APPROXIMATIONS TO THE DISTRIBUTION FUNCTION OF SUMS OF INDEPENDENT CHI RANDOM VARIABLES

BY

HERMAN RUBIN and JAMES ZIDEK

TECHNICAL REPORT NO. 106
AUGUST 3, 1965

THIS RESEARCH WAS SPONSORED BY THE ARMY RESEARCH OFFICE, OFFICE OF NAVAL RESEARCH, AND AIR FORCE OFFICE OF SCIENTIFIC RESEARCH BY CONTRACT NO. Nonr-225(52) (NR 342-022)

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

APPROXIMATIONS TO THE DISTRIBUTION FUNCTION OF SUMS OF INDEPENDENT CHI RANDOM VARIABLES

by

Herman Rubin and James Zidek

TECHNICAL REPORT NO. 106

August 3, 1965

PREPARED UNDER CONTRACT Nonr-225(52)

(NR-342-022)

FOR

OFFICE OF NAVAL RESEARCH

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

APPROXIMATIONS TO THE DISTRIBUTION FUNCTION OF SUMS
OF INDEPENDENT CHI RANDOM VARIABLES

by

Herman Rubin and James Zidek

This report is concerned with the problem of approximating the distribution function, denoted by F_n, of

$$(1) \qquad \sum_{i=1}^{n} |X_i| \quad , \qquad n = 1,2,\ldots ,$$

where the $\{X_i\}_{i=1}^{\infty}$ are independent standard normal random variables. In the first of its three sections, we present three well-known approximations adapted to the problem at hand. These are the Edgeworth (cf., Cramér [1], page 228), the Cramér [2], and a saddlepoint approximation. In addition, a second very effective saddlepoint approximation is derived which, as far as is known by the authors, has not appeared previously in the literature. The second section is devoted to a discussion of the problem of calculating the moment generating function, say M, of F_1 for complex values of its argument. This is of concern, for example, in the problem of numerically inverting M^n in order to obtain the values of F_n corresponding to prescribed values of its argument. Finally, in section three, we tabulate calculations which provide a comparison of the approximations given in section one for the cases $n = 10$ and 40. To facilitate this comparison, the corresponding values (to eight decimal places) of F_{10} and F_{40} are given.

1. Four Approximation to F_n.

Let f denote the probability density function for the distribution of $|X_1|$. Then

(2) $$f(x) = \sqrt{\frac{2}{\pi}}\, e^{-x^2/2} \quad , \qquad x \geq 0$$

$$= 0 \quad , \qquad x < 0 \;.$$

Also, if $M(z) = E(e^{zX_1})$, it is easily shown that

(3) $$M(z) = 2e^{z^2/2}\,\Phi(z) \quad ,$$

where z is complex and

(4) $$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-v^2/2}\, dv \;.$$

Let α_r $(r = 1,2,\ldots)$ and σ denote the r^{th} cumulant and standard deviation, respectively, of $|X_1|$. Then

(5)
$$\alpha_1 = \mu_1 \simeq 0.79788\ 45608\ 03$$
$$\alpha_2 = 1-\mu_1^2 \simeq 0.36338\ 02276\ 32$$
$$\alpha_3 = \mu_1(\mu_1^2 - \alpha_2) \simeq 0.21801\ 36141\ 45$$
$$\alpha_4 = 2\mu_1^2(2-3\mu_1^2) \simeq 0.11477\ 0682054$$
$$\alpha_5 = \mu_1(3-20\mu_1^2 + 24\mu_1^4) \simeq 0.00443\ 76884\ 6262$$
$$\sigma \cong 0.60281\ 02749\ 89 \;.$$

Now, using the Edgeworth expansion of F_n we obtain, as $n \to \infty$,

$$1-F_n(x) \sim 1-\Phi(w_n) + \frac{1}{\sqrt{n}} \cdot \frac{\lambda_3}{3!} \cdot \Phi^{(3)}(w_n)$$

$$(6) \qquad -\frac{1}{n}\left[\frac{\lambda_4}{4!} \cdot \Phi^{(4)}(w_n) + \frac{10}{6!} \cdot \lambda_3^2 \cdot \Phi^{(6)}(w_n)\right]$$

$$+\left(\frac{1}{\sqrt{n}}\right)^3 \left[\frac{\lambda_5}{5!} \cdot \Phi^{(5)}(w_n) + \frac{35}{7!} \cdot \lambda_3\lambda_4 \Phi^{(7)}(w_n) + \frac{280}{9!} \cdot \lambda_3^3 \cdot \Phi^{(9)}(w_n)\right] .$$

where

$$(7) \qquad w_n = (x - n\mu_1)/(\sqrt{n}\sigma)$$

$$\lambda_n = \alpha_n/\sigma^n$$

$(n = 1,2,\ldots)$ and $\Phi^{(r)}$ denotes the r^{th} derivative of $\Phi (r = 0,1,2,\ldots)$. From equations (5) and (6),

$$1-F_n(x) \sim 1-\Phi + \frac{1}{\sqrt{n}} \cdot (.16587\ 86244\ 05) \cdot \Phi^{(3)}$$

$$(8) \qquad - \frac{1}{n}\left[(.03621\ 57209\ 836)\Phi^{(4)} + (.01375\ 78590\ 173)\Phi^{(6)}\right]$$

$$-\left(\frac{1}{\sqrt{n}}\right)^3 \left[(.00046\ 45926\ 47496)\Phi^{(5)} - (.00600\ 74139\ 7860)\Phi^{(7)}\right.$$

$$\left. - (.00076\ 07115\ 7618)\Phi^{(9)}\right] ,$$

where, for convenience, the argument, w_n, has been suppressed on the right of this asymptotic equality.

