Boot-Strapping in Simple Regression

Amal Agarwal, Sagar Chordia, Tuhin Sarkar

Table 1 : Simulation regression results of Model-based error normal distribution ($\varepsilon_i \sim N(0, 1)$, B = 2000, α = 0.05)

β_1	Statistic	OLS	LAD	LMS
Confidence Interval	number containing β_{10}	734	1329	1524
Lower Limit of Confidence Interval	minimum	35.84046	35.54002	35.15892
	maximum	37.02594	37.04596	37.3829
	mean	36.51996	36.3844	36.33775
	median	36.53213	36.40568	36.34371
	standard deviation	0.179669	0.292927	0.346871
Upper Limit of Confidence Interval	minimum	36.62536	36.58277	35.15892
	maximum	37.78652	38.06254	38.37167
	mean	37.17265	37.30813	37.36648
	median	37.16744	37.29341	37.36286
	standard deviation	0.179017	0.274441	0.302602
$\hat{\beta}_1$	minimum	36.24005	36.07642	35.74465
	maximum	37.38269	37.51898	37.85468
	mean	36.84631	36.84626	36.85212
	median	36.84825	36.85237	36.85237
	standard deviation	0.176028	0.27849	0.31962

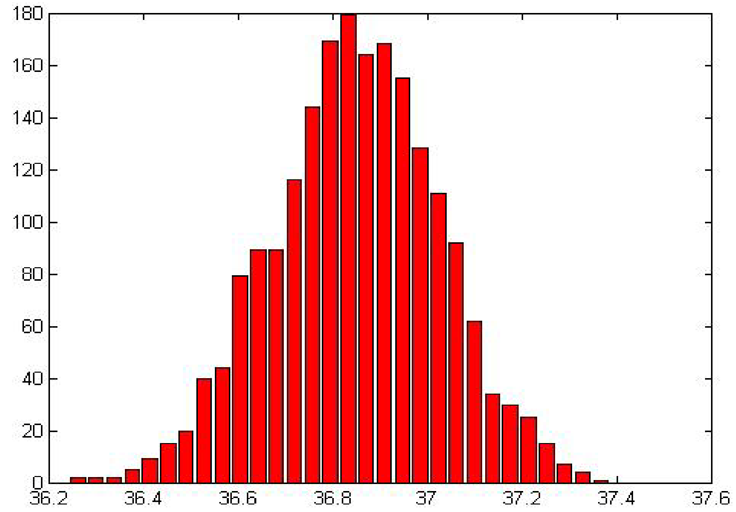

Figure 1. Histogram of coefficient β_1 estimation by OLS method for model-based Bootstrap model (error is normal distribution)

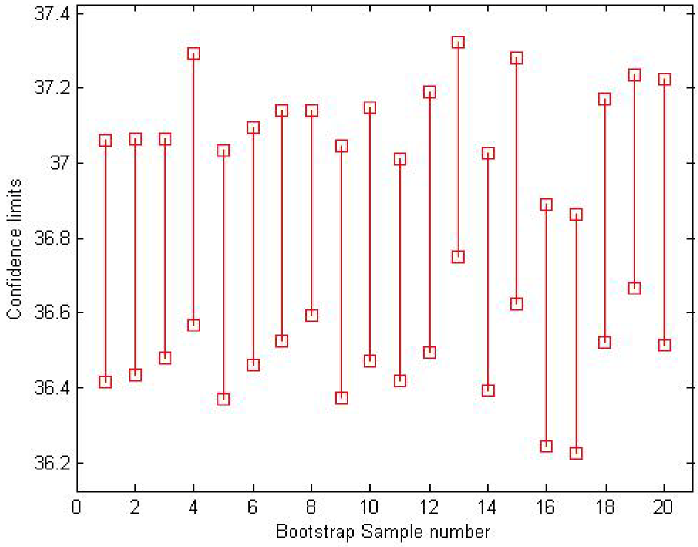

Figure 2. Confidence interval of coefficient $\boldsymbol{\beta_1}$ estimation by OLS method for model-based Bootstrap model (error is normal distribution and $\boldsymbol{B}$ = 2000)

Table 2 : Simulation regression results of Model-based error uniform distribution ($\boldsymbol{\varepsilon_i \sim U(-\sqrt{12}, \sqrt{12})}$, $\boldsymbol{B}$ = 2000, α = 0.05)

$\boldsymbol{\beta_1}$	**statistic**	**OLS**	**LAD**	**LMS**
Confidence Interval	number containing β_{10}	825	2000	1683
Lower Limit of Confidence Interval	minimum	36.82132	36.91626	36.62037
	maximum	37.11288	37.29047	37.49834
	mean	36.99008	37.09256	37.03786
	median	36.99151	37.09249	37.03595
	standard deviation	0.045735	0.057817	0.08292
Upper Limit of Confidence Interval	minimum	37.03052	36.85465	36.62037
	maximum	37.32048	37.26026	37.57
	mean	37.1553	37.0529	37.11343
	median	37.15331	37.05355	37.11116
	standard deviation	0.044622	0.062112	0.081627
$\hat{\beta}_1$	minimum	36.92759	36.88545	36.66743
	maximum	37.21668	37.27537	37.53417
	mean	37.07269	37.07273	37.07564
	median	37.07169	37.07306	37.07306
	standard deviation	0.043174	0.059905	0.082197

Table 3 : Cases simulation regression results (B = 2000, α = 0.05, ρ = 0.7)

$\boldsymbol{\beta_1}$	**Statistic**	**OLS**	**LAD**	**LMS**
Confidence Interval	Number containing β_{10}	511	1771	1833
Lower Limit of Confidence Interval	minimum	-1.33659	-1.80676	-3.86731
	maximum	1.121802	1.472863	1.48619
	mean	0.472889	0.308815	-1.07647
	median	0.485913	0.383328	-0.98229
	standard deviation	0.226764	0.463216	0.661598
Upper Limit of Confidence Interval	minimum	0.36316	0.367907	-3.86731
	maximum	2.832343	3.414443	6.435961
	mean	1.503245	1.63573	2.751256
	median	1.481236	1.558748	2.70095
	standard deviation	0.298448	0.451254	0.979869
$\hat{\beta}_1$	minimum	-0.4577	-0.71943	-1.66747
	maximum	1.920027	2.274507	3.337119
	mean	0.988067	0.972273	0.837392
	median	0.978469	0.98243	0.952451
	standard deviation	0.243123	0.431192	0.733461

Code for Bootstrapping in Simple Regression

(To be run in Matlab version≥ 7.10.0.499 (R2010a))

