Departamento de Estadística e I.O.

Máster en Estadística Aplicada

UNIVERSIDAD
DE GRANADA

MODELOS DE RESPUESTA DISCRETA
APLICACIONES BIOSANITARIAS

Tema 4
Modelos logit de respuesta multinomial

Profesores

Ana María Aguilera del Pino

Manuel Escabias Machuca

Título original: Modelos de Respuesta Discreta. Aplicaciones Biosanitarias.
Tema 4: Modelos logit de respuesta multinomial

Índice general

Capítulo 4

Modelos logit de respuesta multinomial

En este capítulo se generalizarán los modelos logit para el caso en que la variable respuesta es una variable categórica politómica con más de dos categorías de respuesta, distinguiendo entre respuestas puramente nominales y ordinales.

4.1. Introducción

Consideremos una v.a. de respuesta politómica Y con más de dos categorías de respuesta denotadas por $Y_1, Y_2, \ldots, Y_S$. Análogamente a la regresión logística de respuesta binaria, el objetivo es explicar la probabilidad de cada categoría de respuesta en términos de un conjunto de covariables $X_1, X_2, \ldots, X_R$ observadas sin error. Es decir, ajustar un modelo de la forma

$$p_s(x) = f_s(x) \quad \forall s = 1, \ldots, S,$$

definiendo $p_s(x) = P[Y = Y_s/X = x]$ para cada vector x de valores observados de las variables explicativas.

Recordemos que en el caso de una variable de respuesta binaria su distribución condicionada a cada combinación de valores observados de las covariables era una Bernouilli. Cuando la variable de respuesta es politómica, la distribución de Bernouilli se convierte en una multinomial de parámetros las probabilidades de cada una de las categorías de respuesta. Es decir, $(Y/X = x) \to M(1; p_1(x), \ldots, p_S(x))$, verificando que $\sum_{s=1}^{S} p_s(x) = 1$.

En el caso de una variable de respuesta binaria, el modelo logit se ha construido como un modelo lineal para el logaritmo de la ventaja de respuesta $Y = 1$ frente a respuesta $Y = 0$. Cuando la respuesta es politómica se puede definir una transformación logit para comparar cada par de categorías de respuesta. En total $\begin{pmatrix} S \\ 2 \end{pmatrix}$ transformaciones logit del tipo

$$\ln \left[\frac{\dfrac{p_t(x)}{p_t(x)+p_s(x)}}{\dfrac{p_s(x)}{p_t(x)+p_s(x)}} \right] = \ln \left[\frac{p_t(x)}{p_s(x)} \right], \quad \forall t, s = 1, \ldots, S \ (t \neq s),$$

que representan el logaritmo de la ventaja de respuesta Y_t frente a Y_s condicionado a las observaciones que caen en uno de ambos niveles. Sin embargo, este conjunto de transformaciones logit es redundante y para construir el modelo logit de respuesta multinomial bastará con considerar $(S-1)$ transformaciones logit básicas.

4.2. Modelos logit generalizados de respuesta nominal

Cuando la variable de respuesta es nominal el modelo logit se formula en base a las *transformaciones logit generalizadas* definidas con respecto a una categoría de referencia. Tomando como categoría de referencia la última (Y_S) (BMDP), las transformaciones logit generalizadas se definen como

$$L_s(x) = \ln \left[\frac{p_s(x)}{p_S(x)} \right] \quad \forall s = 1, \ldots, S-1,$$

donde $L_s(x)$ representa el logaritmo de la ventaja de respuesta Y_s dado que la respuesta cae en la categoría Y_s o en la Y_S.

Observemos que cualquier transformación logit para un par de categorías se puede obtener a partir de sus transformaciones logit generalizadas asociadas en la forma

$$\ln \left[\frac{p_t(x)}{p_s(x)} \right] = L_t(x) - L_s(x) \quad \forall t, s.$$

4.2.1. Formulación e interpretación

Análogamente al caso de respuesta binaria, el modelo logit generalizado de respuesta nominal se formula como un modelo lineal para cada uno de las transformaciones logit generalizadas. En el caso de R variables explicativas $X_1, \ldots, X_R$ el modelo es de la forma

$$L_s(x) = \sum_{r=0}^{R} \beta_{rs} x_r = x' \beta_s \quad \forall s = 1, \ldots, S-1,$$

para cada vector de valores observados de las variables explicativas $x = (x_0, x_1, \ldots, x_R)'$ con $x_0 = 1$, y $\beta_s = (\beta_{0s}, \beta_{1s}, \ldots, \beta_{Rs})'$ el vector de parámetros asociado a la categoría Y_s.

