The Percentile Points of Distributions Having Known Cumulants

Sir Ronald A. Fisher
and
E. A. Cornish

In an earlier study of the uses of moments and cumulants in the specification of statistical distributions, the authors developed explicit asymptotic expansions, expressing any desired percentile point of such distributions in terms of known cumulants.

The general formulae are now presented as far as the sixth adjustment, based on the eighth cumulant, and also numerical tables showing the coefficients of all terms for ten chosen levels of significance over the range 0.5 to 0.0005 (single tail), together with the first five Hermite polynomials and tables for the common tests of significance, χ^2, t and z, at the same levels.

I. Introductory

In 1937, in a study of the uses of moments and cumulants in the specification of statistical distributions, the authors (1937)* were led to develop explicit asymptotic expansions, expressing any desired percentile point of such distributions in terms of known cumulants. The general formulae were presented so far as the fourth adjustment, based on the sixth cumulant, and also numerical tables showing the coefficients of all terms for nine chosen levels of significance over the range 0.25 to 0.0005 (single tail), together with the first five Hermite polynomials at the same levels.

As an illustrative example, the cumulants of the z distribution were expressed in terms of $1/n_1$ and $1/n_2$, the reciprocals of the two numbers of degrees of freedom, and the rapid convergence at the 5% point exhibited for the case $n_1 = 24$ and $n_2 = 60$. In the intervening period, the formulae have frequently been found useful, either for calculations of higher accuracy in the case of functions already tabulated, or for values outside their range, or especially with tables of multiple entry to supply intermediate values more accurate than can be obtained by interpolation (Fisher 1941, Goldberg and Levine 1946), or for cases where no tables existed (Johnson and Welch 1939).

II. The Method of Expansion

The several steps in the method of expansion are set out in Sections 7 and 8 of our previous paper. Here a brief outline of these steps will be sufficient, as the principal purpose of this note is to extend the formulae and tables to the sixth corrective term, so widening the range of useful application.

If the element of frequency in the distribution of a variate ξ is $f(\xi)d\xi$, the

* Attention is drawn to this reference, as it has been repeatedly misquoted.

 Technometrics, 2: 209-225, (1960).

effect of the operator

$$\exp\left\{\frac{a_r}{r!}\left(-\frac{d}{d\xi}\right)^r\right\},$$

when acting on the frequency function, is to increase the rth cumulant by a_r, but leaves the distribution otherwise unchanged. This important operational property of the cumulants is basic to the method, the essential steps of which are as follows:

(i) If the cumulants $\kappa_1, \kappa_2, \cdots$ of the distribution of ξ are expressible in power series of the reciprocal of some number n, the frequency element may be represented as

$$\exp\left\{-av^{1/2}\frac{d}{d\xi}+\frac{1}{2}bv\frac{d^2}{d\xi^2}-\frac{1}{6}cv^{3/2}\frac{d^3}{d\xi^3}+\frac{1}{24}dv^2\frac{d^4}{d\xi^4}-\frac{1}{120}ev^{5/2}\frac{d^5}{d\xi^5}\right.$$
$$\left.+\frac{1}{720}fv^3\frac{d^6}{d\xi^6}-\cdots\right\}\frac{1}{\sqrt{2\pi v}}e^{-(\xi-m)^2/2v}\,d\xi \tag{1}$$

where the coefficients a and c are of order $n^{-1/2}$, b and d of order n^{-1}, e of order $n^{-3/2}$, f of order n^{-2} $\cdots$, and are related, respectively, to κ_1 and κ_3, κ_2 and $\kappa_4, \kappa_5, \kappa_6, \cdots$, and m and v are the mean and variance of a normal distribution chosen for convenience.

(ii) Expanding the operator and integrating, the frequency less than $m + \xi v^{1/2}$ may be expressed in terms of the corresponding normal probability integral and a series of adjustments of decreasing order of magnitude.

(iii) If x is the normal deviate at some chosen level of probability, and ξ the corresponding deviate of the distribution under consideration, the difference $\xi - x$ may be found by equating the expression for the probability that the variate has a value less than ξ to

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\xi}e^{-\xi^2/2}\,d\xi-\frac{1}{\sqrt{2\pi}}e^{-\xi^2/2}\left\{(\xi-x)-\frac{1}{2}(\xi-x)^2\xi\right.$$
$$\left.+\frac{1}{6}(\xi-x)^3(\xi^2-1)-\frac{1}{24}(\xi-x)^4(\xi^3-3\xi)+\cdots\right\} \tag{2}$$

in which the coefficients are the Hermite polynomials. By considering the terms of each order of magnitude in succession, we may develop an expansion for $\xi - x$ in terms of successive polynomials in ξ.

