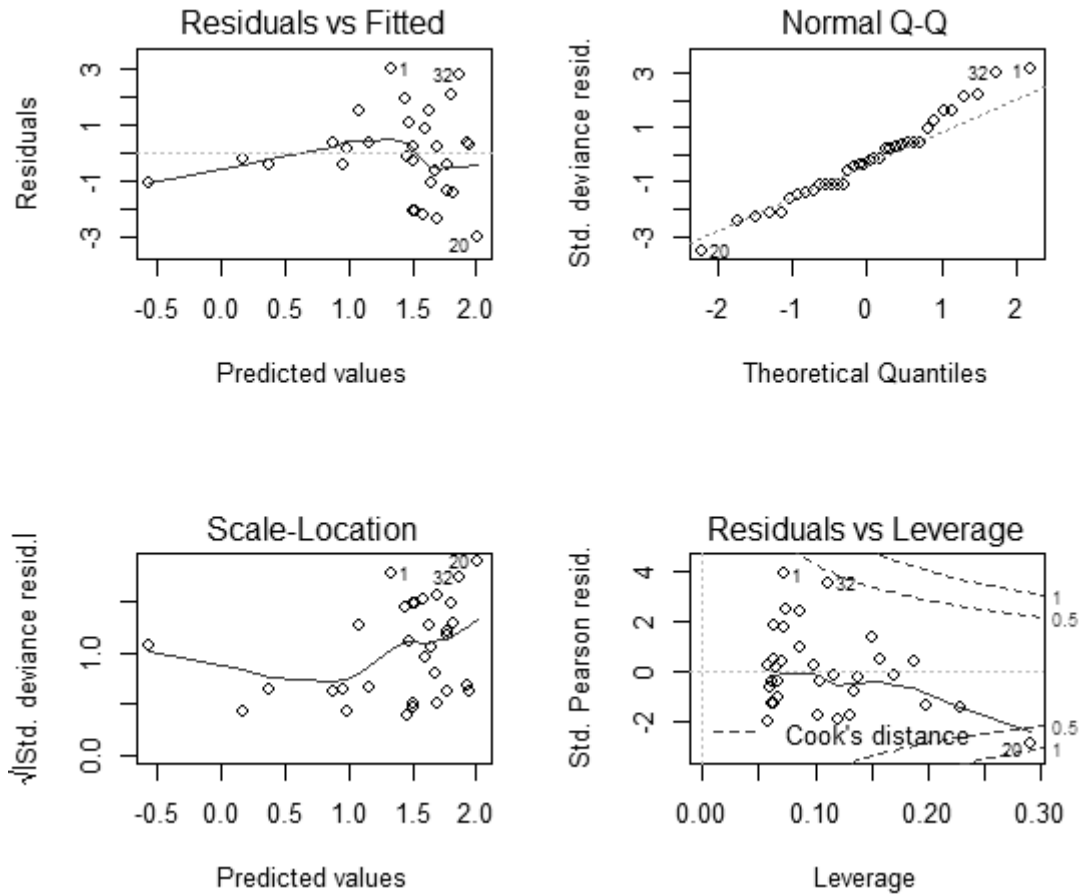


Figura 2: Gráficos diagnósticos del modelo ajustado

El modelo obtenido pone el énfasis en el la clase de palabras utilizadas y sus diferencias en el TR

4 DISCUSIÓN

A tenor de los resultados obtenidos, hemos comprobado que existen dos procesos diferenciados en función de la clase de palabras que se presenta a los sujetos. La interacción $C:T$ donde el tipo de palabra es el efecto principal que intereractúa con el tiempo de exposición del estímulo.

Las palabras de clase abierta se reconocen más rápidamente (ya que tienen un TR menor), y se cometen menos falsas alarmas que con las palabras de clase cerrada (un 50% menos). No hay diferencias entre CA y CC cuando se trata de aciertos y rechazos incorrectos.

Por otra parte, como hemos utilizado el equipo de registro de potenciales evocados para el registro del los TR, se obtuvo al igual que en el experimento anterior los potenciales evocados y las imágenes de activación eléctrica de la corteza cerebral (ver figura 3) donde podemos comprobar que tanto el voltaje como la localización en la corteza de los sujetos al realizar la tarea con palabras de CA son distintos a los de CC. Además, existen diferencias individuales que pueden ser estudiadas por si hubiese algún tipo de patrón asociado a variables psicológicas. Estos resultados son consistentes con los obtenidos por Henson, Hornberger y Rugg, (2005) y Hoppstädter, Baeuchl, Diener, *et al.*, (2015).



Figura 3: Gráfico de actividad eléctrica de la corteza cerebral al realizar las dos tareas.

En este experimento, de nuevo se asume un modelo cognitivo como un modelo más eficiente en memoria de palabras con significado frente al conductista, ya que los

sujetos reconocen mejor y más rápido las palabras de CA (que tienen significado) que las de CC (que no tienen significado) (ver anexo 1). Como línea futura de investigación cabe citar la necesidad de estudio de las características de los registros de la figura 3.

Anexo 1 (Experimento 3)

(Pelegrina, Emberley, et al., 2011)

Palabras de clase abierta			Palabras de clase cerrada		
villa	edad	lado	tuya	donde	algo
toro	falta	agua	acá	según	usted
regla	grupo	hora	mía	fuera	ahora
santa	joven	amor	ayer	otra	eso
odio	campo	tarde	abajo	cual	ello
torre	autor	forma	acaso	allí	bien
red	duda	caso	jamás	aún	así
rosa	arte	fin	cerca	tal	desde
labor	fondo	año	esta	aquí	tan
nota	aire	cosa	cuya	mucho	otro
ruido	cara	casa	lejos	antes	solo
brazo	gente	día	allá	tanto	como
carta	hijo		junto	poco	
color	calle		quizás	otro	

DISCUSIÓN GENERAL

Defiende tu derecho a pensar, incluso pensar de manera errónea,
es mejor que no pensar. Hipatia de Alejandría.

Un modelo estadístico es una ecuación que reproduce los fenómenos que observamos de la forma más exacta posible. Para ello tiene en cuenta los datos suministrados y la influencia que el azar tiene en estas observaciones, además, el modelo es diferente cada vez que se modifica la información. Estos modelos pueden aplicarse en el ámbito de la psicología, ya que el modelo matemático, a diferencia de la teoría, presenta una estrecha vinculación con los datos empíricos, y de este modo, el modelo se convierte en una herramienta teórica de fácil manejo que permite la introducción de ajustes a fin de describir adecuadamente los datos experimentales. La psicología matemática hoy en día está ampliamente extendida y se están desarrollando aplicaciones de modelado estadístico y simulación. Debemos señalar también, que la Teoría de Detección de Señales y los modelos expuestos en este trabajo sobre memoria de reconocimiento han conseguido una formalización bien fundamentada aplicada en diversos contextos.

La base de gran parte de las pruebas estadísticas utilizadas en psicología se encuentran en el Modelo Lineal Clásico (MLC). Este modelo supone que su estructura refleja los elementos explicativos de un fenómeno por medio de relaciones funcionales probabilísticas entre variables, el cual hemos utilizado en el segundo experimento de este trabajo desde la perspectiva del modelado estadístico, donde se ha ajustado un modelo lineal múltiple en el que hemos encontrado diferencias en la latencia del P300 en función del tipo de memoria y la presentación de una señal o un ruido. Además, aparecen la interacción entre el tipo de memoria y la señal-ruido. El más reciente, MLG (Nelder y Wedderburn, 1972) es una extensión del MLC. Este modelo ha supuesto un gran cambio en la estadística, siendo una solución válida para modelos de dependencia con datos categóricos y cuantitativos conjuntamente.

De forma complementaria con las habituales pruebas de contraste de hipótesis y las pruebas de significación, el modelado estadístico, que empieza a despuntar a partir de los años 60, se centra en la estimación de parámetros y el ajuste de modelos de probabilidad a los datos empíricos (Ato y López, 1996). Esto nos ha permitido en nuestro trabajo integrar en el modelo todos los elementos necesarios para el análisis: significación estadística y la estructura de la relación entre las variables. La APA en 2013 publicó la segunda edición de “Beyond Significance Testing: Reforming Data Analysis Methods in Behavioral Research” donde insta a los investigadores a utilizar las nuevas técnicas estadísticas que están evolucionando y apareciendo desde mediados del siglo XX, para solventar las limitaciones de los test estadísticos clásicos.

