THE NON-CENTRAL χ^2- AND F-DISTRIBUTIONS AND THEIR APPLICATIONS†

BY P. B. PATNAIK, *University College, London*

1. INTRODUCTORY

In the Neyman-Pearson theory of testing statistical hypotheses, the efficiency of a statistical test is to be judged by its power of detecting departures from the null hypothesis. Thus besides knowing the random sampling distribution of a given statistic T under this hypothesis, say H_0, it is also necessary to know the distribution of T under admissible hypotheses alternative to H_0. Hence the power function of the test is obtained. In the case of the well-known tests using χ^2, t and F, the evaluation of their power functions involves the use of what have been called non-central distributions. For example, if we are applying the t-test to examine if a sample has come from a normal population with mean $\mu = 0$ (H_0), we know that under H_0, t has a 5 % chance of exceeding the 5 % point of its distribution. But in order to compute the power of the test we wish to know the chance that t exceeds this point when μ has alternative values, not equal to zero. This chance is given by the non-central t-integral. This distribution has been studied by Fisher (1931), Neyman (1935), Neyman & Tokarska (1936) and Johnson & Welch (1939). In a similar way, the non-central χ^2- and F-distributions arise in consideration of the power functions of the χ^2- and variance-ratio tests.

The power function may be used either to determine the extent of the departures from H_0 in a given direction, which will be detected as significant (at a prescribed level) with a given probability, or it may be used to determine in advance the size of experiment necessary to ensure that a worth-while difference will be established as significant, if it exists. But apart from its value in this connexion, the study of non-central distributions is of considerable interest. The mathematical forms of these distributions of t, χ^2 and F have been long known, but their use without extensive tabling has not been easy. The present paper is therefore concerned with two lines of investigation:

(*a*) The derivation of certain approximations to the probability integrals of (i) non-central χ^2, and (ii) the ratio of non-central χ^2 to an independent central χ^2, which we have termed non-central F. These approximations, depending on tabled functions, permit easy calculation.

(*b*) Discussion of the ways in which these distributions may be used in connexion with the power functions of statistical tests.

2. THE NON-CENTRAL χ^2-DISTRIBUTION

2·1. *Geometrical derivation*

As is well known, the statistic χ^2 is defined as the sum of squares of (say) n independent random deviates, ξ_i, all drawn from a normal population with mean, 0, and standard deviation, σ, viz.

$$\chi^2 = \sum_{i=1}^{n} \xi_i^2/\sigma^2.$$

† Part of a thesis approved for the degree of Ph.D. of the University of London.

If, however, the mean ξ_i is a_i and we write

$$x_i = \xi_i - a_i,$$

then we have the non-central χ^2 defined by

$$\chi'^2 = \sum_{i=1}^{n} (x_i + a_i)^2/\sigma^2.$$

The probability distribution of χ'^2 has been obtained by Fisher (1928) as a particular case of the distribution of the multiple correlation coefficient. A purely analytical proof was given by Tang (1938). As χ'^2 is a generalized form of χ^2 it may be of interest to compare its geometrical representation with the familiar geometry of χ^2. We therefore give a direct geometrical derivation of the χ'^2-distribution.

Without loss of generality we shall assume in what follows that $\sigma = 1$, so that the probability law of x is given by

$$p(x) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}x^2}. \tag{1}$$

Then
$$\chi'^2 = \sum_{i=1}^{n} \xi_i^2.$$

In the n-dimensional space of the ξ's, suppose O is the origin, P the point $(\xi_1, \ldots, \xi_n)$, A the point $(a_1, \ldots, a_n)$, $\angle POA = \theta$ and M the foot of the perpendicular from P on OA as shown in Fig. 1. Then

$$OP^2 = \chi'^2, \quad OA^2 = \sum_{i=1}^{n} a_i^2 = \lambda, \quad \text{say}.$$

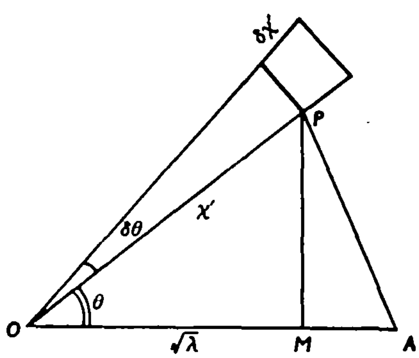

Fig. 1

From (1), the probability density at P is proportional to

$$\exp\left[-\frac{1}{2}\sum_{i=1}^{n}(\xi_i - a_i)^2\right] = \exp\left[-\tfrac{1}{2}PA^2\right] = \exp\left[-\tfrac{1}{2}(\chi'^2 + \lambda - 2\chi'\sqrt{\lambda}\cos\theta)\right]. \tag{2}$$

If we keep OP and θ fixed, P describes an $(n-1)$-dimensional sphere of radius $PM = \chi'\sin\theta$ with its surface area proportional to $(\chi'\sin\theta)^{n-2}$. If χ' is increased to $\chi' + d\chi'$ and θ to $\theta + d\theta$, then a disk of area $\chi' d\chi' d\theta$ moves round this surface and hence covers a volume proportional to

$$(\chi'\sin\theta)^{n-2}\chi' d\chi' d\theta.$$

To obtain the distribution of χ' alone, we integrate out θ. Thus

$$p(\chi')\,d\chi' = C\int_0^{\pi} e^{-\frac{1}{2}(\chi'^2+\lambda-2\chi'\ \lambda\cos\theta)}(\chi'\sin\theta)^{n-2}\chi'\,d\theta\,d\chi',$$

which is equivalent to

$$p(\chi'^2)\,d\chi'^2 = \frac{C}{2} e^{-\frac{1}{2}(\chi'^2+\lambda)} (\chi'^2)^{\frac{1}{2}n-1}\,d\chi'^2 \times \int_0^{\frac{1}{2}\pi} (e^{-\sqrt{\lambda}\chi'\cos\theta} + e^{\sqrt{\lambda}\chi'\cos\theta}) \sin^{n-2}\theta\,d\theta. \tag{3}$$

Expanding the integrand and integrating term by term, we find

$$p(\chi'^2) = \tfrac{1}{2} C\, e^{-\frac{1}{2}(\chi'^2+\lambda)} (\chi'^2)^{\frac{1}{2}n-1} \frac{\Gamma(\frac{1}{2})\,\Gamma[\frac{1}{2}(n-1)]}{\Gamma(\frac{1}{2}n)} \left\{1 + \frac{1}{n}\left(\frac{\chi'^2\lambda}{2}\right) + \frac{1}{n(n+2)\,2!}\left(\frac{\chi'^2\lambda}{2}\right)^2 + \dots\right\}.$$

If zero is substituted for λ, this reduces to the ordinary χ^2-distribution which therefore gives us the value of C.

We then have
$$p(\chi'^2) = \frac{e^{-\frac{1}{2}\chi'^2} e^{-\frac{1}{2}\lambda}}{2^{\frac{1}{2}n}} \sum_{j=0}^{\infty} \frac{(\chi'^2)^{\frac{1}{2}n+j-1}\lambda^j}{\Gamma(\frac{1}{2}n+j)2^{2j}\,.\,j!}. \tag{4}$$

2·2. *Derivation through a transformation of variates*

Next we will show that it is possible to effect a variate transformation so as to transform χ'^2 into a sum of $(n-1)$ central squares and a single non-central square and then derive its distribution. Make the following orthogonal transformation:

$$\left.\begin{aligned} y_1 &= c_{11}\xi_1 + c_{12}\xi_2 + \dots + c_{1n}\xi_n, \\ &\dots\dots\dots\dots\dots\dots \\ y_n &= c_{n1}\xi_1 + c_{n2}\xi_2 + \dots + c_{nn}\xi_n. \end{aligned}\right\} \tag{5}$$

$$\left.\begin{array}{ll} \text{Then} & \sum_1^n \xi_i^2 = \sum_1^n y_j^2. \\ \text{Generally, if} & \mathscr{E}(y_j) = c_{j1}a_1 + c_{j2}a_2 + \dots + c_{jn}a_n = b_j \quad (j = 1 \text{ to } n), \\ \text{we have} & \sum_1^n a_i^2 = \sum_1^n b_j^2, \\ \text{and} & \sum_1^n a_i\xi_i = \sum_1^n b_j y_j. \end{array}\right\} \tag{6}$$

Now we can make

$$b_1 = b_2 = \dots = b_{n-1} = 0 \quad \text{and} \quad b_n = \sqrt{(\Sigma a_i^2)} = \sqrt{\lambda}.$$

Thus $\chi'^2 = \sum_1^n \xi_i^2$ is distributed as $\sum_1^{n-1} y_j^2 + y_n^2$, the sum of the squares of $(n-1)$ normal variates with mean zero and the square of a single normal variate with mean $\sqrt{\lambda}$, the S.D.'s being unity.

Writing
$$\chi'^2 = w, \quad \sum_1^{n-1} y_j^2 = u \quad \text{and} \quad y_n^2 = v,$$

we see that u has a χ^2-distribution with $(n-1)$ degrees of freedom, that is,

$$p(u) = \frac{e^{-\frac{1}{2}u} u^{\frac{1}{2}(n-3)}}{2^{\frac{1}{2}(n-1)}\Gamma[\frac{1}{2}(n-1)]},$$

and that v follows the law

$$\begin{aligned} p(v) &= \frac{v^{-\frac{1}{2}}}{2\sqrt{(2\pi)}}\{e^{-\frac{1}{2}(\sqrt{v}-\sqrt{\lambda})^2} + e^{-\frac{1}{2}(-\sqrt{v}-\sqrt{\lambda})^2}\} \\ &= \frac{1}{2^{\frac{1}{2}}\Gamma(\frac{1}{2})} e^{-\frac{1}{2}(v+\lambda)} v^{-\frac{1}{2}}\left(1 + \frac{v\lambda}{2!} + \frac{(v\lambda)^2}{4!} + \dots\right). \end{aligned}$$

Hence, replacing v by $(w-u)$ in the joint probability law $p(u,v)$, we have

$$p(u,w) = \frac{e^{-\frac{1}{2}w} e^{-\frac{1}{2}\lambda} w^{\frac{1}{2}(n-4)}}{2^{\frac{1}{2}n}\Gamma(\frac{1}{2})\Gamma[\frac{1}{2}(n-1)]}\left(\frac{u}{w}\right)^{\frac{1}{2}(n-3)}\left\{\left(1-\frac{u}{w}\right)^{-\frac{1}{2}} + \frac{w\lambda}{2!}\left(1-\frac{u}{w}\right)^{\frac{1}{2}} + \dots\right\}.$$

Whence integrating with respect to u from 0 to w, we obtain

$$p(w) = \frac{e^{-\frac{1}{2}w}\, e^{-\frac{1}{2}\lambda} w^{\frac{1}{2}(n-2)}}{2^{\frac{1}{2}n}\,\Gamma(\frac{1}{2})\,\Gamma[\frac{1}{2}(n-1)]}\left\{B\left(\frac{n-1}{2},\frac{1}{2}\right)+\frac{w\lambda}{2!}\,B\left(\frac{n-1}{2},\frac{3}{2}\right)+\ldots\right\},$$

that is,
$$p(\chi'^2) = \frac{e^{-\frac{1}{2}\chi'^2}\, e^{-\frac{1}{2}\lambda}(\chi'^2)^{\frac{1}{2}n-1}}{2^{\frac{1}{2}n}\,\Gamma(\frac{1}{2}n)}\left\{1+\frac{1}{n}\left(\frac{\chi'^2\lambda}{2}\right)+\frac{1}{n(n+2)\,.\,2!}\left(\frac{\chi'^2\lambda}{2}\right)^2+\ldots\right\}, \tag{7}$$

which is seen to be the same as (4).

In this distribution of χ'^2, n may be called the number of degrees of freedom and λ, which is equal to the sum of the squares $\sum\limits_1^n a_i^2$, the non-central parameter.

2·3. *Conditional distribution of χ'^2 under linear constraints*

Suppose the ξ's are subject to k ($<n$) linear constraints. These can be transformed into an orthogonal set represented, say, by the equations

$$\sum_{i=1}^n c_{ji}\xi_i = \rho_j \quad (j = 1, \ldots, k), \tag{8}$$

where
$$\sum_{i=1}^n c_{ji}^2 = 1, \quad \sum_{i=1}^n c_{ji}c_{li} = 0 \quad (j \neq l).$$

We make an orthogonal transformation of variates defined by the equations (5), so that $\Sigma\xi_i^2$ transforms to Σy_j^2 and the k constraints of (8) become simply $y_1 = \rho_1, \ldots, y_k = \rho_k$. To find the distribution of Σy_j^2 subject to these conditions, we first see that, in virtue of the relations in (6), the joint probability law of the ξ's

$$p(\xi_1, \ldots, \xi_n) = C\exp\left\{-\frac{1}{2}\sum_1^n(\xi_i - a_i)^2\right\}$$

transforms into
$$p(y_1, \ldots, y_n) = C\exp\left\{-\frac{1}{2}\sum_1^n(y_j - b_j)^2\right\}.$$

When $y_1, \ldots, y_k$ take respectively the constant values $\rho_1, \ldots, \rho_k$, we have the conditional probability law

$$p(y_{k+1}, \ldots, y_n \mid \rho_1, \ldots, \rho_k) = C_1\exp\left\{-\frac{1}{2}\sum_{k+1}^n (y_j - b_j)^2\right\}. \tag{9}$$

It can be shown from (9), as in § 2·1, that the sum of the non-central squares $(y_{k+1}^2 + \ldots + y_n^2)$ is distributed as a χ'^2 with $(n-k)$ degrees of freedom and parameter

$$\lambda = b_{k+1}^2 + \ldots + b_n^2.$$

From (6) we see that
$$y_{k+1}^2 + \ldots + y_n^2 = \Sigma\xi_i^2 - (\rho_1^2 + \ldots + \rho_k^2)$$

and
$$b_{k+1}^2 + \ldots + b_n^2 = \Sigma a_i^2 - (b_1^2 + \ldots + b_k^2)$$
$$= \sum_{i=1}^n a_i^2 - \sum_{j=1}^k\left(\sum_{i=1}^n a_i c_{ji}\right)^2. \tag{10}$$

Hence $\left(\sum\limits_1^n \xi_i^2 - \sum\limits_1^k \rho_j^2\right)$ is distributed as a χ'^2 with $(n-k)$ degrees of freedom and parameter λ, given by the expression in (10).