(iv) The expansion for $\xi - x$ is converted to a much more useful expansion in terms of x, the values of which are known in advance, so obtaining the percentile deviate ξ explicitly in terms of the normal deviate.

The adjustments to the normal deviate having the required probability integral are set out in tabular form below. Adjustments V and VI are new with this paper.

Table I gives the numerical values of the first seven Hermite polynomials, over the same range of percentiles as given previously, and Table II gives the numerical values of the polynomials in the several adjustment terms.

Table I—*Hermite Polynomials*

p	x	$x^2 - 1$	$x^3 - 3x$	$x^4 - 6x^2 + 3$	$x^5 - 10x^3 + 15x$	$x^6 - 15x^4 + 45x^2 - 15$	$x^7 - 21x^5 + 105x^3 - 105x$
.5	0	−1.0000 0000 0000	0	3.0000 0000 0000	0	−15.0000 0000 0000	0
.25	0.6744 8975 0296	−0.5450 6357 6746	−1.7166 1929 6367	0.4773 4860 9677	7.1884 4393 0012	2.4617 8870 2987	−41.4702 1233 2512
.10	1.2815 5156 5545	0.6423 7441 5150	−1.7398 6719 3688	−4.1568 5277 1360	1.6322 4759 7878	22.8760 7332 1217	19.5233 8199 1054
.05	1.6448 5362 6951	1.7055 4345 4095	−0.4843 3791 7511	−5.9132 9534 2574	−7.7891 5362 1425	16.7544 5912 7787	74.2935 5459 2500
.025	1.9599 6398 4540	2.8414 5882 0694	1.6492 2898 3034	−5.2919 4705 3076	−16.9689 4156 4258	−6.7987 7905 6330	88.4882 8729 6296
.01	2.3263 4787 4041	4.4118 9443 1054	5.6109 0548 2094	−0.1827 6525 3449	−22.8687 9748 7187	−52.2869 5214 8942	15.5751 4495 1354
.005	2.5758 2930 3549	5.6348 9660 1021	9.3628 7318 0281	7.2124 7330 0116	−18.8733 9264 3621	−84.6770 0432 9402	−104.8731 5322 6683
.0025	2.8070 3376 8344	6.8794 3857 6622	13.6967 4885 5139	17.8089 2082 3031	−4.7967 5329 2545	−102.5092 5258 5743	−258.9664 1382 0596
.001	3.0902 3230 6168	8.5495 3570 6083	20.2395 8682 9338	36.8964 1796 5259	33.0601 5546 0763	−82.3185 2937 4518	−452.7443 1163 3935
.0005	3.2905 2673 1492	9.8275 6617 0663	25.7568 1572 7087	55.2707 9215 6104	78.8427 5615 2042	−16.9197 6407 7727	−528.7314 7290 0548

