THE MULTIVARIATE SADDLEPOINT METHOD AND CHI-SQUARED FOR THE MULTINOMIAL DISTRIBUTION

BY I. J. GOOD

Admiralty Research Laboratory, Teddington, Middlesex[1]

1. Introduction. This paper is largely a continuation of Good [5], but the above title is more descriptive than the previous title would be. The contents are:

(i) Further discussion of the saddlepoint theorem for coefficients in a power of a power series, especially in more than one variable.

(ii) A generalization of Faá di Bruno's formula for the repeated differentiation of a function of a function.

(iii) Some discussion of the relationship between moments and cumulants, especially bivariate ones.

(iv) Some exemplification, but not a systematic exposition, of multivariate notations in analysis, which are less familiar than those used in algebra.

(v) Corrections to the previous paper [5].

(vi) The results of some numerical trials of the method of calculating the distribution of chi-squared for an equiprobable multinomial distribution.

2. Further Formalism and Discussion of the Saddlepoint Theorems. In this section I shall discuss certain formal aspects of the saddlepoint theorems given in Daniels [3] and Good [5]. (In order to minimize repetition, I shall assume that the reader has a copy of [5] ready to hand.) A part of the formalism involves the use of Hermite polynomials in one or more variables. When there is only one variable, Hermite functions are shown to be relevant, for example, by Jeffreys and Jeffreys [6], Para. 23.09. But that context is rather different from ours, and the method of proof, by partial integration, does not appear to be applicable when there is more than one variable.

The formalism will shed further light on why it is desirable to make use of a saddlepoint of the integrand (or of a function closely related to the integrand).

I wish to emphasize that the discussion is formal, and I have not investigated general conditions of validity and bounds for errors. In any specific application some attempt should be made to estimate the error, either analytically or by means of numerical experiments.

I shall take the opportunity of correcting some slips in [5].

Let $\mathbf{M}$ be the column vector whose components are $(M_1, \cdots, M_l)$, and let transposition be denoted by a "prime" or "dash", so that $\mathbf{M}'$ is the corresponding row vector. If $\boldsymbol{\theta}$ is another (1-dimensional column) vector, then $\mathbf{M}'\boldsymbol{\theta}$ represents the scalar product $M_1\theta_1 + \cdots + M_l\theta_l$, in accordance with the usual notation for matrix multiplication, a notation that will be used more generally. Let $c(\mathbf{M}, t) = c(M_1, \cdots, M_l, t)$ be the coefficient of $\mathbf{z}^{\mathbf{M}} = z_1^{M_1} \cdots z_l^{M_l}$ in $(f(\mathbf{z}))^t =$

Received September 29, 1959; revised October 29, 1960.

[1] I am indebted to the Admiralty for permission to publish this paper.

$(f(z_1, \cdots, z_l))^t$, where t is a positive integer, and where M_j/t and t/M_j are bounded ($j = 1, 2, \cdots, l$). The notation $\mathbf{x}^{\mathbf{M}}$ for a "scalar indicial" will be used more generally, for example, $\boldsymbol{\rho}^{\mathbf{M}}$ denotes $\rho_1^{M_1} \cdots \rho_l^{M_l}$. In accordance with a convention often used by physicists, $d\mathbf{z}$, for example, will denote $dz_1\, dz_2 \cdots dz_l$. (The naturalness of this notation is illustrated by the suggestive notation $d\mathbf{z}/d\boldsymbol{\xi}$ for a Jacobian.) Our main purpose in this section is to develop formulae for $c(\mathbf{M}, t)$. We start with

$$c(\mathbf{M}, t) = \frac{1}{(2\pi i)^l} \oint \cdots \oint (f(\mathbf{z}))^t \frac{d\mathbf{z}}{z_1^{M_1+1} \cdots z_l^{M_l+1}} \tag{1}$$

$$\equiv \frac{1}{(2\pi)^l \boldsymbol{\rho}^{\mathbf{M}}} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} (f(\rho_1 e^{i\theta_1}, \cdots, \rho_l e^{i\theta_l}))^t e^{-\mathbf{M}'\boldsymbol{\theta}}\, d\boldsymbol{\theta}. \tag{2}$$

Let $\mathbf{r} = (r_1, \cdots, r_l)'$ be a vector each component of which is a non-negative integer, and let $|\mathbf{r}| = r_1 + \cdots + r_l$, $\mathbf{r}! = r_1! \cdots r_l!$ Similar notation will be used for $\mathbf{s}$ and $\mathbf{n}$.

Consider an (artificial) probability distribution of a random vector $\mathbf{X}$, such that the probability is $p_{\mathbf{n}} = p_{n_1, \cdots, n_l}$ that $\mathbf{X} = \mathbf{n}$; where $p_{\mathbf{n}}$ has the probability generating function $\mathbf{x}^{-\mathbf{M}/t} f(\rho_1 x_1, \cdots, \rho_l x_l)/f(\boldsymbol{\rho})$. The corresponding moment generating function is

$$e^{-\mathbf{M}'\boldsymbol{\xi}/t} f(\rho_1 e^{\xi_1}, \cdots, \rho_l e^{\xi_l})/f(\boldsymbol{\rho}).$$

Let the corresponding cumulants be $\kappa_{\mathbf{r}} = \kappa_{r_1, \cdots, r_l}$, so that, if $|\mathbf{r}| \geqq 0$,

$$\kappa_{\mathbf{r}} = \prod_{j=1}^{l} \left(\frac{\partial}{\partial \xi_j}\right)^{r_j} \left\{-\frac{\mathbf{M}'\boldsymbol{\xi}}{t} - \log f(\boldsymbol{\rho}) + \log f(\rho_1 e^{\xi_1}, \cdots, \rho_l e^{\xi_l})\right\}\Bigg|_{\boldsymbol{\xi}=\mathbf{0}}. \tag{3}$$

The "order" of the cumulant $\kappa_{\mathbf{r}}$ is defined as $|\mathbf{r}|$. Note that $\kappa_{\mathbf{0}} = 0$, and that, if $|\mathbf{r}| = 1$, the cumulants take the values k_j, where

$$\begin{aligned} k_j &= -\frac{M_j}{t} + \rho_j \frac{\partial}{\partial \rho_j} \log f(\boldsymbol{\rho}) \\ &= -\frac{M_j}{t} + \frac{\partial}{\partial \xi_j} \log f(\rho_1 e^{\xi_1}, \cdots, \rho_l e^{\xi_l})\Big|_{\boldsymbol{\xi}=\mathbf{0}} \qquad (j = 1, 2, \cdots, l). \end{aligned} \tag{4}$$