The Cramér approximation gives

$$
(9) \qquad 1 - F_n(x) \sim \begin{cases} (1-\Phi(w_n)) \exp\left[\frac{w_n^3}{\sqrt{n}} \lambda\left(\frac{w_n}{\sqrt{n}}\right)\right] & x \geq n\mu_1 \\ \Phi(w_n) \exp\left[\frac{w_n^3}{\sqrt{n}} \lambda\left(\frac{w_n}{\sqrt{n}}\right)\right] & x < n\mu_1 \end{cases} ,
$$

where, if the series

$$
(10) \qquad \sigma z = \sum_{r=2}^{\infty} \frac{\alpha_r}{(r-1)!} h^r
$$

is inverted to obtain h as a power series, $h(z)$, in z,

$$
(11) \qquad \begin{aligned} \lambda(z) &= \frac{\sqrt{n}}{z^3} \left\{ \frac{z^2}{2} - \sum_{r=2}^{\infty} \frac{(r-1)\alpha_r}{r!} h^r(z) \right\} \\ &= \frac{\lambda_3}{6} + \left(\frac{\lambda_4}{24} - \frac{\lambda_3^2}{8}\right) z + \left(\frac{\lambda_5}{120} - \frac{\lambda_3\lambda_4}{12} + \frac{\lambda_3^3}{8}\right) z^2 + \ldots \\ &\simeq 0.16587\ 86244\ 05 - 0.08760\ 50101\ 719\ z \\ &\qquad + 0.05161\ 09001\ 629\ z^2 + \ldots \quad . \end{aligned}
$$

We now derive two saddlepoints approximations to F_n using the Gurland [3] inversion formula which asserts

$$
(12) \qquad 1 - F_n(x) = \frac{1}{2} + \lim_{T\to\infty} \lim_{\eta\to 0+} \left\{ \int_{\eta}^{T} + \int_{-T}^{-\eta} \right\} \frac{e^{-ixt}}{2\pi it} M^n(it)\, dt,
$$

where $i = \sqrt{-1}$. Equation (12) is equivalent to

$$
(13) \qquad 1 - F_n(x) = \tfrac{1}{2}(1 - \mathrm{sign}(c)) + \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left[-(c+iu)x+n \log M(c+iu)\right] \frac{du}{c+iu}
$$

for every real number c, and this is the form in which we shall make use of it. Since,

$$
\frac{d}{dh} M(h) = \sqrt{\frac{2}{\pi}}\,[1 + h\Phi(h)/\varphi(h)] \quad ,
$$

where $\varphi(h) = \Phi^{(1)}(h)$, the critical point, say $c_n (n=1,2,\ldots)$ of the integrand is the solution of the equation

$$
(14) \qquad \varphi(h)/\Phi(h) + h = x/n \quad .
$$

Equation (14) can be solved numerically using Newton's method with initial approximation $c_{n0} = x/n$, and k^{th} iterate

$$
(15) \qquad c_{nk} = c_{n,k-1} - [g(c_{n,k-1}) + c_{n,k-1} - x/n] \,/\, [1-c_{n,k-1}\, g(c_{n,k-1}) - g^2(c_{n,k-1})] ,
$$

$(k = 1,2,\ldots)$, where $g(x) = \varphi(x)/\Phi(x)$. In Table 1, we give the values of c_n which were obtained by this method in the case $n = 20$ for various values of x. For simplicity we shall hereafter write $c = c_n$.

TABLE 1

SOLUTIONS OF EQUATION (14) OBTAINED USING NEWTON'S METHOD

x	c
2	-9.80191
8	-1.77896
10	-1.13115
14	-0.29396
16	0.00581
20	0.48106
22	0.67787
26	1.01960

Now, for values of $|u|$ in some neighbourhood of the origin

(16) $$-(c + iu)x + n \log M(c + iu) = (-cx + n \log M(c)) - nu^2 \psi^{(2)}(c)/2 + n \sum_{r=3}^{\infty} \frac{\psi^{(r)}(c)}{r!} (iu)^r ,$$

where $\psi(z) = \log M(z)$. Let

(17) $$\begin{cases} \sigma^* = [\psi^{(2)}(c)]^{+1/2} \\ b_r = \psi^{(r)}(c)\, i^r / (r! \cdot \sigma^{*r}) \\ a_r = (-i)^r / (c\sigma^*)^r \end{cases}$$