TABLE II—*Numerical values of the polynomials in the general formula*

	p									
	.5	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
c	−0.16667	−0.09084	0.10706	0.28426	0.47358	0.73532	0.93915	1.14657	1.42492	1.63793
b	0	0.33724	0.64078	0.82243	0.97998	1.16317	1.28791	1.40352	1.54512	1.64526
d	0	−0.07153	−0.07249	−0.02018	0.06872	0.23379	0.39012	0.57070	0.84332	1.07320
c^2	0	0.07663	0.06106	−0.01878	−0.14607	−0.37634	−0.59171	−0.83890	−1.21026	−1.52234
bc	0.16667	0.09084	−0.10706	−0.28426	−0.47358	−0.73532	−0.93915	−1.14657	−1.42492	−1.63793
e	0.02500	0.00398	−0.03464	−0.04928	−0.04410	−0.00152	0.06010	0.14841	0.30747	0.46059
cd	−0.08333	0.00282	0.14644	0.17532	0.10210	−0.17621	−0.53531	−1.02868	−1.89358	−2.71243
c^3	0.05247	−0.01428	−0.11629	−0.11899	−0.02937	0.25195	0.59757	1.06301	1.86790	2.62337
b^2	0	−0.08431	−0.16019	−0.20561	−0.24500	−0.29079	−0.32198	−0.35088	−0.38628	−0.41132
bd	0	0.10729	0.10874	0.03027	−0.10308	−0.35068	−0.58518	−0.85605	−1.26497	−1.60980
bc^2	0	−0.19158	−0.15265	−0.04696	0.36517	0.94084	1.47928	2.09726	3.02565	3.80584
f	0	0.00998	0.00227	−0.01082	−0.02357	−0.03176	−0.02621	−0.00666	0.04592	0.10950
ce	0	−0.05126	0.01086	0.09462	0.16106	0.16058	0.05366	−0.17498	−0.70466	−1.30531
d^2	0	−0.03285	0.00776	0.05985	0.09659	0.07888	−0.01226	−0.19116	−0.59062	−1.03555
c^2d	0	0.14764	−0.10858	−0.39517	−0.55856	−0.32621	0.35696	1.60445	4.29316	7.23307
c^4	0	−0.06898	0.09585	0.25623	0.31624	0.07286	−0.46535	−1.39199	−3.32716	−5.40702
b^2c	−0.16667	−0.09084	0.10706	0.28426	0.47358	0.73532	0.93915	1.14657	1.42492	1.63793
be	−0.05000	−0.00796	0.06928	0.09855	0.08820	0.00305	−0.12021	−0.29682	−0.61494	−0.92118
bcd	0.25000	−0.00846	−0.43931	−0.52597	−0.30631	0.52864	1.60592	3.08604	5.68074	8.13729
bc^3	−0.20988	0.05714	0.46515	0.47598	0.11748	−1.00781	−2.39028	−4.25205	−7.47158	−10.49350
g	−0.00298	0.00049	0.00454	0.00332	−0.00135	−0.01037	−0.01680	−0.02034	−0.01633	−0.00336
cf	0.02083	−0.00791	−0.03371	−0.01141	0.04024	0.12188	0.16262	0.15484	0.01974	−0.21672
de	0.03333	−0.01168	−0.04871	−0.01011	0.07080	0.18338	0.21571	0.14718	−0.18946	−0.70228
c^2e	−0.08333	0.04892	0.12788	−0.03219	−0.31878	−0.67431	−0.71959	−0.37377	1.00984	3.01445
cd^2	−0.11111	0.05854	0.14979	−0.06496	−0.42767	−0.82209	−0.76354	−0.11687	2.10109	5.17256
c^3d	0.23457	−0.17412	−0.28497	0.38896	1.39396	2.27932	1.73496	−0.69355	−8.16706	−18.12350
c^5	−0.09091	0.08374	0.07657	−0.28485	−0.76311	−1.05359	−0.53986	1.05630	5.56273	11.36150

TABLE II—*Continued*

	p									
	.5	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
b^3	0	0.04216	0.08010	0.10280	0.12250	0.14540	0.16099	0.17544	0.19314	0.20566
b^2d	0	−1.13411	−0.13593	−0.03784	0.12885	0.43835	0.73147	1.07006	1.58122	2.01225
b^2c^2	0	0.33526	0.26714	−0.08217	−0.63904	−1.64648	−2.58873	−3.67020	−5.29489	−6.66023
bf	0	−0.02496	−0.00567	0.02705	0.05892	0.07941	0.06553	0.01666	−0.11479	−0.27376
bce	0	0.17941	−0.03802	−0.33116	−0.56370	−0.56204	−0.18780	0.61244	2.46632	4.56859
bd^2	0	0.11498	−0.02716	−0.20949	−0.33807	−0.27606	0.04292	0.66905	2.06717	3.62442
bc^2d	0	−0.66436	0.48861	1.77828	2.51350	1.46795	−1.60634	−7.22002	−19.31920	−32.54883
bc^4	0	0.37938	−0.52718	−1.40927	−1.73933	−0.40075	2.55940	7.65594	18.29937	29.73862
h	0	−0.00103	0.00048	0.00184	0.00219	0.00039	−0.00260	−0.00642	−0.01123	−0.01311
cg	0	0.00933	−0.00937	−0.02175	−0.01828	0.01651	0.06034	0.10752	0.14686	0.13160
df	0	0.01700	−0.01657	−0.03576	−0.02362	0.04640	0.12406	0.19389	0.20794	0.09578
e^2	0	0.01023	−0.01020	−0.02156	−0.01387	0.02788	0.07209	0.10787	0.09883	0.00803
c^2f	0	−0.04600	0.07442	0.12464	0.04811	−0.25595	−0.56524	−0.81931	−0.79570	−0.24501
cde	0	−0.14366	0.22074	0.34077	0.05220	−0.91516	−1.77648	−2.26940	−1.31805	1.58689
d^3	0	−0.03120	0.04431	0.06346	−0.00731	−0.21997	−0.39141	−0.45485	−0.12354	0.67346
c^3e	0	0.15345	−0.35265	−0.42236	0.15607	1.76170	3.04573	3.56562	1.33265	−4.18956
c^2d^2	0	0.30132	−0.63747	−0.68830	0.52403	3.60694	5.77533	5.99761	−0.47341	−13.97518
c^4d	0	−0.36979	1.09609	0.79893	−1.77731	−7.45279	−10.76740	−9.59341	6.35417	36.18663
c^6	0	0.10100	−0.40194	−0.12422	1.07691	3.38425	4.37679	3.02544	−5.79725	−20.92219

