THE LIMITING DISTRIBUTION OF FUNCTIONS OF SAMPLE MEANS AND APPLICATION TO TESTING HYPOTHESES

P. L. HSU

NATIONAL UNIVERSITY OF PEKING

Introduction

In 1935 J. L. Doob published a paper [2][1] in which he derived the limiting distribution of a function of four sample means from one homogeneous sample. This work is susceptible to an easy generalization and supplies a powerful weapon with which to find the limiting distribution of a vast number of statistics. But since publication its importance seems to have been overlooked. A generalization of Doob's theorem to any number of sample means was given by the author [7].

In the first part of this paper two theorems are proved which embody a further generalization of Doob's result to the case of several samples of different sizes, and numerous examples are given to illustrate their wide applicability.

These examples are confined to the limiting distributions of given statistics, but in the second part a much more important constructive application is made. Two hypotheses of a general character, concerning one sample and several samples respectively, are formulated, and a systematic method of constructing a test function for each hypothesis included in the two general ones is given. The construction is done in such a manner that, as a consequence of the results obtained in the first part, (i) the test function has for its limiting distribution the χ^2 distribution with a known degree of freedom when the hypothesis tested is true, and (ii) the power of the test tends in general to unity as its limit. Special hypotheses and their large sample tests are treated as examples in the second part of the paper.

I

The limiting distribution of functions of sample means

1. *The mathematical model of* k *samples.*—Let there be given k random vectors of m components each,

$$\mathbf{u}_a = [U_{1a}, U_{2a}, \cdots, U_{ma}], \qquad a = 1, \cdots, k, \tag{1}$$

possessing finite second moments. Let

$$E(U_{ia}) = \mu_{ia}, \qquad E(U_{ia}U_{ja}) - \mu_{ia}\mu_{ja} = \eta_{ija}. \tag{2}$$

[1] Boldface numbers in brackets refer to references at the end of the paper (p. 402).

By a sample of size N_a of $\mathbf{u}_a$ is meant a system of N_a mutually independent random vectors,

$$\mathbf{u}_{ar} = [U_{1ar}, U_{2ar}, \cdots, U_{mar}], \qquad r = 1, \cdots, N_a, \tag{3}$$

each of which is distributed the same as U_a. Thus the k vectors (1) give rise to k samples, namely, the vectors (3) wherein a takes the values $1, \cdots, k$. The total number of such vectors is

$$N = N_1 + N_2 + \cdots + N_k.$$

We shall also assume that two vectors belonging to two different samples are always independent. Then about the distribution of the N vectors $\mathbf{u}_{ar}$ we know the following facts: (i) $\mathbf{u}_{ar}$ and $\mathbf{u}_{\beta s}$ are independent if either $a \neq \beta$ or $r \neq s$; (ii) for every fixed a the vectors $\mathbf{u}_{ar}$ $(r = 1, \cdots, N_a)$ are equi-distributed; (iii) each $\mathbf{u}_{ar}$ has finite moments of the first two orders given by (2).

2. *Sample mean and normalized sample mean.*—If U is any random variable and if $U_1, \cdots, U_n$ are a sample of size n, we shall term the quantities $\bar{U} = \frac{1}{n}(U_1 + \cdots + U_n)$ and $n^{\frac{1}{2}}\{\bar{U} - E(U)\}$ the sample mean and the normalized sample mean of U respectively. Thus the samples (3) give rise to the sample means

$$\bar{U}_{ia} = \frac{1}{N_a} \sum_{r=1}^{N_a} U_{iar}$$

and the normalized sample means

$$Z_{ia} = N^{\frac{1}{2}}_a (\bar{U}_{ia} - \mu_{ia}).$$

Hence

$$\bar{U}_{ia} = \mu_{ia} + N_a^{-\frac{1}{2}} Z_{ia}. \tag{4}$$

We recall here the well-known central limit theorem:[2]

As $N_a \to \infty$, *the distribution law of the vector* $[Z_{1a}, \cdots, Z_{ma}]$ *tends to the* m-*dimensional normal law with zero means and the dispersion matrix* $[\eta_{ija}]$.

3. *The statistic* T.—Consider a function of mk real variables,

$$f(x_{11}, \cdots, x_{m1}; \cdots; x_{1k}, \cdots, x_{mk}), \tag{5}$$

defined in the whole mk-dimensional space and possessing continuous derivatives of every kind of order two or three, as the case may be, in the neighborhood

$$| x_{ia} - \mu_{ia} | \leqq \delta, \qquad i = 1, \cdots, m; a = 1, \cdots, k. \tag{6}$$

[2] Cf. Cramér [1], Chap. 10, theorem 20-a.

Write

$$f_{i\alpha} = \frac{\partial f}{\partial x_{i\alpha}}, \qquad f_{ij\alpha\beta} = \frac{\partial^2 f}{\partial x_{i\alpha}\partial x_{j\beta}}, \qquad f_{ijh\alpha\beta\gamma} = \frac{\partial^3 f}{\partial x_{i\alpha}\partial x_{j\beta}\partial x_{h\gamma}},$$

$$a = f(\mu_{11}, \cdots, \mu_{m1}; \cdots; \mu_{1k}, \cdots, \mu_{mk}),$$

$$b_{i\alpha} = f_{i\alpha}(\mu_{11}, \cdots, \mu_{m1}; \cdots; \mu_{1k}, \cdots, \mu_{mk}),$$

$$c_{ij\alpha\beta} = f_{ij\alpha\beta}(\mu_{11}, \cdots, \mu_{m1}; \cdots; \mu_{1k}, \cdots, \mu_{mk}).$$

If in f each argument $x_{i\alpha}$ is replaced by $\bar{U}_{i\alpha}$, the result is a statistic,

$$T = f(\bar{U}_{11}, \cdots, \bar{U}_{m1}; \cdots; \bar{U}_{1k}, \cdots, \bar{U}_{mk}). \tag{7}$$

By (4) we have

$$T = f(\mu_{11} + N_1^{-\frac{1}{2}} Z_{11}, \cdots; \cdots; \cdots, \mu_{mk} + N_k^{-\frac{1}{2}} Z_{mk}). \tag{8}$$

The main purpose of the first part of this paper is to derive the limiting distribution of T when the sample sizes become infinite simultaneously. It is necessary to impose a restriction on the manner in which these sizes grow. We put

$$N_\alpha = N g_\alpha, \qquad \alpha = 1, \cdots, k; \qquad g_1 + \cdots + g_k = 1, \tag{9}$$

regard the g_α as fixed, and allow N to grow indefinitely. The method is based on the Taylor expansion of (8) in the neighborhood of

$$| N_\alpha^{-\frac{1}{2}} Z_{i\alpha} | \leqq \delta, \qquad i = 1, \cdots, m; \ \alpha = 1, \cdots, k. \tag{10}$$

If all the second derivatives exist and are continuous in (6), then in (10) we have

$$T = a + N^{-\frac{1}{2}} R + N^{-1} \sum_{i,j,\alpha,\beta} \varphi_{ij\alpha\beta} Z_{i\alpha} Z_{j\beta}, \tag{11}$$

where

$$R = \sum_{i,\alpha} g_\alpha^{-\frac{1}{2}} b_{i\alpha} Z_{i\alpha},$$

$$\varphi_{ij\alpha\beta} = \tfrac{1}{2}(g_\alpha g_\beta)^{-\frac{1}{2}} f_{ij\alpha\beta}(\mu_{11} + \theta N_1^{-\frac{1}{2}} Z_{11}, \cdots; \cdots; \cdots, \mu_{mk} + \theta N_k^{-\frac{1}{2}} Z_{mk}), \qquad |\theta| \leqq 1.$$

Again, if all the third derivatives exist and are continuous in (6), then in (10) we have

$$T = a + N^{-\frac{1}{2}} R + N^{-1} S + N^{-\frac{3}{2}} \sum_{i,j,h,\alpha,\beta,\gamma} \varphi_{ijh\alpha\beta\gamma} Z_{i\alpha} Z_{j\beta} Z_{h\gamma}, \tag{12}$$

where

$$S = \tfrac{1}{2} \sum_{i,j,\alpha,\beta} (g_\alpha g_\beta)^{-\frac{1}{2}} c_{ij\alpha\beta} Z_{i\alpha} Z_{j\beta},$$

$$\varphi_{ijh\alpha\beta\gamma} = \tfrac{1}{6} (g_\alpha g_\beta g_\gamma)^{-\frac{1}{2}} f_{ijh\alpha\beta\gamma} (\mu_{11} + \theta N_1^{-\frac{1}{2}} Z_{11}, \cdots ; \cdots ; \cdots , \mu_{mk} + \theta N_k^{-\frac{1}{2}} Z_{mk}), \qquad |\theta| \leqq 1.$$

4. *The limiting distributions of* R *and* S.—In view of the central limit theorem (sec. 2) and the independence of the vectors $[Z_{1\alpha}, \cdots, Z_{m\alpha}]$ for different values of α, we obtain immediately the following lemma:

Lemma 1. *As* $N \to \infty$, *the distribution law of* R *tends to the limit*

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x}{\sigma}} e^{-\frac{1}{2}y^2} dy, \tag{13}$$

where

$$\sigma^2 = \sum_{i,j,\alpha} g_\alpha^{-1} b_{i\alpha} b_{j\alpha} \eta_{ij\alpha}, \tag{14}$$

provided $\sigma^2 \neq 0$. *If* $\sigma^2 = 0$, *then* R $= 0$ *with unit probability.*

On the same ground we conclude:

Lemma 2. *As* $N \to \infty$, *the distribution law of* S *tends to a limit which is the distribution law of the quadratic form*

$$\sum_{i,j,\alpha,\beta} (g_\alpha g_\beta)^{-\frac{1}{2}} c_{ij\alpha\beta} W_{i\alpha} W_{j\beta}, \tag{15}$$

where the $W_{i\alpha}$ *are normal variates having zero means and the same second moments as the variables* $U_{i\alpha} - \mu_{i\alpha}$.

