ERRORS IN NORMAL APPROXIMATIONS TO THE t, τ, AND SIMILAR TYPES OF DISTRIBUTION[1,2]

BY J. T. CHU[3]

University of North Carolina and *Case Institute of Technology*

1. Summary. For the cdf's of Student's (t) and Thompson's (τ) distributions, upper and lower bounds are obtained in terms of the normal cdf. It is then shown that, in using the normal approximation for the cdf's of these distributions, the proportional errors are uniformly smaller than $1/n$ for all $n \geqq 8$ and 13, respectively, where n is the number of degrees of freedom. Similar methods may be used to derive bounds for cdf's of similar types. Examples are given.

2. Introduction. Let $F_n(x)$, $n = 1, 2, \cdots$, be a sequence of cdf's (cumulative distribution functions) such that for every fixed x, $F_n(x) \to F(x)$ as $n \to \infty$, where $F(x)$ is a cdf independent of n. From a practical point of view, it is desirable to know how large n has to be in order that $D_n(x) = | F_n(x) - F(x) |$ be small enough so that $F(x)$ may be used as an approximation to $F_n(x)$, although approximations are often used in practice without much knowledge about the magnitudes of the errors. The function $D_n(x)$ may, of course, vary considerably for different values of n and x. But the most interesting kinds of $D_n(x)$'s are probably those which tend rapidly to 0, uniformly in x. In such cases, $F(x)$ provides for all n's greater than some minimum, and for all x's, a satisfactory approximation for $F_n(x)$. Generally, however, even though there is ample numerical evidence that as n increases $D_n(x)$ rapidly becomes uniformly small, it may not be easy to obtain a mathematical proof.

There are, on the other hand, types of sequences of cdf's for which we are able to confirm rigorously that they do tend rapidly to normality. Suppose that a cdf has one of the following forms:

$$F_n(x) = C_n \int_{-\infty}^{x} (1 \pm z^2/n)^{\mp m/2}\, dz,$$

where C_n and m depend only on n, a positive integer. (If the integrand is $1 - z^2/n$, it should be replaced by 0 when $z^2 \geqq n$.) By simple transformations of the variable of integration, upper and lower bounds are found for $F_n(x)$ in terms of $\Phi(x)$, the unit-normal cdf specified by (6). These bounds may sometimes be further simplified by obtaining bounds on C_n. If $m/n \to 1$ as $n \to \infty$, then $\sqrt{2\pi}C_n \to 1$, and for every fixed x, both bounds for $F_n(x)$, and consequently

Received April 6, 1955.

[1] Most of the work was done at Chapel Hill under the sponsorship of the Office of Naval Research.

[2] Presented by title, "Errors in normal approximations to certain types of distribution functions," at the 1955 Western Regional Meeting of the IMS.

[3] Now with the Department of Mathematics, Case Institute of Technology.

$F_n(x)$ itself, tend to $\Phi(x)$. If m/n and $\sqrt{2\pi}C_n$ tend rapidly to 1, then $F_n(x)$ tends rapidly to $\Phi(x)$ for all x. Therefore, in this case, the error in using $\Phi(x)$ as an approximation to $F_n(x)$ rapidly becomes uniformly small as n increases. In Section 4 applications are given to sequences of cdf's corresponding to the Student's t-distribution, the τ-distribution of W. R. Thompson [10], and the distributions of the partial and total correlation coefficients when the variates involved are independently and normally distributed. For most of these cdf's, we are able to show that the error bounds in using the normal approximation are small, although the actual errors may be even smaller.

An application to the χ^2-distribution is given in Section 4.D. Similar methods were used by the author [1] to derive upper and lower bounds for the cdf of the sample median $\tilde{x}$ in terms of its asymptotic distribution function (which is normal). There, we also showed that if the parent distribution is normal, then, even for samples of moderate sizes, the error is small in using the normal approximation to the cdf of $\tilde{x}$. The cdf of $\tilde{x}$ can be reduced to one of the forms given above by several transformations. But some different arguments are also needed in order to get the bounds obtained in [1].