Equivalentemente, el modelo se puede escribir de la siguiente forma para las probabilidades de respuesta

$$\begin{aligned} p_s(x) &= \frac{\exp\left(\sum_{r=0}^{R} \beta_{rs} x_r\right)}{1 + \sum_{s=1}^{S-1} \exp\left(\sum_{r=0}^{R} \beta_{rs} x_r\right)} \quad \forall s = 1, \ldots, S-1, \\ p_S(x) &= \frac{1}{1 + \sum_{s=1}^{S-1} \exp\left(\sum_{r=0}^{R} \beta_{rs} x_r\right)}. \end{aligned} \tag{4.1}$$

Para demostrar esta equivalencia basta con razonar como sigue:

$$\sum_{s=1}^{S-1} \frac{p_s(x)}{p_S(x)} = \sum_{s=1}^{S-1} \exp\left(x'\beta_s\right),$$

de donde se tiene que

$$\frac{1 - p_S(x)}{p_S(x)} = \sum_{s=1}^{S-1} \exp\left(x'\beta_s\right),$$

y por lo tanto

$$p_S(x) = \frac{1}{1 + \sum_{s=1}^{S-1} \exp\left(\sum_{r=0}^{R} \beta_{rs} x_r\right)},$$

lo que implica a su vez

$$p_s(x) = \frac{\exp\left(\sum_{r=0}^{R} \beta_{rs} x_r\right)}{1 + \sum_{s=1}^{S-1} \exp\left(\sum_{r=0}^{R} \beta_{rs} x_r\right)} \quad \forall s = 1, \ldots, S-1.$$

El modelo (4.1) se puede formular a su vez de una forma más resumida como sigue

$$p_s(x) = \frac{\exp\left(\sum_{r=0}^{R} \beta_{rs} x_r\right)}{\sum_{s=1}^{S} \exp\left(\sum_{r=0}^{R} \beta_{rs} x_r\right)} \quad \forall s = 1, \ldots, S, \tag{4.2}$$

definiendo $\beta_{rS} = 0 \ \ \forall r = 0, 1, \ldots, R$.

En el caso de una única variable explicativa cuantitativa X, el modelo es de la forma

$$L_s(x) = \alpha_s + \beta_s x, \quad \forall s = 1, \ldots, S-1,$$

para cada valor observado x de la variable explicativa X. La exponencial de los parámetros β_s asociados a cada categoría de respuesta se interpreta como sigue en términos de cocientes de ventajas:

$$\begin{aligned}
\theta_s(\Delta X = 1) &= \frac{\dfrac{p_s(x+1)}{p_S(x+1)}}{\dfrac{p_s(x)}{p_S(x)}} = \frac{\exp\left(\alpha_s + \beta_s(x+1)\right)}{\exp\left(\alpha_s + \beta_s x\right)} \\
&= \exp\left(\beta_s\right) \ \ \forall s = 1, \ldots, S-1,
\end{aligned}$$

siendo $\theta_s(\Delta X = 1)$ el cociente de ventajas de respuesta Y_s frente a la última Y_S cuando se incrementa en una unidad la variable X.

En el modelo logit generalizado múltiple los cocientes de ventajas se definen incrementando una de las variables y controlando fijas las demás

$$\begin{aligned}
\theta_s(\Delta X_l = 1/X_r = x_r, \ r \neq l) &= \frac{\dfrac{P[Y = Y_s/X_l = x_l + 1, X_r = x_r, \ r \neq l]}{P[Y = Y_S/X_l = x_l + 1, X_r = x_r, \ r \neq l]}}{\dfrac{P[Y = Y_s/X_l = x_l, X_r = x_r, \ r \neq l]}{P[Y = Y_S/X_l = x_l, X_r = x_r, \ r \neq l]}} \\
&= \exp\left(\beta_{ls}\right) \ \ \forall s = 1, \ldots, S-1,
\end{aligned}$$

siendo $\theta_s(\Delta X_l = 1/X_r = x_r, \ r \neq l)$ el cociente de ventajas de respuesta Y_s en lugar de la última Y_S cuando se incrementa en una unidad la variable X_l y las demás se controlan fijas.