It turns out that the limiting distribution of S is the distribution of a certain quadratic form in normal variates. In most of the actual cases that we encounter this form is semi-definite. Hence we shall complete the solution of the limiting distribution of S by a lemma, given in the next section, about the distribution of semi-definite quadratic forms in normal variates.

5. *Distribution of semi-definite quadratic forms in normal variates.*—Suppose that a semi-definite form Q in normal variates with zero means is reduced in any manner to a sum of squares,[3]

$$Q = W_1^2 + W_2^2 + \cdots + W_q^2, \tag{16}$$

where the W's are themselves normal variates with zero means. Let

$$E(W_i W_j) = \omega_{ij}. \tag{17}$$

Let the dispersion matrix $[\omega_{ij}]$ be of rank $\rho > 0$ and let its non-vanishing latent roots, which are necessarily positive, be $\lambda_1, \cdots, \lambda_\rho$. Then it is always possible to apply an orthogonal transformation on $W_1, \cdots, W_q$ to get a new

[3] If Q is negative, we have only to give the right-hand side of (16) a minus sign.

set of normal variates with zero means, $W'_1, \cdots, W'_q$, such that $Q = W'^2_1 + \cdots + W'^2_q$ and

$$(18) \qquad E(W'_iW'_j) = 0, (i \neq j); \qquad E(W'^2_i) = \lambda_i, (i = 1, \cdots, \rho); \quad E(W'^2_i) = 0, (i = \rho + 1, \cdots, q).$$

The last equation of (18) implies that all the W'_i $(i > \rho)$ vanish with unit probability. Hence Q is essentially equal to $W'^2_1 + \cdots + W'^2_\rho$ and so its distribution law is

$$(19) \quad (2\pi)^{-\frac{1}{2}\rho}(\lambda_1 \cdots \lambda_\rho)^{-\frac{1}{2}} \int_{y_1^2 + \cdots + y_\rho^2 \leq x} \exp\left(-\frac{y_1^2}{2\lambda_1} - \cdots - \frac{y_\rho^2}{2\lambda_\rho}\right) dy_1 \cdots dy_\rho.$$

If, further, the relations

$$(20) \qquad \sum_{h=1}^{q} \omega_{ih}\,\omega_{jh} = \omega_{ij}, \qquad i,j = 1, \cdots, q,$$

are satisfied by the ω_{ij}, then $[\omega_{ij}]^2 = [\omega_{ij}]$ and so all the λ_i are unity. Then (19) reduces to the familiar χ^2 distribution with ρ degrees of freedom,

$$\left\{2^{\frac{1}{2}\rho}\,\Gamma(\tfrac{1}{2}\rho)\right\}^{-1} \int_0^x y^{\frac{1}{2}\rho - 1} e^{-\frac{1}{2}y} dy.$$

In this case it is also easy to find ρ. In fact, $\rho = \Sigma\lambda_i = \Sigma\omega_{ii}$.

We have therefore established the following lemma:

Lemma 3. *The distribution law of* Q *is* (19) *in general. If, in particular, the relations* (20) *are satisfied, then the distribution is the* χ^2 *distribution with* ρ *degrees of freedom, where* $\rho = \omega_{11} + \cdots + \omega_{qq}$.

6. *Limiting distribution of* T.—We shall use $\bar{E}$ to denote the negation of an event E, $(E_1; E_2)$ the conjunction of two events E_1 and E_2, and $P(E)$ the probability of E.

Theorem 1. *If the function* f *in* (5) *possesses continuous second derivatives of every kind in the neighborhood* (6), *then the limiting distribution of* $N^{\frac{1}{2}}(T-a)$ *is the same as the limiting distribution of* R. *Consequently this limit is the normal law* (13), *provided the quantity* σ^2 *in* (14) *does not vanish.*

Proof. We have seen that, when $Z_{i\alpha}$ satisfy the inequalities (10), T may be expressed as in (11), namely,

$$(21) \qquad T = a + N^{-\frac{1}{2}}R + N^{-1}R_1,$$

where $N^{-1}R_1$ denotes the last term in (11). Let us denote by E the event that all the inequalities (10) are true, and by $F(x)$ the distribution law of $N^{\frac{1}{2}}(T-a)$. Then

$$F(x) = P\{N^{\frac{1}{2}}(T-a) \leq x; E\} + P\{N^{\frac{1}{2}}(T-a) \leq x; \bar{E}\}.$$

But

$$(22)\quad P(E_1;\bar{E}) \leqq P(\bar{E}) \leqq \sum_{i,a} P(Z_{ia}^2 \geqq N_a\delta^2) \leqq \frac{1}{N\delta^2}\sum_{i,a} g_a^{-1}\eta_{iia} = o(1)$$

for every event E_1; hence

$$(23)\quad F(x) = P\{N^{\frac{1}{2}}(T-a) \leqq x; E\} + o(1).$$

Using (21), we have

$$N^{\frac{1}{2}}(T-a) = R + N^{-\frac{1}{2}}R_1 \text{ in conjunction with E.}$$

Besides, in the neighborhood (10) the functions $\varphi_{ija\beta}$ are continuous and therefore bounded. Let A be a common upper bound of the absolute values of all these functions. Then we have

$$|R_1| \leqq A\left(\sum_{i,a}\left|Z_{ia}\right|\right)^2 \text{ in conjunction with } E.$$

Hence

$$P\left\{R + N^{-\frac{1}{2}}A\left(\sum_{i,a}\left|Z_{ia}\right|\right)^2 \leqq x; E\right\} \leqq P\left\{N^{\frac{1}{2}}(T-a) \leqq x; E\right\} \leqq P\left\{R - N^{-\frac{1}{2}}A\left(\sum_{i,a}\left|Z_{ia}\right|\right)^2 \leqq X; E\right\}.$$

Using (22), we get

$$(24)\quad P\left\{R + N^{-\frac{1}{2}}A\left(\sum_{i,a}\left|Z_{ia}\right|\right)^2 \leqq x\right\} + o(1) \leqq F(x) \leqq P\left\{R - N^{-\frac{1}{2}}A\left(\sum_{i,a}\left|Z_{ia}\right|\right)^2 \leqq x\right\} + o(1).$$

Now it has been shown by Doob [**2**] that if X has a limiting distribution and if Y tends to zero in probability, then $X + Y$ has the same distribution as X. This theorem may be applied to the two extreme terms in (24), because evidently $N^{-\frac{1}{2}}(\Sigma|Z_{ia}|)^2$ tends to zero in probability. Hence both these terms are equal to $P(R \leqq x) + o(1)$ and consequently

$$F(x) = P(R \leqq x) + o(1), \quad \text{q.e.d.}$$

Theorem 2. *If the function* f *in* (5) *possesses continuous third derivatives of every kind in the neighborhood* (6), *and if quantity* σ^2 *in* (14) *vanishes, then the limiting distribution of* N(T–a) *is the same as the limiting distribution of* S. *Consequently this limit is the distribution law of the quadratic form* (15).

We shall merely sketch the proof, which is similar to that of theorem 1. Denoting the distribution law of $N(T-a)$ by $F_1(x)$ we have, as analogy of (23),

$$F_1(x) = P\{N(T-a) \leqq x; E\} + o(1).$$

In conjunction with E, T may be expressed as in (12), whereby the second term may be dropped, since now $\sigma^2 = 0$ and so R is essentially zero. Hence

$$N(T-a) = S + N^{-1/2}S_1,$$

where $N^{-1/2}\, S_1$ denotes the last term in (12). As before, we have

$$\left|S_1\right| \leqq B\left(\sum_{i,a}\left|Z_{ia}\right|\right)^3$$

in conjunction with E, where B is some constant. Then we obtain the analogy of (24),

$$P\left\{S + N^{-1/2}B\left(\sum_{i,a}\left|Z_{ia}\right|\right)^3 \leqq x\right\} + o(1) \leqq F_1(x)$$
$$\leqq P\left\{S - N^{-1/2}B\left(\sum_{i,a}\left|Z_{ia}\right|\right)^3 \leqq x\right\} + o(1),$$

which leads as before to the result

$$F_1(x) = P(S \leqq x) + o(1), \quad \text{q.e.d.}$$

With the help of lemma 3 the limiting distribution of T is completely solved, provided the quadratic form S is semi-definite.

Let us summarize the results contained in theorems 1 and 2: In order to obtain the limiting distribution of T, which is a function of the sample means $\bar{U}_{ia}$, make the substitution (4) and compute the Taylor expansion in powers of $N^{-1/2}$ to three terms,

$$a + N^{-1/2}R + N^{-1}S. \tag{25}$$

If the quantity σ^2 in (14) does not vanish, the limiting distribution of $N^{1/2}(T-a)$ is the normal distribution with mean zero and variance σ^2. If $\sigma^2 = 0$, then $N(T-a)$ has the same limiting distribution as that of S, and this latter is the distribution of a certain quadratic form in normal variates. If the form in question is semi-definite, the explicit formula of its distribution law is given in lemma 3.