Adjustment	Coefficient	Polynomial in x	Divisor
VI	b^3	x	16
	b^2d	$5(x^3 - 3x)$	64
	b^2c^2	$-35(2x^3 - 5x)$	288
	bf	$-(x^5 - 10x^3 + 15x)$	288
	bce	$7(2x^5 - 17x^3 + 21x)$	360
	bd^2	$7(3x^5 - 24x^3 + 29x)$	768
	bc^2d	$-(14x^5 - 103x^3 + 107x)$	64
	bc^4	$11(252x^5 - 1688x^3 + 1511x)$	15552
	h	$x^7 - 21x^5 + 105x^3 - 105x$	40320
	cg	$-(2x^7 - 37x^5 + 160x^3 - 135x)$	5040
	df	$-(x^7 - 17x^5 + 69x^3 - 57x)$	1152
	e^2	$-(2x^7 - 33x^5 + 132x^3 - 108x)$	3600
	c^2f	$18x^7 - 293x^5 + 1100x^3 - 795x$	5184
	cde	$18x^7 - 273x^5 + 974x^3 - 695x$	1440
	d^3	$9x^7 - 131x^5 + 451x^3 - 321x$	3072
	c^3e	$-(396x^7 - 5708x^5 + 18755x^3 - 11811x)$	19440
	c^2d^2	$-(594x^7 - 8193x^5 + 26006x^3 - 16367x)$	13824
	c^4d	$5148x^7 - 67004x^5 + 195259x^3 - 109553x$	62208
	c^6	$-(154440x^7 - 1887684x^5 + 5033714x^3 - 2542637x)$	4199040

(a) χ^2 *distribution*

If n is the number of degrees of freedom

$$
\begin{aligned}
\chi^2 = {} & n \\
& + \sqrt{n}\,(x\sqrt{2}) \\
& + \tfrac{2}{3}(x^2 - 1) \\
& + \frac{1}{\sqrt{n}}\left(\frac{x^3 - 7x}{9\sqrt{2}}\right) \\
& - \frac{1}{n}\left(\frac{6x^4 + 14x^2 - 32}{405}\right) \\
& + \frac{1}{n\sqrt{n}}\left(\frac{9x^5 + 256x^3 - 433x}{4860\sqrt{2}}\right) \\
& + \frac{1}{n^2}\left(\frac{12x^6 - 243x^4 - 923x^2 + 1472}{25515}\right) \\
& - \frac{1}{n^2\sqrt{n}}\left(\frac{3753x^7 + 4353x^5 - 289517x^3 - 289717x}{9185400\sqrt{2}}\right).
\end{aligned}
\tag{3a}
$$

(b) *t distribution*

If n is the number of degrees of freedom

$$
\begin{aligned}
t = {} & x \\
& + \frac{1}{n}\left(\frac{x^3 + x}{4}\right) \\
& + \frac{1}{n^2}\left(\frac{5x^5 + 16x^3 + 3x}{96}\right) \\
& + \frac{1}{n^3}\left(\frac{3x^7 + 19x^5 + 17x^3 - 15x}{384}\right) \\
& + \frac{1}{n^4}\left(\frac{79x^9 + 776x^7 + 1482x^5 - 1920x^3 - 945x}{92160}\right) \\
& + \frac{1}{n^5}\left(\frac{27x^{11} + 339x^9 + 930x^7 - 1782x^5 - 765x^3 + 17955x}{368640}\right).
\end{aligned}
\tag{3b}
$$

The number of terms and the orders of magnitude differ in this expansion because it has been derived from the original expansion of Student's integral in powers of n^{-1} as given by Fisher (1926).

(c) *z distribution*

If n_1 and n_2 are the degrees of freedom, and

$$\sigma = \frac{1}{n_1} + \frac{1}{n_2} \quad \delta = \frac{1}{n_1} - \frac{1}{n_2},$$

