Qs. 2(b)

AI23MTECH14004

AI23MTECH12002

2.(b)

Suppose, Y is the dependent variable, there are K ordered categories (1 to K) and p independent variable $x_1 \ldots x_p$.

As per the proportional odds model,

it assumes that on observations cumulative odds are falling into a category with a value $\leq k$ can be described as a linear combination of dependent variables. [Where k ranging from 1 to $K-1$]

$$\log\left[\frac{\mathbb{P}(Y \leq k)}{1-\mathbb{P}(Y \leq k)}\right] = \alpha_k + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p$$

[Where → $\mathbb{P}(Y \leq k)$ is the cumulative probability that $Y \leq K$.

→ α_K is the threshold parameter associated with category k

→ $\beta_1, \beta_2 \ldots \beta_p$ are the coefficients of indep variables.

Cumulative logistic distribution

$$\mathbb{P}(Y_i \le k) = \frac{1}{1 + \text{Exp}(-(\alpha_k + \beta x_{i,1} + \beta_2 x_{i,2} + \dots \beta_p x_{i,p}))}$$

$$\text{link}(y_{ij}) = \frac{\theta_j - \beta^T x_i}{\tau_i}$$

[$\beta^T x_i \rightarrow$ location of ith row

$\tau_i \rightarrow$ scalar for ith row]

$\therefore \log \tau_i = \gamma^T (x_i - \bar{x})$

[$\tau \rightarrow$ vector of unknown parameters]

So, For cumulative probabilities, we can define.

$R_1 = n_1$ $\qquad$ $z_1 = R_1/n$

$R_2 = n_1 + n_2$ $\qquad$ $z_2 = R_2/n$

$\vdots$ $\qquad$ $\vdots$

$R_K = \Sigma n_j = n$ $\qquad$ $z_K = R_K/n = 1$

$\therefore$ Considering parameters of the function, likelihood can be-

$$\left\{\left(\frac{y_2}{y_1}\right)^{R_1}\left(\frac{y_2-y_1}{y_2}\right)^{R_2-R_1}\right\}\left\{\left(\frac{y_2}{y_3}\right)^{R_2}\left(\frac{y_3-y_2}{y_3}\right)^{R_3-R_2}\right\}$$

$$\cdots \left\{\left(\frac{y_{k-1}}{y_k}\right)^{R_{k-1}}\left(\frac{y_k-y_{k-1}}{y_k}\right)^{R_k-R_{k-1}}\right\}$$

(product of $k-1$ quantities)

These factors are the probability that the first two cells divide in the ratio (R_1) $R_1 : R_2 - R_1$

The probability R_3 in cell 3 relative to cells 1 and 2 combined is

$$R_3 - R_2 : R_2$$

$\therefore$ Now we can define

$$\phi_j = \log\left(\frac{y_i}{y_{j+1} - y_j}\right) = \text{logit}\left(\frac{y_j}{y_{j+1}}\right)$$

$$g(\phi) = \log\left(1 + \operatorname{Exp}(\phi)\right)$$

$$= \log\left(I + \operatorname{Exp}\left(\log\left(\frac{y_j}{y_{j+1} - y_j}\right)\right)\right)$$

$$= \log\left(I + \frac{y_i}{y_{j+1} - y_j}\right)$$

$$= \log\left(\frac{y_{j+1} - y_j + y_i}{y_{j+1} - y_j}\right)$$

$$= \log\left(\frac{y_{j+1}}{y_{j+1} - y_j}\right)$$

$\therefore$ log likelihood would be —

$$= n\left\{\left[z_1 \phi_1 - z_2\, g(\phi_1)\right] + \left[z_2 \phi_2 - z_3\, g(\phi_2)\right] + \dots \left[z_{k-1} \phi_{k-1} - g(\phi_{k-1})\right]\right\}$$