1953

Asymptotic distributions for tests of contingency hypotheses

Frank Richard Marshall
The University of Montana

ASYMPTOTIC DISTRIBUTIONS FOR TESTS
OF CONTINGENCY HYPOTHESES

by

FRANK RICHARD MARSHALL

B. A. Carroll College, 1952

Presented in partial fulfullment of the requirements for the degree of

Master of Arts

MONTANA STATE UNIVERSITY

1953

Approved by:

Chairman, Board of Examiners

Dean, Graduate School

Aug 17 1953
Date

ACKNOWLEDGMENTS

The author is especially indebted to Professor George Marsaglia for the suggestion of the subject of this dissertation and for his generous assistance and constant guidance throughout the entire work.

F. R. M.

1. Introduction.

The tests of the hypotheses for "goodness of fit" and for independence in contingency tables are among the most well-known and most widely used in applied statistics. Unfortunately, while the number of "cook-book" applications is high, rigorous treatments of the subjects are quite scarce. It is the purpose of this thesis to present a rigorous development of these two theories and of the fundamental results upon which they depend. As the development of all the requisite theorems would require elaboration of the text-book proportion, only those definitions and theorems which are new or not well represented in the literature have been given in detail. Certain other results that are well known and readily available are merely stated as needs require. In addition, many of the methods of probability theory are assumed without specific mention; for example, the one-to-one correspondence between a density function and its moment generating function, providing the latter exists in the neighborhood of the origin.

2. Preliminary Results.

Theorem 1: *If the n.n matrix A is symmetric there exists an orthogonal matrix P such that*

$$PAP' = D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ & \cdots & & \\ 0 & 0 & \cdots & d_n \end{pmatrix} \tag{1}$$

where the d_i are the characteristic roots of A. We note that, from equations (1)

$$(2) \qquad D^{-1} = (PAP')^{-1} = PA^{-1}P' = \begin{pmatrix} 1/d_1 & 0 & \cdots & 0 \\ & \cdots & & \\ 0 & 0 & & 1/d_n \end{pmatrix}$$

Theorem 2: Let A and B be matrices of rank r and s respectively such that the product AB is defined and has rank t. Then

$$t \leq \inf \{r, s\}$$

Theorem 3: If the m·n matrix B is of rank r, and if $\mathcal{J} = (z_1, \ldots, z_n)$, then the number of linearly independent solutions to

$$\mathcal{J} B = (0, \ldots, 0)$$

is m - r.

The following theorem does not appear in the literature, so that a proof is given here.

Theorem 4: Let A be a matrix of order m·n and of rank r. Then AA' is of rank r.

Proof: Since (A'A)' = A'A, we see that A'A is symmetric and apply theorem 1 to find P so that

$$P'A'AP = (AP)'(AP) = D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ & \cdots & & \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

Let the columns of P be $(P_1, \ldots, P_n)$, then the columns of AP are $(AP_1, \ldots, AP_n)$. Now

$$d_k = (AP_k)'(AP_k) = \sum_{j=1}^{m} \left(\sum_{i=1}^{n} a_{ji}p_{ik}\right)^2$$

hence

$$d_k = 0$$

implies that

$$AP_k = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

Now since P^{-1} exists, $P_1, \ldots, P_n$ are independent, and by theorem 3, the number of linearly independent solutions to

$$AP_k = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

cannot exceed $n - r$. The number of d's such that $d = 0$ is $n - s$, where s is the rank of A'A, hence

$$n - s \leq n - r$$

and

$$r \leq s$$

but, by theorem 2

$$r \geq s$$

hence

$$r = s$$

where r is the rank of A and s is the rank of A'A.

Preliminary remarks to theorem 5: Let $x_1, \ldots, x_n$ be random variables with $E(x_i) = u_i$, $i = 1, \ldots, n$. Let ξ be the vector $(x_1, \ldots, x_n)$, and define $E(\xi)$ as the vector $(u_1, \ldots, u_n)$. In particular, if $\alpha = ((x_1 - u_1), \ldots, (x_n - u_n))$

$$E(\alpha) = (0, \ldots, 0).$$

Then the covariance matrix of the vector ξ is

$$S = \text{cov}(\xi) = (\sigma_{ij}) = (E(x_i - u_i)(x_j - u_j))$$

<u>Theorem 5: Let</u> $x_1, \ldots, x_n$ <u>be random variables, let</u> $\xi = (x_1, \ldots, x_n)$ <u>and let M be an</u> n.n <u>matrix. Define the random variables in</u> $\eta = (y_1, \ldots, y_n)$ <u>by</u>

$$\eta = \xi M.$$

<u>Then</u>

(5.1) $$E(\eta) = E(\xi)M$$

(5.2) $$\text{cov}(\eta) = M\text{cov}(\xi)M'$$

<u>Lemma 1</u>: $\int_{-\infty}^{\infty} e^{-1/2\, cx^2}\, dx = \sqrt{\frac{2\pi}{c}}$

Proof: Let

$$\int_{-\infty}^{\infty} e^{-1/2\, y^2}\, dy = I$$

then

$$\int_{-\infty}^{\infty} e^{-1/2\, y^2}\, dy \int_{-\infty}^{\infty} e^{-1/2\, x^2}\, dx = I^2$$

Let $y = r \sin\theta$, $z = r \cos\theta$

then

$$0 \le r \le \infty$$

$$0 \le \theta \le 2\pi ,$$

the Jacobian of the transformation is r, and

$$I^2 = \int_0^{2\pi} \int_0^{\infty} r e^{-1/2\, r^2} \, dr \, d\theta$$

$$= 2\pi \int_0^{\infty} r e^{-1/2\, r^2} \, dr$$

$$= -2\pi e^{-1/2\, r^2} \Big|_0^{\infty}$$

$$= 2\pi$$

hence

(3) $$I = \int_{-\infty}^{\infty} e^{-1/2\, y^2} \, dy = \sqrt{2\pi} \quad .$$

Let $y = \sqrt{c}x$, then

$$\int_{-\infty}^{\infty} e^{-1/2\, cx^2} \, dx$$

$$= \frac{1}{\sqrt{c}} \int_{-\infty}^{\infty} e^{-1/2\, y^2} \, dy$$

$$= \sqrt{2\pi / c}$$

Lemma 2: $\int_{-\infty}^{\infty} e^{-1/2\, cy^2 + ay}\, dy = e^{a^2/2c} \sqrt{2\pi/c}$

