Lecture 4: Selection.
Sam Kortum - Fall 2002
Greene (ch. 20), Heckman
This Draft: November 11, 2002.

Much of the traditional literature on Truncation, Censoring, and Selection relies heavily on properties of the Normal distribution.

1 Normal Distribution

If \( X \) has a Standard Normal distribution its density is
\[
\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad -\infty < x < \infty.
\]
Note that \( \phi'(x) = -\phi(x)x \) and \( \phi(-x) = \phi(x) \). The associated (cumulative) distribution function is
\[
\Pr[X \leq x] = \int_{-\infty}^{x} \phi(t) dt = \Phi(x).
\]
Note that \( \Phi'(x) = \phi(x) \) and \( \Phi(-x) = 1 - \Phi(x) \). Letting \( Y = \mu + \sigma X \), we get
\[
\Pr[Y \leq y] = \Pr[\mu + \sigma X \leq y] = \Pr[X \leq \frac{y - \mu}{\sigma}] = \int_{-\infty}^{\frac{y - \mu}{\sigma}} \phi(t) dt = \Phi(\frac{y - \mu}{\sigma}).
\]
Applying Leibnitz’ rule to the second to the integral above, the density of \( Y \) is
\[
f(y) = \frac{1}{\sigma} \phi\left(\frac{y - \mu}{\sigma}\right) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{y - \mu}{\sigma}\right)^2}, \quad -\infty < y < \infty
\]

2 Truncated Normal Distribution

Now suppose we condition on \( Y \in A = [a_1, a_2] \), where \( -\infty < a_1 < a_2 < \infty \). The probability of \( Y \) falling into this interval is \( \Phi\left(\frac{a_2 - \mu}{\sigma}\right) - \Phi\left(\frac{a_1 - \mu}{\sigma}\right) \). Thus the conditional density of \( Y \) is
\[
f(y | A) = \frac{1}{\sigma} \phi\left(\frac{y - \mu}{\sigma}\right) \frac{1}{\Phi\left(\frac{a_2 - \mu}{\sigma}\right) - \Phi\left(\frac{a_1 - \mu}{\sigma}\right)}, \quad a_1 \leq y \leq a_2
\]
We want to derive the mean and variance of this distribution.
2.1 Moment Generating Function

The MGF is

\[ M(t) = E[e^{tY} | Y \in A] = \int_{a_1}^{a_2} e^{ty} f(y) dy = e^{\mu t + \sigma^2/2} \frac{\Phi(\frac{a_2 - \mu}{\sigma}) - \Phi(\frac{a_1 - \mu}{\sigma})}{\Phi(\frac{a_2 - \mu}{\sigma}) - \Phi(\frac{a_1 - \mu}{\sigma})} \]

The last equality follows from

\[
\frac{1}{\sigma \sqrt{2\pi}} \int_{a_1}^{a_2} e^{ty} \frac{1}{\sigma} \left( \frac{y - \mu}{\sigma} \right)^2 dy = \frac{1}{\sigma \sqrt{2\pi}} \int_{a_1}^{a_2} e^{\frac{y}{2\sigma} - \frac{1}{2}(\sigma^2 t + \mu)^2 - \frac{1}{2}(\sigma^2 t + \mu)^2} dy = e^{\mu t + \sigma^2/2} \int_{a_1}^{a_2} e^{\frac{y}{2\sigma} - \frac{1}{2}(\sigma^2 t + \mu)^2} dy = e^{\mu t + \sigma^2/2} \left[ \Phi(\frac{a_2 - \mu}{\sigma}) - \Phi(\frac{a_1 - \mu}{\sigma}) \right].
\]

where \( \mu' = \sigma^2 t + \mu \).

2.2 Expected Value

Putting the MGF to work:

\[ E[Y | Y \in A] = M'(t)|_{t=0} = \mu - \sigma \frac{\phi(\alpha_2) - \phi(\alpha_1)}{\Phi(\alpha_2) - \Phi(\alpha_1)}. \]

where \( \alpha_k = \frac{a_k - \mu}{\sigma} \). Letting \( a_2 \) tend to infinity,

\[ E[Y | Y > a_1] = \mu + \sigma \frac{\phi(\alpha_1)}{1 - \Phi(\alpha_1)} = \mu + \sigma \lambda(\alpha_1), \]

where \( \lambda(\alpha) > 0 \) is the hazard function. The hazard function of the Normal distribution is often called the inverse Mills ratio in the micro-econometrics literature.