Los modelos que surgen del modelado estadístico tratan de explicar la variación de una respuesta a partir de la relación conjunta de dos fuentes de variabilidad, una de carácter determinista y otra aleatoria (Judd y McClelland, 1989). En esto es semejante el MLC, ya que el modelo trata de representar lo mejor posible los datos, reduciendo el error al máximo. En el caso del primer experimento se ha ajustado un modelo loglineal en el que se ha encontrado asociación entre la escala de estimación y la respuesta de los sujetos y también entre la señal-ruido y la respuesta de los sujetos. Esto ha sido mostrado en los experimentos anteriores, considerando los modelos de respuesta como elemento a modelar y no simplemente los resultados de las respuestas, siendo esto último una aportación del MLG y del modelado estadístico.

El modelado estadístico consiste, por tanto, en la construcción, formulación y ajuste de modelos a los datos empíricos respondiendo a tres criterios: la bondad de ajuste, la parsimonia y la integración teórica. Dicho de otro modo, buscamos el modelo

que explique la variable de respuesta con el menor error y con el menor número de variables explicativas. Esto se realizará siguiendo las etapas expuestas en este trabajo: especificar el modelo teórico, estimar los parámetros del modelo, seleccionar el modelo, evaluarlo e interpretarlo. Podemos aceptar o no el modelo, en este caso reiniciaríamos el proceso.

En efecto, en el tercer experimento se ha ajustado un modelo de Poisson Múltiple, en el que hemos encontrado diferencias en las respuestas de los sujetos en función del tipo de palabra que se les ha presentado, la señal-ruido, además se ha comprobado que los tiempos de reacción son diferentes en función de tipo de palabra. Esto nos ha permitido modelar el estímulo y su probabilidad, de acuerdo con la estadística Bayesiana que la TDS incorpora en las probabilidades *a priori*. Por consiguiente, desde el modelado aquí propuesto se presupone en cierto modo la consideración de una “memoria estadística”

Para la estimación de los parámetros del modelo, debido a su complejidad es necesario utilizar un software estadístico que nos permita realizar el modelado estadístico. Aunque la cantidad de programas estadísticos es bastante amplia (*SAS*, *SPSS*, *STATA*, *StatGraphics*, *MATLAB*, etc.), en modelado estadístico se ha utilizado mucho el programa *GLIM* (Nelder, 1974), su última versión *GLIM4* ha sido ampliamente utilizado. Hoy en día tenemos a nuestro alcance varios programas estadísticos gratuitos que nos permiten realizar modelado estadístico, uno de ellos es *lem* (Vermunt, 1997), de sencillo manejo y muy apto para la enseñanza del modelado. Pero el programa que hoy en día tiene la mayor potencia de cálculo y nos proporciona más información sobre nuestros modelos es *R* (Ihaka y Gentleman, 1996). Actualmente

tiene 7469 paquetes estadísticos que nos permiten realizar casi cualquier prueba estadística conocida, además cuenta con abundante documentación y tutoriales, para realizar nuestros cálculos. Con *R* se trabaja con decisiones y pasos sucesivos, de manera similar a lo que hacemos con el modelado estadístico: construimos modelos, calculamos la desviación para evaluar el modelo y entonces, aceptamos inicialmente o lo rechazamos.

En definitiva, como hemos visto anteriormente, se han aplicado técnicas de modelado estadístico en el contexto de la memoria de reconocimiento, con el objetivo de comprobar su utilidad analítica y además conocer si proporcionan información adicional a las técnicas tradicionalmente utilizadas en este campo, como los índices de discriminación de la TDS. Hasta nuestros trabajos anteriores en este sentido no hemos constatado modelos cognitivos ni de la neurociencia en este contexto, aunque el modelado es muy importante en las diferentes ingenierías, lo cual permite prever un desarrollo en la neurociencia.

En resumen, el modelado estadístico nos permite obtener el modelo mejor ajustado de los datos, utilizando variables cualitativas y cuantitativas, superando limitaciones de los métodos de análisis más extendidos como la independencia, la homogeneidad y la distribución de los datos (Agresti, 2002). Pero también, la posibilidad de modelar los datos obtenidos, y no solo, la obtención de diferencias significativas o no entre las distintas condiciones experimentales.

Así pues, los datos obtenidos en nuestras investigaciones nos permiten realizar un “análisis adaptado a los datos, los estímulos y las condiciones del diseño”, aspecto a

tener muy en cuenta, debido a la naturaleza de los datos obtenidos, ya que, a menudo trabajamos con variables que no se distribuyen según los supuestos, condiciones de aplicación o requisitos de la estadística clásica. Este es el caso, entre otros muchos, de los datos registrados en este trabajo en particular y en neurociencia en general, ya que, a pesar de poder realizar medias psicofísicas exhaustivas, también deben utilizar de forma conjunta y estructural variables categóricas en sus investigaciones.

Como ejemplo cabe señalar que otros autores ya han realizado intentos en este sentido. Así, en el primer experimento se planteó la posibilidad de utilizar estas técnicas de modelado en memoria de reconocimiento utilizando los índices de la detección de señales, que son medidas contrastadas para la medición de las respuestas de los sujetos en este tipo de condiciones. DeCarlo (1998) establece una relación entre los MLG y la TDS a partir de la regresión logística como modelo de distribución de los datos y proporciona un punto de vista complementario entre los modelos de regresión y detección de señales.

Por ejemplo, los resultados de ambos procedimientos apuntan en la misma dirección, el índice de sensibilidad es bajo y el modelado indica lo mismo, de hecho, en la condición IRL los modelos obtienen un mejor ajuste que en la condición IRI, que obtiene una sensibilidad menor. De hecho, para cada una de las condiciones se han ajustado modelos iguales. Pero, además, estos modelos nos proporcionan información adicional a los índices, ya que en ambas condiciones aparece una relación entre la seguridad de la respuesta de los sujetos y el contestar si o no a los estímulos (palabras). Así, los sujetos están más seguros de sus respuestas cuando responden si, tanto en los aciertos como en las falsas alarmas.

Como ya hemos indicado estos resultados indican la existencia de un “efecto general de escala” asociado a la respuesta de los sujetos.

Una vez que hemos comprobado que la aplicación del modelado estadístico supone la aportación de información complementaria a los índices de la TDS, el siguiente paso fue comprobar su eficacia en un nuevo experimento donde utilizamos palabras en vez de frases para comprobar las diferencias entre la memoria “implícita” y “explícita” y añadimos el registro electrofisiológico de los potenciales evocados relacionados con eventos (ERP). La aplicación de modelos matemáticos en potenciales evocados no es nuevo (Hunt, 1985) y aporta una medida psicofisiológica muy estable (Duncan-Johnson, 1979, Kutas, *et al.*, 1977). En esta ocasión encontramos que los sujetos respondían de forma diferenciada a los estímulos operacionalizados como memoria “implícita” y “explícita”. Lo que concuerda con lo comprobado por Chenchille, *et al.*, (1999).

Finalmente, y en relación con la neurociencia el modelo ajustado nos muestra que el potencial P300 tiene menor latencia cuando se presenta la señal en la condición de memoria “implícita”, y mayor cuando se presenta el ruido. En cambio, con la condición de memoria “explícita” la latencia del P300 es mayor cuando se presenta la señal que cuando se presenta el ruido. También aquí hemos observado las posibilidades amplias del modelo en el ámbito de las diferencias individuales.

Estos resultados nos impulsan a seguir esta línea de investigación aplicando y extendiendo el modelado estadístico para completar el modelo de memoria correspondiente. En este contexto, el modelado estadístico tiene una función primordial

al ajustar modelos cualitativamente distintos. Por consiguiente, tenemos nuestras dudas de cómo denominar a los modelos resultantes, pero tenemos certeza, y la podemos obtener, sobre la presencia de modelos distintos en un sentido cualitativo y cuantitativo en memoria.

En cuanto al tipo de palabras, de nuevo el uso del modelado estadístico nos ha proporcionado una visión amplia sobre los datos del experimento, permitiéndonos extraer el máximo de información posible. En efecto, se quería comprobar un modelo donde incluimos las palabras de clase abierta (CA) y clase cerrada (CC) junto a los tiempos de reacción. Los resultados obtenidos por el modelo ajustado son bastante satisfactorios, donde encontramos que los sujetos reconocen más rápido las palabras de CA, y discriminan entre la señal y el ruido en los dos tipos de palabras, y con las palabras de CA se cometen menos falsas alarmas. Además, en ambos tipos de palabras hay muchos aciertos y pocos rechazos incorrectos. Hallazgos similares a los obtenidos por Curran, *et al.*, 2001, de Zubizaray, *et al.*, 2005, Heathcote, 2006. El tipo de clasificación de las palabras según sus características nos proporciona nuevos elementos sobre su representación dentro del modelado.