1. Model-based Bootstrap Regression

a) Normal Error:

```
% Model Based Bootstrap Regression Analysis with normal error
% Disclaimer: Wait for atleast 7 minutes for the program to give results
% Note the tables are automatically created as excel files in the MATLAB
folder inside the documents (can be set while installation)
clear all;
clc;

n = 14; % sample size
B = 2000;   % number of bootstrap samples
alpha = 0.05;   % level of significance
X_1 = ones(n,1);
X_2 = 0+10*rand(n,1);
X = [X_1, X_2];

% Original coefficients
beta0 = 0;
beta1 = 37;

% Original Model
er = normrnd(0,1,n,1);
Y = beta0 + beta1*(X_2) + er;

% regression on original data
[b_OLS,bint,r,rint,stats] = regress(Y,X);
Y_cap = b_OLS(1) + b_OLS(2)*(X_2);
sigma_cap_OLS = sqrt((sum((Y-(b_OLS(1)+b_OLS(2)*X_2)).^2))/(n-2));

% Bootstrap Samples
H = X*(inv(X'*X))*X';
var_r = zeros(n,1);
for i=1:n
    var_r(i) = stats(4)*(1-H(i,i));
end
rs = r./var_r;  % Revised Studentized Residuals
rs_r = rs - mean(rs);   % Cntralized Residuals
f = @(x) x;
bootstat = bootstrp(B,f,rs_r);

% b_OLS_b contains OLS estimated regression parameters in bootstrap samples
Y_b = zeros(n,B);
b_OLS_b = zeros(2,B);
bint_b = zeros(2,2,B);
r_b = zeros(n,B);
rint_b = zeros(n,2,B);
stats_b = zeros(4,B);
sigma_cap_OLS_b = zeros(1,B);
theta_star_OLS_b = zeros(1,B);
```

```
for j=1:B
    Y_b(:,j) = b_OLS(1) + b_OLS(2)*X_2 + (bootstat(j,:))';
    [b_OLS_b(:,j),bint_b(:,:,j),r_b(:,j),rint_b(:,:,j),stats_b(:,j)] =
regress(Y_b(:,j),X);
    sigma_cap_OLS_b(j) = sqrt((sum((Y_b(:,j)-
(b_OLS(1)+b_OLS(2)*X_2)).^2))/(n-2));
    theta_star_OLS_b(j) = (b_OLS_b(2,j)-
(b_OLS(2)))/((sigma_cap_OLS_b(j))/(sqrt(sum((X_2-mean(X_2)).^2))));
end
% histogram plot of beta1 estimated by OLS
[ne,xc] = hist(b_OLS_b(2,:),30,'-r');
bh = bar(xc,ne);
set(bh,'facecolor',[1 0 0]);

% Analysis on Beta1 estimate by OLS
min_b1_OLS_b = min(b_OLS_b(2,:));
max_b1_OLS_b = max(b_OLS_b(2,:));
mean_b1_OLS_b = mean(b_OLS_b(2,:));
median_b1_OLS_b = median(b_OLS_b(2,:));
std_b1_OLS_b = std(b_OLS_b(2,:));

% Confidence Intervals for beta1 using OLS estimated beta1
theta_star_OLS_b_sorted = sort(theta_star_OLS_b);
omega_OLS = theta_star_OLS_b_sorted((1-(alpha/2))*B);
% In original sample
CI_L_beta1_OLS = b_OLS(2)-((omega_OLS*sigma_cap_OLS)/(sqrt(sum((X_2-
mean(X_2)).^2))));
CI_U_beta1_OLS = b_OLS(2)+((omega_OLS*sigma_cap_OLS)/(sqrt(sum((X_2-
mean(X_2)).^2))));
% In Bootstrap samples
CI_L_beta1_OLS_b = b_OLS_b(2,:)-
((omega_OLS*sigma_cap_OLS_b)./(sqrt(sum((X_2-mean(X_2)).^2))));
CI_U_beta1_OLS_b =
b_OLS_b(2,:)+((omega_OLS*sigma_cap_OLS_b)./(sqrt(sum((X_2-
mean(X_2)).^2))));

% Analysis on Confidence Intervals for beta1 estimated by OLS
% CI Lower Limit
min_CI_L_beta1_OLS_b = min(CI_L_beta1_OLS_b);
max_CI_L_beta1_OLS_b = max(CI_L_beta1_OLS_b);
mean_CI_L_beta1_OLS_b = mean(CI_L_beta1_OLS_b);
median_CI_L_beta1_OLS_b = median(CI_L_beta1_OLS_b);
std_CI_L_beta1_OLS_b = std(CI_L_beta1_OLS_b);
% CI Upper Limit
min_CI_U_beta1_OLS_b = min(CI_U_beta1_OLS_b);
max_CI_U_beta1_OLS_b = max(CI_U_beta1_OLS_b);
mean_CI_U_beta1_OLS_b = mean(CI_U_beta1_OLS_b);
median_CI_U_beta1_OLS_b = median(CI_U_beta1_OLS_b);
std_CI_U_beta1_OLS_b = std(CI_U_beta1_OLS_b);

% Bootstrap sample number for which beta1 lies in the confidence limits by
OLS
diff_OLS = Inf*ones(1,B);
for j=1:B
    if beta1<=CI_U_beta1_OLS_b(j) && beta1>=CI_L_beta1_OLS_b(j)
        diff_OLS(j) = abs(((CI_L_beta1_OLS_b(j)+CI_U_beta1_OLS_b(j))/2)-
beta1);
    end
end
for j=1:B
```

```
    if diff_OLS(j)==min(diff_OLS)
        j_min=j;
    end
end
beta1_CI_check_OLS = j_min;

% Confidence Intervals plot for beta1 by OLS
B_plot = [(1:20)',(1:20)'];
y1=min(CI_L_beta1_OLS_b(1:20))-0.1;
y2=max(CI_U_beta1_OLS_b(1:20))+0.1;
CI_plot = [CI_L_beta1_OLS_b(1:20)',CI_U_beta1_OLS_b(1:20)'];
figure, plot(B_plot(1,:), CI_plot(1,:), '-rs',...
             B_plot(2,:), CI_plot(2,:), '-rs',...
             B_plot(3,:), CI_plot(3,:), '-rs',...
             B_plot(4,:), CI_plot(4,:), '-rs',...
             B_plot(5,:), CI_plot(5,:), '-rs',...
             B_plot(6,:), CI_plot(6,:), '-rs',...
             B_plot(7,:), CI_plot(7,:), '-rs',...
             B_plot(8,:), CI_plot(8,:), '-rs',...
             B_plot(9,:), CI_plot(9,:), '-rs',...
             B_plot(10,:), CI_plot(10,:), '-rs',...
             B_plot(11,:), CI_plot(11,:), '-rs',...
             B_plot(12,:), CI_plot(12,:), '-rs',...
             B_plot(13,:), CI_plot(13,:), '-rs',...
             B_plot(14,:), CI_plot(14,:), '-rs',...
             B_plot(15,:), CI_plot(15,:), '-rs',...
             B_plot(16,:), CI_plot(16,:), '-rs',...
             B_plot(17,:), CI_plot(17,:), '-rs',...
             B_plot(18,:), CI_plot(18,:), '-rs',...
             B_plot(19,:), CI_plot(19,:), '-rs',...