(18) $$K(c,n,x) = \frac{1}{\sqrt{2\pi}} \exp[-cx + n\,\psi(c)] ,$$

(19) $$I(x,n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp[-(c + iu) + n\psi(c + iu)] \frac{du}{c+iu} ,$$

$$(r = 1,2,\ldots;\ n = 1,2,\ldots) .$$

With this notation,

$$(20)\qquad \left(1 + \frac{iy}{c\sqrt{n}\,\sigma^*}\right)^{-1} \exp\left[n \sum_{r=3}^{\infty} b_r \left(\frac{y}{\sqrt{n}}\right)^r\right] = \sum_{m=0}^{\infty} d_m(y) \left(\frac{1}{\sqrt{n}}\right)^m ,$$

where

$$d_0(y) = 1$$

$$(21)\qquad d_2(y) = a_2 y^2 + (b_4 + a_1 b_3) y^4 + \frac{1}{2} b_3^2\, y^6$$

$$d_4(y) = a_4 y^4 + (b_6 + a_1 b_5 + a_2 b_4 + a_3 b_3) y^6 + (\tfrac{1}{2} b_4^2 + b_3 b_5 + a_1 b_3\, b_4 + \tfrac{1}{2}\, a_2 b_3^2) y^8 + (\tfrac{1}{2} b_4 b_3^2 + \tfrac{1}{6}\, a_1 b_3^3) y^{10} + \tfrac{1}{24}\, b_3^4 y^{12} ,$$

and in general $d_{2k-1}(y)$ is an odd polynomial in y, $d_{2k}(y)$ an even polynomial in y, $k = 1,2, \ldots$.

According to an argument given by Daniels [4],

$$(22)\qquad I(x,n) \sim K(c,n,x) \,/\, (c\sqrt{n}\,\sigma^*) \sum_{m=0}^{\infty} d_{2m} \left(\frac{1}{\sqrt{n}}\right)^{2m} ,$$

where

$$d_0 = 1$$

$$d_2 = a_2 + 3(b_4 + a_1 b_3) + 15\ b_3^2\ /2$$

(23)
$$d_4 = 3a_4 + 15(b_6 + a_1 b_5 + a_2 b_4 + a_3 b_3)$$
$$+ 105(\tfrac{1}{2}\ b_4^2 + b_3 b_5 + a_1 b_3 b_4 + \tfrac{1}{2}\ a_2 b_3^2)$$
$$+ 945(\tfrac{1}{2}\ b_4 b_3^2 + \tfrac{1}{6}\ a_1 b_3^3) + 10395\ b_3^4\ /24$$

and, in general,

(24)
$$d_{2m} = \int_{-\infty}^{\infty} \varphi(y)\ d_{2m}(y)\ dy$$

$m = 0,1,2,\dots$. After simplification, equations (23) become, with $\psi^{(r)}(c) = U_r$, $r = 1,2,\dots,$

$$d_0 = 1$$

$$d_2 = -1/(\sigma^* c^2) + [\ \tfrac{1}{8}\ U_4 - \tfrac{1}{2c}\ U_3]/\ \sigma^{*2}$$
$$-\ 5\ U_3^2\ /\ (24\sigma^{*3})$$

(25)
$$d_4 = 3/(\sigma^{*2} c^4)$$
$$+ \frac{5}{8\sigma^{*3}}\ (\ -\ \frac{1}{30}\ U_6 + \frac{1}{5c}\ U_5 - \frac{1}{c^2}\ U_4 + \frac{4}{c^3}\ U_3)$$
$$+ \frac{35}{24\sigma^{*4}}\ (\frac{1}{16}\ U_4^2 + \frac{1}{10}\ U_3 U_5 - \frac{1}{2c}\ U_3 U_4 + \frac{1}{c^2}\ U_3^2)$$
$$+ \frac{35}{16\sigma^{*5}}\ (-\ \frac{1}{4}\ U_4 U_3^2 + \frac{1}{3c}\ U_3^3) + \frac{385}{48\sigma^{*6}}\ U_3^4 \quad .$$

Observe that equation (14) implies

$$\text{(26)} \qquad K(c,n,x) = (e^{-xc}/\sqrt{2\pi})\left[\sqrt{\frac{2}{\pi}}\left(\frac{x}{n} - c\right)^{-1}\right]^{n}$$

and (after considerable simplification) on letting $u = x/n$,

$$\text{(27)} \qquad \begin{aligned}
U_1 &= u \\
U_2 &= cu + 1-u^2 = \sigma*^2 \\
U_3 &= uc^2 + c(1-3u^2) - u(1 - 2u^2) \\
U_4 &= uc^3 + c^2\ (1-7u^2) - uc\ (5-12u^2) + u^2(4-6u^2) \\
U_5 &= uc^4 + c^3(1-15u^2) - uc^2(16-50u^2) \\
&\quad - c(3-35u^2 + 60u^4) + u(3-20u^2 + 24u^4) \\
U_6 &= uc^5 + c^4(1-31u^2) - uc^3(42-180u^2) \\
&\quad - c^2(13-191u^2 + 390u^4) + cu(41-270u^2 + 360u^4) \\
&\quad - u^2(28-120u^2 + 120u^4) \quad .
\end{aligned}$$