Proof:
$$-c/2\, y^2 + ay$$

$$= -c/2 \left[y^2 - 2a/c\, y + a^2/c^2 - a^2/c^2 \right]$$

$$= -c/2 \left[(y - a/c)^2 - a^2/c^2 \right]$$

then

$$\int_{-\infty}^{\infty} e^{-1/2\, cy^2 + ay}\, dy$$

$$= \int_{-\infty}^{\infty} e^{-c/2\, (y - a/c)^2 + a^2/2c}\, dy$$

$$= e^{a^2/2c} \int_{-\infty}^{\infty} e^{-c/2\, (y - a/c)^2}\, dy$$

Let $x = y - a/c$, then, by lemma 1

$$e^{a^2/2c} \int_{-\infty}^{\infty} e^{-c/2\, (y - a/c)^2}\, dy$$

$$= e^{a^2/2c} \int_{-\infty}^{\infty} e^{-c/2\, x^2}\, dx$$

$$= e^{a^2/2c} \sqrt{2\pi/c}$$

3. The Multinomial Distribution.

The random variables $x_1, \ldots, x_k$ have the multinomial distribution if their density function has the form

$$f(x_1, \ldots, x_k) = \frac{n!}{x_1!x_2! \ldots x_k!} p_1^{x_1} \ldots p_k^{x_k}$$

where $\sum_{i=1}^{k} x_i = n$, and $\sum_{i=1}^{k} p_i = 1$. In order that $f(x_1, \ldots, x_k)$ be a density function for discrete variables, it is necessary that

$$\sum_{\substack{\Sigma x_i = n \\ 0 \leq x_i \leq n}} \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} \ldots p_k^{x_k} = 1$$

We see that this condition is satisfied since, by the multinomial theorem

(4) $$\sum_{\substack{\Sigma x_i = n \\ 0 \leq x_i \leq n}} \frac{n!}{x_1! \ldots x_k!} p_1^{x_1} \ldots p_k^{x_k} = (p_1 + \ldots + p_k)^n$$

and by assumption

$$\sum_{i=1}^{k} p_i = 1$$

We now develop the moment generating function $m(t_1, \ldots, t_k)$ for the multinomial distribution.

$$m(t_1, \ldots, t_k) = E\left(e^{\sum_{i=1}^{k} t_i x_i}\right)$$

$$= \sum_{\substack{\Sigma x_i = n \\ 0 \le x_i \le n}} e^{t_1 x_1} \cdots e^{t_k x_k} \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$

$$= \sum_{\substack{\Sigma x_i = n \\ 0 \le x_i \le n}} \frac{n!}{x_1! \cdots x_k!} (e^{t_1} p_1)^{x_1} \cdots (e^{t_k} p_k)^{x_k}$$

hence, by (4)

(5) $$m(t_1, \ldots, t_k) = (e^{t_1} p_1 + \cdots + e^{t_k} p_k)^n$$

4. <u>The Multivariate Normal Distribution.</u>

<u>The variables</u> $x_1, \ldots, x_n$ <u>have a joint normal distribution if their density function has the form</u>

$$f(x_1, \ldots, x_n) = c e^{-1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(x_i - u_i)(x_j - u_j)}$$

where the matrix $A = (a_{ij})$ is positive definite and symmetric and the constant c is such that

$$c\int_{R_n}\int\cdots\int e^{-1/2\,\xi A \xi'}\,d\xi = 1 \tag{6}$$

where

$$\xi = (x_1 - u_1), \ldots, (x_n - u_n) ,$$

and

$$f(x_1, \ldots, x_n) = ce^{-1/2\,\xi A \xi'}.$$

Now, by theorem 1, there exists an orthogonal matrix P such that

$$PAP' = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ & \cdots & \cdot & \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

then

$$|P|^2 = 1$$

(7)

$$|A| = d_1 d_2 \cdots d_n$$

We transform the integral in (6) according to the relation $\xi = \eta P$ where $\eta = (y_1, \ldots, y_n)$, the Jacobian of the transformation being $|P| = \pm 1$, so that its absolute value is 1. Thus, by equations (7) and lemma 1,

$$\int_{R_n}\int\cdots\int e^{-1/2\,\xi A \xi'}\,dx_1 \cdots dx_n$$

$$= \int_{R_n}\int\cdots\int e^{-1/2\,\eta PAP' \eta'}\,dy_1 \cdots dy_n$$

$$= \int_{R_n}\int \cdots \int e^{-1/2 \sum_{i=1}^{n} y_i^2 d_i} dy_1 \cdots dy_n$$

$$= \int_{-\infty}^{\infty} e^{-1/2\, y_1^2 d_1} dy_1 \cdots \int_{-\infty}^{\infty} e^{-1/2\, y_n^2 d_n} dy_n$$

$$= (2\pi)^{n/2}/|A|^{1/2} .$$

Hence

$$(8) \qquad |A|^{1/2}/(2\pi)^{n/2} \int_{R_n}\int \cdots \int e^{-1/2\, \xi A \xi'} d\xi = 1$$

and this determines the constant c.

We now develop the moment generating function $m(t_1, \ldots, t_n)$ for the multivariate normal distribution.

Suppose that $x_1, \ldots, x_n$ are random variables such that $E(x_i) = 0$, $i = 1, \ldots, n$. Let $\mathcal{T} = (t_1, \ldots, t_n)$, then

$$m(t_1, \ldots, t_n)$$

$$= c \int_{R_n}\int \cdots \int e^{t_1 x_1} \cdots e^{t_n x_n} e^{-1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j} dx_1 \cdots dx_n$$

$$= c \int_{R_n}\int \cdots \int e^{\mathcal{T} \xi'} e^{-1/2\, \xi A \xi'} d\xi .$$

Let $\eta = (y_1, \ldots, y_n)$ and P be such that

$$PAP' = D = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ & \ldots & & \\ 0 & 0 & \ldots & d_n \end{pmatrix}$$

