Limited Dependent Variable Models

Semester 2/2013

Part 2

Chayanee Chawanote

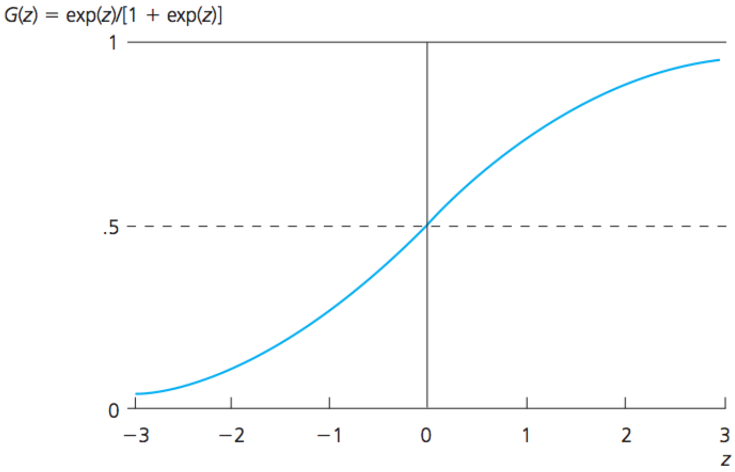
Binary response models

- The response probability: $P(y=1|\mathbf{x}) = P(y=1|x_1, x_2, ..., x_k)$
- In the LPM, we assume that the response probability is linear in a set of parameters.
- For general setting, define the response probability as a function taking on values between 0 and 1: $0 < G(z) < 1$
  $P(y=1|\mathbf{x}) = G(\beta_0 + \beta_1 x_1 + ... + \beta_k x_k) = G(\beta_0 + \mathbf{x}\beta)$ (1)
- The two most used function G that ensure the probabilites between 0 and 1:
  - the logistic function $\rightarrow$ logit model
  - the standard normal cumulative distribution function $\rightarrow$ probit model

- Logit model: G is the logistic function
  $G(z) = \frac{exp(z)}{1+exp(z)} = \Lambda(z)$ (2)
- Probit model: G is the standard normal cumulative distribution function (cdf)
  $G(z) = \Phi(z) \equiv \int_{-\infty}^{z} \phi(v)dv$, (3)
  where $\phi(z)$ is the standard normal density:
  $\phi(z) = \sqrt{2\pi} exp(\frac{-z^2}{2})$ (4)
- G(z) in both (2) and (3) are increasing functions with $G(z) \to 0$ as $z \to -\infty$, and $G(z) \to 1$ as $z \to \infty$.

Logistic function

Graph of the logistic function $G(z) = \exp(z)/[1 + \exp(z)]$.



Latent variable model

- Logit and probit models can be derived from an underlying latent variable model.
- Let y* be an unobserved, or latent, variable determined by $y* = \beta_0 + \mathbf{x}\beta + e$, $y = 1[y* > 0]$ (5)
- 1[.] is an indicator function, which is $= 1$ if the event in the brackets is true, and $= 0$ otherwise.
- Assume e is independent of $\mathbf{x}$ and e either has the standard logistic distribution or the standard normal distribution.
  - e is symmetrically distributed about zero
  - $1 - G(-z) = G(z)$ for all real numbers z
- Probit is more popular since economists tend to favor the normality assumption for e

Latent variable model

- From (5), we can derive the response probability for y:
  $P(y = 1|\mathbf{x}) = P(y* > 0|\mathbf{x}) = P[e > -(\beta_0 + \mathbf{x}\beta)|\mathbf{x}]$
  $= 1 - G[-(\beta_0 + \mathbf{x}\beta)] = G(\beta_0 + \mathbf{x}\beta)$ (6)
- The direction of the effect of $x_j$on $E(y * |x) = \beta_0 + \mathbf{x}\beta$ and on $E(y|x) = P(y = 1|x) = G(\beta_0 + \mathbf{x}\beta)$ is always the same.
- For the magnitude, we care about the effect of x_j on the probability of success $P(y = 1|x)$.

Maximum likelihood estimation of logit and probit models

- Suppose we have a random sample of size n.
- To obtain MLE, conditional on the explanatory variables, we need the density of y given $\mathbf{x}_j$:
  $f(y|\mathbf{x_j};\beta) = [G(\mathbf{x}_i\beta)]^y[1-G(\mathbf{x}_i\beta)]^{1-y}, y = 0, 1$ (7)
- Then, the log-likelihood function for observation i is:
  $\ell_i(\beta) = y_i log[G(\mathbf{x}_i\beta)] + (1-y_i)log[1-G(\mathbf{x}_i\beta)]$ (8)
- The log-likelihood for a sample size of n: $\ell(\beta) = \sum_{i=1}^{n} \ell_i(\beta)$ (9)
- Maximum likelihood estimator (MLE): $\hat{\beta}$ maximizes the log-likelihood in (9)
- Under general conditions, the MLE is consistent, asymptotically normal, and asymptotically efficient.

Goodness of fit

- For binary response model, we calculate pseudo R-squared $= 1 - \ell_{ur}/\ell_0$
- ℓ_{ur} = log-likelihood function for the estimated model, ℓ_0 = log-likelihood function with only intercept model.
- The log-likelihood function is always a negative number, and $|\ell_{ur}| \leq |\ell_0|$.
- If the covariates have no explanatory power, then $\ell_{ur}/\ell_0 = 1$ and pseudo $R^2 = 0$
- If $\ell_{ur} = 0$, the pseudo $R^2 = 1$

Interpretation of probits and logits

- We care about the effect of x on P(y=1|x), that is $\partial p / \partial x$.
- Probit and logit are non-linear, partial effect on p(x) is from $\frac{\partial p(x)}{\partial x_j} = g(\beta_0 + \mathbf{x}\beta)\beta_j$, where $g(z) \equiv \frac{dG}{dz}(z)$
  where G is the cdf, g is a probability density function
- $G(\cdot)$ is a strictly increasing cdf, so $g(z) > 0$ for all z
  - The partial effect always has the same sign as β_j
- So, we can only directly compare sign and significance (based on a standard t test) of coefficients

Hypotheses testing

- For each coefficient, we can construct (asymptotic) t tests as usual: t $= \hat{\beta}_j / se(\hat{\beta}_j)$
- For multiple hypotheses, or tests of multiple exclusion restrictions, we use the likelihood ratio (LR) test
- LR test is based on the difference in the log-likelihood functions for the unrestricted and restricted models, whether the log-likelihood becomes significantly smaller (no larger) when dropping variables.
- LR statistic: $LR = 2(\ell_{ur} - \ell_r)$
- Under H_0, if we are testing q exclustion restrictions, $LR \sim \chi^2_q$