If $|\mathbf{r}| \geqq 2$,

$$\kappa_{\mathbf{r}} = \prod_j \left(\frac{\partial}{\partial \xi_j}\right)^{r_j} \log f(e^{\xi_1}, \cdots, e^{\xi_l})\Big|_{\xi_j = \log \rho_j} \qquad (j = 1, 2, \cdots, l) \tag{5}$$

$$= \prod_j \left(\frac{\partial}{\partial \xi_j}\right)^{r_j} \log f(\rho_1 e^{\xi_1}, \cdots, \rho_l e^{\xi_l})\Big|_{\boldsymbol{\xi}=\mathbf{0}} \tag{6}$$

$$= \prod_j \left(\rho_j \frac{\partial}{\partial \rho_j}\right)^{r_j} \log f(\boldsymbol{\rho}). \tag{7}$$

Formally,

$$c(\mathbf{M}, t) =_{\mathbf{F}} \frac{(f(\boldsymbol{\rho}))^t}{(2\pi)^l \boldsymbol{\rho}^{\mathbf{M}}} \int \cdots \int \exp\left\{t \sum_{\mathbf{r}} \frac{\kappa_{\mathbf{r}}}{\mathbf{r}!} (i\boldsymbol{\theta})^{\mathbf{r}}\right\} d\boldsymbol{\theta}, \tag{8}$$

where $=_F$ means "equals formally". (I am here omitting the ranges of integration since the argument is only formal. Note that in [5] the factor $(f(\varrho))^t$ was omitted in error twice on p. 872 and twice on page 869, and the factor $1/t$ was omitted once on page 872. Also, on page 872, "1+" in the heavy exponential should be deleted. These slips had no real effect on the argument or results.)

To evaluate $c(\mathbf{M}, t)$ approximately we should like to know the saddlepoints of $f(\mathbf{z})\mathbf{z}^{-\mathbf{M}}$. Certainly there is a saddlepoint at $\mathbf{z} = \varrho$ if equations (6.20) of [5] are satisfied. These equations assert the vanishing of the first-order cumulants of our artificial distribution, i.e., they assert $\mathbf{k} = \mathbf{0}$. We do not need to prove that there are no other saddlepoints provided that we can cope directly with our integral, (8), with respect to $\boldsymbol{\theta}$. What we do know, by [5] p. 874, is that there is at most one "real and positive" saddlepoint, i.e., that equations (6.20) ($\mathbf{k} = \mathbf{0}$) have at most one solution with $\rho_1 > 0, \cdots, \rho_l > 0$. But $\boldsymbol{\theta} = \mathbf{0}$ is not necessarily the only important point in the region of integration. (There was some carelessness in [5] concerning this question.) In the next section we shall have an example in which the region of integration with respect to $\boldsymbol{\theta}$ contains two points of equal importance.

Let us now continue with the formal procedure. We have

$$(9) \qquad c(\mathbf{M}, t) =_F \frac{(f(\varrho))^t}{(2\pi)^l \varrho^{\mathbf{M}}} \int \cdots \int e^{it\mathbf{k}'\boldsymbol{\theta} - \frac{1}{2}t^2\boldsymbol{\theta}'\mathbf{K}\boldsymbol{\theta}} \exp\left\{t \sum_{\mathbf{r}}^{|\mathbf{r}| \geq 3} \frac{\kappa_{\mathbf{r}}}{\mathbf{r}!} (i\boldsymbol{\theta})^{\mathbf{r}}\right\} d\boldsymbol{\theta},$$

where $\mathbf{K}$ is the matrix of second-order cumulants

$$(10) \qquad \mathbf{K} = \left(\rho_i \frac{\partial}{\partial \rho_i}\left(\rho_j \frac{\partial}{\partial \rho_j}\right) \log f(\varrho)\right).$$

(I am taking the liberty of using i as a suffix, besides as $\sqrt{(-1)}$.) The determinant of $\mathbf{K}$ is Δ, the Hessian of $\log f(e^{\xi_1}, \cdots, e^{\xi_l})$ (where now $\xi_1 = \log \rho_1$, etc.) and is positive (see [5], p. 874).

If we now imagine the last exponential factor in (9) to be expanded as a (multiple) power series, the various terms of the integrand can be obtained *by partial differentiation* of the first exponential factor with respect to the first-order cumulants. By interchanging the order of differentiation and integration we get

$$c(\mathbf{M}, t) \sim_F \frac{(f(\varrho))^t}{(2\pi)^l \varrho^{\mathbf{M}}} \exp\left\{t \sum_{\mathbf{r}}^{|\mathbf{r}| \geq 3} \frac{\kappa_{\mathbf{r}}}{\mathbf{r}!} \left(\frac{d}{td\mathbf{k}}\right)^{\mathbf{r}}\right\} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{it\mathbf{k}'\boldsymbol{\theta} - \frac{1}{2}t\boldsymbol{\theta}'\mathbf{K}\boldsymbol{\theta}}\, d\boldsymbol{\theta},$$

where

$$\left(\frac{d}{td\mathbf{k}}\right)^{\mathbf{r}} = t^{-|\mathbf{r}|} \left(\frac{\partial}{\partial k_1}\right)^{r_1} \cdots \left(\frac{\partial}{\partial k_l}\right)^{r_l}.$$

But

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{it\mathbf{k}'\boldsymbol{\theta} - \frac{1}{2}t\boldsymbol{\theta}'\mathbf{K}\boldsymbol{\theta}}\, d\boldsymbol{\theta} = \frac{(2\pi)^{\frac{1}{2}l}}{t^{\frac{1}{2}l}\Delta^{\frac{1}{2}}} e^{-\frac{1}{2}t\mathbf{k}'\mathbf{K}^{-1}\mathbf{k}}$$

(see, for example, Cramér [2], p. 119). Therefore

$$(11)\qquad c(\mathbf{M}, t) \sim_{\mathrm{F}} \frac{(f(\boldsymbol{\varrho}))^t}{(2\pi t)^{\frac{1}{2}l}\boldsymbol{\varrho}^{\mathbf{M}}\Delta^{\frac{1}{2}}} \exp\left\{t \sum_{\mathbf{r}}^{|\mathbf{r}|\geqq 3} \frac{\kappa_{\mathbf{r}}}{\mathbf{r}!}\left(\frac{d}{td\mathbf{k}}\right)^{\mathbf{r}}\right\} e^{-\frac{1}{2}t\mathbf{k}'\mathbf{K}^{-1}\mathbf{k}}.$$

Let us now introduce a function $\nu_{\mathbf{r}}$, of $\mathbf{r}$ and t, defined by the identity

$$(12)\qquad \begin{aligned} \sum_{\mathbf{r}}^{|\mathbf{r}|\geqq 0} \frac{\nu_{\mathbf{r}}}{\mathbf{r}!}\boldsymbol{\xi}^{\mathbf{r}} &= \exp\left(t \sum_{\mathbf{r}}^{|\mathbf{r}|\geqq 3} \frac{\kappa_{\mathbf{r}}}{\mathbf{r}!}\boldsymbol{\xi}^{\mathbf{r}}\right) \\ &= \left(\frac{f(\rho_1 e^{\xi_1}, \cdots, \rho_l e^{\xi_l})}{f(\boldsymbol{\varrho})}\right)^t e^{-\mathbf{M}'\boldsymbol{\xi}-t\mathbf{k}'\boldsymbol{\xi}-\frac{1}{2}t\boldsymbol{\xi}'\mathbf{K}\boldsymbol{\xi}}. \end{aligned}$$