Make the substitutions $\xi = \eta P$ and $\tau = \zeta P$ where $\zeta = (z_1, \ldots, z_n)$. Then, by equation (8), lemma 2 and theorem 1,

$$c \int_{R_n}\int \ldots \int e^{\tau\xi'} e^{-1/2\, \xi A \xi'}\, d\xi$$

$$= c \int_{R_n}\int \ldots \int e^{\zeta PP'\eta'} e^{-1/2\, \eta PAP' \eta'}\, d\eta$$

$$= c \int_{R_n}\int \ldots \int e^{\zeta\eta'} e^{-1/2\, \eta D \eta'}\, d\eta$$

$$= c \int_{R_n}\int \ldots \int e^{\sum_{i=1}^{n} z_i y_i} e^{-1/2 \sum_{i=1}^{n} d_i y_i^2}\, dy_1 \ldots dy_n$$

$$= c \int_{R_n}\int \ldots \int e^{\sum_{i=1}^{n} -1/2\, d_i y_i^2 + z_i y_i}\, dy_1 \ldots dy_n$$

$$= c\, (2\pi)^{n/2} / \sqrt{d_1} \ldots \sqrt{d_n}\; e^{1/2\, z_1^2/d_1} \ldots e^{1/2\, z_n^2/d_n}$$

$$= e^{1/2 \sum_{i=1}^{n} z_i^2/d_i}$$

$$= e^{1/2\, \zeta D^{-1} \zeta'}$$

$$= e^{1/2\, \tau P' P A^{-1} P' P \tau'}$$

$$= e^{1/2\, \tau A^{-1} \tau'}$$

Thus, if $E(\xi) = (0, \ldots, 0)$,

$$m(t_1, \ldots, t_n) = e^{1/2\, \tau A^{-1} \tau'} \tag{9}$$

Now suppose $E(x_i) = u_i$, $i = 1, \ldots, n$. Let $\xi = (x_1 - u_1), \ldots, (x_n - u_n)$, so that by (6)

$$m(t_1, \ldots, t_n)$$

$$= c \int_{R_n}\int \cdots \int e^{\sum_{i=1}^{n} t_i x_i} e^{-1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(x_i - u_i)(x_j - u_j)} dx_1 \ldots dx_n$$

$$= e^{\sum_{i=1}^{n} u_i t_i} c \int_{R_n}\int \cdots \int e^{\sum_{i=1}^{n} t_i (x_i - u_i)}$$

$$e^{-1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(x_i - u_i)(x_j - u_j)} dx_1 \ldots dx_n$$

$$= e^{\sum_{i=1}^{n} u_i t_i} c \int_{R_n}\int \cdots \int e^{\tau \xi'} e^{-1/2\, \xi A \xi'} \, d\xi$$

$$= e^{\sum_{i=1}^{n} u_i t_i} e^{1/2\, \tau A^{-1} \tau'}$$

Thus, if $E(\xi) = (u_1, \ldots, u_n)$

(10) $$m(t_1, \ldots, t_n) = e^{\sum_{i=1}^{n} u_i t_i} e^{1/2\, \tau A^{-1} \tau'}$$

We now show that $\mathrm{cov}(\xi) = A^{-1}$. From the properties of the moment generating function

$$\sigma_{hk} = \left(\partial^2 m/\partial t_h \partial t_k\right)_{t_i=0} - \left(\partial m/\partial t_h\right)\left(\partial m/\partial t_k\right)_{t_i=0}$$

In this case, letting $A^{-1} = (a^{ij})$, from equation (10)

$$(\partial m/\partial t_h)_{t_i=0}$$

$$= e^{\sum_{i=1}^{n} u_i t_i} e^{1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} a^{ij} t_i t_j} \left(u_h + \sum_{i=1}^{n} a^{ih} t_i\right)$$

$$= u_h$$

Furthermore

$$(\partial^2 m/\partial t_h \partial t_k)_{t_i=0}$$

$$= e^{\sum_{i=1}^{n} u_i t_i} \, e^{1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} a^{ij} t_i t_j} \left(u_h \left(u_k + \sum_{i=1}^{n} a^{ih} t_i\right) + \left(a^{hk} + \right.\right.$$

$$\left.\left. \sum_{i=1}^{n} a^{ih} t_i \sum_{i=1}^{n} a^{ik} t_i + u_k \sum_{i=1}^{n} a^{ik} t_i\right)\right)$$

$$= u_h u_k + a^{hk}$$

then

$$\sigma_{hk}$$

$$= u_h u_k + a^{hk} - u_h u_k$$

$$= a^{hk} .$$

Hence

(11) $$\mathrm{cov}(\xi) = (\sigma_{ij}) = (a^{ij}) = A^{-1}$$

5. The Chi-Square Distribution.

The chi-square distribution is well known and extensively discussed in the literature, however, we will be concerned with the moment generating function of the chi-square distribution rather than with the density function, hence we will define the chi-square distribution in terms of its moment generating function.

A random variable y has the chi-square distribution with n degrees of freedom if its moment generating function is

$$m(t) = (1-2t)^{-n/2}$$

We will demonstrate an example of a random variable with the chi-square distribution. Let $x_1, \ldots, x_n$ be independent standard normal variables, let

$$y = \sum_{i=1}^{n} x_i^2$$

then

$$m_y(t) = E(e^{ty})$$

$$= e^{t \sum_{i=1}^{n} x_i^2} \, 1/(2\pi)^{n/2} \int_{R_n}\int \cdots \int e^{-1/2 \sum_{i=1}^{n} x_i^2} dx_1 \ldots dx_n$$

$$= 1/\sqrt{2\pi}\int_{-\infty}^{\infty} e^{-1/2\,(1-2t)x_1} dx_1 \ldots 1/\sqrt{2\pi}\int_{-\infty}^{\infty} e^{-1/2\,(1-2t)x_n} dx_n$$

which product of integrals is, by lemma 2, equal to

$$(1-2t)^{-n/2}$$

and we summarize in the following lemma.

Lemma 3: Let $x_1, \ldots, x_n$ be independent standard normal variables, let $y = \sum_{i=1}^{n} x_i^2$, then the random variable y has the chi-square distribution with n degrees of freedom.

Theorem 6: If $x_1, \ldots, x_n$ are jointly normal,

$$f(x_1, \ldots, x_n) = |A|^{1/2}/(2\pi)^{n/2} e^{-1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(x_i - u_i)(x_j - u_j)}$$

then the quantity

$$y = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(x_i - u_i)(x_j - u_j)$$

has the chi-square distribution with n degrees of freedom.