Cuando alguna de las variables explicativas es categórica se introduce en el modelo mediante sus variables del diseño asociadas. Como ejemplo consideraremos el caso de una variable explicativa categórica A de categorías $A_1, \ldots, A_I$. Recordemos que a una variable de este tipo se asocian un total de $(I-1)$ variables de diseño denotadas por X_m^A $(m = 2, \ldots, I)$. El modelo logit generalizado de respuesta nominal Y es entonces un modelo lineal para cada logit generalizado en términos de dichas variables de diseño

$$L_{s/i} = \ln \left[\frac{p_{s/i}}{p_{S/i}} \right] = \beta_{0s} + \sum_{m=2}^{I} \tau_{ms}^{A} X_{im}^{A} \quad i = 1, \ldots, I; s = 1, \ldots, S-1,$$

donde $p_{s/i}$ es la probabilidad de respuesta Y_s en la categoría A_i, es decir, $p_{s/i} = P[Y = Y_s/A = A_i]$.

Utilizando el método parcial para codificar las variables del diseño que asigna un uno a la categoría asociada a cada variable y un cero al resto, y tomando como categoría de referencia la primera, el modelo anterior es de la forma

$$L_{s/i} = \beta_{0s} + \tau_{is} \quad i = 1, \ldots, I; s = 1, \ldots, S-1,$$

definiendo $\tau_{1s} = 0, \quad \forall s = 1, \ldots, S-1$.

La interpretación del modelo en términos de cocientes de ventajas es en este caso como sigue:

$$\begin{aligned} \theta_{s/i1} &= \frac{\frac{p_{s/i}}{p_{S/i}}}{\frac{p_{s/1}}{p_{S/1}}} = \frac{\exp(\beta_{0s} + \tau_{is})}{\exp(\beta_{0s})} \\ &= \exp(\tau_{is}) \quad \forall s = 1, \ldots, S-1, \ \forall i = 2, \ldots, I, \end{aligned}$$

siendo $\theta_{s/i1}$ el cociente de ventajas de respuesta Y_s frente a la última Y_S para la categoría A_i de A respecto de la primera A_1.

4.2.2. Estimación por máxima verosimilitud

Supongamos que disponemos de una muestra aleatoria de tamaño N con Q combinaciones diferentes de valores de las variables explicativas $X_1, \ldots, X_R$.

En cada combinación de valores de las variables explicativas, denotada por $x_q = (x_{q0}, x_{q1}, \ldots, x_{qR})'$ con $x_{q0} = 1 \quad \forall q = 1, \ldots, Q$; se dispone de una muestra aleatoria de n_q observaciones independientes de la variable de respuesta politómica Y, de entre las cuales denotaremos por $y_{s/q}$ al número de observaciones que caen en la categoría de respuesta $Y_s \ \ \forall s = 1, \ldots, S$. Por lo tanto, se verifica que $\sum_{s=1}^{S} y_{s/q} = n_q$ y $\sum_{q=1}^{Q} n_q = N$.

Entonces los vectores $(y_{1/q}, \ldots, y_{S/q})' \ \ \forall q = 1, \ldots, Q$ tienen distribuciones de probabilidad multinomiales independientes, $M(n_q; p_{1/q}, \ldots, p_{S/q})$ verificando que $\sum_{s=1}^{S} p_{s/q} = 1$, donde $p_{s/q} = P[Y = Y_s/X = x_q]$.