En la introducción de esta tesis se planteó como objetivo principal aplicar el modelado estadístico en tareas de memoria de reconocimiento, y los modelos estadísticos que se han obtenido en este trabajo han respondido a nuestros objetivos. Sin embargo, no debemos detenernos aquí, los conocimientos en técnicas estadísticas se van renovando y actualizando. Por consiguiente, las perspectivas futuras de estudios sobre procesos cognitivos implican optar por las más adecuadas técnicas estadísticas que resulten en análisis más robustos que describan los datos y que cuantifiquen estos

procesos. Y es en particular aquí donde la estadística juega un rol importante para hacer frente a la naturaleza de los procesos cognitivos.

En definitiva, creemos que el modelado estadístico es una técnica muy apropiada para la investigación en procesos cognitivos y en la futura neurociencia., porque supone un complemento muy útil a las medidas de detección de señales, psicofísicas y neuropsicológicas.

Para investigaciones futuras tenemos un amplio campo de posibilidades abierto, ya que el registro de potenciales evocados que se ha realizado para este trabajo incluye la latencia de potenciales (podemos utilizar también lo potenciales sensitivos, el N400, P500, etc.), la forma de la onda (modelos regresión para estimar la forma de la onda en Smith y Kutas, 2015), los diferentes electrodos utilizados y su ubicación, el tiempo de reacción (Jones, Curran, *et al.*, 2013), el voltaje (mapeados de actividad eléctrica cerebral en Duffy, 2004), el tipo de onda cerebral predominante (alfa, beta, theta y delta), las imágenes de activación eléctrica cortical (que nos permiten observar patrones de activación cortical en Hoppstädter, *et al.*, 2015), y mediante el módulo *LORETA* que incluye el programa Analyzer podemos obtener imágenes de resonancia magnética funcional de baja resolución que nos permiten conocer aproximadamente el origen de la actividad eléctrica cognitiva (que pueden ser analizadas mediante análisis estadístico espacial). Y también, desde la perspectiva de la TDS podemos modelar con curvas ROC.

Además, en el estudio de los procesos cognitivos, Sarle, (1994), propone que se puede establecer una similitud entre modelos estadísticos y modelos de redes

neuronales. Y nos muestra la equivalencia entre algunos modelos estadísticos y modelos de redes neuronales como, por ejemplo, la regresión lineal múltiple sería equivalente al modelo de red neuronal “perceptrón simple con función lineal”, la regresión logística es equivalente al modelo “perceptrón simple con función logística”, la regresión no lineal múltiple es equivalente al “perceptrón multicapa con función lineal en la salida”. De esta manera se pone de manifiesto que la mayoría de redes neuronales aplicadas al análisis de datos son similares y, en algunos casos, equivalentes a modelos estadísticos ya conocidos.

Además, y en relación con lo expuesto en este trabajo, un perceptrón simple puede ser considerado como un modelo lineal generalizado, debido a la equivalencia entre el concepto de función de enlace en un modelo lineal generalizado y la función de activación de la neurona de salida en un perceptrón. La diferencia importante entre ambos modelos radica en el método de estimación de los coeficientes utilizados para minimizar la función de error. Mientras el perceptrón normalmente estima los parámetros del modelo mediante el criterio de mínimos cuadrados, el modelo lineal generalizado ajusta el modelo mediante el método de máxima verosimilitud.

Esto nos muestra que la aparente desconexión entre las diferentes técnicas de análisis de datos no lo es, sino que son diferentes enfoques y técnicas que nos permiten analizar nuestros datos, y que utilizadas en conjunto nos proporcionan información que se va complementando y nos permite un conocimiento más exhaustivo sobre estos complejos procesos cognitivos. Aunque debemos tener en cuenta que la utilización de estas técnicas sin un substrato teórico que nos guie, puede llevarnos a la formulación de modelos estadísticamente “correctos”, pero no aplicables o no interpretables. En este

trabajo hemos tratado de que se encuentren la estadística y la teoría psicológica para contribuir a la unificación de los modelos de forma parsimoniosa, pero en toda su extensión. En resumen, el modelado une en una misma estructura el modelo sustantivo y el modelo estadístico-matemático, optando por el modelo más parsimonioso y considerando conjuntamente todas las variables plausibles del modelo.

CONCLUSIÓN

Ninguna investigación humana puede ser denominada ciencia si no pasa a través de pruebas matemáticas. Leonardo Da Vinci.

En el presente trabajo se han expuesto diferentes modelos estadísticos aplicados a datos empíricos, para obtener toda la información posible sobre ellos. El objetivo principal de este trabajo ha consistido, por una parte, en aumentar nuestro conocimiento acerca de los modelos asociativos y de los modelos cognitivos, en este caso, la memoria de reconocimiento, y por otra parte, presentar el modelado estadístico como una técnica de análisis complementaria a los análisis clásicos que se suelen aplicar. Además, esto puede llevarnos a un diálogo (o posible debate) sobre la idoneidad de estos modelos y su aplicación a otros ámbitos.

Las respuestas de los sujetos en pruebas de memoria pueden ser dicotómicas (aciertos y errores), politómicas (escalas subjetivas de estimación con una gran variedad de posibilidades) y a partir de ellas y de acuerdo con los estímulos, se elaboran tablas de contingencia complejas, donde los modelos que hemos venido tratando encajan perfectamente porque se adaptan de forma literal a los datos, sean los que sean: variables independientes o de estímulo, variables de respuesta, variables de confundido, etc.

En este caso, la aplicación del modelado nos ha proporcionado información sobre los modelos matemáticos de contenido sustantivo, en todos los datos en los que se han aplicado. Esto nos ha permitido obtener información adicional a la de los métodos de análisis tradicionales.

En el primer experimento hemos comprobado que existe asociación entre la seguridad de la respuesta de los sujetos y el tipo de respuesta que daban asumiendo un modelo cognitivo de reconocimiento. En el segundo, hemos comprobado que pueden

existir dos procesos diferenciados de memoria en función de los estímulos que se presentan, apoyándonos en registros electrofisiológicos lo que apoyaría un modelo asociacionista de la memoria. En tercer lugar, hemos comprobado que los sujetos son capaces de discriminar estímulos diferentes que han memorizado previamente, diferenciándose esta discriminación en el tiempo de reacción de la respuesta, lo que nos lleva a pensar que el sujeto utiliza un modelo cognitivo de reconocimiento.

Podemos concluir que nuestra investigación nos ha proporcionado evidencias e información formalizada matemáticamente sobre la relación entre las variables y condiciones cada investigación sobre la memoria de reconocimiento, desde diferentes perspectivas, representadas por los experimentos en los que hemos tratado de reflejar de la mejor manera posible las características de este complejo proceso cognitivo.

Finalmente indicar que es posible que el modelado se incluya en el futuro en una teoría estadístico-matemática más amplia en torno a la actual teoría de conjuntos que tendría en cuenta los elementos del modelado y su desarrollo futuro. Incorporando conjuntos amplios de variables y redes: por ejemplo, mediante los conceptos de retículo, red o lattice entre otros. Pero estos conceptos se nos escapan a los propósitos del presente trabajo. No obstante, y para terminar enumeramos una sinopsis de los elementos más relevantes de nuestro trabajo actual y sus posibles extensiones en el futuro.

Elementos considerados en el mismo modelado:

Historia previa del estímulo.

Estímulos con sus probabilidades a priori.

Condiciones experimentales.

Diseño.

Diseño restringido.

Plan de investigación.

Respuestas.

Constructos operacionalizados mediante variables.

Grupos de variables a criterio de la teoría.

Modelos de acuerdo con los objetivos de la investigación.

Elegir el modelo estadístico mejor ajustado y más parsimonioso.

Elementos a considerar en el futuro:

Simulación basada en las estructuras anteriores

Modelos de redes y neuronales.

REFERENCIAS

Un científico es un hombre tan endeble y humano como cualquiera; sin embargo, la búsqueda científica puede ennoblecerle, incluso en contra de su voluntad. Isaac Asimov.

