Selected Continuous Probability Distributions

Lloyd Milligan

Ralph H. Johnson VA Medical Center
Charleston, SC 29401-5799

Abstract. Statistical analysis generally entails two steps. First, a statistic is computed from data. Second, the probability of the statistic is determined, based on its known distribution. Computer or calculator programs often omit the probability computation, and the user must look up the value in a table. This paper exhibits MUMPS code for three of the most commonly used continuous distributions: χ^2, Student's t, and the F distribution. The methods described may be used to augment existing MUMPS statistics programs, or they may be generalized to other probability distributions of interest.

The Gamma Distribution

The gamma probability distribution arises naturally from the Poisson, a discrete distribution. Suppose, for example, that the number of calls arriving at a telephone exchange during a time interval t has the Poisson distribution:

$$\text{(E1)} \qquad g(k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \qquad (k = 0, 1, 2 \ldots).$$

We can invert the focus of interest from "number of calls" to "time," and ask what is the probability of waiting less than t minutes for n calls. This is equivalent to the probability of n or more calls arriving in t minutes. The *distribution* function for waiting time to the nth call is obtained by summing $g(k)$ from $k = n$ to ∞:

$$\text{(E2)} \qquad F(x) = \sum_{k=n}^{\infty} \frac{(\lambda x)^k}{k!} e^{-\lambda x} \qquad (x \geq 0).$$

Differentiating (E2) obtains the *density* function of the gamma distribution with parameters $(n-1)$ and $1/\lambda$:

$$\text{(E3)} \qquad f(x) = \frac{\lambda^n x^{n-1} e^{-\lambda x}}{\Gamma(n)} \qquad (x \geq 0). \qquad [1]$$

In (E3), $\Gamma(n)$ is the gamma (factorial) function:

$$\text{(E4)} \qquad \Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt \qquad (\mathcal{R}z > 0) \qquad [2]$$

Note that $\Gamma(n+1) = n!$ for non-negative integers n.

The gamma function appears in several continuous distributions of interest including χ^2, Student's t, and the F distribution. Therefore, it is useful to have a MUMPS extrinsic function that computes $\Gamma(x)$, for real $x > 0$. Because $\Gamma(x)$ grows rapidly, values in tables typically extend only to about $\Gamma(100)$. MUMPS numeric precision limits practical computation to small arguments. In fact, $\Gamma(16)$ exceeds 12-digits, the maximum length guaranteed by the language standard. However, the actual precision provided by implementations is usually greater than the standard requires.

An extrinsic function for $\Gamma(x)$ is included in Listing 1. [3] Briefly the method employs approximating polynomials for $x \leq 2$, and the recurrence relation $\Gamma(x+1) = x\Gamma(x)$ for $x > 2$.

Integration versus Approximation

In a previous paper we exhibited a MUMPS extrinsic function for the trapezoidal rule, which yields satisfactory accuracy when computing the definite integral of selected probability density functions. [4] This approach can be used when approximating polynomials are not readily available. However, for some functions and domain values many iterations are needed to obtain a reasonable degree of accuracy, resulting in a significant performance penalty. In general, non-iterative approximations are preferred, especially when computing multiple probabilities.

Chi-Square

Sums of squares of ν independent random variables, each having the normal distribution with mean zero and unit variance have the χ^2 (chi-square) distribution with ν degrees of freedom. This paper omits the density and distribution functions for χ^2, which can be found in numerous sources.

A series expansion is used to compute the probability of χ^2 (see Listing 1). [5] Equation (E5) displays the series:

$$\text{(E5)} \quad P(\chi^2|\nu) = (\frac{1}{2}\chi^2)^{\nu/2} \frac{e^{-\chi^2/2}}{\Gamma(\frac{\nu+2}{2})} \cdot \left(1 + \sum_{r=1}^{\infty} \frac{\chi^{2r}}{(\nu+2)(\nu+4)\cdots(\nu+2r)}\right) \qquad [2]$$

For large degrees of freedom an approximation to the normal distribution is used:
(E6)

$$Q(\chi^2|\nu) \approx Q(x_2), \qquad x_2 = \frac{(\chi^2/\nu)^{1/3} - \left(1 - \frac{2}{9\nu}\right)}{\sqrt{2/9\nu}} \qquad (\nu > 30) \qquad [2]$$