Another type of bound (also in terms of Φ) is derived for the cdf's of the t- and τ-distributions (see Equations (24) through (27)) by using the interrelationships between these cdf's and their bounds obtained by the methods described previously. In Section 5 some numerical comparisons are given of the two types of bounds and of two kinds of approximations (the normal and Hendrick's [6]) for the cdf of the t-distribution.

3. Lemmas.

LEMMA 1.

$$1 + x \leqq e^x, \qquad \textit{for all real } x, \tag{1}$$

$$1 + x \gtreqless e^{x - x^2/2}, \qquad \textit{according as } x \gtreqless 0. \tag{2}$$

If $x \geqq 0$, *then*

$$x/(1 - e^{-x^2})^{1/2} \geqq 1, \tag{3}$$

and

$$xe^{-x^2}/(1 - e^{-x^2})^{1/2} \leqq 1. \tag{4}$$

PROOF. The function $e^x - x - 1$ has its minimum 0 at $x = 0$, hence we have (1); (2) holds because $\log(1 + x) - x + x^2/2$ is monotonically increasing for all $x > -1$. Substituting $-x^2$ for x in (1), we have (3), and (4) follows from the fact that the LHS (left-hand side) tends to 1 as $x \to 0$ and is a monotonically decreasing function of x. (Differentiate twice).

LEMMA 2. *Let*

$$b_n(c) = \frac{1 \times 3 \cdots (2n - 1)}{2 \times 4 \cdots (2n)} \sqrt{n + c}, \qquad n = 1, 2, \cdots,$$

where $c \geqq -1$. *Then,*

$$\sqrt{\pi} b_n(c) < 1, \qquad \text{if } c \leqq \tfrac{1}{4}, \tag{5}$$
$$> 1, \qquad \text{if } c \geqq \tfrac{2}{7}.$$

PROOF. $b_n(0)$ is known ([11], p. 351) as the Wallis product and tends to $1/\sqrt{\pi}$ as $n \to \infty$. Obviously $b_n(c)$ tends to the same limit for every fixed c. By examining the square of the ratio $b_{n+1}(c)/b_n(c)$, it can be shown that $b_n(c)$ is a strictly increasing function of n if $c \leqq \frac{1}{4}$, and is a strictly decreasing function of n if $c \geqq \frac{2}{7}$ and $n \geqq 2$. Hence, we have (5).

If we have a chain of inequalities, as in (7) below, of the form $A_1 \leqq A_2 \leqq A_3 \leqq A_4 \leqq A_5$, where the A's are functions of m, n, x, or other such quantities, the particular inequality $A_i \leqq A_j$ $(i < j)$ will be denoted by (7. ij).

LEMMA 3. *Let*

$$\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} e^{-t^2/2}\, dt, \tag{6}$$

and $\Phi_0(x) = \Phi(x) - \frac{1}{2}$. *Then,* (7.12) *and* (7.23) *hold for all* $m, n > 0$, *and* $0 \leqq x \leqq \sqrt{n}$; (7.34) *holds for all* $m, n > 0$, *and* $0 \leqq x < \infty$; *and* (7.45) *for all* $m > 3$, $n > 0$, *and* $0 \leqq x < \infty$.

$$\begin{aligned}\sqrt{n/(m+2)}\,\Phi_0(x\sqrt{(m+2)/n}) &\leqq (2\pi)^{-1/2} \int_0^x (1 - z^2/n)^{m/2}\, dz \\ &\leqq \sqrt{n/m}\,\Phi_0(x\sqrt{m/n}) \leqq (2\pi)^{-1/2} \int_0^x (1 + z^2/n)^{-m/2}\, dz \\ &\leqq \sqrt{n/(m-3)}\,\Phi_0(x\sqrt{(m-3)/n}\,)\end{aligned} \tag{7}$$