- Abrahamson, I.G. y Levitt, H. (1969). Statistical analysis of data from experiments in human signal detection. *Journal of Mathematical Psychology*, 6, 391-417.
- Achen, C.H. (1982). *Interpreting and Using Regression*. Newbury Park, CA: Sage Publications.
- Aggarwal, C.C. (2013). *Outlier Analysis*. New York: Springer-Verlag.
- Aggleton, J.P. y Brown, M.W. (1999). Episodic memory, amnesia, and the hippocampal-anterior thalamic axis. *Behavioral and Brain Sciences*, 22(3), 425-444.
- Aggleton, J. P., McMackin, D., Carpenter, K., *et al.* (2000). Differential cognitive effects of colloid cysts in the third ventricle that spare or compromise the fornix. *Brain*, 123(1), 800-815.
- Agresti, A. (1984). *Analysis of Ordinal Categorical Data*. New York: J. Wiley & Sons.
- Agresti, A. (2007). *An Introduction to Categorical Data Analysis (2nd Edition)*. Hoboken, NJ: Wiley-Interscience.
- Agresti, A. (1990). *Categorical Data Analysis*. Hoboken, NJ: Wiley-Interscience.
- Agresti, A. (2013). *Categorical Data Analysis (3rd Edition)*. Hoboken, NJ: Wiley-Interscience.
- Agresti, A. (1999). On logit confidence intervals for the odds ratio with small samples. *Biometrics*, 55, 597-602.
- Aitkin, M. (1980). A note on the selection of log-linear models. *Biometrics*, 36(1), 173-178.
- Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of Econometrics* , 16, 3-14.
- Algom, D. (1992). *Advances in Psychology Vol.92: Psychophysical Approaches to Cognition*. Amsterdam: North-Holland.

- Allan, L.G., Siegel, S.D. y Tangen, J.M. (2005). A signal detection analysis of contingency data. *Learning and Behavior*, 33, 250-263.
- Altman, D.G. y Bland, J.M. (1994). Diagnostic tests 1: sensitivity and specificity. *British Medical Journal*, 308, 1552.
- Andersen, E.B. (1997). *Introduction to the Statistical Analysis of Categorical Data*. Berlin: Springer-Verlag.
- Arbel, Y., Spencer, K.M. y Donchin, E. (2011). The N400 and the P300 are not all that independent. *Psychophysiology*, 48, 861-875.
- Arminger, G., Clogg, C.C. y Sobel, M.E. (1995). *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. New York: Springer-Verlag.
- Arnau, J. (1986). Psicología experimental cognitiva: modelos básicos de procesamiento de la información. *Anuario de Psicología*, 35, No.2, 5-16.
- Arnau, J. (1977). Utilización de modelos matemáticos en psicología. *Anuario de Psicología*, 17, 5-17.
- Atkinson, R.C. y Juola, J.F. (1974). Search and decision processes in recognition memory. En D.H. Krantz, R.C. Atkinson, R.D. Luce, y P. Suppes (Eds.). *Contemporary Developments in Mathematical Psychology: Vol. 1. Learning, Memory & Thinking*. San Francisco: Freeman.
- Atkinson, R.C., Hertmann, D.J. y Wescourt, K. T. (1974). Search processes in recognition memory. En Solso (Ed.), *Theories in Cognitive Psychology: The Loyola Symposium*. Potomac, MD: Erlbaum.
- Ato, M. y López, J.J. (1996). *Análisis Estadístico para Datos Categóricos*. Madrid: Síntesis.
- Ato, M., Losilla, J.M., Navarro, J.B., Palmer, A. y Rodrigo, M.F. (2005). *Análisis de Datos: Modelo Lineal Generalizado*. Tarrasa: CBS.

- Azen, R. y Walker, C.M. (2011). *Categorical Data Analysis for the Behavioral and Social Sciences*. New York: Routledge.
- Baddeley, A.D. (1992). Working memory: the interface between memory and cognition. *Journal of Cognitive Neuroscience*, 4(3), 281-288.
- Baddeley, A.D. (1982). *Psicología de la Memoria*. Barcelona: Debate.
- Ballesteros, S., Reales, J.M. y Manga, D. (1999). Memoria implícita y memoria explícita intramodal e intermodal: influencia de las modalidades elegidas y del tipo de estímulos. *Psicothema*, 11(4), 831-851.
- Balota, D.A. y Spieler, D.H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: beyond measures of central tendency. *Journal of Experimental Psychology*, 128(1), 32-55.
- Bassi, F., Hagenaars, J.A., Croon, M., & Vermunt, J.K. (2000). Estimating true changes when categorical panel data are affected by uncorrelated and correlated classification errors. *Sociological Methods & Research*, 29(2), 230-268.
- Belsley, D.A., Kuh, E. y Welsch, R.E. (2004). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. Hoboken, NJ: Wiley-Interscience.
- Berkson, J. (1944). Application of the logistic function to bio-assay. *Journal of the American Statistical Association*, 39(227), 357-365.
- Bernbach, H.A. y Kupchak, P.G. (1972). Recognition and recall in short-term memory. *Journal of Mathematical Psychology*, 9, 231-242.
- Berry, W.D. (1993). *Understanding Regression Assumptions*. Newbury Park, CA: Sage Publications.
- Bhat, S., Agarwal, A., Gray, A. y Vuduc, R. (2010). Toward interactive statistical modeling. *Procedia Computer Science*, 1, 1835-1844.

- Biassou, N., Obler, L.K., Nespoulous, J.L., Dordain, M. y Harris, K.S. (1997). Dual processing of open-and closed-class words. *Brain and Language*, 57, 360-373.
- Bishop, Y.M., Fienberg, S.E. y Holland, P.W. (1975). *Discrete Multivariable Analysis: Theory and Practice*. Cambridge, MA: The MIT Press.
- Bjork, E.L. & Bjork, R.A. (1996). *Memory: Handbook of Perception and Cognition (2nd Edition)*. San Diego, CA: Academic Press.
- Bozdogan, H. (2000). Akaike's Information Criterion and recent developments in information complexity. *Journal of Mathematical Psychology*, 44(1), 62-91.
- Bradley, R.A. y Terry, M.E. (1952). Rank analysis of incomplete block designs: I. the method of paired comparisons. *Biometrika*, 39(3-4), 324-345.
- Breiman, L. (2001). Statistical modeling: the two cultures. *Statistical Science*, 16(3), 199-231.
- Brown, M.B. y Benedetti, J.K. (1977). On the mean and variance of the tetrachoric correlation coefficient. *Psychometrika*, 42(3), 347-355.
- Calcagno, V. y de Mazancourt, C. (2010). glmulti: an R package for easy automated model selection with (Generalized) Linear Models. *Journal of Statistical Software*, 34(12), 1-29 .
- Caldwell, J.I. y Masson, M.E. (2001). Conscious and unconscious influences of memory for object location. *Memory & Cognition*, 29(2), 285-295.
- Callaway, E., Tueting, P. y Koslow, S.H. (1978). *Event-Related Brain Potentials in Man*. New York: Academic Press.
- Cameron, A.C. y Trivedi, P.K. (2013). *Regression Analysis of Count Data (2nd Edition)*. New York: Cambridge University Press.
- Carreiras, M., Álvarez, C.J. y de Vega, M. (1993). Syllabe frequency and visual word recognition in spanish. *Journal of Memory & Language*, 32, 766-780.

- Caussinus, H. (1965). Contribution à l'analyse statistique des tableaux de corrélation. *Annales de la Faculté des Sciences de l'Université de Toulouse*, 29, 77-183.
- Chatterjee, S. y Simonoff, J.S. (2013). *Handbook of Regression Analysis*. Hoboken, NJ: J. Wiley & Sons.
- Chechile, R.A., Sloboda, L.N. y Chamberland, J.R. (2012). Obtaining separate measures for implicit and explicit memory. *Journal of Mathematical Psychology*, 56(1), 35-53.
- Christensen, R. (1997). *Log-Linear Models and Logistic Regression (2nd Edition)*. New York: Springer-Verlag.
- Chwilla, D.J., Brown, C.M. y Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology*, 32, 274-285.
- Claeskens, G. y Hjort, N.L. (2008). *Model Selection and Model Averaging*. Cambridge: Cambridge University Press.
- Clogg, C.C. (1982). Some models for the analysis of association in multiway cross-classifications having ordered categories. *Journal of the American Statistical Association*, 77(380), 803-815.
- Clogg, C.C. y Shihadeh, E.S. (1994). *Statistical Models for Ordinal Variables*. Thousand Oaks, CA: Sage.
- Collett, D. (1991). *Modelling Binary Data*. London: Springer-Verlag.
- Cook, R.D. (1977). Detection of influential observation in linear regression. *Technometrics*, 19(1), 15-18.
- Cook, R.D. y Weisberg, S. (1999). *Applied Regression including Computing and Graphics*. New York: Wiley-Interscience.
- Cook, R.D. y Weisberg, S. (1982). *Residuals and Influence in Regression*. New York: Chapman & Hall.

