National Bureau of Standards
Statistical Engineering Division
Gaithersburg, MD 20899

N-86-4
November 1986

An Algorithm for Computing the Doubly Noncentral F C.D.F. to a Specified Accuracy

Charles P. Reeve

Let X_1 and X_2 be independent noncentral chi-squared random variables with degrees of freedom ν_1 and ν_2 (both >0) and noncentrality parameters λ_1 and λ_2 (both ≥ 0) respectively. Then the random variable

$$Y = (X_1/\nu_1)/(X_2/\nu_2) \tag{1}$$

is said to have the doubly noncentral F distribution, indicated by $Y \sim F''(\nu_1,\nu_2,\lambda_1,\lambda_2)$. This distribution has been used in the evaluation of the power function of analysis of variance tests in which interaction or bias effects occur, as in Scheffé [11]. Numerical examples of this usage are given in Bulgren [2] and Tiku [13]. It has also been used in engineering problems in the context of information theory as discussed in Price [8].

Exact formulas for the F" cumulative distribution function (c.d.f.) are given in Tiku [12] and [2] using the beta c.d.f., in Tiku [14] using Laguerre polynomials, and in [8] for special cases of ν_1 and ν_2. Approximations are given in Johnson and Kotz [4] and Tiku [12,13] using the central F c.d.f., and in Mudholkar, Chaubey, and Lin [5] using Edgeworth series expansions.

The author has been unable to find any published algorithms for computing exact values of the F" c.d.f. although computer programs have obviously been used in generating published tables in [2,5,12,14]. The purpose of this note is to present an efficient algorithm for computing the F" c.d.f. to a specified accuracy using exact formulas.

The algorithm uses the series representation in eq. (2.2) of [2] which, unfortunately, contains typographical errors. With these corrected the c.d.f. of Y can be re-written

$$F_Y(x) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} A_i B_j I(u,\nu_1/2+i,\nu_2/2+j) \tag{2}$$

where $A_i = (\lambda_1/2)^i e^{-\lambda_1/2}/\Gamma(i+1)$, $B_j = (\lambda_2/2)^j e^{-\lambda_2/2}/\Gamma(j+1)$, $u = \nu_1 x/(\nu_1 x+\nu_2)$, and $x \geq 0$. The A_i and B_j are Poisson probabilities, and

$I(u,a,b) = \int_0^u t^{a-1}(1-t)^{b-1}dt/B(a,b)$ is the c.d.f. of the beta distribution (also called the incomplete beta ratio) where $0 \leq u \leq 1$, $a>0$, $b>0$, and

$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$. For computational purposes the two infinite series must be truncated, thus (2) is re-expressed as

$$F_Y(x) = \sum_{i=I'}^{I''} \sum_{j=J'}^{J''} A_i B_j I(u,\nu_1/2+i,\nu_2/2+j) + R \qquad (3)$$

where I', I", J', and J" are non-negative integers and R is the remainder. If the beta c.d.f.'s are computed without error, it can easily be shown that choosing I', I", J', and J" such that $\sum_{i=I'}^{I''} A_i \geq 1-\varepsilon/2$ and $\sum_{j=J'}^{J''} B_j \geq 1-\varepsilon/2$ yields $R \leq \varepsilon$ provided $\varepsilon > 0$. Therefore, ε serves as an absolute error bound on $F_Y(x)$. For maximum computational efficiency, the number of terms in each sum is minimized by indexing i and j over the largest of the Poisson probabilities A_i and B_j respectively. The final task is to compute the $(I''-I'+1)(J''-J'+1)$ beta c.d.f.'s and the double summation. An efficient procedure for doing this is to compute only $I(u,\nu_1/2+I',\nu_2/2+J^*)$ and $I(u,\nu_1/2+I^*,\nu_2/2+J')$ directly, indicated by the symbols "x" and "y" in figure 1. The remaining beta c.d.f's are computed using the recurrence relations

$$I(x,a,b) = I(x,a,b+1) - x^a(1-x)^b/[bB(a,b)], \qquad (4a)$$

$$I(x,a,b) = I(x,a+1,b) + x^a(1-x)^b/[aB(a,b)], \text{ and} \qquad (4b)$$

$$I(x,a,b) = xI(x,a-1,b) + (1-x)I(x,a,b-1) \qquad (4c)$$

as found in Abramowitz and Stegun [1]. Subject to the restrictions $J' \leq J^* \leq J''$

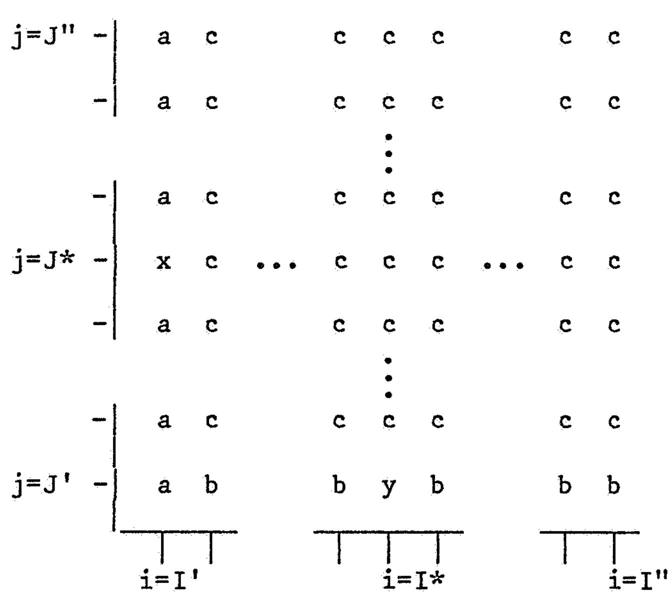

Figure 1

Computation of the $I(u,\nu_1/2+i,\nu_2/2+j)$