Let us also introduce a formal symbol $\boldsymbol{\nu}$, to be manipulated as if it were a vector with l components, and an operator $[\cdots\}$ which has the effect of replacing $\boldsymbol{\nu}^{\mathbf{r}}$ by $\nu_{\mathbf{r}}$. Then the part of (11) beginning "exp" can be written

$$\begin{aligned} \sum \frac{\nu_{\mathbf{r}}}{\mathbf{r}!}\left(\frac{d}{td\mathbf{k}}\right)^{\mathbf{r}} e^{-\frac{1}{2}t\mathbf{k}'\mathbf{K}^{-1}\mathbf{k}} &=_{\mathrm{F}} \left[\sum \frac{\boldsymbol{\nu}^{\mathbf{r}}}{\mathbf{r}!}\left(\frac{d}{td\mathbf{k}}\right)^{\mathbf{r}} e^{-\frac{1}{2}t\mathbf{k}'\mathbf{K}^{-1}\mathbf{k}}\right\} \\ &=_{\mathrm{F}} \left[\exp - \tfrac{1}{2}t\left(\mathbf{k}' + \frac{\boldsymbol{\nu}'}{t}\right)\mathbf{K}^{-1}\left(\mathbf{k} + \frac{\boldsymbol{\nu}}{t}\right)\right\}, \end{aligned}$$

by Taylor's theorem in several variables. We have then

$$(13)\qquad c(\mathbf{M}, t) \sim_{\mathrm{F}} \frac{(f(\boldsymbol{\varrho}))^t}{(2\pi t)^{\frac{1}{2}l}\boldsymbol{\varrho}^{\mathbf{M}}\Delta^{\frac{1}{2}}} \left[\exp - \frac{1}{2t}(t\mathbf{k}' + \boldsymbol{\nu}')\mathbf{K}^{-1}(t\mathbf{k} + \boldsymbol{\nu})\right\}.$$

Now the Hermite polynomial in l variables may be defined as

$$H_{\mathbf{r}}(\mathbf{x} \mid \mathbf{C}) = (-1)^{|\mathbf{r}|} \exp(\tfrac{1}{2}\mathbf{x}'\mathbf{C}\mathbf{x})\,(d/d\mathbf{x})^{\mathbf{r}} \exp(-\tfrac{1}{2}\mathbf{x}'\mathbf{C}\mathbf{x}),$$

where $\mathbf{C}$ is an l by l symmetric matrix. (See, for example, Erdélyi, *et al.*, [4], p. 285.) When $l = 1$, I shall adopt the convention

$$H_n(x) = (-1)^n e^{x^2}(d/dx)^n e^{-x^2} = H_n(x \mid 2).$$

We have then

$$(14)\qquad c(\mathbf{M}, t) \sim_{\mathrm{F}} \frac{(f(\boldsymbol{\varrho}))^t}{(2\pi t)^{\frac{1}{2}l}\boldsymbol{\varrho}^{\mathbf{M}}\Delta^{\frac{1}{2}}} e^{-\frac{1}{2}t\mathbf{k}'\mathbf{K}^{-1}\mathbf{k}} \sum_{\mathbf{r}} \frac{(-1)^{|\mathbf{r}|}\nu_{\mathbf{r}}}{\mathbf{r}!t^{\frac{1}{2}|\mathbf{r}|}} H_{\mathbf{r}}(\mathbf{k}t^{\frac{1}{2}} \mid \mathbf{K}^{-1}).$$

In particular, when $\boldsymbol{\varrho}$ is a saddlepoint of $f(\mathbf{z})\mathbf{z}^{-\mathbf{M}}$, so that $\mathbf{k} = \mathbf{0}$, we have

$$(15)\qquad c(\mathbf{M}, t) \sim_{\mathrm{F}} \frac{(f(\boldsymbol{\varrho}))^t}{(2\pi t)^{\frac{1}{2}l}\boldsymbol{\varrho}^{\mathbf{M}}\Delta^{\frac{1}{2}}} \left[\exp - \frac{1}{2t}\boldsymbol{\nu}'\mathbf{K}^{-1}\boldsymbol{\nu}\right\},$$

$$(16)\qquad c(\mathbf{M}, t) \sim_{\mathrm{F}} \frac{(f(\boldsymbol{\varrho}))^t}{(2\pi t)^{\frac{1}{2}l}\boldsymbol{\varrho}^{\mathbf{M}}\Delta^{\frac{1}{2}}} \sum_{\mathbf{r}} \frac{(-1)^{|\mathbf{r}|}\nu_{\mathbf{r}}}{\mathbf{r}!t^{\frac{1}{2}|\mathbf{r}|}} H_{\mathbf{r}}(\mathbf{0} \mid \mathbf{K}^{-1}).$$

(It is perhaps opportune to remind the reader at this point that $\nu_{\mathbf{r}}$ depends on t.) Now the Hermite polynomials have the generating function

$$\sum \frac{\mathbf{a}^{\mathbf{r}}}{\mathbf{r}!} H_{\mathbf{r}}(\mathbf{x} \mid \mathbf{C}) = \exp\{\tfrac{1}{2}\mathbf{x}'\mathbf{C}\mathbf{x} - \tfrac{1}{2}(\mathbf{x}' - \mathbf{a}')\mathbf{C}(\mathbf{x} - \mathbf{a})\}$$

(see, for example, Erdélyi, *et al.*, [4], p. 285), and in particular,

$$\sum \frac{\mathbf{a}^{\mathbf{r}}}{\mathbf{r}!} H_{\mathbf{r}}(\mathbf{0} \mid \mathbf{C}) = e^{-\frac{1}{2}\mathbf{a}'\mathbf{C}\mathbf{a}} = \sum_{n=0}^{\infty} \frac{(-\frac{1}{2}\mathbf{a}'\mathbf{C}\mathbf{a})^n}{n!}.$$

Therefore $H_{\mathbf{r}}(\mathbf{0} \mid \mathbf{C}) = 0$ if $|\mathbf{r}|$ is odd, and, when $|\mathbf{r}|$ is even,

$$\frac{1}{\mathbf{r}!} H_{\mathbf{r}}(\mathbf{0} \mid \mathbf{C}) = \mathcal{C}(\mathbf{a}^{\mathbf{r}})(-\tfrac{1}{2}\mathbf{a}'\mathbf{C}\mathbf{a})^{\frac{1}{2}|\mathbf{r}|}, \tag{17}$$

where $\mathcal{C}(\cdots)$ means "*the coefficient of* $\cdots$ *in*". So

$$c(\mathbf{M}, t) \sim_{\mathrm{F}} \frac{(f(\boldsymbol{\rho}))^t}{(2\pi t)^{\frac{1}{2}l} \boldsymbol{\rho}^{\mathbf{M}} \Delta^{\frac{1}{2}}} \sum_{\mathbf{r}}^{|\mathbf{r}|\text{ even}} \frac{\nu_{\mathbf{r}}}{(\frac{1}{2}|\mathbf{r}|)!} \left(-\frac{1}{2t}\right)^{\frac{1}{2}|\mathbf{r}|} \mathcal{C}(\mathbf{a}^{\mathbf{r}})(\mathbf{a}'\,\mathbf{K}^{-1}\,\mathbf{a})^{\frac{1}{2}|\mathbf{r}|}, \tag{18}$$

in which, by the way, $\nu_{\mathbf{r}} = 0$ when $|\mathbf{r}| = 2$.