La función de verosimilitud de los datos en entonces

$$\prod_{q=1}^{Q} \left(\frac{n_q!}{\prod_{s=1}^{S} (y_{s/q})!} \prod_{s=1}^{S} p_{s/q}^{y_{s/q}} \right),$$

de modo que el núcleo de la log-verosimilitud viene dado por

$$K = \sum_{q=1}^{Q} \sum_{s=1}^{S} y_{s/q} \ln{(p_{s/q})}.$$

Bajo el modelo logit generalizado de respuesta nominal (4.2), el núcleo de la log-verosimilitud se reduce a

$$\begin{aligned} K &= \sum_{q=1}^{Q} \sum_{s=1}^{S} y_{s/q} \left(\sum_{r=0}^{R} \beta_{rs} x_{qr} \right) - \sum_{q=1}^{Q} \left(\sum_{s=1}^{S} y_{s/q} \right) \ln \left(\sum_{s=1}^{S} \exp \left(\sum_{r=0}^{R} \beta_{rs} x_{qr} \right) \right) \\ &= \sum_{q=1}^{Q} \sum_{s=1}^{S} y_{s/q} \left(\sum_{r=0}^{R} \beta_{rs} x_{qr} \right) - \sum_{q=1}^{Q} n_q \ln \left(\sum_{s=1}^{S} \exp \left(\sum_{r=0}^{R} \beta_{rs} x_{qr} \right) \right). \end{aligned}$$

Derivando respecto de los parámetros como sigue

$$\frac{\Delta K}{\beta_{rs}} = \sum_{q=1}^{Q} y_{sq} x_{qr} - \sum_{q=1}^{Q} n_q x_{qr} \frac{\exp \left(\sum_{r=0}^{R} \beta_{rs} x_{qr} \right)}{\sum_{s=1}^{S} \exp \left(\sum_{r=0}^{R} \beta_{rs} x_{qr} \right)},$$

se obtienen las ecuaciones de verosimilitud cuya forma matricial es

$$X'_{((R+1)\times Q)} y_{s_{(Q\times 1)}} = X'_{((R+1)\times Q)} \hat{m}_{s_{(Q\times 1)}} \quad \forall s = 1, \ldots, S-1,$$

siendo

$$\begin{aligned} y_s &= (y_{s/1}, \ldots, y_{s/Q})' \\ \hat{m}_s &= (\hat{m}_{s/1}, \ldots, \hat{m}_{s/Q})' \end{aligned}$$

denotando por $\hat{m}_{s/q}$ a la frecuencia esperada de respuesta Y_s en la combinación x_q de valores observados de las variables explicativas, estimada bajo el modelo y definida por $\hat{m}_{s/q} = n_q \hat{p}_{s/q}$.

Observemos que para obtener los estimadores de máxima verosimilitud de los parámetros es necesario resolver simultáneamente $(S-1)$ sistemas de $(R+1)$ ecuaciones no lineales cada uno. La solución se obtiene de nuevo recurriendo a métodos iterativos como el de Newton-Raphson.

El estimador de los parámetros $\hat{\beta}$ es un vector columna de dimensión $(R+1) \times (S-1)$ definido por

$$\hat{\beta} = (\hat{\beta}'_1, \hat{\beta}'_2, \ldots, \hat{\beta}'_{S-1})',$$

donde $\hat{\beta}'_s$ es el estimador de máxima verosimilitud del vector de parámetros asociado a cada categoría de respuesta Y_s $(s = 1, \ldots, S-1)$.

La matriz de covarianzas de $\hat{\beta}$ es la inversa de la matriz de información de Fisher. Para obtenerla calcularemos en primer lugar la matriz de covarianzas de cada vector de parámetros $\hat{\beta}_s$. Para ello es necesario calcular las siguientes derivadas segundas de K con $l \neq r$:

$$\frac{\Delta^2 K}{\Delta\beta_{ls}\Delta\beta_{rs}} = -\sum_{q=1}^{Q} n_q x_{qr} x_{ql} \frac{\exp\left(\sum_{r=0}^{R}\beta_{rs}x_{qr}\right)\left[\sum_{s=1}^{S}\exp\left(\sum_{r=0}^{R}\beta_{rs}x_{qr}\right) - \exp\left(\sum_{r=0}^{R}\beta_{rs}x_{qr}\right)\right]}{\left[\sum_{s=1}^{S}\exp\left(\sum_{r=0}^{R}\beta_{rs}x_{qr}\right)\right]^2},$$

de modo que la matriz de covarianzas es de la forma

$$Cov(\hat{\beta}_s) = \left[-E\left(\frac{\Delta^2 K}{\Delta\beta_{ls}\Delta\beta_{rs}}\right)\right]^{-1} = \left[X' Diag[n_q p_{s/q}(1-p_{s/q})]X\right]^{-1}.$$

