CDF (BIVNORM or CHISQ or DICKEYF or F or NORMAL or T or WTDCHI, 
DF=degrees of freedom for CHISQ or T, DF1=numerator degrees of freedom for F, 
DF2=denominator degrees of freedom for F, 
NLAGS=number of lags for augmented Dickey-Fuller test, 
NOB=number of observations for unit root or cointegration test, 
NVAR=number of variables for cointegration test, 
RHO=correlation coefficient for BIVNORM, 
EIGVAL=vector of eigenvalues for WTDCHI, 
LOWTAIL or UPTAIL or TWOTAIL, 
CONSTANT, TREND, TSQ, INVERSE, PRINT) 
test statistic  [significance level];  
or  
significance level  [critical value];  (for INVERSE)  
or  
x value  y value  [significance level];  (for BIVNORM) 

Function:

CDF calculates and prints tail probabilities (p-value or significance level) or critical values for several cumulative distribution functions. This is useful for hypothesis testing.

Usage:

CDF followed by the value of a scalar test statistic is the simplest form of the command. In this case, a two-tailed probability for the normal distribution will be calculated and printed. If the INVERSE option is used, the first argument must be a probability level; a critical value will be calculated. Arguments need not be scalars; they can be series or matrices. Other distributions or tail areas may be selected through the options. For hypothesis testing with regression diagnostics, see the REGOPT(PVPRINT) command.

Options:

BIVNORM/CHISQ/DICKEYF/F/NORMAL/T/WTDCHI specifies the bivariate normal, chi-squared, Dickey-Fuller, F, standard normal, student’s t, and weighted chi-squared distributions, respectively.

DF= the degrees of freedom for the chi-squared or student’s t distribution, or the number of observations for Dickey-Fuller (the NOB= option is a better way of specifying this).

DF1= the numerator degrees of freedom for the F distribution.

DF2= the denominator degrees of freedom for the F distribution.

RHO= the correlation coefficient for the bivariate normal distribution.

EIGVAL= vector of eigenvalues for the weighted chi-squared distribution.

NLAGS= the number of lagged differences in the augmented Dickey-Fuller test. This number is used to compute the approximate finite sample P-value or critical value. The default is zero (assume an unaugmented test). The NOB=
CDF

option must also be specified for the finite sample value.

**NOB=** the number of observations for the augmented Dickey-Fuller or Engle-Granger tests. This number is used to compute the approximate finite sample P-value or critical value. The default is zero (to compute asymptotic instead of finite sample value).

**NVAR=** the number of variables for an Engle-Granger/Dickey-Fuller cointegration test. The default is 1 (plain unit root test), and the maximum is 6.

**CONSTANT/NOCONST** specifies whether a constant term (C) was included in the regression for Dickey-Fuller. NOCONST is only valid for NVAR=1.

**TREND/NOTREND** specifies whether a trend term (1,2,...,T) was included in the regression for Dickey-Fuller.

**TSQ/NOTSQ** specifies whether a squared trend term (1,4,9,...) was included in the regression for Dickey-Fuller.

**LOWTAIL/TWOTAIL/UPTAIL** specifies the area of integration for the density function. TWOTAIL is the default for most symmetric distributions (normal and t), UPTAIL is the default for chi-squared and F, and LOWTAIL is the default for bivariate normal and Dickey-Fuller. TWOTAIL is not defined for bivariate normal.

**INVERSE/NOINVERSE** specifies the inverse distribution function (input is significance level, output is critical value). Normally the input is a test statistic and the output is a significance level. INVERSE is not defined for bivariate normal.

**PRINT/NOPRINT** turns on printing of results. PRINT is true by default if there is no output specified.

**Output:**

If the PRINT option is on, the input and output values will be printed, along with the degrees of freedom. If a second argument is supplied, it will be filled with the output values and stored (see examples 4 and 6). Input and output arguments may be any numeric TSP variables.

**Examples:**

1. To compute the significance level of a Hausman test statistic with 5 degrees of freedom:

   `CDF(CHISQ,DF=5) HAUS;`
   produces the output--  CHISQ(5) Test Statistic:  7.235999 , Upper tail area: .20367

2. Significance level of the test for AR(1) with lagged dependent variable(s):

   `CDF @DHALT;`  or  `REGOPT(PVPRINT) DHALT;` before the regression

3. Two-tailed critical value for the normal distribution:

   `CDF(INV) .05;`
   produces the output--  NORMAL Critical value:  1.959964 , Two-tailed area: .05000

4. Several critical values for the normal distribution:

   `READ PX; .1 .05 .01 ;`
CDF(INV,NORM) PX CRIT;
PRINT PX,CRT;

5. F critical values:

CDF(INV,F,DF1=3,DF2=10) .05;
produces the output-- F(3,10) Critical value: 3.708265 , Upper tail area: .05000

6. Bivariate normal:

CDF(BIVNORM,RHO=.5,PRINT) -1 -2 PBIV;
produces the output-- BIVNORM Test Statistic: -1.0000 , -2.0000 , Lower tail area: .01327

7. Dickey-Fuller unit root test (same as UNIT(MAXLAG=0,NOWS) Y;) :

TREND TIME;
SMPL 2,50;
DY = Y-Y(-1);
OLSQ DY TIME C Y(-1);
CDF(DICKEYF) @T(3);