The approximation given in (E6) is accurate to about three decimal places, and is employed in the MUMPS implementation of $P(\chi^2|\nu)$ for the case $\nu \geq 60$. [5]

Chi-Square Computational Example. The χ^2 statistic is commonly used with frequency data. In this context χ^2 is computed as follows:

(E7)
$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

where O stands for observed, E for expected frequencies, and n for the number of classes. The MUMPS routine `^AGGSC`, included with the sources accompanying this paper, computes χ^2 for one-way and multiway frequency tables.

Suppose an experiment is conducted using the following MUMPS code:

```
f i=1:1:100 s %=$r(5),x(%)=$s($d(x(%)):1+x(%),1:1)
```

Also, suppose this experiment has the following outcome: x_0=19, x_1=19, x_2=18, X_3=23 and x_4=21. Are the results compatible with the hypothesis that $random produces a uniform distribution? Using equation (E7) directly or routine `^AGGSC` to compute χ^2 obtains the value $\chi^2 = 0.8$ with $\nu = 4$ degrees of freedom. Invoking `$$Q^ZZNEWPR(.8,4)` obtains the probability .93845. (Note: that with 4 degrees of freedom, a χ^2 value of at least 9.5 is needed to reach the .05 level of significance.) [6]

Student's t

The ratio $X/\sqrt{\chi^2/\nu}$, where X is a unit normal random variable and χ^2 is an independent random variable distributed as chi-square with ν degrees of freedom, has the distribution known as Student's t. [1] [2] (E8) gives the density function:

(E8)
$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2} \qquad (-\infty < x < \infty) \qquad [1]$$

This paper omits the related distribution function (given in note [2]), which is not used directly in the MUMPS code for Student's t. However, the following approximation for large ν is used:

$$\text{(E9)} \qquad A(t|\nu) \approx 2P(x) - 1, \qquad x = \frac{t\left(1 - \frac{1}{4\nu}\right)}{\sqrt{1 + \frac{t^2}{2\nu}}} \qquad [2]$$

A MUMPS extrinsic function to compute $P(t|\nu)$ accompanies this paper. This function integrates (E8) for $\nu \leq 15$ and uses the approximation (E9) for larger degrees of freedom. [7]

Student's t computational example. Issues such as the suitability of statistical tests to particular purposes, the assumptions underlying tests, and so forth, are not considered in this paper. In general the t statistic is used in two ways: 1) to test the significance of the difference between an obtained mean and a theoretical mean, and 2) to test the difference between two sample means. The t-test routine `^AGGST` in the sources accompanying this paper supports both of these applications (line tags `T1` and `T2` respectively). Another variant of t is used with correlated samples. Routine `^AGGSR`, also included with the sources, computes t-test for related measures.

To illustrate computing the probability of t, an example comparing a sample mean to a population mean is used. (E10) provides an estimate of t for this case, with $n - 1$ degrees of freedom:

$$\text{(E10)} \qquad t = \frac{\bar{X} - m}{s_{\bar{x}}} \quad where \quad s_{\bar{x}} = \frac{s}{\sqrt{n}} \qquad [8]$$

Suppose 10 sample observations have been obtained and we wish to know whether the sample mean differs significantly from a population mean of 3 ($\alpha < .05$). The sample values are as follows: 2, 4, 4, 3, 5, 5, 4, 3, 5, and 5. Using `T1^AGGST` we enter the sample data and obtain the value $t = 3.0$ with 9 degrees of freedom. To compute the probability of t, it is necessary to invoke `$$CumT^ZZNEWPR(3,9)` which returns $p = .493$. Because no hypothesis has been expressed about whether the sample mean will be greater or less than the population mean, a two-tailed test is appropriate. Therefore $Q(3|9) = 1-(2 \times .493) = .014$, which is clearly less than the chosen α-level.