Thus,

$$\text{(28)} \qquad \begin{aligned}
1 - F_n(x) &\sim \frac{1}{2}(1 - \text{sign}\ (c_n)) \\
&\quad + \{e^{-xc_n}/(\sqrt{2\pi}\ c\sqrt{nU_2})\}\left\{\sqrt{\frac{2}{\pi}}\left(\frac{x}{n} - c_n\right)^{-1}\right\}^{n}\left[1 + \frac{d_2}{n} + \frac{d_4}{n^2}\right] ,
\end{aligned}$$

where $c_n (n = 1,2,\ldots)$ is obtained from equation (14), d_2 and d_4 from equations (25) with the aid of equations (27).

Let us return now to equation (19) and by a somewhat more delicate argument obtain a second saddlepoint approximation to $1 - F_n$. From this equation together with equations (16), (17), (18) and the argument of Daniels which was previously referred to, we have on letting $\rho = \rho_n = c_n\sqrt{n\, U_2}$, $b_r' = b_r/i^r$ $(r = 0, 1, 2, \ldots)$,

$$(29) \qquad 1 - F_n(x) \sim \frac{1}{2}(1 - \text{sign}(c)) + K(c,n,x) \int_{-\infty}^{\infty} \frac{\varphi(u)}{\rho+iu} \exp\left[\sum_{r=3}^{\infty} b_r' \frac{(iu)^r}{(\sqrt{n})^{r-2}}\right] du .$$

Now

$$(30) \qquad \exp\left[\sum_{r=3}^{\infty} b_r' \frac{(iu)^r}{(\sqrt{n})^{r-2}}\right] = \sum_{k=0}^{\infty} g_k(iu) \left(\frac{1}{\sqrt{n}}\right)^k ,$$

where

$$g_0(y) = 1$$

$$g_1(y) = b_3' y^3$$

$$(31) \qquad g_2(y) = b_4' y^4 + \frac{1}{2} b_3'^2 y^6$$

$$g_3(y) = b_5' y^5 + b_3' b_4' y^7 + \frac{1}{6} b_3'^3 y^9$$

$$g_4(y) = b_6' y^6 + \left(\frac{1}{2} b_4'^2 + b_3' b_5'\right) y^8 + \frac{1}{2} b_3'^2 b_4' y^{10} + \frac{1}{24} b_3'^4 y^{12} .$$

Define $Q_k(\rho)$ $(k = 0, 1, \ldots)$ by

$$(32) \qquad Q_k(\rho) = \int_{-\infty}^{\infty} \frac{\varphi(u)}{\rho+iu} (iu)^k \, du .$$

Then

$$Q_0(\rho) = \frac{1}{\varphi(\rho)} \left[\frac{1}{2}(1 + \text{sign}(c)) - \Phi(\rho)\right] \tag{33}$$

and the $\{Q_k\}$ satisfy the recurrence formulae,

$$\begin{aligned} Q_{2k-1}(\rho) &= (-1)^{k-1} (2k-3)(2k-5) \quad \ldots \; 3.1 - \rho Q_{2k-2}(\rho) \\ Q_{2k}(\rho) &= -\rho Q_{2k-1}(\rho) \quad , \qquad (k = 1,2,\ldots) \quad . \end{aligned} \tag{34}$$

Thus,

$$\begin{aligned} 1 - F_n(x) &\sim \frac{1}{2} (1 - \text{sign}(c)) \\ &+ (e^{-xc}/\sqrt{2\pi}) \left[\sqrt{\frac{2}{\pi}} \left(\frac{x}{n} - c\right)^{-1}\right]^n \sum_{k=0}^{4} g_k(\rho) \left(\frac{1}{\sqrt{n}} \right)^k \quad , \end{aligned} \tag{35}$$

where

$$\begin{aligned} g_0(\rho) &= Q_0(\rho) \\ g_1(\rho) &= \frac{1}{6} U_3 \left(\sqrt{U_2}\right)^{-3} \cdot Q_3(\rho) \\ g_2(\rho) &= \frac{1}{24} U_4 U_2^{-2} Q_4(\rho) + \frac{1}{72} U_3^2 U_2^{-3} Q_6(\rho) \\ g_3(\rho) &= \frac{1}{120} U_5 U_2^{-5} Q_5(\rho) + \frac{1}{144} U_3 U_4 \left(\sqrt{U_1}\right)^{-7} Q_7(\rho) \\ &\quad + \frac{1}{1296} U_3^3 \left(\sqrt{U_2}\right)^{-9} Q_9(\rho) \\ g_4(\rho) &= \frac{1}{720} U_6 U_2^{-3} Q_6(\rho) + \left(\frac{1}{1152} U_4^2 + \frac{1}{720} U_3 U_5\right) U_2^{-4} Q_8(\rho) \\ &\quad + \frac{1}{1728} U_3^2 U_4 U_2^{-5} Q_{10}(\rho) + \frac{1}{31104} U_3^4 U_2^{-6} Q_{12}(\rho) \quad , \end{aligned} \tag{36}$$

where the $\{Q_i\}_3^{12}$ are readily obtained using recurrence relations (34), Q_0 is obtained from equation (33) and the $\{U_i\}_{i=1}^5$ are given by equation (27).

