BOX-COX NORMALITY PLOT

PURPOSE
Generates a Box-Cox normality plot.

DESCRIPTION
A Box-Cox normality plot is a graphical data analysis technique for determining the transformation (from the Box-Cox transformation family) that yields a transformed variable that is “closest” to being normally distributed. The Box-Cox transformation family is essentially the power-transformation family (adjusted to include log transformations). The form for the family is:

\[ T(y) = \frac{y^\lambda - 1}{\lambda} \]  (EQ 2-3)

For each of selected members of the Box-Cox family, the transformation is carried out, a normal probability plot is computed, and the linearity of the normal probability plot is summarized via the correlation coefficient. The resulting normality plot thus consists of:

- Vertical axis = normal probability plot correlation coefficient;
- Horizontal axis = Box-Cox lambda parameter.

The value of the lambda parameter (on the horizontal axis) which corresponds to the maximum of the normal probability plot correlation coefficient curve (on the vertical axis) is, of course, of interest—it indicates the best-transformation member of the family. The normality technique is applicable for general transformation families. The current DATAPLOT implementation has the normality plot implemented only for the Box-Cox family (the most important and common of the various transformation families).

SYNTAX
BOX-COX NORMALITY PLOT <y> [<SUBSET/EXCEPT/FOR qualification>]
where <y> is a response variable;
and where the <SUBSET/EXCEPT/FOR qualification> is optional.

EXAMPLES
BOX-COX NORMALITY PLOT Y

DEFAULT
None

SYNONYMS
BOX COX NORMALITY PLOT

RELATED COMMANDS
LINES = Sets the types for plot lines.
CHARACTERS = Sets the types for plot characters.
BOX-COX LINEARITY PLOT = Generates a Box-Cox linearity plot.
BOX-COX HOMOSCED PLOT = Generates a Box-Cox homoscedasticity plot.
LET = Transforms variables (and many other options).
PROBABILITY PLOT = Generates a probability plot.
PPCC PLOT = Generates a probability plot correlation coefficient plot.
PLOT = Generates a data or function plot.

APPLICATION
Exploratory Data Analysis

IMPLEMENTATION DATE
Pre-1987
PROGRAM
    SKIP 25
    READ AUTO83B.DAT Y1

    TITLE AUTOMATIC
    MULTIPLY 2 2; MULTIPLY CORNER COORDINATES 0 0 100 100
    HISTOGRAM Y1
    Y1LABEL CORRELATION COEFFICIENT
    X1LABEL LAMBDA
    BOX-COX NORMALITY PLOT Y1
    X1LABEL; Y1LABEL
    LET TEMP = XPLOT
    LET A TEMP = MAXIMUM YPLOT
    RETAIN TEMP SUBSET YPLOT = A TEMP
    LET LAMBDA = TEMP(1)
    LET YNEW = (Y1**LAMBDA - 1)/LAMBDA
    HISTOGRAM YNEW
    END OF MULTIPLY

HISTOGRAM Y1
    HISTOGRAM YNEW
    BOX-COX NORMALITY PLOT Y1