8. Augmented Dickey-Fuller unit root test, with finite sample P-value
(same as UNIT(MAXLAG=3,NOWS,FINITE) Y;) :

TREND TIME;
SMPL 2,50;
DY = Y-Y(-1);
SMPL 5,50;
OLSQ DY TIME C Y(-1) DY(-1)-DY(-3);
CDF(DICKEYF,NOB=NOB,NLAGS=3) @T(3);

9. Engle-Granger cointegration test (same as COINT(NOUNIT,MAXLAG=0,ALLORD) Y1-Y4; ) :

TREND TIME;
OLSQ Y1 TIME C Y2 Y3 Y4; EGTEST;
OLSQ Y2 TIME C Y1 Y3 Y4; EGTEST;
OLSQ Y3 TIME C Y1 Y2 Y4; EGTEST;
OLSQ Y4 TIME C Y1 Y2 Y3; EGTEST;
PROC EGTEST;
SMPL 2,50;
DU = @RES-@RES(-1);
OLSQ DU @RES(-1);
CDF(DICKEYF,NVAR=4) @T;
SMPL 1,50;
ENDPROC;
10. Verify critical values for Durbin-Watson statistic, for regression with 10 observations and 2 RHS variables:

CDF

\[
\text{SMPL 1,10; OLSQ Y C X1;}
\]
\[
\text{SET PI = 4*ATAN(1); SET F = PI/(2*@NOB); TREND I;}
\]
\[
\text{EIGB = 4*SIN(I*F)**2;}
\]
\[
\text{SELECT I <= (@NOB-@NCOEF); ? use largest lest eigenvalues for dL}
\]
\[
\text{CDF(WTDCHI,EIG=EIGB) .879; ? dL for 5% (n=10, k'=1)}
\]
\[
\text{SELECT (@NCOEF <= I) & (I <= (@NOB-1)); ? use smallest eigenvalues for dU}
\]
\[
\text{MMAKE dU 1.320 1.165 1.001; ? dU for 5%, 2.5%, 1% (n=10, k'=1)}
\]
\[
\text{CDF(WTDCHI,EIG=EIGB,PRINT) dU;}
\]

11. Reproduce exact P-value for Durbin-Watson statistic (can already be done by REGOPT):

\[
\text{SMPL 1,10; ? data from Judge, et al (1988) example: DW = 1.037, P-value = .0286}
\]
\[
\text{READ Y X1; 4 2 7 4 7.5 6 4 3 2 5 3 4.5 4 7.5 8 5 6 ;}
\]
\[
\text{REGOPT(DWPVAL=EXACT);}
\]
\[
\text{OLSQ Y C X1;}
\]
\[
\text{MMAKE X @RNMS;}
\]
\[
\text{MAT XPXI = (X'X)^*;}
\]
\[
\text{TREND OBS; SELECT OBS > 1;}
\]
\[
\text{DC = 0; DX1 = X1 - X1(-1);}
\]
\[
\text{MMAKE DX DC DX1;}
\]
\[
\text{MMAKE BVEC 2 -1; MFORM(BAND,NROW=@NOB) DDP = BVEC;}
\]
\[
\text{MAT DMDP = DDP - DX*XPXI*DX'}
\]
\[
\text{? Eigenvalues of DMD' = D*D' - DX*(X'X)^*(DX)'}
\]
\[
\text{? (same as nonzero eigenvalues of MA, because A = D'D)}
\]
\[
\text{MAT ED = EIGVAL(DMDP);}
\]
\[
\text{CDF(WTDCHI,EIG=ED) @DW;}
\]

Method:

BIVNORM: ACM Algorithm 462. Inverse is not supplied because it is not unique unless x or y is known, etc.

CHISQ: DCDFLIB method: Abramovitz-Stegun formula 26.4.19 converts it to Incomplete Gamma, and then use DiDinato and Morris(1986). Inverse by iteration (trying values of x to yield p -- faster methods are also known).

F and T: DCDFLIB method: Abramovitz-Stegun formula 26.6.2 converts it to Incomplete Beta, and then use DiDinato and Morris (1993), i.e. ACM Algorithm 708. Inverse by iteration.


DICKEYF: Asymptotic values from Tables 3 and 4 in MacKinnon (1994). Finite sample critical values from Cheung and Lai (1995) [augmented Dickey-Fuller] or MacKinnon (1991) [Engle-Granger]. To convert these to finite sample P-values, a logistic interpolation is used with the .05 size and either the .01 or .10 size (whichever is closer to the observed test statistic). Such interpolated P-values are fine for testing at the .01, .05, or .10 sizes, but would be highly speculative outside this range.

WTDCHI: If \( w_i \) are the eigenvalues, \( c_i \) are chi-squared(1) variables, and \( d \) is the test statistic, WTDCHI computes the following probability: \( \text{Pr}[\sum(w_i \cdot c_i) / \sum(c_i) < d] = \text{Pr}[\sum((w_i - d) \cdot c_i) < 0] \). This is useful for computing P-values for the Durbin-Watson statistic, other ratios of quadratic forms in normal variables, and certain non-nested tests (for example, Vuong(1989)). The Pan method is used when the number of eigenvalues is less than 90; otherwise the Imhoff method is used. If the absolute values of the smallest eigenvalues are less than 1D-12, they are not used; otherwise duplicate eigenvalues are not checked for. The inverse of this distribution is not implemented.
References:


