APPROXIMATE FORMULAS FOR THE PERCENTAGE POINTS AND NORMALIZATION OF t AND χ^2 [1]

By Henry Goldberg[2] and Harriet Levine

Statistical Research Group, Columbia University

1. Introduction. The χ^2 Distribution and Student's t-distribution are functions of a parameter n (degrees of freedom) and approach the normal distribution as n approaches infinity. The normal distribution is a good approximation to these distributions for large n. For small or moderate n, a better approximation may be obtained by using a function of t(or χ^2) which approaches the normal distribution more rapidly as n increases. Hotelling and Frankel [7] pointed out that an additional advantage of the normalization of a distribution is that further statistical tests are possible with the normalized variate. Normalizing t(or χ^2) is equivalent to transforming it into a function which is normally distributed to a required degree of approximation; that is, a normally distributed variate of zero mean and unit variance is expressed as a function of t(or χ^2) in powers of $1/n$.

The reverse problem of expressing t(or χ^2) as a function of a normally distributed variate of zero mean and unit variance in powers of $1/n$ is also of practical importance in connection with significance tests for which the significance levels, or percentage points, of the t and χ^2 distributions are required.

Cornish and Fisher [1] (see also [2]) have given a method for the normalization of distributions which approach normality as the number of degrees of freedom, n, increases and whose cumulants are expressed in power series of $1/n$, so that the order of magnitude of the rth cumulant is that of $n^{-(r-1)}$. A method has also been given for expressing a variate with such a distribution as a function of a normally distributed variate of zero mean and unit variance in powers of $1/n$.

It is the purpose of this note to apply the Cornish-Fisher method (1) to the derivation of asymptotic formulas for the percentage points of the t and χ^2 distributions and (2) to the normalization of these distributions. Tables are given which indicate the accuracy of these approximations and compare them with other approximations. Tables are also given to facilitate the calculation of the approximations for the percentage points of t and χ^2.

---

[1] This paper reports work done in the Statistical Research Group, Division of War Research, Columbia University under contract OEMsr-618 with the Applied Mathematics Panel, National Defense Research Committee, Office of Scientific Research and Development. The work was first reported in an unpublished memorandum, "Application of the Cornish-Fisher method to an approximation of the significance levels of t and χ^2" (SRG number 507, April 28, 1945).

[2] Henry Goldberg died April 19, 1945.

2. The Cornish-Fisher method.[3] Consider the random variable y with probability distribution function $f(y)$, expected value $E(y)$, and variance $\sigma^2(y)$. Let K_r denote the rth cumulant of y and a_r denote the rth relative cumulant of y; i.e., $a_r = \frac{K_r}{K_2^{r/2}}$. Let x denote a normally distributed variate with zero mean and unit variance.

For every p, $(0 \leq p \leq 1)$, let y_p be defined by

$$\int_{-\infty}^{y_p} f(y)\,dy = p$$

and x_p by

$$\int_{-\infty}^{x_p} \frac{1}{\sqrt{2\pi}} e^{-(r^2/2)}\,dr = p.$$

That is, corresponding to every y_p, there is an x_p having the same probability integral (p). The Cornish-Fisher Method for expressing a normally distributed variate with zero mean and unit variance as a function of a standardized variate with the same probability integral gives

$$x_p \sim b_0 + b_1 z_p + b_2 z_p^2 + b_3 z_p^3 + b_4 z_p^4 + b_5 z_p^5 + \cdots \tag{1}$$

where z_p is the standardized variate corresponding to y_p; i.e.,

$$z_p = \frac{y_p - E(y)}{\sigma(y)}$$

and the b_i are defined in terms of the relative cumulants.

Cornish and Fisher give also the following expansion for a standardized variate as a function of a normally distributed variate:

$$z_p \sim c_0 + c_1 x_p + c_2 x_p^2 + c_3 x_p^3 + c_4 x_p^4 + c_5 x_p^5 + \cdots \tag{2}$$

where the c_i are defined in terms of the relative cumulants.