then

$$
\begin{aligned}
z = {} & \sqrt{\frac{\sigma}{2}}\,(x) \\
& - \delta\left(\frac{x^2 + 2}{6}\right) \\
& + \sqrt{\frac{\sigma}{2}}\left\{\sigma\left(\frac{x^3 + 3x}{24}\right) + \frac{\delta^2}{\sigma}\left(\frac{x^3 + 11x}{72}\right)\right\} \\
& - \left\{\delta\sigma\left(\frac{x^4 + 9x^2 + 8}{120}\right) - \frac{\delta^3}{\sigma}\left(\frac{3x^4 + 7x^2 - 16}{3240}\right)\right\} \\
& + \sqrt{\frac{\sigma}{2}}\left\{\sigma^2\left(\frac{x^5 + 20x^3 + 15x}{1920}\right) + \delta^2\left(\frac{x^5 + 44x^3 + 183x}{2880}\right)\right. \\
& \left. + \frac{\delta^4}{\sigma^2}\left(\frac{9x^5 - 284x^3 - 1513x}{155520}\right)\right\} + \left\{\delta\sigma^2\left(\frac{4x^6 - 25x^4 - 177x^2 + 192}{20160}\right)\right. \\
& \left. + \delta^3\left(\frac{4x^6 + 101x^4 + 117x^2 - 480}{90720}\right) - \frac{\delta^5}{\sigma^2}\left(\frac{12x^6 + 513x^4 + 841x^2 - 2560}{1632960}\right)\right\}
\end{aligned}
\tag{3c}
$$

$$- \sqrt{\frac{\sigma}{2}}\left\{\sigma^3\left(\frac{x^7 + 7x^5 + 7x^3 + 105x}{21504}\right)\right.$$

$$+ \delta\sigma^2\left(\frac{801x^7 + 10511x^5 + 30151x^3 + 62241x}{4838400}\right)$$

$$- \frac{\delta^4}{\sigma}\left(\frac{477x^7 + 4507x^5 - 82933x^3 - 264363x}{43545600}\right)$$

$$\left. + \frac{\delta^6}{\sigma^3}\left(\frac{3753x^7 + 55383x^5 - 368897x^3 - 1213927x}{1175731200}\right)\right\}.$$

As an illustration of the accuracy, we may use the example given previously. When $n_1 = 24$ and $n_2 = 60$, the 5% value of z is 0.26534844, and the asymptotic expansion yields the following values:

Order of magnitude	Successive terms	Successive totals	Successive errors
0	0.2809 1224	0.2809 1224	0.0155 6380
1	— 196 0643	2613 0581	— 40 4263
2	44 6851	2657 7432	4 2588
3	— 4 8004	2652 9428	— 5416
4	5645	2653 5073	229
5	— 154	2653 4919	75
6	— 102	2653 4817	— 27

The numerical values of the polynomials in x occurring in the above formulae are given in Table III.

IV. Examples of the Types of Problem to Which the Expansion Has Been Applied

(a) *The asymptotic approach to Behrens' integral*

Fisher (1926) developed the ordinate and integral of Student's distribution in a series of powers of n^{-1}, giving the polynomial coefficients so far as the fifth adjustment. The purpose of this expansion was to supply sufficiently accurate values of the probabilities corresponding to any values of t for values of n beyond the range which it was proposed to tabulate.

With Behrens' extension of Student's test there were even stronger reasons for using a similar method. The direct calculations carried out by Sukhatme (1938) are very much more laborious than those needed for Student's integral. At any single level of significance, various values are required for three parameters, provided by the two numbers of degrees of freedom of the two samples, and the estimated ratio of the variances of the two means. For functions of many variables, there is a great advantage in the use of explicit formulae in which the several variables may be substituted, and there is much to be gained by extending the use of such formulae over regions too extensive for complete tabulation. Finally, it should be noted that the logical situation in which we would prefer to rely on the separate estimates of variances from the two samples rather than on any process of pooling these estimates, is of more frequent occurrence with large samples than with small, and is particularly applicable to cases,

TABLE III—*Numerical values of the polynomials in the expansions for χ^2, t and z.*

(a) χ^2

p									
.5	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
0	0.9538726	1.8123876	2.3261743	2.7718076	3.2899527	3.6427727	3.9697452	4.3702484	4.6535075
−0.666667	−0.363376	0.428250	1.137029	1.894306	2.941263	3.756598	4.586292	5.699690	6.551711
0	−0.346842	−0.539450	−0.554981	−0.486382	−0.290266	−0.073888	0.193953	0.619906	0.989534
0.07901	0.06022	−0.01772	−0.12296	−0.27240	−0.54197	−0.80252	−1.11315	−1.60211	−2.03211
0	−0.0309	0.0022	0.0779	0.1948	0.4116	0.6228	0.8752	1.2735	1.6249
0.0577	0.0393	−0.0253	−0.1006	−0.1952	−0.3425	−0.4642	−0.5886	−0.7467	−0.8535
0	0.012	0.073	0.122	0.170	0.203	0.183	0.100	−0.145	−0.469

The values in the columns are in the same order as the polynomials in expansion 3a. Sufficient figures have been retained to ensure accuracy in the fourth decimal place for $n > 30$, except for $x\sqrt{2}$ which should be taken more accurately for $n > 1600$.