             B_plot(20,:), CI_plot(20,:), '-rs');
title ('Confidence interval of coefficient Beta1 estimation by OLS method
for model-based Bootstrap model(error is normal distribution and B =
2000)');
axis([0 21 y1 y2]);
xlabel('Bootstrap Sample number');
ylabel('Confidence limits');

% LAD estimation on original sample
beta0_LAD = zeros(n,n);
beta1_LAD = zeros(n,n);
d = zeros(n,n);
for i = 1:n
    for j = 1:n
        if i~=j
            beta1_LAD(i,j) = (Y(j)-Y(i))/(X_2(j)-X_2(i));
            beta0_LAD(i,j) = Y(j)-beta1_LAD(i,j)*X_2(j);
            d(i,j) = sum(abs(Y-(beta0_LAD(i,j)+beta1_LAD(i,j)*X_2)));
        end
    end
end
d_min = Inf;
for i=1:n
    for j=1:n
        if i~=j
            if d_min>d(i,j);
                d_min=d(i,j);
                i_min = i;
                j_min = j;
            end
        end
```

```
    end
end
b0_LAD = beta0_LAD(i_min,j_min);
b1_LAD = beta1_LAD(i_min,j_min);

% LAD estimation on Bootstrap Samples
beta0_LAD_b = zeros(n,n,B);
beta1_LAD_b = zeros(n,n,B);
d_b = zeros(n,n,B);
for k = 1:B
    for i = 1:n
        for j = 1:n
            if i~=j
                beta1_LAD_b(i,j,k) = (Y_b(j,k)-Y_b(i,k))/(X_2(j)-X_2(i));
                beta0_LAD_b(i,j,k) = Y_b(j,k)-beta1_LAD_b(i,j,k)*X_2(j);
                d(i,j,k) = sum(abs(Y_b(:,k)-
(beta0_LAD_b(i,j,k)+beta1_LAD_b(i,j,k)*X_2)));
            end
        end
    end
end
d_min = Inf*ones(1,B);
i_min = zeros(1,B);
j_min = zeros(1,B);
for k=1:B
    for i=1:n
        for j=1:n
            if i~=j
                if d_min(k)>d(i,j,k);
                    d_min(k) = d(i,j,k);
                    i_min(k) = i;
                    j_min(k) = j;
                end
            end
        end
    end
end
b0_LAD_b = zeros(1,B);
b1_LAD_b = zeros(1,B);
for k=1:B
    b0_LAD_b(k) = beta0_LAD_b(i_min(k),j_min(k),k);
    b1_LAD_b(k) = beta1_LAD_b(i_min(k),j_min(k),k);
end

% Analysis on Beta1 estimate by LAD
min_b1_LAD_b = min(b1_LAD_b(1,:));
max_b1_LAD_b = max(b1_LAD_b(1,:));
mean_b1_LAD_b = mean(b1_LAD_b(1,:));
median_b1_LAD_b = median(b1_LAD_b(1,:));
std_b1_LAD_b = std(b1_LAD_b(1,:));

% Confidence Intervals for beta1 using LAD estimated beta1
sigma_cap_LAD = sqrt((sum((Y-(b0_LAD+b1_LAD*X_2)).^2))/(n-2));
sigma_cap_LAD_b = zeros(1,B);
theta_star_LAD_b = zeros(1,B);
for j=1:B
    sigma_cap_LAD_b(j) = sqrt((sum((Y_b(:,j)-(b0_LAD+b1_LAD*X_2)).^2))/(n-
2));
    theta_star_LAD_b(j) = (b1_LAD_b(j)-
(b1_LAD))/((sigma_cap_LAD_b(j))/(sqrt(sum((X_2-mean(X_2)).^2))));
end
```

```
theta_star_LAD_b_sorted = sort(theta_star_LAD_b);
omega_LAD = theta_star_LAD_b_sorted((1-(alpha/2))*B);
% In original sample
CI_L_beta1_LAD = b1_LAD-((omega_LAD*sigma_cap_LAD)/(sqrt(sum((X_2-
mean(X_2)).^2))));
CI_U_beta1_LAD = b1_LAD+((omega_LAD*sigma_cap_LAD)/(sqrt(sum((X_2-
mean(X_2)).^2))));
% In Bootstrap samples
CI_L_beta1_LAD_b = b1_LAD_b(1,:)-
((omega_LAD*sigma_cap_LAD_b)./(sqrt(sum((X_2-mean(X_2)).^2))));
CI_U_beta1_LAD_b =
b1_LAD_b(1,:)+((omega_LAD*sigma_cap_LAD_b)./(sqrt(sum((X_2-
mean(X_2)).^2))));

% Analysis on Confidence Intervals for beta1 estimated by LAD
% CI Lower Limit
min_CI_L_beta1_LAD_b = min(CI_L_beta1_LAD_b);
max_CI_L_beta1_LAD_b = max(CI_L_beta1_LAD_b);
mean_CI_L_beta1_LAD_b = mean(CI_L_beta1_LAD_b);
median_CI_L_beta1_LAD_b = median(CI_L_beta1_LAD_b);
std_CI_L_beta1_LAD_b = std(CI_L_beta1_LAD_b);
% CI Upper Limit
min_CI_U_beta1_LAD_b = min(CI_U_beta1_LAD_b);
max_CI_U_beta1_LAD_b = max(CI_U_beta1_LAD_b);
mean_CI_U_beta1_LAD_b = mean(CI_U_beta1_LAD_b);
median_CI_U_beta1_LAD_b = median(CI_U_beta1_LAD_b);
std_CI_U_beta1_LAD_b = std(CI_U_beta1_LAD_b);

% Bootstrap sample number for which beta1 lies in the confidence limits by
LAD
diff_LAD = Inf*ones(1,B);
for j=1:B
    if beta1<=CI_U_beta1_LAD_b(j) && beta1>=CI_L_beta1_LAD_b(j)
        diff_LAD(j) = abs(((CI_L_beta1_LAD_b(j)+CI_U_beta1_LAD_b(j))/2)-
beta1);
    end
end
for j=1:B
    if diff_LAD(j)==min(diff_LAD)
        j_min=j;
    end
end
beta1_CI_check_LAD = j_min;

% LMS estimation on original sample
d_cap = Inf;
% X_2_sorted = sort(X_2);
beta0_LMS = zeros(n,n,n);
beta1_LMS = zeros(n,n,n);
d = zeros(n,n,n);
for i=1:n
    for j=1:n
        for k=1:n
            beta1_LMS(i,j,k) = (Y(i)-Y(k))/(X_2(i)-X_2(k));
            beta0_LMS(i,j,k) = Y(j)+Y(k)-beta1_LMS(i,j,k)*(X_2(j)+X_2(k));
            d(i,j,k) = median((Y-
(beta0_LMS(i,j,k)+beta1_LMS(i,j,k)*X_2)).^2);
        end
    end
end
for i=1:n
```

```
    for j=1:n
        for k=1:n
            if d_cap>d(i,j,k);
                d_cap=d(i,j,k);
                i_cap = i;
                j_cap = j;
                k_cap = k;
            end
        end
    end
end
b0_LMS = beta0_LMS(i_cap,j_cap,k_cap);
b1_LMS = beta1_LMS(i_cap,j_cap,k_cap);

% LMS estimation on Bootstrap Samples