2. <u>The Calculation of</u> $M(z)$.

We consider now, the problem of computing $M(z)$ as defined in equation (3). This computation is necessary to the numerical inversion procedure by means of which one obtains, using equation (13), exact values of F_n for specified values of its argument. The difficulty here is that excessive and uncontrollable round off errors occur, even for moderate values of $|z|$ when $\Phi(z)$ is computed using its Taylor expansion. Some examples illustrating this remark are presented below in Table 2.

TABLE 2

VALUES OF $M(z)$ AND THOSE $(\tilde{M}(z))$ COMPUTED USING THE TAYLOR EXPANSION OF Φ FOR SEVERAL VALUES OF z

z	$Re(z)$	-1.5	-1.5
	$Im(z)$	5	.5
$M(z)$	$Re\{M(z)\}$	0.04894 653	0.39447 03827 73
	$Im\{M(z)\}$	0.14998 03	0.08722 46093 6
$\tilde{M}(z)$	$Re\{\tilde{M}(z)\}$	0.04894 631	0.39447 03827 73
	$Im\{\tilde{M}(z)\}$	0.14997 98	0.08722 46093 5

z	$Re(z)$	1	1.5
	$Im(z)$	2.5	5
$M(z)$	$Re\{M(z)\}$	0.27320 08274 0	-0.04893 86
	$Im\{M(z)\}$	0.36194 01413 0	0.15000 18
$\tilde{M}(z)$	$Re\{\tilde{M}(z)\}$	0.27320 08274 2	-0.04893 84
	$Im\{\tilde{M}(z)\}$	0.36194 01412 8	0.15000 14

The solution to this problem which we propose here involves the complex forms of Shenton's [5] continued fraction for small values of $|z|$ and Laplace's continued fraction (cf., Kendall and Stuart [6], p. 138) otherwise. Since

(37) $$M(z) = 2e^{z^2/2} - M(-z) \quad ,$$

there is no loss of generality in considering the case $R(z) \geq 0$, alone. The notation

(38) $$\cfrac{a_1}{b_1 + \cfrac{a_2}{b_2 + \cfrac{a_3}{b_3 + \dots}}} = \frac{a_1}{b_1+}\;\frac{a_2}{b_2+}\;\frac{a_3}{b_3+}\;\cdots$$

will be adopted here.

Using the Shenton fraction we obtain

(39) $$M(z) = e^{z^2/2} + \sqrt{\frac{2}{\pi}}\left(\frac{z}{1-}\;\frac{z^2}{3+}\;\frac{2z^2}{5-}\;\frac{3z^2}{7+}\;\cdots\right), \quad Re(z) > 0$$

and using that of Laplace

(40) $$M(z) = 2e^{z^2/2} + \sqrt{\frac{2}{\pi}}\left(\frac{1}{z+}\;\frac{1}{z+}\;\frac{2}{z+}\;\frac{3}{z+}\;\cdots\right), \quad Re(z) > 0 \quad .$$

Equation (39) may be rewritten with $t = 1/z^2$ as

(41) $$M(z) = e^{z^2/2} + \sqrt{\frac{2}{\pi}}\left(\frac{z}{1-}\;\frac{1/3}{t+}\;\frac{2/(3.5)}{1-}\;\frac{3/(5.7)}{t+}\;\cdots\right), \quad Re(z) > 0 \quad .$$

The $2n^{th}$ approximant to the continued fraction in this expression is

(42) $$\frac{z}{1-}\;\frac{1/3}{t+}\;\cdots\;\frac{(2n-1)/[(4n-3)(4n-1)]}{t}$$

$(n = 1,2,\dots)$

while the remainder satisfies

(43) $$\frac{2n/[(4n-1)(4n+1)]}{1-} \quad \frac{(2n+1)/[(4n+1)(4n+3)]}{t+} \cdots$$

$$\sim \frac{1/8n}{1-} \quad \frac{1/8n}{t+} \quad \frac{1/8n}{1-} \quad \cdots$$

as $n \to \infty$. The continued fraction appearing on the right of (43) represents the function $u(t)$ which satisfies the equation

(44) $$u = a_n/[1 - \frac{a_n}{t+u}] \quad ,$$

that is,

(45) $$u(t) = (a_n - \frac{t}{2}) + \sqrt{(\frac{t}{2})^2 + a_n^2} \quad ,$$

where $a_n = 1/8n$.

Let

(46)
$$R_n(t) = \mathrm{Re}[(\frac{t}{2})^2 + a_n^2]$$
$$I_n(t) = \mathrm{Im}[(\frac{t}{2})^2 + a_n^2]$$
$$R_n'(t) = [\frac{1}{2}(R_n(t) + \{R_n^2(t) + I_n^2(t)\}^{\frac{1}{2}})]^{\frac{1}{2}} \quad ,$$

$(n = 1,2,\ldots)$.