When $l = 2$, and when (ρ, ρ') is a saddlepoint (i.e., equations (6.8) and (6.9) of [5] are satisfied), we have, from (18),

$$c(M, N, t) \sim_{\mathrm{F}} \frac{(f(\rho, \rho'))^t}{2\pi t \rho^M \rho'^N \Delta^{\frac{1}{2}}} \sum_{h,i,j=0}^{\infty} \frac{(-1)^{h+j}}{h!i!j!} \left(\frac{\kappa_{02}}{2t\Delta}\right)^h \left(\frac{\kappa_{11}}{t\Delta}\right)^i \left(\frac{\kappa_{20}}{2t\Delta}\right)^j \nu_{i+2j,2h+i}. \tag{19}$$

When $l = 1$, and when ρ is a saddlepoint (i.e., equation (6.1) of [5] is satisfied), we have

$$c(M, t) \sim_{\mathrm{F}} \frac{(f(\rho))^t}{\sigma \rho^M (2\pi t)^{\frac{1}{2}}} \sum_{s=0}^{\infty} \frac{\nu_{2s}}{s!} \left(-\frac{1}{2\sigma^2 t}\right)^s. \tag{20}$$

There are various methods for calculating the $\nu_{\mathbf{r}}$'s. The most convenient one will depend on the function f and on the computational resources. These methods will depend on the relationships between one or more pairs of the cumulants, $\kappa_{\mathbf{r}}$, the moments, $\mu_{\mathbf{r}}' = E(\mathbf{X}^{\mathbf{r}})$, the moments $\mu_{\mathbf{r}} = E(\mathbf{X} - E\mathbf{X})^{\mathbf{r}}$ about the mean (which are equal to the moments when we are using the saddlepoint method proper), and perhaps the factorial moments,

$$\mu'_{[\mathbf{r}]} = (d/d\mathbf{x})^{\mathbf{r}}(\mathbf{x}^{-\mathbf{M}/t} f(\rho_1 x_1, \cdots, \rho_l x_l)/f(\boldsymbol{\rho})) \mid_{x_1 = \cdots = x_l = 1}. \tag{21}$$

Ordinary moments can be expressed in terms of factorial moments, using Stirling numbers of the second kind. Factorial moments can be expressed in terms of ordinary moments, using Stirling numbers of the first kind. (See Kendall [8], p. 57, and Riordan [12], pp. 33 and 48.) For the case $l = 1$, a table of relationships between moments and cumulants, up to $r = 10$, is given by Kendall [8], pp. 62–64. The $\nu_{\mathbf{r}}$'s can be obtained from the formulae that express the moments in terms of the cumulants by multiplying the cumulants by t and putting those of order less than 3 equal to zero. In this manner I have again checked formula (6.2) of [5].

It was pointed out by Lukacs [9], that, for $l = 1$, the relationships between moments and cumulants can be obtained from Faà di Bruno's formula for the repeated differentiation of a function of a function. When $l > 1$ we can either use the rules given by Kendall [7], or the generalization of Faà di Bruno's formula:

$$(22) \qquad \frac{1}{\mathbf{r}!}\left(\frac{d}{d\mathbf{x}}\right)^{\mathbf{r}} \varphi\,(\boldsymbol{\psi}\,(\mathbf{x})) = \sum \frac{\prod_{\mathbf{s}} f_{\mathbf{s}}^{\mathbf{i_s}}}{\prod_{\mathbf{s}} \mathbf{i_s}\,!}\left(\frac{\partial}{\partial \mathbf{y}}\right)^{\mathbf{j}} \varphi(\mathbf{y}),$$

a formula that will now be explained, and then proved.

(i) $\mathbf{r}$ is a multipartite number, i.e., a vector whose components are non-negative integers.

(ii) The number of components of $\boldsymbol{\psi}$ need not be equal to the number of independent variables $x_1, x_2, \cdots$, i.e., to the number of components of $\mathbf{x}$.

(iii) $\mathbf{y} = \boldsymbol{\psi}(\mathbf{x})$. The notation of partial differentiation on the right implies that $\mathbf{y}$ is not supposed to be expressed in terms of $\mathbf{x}$ before the differentiations are performed.

(iv) $\mathbf{i_s}$ is a function of the vector $\mathbf{s}$ and is itself a vector of dimensionality that of $\boldsymbol{\psi}$ and has non-negative integer components.

(v) $\mathbf{s} \neq \mathbf{0}$.

(vi) $\mathbf{j}$ is an abbreviation for $\sum_{\mathbf{s}} \mathbf{i_s}$.

(vii) $\mathbf{f_s} = (1/\mathbf{s}!)\,(d/d\mathbf{x})^{\mathbf{s}}\boldsymbol{\psi}(\mathbf{x})$.

(viii) By the time the summation sign is to be interpreted, $\mathbf{s}$ has already become a dummy variable, i.e., the summand is not a function of $\mathbf{s}$. The summation is to be performed over all selections of the function $\mathbf{i_s}$ for which $\sum_{\mathbf{s}} \mathbf{s}|\mathbf{i_s}| = \mathbf{r}$; in other words, when $\boldsymbol{\psi}$ is a scalar function, over all partitions of the multipartite number $\mathbf{r}$.

(ix) In conformity with the notation for the factorial of a vector, $\prod_{\mathbf{s}} \mathbf{i_s}!$ means

$$\textstyle\prod_{\mathbf{s}} i_{\mathbf{s}}^{(1)}!\, i_{\mathbf{s}}^{(2)}! \cdots,$$

where $i_{\mathbf{s}}^{(1)}, i_{\mathbf{s}}^{(2)}, \cdots$ are the components of $\mathbf{i_s}$.