In what follows, when we are dealing with cases of a single sample ($k = 1$), we shall drop the index a from all the letters.

7. *Probabilities of events.*—Consider a set of events, $E_1, \cdots, E_m$, forming a complete disjunction and having the probabilities $p_1, \cdots, p_m$. Let X_i be

the random variable such that $X_i = 1$ or zero according as E_i happens or does not happen. Then we have a random vector $[X_1, \cdots, X_m]$ with

$$(26) \qquad E(X_i) = p_i, \qquad E(X_i^2) = p_i, \qquad E(X_i X_j) = 0, \qquad i \neq j.$$

A sample of size N corresponds to N trials of experiment, and the sample means $\overline{X}_1, \cdots, \overline{X}_m$ are the relative frequencies $n_1/N, \cdots, n_m/N$, where n_i denotes the number of happenings of E_i in N trials. The quantity σ^2 in (14) has a simple expression. We have, by (26),

$$(27) \qquad \sigma^2 = \sum_i b_i^2 p_i(1-p_i) - \sum_{i \neq j} b_i b_j p_i p_j = \sum_i p_i b_i^2 - \left(\sum_i p_i b_i\right)^2 = \sum_i p_i(b_i - \sum_i p_i b_i)^2.$$

8. *Example 1: The* χ^2 *statistic.*—This classical statistic is defined as

$$T_1 = \sum_{i \neq 1}^{m} \frac{(n_i - p_i^0 N)^2}{p_i^0 N}$$

and is used to test the hypothesis that $p = p_i^0$, $(i = 1, \cdots, m)$. As explained in section 7, we have $n_i = N\overline{X}_i$. Hence

$$\frac{T_1}{N} = \sum_{i=1}^{m} \frac{(\overline{X}_i - p_i^0)}{p_i^0}.$$

The expansion (25) of T_1/N is

$$a + N^{-\frac{1}{2}} \sum_{i=1}^{m} b_i Z_i + N^{-1} \sum_{i=1}^{m} \frac{Z_i^2}{p_i^0},$$

where

$$a = \sum_{i=1}^{m} \frac{(p_i - p_i^0)^2}{p_i^0}, \qquad b_i = \frac{p_i - p_i^0}{p_i^0}.$$

If the hypothesis is false, $p_i \neq p_i^0$ for some i. Then, by (27), the quantity σ^2 in (14) takes the value

$$\sigma_1^2 = \sum_{i=1}^{m} p_i \left(b_i - \sum_{i=1}^{m} p_i b_i\right)^2 \neq 0.$$

For, if $\sigma_1^2 = 0$, b_i would be independent of i and so $p_i = \lambda p_i^0$ for all i. Since $\Sigma p_i = 1 = \Sigma p_i^0$, we would have $p_i = p_i^0$ for all i, contrary to our assumption. Hence the limiting distribution of $N^{\frac{1}{2}}(N^{-1}T_1 - a)$ is the normal distribution with mean zero and variance σ_1^2.

If the hypothesis is true, then a and all the b_i vanish. Hence the limiting distribution of T_1 is the same as that of

$$\sum_{i=1}^{m} \left(\frac{Z_i}{\sqrt{p_i^0}}\right)^2.$$

This limit is the distribution law of $W_1^2 + \cdots + W_m^2$, where $[W_1, \cdots, W_m]$ is a normal vector having zero means and the same dispersion matrix as the vector

$$\left[\frac{X_1 - p_1^0}{\sqrt{p_1^0}}, \cdots, \frac{X_m - p_m^0}{\sqrt{p_m^0}}\right].$$

Hence

$$\omega_{ij} = E(W_iW_j) = 1 - p_i^0, (i = j), = -\sqrt{p_i^0 p_j^0}, (i \neq j).$$

It is easy to verify that the relations (20) are satisfied, and that $\Sigma\omega_{ii} = m - 1$. Hence the limiting distribution of T_1 is the χ^2 distribution with m-1 degrees of freedom.

9. *Example 2: The mean square contingency.*—Let $E_1, \cdots, E_s$ and $E'_1, \cdots, E'_t$ be two sets of events, each forming a complete disjunction. Then the st events $E_{ij} = (E_i; E'_j)$ form a complete disjunction. Let

$$P(E_{ij}) = p_{ij}, \qquad P(E_i) = \sum_j p_{ij} = p_i, \qquad P(E'_j) = \sum_i p_{ij} = p'_j.$$

Let n_{ij} be the number of occurrences of E_{ij} in N trials, and let

$$n_i = \sum_j n_{ij}, \qquad n'_j = \sum_i n_{ij}.$$

The mean square contingency is defined as

$$T_2 = N \sum_{i=1}^{s} \sum_{j=1}^{t} \frac{\left(n_{ij} - \frac{n_i n'_j}{N}\right)^2}{n_i n'_j}.$$

It is used to test the hypothesis of complete independence of the two sets of events, that is, that $p_{ij} = p_i p'_j$ for all i and j.

We define st random variables X_{ij} such that $X_{ij} = 1$ or zero according as E_{ij} happens or does not happen. A sample of size N gives the sample means

$$\overline{X}_{ij} = \frac{n_{ij}}{N}.$$

Let also

$$\overline{X}_i = \sum_j \overline{X}_{ij} = \frac{n_i}{N}, \qquad \overline{X}'_j = \sum_i \overline{X}_{ij} = \frac{n'_j}{N}.$$

Then

$$\frac{T_2}{N} = \sum_{i,j} \frac{(\overline{X}_{ij} - \overline{X}_i\overline{X}'_j)^2}{\overline{X}_i\overline{X}'_j}.$$

Upon substituting $p_{ij} + N^{-\frac{1}{2}}Z_{ij}$ for $\overline{X}_{ij}$ we obtain the three-term expansion of T_2/N,

$$a + N^{-\frac{1}{2}}R + N^{-1}S,$$

where

$$a = \sum_{i,j} \frac{(p_{ij} - p_i p'_j)^2}{p_i p'_j},$$

$$R = \sum_{i,j} \frac{p_{ij} - p_i p'_j}{p_i^2 p'^2_j} \left\{ 2p_i p'_j Z_{ij} - (p_{ij} + p_i p'_j)(p_i Z'_j + p'_j Z_i) \right\},$$

$$S = \sum_{i,j} \left\{ \frac{(Z_{ij} - p_i Z'_j - p'_j Z_i)^2}{p_i p'_j} + (p_{ij} - p_i p'_j) Q_{ij} \right\},$$

$$Z_i = \sum_j Z_{ij}, \qquad Z'_j = \sum_i Z_{ij},$$

and the Q_{ij} are certain quadratic forms in the Z_{ij}.

We have

$$R = \sum_{i,j} b_{ij} Z_{ij},$$

where

$$b_{ij} = \frac{2p_{ij}}{p_i p_j} - \frac{1}{p_i^2} \sum_{\nu=1}^{t} \frac{p^2_{i\nu}}{p'_\nu} - \frac{1}{p'^2_j} \sum_{\mu=1}^{s} \frac{p^2_{\mu j}}{p_\mu}.$$

According to (27) the quantity σ^2 in (14) has the value

$$\sigma_2^2 = \sum_{i,j} p_{ij} \left(b_{ij} - \sum_{i,j} p_{ij} b_{ij} \right)^2.$$

But

$$\sum_{i,j} p_{ij} b_{ij} = 2 \sum_{i,j} \frac{p^2_{ij}}{p_i p'_j} - \sum_{i,\nu} \frac{p^2_{i\nu}}{p_i p'_\nu} - \sum_{\mu,j} \frac{p^2_{\mu j}}{p_\mu p'_j} = 0.$$

Hence

$$\sigma_2^2 = \sum_{i,j} p_{ij} b^2_{ij}.$$

If the hypothesis is false, $p_{ij} \neq p_i p'_j$ for some (i, j). Then $\sigma_2^2 \neq 0$. For σ_2^2 can vanish only when all the $b_{ij} = 0$, and this implies that

$$0 = \sum_{i,j} p_i p'_j b_{ij} = 2 - \sum_{i,\nu} \frac{p^2_{i\nu}}{p_i p'_\nu} - \sum_{\mu,j} \frac{p^2_{\mu j}}{p_\mu p'_j} = 2 - 2 \sum_{i,j} \frac{p^2_{ij}}{p_i p'_j} = -2 \sum_{i,j} \frac{(p_{ij} - p_i p'_j)^2}{p_i p'_j},$$

that is, $p_{ij} = p_i p'_j$ for all (i,j). Therefore in this case the limiting distribution of $N^{\frac{1}{2}}(N^{-1}T - a)$ is the normal distribution about zero with variance σ_2^2.

If the hypothesis is true, $p_{ij} = p_i p'_j$ for all (i, j). Then a and all the b_{ij} vanish. Hence the limiting distribution of T_2 is the distribution of

$$\sum_{i,j} W_{ij}^2,$$

where the W_{ij} are normal variates with zero means and the same second moments as the system

$$\frac{X_{ij}-p_{ij}-p_i(X'_j-p'_j)-p'_j(X_i-p_i)}{\sqrt{p_ip'_j}} \qquad X_i = \sum_j X_{ij}, \qquad X'_j = \sum_i X_{ij} .$$

Direct computation gives the values of these moments.:

$$\omega_{ij\mu\nu} = E(W_{ij}W_{\mu\nu}) = \begin{cases} (1-p_i)(1-p'_j) & \text{if } i=\mu, j=\nu, \\ -(1-p_i)\sqrt{p'_jp'_\nu} & \text{if } i=\mu, j\neq\nu, \\ -(1-p'_j)\sqrt{p_ip_\mu} & \text{if } i\neq\mu, j=\nu, \\ \sqrt{p_ip_\mu p'_jp'_\nu} & \text{if } i\neq\mu, j\neq\nu. \end{cases}$$

It may easily be verified that relations (20) are satisfied:

$$\sum_{g,h} \omega_{ijgh}\omega_{\mu\nu gh} = \omega_{ij\mu\nu}, \qquad i, \mu = 1, \cdots, s; \qquad j, \nu = 1, \cdots, t,$$

and that $\Sigma\omega_{ijij} = (s-1)(t-1)$. Hence the limiting distribution of T_2 is the χ^2 distribution with $(s-1)(t-1)$ degrees of freedom.