The F Distribution

If X_1^2 and X_2^2 are independent chi-square distributed random variables with ν_1 and ν_2 degrees of freedom, then the ratio of X_1^2/ν_1 and X_2^2/ν_2

has the F distribution with ν_1 and ν_2 degrees of freedom. F is sometimes referred to as the variance ratio, and is widely used in statistical analyses where effects are represented by mean squares. The density function of F is as follows:

(E11)

$$f(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(\frac{m}{n}\right)^{m/2} x^{(m/2)-1}\left(1+\frac{m}{n}x\right)^{-(m+n)/2} \quad (0 \le x < \infty) \qquad [1]$$

The distribution function is omitted but can be found in note [2]. To compute $P(F|\nu_1, \nu_2)$ using MUMPS, several cases are distinguished and separate methods are employed, based on the degrees of freedom. [9]

Computational Example for F.

A 3-factor factorial design example from Bruning and Kintz [10] is included in the `^AGGSI*` initialization routines accompanying this paper. The example (worked in note [10], pages 31-38) can be computed using the routine `^AGGSFA`. Working this example leads to an analysis-of-variance summary table with sums of squares, degrees of freedom, and mean squares for each source of variance (see note [10], page 38, or the output of `^AGGSFA`). Compute the probability of $F = 121.1$ associated with the "rate" variable. Invoking `$$QF^ZZNEWPR(121.1,1,40)` obtains the result 0.00000. This does not mean that the probability is zero, but that Q is less than .00001, the smallest value returned by the `$$QF()` extrinsic function. Arguing $F = 14.32$ for the "list∗rate" effect obtains the result 0.00050. Both of these computations belong to the case where ν_2 is even, and therefore are computed by expanding a series.

The Routine ^ZZNEWPR

The functions discussed in this paper are coded in the MUMPS routine `^ZZNEWPR`. Several simple mathematical utilities are included: minimum value, absolute value and harmonic mean. Additionally, two methods of numeric integration are included and are discussed in note [4]. Briefly, the conventions used in routine `^ZZNEWPR` are as follows:

- Mathematical constants are given as extrinsic functions without parameters: `$$e()`, `$$PI()`, and so forth.
- Line tags of the form `Pn` where `n` = 1,2,3 evaluate products (powers), and tags of the form `Sn` return sums (of the corresponding products).
- The `$$POWER()` function is implementation specific. Users of MUMPS versions that do not implement exponentiation, or a

power function `$ZCALL`, will need to substitute a MUMPS power function such as the math library prototype (MDC: X11/SC13/-91-...).

- Tags corresponding to probability density functions are named for the function, and the distribution function is named "Cum" followed by an abbreviation of the function name.
- Numeric integration of the probability density function is not usually practical. Therefore other methods are used for the probability of primary interest, $Q = 1 - P$. These extrinsic functions typically begin with "Q" and approximations begin with "A."
- Polynomials are evaluated using the "reverse Polish" method, that is without resorting to the `$$POWER()` function.

The methods used in `^ZZNEWPR` can be significantly improved. These extrinsic functions are intended to illustrate *functionality* of estimating the probability of a statistic. Better algorithms and approximations can be found in the literature and coded in MUMPS as the need justifies.

Notes and References

[1] James R. McCord III and Richard M. Moroney, Jr. *Introduction to Probability Theory.* The MacMillan Co. New York, 1965. The telephone exchange example and derivation of the gamma distribution are taken directly from this source.

[2] *Handbook of Mathematical Functions.* Milton Abramowitz and Irene A. Stegun, Eds., Dover Publications, Inc., New York, 1972.

[3] The extrinsic function for $\Gamma(x)$ is found at tag `Gamma^ZZNEWPR`. Note that tags in MUMPS are case sensitive. To test the Gamma function write `$$Gamma^ZZNEWPR(3.5)`. The correct answer is 3.32338.

[4] Lloyd Milligan. Elementary Probability Functions. *MUMPS Computing*, Vol 22, No 4, Sep 1992 pp 54-59. This article exhibits an extrinsic function for the trapezoidal rule and then applies this function to compute the Normal probability. A minor error appears in the code for the trapezoidal rule. Line `TRAPZ+5` should begin: `S @VAR=UB,@("SUM="` .... The error has negligible effect. However, a more efficient method of computing Normal probabilities is given with the sources accompanying the current paper at tag `CumNorm^ZZNEWPR`, which uses a 6-term polynomial from note [2] with $|\epsilon(x)| < 1.5 \cdot 10^{-7}$.