In particular, if there is only a single constraint on the ξ's, given by

$$\sum_1^n c_i \xi_i = \rho, \quad \sum_1^n c_i^2 = 1, \tag{11}$$

then $\left(\sum_1^n \xi_i^2 - \rho^2\right)$ follows a χ'^2-distribution with $(n-1)$ degrees of freedom and

$$\lambda = \sum_1^n a_i^2 - \left(\sum_1^n a_i c_i\right)^2. \tag{12}$$

3. Approximations to the χ'^2-distribution

3·1. *The χ^2-approximation*

Fisher (1928) has shown that the distribution function of χ'^2 given by (4) can be expressed in terms of a Bessel function with imaginary argument. When n, the number of degrees of freedom, is odd, this can be reduced to elementary functions. When n is even, we see that the probability integral

$$\int_{\chi'^2}^{\infty} p(\chi'^2)\, d\chi'^2$$

can be expressed as a double Poisson sum. However, in both cases, the labour of calculating the probability integral is considerable.

In his paper, Fisher has given a table of the upper 5 % significance points of the χ'^2-distribution for $n = 1$ to 7 and $\sqrt{\lambda} = 0$ (0·2) 5·0. Garwood† has an unpublished table of the lower 5 % points for the same range of values of n and λ. No tables of the probability integral are available. It may therefore be useful to have an easy method of determining the probability integral and percentage points sufficiently accurately for any given values. For this purpose we shall consider several approximations to the distribution of χ'^2.

The characteristic function of this distribution is easily seen to be

$$\phi(t) = \exp\left\{\frac{\lambda i t}{1-2it}\right\} \Big/ (1-2it)^{\frac{1}{2}n}.$$

Hence we have the following cumulants:

$$\left.\begin{aligned} \kappa_1 &= n+\lambda, & \kappa_2 &= 2(n+2\lambda), \\ \kappa_3 &= 8(n+3\lambda), & \kappa_4 &= 48(n+4\lambda), \end{aligned}\right\} \tag{13}$$

the general rth cumulant being

$$\kappa_r = 2^{r-1}(r-1)!\,(n+r\lambda).$$

In the β_1, β_2 diagram, it was found that the point computed from the above κ's moved close to and above the Type III line, and this suggested that we might fit a Type III distribution from the first two moments. This is given by

$$f(y) = \frac{e^{-\frac{1}{2}y}\, y^{\frac{1}{2}\nu-1}}{2^{\frac{1}{2}\nu}\,\Gamma(\frac{1}{2}\nu)}, \tag{14}$$

where

$$y = \chi'^2/\rho,$$

$$\rho = \frac{n+2\lambda}{n+\lambda} = 1 + \frac{\lambda}{n+\lambda}, \quad \nu = \frac{(n+\lambda)^2}{(n+2\lambda)} = n + \frac{\lambda^2}{n+2\lambda}. \tag{15}$$

This means that we are representing the distribution of (χ'^2/ρ) by that of χ^2 with ν degrees of freedom, ν being in general a fraction.

† I am grateful to Dr F. Garwood for kindly making his table available to me for reference.

In what follows we shall write x for χ'^2, $p(x)$ for the true distribution of χ'^2 with n degrees of freedom and parameter λ and $f(x)$ for the approximation to $p(x)$ obtained by assuming that $x/\rho = y$ is distributed as χ^2 with ν degrees of freedom.

Then the probability integral $$\int_0^x p(x)\,dx = \int_0^y p(y)\,dy$$

is approximately given by $$\int_0^y f(y)\,dy.$$

This integral can be expressed in the notation of the tables of the *Incomplete* Γ-*function* (K. Pearson, 1922) as $I(u,p)$, where

$$u = \frac{y}{\sqrt{(2\nu)}} = \frac{x}{\sqrt{[2(n+2\lambda)]}}, \quad p = \frac{\nu}{2} - 1 = \frac{(n+\lambda)^2}{2(n+2\lambda)} - 1, \tag{16}$$

and could be evaluated by interpolation in these tables. For interpolation u-wise the second differences with Everett interpolation coefficients may be used, while linear interpolation p-wise seems adequate.

The approximations to the probability integral so obtained for certain values of n, λ and x are shown in Table 1 for comparison with the exact values. In some of these cases x is the upper 5% point (Fisher) or the lower 5% point (Garwood), so that the exact values are 0·95 or 0·05. The others are directly computed. For many purposes, especially in connexion with power functions, the degree of accuracy given by this method may be considered quite adequate.

Table 1. *Showing exact and approximate values of the χ'^2-integrals,* $\int_0^x p(x)\,dx$

n	λ	x	Approx.	Exact
4	4	1·765	0·0399	0·0500
	4	10·000	0·7191	0·7118
	4	17·309	0·9492	0·9500
	4	24·000	0·9913	0·9925
	10	10·000	0·3178	0·3148
7	1	4·000	0·1621	0·1628
	1	16·004	0·9499	0·9500
	16	10·257	0·0430	0·0500
	16	24·000	0·5947	0·5898
	16	38·970	0·9482	0·9500
12	6	24·000	0·8187	0·8174
	18	24·000	0·2936	0·2901
16	8	30·000	0·7895	0·7880
	8	40·000	0·9626	0·9632
	32	30·000	0·0590	0·0609
	32	60·000	0·8329	0·8316
24	24	36·000	0·1556	0·1567
	24	48·000	0·5333	0·5296
	24	72·000	0·9656	0·9667

To find the percentage points of the χ'^2 distribution, we first interpolate in the appropriate percentage point tables of the χ^2 (e.g. Thompson, 1941) for ν degrees of freedom and then multiply the interpolate by ρ. Four-point Lagrangian interpolation formulae may be used. The approximate upper and lower 5 % points obtained by this method for certain values of n and λ are given in Table 2, along with the exact values. Clearly the accuracy is not as good for the lower points as for the upper ones. Although the comparisons have had to be confined only to small values of n, since Fisher and Garwood have only given exact percentage points up to $n = 7$, from the closeness of the probability integral approximation (Table 1) we could still expect that the approximation to the percentage points would be fairly close for higher n.

These approximations based on the χ^2 fit will be referred in subsequent sections as the *first approximation*.

Table 2. *Showing exact and approximate values of the percentage points of the χ'^2-distribution*

n	λ	Upper 5 % point		Lower 5 % point	
		Approx.	Exact	Approx.	Exact
2	1	8·63	8·64	0·20	0·17
	4	14·72	14·64	0·94	0·65
	16	33·35	33·06	6·89	6·32
	25	45·66	45·31	12·68	12·08
4	1	11·72	11·71	0·93	0·91
	4	17·38	17·31	1·95	1·77
	16	35·69	35·43	8·36	7·88
	25	47·94	47·61	14·26	13·73
7	1	16·01	16·00	2·51	2·49
	4	21·28	21·23	3·78	3·66
	16	39·16	38·97	10·64	10·26
	25	51·34	51·06	16·68	16·23

3·2. *The normal approximation*

It is known that, for $n > 30$, Fisher's approximation, that $\sqrt{(2\chi^2)}$ is distributed as a normal variate $N(\sqrt{(2n-1)}, 1)$,† will give fairly close values to the probability integral and percentage points of the χ^2-distribution. It can be shown that a similar normal approximation is available for the χ'^2-distribution for large values of n or λ.

First we shall show that χ' approaches normality with greater rapidity than χ'^2.

If x is written for χ'^2, and x_0 is mean x, we have by Taylor's theorem

$$x^{\frac{1}{2}} = x_0^{\frac{1}{2}} + \tfrac{1}{2}(x-x_0)\,x_0^{-\frac{1}{2}} - \tfrac{1}{8}(x-x_0)^2\,x_0^{-\frac{3}{2}} + \tfrac{1}{16}(x-x_0)^3\,x_0^{-\frac{5}{2}} + \ldots,$$

$$x^{\frac{3}{2}} = x_0^{\frac{3}{2}} + \tfrac{3}{2}(x-x_0)\,x_0^{\frac{1}{2}} + \tfrac{3}{8}(x-x_0)^2\,x_0^{-\frac{1}{2}} - \tfrac{1}{16}(x-x_0)^3\,x_0^{-\frac{3}{2}} + \ldots.$$

† Here and below the notation $N(a, b)$ is used to indicate that a variable is normally distributed with mean a and standard deviation b

By taking expectations on both sides and substituting from (13) the moments of $x = \chi'^2$, we get μ_1' and μ_3' of χ'. Also

$$\mu_2'(\chi') = \mu_1'(\chi'^2), \quad \mu_4'(\chi') = \mu_2'(\chi'^2).$$

Hence we derive the following moments:

$$\mu_1' = (n+\lambda)^{\frac{1}{2}} - \frac{n+2\lambda}{4(n+\lambda)^{\frac{3}{2}}} + \frac{1}{2}\frac{n+3\lambda}{(n+\lambda)^{\frac{5}{2}}} - \frac{15}{32}\frac{(n+2\lambda)^2}{(n+\lambda)^{\frac{7}{2}}} + \dots,$$

$$\mu_2' = (n+\lambda),$$

$$\mu_3' = (n+\lambda)^{\frac{3}{2}} + \frac{3}{4}\frac{n+2\lambda}{(n+\lambda)^{\frac{1}{2}}} - \frac{1}{2}\frac{n+3\lambda}{(n+\lambda)^{\frac{3}{2}}} + \frac{9}{32}\frac{(n+2\lambda)^2}{(n+\lambda)^{\frac{5}{2}}} + \dots,$$

$$\mu_4' = 2(n+2\lambda) + (n+\lambda)^2,$$

from which we obtain

$$\mu_1' = (n+\lambda)^{\frac{1}{2}} - \frac{n+2\lambda}{4(n+\lambda)^{\frac{3}{2}}} + \dots, \qquad \mu_2 = \frac{n+2\lambda}{2(n+\lambda)} + \dots,$$

$$\mu_3 = \frac{n+3\lambda}{(n+\lambda)^{\frac{3}{2}}} - \frac{3}{4}\frac{(n+2\lambda)^2}{(n+\lambda)^{\frac{5}{2}}} + \dots, \quad \mu_4 = \tfrac{3}{4} + O[(n+\lambda)^{-2}].$$

Hence $$\gamma_1 = \frac{\mu_3}{\mu_2^{\frac{3}{2}}} = \frac{n^2+4n\lambda}{\sqrt{2}(n+\lambda)(n+2\lambda)^{\frac{3}{2}}} + \dots, \quad \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = O[(n+\lambda)^{-1}].$$

Comparing these with the corresponding coefficients of the χ'^2-distribution, viz.

$$\gamma_1 = \frac{\sqrt{8}(n+3\lambda)}{(n+2\lambda)^{\frac{3}{2}}} + \dots, \quad \gamma_2 = \frac{12(n+4\lambda)}{(n+2\lambda)^2} + \dots,$$

we see that χ' approaches normality faster than χ'^2.

From the above it follows that $\sqrt{(2\chi'^2)}$ has mean $\sqrt{\{2(n+\lambda) - (n+2\lambda)/(n+\lambda)\}}$ to order $(n+\lambda)^{-\frac{1}{2}}$ and variance $(n+2\lambda)/(n+\lambda)$ to order $(n+\lambda)^{-1}$. We can therefore regard

$$\sqrt{\left\{\frac{2\chi'^2(n+\lambda)}{n+2\lambda}\right\}}$$

as distributed normally with mean

$$\sqrt{\left\{\frac{2(n+\lambda)^2}{n+2\lambda} - 1\right\}}$$

and variance unity.

This result may also be derived by taking the χ^2-approximation to the χ'^2-distribution and then using the known result that for large ν, $\sqrt{(2\chi^2)}$ is distributed as $N[\sqrt{(2\nu-1)}, 1]$. For, substituting χ'^2/ρ for χ^2 and the expressions in (15) for ρ and ν, we reach the same normal approximation.

Since $\nu > n$ from (15), it can be seen that the normal approximation to χ' with n degrees of freedom will be better than the normal approximation to χ with the same degrees of freedom. Thus, for example, if $n = 25$, we have

$\lambda =$	0	10	20	30	40
$\nu =$	25	27·22	31·15	35·59	40·24.

Hence for sufficiently large values of n and λ, the probability integral and percentage points may be obtained from the normal tables. Table 3 gives a comparison of some values of the probability integral, thus calculated, with the exact values.

Table 3. *Values of the χ'^2-integral on the normal approximation*

n	λ	ν	x	From χ^2	From normal	Exact
16	32	28·8	30	0·0590	0·0638	0·0609
16	32	28·8	60	0·8329	0·8320	0·8316
24	24	32·0	36	0·1556	0·1515	0·1567
24	24	32·0	72	0·9656	0·9686	0·9667

3·3. *Closer approximations to the χ'^2-distribution*

The probability function of χ'^2 can be represented in the form of a series with the fitted probability function of $(\rho\chi^2)$ as the leading term and, from these mathematical expansions, closer approximations to the probability integral and percentage points may be obtained. Two methods will be briefly considered.