Proof: Let $\xi = (x_1 - u_1), \ldots, (x_n - u_n)$, then

$$y = \xi A \xi'$$

and, by (11), $\operatorname{cov}(\xi) = A^{-1}$. Choose P orthogonal so that

$$P'AP = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

and let $\zeta = \xi P$ where $\zeta = (z_1, \ldots, z_n)$, then the z_i, as linear combinations of normal random variables, are normally distributed. Further, by theorem 5,

$$E(\zeta) = E(\xi)P = 0$$

and

$$\operatorname{cov}(\zeta) = PA^{-1}P' = \begin{pmatrix} 1/d_1 & 0 & \cdots & 0 \\ & \cdots & \cdots & \\ 0 & 0 & \cdots & 1/d_n \end{pmatrix}$$

Hence $z_1/d_1^{-1/2}, \ldots, z_n/d_n^{-1/2}$ are independent standard normal variables and by lemma 3

$$v = \sum_{i=1}^{n} z_i^2/d_i^{-1}$$

has the chi-square distribution with n degrees of freedom.

But

$$\sum_{i=1}^{n} z_i^2 d_i$$

$$= \zeta P'AP\zeta'$$

$$= \xi PP'APP'\xi'$$

$$= \xi A \xi'$$

hence

$$\xi A \xi' = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}(x_i - u_i)(x_j - u_j)$$

has the chi-square distribution with n degrees of freedom.

6. The Multinomial Distribution with Specified Parameters.

Let $x_1, \ldots, x_k$ have the multinomial distribution with parameters $p_1, \ldots, p_k$; $\sum_{i=1}^{k} x_i = n$, $\sum_{i=1}^{k} p_i = 1$. In applications involving the multinomial distribution, the simplest hypothesis that may be tested is that which specifies the probabilities $p_1, \ldots, p_k$. Accordingly, we seek a function of $x_1, \ldots, x_k$ and $p_1, \ldots, p_k$, whose distribution is one of the well known types. To this end, let $y_i = x_i - np_i/\sqrt{n}$.

We shall show that the joint distribution of $y_1, \ldots, y_k$ is normal. We first establish some lemmas.

Lemma 4: Let $f(n)$ be a function of the integer n, such that $\lim_{n \to \infty} nf(n) = 0$, and let a and b be independent of n. Then

$$\lim_{n \to \infty} \left[-\sqrt{n}a + n \log (1 + a/\sqrt{n} + b/n + f(n)) \right] = b - a^2/2$$

Proof: Let $z = (a/\sqrt{n} + b/n + f(n))$, then by the MacLaurin expansion of $\log (1 + z)$,

$$n \log (1 + a/\sqrt{n} + b/n + f(n))$$

$$= n\left[(a/\sqrt{n} + b/n + f(n)) - 1/2(a^2/n + b^2/n^2 + (f(n))^2 + 2ab/n^{3/2} + 2af(n)/\sqrt{n} + 2bf(n)/n + \ldots\right] .$$

Then

$$\lim_{n\to\infty}\left[-\sqrt{n}a + n\log(1 + a/\sqrt{n} + b/n + f(n))\right]$$

$$= -\sqrt{n}a + \sqrt{n}a + b - a^2/2 + O(nf(n))$$

$$= b - a^2/2$$

<u>Lemma 5</u>: <u>Let</u> C <u>be the covariance matrix of</u> $x_1, \ldots, x_{k-1}$ <u>where</u> $x_1, \ldots, x_{k-1}, x_k$ <u>are multinomially distributed with parameters</u> $p_1, \ldots, p_{k-1}, n$; <u>and let</u> $y_i = x_i - np_i/\sqrt{n}$, <u>and let</u> $\eta = (y_1, \ldots, y_{k-1})$. <u>Then</u>

$$(12)\quad C^{-1} = \begin{pmatrix} \frac{1}{p_1} & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \frac{1}{p_{k-1}} \end{pmatrix} + \begin{pmatrix} \frac{1}{p_k} & \frac{1}{p_k} & \cdots & \frac{1}{p_k} \\ \cdot & \cdot & \cdots & \cdot \\ \frac{1}{p_k} & \frac{1}{p_k} & & \frac{1}{p_k} \end{pmatrix}$$

$$= (\delta_{ij}/p_i + 1/p_k),$$

and

$$(13)\qquad \eta C^{-1}\eta' = \sum_{i=1}^{k} (x_i - np_i)^2/np_i .$$

We shall show that C times the matrix in equation (12) yields the identity matrix. We have

$$CC^{-1} = (\delta_{ij}p_i - p_ip_j)(\delta_{ij}/p_i - 1/p_k)$$

then

$$a_{ss} = -\left[\frac{p_sp_1}{p_k} + \cdots + \frac{(p_s^2 - p_s)(p_s + p_k)}{p_sp_k} + \frac{p_sp_{s+1}}{p_k} + \cdots + \frac{p_sp_{k-1}}{p_k}\right]$$

$$= \frac{p_s^2 \sum_{i=1}^{k-1} p_i + p_s^2p_k - p_s^2 - p_sp_k}{-p_sp_k}$$

$$= \frac{p_s^2(1 - p_k) + p_s^2p_k - p_s^2 - p_sp_k}{-p_sp_k}$$

$$= 1.$$

$$\underset{r \neq s}{a_{rs}}$$

$$= -\left[\frac{p_rp_1}{p_k} + \cdots + \frac{p_r^2 - p_r}{p_k} + \cdots + \frac{p_rp_s(p_s + p_k)}{p_sp_k} + \cdots + \frac{p_rp_{k-1}}{p_k}\right]$$

$$= \frac{p_rp_s(1 - p_k) - p_rp_s + p_rp_sp_k}{-p_sp_k}$$

$$= 0.$$

Hence

$$CC^{-1} = I.$$

To establish (13):

$$\sum_{i=1}^{k} \frac{(x_i - np_i)^2}{np_i}$$

$$= \sum_{i=1}^{k} \frac{(x_i^2 - 2nx_ip_i + n^2p_i^2)}{np_i}$$

$$= \sum_{i=1}^{k} \frac{x_i^2}{np_i} - 2\sum_{i=1}^{k} x_i + n\sum_{i=1}^{k} p_i$$

$$= \sum_{i=1}^{k} \frac{x_i^2}{np_i} - 2n + n$$

$$= \sum_{i=1}^{k} \frac{x_i^2}{np_i} - n$$

$$? C^{-1} = \left(\frac{x_1 - np_1}{\sqrt{n}} \cdots \frac{x_{k-1} - np_{k-1}}{\sqrt{n}} \right) \begin{pmatrix} \frac{p_1 + p_k}{p_1p_k} & \frac{1}{p_k} & \cdots & \frac{1}{p_k} \\ \cdot & \cdot & \cdots & \cdot \\ \frac{1}{p_k} & \frac{1}{p_k} & \cdots & \frac{p_{k-1}+p_k}{p_{k-1}p_k} \end{pmatrix}$$