10. *Example 3*: *"Student's" t-statistic.*—Let X be a random variable having

$$E(X) = \xi, \qquad E\{(X-\xi)^2\} = 1, \quad E\{(X-\xi)^3\} = \alpha_3, \quad E\{(X-\xi)^4\} = \alpha_4 < \infty.$$

Let $X_1, \cdots, X_N$ be a sample. Then "Student's" is defined, except a factor depending on N, as

$$T_3 = \frac{\bar{X}}{V^{\frac{1}{2}}},$$

where $\bar{X}$ and V are the mean and the variance of the sample.

Consider the random variables

$$U_1 = X, \qquad U_2 = X^2.$$

They have the means

$$E(U_1) = \xi, \qquad E(U_2) = 1 + \xi^2,$$

and the sample means

$$\bar{U}_1 = \bar{X}, \qquad \bar{U}_2 = \frac{1}{N}\sum_r X'^2_r = V + \bar{U}_1^2.$$

Hence

$$T_3 = \bar{U}_1(\bar{U}_2 - \bar{U}_1^2)^{-\frac{1}{2}}.$$

The two terms of the Taylor expansion are

$$\xi + N^{-\frac{1}{2}}\{(1+\xi^2)Z_1 - \tfrac{1}{2}\xi Z_2\}.$$

The quantity σ^2 in (14) is the expectation of the square of

$$(1+\xi^2)(X-\xi) - \tfrac{1}{2}\xi(X^2-1-\xi^2) \tag{28}$$

and has the value

$$\sigma_3^2 = \frac{1}{4}(\alpha_4 - 1)\xi^2 - \alpha_3\xi + 1.$$

If (28) does not vanish with unit probability, then $\sigma_3^2 \neq 0$ and the limiting distribution of $N^{\frac{1}{2}}(T_3 - \xi)$ is the normal distribution about zero with the variance σ_3^2.

If (28) is essentially zero, then $\sigma_3^2 = 0$ and therefore $\xi \neq 0$. The random variable X can take precisely two values, namely,

$$a = \frac{1+\xi^2+\sqrt{1+\xi^2}}{\xi} \quad \text{and} \quad b = \frac{1+\xi^2-\sqrt{1+\xi^2}}{\xi}.$$

Let

$$P(X=a) = p, \qquad P(X=b) = 1-p.$$

Then we must have

$$\xi = pa + (1-p)b = \frac{1+\xi^2+(2p-1)\sqrt{1+\xi^2}}{\xi},$$

whence

$$p = \frac{\sqrt{1+\xi^2}-1}{2\sqrt{1+\xi^2}}.$$

Among the N numbers $X_1, \cdots, X_N$, let n have the value a and $N-n$ have the value b.

Then

$$T_3 = \frac{b + (a-b)\dfrac{n}{N}}{|a-b|\left(\dfrac{n}{N} - \dfrac{n^2}{N^2}\right)^{\frac{1}{2}}}.$$

As explained in section 7, n/N is the sample mean of a random variable which takes the value one with probability p and zero with probability $1 - p$. On substituting $p + N^{-\frac{1}{2}}Z$ for n/N, we obtain the expansion

$$\xi + \frac{2(1+\xi^2)^2}{N\xi^3} Z^2.$$

Hence the limiting distribution of $N(T_3 - \xi)$ is the same as that of $2\xi^{-3}(1+\xi^2)^2Z^2$. But limiting distribution of Z is the normal distribution with mean zero and the variance $p(1-p) = \frac{1}{4}\xi^2(1+\xi^2)^{-1}$. Hence the limiting distribution of $2N\xi(1+\xi^2)^{-1}(T_3 - \xi)$ is the χ^2 distribution with one degree of freedom.

11. *Example 4: The ratio of moments.*—Let X be a random variable having

$E(X) = 0, \quad E(X^2) = 1, \quad E(X^i) = \alpha_i, \quad \alpha_{2m} < \infty$ from some integer $m \geqq 3$, and $X_1, \cdots, X_N$ be a sample. Consider the statistic

$$T_4 = \frac{S_m}{S_2^{\frac{m}{2}}},$$

where

$$S_i = \frac{1}{N}\sum_{r=1}^{N}(X_r - \bar{X})^i.$$

When $m=3$ and $m=4$, T_4 becomes the familiar b_1 and b_2 of K. Pearson.

The random variables

$$U_i = X^i, \qquad i = 1, \cdots, m,$$

have the means α_i and the sample means

$$\bar{U}_i = \frac{1}{N}\sum_{r=1}^{N} X_r^{\,i}.$$

We have

$$T_4 = \frac{\bar{U}_m - m\,\bar{U}_1\bar{U}_{m-1} + \cdots}{(\bar{U}_2 - \bar{U}_1^2)^{\frac{1}{2}m}}.$$

Making the substitution

$$\bar{U}_1 = N^{-\frac{1}{2}}Z_1, \qquad \bar{U}_2 = 1 + N^{-\frac{1}{2}}Z_2, \qquad \bar{U}_i = \alpha_i + N^{-\frac{1}{2}}Z_i, \qquad i = 3, \cdots, m,$$

and computing the two-term expansion we obtain

$$\alpha_m + N^{-\frac{1}{2}}(Z_m - \tfrac{1}{2}m\,\alpha_m Z_2 - m\,\alpha_{m-1}Z_1).$$

The quantity σ^2 in (14) is the expectation of the square of

$$X^m - \alpha_m - \tfrac{1}{2}m\,\alpha_m(X^2 - 1) - m\,\alpha_{m-1}X, \tag{29}$$

and has the value

$$\sigma_4^2 = a_{2m} - m a_m a_{m+2} - 2m a_{m-1} a_{m+1} - \frac{1}{4}(m-2)^2 a_m^2 + \frac{1}{4} m^2 a_4 a_m^2 + m^2 a_3 a_{m-1} a_m + m^2 a^2_{m-1}.$$

Hence, if (29) is not essentially zero, the limiting distribution of $N^{\frac{1}{2}}(T_4 - a_m)$ is the normal distribution about zero with the variance σ_4^2.

12. *Functions of variances and covariances; a simplification.*—We are going to study a pair of statistics, denoted by T_5 and T_6, which are formed of one homogeneous multivariate sample and are functions of the variances and covariances.

Let

$$(30) \qquad [X_1, \cdot \cdot \cdot, X_p]$$

be a random vector having finite fourth moments and not satisfying any linear or quadratic relation with unit probability. Let

$$E(X_i) = \xi_i, \qquad E(X_iX_j) - \xi_i\xi_j = \sigma_{ij}.$$

Let

$$[X_{1r}, \cdot \cdot \cdot, X_{pr}], r = 1, \cdot \cdot \cdot, N,$$

be a sample of size N and let

$$\bar{X}_i = \frac{1}{N}\sum_{r=1}^{N} X_{ir}, \qquad v_{ij} = \frac{1}{N}\sum_{r=1}^{N} X_{ir}X_{jr} - \bar{X}_i\bar{X}_j.$$

Let T be any statistic which is a function of the v_{ij} only:

$$(31) \qquad T = F(v_{11}, v_{12}, \cdot \cdot \cdot, v_{p-1,p}, v_{pp}).$$

The $\frac{1}{2}p(p+1)$ random variables

$$U_i = X_i - \xi_i, \qquad U_{ij} = (X_i - \xi_i)(X_j - \xi_j), \qquad i \leqq j,$$

have the means

$$E(U_i) = 0, \qquad E(U_{ij}) = \sigma_{ij},$$

and the sample means

$$\bar{U} = \bar{X}_i - \xi_i, \qquad \bar{U}_{ij} = v_{ij} + \bar{U}_i\bar{U}_j.$$

Hence

$$T = F(\bar{U}_{11} - \bar{U}_1^2, \bar{U}_{12} - \bar{U}_1\bar{U}_2, \cdot \cdot \cdot, \bar{U}_{pp} - \bar{U}_p^2).$$

On substituting $N^{-\frac{1}{2}}Z_i$ for $\bar{U}_i$ and $\sigma_{ij}+N^{-\frac{1}{2}}Z_{ij}$ for $\bar{U}_{ij}$ we obtain the three-term expansion of T,

$$(32) \qquad A + N^{-\frac{1}{2}} \sum_{i \leqq j} B_{ij} Z_{ij} + N^{-1}\left(\tfrac{1}{2} \sum_{\substack{i \leqq j \\ \mu \leqq \nu}} C_{ij\mu\nu} Z_{ij} Z_{\mu\nu} - \sum_{i \leqq j} B_{ij} Z_i Z_j \right),$$

where

$$A = F(\sigma_{11}, \sigma_{12}, \cdots, \sigma_{pp}),$$

$$B_{ij} = \frac{\partial}{\partial \sigma_{ij}} F(\sigma_{11}, \sigma_{12}, \cdots, \sigma_{pp}),$$

$$C_{ij\mu\nu} = \frac{\partial^2}{\partial \sigma_{ij} \partial \sigma_{\mu\nu}} F(\sigma_{11}, \sigma_{12}, \cdots, \sigma_{pp}).$$