First method

The cumulants of the distribution $f(x)$, as defined on p. 207 above, are seen to be

$$\left.\begin{aligned} \kappa_1^* &= n+\lambda, & \kappa_2^* &= 2(n+2\lambda), \\ \kappa_3^* &= \frac{8(n+2\lambda)^3}{n+\lambda}, & \kappa_4^* &= \frac{48(n+2\lambda)^3}{(n+\lambda)^2}, \end{aligned}\right\} \tag{17}$$

the rth cumulant being
$$\kappa_r^* = 2^{r-1}(r-1)!\frac{(n+2\lambda)^{r-1}}{(n+\lambda)^{r-2}}.$$

Comparing these with the corresponding cumulants of $p(x)$ in (13), we find $\kappa_r^* > \kappa_r$ for $r > 2$. Let us write

$$\kappa_3 - \kappa_3^* = c_3, \quad \kappa_4 - \kappa_4^* = c_4, \ldots. \tag{18}$$

Then the corresponding differences of cumulants of $p(y)$ and $f(y)$ as defined on p. 207, will be

$$c_3/\rho^3, \quad c_4/\rho^4, \ldots.$$

By the application of the Edgeworth operator to $f(y)$ we have

$$\begin{aligned} p(y) &= \exp\left\{-\frac{c_3}{6\rho^3}\frac{d^3}{dy^3}+\frac{c_4}{24\rho^4}\frac{d^4}{dy^4}+\ldots\right\}f(y) \\ &= \left[1+\left\{-\frac{c_3}{6\rho^3}D^3+\frac{c_4}{24\rho^4}D^4-\ldots\right\}+\frac{1}{2!}\left\{\left(\frac{c_3}{6\rho^3}\right)^2 D^6+\left(\frac{c_4}{24\rho^4}\right)^2 D^8-\ldots\right\}+\ldots\right]f(y). \end{aligned}$$

Hence the probability integral $\int_0^v p(y)\,dy$ is given by

$$\begin{aligned} \int_0^v f(y)\,dy+\Bigg[&\left\{-\frac{c_3}{6\rho^3}f''(y)+\frac{c_4}{24\rho^4}f'''(y)+\ldots\right\} \\ &+\frac{1}{2!}\left\{\left(\frac{c_3}{6\rho^3}\right)^2 f^{(5)}(y)+\left(\frac{c_4}{24\rho_4}\right)^2 f^{(7)}(y)+\ldots\right\}+\ldots\Bigg]. \end{aligned} \tag{19}$$

Since the higher derivatives of $f(y)$ become smaller in value for a given y, we retain only the first term in the square brackets of (19) and get a second approximation to the probability integral in the form

$$\int_0^v f(y)\,dy - \frac{c_3}{6\rho^3}\frac{d^3}{dy^3}\int_0^v f(y)\,dy,$$

which can be written as
$$I(u,p)-\frac{c_3}{6\rho^3[\sqrt{(2\nu)}]^3}\frac{d^3I}{du^3}. \tag{20}$$

When using the expression (20) for the evaluation of the integral, the computation of the first term $I(u,p)$ will, in general, require interpolation in the tables of the Incomplete Γ-function. We shall now show that by a suitable modification of the Everett interpolation formula, the second term in (20) can be accounted for and the whole expression computed in one calculation.

If u_1, u_2 are the tabulated values between which u lies and Δ_1'', Δ_2'' the tabulated second differences, we have as an approximation

$$\frac{d^3I}{du^3}\sim(\Delta_2''-\Delta_1'')\,10^3,$$

the interval for u being 0·1 in the tables. Suppose q is the fraction $(u-u_1)/(u_2-u_1)$, E_1'', E_2'' the second-order Everett interpolation coefficients corresponding to q and $k=\dfrac{c_3}{6\rho^3}\dfrac{10^3}{[\sqrt{(2\nu)}]^3}$. Then (20) becomes

$$I(u_1,p)(1-q)+I(u_2,p)\,q+\Delta_1''(E_1''+k)+\Delta_2''(E_2''-k). \tag{21}$$

If p is not a tabled value but lies between p_1 and p_2, then we evaluate the above expression for p_1 and for p_2 and then interpolate linearly for p.

Second method

It is well known that by using the Edgeworth form of the Gram-Charlier Type A series, a frequency function can be normalized if it approaches normality asymptotically and if its cumulants are in increasing order of some quantity, n^{-1}.

Goldberg & Levine (1946) have shown that by the method of normalization the percentage points of the χ^2-distribution could be obtained to a fairly good degree of accuracy. A similar method might be applied usefully to the χ'^2-distribution. However, a modified form of expansion with the fitted χ^2-function as the first term will be found more suitable.

Let us standardize the variate x (written for χ'^2) by introducing

$$\xi=\frac{x-(n+\lambda)}{\sqrt{(2n+4\lambda)}}.$$

Then, using the same notation as before, the cumulants of the distribution $p(\xi)$ are

$$0,\quad 1,\quad \kappa_3/\kappa_2^{\frac{3}{2}},\quad \kappa_4/\kappa_2^2,\ \ldots.$$

Since $f(x)$ has the same mean and standard deviation as $p(x)$, we get for the cumulants of $f(\xi)$

$$0,\quad 1,\quad \kappa_3^*/\kappa_2^{\frac{3}{2}},\quad \kappa_4^*/\kappa_2^2,\ \ldots.$$

These cumulants, from the third onwards, are of orders $-\frac{1}{2}$, -1, $-\frac{3}{2}$, ... in both n and λ. Now let

$$\alpha=\alpha(\xi)=e^{-\frac{1}{2}\xi^2}/\sqrt{(2\pi)},$$

and let ξ_3, ξ_4, ... be the Hermite polynomials of orders 3, 4, ... Then we have, arranging the terms in order of magnitude of n (Kendall, I, 1945, § 6·32),

$$p(\xi)=\alpha(\xi)+\frac{1}{6}\frac{\kappa_3}{\kappa_2^{\frac{3}{2}}}\alpha\xi_3+\alpha\left(\frac{1}{24}\frac{\kappa_4}{\kappa_2^2}\xi_4+\frac{1}{72}\frac{\kappa_3^2}{\kappa_2^3}\xi_6\right)+\ldots. \tag{22}$$

There is a similar expansion for $f(\xi)$ with κ_r^* in place of κ_r $(r>2)$.

Now we subtract formally this second series from the first, term by term, and transfer $f(\xi)$ to the right-hand side. We then obtain

$$p(\xi) = f(\xi) + \alpha(\xi)\left[\frac{1}{6}\frac{c_3}{\kappa_2^{\frac{3}{2}}}\xi_3 + \left(\frac{1}{24}\frac{c_4}{\kappa_2^2}\xi_4 + \frac{1}{72}\frac{c_{33}}{\kappa_2^3}\xi_6\right) + \dots\right], \tag{23}$$

where $c_3 c_4$ have the same meanings as in (18) and c_{rhl} is written for $(\kappa_r\kappa_h\kappa_l - \kappa_r^*\kappa_h^*\kappa_l^*)$.

We know that the infinite series in (23) is not uniformly convergent. We can still integrate it formally term by term and make use of the first few terms to get a better approximation than that given by the integral of $f(\xi)$ alone. Thus retaining terms up to $O(n^{-\frac{3}{2}})$, we derive an approximation to the probability integral

$$\int_0^x p(x)\,dx = \int_0^\xi p(\xi)\,d\xi$$

in the form

$$\int_0^\xi f(\xi)\,d\xi + \alpha(\xi)\left[-\frac{1}{6}\frac{c_3}{\kappa_2^{\frac{3}{2}}}\xi_2 - \left(\frac{1}{24}\frac{c_4}{\kappa_2^2}\xi_3 + \frac{1}{72}\frac{c_{33}}{\kappa_2^3}\xi_5\right) - \left(\frac{1}{120}\frac{c_5}{\kappa_2^{\frac{5}{2}}}\xi_4 + \frac{1}{144}\frac{c_{34}}{\kappa_2^{\frac{7}{2}}}\xi_6 + \frac{1}{1296}\frac{c_{333}}{\kappa_2^{\frac{9}{2}}}\xi_8\right)\right]. \tag{24}$$

The first term in (24) is our first approximation of § 3·1 and the rest give a correction to it which is seen to result in a considerable improvement (see Table 4). For evaluating this expression, the values of the Hermite polynomials may be taken from Jorgensen's tables (1916) if ξ is an argument tabled there; otherwise they have to be directly calculated. $\alpha(\xi)$ may be found (without need for interpolation) from *Tables of the Probability Functions*, Vol. 2 (Federal Works Agency, New York, 1942).

The coefficients in (24) involve only differences of the cumulants and so are smaller than the corresponding coefficients in (22). Thus a closer approximation is likely to result from (24) than from the same order of terms in (22).

For the percentage points, we employ the inversion of the Gram-Charlier series obtained by Cornish & Fisher (1937). If x, x' and ξ are respectively the percentage points of the distributions $p(x)$, $f(x)$ and $\alpha(\xi)$, then for a given probability level, we have

$$\frac{x-(n+\lambda)}{\sqrt{(2n+4\lambda)}} = \xi + \frac{1}{6}\frac{\kappa_3}{\kappa_2^{\frac{3}{2}}}(\xi^2-1) + \left\{\frac{1}{24}\frac{\kappa_4}{\kappa_2^2}(\xi^3-2\xi) - \frac{1}{36}\kappa_3^2(2\xi^3-5\xi)\right\} + \dots.$$

$\dfrac{x'-(n+\lambda)}{\sqrt{(2n+4\lambda)}}$ has a similar expansion with κ_r^* in place of $\kappa_r\,(r>2)$. By differencing as before we obtain an expression for x in terms of x' and ξ. Retaining terms up to $O(n^{-\frac{3}{2}})$, we find

$$x = x' + \sqrt{(2n+4\lambda)}\left[\frac{1}{6}\frac{c_3}{\kappa_2^{\frac{3}{2}}}(\xi^2-1) + \left\{\frac{1}{24}\frac{c_4}{\kappa_2^2}(\xi^3-3\xi) - \frac{1}{36}\frac{c_{33}}{\kappa_2^3}(2\xi^3-5\xi)\right\} + \left\{\frac{1}{120}\frac{c_5}{\kappa_2^{\frac{5}{2}}}(\xi^4-6\xi^2+3) \right.\right.$$
$$\left.\left. - \frac{1}{24}\frac{c_{34}}{\kappa_2^{\frac{7}{2}}}(\xi^4-5\xi^2+2) + \frac{1}{324}\frac{c_{333}}{\kappa_2^{\frac{9}{2}}}(12\xi^4-53\xi^2+17)\right\}\right]. \tag{25}$$

In this, x' is our first approximation, and the correction improves it considerably even at the lower end of the distribution. The values of the expressions in ξ in (25) are directly available for several probability levels from the table in Cornish & Fisher's paper.

The approximate values of the probability integral of the χ'^2-distribution obtained by these methods in a few cases are given in Table 4. Table 5 shows the approximate upper and lower 5 % points evaluated by method II.

Comparing the two methods for the probability integral, the second one, employing terms of the Gram-Charlier series up to $O(n^{-\frac{3}{2}})$, gives greater accuracy and is to be preferred, although from the point of view of labour and time involved, the first method is simpler and easier to apply. With respect to the percentage points, the method using the Cornish-Fisher inversion appears to be quite good, particularly at the upper points, but it does involve a certain amount of labour.

Table 4. *Closer approximations to the χ'^2-integral*

n	λ	x	1st approx.	2nd approx. method		Exact
				I	II	
4	4	10·00	0·7191	0·7209	0·7119	0·7118
4	4	24·00	0·9913	0·9917	0·9913	0·9925
7	16	24·00	0·5947	0·5938	0·5869	0·5898
7	16	38·97	0·9482	0·9504	0·9502	0·9500
16	8	20·00	0·3380	0·3345	0·3368	0·3369
16	8	40·00	0·9626	0·9632	0·9631	0·9632

Table 5. *Closer approximation to the χ'^2-percentage points, using method II*

n	λ	Upper 5 % point			Lower 5 % point		
		1st approx.	2nd approx.	Exact	1st approx.	2nd approx.	Exact
2	4	14·72	14·67	14·64	0·945	0·574	0·646
2	16	33·35	33·06	33·06	6·891	6·526	6·322
4	4	17·38	17·33	17·31	1·954	1·731	1·765
4	16	35·69	35·42	35·43	8·363	8·017	7·884
7	4	21·28	21·27	21·23	3·789	3·750	3·664
7	16	39·16	38·97	38·97	10·637	10·267	10·257

4. APPLICATIONS OF THE χ'^2-DISTRIBUTION

4·1. *The power function of the χ^2-test*

There are several possible applications of the non-central χ^2-distribution in statistics. We shall consider only a few of them. We will show here how this distribution arises in the study of power functions of the χ^2-tests and how the approximations of § 3 are useful in this connexion.

Suppose $\xi_1, \xi_2, \ldots, \xi_n$ are n independent observations in a sample. If we make the null hypothesis H_0, that the ξ_i have been drawn from a normal population with mean zero and S.D. unity, then if H_0 is true, the statistic $\chi^2 = \Sigma\xi_i^2$ will exceed χ^2_α, the α-significance point of the χ^2-distribution, based on n degrees of freedom, in a proportion α of the cases.

The power of the χ^2-test is given by the probability that $\Sigma\xi^2$ exceeds χ^2_α under some alternative hypothesis. If as an alternative to H_0, we suppose that the ξ_i have been drawn from normal populations having unit S.D. but different means a_i, then $\Sigma\xi_i^2$ will follow the non-central χ^2-distribution with n degrees of freedom and parameter $\lambda = \Sigma a_i^2$. Denoting this by $p_n(\chi'^2 \mid \lambda)$, the power function is given by

$$\int_{\chi_\alpha^2}^{\infty} p_n(\chi'^2 \mid \lambda)\, d\chi'^2 \equiv \beta(n, \lambda, \alpha). \tag{26}$$

Thus the power is a function of the single parameter λ and we may write the null hypothesis as $H_0(\lambda = 0)$ and an alternative as $H_1(\lambda)$, where H_1 is a composite hypothesis including the family of alternatives for which $\Sigma a_i^2 = \lambda$.

