

Risposte per una selezione di problemi con numerazione dispari

Capitolo 15

- 15.1** *a.* La probabilità stimata di votare Repubblicano aumenta all'aumentare del reddito.
b. (i) $e^{-1.0+0.02(10)}/[1+e^{-1.0+0.02(10)}] = 0.31$, (ii) 0.73 *c.* (i) $1.00/0.02 = 50$ mila, (ii) sopra 50 mila *d.* $\hat{\beta}\pi(1-\pi) = 0.02(0.5)(0.5) = 0.005$ *e.* L'odds moltiplicato per $e^{0.02} = 1.02$ per ogni incremento di mille dollari del reddito.
- 15.3** *a.* $2.043/0.282 = 7.2$ *b.* $\hat{P}(y=1) = e^{2.043-0.282(20)}/[1+e^{2.043-0.282(20)}] = 0.027$
c. $\hat{P}(y=1) = 0.847 - 0.051(20) = -0.17$; no, la probabilità non può essere negativa. *d.* $z = -0.282/0.101 = -2.80$, Wald = $(-2.80)^2 = 7.8$, $P = 0.005$, forte evidenza di un'associazione negativa tra WAIS e senilità.
- 15.5** $\hat{P}(y=1) = 0.60$ quando i guadagni della moglie = \$20,000 e aumenta a 0.95 quando i guadagni della moglie = \$100,000.
- 15.7** *a.* La probabilità che il figlio ottenga un diploma di scuola superiore aumenta all'aumentare del livello di istruzione della madre. L'odds stimato moltiplicato per $e^{0.09} = 1.09$ per ogni aumento di un'unità nel livello di istruzione della madre, controllando per le altre variabili. *b.* La probabilità di ottenere il diploma è più bassa quando la madre è impiegata. L'odds stimato quando la madre è impiegata è uguale a $e^{-0.92} = 0.40$ volte l'odds rispetto a quando la madre non è impiegata. *c.* $e^{0.21} = 1.23$, che corrisponde al 23% di incremento nell'odds.
- 15.9** *a.* Sia $r = 1$ per nero, 0 per bianco, sia $a = 1$ per AZT = sì e 0 for AZT = no. L'equazione di previsione è $\text{logit}[\hat{P}(y=1)] = -1.074 - 0.7195a + 0.0555r$. *b.* La probabilità stimata dei sintomi dell'AIDS diminuisce con l'uso di AZT (data la razza) ed è leggermente superiore per i neri, dato l'uso di AZT. *c.* $\hat{P}(y=1) = e^{-1.074}/[1+e^{-1.074}] = 0.25$ *d.* $e^{-0.720} = 0.49$, per cui gli odds stimati dei sintomi dell'AIDS per quelli che usano AZT sono 0.49 volte gli odds stimati per quelli che non usano AZT, controllando per la razza. *e.* La statistica di Wald è $(-0.720/0.279)^2 = 6.65$, $df = 1$ ($P = .01$), ovvero una forte evidenza che i sintomi dell'AIDS sono meno probabili per coloro che utilizzando AZT, controllando la razza.
- 15.13** *a.* Ci sono 3 categorie per la felicità e 2 probabilità cumulative nel modello. Il logit per ogni probabilità cumulativa ha il suo termine di intercetta. *b.* $\hat{\beta} = 0.418$, per cui la felicità tende ad aumentare con l'aumentare del reddito. *c.* Statistica di Wald = $(0.418/0.223)^2 = 3.53$, con $df = 1$, P -value = 0.06 per $H_a: \beta = 0$ e 0.03 per $H_a: \beta > 0$ *d.* $\chi^2 = 3.8$, $df = 4$, P -value = 0.43. Il modello logit cumulativo considera la variabile risposta ordinale e il predittore come variabile quantitativa con una più forte evidenza dell'effetto, rispetto al test chi-quadro di indipendenza di Pearson, che tratta entrambe le variabili come nominali.
- 15.15** *a.* $\hat{\beta} = -0.54$ per genere, 0.77 per posizione, -0.82 per le cinture di sicurezza. Controllando per gli altri predittori, la probabilità di una lesione più grave è più bassa per i maschi, più alta nelle zone rurali e più bassa per coloro che indossano

