





/\* Bootstrap distribution of the sample mean. Example taken from

R. Wicklin (2010), Statistical Programming with SAS/IML Software,

SAS Press: Cary, NC, pp. 350-356.

\*/

%let DSName = Sashelp.Cars;

%let VarName = MPG\_City;

%let alpha = 0.05; /\* significance; (1-alpha)100% conf limits \*/

proc iml;

use &DSName;

read all var {&VarName} into x;

close &DSName;

/\* Resample B times from the data (with replacement)

to form B bootstrap samples. \*/

B = 5000; /\* number of bootstrap samples \*/

call randseed(12345);

xBoot = Sample(x, B||nrow(x)); /\* each column is a resample \*/

/\* Compute the statistic on each bootstrap resample \*/

s = T( mean(xBoot) ); /\* mean of each resample \*/

title "Bootstrap distribution of the mean";

if num(symget("SYSVER"))>=9.4 then do;

call Histogram(s) density="Kernel"; /\* graph bootstrap distrib \*/

end;

Mean = mean(x); /\* sample mean of original data \*/

/\* Analyze the bootstrap distribution \*/

MeanBoot = s[:]; /\* a. mean of bootstrap dist \*/

StdErrBoot = std(s); /\* b. estimate of std error \*/

prob = &alpha/2 || 1-&alpha/2; /\* lower/upper percentiles \*/

call qntl(CIBoot, s, prob); /\* c. quantiles of bootstrap dist\*/

pct = putn(1-&alpha, "PERCENT5.");

print Mean MeanBoot StdErrBoot

(CIBoot`)[c=("Lower "+pct+" CL" || "Upper "+pct+" CL")];

quit;

title;

/\* By the Central Limit Theorem, the sampling distribution of

the mean is approximately normally distributed.

If desired, compare the bootstrap estimates with

estimates of the SEM and CLM that assume normality. \*/

/\*

proc means data=&DSName mean stderr clm alpha=&alpha;

var &VarName;

run;

\*/