... PROBABILITY PLOT

PURPOSE

Generates a probability plot for one of 38 distributions.

DESCRIPTION

A probability plot is a graphical data analysis technique for determining how well the specified distribution fits the data set. Linearity in the probability plot is indicative of a good distributional fit. The probability plot consists of:

Vertical axis = ordered observations;
Horizontal axis = order statistic medians.

DATAPLOT has extensive probability plot capabilities--38 distributions/distributional families are available. When distributional families are specified, then the LET command is used before the PROBABILITY PLOT command to specify fully which member of the distributional family is desired. For example,

```
LET GAMMA = 5.3
WEIBULL PROBABILITY PLOT Y
```

The name of the distributional parameter for families is given in the list below.

SYNTAX 1

<dist> PROBABILITY PLOT <x> <SUBSET/EXCEPT/FOR/qualification>
where <x> is the variable of raw data values under analysis;
<dist> is one of the following distributions:

UNIFORM	
SEMI-CIRCULAR	
TRIANGULAR	(C, defaults to 0)
NORMAL	
LOGISTIC	
DOUBLE EXPONENTIAL	
CAUCHY	
TUKEY LAMBDA	(LAMBDA)
LOGNORMAL	
HALFNORMAL	
T	(NU)
CHI-SQUARED	(NU)
F	(NU1, NU2)
EXPONENTIAL	
GAMMA	(GAMMA)
BETA	(ALPHA, BETA)
WEIBULL	(GAMMA)
EXTREME VALUE TYPE 1	
EXTREME VALUE TYPE 2	(GAMMA)
PARETO	(GAMMA)
BINOMIAL	(N, P)
GEOMETRIC	(P)
POISSON	(LAMBDA)
NEGATIVE BINOMIAL	(N, K, P)
WALD	(GAMMA)
INVERSE GAUSSIAN	(GAMMA)
RIG	(GAMMA)
FL	(GAMMA)
GENERALIZED PARETO	(GAMMA)
DISCRETE UNIFORM	(N)
NON-CENTRAL T	(NU, LAMBDA)
NON-CENTRAL F	(NU1, NU2, LAMBDA)
NON-CENTRAL CHI-SQUARE	(NU, LAMBDA)
NON-CENTRAL BETA	(ALPHA, BETA, LAMBDA)

DOUBLY NON-CENTRAL F	(NU1, NU2, LAMBDA1, LAMBDA2)
DOUBLY NON-CENTRAL T	(NU, LAMBDA1, LAMBDA2)
HYPERGEOMETRIC	(K, N, M)
VON MISES	(B)

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

This syntax is used for raw data.

SYNTAX 2

<dist> PROBABILITY PLOT <y> <x> <SUBSET/EXCEPT/FOR/qualification>
where <y> is the variable of pre-computed frequencies;
<x> is the variable of distinct values for the variable under analysis;
<dist> is as above;
and where the <SUBSET/EXCEPT/FOR qualification> is optional.

This syntax is used for pre-computed frequencies.

EXAMPLES

NORMAL PROBABILITY PLOT X
CAUCHY PROBABILITY PLOT X
TUKEY LAMBDA PROBABILITY PLOT X
LOGNORMAL PROBABILITY PLOT X
WEIBULL PROBABILITY PLOT X
EXTREME VALUE TYPE 1 PROBABILITY PLOT X
POISSON PROBABILITY PLOT X
NORMAL PROBABILITY PLOT F X
CAUCHY PROBABILITY PLOT F X
TUKEY LAMBDA PROBABILITY PLOT F X
LOGNORMAL PROBABILITY PLOT F X
WEIBULL PROBABILITY PLOT F X
EXTREME VALUE TYPE 1 PROBABILITY PLOT F X
POISSON PROBABILITY PLOT F X

NOTE 1

For distributions that have a family of parameters, the PPCC PLOT can be used to find the optimal value of the parameter to use for generating the probability plot.

NOTE 2

The PROBABILITY PLOT command fits a least squares line to the resulting probability plot and automatically saves the following internal parameters:

PPCC	=	the correlation coefficient between the vertical and horizontal axis variables
PPA0	=	the intercept of the fitted line
PPA1	=	the slope of the fitted line
SDPPA0	=	standard deviation of PPA0
SDPPA1	=	standard deviation of PPA1
PPRESSD	=	residual standard deviation from fitted line
PPRESDF	=	residual degrees of freedom from fitted line

These parameters can be printed or used in subsequent computations if desired.

NOTE 3

The Weibull, extreme value type II, and generalized Pareto distributions can be based on either the minimum or maximum order statistic. The command SET MINMAX <1/2> is required before the PROBABILITY PLOT command for these distributions. A value of 1 specifies the minimum order statistic and a value of 2 specifies the maximum order statistic. Currently, the generalized Pareto distribution is only supported for the maximum order statistic (i.e., enter SET MINMAX 2).

DEFAULT

None

SYNONYMS

EV2 and FRECHET are synonyms for EXTREME VALUE TYPE 2.

