Lecture Notes
On Binary Choice Models:
Logit and Probit

Thomas B. Fomby
Department of Economics
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

\[ L = \prod_{i=1}^{N} P_i^{y_i} (1 - P_i)^{1-y_i}. \]  \hspace{1cm} (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 \( F(X_i'\beta) \) denote either of these cumulative distribution functions. Then, the likelihood function of both models is

\[ L = \prod_{i=1}^{N} F(X_i'\beta)^{y_i} (1 - F(X_i'\beta))^{1-y_i}. \]  \hspace{1cm} (2)

Then, the log-likelihood function is

\[ \ln L = l = \sum_{i=1}^{N} [y_i \ln F(X_i'\beta) + (1-y_i)(\ln(1 - F(X_i'\beta))]. \]  \hspace{1cm} (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.
\[ \tilde{\beta}_{n+1} = \tilde{\beta}_n - \left[ \frac{\partial^2 l}{\partial \beta \partial \beta'} \right]^{-1}_{\beta=\tilde{\beta}_n} \left[ \frac{\partial l}{\partial \beta} \right]_{\beta=\tilde{\beta}_n} \]  

(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

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

(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

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

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

\[ F(t) = \frac{1}{1 + e^{-t}} \]  

(6)

is the logistic cdf and the logistic pdf is

\[ f'(t) = f(t) = \frac{e^{-t}}{(1 + e^{-t})^2} \]  

(7)

Also, note that

\[ 1 - F(t) = \frac{e^{-t}}{1 + e^{-t}} = F(-t) \]  

(8-1)

\[ \frac{f(t)}{F(t)} = 1 - F(t) \]  

(8-2)

\[ f'(t) = -f(t)F(t)(1 - e^{-t}) \]  

(8-3)

Using these results it can be shown for the logit model,
The Hessian can be shown to be
\[
\frac{\partial^2 l}{\partial \beta \partial \beta'} = -\sum_{i=1}^{N} \frac{\exp(-X_i'\beta)}{[1 + \exp(-X_i'\beta)]^2} X_i X_i'
\]
\[
= -\sum_{i=1}^{N} f(X_i'\beta) X_i X_i'
\]
(10)

Note that this \(X_i X_i'\) matrix is p.d. for all \(\tilde{\beta}\).

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

For the probit model, \(P_t = F(X_t'\beta)\) where
\[
f(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} t^2\right)
\]
(11)
is the probit pdf and the probit cdf is
\[
F(t) = \int_{-\infty}^{t} f(v) dv
\]
(12)

Also, note that
\[
f'(t) = -tf(t) \quad (13-1)
\]
\[
F(-t) = 1 - F(t) \quad (13-2)
\]

Then, the score vector for the probit model is
\[
\frac{\partial l}{\partial \beta} = \sum_{i=1}^{N} \left[ y_i \frac{f(X_i'\beta)}{F(X_i'\beta)} - (1 - y_i) \frac{f(X_i'\beta)}{1 - F(X_i'\beta)} \right] X_i
\]
(14)

The probit Hessian is then
\[
\frac{\partial^2 l}{\partial \beta \partial \beta'} = -\sum_{i=1}^{N} f(X_i'\beta) \left[ y_i \frac{f(X_i'\beta) + X_i'\beta F(X_i'\beta)}{F(X_i'\beta)^2} + (1 - y_i) \frac{f(X_i'\beta) - X_i'\beta(1 - F(X_i'\beta))}{[1 - F(X_i'\beta)]^2} \right] X_i X_i'
\]
Estimation of Marginal Effects in the Logit and Probit Models

The analysis of marginal effects requires that we examine
\[
\frac{\partial P_i}{\partial X_{ij}} = f(X_i'\beta_j), \quad i = 1,2,\ldots,N, \; j = 1,2,\ldots,K.
\]
\[
\frac{\partial \hat{P}_i}{\partial X_{ij}} \bigg|_{X = \bar{X}} = f(X_i'\hat{\beta}_{ML})\hat{\beta}_{ML,j}, \quad i = 1,2,\ldots,N, \; j = 1,2,\ldots,K
\]

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