If $B_{ij} \neq 0$ for some (i, j) then the term

$$\sum_{i \leqq j} B_{ij} Z_{ij},$$

being the normalized sample mean of

$$\sum_{i \leqq j} B_{ij} \left\{ (X_i - \xi_i)(X_j - \xi_j) - \sigma_{ij} \right\},$$

cannot vanish with unit probability. Therefore it is sufficient to have the two-term expansion,

$$(33) \qquad A + N^{-\frac{1}{2}} \sum_{i \leqq j} B_{ij} Z_{ij}.$$

If $B_{ij} = 0$ for all (i, j), then (32) becomes

$$(34) \qquad A + \frac{1}{2N} \sum_{\substack{i \leqq j \\ \mu \leqq \nu}} C_{ij\mu\nu} Z_{ij} Z_{\mu\nu}.$$

But (33) and (34) are precisely the expansions that we shall obtain if we make the direct substitution $v_{ij} = \sigma_{ij} + N^{-\frac{1}{2}} Z_{ij}$ in F. Hence we have the following rule of simplification:

In order to obtain the limiting distribution of (31), *make the substitution* $v_{ij} = \sigma_{ij} + N^{-\frac{1}{2}} Z_{ij}$ *and then follow the steps described at the end of section* 6.

This rule of simplification can be extended immediately to the case of k samples.

13. *Example 5: The hypothesis of independence and Wilks's test function.*— Consider again the random vector (30). The hypothesis of independence is the following: $X_1, \cdots, X_p$ are classified into κ mutually independent sets consisting of $s_1, \cdots, s_\kappa$ members respectively:

$$(35) \quad [X_1, \cdots, X_{s_1}], [X_{s_1+1}, \cdots, X_{s_1+s_2}], \cdots, [X_{s_1+\cdots+s_{\kappa-1}+1}, \cdots, X_p]$$

This hypothesis was first studied by Wilks [**10**], who applied the principle of likelihood on the assumption of normality of (30) and obtained the test function which we now define.

Let the matrices

$$V = [v_{ij}], \qquad M = [\sigma_{ij}]$$

be so partitioned that

$$V = \begin{bmatrix} V_{11} V_{12} \cdots V_{1\kappa} \\ V_{21} V_{22} \cdots V_{2\kappa} \\ \text{-----------} \\ V_{\kappa 1} V_{\kappa 2} \cdots V_{\kappa\kappa} \end{bmatrix}, \qquad M = \begin{bmatrix} M_{11} M_{12} \cdots M_{1\kappa} \\ M_{21} M_{22} \cdots M_{2\kappa} \\ \text{-----------} \\ M_{\kappa 1} M_{\kappa 2} \cdots M_{\kappa\kappa} \end{bmatrix}$$

where $V_{\mu\nu}$ and $M_{\mu\nu}$ have s_μ rows and s_ν columns, $(\mu, \nu = 1, \cdots, \kappa)$. Let also

$$V_1 = \begin{bmatrix} V_{11} & & O \\ & \ddots & \\ O & & V_{\kappa\kappa} \end{bmatrix}, \qquad M_1 = \begin{bmatrix} M_{11} & & O \\ & \ddots & \\ O & & M_{\kappa\kappa} \end{bmatrix}.$$

Then Wilks's test function is

$$T_5 = \frac{|V|}{|V_1|}.$$

Let us now study the limiting distribution of T_5. Suppose first that not only is the hypothesis of independence false but actually some of the covariances σ_{ij} lying in the matrices $M_{\mu\nu}$, $(\mu \neq \nu)$, are not zero. Following the rule of simplification in section 12 we make the substitution $v_{ij} = \sigma_{ij} + N^{-\frac{1}{2}} Z_{ij}$ in T_5 and obtain the two-term expansion,

$$a + aN^{-\frac{1}{2}} \sum_{i,j=1}^{p} (\alpha_{ij} - \beta_{ij}) Z_{ij},$$

where

$$a = \frac{|M|}{|M_1|},$$

α_{ij} is the element (i, j) of the matrix M^{-1} and β_{ij} that of M_1^{-1}. By our assumption $M \neq M_1$, hence $\alpha_{ij} - \beta_{ij}$ cannot vanish for all (i,j). The quantity σ^2 in (14) is the expectation of the square of

$$(36) \qquad a \sum_{i,j} (\alpha_{ij} - \beta_{ij})(U_{ij} - \sigma_{ij}) = a \sum_{i,j} (\alpha_{ij} - \beta_{ij}) U_{ij}$$

and has the value

$$\sigma_5^2 = a^2 \sum_{i,j,\mu,\nu} (\alpha_{ij} - \beta_{ij})(\alpha_{\mu\nu} - \beta_{\mu\nu}) \sigma_{ij\mu\nu},$$

where

$$\sigma_{ij\mu\nu} = E\{(X_i - \xi_i)(X_j - \xi_j)(X_\mu - \xi_\mu)(X_\nu - \xi_\nu)\}\,. \tag{37}$$

Since (36) cannot be essentially zero, the limiting distribution of $N^{\frac{1}{2}}(T_5 - a)$ is the normal distribution about zero with the variance $\sigma_5^2 \neq 0$.

Suppose next that the hypothesis is true, so that the sets (30) are independent. Then we can assume without loss of generality that $\sigma_{ii} = 1$ and $\sigma_{ij} = 0$ for $i \neq j$, $(i, j = 1, \cdots, p)$. For, if these are not true, we can subject each set in (30) to a linear transformation so that for the new variables the variances are one and the covariances are zero. The new sets of variables are still independent whereas T_5 is invariant under such a transformation. Hence our problem reduces to finding the limiting distribution of T_5 under the assumption that

$$M = I \tag{38}$$

and that the sets (30) are independent.

Remembering the rule of simplification (sec. 12) we make the substitution $v_{ii} = 1 + N^{-\frac{1}{2}}Z_{ii}$, $v_{ij} = N^{-\frac{1}{2}}Z_{ij}$, $(i \neq j)$, in T_5 and obtain

$$T_5 = \frac{|I + N^{-\frac{1}{2}}Z\,|}{|I + N^{-\frac{1}{2}}Z_1|}, \tag{39}$$

where Z and Z_1 are the matrices obtained on replacing each v_{ij} by Z_{ij} in V and V_1. The three-term expansion of (39) is the same as that of

$$\frac{1 + N^{-\frac{1}{2}}b + N^{-1}c}{1 + N^{-\frac{1}{2}}b_1 + N^{-1}c_1},$$

where b is the sum of the diagonal elements of Z, c is the sum of the two-rowed principal minors of Z, b_1 and c_1 are the same functions of the elements of Z_1. Since evidently $b = b_1$, we have the expansion

$$1 - \frac{1}{N}(c_1 - c)\,.$$

Obviously

$$c_1 - c = \sum_{i<j}{}' Z_{ij}^2\,, \tag{40}$$

where Σ' denotes summation extended to those (i, j) for which the position of Z_{ij} in Z is the position of a zero in Z_1. Hence the limiting distribution of $N(1 - T_5)$ is the same as that of (40), that is, the distribution of

$$\sum_{i<j}{}' W_{ij}^2, \tag{41}$$

where the W_{ij} are normal varieties having zero means and the same second moments as the set U_{ij}. Under the assumption of independence and (38) we have, for all the W_{ij} in (41),

$$E(W_{ij}^2) = \sigma_{iijj} = \sigma_{ii}\sigma_{jj} = 1,$$

$$E(W_{ij}W_{\mu\nu}) = \sigma_{ij\mu\nu} = \sigma_{i\mu}\sigma_{j\nu} = 0, \qquad (i,j) \neq (\mu,\nu).$$

Hence the W_{ij} in (41) are independent unit normal variates. Thus the limiting distribution of $N(1 - T_5)$ is the χ^2 distribution whose degree of freedom is the number of terms in the sum (41), namely, $\Sigma s_i s_j$ ($i \leqq j; i, j = 1, \cdots, \kappa$).

14. *Example 6: The hypothesis of independence and homoscedasticity and the likelihood ratio test.*—Consider the following hypothesis: The random variables X_i in (30) are independent and $\sigma_{11} = \sigma_{22} = \cdots = \sigma_{pp}$. If we regard the distribution of (30) as normal and apply the principle of likelihood we easily obtain the test function

$$T_6 = \frac{|V|^{\frac{1}{p}}}{\frac{1}{p}(v_{11} + \cdots + v_{pp})},$$

which is the ratio of the geometric mean of the latent roots of V to their arithmetic mean.

