National Bureau of Standards
Statistical Engineering Division
Gaithersburg, MD 20899

N-86-5
December 1986

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

Charles P. Reeve

Let Z be a normally distributed random variable with mean δ and variance 1, and X be a noncentral chi-squared random variable, independent of Z, with degrees of freedom $\nu>0$ and noncentrality parameter $\lambda \geq 0$. Then the random variable

$$Y = Z/\sqrt{X/\nu} \qquad (1)$$

has the doubly noncentral t distribution, indicated by $Y \sim t''(\nu,\delta,\lambda)$. This distribution was introduced by Robbins [14] as the distribution of Student's t statistic when the observations have unequal population means. It was later used by Patnaik [11] in testing hypotheses concerning the standardized means of nonhomogeneous normal populations.

Krishnan [7] gives a series representation for the t" cumulative distribution function (c.d.f.) in terms of incomplete beta functions. Alternative series representations are given by Bulgren and Amos [2], Bulgren [3], Carey [4], and [7]. Approximations are given by Johnson and Kotz [6] and Mulholkar and Chaubey [8]. Numerical examples of usage are given in [3] and [7].

The author has been unable to find any published algorithms for computing exact values of the t" c.d.f. although computer programs have obviously been used in generating published tables in [2,3,4,7,8]. The purpose of this note is to present an efficient algorithm for computing the t" c.d.f. to a specified accuracy using exact formulas.

The algorithm uses the series representation in eq. (4) of [7] which can be re-written

$$F_Y(x) = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} A_j B_i I(u,1/2+i/2,\nu/2+j)/2 + \sum_{j=0}^{\infty} A_j \sum_{i=0}^{\infty} B_i(-1)^i/2 \qquad (2)$$

where $A_j = (\lambda/2)^j e^{-\lambda/2}/\Gamma(j+1)$, $B_i = (\delta/\sqrt{2})^i e^{-\delta^2/2}/\Gamma(i/2+1)$, $u = x^2/(x^2+\nu)$, and $x \geq 0$. When $x<0$ the c.d.f. is computed from the relation $F_Y(x;\nu,\delta,\lambda) = 1 - F_Y(-x;\nu,-\delta,\lambda)$. The A_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 $0<u<1$, $a>0$, $b>0$, and $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$. The quantity $(\delta/\sqrt{2})^i$ in B_i is erroneously given as $(\delta^2/2)^{i/2}$ in [7]. In the latter form the quantity would be (incorrectly) positive when δ is negative and i is an odd integer.

Each summation over i in (2) can be split into two summations over even and odd values of i. For i=0,1,2,... let $B_i^e = B_{2i} = (\delta^2/2)^i e^{-\delta^2/2}/\Gamma(i+1)$ and $B_i^o = B_{2i+1} = (\delta/\sqrt{2})(\delta^2/2)^i e^{-\delta^2/2}/\Gamma(i+3/2)$. Then (2) takes the form

$$F_Y(x) = \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} A_j B_i^e I(u,1/2+i,\nu/2+j)/2 + \sum_{j=0}^{\infty}\sum_{i=0}^{\infty} A_j B_i^o I(u,1+i,\nu/2+j)/2 + \{1 - \sum_{i=0}^{\infty} B_i^o\}/2 \qquad (3)$$

since $\sum_{j=0}^{\infty} A_j = \sum_{i=0}^{\infty} B_i^e = 1$. For computational purposes the infinite series must be truncated, thus (3) is re-expressed as

$$F_Y(x) = \sum_{j=J'}^{J''}\sum_{i=I_e'}^{I''} A_j B_i^e I(u,1/2+i,\nu/2+j)/2 + \sum_{j=J'}^{J''}\sum_{i=I_o'}^{I''} A_j B_i^o I(u,1+i,\nu/2+j)/2 + \{1 - \sum_{i=I_o'}^{I''} B_i^o\}/2 + R \qquad (4)$$

where I_e', I_o', 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_e', I_o', I'', J', and J'' such that $\sum_{j=J'}^{J''} A_j \geq 1-2\varepsilon/3$, $\sum_{i=I_e'}^{I''} B_i^e \geq 1-2\varepsilon/3$, and $I_o' = \max\{I_e'-1,0\}$ yields $R<\varepsilon$ 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 j and i over the largest of the Poisson probabilities A_j and B_i^e respectively. It then follows that i also indexes over the B_i^o which are largest in absolute value.

The final task is to compute the $(2I''-I_e'-I_o'+2)(J''-J'+1)$ beta c.d.f.'s and the summations. An efficient procedure for doing this is to first compute only $I(u,1/2+I_e',\nu/2+J_e^*)$ and $I(u,1/2+I_e^*,\nu/2+J')$ directly, indicated by the symbols "x" and "y" in figure 1. The remaining beta c.d.f's are computed using the