d_cap_b = Inf*ones(1,B);
X_2_sorted = sort(X_2);
beta0_LMS_b = zeros(n,n,n,B);
beta1_LMS_b = zeros(n,n,n,B);
d = zeros(n,n,n,B);
for l = 1:B
    for i = 1:n
        for j = 1:n
            for k=1:n
                beta1_LMS_b(i,j,k,l) = (Y_b(i,l)-Y_b(k,l))/(X_2(i)-X_2(k));
                beta0_LMS_b(i,j,k,l) = Y_b(j,l)+Y_b(k,l)-
beta1_LMS_b(i,j,k,l)*(X_2(j)+X_2(k));
                d(i,j,k,l) = median((Y_b(:,l)-
(beta0_LMS_b(i,j,k,l)+beta1_LMS_b(i,j,k,l)*X_2)).^2);
            end
        end
    end
end
i_cap_b = zeros(1,B);
j_cap_b = zeros(1,B);
k_cap_b = zeros(1,B);
for l=1:B
    for i=1:n
        for j=1:n
            for k=1:n
                if d_cap_b(l)>d(i,j,k,l);
                    d_cap_b(l) = d(i,j,k,l);
                    i_cap_b(l) = i;
                    j_cap_b(l) = j;
                    k_cap_b(l) = k;
                end
            end
        end
    end
end
b0_LMS_b = zeros(1,B);
b1_LMS_b = zeros(1,B);
for l=1:B
    b0_LMS_b(l) = beta0_LMS_b(i_cap_b(l),j_cap_b(l),k_cap_b(l),l);
    b1_LMS_b(l) = beta1_LMS_b(i_cap_b(l),j_cap_b(l),k_cap_b(l),l);
end

% Analysis on Beta1 estimate by LMS
min_b1_LMS_b = min(b1_LMS_b(1,:));
max_b1_LMS_b = max(b1_LMS_b(1,:));
mean_b1_LMS_b = mean(b1_LMS_b(1,:));
```

```
median_b1_LMS_b = median(b1_LMS_b(1,:));
std_b1_LMS_b = std(b1_LMS_b(1,:));

% Confidence Intervals for beta1 using LMS estimated beta1
sigma_cap_LMS = sqrt((sum((Y-(b0_LMS+b1_LMS*X_2)).^2))/(n-2));
sigma_cap_LMS_b = zeros(1,B);
theta_star_LMS_b = zeros(1,B);
for j=1:B
    sigma_cap_LMS_b(j) = sqrt((sum((Y_b(:,j)-(b0_LMS+b1_LMS*X_2)).^2))/(n-
2));
    theta_star_LMS_b(j) = (b1_LMS_b(j)-
(b1_LMS))/((sigma_cap_LMS_b(j))/(sqrt(sum((X_2-mean(X_2)).^2))));
end
theta_star_LMS_b_sorted = sort(theta_star_LMS_b);
omega_LMS = theta_star_LMS_b_sorted((1-(alpha/2))*B);
% In original sample
CI_L_beta1_LMS = b1_LMS-((omega_LMS*sigma_cap_LMS)/(sqrt(sum((X_2-
mean(X_2)).^2))));
CI_U_beta1_LMS = b1_LMS+((omega_LMS*sigma_cap_LMS)/(sqrt(sum((X_2-
mean(X_2)).^2))));
% In Bootstrap samples
CI_L_beta1_LMS_b = b1_LMS_b(1,:)-
((omega_LMS*sigma_cap_LMS_b)./(sqrt(sum((X_2-mean(X_2)).^2))));
CI_U_beta1_LMS_b =
b1_LMS_b(1,:)+((omega_LMS*sigma_cap_LMS_b)./(sqrt(sum((X_2-
mean(X_2)).^2))));

% Analysis on Confidence Intervals for beta1 estimated by LMS
% CI Lower Limit
min_CI_L_beta1_LMS_b = min(CI_L_beta1_LMS_b);
max_CI_L_beta1_LMS_b = max(CI_L_beta1_LMS_b);
mean_CI_L_beta1_LMS_b = mean(CI_L_beta1_LMS_b);
median_CI_L_beta1_LMS_b = median(CI_L_beta1_LMS_b);
std_CI_L_beta1_LMS_b = std(CI_L_beta1_LMS_b);
% CI Upper Limit
min_CI_U_beta1_LMS_b = min(CI_U_beta1_LMS_b);
max_CI_U_beta1_LMS_b = max(CI_U_beta1_LMS_b);
mean_CI_U_beta1_LMS_b = mean(CI_U_beta1_LMS_b);
median_CI_U_beta1_LMS_b = median(CI_U_beta1_LMS_b);
std_CI_U_beta1_LMS_b = std(CI_U_beta1_LMS_b);

% Bootstrap sample number for which beta1 lies in the confidence limits by
LMS
diff_LMS = Inf*ones(1,B);
for j=1:B
    if beta1<=CI_U_beta1_LMS_b(j) && beta1>=CI_L_beta1_LMS_b(j)
        diff_LMS(j) = abs(((CI_L_beta1_LMS_b(j)+CI_U_beta1_LMS_b(j))/2)-
beta1);
    end
end
for j=1:B
    if diff_LMS(j)==min(diff_LMS)
        j_min=j;
    end
end
beta1_CI_check_LMS = j_min;

% generating table in excel
filename = 'table1.xlsx';
A = {'Beta_1','statistic','OLS','LAD','LMS';...
    'Confidence Interval','Number containing beta1',...
```

```
     beta1_CI_check_OLS,beta1_CI_check_LAD,beta1_CI_check_LMS;...
    'Lower Limit of Confidence Interval',...

'minimum',min_CI_L_beta1_OLS_b,min_CI_L_beta1_LAD_b,min_CI_L_beta1_LMS_b;..
.

'','maximum',max_CI_L_beta1_OLS_b,max_CI_L_beta1_LAD_b,max_CI_L_beta1_LMS_b
;...

'','mean',mean_CI_L_beta1_OLS_b,mean_CI_L_beta1_LAD_b,mean_CI_L_beta1_LMS_b
;...

'','median',median_CI_L_beta1_OLS_b,median_CI_L_beta1_LAD_b,median_CI_L_bet
a1_LMS_b;...
    '','standard
deviation',std_CI_L_beta1_OLS_b,std_CI_L_beta1_LAD_b,std_CI_L_beta1_LMS_b;
    'Upper Limit of Confidence Interval',...

'minimum',min_CI_U_beta1_OLS_b,min_CI_U_beta1_LAD_b,min_CI_L_beta1_LMS_b;..
.

'','maximum',max_CI_U_beta1_OLS_b,max_CI_U_beta1_LAD_b,max_CI_U_beta1_LMS_b
;...

'','mean',mean_CI_U_beta1_OLS_b,mean_CI_U_beta1_LAD_b,mean_CI_U_beta1_LMS_b
;...

'','median',median_CI_U_beta1_OLS_b,median_CI_U_beta1_LAD_b,median_CI_U_bet
a1_LMS_b;...
    '','standard
deviation',std_CI_U_beta1_OLS_b,std_CI_U_beta1_LAD_b,std_CI_U_beta1_LMS_b;
    'Estimated Beta_1',...
    'minimum',min_b1_OLS_b,min_b1_LAD_b,min_b1_LMS_b;...
    '','maximum',max_b1_OLS_b,max_b1_LAD_b,max_b1_LMS_b;...
    '','mean',mean_b1_OLS_b,mean_b1_LAD_b,mean_b1_LMS_b;...
    '','median',median_b1_OLS_b,median_b1_LAD_b,median_b1_LMS_b;...
    '','standard deviation',std_b1_OLS_b,std_b1_LAD_b,std_b1_LMS_b;};
sheet = 1;
xlRange = 'A1';
xlswrite(filename,A,sheet,xlRange);
```