- Cooms, C.H., Dawes, R.M. y Tversky, A. (1981). *Introducción a la Psicología Matemática*. Madrid: Alianza.
- Cordeiro, G.M. (2004). On Pearson's residuals in Generalized Linear Models. *Statistics & Probability Letters*, 66, 213-219.
- Cornfield, J. (1966). A bayesian test of some classical hypotheses: with applications to sequential clinical trials. *Journal of the American Statistical Association*, 61(315), 577-594.
- Cowan, N. (2005). *Working Memory Capacity*. New York : Psychology Press.
- Cowan, N., Baddeley, A.D., Elliott, E.M. y Norris, J. (2003). List composition and the word length effect in immediate recall: a comparison of localist and globalist assumptions. *Psychological Bulletin*, 10(1), 74-79.
- Cox, D.R. (1970). *The Ananlysis of Binary Data*. London: Methuen.
- Crawley, M.J. (2013). *The R Book (2nd Edition)*. Chichester: J. Wiley & Sons.
- Cronbach, L.J. (1987). Statistical tests for moderator variables: Flaws in analyses recently proposed. *Psychological Bulletin*, 102, 414-417.
- Cumming, G., Fidler, F, Leonard, M., *et al.* (2007). Statistical reform in psychology: is anything changing? *Psychological Science*, 18(3), 230-232.
- Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition*, 28(6), 923-938.
- Curran, T., Schacter, D.L., Johnson, M.K. y Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. *Journal of Cognitive Neuroscience*, 13(2), 201-216.
- Czado, C. (1997). On selecting parametric link transformation families in Generalized Linear Models. *Journal of Statistical Planning & Inference*, 61, 125-139.

- Davies, G.M. y Wright, D.B. (2010). *Current Issues in Applied Memory Research*. Hove: Psychology Press.
- de Zubicaray, G.I., McMahon, K.L., Eastburn, M.M., *et al.* (2005). fMRI evidence of word frequency and strength effects in recognition memory. *Cognitive Brain Research*, 24, 587-598.
- DeCarlo, L.T. (1998). Signal Detection Theory and Generalized Linear Models. *Psychological Methods*, 3(2), 186-205.
- DeMaris, A. (1992). *Logit Modeling: Practical Applications*. Newbury Park, CA: Sage Publications.
- Dere, E., Easton, A., Nadel, L. & Huston, J.P. (2008). *Handbook of Episodic Memory*. Amsterdam: Elsevier Science.
- Ditraglia, G.M. & Polich, J. (1991). P300 and the introverted-extraverted personality types. *Psychophysiology*, 28(2), 177-184.
- Dobson, A.J. (2002). *An Introduction to Generalized Linear Models (2nd Edition)*. Boca Raton, FL: Chapman & Hall.
- Dobson, A.J. (1983). *Introduction to Statistical Modelling*. London: Chapman & Hall.
- Donchin, E., Karis, D., Bashore, T.R., Coles, M.G. y Gratton, G. (1986). Cognitive psychophysiology and human information processing. En M.G. Coles, E. Donchin, y S.W. Porges (Eds.), *Psychophysiology: Systems, Processes, and Applications*. New York: Guilford Press.
- Draper, N. y Smith, H. (1998). *Applied Regression Analysis (3rd Edition)*. New York: J. Wiley & Sons.
- Duncan-Johnson, C.C. y Donchin, E. (1979). The time constant in P300 recording. *Psychophysiology*, 16(1), 53-55.

- Dunn, O.J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56(293), 54-64
- Earhard, B. (1982). Determinants of the word-frequency effect in recognition memory. *Memory and Cognition*, 10(2), 115-124.
- Eimer, M. (1998). Methodological issues in Event-related Brain Potential research. *Behavior Research Methods*, 30(1), 3-7.
- Egan, J.P. (1958). *Recognition memory and the operating characteristic*. Indiana University: Hearing and Communication Laboratory. Technical Note AFCRC-TN-58-51
- Eliason, S.R. (1993). *Maximum Likelihood Estimation: Logic and Practice*. Newbury Park, CA: Sage Publications.
- Elosua, P. (2009). ¿Existe vida más allá del SPSS? Descubre R. *Psicothema*, 21(4), 652-655.
- Everitt, B.S. y Howell D.C. (2005). *Encyclopaedia of Statistics in Behavioural Science (Vols.1-4)*. Chichester: J. Wiley & Sons.
- Fahrmeir, L. y Tutz, G. (2001). *Multivariate Statistical Modelling Based on Generalized Linear Models (2nd Edition)*. New York: Springer-Verlag.
- Fahrmeir, L., Kneib, T., Lang, S. y Marx, B. (2013). *Regression: Models, Methods and Applications*. New York: Springer-Verlag.
- Fienberg, S.E. y Rinaldo, A. (2007). Three centuries of categorical data analysis: log-linear models and maximum likelihood estimation. *Journal of Statistical Planning & Inference*, 137, 3430-3445.
- Fienberg, S.E. y Rinaldo, A. (2012). Maximum likelihood estimation in log-linear models. *The Annals of Statistics*, 40(2), 996-1023.

- Findley, D.F. (1991). Counterexamples to parsimony and BIC. *Annals of the Institute of Statistical Mathematics*, 43, 509-514.
- Forster, M.R. (2000). Key concepts in model selection: performance and generalizability. *Journal of Mathematical Psychology*, 44, 205-231.
- Fox, J. (2000). *Multiple and Generalized Nonparametric Regression*. Thousand Oaks, CA: Sage Publications.
- Fox, J. (1991). *Regression Diagnostics: An Introduction*. Newbury Park, CA: Sage Publications
- Freeman, E., Heathcote, A., Chalmers, K. y Hockley, W. (2010). Item effects in recognition memory for words. *Journal of Memory & Language*, 62, 1-18.
- Frees, E.W. (2004). *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*. Cambridge: Cambridge University Press.
- Freund, R.J., Wilson, W.J. y Sa, P. (2006). *Regression Analysis: Statistical Modeling of a Response Variable (2nd Edition)*. Burlington, MA: Academic Press.
- Gardener, M. (2013). *The Essential R Reference*. Indianapolis, IN: J. Wiley & Sons.
- Gardiner, J.M. y Richardson-Klavehn, A. (2000). Remembering and knowing. En E. Tulving, E. Ferguson e I.M. Craik (Eds.). *The Oxford Handbook of Memory*. New York: Oxford Univ. Press.
- Gill, J. (2001). *Generalized Linear Models: A Unified Approach*. Thousand Oaks, CA: Sage Publications.
- Glonek, G.F. y McCullagh, P. (1995). Multivariate logistic models. *Journal of the Royal Statistical Society*. 57(3), 533-546.
- Goldstein, A., Spencer, K.M. y Donchin, E. (2002). The influence of stimulus deviance and novelty on the P300 and novelty P3. *Psychophysiology*, 39, 781-790.

- Goodman, L.A. (1971). The analysis of multidimensional contingency tables: stepwise procedures and direct estimation methods for building models for multiple classifications. *Technometrics*, 13(1), 33-61.
- Graf, P. y Mandler, G. (1984). Activation makes words more accessible, but not necessarily more retrievable. *Journal of Verbal Learning and Verbal Behavior*, 23(5), 553-568.
- Graybill, F.A. e Iyer, H.K. (1994). *Regression Analysis: Concepts and Applications*. Belmont, CA: Duxbury Press.
- Green, D.M. & Swets, J.A. (1988). *Signal Detection Theory and Psychophysics*. Los Altos, CA: Peninsula Publishing.
- Grimnes, S. & Martinsen, O.G. (2008). *Bioimpedance and Bioelectricity Basics (2nd Edition)*. London: Academic Press.
- Grünwald, P.D. (2000). Model selection based on Minimum Description Length. *Journal of Mathematical Psychology*, 44, 133-152.
- Gürtler, N. y Henze, N. (2000). Recent and classical goodness-of-fit tests for the Poisson distribution. *Journal of Statistical Planning & Inference*, 90, 207-225.
- Haberman, S.J. (1974). *The Analysis of Frequency Data*. Chicago: University of Chicago Press.
- Haight, F.A. (1967). *Handbook of the Poisson Distribution*. New York: J. Wiley & Sons.
- Haight, F.A. (1966). Some statistical problems in connection with word association data. *Journal of Mathematical Psychology*, 3, 217-233.
- Haist, F., Shimamura, A.P. y Squire, L.R. (1992). On the relationship between recall and recognition memory. *Journal of Experimental Psychology*, 18(4), 691-702.

