

**Departamento de Estadística e I.O.**

**Máster en Estadística Aplicada**



**UNIVERSIDAD  
DE GRANADA**

**MODELOS DE RESPUESTA DISCRETA  
APLICACIONES BIOSANITARIAS**

**Tema 4**

**Guía de trabajo autónomo**

**Profesores**

**Ana María Aguilera del Pino**

**Manuel Escabias Machuca**

Título original: Modelos de Respuesta Discreta. Aplicaciones Biosanitarias.  
Guía de Trabajo Autónomo del Tema 4

© Los profesores

Todos los derechos reservados. Esta publicación es de uso personal del alumno y no puede ser reproducida, ni registrada, ni transmitida en ninguna forma ni por ningún medio, sin el permiso de los autores

# Índice general

<b>4. Guía de Trabajo Autónomo del Tema 4</b>	<b>1</b>
4.1. Justificación del Tema 4 . . . . .	1
4.2. Objetivos del Tema 4 . . . . .	2
4.3. Actividades del Tema 4 . . . . .	2

# Capítulo 4

## Guía de Trabajo Autónomo del Tema 4

Para asimilar los contenidos del Tema 4 el alumno debe estudiar los apuntes de teoría del tema y realizar las actividades de trabajo autónomo de esta guía. Estas actividades tienen como objetivo facilitar el autoaprendizaje del alumno y la consecución de los objetivos del tema. Es conveniente que el alumno las trabaje individualmente y plantee en el foro sus posibles dudas e inquietudes relacionadas con las mismas.

### 4.1. Justificación del Tema 4

Este capítulo está dedicado al estudio de modelos de respuesta múltiple en función de variables explicativas tanto cuantitativas como cualitativas. Se distinguirá entre variable de respuesta nominal y ordinal y se generalizará la transformación logística para adaptarse a cada caso. Para el caso de respuesta nominal se contruirán modelos lineales para el logaritmo de la ventaja de cada respuesta en lugar de una de referencia prefijada. Si la respuesta es ordinal se formularán modelos lineales para los logaritmos de las ventajas acumulativas de estar por encima en lugar de por debajo de cada respuesta de acuerdo a la escala ordinal asociada. La interpretación de las exponenciales de los parámetros será de nuevo en términos de cocientes de ventajas. El problema de estimación de parámetros, contrastes de hipótesis y validación y selección de modelos, se llevará a cabo de forma similar a las metodologías estudiadas en el Tema 2. Para poner en práctica los conocimientos adquiridos se realizarán ajustes de modelos logit nominales y ordinales con datos reales con variables explicativas tanto cuantitativas como cualitativas, mediante el uso del software libre R.

## 4.2. Objetivos del Tema 4

- Estudiar la formulación del modelo logit generalizados de respuesta nominal a partir de una o varias variables relacionadas.
- Formular modelos logit de respuesta ordinal a partir de una o varias variables relacionadas.
- Interpretar la relación entre la variable de respuesta y cada variable categórica en términos de cocientes de ventajas.
- Estimación de parámetros, selección de variables, inferencia sobre los parámetros estimados, validación de modelos e interpretación de resultados.
- Implementar los modelos logit nominales y ordinales estudiados con el software R para poder realizar aplicaciones con datos biomédicos o en cualquier otro campo de interés.

## 4.3. Actividades del Tema 4

1. Haga una lectura comprensiva de los apuntes del Capítulo 4 para obtener una visión global del tema.
2. Explique detalladamente la interpretación de las exponenciales de los parámetros asociados al modelo logit de respuesta nominal con una variable explicativa cuantitativa y una variable explicativa categórica codificada con el método parcial.
3. Explique detalladamente la interpretación de las exponenciales de los parámetros asociados al modelo logit acumulativo de respuesta ordinal con una variable explicativa cuantitativa y una variable explicativa categórica codificada con el método parcial.
4. Elabore sus propios apuntes sobre como se realiza la estimación, inferencia y selección paso a paso de modelos logit con variables explicativas categóricas con R, y utilícelos para resolver los ejercicios que se proponen a continuación. Plantee cualquier tipo de duda o comentario sobre su resolución en el foro de este tema.
5. Con objeto de estudiar la respuesta al tratamiento con quimioterapia del cáncer de pulmón, se ha llevado a cabo un ensayo clínico en el que los pacientes han sido asignados aleatoriamente a dos tipos de

tratamiento. La terapia secuencial en la que el paciente siempre recibe el mismo tipo de quimioterapia y la terapia alternante en la que se alternan tres tipos diferentes. Los datos resultantes del estudio figuran en la siguiente tabla:

Terapia	Sexo	Respuesta a la quimioterapia			
		Enfermedad Progresiva	Sin Cambio	Remisión Parcial	Remisión Completa
Secuencial	Masculino	28	45	29	26
	Femenino	4	12	5	2
Alternante	Masculino	41	44	20	20
	Femenino	12	7	3	1