Then (cf. Ahlfors [7], p. 3),

$$[R_n^2(t) + I_n^2(t)]^{\frac{1}{2}} = \pm [R_n'(t) + \frac{i}{2} I_n(t)/R_n'(t)]$$

if $R_n'(t) \neq 0$ and 0 otherwise. The square root in this equation has branch points at $\pm 2ai$, and the function obtained by choosing either sign is a branch of the square root. Rather than fix the sign, we take

$$(47)\quad [R_n^2(t)+I_n^2(t)]^{\frac{1}{2}} = \begin{cases} \text{sign}\{Re(t)\}[R_n'(t) + \frac{i}{2}I_n(t)/R_n'(t)], & Re(t) \neq 0, R_n'(t) \neq 0, \\ R_n'(t) + \frac{i}{2}I_n(t)/R_n'(t), & Re(t) = 0,\ R_n'(t) \neq 0 \\ = 0 & R_n'(t) = 0 \end{cases}$$

to obtain a continuous approximation to the continued fraction.

Using equation (45) we obtain the following asymptotic approximation to the continued fraction of equation (41):

$$(48)\qquad \frac{z}{1-}\ \frac{1}{3t+}\ \frac{2}{5-}\ \cdots\ (2n-1)/[(4n-1)\{\tfrac{t}{2} + a_n + \sqrt{(\tfrac{t}{2})^2 + a_n^2}\}],\quad Re(z) > 0,\quad (n = 1,2,\ldots)$$

which, as it happens, gives satisfactory results in the case $Re(z) = 0$.

By a similar argument we obtain the following approximation to the continued fraction of equation (40):

$$(49)\qquad \frac{1}{z+}\ \frac{1}{z+}\ \frac{2}{z+}\ \cdots\ \frac{2(n-1)}{z+\sqrt{z^2+4n}},\qquad Re(z) > 0,\quad (n = 2,3,\ldots).$$

We shall make, in return, a modification of this last approximation. Observe that if

$$(50)\qquad \begin{aligned} Hh_n(z) &= \int_0^\infty \frac{t^{n-1}}{(n-1)!}\, e^{-\frac{1}{2}(t+z)^2}\,dt, \qquad n = 1,2,\ldots,\quad Re(z) > 0 \\ Hh_0(z) &= e^{-z^2/2}\ , \end{aligned}$$

then

$$(51)\qquad Hh_n(z) = (n+1)\, Hh_{n+2}(z) + zHh_{n+1}(z) \qquad (n = 0,1,2,\ldots) .$$

The substitution, $Q_n = Hh_{n-1}/Hh_n$, yields

$$(52)\qquad Q_n(z) = z + n/Q_{n+1}(z) \quad , \qquad n = 1,2, \ldots .$$

Thus

$$(53)\qquad \begin{aligned} Q_1(z) &= \varphi(z)/(1 - \Phi(z)) \\ &= z + 1/Q_2(z) \\ &= z + \frac{1}{z+}\ \frac{2}{z+} \cdots \frac{(n-1)}{Q_n(z)} \quad , \end{aligned} \qquad (n = 2,3,\ldots) .$$

In (49) replace $4n$ by a_n', where the $\{a_n'\}_{n=2}^{\infty}$ are chosen so that the approximation $\frac{1}{2}(z + \sqrt{z^2 + a_n'})$ to $Q_n(z)$ is exact at $z = 0$, that is,

$$(54)\qquad a_n' = 8\, \Gamma^2(\tfrac{n+1}{2}) \,/\, \Gamma^2(\tfrac{n}{2}), \qquad n = 2,3,\ldots ,$$

where Γ denotes the Gamma-function.

Let $R_n(z)$ and $I_n(z)$ denote $\mathrm{Re}(z^2 + a_n')$ and $\mathrm{Im}(z^2 + a_n')$, respectively. Then, if

$$(55)\qquad R_{kn}(z) = [\tfrac{1}{2}((-1)^{k-1} R_n(z) + \{R_n^2(z) + I_n^2(z)\}^{\frac{1}{2}})]^{\frac{1}{2}} \quad ,$$

$k = 1,2$; $n = 2,3,\ldots,$ we take, as a result of the same considerations which gave rise to equation (47),

$$
(56) \quad [R_n(z)+iI_n(z)]^{\frac{1}{2}} = \begin{cases} \operatorname{sign}\{I_n(z)\}[I_n(z)/\{2R_{2n}(z))\} + iR_{2n}(z)] , \\ \qquad\qquad R_n(z)<0,\ I_n(z)\neq 0,\ R_{2n}(z)\neq 0, \\ I_n(z)/\{2R_{2n}(z)\}+ iR_{2n}(z) , \\ \qquad\qquad R_n(z)< 0,\ I_n(z)=0,\ R_{2n}(z)\neq 0, \\ R_{1n}(z) + iI_n(z)/\{2R_{1n}(z)\} , \quad R_n(z)> 0,\ R_{1n}(z)\neq 0, \\ 0 \qquad R_{2n}(z) = 0 \text{ or } R_{1n}(z) = 0 . \end{cases}
$$