Suppose first that not only is the hypothesis false but the relations

$$(42) \qquad \sigma_{11} = \cdots = \sigma_{pp}, \qquad \sigma_{ij} = 0, \qquad i \neq j,$$

are not all true. Making the substitution $v_{ij} = \sigma_{ij} + N^{-\frac{1}{2}} Z_{ij}$ we get the two-term expansion

$$a + aN^{-\frac{1}{2}} \sum_{i,j=1}^{p} b_{ij} Z_{ij}, \qquad Z_{ji} = Z_{ij},$$

where

$$a = \frac{|M|^{\frac{1}{p}}}{\frac{1}{p}(\sigma_{11} + \cdots + \sigma_{pp})},$$

$$b_{ii} = \frac{1}{p} a_{ii} - \frac{1}{\sigma_{11} + \cdots + \sigma_{pp}}, \qquad b_{ij} = \frac{1}{p} a_{ij}, \qquad i \neq j,$$

and the a_{ij} are the elements of M^{-1}. Since some of relations (42) are not true, the b_{ij} cannot all vanish. The quantity σ^2 in (14) is the expectation of the square of

$$(43) \qquad a \sum_{i,j} b_{ij} (U_{ij} - \sigma_{ij}) = a \sum_{i,j} b_{ij} U_{ij}$$

and has the value

$$\sigma_6^2 = a^2 \sum_{i,j,\mu,\nu} b_{ij} b_{\mu\nu} \sigma_{ij\mu\nu},$$

where $\sigma_{ij\mu\nu}$ is defined in (37). Since (43) cannot be essentially zero, the limiting distribution of $N^{\frac{1}{2}}(T_6 - a)$ is the normal distribution about zero with the variance $\sigma_6^2 \neq 0$.

Suppose next that the hypothesis is true, so that $X_1, \cdots, X_p$ are independent and have a common variance η. Making the substitution $v_{ii} = \eta + N^{-\frac{1}{2}}Z_{ii}$, $v_{ij} = N^{-\frac{1}{2}}Z_{ij}, (i \neq j)$, in T_6 we get the three-term expansion

$$1 - \frac{1}{pN\eta^2}\left\{\frac{1}{2}\sum_i Z_{ii}^2 - \frac{1}{2p}\left(\sum_i Z_{ii}\right)^2 + \sum_{i<j} Z_{ij}^2\right\}.$$

Hence $pN(1 - T_6)$ has the same limiting distribution as that of

$$\frac{1}{2\eta^2}\sum_i\left(Z_{ii} - \frac{1}{p}\sum_i Z_{ii}\right)^2 + \frac{1}{\eta^2}\sum_{i\wedge j} Z_{ij}^2 = \sum_{i\leqq j} Y_{ij}^2,$$

where

$$Y_{ii} = \frac{1}{\sqrt{2}\eta}\left(Z_{ii} - \frac{1}{p}\sum_i Z_{ii}\right), \qquad Y_{ij} = \frac{1}{\eta}Z_{ij}, \qquad i<j.$$

The Y's are the normalized sample means of the following system of variates:

(44) $$\frac{1}{\sqrt{2}\eta}\left(U_{ii} - \frac{1}{p}\sum_i U_{ii}\right), \qquad \frac{1}{\eta}U_{ij}, \qquad i<j.$$

The limiting distribution in question is the distribution of

(45) $$\sum_{i\leqq j} W_{ij}^2,$$

where the W_{ij} are normal variates with zero means and the same second moments as (44). Under the assumption that the hypothesis is true, we have

$$E(U_{ii}^2) = \mu_{4i} - \eta^2, \qquad E(U_{ij}^2) = \eta^2, \qquad i<j,$$
$$E(U_{ij}U_{\mu\nu}) = 0, \qquad (i,j) \neq (\mu,\nu),$$

where

$$\mu_{4i} = E\{(X_i - \xi_i)^4\}.$$

Then it is easy to compute the second moments of the W_{ij}:

$$E(W_{ii}^2) = \frac{1}{2p^2}\sum_i \frac{\mu_{4i}}{\eta^2} + \frac{p-2}{2p}\frac{\mu_{4i}}{\eta^2} - \frac{p-1}{2p},$$

$$E(W_{ii}W_{jj}) = \frac{1}{2p^2}\sum_i \frac{\mu_{4i}}{\eta^2} - \frac{1}{2p}\left(\frac{\mu_{4i} + \mu_{4i}}{\eta^2}\right) + \frac{1}{2p}, \qquad i \neq j,$$

$$E(W_{ii}W_{\mu\nu}) = 0, \qquad \mu < \nu,$$

$$E(W_{ij}^2) = 1, \qquad i < j,$$

$$E(W_{ij}W_{\mu\nu}) = 0, \qquad i<j, \mu < \nu; (i,j) \neq (\mu,\nu).$$

It is thus seen that the distribution of (45) is the composition of two independent parts: the part contributed by ΣW_{ij}^2, $(i < j)$, and that contributed by ΣW_{ii}^2. The former is χ^2 distribution with $\frac{1}{2}p(p-1)$ degrees of freedom; for the latter we apply lemma 3. Let $\omega_{ij} = E(W_{ii}W_{jj})$ and $\lambda_1, \cdots, \lambda_{p-1}$ be the non-vanishing latent roots of (ω_{ij}), which is of rank $p-1$. Then the distribution of (45), that is, the limiting distribution of $pN(1-T_6)$ is

$$\frac{1}{(2\pi)^{\frac{1}{2}(p-1)}2^{\frac{1}{4}p(p-1)}\Gamma\left(\frac{1}{4}p(p-1)\right)\left(\lambda_1\cdots\lambda_{p-1}\right)^{\frac{1}{2}}}\times$$

$$\int_{y_1^2+\ldots+y_{p-1}^2+z\leq x} z^{\frac{1}{4}p(p-1)-1}\exp\left(-\frac{y_1^2}{2\lambda_1}-\cdots-\frac{y_{p-1}^2}{2\lambda_{p-1}}-\frac{z}{2}\right)dy_1\cdots dy_{p-1}dz.$$

A sufficient condition for this distribution to be the χ^2 distribution with $\frac{1}{2}(p+2)(p-1)$ degrees of freedom is that $\mu_{4i} = 3\eta^2$ for all i, for then $\lambda_1 = \cdots = \lambda_{p-1} = 1$.

It may be noticed that, although the limiting distribution of $N(1-T_5)$ is always the χ^2 distribution when the hypothesis of independence is true, regardless of the distribution of (30), the limiting distribution of $pN(1-T_6)$, even when the hypothesis tested is true, still depends on the fourth moments of (30), and becomes the χ^2 distribution under the condition that

$$E\{(X_i-\xi_i)^4\} = 3\,[E\{(X_i-\xi_i)^2\}]^2 \text{ for all } i.$$

15. *Problems of* k *samples and the statistics* L *and* L₁.—Let

$$X_1, \cdots, X_k$$

be k random variables having

$$E(X_a) = \xi_a, \qquad E\{(X_a-\xi_a)^2\} = \eta_a \neq 0, \qquad \eta_a^{-\frac{3}{2}}E\{(X_a-\xi_a)^3\} = a_a,$$
$$\eta_a^{-2}E\{(X_a-\xi_a)^4\} = b_a < \infty.$$

Let

$$X_{a1}, \cdots, X_{aN_a}, \qquad a = 1, \cdots, k,$$

be k samples of sizes $N_1, \cdots, N_k$. Consider the following two hypotheses:

$$H:\ \xi_1 = \cdots = \xi_k \text{ and } \eta_1 = \cdots = \eta_k,$$

$$H':\ \eta_1 = \cdots = \eta_k.$$

With the help of their method of likelihood ratio applied to normal distributions, Neyman and Pearson [9] obtain the following test functions for H and H' respectively:

$$L = \left(\prod_{a=1}^{k} Y_a^{g_a}\right)\left\{\sum_{a=1}^{k} g_a Y_a + \sum_{a=1}^{k} g_a(\bar{X}_a - \bar{X})^2\right\}^{-1},$$

$$L_1 = \left(\prod_{a=1}^{k} Y_a^{g_a}\right)\left(\sum_{a=1}^{k} g_a Y_a\right)^{-1},$$

where the g_a are defined in (9),

$$\bar{X}_a = \frac{1}{N_a}\sum_{r=1}^{N_a} X_{ar}, \qquad \bar{X} = \sum_{a=1}^{k} g_a \bar{X}_a, \qquad Y_a = \frac{1}{N}\sum_{r=1}^{N_a}(X_{ar} - \bar{X}_a)^2.$$

We shall call L and L_1 respectively T_7 and T_8 and find their limiting distributions when the sample sizes N_a become infinite in the manner specified in section 3, from arbitrary parent distributions.