Sources accompanying the current paper also include Simpson's rule:

$$\int_a^b f(x)dx \approx \frac{h}{3}(y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \cdots + 4y_{n-1} + y_n)$$

$$where \quad h = (b-a)/n.$$

Simpson's rule sums segments of parabolas instead of trapezoids to approximate the area. The parameter list for Simpson's rule is identical to that for the trapezoidal rule. Invoke `$$SIMPSONS^ZZNEWPR(EXPR,VAR,LB,UB,N)`, where EXPR contains a MUMPS expression for the function to be integrated, VAR contains the name of the variable of integration, LB and UB are the lower and upper bounds respectively, and N is the number of iterations.

[5] Two accessory functions are used in addition to the `$$Gamma()` extrinsic function in computing (E5). First the fraction $\chi^{2r}/(\nu+2)(\nu+4)\cdots(\nu+2r)$, essentially a "power" function, is evaluated at the tag `P1^ZZNEWPR`. The second accessory function evaluates the sum over r at `S1^ZZNEWPR`. Finally, the probability $P(\chi^2|\nu)$ is computed at `P^ZZNEWPR`, which also invokes the `$$Gamma()` extrinsic function.

The approximation for large ν, (E6), is implemented at `X2^ZZNEWPR`, and invoked for $\nu \geq 60$ from `Q^ZZNEWPR`. The result is justified to three places, which serves to clue the user that the x_2 approximation has been used. Of course, `$$Q(CS,NU)` returns a *probability* after arguing `$$X1(CS,NU)` to the cumulative normal function.

[6] In a production application the extrinsic function to compute the probability associated with a statistic, for example $Q(\chi^2|\nu)$, would be called directly from the statistics routine. However, the statistics routines distributed with this paper are old (*ca.* 1984) and the user can integrate probabilities or make other modifications to these routines as needed.

[7] The tag `CumT^ZZNEWPR` returns $P(\{0 \leq x \leq t\}|\nu)$. To compute $Q(t|\nu)$ use 1-(2*`$$CumT()`) for a 2-tailed test, or 1 - (.5+`$$CumT()`) for a 1-tailed test. The approximation given in (E9) assures accuracy to at least three decimal places for $\nu > 15$. The definite integral of the density function (E8) is computed for $\nu \leq 15$ to the same degree of accuracy. Numeric integration is slow, as previously noted. However, the t-distribution is normally used in contexts wherein only one or a few statistics are computed.

Finally, an interesting inverse function $Ap(x,\nu) \rightarrow t$ is given at the tag `Ap^ZZNEWPR`. [2] Four terms of the asymptotic expansion suffice to give three places accuracy when $\nu > 3$ and $Q > .005$.

[8] Allen L. Edwards. *Experimental Design in Psychological Research.* Holt, Rinehart and Winston, New York 1963.

[9] Series expansions for $Q(F|\nu_1,\nu_2)$ are given in note [2] for the cases where ν_1 is even, ν_2 is even, and ν_1,ν_2 are odd. When $\nu_1 = \nu_2 = 1$ the latter case reduces to the following:

$$Q(F) = 1 - \frac{2\tan^{-1}\sqrt{F}}{\pi}$$

This special case is coded as Q11^ZZNEWPR in the sources, but requires the inverse tangent as an accessory function. [11] The series for even degrees of freedom are relatively simple. For example, the series expansion for the case where ν_1 is even is as follows:

$$Q(F|\nu_1,\nu_2) = x^{\nu_2/2}\left[1+\frac{\nu_2}{2}(1-x)+\frac{\nu_2(\nu_2+2)}{2\cdot 4}(1-x)^2+\dots\right.$$
$$\left.+\frac{\nu_2(\nu_2+2)\cdots(\nu_2+\nu_1-4)}{2\cdot 4\cdots(\nu_1-2)}(1-x)^{\frac{\nu_1-2}{2}}\right]$$
$$\text{where,}\qquad x=\frac{\nu_2}{\nu_2+\nu_1 F}\qquad [2]$$

The case where ν_2 is even is similar. The products are expanded at tags P2 and P3^ZZNEWPR respectively, and the sums are expanded at S2 and S3^ZZNEWPR. The series expansion for the odd degrees of freedom case involves powers of trigonometric functions, and is not coded.