Proof of (22): By repeated applications of Taylor's theorem in several variables, together, in the last step, with some applications of the multinomial theorem, we have

$$\begin{aligned}\sum \frac{\mathbf{w^r}}{\mathbf{r}!}\left(\frac{d}{d\mathbf{x}}\right)^{\mathbf{r}} \varphi(\boldsymbol{\psi}(\mathbf{x})) &= \varphi(\boldsymbol{\psi}(\mathbf{x} + \mathbf{w})) \\ &= \varphi\left(\boldsymbol{\psi}(\mathbf{x}) + \sum_{\mathbf{s}}^{|\mathbf{s}|\geq 1} \mathbf{w^s f_s}\right) \\ &= \sum_{\mathbf{r}} \frac{1}{\mathbf{r}!}\left(\sum_{\mathbf{s}}^{|\mathbf{s}|\geq 1} \mathbf{w^r f_s}\right)^{\mathbf{r}}\left(\frac{\partial}{\partial \mathbf{y}}\right)^{\mathbf{r}} \varphi(\mathbf{y}) \\ &= \sum_{\mathbf{i_s}} \frac{\prod_{\mathbf{s}} (\mathbf{w^s f_s})^{\mathbf{i_s}}}{\prod_{\mathbf{s}} \mathbf{i_s}!}\left(\frac{\partial}{\partial \mathbf{y}}\right)^{\mathbf{j}} (\mathbf{y}),\end{aligned}$$

and the result follows on equating coefficients of $\mathbf{w^r}$.

In Faà di Bruno's formula both ψ and x are scalars. Riordan [11], who gives earlier references, points out that the generalization to the case where there are several independent variables is purely a matter of notation. Here I have taken $\boldsymbol{\psi}$ as a vector, for good measure.

By taking $\varphi = \exp$, and ψ as the scalar function $\psi(\mathbf{x}) = \sum (\kappa_{\mathbf{r}}/\mathbf{r}!)\mathbf{x}^{\mathbf{r}}$, we get

$$\frac{1}{\mathbf{r}!}\mu_{\mathbf{r}} = \sum_{i_{\mathbf{s}}} \frac{1}{\prod_{\mathbf{s}} i_{\mathbf{s}}!} \prod_{\mathbf{s}} \left(\frac{\kappa_{\mathbf{s}}}{\mathbf{s}!}\right)^{i_{\mathbf{s}}}. \tag{23}$$

By taking $\varphi = \log$, and $\psi(\mathbf{x}) = \sum (\mu_{\mathbf{r}}/\mathbf{r}!)\mathbf{x}^{\mathbf{r}}$, we get

$$\frac{1}{\mathbf{r}!}\kappa_{\mathbf{r}} = \sum_{i_{\mathbf{s}}} \frac{(-1)^{j-1}(j-1)!}{\prod_{\mathbf{s}} i_{\mathbf{s}}} \prod_{\mathbf{s}} \left(\frac{\mu_{\mathbf{s}}}{\mathbf{s}!}\right)^{i_{\mathbf{s}}}. \tag{24}$$

(Cf., Lukacs [9], for the case $l = 1$.) In both formulae the summation is over all $i_{\mathbf{s}}$'s for which $i_{\mathbf{s}} \geqq 0$, $\sum_{\mathbf{s}} \mathbf{s} i_{\mathbf{s}} = \mathbf{r}$, and $|\mathbf{s}| \geqq 2$, and j means $\sum_{\mathbf{s}} i_{\mathbf{s}}$. If the moments, $\mu_{\mathbf{s}}$, are replaced by $\mu_{\mathbf{s}}'$, then the condition $|\mathbf{s}| \geqq 2$ in (23) and (24) is to be replaced by $|\mathbf{s}| \geqq 1$. To get the $\nu_{\mathbf{r}}$'s we replace the cumulants by $t\kappa_{\mathbf{s}}$ in (23), and put $\kappa_{\mathbf{s}} = 0$ when $|\mathbf{s}| = 2$.

For example, when $l = 2$, we have

$$\begin{aligned}
&\kappa_{rs} = \mu_{rs} \quad \text{if} \quad 2 \leqq r + s \leqq 3,\\
&\kappa_{40} = \mu_{40} - 3\mu_{20}^2, \ \kappa_{31} = \mu_{31} - 3\mu_{20}\mu_{11}, \ \kappa_{22} = \mu_{22} - \mu_{20}\mu_{02} - 2\mu_{11}^2, \text{ etc.}\\
&\mu_{40} = \kappa_{40} + 3\kappa_{20}^2, \ \mu_{31} = \kappa_{31} + 3\kappa_{20}\kappa_{11}, \ \mu_{22} = \kappa_{22} + \kappa_{20}\kappa_{02} + 2\kappa_{11}^2, \text{ etc.}\\
&\kappa_{60} = \mu_{60} - 15\mu_{40}\mu_{20} - 10\mu_{30}^2 + 30\mu_{20}^3\\
&\mu_{60} = \kappa_{60} + 15\kappa_{40}\kappa_{20} + 10\kappa_{30}^2 + 15\kappa_{20}^3\\
&\kappa_{51} = \mu_{51} - 5\mu_{40}\mu_{11} - 10\mu_{31}\mu_{20} - 10\mu_{30}\mu_{21} + 30\mu_{20}^2\mu_{11}\\
&\mu_{51} = \kappa_{51} + 5\kappa_{40}\kappa_{11} + 10\kappa_{31}\kappa_{20} + 10\kappa_{30}\kappa_{21} + 15\kappa_{20}^2\kappa_{11}\\
&\kappa_{42} = \mu_{42} - \mu_{40}\mu_{02} - 8\mu_{31}\mu_{11} - 6\mu_{22}\mu_{20} - 4\mu_{30}\mu_{12} - 6\mu_{21}^2 + 6\mu_{20}^2\mu_{02} + 24\mu_{20}\mu_{11}^2\\
&\mu_{42} = \kappa_{42} + \kappa_{40}\kappa_{02} + 8\kappa_{31}\kappa_{11} + 6\kappa_{22}\kappa_{20} + 4\kappa_{30}\kappa_{12} + 6\kappa_{21}^2 + 3\kappa_{20}^2\kappa_{02} + 12\kappa_{20}\kappa_{11}^2\\
&\kappa_{33} = \mu_{33} - 3\mu_{31}\mu_{02} - 9\mu_{22}\mu_{11} - 3\mu_{13}\mu_{20} - \mu_{30}\mu_{03} - 9\mu_{21}\mu_{12} + 18\mu_{20}\mu_{02}\mu_{11}\\
&\qquad\qquad + 12\mu_{11}^2\\
&\mu_{33} = \kappa_{33} + 3\kappa_{31}\kappa_{02} + 9\kappa_{22}\kappa_{11} + 3\kappa_{13}\kappa_{20} + \kappa_{30}\kappa_{03} + 9\kappa_{21}\kappa_{12} + 9\kappa_{20}\kappa_{02}\kappa_{11} + 6\kappa_{11}^2
\end{aligned} \tag{25}$$

The other three pairs of bivariate formulae of order six can be written down by symmetry. For our application the formulae of order five, and other odd orders, are irrelevant. The above formulae for the moments and cumulants check one another. A further check is that when a cumulant of order r is expressed linearly in terms of moments about the mean, the sum of the coefficients is equal to $r!\,\mathcal{C}(x^r) \log(e^x - x)$, whatever be the value of 1. Similarly, when a moment (of order r) about the mean is expressed linearly in terms of the cumulants, then the sum of the coefficients is equal to $r!\,\mathcal{C}(x^r) \exp(e^x - 1 - x)$. (These two assertions are readily proved by putting all the moments about the mean equal to 1 first, and second putting all the cumulants equal to 1.) Up to order ten we can therefore check against the case $l = 1$, given by Kendall [8].

