Limited dependent variable models

Hedibert Freitas Lopes

The University of Chicago Booth School of Business
5807 South Woodlawn Avenue, Chicago, IL 60637
http://faculty.chicagobooth.edu/hedibert.lopes
hlopes@ChicagoBooth.edu
Outline

1. Tobit model
   Likelihood functions
   Full conditional distributions
   Linear tobit model

2. Probit/ordered probit models
   Full conditional distributions

3. Multinomial probit models
   SUR
   Prior and full conditionals
Limited dependent variable models

Tobit model
Linear Tobit model
Probit/ordered probit models
Multinomial probit model
When the dependent variable is censored, values in a certain range are all transformed to a single value (Greene, pp.761)

- Household purchases of durable goods (Tobin, 1958)
- The number of extramarital affairs
- The number of hours worked by a woman in the labor force
- The number of arrests after release from prison
- Household expenditure on various commodity groups
- Vacation expenditures
Tobin model


Only $y_i$ is observed.

Censoring rule (truncation):

$$y_i = \begin{cases} y_i^* & y_i^* > 0 \\ 0 & y_i^* \leq 0 \end{cases}$$

Model structure:

$$y_i^* \sim F(y_i^*|x_i, \theta) \quad \text{with p.d.f. } f(y_i^*|x_i, \theta)$$
Likelihood functions

The likelihood function of $\theta$ is

$$L(\theta; y, x) = \left\{ \prod_{i: y_i > 0} f(y_i|x_i, \theta) \right\} \{ F(0|x_i, \theta) \}^k$$

where $k$ is the number of observations where $y_i = 0$.

The complete likelihood is

$$p(y_i|x_i, \theta) = \int p(y_i, y_i^*|x_i, \theta) dy_i^*$$

$$= \int p(y_i|x_i, \theta, y_i^*) f(y_i^*|x_i, \theta) dy_i^*$$

$$= p(y_i|x_i, \theta, y_i^* > 0)(1 - F(0|x_i, \theta))$$

$$+ p(y_i|x_i, \theta, y_i^* \leq 0) F(0|x_i, \theta)$$
Full conditional distributions

Given $y^*$, sampling $\theta$ is, in general, straightforward:

$$
\pi(\theta|y^*, x) \propto \pi(\theta) \prod_{i=1}^{n} f(y^*_i | x_i, \theta)
$$

The data augmentation argument leads to:

$$(y^*_i | x_i, y_i, \theta) \sim y_i 1(y_i > 0) + \left[ \frac{f(y^*_i | x_i, \theta)}{F(0 | x_i, \theta)} 1(y^*_i \leq 0) \right] 1(y_i = 0)$$

In other words, $y^*_i = y_i$ if $y_i > 0$ and

$$(y^*_i | x_i, y_i, \theta) \sim \frac{f(y^*_i | x_i, \theta)}{F(0 | x_i, \theta)} y^*_i \leq 0$$

if $y_i = 0$. 
Linear tobit model

In this case,
\[ y_i^* \sim N(x_i'\beta, \sigma^2) \]

Prior distribution: \( \pi(\beta, \sigma^2) = \pi(\beta)\pi(\sigma^2) \)

\[ \beta \sim N(\beta_0, V_0) \]
\[ \sigma^2 \sim IG(\nu_0/2, \nu_0s_0^2/2) \]

Full conditional distributions:
\[ \beta \sim N(V_{\beta}(V_0^{-1}\beta_0 + X'y^*), V_{\beta}) \]
\[ \sigma^2 \sim IG \left(\frac{\nu_0 + n}{2}, \frac{\nu_0s_0^2 + (y^* - X\beta)'(y^* - X\beta)}{2}\right) \]
\[ y_i^* \sim y_i1(y_i > 0) + N_{[-\infty,0]}(x_i'\beta, \sigma^2)1(y_i = 0) \]

where \( V_{\beta}^{-1} = V_0^{-1} + \sigma^{-2}(X'X)^{-1} \).
Example

Let $y_i^* | x_i \sim N(\beta x_i, \sigma^2)$ and $x_i \sim N(a, 1)$, such that

$$y_i^* \sim N(\beta a, \sigma^2 + \beta^2)$$

and

$$Pr(y = 0 | a, \beta = 2, \sigma^2 = 0.25) = \Phi(-2a/2.062)$$
$\pi(\beta | y, x)$ and $\pi(\sigma^2 | y, x)$

$\nu_0 = 0$ and $V_0^{-1} = 0$ (non-informative prior)

(without truncation) and (with truncation)
Probit/ordered probit models

In the ordered probit models, \( y_i \in \{1, 2, \ldots, J\} \) and

\[
y_i = \begin{cases} 
1 & \gamma_0 < y_i^* \leq \gamma_1 \\
\vdots & \vdots \\
J & \gamma_{J-1} < y_i^* \leq \gamma_J
\end{cases}
\]

where \( y_i^* \sim F(x_i, \theta) \) with p.d.f. \( f(y_i^*|x_i, \theta) \).