Letting \( a_1 \) tend to minus infinity,

\[ E[Y | Y < a_2] = \mu - \sigma \frac{\phi(\alpha_2)}{\Phi(\alpha_2)} = \mu - \sigma \lambda(-\alpha_2). \]

Letting \( a_2 \) tend to infinity as well, of course, we get \( E[Y] = \mu \). Relative to this non-truncated case, truncation from below raises the mean \( E[Y | Y > a_1] > E[Y] \). Truncation from above lowers the mean \( E[Y | Y < a_2] < E[Y] \).
2.3 Variance

Putting the MGF to work again:

\[ E[Y^2|Y \in A] = M''(t)|_{t=0} = \sigma^2 + \mu^2 + \sigma^2 \phi'(\alpha_2) - \phi'(\alpha_1) - 2\mu \sigma \phi(\alpha_2) - \phi(\alpha_1). \]

Therefore,

\[ Var[Y|Y \in A] = E[Y^2|Y \in A] - E[Y|Y \in A]^2 = \sigma^2 \left\{ 1 - \frac{\alpha_2 \phi(\alpha_2) - \alpha_1 \phi(\alpha_1)}{\Phi(\alpha_2) - \Phi(\alpha_1)} - \left[ \frac{\phi(\alpha_2) - \phi(\alpha_1)}{\Phi(\alpha_2) - \Phi(\alpha_1)} \right]^2 \right\}. \]

Letting \( a_2 \) tend to infinity,

\[ Var[Y|Y > a_1] = \sigma^2 \left\{ 1 + \frac{\alpha_1 \phi(\alpha_1)}{1 - \Phi(\alpha_1)} - \left[ \frac{\phi(\alpha_1)}{1 - \Phi(\alpha_1)} \right]^2 \right\} = \sigma^2 \left[ 1 + \alpha_1 \lambda(\alpha) - \lambda(\alpha_1)^2 \right] = \sigma^2 \left[ 1 - \delta(\alpha_1) \right], \]

where \( \delta(\alpha) = \lambda(\alpha)[\lambda(\alpha) - \alpha] \). Notice that \( \delta(\alpha) = \lambda'(\alpha) \). It can be shown that \( 0 < \delta(\alpha) < 1 \). [Derivative is \( \delta' = \lambda'[\lambda - \alpha] + [\lambda' - 1] \lambda = \lambda[(\lambda - \alpha)^2 + \lambda(\lambda - \alpha) - 1] \). Thus \( \delta'(\alpha^*) = 0 \) implies \( 1 > 1 - (\lambda - \alpha^*)^2 = \lambda(\lambda - \alpha^*) = \delta(\alpha^*) \). Since \( \lim_{\alpha \to -\infty} \lambda(\alpha)\alpha = 0 \) we have \( \lim_{\alpha \to -\infty} \delta(\alpha) = 0 \). To be completed.]

Letting \( a_1 \) tend to minus infinity,

\[ Var[Y|Y < a_2] = \sigma^2 \left\{ 1 - \frac{\alpha_2 \phi(\alpha_2)}{\Phi(\alpha_2) - \Phi(\alpha_1)} - \left[ \frac{\phi(\alpha_2)}{\Phi(\alpha_2) - \Phi(\alpha_1)} \right]^2 \right\} = \sigma^2 \left[ 1 - \alpha_2 \lambda(-\alpha_2) - \lambda(-\alpha_2)^2 \right] = \sigma^2 \left[ 1 - \delta(-\alpha_2) \right]. \]

Letting \( a_2 \) tend to infinity as well, of course, we get \( Var[Y] = \sigma^2 \). Relative to the non-truncated case, note how the variance shrinks toward zero with truncation either from above or from below.

3 Bivariate Normal

Suppose \( U_1 \) and \( U_2 \) are independent random variables, each drawn from the Standard Normal density \( \phi(u) \).