3. (a) Let $Y_i \sim Cat(\theta_{i1}, \theta_{i2}, \theta_{i3})$, and consider the multinomial logit models

$$\log\left(\frac{\theta_{i2}}{\theta_{i1}}\right) = \mathbf{x}_i'\boldsymbol{\beta}_2,$$
$$\log\left(\frac{\theta_{i3}}{\theta_{i1}}\right) = \mathbf{x}_i'\boldsymbol{\beta}_3,$$

where $\theta_{i1} + \theta_{i2} + \theta_{i3} = 1$. Show that

$$\theta_{i1} = \frac{1}{1 + e^{\mathbf{x}_i'\boldsymbol{\beta}_2} + e^{\mathbf{x}_i'\boldsymbol{\beta}_3}},$$
$$\theta_{i2} = \frac{e^{\mathbf{x}_i'\boldsymbol{\beta}_2}}{1 + e^{\mathbf{x}_i'\boldsymbol{\beta}_2} + e^{\mathbf{x}_i'\boldsymbol{\beta}_3}},$$
$$\theta_{i3} = \frac{e^{\mathbf{x}_i'\boldsymbol{\beta}_3}}{1 + e^{\mathbf{x}_i'\boldsymbol{\beta}_2} + e^{\mathbf{x}_i'\boldsymbol{\beta}_3}}.$$

(2 points)

Solution:

The equations $\log\left(\frac{\theta_{i2}}{\theta_{i1}}\right) = \mathbf{x}_i'\boldsymbol{\beta}_2$ and $\log\left(\frac{\theta_{i3}}{\theta_{i1}}\right) = \mathbf{x}_i'\boldsymbol{\beta}_3$ are equivalent to the equations

$$\frac{\theta_{i2}}{\theta_{i1}} = \exp(\mathbf{x}_i'\boldsymbol{\beta}_2), \qquad \frac{\theta_{i3}}{\theta_{i1}} = \exp(\mathbf{x}_i'\boldsymbol{\beta}_3),$$

and hence

$$\theta_{i2} = \theta_{i1}\exp(\mathbf{x}_i'\boldsymbol{\beta}_2), \qquad \theta_{i3} = \theta_{i1}\exp(\mathbf{x}_i'\boldsymbol{\beta}_3).$$

Since $\theta_{i1} + \theta_{i2} + \theta_{i3} = 1$, we have $\theta_{i1}(1 + \exp(\mathbf{x}_i'\boldsymbol{\beta}_2) + \exp(\mathbf{x}_i'\boldsymbol{\beta}_3)) = 1$, i.e.,

$$\theta_{i1} = \frac{1}{1 + \exp(\mathbf{x}_i'\boldsymbol{\beta}_2) + \exp(\mathbf{x}_i'\boldsymbol{\beta}_3)},$$

and then

$$\theta_{i2} = \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta}_2)}{1 + \exp(\mathbf{x}_i'\boldsymbol{\beta}_2) + \exp(\mathbf{x}_i'\boldsymbol{\beta}_3)},$$
$$\theta_{i3} = \frac{\exp(\mathbf{x}_i'\boldsymbol{\beta}_3)}{1 + \exp(\mathbf{x}_i'\boldsymbol{\beta}_2) + \exp(\mathbf{x}_i'\boldsymbol{\beta}_3)}.$$

(b) Let the random variable Y_i be defined on ordinal scale with m distinctive possible outcomes. Let the possible outcomes have natural order "1" < "2" < "3". Consider cumulative proportional odds logit model

$$\log\left(\frac{P(Y_i \leq k)}{1 - P(Y_i \leq k)}\right) = \text{logit}\left(P(Y_i \leq k)\right) = \beta_{0k} + \beta_1 x_{i1}, \qquad k = 1, 2.$$

Solve the probabilities $P(Y_i = 1), P(Y_i = 2), P(Y_i = 3)$ as functions of parameters β_{0k}, β_1.