An approximation for ν_1 and ν_2 large is also given in [2]:

$$Q(F|\nu_1,\nu_2)\approx Q(x),\qquad x=\frac{F^{1/3}\left(1-\frac{2}{9\nu_2}\right)}{\sqrt{\frac{2}{9\nu_1}+F^{2/3}\frac{2}{9\nu_2}}}$$

In practice, $\nu_1 > 8$ and $\nu_2 > 10$ suffice for accuracy to three decimal places. This approximation is coded at tag ApQF^ZZNEWPR. Line tag QF^ZZNEWPR serves as an entry point for computing the probability $1-P(F|\nu_1,\nu_2)$, automatically selecting the method appropriate to the degrees of freedom. The best cases are the series for even degrees of freedom and $Q(F|1,1)$. Next best is the approximation for large degrees of freedom. The worst is the case of small odd degrees of freedom, which is handled by harmonic interpolation.

[10] James L. Bruning and B.L. Kintz. *Computational Handbook of Statistics: Second Edition.* Scott, Foresman and Company, Glenview Illinois, 1977.

[11] For the inverse tangent we use the function $$ATANRAD^MATHALT, coded by Carl Bauer and Deidre Waclaw.

USER MAR 5,1993@14:11 Listing 1 - Probability Functions

```
ZZNEWPR * *  212 LINES,  8246 BYTES

ZZNEWPR;CHA/LM;03:49 PM  29 Oct 1992
       ;;
       ;;Reference HMF = Handbook of Mathematical Functions, edited by:
       ;;                Milton Abramowitz and Irene A. Stegun
       ;;                Dover Publications, New York, 1972.
       Q
       ;;
gamma();;Euler's constant
       Q .5772156649015328606065l2
e()    ;;
       Q 2.718281828459045235360287
       ;;
PI()   ;;
       Q 3.14159265358979323846264 3
       ;;
C1()   ;;Reciprocal of (square root 2 * Gamma 1/2)
       Q .398942280380143
MIN(X,Y);;
       Q $S(X<Y:X,1:Y)
       ;;
ABS(X) ;;Absolute Value
       Q $S(X<0:-X,1:X)
       ;;
HM(A,B);;Harmonic mean of 2 quantities
       Q 2*A*B/(A+B)
       ;;
TRAPZ(EXPR,VAR,LB,UB,N);;Trapezoidal Rule
       N DX,SUM,@VAR
       S @VAR=LB,@("SUM="_EXPR),DX=UB-LB/N Q:'DX 0
       F @VAR=LB+DX:DX:UB-DX D
       .S @("SUM="_EXPR_"*2+SUM")
       S @VAR=UB,@("SUM="_EXPR_"+SUM*DX/2")
       Q SUM
       ;;
SIMPSONS(EXPR,VAR,LB,UB,N);;Simpson's Rule
       N %,DX,SUM,@VAR
       S %=0,@VAR=LB,@("SUM="_EXPR),DX=UB-LB/N Q:'DX 0
       F @VAR=LB+DX:DX:UB-DX D
       .S %='%,@("SUM="_EXPR_"*2*(%+1)+SUM")
       S @VAR=UB,@("SUM="_EXPR_"+SUM*DX/3")
       Q SUM
       ;;
P1(CS,NU,R);;Product: CS^R/((NU+2)*(NU+4)*...*(NU+2R))
       Q:R<1 "" N %,I S %=1 F I=2:2:2*R S %=%*CS/(NU+I)
       Q %
       ;;
S1(CS,NU);;Sum of $$P1(CS,NU,R) from R=1 to infinity.
       ;;error < .00000001
       N %,S,I S (%,S)=0 F I=1:1 S %=%+$$P1(CS,NU,I) Q:%-S<.00000001  S S=%
       Q %
       ;;