En segundo lugar obtendremos las matrices de covarianzas cruzadas entre cada par de estimadores $\hat{\beta}_s$ y $\hat{\beta}_t$ $(t \neq s)$. Para ello calcularemos en primer lugar las siguientes derivadas segundas de K con $l \neq r$ y $s \neq t$:

$$\frac{\Delta^2 K}{\Delta\beta_{lt}\Delta\beta rs} = -\sum_{q=1}^{Q} n_q x_{qr} x_{ql} \frac{-\exp\left(\sum_{r=0}^{R}\beta_{rs}x_{qr}\right)\exp\left(\sum_{r=0}^{R}\beta_{rt}x_{qr}\right)}{\left[\sum_{s=1}^{S}\exp\left(\sum_{r=0}^{R}\beta_{rs}x_{qr}\right)\right]^2},$$

que proporcionan la siguiente expresión para la matriz de covarianzas:

$$Cov(\hat{\beta}_s, \hat{\beta}_t) = \left[-E\left(\frac{\Delta^2 K}{\Delta\beta_{lt}\Delta\beta_{rs}}\right)\right]^{-1} = \left[-X' Diag[n_q p_{s/q} p_{t/q}]X\right]^{-1}.$$

La matriz de covarianzas del estimador $\hat{\beta}$ es entonces de la forma

$$Cov(\hat{\beta}) = \begin{pmatrix} Cov(\hat{\beta}_1) & Cov(\hat{\beta}_1, \hat{\beta}_2) & \ldots & Cov(\hat{\beta}_1, \hat{\beta}_{S-1}) \\ Cov(\hat{\beta}_1, \hat{\beta}_2) & Cov(\hat{\beta}_2) & \ldots & Cov(\hat{\beta}_2, \hat{\beta}_{S-1}) \\ \ldots & \ldots & \ldots & \ldots \\ Cov(\hat{\beta}_1, \hat{\beta}_{S-1}) & Cov(\hat{\beta}_2, \hat{\beta}_{S-1}) & \ldots & Cov(\hat{\beta}_{S-1}) \end{pmatrix}$$

4.3. Modelos logit de respuesta ordinal

Para el caso de una variable respuesta cualitativa ordinal se definen a continuación dos tipos diferentes de transformaciones logit que tienen en cuenta el orden entre las distintas categorías de respuesta y dan lugar a modelos logit más simples desde el punto de vista de la interpretación.

4.3.1. Modelos logit acumulativos

Estos modelos se formulan como modelos lineales en términos de las variables explicativas para las transformaciones logit acumulativas definidas en la siguiente forma:

$$\begin{aligned} L_s^1(x) &= \ln\left(\frac{P[Y > Y_s/X = x]}{1 - P[Y > Y_s/X = x]}\right) = \ln\left(\frac{P[Y > Y_s/X = x]}{P[Y \leq Y_s/X = x]}\right) \\ &= \ln\left(\frac{1 - F_s(x)}{F_s(x)}\right), \ \forall s = 1, \ldots, S-1, \end{aligned}$$

definiendo la función de distribución asociada a la variable Y ordinal en la forma

$$F_s(x) = P[Y \leq Y_s/X = x] = \sum_{l=1}^{s} p_s(x),$$

y denotando por x al vector de observaciones de las variables explicativas.

Observemos que $L_s^1(x)$ es el logaritmo de la ventaja a favor de clasificarse en una categoría superior a s en lugar de en una inferior o igual a s. Es decir, logaritmo de la ventaja de clasificarse por encima de la categoría Y_s.

Hay libros como los de Agresti en los que las transformaciones logit se definen como $\ln\left[F_s(x)/(1 - F_s(x)\right]$ que son las opuestas de las definidas anteriormente y llevan a modelos equivalentes. La formulación aquí presentada es la implementada por los programas PR de BMDP y SPSS que son los utilizados para resolver los aspectos computacionales.

