An introduction to Quasi-MLE Poisson

Xiang Ao

January 22, 2009

1 Poisson Model

If the intereseted variable is a count data variable, it is natural to model it as a Poisson process. The number of occurance follows a Poisson process:

$$\Pr[Y = y] = \frac{e^{-\mu}\mu^y}{y!} \tag{1}$$

One of Poisson's properties is that it has equal variance as its mean:

$$\mathrm{E}[Y] = \mathrm{Var}[Y] = \mu \tag{2}$$

In a count data model such as Poisson regression model, we are modeling log of the expected counts:

$$\log(\mathrm{E}(Y)) = \mathbf{X_i'}\beta \tag{3}$$

The log-likelihood function is

$$\log L(\beta) = \sum_{i=1}^{N} -\exp(X_i'\beta) + y_i X_i'\beta - \ln y_i! \tag{4}$$

To maximize it, the first order condition is

$$\sum_{i=1}^{N}(y_i - \exp(X_i'\beta))X_i' = 0 \tag{5}$$

We can use Newton-Raphson or other optimization algorithms to find the solution for the first order condition.

Note that in equation 5, if X_i' includes a constant, then $(y_i - \exp(X_i'\beta))$ sums to zero. If $\mathrm{E}[y_i|X_i] = \exp(X_i'\beta)$, then the summation on the left-hand side has expectation of zero. Hence the only specification needed to apply equation 5 is the conditional expectation of Y given X. Even the data is not Poisson-distributed, the estimator by equation 5 is still consistent. Therefore, this estimator is called quasi_ML (QML) Poisson estimator.

Although the QML Poisson estimator is consistent under relatively weak condition, the regular variance estimator for the coefficients are not valid anymore.

If a stronger assumption is made that the data follows a Poisson distribution, then the error term has a variance which equals to the mean of Y. The estimator $\hat{\beta}_P$ (ML estimator) follows a normal distribution asymptotically with a variance matrix

$$\text{Var}_{ML}[\hat{\beta}_P] = (\sum_{i=1}^{N} \mu_i X_i X_i')^{-1} \tag{6}$$

On the other hand, the variance matrix for the QML Poisson estimator $\hat{\beta}_{QML}$ is

$$\text{Var}_{QML}[\hat{\beta}_P] = (\sum_{i=1}^{N} \mu_i X_i X_i')^{-1} (\sum_{i=1}^{N} \omega_i X_i X_i')^{-1} (\sum_{i=1}^{N} \mu_i X_i X_i')^{-1} \tag{7}$$

where $\omega_i = \text{Var}[y_i|X_i]$ is the conditional variance of y_i.

2 Implementation

Poisson estimation is implemented in almost every statistical package. However, some of them may not work if you have a continuous dependent variable.

Stata's implementation of Poisson model: poisson and xtpoisson do take continuous dependent variable. However, if you intend to use it as QMLE-Poisson, standard errors need to be adjusted. Those two procedures do not adjust for standard errors. A user-written program called xtpqml calls for xtpoisson and it calculates robust standard error which is suggested by Wooldridge (1999).