CORRELATION

PURPOSE

Compute the correlation coefficient between two variables.

DESCRIPTION

This coefficient is a measure of the linear relationship between two variables. It is computed as:

$$r = \frac{\sum_{i=1}^{N} (x - \bar{x})(y - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x - \bar{x})^2 \sum_{i=1}^{N} (y - \bar{y})^2}} \quad \textbf{(EQ 2-5)}$$

A perfect linear relationship yields a correlation coefficient of +1 (or -1 for a negative relationship) and no linear relationship yields a correlation coefficient of 0. The two variables must have the same number of elements.

SYNTAX

LET <par> = CORRELATION <y1> <y2>   <SUBSET/EXCEPT/FOR qualification>
where <y1> is the first response variable;
       <y2> is the second response variable;
       <par> is a parameter where the computed correlation is stored;
and where the <SUBSET/EXCEPT/FOR qualification> is optional.

EXAMPLES

LET A = CORRELATION Y1 Y2
LET A = CORRELATION Y1 Y2 SUBSET TAG > 2

DEFAULT

None

SYNONYMS

None

RELATED COMMANDS

CORRELATION PLOT	=	Generate an autocorrelation or cross-correlation plot.
COVARIANCE	=	Compute the covariance between two variables.
RANK CORRELATION	=	Compute the rank correlation between two variables.
AUTOCORRELATION	=	Compute the lag 1 autocorrelation of a variable.
LINEAR SLOPE PLOT	=	Generate a linear slope versus subset plot.
FIT	=	Perform a least squares fit (including a linear fit between two variables).

REFERENCE

Consult any introductory statistics text.

APPLICATIONS

Linear Regression

IMPLEMENTATION DATE

Pre-1987

PROGRAM

SKIP 25
READ SNAIL.DAT Y X
LET A = CORRELATION Y X

The computed correlation is 0.9389.