In summary, we have obtained the following approximations to $M(z)$,

$$
(57) \qquad M(z) \sim e^{z^2/2} + \sqrt{\frac{2}{\pi}} (\frac{z}{1-} \frac{1}{3t+} \frac{2}{5-} \cdots
$$

$$
(2n-1)/[(4n-1)\{\frac{t}{2} + \frac{1}{8n} + \sqrt{(\frac{t}{2})^2 + \frac{1}{64n^2}}\}], \quad \operatorname{Re}(z) \geq 0
$$

and

$$
(58) \qquad M(z) \sim 2e^{z^2/2} + \sqrt{\frac{2}{\pi}} (\frac{1}{z+} \frac{1}{z+} \frac{2}{z+} \cdots
$$

$$
2(n-1)/[z + \sqrt{z^2 + 8\Gamma^2(\frac{n+1}{2} / \Gamma^2(\frac{n}{2})}]), \quad \operatorname{Re}(z) > 0
$$

where the determination of the square roots involved is as given in equations (47) and (56).

These approximations were computed and compared on a grid for z comprised of 231 points spread over the region.

(59) $$D = \{z: -5 \leq \mathrm{Re}(z) \leq 5, \qquad 0 \leq \mathrm{Im}(z) \leq 5\} \ .$$

On the basis of this, Table 3 has been completed and lists for all subregions of D, a suitable approximation to M. The approximations given in equations (57) and (58) are denoted by S_r and C_r $(r=1,2,\ldots)$, respectively, where in each case, r designates a value of n sufficiently large as to give an accuracy of at least 10 significant figures for that subregion.

TABLE 3

APPROXIMATIONS TO $M(z)$, $z \in \underset{\smile}{D}$ WHICH ARE ACCURATE TO AT LEAST TEN SIGNIFICANT PLACES

REGIONS R(z)	REGIONS I(z)	APPROXIMATION
(3.5, 5]	[0, 5]	C_{10}
(-4.25, -3.75]	(4.25, 5]	C_{10}
(-4.75, -4.25]	(3.25, 5]	C_{10}
[-5, -4.75]	(2.25, 5]	C_{10}
(2.25, 5]	(0, 5]	C_{20}
(-2.25, -1.75]	(3.75, 5]	C_{20}
(-2.75, -2.25]	(2.25, 5]	C_{20}
(-3.75, -2.75]	[0, 5]	C_{20}
(-4.25, -3.75]	[0, 4.25]	C_{20}
(-4.75, -4.25]	[0, 3.25]	C_{20}
[-5 -4.75]	[0, 2.25]	C_{20}
(-2.75, -2.25]	[0, 2.25]	C_{30}
(.75, 1.75]	[0, .25]	S_5
(-.25, .75]	[0, 1.75]	S_5
(-1.25, -,25]	[0, .75]	S_5
(-1.75, -1.25]	[0, .25]	S_5
(.75, 1.75]	(.25, 3.25]	S_{10}
(-.25, .75]	(1.75, 3.25]	S_{10}
(-1.25, -.25]	(.75, 2.75]	S_{10}
(-1.75, -1.25]	(.25, 2.25]	S_{10}
(-2.25, -1.75]	[0, 1.25]	S_{10}
(1.75, 2.25]	[0, 2.75]	S_{10}
(1.75, 2.25]	(2.75, 5]	S_{15}
(.75, 1.75]	(3.25, 5]	S_{15}
(-.25, .75]	(3.25, 5]	S_{15}
(-1.25, -.25]	(2.75, 5]	S_{15}
(-1.75, -1.25]	(2.25, 5]	S_{15}
(-2.25, -1.75]	(1.25, 3.75]	S_{15}

3. An Evaluation of Proposed Approximations.

Table 4, below, summarizes illustrative calculations which were carried out using the four approximations to $1 - F_n(x)$ discussed in Section 1 for the cases $n = 10$ and 40, for various values of x. Also given are the corresponding exact values of $1 - F_n$ which serve as a basis for the comparisons of the four alternatives.

An examination of this table reveals that the saddlepoint approximation given in (35) yields substantially better results in all cases considered than those provided by any of the other three methods.

TABLE 4

A COMPARISON OF VARIOUS APPROXIMATIONS TO $1 - F_{40}$ */

$n = 40$

x	$1 - F_{40}(x)$	EDGEWORTH	CRAMÉR	SADDLEPOINT 1 (See (28)	SADDLEPOINT 2 See (35)
45	.00068290	.00029954	.00067782	.00074527	.00069266
		.00061180		.00066672	.00068382
		.00068880		.00069088	.00068288
		.00069780			.00068290
					.00068290
43	.00303962	.00182203	.00309701	.00340155	.00308900
		.00296077		.00291545	.00304341
		.00305561		.00311715	.00303959
		.00304550			.00303964
					.00303961
41	.01140206	.00858968	.01159390	.01325463	.01161291
		.01145237		.01051556	.01141441
		.01140474		.01215552	.01140201
		.01134306			.01140212
					.01140200

*/ Whenever several entries are tabulated for a given approximation and value of the argument, they represent the sequence of approximations obtained by taking for the k^{th} $(1 \leq k \leq 5)$, the sum of the first k terms of the original approximation as given in Section 1.