$$a_{1s} = \sum_{\substack{i=1 \\ i \neq s}}^{k-1} \frac{x_i - np_i}{\sqrt{n}p_k} + \frac{x_s p_s + x_s p_k - np_s(p_s + p_k)}{\sqrt{n}p_s p_k}$$

$$= \frac{p_s(n - x_k) - np_s + x_s p_k}{\sqrt{n}p_s p_k}$$

$$= \frac{x_s p_k - x_k p_s}{\sqrt{n}p_s p_k}$$

Then

$$\eta C^{-1} \eta'$$

$$= \left(\frac{x_1 p_k - x_k p_1}{\sqrt{n}p_1 p_k} \cdots \frac{x_{k-1} p_k - x_k p_{k-1}}{\sqrt{n}p_{k-1} p_k} \right) \begin{pmatrix} \frac{x_1 - np_1}{\sqrt{n}} \\ \vdots \\ \frac{x_{k-1} - np_{k-1}}{\sqrt{n}} \end{pmatrix}$$

$$= \sum_{i=1}^{k-1} \frac{x_i^2 p_k - nx_i p_i p_k - x_k x_i p_i + nx_k p_i^2}{np_i p_k}$$

$$= \sum_{i=1}^{k-1} \frac{x_i^2}{np_i} - \sum_{i=1}^{k-1} x_i - \frac{x_k}{np_k} \sum_{i=1}^{k-1} x_i + \frac{x_k}{p_k} \sum_{i=1}^{k-1} p_i$$

$$= \sum_{i=1}^{k-1} \frac{x_i^2}{np_i} - \sum_{i=1}^{k-1} x_i - \frac{x_k}{np_k} (n - x_k) + \frac{x_k}{p_k} (1 - p_k)$$

$$= \sum_{i=1}^{k-1} \frac{x_i^2}{np_i} - (n - x_k) - \frac{x_k}{p_k} + \frac{x_k^2}{np_k} + \frac{x_k}{p_k} - x_k$$

$$= \sum_{i=1}^{k} \frac{x_i^2}{np_i} - n.$$

Hence

$$\eta C^{-1} \eta' = \sum_{i=1}^{k} \frac{x_i^2}{np_i} - n = \sum_{i=1}^{k} \frac{(x_i - np_i)^2}{np_i}$$

Theorem 7: Let $x_1, \ldots, x_k$ have a joint multinomial distribution with parameters $p_1, \ldots, p_k$; $\sum_{i=1}^{k} x_i = n$; $\sum_{i=1}^{k} p_i = 1$, and let $y_i = (x_i - np_i) / \sqrt{n}$. Then the random variables $y_1, \ldots, y_{k-1}$ are asymptotically jointly normally distributed and the quantity $\sum_{i=1}^{k} (x_i - np_i)^2/np_i$ is asymptotically

<u>chi-square</u> <u>distributed</u> <u>with</u> k-1 <u>degrees</u> <u>of</u> <u>freedom</u>.

Proof: We seek the moment generating function $M(t_1, \ldots, t_{k-1})$ of $y_1, \ldots, y_{k-1}$. We have

$$M(t_1, \ldots, t_{k-1})$$

$$= E\left[e^{y_1 t_1} \cdots e^{y_{k-1}t_{k-1}}\right]$$

$$= \sum_{\substack{\Sigma a_i = n \\ 0 \le a_i \le n}} e^{(a_1 - np_1/\sqrt{n})t_1} \cdots e^{(a_{k-1} - np_{k-1}/\sqrt{n})t_{k-1}} \frac{n!}{a_1! \cdots a_k!} p_1^{a_1} \cdots p_k^{a_k}$$

$$= \sum_{\substack{\Sigma a_i = n \\ 0 \le a_i \le n}} e^{-n \sum_{i=1}^{k-1} p_i t_i/\sqrt{n}} \frac{n!}{a_1! \cdots a_k!} \left(p_1 e^{t_1/\sqrt{n}}\right)^{a_1} \cdots \left(p_{k-1} e^{t_{k-1}/\sqrt{n}}\right)^{a_{k-1}} p_k^{a_k}$$

$$= e^{-\sqrt{n} \sum_{i=1}^{k-1} p_i t_i} \left(p_1 e^{t_1/\sqrt{n}} + \cdots + p_{k-1} e^{t_{k-1}/\sqrt{n}} + p_k\right)^n.$$

This last quantity, by use of the MacLaurin expansion for each $p_i e^{t_i/\sqrt{n}}$,

$$p_i e^{t_i/\sqrt{n}} = p_i \left(1 + t_i/\sqrt{n} + t_i^2/2!n + t_i^3/3!n^{3/2} + \cdots\right),$$

becomes

$$e^{-\sqrt{n}\sum_{i=1}^{k-1} p_i t_i}\Big[(p_1 + p_1 t_1/\sqrt{n} + p_1 t_1^2/2n + p_1 t_1^3/6n^{3/2} + \cdots)$$

$$+ (p_2 + p_2 t_2/\sqrt{n} + p_2 t_2^2/2n + \cdots) + \cdots$$

$$+ (p_{k-1} + p_{k-1} t_{k-1}/\sqrt{n} + p_{k-1} t^2_{k-1}/2n + \cdots) + p_k\Big]^n$$

$$= e^{-\sqrt{n}\sum_{i=1}^{k-1} p_i t_i}\left[1 + \sum_{i=1}^{k-1} p_i t_i/\sqrt{n} + \sum_{i=1}^{k-1} p_i t_i^2/2n + f(n)\right]^n$$

where $\lim_{n \to \infty} nf(n) = 0.$

Hence

$$\operatorname{Log} M(t_1, \ldots, t_{k-1})$$

$$= -\sqrt{n}\sum_{i=1}^{k-1} p_i t_i + n \log\left[1 + \sum_{i=1}^{k-1} p_i t_i/\sqrt{n} + \sum_{i=1}^{k-1} p_i t_i^2/2n + f(n)\right]$$

and by lemma 4,

$$\lim_{n \to \infty} \log M = \sum_{i=1}^{k-1} p_i t_i^2/2 - \Big(\sum_{i=1}^{k-1} p_i t_i\Big)^2/2 \;,$$

and

$$\lim_{n \to \infty} M(t_1, \ldots, t_{k-1})$$

$$= e^{1/2 \sum_{i=1}^{k-1} p_i t_i^2 - \sum_{i=1}^{k-1}\sum_{j=1}^{k-1} p_i p_j t_i t_j}$$

$$= e^{1/2 \sum_{i=1}^{k-1} \sum_{j=1}^{k-1} (\delta_{ij}p_i - p_ip_j)t_it_j}$$

$$= e^{1/2\, T C T'}$$

where $T = (t_1, \ldots, t_{k-1})$ and $C = (\delta_{ij}p_i - p_ip_j)$.