PROOF. It is easy to see that (7.23) and (7.34) are immediate consequences of (1). Now use the transformation

$$v(z) = [n \log (1 + z^2/n)]^{1/2},$$

so that

$$\int_0^x (1 + z^2/n)^{-m/2}\, dz = \int_0^{v(x)} \exp\,[-\,(m-3)v^2/2n]\,h(v/\sqrt{n})\, dv,$$

where $h(x)$ is the LHS of (4). By (2) and (4), $h(v/\sqrt{n}) \leqq 1$ and $v(x) \leqq x$. Hence we have (7.45). Finally, (7.12) can be obtained in a similar way by using (1) and (3) after applying to the integral of (7.12) the transformation

$$u(z) = [-n \log (1 - z^2/n)]^{1/2}.$$

LEMMA 4. *Suppose* n_0 *is a fixed integer, and for every integer* $n \geqq n_0$,

$$F_n(x) = C_n \int_{-\infty}^{x} (1 \pm z^2/n)^{\mp m/2}\, dz \tag{8}$$

is a cdf, *where* C_n *and* m *depend only on* n *and* $\lim_{n\to\infty} m/n = 1$. (*If the integrand is* $(1 - z^2/n)^{m/2}$, *it should be replaced by* 0 *whenever* $|z| \geqq \sqrt{n}$.) *Then, for every fixed* x,

$$\lim_{n\to\infty} F_n(x) = \Phi(x), \tag{9}$$

where $\Phi(x)$ *is defined by* (6).

PROOF. By Lemma 3, we have $\lim_{n\to\infty} C_n = 1/\sqrt{2\pi}$. Using the same lemma once again, we obtain (9).

4. Normal approximations. We showed in Lemma 4 that if a cdf is of one of the types (8), then it tends to Φ as $n \to \infty$. In this section we shall use Lemmas 2 and 3 to prove that for several well-known sequences of cdf's of these types, the "speed" of approaching the limiting cdf is "uniformly rapid." Therefore, in using $\Phi(x)$ as an approximation to these cdf's, the error is small for all values of x, if n is greater than a certain minimum.

A. *t-distribution.* The cdf of the t-distribution with n d.f. (degrees of freedom), $n = 1, 2, \cdots$, is given by

$$F_n(x) = \int_{-\infty}^{x} a_n(1 + z^2/n)^{-(n+1)/2}\, dz, \tag{10}$$

where

$$a_n = (n\pi)^{-1/2}\Gamma\left(\frac{n+1}{2}\right)/\Gamma(n/2). \tag{11}$$

It is well known that as $n \to \infty$, $F_n(x) \to \Phi(x)$ for every fixed x, and that the "speed" of approaching the limit is rather fast. In fact, the normal approximation is often used in practice when $n \geqq 30$. We shall derive for $F_n(x)$ upper and lower bounds in terms of $\Phi(x)$, then show that the proportional error in using $\Phi(x)$ as an approximation to $F_n(x)$ is less than $1/n$ for all x and all $n \geqq 8$.

Applying (7.45) to $F_n(y) - \frac{1}{2}$ and $\frac{1}{2} - F_n(-x)$ and using the fact that $\Phi(-x) = 1 - \Phi(x)$, it can be shown easily that for arbitrary $x, y \geqq 0$, and $n \geqq 3$,

$$\begin{aligned} &F_n(y) - F_n(-x) \\ &\qquad \leqq a_n\sqrt{2\pi n/(n-2)}[\Phi(y\sqrt{(n-2)/n}) - \Phi(-x\sqrt{(n-2)/n})]. \end{aligned} \tag{12}$$

From (7.34) we obtain, in a similar way,

$$\begin{aligned} &F_n(y) - F_n(-x) \\ &\qquad \geqq a_n\sqrt{2\pi n/(n+1)}[\Phi(y\sqrt{(n+1)/n}) - \Phi(-x\sqrt{(n+1)/n})]. \end{aligned} \tag{13}$$

Using $\Gamma(x+1) = x\Gamma(x)$ and $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, it can be seen that for any $c \geqq -1$, $a_{2m} = b_m(c)\sqrt{m/(2m+2c)}$ and $a_{2m+1} = \sqrt{(m+c)/(2m+1)}/(\pi b_m(c))$, where

