11.05.2018

Multinomial distribution

$$X = (x_1, x_2 \ldots, x_K) \sim \text{Multinomial}(n, \phi = (\phi_1, \phi_2 \ldots, \phi_K))$$

pmf: $$P(X/\theta) = \frac{n!}{x_1! \cdots x_k!} \phi_1^{x_1} \cdot \phi_2^{x_2} \cdots \phi_K^{x_K}$$

$$\phi_K = 1 - (\phi_1 + \phi_2 + \cdots + \phi_{K-1})$$

$$P(X|\theta) = \frac{n!}{x_1! \cdots x_K!} \exp(\log(\phi_1^{x_1} \cdot \phi_2^{x_2} \cdots, \phi_K^{x_K}))$$

$$= \frac{n!}{x_1! \cdots x_K!} \exp\Big(x_1 \log \phi_1 + x_2 \log \phi_2 \ldots + x_{K-1} \log \phi_{K-1} + (n - \sum_{j=1}^{K-1} x_j) \log(1 - \sum_{j=1}^{K-1} \phi_j)\Big)$$

$$= \frac{n!}{x_1! \cdots x_K!} \exp\Big(x_1 \log \phi_1 + \cdots + x_{K-1} \log \phi_{K-1} + n \log(1 - \sum_{j=1}^{K-1} \phi_j) - \sum_{j=1}^{K-1} x_j (\log(1 - \sum_{j=1}^{K-1} \phi_j))\Big)$$

$$= \frac{n!}{x! \cdots x_K!} \exp\Big(x_1 \log \frac{\phi_1}{\phi_K} + \cdots + x_{K-1} \log \frac{\phi_{K-1}}{\phi_K} + n \log(1 - \phi_K)\Big)$$

$$b(x) = \frac{n!}{x_1! \cdots x_K!} \qquad a(\eta) = -n \log(1 - \phi_K)$$

~~T(y) =~~

$$\eta = \begin{bmatrix} \log \frac{\phi_1}{\phi_K} \\ \vdots \\ \log \frac{\phi_{K-1}}{\phi_K} \end{bmatrix}_{(K-1) \times 1}$$

$$T(1) = \begin{bmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(K-1) \times 1}, \quad T(2) = \begin{bmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{bmatrix}, \quad T(K-1) = \begin{bmatrix} 0 \\ \vdots \\ x_{K-1} \end{bmatrix}$$

$$T(K) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Scanned by CamScanner

$$T = \begin{bmatrix} x_1 & 0 & & 0 \\ 0 & x_2 & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & x_{k-1} \end{bmatrix}_{(k-1)\times(k-1)}$$

$$\eta^T \cdot y = \underbrace{\begin{bmatrix} \log\frac{\phi_1}{\phi_k} & \log\frac{\phi_2}{\phi_k} & \cdots & \log\frac{\phi_{k-1}}{\phi_k} \end{bmatrix}}_{1\times(k-1)} \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & x_{k-1} \end{bmatrix}_{(k-1)\times(k-1)}$$

$\eta^T \cdot y \rightarrow 1\times(k-1)$

$\hookrightarrow \log\frac{\phi_1}{\phi_k} x_1 + \log\frac{\phi_2}{\phi_k} \cdot x_2 \cdots + \log\frac{\phi_{k-1}}{\phi_k} x_{k-1}$

How are the constants defined for multinomial regression ?

$y = (x_1=1, x_2=1, \ldots, x_k=1) \sim \text{Multinomial}(\phi_1, \phi_2, \ldots, \phi_k)$

$n = 1$

$b(y) = 1$

$$\eta = \begin{bmatrix} \log\frac{\phi_1}{\phi_k} \\ \vdots \\ \log\frac{\phi_{k-1}}{\phi_k} \end{bmatrix}_{(k-1)\times 1} \qquad T = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}_{(k-1)\times(k-1)}$$

$a(\eta) = -\log(\phi_k)$

Scanned by CamScanner

$$\eta = \begin{bmatrix} \log \frac{\phi_1}{\phi_k} \\ \vdots \\ \log \frac{\phi_{k-1}}{\phi_k} \end{bmatrix} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_{k-1} \end{bmatrix}_{(k-1) \times 1}$$

$$\log \frac{\phi_1}{\phi_k} = \eta_1 \Rightarrow \phi_1 = \phi_k e^{\eta_1}$$

$$\log \frac{\phi_i}{\phi_k} = \eta_i \Rightarrow \phi_i = \phi_k e^{\eta_i}$$

$$\vdots$$

$$\log \frac{\phi_{k-1}}{\phi_k} = \eta_{k-1} \Rightarrow \phi_{k-1} = \phi_k e^{\eta_{k-1}}$$

$$\phi_1 + \dots + \phi_i + \dots + \phi_{k-1} = \phi_k (e^{\eta_1} + \dots + e^{\eta_{k-1}})$$

$$\phi_k = 1 - \sum_{j=1}^{k-1} \phi_j$$

$$1 - \phi_k = \phi_k (e^{\eta_1} + \dots + e^{\eta_{k-1}})$$

$$1 = \phi_k (e^{\eta_1} + \dots + e^{\eta_{k-1}} + 1)$$

$$\phi_k = \frac{1}{1 + \sum_{j=1}^{k-1} e^{\eta_j}}$$

so, $$\phi_i = \frac{e^{\eta_i}}{1 + \sum_{j=1}^{k-1} e^{\eta_j}}$$

Derive Generalized linear model for multinomial distribution?

Assumption-1: $y|x;\theta \sim \text{Multinomial}(\phi_1, \phi_2, \ldots, \phi_k)$

Assume-2: $h_\theta(x) = E[T(y)|x;\theta] = E[T(y)] = \phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{k-1} \end{bmatrix}$

Assume-3: $\eta = \theta^T x$

$$\phi_i = \frac{e^{\eta_i}}{1+\sum_{j=1}^{k-1} e^{\eta_j}} \quad \forall i \in \{1, \ldots, k-1\}$$

$$\phi_i = \frac{e^{\theta_i^T x}}{1+\sum_{j=1}^{k-1} e^{\theta_j^T x}}$$

$$h_\theta(x) = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{k-1} \end{bmatrix} = \begin{bmatrix} \dfrac{e^{\theta_1^T x}}{1+\sum_{j=1}^{k-1} e^{\theta_j^T x}} \\ \vdots \\ \dfrac{e^{\theta_{k-1}^T x}}{1+\sum_{j=1}^{k-1} e^{\theta_j^T x}} \end{bmatrix}_{(k-1)\times 1}$$