P(CS,NU);;Probability of Chi Square, given NU degrees of freedom.
```

```
ZZNEWPR * *  (cont.)

       ;;HMF: 26.4.6, a series expansion for P(CS|NU).
       Q 1+$$S1(CS,NU)*$$POWER(CS/2,NU/2)*$$EXP(-CS/2)/$$Gamma(NU+2/2)
       ;;
X2(CS,NU);;Approximation to the Chi-Square Distribution, NU>30.
       ;;HMF: 26.4.14
       Q 2/9/NU-1+$$POWER(CS/NU,1/3)/$$POWER(2/9/NU,.5)
       ;;
Q(CS,NU);;
       Q:NU<60 $J(1-$$P(CS,NU),0,5)
       Q $J($$CumNorm(-$$X2(CS,NU)),0,3)
       ;;
Poly(X);;Polynomial approximation of Gamma(1+X), 0<=X<=1, (HMF, 6.1.36).
       ;;
       Q:'$D(X)!(X<0)!(X>1) "" N B,G
       S B(1)=-.577191652,B(2)=.988205891,B(3)=-.897056937,B(4)=.918206857
       S B(5)=-.756704078,B(6)=.482199394,B(7)=-.193527818,B(8)=.035868343
       S B(0)=1,G=B(8) F B=8:-1:1 S G=G*X+B(B-1)
       Q $J(G,0,6) ;Absolute value of error(X) <= .0000003
       ;;
Gamma(X);;The Gamma (Factorial) Function, based on the recurrence
       ;;relation: Gamma(z+1) = z*Gamma(z), 0<X, (HMF, 6.1.15).
       N % S %=$S(X>2:X-1*$$Gamma(X-1),X>1:$$Poly(X-1),X>0:1/$$RecipG(X),1:"")
       Q $J(%,0,6-$L($P(%,".")))
       ;;
FastG(X);;Selected values (integers and half integers up to 10)
       Q $S(X=.5:1.77245,X=1:1,X=1.5:.88623,X=2:1,X=2.5:1.32935,X=3:2,X=3.5:3.3
    ...2338,X=4:6,X=4.5:11.6318,X=5:24,X=5.5:52.3431,X=6:120,X=6.5:287.887,X=7
    ...:720,X=7.5:1871.27,X=8:5040,X=8.5:14034.5,X=9:40320,X=9.5:119293,X=10:3
    ...62880,1:"")
       ;;
RecipG(X);;1/Gamma(X), from series expansion for 1/Gamma(z),
       ;;(HMF, 6.1.34, reproduced from H.T.Davis, Tables of higher
       ;;mathematical functions, 2 vols., Principia Press, Bloomington, 1935.)
       Q:'$D(X) "" N K,R
       S K(1)=1,K(2)=.5772156649015329,K(3)=-.6558780715202538
       S K(4)=-.0420026350340952,K(5)=.1665386113822915
       S K(6)=-.0421977345555443,K(7)=-.009621971527877
       S K(8)=.007218943246663,K(9)=-.0011651675918591
       S K(10)=-.0002152416741149,K(11)=.0001280502823882
       S K(12)=-.0000201348547807,K(13)=-.0000012504934821
       S K(14)=.000001133027232,K(15)=-.0000002056338417
       S K(16)=.000000006116095,K(17)=.0000000050020075
       S K(18)=-.0000000011812746,K(19)=.0000000001043427
       S K(20)=.0000000000077823,K(21)=-.0000000000036968
       S K(22)=.00000000000051,K(23)=-.0000000000000206
       S K(24)=-.0000000000000054,K(25)=.0000000000000014
       S K(26)=.0000000000000001
       S K(0)=0,R=K(26) F K=26:-1:1 S R=R*X+K(K-1)
       Q $J(R,0,10)
       ;**********************************************************************
       ;**********************************************************************