**Note:** In probit models, \( y_i \in \{1, 2\} \) and \( \gamma_1 = 0 \).
Full conditional distributions

Assuming that $\pi(\theta)$ and $\pi(\gamma)$ are uniform in $-\infty = \gamma_0 \leq \cdots \leq \gamma_J = \infty$, then

\[
\pi(\theta | y^*, X) \propto \pi(\theta) \prod_{i=1}^{n} f(y_i^* | x_i, \theta)
\]

\[
(y_i^* | x_i, y_i, \theta) \sim \sum_{j=1}^{J} \frac{f(y_i^* | x_i, \theta)1(y_i = j)}{F(\gamma_j | x_i, \theta) - F(\gamma_{j-1} | x_i, \theta)}
\]

\[
(\gamma_j | \gamma(-j), \theta, y^*) \sim U(\bar{\gamma}_{j-1}, \bar{\gamma}_{j+1})
\]

where

\[
\gamma(-j) = (\gamma_1, \cdots, \gamma_{j-1}, \gamma_{j+1}, \cdots, \gamma_J)
\]

\[
\bar{\gamma}_{j-1} = \max \{ \max \{y_i^* : y_i = j\}, \gamma_{j-1}\}
\]

\[
\bar{\gamma}_{j+1} = \min \{ \min \{y_i^* : y_i = j+1\}, \gamma_{j+1}\}
\]
Multinomial probit models

In multinomial probit models\(^1\), \(U_{ij}\) is the utility of individual \(i\) when she chooses alternative \(j\), and

\[
y_{ij}^* = U_{ij} - U_{i0} = x_{ij}'\beta + \varepsilon_{ij}
\]

while the econometrician only observes \(y_i\) such that

\[
y_i = \begin{cases} 
  0 & \text{max}(y_i^*) < 0 \\ 
  1 & y_{i1}^* = \text{max}(y_i^*) \geq 0 \\ 
  \vdots & \vdots \\ 
  J - 1 & y_{i,J-1}^* = \text{max}(y_i^*) \geq 0 \\ 
  J & y_{i,J}^* = \text{max}(y_i^*) \geq 0 
\end{cases}
\]

where \(y_i \in \{0, 1, \ldots, J\}\) and \(y_i^* = (y_{i1}^*, \ldots, y_{iJ}^*)\).

\(^1\)See all reference in the end.
Stacking up the equations

\[
\begin{pmatrix}
y_{i1}^* \\
y_{i2}^* \\
\vdots \\
y_{iJ}^*
\end{pmatrix}
= \begin{pmatrix}
x_{i1}' & 0 & \cdots & 0 \\
0 & x_{i2}' & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & x_{iJ}'
\end{pmatrix}
\begin{pmatrix}
\beta_1 \\
\beta_2 \\
\vdots \\
\beta_J
\end{pmatrix}
+ \begin{pmatrix}
\epsilon_{i1} \\
\epsilon_{i2} \\
\vdots \\
\epsilon_{iJ}
\end{pmatrix}
\]

or,

\[
y_{i}^* = X_i \beta + \epsilon_i
\]

with \( \epsilon_i \sim N(0, \Omega) \).

Also,

\[
y^* = X \beta + \epsilon
\]

where \( y^* = (y_1^*, \ldots, y_N^*) \), \( X = (X'_1, \ldots, X'_N) \) and \( \epsilon = (\epsilon_1, \ldots, \epsilon_N) \).
Prior and full conditionals

Prior distribution

\[ \pi(\beta, \Omega) = \pi(\beta) \pi(\Omega) \]

\[ \beta \sim \text{Normal} \]

\[ \Omega \sim \text{Inverse Wishart} \]

Full conditional distributions

\[ \beta \sim \text{Normal} \]

\[ \Omega \sim \text{Inverse Wishart} \]

\[ y_i^* \sim \begin{cases} N(X_i\beta, \Omega)1(\max(y_i^*) < 0) & y_i = 0 \\ N(X_i\beta, \Omega)1(\max(y_i^*) = y_{ij}^* \geq 0) & y_i = j \end{cases} \]
References


