Generalized logit models for nominal multinomial responses

Categorical Data Analysis, Summer 2015

Local odds ratios

		Y: 1	2	3	4	
X	1	π_{11}	π_{12}	π_{13}	π_{14}	π_{1+}
	2	π_{21}	π_{22}	π_{23}	π_{24}	π_{2+}
	3	π_{31}	π_{32}	π_{33}	π_{34}	π_{3+}

- Odds of $Y = 4$ versus $Y = 2$ when $X = 1$ is $(\pi_{14}/\pi_{1+})\,/\,(\pi_{12}/\pi_{1+}) = \pi_{14}/\pi_{12}$
- Odds of $Y = 4$ versus $Y = 2$ when $X = 3$ is $(\pi_{34}/\pi_{3+})\,/\,(\pi_{32}/\pi_{3+}) = \pi_{34}/\pi_{32}$
- Local odds ratio = $(\pi_{14}/\pi_{12})\,/\,(\pi_{34}/\pi_{32}) = (\pi_{14}\pi_{34})\,/\,(\pi_{12}\pi_{32})$
- *Interpretation*: If local OR = 2, "There is a two-fold increase in the odds of a response, Y, in class 4 versus class 2 when comparing $X = 1$ to $X = 3$."

Multinomial regression models for nominal response

- Let Y be a categorical response variable with J categories ($J > 2$)
- We desire a model for multinomial responses similar to a logistic regression model
  - Y could be the location of a colorectal tumor (proximal, distal or rectal)
  - $\mathbf{X}$ could be the covariate classes defined by a subject's race (AA or non-AA) and gender (male or female)
- Let $\pi_j(\mathbf{x}) = P(Y = j|\mathbf{x})$ for some fixed setting of the $\mathbf{x}$ explanatory variables, with $\sum_j \pi_j(\mathbf{x}) = 1$
- At this fixed setting of $\mathbf{x}$ we treat the counts at the J categories of Y as multinomial with probabilities $\{\pi_1(\mathbf{x}), \ldots, \pi_J(\mathbf{x})\}$.

Baseline-category logits

- We select one of the J categories of Y as the baseline (or reference) category
- Without loss of generality, order the categories of Y so the Jth level coincides with this baseline category
- Define the *generalized logit* (relative to the baseline category) as

$$g_j(\mathbf{x}) = \log\left[\frac{\pi_j(\mathbf{x})}{\pi_J(\mathbf{x})}\right] = \alpha_j + \boldsymbol{\beta}_j'\mathbf{x},\ j = 1, \ldots, J-1$$

- *This model defines $J-1$ sets of model parameters, one for each of the $J-1$ generalized logits.*
- Therefore, for each logit we have
  - A separate intercept (α_j)
  - A separate set of regression parameters ($\boldsymbol{\beta}_j$)

Multinomial likelihood

- Consider subject i's contribution to the log-likelihood

$$L_i = \log\left(\prod_{j=1}^{J} \pi_{ij}^{z_{ij}}\right)$$

  - $\pi_{ij} = P(Y_i = j)$
  - $z_{ij} = 1$ if $Y_i = j$ and $z_{ij} = 0$ if $Y_i \neq j$
  - $\mathbf{z}_i = (z_{i1}, \ldots, z_{iJ})$ is a vector of a single 1 and the rest 0

$$\begin{aligned} L_i &= \sum_{j=1}^{J} z_{ij} \log \pi_{ij} = \sum_{j=1}^{J-1} z_{ij} \log \pi_{ij} + z_{iJ} \log \pi_{iJ} \\ &= \sum_{j=1}^{J-1} z_{ij} \log \pi_{ij} + \left(1 - \sum_{j=1}^{J-1} z_{ij}\right) \log \pi_{iJ} \\ &= \sum_{j=1}^{J-1} z_{ij} \log \frac{\pi_{ij}}{\pi_{iJ}} + \log \pi_{iJ} \end{aligned}$$

Multinomial likelihood (cont.)

Conclusions:

1. The multinomial distribution is a member of the multivariate exponential dispersion family
2. The baseline-category logits are the natural parameters for the multinomial distribution
3. The baseline-category logit functions are the canonical link functions for the multinomial GLM

Inverting generalized logits to obtain probabilities

Recall that for covariate pattern $\mathbf{x}$, we define the generalized logit as

$$g_j(\mathbf{x}) = \log\left[\frac{\pi_j(\mathbf{x})}{\pi_J(\mathbf{x})}\right] = \alpha_j + \boldsymbol{\beta}_j'\mathbf{x},\ j = 1, \ldots, J-1.$$

These can be solved for the individual $\pi_j(\mathbf{x})$ yielding

$$\begin{aligned}
(i)\ \ \pi_j(\mathbf{x}) &= P(Y = j|\mathbf{x}) = \frac{\exp\{g_j(\mathbf{x})\}}{1 + \sum_{j=1}^{J-1} \exp\{g_j(\mathbf{x})\}},\ j = 1, \ldots, J-1 \\
(ii)\ \ \pi_J(\mathbf{x}) &= P(Y = J|\mathbf{x}) = \frac{1}{1 + \sum_{j=1}^{J-1} \exp\{g_j(\mathbf{x})\}}
\end{aligned}$$

Example

Let Y be a three-level categorical variable with the third level identified as the reference category.