(2 points)

Solution:

Now

$$P(Y_i = 1) = P(Y_i \leq 1) = \frac{\exp(\beta_{01} + \beta_1 x_{i1})}{1 + \exp(\beta_{01} + \beta_1 x_{i1})},$$
$$P(Y_i \leq 2) = \frac{\exp(\beta_{02} + \beta_1 x_{i1})}{1 + \exp(\beta_{02} + \beta_1 x_{i1})}.$$

Since $P(Y_i \leq 2) = P(Y_i = 1) + P(Y_i = 2)$, we have

$$P(Y_i = 2) = \frac{\exp(\beta_{02} + \beta_1 x_{i1})}{1 + \exp(\beta_{02} + \beta_1 x_{i1})} - \frac{\exp(\beta_{01} + \beta_1 x_{i1})}{1 + \exp(\beta_{01} + \beta_1 x_{i1})}.$$

Further, since $P(Y_i \leq 3) = 1 - P(Y_i \leq 2)$, we have

$$P(Y_i = 3) = 1 - \frac{\exp(\beta_{02} + \beta_1 x_{i1})}{1 + \exp(\beta_{02} + \beta_1 x_{i1})}.$$

(c) Let Y_i be such a random variable that for known n_i value, the product $n_i Y_i$ follows the binomial distribution $n_i Y_i \sim Bin(n_i, \mu_i)$. Derive with help of the properties of the binomial distribution what are the expected value $\mathrm{E}(Y_i)$ and the variance $\mathrm{Var}(Y_i)$ of the random variable Y_i.

(2 points)

Solution:

If we denote $Z_i = n_i Y_i \sim Bin(n_i, \mu_i)$, the it is well known that

$$\mathrm{E}(Z_i) = n_i \mu_i, \qquad \mathrm{Var}(Z_i) = n_i \mu_i (1 - \mu_i).$$

Since $Y_i = \frac{Z_i}{n_i}$, we have

$$\mathrm{E}(Y_i) = \mathrm{E}\left(\frac{Z_i}{n_i}\right) = \frac{1}{n_i} \cdot \mathrm{E}\left(Z_i\right) = \frac{1}{n_i} \cdot n_i \mu_i = \mu_i,$$
$$\mathrm{Var}(Y_i) = \mathrm{Var}\left(\frac{Z_i}{n_i}\right) = \left(\frac{1}{n_i}\right)^2 \cdot \mathrm{Var}\left(Z_i\right) = \frac{1}{n_i^2} \cdot n_i \mu_i (1 - \mu_i) = \frac{\mu_i (1 - \mu_i)}{n_i}.$$

Another way to obtain the results is to see that the density function of the random variable Y_i can be written as

$$f(y_i | \mu_i, n_i) = \binom{n_i}{n_i y_i} \pi_i^{n_i y_i} (1 - \mu_i)^{(n_i - n_i y_i)}, \quad n_i y_i = 0, 1, 2, \ldots, n_i,$$
$$= \exp\left(\frac{y_i \Theta_i - \log(1 + e^{\Theta_i})}{\frac{1}{\phi}} + c(y_i, \phi)\right)$$

where

$$\Theta_i = \log\left(\tfrac{\mu_i}{1-\mu_i}\right),\ \phi = n_i, a(\phi) = \frac{1}{\phi},\ b(\Theta_i) = \log(1+e^{\Theta_i}),\ c(y_i, \phi) = \binom{\phi}{\phi y_i}.$$

Hence the expected value $\mathrm{E}(Y_i)$ and the variance $\mathrm{Var}(Y_i)$ of the random variable Y_i are

$$\begin{aligned}
\mathrm{E}(Y_i) &= b'(\Theta_i) = \frac{e^{\Theta_i}}{1+e^{\Theta_i}} = \mu_i,\\
\mathrm{Var}(Y_i) &= b''(\Theta_i)a(\phi) = \frac{e^{\Theta_i}}{(1+e^{\Theta_i})^2} \cdot \frac{1}{\phi} = \mu_i(1-\mu_i) \cdot \frac{1}{n_i} = \frac{\mu_i(1-\mu_i)}{n_i}.
\end{aligned}$$