Toulouse School of Economics

M1 INTERMEDIATE ECONOMETRICS

Classical regression and Tobit

Koen Jochmans   François Poinas

2024 — 2025

Overview

This deck of slides goes over the classical linear regression model and the tobit model from a maximum-likelihood perspective.

The relevant sections in Hansen are H5.5 and H27.1-H27.6.

Normal model

The classical linear regression model is

$$Y|X \sim N(X'\beta, \sigma^2).$$

The density of Y conditional on X is

$$\frac{1}{\sigma}\phi\left(\frac{Y - X'\beta}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2}\frac{(Y - X'\beta)^2}{\sigma^2}\right).$$

The likelihood function for this problem is

$$L_n(\beta, \sigma^2) = \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2}\frac{(Y - X'\beta)^2}{\sigma^2}\right).$$

The log-likelihood is

$$\ell_n(\beta, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2}\sum_{i=1}^{n}\frac{(Y_i - X_i'\beta)^2}{\sigma^2}.$$

Here, the first term is an inessential constant.

The first-order conditions are

$$\frac{\partial \ell_n(\beta, \sigma^2)}{\partial \beta} = \sum_{i=1}^{n}\frac{X_i(Y_i - X_i'\beta)}{\sigma} = 0,$$

and

$$\frac{\partial \ell_n(\beta, \sigma^2)}{\partial \sigma^2} = \frac{1}{2}\left(\frac{\sum_{i=1}^{n}(Y_i - X_i'\beta)^2}{\sigma^4} - \frac{n}{\sigma^2}\right) = 0.$$

We can solve these equations sequentially, giving

$$\hat{\beta}_{\text{mle}} = \left(\sum_{i=1}^{n} X_i X_i'\right)^{-1} \left(\sum_{i=1}^{n} X_i Y_i\right),$$

which corresponds to the OLS estimator, and

$$\hat{\sigma}^2_{\text{mle}} = \frac{1}{n} \sum_{i=1}^{n} (Y_i - X_i' \hat{\beta}_{\text{mle}})^2,$$

which is again standard (albeit not degree-of-freedom corrected) from regression analysis.

Censoring

A variable Y^* is left-censored at zero when we observe

$$Y = \max(Y^*, 0).$$

Right censoring is analogous and leads to top-coded variables instead.

Take

$$Y^* = X'\beta + e, \qquad e|X \sim N(0, \sigma^2).$$

Then censoring below zero leads to $Y|X$ having a mass point at zero.

Not obvious how to construct an estimator for β and/or σ^2 here.

Ignoring censoring

A naive approach to ‘dealing’ with censoring is to retain only data that is not censored.

However,

$$\mathbb{E}(Y^*|X, Y^* > 0) = X'\beta + \mathbb{E}(e|X, Y^* > 0) = X'\beta + \mathbb{E}(e|X, e > -X'\beta)$$

and by normality,

$$\mathbb{E}(e|X, e > -X'\beta) = +\sigma \frac{\phi(-X'\beta/\sigma)}{1 - \Phi(-X'\beta/\sigma)} = \sigma \frac{\phi(X'\beta/\sigma)}{\Phi(X'\beta/\sigma)},$$

so that

$$\mathbb{E}(Y^*|X, Y^* > 0) = X'\beta + \sigma \frac{\phi(X'\beta/\sigma)}{\Phi(X'\beta/\sigma)} \neq X'\beta$$

Here we have used that, for any (a, u) with $a < u$,

$$\mathbb{P}(e \leq u | e > a) = \frac{\mathbb{P}(e \leq u) - \mathbb{P}(e \leq a)}{1 - \mathbb{P}(e \leq a)} = \frac{\Phi(u/\sigma) - \Phi(a/\sigma)}{1 - \Phi)(a/\sigma)},$$

so that the conditional density is

$$\frac{(1/\sigma)\, \phi(u/\sigma)}{1 - \Phi(a/\sigma)},$$

to obtain

$$\mathbb{E}(e|e > a) = \frac{\int_a^\infty (u/\sigma)\phi(u/\sigma)\, du}{1 - \Phi(a/\sigma)} = \sigma \frac{\phi(a/\sigma)}{1 - \Phi(a/\sigma)}.$$

The last step uses the change of variable $z = u/\sigma$ and the fact that $\phi'(z) = -z\phi(z)$.

The likelihood function

We have that

$$\mathbb{P}(Y=0|X)=\mathbb{P}(Y^*<0|X)=\Phi(-X'\beta/\sigma)$$

while, for any $y>0$

$$\mathbb{P}(Y\leq y|X)=\mathbb{P}(Y=0|X)+\mathbb{P}(Y\leq y|X,Y>0)\,\mathbb{P}(Y>0|X)$$

with

$$\mathbb{P}(Y\leq y|X,Y>0)\,\mathbb{P}(Y>0|X)=\mathbb{P}(Y^*\leq y|X=x)-\mathbb{P}(Y^*\leq 0|X)$$

so that for $y>0$

$$\mathbb{P}(Y\leq y|X)=\mathbb{P}(Y^*\leq y|X)=\Phi\left(\frac{y-X'\beta}{\sigma}\right)$$

with density

$$\frac{1}{\sigma}\phi\left(\frac{y-X'\beta}{\sigma}\right)$$

The likelihood function thus is

$$L_n(\beta,\sigma^2) = \prod_{i=1}^{n} \Phi\left(\frac{-X_i'\beta}{\sigma}\right)^{\{Y_i=0\}} \left(\frac{1}{\sigma}\phi\left(\frac{Y_i - X_i'\beta}{\sigma}\right)\right)^{\{Y_i>0\}}$$

which has a probit component and a normal-regression component.

A useful reparametrization has $\gamma = \beta/\sigma$ and $\delta = 1/\sigma$, which aids with numerical optimisation.

In general, the likelihood function is invariant to one-to-one reparametrizations.

Marginal effects

In the tobit model β captures average partial effects of a change in X on Y^*.

The marginal effect on Y is nonlinear. We have

$$\frac{\partial \mathbb{E}(Y|X)}{\partial X} = \frac{\partial \mathbb{E}(Y|X, Y > 0)\, \mathbb{P}(Y > 0|X)}{\partial X}.$$

By an application of the chain rule, using the calculations from above, we obtain

$$\beta\, \Phi\left(\frac{X'\beta}{\sigma}\right).$$

The average marginal effect, for example, then is

$$\mathbb{E}\left(\beta\, \Phi\left(\frac{X'\beta}{\sigma}\right)\right).$$