1

1/05/16 ALSM 2 EXAM NOTES: MULTINOMIAL DISTRIBUTION

Multinomial Distribution
- Generalisation/Extension to Binomial Distribution
- Binomial is a joint probability distribution

$P(y_1, \ldots, y_J \mid \theta_1, \ldots, \theta_J, n) = \frac{n!}{y_1! \ldots y_J!} \theta_1^{y_1} \theta_2^{y_2} \ldots \theta_J^{y_J}$

J: # of categories
θ_i are the respective probabilities of the categories $\theta_1 + \ldots + \theta_J = 1$
$n = y_1 + y_2 + \ldots + y_J$

When $J = 2$ $P(y_1, y_2 \mid \theta_1, \theta_2) = \frac{n!}{y_1! y_2!} \theta_1^{y_1} \theta_2^{y_2}$ $n = y_1 + y_2$ $\theta_1 + \theta_2 = 1$

$P(y_i \mid \theta_i, n) = \frac{n!}{y!(n-y)!} \theta^y (1-\theta)^{n-y}$ Binomial Distribution

- Even if multinomial is not a member of the exponential family, we can control $\vec{\theta}_{\vec{y}}$ collected over N groups via a set of parameters $\vec{\beta}$

Nominal Logistic Regression
The outcomes of experiments are in J categories and there is no natural order amongst the response categories. One category is arbitrarily chosen as a reference category e.g. θ_1. Then the logits for the other categories are defined by:

$Logit(\theta_j) = \left[\frac{\theta_j}{\theta_1}\right] = x^T \beta_j \quad \forall j = 2, \ldots, J$

having the constraint $\Sigma_{j=1}^{J} \theta_j = 1$
When the estimates $\hat{\beta}_j$ are computed, then

$\hat{\theta}_j = \hat{\theta}_1 \exp(x^T \hat{\beta}_j) \quad \forall j = 2, \ldots, J$

$\hat{\theta}_1 = \frac{1}{1 + \Sigma_{j=2}^{J} \exp(x^T \hat{\beta}_j)}$

or $\hat{\theta}_j = \frac{\exp(x^T \hat{\beta}_j)}{1 + \Sigma_{j=2}^{J} \exp(x^T \hat{\beta}_j)} \quad \forall j = 2, \ldots, J$ Softmax function → converting output to probability

$\hat{\theta}_1 + \hat{\theta}_1 \exp(x^T \hat{\beta}_2) + \ldots + \hat{\theta}_1 \exp(x^T \hat{\beta}_J) = 1$

Alligators Example

Linear model linking proportions to explanatory variables:

$\log\left[\frac{\theta_j}{\theta_1}\right] = \beta_j^{\text{intercept}} + \beta_j^{\text{sex}} + \beta_j \text{Size} + \beta_j L1 + \beta_j L2 + \beta_j L3 \quad \forall j = 2, \dots, 5$

$\theta_j = \frac{\exp(x^T \hat{\beta}_j)}{1 + \sum_{j=2}^{T} \exp(x^T \hat{\beta}_j)} \quad \forall j = 2, \dots, 5$ with $x = [1, \text{sex}, \text{size}, L1, L2, L3]$

$\beta_j = [\beta_j^{\text{intercept}}, \beta_j^{\text{sex}}, \beta_j^{\text{size}}, \beta_j^{L1}, \beta_j^{L2}, \beta_j^{L3}]$

$\theta_1 = \frac{1}{1 + \sum_{j=2}^{T} \exp(x^T \hat{\beta}_j)}$

Function to maximise

$L = \prod_{i=1}^{16} P(y_{1i}, y_{2i}, y_{3i}, y_{4i}, y_{5i} \mid \theta_{1i}, \theta_{2i}, \theta_{3i}, \theta_{4i}, \theta_{5i})$

$n_i = y_{1i} + y_{2i} + y_{3i} + y_{4i} + y_{5i}$ → group sizes can be different

$\theta_{1i} = \frac{1}{1 + \sum_{j=2}^{2} \exp(x_i^T \hat{\beta})}$

$\theta_{ji} = \frac{\exp(x^T_i \hat{\beta}_i)}{1 + \sum_{j=2}^{T} \exp(x^T_i \hat{\beta}_j)}$

$L = \prod_{i=1}^{16} \frac{n_i}{y_{1i}! y_{2i}! y_{3i}! y_{4i}! y_{5i}!} \theta_{1i}^{y_{1i}} \theta_{2i}^{y_{2i}} \theta_{3i}^{y_{3i}} \theta_{4i}^{y_{4i}} \theta_{5i}^{y_{5i}} \quad df = 4 \times 16$

saturated model

$L(\theta_{1i}, \theta_{2i}, \theta_{3i}, \theta_{4i}, \theta_{5i}) \quad df = 6 \times 4$

$L(\{\beta_j\}\ j = 2, \dots, 5)$

ODDS RATIO

i = group $\quad j$ = vote choice

OR (example) = $\theta_{j=2, i=1} / \theta_{j=1, i=1}$ ratio for group 1 women (18-23) Sex = 0 Age1 = 0 Age2 = 0

$\theta_{j=2, i=4} / \theta_{j=1, i=4}$ ratio for group 2 men (18-23) Sex = 1 Age1 = 0 Age2 = 0

Category 2 = y_2 $\quad j=1$ reference category

How to use exp[cov...] to get OR ratio.

1.5 ⇒ indicates the option is more important to women than men.

- Assumed use of Softmax function