POWER(X,Y);;Implementation-specific
```

```
ZZNEWPR * *  (cont.)

        Q X**Y ;DTM
        ;;
EXP(X) ;;
        Q $$POWER($$e,X)
        ;;
DensityG(X,ALPHA,BETA);;Density function of the Gamma distribution
        Q $S(X>0:$$POWER(X,ALPHA)*$$EXP(-X/BETA)/$$POWER(BETA,ALPHA+1)/$$Gamma(A
    ...LPHA+1),1:0)
        ;;
ChiSqr(X,N);;Density function of the Chi Square distribution with 2*N df.
        Q $S(N=.5:$$C1*$$POWER(X,-.5)*$$EXP(-X/2),N=1:.5*$$EXP(-X/2),N=1.5:$$C1*
    ...$$POWER(X,.5)*$$EXP(-X/2),N=2:.25*X*$$EXP(-X/2),N>2:$$POWER(X,N-1)/$$PO
    ...WER(2,N)*$$EXP(-X/2)/$$Gamma(N))
        ;;
t(X,N) ;;Density function of the t distribution with N df.
        Q $$Gamma(N+1/2)/$$Gamma(N/2)/$$POWER(N*$$PI,.5)*$$POWER(X*X/N+1,N+1/-2)
        ;;
F(X,M,N);;Density function of the F distribution with (M,N) df.
        Q $$Gamma(M+N/2)/$$Gamma(M/2)/$$Gamma(N/2)*$$POWER(M/N,M/2)*$$POWER(X,M/
    ...2-1)*$$POWER(M/N*X+1,M+N/-2)
        ;;
ApQF(F,nu1,nu2);;Normal approximation to Q(F) for large degrees of freedom.
        ;;HMF: 26.6.15
        Q 1-$J($$CumNorm(-2/9/nu2+1*$$POWER(F,1/3)-1+(2/9/nu1)/$$POWER(2/9/nu1+(
    ...$$POWER(F,2/3)*2/9/nu2),.5)),0,3)
        ;;
P2(x,nu1,nu2,I);;I-th product in series expansion of Q(F|nu1,nu2), (nu1, even)
        N %,P S P=1
        F %=1:1:I S P=2*%-2+nu2*P/2/%*(1-x)
        Q P
        ;;
P3(x,nu1,nu2,I);;I-th product in series expansion of Q(F|nu1,nu2), (nu2, even)
        N %,P S P=1
        F %=1:1:I S P=2*%-2+nu1*P/2/%*x
        Q P
        ;;
S2(F,nu1,nu2);;Series expansion of Q(F|nu1,nu2), (nu1, even)
        N %,S,X S S=1,X=nu2/(nu1*F+nu2)
        F %=1:1:nu1-2/2 S S=S+$$P2(X,nu1,nu2,%)
        Q $J($$POWER(X,nu2/2)*S,0,5)
        ;;
S3(F,nu1,nu2);;Series expansion of Q(F|nu1,nu2), (nu2, even)
        N %,S,X S S=1,X=nu2/(nu1*F+nu2)
        F %=1:1:nu2-2/2 S S=S+$$P3(X,nu1,nu2,%)
        Q $J(1-($$POWER(1-X,nu1/2)*S),0,5)
        ;;
Q11(F) ;;A special case: Q(F|1,1), Requires inverse tangent.
        Q $J(1-(2*$$ATANRAD^MATHALT($$POWER(F,.5))/$$PI),0,5)
        ;;
QX1(F,nu1,nu2);;Approximation using harmonic mean, (nu1,nu2 odd, nu1 > 1)
        Q $$HM($$S2(F,nu1-1,nu2),$$S2(F,nu1+1,nu2))
        ;;
```

```
ZZNEWPR * *  (cont.)

Q1X(F,nu1,nu2);;Approximation using harmonic mean, (nu1,nu2 odd, nu2 > 1)
       Q $$HM($$S3(F,nu1,nu2-1),$$S3(F,nu1,nu2+1))
       ;;
QXX(F,nu1,nu2);;Approximation using arithmetic mean of harmonic means,
       ;;(nu1,nu2 odd, nu1,nu2 > 1)
       Q $$QX1(F,nu1,nu2)+$$Q1X(F,nu1,nu2)/2
       ;;
QF(F,nu1,nu2);;Various cases: series expansions, approximation,
       ;; interpolation