It was shown in § 3·1 that the χ'^2-distribution is fairly well approximated by a Type III distribution fitted from its first two moments. The power function β could therefore be evaluated quickly and fairly accurately by the method of the first approximation. When greater accuracy is needed, one of the other methods described in § 3·3 may be used.

We give here a table (Table 6) of values of the power of the χ^2-test applied at the significance level $\alpha = 0{\cdot}05$, obtained by the second method of § 3·3. The accuracy of these values in different parts of the table can be judged from the closeness between the approximate and exact values of the probability integral shown in Tables 1 and 4. In some of the cases tabled there, the limit x was chosen near to the 5 % point of the corresponding χ^2, so as to give a value of

$$1 - \int_0^x p_n(x \mid \lambda)\, dx$$

in the neighbourhood of the power β. It is believed that, in general, there is three-figure accuracy in Table 6.

Table 6. *The power function of the χ^2-test using a 5% significance level; values of $\beta(n, \lambda, \alpha)$,* where $\alpha = 0{\cdot}05$

n \ λ	2	4	6	8	10	12	14	16	18	20
2	0·234	0·416	0·585	0·719	0·819	0·885	0·929	0·956	0·973	0·983
3	0·195	0·357	0·518	0·655	0·762	0·841	0·897	0·935	0·958	0·974
4	0·171	0·320	0·470	0·605	0·719	0·803	0·867	0·913	0·943	0·963
5	0·157	0·292	0·432	0·565	0·678	0·769	0·839	0·891	0·927	0·952
6	0·146	0·270	0·404	0·531	0·644	0·738	0·813	0·870	0·911	0 940
7	0·138	0·251	0·378	0·502	0·614	0·710	0·788	0·849	0·895	0·928
8	0·131	0·238	0·357	0·477	0·588	0·685	0·765	0·830	0·879	0·916
9	0·125	0·225	0·339	0·455	0·564	0·661	0·744	0·811	0·863	0·903
10	0·121	0·215	0·323	0·435	0·542	0·640	0·724	0·793	0·848	0·891
12	0·113	0·198	0·297	0·402	0·505	0·601	0·686	0·759	0·818	0·866
14	0·108	0·185	0·276	0·374	0·473	0·567	0·653	0·728	0·791	0·842
16	0·103	0·174	0·259	0·351	0·446	0·538	0·623	0·699	0·764	0·819
18	0·099	0·165	0·244	0·332	0·422	0·512	0·596	0·673	0·740	0·796
20	0·096	0·158	0·232	0·315	0·402	0·489	0·572	0·648	0·716	0·775

When n or λ is so large that $\nu = n + \lambda^2/(n+2\lambda)$ is over 30, we may use the normal approximation of § 3·2 for obtaining the power function more quickly than by the method of the χ^2-approximation.

The above table can be used in a variety of ways: (*a*) For given λ and n, we may ask what is the chance of establishing significance at the 5 % level? (*b*) For given n, we may ask how large λ must be to have, say, a 90 % chance ($\beta = 0{\cdot}90$) of establishing significance at the 5 % level when a real difference in the α_i exists? (c) For given λ, we may ask how many observations are necessary to have a chance β of establishing significance?

An alternative graphical approach to the inverse problems (b) and (c) is indicated in § 7·3, p. 228 below.

4·2. *Application to the χ^2-test for the goodness of fit*

The χ^2-test for goodness of fit is concerned with the comparison of observed frequencies with those expected under a given hypothesis. The latter may be the theoretical frequencies of a continuous distribution or may be obtained by taking integrals of a continuous frequency distribution over a set of class intervals. Denote the observed frequencies by n_i and the expected frequencies by $N\pi_i$ $(i = 1, 2, \ldots, k)$, where k is the number of groups and N the total number of observations in the sample. Then

$$\sum_{i=1}^{k} n_i = \sum_{i=1}^{k} N\pi_i = N. \tag{27}$$

As is well known, the distribution of

$$\phi^2 = \sum_{i=1}^{k} \frac{(n_i - N\pi_i)^2}{N\pi_i}, \tag{28}$$

when the $N\pi_i$ are the true population expectations, may be related as an approximation to that of the sum of squares of normal variables. To link up also with the non-central theory discussed in §§ 2·1–2·3, the following approach may be adopted, although it must be realized that the conclusions reached are not exact. As in all problems concerning ϕ^2, it is generally only possible to assess the degree of error involved, in samples of finite size, by specific numerical comparisons.

As shown originally by K. Pearson (1900, 1916), the variances and co-variances of the k frequencies n_i, restricted by the condition (27), are precisely those holding in the section

$$X_1 + X_2 + \ldots + X_k = 0 \tag{29}$$

of the k-dimensioned normal probability distribution whose probability density at

$$(X_1, X_2, \ldots, X_k)$$

is

$$p(X_1, X_2, \ldots, X_n) = \text{constant} \times \exp\left[-\tfrac{1}{2}\sum_i \frac{X_i^2}{N\pi_i}\right]. \tag{30}$$

Thus, provided that the expectations $N\pi_i$ are large enough to prevent serious inaccuracy from discontinuity effects or boundary limitations, relationships between the n_i may be treated as relationships, within the prime (29), between normal variables X_i which in the k-dimensioned space are distributed independently with zero means and variances $N\pi_i$. With these limitations, we may write

$$x_i = \frac{X_i}{\sqrt{(N\pi_i)}} = \frac{n_i - N\pi_i}{\sqrt{(N\pi_i)}} \quad (i = 1, \ldots, k). \tag{31}$$

The distribution of the ϕ^2 defined in (28) can then be derived from the results given in § 2·3. The condition $\Sigma n_i = N$ may be written

$$\sum_i \sqrt{\pi_i} \frac{n_i - N\pi_i}{\sqrt{(N\pi_i)}} = 0 \tag{32}$$

corresponding to $\sum_i c_i x_i = \rho = 0$, where $\sum_i c_i^2 = 1$. Hence ϕ^2 will be approximately distributed as χ^2 with $k-1$ degrees of freedom.

Having in mind the question of the power of the test, we may next ask what will be the distribution of ϕ^2 if the frequencies $N\pi_i$ inserted into the expression (28) are not the true expectations? Suppose that Np_i are the true expectations; both $\sum_i p_i$ and $\sum_i \pi_i$ will be unity.

In the notation of § 2 we now have

$$\xi_i = \frac{n_i - N\pi_i}{\sqrt{(Np_i)}}, \quad x_i = \frac{n_i - Np_i}{\sqrt{(Np_i)}}, \quad a_i = \frac{N(p_i - \pi_i)}{\sqrt{(Np_i)}}, \tag{33}$$

while
$$\sum_i \sqrt{p_i} \frac{(n_i - N\pi_i)}{\sqrt{(Np_i)}} = \sum_i c_i \xi_i = 0. \tag{34}$$

It follows that approximately

$$\phi'^2 = \sum_i \xi_i^2 = \sum_i \frac{(n_i - N\pi_i)^2}{Np_i} \tag{35}$$

will be distributed as a non-central χ^2 with $k-1$ degrees of freedom and

$$\lambda' = \sum_i (a_i^2) = N \sum_i \frac{(p_i - \pi_i)^2}{p_i}. \tag{36}$$

The sum of squares we need is the ϕ^2 of (28), not the ϕ'^2 of (35). By introducing a further approximation we may, however, conclude that $\phi^2 = \sum_i (n_i - N\pi_i)^2/N\pi_i$ is distributed as non-central χ^2 with $k-1$ degrees of freedom, and

$$\lambda = N \sum_i \frac{(p_i - \pi_i)^2}{\pi_i}. \tag{37}†$$

The approximation involved should not be serious if the differences $\delta_i = N\pi_i - Np_i$ are small compared to $N\pi_i$; for

$$\begin{aligned}\phi'^2 = \sum_i \frac{(n_i - N\pi_i)^2}{Np_i} &= \sum_i \frac{(n_i - N\pi_i)^2}{N\pi_i} \left\{1 - \frac{\delta_i}{N\pi_i}\right\}^{-1} \\ &= \phi^2 + \sum_i \delta_i \frac{(n_i - N\pi_i)^2}{(N\pi_i)^2} + \sum_i \delta_i^2 \frac{(n_i - N\pi_i)^2}{(N\pi_i)^3} + \dots.\end{aligned}$$

Since the multipliers δ_i in the second term may be positive or negative and $\Sigma\delta_i = 0$, this term will generally be small; the further terms, containing successive powers of $\delta_i/(N\pi_i)$, will also be of diminishing importance.

This result makes it possible to determine the power of the goodness of fit test of any simple (completely specified) hypothesis H_0 (specifying probabilities π_i) with respect to a simple alternative hypothesis H_1 (specifying probabilities p_i). Hence, for any given class of alternatives H, we can determine the power function. In so far as the 5 % significance level is used, the power may be determined from Table 6, p. 214, using the λ of equation (37) and degrees of freedom $k-1$. Otherwise, we can use the χ^2-approximation to the χ'^2-distribution developed in § 3·1. Thus the power is

$$\int_{\chi_\alpha^2}^{\infty} p_{k-1}(\chi'^2 \mid \lambda)\, d\chi'^2 = \int_{\chi_\alpha^2/\rho}^{\infty} p_\nu(\chi^2)\, d\chi^2, \tag{38}$$

where
$$\rho = \frac{k-1+2\lambda}{k-1+\lambda}, \quad \nu = \frac{(k-1+\lambda)^2}{k-1+2\lambda}, \quad \lambda = N\left(\sum_i \frac{p_i^2}{\pi_i} - 1\right). \tag{39}$$

† In making the approximation, we have associated the λ of (37) with the distribution of ϕ^2 rather than the λ' of (36), but this step perhaps needs fuller justification.

For comparison of this approximate distribution with the exact one, we proceed now to find the exact moments of ϕ^2. It is known (e.g. Haldane, 1937) that under H_1 the expectations of the powers of the observed frequency n_i are

$$\left.\begin{aligned}\mathscr{E}(n_i) &= Np_i,\\ \mathscr{E}(n_i^2) &= N_2p_i^2+Np_i,\\ \mathscr{E}(n_i^4) &= N_4p_i^4+6N_3p_i^3+7N_2p_i^2+Np_i,\\ \mathscr{E}(n_i^2n_j^2) &= N_4p_i^2p_j^2+N_3(p_i^2p_j+p_j^2p_i)+N_2p_ip_j,\\ &\text{etc.,}\end{aligned}\right\} \tag{40}$$

where
$$N_r = N!/(N-r)!.$$

Writing ϕ^2 in (28) in the form $\quad \phi^2 = \frac{1}{N}\Sigma(n_i^2/\pi_i)-N,$

we have
$$\begin{aligned}\mathscr{E}(\phi^2) &= \frac{1}{N}\Sigma\frac{\mathscr{E}(n_i^2)}{\pi_i}-N\\ &= \frac{1}{N}\left\{\Sigma\frac{N_2p_i^2}{\pi_i}+\Sigma\frac{Np_i}{\pi_i}\right\}-N.\end{aligned}$$

Hence
$$\mu_1' = (N-1)\,\Sigma(p_i^2/\pi_i)+\Sigma(p_i/\pi_i)-N. \tag{41}$$

Again,
$$\mathscr{E}[(\phi^2)^2] = \mathscr{E}\left[\frac{1}{N}\Sigma\frac{n_i^4}{\pi_i^2}+\frac{1}{N^2}\sum_{i\neq j}\frac{n_i^2n_j^2}{\pi_i\pi_j}-2\Sigma\frac{n_i^2}{\pi_i}+N^2\right],$$

from which on substitution and simplication we obtain

$$\begin{aligned}\mu_2 = N^{-1}\{&(N-1)(6-4N)[\Sigma(p_i^2/\pi_i)]^2+4(N-1)(N-2)\,\Sigma(p_i^3/\pi_i^2)\\ &-4(N-1)\,\Sigma(p_i^2/\pi_i)\,\Sigma(p_i/\pi_i)+6(N-1)\,\Sigma(p_i^2/\pi_i^2)\\ &-[\Sigma(p_i/\pi_i)]^2+\Sigma(p_i/\pi_i^2)\}.\end{aligned} \tag{42}$$

In a similar way the third moment has also been obtained but the expression is so long and so difficult to evaluate numerically that it may not be of much value for comparison purposes.

When $p_i = \pi_i$ the above expressions reduce to those derived by Haldane (1937) for the exact moments of the distribution of ϕ^2 under the null hypothesis.

The approximation to the distribution of ϕ^2 obtained, using the simplification of § 3·1, will have the following first two moments:

$$\left.\begin{aligned}\mu_1' &= \nu+\lambda = k-1+\lambda = k-1+N[\Sigma(p_i^2/\pi_i)-1],\\ \mu_2 &= 2(\nu+2\lambda) = 2(k-1)+4\lambda = 2(k-1)+4N[\Sigma(p_i^2/\pi_i)-1],\end{aligned}\right\} \tag{43}$$

using the expression for λ in (37).

A comparison of these approximate moments with the exact ones, (41) and (42), appears to be only possible numerically. Some comparisons have been made, including a check-up on the whole distribution by a random sampling experiment. In the cases taken, the approximation appeared satisfactory for practical purposes but some further investigation is in hand. The results will be published in a subsequent paper.

4·3. *Uses of the power function of the χ^2 goodness of fit test*

We have seen in § 4·2 that, to the approximation involved, the power of the χ^2-test for H_0 with regard to an alternative H_1 is a function of $k-1$, λ, α and can be written $\beta(k-1,\lambda,\alpha)$, where k is the number of groups, α the significance level at which the test is applied and

$$\lambda = N\left(\sum_{i=1}^{k}\frac{p_i^2}{\pi_i}-1\right) = N\Delta(H_0,H_1).$$

This shows that λ is a function of π_i and p_i, and can be regarded as a measure of 'discrepancy' between the two distribution functions specified by H_0 and H_1.