(b) t

p									
.5	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
0	0.24533	0.84658	1.52377	2.37227	3.72907	4.91655	6.23122	8.15013	9.72973
0	0.0795	0.5709	1.4202	2.8225	5.7197	8.8348	12.8509	19.6925	26.1330
0	−0.005	0.259	0.983	2.556	6.719	12.144	20.221	36.154	53.169
0	0	0.1	0.4	1.6	5.6	12.1	23.2	48.6	79.4

The values in the columns are in the same order as the polynomials in expansion 3b. Sufficient figures have been retained to ensure accuracy in the fourth decimal place for $n > 30$.

TABLE III—*Continued*

(c) z

p									
.5	.25	.10	.05	.025	.01	.005	.0025	.001	.0005
0	0.67448975	1.28155157	1.64485363	1.95996398	2.32634787	2.57582930	2.80703377	3.09023231	3.29052673
0.33333333	0.40915607	0.60706240	0.78425724	0.97357647	1.23531574	1.43914943	1.64657310	1.92492262	2.13792770
0	0.0970966	0.2478934	0.3910327	0.5587089	0.8153747	1.0340770	1.2724563	1.6158742	1.8958323
0	0.1073089	0.2250258	0.3131057	0.4040101	0.5302747	0.6308956	0.7360447	0.8819839	0.9975582
0.0666667	0.1025116	0.2123230	0.3305821	0.4777495	0.7166304	0.9311327	1.1750042	1.5428288	1.8557024
−0.004938	−0.003764	0.001108	0.007685	0.017025	0.033873	0.050157	0.069572	0.100132	0.127007
0	0.008539	0.033737	0.065478	0.108805	0.184807	0.257207	0.343093	0.478317	0.597758
0	0.047595	0.114789	0.176687	0.249610	0.363825	0.464148	0.576788	0.745061	0.887356
0	−0.00711	−0.01611	−0.02343	−0.03114	−0.04168	−0.04971	−0.05761	−0.06765	−0.07475
0.00952	0.00529	−0.00736	−0.01938	−0.03126	−0.04286	−0.04537	−0.03958	−0.01462	0.02094
−0.00529	−0.00447	0	0.00722	0.01859	0.04128	0.06515	0.09556	0.14695	0.19516
0	0	0	0	0	0	0.0178	0.0256	0.0384	0.0502
0	0.00344	0.00833	0.01491	0.02660	0.05478	0.09004	0.14149	0.24158	0.34748
0	0.0109	0.0381	0.0804	0.1534	0.3174	0.5105	0.7799	1.2814	1.7939
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0

The values in the columns are in the same order as the polynomials in expansion 3c. Sufficient figures have been retained to ensure accuracy in the sixth decimal place for $n_1 > 24$ and $n_2 > 60$.

such as arise in Physics and Astronomy, in which we wish to compare estimates of the value of the same quantity (a) from relatively ample data of low intrinsic accuracy, and (b) from a small series of observations of relatively high precision. When, as often happens, the estimates of precision of the means obtained in these two ways are of the same order of magnitude, the only satisfactory test is that based on Behrens' solution. The asymptotic expansion is particularly suitable for evaluating the percentiles for this special application. There were thus four manifest advantages of the asymptotic approach to Behrens' integral (Fisher 1941):

(i) a check on Sukhatme's values, obtained by a completely independent method, and applicable at least for the higher values of n_1 and n_2,

(ii) greater accuracy than could be obtained for percentiles from Sukhatme's table for values of n_1 and n_2 greater than 12,

(iii) a wider range of levels of significance in the region to which the asymptotic expansion is applicable,

(iv) the theoretical guidance offered by the algebraic form of the leading terms of the expansion.

(b) *The fiducial distribution of the binomial parameter*, p

When discussing the application of the fiducial argument to discontinuous observations, Fisher (1959) found that the mean of the fiducial distribution of p, the parameter of the binomial distribution, for given observational frequencies a, b out of N, was

$$\bar{p} = \frac{a}{N} + \frac{b-a}{2N^2} - \frac{3(b-a)}{2N^3} + \frac{15(b-a)}{2N^4} - \cdots \tag{4}$$

if χ were taken to be normally distributed.

On the other hand, the mean of the Bayesian distribution *a posteriori*, using the Bayesian probability *a priori*

$$\frac{1}{\pi\sqrt{pq}}\,dp,$$

was

$$\frac{a+\frac{1}{2}}{N+1} = \frac{a}{N} + \frac{b-a}{2N^2} - \frac{b-a}{2N^3} + \frac{b-a}{2N^4} - \cdots . \tag{5}$$

Asymptotic agreement between these means appears when allowance is made for the effects of departure from normality in the binomial distribution, which are appreciable in the expression (4).

