Practice III

Federica Romei

November 12, 2013

1. Write an "main" m file where you define M, numbers of Montecarlo simulation, N numbers of random draws, β that is a vector 1*2 , γ a vector 2*1 and A a 2*2 symmetric matrix where in the main diagonal you have all 1 and in the other position a number greater than zero and lower than 1 . Assume you want to estimate the following model

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

where

- $\{y_i, x_i\}$ are i.i.d;
- $E[x_i, u_i] = f(x_i)$;
- $E[x_i' x_i]$ is full rank;
- $Var(x_i u_i) = f(x_i) < \infty$.

and you know that exists z_i s.t.

- $\{z_i, x_i\}$ are i.i.d;
- $E[z_i, u_i] = 0$;
- $E[x_i' z_i]$ is full rank;
- $E(x_i \epsilon_i) = 0$.
- $Var(x_i \epsilon_i) = f(z_i) < \infty$.

and that you can rewrite:

$$x_i = z_i \gamma + \epsilon_i.$$

Do a function that estimate the β^{IV} and the t^2 of the first stage regression. Let's do step by step

(a) Generate 2 indipendent random vectors from a normal distribution of dimension N *1, u and ϵ.

(b) Use the command chol (cholesky decomposition http://en.wikipedia.org/wiki/Cholesky_decomposition) to generate a covariance among the errors as

$$ERR = [u\epsilon] * chol(A).$$

Now the first column of ERR will be the error of the second stage estimation and the second column the error of the first stage. (Order is not important).

(c) Generate your z as a vector of ones and a vector of random variables and with the γ and error generate the x;

(d) Use the x and errors to generate the y;

(e) From now on you forget about the errors ;

(f) Estimate $\hat{\gamma}$ and t^2 to the square on the coefficient of z random variable under heteroschedasticity;

(g) Store the value of t^2 for every Montecarlo simulation;

(h) Estimate β^{IV} and store for every simulation.

(i) In the main file plot the distribution of t^2 and β^{IV}.

Solution:

```
function [beta_est t_first]=beta_estimation(N,M,A,gamma,beta)
% Initialize the matrix
beta_est=NaN(2,M);
t_first=NaN(M,1);

for m=1:M;
    %% GENERATE THE ERRORS
    % It is important that the errors of x and y are correlated
    u=normrnd(0,1,N,1);% We set the error as normal mean 0 and variance 1
    eps=normrnd(0,1,N,1);% We set the error as normal mean 0 and variance 1
    %% Generate the Instruments
    z=normrnd(3,1,N,1); % Instrument

    %% Correlation among the errors
    errors=[u eps]*chol(A); % We correlate the errors

    %% Vector of Instruments
    z_in=[ones(N,1) z]; % We build the matrix of instruments

    %% Gen the x
    x_g=z_in*gamma'+errors(:,2);% We build x



    x=[ones(N,1) x_g]; % We build the matrix of variables

    %% Gen the y
    y=x*beta'+errors(:,1); % We build y

    %% From here we are HUMAN!
    %% Estimation of gamma_hat and his Variance
    gamma_hat = (z_in'*z_in)\z_in'*x_g;

    er_hat=  x_g - z_in*gamma_hat;
```

```
    ZE=z_in.*(er_hat *ones(1,size(z_in,2)));

    Sigma_het (:,:)=(z_in'*z_in)\(ZE'*ZE)/(z_in'*z_in);

    %% Do the F test
    t_first(m,1)= (gamma_hat(2)/sqrt(Sigma_het(2,2)))^2;

    %% Estimate Beta_iv
    beta_iv=(z_in'*x)\z_in'* y; % We find the betas




    beta_est(:,m)= beta_iv; % we store



end
```

1 Problem Set 3

2. Assume that
$$y = \beta_0 + \beta_1 x + \beta_2 w + u$$
where $cov(x, z) \neq 0$. Assume that a collegue of yours, Paul suggests to estimate β_1 using this model:
$$y = \gamma_0 + \gamma_1 x + \epsilon$$
and that he claims that $\gamma_1 = \beta_1$.
   (a) Do you agree with Paul? Show to which value γ_1 converge to both analitically and with a Montecarlo simulation on Matlab.
   (b) What about the inference? Show on Matlab with one example if the inference works under the model of Paul;
   (c) Under which conditions you can agree with Paul?

3.

4. Assume that you want to estimate
$$y_i = \beta_0 + \beta x_i^* + u_i$$
but you can observe only:
$$x_i = x_i^* + \epsilon_i$$
   (a) Show that
$$\hat{\beta} \xrightarrow{D} \beta \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\epsilon^2}$$
   where $\sigma_x = Var(x^*)$.