The power function can be used to answer several questions connected with the test of goodness of fit: (*a*) For given sample size N and number of groups k, we may ask what is the chance of establishing the inadequacy of the hypothesis H_0, using a given significance level? (*b*) For given k, we may ask how many observations are necessary to give a chance of, say, 90 % of establishing significance at the 5 % level? (*c*) For given k and N, we may ask how large a departure of H_1 from H_0 (measured by $\Delta(H_0, H_1)$) will be detected with a given chance?

We shall illustrate these applications by an example from genetics. Consider the intercross

$$\frac{\mathbf{AB}}{\mathbf{ab}} \times \frac{\mathbf{AB}}{\mathbf{ab}},$$

where **A** and **B** are two independent factors, the recessive genes of which are represented by **a** and **b**. The offspring are of the four types [**AB**], [**Ab**], [**aB**], [**ab**] with frequencies in the proportions 9, 3, 3, 1. We test whether the experiment is to confirm this theory or to reject it in favour of a definite alternative giving frequencies proportional to 9, 3, $3r$, r (r being less than 1). This happens when the two classes of offspring containing the two recessive genes (**a**, **a**) are less viable than those containing only one dominant gene, so that only a fraction of the offspring survive.

Here, the expected frequencies are

$$\pi_i: 9/16,\ 3/16,\ 3/16,\ 1/16.$$

$$p_i: 9/4(3+r),\ 3/4(3+r),\ 3r/4(3+r),\ r/4(3+r).$$

Hence
$$\lambda = N\left(\frac{4(3+r^2)}{(3+r)^2} - 1\right), \tag{44}$$

where N is the number of offspring studied. Then

$$\Delta \equiv \Delta(H_0, H_1) = \frac{4(3+r^2)}{(3+r)^2} - 1. \tag{45}$$

Let us now consider the three situations where the power-function idea could be applied.

(*a*) Suppose we have 100 observations. Using the χ^2-test at the 5% level to test the null hypothesis ($r=1$), the chance of establishing differential viability when $r = \frac{1}{2}$ is obtained by evaluating λ from (37) and then entering Table 6 (p. 214) with this λ and $n = k-1 = 3$. Here $\lambda = 300/49$ and so the power $\beta = 0{\cdot}52$.

(*b*) Suppose we want a 90 % chance of detecting that $r = \frac{1}{2}$, using the 5 % significance level. We find from Table 6 that $\lambda = 14{\cdot}1$ and hence, putting $r = \frac{1}{2}$ in (45), obtain $\Delta = 3/49$. Then from (44) we find that we shall need a sample of $N = 230$.

(*c*) Again, if $N = 100$, $\alpha = 0{\cdot}05$, we may ask how small r must be to give a 50 : 50 chance for establishing significance? We find λ as before and solve (44) for r. Thus taking $\beta = 0{\cdot}50$, then $\lambda = 5{\cdot}8$ and $r = 0{\cdot}51$.

4·4. *A closer approximation to the power function of the χ^2 goodness of fit test*

In § 4·2, when deriving the χ'^2-approximation to the distribution of

$$\phi^2 = \Sigma\frac{(n_i - N\pi_i)^2}{N\pi_i},$$

we made the assumption that π_i and p_i, the proportions of the expected frequencies under the hypotheses H_0 and H_1 do not differ very much, so that we could regard $(n_i - Np_i)/\sqrt{(N\pi_i)}$ as a normal deviate with zero mean and unit variance. We will now consider the distribution of ϕ^2 without making such an assumption and use it for obtaining a better approximation to the power function.

We can write ϕ^2 in the form

$$\phi^2 = \Sigma \frac{p_i}{\pi_i}\left(\frac{n_i - Np_i}{\sqrt{(Np_i)}} + \frac{Np_i - N\pi_i}{\sqrt{(Np_i)}}\right)^2, \tag{46}$$

the summation being from $i = 1$ to k. Now, under H_1 the quantities $(n_i - Np_i)/\sqrt{(Np_i)}$ are distributed approximately normally, as $N(0, 1)$, subject to the constraint $\Sigma n_i = N$. Hence ϕ^2 in (46) can be regarded as the weighted sum of k normal deviates having different expectations and satisfying the condition $\Sigma n_i = N$.

We have obtained in the Appendix (pp. 231–2 below) the characteristic function of the distribution of such a statistic, viz. $\Sigma v_j(x_j + a_j)^2$ subject to the condition $\Sigma c_j(x_j + a_j) = \rho$. Making the appropriate substitution in (6) of the Appendix, we have the characteristic function of ϕ^2:

$$\left(\Sigma \frac{p}{1 - 2itp/\pi}\right)^{-\frac{1}{2}} \prod_1^k (1 - 2itp/\pi)^{-\frac{1}{2}} \times \exp\left\{N\Sigma\left(\frac{it(p-\pi)^2/\pi}{1 - 2itp/\pi}\right) - \frac{N}{2}\left(\Sigma \frac{p - \pi}{1 - 2itp/\pi}\right)^2 \left(\Sigma \frac{p}{1 - 2itp/\pi}\right)^{-1}\right\}, \tag{47}$$

where the subscripts of p_i and π_i are dropped. From this the expressions for the first three moments are derived. Thus

$$\left.\begin{aligned} \mu_1' &= (N-1)\,\Sigma(p_i^2/\pi_i) + \Sigma(p_i/\pi_i) - N, \\ \mu_2 &= 4(N-1)\,\Sigma(p_i^3/\pi_i^2) - 2(2N-1)\,[\Sigma(p_i^2/\pi_i)]^2 + 2\Sigma(p_i^2/\pi_i^2), \\ \mu_3 &= 24(N-1)\,\Sigma(p_i^4/\pi_i^3) - 24(2N-1)\,[\Sigma(p_i^2/\pi_i)]\,[\Sigma(p_i^3/\pi_i^2)] \\ &\quad + 8(3N-1)\,[\Sigma(p_i^2/\pi_i)]^3 + 8\,\Sigma(p_i^3/\pi_i^3). \end{aligned}\right\} \tag{48}$$

It will be seen that the only assumption made here, that $(n_i - Np_i)/\sqrt{(Np_i)}$ is distributed as $N(0, 1)$ under H_1, is parallel to the assumption on which the χ^2-test of goodness of fit is based, namely, that $(n_i - N\pi_i)/\sqrt{(N\pi_i)}$ is distributed as $N(0, 1)$ under H_0, which is justified when $N\pi_i$ are not too small. So, when Np_i are not too small we can expect the moments in (48) to agree well with the true moments (the first two of which are given in (41) and (42)). Obviously the expressions for μ_1' are identical. The values of μ_2 in the cases examined in the investigation referred to on p. 217 were found to be very close.

We may now obtain a representation of the distribution of ϕ^2 under H_1 as a Type III having the first two moments of (48), that is, assume ϕ^2/ρ as distributed as χ^2 with ν degrees of freedom, where $\rho = \frac{1}{2}\mu_2/\mu_1'$, $\nu = 2\mu_1'^2/\mu_2$. Clearly this will be a better approximation than that of the Type III fitted from the μ_1', μ_2 given in (43), and the power function based on this will be closer to the exact one than that based on (38) and (39). But, although there is gain in accuracy, the simplicity of the approximate method is lost. We may similarly consider fitting a Type III distribution, using the true μ_1' and μ_2, but the labour of computation of μ_2, given in (42), appears to be prohibitive.

5. CONDITIONAL POWER FUNCTIONS

In § 4 we have considered the power function of the χ^2 goodness of fit test when the null hypothesis is fully specified, i.e. is a simple hypothesis. But often we are interested in testing whether an observed sample has come from a certain type of population, so that we are given only the form of the population law, not the values of its parameters, say $\theta_1, \theta_2, \ldots, \theta_r$. H_0 is then a composite hypothesis. Sometimes, also, we have to test the hypothesis that several samples are from the same population, without specifying anything about it. In these cases we obtain estimates of the unspecified parameters, say $T_1, T_2, \ldots, T_r$, from the sample and hence calculate the expected cell frequencies $\hat{m}_i$. Then, if the method of estimation is efficient†,

$$\phi^2 = \sum_i (n_i - \hat{m}_i)^2 / \hat{m}_i \tag{49}$$

is known still to follow approximately a χ^2-distribution with $k-r-1$ degrees of freedom.

Suppose now that as alternative to the composite hypothesis H_0, there is a simple hypothesis H_1. The question then arises: By estimating the m_i on the assumption that H_0 is true and applying the χ^2-test, what chance have we of rejecting H_0, when, in fact, H_1 is true?

Some consideration has been given to this problem, and it seems possible to obtain a solution by making use, as a first step, of what David (1947, p. 339) has termed the conditional power function. This gives the chance of rejecting H_0 when the test is confined to a restricted set, S, of samples which provide the same values, say $T_1^{(s)}, T_2^{(s)}, \ldots, T_r^{(s)}$ for the estimated parameters. Thus, if the process of fitting involves estimating two parameters from the sample mean and variance, samples of a set would be those having a common mean and variance. Again, in testing for independence in a contingency table, the conditional power function would be obtained for a set of samples giving the same marginal totals (see Patnaik, 1948). The development of this method will be left for a later communication.

6. The non-central F-distribution and approximations to it

Suppose two independent variates, $\chi_1'^2$ and χ_2^2, follow respectively a non-central χ^2-distribution with degrees of freedom ν_1 and parameter λ and a χ^2-distribution with degrees of freedom ν_2. Then the ratio

$$F' = \frac{\chi_1'^2/\nu_1}{\chi_2^2/\nu_2}$$

will have the following probability distribution:

$$p(F') = \sum_{j=0}^{\infty} \left[\frac{e^{-\frac{1}{2}\lambda}(\frac{1}{2}\lambda)^j}{j!\, B(\frac{1}{2}\nu_1 + j, \frac{1}{2}\nu_2)} \left(\frac{\nu_1}{\nu_2}\right)^{\frac{1}{2}\nu_1 + j} . F'^{\frac{1}{2}\nu_1 - 1 + j} \left(1 + \frac{\nu_1}{\nu_2} F'\right)^{-\frac{1}{2}(\nu_1+\nu_2) - j} \right], \tag{50}$$

which may be termed the distribution of non-central F or of F'. This corresponds to Fisher's distribution C (1928). Wishart (1932) considered it in the form of the distribution of the correlation ratio

$$E^2 = \frac{\nu_1 F'}{\nu_2 + \nu_1 F'}.$$

Later, Tang (1938) derived the same from that of χ'^2.

If in (50) we put $\nu_1 = 1$, then it reduces to the distribution of non-central t^2. Denoting the non-central t by t', we have

$$t' = \frac{z + \delta}{\sqrt{w}},$$

where z is a normal deviate with expected value zero and w is an unbiased estimate of its variance. Neyman (1935), Neyman & Tokarska (1936) and Johnson & Welch (1939) have

† I. e. gives a solution not very different from the maximum likelihood or minimum χ^2 solutions, which are nearly identical in large samples.

dealt with this distribution in detail and studied its various applications. We will not therefore consider here in particular this special case of the F'-distribution.

Taking the general form (50), we may, by analogy, call ν_1; ν_2 the degrees of freedom and λ the non-central parameter. It can be seen that when ν_2 tends to infinity the distribution of F' reduces to that of $\chi_1'^2/\nu_1$.

The characteristic function is obtained as an infinite sum of confluent hypergeometric functions

$$\phi(t) = e^{-\frac{1}{2}\lambda}\left[H\left(\frac{\nu_1}{2}, -\frac{\nu_2}{2}, -\frac{\nu_2}{\nu_1}it\right)+\frac{\lambda}{2}H\left(\frac{\nu_1}{2}+1, -\frac{\nu_2}{2}, -\frac{\nu_2}{\nu_1}it\right)\right.$$
$$\left.+\left(\frac{\lambda}{2}\right)^2\frac{1}{2!}H\left(\frac{\nu_1}{2}+2, -\frac{\nu_2}{2}, -\frac{\nu_2}{\nu_1}it\right)+\ldots\right],$$

in which the function, $H(a,b,x)$, is the sum of the series

$$1+\frac{a}{b}x+\frac{a(a+1)}{2!\,b(b+1)}x^2+\ldots.$$

Thence we derive the following expressions for the first four moments about the origin:

$$\left.\begin{aligned}
\mu_1' &= \frac{\nu_2(\nu_1+\lambda)}{(\nu_2-2)\,\nu_1},\\
\mu_2' &= \frac{\nu_2^2}{(\nu_2-2)(\nu_2-4)\,\nu_1^2}[(\nu_1+\lambda)^2+2(\nu_1+2\lambda)],\\
\mu_3' &= \frac{\nu_2^3}{(\nu_2-2)(\nu_2-4)(\nu_2-6)\,\nu_1^3}[(\nu_1+\lambda)^3+6(\nu_1+\lambda)(\nu_1+2\lambda)+8(\nu_1+3\lambda)],\\
\mu_4' &= \frac{\nu_2^4}{(\nu_2-2)(\nu_2-4)(\nu_2-6)(\nu_2-8)\,\nu_1^4}[(\nu_1+\lambda)^4+12(\nu_1+\lambda)^2(\nu_1+2\lambda)\\
&\qquad +44(\nu_1+2\lambda)^2+48(\nu_1+4\lambda)-32\lambda^2],
\end{aligned}\right\} \tag{51}$$

of which the first two were obtained by Wishart by a different method.