3. An approximation for the percentage points of Student's t-distribution. The standardized variate $z = t\left(\frac{n-2}{n}\right)^{\frac{1}{2}}$ can be expressed as a function of the normal variate, x, in powers of $1/n$ by using the Cornish-Fisher equation (2). Omitting terms of degree greater than two in $1/n$ gives, after simplification, the following asymptotic expansion for t:

$$t \sim x + \frac{x^3 + x}{4n} + \frac{5x^5 + 16x^3 + 3x}{96n^2} + \cdots. \tag{3}$$

[3] Churchill Eisenhart suggested the use of the Cornish-Fisher Method for obtaining percentage points of the chi-square distribution not given in existing tables, a problem which arose in several connections, including the computation of a table of factors for tolerance limits for normal distributions according to two formulas devised in the Statistical Research Group, one by A. Wald and J. Wolfowitz and the other by Albert H. Bowker, both of which are published elsewhere in this issue of the *Annals of Math. Stat.* The table will be included in a volume by the Statistical Research Group, *Techniques of Statistical Analysis*, to be published by the McGraw-Hill Book Company in 1946; its preparation, including the work reported in the present paper, was directed by Albert H. Bowker; the Statistical Research Group was directed by W. Allen Wallis.

For simplicity, the subscript p which appears in the Cornish-Fisher equation (2) has been dropped. It should be understood, however, that the x and t used in expansion (3) have the same probability integral. It is interesting to note that the first two terms were derived by Peiser [4].

TABLE 1

*Table of Polynomials Required for the Approximation for the Percentage Points of the t-distribution**

Probability Integral (p)	$x_p = x$	$f_1(x)$	$f_2(x)$
.999	3.090232	8.150129	19.692529
.9975	2.807034	6.231221	12.850916
.995	2.575829	4.916548	8.834762
.99	2.326348	3.729074	5.719746
.975	1.959964	2.372271	2.822499
.95	1.644854	1.523769	1.420203
.90	1.281552	.846585	.570891
.75	.674490	.245335	.079490

* This table can be used for determining x, $f_1(x)$ and $f_2(x)$ corresponding to the complements of the selected values of p by using the relations

$$x_{1-p} = -x_p$$
$$f_1(-x) = -f_1(x)$$
$$f_2(-x) = -f_2(x).$$

To facilitate the use of the approximation, tables of the required polynomials in x have been computed for selected probability integrals. The approximation can be written

$$t \sim x + \frac{f_1(x)}{n} + \frac{f_2(x)}{n^2} + \cdots$$

where

$$f_1(x) = \frac{x^3 + x}{4}$$

and

$$f_2(x) = \frac{5x^5 + 16x^3 + 3x}{96}.$$

Table 1 gives values of x_p (or x), $f_1(x)$ and $f_2(x)$ for selected values of the probability integral p. Table 2 gives approximate and exact percentage points of t for selected values of p and degrees of freedom. The exact values were taken from Merrington [5]. Table 2 shows the high degree of accuracy of the three

TABLE 2

Comparative Table of Approximate and Exact Values of the Percentage Points of the t-distribution