b) Uniform Error:

Same code as above except replace the “original model” block as:

```
% Original Model
er = -sqrt(12)+2*sqrt(12)*rand(n,1);
Y = beta0 + beta1*(X_2) + er;
```

2. Cases Bootstrap Regression

```
% Cases Bootstrap Regression Analysis
% Disclaimer: Wait for atleast 7 minutes for the program to give results
% Note: The tables are automatically created as excel files in the MATLAB
folder inside the documents (can be set while installation)
clear all;
clc;

n = 14; % sample size
B = 2000;   % number of bootstrap samples
alpha = 0.05;   % level of significance

% Original coefficients
beta0 = 1;
beta1 = 0.7;

% Original Model
mu = [0,1];
SIGMA = [1, 0.7; 0.7, 1];
xy = mvnrnd(mu,SIGMA,n);
X_1 = ones(n,1);
X_2 = xy(:,1);
X = [X_1, X_2];
Y = xy(:,2);

% regression on original data
[b_OLS,bint,r,rint,stats] = regress(Y,X);
Y_cap = b_OLS(1) + b_OLS(2)*(X_2);
sigma_cap_OLS = sqrt((sum((Y-(b_OLS(1)+b_OLS(2)*X_2)).^2))/(n-2));

% Bootstrap Samples
f = @(x) x;
bootstat_xy = bootstrp(B,f,xy);

% b_OLS_b contains OLS estimated regression parameters in bootstrap samples
X_2_b = zeros(n,B);
X_b = zeros(n,2,B);
Y_b = zeros(n,B);
b_OLS_b = zeros(2,B);
bint_b = zeros(2,2,B);
r_b = zeros(n,B);
rint_b = zeros(n,2,B);
stats_b = zeros(4,B);
sigma_cap_OLS_b = zeros(1,B);
theta_star_OLS_b = zeros(1,B);
for j=1:B
    X_2_b(:,j) = (bootstat_xy(j,1:n))';
    X_b(:,:,j) = [X_1, X_2_b(:,j)];
    Y_b(:,j) = (bootstat_xy(j,n+1:2*n))';
    [b_OLS_b(:,j),bint_b(:,:,j),r_b(:,j),rint_b(:,:,j),stats_b(:,j)] =
regress(Y_b(:,j),X_b(:,:,j));
    sigma_cap_OLS_b(j) = sqrt((sum((Y_b(:,j)-
(b_OLS(1)+b_OLS(2)*X_2)).^2))/(n-2));
    theta_star_OLS_b(j) = (b_OLS_b(2,j)-
(b_OLS(2)))/((sigma_cap_OLS_b(j))/(sqrt(sum((X_2-mean(X_2)).^2))));
end
% histogram plot of beta1 estimated by OLS
[ne,xc] = hist(b_OLS_b(2,:),30,'-r');
```

```
bh = bar(xc,ne);
set(bh,'facecolor',[1 0 0]);

% Analysis on Beta1 estimate by OLS
min_b1_OLS_b = min(b_OLS_b(2,:));
max_b1_OLS_b = max(b_OLS_b(2,:));
mean_b1_OLS_b = mean(b_OLS_b(2,:));
median_b1_OLS_b = median(b_OLS_b(2,:));
std_b1_OLS_b = std(b_OLS_b(2,:));

% Confidence Intervals for beta1 using OLS estimated beta1
theta_star_OLS_b_sorted = sort(theta_star_OLS_b);
omega_OLS = theta_star_OLS_b_sorted((1-(alpha/2))*B);
% In original sample
CI_L_beta1_OLS = b_OLS(2)-((omega_OLS*sigma_cap_OLS)/(sqrt(sum((X_2-
mean(X_2)).^2))));
CI_U_beta1_OLS = b_OLS(2)+((omega_OLS*sigma_cap_OLS)/(sqrt(sum((X_2-
mean(X_2)).^2))));
% In Bootstrap samples
CI_L_beta1_OLS_b = zeros(1,B);
CI_U_beta1_OLS_b = zeros(1,B);
for j=1:B
    CI_L_beta1_OLS_b(j) = b_OLS_b(2,j)-
((omega_OLS*sigma_cap_OLS_b(j))./(sqrt(sum((X_2_b(:,j)-
mean(X_2_b(:,j))).^2))));
    CI_U_beta1_OLS_b(j) =
b_OLS_b(2,j)+((omega_OLS*sigma_cap_OLS_b(j))./(sqrt(sum((X_2_b(:,j)-
mean(X_2_b(:,j))).^2))));
end
% Analysis on Confidence Intervals for beta1 estimated by OLS
% CI Lower Limit
min_CI_L_beta1_OLS_b = min(CI_L_beta1_OLS_b);
max_CI_L_beta1_OLS_b = max(CI_L_beta1_OLS_b);
mean_CI_L_beta1_OLS_b = mean(CI_L_beta1_OLS_b);
median_CI_L_beta1_OLS_b = median(CI_L_beta1_OLS_b);
std_CI_L_beta1_OLS_b = std(CI_L_beta1_OLS_b);
% CI Upper Limit
min_CI_U_beta1_OLS_b = min(CI_U_beta1_OLS_b);
max_CI_U_beta1_OLS_b = max(CI_U_beta1_OLS_b);
mean_CI_U_beta1_OLS_b = mean(CI_U_beta1_OLS_b);
median_CI_U_beta1_OLS_b = median(CI_U_beta1_OLS_b);
std_CI_U_beta1_OLS_b = std(CI_U_beta1_OLS_b);

% Bootstrap sample number for which beta1 lies in the confidence limits by
OLS
diff_OLS = Inf*ones(1,B);
for j=1:B
    if beta1<=CI_U_beta1_OLS_b(j) && beta1>=CI_L_beta1_OLS_b(j)
        diff_OLS(j) = abs(((CI_L_beta1_OLS_b(j)+CI_U_beta1_OLS_b(j))/2)-
beta1);
    end
end
for j=1:B
    if diff_OLS(j)==min(diff_OLS)
