Practice3 Heteroschedasticity

Federica Romei

November 19, 2012

1. Set M, numbers of Montecarlo simulation equal to 100 and N, numbers of random draws equal to 100. Define

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

where

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

(a) Estimate β_1 assuming that $Var(x_i u_i) = f(x_i) = \sigma^2 E(x_i' x_i)$ where σ^2 is the variance of u_i using heteroschedastic and homoschedastic variance estimator. Build a 95% and a 90% c.i. and count the number of times that β is within the interval ;

(b) Do the same under the assumption that $Var(x_i u_i) = f(x_i)$. (*HINT: Write* $u_i = u_i * \sqrt{x_i^2}$). What does it change? What do you learn?

Solution:

```
\begin{footnotesize}
clear;
clc;
%% Second Practice
%% Answer to the questions:
%
g=1; % Define the numbers of random draws and Montecarlo simulations
test=5; % This define the C.I.
s='homo'; % Define homoschedasticity or heteroschedasticity
if strcmp('heteroschedastic', s)
    display('Heteroschedasticity')
```

```
else
    display('Homoschedasticity')
end
if g==1;
N=1000; % Number of sample draw
M=1000; % Number of time we do the loop- Number of MonteCarlo
else
    N=100;
    M=100;
end
k=1; % Number of coefficient we want to estimate excluding the constant

%% Initialize the Vector and Matrices
   beta=NaN(2,1);
   Sigma=NaN(2,2);



%% IF:
% We do not want to loose time changing everywhere the name of
% distribution.



    distribution='norm';
    Parameters = [0,1];

%% Beta
% We decide which are going to be the exact value of beta. The first is the
% constant while the second is the one that multiply x-vector.
   Real_Beta= [1 ; 5];

   %% Initialize the Confidence Interval

   Intervarl_hom=NaN(M,2);
   Intervarl_het =NaN(M,2);

   %%
```

```
   if test==5;
   Number=1.96;
   else
       Number=1.64;
   end

   %% Loop
for m=1:M
    %% Matrix of x:
    % We create a matrix of x where in the first column there will be all
    % ones while in the second column there will be data distributed as a
    % Normal with mean  5 and wariance equal to 1.

 x=[ones(N,1) normrnd(0,1,N,1)]; % X-Data-Vector
 %% u-vector:
 % U vector should be a vector N*1 with zero mean.
 % We want that E[u|x]=0 then E[u]=0 too. Indeed we can apply Law of
 % Iterated Mean to prove that E_x[E[u|x]]=E[u].
    % We know that if E[u|x]=0 -> E[u]=0 that is not equivalent to say if
    % E[u]=0 ->E[u|x]=0. But we know that if E[u]different from 0
    % ->E[u|x] different from zero too.
 u= randraw(distribution, Parameters, [N 1]);
 if strcmp('heteroschedastic', s)
    u= u .*sqrt(x(:,2).^2); % Heteroschedastic Variance
 else
     u=u; % Homoschedastic Case
 end



%% Y
% We build y as y=x * beta+ u where u is orthogonal to x (independent).

y = x* Real_Beta+ u;
```

```
% We estimate beta as :
% \hat{Beta}= (x'x)^{-1}*x'y
beta= inv(transpose(x)*x)*transpose(x)*y;

 %% Estimation of u_hat
u_hat= y - x *beta;

%% Variance
%% Estimation of u_hat variance
variance=(u_hat'*u_hat)/(N-k-1);




%% Variance
% Under Homoschedasticity we estimate the variance matrix of variance-covariance as:
% VAR(\hat_{\beta})=( \frac{1}{N-k-1}\sum_{i=0}^{N} \hat{u}_i^2) * (x'x)^{-1}
Sigma_hom(:,:)=variance*inv(transpose(x)*x);
%% Heteroschedastic case
XE=x.*[u_hat u_hat];%  TEHERE IS a BETTER WAY TO WRITE THIS- UP TO YOU!
Sigma_het (:,:)=(x'*x)\(XE'*XE)/(x'*x);


%% Build the C.I.
Intervarl_hom(m,:)=[(beta(2,1)-Number*sqrt(Sigma_hom(2,2)))...
 (beta(2,1)+Number*sqrt(Sigma_hom(2,2)))];
Intervarl_het(m,:)=[(beta(2,1)-Number*sqrt(Sigma_het(2,2)))...
 (beta(2,1)+Number*sqrt(Sigma_het(2,2)))];




end

%% Count the Numbers of time that Real Beta is in the C.I.
%i=0;
```

```
%j=0;

Hom=find(Intervarl_hom(:,2)<Real_Beta(k+1,1) | Intervarl_hom(:,1)>(Real_Beta(k+1,1))
Hom_Prob=1-length(Hom)/ size(Intervarl_hom,1);



Het=find(Intervarl_het(:,2)<Real_Beta(k+1,1) | Intervarl_het(:,1)>(Real_Beta(k+1,1))
Het_Prob=1-length(Het)/ size(Intervarl_het,1);

%% Display the Result
display('Homoschedastic C. I.')
display(Hom_Prob)
display('Heteroschedastic C. I.')
display(Het_Prob)
\end{footnotesize}
```

1 Problem Set 3

2. A stardard stata output gives to you the following informations: 1

```
Linear regression                                     Number of obs =     706
                                                      F(  1,   704) =   65.69
                                                      Prob > F      =  0.0000
                                                      R-squared     =  0.1033
                                                      Root MSE      =  421.14

------------------------------------------------------------------------------
             |               Robust
       sleep |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      totwrk |  -.1507458    .0185992    -8.10   0.000    -.1872623   -.1142294
       _cons |   3586.377    41.98156    85.43   0.000     3503.953    3668.801
------------------------------------------------------------------------------
```

.

   (a) Write a function in Matlab with the same output (exept for F, $Prob > F$ and $Root \quad MSE$).(Remember that it should be as flexible as possible).

3. (a) Show analytically and graphically that if n, observations is equal to $k + 1$, numbers of coefficients, then $R^2 = 1$.