16. *Example 7: The L-statistic.*—The random variables

$$U_{1a} = X_a - \xi_a, \qquad U_{2a} = (X_a - \xi_a)^2,$$

have the means

$$E(U_{1a}) = 0, \qquad E(U_{2a}) = \eta_a,$$

and sample means

$$\bar{U}_{1a} = \bar{X}_a - \xi_a, \qquad \bar{U}_{2a} = Y_a + \bar{U}_{1a}^2.$$

Hence

$$(46) \quad T_7 = L = \left\{\prod_{a=1}^{k}(\bar{U}_{2a} - \bar{U}_{1a}^2)^{g_a}\right\}\left\{\bar{U}_2 - \bar{U}_1^2 - 2\sum_{a=1}^{k} g_a(\xi_a - \bar{\xi})\bar{U}_{1a} + \sigma_\xi^2\right\}^{-1},$$

where

$$\bar{U}_i = \sum_a g_a U_{ia}, \qquad i = 1, 2,$$

$$\bar{\xi} = \sum_a g_a \xi_a, \qquad \sigma_\xi^2 = \sum_a g_a(\xi_a - \bar{\xi})^2.$$

If the hypothesis H is false, we make the substitution

$$\bar{U}_{1a} = N_a^{-\frac{1}{2}} Z_{1a}, \qquad \bar{U}_{2a} = \eta_a + N_a^{-\frac{1}{2}} Z_{2a},$$

in T_7 and obtain the two-term expansion

$$a + aN^{-\frac{1}{2}}\sum_a (A_a Z_{1a} + B_a Z_{2a}),$$

where

$$a = \left(\prod_a \eta_a^{g_a}\right)\left(\sum_a g_a\eta_a + \sigma_\xi^{12}\right)^{-1},$$

$$A_a = \frac{2g_a^{\frac{1}{2}}(\xi_a - \bar{\xi})}{\sum_a g_a\eta_a + \sigma_\xi^2}, \qquad B_a = g_a^{\frac{1}{2}}\left(\frac{1}{\eta_a} - \frac{1}{\sum_a g_a\eta_a + \sigma\xi^2}\right).$$

The A_a and B_a cannot all vanish, for otherwise H would be true. Let

$$V_a = A_a(X_a - \xi_a) + B_a\{(X_a - \xi_a)^2 - \eta_a\}.$$

The quantity σ^2 in (14) is equal to $a^2\Sigma E(V_a^2)$, $(a = 1, \cdots, k)$, and has the value

$$\sigma_7^2 = a^2\sum_a \{A_a^2\eta_a + 2A_aB_aa_a\eta_a^{\frac{1}{2}} + B_a\eta_a^2(b_a - 1)\}.$$

Suppose that one at least of the V_a is not essentially zero. Then the limiting distribution of $N^{\frac{1}{2}}(T_7 - a)$ is the normal distribution about zero with the variance $\sigma_7^2 \neq 0$.

If the hypothesis H is true, so that

$$\xi_a = \xi, \qquad \eta_a = \eta, \qquad a = 1, \cdots, k,$$

then (46) becomes

$$T_7 = \left\{\prod_{a=1}^{k} (\bar{U}_{2a} - \bar{U}_{1a}^2)g_a\right\}(\bar{U}_{2a} - \bar{U}_1^2)^{-1}.$$

Making the substitution $\bar{U}_{1a} = N_a^{-\frac{1}{2}}Z_{1a}$ and $\bar{U}_{2a} = \eta + N^{-\frac{1}{2}}Z_{2a}$ we obtain the three-term expansion

$$1 - \frac{1}{N}\left\{\frac{1}{\eta}\prod_a Z_{1a}^2 - \frac{1}{\eta}\left(\sum_a g_a^{\frac{1}{2}}Z_{1a}\right)^2 + \frac{1}{2\eta^2}\sum_a Z_{2a}^2 - \frac{1}{2\eta^2}\left(\sum_a g_a^{\frac{1}{2}}Z_{2a}\right)^2\right\}.$$

Hence the limiting distribution of $N(1 - T_7)$ is the distribution of

$$Q = \sum_a W_{1a}^2 - \left(\sum_a g_a^{\frac{1}{2}}W_{1a}\right)^2 + \sum_a W_{2a}^2 - \left(\sum_a g_a^{\frac{1}{2}}W_{2a}\right)^2,$$

where W_{ia}, W_{2a} are normal variates with zero means and the following second moments:

$$E(W_{ia}W_{j\beta}) = 0, \qquad a \neq \beta; i, j = 1, 2,$$

$$E(W_{1a}^2) = \frac{1}{\eta}E(U_{1a}^2) = 1, \qquad E(W_{2a}^2) = \frac{1}{2\eta^2}E\{(U_{2a}^2 - \eta)^2\} = \frac{1}{2}(b_a - 1).$$

$$E(W_{1a}W_{2a}) = \frac{1}{\sqrt{2\eta^{\frac{3}{2}}}}E(U_{1a}^3) = \frac{1}{\sqrt{2}}a_a.$$

Let us treat in detail the particular case where the a_α and the b_α are independent of α:

$$\frac{1}{\sqrt{2}} a_\alpha = A, \qquad \frac{1}{2}(b_\alpha - 1) = B, \qquad \alpha = 1, \cdots, k.$$

It is possible to perform an orthogonal transformation to each of the vectors $[W_{i1}, \cdots, W_{ik}]$:

$$[W_{i1}, \cdots, W_{ik}] \to [W'_{i1}, \cdots, W'_{ik}] \qquad i = 1, 2),$$

so that

$$E(W'_{i\alpha}W'_{j\beta}) = 0, \qquad \alpha \neq \beta; i, j = 1, 2,$$

$$E(W'^2_{1\alpha}) = 1, \quad E(W'_{1\alpha}W'_{2\alpha}) = A, \quad E(W'^2_{2\alpha}) = B, \quad Q = \sum_{\alpha=1}^{k-1}(W'^2_{1\alpha} + W'^2_{2\alpha}).$$

We now apply lemma 3. The dispersion matrix being

$$\begin{bmatrix} I & AI \\ AI & BI \end{bmatrix},$$

its λ-equation is easily reduced to

$$\{\lambda^2 - (1 + B)\lambda + B - A^2\}^{k-1} = 0. \tag{47}$$

If each X_α can take essentially two values, then $B = A^2$ and so the only non-vanishing root of (47) is $1 + B$ of multiplicity $k - 1$. In this case $Q/(1 + B)$ has the χ^2 distribution with $k - 1$ degrees of freedom. In the contrary case, $B > A^2$ and the equation (47) has the roots

$$\gamma_1 = \tfrac{1}{2}[1+B+\{(1-B)^2+4A^2\}^{\frac{1}{2}}], \qquad \gamma_2 = \tfrac{1}{2}[1+B-\{(1-B)^2+4A^2\}^{\frac{1}{2}}],$$

both of multiplicity $k - 1$. The distribution law of Q is then

$$(4\gamma_1\gamma_2)^{-\frac{1}{2}(k-1)}\left\{\Gamma\left(\frac{k-1}{2}\right)\right\}^{-k} \int_R (y_1y_2)^{\frac{1}{2}(k-3)} \exp\left(-\frac{y_1}{2\gamma_1} - \frac{y_2}{2\gamma_2}\right) dy_1 dy_2, \tag{48}$$

where R is the region $0 \leqq y_1, 0 \leqq y_2, y_1 + y_2 \leqq x$. The necessary and sufficient condition for $\gamma_1 = \gamma_2$ is that $B = 1$ and $A = 0$, that is, $b_\alpha = 3$ and $a_\alpha = 0$ for all α. If this condition is satisfied, then $\gamma_1 = \gamma_2 = 1$ and (48) becomes the χ^2 distribution with $2k - 2$ degrees of freedom.

17. *Example 8: The* L_1-*statistic.*—We have

$$T_8 = L_1 = \left(\prod_{\alpha=1}^{k} Y_\alpha^{g_\alpha}\right) \left(\sum_{\alpha=1}^{k} g_\alpha Y_\alpha\right)^{-1}.$$

Following the rule of simplification in section 12 we make the substitution $Y_\alpha = \eta_\alpha + N_\alpha^{-\frac{1}{2}} Z_\alpha$ and expand the result.

If the hypothesis H' is false, the two-term expansion is

$$a + aN^{-\frac{1}{2}} \sum_{\alpha=1}^{k} A_\alpha Z_\alpha ,$$

where

$$a = \left(\prod_\alpha \eta_\alpha^{g_\alpha}\right) \left(\sum_\alpha g_\alpha \eta_\alpha\right)^{-1},$$

$$A_\alpha = g_\alpha^{\frac{1}{2}} \left(\frac{1}{\eta_\alpha} - \frac{1}{\sum_\alpha g_\alpha \eta_\alpha}\right).$$

The A_α cannot all vanish, for otherwise H' would be true. Let

$$V_\alpha = A_\alpha \{(X_\alpha - \xi_\alpha)^2 - \eta_\alpha\} .$$

The quantity σ^2 in (14) has the value

$$\sigma_8{}^2 = a^2 \sum_\alpha E(V_\alpha{}^2) = a^2 \sum_\alpha A_\alpha{}^2 \eta_\alpha{}^2 (b_\alpha - 1) .$$

If the V_α are not all essentially zero, the limiting distribution of $N^{\frac{1}{2}}(T_8 - a)$ is the normal distribution about zero with the variance $\sigma_8{}^2 \neq 0$.

If the hypothesis H' is true, then $\eta_\alpha = \eta$, $(\alpha = 1, \cdots, k)$. Making the substitution $Y_\alpha = \eta + N_\alpha{}^{-\frac{1}{2}} Z_\alpha$ in T_8 we obtain the three-term expansion

$$1 - \frac{1}{2N\eta^2}\left\{\sum_\alpha Z_\alpha{}^2 - \left(\sum_\alpha g_\alpha{}^{\frac{1}{2}} Z_\alpha\right)^2\right\}.$$

Hence the limiting distribution of $N(1 - T_8)$ is the distribution of $\Sigma W_\alpha{}^2 - (\Sigma g_\alpha{}^{\frac{1}{2}} W_\alpha)^2$, $(\alpha = 1, \cdots, k)$, where the W_α are independent normal variates with zero means and $E(W_\alpha{}^2) = (2\eta^2)^{-1} E[\{(X_\alpha - \xi_\alpha)^2 - \eta\}^2] = \frac{1}{2}(b_\alpha - 1)$. In particular, if each $b_\alpha = 3$, the limiting distribution of $N(1 - T_8)$ is the χ^2 distribution with $k - 1$ degrees of freedom.