Methods of evaluating the probability integral of the F'-distribution have been worked out by Wishart and Tang. They involve a considerable amount of labour. Following the procedure adopted in the case of χ'^2, it may be possible to obtain a quick, though approximate, method by fitting an F-distribution with the exact first two moments of F'. If we regard F'/k as following an F-distribution with ν and ν_2 degrees of freedom, then, equating the expressions for μ_1' and μ_2, we have

$$\frac{\nu_2(\nu_1+\lambda)}{k(\nu_2-2)\,\nu_1} = \frac{\nu_2}{\nu_2-2},$$
$$\frac{\nu_2^2}{k^2(\nu_2-2)(\nu_2-4)\,\nu_1^2}[(\nu_1+\lambda)^2+2(\nu_1+2\lambda)] = \frac{\nu_2^2}{(\nu_2-2)(\nu_2-4)}\frac{\nu+2}{\nu},$$

which give the scale factor and the modified degrees of freedom, viz.

$$k = \frac{\nu_1+\lambda}{\nu_1}, \quad \nu = \frac{(\nu_1+\lambda)^2}{\nu_1+2\lambda}. \tag{52}$$

The same result will follow if we approximate the distribution of $\chi_1'^2$ (the numerator in F') by a Type III from the first two moments as in § 3·1.

Using the above approximation, the probability integral

$$\int_0^{F'} p_{\nu_1,\nu_2}(F' \mid \lambda)\,dF'$$

is approximately equal to
$$\int_0^{F'/k} p_{\nu,\nu_2}(F)\,dF,$$
where k and ν are defined in (52). This can be expressed in the form of an Incomplete B-function, viz.
$$I_x\left(\frac{\nu}{2},\frac{\nu_2}{2}\right),$$
where
$$x = \frac{\nu F'/k}{\nu_2+\nu F'/k}.$$
For given values of ν_1, ν_2, λ and F', we can therefore evaluate the integral from the *Tables of the Incomplete B-function* (K. Pearson, 1934). When ν_2 is even or, if odd, is less than 22, we need interpolate only for x and $\frac{1}{2}\nu(=p)$. Four-point Lagrangian interpolation p-wise and linear interpolation x-wise will be necessary.

Tang's tables of P_{II} (the error of the second kind) (1938) give exact values of the integral of the E^2-distribution, which, put in the F'-form, is
$$\int_0^{F_\alpha} p_{\nu_1,\nu_2}(F' \mid \lambda)\,dF', \tag{53}$$

Table 7. *Approximate and exact values of the F'-integral,* $\int_0^x p_{\nu_1,\nu_2}(F' \mid \lambda)\,dF'$

ν_1	ν_2	λ	x	Approx.	Exact
3	10	4	3·708	0·752	0·745
		4	6·552	0·919	0·918
		16	3·708	0·203	0·206
		16	6·552	0·520	0·517
3	20	4	3·098	0·706	0·700
		4	4·938	0·889	0·887
		16	3·098	0·119	0·126
		16	4·938	0·350	0·347
5	10	6	3·326	0·731	0·731
		6	5·636	0·913	0·914
		24	3·326	0·157	0·158
		24	5·636	0·463	0·461
5	20	6	2·711	0·665	0·664
		6	4·103	0·869	0·870
		24	2·711	0·064	0·069
		24	4·103	0·244	0·245
8	10	9	3·072	0·715	0·714
		9	5·057	0·909	0·908
		36	3·072	0·117	0·119
		36	5·057	0·409	0·408
8	30	9	2·266	0·581	0·578
		9	3·173	0·815	0·813
		36	2·266	0·014	0·017
		36	3·173	0·085	0·088

F_α being the α-percentage point of the F-distribution with ν_1, ν_2 degrees of freedom. Two levels of α were chosen for the tables, namely, 0·05 and 0·01, and the range of ν_1 is 0 to 8. The tables have to be entered with $\phi = \sqrt{[\lambda/(\nu_1+1)]}$. Since ϕ is at intervals of 0·5, the corresponding intervals for λ are very wide, which therefore makes interpolation unsatisfactory.

Table 7 gives the values of the integral (53) calculated by the approximate method indicated above, for certain cases where Tang's exact values are available. The comparison shows that, in general, we have two-figure accuracy, while the error in the third place appears to be quite small near the tails.†

It is to be noted that the table compares the integral at only two points, the 5 and 1 % points of the corresponding F-distribution. Due to the lack of exact values it has not been possible to judge the closeness at other points. However, some idea of the general accuracy could be had by comparing the true and approximate figures for different λ's with the same ν_1, ν_2 and $x(=F_\alpha)$.

It can be easily shown (see Hartley, 1948) that the maximum error in the F'-integral due to our approximation will not exceed the maximum error in the corresponding χ'^2-integral, that is, in

$$\int_0^{\chi'^2} p_{\nu_1}(\chi'^2 \mid \lambda)\, d\chi'^2.$$

Table 1 on p. 207 gives an idea of the magnitude of the errors in the χ'^2-integral, and so we can say that the errors in the F'-integral will not be of a higher order.

The percentage points of F' can be obtained by interpolation in the F-tables (Merrington & Thompson, 1943), for the fractional ν and ν_2 and multiplying the interpolate by k in (52).

Closer approximations to the probability integral and percentage points may be derived by the method based on the Gram-Charlier series, analogous to the second method of § 3·3.

7. THE POWER FUNCTION OF THE ANALYSIS OF VARIANCE TESTS

7·1. *Evaluation of the power function*

The test of a general linear hypothesis may be formulated as follows: Suppose $x_1, x_2, \ldots, x_N$ be N normal variates with means $\xi_1, \xi_2, \ldots, \xi_n$ and the same S.D., σ. ξ_i is a linear function of $s<N$ parameters, $\theta_1, \theta_2, \ldots, \theta_s$. Thus

$$\xi_i = a_{i1}\theta_1 + a_{i2}\theta_2 + \ldots + a_{is}\theta_s.$$

The linear hypothesis specifies, say, r of these parameters, i.e.

$$\theta_1 = \theta_1^0, \quad \theta_2 = \theta_2^0, \quad \ldots, \quad \theta_r = \theta_r^0. \tag{54}$$

It is possible by a suitable transformation of variates (see Tang, 1938) of the form

$$y_j = c_{j1}x_1 + c_{j2}x_2 + \ldots + c_{jN}x_N$$

to transform
$$T^2 = \sum_{i=1}^{N} (x_i - \xi_i)^2$$

into
$$T^2 = \sum_{j=1}^{N-s} y_j^2 + \sum_{j=N-s+1}^{N-s+r} (y_j - \eta_j)^2 + \sum_{j=N-s+r+1}^{N} (y_j - \eta_j)^2,$$

where η_j in the second sum is a linear function of $\theta_1^0, \theta_2^0, \ldots, \theta_r^0$ and η_j in the third sum is a linear function of all the θ's, while the a's and c's enter as coefficients.

† [Further exploration shows that the differences between the approximate and true values are systematic, with regular fluctuations. Use is being made of this fact to prepare certain rather more extensive tables of the power function. ED.]

To test the hypothesis (54) we consider the criterion

$$\left\{\frac{T^2_{\min.}(\theta_1^0,\ldots,\theta_r^0,\theta_{r+1},\ldots,\theta_s)}{T^2_{\min.}(\theta_1,\ldots,\theta_r,\ldots,\theta_s)}-1\right\} = \sum_{j=N-s+1}^{N-s+r}(y_j-\eta_j)^2\Big/\sum_{j=1}^{N-s}y_j^2. \tag{55}$$

If the hypothesis specifies such values for $\theta_1, \theta_2, \ldots, \theta_r$ that η_j's in (55) vanish, then the numerator and denominator are the sums of r and $N-s$ central squares respectively. So, the ratio of the mean squares follows an F-distribution. On the other hand, if the η_j's do not all vanish, we have the ratio of a sum of r non-central squares to the sum of $N-s$ central squares; hence, the ratio of the mean squares is distributed as non-central F, the parameter λ being $\Sigma\eta_j^2$ which can be expressed in terms of $\theta_1^0, \ldots, \theta_r^0$ (see Tang, p. 137). Thus we get the F-test of the analysis of variance and obtain the power function of this test with respect to an alternative hypothesis as an F'-integral.

We shall now consider the question of evaluating the power of the analysis of variance test by taking as an illustration the simple case of k groups of observations

$$x_{ti}(i = 1, \ldots, n;\ t = 1, \ldots, k).$$

Suppose

$$x_{ti} = A + B_t + z_{ti}, \tag{56}$$

where A is the general mean, B_t the deviation of the mean of the tth group from the general mean so that $\Sigma B_t = 0$ and z_{ti}'s are random residuals, distributed normally with mean zero and S.D. $= \sigma_0$. The expressions for the mean squares between groups and within groups follow from the set-up (56):

$$v = \frac{1}{k-1}\sum_{t=1}^{k}(\bar{x}_{t.}-\bar{x}_{..})^2 = \frac{1}{k-1}\sum_{t=1}^{k}n(\bar{z}_{t.}-\bar{z}_{..}+B_t)^2,$$

$$v_0 = \frac{1}{k(n-1)}\sum_{t=1}^{k}\sum_{i=1}^{n}(x_{ti}-\bar{x}_{t.})^2 = \frac{1}{k(n-1)}\sum_{t=1}^{k}\sum_{i=1}^{n}(z_{ti}-\bar{z}_{t.})^2,$$

where the symbols have the usual meanings. Since $(\bar{z}_{t.}-\bar{z}_{..})$ is a normal deviate with zero mean and variance σ_0^2/n, we see that v is the sum of k non-central squares subject to the linear constraint

$$\sum_{t=1}^{k}(\bar{z}_{t.}-\bar{z}_{..}+B_t) = 0.$$

Since further $\Sigma B_t = 0$, we find from the result of § 2·3 that v is distributed as $\sigma_0^2\chi_1'^2/(k-1)$, where $\chi_1'^2$ has $(k-1)$ degrees of freedom and parameter

$$\lambda = n\Sigma B_t^2/\sigma_0^2.$$

Writing

$$S^2 = (\Sigma B_t^2)/k \tag{57}$$

for the variability between the groups, we have

$$\lambda = knS^2/\sigma_0^2. \tag{58}$$

Now v_0 follows the distribution of $\sigma_0^2\chi_2^2/[k(n-1)]$, where χ_2^2 has $k(n-1)$ degrees of freedom. Hence v/v_0 is distributed as

$$\frac{1}{k-1}\chi_1'^2\Big/\frac{1}{k(n-1)}\chi_2^2,$$

i.e. as F' with $\nu_1 = k-1$, $\nu_2 = k(n-1)$ and λ given by (58).

In this example we desire to test for any possible difference between the averages of the groups, so that our null hypothesis is

$$B_1 = B_2 = \ldots = B_k = 0. \tag{59}$$

Then, from (57), S^2 and therefore λ is zero. Hence v/v_0 follows an F-distribution and we get an F-test. Thus the test of the hypothesis in (59) is based on the critical region

$$\frac{v}{v_0} \geqslant F_\alpha, \tag{60}$$

where α is the significance level at which we are testing.

Let us consider an alternative hypothesis that the B_t's are not all zero. Then it is known that the power function, that is, the probability that $(v/v_0) \geqslant F_\alpha$, depends only on the single parameter

$$\frac{S^2}{\sigma_0^2} = \frac{\Sigma B_t^2}{k\sigma_0^2}.$$

Hsu (1941) has shown that amongst all critical regions of size α, whose power functions depend on the single parameter (S^2/σ_0^2), the critical region of (60) is the most powerful.

Thus we specify the hypothesis alternative to the null hypothesis (59) by the single parameter S^2/σ_0^2 in place of the individual parameters, the B_t's. In certain situations, as, for instance, in a manufacturing process, we are more interested in detecting the over-all variability in a set of machines than in detecting the deviation of each particular machine from the general machine average. Then the power function will be useful in measuring the chance of detecting this over-all variability by means of the F-test.

The power function of the analysis of variance tests has been considered by Tang (1938) and Hsu (1941). The rather restricted scope of Tang's tables has already been mentioned in § 6. The labour involved in computing the exact values of the power is very heavy, and no tabling on an extensive scale has so far been found possible. However, with the approximations to the F'-distribution derived in § 6, we may obtain easily a sufficiently accurate value for the power function of the test of any linear hypothesis.

Returning to the case of k groups and kn observations, we have the power function given by

$$\beta\left(\frac{S^2}{\sigma_0^2}\right) = \int_{F_\alpha}^{\infty} p_{\nu_1,\nu_2}(F' \mid \lambda)\, dF',$$

where F_α is the α percentage point of the F-distribution with degrees of freedom ν_1, ν_2. Following the procedure of § 6, this integral approximately equals

$$\int_{F_\alpha \nu_1/(\nu_1+\lambda)}^{\infty} p_{\nu,\nu_2}(F)\, dF, \tag{61}$$

where
$$\nu = \frac{(\nu_1+\lambda)^2}{\nu_1+2\lambda}.$$

Therefore, to this approximation, we have

$$\left.\begin{aligned} \beta\left(\frac{S^2}{\sigma_0^2}\right) &= I_x(\tfrac{1}{2}\nu_2, \tfrac{1}{2}\nu) \\ \text{in which} \quad x &= \frac{(\nu_1+2\lambda)\,\nu_2}{(\nu_1+2\lambda)\,\nu_2 + (\nu_1+\lambda)\,\nu_1 F_\alpha}. \end{aligned}\right\} \tag{62}$$

7·2. *The difference between systematic and random effects*

Next we shall consider two alternatives that arise in practical situations—the random and systematic set-ups (see Daniels, 1939) which may best be described in terms of two examples:

If the groups in the previous illustration correspond to villages and the observations are the yields of fields in a crop survey, then we can regard the k villages as a random sample from a population of villages and the random set-up represented by

$$x_{it} = A + y_t + z_{it} \tag{63}$$

becomes relevant. Here, A is the general mean, y_t's are the group means which are independent random variables with expected value zero and S.D. $= \sigma$, and z_{it}'s the random residuals having mean zero and S.D. $= \sigma_0$.

