Econometrics 2 (Fall 2021)

Homework 1: Probit

Due Wednesday on Sept. 01, 2021.

This code estimates a probit model using Maximum Likelihood. The latent variable model is:

$$y_i = \beta_0 + \beta_1 x_i + u_i \ ,$$

$$z_i = \begin{cases} 1 & \text{if } y_i > 0 \\ 0 & otherwise \end{cases}$$

with $u_i \sim N(0, \sigma)$ **and** $\sigma = 1$.

1. Set the parameters.

There are 100 simulations with 300 observations per simulation. Set $\beta_0 = 0.5, \beta_1 = 3$ and $\sigma = 1$.

```
close all
clear
clc
global x z N

N = 300;                                                    % Number of observations.
beta0 = 0.5;                                                % Intercept.
beta1 = 3;                                                  % Coefficient on X.
sigma = 1;                                                  % Standard deviation.

sim = 100;                                                  % Number of simulations.
results_mat = zeros(sim, 2);                                % Results matrix.
```

2. Maximum Likelihood Estimation.

In each simulation, generate draw the error terms, *U*, from the standard normal distribution and generate the data, X, Y and *Z*. Estimate the model using Maximum Likelihood and record the estimates.

```
x = ((1:N)'./N).*normrnd(0,1,N,1);                          % Generate X.

for s = 1:sim

    % Generate the data.

    u = normrnd(0,sigma,N,1);                               % Generate U.
    y = beta0*ones(size(x,1),1) + beta1*x + u;              % Generate Y.
    z = double((y > 0));                                    % Generate Z.
```

```
    % Estimation using ML.

    b0 = [1 1];                                                    % Initial values.

    options = optimset('Display','off');                           % Turn off display.
    [b_mle, ~, ~, ~, ~, hess] = fminunc('logl_prob', b0, options);  % Minimization.

    % Store estimates.

    results_mat(s, 1:size(results_mat,2)) = b_mle';                % Store results.
    vmat = inv(hess);

end
```

3. Display the results of the last simulation.

The estimate and standard error (in parenthesis) of β_0 in the last simulation is:

```
fprintf(' %0.4f\n(%0.4f)\n',b_mle(1),sqrt(vmat(1,1)))
```

The estimate and standard error (in parenthesis) of β_1 in the last simulation is:

```
fprintf(' %0.4f\n(%0.4f)\n',b_mle(2),sqrt(vmat(2,2)))
```

4. Empirical results.

The average and standard deviation (in parenthesis) of β_0 is:

```
fprintf(' %0.4f\n(%0.4f)\n', mean(results_mat(:,1)), std(results_mat(:,1)))
```

The average and standard deviation (in parenthesis) of β_1 is:

```
fprintf(' %0.4f\n(%0.4f)\n', mean(results_mat(:,2)), std(results_mat(:,2)))
```

5. Plot the estimated coefficients from all simulations.

Plot a histogram of β_0.

```
figure(1)
hold on
histogram(results_mat(:,1),5)
xlabel('$\beta_{0}$','interpreter','LaTex'); ylabel('Frequency')
title('Plot of $\beta_{0}$','interpreter','LaTex')
hold off
```

Plot a histogram of β_1.

```
figure(2)
hold on
histogram(results_mat(:,2),5)
xlabel('$\beta_{1}$','interpreter','LaTex'); ylabel('Frequency')
title('Plot of $\beta_{1}$','interpreter','LaTex')
hold off
```