An Asymptotic Expansion for the
Upper Percentage Points of the χ^2-Distribution

By Henry E. Fettis

Abstract. An asymptotic development is given for estimating the value of the variable χ for which the χ^2-distribution

$$Q(\chi^2, \nu) = \frac{1}{\Gamma(\nu/2)} \int_{\chi^2/2}^{\infty} t^{\nu/2-1} e^{-t}\,dt$$

assumes a preassigned value α, in the region where the quantity $\eta = -\ln[\Gamma(\nu/2)\alpha]$ satisfies

$$\eta \gg \ln \eta.$$

This development generalizes a similar one given by Blair and coauthors [2] for the case $\nu = 1$. It is also shown how the estimates thus obtained may be used in conjunction with various iterative schemes to give more accurate values.

Introduction. The χ^2 , or incomplete gamma, distribution which occurs in many statistical problems is defined by

(1) $$Q(\chi^2, \nu) = \frac{1}{\Gamma(\nu/2)} \int_{\chi^2/2}^{\infty} t^{\nu/2-1} e^{-t}\,dt.$$

In this connection, a problem frequently encountered is that of finding "percentage points", and this in turn is equivalent to the determination of the value of the variable χ for which

(2) $$Q(\chi^2, \nu) = \alpha,$$

where α is a given quantity $\leqslant 1$.

A natural approach to this problem is to start with a reasonable guess for χ, and to use some iterative method to converge on the correct value.

The second order method of Newton (see, e.g. [4]) uses the formula

(3) $$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

with $f(\chi) = Q(\chi^2, \nu) - \alpha$. When applying this to the present problem, it is more expedient to use $u = \chi^2/2$ as the independent variable. Thus,

(4) $$f(u) = Q(u, \nu) - \alpha,$$

---

(5)
$$f'(u) = -\frac{e^{-u}u^{v/2-1}}{\Gamma(v/2)}.$$

In order to apply this method effectively, it is necessary to have at one's disposal a reasonable starting value. For small values of "u", such a value can be obtained by inverting the power series for $(1 - Q)$ in a manner similar to the one followed by Strecok [6] for the error function $(v = 1)$. However, this method does not give good results for large values of "u", due to the slow convergence of the inverted series. Again, for the case $v = 1$, other investigators, notably Philip [5], Strecok [6], Blair et al. [2], have developed various degrees of asymptotic approximations for this region. The most effective one of these appears to be that of Blair and coauthors [2]. In the present work, we obtain a generalization of this asymptotic development for arbitrary values of v.

In the region of large "u", it is convenient to work with $Q(\chi^2, v)$ in the form

(6)
$$\Gamma(\gamma/2)Q(\chi^2, v) = e^{-u}q(u, v),$$

where, as before, $u = \chi^2/2$. The function $q(u)$ is a well behaved one in the "tail-end" region, and can be represented (see [1, Art. 26]) either by the continued fraction

(7)
$$u^{v/2}\left\{\cfrac{1}{u + 1 - v/2 \atop 1 + \cfrac{1}{u + \cfrac{2 - v/2}{1 + \cfrac{2}{u + \cfrac{3 - v/2}{1 + \cdots +}}}}}\right\}$$

or by the asymptotic series

(8)
$$q(u, v) \sim u^{v/2-1} \sum_{k=0}^{\infty} \frac{(1 - v/2)_k}{(-u)^k}; \quad (\alpha)_k = \frac{\Gamma(\alpha + k)}{\Gamma(\alpha)},$$

which, in the case where v is an even integer, reduces to the polynomial

(9)
$$q(u, v) = \Gamma\left(\frac{v}{2}\right)\left[\sum_{k=0}^{v/2-1} \frac{u^k}{k!}\right].$$

The problem is, then, to solve for "u" the equation

(10)
$$\Gamma(v/2)Q(u, v) = e^{-u}q(u, v),$$

or, equivalently,

(11)
$$u = -\ln[\Gamma(v/2)Q(u, v)] + \ln q(u),$$

with v and $Q(u, v)$ given. If we now use the asymptotic form of $q(u)$ as given by Eq. (8), and apply the following formula (easily verified either by differentiation, or with the aid of the multinomial theorem [7])

