

BOOTSTRAP ... PLOT**PURPOSE**

Generates a bootstrap plot for one of 30+ statistics.

DESCRIPTION

The bootstrap is a non-parametric method for calculating a sampling distribution for a statistic. Given a sample data set and a desired statistic (e.g., the mean), the bootstrap works by computing the desired statistic for a subsample of the data set. The subsampling is done with replacement and the sample size is equal to the size of the original data set (since the sampling is done with replacement, there can be duplicates). The **BOOTSTRAP SAMPLE** command (documented under the **SUPPORT COMMANDS** chapter) specifies the number of bootstrap samples taken. To calculate a bootstrap for a statistic not listed below, check the documentation for the **BOOTSTRAP SAMPLE** command (in Volume II under the **LET** subcommands).

For the bootstrap plot, the vertical axis contains the computed value of the statistic and the horizontal axis contains the sample number (for $k = 1, 2, \dots, N$). The number of response variables depends on the number of variables required to compute the statistic (e.g., the **MEAN** uses one while the **LINEAR SLOPE** uses two). The bootstrap plot is typically followed by some type of distributional plot such as a histogram.

SYNTAX 1

BOOTSTRAP <stat> **PLOT** <y> <SUBSET/EXCEPT/FOR qualification>

where <y> is a response variable;

<stat> is one of the following statistics:

MEAN, MIDMEAN, MEDIAN, TRIMMED MEAN, WINDSORIZED MEAN,
 SUM, PRODUCT, SIZE (or NUMBER or COUNT), MINIMUM, MAXIMUM,
 AVERAGE ABSOLUTE DEVIATION (AAD), MEDIAN ABSOLUTE DEVIATION (MAD),
 STANDARD DEVIATION, VARIANCE, STANDARD DEVIATION OF MEAN, VARIANCE OF MEAN,
 RELATIVE STANDARD DEVIATION, RELATIVE VARIANCE,
 RANGE, MIDRANGE, LOWER HINGE, UPPER HINGE, LOWER QUARTILE, UPPER QUARTILE,
 <FIRST/SECOND/THIRD/FOURTH/FIFTH/SIXTH/SEVENTH/EIGHTH/NINTH> DECILE,
 SKEWNESS, KURTOSIS,
 AUTOCORRELATION, AUTOCOVARANCE, SINE FREQUENCY, SINE AMPLITUDE,
 TAGUCHI SN0 (or SN), TAGUCHI SN+ (or SNL), TAGUCHI SN- (or SNS), TAGUCHI SN00 (or SN2);

and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax is used for statistics requiring one response variable.

SYNTAX 2

BOOTSTRAP <stat> <y1> <y2> <SUBSET/EXCEPT/FOR qualification>

where <y1> is the first response variable;

<y2> is the second response variable;

<stat> is one of the following statistics:

LINEAR INTERCEPT, LINEAR SLOPE, LINEAR RESSD, LINEAR CORRELATION;

and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax is used for statistics requiring two response variables to compute.

EXAMPLES

BOOTSTRAP MEAN PLOT Y

BOOTSTRAP LINEAR SLOPE PLOT Y1 X1

NOTE 1

The **SEED** command can be used to specify the seed for generating the random bootstrap samples.

NOTE 2

The jackknife is a similar technique. However, it uses a different resampling scheme. See the **JACKKNIFE PLOT** command for details. The bootstrap is generally considered to provide better results than the jackknife, but it involves more computation.

DEFAULT

None

SYNONYMS

None

RELATED COMMANDS

- LINE = Sets the type for plot lines.
- HISTOGRAM = Generates a histogram.
- BOOTSTRAP SAMPLE = Set the sample size for the bootstrap.
- JACKNIFE PLOT = Generate a jackknife plot.
- PLOT = Generates a data or function plot.

REFERENCE

"A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation," Efron and Gong, The American Statistician, February, 1983.