EV1 and GUMBEL are synonyms for EXTREME VALUE TYPE 1.

FATIGUE LIFE is a synonym for FL.

RECIPROCAL INVERSE GAUSSIAN is a synonym for RIG.

IG is a synonym for INVERSE GAUSSIAN.

LAPLACE is a synonym for DOUBLE EXPONENTIAL.

RELATED COMMANDS

FREQUENCY PLOT	=	Generates a frequency plot.
HISTOGRAM	=	Generates a histogram.
PIE CHART	=	Generates a pie chart.
PERCENT POINT PLOT	=	Generates a percent point plot.
PPCC PLOT	=	Generates probability plot correlation coefficient plot.
PLOT	=	Generates a data or function plot.

APPLICATIONS

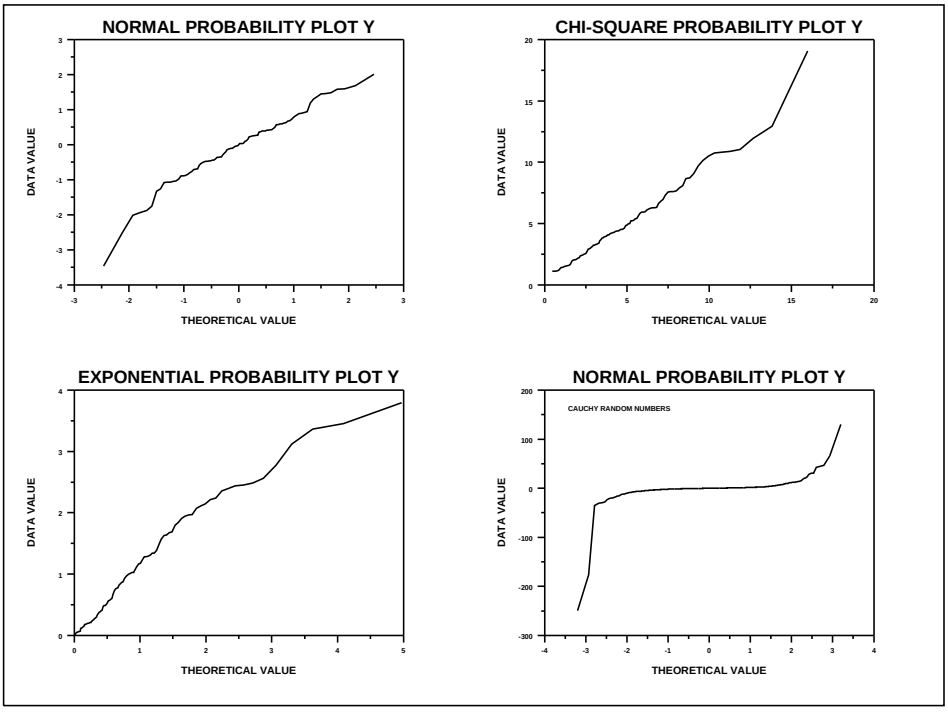
Distributional Analysis

IMPLEMENTATION DATE

Pre-1987 (the saving of the various internal parameters was implemented 93/12, many distributions were added after 1987)

PROGRAM 1

```
MULTIPLOT 2 2; MULTIPLOT CORNER COORDINATES 0 0 100 100
TITLE AUTOMATIC; X1LABEL THEORETICAL VALUE; Y1LABEL DATA VALUE
.
LET Y = NORMAL RANDOM NUMBERS FOR I = 1 1 100
NORMAL PROBABILITY PLOT Y
.
LET NU = 5
LET Y = CHI-SQUARE RANDOM NUMBERS FOR I = 1 1 100
CHI-SQUARE PROBABILITY PLOT Y
.
LET Y = EXPONENTIAL RANDOM NUMBERS FOR I = 1 1 100
EXPONENTIAL PROBABILITY PLOT Y
.
LET Y = CAUCHY RANDOM NUMBERS FOR I = 1 1 1000
LEGEND 1 CAUCHY RANDOM NUMBERS
NORMAL PROBABILITY PLOT Y
END OF MULTIPLOT
```



PROGRAM 2

```
. ALASKA PIPELINE RADIOGRAPHIC DEFECT BIAS CURVE
. PERFORM A LINEAR REGRESSION
SKIP 25
READ BERGER1.DAT TRUE MEAS
CAPTURE FIT_1_OUT.DAT
FIT MEAS TRUE
END OF CAPTURE
.
MULTIPLOT 2 2 ; MULTIPLOT CORNER COORDINATES 0 0 100 100
TITLE ORIGINAL DATA
X1LABEL TRUE DEPTH (IN .001 INCH)
Y1LABEL MEASURED DEPTH
CHARACTERS X
LINES BLANK
PLOT MEAS TRUE
TITLE PREDICTED VALUES
PLOT MEAS PRED VS TRUE
TITLE RESIDUALS
Y1LABEL
PLOT RES VS TRUE
X1LABEL
TITLE NORMAL PROBABILITY PLOT
NORMAL PROBABILITY PLOT RES
END OF MULTIPLOT
```