$$\ln(1 + a_1x + a_2x^2 + a_3x^3 + \cdots +)$$

$$\text{(12)} \qquad = a_1x + (a_2 - a_1^2/2)x^2 + (a_3 - a_2a_1 + a_1^3/3)x^3$$

$$+ (a_4 - a_3a_1 + a_2a_1^2 - a_2^2/2 - a_1^4/4)x^4 + \cdots + ,$$

Eq. (11) may be written in the form

$$\text{(13)} \qquad u \sim \eta + (v/2 - 1)\ln u + \frac{P_1}{u} + \frac{P_2}{u^2} + \frac{P_3}{u^3} + \frac{P_4}{u^4} + \cdots +,$$

where

$$\eta = -\ln[\Gamma(v/2)Q(u, v)],$$

$$P_1 = (v/2 - 1),$$

$$\text{(14)} \qquad P_2 = (v/2 - 1)\left(\frac{v - 6}{4}\right),$$

$$P_3 = (v/2 - 1)\left(\frac{v^2 - 16v + 52}{13}\right),$$

$$P_4 = (v/2 - 1)\left(\frac{v^3 - 30v^2 + 244v - 568}{32}\right).$$

This suggests a development of the form

$$\text{(15)} \qquad u \sim \eta + a_1 + \frac{a_2}{\eta} + \frac{a_3}{\eta^2} + \frac{a_4}{\eta^3} + \frac{a_5}{\eta^4} + \cdots + .$$

We now assume that u is sufficiently large that

$$\eta \gg \ln(\eta)$$

and substitute Eq. (15) into (13), using the approximations

$$\frac{1}{u} \cong \frac{1}{\eta}\left[1 - \frac{a_1}{\eta} + \frac{(a_1^2 - a_2)}{\eta^2} + \frac{(2a_1a_2 - a_3 - a_1^3)}{\eta^3} + \cdots +\right],$$

$$\frac{1}{u^2} \cong \frac{1}{\eta^2}\left[1 - \frac{2a_1}{\eta} + \frac{(3a_1^2 - 2a_2)}{\eta^2} + \cdots +\right],$$

$$\text{(16)} \qquad \frac{1}{u^3} \cong \frac{1}{\eta^3}\left[1 - \frac{3a_1}{\eta} + \cdots +\right],$$

$$\frac{1}{u^4} \cong \frac{1}{\eta^4}[1 + \cdots +],$$

$$\ln u \cong \ln \eta + \frac{a_1}{\eta} + \frac{(a_2 - a_1^2/2)}{\eta^2} + \frac{(a_3 - a_2a_1 + a_1^3/3)}{\eta^3}$$

$$+ \frac{(a_4 - a_3a_1 + a_2a_1^2 - a_2^2/2 - a_1^2/4)}{\eta^4} + \cdots + .$$

Equating like powers of $(1/\eta)$, we get

$$a_1 = (\nu/2 - 1)\ln \eta,$$

$$a_2 = (\nu/2 - 1)a_1 + P_1,$$

$$a_3 = (\nu/2 - 1)(a_2 - a_1^2/2) - a_1 P_1 + P_2,$$

(17)

$$a_4 = (\nu/2 - 1)(a_3 - a_2 a_1 + a_1^3/3) + (a_1^2 - a_2)P_1 - 2a_1 P_2 + P_3,$$

$$a_5 = (\nu/2 - 1)(a_4 - a_3 a_1 + a_2 a_1^2 - a_2^2/2 - a_1^4/4)$$

$$+ (2a_2 a_1 - a_3 - a_1^3)P_1 + (3a_1^2 - 2a_2)P_2 - 3a_1 P_3 + P_4,$$

and finally, by successive resubstitutions,

$$a_1 = (\nu/2 - 1)\ln \eta,$$

$$a_2 = (\nu/2 - 1)[a_1 + 1],$$

$$a_3 = (\nu/2 - 1)[-a_1^2/2 + (\nu - 4)a_1/2 + (3\nu - 10)/4],$$

(18) $a_4 = (\nu/2 - 1)[a_1^3/3 + (-3\nu + 10)a_1^2/4 + (\nu^2 - 12\nu + 28)a_1/4$

$$+ (11\nu^2 - 92\nu + 188)/24],$$

$$a_5 = (\nu/2 - 1)[-a_1^4/4 + (11\nu - 34)a_1^3/12 + (-3\nu^2 + 26\nu - 52)a_1^2/4$$

$$+ (\nu^3 - 25\nu^2 + 144\nu - 244)a_1/8 + (25\nu^3 - 390\nu^2 + 1908\nu - 3032)/96].$$

For $\nu = 1$, the above coefficients become identical to those obtained by Blair et al. [2].