Probability Integral (p)	Degrees of Freedom	Approximate Percentage Point			Exact Percentage Point
		Normal	2 Term	3 Term	
.9975	1	2.8070	9.0383	21.8892	127.32
	2		5.9226	9.1354	14.089
	10		3.4302	3.5587	3.5814
	20		3.1186	3.1507	3.1534
	40		2.9628	2.9708	2.9712
	60		2.9109	2.9145	2.9146
	120		2.8590	2.8599	2.8599
.9950	1	2.5758	7.4924	16.3271	63.657
	2		5.0341	7.2428	9.9248
	10		3.0675	3.1558	3.1693
	20		2.8217	2.8437	2.8453
	40		2.6987	2.7043	2.7045
	60		2.6578	2.6602	2.6603
	120		2.6168	2.6174	2.6174
.9750	1	1.9600	4.3322	7.1547	12.706
	2		3.1461	3.8517	4.3027
	10		2.1972	2.2254	2.2281
	20		2.0786	2.0856	2.0860
	40		2.0193	2.0210	2.0211
	60		1.9995	2.0003	2.0003
	120		1.9797	1.9799	1.9799
.9500	1	1.6449	3.1686	4.5888	6.3138
	2		2.4067	2.7618	2.9200
	10		1.7972	1.8114	1.8125
	20		1.7210	1.7246	1.7247
	40		1.6829	1.6838	1.6839
	60		1.6702	1.6706	1.6707
	120		1.6576	1.6577	1.6577
.7500	1	0.6745	.9198	.9993	1.0000
	2		.7972	.8170	.8165
	10		.6990	.6998	.6998
	20		.6868	.6870	.6870
	40		.6806	.6807	.6807
	60		.6786	.6786	.6786
	120		.6765	.6765	.6766

term approximation for $n \geq 10$ and the superiority of this approximation over the two-term approximation derived by Peiser.

4. An approximation for the percentage points of the χ^2 distribution. The standardized variate $z = \frac{\chi^2 - n}{\sqrt{2n}}$ can be expressed as a function of the normal variate, x, in powers of $1/n$ by using the Cornish-Fisher equation (2). Retain-

TABLE 3

*Table of Polynomials Required for the Approximation for the Percentage Points of the χ^2 distribution**

Probability Integreal (p)	$G_1(x)$	$G_2(x)$	$G_3(x)$	$G_4(x)$	$G_5(x)$
.999	4.370248	5.699690	.619006	−1.602112	1.273498
.9975	3.969745	4.586292	.193953	−1.113149	.875184
.995	3.642773	3.756598	−.073888	−.802518	.622768
.99	3.289953	2.941263	−.290266	−.541971	.411597
.975	2.771808	1.894306	−.486382	−.272398	.194832
.95	2.326174	1.137029	−.554981	−.122957	.077898
.90	1.812388	.428250	−.539450	−.017722	.002186
.75	.953873	−.363376	−.346842	.060220	−.030881

* This table can be used for determining the $G_i(x)$ for values of x corresponding to the complements of the selected values of p by using the relations

$$x_{1-p} = -x_p$$
$$G_i(-x) = (-1)^i G_i(x), \text{ for } i = 1, \dots, 5.$$

ing terms in $n^{-3/2}$ gives, after simplification, the following asymptotic expansion for χ^2:

$$\chi^2 \sim n + G_1(x)n^{\frac{1}{2}} + G_2(x) + \frac{G_3(x)}{n^{\frac{1}{2}}} + \frac{G_4(x)}{n} + \frac{G_5(x)}{n^{\frac{3}{2}}} + \cdots \tag{4}$$

where

$$G_1(x) = \sqrt{2}x$$

$$G_2(x) = \tfrac{2}{3}(x^2 - 1)$$

$$G_3(x) = \frac{1}{9\sqrt{2}}(x^3 - 7x)$$

$$G_4(x) = -\frac{1}{405}(6x^4 + 14x^2 - 32)$$

$$G_5(x) = \frac{1}{4860\sqrt{2}}(9x^5 + 256x^3 - 433x).$$

TABLE 4

Comparative Table of Various Approximate and Exact Values of the Percentage Points of the χ^2 Distribution