- le cinture di sicurezza. *b.* $e^{-0.824} = 0.44$; per quelli che indossano le cinture di sicurezza, l'odd stimato di infortunio più grave rispetto a qualsiasi altra categoria sono 0.44 volte l'odd stimato per coloro che non indossano le cinture di sicurezza.
- c.* L'intervallo $(-0.878, -0.770)$ per i coefficienti β dell'uso delle cinture di sicurezza ha punti estremi elevati a potenza (0.42, 0.46), che forma l'intervallo per questo odds ratio. *d.* Statistica di Wald = 891.5, $df = 1$. Estremamente forte evidenza dell'effetto.
- 15.17 *a.* $e^{0.3} = 1.35$, per cui l'odd stimato di votare Repubblicano (dato che si vota sia per il Repubblicano che per l'Indipendente) moltiplica 1.35 per ogni aumento di \$10,000 del reddito annuo. *b.* $\log[\hat{P}(y=2)/\hat{P}(y=1)] = (1.0 + 0.3x) - (3.3 - 0.2x) = -2.3 + 0.5x$. Dato che si vota per il candidato Repubblicano o Democratico, l'odd di votare Repubblicano aumenta con il reddito, per un fattore moltiplicativo di $e^{0.5} = 1.65$ per ogni aumento di \$10,000 nel reddito annuo.
- 15.23 *a.* (W, X, Y, Z) . *b.* (XY, W, Z) . *c.* (WX, WY, WZ, XY, XZ, YZ) .
- 15.25 *a.* $\text{logit}[P(y=1)] = \alpha + \beta_1 a + \beta_2 s$, dove a e s sono variabili dummy per l'uso di alcool e sigarette (1 per sì e 0 per no) *b.* Il modello si adatta bene.
- 15.29 Il modello logistico con effetti additivi per età e genere si adatta bene (χ^2 Pearson = 0.1, $df = 2$). L'odd stimato delle femmine è $e^{0.38} = 1.46$ volte quello dei maschi, data l'età. Gli odd stimati sono notevolmente più alti per quelli di almeno 19 anni, rispetto agli altri due gruppi di età, dato il genere.
- 15.31 $e^{2.34} = 10.4$, ma questo è un rapporto di *odd*, non probabilità (che l'interpretazione "più verosimile" suggerisce). L'odd che un maschio sia un cacciatore è pari a 10.4 volte l'odd per una femmina, controllando per le altre variabili.
- 15.33 L'effetto dell'età aumenta e diminuisce. Se l'età è presente nel modello come variabile quantitativa (piuttosto che categoriale), questo effetto potrebbe essere descritto da un termine quadratico per l'età.
- 15.35 Per ogni livello della razza delle vittime, gli imputati neri avevano una probabilità più alta di essere condannati alla pena di morte. Sommando le due tabelle parziali (e, quindi, ignorando invece di controllare la razza delle vittime), gli imputati bianchi avevano una probabilità più alta di essere condannati alla pena di morte. L'associazione cambia direzione, verificandosi il paradosso di Simpson.
- 15.37 *a.* (XZ, YZ) . *b.* (XZ, YZ) . *c.* (XZ, YZ) . *d.* (XY, XZ, YZ) . *e.* (XYZ) , c'è interazione, l'effetto di X su Y varia al variare del livello di Z .
- 15.39 $\log(f_e) = \log(r_i) + \log(c_j) - \log(n)$, che ha l'espressione $\log(f_e) = \alpha + \beta_i + \gamma_j$ con effetti principali per le due classificazioni.

Capitolo 16

- 16.1 Può utilizzare varie strutture di correlazione per misure ripetute su un soggetto e non ha bisogno di eliminare i soggetti per i quali alcune osservazioni sono mancanti.
- 16.5 Il tasso di rischio stimato per i maschi è 0.94 volte il tasso delle femmine. Per saggiare l'effetto della razza, $z = 0.353/0.0423 = 8.3$, $P < 0.0001$; vi è un'evidenza estremamente forte che il tasso è più alto per i neri.
- 16.7 Il tasso stimato di licenziamento per i neri era di 2.13 volte il tasso stimato per i bianchi, controllando per gli altri predittori.
- 16.9 G , L , T e C hanno effetti diretti su B e G ha anche effetti indiretti attraverso i suoi effetti su L , T e C . Si dovrebbero adattare i modelli bivariati per L e G , T e G e C e G , nonché il modello di regressione multipla con G , L , T e C che influenzano B .
- 16.11 Sì, l'effetto della percentuale di diplomati della scuola superiore diminuisce considerevolmente (e non è statisticamente significativo) dopo aver controllato per la percentuale di urbanizzazione.
- 16.19 y_t è condizionatamente indipendente da y_{t-2} , dato y_{t-1} .