The approximation obtained from Eqs. (17), (18) may be used as a starting value, u_0, to obtain improved values, either by the following iterative sequence derived from Eq. (11) (see, e.g. [3]):

(19) $$u_{n+1} = -\ln[\Gamma(\nu/2) \cdot \alpha] + \ln q(u_n)$$

or by a higher order method such as Newton's (Eq. (3)) applied either directly to Eq. (4), or perhaps more conveniently to Eq. (11). In the latter case, we have

(20) $$f(u) = \ln[\Gamma(\nu/2) \cdot \alpha] + u - \ln q(u),$$

(21) $$f'(u) = 1 - q'(u)/q(u),$$

and since

(22) $$q'(u) = q(u) - u^{\nu/2 - 1},$$

the expression for $f'(u)$ (Eq. (21)) may be simplified to

(23) $$f'(u) = \frac{u^{\nu/2 - 1}}{q(u)},$$

so that $f(u)$ and $f'(u)$ can both be obtained from essentially the same continued fraction (or polynomial if v is an even integer).

Numerical Example. For simplicity, the case where $v = 4$ is considered. From Eq. (18), we get

$$a_1 = 2.220326,$$
$$a_2 = 3.220326,$$
$$a_3 = -1.964926,$$
$$a_4 = -1.203292,$$
$$a_5 = 6.448730.$$

Thus, for, e.g. $\alpha = .0001$, we find $\eta = 9.210340$, and $u \cong 11.7565$ as a first approximation. This estimate may be improved by direct iteration, employing Eq. (11), with $q(u) = 1 + u$, or by Newton's method (Eq. (3)), with

$$f(u) = \ln \alpha + u - \ln(1 + u), \quad f'(u) = \frac{u}{1 + u}.$$

Using the first-mentioned method with $u_0 = 11.7565$ as a starting value, we get

$$u_1 = \ln(1 + u_0) - \ln(.0001) = 11.7563813,$$
$$u_2 = 11.7563720,$$
$$u_3 = 11.7563713,$$
$$u_4 = 11.7563712 = u_{5'},$$

while with the Newton method we find

$$f(u_0) = \ln \alpha + u_0 - \ln(1 + u_0) = 1.18682 \times 10^{-4},$$

$$f'(u_0) = .92160859,$$

$$u_1 = 11.7565 - .00012878 = 11.75637122 = u_2.$$

The more rapid convergence of the Newton method is not surprising, since it is quadratic, and the initial approximation is already correct to within one unit of the sixth significant figure. The direct iteration algorithm, on the other hand, is only linearly convergent, but is attractive due to its simplicity and adaptability to a desktop automatic calculator. For larger values of α, more iterations will be required in both of the previously mentioned methods. In this region, a scheme with higher order convergence such as Halley's third order method as suggested by Chen* may be used to advantage

$$(24) \qquad u_k - u_{k+1} = \frac{f(u_k)}{f'(u_k) - \dfrac{f(u_k)f''(u_k)}{2f'(u_k)}}$$

---

*Tien Chi-Chen—private communication (IBM Corp., San Jose, California).

In the above, the presence of the second derivative poses no essentially new problem since, from Eqs. (22), (23), we have

$$(25) \qquad f''(u) = [f'(u)]^2 - \left[1 - \frac{\nu/2 - 1}{u}\right] f'(u).$$

1885 California, Apt. 62
Mountain View, California 94041

1. M. ABRAMOWITZ & I. STEGUN, Editors, *Handbook of Mathematical Functions, with Formulas, Graphs and Tables,* Dover, New York, 1966.

2. J. M. BLAIR, C. A. EDWARDS & J. H. JOHNSON, "Rational Chebyshev approximations for the inverse of the error function," *Math. Comp.,* v. 30, 1976, pp. 827—830.

3, H. E. FETTIS, "A stable algorithm for computing the inverse error function in the 'tail-end' region," *Math. Comp.,* v. 28, 1975, pp. 585—587.

4. W. GANDER, "A machine independent algorithm for computing percentage points of the χ^2-distribution," *Z. Angew. Math. Phys.,* v. 28, 1977, pp. 1133—1136.

5. J. R. PHILIP, "The function inverfc θ," *Austral. J. Phys.,* v. 13, 1960, pp. 13—20.

6. A. J. STRECOK, "On the calculation of the inverse of the error function," *Math. Comp.,* v. 22, 1968, pp. 144—158.

7. M. ZYCZKOWSKI, "Potenzieren von verallgemeinerten Potenzreihen mit beliebigen Exponent," *Z. Angew. Math. Phys.,* v. 12, 1961, pp. 572—576.