Approximation	Degrees of Freedom	Probability Integral (p)									
		.005	.01	.05	.10	.25	.75	.90	.95	.99	.995
Exact Value	1	.0000	.0002	.0039	.0158	.1015	1.3233	2.7055	3.8415	6.6349	7.8794
Cornish-Fisher		*	*	.1650	.1354	.1207	1.2730	2.6857	3.8632	6.8106	8.1457
Peiser		1.1877	.9416	.3658	.1553	.0296	1.2437	2.7012	3.9082	6.9409	8.3255
Wilson-Hilferty		*	*	.0000	.0052	.0972	1.3156	2.6390	3.7468	6.5858	7.9048
Fisher		1.2416	.8796	.2079	.0396	.0530	1.4020	2.6027	3.4976	5.5323	6.3933
Exact Value	2	.0100	.0201	.1026	.2107	.5754	2.7726	4.6052	5.9915	9.2103	10.5966
Cornish-Fisher		.0357	.0773	.1507	.2370	.5739	2.7595	4.6018	6.0004	9.2632	10.6749
Peiser		.6572	.4938	.2398	.2466	.5329	2.7403	4.6099	6.0343	9.3887	10.8560
Wilson-Hilferty		.0001	.0029	.0790	.1968	.5857	2.7628	4.5590	5.9369	9.2205	10.6729
Fisher		.3560	.1766	.0038	.1015	.5592	2.8957	4.5409	5.7017	8.2353	9.2789
Exact Value	10	2.1558	2.5582	3.9403	4.8652	6.7372	12.5489	15.9871	18.3070	23.2093	25.1882
Cornish-Fisher		2.1606	2.5621	3.9418	4.8657	6.7369	12.5484	15.9872	18.3077	23.2120	25.1921
Peiser		2.2605	2.6293	3.9565	4.8676	6.7299	12.5434	15.9889	18.3175	23.2532	25.2527
Wilson-Hilferty		2.0937	2.5122	3.9315	4.8695	6.7506	12.5386	15.9677	18.2918	23.2394	25.2523
Fisher		1.5897	2.0656	3.6830	4.7350	6.7874	12.6675	15.9073	18.0225	22.3463	24.0452
Exact Value	20	7.4339	8.2604	10.8508	12.4426	15.4518	23.8277	28.4120	31.4104	37.5662	39.9968
Cornish-Fisher		7.4020	8.2614	10.8511	12.4427	15.4517	23.8276	28.4120	31.4106	37.5670	40.0309
Peiser		7.4491	8.2930	10.8582	12.4436	15.4483	23.8249	28.4129	31.4159	37.5895	40.0641
Wilson-Hilferty		7.3835	8.2257	10.8470	12.4480	15.4619	23.8194	28.3989	31.4017	37.5914	40.0461
Fisher		6.7314	7.6779	10.5807	12.3179	15.5153	23.9397	28.3245	31.1249	36.7340	38.9035

* Computed percentage point is negative.

TABLE 4 (CONT.)