APPLICATIONS

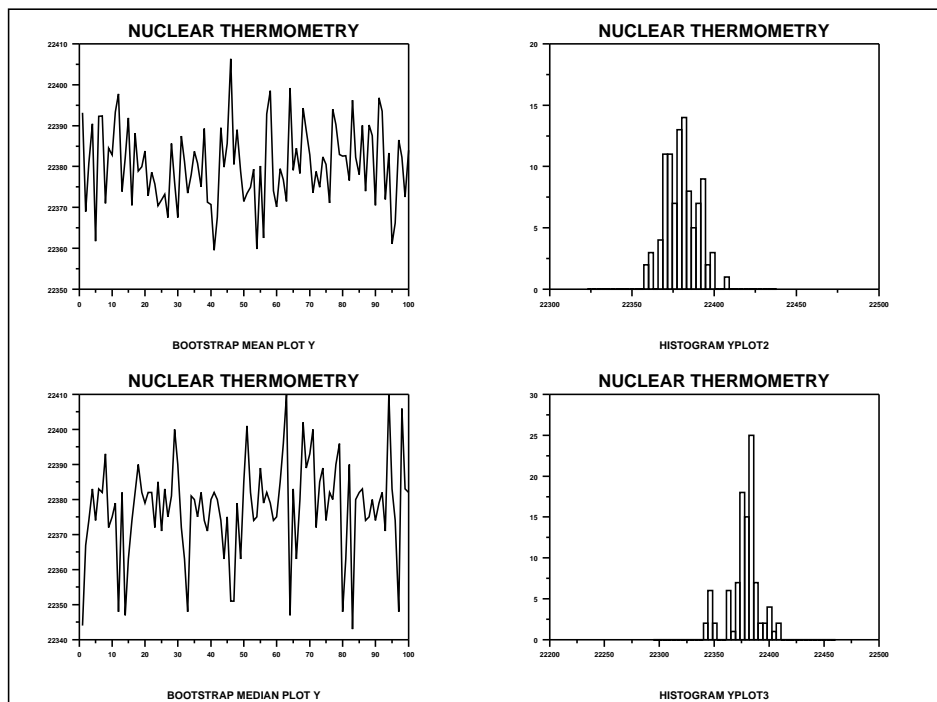
Sample Distribution of a Statistic

IMPLEMENTATION DATE

89/2

PROGRAM 1

```
. PURPOSE--CARRY OUT BOOTSTRAP ANALYSIS OF HARVEY MARSHAK
. NUCLEAR THERMOMETRY DATA. (LOCATION ESTIMATION ANALYSIS)
SKIP 25; READ MARSHAK.DAT Y
TITLE NUCLEAR THERMOMETRY; X3LABEL AUTOMATIC
MULTIPLY 2 2; MULTIPLY CORNER COORDINATES 0 0 100 100
BOOTSTRAP MEAN PLOT Y
LET YPLOT2 = YPLOT
HISTOGRAM YPLOT2
BOOTSTRAP MEDIAN PLOT Y
LET YPLOT3 = YPLOT; HISTOGRAM YPLOT3
END OF MULTIPLY
```



PROGRAM 2

```

.PURPOSE--CARRY OUT BOOTSTRAP ANALYSIS OF BERGER ALASKA PIPELINE DATA
.LINEAR FIT SLOPE ANALYSIS
.
SKIP 25
READ BERGER1.DAT Y X
MULTILOT 2 2
MULTILOT CORNER COORDINATES 0 0 100 100
TITLE ALASKA PIPELINE
CHAR X
LINES
X3LABEL AUTOMATIC
PLOT Y X
LET Y2 = LOG(Y)
LET X2 = LOG(X)
PLOT Y2 X2
BOOTSTRAP LINEAR SLOPE PLOT Y2 X2
LET YPLOT2 = YPLOT
LET S = STANDARD DEVIATION YPLOT2
X2LABEL SD(SLOPE) = 0.022983
HIST YPLOT2
END OF MULILOT
    
```