Up to order six the formulae for the bivariate $\nu_{\mathbf{r}}$'s in terms of the bivariate cumulants are

$$(26) \qquad \begin{aligned} \nu_{00} &= 1,\ \nu_{rs} = 0 \quad \text{if} \quad 1 \leqq r + s \leqq 2 \\ \nu_{rs} &= t\kappa_{sr} \quad \text{if} \quad 3 \leqq r + s \leqq 4 \\ \nu_{60} &= t\kappa_{60} + 10t^2\kappa_{30}^2 \\ \nu_{51} &= t\kappa_{51} + 10t^2\kappa_{30}\kappa_{21} \\ \nu_{42} &= t\kappa_{42} + 4t^2\kappa_{30}\kappa_{12} + 6t^2\kappa_{21}^2 \\ \nu_{33} &= t\kappa_{33} + t^2\kappa_{30}\kappa_{03} + 9t^2\kappa_{21}\kappa_{12}\,. \end{aligned}$$

We can write down ν_{24}, ν_{15}, ν_{06}, by symmetry.

For the case $l = 2$, the sum in formula (18) can now be written

$$(27) \quad 1 + \frac{1}{8t}\sum_{r,s}^{r+s=4} \kappa_{rs}\,\mathcal{C}(a_1^r\,a_2^s)(\mathbf{a}'\mathbf{K}^{-1}\mathbf{a})^2 - \frac{1}{48t^3}\sum_{r,s}^{r+s=6} \nu_{rs}\,\mathcal{C}(a_1^r\,a_2^s)(\mathbf{a}'\mathbf{K}^{-1}\mathbf{a})^3 + \cdots.$$

The last term given explicitly contains terms of order t^{-1}, since the ν_{rs}'s of order six contain terms of order t^2, but no omitted term does so. It is possible that, for any finite value of t, we should sometimes get a more accurate result by using the terms shown explicitly here than by using the terms shown explicitly in [5], formula (6.10). The difference between the explicit parts of these formulae is

$$(28) \qquad \frac{1}{48t^2}\sum_{r,s}^{r+s=6} \kappa_{rs}\,\mathcal{C}(a_1^r\,a_2^s)(\mathbf{a}'\mathbf{K}^{-1}\mathbf{a})^3.$$

To conclude this section I should like to summarize some advantages of using a saddlepoint method proper.

(i) It is more difficult to justify formula (14) for the more general method (in which the first-order cumulants of our artificial random variable do not all vanish), because the modulus of the integrand in (9) is liable not to decrease rapidly enough when we move away from its maximum.

(ii) It is only when $\mathbf{k} = \mathbf{0}$ that the series (14) consists of terms of smaller and smaller order. For when $\mathbf{k} \neq \mathbf{0}$ it can be shown that $H_{\mathbf{r}}(\mathbf{k}t^{\frac{1}{2}} \mid \mathbf{C})$ is of order as large as $t^{\frac{1}{2}|\mathbf{r}|}$.

(iii) When $\mathbf{k} = \mathbf{0}$ the Hermite polynomials vanish for odd values of $|\mathbf{r}|$, and also simplify for even values of $|\mathbf{r}|$.

Nevertheless it seems worthwhile to notice the existence of the formulae with $\mathbf{k} \neq \mathbf{0}$, since

(a) Saddlepoints do not always exist. An example is given in the next section in which there is at any rate no real saddlepoint.

(b) Even when a saddlepoint exists it is often numerically laborious to compute it.

(c) The more general formulae are of some mathematical interest and enable one to see the saddlepoint method in a more general context.

3. Chi-squared for the Equiprobable Multinomial Distribution. In [5] I gave a method of obtaining a saddlepoint approximation for the probability "density" of

$$\chi^2 = tN^{-1} \textstyle\sum_r (n_r - N/t)^2,$$

or equivalently of $S = \sum n_r^2$, for a t-category equiprobable multinomial distribution of sample size N, where the cell entries are $n_0, n_1, \cdots, n_{t-1}$. The purpose of the present section is to continue the discussion of this method. The calculation required the solution of two equations for a saddlepoint (ρ, ρ') (equations (8.3) and (8.4) of [5].) In discussions between Mr. Peter John Taylor and myself, arising out of attempts to solve the equations on an electronic computer (Pegasus at the Admiralty), we discovered that these equations do not have a solution when $\chi^2 > t$, and the saddlepoint method appears to break down. This is unfortunate since $\chi^2 > t$ is much the more interesting case in most applications. On page 877 of [5] I erroneously supposed that there is always a solution.

There is, however, a way round the difficulty.

Let the probability that $S = M$ (i.e., that $\chi^2 = tMN^{-1} - N$) be denoted by $p(M \mid N, t)$. Then

$$p(M \mid N, t) = \mathcal{C}(x^M y^N) N!\, t^{-N} (f_L)^t, \tag{29}$$

where

$$f_L = f_L(x, y) = \sum_{n=0}^{L} \frac{x^{n^2} y^n}{n!},$$

and $L \geqq \min(N, M^{\frac{1}{2}})$. (There is little inaccuracy in taking L smaller provided that $P(\max n_i > L)$ is negligible. The probability can be estimated as in [5].)

In [5] I took $L = \infty$. By taking L finite it turns out that the saddlepoint equations, namely

$$\sum_0^L \frac{n^2 \rho^{n^2} \rho'^n}{n!} = \frac{M}{t} \sum_0^L \frac{\rho^{n^2} \rho'^n}{n!}, \tag{30}$$

$$\sum_0^L \frac{n \rho^{n^2} \rho'^n}{n!} = \frac{N}{t} \sum_0^L \frac{\rho^{n^2} \rho'^n}{n!}, \tag{31}$$

always have a solution when $\chi^2 > t$, except perhaps in the trivial case $M = N^2$. (The solution is unique in virtue of [5], p. 874.) This statement is a special case of the following more general one.