Approximation	Degrees of Freedom	Probability Integral (p)									
		.005	.01	.05	.10	.25	.75	.90	.95	.99	.995
Exact Value	40	20.7065	22.1643	26.5093	29.0505	33.6603	45.6160	51.8050	55.7585	63.6907	66.7659
Cornish-Fisher		20.6835	22.1645	26.5094	29.0505	33.6603	45.6160	51.8051	55.7585	63.6909	66.7896
Peiser		20.7060	22.1797	26.5128	29.0510	33.6586	45.6146	51.8055	55.7613	63.7029	66.8072
Wilson-Hilferty		20.6690	22.1394	26.5080	29.0555	33.6676	45.6097	51.7963	55.7534	63.7104	66.8024
Fisher		19.9230	21.5289	26.2330	28.9305	33.7325	45.7225	51.7119	55.4726	62.8830	65.7119
Exact Value	60	35.5346	37.4848	43.1879	46.4589	52.2938	66.9814	74.3970	79.0819	88.3794	91.9517
Cornish-Fisher		35.5155	37.4850	43.1880	46.4589	52.2938	66.9814	74.3970	79.0820	88.3795	91.9709
Peiser		35.5303	37.4949	43.1902	46.4592	52.2927	66.9805	74.3973	79.0838	88.3877	91.9829
Wilson-Hilferty		35.5034	37.4647	43.1874	46.4633	52.2998	66.9762	74.3900	79.0782	88.3961	91.9820
Fisher		34.7185	36.8285	42.9095	46.3411	52.3697	67.0853	74.3013	78.7960	87.5834	90.9164
Exact Value	80	51.1720	53.5400	60.3915	64.2778	71.1445	88.1303	96.5782	101.879	112.329	116.321
Cornish-Fisher		51.1555	53.5401	60.3915	64.2778	71.1445	88.1303	96.5782	101.879	112.329	116.338
Peiser		51.1664	53.5475	60.3931	64.2781	71.1437	88.1295	96.5784	101.881	112.335	116.347
Wilson-Hilferty		51.1448	53.5227	60.3912	64.2819	71.1497	88.1256	96.5723	101.876	112.344	116.348
Fisher		50.3375	52.8718	60.1120	64.1614	71.2225	88.2325	96.4809	101.594	111.540	115.297
Exact Value	100	67.3276	70.0648	77.9295	82.3581	90.1332	109.141	118.498	124.342	135.807	140.169
Cornish-Fisher		67.3276	70.0649	77.9295	82.3581	90.1332	109.141	118.498	124.342	135.807	140.184
Peiser		67.3363	70.0708	77.9308	82.3583	90.1326	109.141	118.498	124.343	135.812	140.192
Wilson-Hilferty		67.3032	70.0494	77.9294	82.3618	90.1378	109.137	118.493	124.340	135.820	140.193
Fisher		66.4809	69.3888	77.6493	82.2427	90.2126	109.242	118.400	124.056	135.023	139.154

As before, the subscript p which appears in the Cornish-Fisher equation (2) has been dropped. The x and χ^2 which are used in expansion (4) have the same probability integral. The first four terms were derived by Peiser [4].

Table 3 gives values of the $G_i(x)$ for selected values of the probability integral p. Table 4 compares various approximations with the exact percentage

TABLE 5

Comparative Table of Approximate and Exact Values of the Probability Integral of t

t	Probability Integral of t							
	$n = 1$		$n = 2$		$n = 10$		$n = 20$	
	Approximate	Exact	Approximate	Exact	Approximate	Exact	Approximate	Exact
0.1	.5311	.5317	.5351	.5353	.5388	.5388	.5393	.5393
1	.7734	.7500	.7917	.7887	.8296	.8296	.8354	.8354
3	1.0000	.8976	1.0000	.9523	.9954	.9933	.9967	.9965
5	1.0000	.9372	1.0000	.9811	1.0000	.9997	1.0000	1.0000
6	1.0000	.9474	1.0000	.9867	1.0000	.9999	1.0000	1.0000

TABLE 6

Comparative Table of Approximate and Exact Values of the Probability Integral of χ^2

χ^2	Probability Integral of χ^2							
	$n = 2$		$n = 10$		$n = 20$		$n = 29$	
	Approximate	Exact	Approximate	Exact	Approximate	Exact	Approximate	Exact
1	.3963	.3935	.0010	.0002	.0000	.0000	.0000	.0000
5	.9646	.9179	.1098	.1088	.0004	.0003	.0000	.0000
10	1.0000	.9933	.5594	.5595	.0323	.0318	.0005	.0004
20	1.0000	1.0000	.9768	.9707	.5420	.5421	.1071	.1071
30	1.0000	1.0000	1.0000	.9991	.9305	.9301	.5860	.5860
50							.9916	.9910

points of χ^2 for selected values of p and degrees of freedom. The Peiser four-term approximation, the Wilson-Hilferty approximation,

$$\chi_p^2 = n\left(1 - \frac{2}{9n} + x_p\sqrt{\frac{2}{9n}}\right)^3$$

and the Fisher approximation,

$$\chi_p^2 = \tfrac{1}{2}(x_p + \sqrt{2n - 1})^2$$

are given for comparison. The exact values were taken from Thompson [6]. Table 4 shows the high degree of accuracy, and the general superiority of the Cornish-Fisher approximation, for $n \geq 10$. For low probabilities (.005) the Peiser approximation is often better than the full series; for small n, (1, 2), the Wilson-Hilferty approximation is often better.

