Multicategory Logit Models (Ch. 6) 3/31/22

Review: (Sec. 1.2) Multinomial distribution

Response variable Y can take on one of J categories:

$$Y = \begin{cases} 1 & \text{w/prob. } \pi_1 \\ 2 & \text{w/prob. } \pi_2 \\ 3 & \vdots \\ \vdots & \\ J & \text{w/prob } \pi_J \end{cases} \qquad \sum_{i=1}^{J} \pi_i = 1$$

→ Multinomial

→ n trials

$Z = (\#\text{cat. }1, \#\text{cat }2, \ldots, \#\text{cat } J)$

— Unordered (Nominal) → Baseline – Logit Models (Sec. 6.1)

— Ordered (Ordinal) → Cumulative Logit Models (Sec. 6.2)

Baseline – Category Logit Models

Choose one category to be the "baseline" category – w/o loss of generality → J

Model: Consists of $J-1$ equations:

$j = 1, 2, \ldots, J-1$

$$\log\left(\frac{\pi_j}{\pi_J}\right) = \beta_{0j} + \beta_{1j} X_1 + \cdots + \beta_{pj} X_p$$
$$= \beta_j^T \underline{x}$$

Note: $j = J \rightarrow \log\left(\frac{\pi_J}{\pi_J}\right) = \log(1) = 0$

For convenience, define $\beta_{0J} = \cdots = \beta_{pJ} = 0$

Notes: ① $\dfrac{\pi_j}{\pi_J}$ is not the odds of categ. j

"conditional odds"

$\dfrac{\pi_j}{1-\pi_j}$

$\log\left(\dfrac{\pi_j}{\pi_J}\right)$ = "logit" of $P(Y=j \mid Y=j \text{ or } Y=J)$

"log odds"

$$= \log\left(\frac{P(Y=j \mid Y=j \text{ or } Y=J)}{1-P(Y=j \mid Y=j \text{ or } Y=J)}\right)$$

② If $J=2 \longrightarrow$ logistic regression.

Interpret parameters? Model w/ one predictor:

$$\log\left(\frac{\pi_j(x)}{\pi_J(x)}\right) = \alpha_j + \beta_j x \qquad j=1,\ldots,J-1$$

If x increases by 1 unit:

$$\frac{\pi_j(x+1)}{\pi_J(x+1)} = e^{\alpha_j+\beta_j(x+1)} = e^{\alpha_j+\beta_j x}\cdot \underline{e^{\beta_j}}$$

Multiplicative change in the conditional odds of j to J when x increases by 1 unit.

– Same w/ multiple predictors except holding all other variables constant (at zero if interaction term present)

How do we compare π_a to π_b $(b \neq J)$

$$\log\left(\frac{\pi_a}{\pi_b}\right) = \log\left(\frac{\pi_a/\pi_J}{\pi_b/\pi_J}\right)$$

$$= \log\left(\frac{\pi_a}{\pi_J}\right) - \log\left(\frac{\pi_b}{\pi_J}\right)$$

$$= \alpha_a + \beta_a x - (\alpha_b + \beta_b x)$$

$$= (\alpha_a - \alpha_b) + (\beta_a - \beta_b) x$$

Interpret $\alpha_a - \alpha_b$: $\frac{\pi_a}{\pi_b} = e^{(\alpha_a-\alpha_b)+(\beta_a-\beta_b)x}$

If $x=0$, the conditional odds of a to b are $e^{\alpha_a-\alpha_b}$.

Interpret $\beta_a - \beta_b$: The conditional odds of a to b are multiplied by $e^{\beta_a-\beta_b}$ when x increases by 1 unit.

How to model π_j ?

Recall: Logistic regression — W = odds

$$W = \frac{\pi}{1-\pi} \iff \pi = \frac{W}{1+W}$$

If $\text{logit}(\pi) = \alpha+\beta x$, then $\frac{\pi}{1-\pi} = e^{\alpha+\beta x} = W$

$$\iff \pi = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$$

↳ logistic curve

With multinomial models: Since $\sum_{j=1}^{J} \pi_j = 1$,

$$\underbrace{\sum_{i=1}^{J} e^{\alpha_i+\beta_i x}}_{\frac{\pi_i}{\pi_J}} = \sum_{i=1}^{J} \frac{\pi_i}{\pi_J} = \frac{1}{\pi_J} \underbrace{\sum_{i=1}^{J} \pi_i}_{1} = \frac{1}{\pi_J}$$

$$\Rightarrow \pi_J = \frac{1}{\sum_{i=1}^{J} e^{\alpha_i+\beta_i x}} = \frac{1}{1+\sum_{i=1}^{J-1} e^{\alpha_i+\beta_i x}}$$

To find $\pi_k = P(Y=k)$, we know

$$\frac{\pi_k}{\pi_J} = e^{\alpha_k+\beta_k x} \Leftrightarrow e^{\alpha_k+\beta_k x}\pi_J$$

$$\pi_k = \frac{e^{\alpha_k+\beta_k x}}{\sum_{i=1}^{J} e^{\alpha_i+\beta_i x}}$$

Noting that $\alpha_J = \beta_J = 0 \Rightarrow$ $$\pi_k = \frac{e^{\alpha_k+\beta_k x}}{1+\sum_{i=1}^{J-1} e^{\alpha_i+\beta_i x}}$$