Suppose $M \neq N^2$. Let non-negative integers be defined by the inequalities $\mu^2 < M/t \leqq (\mu + 1)^2$, $\nu < N/t \leqq \nu + 1$. *Then the simultaneous equations* (30) *and* (31) *have a solution if* (i) $\mu > \nu$, *and also if* (ii) $\mu = \nu$ *and*

$$\frac{\chi^2}{t} \cdot \frac{N}{t} > \left(\frac{N}{t} - \nu\right)\left(\nu + 1 - \frac{N}{t}\right); \tag{32}$$

but they do not have a solution if (iii) $\mu = \nu$ *and*

$$\frac{\chi^2}{t} \cdot \frac{N}{t} < \left(\frac{N}{t} - \nu\right)\left(\nu + 1 - \frac{N}{t}\right). \tag{33}$$

(For example, there is no finite solution if $N < t$, $\chi^2 < t - N$.) This statement covers essentially all cases since it is impossible for μ to be less than ν.

Outline of proof. For each $\rho > 0$, equation (30) has a positive solution for ρ'. Call it $\rho_1' = \rho_1'(\rho)$. Similarly equation (31) has a solution $\rho_2' = \rho_2'(\rho)$. The condition for the existence of a joint solution is that $\rho_1' - \rho_2'$ changes sign when ρ increases from 0 to ∞. (ρ_1' and ρ_2' are continuous functions of ρ.)

Equation (30) can be written in the form

$$\frac{M}{t} + \cdots + \frac{\frac{M}{t} - \mu^2}{\mu!}\rho^{\mu^2}\rho'^{\mu} = \frac{(\mu+1)^2 - \frac{M}{t}}{(\mu+1)!}\rho^{(\mu+1)^2}\rho'^{\mu} + \cdots + \frac{L^2 - \frac{M}{t}}{L!}\rho^{L^2}\rho'^{L},$$

in which the terms on both sides are all non-negative. When ρ is very small it turns out that we can approximate the relationship between ρ and ρ_1' by retaining only the last term on the left and the first term on the right (or the second one if M/t is an integer that is a perfect square). But when $\rho \to \infty$ it turns out that we need retain only the first term on the left and the last one on the right (even if $L = N$, $t = 1$, provided that $M \neq N^2$). (If L were infinite there would not be a last term on the right.) We find

$$\rho_1' \sim \left(\frac{L!}{L^2 - M/t} \cdot \frac{M}{t}\right)^{\frac{1}{L}} \rho^{-L} \quad \text{as} \quad \rho \to \infty,$$

and similarly

$$\rho_2' \sim \left(\frac{L!}{L - N/t} \cdot \frac{N}{t}\right)^{\frac{1}{L}} \rho^{-L} \quad \text{as} \quad \rho \to \infty.$$

Also

$$\rho_1' \sim (\mu + 1)\,\frac{\frac{M}{t} - \mu^2}{(\mu+1)^2 - \frac{M}{t}}\,\rho^{-2\mu-1} \quad \text{as} \quad \rho \to 0,$$

if

$$M/t \neq (\mu + 1)^2;$$

$$\rho_2' \sim (\nu + 1)\,\frac{\frac{N}{t} - \nu}{\nu + 1 - \frac{N}{t}}\,\rho^{-2\nu-1} \quad \text{as} \quad \rho \to 0,$$

if

$$N/t \neq \nu + 1;$$

$$\rho_1' \sim A_1\rho^{-2\mu-2} \quad \text{as} \quad \rho \to 0,$$

if $$M/t = (\mu + 1)^2 \text{ (where } A_1 \text{ is independent of } \rho),$$

$$\rho_2' \sim A_2 \rho^{-2\nu-2} \quad \text{as} \quad \rho \to 0,$$

if $$N/t = \nu + 1 \text{ (where } A_2 \text{ is independent of } \rho).$$

We now see easily that $\rho_1' < \rho_2'$ when $\rho \to \infty$, if $M < N^2$.

When $\rho \to 0$, then $\rho_1' \gg \rho_2'$ if $\mu > \nu$. Now

$$\frac{M}{t} = \left(\frac{N}{t}\right)^2 + \frac{\chi^2}{t} \cdot \frac{N}{t}, \tag{34}$$

so if $N/t = \nu + 1$, we have $\mu \geqq \nu + 1 > \nu$. Therefore $\rho_1' \gg \rho_2'$ if N/t is an integer. If N/t is not an integer, then $\rho_1' \gg \rho_2'$ if $M/t = (\mu + 1)^2$ (since $\mu \geqq \nu$). Finally if $N/t \neq \nu + 1$, $M/t \neq (\mu + 1)^2$, and $\mu = \nu$; then $\rho_1' \gtrless \rho_2'$ when $\rho \to 0$ provided that

$$\frac{\dfrac{M}{t} - \mu^2}{(\mu + 1)^2 - \dfrac{M}{t}} \gtrless \frac{\dfrac{N}{t} - \mu}{\mu + 1 - \dfrac{N}{t}},$$

and this reduces to the asserted condition in virtue of the relation (34).

We may observe that, if $L = \infty$, equations (30) and (31), which are now equations (8.3) and (8.4) of [5], can be solved when $\rho = 1$, and give

$$\rho_1'(1) = -\frac{1}{2} + \left(\frac{M}{t} + \frac{1}{4}\right)^{\frac{1}{2}}, \qquad \rho_2'(1) = \frac{N}{t},$$

and so

$$\rho_1'(1) > \rho_2'(1) \quad \text{if} \quad \chi^2 > t$$

$$\rho_1'(1) < \rho_2'(1) \quad \text{if} \quad \chi^2 < t.$$

We may therefore expect that, even if L is finite, the value of ρ satisfying equations (30) and (31) usually exceeds 1 when $\chi^2 > t$ and is usually less than 1 (if it exists) when $\chi^2 < t$. Of course, when $L = \infty$, values of ρ exceeding 1 are not legitimate, and I suspect that $f(x, y)$ cannot be continued analytically across the boundary $|x| = 1$.

Mr. P. J. Taylor has kindly written a Pegasus Autocode program for the solution of equations (30) and (31). Using this program, pairs of values of ρ and ρ' were obtained with $N = t = 10$, $L = 11$ ($L = 10$ would have been adequate) and $M = 28(2)46$. As evidence of the correctness of the program, I quote two pairs of results. For $M = 28$, we obtained $\rho = 1.1552500$, $\rho' = 0.5914205$; and for $M = 46$, $\rho = 1.2158011$, $\rho' = 0.4119696$. (The seventh places of decimals are unreliable.) The method of solution was to start with a trial value of ρ (either guessed or derived from the preceding value of M), then to solve equation (31) for ρ', then use this value of ρ' to solve equation (30) for ρ, and so on; the whole procedure being greatly speeded up by assuming the consecutive differences in the values of ρ to form a geometrical progression. This procedure can be seen to converge by drawing the graphs of $\rho_1'(\rho)$ and $\rho_2'(\rho)$. (The procedure would diverge if equations (30) and (31) were taken in the opposite order.) The solution of each equation separately was obtained by a crude method of bisection, the time being halved by making use of some fairly readily estimated upper and lower bounds. (Each bisection gives one additional binary place.)

Had our intention been to produce tables we would have taken L considerably larger than N, so that the results would be applicable for considerably larger values of N, M, and t, having the same ratios, or as trial values when the ratios were approximately the same. (This point will be exemplified below.)