5. Normalization of t and χ^2. The Cornish-Fisher equation (1) applied to the t-distribution or, alternatively, a formal reversion of the power series (3) gives the asymptotic expansion

$$x \sim t\left[1 - \frac{t^2 + 1}{4n} + \frac{13t^4 + 8t^2 + 3}{96n^2} + \cdots\right]. \tag{5}$$

Expansion (5) agrees with the first three terms of an expansion derived by Hotelling and Frankel [7].

Applying the Cornish-Fisher equation (1) to the χ^2 distribution gives the expansion

$$\begin{aligned} x \sim \frac{1}{38880\sqrt{2}\,n^{\frac{1}{2}}} \Big\{ & -68649n + [128469\chi^2 + 29056] \\ & - \frac{2}{n}[53553\chi^4 + 2208\chi^2 - 386] + \frac{2}{n^2}[34257\chi^6 + 792\chi^4 + 238\chi^2] \\ & - \frac{1}{n^3}[25221\chi^8 + 304\chi^6] + \frac{3993}{n^4}\chi^{10} + \cdots \Big\}. \end{aligned} \tag{6}$$

6. Accuracy of the normalizations of t and χ^2. The accuracy of the normalization (5) of t may be judged from Table 5, which compares the approximate value of the probability integral with the exact value. The approximate value is the normal probability integral corresponding to the value of x computed from (5) for the given values of t and n. The exact values were obtained from Student's tables [8]. For fixed n, the approximation improves as t decreases from moderate to small values. The approximation appears to improve as t increases from moderate values (about 3) to large values because of the more rapid approach to unity of the probability integral of a normal variate.

The accuracy of the normalization (6) of χ^2 may be judged from Table 6, which compares the approximate value of the probability integral with the exact value. The approximate value is the normal probability integral corresponding to the value of x computed from (6) for the given values of χ^2 and n. The exact values were obtained from the table of Pearson [9].

REFERENCES

[1] E. A. CORNISH AND R. A. FISHER, "Moments and cumulants in the specification of distributions," *Revue de l'Institut International de Stat.*, Vol. 4 (1937), pp. 1–14.

[2] M. G. KENDALL, *The Advanced Theory of Statistics*, Vol. 1, London, 1943, pp. 156–159.

[3] S. S. WILKS, *Mathematical Statistics*, Princeton Univ. Press, 1943, pp. 113–115.
[4] A. M. PEISER, "Asymptotic formulas for significance levels of certain distributions," *Annals of Math. Stat.*, Vol. 14 (1943), pp. 56–62.
[5] M. MERRINGTON, "Table of percentage points of the t-distribution," *Biometrika*, Vol. 32 (1941–42), p. 300.
[6] C. M. THOMPSON, "Table of percentage points of the χ^2-distribution," *Biometrika*, Vol. 32 (1941–42), pp. 188–189.
[7] H. HOTELLING AND L. R. FRANKEL, "The transformation of statistics to simplify their distribution," *Annals of Math. Stat.*, Vol. 9 (1938), pp. 87–96.
[8] E. S. PEARSON AND J. WISHART, *"Student's" Collected Papers*, Cambridge Univ. Press, 1942, pp. 118–120.
[9] K. PEARSON, *Tables for Statisticians and Biometricians*, Part 1, Third Edition, Cambridge Univ. Press, 1930, Table XII, pp. 26–28.
[10] A. H. BOWKER, "Computation of factors for tolerance limits on a normal distribution when the sample is large," *Annals of Math. Stat.*, Vol. 17 (1946).