BOOTSTRAP INDEX

PURPOSE

This command is used in conjunction with the BOOTSTRAP SAMPLE command to perform a bootstrap analysis for a statistic not directly supported by the BOOTSTRAP PLOT command.

The bootstrap is a non-parametric method for estimating the sampling distribution of 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 size of the sample is equal to the size of the original data set. The pair of commands BOOTSTRAP INDEX and BOOTSTRAP SAMPLE compute one subsample. The LOOP command is typically used to generate a large number of samples (values between 100 and 1000 are typical). The desired statistic is calculated for each subsample. The collection of these statistics is used as an estimate of the sampling distribution. A histogram or some other type of distributional plot is usually performed on the computed values of the statistic.

The BOOTSTRAP INDEX command generates a random sample of the integers from 1 to N (where N is the sample size). The BOOTSTRAP SAMPLE command then uses these numbers as indices into the original data. For example, suppose that Y is an array containing the values 106, 212, 314, and 58 and that the BOOTSTRAP INDEX command returned an array with 2 1 2 4, then the bootstrap sample would be 212, 106, 212, and 58 (that is, Y(2), Y(1), Y(2), Y(4)).

SYNTAX

LET <ind> = BOOTSTRAP INDEX FOR I = 1 1 <par>
where <par> is a positive number or parameter that specifies the size of the variable for which the bootstrap analysis is being performed;
and   <ind> is a variable that contains the computed index numbers.

EXAMPLES

LET IND = BOOTSTRAP INDEX FOR I = 1 1 100
LET IND2 = BOOTSTRAP INDEX FOR I = 1 1 N

NOTE 1

It is significantly faster to perform a bootstrap analysis for one of the 35 or so supported statistics (see the documentation for BOOTSTRAP PLOT to see a list). However, the sample program below demonstrates how to do it for an arbitrary statistic.

NOTE 2

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

NOTE 3

The jacknife is a similar technique. However, it uses a different resampling scheme. See the JACKNIFE PLOT command for details.

DEFAULT

None

SYNONYMS

None

RELATED COMMANDS

BOOTSTRAP PLOT	=	Generate a bootstrap plot.
BOOTSTRAP SAMPLE	=	Generate a bootstrap sample.
LOOP	=	Initiate a loop.
JACKNIFE PLOT	=	Generate a jacknife plot.
JACKINFE INDEX	=	Generate a jacknife index.
SEED	=	Specify the seed for random number generation.

REFERENCE

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

APPLICATIONS

Sample Distribution of a Statistic

IMPLEMENTATION DATE
      89/2

PROGRAM
```
SKIP 25
READ BERGER1.DAT X Y
LET N = NUMBER Y
LET NUMBOOT = 100
LET Y2 = LOG(Y)
LET X2 = LOG(X)
FIT Y2 X2
LET PRED2 = PRED
LET RES2 = RES
LET Y0 = 50
LET Y20 = LOG(50)
FEEDBACK OFF
PRINTING OFF
LOOP FOR K = 1 1 NUMBOOT
    LET IND = BOOTSTRAP INDEX FOR I = 1 1 N
    LET RES3 = BOOTSTRAP SAMPLE RES2 IND
    LET Y3 = PRED2 + RES3
    FIT Y3 X
    LET X0(K) = (Y20-A0)/A1
END OF LOOP
FEEDBACK ON
PRINTING ON
LET SDX0 = STANDARD DEVIATION X0
HISTOGRAM X0
```