HYPPDF

PURPOSE

Compute the hypergeometric probability density function.

DESCRIPTION

The hypergeometric distribution is the probability of selecting LL marked items when a random sample of size KK is taken without replacement from a population a population of MM items, NN of which are marked. Marked and unmarked items can also be thought of as successes and failures. It has probability function:

$$p(x) = \frac{\binom{NN}{x}\binom{MM - NN}{KK - x}}{\binom{MM}{KK}} \qquad \textbf{(EQ 8-243)}$$

where (a b) is the combinatorial function a things taken b at a time and has the formula:

$$\binom{a}{b} = \frac{a!}{b!(a-b)!} \qquad \textbf{(EQ 8-244)}$$

The ! symbol is the factorial function.

SYNTAX

LET <y> = HYPPDF(<x>,<kk>,<nn>,<mm>) <SUBSET/EXCEPT/FOR qualification>

where <x> is an integer variable, number, or parameter;

<kk> is a positive integer number, parameter, or variable that is the size of the sample;

<nn> is a positive integer number, parameter, or variable that is number of marked items;

<mm> is a positive integer number, parameter, or variable that is number of items in the population;

<y> is a variable or a parameter (depending on what <x> is) where the computed hypergeometric cdf value is stored;

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

EXAMPLES

LET A = HYPPDF(7,150,400,1200)

LET Y = HYPPDF(X1,150,400,1200)

NOTE

DATAPLOT uses the algorithm AS R77 (which is a refinement of AS 152) from the Applied Statistics series (see the REFERENCE section below). DATAPLOT uses the version of this algorithm from the statlib archive. For small to moderate values of the input parameters, an exact algorithm is used (essentially by taking the log of the factorial function). For large values of the input parameters, a normal approximation is used. One change that DATAPLOT makes in this algorithm is that a binomial approximation rather than a normal approximation is used. The binomial approximation is taken from the Ling and Pratt article (see the REFERENCE section below).

DEFAULT

None

SYNONYMS

None

RELATED COMMANDS

HYPCDF	=	Compute the hypergeometric cumulative distribution function.
HYPPPF	=	Compute the hypergeometric percent point function.
BINCDF	=	Compute the binomial cumulative distribution function.
BINPDF	=	Compute the binomial probability density function.
BINPPF	=	Compute the binomial percent point function.
POIPDF	=	Compute the Poisson probability density function.
POICDF	=	Compute the Poisson cumulative distribution function.
POIPPF	=	Compute the Poisson percent point function.
NBCDF	=	Compute the negative binomial cumulative distribution function.

NBPDF	=	Compute the negative binomial probability density function.
NBPPF	=	Compute the negative binomial percent point function.
GEOCDF	=	Compute the geometric cumulative distribution function.
GEOPDF	=	Compute the geometric probability density function.
GEOPPF	=	Compute the geometric percent point function.

REFERENCE

"A Remark on Algorithm AS 152: Cumulative Hypergeometric Probabilities," Shea, Applied Statistics Journal, Vol. 38, No. 1, 1989.

"Algorithm AS 152: Cumulative Hypergeometric Probabilities," Lund, Applied Statistics, Vol. 31 (pp. 221-223).

"The Accuracy of Piezer Approximations to the Hypergeometric Distribution, with Comparisons to Some Other Approximations," Ling and Pratt, Journal of the American Statistical Association, March, 1984.

"Statistical Distributions," 2nd Edition, Evans, Hastings, and Peacock, Wiley and Sons, 1993, (chapter 20).

"Discrete Univariate Distributions," Johnson and Kotz, Houghton-Mifflin, 1969 (chapter 6).

APPLICATIONS

Data Analysis

IMPLEMENTATION DATE

94/9

PROGRAM

```
XTIC OFFSET 0.5 0.5
LINE BLANK
SPIKE ON
SPIKE THICKNESS 0.3
TITLE AUTOMATIC
Y1LABEL PROBABILITY
X1LABEL X
PLOT HYPPDF(X,95,400,1200) FOR X = 0 1 95
```