$m = 1, 2, \cdots$. Letting c be $\frac{1}{4}$ and $\frac{2}{7}$ in turn, we obtain, by (5), $\sqrt{2\pi}a_n < \sqrt{2n/(2n+1)}$ if $n = 2m$ and $< \sqrt{1 - \frac{3}{7}n}$ if $n = 2m + 1$, $m = 1, 2, \cdots$. In general, for $n \geqq 3$,

$$\sqrt{2\pi}a_n < \sqrt{1 - \tfrac{3}{7}n}. \tag{14}$$

Likewise, letting $c = \frac{1}{2}$ and $\frac{1}{4}$, respectively, we obtain $\sqrt{2\pi}a_n > \sqrt{n/(n+1)}$ if $n = 2m$ and $> \sqrt{1 - \frac{1}{2}n}$ if $n = 2m + 1$, $m = 1, 2, \cdots$. In general, for $n \geqq 1$,

$$\sqrt{2\pi}a_n > \sqrt{n/(n+1)}. \tag{15}$$

(Direct comparison shows that (15) holds for $n = 1$.)

From (12) through (15), we have, for arbitrary $x, y \geqq 0$, and $n \geqq 3$,

$$\begin{aligned} &F_n(y) - F_n(-x) \\ &\quad \leqq \sqrt{(7n-3)/(7n-14)}[\Phi(y\sqrt{(n-2)/n}) - \Phi(-x\sqrt{(n-2)/n})], \end{aligned} \tag{16}$$

$$\begin{aligned} &F_n(y) - F_n(-x) \\ &\quad \geqq (n/(n+1))[\Phi(y\sqrt{(n+1)/n}) - \Phi(-x\sqrt{(n+1)/n})]. \end{aligned} \tag{17}$$

The proportional error in using A as an approximation to B is defined to be

$$E = |(B/A) - 1|. \tag{18}$$

Now, omitting $\sqrt{1 - 2/n} < 1$ and $\sqrt{1 + 1/n} > 1$ in the arguments in the Φ's of (16) and (17), we see that E is not more than the maximum of $\sqrt{(7n-3)/(7n-14)} - 1$ and $1 - n/(n+1)$. For simplicity, we may state that $E < 1/n$ for all $n \geqq 8$. The actual values of E are often much smaller than $1/n$. For example, if $n = 30$, and $y = x = 2.042$, then $F_n(y) - F_n(-x) = 0.95$ while $\Phi(y) - \Phi(-x) = 0.9588$, so $E = 0.0092$. Nevertheless, the bound $1/n$ is independent of x and y, and small enough to justify, in a rigorous manner, the use of the normal approximation, provided that n is not too small. More numerical comparisons are given in Section 5.

B. *Thompson's τ-distribution.* The cdf of the τ-distribution with n d.f. is given ([2], p. 241) by

$$G_n(x) = \int_{-\sqrt{n}}^{x} a_n'(1 - z^2/n)^{(n-3)/2}\,dz, \tag{19}$$

where $|x| \geqq \sqrt{n}$, $a_n' = (n\pi)^{-1/2}\Gamma(n/2)/\Gamma((n-1)/2)$, and $n = 2, 3, \cdots$. For applications of the τ-distribution, the readers are referred to ([2], p. 390) and [10]. Obviously by (11), $a_n' = a_{n-1}\sqrt{1 - 1/n}$. Using (7), then (14) and (15), we obtain for $x, y \geqq 0$, and $n \geqq 4$,

$$G_n(y) - G_n(-x)$$

$$\leqq a_{n-1}\sqrt{2\pi(n-1)/(n-3)}[\Phi(y\sqrt{(n-3)/n}) - \Phi(-x\sqrt{(n-3)/n})] \tag{20}$$

$$\leqq \sqrt{(7n-10)/(7n-21)}[\Phi(y\sqrt{(n-3)/n}) - \Phi(-x\sqrt{(n-3)/n})],$$

$$G_n(y) - G_n(-x) \geqq a_{n-1}\sqrt{2\pi}[\Phi(y\sqrt{(n-1)/n}) - \Phi(-x\sqrt{(n-1)/n})]$$