On the other hand, if the groups correspond to k machines which, from the user's standpoint, constitute the entire population of machines, we cannot regard them as a sample, and so the systematic set-up, in (56), considered on p. 224, is relevant. The null hypothesis in the random set-up is that the parameter $\sigma^2 = 0$, and in the other that $S^2 = 0$ (which is equivalent to (59)). But it is easily seen that both lead to the same F-test for the null hypothesis.

In applying the test, we are on the look out for the existence of alternative conditions, where in one case σ^2 and in the other S^2 is > 0. It will be noted that (S^2/σ_0^2) of the systematic set-up corresponds to (σ^2/σ_0^2) of the random set-up. Both are measures of relative variability between groups and may be termed 'relative group variability'.

It is possible to relate the power function under the random set-up to that under the systematic set-up. If we regard the k groups as a sample from an infinite number of groups, then $\Sigma B_t^2/(k-1)$, i.e. $kS^2/(k-1)$ will be the sample estimate of the population variance σ^2. Thus treating S^2 as a random variable having a probability distribution denoted by $p(S^2/\sigma^2)$, we can obtain the average power over all the S^2's. Thus

$$\beta = \int_0^\infty \beta\left(\frac{S^2}{\sigma_0^2}\right) p(S^2 \mid \sigma^2)\, dS^2$$

gives the power when the random set-up applies.

This power β for given (σ^2/σ_0^2) is directly obtained (see Johnson, 1948) from the F-integral:

$$\int_{F_\alpha/(n\sigma^2/\sigma_0^2+1)}^\infty p_{\nu_1,\nu_2}(F)\, dF = \int_{F_\alpha(\nu_1+1)/(\nu_1+1+\lambda)}^\infty p_{\nu_1,\nu_2}(F)\, dF, \tag{64}$$

where $\qquad \nu_1 = k-1, \quad \nu_2 = k(n-1) \quad \text{and} \quad \lambda = kn\sigma^2/\sigma_0^2.$

This can be put in the form of the Incomplete B-function

$$\left.\begin{aligned} &I_{x'}(\tfrac{1}{2}\nu_2, \tfrac{1}{2}\nu_1),\\ \text{where} \qquad &x' = \frac{(\nu_1+1+\lambda)\,\nu_2}{(\nu_1+1+\lambda)\,\nu_2 + (\nu_1+1)\,\nu_1 F_\alpha}.\end{aligned}\right\} \tag{65}$$

It is interesting to note a result which we believe is true in general and which on intuitional grounds might be expected to hold, namely, if the null hypothesis is not true, then for the same numerical values of the ratios S^2/σ_0^2 and σ^2/σ_0^2, the power of the F-test is greater in the systematic case than in the random. Four particular cases have been examined numerically as follows:

	(*a*)	(*b*)	(*c*)	(*d*)
Number of groups, k	4	4	12	10
Number of observations in each group, n	6	11	6	11

Values of the power have been calculated, using equations (62) and (65), and are plotted in Fig. 2 (*a*)–(*d*) as ordinates against σ^2/σ_0^2 ($= S^2/\sigma_0^2$). We find from these that the systematic power curve lies above the other; further, we note that the curves are closer to one another in (*c*) and (*d*) than in (*a*) and (*b*), a fact which agrees with theory that the two power functions must tend to each other with increasing k. The errors of approximation in calculating the power in the systematic case are likely to be small judged by the comparative Table 7 and should not affect the relative positions of the power curves.

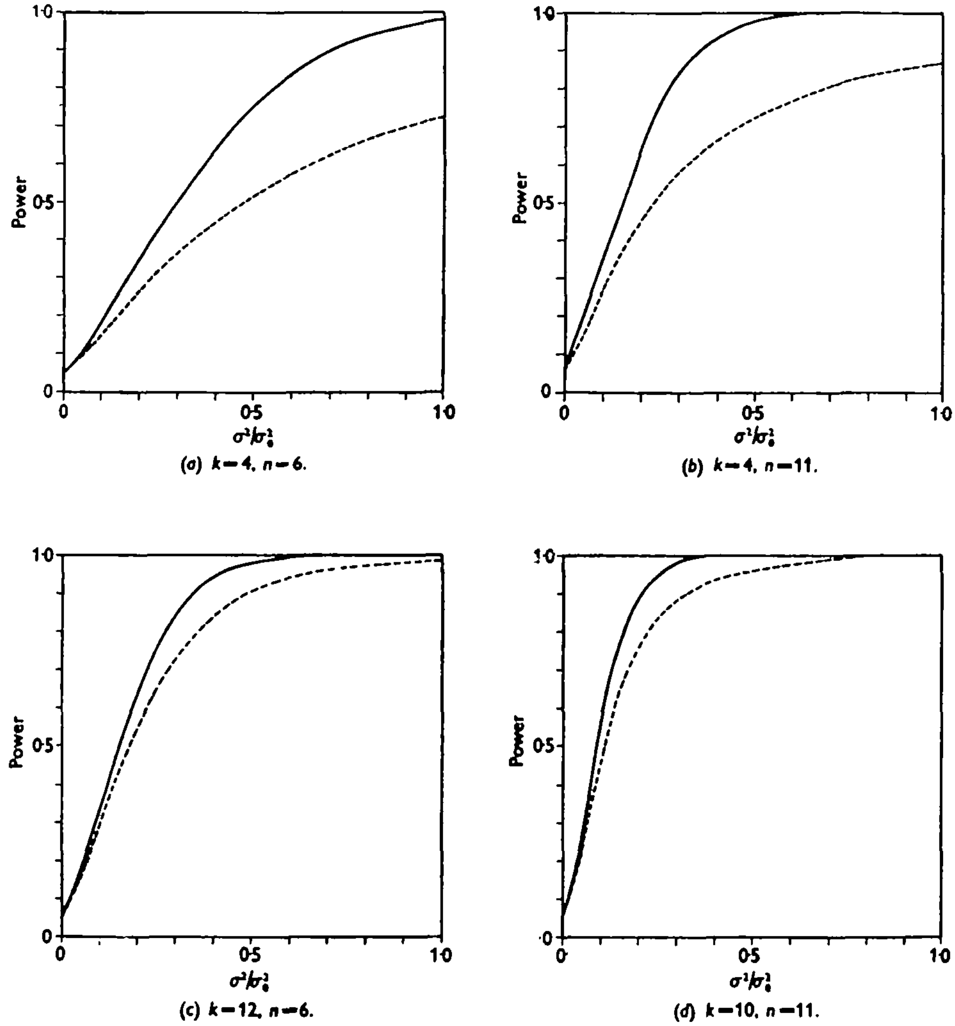

Fig. 2. Power curves for the random and systematic set-ups for k groups with n observations in each: ---- random, —— systematic.

This relation may be interpreted in a different way. Taking case (*a*) above, it will be seen from Fig. 2 (*a*) that we can detect, for instance, a 'systematic' relative group variability of 0·45 with a 70 % chance, while we cannot, with the same chance, detect a variability of magnitude less than 0·9 in the random case. The difference is of course to be expected. For the random set-up, our appreciation of σ^2 is obscured by random variations in both y and z of equation (63); for the systematic set-up, our appreciation of S^2 is only obscured by random fluctuations in the z of equation (56).

7·3. *Applications of the power-function*

We will be concerned here mainly with the systematic set-up and will illustrate the application of our results, taking the simple case of k groups and n observations. The treatment is, however, quite general and could be applied to any designed experiment as outlined in the general statement given at the beginning of § 7·1.

Two types of question may be asked in connexion with the test for differences between groups:

(*a*) What is the extent of departure from the null hypothesis, measured by (S^2/σ_0^2), that could be detected with a given chance?

(*b*) How many observations are we to take in each group so that we could detect a given ratio of between group to within group variability (S^2/σ_0^2) with a prescribed chance?

To answer these questions we have to examine the function $\beta(S^2/\sigma_0^2)$ which may be written in the form

$$\beta(\nu_1, \nu_2, \lambda, \alpha) = e^{-\frac{1}{2}\lambda} \sum_{j=0}^{\infty} \frac{(\frac{1}{2}\lambda)^j}{j!\, B(\frac{1}{2}\nu_1+j, \frac{1}{2}\nu_2)} \int_{\nu_1 F_\alpha/(\nu_2+\nu_1 F_\alpha)}^{1} x^{\frac{1}{2}\nu_1-1+j}(1-x)^{\frac{1}{2}\nu_2-1}\, dx, \quad \text{from (50)},$$

and consider its inverse, i.e. $\lambda \equiv \lambda(\nu_1, \nu_2, \alpha, \beta)$. Generally, λ has to be obtained by inverse interpolation from tables of β such as Tang's. The interval of tabulation of 0·5 for

$$\phi = \sqrt{[\lambda/(\nu_1+1)]}$$

in Tang's tables is not fine enough for interpolation to be satisfactory. Still, they give a trial value of ϕ for which β is calculated and then corrected with the help of the derivative $\partial\beta/\partial\phi$. Following this rather laborious method, Emma Lehmer (1944) has tabled ϕ for $\alpha = 0{\cdot}01$, $0{\cdot}05$ and $\beta = 0{\cdot}7$, $0{\cdot}8$ and for a wide range of ν_1 and ν_2. For these two values of the power we may use her tables to obtain our λ. It would clearly be of value for these tables to be extended.

We may, however, for any set of values of ν_1, ν_2, α and β, get λ approximately with the help of the approximate form of β given in (61). Taking a trial value of λ we can find two consecutive integers λ_1, λ_2 between which λ lies by the following method. From the expression (61) for β we see that λ must satisfy the relation

$$F_\beta(\nu, \nu_2) = \frac{\nu_1}{\nu_1+\lambda} F_\alpha(\nu_1, \nu_2), \tag{66}$$

where the arguments ν, ν_2 and ν_1, ν_2 are the degrees of freedom. Hence the two integers λ_1 and λ_2 would make the right-hand side of (66) just greater and just less than the left-hand side. These can be got by trial and error, taking the α and β percentage points from the F-tables and comparing the two sides. (It is to be noted that ν in (66) involves λ.) For these values of λ_1 and λ_2, β is then evaluated using (62) and by backward interpolation λ is determined.

To deal with inverse problems, such as (*b*) mentioned above, a graphical representation of the relation between ν_1, ν_2 and λ for fixed α and β will be most useful. Following the procedure described above for finding λ, charts have been constructed for $\alpha = 0{\cdot}05$ and for two levels of power, $\beta = 0{\cdot}5$ and $0{\cdot}9$, which are likely to be of practical interest (see Figs. 3 (*a*), (*b*)). The charts give, to the approximation involved in (61), contours of equal power and could be used for determining any one of the three quantities, ν_1, ν_2 and λ, given the other two. When $\nu_2 = \infty$, the F' reduces to χ'^2/ν_1, and hence these charts could also be used for answering the inverse questions connected with the power function of the χ^2-test (see p. 215).

We give here two illustrations of the use of these charts.

Illustration 1. To study the seasonal variation in the frequency of occurrence of a particular dominant alga in a pond, ten samples of 15 c.c. of water are taken from the pond on the first day of each of the five months, April to August. Fifteen drops are taken on slides from each sample after shaking it thoroughly, and the number of algae of the particular form are

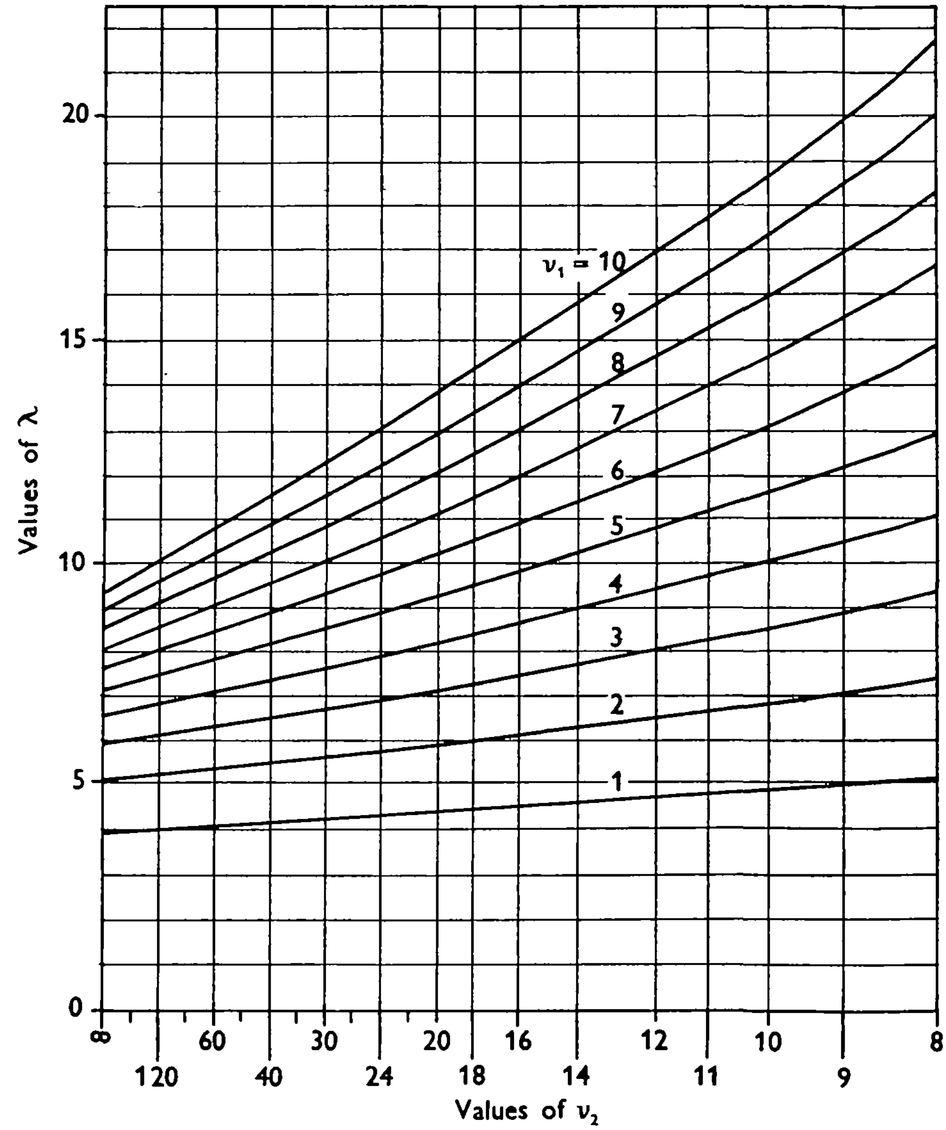

Fig. 3*a*. Contours of equal power for the analysis of variance test with the systematic set-up: $\alpha = 0{\cdot}05$, and a power $\beta(\nu_1, \nu_2, \lambda) = 0{\cdot}5$

counted under the microscope and the total for the fifteen slides is taken as the density for each sample.

