Ordinal Probit, Multinomial logit, Conditional Poisson

Elad Zippory

September, 14, 2015

1 Ordinal Probit

Y^* is the continuous latent that is modelled by $x'\beta + \epsilon$.

$$Y = \begin{cases} 1 & \text{if } y^* < 0 \\ 2 & \text{if } y^* \in (0, c_2) \\ 3 & \text{if } y^* \in (c_2, c_3) \\ 4 & \text{if } y^* > c_3 \end{cases}$$

$$P(y = j) = P(y > (j-1)) - P(y > j)$$

by construction $P(y > j) = \Phi(c_j - x\beta)$
Since the density is

$$P(y = 1) \times \cdots \times P(y = J) = \prod_{i=1}^{n} \prod_{j=1}^{J} P(y_i = j)$$

Using the link function:

$$\prod_{i=1}^{n} \prod_{j=1}^{J} \left[\Phi(c_j - x\beta) - \Phi(c_{j-1} - x\beta)\right]^{z_{i,j}}$$

s.t. $z_{i.j} = 1$ if $y_i = j$ and 0 otherwise. The log likelihood becomes:

$$\sum_{i=1}^{n} \sum_{j=1}^{J} z_{i,j} log \left[\Phi(c_j - x\beta) - \Phi(c_{j-1} - x\beta)\right]$$

s.t. $c_1 < \cdots < c_J$

This model is the Proportional odds ratios as it assumes the same slopes per category. The cumulative logit relaxes that:

$$\frac{P(y \leq \tau_j)}{P(y > J)} = \frac{\pi_1 + \cdots + \pi_j}{\pi_{j+1} + \cdots + \pi_J}$$

and taking the log will get you $x'\beta_j$ instead of $x'\beta$.

2 multinomial logit

We want to get $P(Y=j|x,\beta)=\pi_j \propto exp[x\beta]$ s.t. $0 \leq \pi_j \leq 1$ and $\sum \pi = 1$.
So we construct

$$\pi_{i,j} = \frac{exp(x_i\beta_j)}{\sum_k exp(x_i\beta_k)}$$

which is not identified. So we set category 1 as a benchmark by setting $\beta_1 = [0,\ldots 0]'$.
So,

$$\pi_1 = \frac{1}{1+\sum exp(x_i\beta_k)}$$

and

$$\pi_{j\neq 1} = \frac{exp(x\beta_j)}{1+\sum exp(x_i\beta_k)}$$

$$L(\beta_2,\ldots,\beta_j|y,X) = \prod_{i=1}^{N}\prod_{j=2}^{J}\left[\frac{exp(x_i\beta_j)}{1+\sum_{k=2}^{J} exp(x_i\beta_k)}\right]^{\mathbf{1}(y_i=j)}$$

$$logL(\beta_2,\ldots,\beta_j|y,X) = \sum_{i=1}^{N}\sum_{j=2}^{J}\left[\mathbf{1}(y_i=j)log\left(\frac{exp(x_i\beta_j)}{1+\sum_{k=2}^{J} exp(x_i\beta_k)}\right)\right]$$

3 conditional Poisson

Again, this is getting old. There are J categories and $\pi_1,\ldots,\pi_J$ probabilities. y_j is the number of outcomes for category j s.t. $\sum y_j = n$

The multinomial distribution is:

$$f(y|n) = \frac{n!}{\prod_{j=1}^{J} y_j!}\prod_{j=1}^{J}\pi_j^{y_j}$$

But it is not a direct member of the exponential family. So we are going to use a Poisson with a condition, a conditional poisson, in order to enjoy the properties of the exponential family. The Poisson looks like:

$$f(y) = \prod_{j=1}^{J}\frac{\lambda_j^{y_j}e^{-\lambda_j}}{y_j!}$$

But we are interested in $f(y|n)$. How is n distributed? We know from the multinomial that $\sum y_j = n$. if $y_j \backsim Po(\lambda_j)$ the $n \backsim Po(\sum \lambda_j)$.

$$
\begin{aligned}
f(y|n) &= \frac{\prod \frac{\lambda_j^{y_j} e^{-\lambda_j}}{y_j!}}{\frac{[\sum \lambda_j]^n e^{-\sum \lambda_j}}{n!}} \\
&= \prod \left[\frac{\lambda_j}{\sum \lambda}\right]^{y_j} \frac{n!}{\prod y_j!}
\end{aligned}
$$

There are two conditions here.

1. $\pi_j = \frac{\lambda_j}{\sum \lambda}$
2. $\sum \pi_j = 1$