Hence, with reference to equations (9), we have

$$\lim_{n \to \infty} f(y_1, \ldots, y_{k-1}) = e^{-1/2\, \eta C^{-1} \eta'} .$$

Now according to theorem 6, $\eta C^{-1} \eta'$ has, in the limit, the chi-square distribution with k-1 degrees of freedom, and from lemma 5 we see that

$$\eta C^{-1} \eta' = \sum_{i=1}^{k} (x_i - np_i)^2/np_i .$$

This completes the proof of the theorem.

To illustrate the use of theorem 7, we consider the following examples:

Example 1: Suppose we wish to test the hypothesis that a die is true. A sample of n throws of the die is observed. Let x_i be the number of times the value i occurs in the n throws; $i = 1, 2, \ldots, 6$; $\sum_{i=1}^{6} x_i = n$. Let the probability of the occurrence of the value i be p_i, $\sum_{i=1}^{6} p_i = 1$, then the probability of obtaining x_i occurrences of value i in n throws of the die is

$$\frac{n!}{x_1! \cdots x_6!} p_1^{x_1} \cdots p_6^{x_6} .$$

Thus we see that $x_1, \ldots, x_6$ are jointly multinomially distributed with parameters $p_1, \ldots, p_6$, $\sum_{i=1}^{6} x_i = n$, $\sum_{i=1}^{6} p_i = 1$. The hypothesis that the die is true is the hypothesis that $p_i = 1/6$, $i = 1, \ldots, 6$. By theorem 7 our statistic is $\chi^2 = \sum_{i=1}^{6} (x_i - np_i)^2/np_i$, with 5 degrees of freedom.

Example 2: Suppose that each member of a certain genus of plants possesses one of k mutually exclusive characteristics, where it has been established that these characteristics occur in a fixed ratio. A botanist suspects that a species of plant under observation is a member of this genus. A sample of n members is taken of the species; x_i denotes the number of observations of characteristic i, p_i is the probability associated with characteristic i, $\sum_{i=1}^{k} x_i = n$, $\sum_{i=1}^{k} p_i = 1$. We recognize the distribution of $x_1, \ldots, x_k$ as the joint multinomial distribution with parameters $p_1, \ldots, p_k$. The hypothesis to be tested is that $p_i = \bar{p}_i$ where $\bar{p}_i$ is the proportion of the sample that would have characteristic i if the sample were from the known genus. By theorem 7, we use, as our statistic, $\chi^2 = \sum_{i=1}^{k} (x_i - np_i)^2/np_i$, with k-1 degrees of freedom.

Example 3: A classical example arises from one of Mendel's experiments. He observed, simultaneously, the shape and color of hybrid peas obtained by crossing two types of peas. Among n = 556 peas he observed

Round and yellow - - - - -	315
Round and green - - - - -	108
Angular and yellow - - - -	101
Angular and green - - - -	32

The Mendelian theory of inheritance states that the frequencies should be in the ratios 9 : 3 : 3 : 1. the hypothesis, then, is $p_i = \overline{p}_i$ where $\overline{p}_1 = 9/16$, $\overline{p}_2 = 3/16$, $\overline{p}_3 = 3/16$, $\overline{p}_4 = 1/16$, by theorem 7 we test this hypothesis with $\chi^2 = \sum_{i=1}^{4} (x_i - np_i)^2/np_i$ with 3 degrees of freedom.

7. The Multinomial Distribution with Unspecified Parameters.

We now consider another well known test concerning the multinomial distribution, the test for independence in contingency tables. A contingency table is a rectangular table formed in this fashion: we have characteristics $A_1, \ldots, A_r$ and $B_1, \ldots, B_s$, the ijth entry in the table is the number of items having characteristics A_i and B_j. For example:

Example 1: A survey is taken in which members of designated professions are questioned as to their opinions on a controversial political issue. A sample of size n = 1000 is obtained, the classification by profession and by opinion gives the contingency table:

	Favor	Oppose	Undecided	
Lawyers	207	142	43	392
Doctors	98	114	69	281
Educators	129	150	48	327
	434	406	160	1000

Example 2: A surgeon in a military base wishes to test the effectiveness of a cold preventive serum. One group of servicemen is treated with the serum; a control group is not treated. Both groups live under identical conditions for an observed period of time, at the end of the period the health records are checked and the tabular information for testing the effectiveness of the serum is:

	No colds	One cold	More than one cold	
Group A	x_{11}	x_{12}	x_{13}	$x_{1.}$
Group Control	x_{21}	x_{22}	x_{23}	$x_{2.}$
	$x_{.1}$	$x_{.2}$	$x_{.3}$	N

Example 3: Three instructors have taught the same course for several years. In an effort to determine whether or not the instructors give significantly different percentages of grades, the grade distributions as given by the instructors is

tabulated as follows:

	A	B	C	D	F	
Professor X	42	78	256	38	46	460
Professor Y	70	104	424	48	58	704
Professor Z	30	84	356	40	34	544
	142	266	1036	126	138	1708

Let x_{ij} be the number of elements in the ijth cell, and p_{ij} be the probability that an element will fall in the ijth cell, that is, that an element will have characteristics A_i and B_j where $\sum_{i=1}^{r} \sum_{j=1}^{s} p_{ij} = 1.$

Let $\sum_{j=1}^{s} p_{ij} = p_{i.}$ and $\sum_{i=1}^{r} p_{ij} = p_{.j}$, then $\sum_{i=1}^{r} p_{i.} = 1$ and $\sum_{j=1}^{s} p_{.j} = 1$.

The test for independence in the contingency table is

Hypothesis: $p_{ij} = p_{i.}p_{.j}$

We seek a random variable whose distribution is known if the hypothesis is true. The classical statistic for this test is that first proposed by Karl Pearson:

$$(14) \qquad v = \sum_{i,j} \frac{\left(x_{ij} - \frac{x_{i.}x_{.j}}{n}\right)^2}{\frac{x_{i.}x_{.j}}{n}}$$

where $x_{i.} = \sum_{j=1}^{s} x_{ij}$, $x_{.j} = \sum_{i=1}^{r} x_{ij}$

and $\sum_{i=1}^{r} x_{i.} = n$, $\sum_{j=1}^{s} x_{.j} = n$.