We can now apply Theorem 6.2 of [5], but with the following slight modification.

In the present problem the parity of M is the same as that of N, otherwise $p(M, N, t)$ vanishes. Therefore condition (6.16) of [5] is not satisfied (see [5], pp. 873 and 874). There are here two equally important saddlepoints, namely at (ρ, ρ') and at $(-\rho, -\rho')$. These make numerically equal contributions and the signs agree or disagree according as M and N are of the same or of opposite parities. Therefore the formulae for $c(M, N, t)$ and $p(M, N, t)$ need to be doubled when M and N are of the same parity. This point was overlooked in [5], and formula (8.5) should read

$$P(\chi^2 = t \mid N = t) = \begin{cases} \dfrac{1}{\sqrt{\pi t}}\left(1 + \dfrac{1}{6t} + \cdots\right) & (t \text{ even}) \\ 0 & (t \text{ odd}). \end{cases} \tag{35}$$

In order to calculate formula (6.10) of [5] I used another Pegasus autocode program. The values of ρ and ρ' corresponding to $t = N = 10$, $L = 11$, $M = 28(2)46$ were those obtained in the previous program. For $M = 20$, ρ and ρ' are both approximately equal to 1 (cf. [5], p. 877), and this case gave a good approximate check on the program.

Column (i) of the table gives all the possible values, a, of χ^2, when $N = t = 10$ (and also the impossible values $a = 38$ and $a = 46$). Column (ii) gives the precise probabilities that $\chi^2 = a$. This column was kindly calculated by Mr. P. J. Taylor, by using the formula

$$P(S = M) = \sum \frac{N!}{n_1! \cdots n_t!} \cdot \frac{1}{t^N} \qquad (M = a + 10),$$

summed over all $n_1, \cdots, n_t$ for which $n_1 + \cdots + n_t = N$, $n_1^2 + \cdots + n_t^2 = M$ (I give these exact probabilities in full detail in case the reader wishes to test some other method of approximation.) The number in brackets, following each probability, is the number of partitions of X corresponding to that probability. The smallness of these numbers of partitions suggests that X is likely to be too small for our asymptotic approximations to be very accurate.

Column (iii) gives the gamma-variate approximation (obtained from the tables of Pearson [10]),

$$\frac{1}{2^{4.5}\Gamma(4.5)} \int_{a-1}^{a+1} e^{-\frac{1}{2}\zeta} \zeta^{3.5}\, d\zeta,$$

TABLE

Accurate and Approximate Values of $P(\chi^2 = a)$ when $N = t = 10$

(i) a	(ii) $P(\chi^2 = a)$	(iii) Gamma approx.	(iv) 1st term of (6.10) doubled.	(v) % age correction to allow for next term.	(vi) Two terms of (6.10) doubled.
0	.000362880 (1)				
2	.016329600 (1)				
4	.114307200 (1)	.130			
6	.212284800 (2)	.197			
8	.223851600 (2)	.199			
10	.193369680 (2)	.163	.180		
12	.082555200 (3)	.114			
14	.085730400 (2)	.071			
16	.031752000 (2)	.0422			
18	.015699600 (3)	.0235	.01935	−76	.00457
20	.010160640 (2)	.0125	.01200	−71	.00354
22	.007620480 (1)	.0066	.007645	−63	.00283
24	.002850120 (3)	.00322	.004942	−56	.00217
26	.001383480 (2)	.00159	.003222	−50	.00160
28	.000181440 (1)	.000739	.002109	−45	.00115
30	.000635040 (1)	.000354	.001383	−42	.000809
32	.000725760 (2)	.000163	.000906	−38	.000560
34	.000045360 (1)	.000074	.000593	−35	.000384
36	.000060480 (1)	.0000334	.000387	−33	.000260
38	.000000000 (0)	.0000150			
40	.000001134 (1)				
42	.000062370 (2)				
44	.000025920 (1)				
46	.000000000 (0)				
48	.000001080 (1)				
56	.000003240 (1)				
58	.000000405 (1)				
72	.000000090 (1)				
90	.000000001 (1)				
Total	1.000000000				

in which the range of integration corresponds to the use of the continuity correction apparently first published by Cochran [1], which is analogous to that used for 2×2 contingency tables. Column (iv) gives the values obtained from the leading term of (6.10), after doubling for the reason mentioned above. Column (v) gives the percentage corrections to be made to allow for the term of order $1/t$. These corrections are disappointingly large, and, when above 60%, do not improve the estimates. Column (vi) gives (6.10) to order $1/t$.

The gamma-variate approximation to the *tail-area* probability (with continuity correction) is never wrong by more than a factor of 2, in the range covered by the above table, when the tail-area probability exceeds $\frac{1}{1000}$, even though the expectation in each cell of our multinomial distribution is only 1. It would be interesting to know whether (6.10) of [5] would give better results for larger values of N and t, with $N \leqq t$. The exact values of $p(M, N, t)$ could be calculated from (4.6) of [5], or from a recurrence relation derivable from it. The exact calculation of $p(M, N, t)$ for all $M \leqq M_0$, and all $N \leqq N_0$, would require about $\frac{1}{2}M_0^2 N_0^2 \log_2 t$ multiplications.

REFERENCES

[1] W. G. Cochran, "The χ^2 correction for continuity," *Iowa State Col. J. Sci.*, Vol. 16 (1942), pp. 421–436.

[2] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, 1946.

[3] H. E. Daniels, "Saddlepoint approximations in statistics," *Ann. Math. Stat.*, Vol. 25 (1954), pp. 631–650.

[4] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, Staff of the Bateman manuscript project, *Higher Transcendental Functions*, Vol. II, McGraw-Hill, New York, Toronto, London, 1953.

[5] I. J. Good, "Saddlepoint methods for the multinomial distribution," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 861–881.

[6] Harold Jeffreys and Bertha Swirles Jeffreys, *Methods of Mathematical Physics*, University Press, Cambridge, 1946.

[7] M. G. Kendall, "The derivation of multivariate sampling formulae from univariate formulae by symbolic operation," *Ann. Eugenics*, London, Vol. 10 (1941), pp. 392–402.

[8] Maurice G. Kendall, *The Advanced Theory of Statistics*, Vol. I, Charles Griffin, London, 1945.

[9] Eugene Lukacs, "Applications of Faà di Bruno's formula in mathematical statistics," *Amer. Math. Monthly*, Vol. 62, (1955), pp. 340–348.

[10] Karl Pearson, *Tables of the Incomplete* Γ*-Function*, Biometrika Office, London, 1922, 1934.

[11] John Riordan, "Derivatives of composite functions," *Bull. Amer. Math. Soc.*, Vol. 52 (1946), pp. 664–667.

[12] John Riordan, *An Introduction to Combinatorial Analysis*, John Wiley, New York; Chapman and Hall, London, 1958.