En el caso de una única variable explicativa cuantitativa X observada sin error el modelo logit acumulativo es de la forma

$$L_s^1(x) = \alpha_s + \beta_s x,$$

para cada valor observado x de la variable X.

Equivalentemente, el modelo anterior se puede formular como sigue en forma de curvas de regresión logística para cada respuesta binaria obtenida asignando el valor 1 a $(Y > Y_s)$ y el valor 0 a $(Y \leq Y_s)$:

$$P[Y > Y_s/X = x] = \frac{\exp\left(\alpha_s + \beta_s x\right)}{1 + \exp\left(\alpha_s + \beta_s x\right)}, \ \forall s = 1, \ldots, S-1,$$

de modo que la función de distribución asociada a la variable de respuesta es de la forma

$$F_s(x) = \frac{1}{1 + \exp\left(\alpha_s + \beta_s x\right)}, \ \forall s = 1, \ldots, S-1.$$

Como consecuencia las curvas de respuesta para explicar la probabilidad de cada tipo de respuesta Y_s en función de X es de la forma

$$p_s(x) = F_s(x) - F_{s-1}(x) = \frac{1}{1+\exp(\alpha_s + \beta_s x)} - \frac{1}{1+\exp(\alpha_{s-1} + \beta_{s-1} x)}.$$

Los parámetros β_s se interpretan como sigue en términos de cocientes de ventajas:

$$\begin{aligned}
\theta_s^1(\Delta X = 1) &= \frac{\dfrac{P[Y > Y_s/X = x+1]}{P[Y \leq Y_s/X = x+1]}}{\dfrac{P[Y > Y_s/X = x]}{P[Y \leq Y_s/X = x]}} \\
&= \frac{\exp(\alpha_s + \beta_s(x+1))}{\exp(\alpha_s + \beta_s x)} = \exp(\beta_s) \quad \forall s = 1, \ldots, S-1,
\end{aligned}$$

siendo $\theta_s^1(\Delta X = 1)$ el cociente de ventajas de categoría superior a Y_s cuando se incrementa en una unidad la variable X.

En el caso de R variables explicativas $X_1, \ldots, X_R$ el modelo logit acumulativo es de la forma

$$L_s^1(x) = \sum_{r=0}^{R} \beta_{rs} x_r = x'\beta_s,$$

siendo $x = (x_0, x_1, \ldots, x_R)'$ el vector de valores observados de las variables explicativas con $x_0 = 1$, y $\beta_s = (\beta_{0s}, \beta_{1s}, \ldots, \beta_{Rs})'$ el vector de parámetros asociados a la categoría Y_s, $\forall s = 1, \ldots, S-1$.

Nuevamente las exponenciales de los parámetros asociados a las variables explicativas se interpretan como sigue en términos de cocientes de ventajas:

$$\begin{aligned}
\theta_s^1(\Delta X_l = 1/X_r = x_r,\ r \neq l) &= \frac{\dfrac{P[Y > Y_s/X_l = x_l + 1, X_r = x_r,\ r \neq l]}{P[Y \leq Y_s/X_l = x_l + 1, X_r = x_r,\ r \neq l]}}{\dfrac{P[Y > Y_s/X_l = x_l, X_r = x_r,\ r \neq l]}{P[Y \leq Y_s/X_l = x_l, X_r = x_r,\ r \neq l]}} \\
&= \exp(\beta_{ls}) \quad \forall s = 1, \ldots, S-1,
\end{aligned}$$

siendo $\theta_s^1(\Delta X_l = 1/X_r = x_r,\ r \neq l)$ el cociente de ventajas de categoría superior a Y_s cuando se incrementa en una unidad la variable X_l y las demás se controlan fijas.