        j_min=j;
    end
end
beta1_CI_check_OLS = j_min;

% Confidence Intervals plot for beta1 by OLS
```

```
B_plot = [(1:20)',(1:20)'];
y1=min(CI_L_beta1_OLS_b(1:20))-0.1;
y2=max(CI_U_beta1_OLS_b(1:20))+0.1;
CI_plot = [CI_L_beta1_OLS_b(1:20)',CI_U_beta1_OLS_b(1:20)'];
figure, plot(B_plot(1,:), CI_plot(1,:), '-rs',...
             B_plot(2,:), CI_plot(2,:), '-rs',...
             B_plot(3,:), CI_plot(3,:), '-rs',...
             B_plot(4,:), CI_plot(4,:), '-rs',...
             B_plot(5,:), CI_plot(5,:), '-rs',...
             B_plot(6,:), CI_plot(6,:), '-rs',...
             B_plot(7,:), CI_plot(7,:), '-rs',...
             B_plot(8,:), CI_plot(8,:), '-rs',...
             B_plot(9,:), CI_plot(9,:), '-rs',...
             B_plot(10,:), CI_plot(10,:), '-rs',...
             B_plot(11,:), CI_plot(11,:), '-rs',...
             B_plot(12,:), CI_plot(12,:), '-rs',...
             B_plot(13,:), CI_plot(13,:), '-rs',...
             B_plot(14,:), CI_plot(14,:), '-rs',...
             B_plot(15,:), CI_plot(15,:), '-rs',...
             B_plot(16,:), CI_plot(16,:), '-rs',...
             B_plot(17,:), CI_plot(17,:), '-rs',...
             B_plot(18,:), CI_plot(18,:), '-rs',...
             B_plot(19,:), CI_plot(19,:), '-rs',...
             B_plot(20,:), CI_plot(20,:), '-rs');
title ('Confidence interval of coefficient Beta1 estimation by OLS method
for model-based Bootstrap model(error is normal distribution and B =
2000)');
axis([0 21 y1 y2]);
xlabel('Bootstrap Sample number');
ylabel('Confidence limits');

% LAD estimation on original sample
beta0_LAD = zeros(n,n);
beta1_LAD = zeros(n,n);
d = zeros(n,n);
for i = 1:n
    for j = 1:n
        if i~=j
            beta1_LAD(i,j) = (Y(j)-Y(i))/(X_2(j)-X_2(i));
            beta0_LAD(i,j) = Y(j)-beta1_LAD(i,j)*X_2(j);
            d(i,j) = sum(abs(Y-(beta0_LAD(i,j)+beta1_LAD(i,j)*X_2)));
        end
    end
end
d_min = Inf;
for i=1:n
    for j=1:n
        if i~=j
            if d_min>d(i,j);
                d_min=d(i,j);
                i_min = i;
                j_min = j;
            end
        end
    end
end
b0_LAD = beta0_LAD(i_min,j_min);
b1_LAD = beta1_LAD(i_min,j_min);

% LAD estimation on Bootstrap Samples
beta0_LAD_b = zeros(n,n,B);
```

```
beta1_LAD_b = zeros(n,n,B);
d_b = zeros(n,n,B);
for k = 1:B
    for i = 1:n
        for j = 1:n
            if i~=j
                beta1_LAD_b(i,j,k) = (Y_b(j,k)-Y_b(i,k))/(X_2_b(j,k)-
X_2_b(i,k));
                beta0_LAD_b(i,j,k) = Y_b(j,k)-
beta1_LAD_b(i,j,k)*X_2_b(j,k);
                d(i,j,k) = sum(abs(Y_b(:,k)-
(beta0_LAD_b(i,j,k)+beta1_LAD_b(i,j,k)*X_2_b(:,k))));
            end
        end
    end
end
d_min = Inf*ones(1,B);
i_min = zeros(1,B);
j_min = zeros(1,B);
for k=1:B
    for i=1:n
        for j=1:n
            if i~=j
                if d_min(k)>d(i,j,k);
                    d_min(k) = d(i,j,k);
                    i_min(k) = i;
                    j_min(k) = j;
                end
            end
        end
    end
end
b0_LAD_b = zeros(1,B);
b1_LAD_b = zeros(1,B);
for k=1:B
    b0_LAD_b(k) = beta0_LAD_b(i_min(k),j_min(k),k);
    b1_LAD_b(k) = beta1_LAD_b(i_min(k),j_min(k),k);
end

% Analysis on Beta1 estimate by LAD
min_b1_LAD_b = min(b1_LAD_b(1,:));
max_b1_LAD_b = max(b1_LAD_b(1,:));
mean_b1_LAD_b = mean(b1_LAD_b(1,:));
median_b1_LAD_b = median(b1_LAD_b(1,:));
std_b1_LAD_b = std(b1_LAD_b(1,:));

% Confidence Intervals for beta1 using LAD estimated beta1
sigma_cap_LAD = sqrt((sum((Y-(b0_LAD+b1_LAD*X_2)).^2))/(n-2));
sigma_cap_LAD_b = zeros(1,B);
theta_star_LAD_b = zeros(1,B);
for j=1:B
    sigma_cap_LAD_b(j) = sqrt((sum((Y_b(:,j)-
(b0_LAD+b1_LAD*X_2_b(:,j))).^2))/(n-2));
    theta_star_LAD_b(j) = (b1_LAD_b(j)-