- Hald, A. (2007). *A History of Parametric Statistical Inference from Bernoulli to Fisher 1713-1935*. New York: Springer-Verlag.
- Hald, A. (2003). *A History of Probability and Statistics and their Applications before 1750*. Hoboken, NJ: Wiley-Interscience.
- Handy, T.C. (2009). *Brain Signal Analysis: Advances in Neuroelectric and Neuromagnetic Methods*. Cambridge, MA: The MIT Press.
- Hardy, M.A. (1993). *Regression with Dummy Variables*. Newbury Park, CA: Sage Publications.
- Hasselmo, M.E. (2012). *How We Remember: Brain Mechanisms of Episodic Memory*. Cambridge, MA: The MIT Press.
- Heathcote, A., Ditton, E. y Mitchell, K. (2006). Word frequency and word likeness mirror effects in episodic recognition memory. *Memory and Cognition*, 34(4), 26-838.
- Heinze, H.J., Münte, T.F. y Mangun, G.R. (1994). *Cognitive Electrophysiology*. New York: Springer-Verlag.
- Heinze, H.J., Münte, T.F., Kutas, M., Butler, S.R., *et al.* (1999). Cognitive Event-related Potentials. *EEG and Clinical Neurophysiology*, 52, 91-95.
- Hillyard, S.A. y Kutas, M. (1983). Electrophysiology of cognitive processing. *Annual Review of Psychology*, 34, 33-61.
- Hillyard, S.A., Picton, T.W. y Regan D. (1978). Sensation, percption and attention En E. Callaway, Tueting, P. y Koslow, S. (eds). *Event-Related Brain Potentials in Man*. New York: Academic Press
- Hintzman, D.L., Caulton, D.A. y Levitin, D.J. (1998). Retrieval dynamics in recognition and list discrimination: Further evidence of separate processes of familiarity and recall. *Memory and Cognition*, 26(3), 449-462.

- Hirshman, E. y Master, S. (1997). Modeling the conscious correlates of recognition memory: Reflections on the remember-know paradigm. *Memory and Cognition*, 25(3), 345-351.
- Hirshman, E. y Palij, M. (1992). Rehearsal and the word frequency effect in recognition memory. *Journal of Memory & Language*, 31, 477-484.
- Hoaglin, D.C. y Welsch, R.E. (1978). The hat matrix in regression and ANOVA. *American Statistician*, 32(1), 17-22.
- Hoppstädter M., Baeuchl C., Diener C., Flor H. y Meyer P. (2015). Simultaneous EEG-fMRI reveals brain networks underlying recognition memory ERP old/new effects. *Neuroimage*. 116, 112-22.
- Hosmer, D.W. y Hjort, N.L. (2002). Goodness-of-fit processes for logistic regression: simulation results. *Statistics in Medicine*, 21, 2723-2738.
- Hosmer, D.W., Lemeshow, S. y Sturdivant, R.X. (2013). *Applied Logistic Regression (3rd Edition)*. Hoboken, NJ: J. Wiley & Sons.
- Hunt, E. (1985). Mathematical models of the Event-Related Potential. *Psychophysiology*, 22(4), 395-402.
- Huppert, F.A. y Piercy, M. (1978). The role of trace strength in recency and frequency judgments by amnesic and control subjects. *Quarterly Journal of Experimental Psychology*, 30(2), 347-354.
- Hurst, W. y Volpe, B.T. (1982). Temporal order judgments with amnesia. *Brain and Cognition*, 1(3), 294-306.
- Hutcheson, G.D. y Sofroniou, N. (1999). *The Multivariate Social Scientist: Introductory Statistics Using Generalized Linear Models*. London: Sage.
- Inoue, C. y Bellezza, F.S. (1998). The detection model of recognition using know and remember judgments. *Memory and Cognition*, 26(2), 299-308.

- Jaccard, J. (2001). *Interaction Effects in Logistic Regression*. Thousand Oaks, CA: Sage Publications.
- Jacoby, L.L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30(5), 513-541.
- Jacoby, L.L. y Kelley, C.M. (1992). Unconscious influences of memory: Dissociations and automaticity. En A. Milner y E.M. Rugg (Eds.). *The Neuropsychology of Consciousness*. San Diego: Academic Press.
- Jacoby, L.L., Kelley, C.M. y Dywan, J. (1989). Memory attributions. En H. Roediger III, E. Fergus e I.M. Craik (Eds.). *Varieties of Memory and Consciousness: Essays in Honour of Endel Tulving*. Hillsdale: Erlbaum.
- Jacoby, L.L., Toth, J.P. y Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: measuring recollection. *Journal of Experimental Psychology: General*, 122(2), 139-154.
- Johnson, R. (1993). On the neural generators of the P300 component of the Event-Related Potential. *Psychophysiology*, 30, 90-97.
- Jones, M., Curran, T., Mozer, M.C., & Wilder, M.H. (2013). Sequential effects in response time reveal learning mechanisms and event representations. *Psychological Review*, 120(3), 628-666.
- Judd, C.M. y McClelland, G.H. (1989). *Data Analysis: A Model-Comparison Approach*. Orlando, FL: Harcourt Brace Jovanovich.
- Judd, C.M., McClelland, G.H. y Ryan, C.S. (2008). *Data Analysis: A Model-Comparison Approach (2nd Edition)*. New York: Routledge.
- Kadane, J.B. y Lazar, N.A. (2004). Methods and criteria for model selection. *Journal of American Statistical Association*, 99, 279-290.

- Kam, C.D. & Franzese, R.J. (2007). *Modeling and Interpreting Interactive Hypotheses in Regression Analysis*. Ann Arbor, MI: Michigan University Press.
- Kappenman, E.S. y Luck, S.J. (2010). The effects of electrode impedance on data quality and statistical significance in ERP recordings. *Psychophysiology*, 47, 888-904.
- Katayama J y Polich J. (1998). Stimulus context determines P3a and P3b. *Psychophysiology*, 35, 23-33.
- Kelley, R. y Wixted, J.T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(3), 701-722.
- Kenny, D.A., Kashy, D.A., Mannetti, L., Pierro, A. y Livi, S. (2002). The statistical analysis of data from small groups. *Journal of Personality and Social Psychology*, 83(1), 126-137.
- Kinoshita, S. (1995). The word frequency effect in recognition memory versus repetition priming. *Memory and Cognition*, 23(5), 569-580.
- Kirk, R.E. (1982). *Experimental Design: Procedures for the Behavioral Sciences*. Belmont, CA: Brooks/Cole.
- Kleinbaum, D.G. (1994). *Logistic Regression: A Self-Learning Text*. New York: Springer-Verlag.
- Kleinbaum, D.G. y Klein, M. (2010). *Logistic Regression: A Self-learning Text (3rd Edition)*. New York: Springer-Verlag.
- Kleinbaum, D.G., Kupper, L.L. y Muller, K.E. (1988). *Applied Regression Analysis and other Multivariable Methods (2nd Edition)*. Belmont, CA: Duxbury Press.

- Klimesch, W., Doppelmayr, A., Yonelinas, A., *et al.* (2001). Theta synchronization during episodic retrieval: Neural correlates of conscious awareness. *Cognitive Brain Research*, 12(1), 33-38.
- Kline, R.B. (2013). *Beyond Significance Testing - Statistics Reform in the Behavioral Sciences (2nd Edition)*. Washington: APA.
- Knoke, D. & Burke, P.J. (1980). *Log-Linear Models*. Newbury Park, CA: Sage Publications.
- Knoblauch, K. y Maloney, L.T. (2012). *Modeling Psychophysical Data in R*. New York: Springer-Verlag.
- Krzanowski, W.J. (1998). *An Introduction to Statistical Modelling*. London: Arnold.
- Kuhn, M. y Johnson, K. (2013). *Applied Predictive Modeling*. New York: Springer-Verlag.
- Kutas, M. y Federmeier, K.D. (2011). Thirty years and counting: finding meaning in the N400 component of the Event Related Brain Potential (ERP). *Annual Review of Psychology*, 62, 14.1-14.27.
- Kutas, M. y Van Petten, C. (1994). Psycholinguistics electrified: Event-Related Potential investigations. En M.A. Gernsbacher (Ed.), *Handbook of Psycholinguistics*. San Diego, CA: Academic Press.
- Kutas, M., McCarthy, G. y Donchin, E. (1977). Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. *Science*, 197(4305), 792-795.
- Landis, J.R. y Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.
- Lang, J.B. y Agresti, A. (1994). Simultaneous modelling joint and marginal distributions of multivariate categorical responses. *Journal of the American Statistical Association*, 89, 625-632.