Modelos logit acumulativos de efectos homogéneos o ventajas proporcionales

Un modelo logit acumulativo más parsimonioso y simple de interpretar es el que tiene el mismo parámetro β para todas las transformaciones logit acumulativas. En el caso de una única variable explicativa cuantitativa el modelo es de la forma

$$L_s^1(x) = \alpha_s + \beta x.$$

Este modelo representa el mismo efecto de la variable X sobre los logaritmos de las ventajas de respuesta superior a Y_s para todos los $s = 1, \ldots, S-1$. Por ello recibe el nombre de modelo logit acumulativo de *efectos homogéneos* o de *ventajas proporcionales.*

Su interpretación es mucho más interesante puesto que

$$\theta_s^1(\Delta X = 1) = \exp(\beta),$$

que significa que la ventaja de respuesta superior a Y_s queda multiplicada por $\exp(\beta)$ cuando se incrementa en una unidad la variable X. En este caso las curvas logísticas

$$P[Y > Y_s/X = x] = \frac{\exp(\alpha_s + \beta x)}{1 + \exp(\alpha_s + \beta x)}, \ \forall s = 1, \ldots, S-1,$$

corresponden a la misma curva desplazada en el eje x, y el signo del parámetro β está asociado a su caracter creciente a decreciente. Así, $\beta > 0$ significa que la probabilidad de clasificarse en la escala superior (categorías más altas) de la variables de respuesta aumenta al incrementar la variable explicativa X. Por otro lado, $\beta < 0$ significa que al incrementar X aumenta la probabilidad de clasificarse en la escala inferior de la variable Y (categorías más bajas de la escala ordinal).

Cuando se dispone de R variables explicativas $X_1, \ldots, X_R$ el modelo logit acumulativo de ventajas proporcionales es de la forma

$$L_s^1(x) = \alpha_s + \sum_{r=1}^{R} \beta_r x_r, \quad \forall s = 1, \ldots, S-1,$$

de modo que las exponenciales de los parámetros asociados a las variables explicativas se interpretan como sigue en términos de cocientes de ventajas:

$$\theta_s^1(\Delta X_l = 1/X_r = x_r, \ r \neq l) = \exp(\beta_l).$$

4.3.2. Modelos logit para categorías adyacentes

Análogamente a los anteriores, estos modelos se formulan como modelos lineales en términos de las variables explicativas para las transformaciones logit asociadas a categorías adyacentes que se definen a continuación:

$$L_s^2(x) = \ln\left[\frac{p_s(x)}{p_{s+1}(x)}\right], \quad \forall s = 1, \ldots, S-1.$$

En el caso de una única variable explicativa cuantitativa X el modelo logit para categorías adyacentes es de la forma

$$L_s^2(x) = \alpha_s + \beta_s x, \tag{4.3}$$

para cada valor observado x de la variable X.

Equivalentemente, el modelo anterior se puede formular como sigue para cada una de las probabilidades de respuesta:

$$p_s(x) = \frac{\exp\left(\sum_{l=s}^{S-1}(\alpha_l + \beta_l x)\right)}{1 + \sum_{s=1}^{S-1}\exp\left(\sum_{l=s}^{S-1}(\alpha_l + \beta_l x)\right)}. \tag{4.4}$$

Esta expresión se obtiene a partir de la siguiente relación entre las transformaciones logit generalizadas y las transformaciones logit para categorías adyacentes:

$$L_s(x) = \ln\left[\frac{p_s(x)}{p_S(x)}\right] = \ln\left[\frac{p_s(x)}{p_{s+1}(x)}\frac{p_{s+1}(x)}{p_{s+2}(x)}\cdots\frac{p_{S-1}(x)}{p_S(x)}\right] = \sum_{l=s}^{S-1} L_l^2(x),$$

de modo que si el modelo logit para categorías adyacentes (4.3) se verifica, entonces se verificaría el siguiente modelo logit generalizado

$$L_s(x) = \sum_{l=s}^{S-1} L_l^2(x) = \left(\sum_{l=s}^{S-1}\alpha_l\right) + \left(\sum_{l=s}^{S-1}\beta_l\right)x,$$

que lleva directamente a la expresión (4.4) para las probabilidades de respuesta.

Las exponenciales de los parámetros β_s se interpretan como sigue en términos de cocientes de ventajas:

$$\begin{aligned}\theta_s^2(\Delta X = 1) &= \frac{\dfrac{p_s(x+1)}{p_{s+1}(x+1)}}{\dfrac{p_s(x)}{p_{s+1}(x)}} \\ &= \frac{\exp\left(\alpha_s + \beta_s(x+1)\right)}{\exp\left(\alpha_s + \beta_s x\right)} = \exp\left(\beta_s\right) \quad \forall s = 1,\ldots,S-1,\end{aligned}$$

siendo $\theta_s^2(\Delta X = 1)$ el cociente de ventajas de categoría Y_s en lugar de Y_{s+1} cuando se incrementa en una unidad la variable X.