- a) Ajustar, contrastar e interpretar el modelo logit generalizado de respuesta nominal más adecuado para explicar la probabilidad de cada tipo respuesta al tratamineto con quimioterapia.
  - b) Ajustar, contrastar e interpretar el modelo logit acumulativo de respuesta ordinal más apropiado para explicar la probabilidad de cada tipo respuesta al tratamineto con quimioterapia. Compararlo con el anterior.
  - c) Ajustar el modelo de ventajas proporcionales que contiene también la interacción entre el sexo y el tratamiento. Interpretarlo. ¿Se ajusta mejor que el modelo sin interacción?
6. Con objeto de estudiar la influencia de los acontecimientos (SV) vitales y el estatus socioeconómico (ES) sobre el estado mental (EM), se ha tomado una muestra de 40 individuos. La variable estado mental tiene cuatro categorías (bueno, deterioro leve, deterioro moderado y deteriorado). La variable sucesos vitales es un índice que mide el número y la importancia de acontecimientos vitales como, por ejemplo, el nacimiento de un hijo, un trabajo nuevo, un divorcio o una muerte en la familia, ocurridos en los últimos tres años. La variable estado socioeconómico tiene sólo dos categorías (alto y bajo). Los datos figuran en la siguiente tabla teniendo en cuenta que a las variables categóricas se les han asignado códigos enteros a sus categorías empezando por el uno:

EM	ES	SV	EM	ES	SV
1	1	1	1	1	9
1	1	4	1	1	3
1	2	2	1	1	0
1	2	1	1	1	3
1	1	3	1	1	7
1	2	1	1	2	2
2	1	5	2	2	6
2	1	3	2	2	1
2	1	8	2	1	2
2	2	5	2	1	5
2	1	9	2	2	3
2	1	3	2	1	1
3	2	0	3	1	4
3	2	3	3	2	9
3	1	6	3	2	4
3	2	3	4	1	8
4	1	2	4	1	7
4	2	5	4	2	4
4	2	4	4	1	8
4	2	8	4	2	9

Para predecir el estado mental en función del número de sucesos vitales y el estatus socioeconómico, se seleccionará un modelo logit acumulativo de respuesta ordinal. En base a los resultados del ajuste se pide:

- a) Utilizar el procedimiento de selección "stepwise" basado en contrastes condicionales de razón de verosimilitudes para seleccionar el modelo logit de respuesta múltiple que mejor se ajusta a estos datos. Razonar si el estatus socioeconómico es un factor de confusión para estudiar la relación entre el estado mental y los acontecimientos vitales.
- b) Contrastar la bondad del ajuste del modelo seleccionado mediante el test más adecuado, y analizar sus residuos. Contrastar la significación estadística de sus parámetros, interpretar su exponencial en términos de cocientes de ventajas y calcular intervalos de confianza del 99 % para los cocientes de ventajas. Estimar el cambio el estado mental cuando se aumenta en 3 unidades el índice que mide los suceso vitales y calcular para los cocientes de ventajas resultantes un intervalo de confianza del 99 %. Proporcionar finalmente una conclusión general sobre la relación entre el estado mental y los acontecimientos vitales.

7. Con objeto de estudiar la actitud sobre el aborto (negativa, mixta, positiva) se dispone de datos recolectados durante tres años por el Centro Americano de Investigación de Opinión. Los encuestados fueron identificados por sus años de educación y su grupo religioso (católicos, protestantes del sur y protestantes de otro tipo). La actitud hacia el aborto fue determinada en base a si el encuestado pensaba que el aborto debería ser legal bajo tres circunstancias: (a) alto riesgo de malformaciones fetales, (b) peligro para la salud de la madre, y (c) el embarazo es consecuencia de violación. El objetivo es ajustar un modelo logit de respuesta múltiple adecuado para estos datos, interpretarlo y extraer conclusiones a partir de los estimadores.

Año	Religión	Años Educa.	Actitud		
			Negativa	Mixta	Positiva
1972	Protestante	0-8	9	12	48
		9-12	13	43	197
		12+	4	9	139
	Prot. Sur	0-8	9	17	30
		9-12	6	10	97
		12+	1	8	68
	Católica	0-8	14	12	32
		9-12	18	50	131
		12+	8	13	64
1973	Protestante	0-8	4	16	59
		9-12	6	24	197
		12+	4	11	124
	Prot. Sur	0-8	4	16	34
		9-12	6	29	118
		12+	1	4	82
	Católica	0-8	2	14	32
		9-12	16	45	141
		12+	7	20	72
1974	Protestante	0-8	7	16	49
		9-12	10	26	219
		12+	4	10	131
	Prot. Sur	0-8	1	19	30
		9-12	5	21	106
		12+	2	11	87
	Católica	0-8	3	9	29
		9-12	15	30	149
		12+	11	18	69



- a)* Ajustar un modelo logit acumulativo de efectos homogéneos para predecir la actitud hacia el aborto en función de las tres variables explicativas consideradas. Formular el modelo y estimar la probabilidad de tener una actitud positiva en lugar de negativa o mixta para un católico con 10 años de educación en el año 1972.
- b)* Contrastar la bondad del ajuste del modelo estimado, hacer inferencia sobre sus parámetros, interpretar los parámetros en términos de cocientes de ventajas y validar el modelo mediante un análisis de residuos. Extraer conclusiones globales del efecto de cada variable explicativa sobre la actitud hacia el aborto.
- c)* Utilizar un procedimiento de selección "stepwise" para decidir si las tres variables deben estar en el modelo o se puede simplificar eliminando alguna de ellas.