- Lehman, M. y Malmberg, K.J. (2009). A global theory of remembering and forgetting from multiple lists. *Journal of Experimental Psychology*, 35, 970-988.
- Leuthold, H. y Sommer, W. (1998). Postperceptual effects and P300 latency. *Psychophysiology*, 35, 34-46.
- Li, Y. y Baron, J. (2012). *Behavioral Research Data Analysis with R*. New York: Springer-Verlag.
- Liao, T.F. (1994). *Interpreting Probability Models: Logit, Probit and other Generalized Linear Models*. Thousand Oaks, CA: Sage Publications.
- Lindsey, J.K. (1997). *Applying Generalized Linear Models*. New York: Springer-Verlag.
- Lindsey, J.K. (1995). *Modelling Frequency and Count Data*. Oxford: Clarendon Press.
- Loftus, G.R. (1978). On interpretation of interactions. *Memory & Cognition*, 6(3), 312-319.
- Lopez, E. (1998). Tratamiento de la colinealidad en regresión múltiple. *Psicothema*, 10(2), 491-507.
- Losilla, J.M., Navarro, J.B., Palmer, A., Rodrigo, M.F. y Ato, M. (2005). *Análisis de Datos: Del Contraste de Hipótesis al Modelado Estadístico*. Tarrasa: CBS.
- Luce, R.D. (1959). *Individual Choice Behavior. A Theoretical Analysis*. New York: J. Wiley & Sons.
- Luce, R.D. (1986). *Response Times: Their Role in Inferring*. New York: Oxford University Press
- Luce, R.D. (1999). Where is mathematical modeling in psychology headed? *Theory & Psychology*, 9(6), 723-737.

- Luck, S.J. (2005). *An Introduction to the Event-Related Potential Technique*. Cambridge, MA: The MIT Press.
- Lunneborg, C.E. (1994). *Modeling Experimental and Observational Data*. Belmont, CA: Duxbury Press.
- Macmillan, M.A. y Creelman, C.D. (1991). *Detection Theory. A User's Guide*. Cambridge. Cambridge University Press.
- Mahalanobis, P.C. (1936). On the generalized distances in statistics. *Proceedings National Institute of Science*, 2(1), 49-55.
- Malapeira, J.M., Honrubia, M.L., Viader, M., *et al.* (1995). Análisis de los componentes del tiempo de reacción. *Anuario de Psicología*, 65, 139-151.
- Mallows, C.L. (1973). Some Comments on CP. *Technometrics*, 15(4), 661-675.
- Malmberg, K.J. (2008). Recognition memory: a review of the critical findings and an integrated theory for relating them. *Cognitive Psychology*, 57, 335-384.
- Malmberg, K.J. y Murnane, K. (2002). List composition and the word-frequency effect for recognition memory. *Journal of Experimental Psychology*, 28(4), 616-630.
- Malmberg, K.J., Steyvers, M., Stephens, J.D. y Shiffrin, R.M. (2002). Feature frequency effects in recognition memory. *Memory and Cognition*, 30(4), 607-613.
- Mandler, G. (1991). Your face looks familiar but I can't remember your name: A review of dual process theory. En W. Hockley y E.S. Lewandowsky (Eds.). *Relating Theory and Data: Essays on Human Memory in Honour of Bennet B. Murdock*. Hillsdale: Erlbaum.

- Mayes, A.R. (1988). Amnesia and memory for contextual information. En G.M. Davies, E. Donald, y M. Thomson (Eds.). *Memory in Context: Context in Memory*. New York: John Wiley.
- McCullagh, P. (2002). What is an statistical model? *The Annals of Statistics*, 30(5), 1225-1310.
- McCullagh, P. y Nelder, J.A. (1989). *Generalized Linear Models (2nd Edition)*. London: Chapman & Hall.
- McNicol, D. (1972). *A Primer of Signal Detection Theory*. London: Allen & Unwin.
- Mell, D., Bach, M. y Heinrich, S.P. (2008). Fast stimulus sequences improve the efficiency of Event-related Potential P300 recordings. *Journal of Neuroscience Methods*, 174, 259-264.
- Menard, S. (2002). *Applied Logistic Regression Analysis (2nd Edition)*. Thousand Oaks, CA: Sage Publications.
- Mickes, L., Wais, P.E. y Wixted, J.T. (2009). Recollection is a continuous: implications for dual-process theories of recognition memory. *Psychological Science*, 20(4), 509-515.
- Mickes, L., Wixted, J.T. y Wais, P.E. (2007). A direct test of the unequal-variance Signal Detection Model of recognition memory. *Psychonomic Bulletin & Review*, 14(5), 858-865.
- Mosteller, F. (1951). Remarks on the method of paired comparisons: I. the least squares solution assuming equal SD and equal correlations. *Psychometrika*, 16(1), 3-9.
- Müller, S. y Welsh, A.H. (2009). Model selection in Generalized Linear Models. *Statistica Sinica*, 19, 1155-1170.

- Münste, T.F., Wieringa, B.M., Weyerts, H, Szentkuti, A, *et al.* (2001). Differences in brain potentials to open and closed class words: class and frequency effects. *Neuropsychologia*, 39(1), 91-102.
- Myung, I.J. (2000). The importance of complexity in model selection. *Journal of Mathematical Psychology*, 44, 190-204.
- Nelder, J.A. y Wedderburn, R.W. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society*, 135(3), 370-384.
- Nelson, T.O. (1984). A comparison of current measures of accuracy feeling-of-knowing predictions. *Psychological Bulletin*, 95, 109-133.
- Norman, K.A., Tepe, K., Nyhus, E. y Curran, T. (2008). Event-Related Potential correlates of interference effects on recognition memory. *Psychonomic Bulletin & Review*, 15(1), 36-43.
- Núñez-Peña, M.I. Corral, M.J. y Escera, C. (2004). Potenciales evocados cerebrales en el contexto de la investigación psicológica: una actualización. *Anuario de Psicología*, 35(1), 3-21.
- Nyberg, L., Cabeza, R. y Tulving, E. (1996). PET studies of encoding and retrieval: The HERA model. *Psychonomic Bulletin & Review*, 3(2), 135-148.
- Olivares-Carreño, E.I. e Iglesias-Dorado, J. (2008). Potenciales evocados de larga latencia y procesamiento mnésico de caras y palabras. *Revista de Neurología*, 47(12), 624-630.
- Oliver, J.C., Rosel, J. y Murray, L. (2000). Análisis de medidas repetidas mediante métodos de máxima verosimilitud. *Psicothema*, 12(2), 403-407.
- Pampel, F.C. (2000). *Logistic Regression: A Primer*. Thousand Oaks, CA: Sage Publications.

- Papanicolaou, A.C., Loring, D.W., Raz, N., *et al.* (1985). Relationship between stimulus intensity and the P300. *Psychophysiology*, 22(3), 326-329.
- Pelegrina, M., Arnau, J. y Malapeira, J.M. (1986). La Teoría de Detección de Señales en tareas de reconocimiento de palabras "clase cerrada". *Anuario de Psicología*, 35(2), 54-72.
- Pelegrina, M., Emberley, E., Wallace, A. y Marín, R. (2011). Análisis ROC y z-ROC en memoria de reconocimiento: modelos continuos, de umbral y asimétricos. *Psicothema*, 23(4), 845-850.
- Pelegrina, M., Ruiz-Soler, M., López, E. y Wallace, A. (2000). Análisis de variables mediante curvas ROC y modelos categóricos. *Psicothema*, 12(2), 427-430.
- Pelegrina, M., Ruiz-Soler, M., y Wallace, A. (2000). Índices de discriminación y análisis ROC en el contexto de los Modelos Lineales Generalizados. *Metodología de las Ciencias del Comportamiento*, 2(1), 63-83.
- Pelegrina, M., Wallace, A. y Tejeiro, R. (2003). Modelado estadístico, TDS y curvas ROC en diferencias individuales. *Metodología de las Ciencias del Comportamiento*, 5(1), 43-59.
- Pelegrina, M., Wallace, A., Emberley, E. y Tejeiro, R. (2011). Modelos matemáticos en memoria de reconocimiento de palabras: un análisis ROC multinivel. *Revista Latinoamericana de Psicología*, 43(3), 521-530.
- Peña, D. (2002). *Análisis de Datos Multivariantes*. Madrid: McGraw-Hill.
- Picton, T.W., Bentin, S., Berg, P., Donchin, E., *et al.* (2000). Guidelines for using human Event-related Potentials to study cognition: recording standards and publication criteria. *Psychophysiology*, 37, 127-152.
- Pierce, D.A. y Schafer, D.W. (1986). Residuals in Generalized Linear Models. *Journal of the American Statistical Association*, 81(396), 977-986.