$$\geqq \sqrt{(n-1)/n}[\Phi(y\sqrt{(n-1)/n}) - \Phi(-x\sqrt{(n-1)/n}) \tag{21}$$

$$\geqq (1 - 1/n)[\Phi(y) - \Phi(-x)].$$

The inequality (21.34) is obtained by using the fact that $\Phi_0(ax) \geqq a\Phi_0(x)$ if $0 \leqq a \leqq 1$. Thus, in using $\Phi(y) - \Phi(-x)$ as an approximation to $G_n(y) - G_n(-x)$, the proportional error E, as defined by (18), is not more than the maximum of $\sqrt{(7n-10)/(7n-21)} - 1$ and $1 - (1 - 1/n)$. For $n \geqq 13$, this maximum is $1/n$.

The t- and τ-distributions are closely related. If x has a t-distribution with n d.f., then $y = x\sqrt{(n+1)/(n+x^2)}$ has a τ-distribution with $n + 1$ d.f. Conversely, if y has a τ-distribution with n d.f., then $x = y\sqrt{(n-1)/(n-y^2)}$ has a t-distribution with $n - 1$ d.f. Thus,

$$F_n(x) = G_{n+1}(x\sqrt{(n+1)/(n+x^2)}), \tag{22}$$

$$G_n(x) = F_{n-1}(x\sqrt{(n-1)/(n-x^2)}), \tag{23}$$

where $F_n(x)$ and $G_n(x)$ are defined by (10) and (19). New upper and lower bounds for $F_n(y) - F_n(-x)$ and $G_n(y) - G_n(-x)$ can be obtained. For example, by (22), (20.12), and (21.12), we have

$$F_n(y) - F_n(-x) \leqq a_n\sqrt{2\pi n/(n-2)}$$
$$\times [\Phi(y\sqrt{(n-2)/(n+y^2)}) - \Phi(-x\sqrt{(n-2)/(n+x^2)})], \tag{24}$$

$$F_n(y) - F_n(-x)$$
$$\geqq a_n\sqrt{2\pi}[\Phi(y\sqrt{n/(n+y^2)}) - \Phi(-x\sqrt{n/(n+x^2)})]. \tag{25}$$

Similarly, by (23), (12), and (13), we have

$$G_n(y) - G_n(-x) \leqq a_{n-1}\sqrt{2\pi(n-1)/(n-3)}$$
$$\times [\Phi(y\sqrt{(n-3)/(n-y^2)}) - \Phi(-x\sqrt{(n-3)/(n-x^2)})], \tag{26}$$

$$G_n(y) - G_n(-x) \geqq a_{n-1}\sqrt{2\pi(n-1)/n}$$
$$\times [\Phi(y\sqrt{n/(n-y^2)}) - \Phi(-x\sqrt{n/(n-x^2)})]. \tag{27}$$

Obviously, the upper bound (24) is better than that in (12). But neither of the lower bounds in (25) or (13) is better than the other. (See the tables given in Section 5.) The same may be said about the upper and lower bounds for $F_n(y) - F_n(-x)$, given by (16) and (17), and those obtained by (22), (20.13), and (21.13). Further, the upper bound for $G_n(y) - G_n(-x)$, given by (20.12) is better than the one in (26). But neither of the lower bounds given by (21.12) or (27) is better than the other. For example, the RHS of (27) is at most $B(n) = a_{n-1}\sqrt{2\pi(n-1)/n}$, whereas the RHS of (21.12) is close to $A(n) = 2a_{n-1}\sqrt{2\pi}\Phi_0(\sqrt{n-1})$ if $y = x$ is close to $\sqrt{n}$. $A(5) > B(5)$. So in this case, the RHS of (21.12) > RHS of (27). On the other hand, if $y = x = 1$, then the RHS of (27) is $D(n) = 2a_{n-1}\sqrt{2\pi(n-1)/n}\Phi_0(\sqrt{n/(n-1)})$ and the RHS of (21.12) is $C(n) = 2a_{n-1}\sqrt{2\pi}\Phi_0(\sqrt{(n-1)/n})$. $C(5) < D(5)$. Therefore, in this case, the RHS of (21.12) < RHS of (27).