To test whether there is significant variation in the density of this form of algae from month to month, the analysis of variance test is applied, say, at 5 % level. It will be of interest to know how large should the ratio of the seasonal variability to the variability in the pond be, so that we could detect it with a 90 % chance.

Here, $\nu_1 = k-1 = 4$, $\nu_2 = k(n-1) = 45$. For these, the chart of Fig. 3 (*b*) gives $\lambda = 16{\cdot}8$, from which we find the ratio of between month to within month variability

$$S^2/\sigma_0^2 = \lambda/nk = 0{\cdot}34, \text{ nearly.}$$

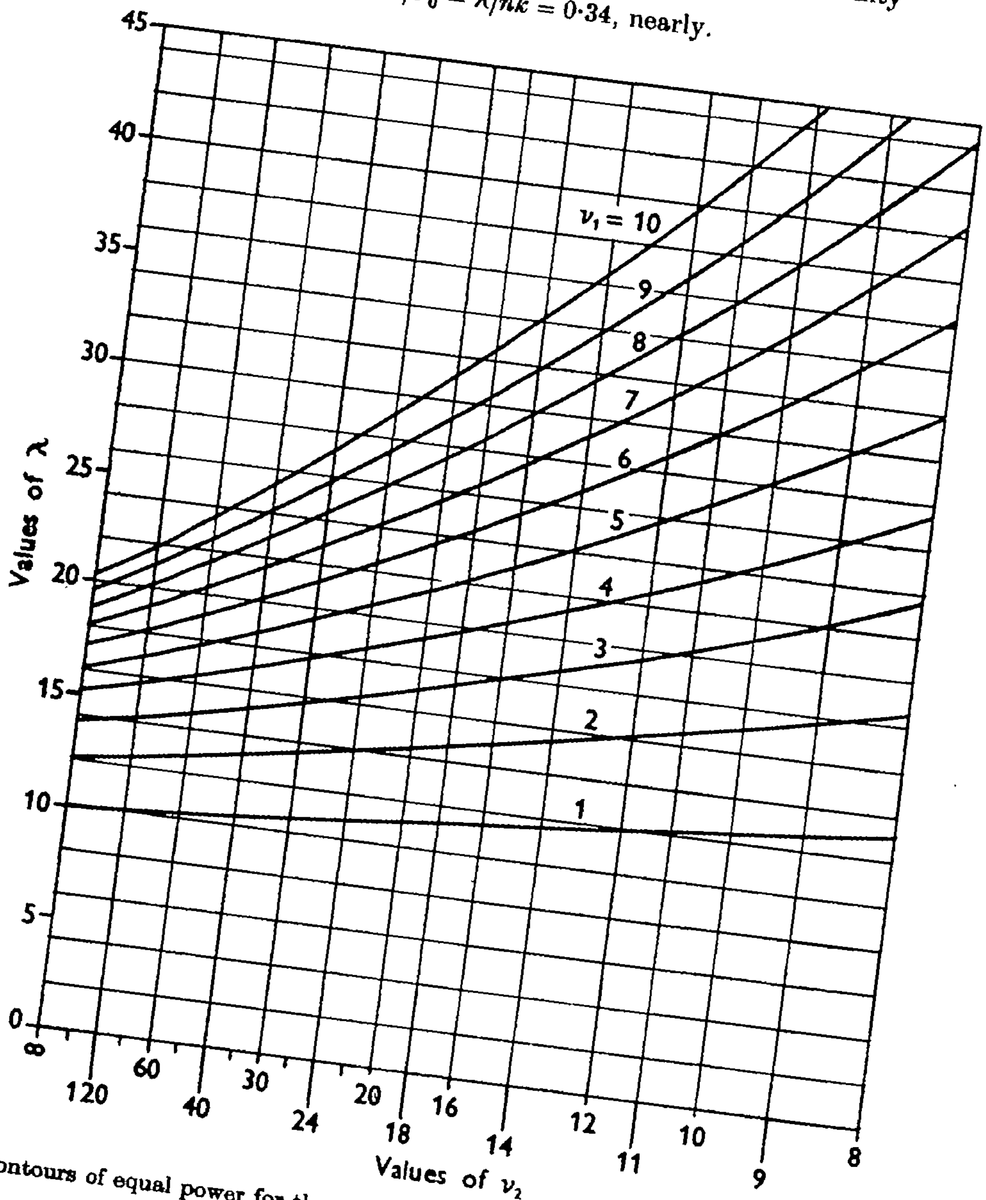

Fig. 3*b*. Contours of equal power for the analysis of variance test with the systematic set-up: $\alpha = 0{\cdot}05$, and a power $\beta(\nu_1, \nu_2, \lambda) = 0{\cdot}9$.

This means that the odds are 9 to 1 on detecting differences at the 5% level if the S.D. of the density of the algae between months was 0·58 of the S.D. within the pond. On the other hand, using Fig. 3 (*a*) we see that there would be a 50 : 50 chance of detecting differences if the S.D. between months is 0·38 of that of a single sample in a month ($S^2/\sigma_0^2 = 0{\cdot}145$).

Illustration 2. There are seven machines producing copper wire for electric cables. It is intended to control the variability in the thickness of the wire due to the machines by taking samples from time to time and testing for differences between the machines. From previous observations we have some idea of the order of variability in the product of a single machine; suppose we do not regard the variability between machines as serious if it does not exceed 0·25 of the within-machine variability. How many samples of wire must we take from each machine to have a 90 % chance of detecting, at the 5 % level for F, a between-machine variability of this magnitude, if it exists?

Since $\dfrac{\lambda}{\nu_2} = \dfrac{n}{n-1}\dfrac{S^2}{\sigma_0^2}$ in virtue of (58), we have now to find n satisfying the relation

$$\frac{\lambda}{\nu_2} = \frac{n}{n-1} \times 0{\cdot}25.$$

Following the contour in chart 3 (b) for $\nu_1 = 6$, we find by inspection a point on it at which the ratio of the co-ordinates is nearly 0·25. This point gives $\nu_2 = 75$ from which we obtain the number of samples required, $n = \nu_2/k+1 = 75/7+1 = 12$, approximately. On the other hand, from 3 (a) we find that we would have a 50 % chance of detection, if $n = 6$.

In conclusion, I should like to acknowledge gratefully the help and guidance I have received from Prof. E. S. Pearson and Dr H. O. Hartley in the course of my investigations.

APPENDIX

Distribution of the sum of squares of independent normal variates with different means and variances

Let $\xi_1, \xi_2, \ldots, \xi_n$ be n independent normal variates with expectations $b_1, b_2, \ldots, b_n$ and variances $v_1, v_2, \ldots, v_n$ respectively. The characteristic function of the statistic

$$\psi^2 = \sum_{j=1}^{n} \xi_j^2 \tag{1}$$

is easily obtained. Introducing $\qquad x_j = (\xi_j - b_j)/\sqrt{v_j}$,

we note that each x_j follows the probability law

$$p(x) = \frac{1}{\sqrt{(2\pi)}} e^{-\frac{1}{2}x^2},$$

and that ψ^2 in (1) can be written as

$$\psi^2 = \Sigma v_j (x_j + a_j)^2, \tag{2}$$

where a_j stands for $b_j/\sqrt{v_j}$. (All summations are from $j = 1$ to n.)

The characteristic function of ψ^2 is given by

$$\phi(t) = \prod_1^n \left[\frac{1}{\sqrt{(2\pi)}} \int_{-\infty}^{\infty} \exp\{it\,v_j(x_j+a_j)^2 - \tfrac{1}{2}x_j^2\}\, dx_j\right]. \tag{3}$$

The integral in (3) is equal to $\quad \sqrt{\left(\dfrac{2\pi}{1-2it\,v_j}\right)} \exp\left\{\dfrac{it\,v_j a_j^2}{1-2it\,v_j}\right\}.$

Hence

$$\phi(t) = \prod_1^n (1-2it\,v_j)^{-\frac{1}{2}} \exp\left\{\frac{\Sigma(itv_j a_j^2)}{1-2it\,v_j}\right\}, \tag{4}$$

from which all the moments of the required distribution can be derived. We may represent this approximately by a χ^2-distribution fitted from the first two moments, $\mu_1' = \Sigma v_j + \Sigma v_j a_j^2$, and $\mu_2 = 2\Sigma v_j^2 + 4\Sigma v_j^2 a_j^2$.

Next we consider the conditional distribution of ψ^2 in (2) subject to a single linear constraint on the x_j's, viz.

$$\Sigma c_j(x_j + a_j) = \rho.$$

The characteristic function of the joint distribution of ψ^2 and ρ is given by

$$\phi(t, t_1) = \Pi\left[\frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{\infty} \exp\{-\tfrac{1}{2}x_j^2 + itv_j(x_j + a_j)^2 + it_1c_j(x_j + a_j)\}\,dx_j\right]. \tag{5}$$

On performing the integrations in (5) we find

$$\phi(t, t_1) = \Pi\left[(1 - 2itv_j)^{-\frac{1}{2}}\exp\left\{\frac{2itv_ja_j^2 + 2it_1c_ja_j - t_1^2c_j^2}{2(1 - 2itv_j)}\right\}\right].$$

The conditional characteristic function of ψ^2, for fixed ρ (Bartlett, 1938), is

$$\phi(t \mid \rho) = \int_{-\infty}^{\infty} e^{-it_1\rho}\phi(t, t_1)\,dt_1 \Big/ \int_{-\infty}^{\infty} e^{-it_1\rho}\phi(0, t_1)\,dt_1$$

$$= \left(\Sigma\left(\frac{c_j^2}{1 - 2itv_j}\right)\right)^{-\frac{1}{2}}\Pi(1 - 2itv_j)^{-\frac{1}{2}}$$

$$\times\exp\left\{\Sigma\left(\frac{itv_ja_j^2}{1 - 2itv_j}\right) - \frac{1}{2}\left(\rho - \Sigma\left(\frac{c_ja_j}{1 - 2itv_j}\right)\right)^2\left(\Sigma\left(\frac{c_j^2}{1 - 2itv_j}\right)\right)^{-1} + \frac{1}{2}\frac{(\rho - \Sigma c_ja_j)^2}{\Sigma c_j^2}\right\}. \tag{6}$$

The moments of the conditional distribution of ψ^2 can then be obtained from (6).

Again, we may fit a Type III to the conditional distribution of ψ^2 by using the first two moments.

REFERENCES

BARTLETT, M. S. (1938). *J. Lond. Math. Soc.* **13**, 62.
CORNISH, E. A. & FISHER, R. A. (1937). *Rev. Inst. Int. Statist.* **5**, 307.
DANIELS, H. E. (1939). *J.R. Statist. Soc. Suppl.* **6**, 186.
DAVID, F. N. (1947). *Biometrika*, **34**, 339.
FISHER, R. A. (1928). *Proc. Roy. Soc.* A, **121**, 654.
FISHER, R. A. (1931). Introduction to the *Brit. Ass. Math. Tables*, **1**, 26.
GOLDBERG, H. & LEVINE, H. (1946). *Ann. Math. Statist.* **17**, 216.
HALDANE, J. B. S. (1937). *Biometrika*, **29**, 133.
HARTLEY, H. O. (1948). *Biometrika*, **35**, 417.
HSU, P. L. (1941). *Biometrika*, **32**, 62.
JOHNSON, N. L. (1948). *Biometrika*, **35**, 80.
JOHNSON, N. L. & WELCH, B. L. (1939). *Biometrika*, **31**, 362.
JORGENSEN, N. R. (1916). *Undersogelser over Frequensflader og Korrelation.* Copenhagen: Busck.
LEHMER, EMMA (1944). *Ann. Math. Statist.* **15**, 388.
MERRINGTON, MAXINE & THOMPSON, CATHERINE M. (1943). *Biometrika*, **33**, 73.
NEYMAN, J. (1935). *J.R. Statist. Soc. Suppl.* **2**, 131.
NEYMAN, J. & TOKARSKA, B. (1936). *J. Amer. Statist. Ass.* **31**, 318.
PATNAIK, P. B. (1948). *Biometrika*, **35**, 157.
PEARSON, K. (1900). *Phil. Mag.* **50**, 157.
PEARSON, K. (1916). *Biometrika*, **11**, 145.
PEARSON, K. (1922). *Tables of the Incomplete Gamma Function.* London: Biometrika.
PEARSON, K. (1934). *Tables of the Incomplete B-Function.* London: Biometrika.
TANG, P. C. (1938). *Statist. Res. Mem.* **2**, 126.
THOMPSON, CATHERINE M. (1941). *Biometrika*, **32**, 188.
WISHART, J. (1932). *Biometrika*, **24**, 441.