- $g_1(\mathbf{x}) = \alpha_1 + \boldsymbol{\beta}_1'\mathbf{x}$
- $g_2(\mathbf{x}) = \alpha_2 + \boldsymbol{\beta}_2'\mathbf{x}$
- *There is no $g_3(\mathbf{x})$ - the third level of Y is the reference category.*

$$\begin{aligned}
\pi_1(\mathbf{x}) &= P(Y = 1|\mathbf{x}) = \frac{\exp\{\alpha_1 + \boldsymbol{\beta}_1'\mathbf{x}\}}{1 + \exp\{\alpha_1 + \boldsymbol{\beta}_1'\mathbf{x}\} + \exp\{\alpha_2 + \boldsymbol{\beta}_2'\mathbf{x}\}} \\
\pi_2(\mathbf{x}) &= P(Y = 2|\mathbf{x}) = \frac{\exp\{\alpha_2 + \boldsymbol{\beta}_2'\mathbf{x}\}}{1 + \exp\{\alpha_1 + \boldsymbol{\beta}_1'\mathbf{x}\} + \exp\{\alpha_2 + \boldsymbol{\beta}_2'\mathbf{x}\}} \\
\pi_3(\mathbf{x}) &= P(Y = 3|\mathbf{x}) = \frac{1}{1 + \exp\{\alpha_1 + \boldsymbol{\beta}_1'\mathbf{x}\} + \exp\{\alpha_2 + \boldsymbol{\beta}_2'\mathbf{x}\}}
\end{aligned}$$

Deriving probabilities

- Consider

$$\log\left(\frac{\pi_j(\mathbf{x})}{\pi_J(\mathbf{x})}\right) = \alpha_j + \beta_j'\mathbf{x}$$

- Exponentiating both sides, we get

$$\frac{\pi_j(\mathbf{x})}{\pi_J(\mathbf{x})} = \exp\left(\alpha_j + \beta_j'\mathbf{x}\right)$$

which is the (local) odds for category j versus category J.

- Multiplying both sides by $\pi_J(\mathbf{x})$, we obtain

$$\pi_j(\mathbf{x}) = \pi_J(\mathbf{x}) \exp\left(\alpha_j + \beta_j'\mathbf{x}\right) \quad (1),$$

- Now, sum both sides over $j = 1, ..., J-1$,

$$\sum_{j=1}^{J-1} \pi_j(\mathbf{x}) = \pi_J(\mathbf{x}) \sum_{j=1}^{J-1} \exp\left(\alpha_j + \beta_j'\mathbf{x}\right) \quad (2),$$

Deriving probabilities (cont.)

- Note that

$$\pi_J(\mathbf{x}) + \sum_{j=1}^{J-1} \pi_j(\mathbf{x}) = \sum_{j=1}^{J} \pi_j(\mathbf{x}) = 1$$

- It follows that

$$\sum_{j=1}^{J-1} \pi_j(\mathbf{x}) = 1 - \pi_J(\mathbf{x}) \quad (3)$$

- Based on (3), we can substitute $1 - \pi_J(\mathbf{x})$ for $\sum_{j=1}^{J-1} \pi_j(\mathbf{x})$ in (2) on Slide 9, resulting in

$$1 - \pi_J(\mathbf{x}) = \pi_J(\mathbf{x}) \sum_{j=1}^{J-1} \exp\left(\alpha_j + \beta_j'\mathbf{x}\right) \quad (4)$$

Deriving probabilities (cont.)