       Q:'(nu1#2) $$S2(F,nu1,nu2) Q:'(nu2#2) $$S3(F,nu1,nu2) ;Series Expansions
    ... (nu1 or nu2 even)
       Q:nu1>8&(nu2>10) $$ApQF(F,nu1,nu2) ;Normal approximation
       Q:nu1=1&(nu2=1) $$Q11(F) ;Special case for 1,1 degrees of freedom.
       ;;Odd degrees of freedom, miscellaneous cases (Interpolation):
       Q:nu1=1 $J($$Q1X(F,nu1,nu2),0,3) ;Weak approximation when df small.
       Q:nu2=1 $J($$QX1(F,nu1,nu2),0,3)
       Q $J($$QXX(F,nu1,nu2),0,3)
       ;;
CumF(F,nu1,nu2);;F distribution with (nu1,nu2) degrees of freedom.
       ;;Note:  Q(F|nu1,nu2) = 1-$$CumF(F,nu1,nu2)
       ;;For small degrees of freedom increase iterations (slow).
       N C1,C2,C3,C4
       S C1=$$Gamma(nu1+nu2/2)/$$Gamma(nu1/2)/$$Gamma(nu2/2)*$$POWER(nu1/nu2,nu
    ...1/2)
       S C2=nu1/2-1,C3=nu1/nu2,C4=nu1+nu2/-2
       Q $J($$SIMPSONS("C1*$$POWER(X,C2)*$$POWER(C3*X+1,C4)","X",0,F,100),0,5)
       ;;
A(t,nu);;Student's t approximation for large degrees of freedom.
       ;;HMF: 26.7.8
       Q -t/(4*nu)+t/$$POWER(t*t/(2*nu)+1,.5)
       ;;
CumT(t,nu);;t distribution with nu degrees of freedom.
       ;;Note: Use 2*$$CumT for 2-tailed test, .5+$$CumT for 1-tailed test.
       ;;Q = 1-(2*$$CumT), or .5-$$CumT.
       Q $J($S(nu>15:$$CumNorm($$A(t,nu))-.5,1:$$TRAPZ("$$t(X,nu)","X",0,t,40))
    ...,0,3)
       ;;
CumChi(ChiSqr,nu);;Chi Square distribution with nu degrees of freedom.
       ;;Note: Q(ChiSqr,nu) = 1-$$CumChi(ChiSqr,nu)
       ;;Note: 100 iterations suffice when nu > 5.
       Q $J($$MIN($$TRAPZ("$$ChiSqr(X,nu/2)","X",0,ChiSqr,100),1),0,5)
       ;;
CumNorm(X);;Polynomial Approximation to Normal Distribution
       ;;HMF: 26.2.19. Use 1-$$CumNorm(X) for one-tailed probability.
       N D,P S D(1)=.049867347,D(2)=.0211410061,D(3)=.0032776263
       S D(4)=.0000380036,D(5)=.0000488906,D(6)=.000005383
       S D(0)=1,P=D(6) F D=6:-1:1 S P=P*X+D(D-1)
       Q $J($$POWER(P,-16)/-2+1,0,5)
       ;;
G1(X)  ;;These polynomials are used in the asymptotic expansion for
       ;;the inverse function: X (Normal) -> t.
       Q X*X*X+X/4
```

```
    USER  MAR   5,1993@14:11                Listing 1 - Probability Functions

ZZNEWPR * *  (cont.)

         ;;
G2(X)    ;;
         Q ((5*X*X+16)*X*X+3)*X/96
         ;;
G3(X)    ;;
         Q (((3*X*X+19)*X*X+17)*X*X-15)*X/384
         ;;
G4(X)    ;;
         Q ((((79*X*X+776)*X*X+1482)*X*X-1920)*X*X-945)*X/92160
         ;;
Ap(X,nu);;Given X, a normal deviate and nu degrees of freedom
         ;;return Student's t to 3-places when nu > 3 and Q > .005
         ;;
         Q $J((($$G4(X)/nu+$$G3(X))/nu+$$G2(X))/nu+$$G1(X))/nu+X,0,3)
         ;;
```