Econ 673
Problem Set #1
Due September 12, 2008

1) Consider the linear model:

$$y_i^* = \beta_0 + \beta_1 X_{1i} + \varepsilon_i; \quad \varepsilon_i \sim N\left(0, \sigma^2\right) \tag{1}$$

Unfortunately, y_i^* is not observed. Instead, you only observe a censored version, y_i, where

$$y_i = \begin{cases} 0 & y_i^* < 0 \\ y_i^* & y_i^* \geq 0 \end{cases} \tag{2}$$

You are interested in evaluating (via Monte Carlo analysis) the properties of the standard OLS estimator of β when using the censored data (i.e., y_i) rather than the uncensored values (i.e., y_i^*

a) Assume the following parameter values for your model:

- $\beta_0 = 0.5$
- $\beta_1 = 1$
- $\sigma = 2$

b) Generate N=200 draws for y_i^* assuming that $X_i \sim U[0,1]$ and using equation (1)

c) Generate corresponding values for y_i using equation (2).

d) Using Matlab (or GAUSS) and no canned routines!, obtain the OLS estimates of the parameters of the mis-specified model:

$$y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{3}$$

Your results should include:

i) OLS estimates of β_0 and β_1, as well as an estimate of σ.

ii) An estimate of the asymptotic variance-covariance matrix for the OLS estimates

The following Matlab functions and commands will be useful:

`inv, diag,`

e) A collegue would like you to construct a 90% confidence interval for the variable $g(\beta)=(\beta_0+5\beta_1)^2$. Use simulation draws from the asymptotic distribution of the estimated parameters to provide them with the desired confidence interval.

f) Repeat steps (b) through (d) above a total of R=1000 times, collecting information about the OLS of β_0 and β_1. Specifically, provide

   i) Summary statistics for the OLS estimates of β_0 and β_1, including mean, standard deviations, minimums and maximums (using Matlab functions **`mean, std, min`** and **`max`**).

   ii) A histogram of the estimates of β_0 and β_1 (using the Matlab functions: **`hist, figure`**, and **`subplot`**).

   You may find it convenient to use a loop in conducting the above Monte Carlo exercise, storing the OLS estimates of β_0 and β_1 after each loop r=1,2…,R. For example:

```
R       = 1000;
theta   = zeros(nobs,2);
for     r = 1:R;
        …your code to obtain OLS estimates of β0 and β1 and iteration r,
          labeled thetahat
        theta(r,:) = thetahat;
end;
```

   What do your results tell you about the direction of the bias in estimating β_0 and β_1 using the censored data?

g) Repeat step (f) using $\beta_0=2$. How have your results changed (and why)?

2) Using the techniques discussed in class, generate R=500 draws from the following

   a) $V_1 \sim TN_{(5,60)}(0,25)$, where $TN_{(a,b)}(\mu,\sigma^2)$ denotes a truncated normal distribution, truncated below at a and above at b.

   b) $$V_2=\begin{cases}1 & \text{with probability } 0.20\\ 2 & \text{with probability } 0.40\\ 3 & \text{with probability } 0.06\\ 4 & \text{with probability } 0.30\\ 5 & \text{with probability } 0.04\end{cases}$$

c) V_3 drawn from an extreme value distribution truncated from below at 0.5 and from above at 1.5.

d) $V_4 \sim T(0.5, 1.5)$, where $T(a, b)$ denotes a triangular distribution with mean a, with pdf:

$$f(x) = \begin{cases} 0 & x \le a - b \\ \dfrac{x + b - a}{b^2} & a - b < x \le a \\ \dfrac{a + b - x}{b^2} & a < x \le a + b \\ 0 & x > a + b \end{cases}$$

For V_4 you must use the acceptance rejection method to obtain the draws of interest.

You can make use of the uniform random number generator **`rand`**. However, do not use any other random number generators. You may also make use of the Matlab function that computes the inverse of the normal cdf, **`norminv`**.

For each variable, provide a histogram of the sequence of draws using the Matlab functions **`hist, figure`**, and **`subplot.`**