- Rearranging terms in (4) on Slide 10, we have

$$1 \quad = \quad \pi_J(\mathbf{x}) \left[1 + \sum_{j=1}^{J-1} \exp\left(\alpha_j + \beta_j'\mathbf{x}\right)\right]$$

- Solving for $\pi_J(\mathbf{x})$, we have

$$\pi_J(\mathbf{x}) = \frac{1}{1 + \sum_{j=1}^{J-1} \exp\left(\alpha_j + \beta_j'\mathbf{x}\right)}$$

- Recalling (1) from Slide 9

$$\pi_j(\mathbf{x}) = \pi_J(\mathbf{x}) \exp\left(\alpha_j + \beta_j'\mathbf{x}\right),$$

we substitute our expression for $\pi_J(\mathbf{x})$ into (1) to obtain

$$\pi_j(\mathbf{x}) = \frac{\exp\left(\alpha_j + \beta_j'\mathbf{x}\right)}{1 + \sum_{j=1}^{J-1} \exp\left(\alpha_j + \beta_j'\mathbf{x}\right)}$$

Obtaining other odds ratios

- Note that the $J - 1$ baseline-category logits uniquely determine all remaining logits comparing any two response levels
- Consider the logit for category j versus j' where $j \neq j'$ and $j' \neq J$

$$\begin{aligned} \log\left(\frac{\pi_j(\mathbf{x})}{\pi_{j'}(\mathbf{x})}\right) &= \log\left(\frac{\pi_j(\mathbf{x})}{\pi_J(\mathbf{x})}\right) - \log\left(\frac{\pi_{j'}(\mathbf{x})}{\pi_J(\mathbf{x})}\right) \\ &= g_j(\mathbf{x}) - g_{j'}(\mathbf{x}) \\ &= (\alpha_j + \beta_j'\mathbf{x}) - (\alpha_{j'} + \beta_{j'}'\mathbf{x}) \end{aligned}$$

CRC tumor location example

- Y_i = tumor location for ith colorectal cancer (CRC) patient (1 = proximal, 2 = distal, 3 = rectal)
- X_{1i} = race for ith CRC patient (1 = AA or 0 = non-AA)
- X_{2i} = gender for ith CRC patient (1 = male or 0 = female)
- Using rectal tumor location as the reference category, we fit the following generalized logits:

$$
\begin{aligned}
g_1(\mathbf{x}_i) &= \alpha_1 + \beta_{11}x_{1i} + \beta_{12}x_{2i} \\
g_2(\mathbf{x}_i) &= \alpha_2 + \beta_{21}x_{1i} + \beta_{22}x_{2i}
\end{aligned}
$$

For the jth logit ($j = 1, 2$)

- α_j is the intercept
- β_{j1} is the effect of subject's race
- β_{j2} is the effect of subject's gender

CRC example: log odds

Covariate classes		proximal / rectal	distal / rectal	proximal / distal
AA	Male	$\alpha_1 + \beta_{11} + \beta_{12}$	$\alpha_2 + \beta_{21} + \beta_{22}$	$(\alpha_1 + \beta_{11} + \beta_{12}) - (\alpha_2 + \beta_{21} + \beta_{22})$
	Female	$\alpha_1 + \beta_{11}$	$\alpha_2 + \beta_{21}$	$(\alpha_1 + \beta_{11}) - (\alpha_2 + \beta_{21})$
non-AA	Male	$\alpha_1 + \beta_{12}$	$\alpha_2 + \beta_{22}$	$(\alpha_1 + \beta_{12}) - (\alpha_2 + \beta_{22})$
	Female	α_1	α_2	$\alpha_1 - \alpha_2$

CRC example: log odds ratios

Calculate the log odds ratio for proximal versus rectal CRC comparing AAs to non-AAs, controlling for subject's gender.

- log odds of proximal versus rectal CRC for AA males is

$$\alpha_1 + \beta_{11} + \beta_{12}$$

- log odds of proximal versus rectal CRC for non-AA males is

$$\alpha_1 + \beta_{12}$$

- log odds ratio of proximal versus rectal CRC for AA males compared to non-AA males is

$$\beta_{11}$$

- Therefore, odds ratio of proximal versus rectal CRC for AA males compared to non-AA males is

$$\exp\{\beta_{11}\}$$

CRC example: log odds ratios (cont.)

Calculate the log odds ratio for proximal versus rectal CRC comparing AAs to non-AAs, controlling for subject's gender.

- log odds of proximal versus rectal CRC for AA females is

$$\alpha_1 + \beta_{11}$$

- log odds of proximal versus rectal CRC for non-AA females is

$$\alpha_1$$

- log odds ratio of proximal versus rectal CRC for AA females compared to non-AA females is

$$\beta_{11}$$

- Therefore, odds ratio of proximal versus rectal CRC for AA females compared to non-AA females is

$$\exp\{\beta_{11}\}$$

CRC example: log odds ratios (cont.)

- Not surprisingly, the odds ratios for proximal versus rectal CRC comparing AAs to non-AAs was the same for males and females
- This is because there was no interaction in the model
- That is to say, we assume homogeneity of the local odds ratios
- Other ORs of interest can be calculated by exponentiating differences of appropriately selected log odds