(b1_LAD))/((sigma_cap_LAD_b(j))/(sqrt(sum((X_2_b(:,j)-
mean(X_2_b(:,j))).^2))));
end
theta_star_LAD_b_sorted = sort(theta_star_LAD_b);
omega_LAD = theta_star_LAD_b_sorted((1-(alpha/2))*B);
% In original sample
```

```
CI_L_beta1_LAD = b1_LAD-((omega_LAD*sigma_cap_LAD)/(sqrt(sum((X_2-
mean(X_2)).^2))));
CI_U_beta1_LAD = b1_LAD+((omega_LAD*sigma_cap_LAD)/(sqrt(sum((X_2-
mean(X_2)).^2))));
% In Bootstrap samples
CI_L_beta1_LAD_b = zeros(1,B);
CI_U_beta1_LAD_b = zeros(1,B);
for j=1:B
    CI_L_beta1_LAD_b(j) = b1_LAD_b(1,j)-
((omega_LAD*sigma_cap_LAD_b(j))./(sqrt(sum((X_2_b(:,j)-
mean(X_2_b(:,j))).^2))));
    CI_U_beta1_LAD_b(j) =
b1_LAD_b(1,j)+((omega_LAD*sigma_cap_LAD_b(j))./(sqrt(sum((X_2_b(:,j)-
mean(X_2_b(:,j))).^2))));
end

% Analysis on Confidence Intervals for beta1 estimated by LAD
% CI Lower Limit
min_CI_L_beta1_LAD_b = min(CI_L_beta1_LAD_b);
max_CI_L_beta1_LAD_b = max(CI_L_beta1_LAD_b);
mean_CI_L_beta1_LAD_b = mean(CI_L_beta1_LAD_b);
median_CI_L_beta1_LAD_b = median(CI_L_beta1_LAD_b);
std_CI_L_beta1_LAD_b = std(CI_L_beta1_LAD_b);
% CI Upper Limit
min_CI_U_beta1_LAD_b = min(CI_U_beta1_LAD_b);
max_CI_U_beta1_LAD_b = max(CI_U_beta1_LAD_b);
mean_CI_U_beta1_LAD_b = mean(CI_U_beta1_LAD_b);
median_CI_U_beta1_LAD_b = median(CI_U_beta1_LAD_b);
std_CI_U_beta1_LAD_b = std(CI_U_beta1_LAD_b);

% Bootstrap sample number for which beta1 lies in the confidence limits by
LAD
diff_LAD = Inf*ones(1,B);
for j=1:B
    if beta1<=CI_U_beta1_LAD_b(j) && beta1>=CI_L_beta1_LAD_b(j)
        diff_LAD(j) = abs(((CI_L_beta1_LAD_b(j)+CI_U_beta1_LAD_b(j))/2)-
beta1);
    end
end
for j=1:B
    if diff_LAD(j)==min(diff_LAD)
        j_min=j;
    end
end
beta1_CI_check_LAD = j_min;

% LMS estimation on original sample
d_cap = Inf;
% X_2_sorted = sort(X_2);
beta0_LMS = zeros(n,n,n);
beta1_LMS = zeros(n,n,n);
d = zeros(n,n,n);
for i=1:n
    for j=1:n
        for k=1:n
            beta1_LMS(i,j,k) = (Y(i)-Y(k))/(X_2(i)-X_2(k));
            beta0_LMS(i,j,k) = Y(j)+Y(k)-beta1_LMS(i,j,k)*(X_2(j)+X_2(k));
            d(i,j,k) = median((Y-
(beta0_LMS(i,j,k)+beta1_LMS(i,j,k)*X_2)).^2);
        end
    end
```

```
end
for i=1:n
    for j=1:n
        for k=1:n
            if d_cap>d(i,j,k);
                d_cap=d(i,j,k);
                i_cap = i;
                j_cap = j;
                k_cap = k;
            end
        end
    end
end
b0_LMS = beta0_LMS(i_cap,j_cap,k_cap);
b1_LMS = beta1_LMS(i_cap,j_cap,k_cap);

% LMS estimation on Bootstrap Samples
d_cap_b = Inf*ones(1,B);
X_2_sorted = sort(X_2);
beta0_LMS_b = zeros(n,n,n,B);
beta1_LMS_b = zeros(n,n,n,B);
d = zeros(n,n,n,B);
for l = 1:B
    for i = 1:n
        for j = 1:n
            for k=1:n
                beta1_LMS_b(i,j,k,l) = (Y_b(i,l)-Y_b(k,l))/(X_2_b(i,l)-
X_2_b(k,l));
                beta0_LMS_b(i,j,k,l) = Y_b(j,l)+Y_b(k,l)-
beta1_LMS_b(i,j,k,l)*(X_2_b(j,l)+X_2_b(k,l));
                d(i,j,k,l) = median((Y_b(:,l)-
(beta0_LMS_b(i,j,k,l)+beta1_LMS_b(i,j,k,l)*X_2_b(:,l))).^2);
            end
        end
    end
end
i_cap_b = zeros(1,B);
j_cap_b = zeros(1,B);
k_cap_b = zeros(1,B);
for l=1:B
    for i=1:n
        for j=1:n
            for k=1:n
                if d_cap_b(l)>d(i,j,k,l);
                    d_cap_b(l) = d(i,j,k,l);
                    i_cap_b(l) = i;
                    j_cap_b(l) = j;
                    k_cap_b(l) = k;
                end
            end
        end
    end
end
b0_LMS_b = zeros(1,B);
b1_LMS_b = zeros(1,B);
for l=1:B
    b0_LMS_b(l) = beta0_LMS_b(i_cap_b(l),j_cap_b(l),k_cap_b(l),l);
    b1_LMS_b(l) = beta1_LMS_b(i_cap_b(l),j_cap_b(l),k_cap_b(l),l);
end

% Analysis on Beta1 estimate by LMS
```

```
min_b1_LMS_b = min(b1_LMS_b(1,:));
max_b1_LMS_b = max(b1_LMS_b(1,:));