- Pitman, E.J. (1936). Sufficient statistics and intrinsic accuracy. *Mathematical Proceedings of the Cambridge Philosophical Society*, 32(4), 567-579.
- Quiroz, Y.T. (2003). N-400: una medida electrofisiológica del procesamiento semántico. *Revista de Neurología*, 36(12), 1176-1180.
- Rotello, C.M., Macmillan, N.A. y Van Tassel, G. (2000). Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory and Language*, 43(1), 67-88.
- Roy, S.S. y Guria, S. (2008). Diagnostics in logistic regression models. *Journal of the Korean Statistical Society*, 37, 89-94.
- Rugg, M.D. y Coles, M.G. (1995). *Electrophysiology of Mind: Event Related Brain Potentials and Cognition*. New York: Oxford University Press.
- Sandi, C., Venero, C. y Cordero, M.I. (2001). Psicobiología de la memoria. En C. Sandi, C. Venero y M.I. Cordero, *Estrés, Memoria y Trastornos Asociados: Implicaciones en el Daño Cerebral y Envejecimiento*. Barcelona: Editorial Ariel.
- Sarle, W.S. (1994). Neural networks and statistical models. En SAS Institute (Ed.), *Proceedings of the 19th Annual SAS Users Group International Conference* (pp.1538-1550). Cary, NC: SAS Institute.
- Schendan, H.E., Ganis, G., y Kutas, M. (1998). Neurophysiological evidence for visual perceptual organization of words and faces within 150ms. *Neuropsychology*, 35, 240-251.
- Sellán, C. (1991). Registro de los potenciales evocados: aproximación a los componentes P300 y N400. *Revista de Psicología General y Aplicada*, 44(3), 317-322.
- Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461-464.

- Shibata, R. (2002). Information criteria for statistical model selection part.3. *Electronics and Communicatons in Japan*, 85(4), 2002.
- Shipley, E.F. (1965). Detection and recognition: experiments and choice models. *Journal of Mathematical Psychology*, 2(2), 277-31.
- Skrondal, A. y Rabe-Hesketh, S. (2003). Multilevel logistic regression for polytomous data and rankings. *Psychometrika*, 68(2), 267-287.
- Simpson, E.H. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society*, 13(2), 238-241.
- Smith, N.J. y Kutas, M. (2015). Regression-based estimation of ERP waveforms: I. The rERP framework. *Psychophysiology*, 52, 157-168.
- Sommer, W. y Matt, J. (1990). Awareness of P300-Related cognitive processes: a signal detection approach. *Psychophysiology*, 27(5), 575-585.
- Song, X-Y. y Lee, S-Y. (2006). Model comparison of Generalized Linear Mixed Models. *Statistics in Medicine*, 25, 1685-1698.
- Squire, L.R. (1987). *Memory and Brain*. New York: Oxford University Press.
- Stanislaw, H. y Todorov, N. (1999). Calculation of Signal Detection Theory measures. *Behavior Research Methods*, 31(1), 137-149.
- Starnes, D.S., Yates, D.S. y Moore, D.S. (2010). *The Practice of Statistics (4th Edition)*. New York: Freeman and Company.
- Stephaneck, P. (1987). Teoría de la Detección de Señales (TDS) y características de la memoria transitoria. *Revista Latinoamericana de Psicología*, 19(3), 381-391.
- Stern, R.M., Ray, W.J. y Quigley, K.S. (2001). *Psychophysiological Recording (2nd Edition)*. New York: Oxford University Press.
- Stevens, J.P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin*, 95(2), 334-344.

- Stigler, S.M. (1986). *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge, MA: Belknap Press.
- Sweatt, J.D. (2010). *Mechanisms of Memory (2nd Edition)*. London: Academic Press.
- Swets, J.A. (1961). Detection theory and psychophysics: a review. *Psychometrika*, 26(1), 49-63.
- Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. *Psychological Bulletin*, 99, 100-117.
- Swick, D. y Knight, R.T. (1999). Contributions of prefrontal cortex to recognition memory: electrophysiological and behavioral evidence. *Neuropsychology*, 13(2), 155-170.
- Tanner, W.P. Y Swet, J.A. (1953). A decision-making theory of visual detection. *Psychological Review*, 61, 401-409.
- Thurstone, L.L. (1927). Psychophysical analysis. *American Journal of Psychology*, 38(3), 368-389.
- Tollefson, M. (2014). *R Quick Syntax Reference*. New York: Apress.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychology*, 26(1), 1-12.
- Tulving, E. y Craik, F.I. (2000). *The Oxford Handbook of Memory*. New York: Oxford University Press.
- Tulving, E. y Markowitsch, H.J. (1998). Episodic and declarative memory: role of the hippocampus. *Hippocampus*, 8(3), 198-204.
- Ulrich, R. y Miller, J. (1993). Information processing models generating lognormally distributed reaction times. *Journal of Mathematical Psychology*, 37(4), 513-525.
- Van Petten, C. (1995). Words and sentences: Event-related Brain Potential measures. *Psychophysiology*, 32, 511-525.

- Van Petten, C. y Senkfor, A.J. (1996). Memory for words and novel visual patterns: repetition, recognition and encoding effects in the Event-related Brain Potential. *Psychophysiology*, 33, 491-506.
- Verfaellie, M. y Treadwell, J.R. (1993). Status of recognition memory in amnesia. *Neuropsychology*, 7(1), 5-13.
- Vermunt, J.K (1997). LEM 1.0: A general program for the analysis of categorical data. Tilburg: Tilburg University.
- Vonesh, E. y Chinchilli, V.M. (1996). *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. Boca Raton, FL: Chapman & Hall.
- Von Eye, A. y Mun, E.Y. (2013). *Log-Linear Modeling - Concepts, Interpretation, and Application*. Hoboken, NJ: J. Wiley & Sons.
- Wagner, A.D., Gabrieli, J.D. y Verfaellie, M. (1997). Dissociations between familiarity processes in explicit recognition and implicit perceptual memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(2), 305-323.
- Wald, A. (1939). Contributions to the theory of statistical estimation and testing hypotheses. *The Annals of Mathematical Statistics* 10(4), 299-326.
- Wald, A. (1945). Sequential tests of statistical hypotheses. *The Annals of Mathematical Statistics*, 16(2), 117-186.
- Wallace, A., Pelegrina, M. y Ruiz-Soler, M.(2002). Aplicación de los Modelos Lineales Generalizados en la evaluación de modelos sobre memoria de reconocimiento. *Metodología de las Ciencias del Comportamiento*, Vol. Esp., 570-572.
- Wei, B-C. y Shih, J-Q. (1994). On statistical model for regression diagnostics. *Annals of the Institute of Statistical Mathematics*, 46(2), 267-278.

- Wheeler, M. A., Stuss, D. T. & Tulving, E. (1997). Toward a theory of episodic memory: the frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*(3), 331-354.
- Wickelgren, W.A. (1970). Time, interference, and rate of presentation in short-term recognition memory for items. *Journal of Mathematical Psychology*, *7*, 219-235.
- Wixted, J.T. (2007). Dual-Process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*(1), 152-176.
- Yonelinas, A.P. (1994). Receiver-operating characteristics in recognition memory: evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1341-1354.
- Yonelinas, A.P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: a formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(6), 1415-1434.
- Yonelinas, A.P. (2001a). Consciousness, control, and confidence: The 3 Cs of recognition memory. *Journal of Experimental Psychology: General*, *130*(3), 361-379.
- Yonelinas, A.P. (2001b). Components of episodic memory: the contribution of recollection and familiarity. *Philosophical Transactions of the Royal Society of London, Biological Sciences*, *356*(1413), 1363-1374.
- Yonelinas, A.P. (2002). The nature of recollection and familiarity: a review of 30 years of research. *Journal of Memory & Language*, *46*, 441-517.

- Yonelinas, A.P. y Jacoby, L.L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology*, 48(4), 516-534.
- Yonelinas, A.P. y Jacoby, L.L. (1996). Noncriterial recollection: Familiarity as automatic, irrelevant recollection. *Consciousness and Cognition*, 5(1-2), 131-141.
- Yonelinas, A.P., Dobbins, I., Szymanski, M.D., *et al.* (1996). Signal-Detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness & Cognition*, 5, 418-441.
- Yonelinas, A.P., Kroll, N.E., Dobbins, I.G. & Soltani, M. (1999). Recognition memory of faces: When familiarity supports associative recognition judgments. *Psychonomic Bulletin & Review*, 6(4), 654-661.
- Yonelinas, A.P., Otten, L.J., Shaw, K.N. & Rugg, M.D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *The Journal of Neuroscience*, 5(11), 3002-3008.
- Zani, A. y Proverbio, A. (2002). *The Cognitive Electrophysiology of Mind and Brain*. Amsterdam: Academic Press.
- Zucchini, W. (2000). An introduction to model selection. *Journal of Mathematical Psychology*, 44, 41-61.

An bhail an tsolais a bheith ar an solas a-gan agus solas taobh istigh. Bealtaine an Shine solas na gréine beannaithe ar tú, agus te do chroí till glows sé cosúil le tine móna mór