C. *Correlation Coefficients.* Let a sample of size $n + 1$ be drawn from a k-variate ($2 \leqq k$) normal distribution with variates $x_1, x_2, \cdots, x_k$. Let $r_{12\cdot3\cdots k}$ be the sample partial correlation coefficient between x_1 and x_2 after elimination of the remaining variates $x_3, \cdots, x_k$. If, actually, $x_1, \cdots, x_k$ are independently distributed, then the pdf of $r_{12\cdot3\cdots k}$ is (see [2], p 412) $\sqrt{n_k}a'_{n_k}(1 - z^2)^{(n_k-3)/2}$, where $|z| \leqq 1$, and $n_k = n - k + 2$. If $k = 2$, then the corresponding pdf is the pdf of the total correlation coefficient r_{12}. The variance of $r_{12\cdot3\cdots k}$ is $1/n_k$. The cdf of $\sqrt{n_k}r_{12\cdot3\cdots k}$ is $G_{n_k}(x)$, where $G_n(x)$ is given by (19). Therefore, the proportional error in using $\Phi(y) - \Phi(-x)$ to approximate $G_{n_k}(y) - G_{n_k}(-x)$ is not more than $1/n_k$. Hotelling ([7], p. 196) stated: "This [the normal approximation] is in ordinary cases the most convenient method of all [methods for evaluating $G_{n_k}(x)$], but no suitable bound for the error is available at present." The bound we obtain here seems acceptable, at least when n is large compared with k.

D. *χ^2-distribution.* It is well known ([2], p. 251) that if x has a χ^2-distribution with n d.f., then both $x_1 = (x - n)/\sqrt{2n}$ and $x_2 = \sqrt{2x} - \sqrt{2n}$ are asymptotically normally distributed with mean 0 and variance 1. According to R. A. Fisher ([5], p. 81), the distribution of $x_3 = \sqrt{2x} - \sqrt{2n-1}$ tends to normality even "faster." Let $F_n(x)$ be the cdf of any of the x_i's. We tried unsuccessfully to derive both upper and lower bounds for $F_n(y) - F_n(-x)$, similar to those given in (16), (17), (20), and (21). It is not difficult, however, to obtain just a lower bound, in terms of Φ_0, for $F_n(y) - F_n(0)$, where $y \geqq 0$; and an upper bound, also in terms of Φ_0, for $F_n(0) - F_n(-x)$, where $x \geqq 0$. The results are simple, but not sufficient to provide complete mathematical justification for using the normal approximation for the distributions of the x_i's.

In the following we shall show briefly how to derive a lower bound and an upper bound, respectively, for $H_n(y) - H_n(0)$ and $H_n(0) - H_n(-x)$, where $H_n(x)$ is the cdf of $\sqrt{2x} - \sqrt{2m}$, $m = n - 1$, $y, x \geqq 0$, and x has a χ^2-distribution with n d.f. Exactly the same technique may be used to derive the corresponding results for the x_i's. But they are less neat, and therefore will be omitted

If $y \geqq 0$, then

$$H_n(y) - H_n(0) = 2^{-n/2}\Gamma^{-1}(n/2) \int_m^{y_m} x^{n/2-1}e^{-x/2}\,dx$$

(28)

$$= \Gamma^{-1}(n/2)(m/2e)^{m/2} \int_0^y \{[1 + z/\sqrt{2m}]\exp[-z/\sqrt{2m} - (1/2)z^2/2m]\}^m\,dz,$$

where $y_m = (y + \sqrt{2m})^2/2$, $z = \sqrt{2x} - \sqrt{2m}$, and $\Gamma^{-1}(x) = 1/\Gamma(x)$. Using (2) and $\Gamma(n + 1) < \sqrt{2\pi}n^{n+1/2} \exp[-n + \frac{1}{12}n]$ (see [11], p. 352), it can be shown that $\Gamma^{-1}(n/2)(m/2e)^{m/2} \geqq 1/\sqrt{2\pi}$ for $n \geqq 4$. Now, applying (2) to the first factor of the integrand in the second integral of (28), we have, for all $y \geqq 0$ and $n \geqq 4$,

