Lecture Notes
On Binary Choice Models:
Logit and Probit

Thomas B. Fomby
Department of Economic
SMU
March, 2010

Maximum Likelihood Estimation of Logit and Probit Models

$$y_i = \begin{cases} 1 & \text{with probability } P_i \\ 0 & \text{with probability } 1 - P_i \end{cases}$$

Consequently, if N observations are available, then the likelihood function is

$$L = \prod_{i=1}^{N} P_i^{y_i} \left(1 - P_i\right)^{1-y_i} . \qquad (1)$$

The logit or probit model arises when P_i is specified to be given by the logistic or normal cumulative distribution function evaluated at $X_i'\beta$. Let $F\left(X_i'\beta\right)$ denote either of theses cumulative distribution functions. Then, the likelihood function of both models is

$$L = \prod_{i=1}^{N} F\left(X_i'\beta\right)^{y_i} \left(1 - F\left(X_i'\beta\right)\right)^{1-y_i} . \qquad (2)$$

Then, the log-likelihood function is

$$\ln L = l = \sum_{i=1}^{N} \left[y_i \ln F\left(X_i'\beta\right) + \left(1 - y_i\right) \ln\left(1 - F\left(X_i'\beta\right)\right)\right]. \qquad (3)$$

Now, the first order conditions arising from equation (3) are nonlinear and non-analytic. Therefore, we have to obtain the ML estimates using numerical optimization methods, eg, the Newton-Raphson method.

This method (which will be explained further later) implies the following recursion.

$$\tilde{\beta}_{n+1} = \tilde{\beta}_n - \left[\frac{\partial^2 l}{\partial\beta\partial\beta'}\right]^{-1}_{\beta=\tilde{\beta}_n} \left[\frac{\partial l}{\partial\beta}\right]_{\beta=\tilde{\beta}_n} \tag{4}$$

In equation (4), $\tilde{\beta}_n$ is the n-th round estimate and the Hessian and score vectors are evaluated at this estimate.

From our previous ML theorem, we know that

$$\sqrt{N}\left(\tilde{\beta}_{ML} - \beta\right) \xrightarrow{asy} N\left(0, -N\left(E\left[\frac{\partial^2 l}{\partial\beta\partial\beta'}\right]^{-1}\right)\right) \tag{5}$$

where $\tilde{\beta}_{ML}$ represents the last iteration of the Newton-Raphson procedure. For finite samples, the asymptotic distribution of $\tilde{\beta}_{ML}$ can be approximated by

$$N\left(\beta, -\left[\frac{\partial^2 l}{\partial\beta\partial\beta'}\right]^{-1}_{\beta=\beta_{ML}}\right).$$

For the logit model, $P_i = F\left(X_i'\beta\right)$ where

$$F(t) = \frac{1}{1+e^{-t}} \tag{6}$$

is the logistic cdf and the logistic pdf is

$$F'(t) = f(t) = \frac{e^{-t}}{\left(1+e^{-t}\right)^2} \tag{7}$$

Also, note that

$$1 - F(t) = \frac{e^{-t}}{1+e^{-t}} = F(-t) \tag{8-1}$$

$$\frac{f(t)}{F(t)} = 1 - F(t) \tag{8-2}$$

$$f'(t) = -f(t)F(t)\left(1 - e^{-t}\right) \tag{8-3}$$

Using these results it can be shown for the logit model,

$$\frac{\partial l}{\partial \beta}=\sum_{i=1}^{N}y_i\frac{1}{1+\exp\left(X_i'\beta\right)}X_i-\sum_{i=1}^{N}\left(1-y_i\right)\frac{1}{1+\exp\left(-X_i'\beta\right)}X_i$$
$$=\sum_{i=1}^{N}\left[y_iF\left(-X_i'\beta\right)-\left(1-y_i\right)F\left(X_i'\beta\right)\right]X_i \qquad (9)$$

The Hessian can be shown to be

$$\frac{\partial^2 l}{\partial\beta\partial\beta'}=-\sum_{i=1}^{N}\frac{\exp\left(-X_i'\beta\right)}{\left[1+\exp\left(-X_i'\beta\right)\right]^2}X_iX_i'$$
$$=-\sum_{i=1}^{N}f\left(X_i'\beta\right)X_iX_i' \qquad (10)$$

Note that this X_iX_i' matrix is p.d. for all $\tilde{\beta}$.

So, iterate $\tilde{\beta}_{n+1}=\tilde{\beta}_n-\left[\frac{\partial^2 l}{\partial\beta\partial\beta'}\right]^{-1}_{\beta=\tilde{\beta}_n}\left[\frac{\partial l}{\partial\beta}\right]_{\beta=\tilde{\beta}_n}$ until $\left|\tilde{\beta}_{n+1}-\tilde{\beta}_n\right|<\varepsilon$.

For the probit model, $P_i=F\left(X_i'\beta\right)$ where

$$f\left(t\right)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}t^2\right) \qquad (11)$$

is the probit pdf and the probit cdf is

$$F\left(t\right)=\int_{-\infty}^{t}f\left(v\right)dv \qquad (12)$$

Also, note that

$$f'\left(t\right)=-tf\left(t\right) \qquad (13\text{-}1)$$
$$F\left(-t\right)=1-F\left(t\right) \qquad (13\text{-}2)$$

Then, the score vector for the probit model is

$$\frac{\partial l}{\partial\beta}=\sum_{i=1}^{N}\left[y_i\frac{f\left(X'\beta_i\right)}{F\left(X'\beta_i\right)}-\left(1-y_i\right)\frac{f\left(X'\beta_i\right)}{1-F\left(X'\beta_i\right)}\right]X_i \qquad (14)$$

The probit Hessian is then

$$\frac{\partial^2 l}{\partial\beta\partial\beta'}=-\sum_{i=1}^{N}f\left(X_i'\beta\right)\left[y_i\frac{f\left(X_i'\beta\right)+X_i'\beta F\left(X_i'\beta\right)}{F\left(X_i'\beta\right)^2}+\left(1-y_i\right)\frac{f\left(X_i'\beta\right)-X'\beta_i\left(1-F\left(X_i'\beta\right)\right)}{\left[1-F\left(X_i'\beta\right)\right]^2}\right]X_iX_i'$$

Estimation of Marginal Effects in the Logit and Probit Models

The analysis of marginal effects requires that we examine

$$\frac{\partial P_i}{\partial X_{ij}} = f\left(X_i'\beta\right)\beta_j, \quad i = 1,2,\cdots,N,\ j = 1,2,\cdots,K\ .$$

$$\left.\frac{\hat{\partial P_i}}{\partial X_{ij}}\right|_{X=\bar{X}} = f\left(\bar{X}_i'\tilde{\beta}_{ML}\right)\tilde{\beta}_{ML,j}, \quad i = 1,2,\cdots,N,\ j = 1,2,\cdots,K$$

Talk about applications of logit and probit : credit scoring, target marketing, bond Rating. Go over example of **German Credit.xls** on class website.