TABLE 4 (contd)

x	$1 - F_{40}(x)$	EDGEWORTH	CRAMÉR	SADDLEPOINT 1 See (28)	SADDLEPOINT 2 See (35)
39	.03587479	.03156626	.03656039	.04436218	.03662763
		.03613245		.02974891	.03590587
		.03584150		.04371489	.03587506
		.03581522			.03587495
					.03587461
24.76	.97458917	.96972849	.97534720	.97012208	.97527441
		.97426376		.97726335	.97457312
		.97454881		.97152270	.97459025
		.97451806			.97458912
					.97458910
22.76	.99465351	.99183393	.99448716	.99406009	.99479098
		.99462428		.99487737	.99464874
		.99466352		.99443831	.99465369
		.99460327			.99465350
					.99465351
20.76	.99925955	.99828323	.99927448	.99920462	.99927765
		.99937764		.99927617	.99925872
		.99937764		.99924173	.99925957
		.99927277			.99925956
					.99925955
18.76	.99993726	.99972032	.99993726	.99993403	.99993872
		1.00001672		.99993785	.99993717
		.99994153		.99993588	.99993726
		.99995055			.99993726
					.99993726

(contd)

TABLE 4 (contd)

n = 10

x	$1 - F_{40}(x)$	EDGEWORTH	CRAMÉR	SADDLEPOINT 1 See (28)	SADDLEPOINT 2 See (35)
13.9	.00270601	.00094763	.00323707	.00303538	.00277470
		.00240150		.00259507	.00271811
		.00280015		.00278765	.00270564
		.00281818			.00270622
					.00270598
12.4	.01612192	.01019966	.01697658	.01916375	.01664558
		.01641306		.01448249	.01618628
		.01610240		.01787318	.01612076
		.01557687			.01612313
					.01612132
10.4	.10723570	.10202246	.11163857	.15429326	.11196924
		.10775032		.04106193	.10749609
		.10707194		.28189662	.10725028
		.10843852			.10724093
					.10723031
5.5	.91051790	.90326388	.91638498	.88517181	.91603617
		.90947218		.93776274	.91048869
		.91023710		.83431773	.91057402
		.91154120			.91052145
					.91052099
3.6	.99589745	.98919333	.99595978	.99554712	.99616174
		.99559060		.99602163	.99589380
		.99594482		.99381403	.99590090
		.99541103			.99589797
					.99589985

REFERENCES

[1] Cramer, H., Mathematical Methods of Statistics. Princeton: Princeton University Press, 1945.

[2] __________, Sur un nouveau théorème - limite de la théorie des probabilites: Actualités Scientifique et Industrielles, No. 736, Hermann et Cie , Paris, 1938, p. 12.

[3] Gurland, J., Inversion formulae for the distribution of ratios, Ann. Math. Stat., vol. 19(1948), pp. 228-237.

[4] Daniels, H. E., Saddlepoint approximations in statistics, Ann. Math. Stat., vol. 25(1954), pp. 631-650.

[5] Shenton, L. R., Inequalities for the normal integral, including a new continued fraction. Biom., 41, pp. 177-189.

[6] Kendall, M. G. and Stuart, A., The Advanced Theory of Statistics, Volume I, London: C. Griffin and Co., 1958.

[7] Ahlfors, L. V., Complex Analysis, New York: McGraw-Hill Book Company, Inc., 1953.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Department of Statistics Stanford University Stanford, California	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

Approximations to the Distribution Function of Sums of Independent Chi Random Variables

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical Report

5. AUTHOR(S) (Last name, first name, initial)

Rubin, Herman and Zidek, James, V.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
August 3, 1965	23	7

8a. CONTRACT OR GRANT NO. Nonr-225(52) b. PROJECT NO. NR-342-022 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 106 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES

Releaseable without limitations on dissemination

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Logistics and Mathematical Sciences Branch Office of Naval Research Washington D. C. 20360

13. ABSTRACT

This report is concerned with the problem of approximating the distribution function, F_n, of $|X_1|+\cdots+|X_n|$, n=1,2,..., where the $\{X_i\}$ are independent standard normal random variables. In Section One, three well known approximations, the Edgeworth, the Cramér, and a Saddlepoint approximation are described. Another saddlepoint approximation is derived. The second section is devoted to a discussion of the problem of calculating the moment generating function of F_1 for complex values of its argument. Finally, the four approximations to F_n presented in Section One are compared for several cases in Section 3 and it is seen that the second saddlepoint approximation yields better results in each case.

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Sums of Independent Chi Variables						
Sums of Random Variables						
Approximations to Distribution Functions						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ______________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ______________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ______________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.