By an extension of the methods used in proving theorem 7, one can demonstrate that (14) has the chi-square distribution with $(r-1)(s-1)$ degrees of freedom. However, we shall make use of a theorem of Cramér[1] to establish our result. We first make some preparatory remarks.

Suppose, for simplicity, the parameters p_{ij} are re-designated $p_1, p_2, \ldots, p_r$. Let $p_1, p_2, \ldots, p_r$ be functions of the variables $\alpha_1, \ldots, \alpha_s$. Suppose the true values of the α's were known, then we could apply theorem 7

$$(15) \qquad \chi^2 = \sum_{i=1}^{r} \frac{\left(x_i - np_i(\alpha_1, \ldots, \alpha_s)\right)^2}{np_i(\alpha_1, \ldots, \alpha_s)}$$

Now suppose the values of the α_j are unknown and must be estimated from the sample, then the p_i are not constants as in theorem 7, but are functions of the sample values of the α_j. We wish to estimate the values of the α_j so that (15) will be a minimum; to this end we will let

$$\frac{\partial \chi^2}{\partial \alpha_j} = 0, \quad j = 1, \ldots, s.$$

[1] H. Cramér, Mathematical Methods of Statistics, Princeton University Press, (1946), pp. 426-434.

and solve the resulting equations for α_j. This process is called the chi-square minimum method of estimation. Thus, from (15) we have

$$\frac{\partial \chi^2}{\partial \alpha_j} = \sum_{i=1}^{r} \left[\frac{-2n(x_i - np_i)}{np_i} - \frac{(x_i - np_i)^2}{np_i^{\,2}} \right] \frac{\partial p_i}{\partial \alpha_j} = 0,$$

then

$$(16) \qquad -1/2 \frac{\partial \chi^2}{\partial \alpha_j} = \sum_{i=1}^{r} \left[\frac{x_i - np_i}{p_i} + \frac{(x_i - np_i)^2}{2np_i^{\,2}} \right] \frac{\partial p_i}{\partial \alpha_j} = 0.$$

We note that, in (16), the second member in brackets approaches zero as n approaches infinity, hence we will consider the effect of large n and modify (16) to

$$(17) \qquad \sum_{i=1}^{r} \left[\frac{x_i - np_i}{p_i} \right] \frac{\partial p_i}{\partial \alpha_j} = 0.$$

Further, if we impose the condition that

$$\sum_{i=1}^{r} p_i(\alpha_1, \ldots, \alpha_s) = c$$

where c is a constant, then

$$\sum_{i=1}^{r} \frac{\partial p_i}{\partial \alpha_j} = 0,$$

hence under this condition equations (17) reduce to

$$(18) \qquad \sum_{i=1}^{r} \frac{x_i}{p_i} \frac{\partial p_i}{\partial \alpha_j} = 0.$$

With these remarks we proceed to Cramer's theorem.

<u>Cramér's Theorem: Suppose that we are given</u> r <u>functions</u> $p_1(\alpha_1, \ldots, \alpha_s), \ldots, p_r(\alpha_1, \ldots, \alpha_s)$ <u>of</u> $s < r$ <u>variables</u> $\alpha_1, \ldots, \alpha_s$ <u>such that, for all points of a non-degenerate interval</u> A <u>in the</u> s<u>-dimensional space of the</u> α_j, <u>the</u> p_i <u>satisfy the following conditions:</u>

a) $\sum_{i=1}^{r} p_i(\alpha_1, \ldots, \alpha_s) = 1.$

b) $p_i(\alpha_1, \ldots, \alpha_s) > c^2 > 0$ for all i.

c) Every p_i has continuous derivatives $\frac{\partial p_i}{\partial \alpha_j}$ and $\frac{\partial^2 p_i}{\partial \alpha_j \partial \alpha_k}$.

d) The matrix $D = \left(\frac{\partial p_i}{\partial \alpha_j}\right)$, where $i = 1, \ldots, r$ and $j = 1, \ldots, s$, is of rank s.

<u>Let the possible results of a certain random experiment</u>

be divided into r mutually exclusive groups, and suppose that the probability of obtaining a result belonging to the ith group is $p_i^o = p_i(\alpha_1^o, \ldots, \alpha_s^o)$ where $\alpha_o = (\alpha_1^o, \ldots, \alpha_s^o)$ is an inner point of the interval A. Let x_i denote the number of results belonging to the ith group, which occur in a sequence of n repetitions of the experiment, so that $\sum_{i=1}^{r} x_i = n$.

The equations (17) of the modified chi-square minimum method then have exactly one system of solutions $\alpha = (\alpha_1, \ldots, \alpha_s)$ such that α converges in probability to α_o as n approaches infinity. The value of χ^2 obtained by inserting these values of the α_j into equations (15) is, in the limit as n approaches infinity, distributed in a chi-square distribution with $r - s - 1$ degrees of freedom.

We now show that Cramér's Theorem may be used to establish the asymptotic distribution of (14).

To review the situation, the variables x_{ij} have a multinomial distribution with parameters p_{ij}.

Let $\sum_{j=1}^{s} p_{ij} = p_{i.}$; $\sum_{i=1}^{r} p_{ij} = p_{.j}$. We are concerned with the hypothesis

$$H: \; p_{ij} = p_{i.}p_{.j}.$$

We apply Cramér's Theorem in the following manner:

Let $\alpha_1 = p_{1.}, \ldots, \alpha_r = p_{r.}, \alpha_{r+1} = p_{.1}, \ldots, \alpha_{r+s} = p_{.s}$.

Our hypothesis then takes the form

$$H: \quad p_{ij} = \alpha_i \beta_j \; .$$

where the p_{ij} are redesignated p_k, $k = 1, 2, \ldots, rs$ and

$\beta_j = \alpha_{r+j}, \quad j = 1, \ldots, s$

where

$$p_{ij} > 0 \text{ for all } ij.$$

Further, we have

$$(19) \quad \sum_{i=1}^{r} \sum_{j=1}^{s} p_{ij} = 1, \quad \sum_{j=1}^{s} p_{ij} = p_{i.} = \alpha_i, \quad \sum_{i=1}^{r} p_{ij} = p_{.j} = \beta_j \; .$$

$$(20) \qquad \sum_{i=1}^{r} \sum_{j=1}^{s} x_{ij} = n, \quad \sum_{j=1}^{s} x_{ij} = x_{i.}, \quad \sum_{i=1}^{r} x_{ij} = x_{.j} \; .$$

Then

$$\alpha_r = 1 - \sum_{i=1}^{r-1} \alpha_i, \quad \beta_s = 1 - \sum_{j=1}^{s-1} \beta_j \; .$$

Hence we have rs functions $p_1, p_2, \ldots, p_{rs}$ of $r - s - 2$ variables $\alpha_1, \ldots, \alpha_{r-1}, \beta_1, \ldots, \beta_{s-1}$.