Direct application of the asymptotic expansion, using the six adjustment terms gave the following expansion for the binomial variate in terms of the normal deviate x

$$a = pN + x\sqrt{Npq} + \tfrac{1}{6}(q-p)(x^2-1) + \frac{1}{72\sqrt{Npq}}\{-x^3 + x - pq(2x^3 - 14x)\} + \cdots$$

which, after inversion, gave*

$$p = \frac{a}{N} - \frac{x\sqrt{ab}}{N^{3/2}} + \frac{(b-a)(2x^2+1)}{6N^{2}} + \left\{\frac{(-2N^2+26ab)x^3+(-7N^2+34ab)x}{72N^{5/2}\sqrt{ab}}\right\} + \cdots \tag{6}$$

whence, substituting its average value for each power of x, the mean of the fiducial distribution is

$$\bar{p} = \frac{a}{N} + \frac{b-a}{2N^2} - \frac{b-a}{2N^3} + \frac{b-a}{2N^4} - \cdots \tag{7}$$

agreeing so far as the fourth term with (5).

The expansion (6) also provides a ready means for comparing the two distributions with respect to other properties. For example, although the means are in agreement, the asymptotic fiducial distribution has the higher variance.

See also the alternative treatment in Fisher (1957).

(c) *Quantitative inheritance*

Panse (1940) has described a statistical technique for the study of quantitative inheritance, in which genetic models, based on data from the F_2 and F_3 generations, are used to represent the constitution of particular characters. The statistical consequences in the population, corresponding to these models were assessed, using the cumulant function of the joint distribution of the F_2 phenotypic value, the mean of the F_3 progeny, and the genotypic variance of F_3 progeny. These functions provided the data for expressing an attribute of the F_3 progeny in terms of the F_2 phenotypic values, and thus the effects of selection in the F_2 phenotype on the mean value of the F_3 progeny could be determined by integration over the F_2 distribution. When the intensity of selection was assigned, the limits of integration were calculable from the asymptotic expansion of the deviate.

V. A Class of Distributions with a Finite Condensation at Zero

The Poisson Series, a discontinuous distribution of positive integers, is well known to have the simple series of cumulants

$$\kappa_r = m,$$

for all values of r. Correspondingly, the cumulative function is

$$K = m(e^{it} - 1),$$

and the characteristic function

$$M = \exp\{m(e^{it} - 1)\}.$$

* For the remaining terms see Fisher (1959)

It is less well known that if, for all values of r,

$$\kappa_r = r!m,$$

or, if a scaling factor be introduced,

$$\kappa_r = r!a^r m,$$

the distribution, derived from

$$K = \frac{mait}{1 - ait}$$

and

$$M = e^{-m}e^{m/(1-ait)}$$

$$= e^{-m} \sum_0^\infty \frac{m^n}{n!}(1 - ait)^{-n}$$

is continuous over the range of positive values, with a finite condensation at zero. For $(1 - ait)^{-n}$ is the characteristic function of the Eulerian distribution

$$\frac{1}{(n-1)!}\left(\frac{X}{a}\right)^{n-1} e^{-X/a}\frac{dX}{a}$$

or of

$$\chi^2 = \frac{2X}{a},$$

for $2n$ degrees of freedom, or of the sum of n random variables, each distributed as

$$e^{-x/a}\frac{dx}{a}, \; x \geq 0, \text{ a positive}$$

Hence the distribution is that of the sum of a number of such variables, when the number is distributed in a Poisson Series of parameter m. The variate is, therefore, never negative, but is zero with finite frequency

$$e^{-m}.$$

Over the range of positive values, the distribution is continuous, and can be expressed as

$$\sqrt{\frac{m}{xa}}\, e^{-m-(x/a)} I_1\left(2\sqrt{\frac{xm}{a}}\right) dx$$

where I_1 is a Bessel function, specified by

$$I_1(u) = \tfrac{1}{2}u + \frac{1}{2^2.4}u^3 + \frac{1}{2^2.4^2.6}u^5 + \cdots$$

The distribution was first recognized (Bennett 1954, Fisher 1954) as characteristic of that of the length of germ plasm still heterogenic at an advanced stage of inbreeding, but its intrinsic incorporation of a finite condensation at zero makes it appropriate to a number of natural phenomena, a good illustration being the rainfall of an arid region. A continuous model for rainfall is unsatisfactory for such regions. For many localities it is preferable to use a model ascribing the total rainfall for a given period, for example, a month or a year, to a number of showers, the number being a random sample from a Poisson series with parameter m, the rainfall of the showers having positive values only, representable by the Eulerian distribution

$$\frac{1}{p!}\,x^p e^{-x}\,dx$$

where p can be small, as in the previous example where it is actually 0, but must be > -1.