Para el caso de R variables explicativas $X_1, \ldots, X_R$ el modelo logit para categorías adyacentes es de la forma

$$L_s^2(x) = \sum_{r=0}^{R} \beta_{rs} x_r = x' \beta_s,$$

siendo $x = (x_0, x_1, \ldots, x_R)'$ el vector de valores observados de las variables explicativas con $x_0 = 1$, y $\beta_s = (\beta_{0s}, \beta_{1s}, \ldots, \beta_{Rs})'$ el vector de parámetros asociados a la categoría Y_s, $\forall s = 1, \ldots, S-1$.

Nuevamente las exponenciales de los parámetros asociados a las variables explicativas se interpretan como sigue en términos de cocientes de ventajas:

$$\begin{aligned} \theta_s^2(\Delta X_l = 1/X_r = x_r,\ r \neq l) &= \frac{\dfrac{P[Y = Y_s/X_l = x_l + 1, X_r = x_r,\ r \neq l]}{P[Y = Y_{s+1} Y_s/X_l = x_l + 1, X_r = x_r,\ r \neq l]}}{\dfrac{P[Y = Y_s/X_l = x_l, X_r = x_r,\ r \neq l]}{P[Y = Y_{s+1}/X_l = x_l, X_r = x_r,\ r \neq l]}} \\ &= \exp\left(\beta_{ls}\right) \ \ \forall s = 1, \ldots, S-1, \end{aligned}$$

siendo $\theta_s^2(\Delta X_l = 1/X_r = x_r,\ r \neq l)$ el cociente de ventajas de categoría Y_s en lugar de Y_{s+1} cuando se incrementa en una unidad la variable X_l y las demás se controlan fijas.

Modelos logit para categorías adyacentes de ventajas paralelas

De nuevo, un modelo logit más parsimonioso y simple de interpretar es el que tiene el mismo parámetro β para todas las transformaciones logit para categorías adyacentes. En el caso de una única variable explicativa cuantitativa el modelo es de la forma

$$L_s^2(x) = \alpha_s + \beta x.$$

Equivalentemente,

$$p_s(x) = \frac{\exp\left(\left(\sum_{l=s}^{S-1} \alpha_s\right) + \beta(S-s)x\right)}{1 + \sum_{s=1}^{S-1} \exp\left(\left(\sum_{l=s}^{S-1} \alpha_s\right) + \beta(S-s)x\right)}.$$

Este modelo recibe el nombre de modelo logit para categorías adyacentes de *ventajas paralelas* (Equalodds en BMDP), y su interpretación es mucho más interesante puesto que

$$\theta_s^2(\Delta X = 1) = e^{\beta},$$

que significa que la ventaja de respuesta Y_s en lugar de Y_{s+1} queda multiplicada por e^{β} cuando se incrementa en una unidad la variable X. Así, $e^{\beta} > 1$ ($\beta > 0$) significa que la probabilidad de clasificarse en cada categoría en lugar de en la siguiente aumenta al incrementar la variable explicativa X. Esto quiere decir que al aumentar X se incrementa la probabilidad de que la respuesta caiga en las categorías inferiores de la escala ordinal. Por otro lado, $e^{\beta} < 1$ ($\beta < 0$) significa que al incrementar X aumenta la probabilidad de clasificarse en la escala superior de la variable Y (categorías más altas de la escala ordinal).

Cuando tenemos R variables explicativas $X_1, \ldots, X_R$ el modelo logit para categorías adyacentes de ventajas paralelas es de la forma

$$L_s^2(x) = \alpha_s + \sum_{r=1}^{R} \beta_r x_r, \quad \forall s = 1, \ldots, S-1.$$

En este caso las exponenciales de los parámetros asociados a las variables explicativas se interpretan como sigue en términos de cocientes de ventajas:

$$\theta_s^2(\Delta X_l = 1/X_r = x_r,\ r \neq l) = \exp(\beta_l).$$