mean_b1_LMS_b = mean(b1_LMS_b(1,:));
median_b1_LMS_b = median(b1_LMS_b(1,:));
std_b1_LMS_b = std(b1_LMS_b(1,:));

% Confidence Intervals for beta1 using LMS estimated beta1
sigma_cap_LMS = sqrt((sum((Y-(b0_LMS+b1_LMS*X_2)).^2))/(n-2));
sigma_cap_LMS_b = zeros(1,B);
theta_star_LMS_b = zeros(1,B);
for j=1:B
    sigma_cap_LMS_b(j) = sqrt((sum((Y_b(:,j)-
(b0_LMS+b1_LMS*X_2_b(:,j))).^2))/(n-2));
    theta_star_LMS_b(j) = (b1_LMS_b(j)-
(b1_LMS))/((sigma_cap_LMS_b(j))/(sqrt(sum((X_2_b(:,j)-
mean(X_2_b(:,j))).^2))));
end
theta_star_LMS_b_sorted = sort(theta_star_LMS_b);
omega_LMS = theta_star_LMS_b_sorted((1-(alpha/2))*B);
% In original sample
CI_L_beta1_LMS = b1_LMS-((omega_LMS*sigma_cap_LMS)/(sqrt(sum((X_2-
mean(X_2)).^2))));
CI_U_beta1_LMS = b1_LMS+((omega_LMS*sigma_cap_LMS)/(sqrt(sum((X_2-
mean(X_2)).^2))));
% In Bootstrap samples
CI_L_beta1_LMS_b = zeros(1,B);
CI_U_beta1_LMS_b = zeros(1,B);
for j=1:B
    CI_L_beta1_LMS_b(j) = b1_LMS_b(1,j)-
((omega_LMS*sigma_cap_LMS_b(j))./(sqrt(sum((X_2_b(:,j)-
mean(X_2_b(:,j))).^2))));
    CI_U_beta1_LMS_b(j) =
b1_LMS_b(1,j)+((omega_LMS*sigma_cap_LMS_b(j))./(sqrt(sum((X_2_b(:,j)-
mean(X_2_b(:,j))).^2))));
end

% Analysis on Confidence Intervals for beta1 estimated by LMS
% CI Lower Limit
min_CI_L_beta1_LMS_b = min(CI_L_beta1_LMS_b);
max_CI_L_beta1_LMS_b = max(CI_L_beta1_LMS_b);
mean_CI_L_beta1_LMS_b = mean(CI_L_beta1_LMS_b);
median_CI_L_beta1_LMS_b = median(CI_L_beta1_LMS_b);
std_CI_L_beta1_LMS_b = std(CI_L_beta1_LMS_b);
% CI Upper Limit
min_CI_U_beta1_LMS_b = min(CI_U_beta1_LMS_b);
max_CI_U_beta1_LMS_b = max(CI_U_beta1_LMS_b);
mean_CI_U_beta1_LMS_b = mean(CI_U_beta1_LMS_b);
median_CI_U_beta1_LMS_b = median(CI_U_beta1_LMS_b);
std_CI_U_beta1_LMS_b = std(CI_U_beta1_LMS_b);

% Bootstrap sample number for which beta1 lies in the confidence limits by
LMS
diff_LMS = Inf*ones(1,B);
for j=1:B
    if beta1<=CI_U_beta1_LMS_b(j) && beta1>=CI_L_beta1_LMS_b(j)
        diff_LMS(j) = abs(((CI_L_beta1_LMS_b(j)+CI_U_beta1_LMS_b(j))/2)-
beta1);
    end
end
for j=1:B
    if diff_LMS(j)==min(diff_LMS)
```

```
        j_min=j;
    end
end
beta1_CI_check_LMS = j_min;

% generating table in excel
filename = 'table 3.xlsx';
A = {'Beta_1','statistic','OLS','LAD','LMS';...
    'Confidence Interval','Number containing beta1',...
     beta1_CI_check_OLS,beta1_CI_check_LAD,beta1_CI_check_LMS;...
    'Lower Limit of Confidence Interval',...

'minimum',min_CI_L_beta1_OLS_b,min_CI_L_beta1_LAD_b,min_CI_L_beta1_LMS_b;..
.

'','maximum',max_CI_L_beta1_OLS_b,max_CI_L_beta1_LAD_b,max_CI_L_beta1_LMS_b
;...

'','mean',mean_CI_L_beta1_OLS_b,mean_CI_L_beta1_LAD_b,mean_CI_L_beta1_LMS_b
;...

'','median',median_CI_L_beta1_OLS_b,median_CI_L_beta1_LAD_b,median_CI_L_bet
a1_LMS_b;...
    '','standard
deviation',std_CI_L_beta1_OLS_b,std_CI_L_beta1_LAD_b,std_CI_L_beta1_LMS_b;
    'Upper Limit of Confidence Interval',...

'minimum',min_CI_U_beta1_OLS_b,min_CI_U_beta1_LAD_b,min_CI_L_beta1_LMS_b;..
.

'','maximum',max_CI_U_beta1_OLS_b,max_CI_U_beta1_LAD_b,max_CI_U_beta1_LMS_b
;...

'','mean',mean_CI_U_beta1_OLS_b,mean_CI_U_beta1_LAD_b,mean_CI_U_beta1_LMS_b
;...

'','median',median_CI_U_beta1_OLS_b,median_CI_U_beta1_LAD_b,median_CI_U_bet
a1_LMS_b;...
    '','standard
deviation',std_CI_U_beta1_OLS_b,std_CI_U_beta1_LAD_b,std_CI_U_beta1_LMS_b;
    'Estimated Beta_1',...
    'minimum',min_b1_OLS_b,min_b1_LAD_b,min_b1_LMS_b;...
    '','maximum',max_b1_OLS_b,max_b1_LAD_b,max_b1_LMS_b;...
    '','mean',mean_b1_OLS_b,mean_b1_LAD_b,mean_b1_LMS_b;...
    '','median',median_b1_OLS_b,median_b1_LAD_b,median_b1_LMS_b;...
    '','standard deviation',std_b1_OLS_b,std_b1_LAD_b,std_b1_LMS_b;};
sheet = 1;
xlRange = 'A1';
xlswrite(filename,A,sheet,xlRange);
```