The advantage of this type of distribution, relevant for the purposes of the water engineer, is that there is a finite probability, namely e^{-m}, of no rain, whereas a continuous distribution would make this probability zero, contrary to experience.

For the asymptotic expansion, when m is sufficiently large, we may take the exact values for the mean (m) and variance ($2m$) and for the measures of non-normality

$$c = \frac{3}{\sqrt{2m}} \qquad f = \frac{90}{m^2}$$

$$d = \frac{6}{m} \qquad g = \frac{315\sqrt{2}}{m^2\sqrt{m}}$$

$$e = \frac{15\sqrt{2}}{m\sqrt{m}} \qquad h = \frac{2520}{m^3}.$$

These yield the six adjustments, to the normal deviate x

$$\text{I}\ \frac{x^2-1}{\sqrt{8m}} \qquad \text{IV}\ -\frac{4x^3-x}{384m^2}$$

$$\text{II}\ \frac{-x}{8m} \qquad \text{V}\ \frac{3x^4+2x^2-11}{480m^2\sqrt{2m}}$$

$$\text{III}\ \frac{x^2-1}{24m\sqrt{2m}} \qquad \text{VI}\ -\frac{96x^5+164x^3-767x}{46080m^3}.$$

The coefficients of Table IV give a rather comprehensive tabulation of the distribution, when m is sufficiently large for the accuracy required, and the levels of significance of interest. These large values of m would be troublesome to use in direct evaluation. For sufficiently small values of m, however, the probability that the variate x exceeds any limit X may be evaluated as

$$P = e^{-m}e^{-X} \sum_{i=0}^{\infty} \frac{X^i}{i!} \sum_{j>i} \frac{m^j}{j!}$$

which may be recognized as the probability that a random Poisson variate with parameter m shall exceed a random Poisson variate with parameter X.

Table IV—*Coefficients of powers of* $(2m)^{-1/2}$

p	x	$\frac{x^2 - 1}{2}$	$-\frac{x}{4}$	$\frac{x^2 - 1}{12}$	$-\frac{4x^3 - x}{96}$	$\frac{3x^4 + 2x^2 - 11}{120}$	$-\frac{96x^5 + 164x^3 - 767x}{5760}$
.5	0	−0.50000	0	−0.08333	0	−0.09167	0
.25	0.67449	−0.27253	−0.16862	−0.04542	−0.00576	−0.07891	−0.09486
.10	1.28155	0.32119	−0.32039	0.05353	−0.07435	0.00314	0.05311
.05	1.64485	0.85277	−0.41121	0.14213	−0.16829	0.13642	−0.10835
.025	1.95996	1.42073	−0.48999	0.23679	−0.29330	0.34128	0.43543
.01	2.32635	2.20595	−0.58159	0.36766	−0.50035	0.73075	1.18428
.005	2.57583	2.81745	−0.64396	0.46957	−0.68527	1.11946	2.03348
.0025	2.80703	3.43972	−0.70176	0.57329	−0.89234	1.59180	3.16056
.001	3.09023	4.27477	−0.77256	0.71246	−1.19741	2.34733	−5.12555
.0005	3.29053	4.91378	−0.82263	0.81896	−1.45024	3.01970	7.00573

References

J. H. Bennett (1954). The distribution of heterogeneity upon inbreeding. J. Roy. stat. soc. B, 16: 88-99.

E. A. Cornish and R. A. Fisher (1937). Moments and cumulants in the specification of distributions. *Rev. de l'Inst. int. de stat. 5:* 307–22.

R. A. Fisher (1926). Expansion of "Student's" integral in powers of n^{-1}. *Metron 5:* 109–12.

R. A. Fisher (1941). The asymptotic approach to Behrens' integral, with further tables for the *d* test of significance. *Ann. eugen. 11:* 141–72.

R. A. Fisher (1954). A fuller theory of "junctions" in inbreeding. *Heredity 8:* 187–97.

R. A. Fisher (1957). The underworld of probability. *Sankhya 18:* 201–10.

R. A. Fisher (1959). "Statistical Methods and Scientific Inference." 2nd ed. Oliver and Boyd, Edin.

H. Goldberg and H. Levine (1946). Approximate formulas for the percentage points and normalization of t and χ^2. *Ann. math. statist. 17:* 216–25.

N. L. Johnson and B. L. Welch (1939). Applications of the non-central t-distribution. *Biometrika 31:* 362-89.

V. G. Panse (1940). A statistical study of quantitative inheritance. *Ann. eugen. 10:* 76–105.

P. V. Sukhatme (1938). On Fisher and Behrens' test of significance for the difference in means of two normal samples. *Sankhya 4:* 39–48.