$$H_n(y) - H_n(0) \geqq \Phi_0(y).$$

Similarly, if $0 \leqq x \leqq \sqrt{2m}$, and $n \geqq 4$, then

$$H_n(0) - H_n(-x) \leqq \Phi_0(x).$$

5. Numerical comparisons. We shall now give some numerical comparisons of two known approximations for the cdf $F_n(x)$ of the t-distribution and the upper and lower bounds for $F_n(x)$ given by (12), (13), (24), and (25). One of the approximations is $\Phi(x)$, and the other, suggested by W. A. Hendricks ([6], p. 216), is $\Phi(x_n)$, where

(29) $$x_n = x(a_n\sqrt{2\pi})\sqrt{2n/(2n + x^2)}.$$

In the tables given below, we choose n to be 10, 30, 60, and 120. For each n, values of x are obtained for which $T = F_n(x) - F_n(-x) = 0.50, 0.75, 0.90, 0.95$, and 0.99, respectively. For each pair of x and n, we compute $A = \Phi(x) - \Phi(-x)$ and $A_1 = \Phi(x_n) - \Phi(-x_n)$; U and L, the RHS of (12) and (13); and U_1 and L_1, the RHS of (24) and (25). We then tabulate the differences between the values of T and the bounds and approximations corresponding to the same n and x. For example, $D_A = A - T$, $D_{U_1} = U_1 - T$, etc. We use [13] and [3] to find the values of the Φ and Γ functions.

Various approximations for T have been suggested. (See, for example, [4], [6], [9], and [12].) Only A and A_1 are tabulated here. It seems that A_1 is a better approximation than A, particularly if the d.f. is small. We also point out that $\Phi(y_n) - \Phi(-x_n) >$ RHS of (25), where y_n is the RHS of (29) with x replaced by y, because $0 < a_n\sqrt{2\pi} < 1$, $a\Phi_0(x) \leqq \Phi_0(ax)$, if $0 \leqq a \leqq 1$, and for every fixed x, $n/(n + x^2)$ is an increasing function of n. Computations indicate that $\Phi(y_n) - \Phi(-x_n) <$ RHS of (24), but we are not able to prove it mathematically. (Using the fact that $\Phi_0(ax) \leqq a\Phi_0(x)$ if $1 \leqq a$, it can be shown that the RHS of (24) $> \Phi(y_n') - \Phi(-x_n')$, where x_n' is the RHS of (29) with $2n$ replaced by n and y_n', the same corresponding to y.)

The tables also indicate, among other things, that L is always closer to T than U, (the actual values of U, computed from (12), are greater than 1 when $T = 0.95$ and $n = 10$, and $T = 0.99$ and $n = 10$, 30, and 60,) and that all the bounds and approximations tend monotonically to T. Again, we are not able to prove or disprove these findings.

TABLES

$T = .50$

n	D_U	D_A	D_L	D_{U_1}	D_{A_1}	D_{L_1}
10	.011	.016	.000	.001	.000	−.007
30	.004	.005	.000	.000	.000	−.002
60	.002	.003	.000	.000	.000	−.001
120	.001	.002	.000	.000	.000	.000

$T = .70$

n	D_U	D_A	D_L	D_{U_1}	D_{A_1}	D_{L_1}
10	.033	.026	−.004	.003	.000	−.019
30	.010	.009	−.001	.001	.000	−.006
60	.005	.004	.000	.001	.000	−.003
120	.003	.002	.000	.000	.000	−.001

$T = .90$

n	D_U	D_A	D_L	D_{U_1}	D_{A_1}	D_{L_1}
10	.076	.030	−.023	.016	−.001	−.038
30	.023	.010	−.007	.006	.000	−.012
60	.011	.005	−.003	.003	.000	−.006
120	.006	.003	−.001	.002	.000	−.003