Clearly, conditions a, b, c of Cramér's Theorem are satisfied.

We now obtain the solutions to equations (18). By

(19) and (20) equations (18) become

$$(21)\qquad \sum_{j=1}^{s}\left(\frac{x_{ij}}{p_{i.}}-\frac{x_{rj}}{p_{r.}}\right)=0 \qquad i=1,\ldots,r-1.$$

$$(22)\qquad \sum_{i=1}^{r}\left(\frac{x_{ij}}{p_{.j}}-\frac{x_{is}}{p_{.s}}\right)=0 \qquad j=1,\ldots,s-1.$$

Then (21) becomes, on summing over j,

$$(23)\qquad \frac{x_{i.}}{p_{i.}}-\frac{x_{r.}}{p_{r.}}=0$$

which, on summing over i, $i = 1, \ldots, r$, becomes

$$\frac{n}{1}-\frac{x_{r.}}{p_{r.}}=0$$

or

$$\frac{x_{r.}}{p_{r.}}=n$$

then, by (23), we have

$$p_{i.}=\frac{x_{i.}}{n}\ .$$

Likewise

$$p_{.j}=\frac{x_{.j}}{n}\ .$$

Hence by the modified chi-square minimum method

$$p_{ij} = \alpha_i\beta_j = p_{i.}p_{.j} = \frac{x_{i.}x_{.j}}{n^2}$$

and equation (15) becomes

$$\chi^2 = \sum_{i,j} \frac{\left(x_{ij} - \frac{x_{i.}x_{.j}}{n}\right)^2}{\frac{x_{i.}x_{.j}}{n}}$$

There remains only to show that the rank of D is $r - s - 2$, where

$$D = \left(\frac{\partial p_i}{\partial \alpha_j}\right)$$

$$= \begin{pmatrix}
\beta_1 & 0 & \cdots & \alpha_1 & 0 & \cdots & 0 \\
\beta_2 & 0 & \cdots & 0 & \alpha_1 & \cdots & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\beta_{s-1} & 0 & \cdots & 0 & 0 & \cdots & \alpha_1 \\
0 & \beta_1 & \cdots & \alpha_2 & 0 & \cdots & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & \beta_{s-1} & \cdots & 0 & 0 & \cdots & \alpha_2 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdots \beta_1 & \alpha_{r-1} & 0 & \cdots & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \cdots \beta_{s-1} & 0 & 0 & \cdots & \alpha_{r-1}
\end{pmatrix}$$

But according to Theorem 4, it will suffice to show that the rank of $D'D$ is $r - s - 2$. One readily obtains upon multiplication, $D'D$

$$
= \begin{pmatrix}
\sum_{j=1}^{s-1}\beta_j^2 & 0 & \cdots 0 & \beta_1\alpha_1 & \beta_2\alpha_1 & \cdots \beta_{s-1}\alpha_1 \\
0 & \sum_{j=1}^{s-1}\beta_j^2 & \cdots 0 & \beta_1\alpha_2 & \beta_2\alpha_2 & \cdots \beta_{s-1}\alpha_2 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
0 & 0 & \sum_{j=1}^{s-1}\beta_j^2 & \beta_1\alpha_{r-1} & \beta_2\alpha_{r-1} & \cdots \beta_{s-1}\alpha_{r-1} \\
\beta_1\alpha_1 & \beta_1\alpha_2 & \cdots \beta_1\alpha_{r-1} & \sum_{i=1}^{r-1}\alpha_i^2 & 0 & \cdots 0 \\
\beta_2\alpha_1 & \beta_2\alpha_2 & \cdots \beta_2\alpha_{r-1} & 0 & \sum_{i=1}^{r-1}\alpha_i^2 & \cdots 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\beta_{s-1}\alpha_1 & \beta_{s-1}\alpha_2 & \cdots \beta_{s-1}\alpha_{r-1} & 0 & 0 & \cdots \sum_{i=1}^{r-1}\alpha_i^2
\end{pmatrix}
$$

By dividing the first row of $D'D$ by α_1, the second row by α_2, . . ., the rth row by β_1, . . ., the $r - s - 2$nd row by β_{s-1}; and by dividing the first column by α_1, . . ., the $r - s - 2$nd column by β_{s-1}, we see that $D'D$ has the same rank as

$$
(24) \qquad M = \begin{pmatrix} A & B \\ B' & C \end{pmatrix}
$$

where A and C are $(r - 1) \times (r - 1)$ and $(s - 1) \times (s - 1)$ diagonal matrices of rank $r - 1$ and $s - 1$, respectively and B is an $(r - 1) \times (s - 1)$ matrix consisting entirely of 1's. From Laplace's development of the determinant of (24), it is clear that $|M| = |A| \cdot |C| \neq 0$ and hence that rank D'D is equal to the rank of D and is $r - s - 2$. We have thus proven our final result:

Theorem 9: Let $x_{11}, x_{12}, \ldots, x_{rs}$ *be a sample from a multinomial population with the mutually exclusive classes* A_iB_j, $i=1, \ldots, r$, $j = 1, \ldots, s$, *in which the probability associated with* A_iB_j *is* p_{ij}. *Let*

$$\chi_c^2 = \sum_{i,j} \frac{\left(x_{ij} - \frac{x_{i.}x_{.j}}{n}\right)^2}{\frac{x_{i.}x_{.j}}{n}} .$$

Then the limiting distribution of χ_c^2 *as* $n \to \infty$ *is the chi-square distribution with* $(r - 1)(s - 1)$ *degrees of freedom.*

BIBLIOGRAPHY

CRAMÉR, H. Mathematical Methods of Statistics. Princeton University Press, Princeton, New Jersey 1946.

MOOD, A. M. Introduction to the Theory of Statistics. McGraw-Hill Book Company, Inc., New York, New York 1950.

WILKS, S. S. Mathematical Statistics. Princeton University Press, Princeton, New Jersey 1944.