Binary regression

TMA4315 H2018
M3w1 13.09.2018

1) Random component:

$P(Y_i = 1) = \pi_i$

$Y_i \sim \text{bin}(n_i = 1, \pi_i)$ individual data

$E(Y_i) = \mu_i = 1 \cdot \pi_i = \pi_i \leftarrow$ parameter of interest

$\text{Var}(Y_i) = \sigma_i^2 = 1 \cdot \pi_i(1-\pi_i) \leftarrow$ no nuisance parameter

M1: binomial is exponential family.

2) Linear predictor: $\eta_i = x_i^T \beta$

3) Link: η_i vs $\mu_i = \pi_i$

$\underset{\text{link}}{\eta_i = g(\pi_i)}$ and $\underset{\text{response}}{h(\eta_i) = \pi_i}$

a) logit: $\eta_i = \ln\left(\frac{\pi_i}{1-\pi_i}\right) \Leftrightarrow \pi_i = \frac{\exp(\eta_i)}{1+\exp(\eta_i)} = \frac{1}{1+\exp(-\eta_i)}$

remark:

$$\frac{\pi_i}{1-\pi_i} = \exp(\beta_0) \cdot \exp(\beta_1 x_{i1}) \cdots \exp(\beta_k \cdot x_{ik})$$

b) probit: $\eta_i = \Phi^{-1}(\pi_i)$ $\pi_i = \Phi(\eta_i)$

c) complementary log-log

$\eta_i = \ln(-\ln(1-\pi_i))$ $\pi_i = 1 - \exp(-\exp(\eta_i))$

Beetle example:

x-axis : η = linear predictor $\beta_0 + \beta_1 \cdot x_1$ ← log10 dose

y-axis : μ? yes, but also $\mu = \pi$ = Prob. success

where is "log10 dose $= x_1$"?

Interpreting the logit:

$$\eta_i = \ln\left(\frac{\pi_i}{1-\pi_i}\right) \; : \; \frac{\pi_i}{1-\pi_i} = e^{\eta_i} = \underbrace{e^{\beta_0} \cdot e^{\beta_1 x_{i1}} \cdots e^{\beta_k x_{ik}}}_{\text{multiplicative model for the odds.}}$$

odds

$$\frac{P(Y_i = 1 \mid x_{i1}, x_{i2}, \dots, x_{ik})}{P(Y_i = 0 \mid x_{i1}, x_{i2}, \dots, x_{ik})}$$

Increase the x_{i1} to $x_{i1}+1$, but keep all other x's fixed

$$\frac{P(Y_i = 1 \mid x_{i1}+1, x_{i2}, \dots, x_{ik})}{P(Y_i = 0 \mid x_{i1}+1, x_{i2}, \dots, x_{ik})} = e^{\beta_0} \cdot \underbrace{e^{\beta_1 \cdot (x_{i1}+1)}}_{e^{\beta_1 x_{i1}} \cdot e^{\beta_1}} \cdot e^{\beta_2 \cdot x_{i2}} \cdots e^{\beta_k \cdot x_{ik}}$$

$$= e^{\beta_1} \cdot \frac{P(Y_i = 1 \mid x_{i1}, \dots, x_{ik})}{P(Y_i = 0 \mid x_{i1}, \dots, x_{ik})}$$

Infont respitory disease hands-on: boy, bottle } reference level

1) sexGirl : $\hat{\beta}_1 = -0.3126$ $\quad (\text{odds boy}) \cdot \underbrace{e^{\hat{\beta}_1}}_{0.73} = \text{odds girl}$

$\exp(\hat{\beta}_1) = 0.73$

foodSuppl: $\hat{\beta}_3 = -0.1725$ $\quad (\text{odds bottle}) \cdot \underbrace{e^{\hat{\beta}_3}}_{0.84} = \text{odd suppl.}$

$\exp(\hat{\beta}_3) = 0.84$

2)

	x_1 sexgirl	x_2 foodSupp	x_3 foodbreast
1: boy & bottle	0	0	0
2: boy & supp	0	1	0
3: breast	0	0	1
4: girl & bottle	1	0	0
5: girl & supp	1	1	0
6: girl & breast	1	0	1

$$\hat{\pi} = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3)}{1 + \exp(\text{—''—})}$$

prob of getting disease

Most favorable = lowest probability of getting disease = 7% for girl & breast

Least favorable = 17% boy and bottle.