$T = .95$

n	D_U	D_A	D_L	D_{U_1}	D_{A_1}	D_{L_1}
10	.050	.024	−.038	.028	−.002	−.041
30	.027	.009	−.011	.010	.000	−.013
60	.013	.005	−.005	.005	.000	−.007
120	.007	.002	−.003	.003	.000	−.003

$T = .99$

n	D_U	D_A	D_L	D_{U_1}	D_{A_1}	D_{L_1}
10	.010	.009	−.061	.010	−.002	−.039
30	.010	.004	−.019	.010	.000	−.012
60	.010	.002	−.010	.009	.000	−.006
120	.007	.001	−.005	.005	.000	−.003

6. Some problems. The referee of this paper mentioned several sequences of distributions, related to the t-distribution and noncentral t-distribution and known, through numerical investigations, to approach rapidly to normality. He asked whether methods of this paper could be used to derive suitable bounds for the errors in using the normal approximation for these distributions. One of them [8] is the distribution of $\bar{x} + ks$, where $\bar{x}$ and s are the sample mean and standard deviation of a sample drawn from a normal distribution, and the others, suggested by J. W. Tukey and ascribed to him by C. P. Winsor in [12], are the distributions of $(n + 2)t/(n + 2 + t)$ and $(7n/5 + 1)t/((7n/5) + 1 + t)$, where t has a t-distribution with n d.f.

The author thinks the question is a very interesting one and hopes some answer will be found in further research. At the present he wishes to mention that the methods used in this paper do have some other applications. For example, by using the transformations $u(z)$ and $v(z)$, given in the proof of Lemma 3, and similar ones, bounds can be obtained, in terms of the cdf of the χ^2-distribution with m d.f., for the cdf of the F-distribution with m and n d.f. and for the distribution of nx, where x has a B-distribution with m and n d.f. The corresponding upper and lower bounds are close to each other if n is large compared with m.

7. Acknowledgment. The author wishes to thank William Kruskal for his comments and suggestions. Thanks are also due to Prof. H. Hotelling for his critical reading of the original manuscript and to Lorrie D. Sylvester for some help on computing the tables.

REFERENCES

[1] J. T. CHU, "On the distribution of the sample median," *Ann. Math. Stat.*, Vol. 26 (1955), pp. 112–116.
[2] H. CRAMÉR, *Mathematical Methods of Statistics, Princeton University Press*, 1946.
[3] H. T. DAVIS, *Tables of the Higher Mathematical Functions*, Vol. 1, The Principia Press Inc., Bloomington, Indiana, 1933.
[4] W. E. DEMING AND R. T. BIRGE, "On the Statistical theory of errors," *Rev. Modern Physics*, Vol. 6 (1934), pp. 119–161.
[5] R. A. FISHER, *Statistical Methods for Research Workers*, 10th ed., Oliver and Boyd, Edinburgh, 1948.
[6] W. A. HENDRICKS, "An approximation to Student's distribution," *Ann. Math. Stat.*, Vol. 7 (1936), pp. 210–221.
[7] H. HOTELLING, "New light on the correlation coefficient and its transforms," *J. Roy. Stat. Soc.*, Series B, Vol. 15 (1953), pp. 193–232.
[8] W. J. JENNETT AND B. L. WELCH, "The control of proportion defective as judged by a single quality characteristic varying on a continuous scale," *J. Roy. Stat. Soc.*, Suppl., Vol. 6 (1939), pp. 80–88.
[9] "STUDENT," "The probable error of a mean," *Biometrika*, Vol. 6 (1908), pp. 1–25.
[10] W. R. THOMPSON, "On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation," *Ann. Math. Stat.*, Vol. 6 (1935), pp. 214–219.
[11] J. V. USPENSKY, *Introduction to Mathematical Probability*, McGraw-Hill, New York, 1937.
[12] C. P. WINSOR, "Biometry," *Medical Physics*, Vol. 1, edited by O. Glasser, The Year Book Publishers, Inc., Chicago, pp. 89–110.
[13] *Tables of Normal Probability Functions*, National Bureau of Standards, Washington, D. C., 1953.