Likelihood and derivations thereof

AIM: $\hat{\beta}$

$$
\begin{tikzcd}
L(\beta) \arrow[r] & \ell(\beta) = \ln L(\beta) \arrow[r] & s(\beta) = \frac{\partial \ell}{\partial \beta} \arrow[r] \arrow[d] \arrow[dr] & H(\beta) = -\frac{\partial s}{\partial \beta^T} = -\frac{\partial^2 \ell}{\partial \beta \, \partial \beta^T} \\
 & & s(\hat{\beta}) = 0 & F(\beta) = \mathrm{Cov}(s(\beta)) = E(H(\beta))
\end{tikzcd}
$$

likelihood — log-likelihood — score vector — observed Fisher info — expected Fisher info

1) $L(\beta) = \prod_{i=1}^{n} \pi_i^{y_i} (1-\pi_i)^{1-y_i}$ $\qquad (\pi_i \to \eta_i \to \beta)$

2) $\ell(\beta) = \sum_{i=1}^{n} \underbrace{\left(y_i \cdot \ln \pi_i + (1-y_i) \ln(1-\pi_i) \right)}_{\ell_i}$

$$= \sum_{i=1}^{n} \left(y_i \underbrace{\ln\left(\frac{\pi_i}{1-\pi_i}\right)}_{\eta_i} + \ln(\underbrace{1-\pi_i}_{\frac{1}{1+e^{\eta_i}}}) \right)$$

$$= \sum_{i=1}^{n} \left(y_i \eta_i + \ln\left(\frac{1}{1+e^{\eta_i}}\right)\right) = \sum_{i=1}^{n} \underbrace{\left(y_i \eta_i - \ln\left(1+e^{\eta_i}\right)\right)}_{\ell_i}$$

3) score function

$$\underset{p\times 1}{S(\beta)} = \frac{\partial l}{\partial \beta} = \sum_{i=1}^{n} \underbrace{\frac{\partial l_i}{\partial \beta}}_{s_i(\beta)} \qquad \begin{bmatrix} \frac{\partial l_i}{\partial \beta_0} \\ \frac{\partial l_i}{\partial \beta_1} \\ \vdots \\ \frac{\partial l_i}{\partial \beta_k} \end{bmatrix} = \sum_{i=1}^{n} s_i(\beta)$$

$$s_i(\beta) = \underbrace{\frac{\partial l_i(\beta)}{\partial \eta_i}}_{1\times 1} \underbrace{\frac{\partial \eta_i}{\partial \beta}}_{p\times 1} \qquad \frac{\partial \eta_i}{\partial \beta} = \frac{\partial\, x_i^T\beta}{\partial \beta} = \underset{p\times 1}{x_i}$$

$$\frac{\partial l_i(\beta)}{\partial \eta_i} = \frac{\partial}{\partial \eta_i}\left(y_i\eta_i - \ln(1+e^{\eta_i})\right) = y_i - \underbrace{\frac{e^{\eta_i}}{1+e^{\eta_i}}}_{\pi_i}$$

$$= y_i - \pi_i$$

$$s_i(\beta) = (y_i - \pi_i)\cdot x_i$$

$$S(\beta) = \sum_{i=1}^{n} (y_i - \pi_i)\cdot x_i$$

$$S(\hat{\beta}) = 0 \iff \sum_{i=1}^{n} \underbrace{\left(y_i - \frac{e^{x_i^T\beta}}{1+e^{x_i\beta}}\right)}_{\text{scalar}} \underset{p\times 1}{x_i} = 0$$

p non-linear equation

Next: $S(\beta)$ can be seen as a random vector:

$$S(\beta) = \sum_{i=1}^{n} (Y_i - \pi_i)\cdot x_i$$ function of Y_i's

What is $E(\beta)$ and $Cov(S(\beta))$?