II

Application to testing hypotheses

18. Lemma 4. *Let* $w = [W_1, \cdots, W_l]$ *be a normally distributed vector such that each* $E(W_\lambda) = 0$ *and the dispersion matrix* Φ *is non-singular. Let* C *be any real matrix of order* $h \times l$, $(h < l)$, *and rank* h. *Then the quadratic form*

$$(49) \qquad X = -\begin{vmatrix} \Phi & C' & w' \\ C & O & o' \\ w & o & o \end{vmatrix} : \begin{vmatrix} \Phi & C' \\ C & O \end{vmatrix}$$

has the χ^2 *distribution with* $l - h$ *degrees of freedom.*

This lemma becomes familiar when $h = 0$, for then C does not appear and X is the quadratic form $w\, \Phi^{-1}\, w'$.

PROOF. By (49) we have

$$X = [w, o]\begin{bmatrix}\Phi & C' \\ C & O\end{bmatrix}^{-1}\begin{bmatrix}w' \\ o'\end{bmatrix} = w\Phi^{-1}w' - w\Phi^{-1}C'(C\Phi^{-1}C')^{-1}C\Phi^{-1}w' .$$

Since Φ is positive definite, there is a real non-singular G such that $G\Phi^{-1}G' = I$. Hence

$$X = yy' - yB'(BB')^{-1}By' , \tag{50}$$

where $y = wG^{-1}$ is again a normally distributed vector and where $B = C\Phi^{-1}G'$ has the same order and rank as C.

The components of y are independent unit normal variates, because the dispersion matrix is $E(y'y) = E(G'^{-1}w'wG^{-1}) = G'^{-1}\Phi G^{-1} = I$. The matrix of the quadratic form (50), $I - B'(BB')^{-1}B = A$ say, has the property that $A^2 = A$. Hence the latent roots of A are either zero or unity. This shows that by an orthogonal transformation X can be reduced to a sum of squares. Hence the χ^2 distribution is established. The number of such squares is $trA = l - trB'(BB')^{-1}B = l - tr(BB')^{-1}BB'$, which gives the degree of freedom, q.e.d.

19. *The case of one sample: the hypothesis* **H**.—Let

$$[U_1, \cdots, U_m]$$

be a random vector, possessing finite second moments and a non-singular dispersion matrix. Let

$$E(U_i) = \mu_i , \qquad E(U_iU_j) - \mu_i\mu_j = \eta_{ij} .$$

Let

$$[U_{1r}, \cdots, U_{mr}] , \qquad r = 1, \cdots, N ,$$

be a sample of size N, and $\bar{U}_i$, Z_i be respectively the sample means and the normalized sample means.

We call the hypothesis **H** the following hypothesis,

$$\mathbf{H}: \ f_\lambda(\mu_1, \cdots, \mu_m) = \sum_{q=1}^{h} c_{q\lambda}\rho_q , \qquad \lambda = 1, \cdots, l , \tag{51}$$

which asserts that l given functions of m populational constants are expressible as linear combinations of h unspecified parameters ρ_q with known coefficients $c_{q\lambda}$. The three integers m, l, and h shall satisfy the relation

$$0 \leqq h < l \leqq m .$$

If $h = 0$, the right-hand side of (51) means zero.

Concerning f_λ and $c_{q\lambda}$, we make the following assumptions:

(i) Each $f_\lambda(x_1, \cdots, x_m)$ is defined in the whole m-dimensional space and possesses continuous third derivatives of every kind in the neighborhood of $(\mu_1, \cdots, \mu_m)$.

(ii) The matrix

$$F = \begin{bmatrix} f_1^{(1)} & \cdots & f_1^{(m)} \\ — & — & — \\ f_l^{(1)} & \cdots & f_l^{(m)} \end{bmatrix}$$

where

$$f_\lambda^{(i)} = \frac{\partial}{\partial \mu_i} f_\lambda(\mu_1, \cdots, \mu_m)$$

is of rank l for all $\mu_1, \cdots, \mu_m$ satisfying (51).

(iii) The matrix

$$C = \begin{bmatrix} c_{11} & \cdots & c_{1l} \\ — & — & — \\ c_{h1} & \cdots & c_{hl} \end{bmatrix}$$

is of rank h.

20. *The unstudentized statistic* D.—Writing

$$Y_\lambda = f_\lambda(\bar{U}_1, \cdots, \bar{U}_m), \qquad y = [Y_1, \cdots, Y_l],$$

$$\varphi_{\lambda\nu} = \sum_{i,j=1}^{m} \eta_{ij} f_\lambda^{(i)} f_\nu^{(j)}, \qquad \Phi = [\varphi_{\lambda\nu}] = F[\eta_{ij}]F',$$

we define the statistic D as

$$D = -\begin{vmatrix} \Phi & C' & y' \\ C & O & 0' \\ y & 0 & 0 \end{vmatrix} : \begin{vmatrix} \Phi & C' \\ C & O \end{vmatrix}. \tag{52}$$

We shall show that, when the hypothesis **H** is true, the limiting distribution of ND is the χ^2 distribution with $l - h$ degrees of freedom. For this purpose we follow the procedure described at the end of section 6. Expanding

$$Y_\lambda = f_\lambda(\mu_1 + N^{-\frac{1}{2}} Z_1, \cdots, \mu_m + N^{-\frac{1}{2}} Z_m)$$

to two terms and using (51) we obtain

$$\sum_{q=1}^{h} c_{q\lambda}\rho_q + N^{-\frac{1}{2}} R_\lambda, \tag{53}$$

where

$$R_\lambda = \sum_{i=1}^{m} f_\lambda^{(i)} Z_i.$$

When (53) is substituted for Y_λ in (52), the terms $\theta_\lambda = \Sigma_q c_{q\lambda}\rho_q$ may be canceled, because $[\theta_1, \cdots, \theta_l]$ is a linear combination of the rows of C. Then we get

$$(54) \qquad -\frac{1}{N}\begin{vmatrix} \Phi & C' & r' \\ C & O & 0' \\ r & 0 & 0 \end{vmatrix} : \begin{vmatrix} \Phi & C' \\ C & O \end{vmatrix},$$

where

$$r = [R_1, \cdots, R_l].$$

The expansion (54) represents the three-term expansion (25) of D. Hence the limiting distribution of ND is the distribution of

$$-\begin{vmatrix} \Phi & C' & w' \\ C & O & 0' \\ w & 0 & 0 \end{vmatrix} : \begin{vmatrix} \Phi & C' \\ C & O \end{vmatrix},$$

where w is a normally distributed vector having zero means and the same dispersion matrix as the system

$$\sum_{i=1}^{m} f_\lambda^{(i)}(U_i - \mu_i), \qquad \lambda = 1, \cdots, l.$$

This dispersion matrix is Φ, which is non-singular under our assumptions. The result now follows from lemma 4.

21. *Studentization of* D.—In order to construct a test function for the hypothesis **H**, we still have to studentize D, that is, to replace the unknown populational constants in Φ by quantities computable from the sample. For this purpose we may replace the set $\varphi_{\lambda\nu}$ by any functions $\psi_{\lambda\nu}$ of sample means (so that the procedure described at the end of section 6 may be applied), provided that when each argument of $\psi_{\lambda\nu}$ is replaced by its expectation the result is $\varphi_{\lambda\nu}$. The studentized statistic

$$ND_1 = -N\begin{vmatrix} \Psi & C' & y' \\ C & O & 0' \\ y & 0 & 0 \end{vmatrix} : \begin{vmatrix} \Psi & C' \\ C & O \end{vmatrix}, \qquad \Psi = (\psi_{ij}),$$

thus obtained has the same limiting distribution as ND when the hypothesis **H** is true, for evidently the expansion (54) is not affected through the replacement of $\varphi_{\lambda\nu}$ by $\psi_{\lambda\nu}$. In their generality the functions $\psi_{\lambda\nu}$ cannot be specified by any fixed rule. In concrete cases, as manifested by the examples given below, most natural functions playing the roles of the $\psi_{\lambda\nu}$ suggest themselves.

A practical difficulty in significance tests is that there are many conceivable composite hypotheses for which one does not know how to construct a test criterion even to satisfy the single requirement of exactness. The hypothesis **H** has many special cases of this kind. When the sample is large, ND_1 may be employed as a test function as its distribution in the limit is known and is independent of any nuisance parameters. The actual test consists in computing ND_1 and referring to the χ^2 distribution with $l - h$ degrees of freedom, large values of ND_1 being significant. A further justification of the test is that its power tends generally to unity as its limit, as will be shown in the next section.

22. *Power of the test.*—If the hypothesis $\mathbf{H}$ is false, then in the expansion of D_1 the constant term is

$$a = -\begin{vmatrix} \Phi & C' & a' \\ C & O & 0' \\ a & 0 & 0 \end{vmatrix} : \begin{vmatrix} \Phi & C' \\ C & O \end{vmatrix} > 0 ,$$

where

$$a = [a_1, \cdots, a_l] , \qquad a_\lambda = f_\lambda(\mu_1, \cdots, \mu_m) ,$$

and the term with $N^{-\frac{1}{2}}$ is not in general essentially zero. Hence $N^{\frac{1}{2}}(D_1 - a)$ tends to be normally distributed about zero with a dispersion $\sigma^2 > 0$. If the test criterion based on D_1 is to reject $\mathbf{H}$ when $ND_1 \geqq c$, then the power of the test is asymptotically equal to

$$\frac{1}{\sqrt{2\pi}\,\sigma} \int_{N^{\frac{1}{2}}(N^{-1}c-a)}^{\infty} e^{-\frac{1}{2}y^2} dy ,$$

which tends to one as $N \to \infty$.

In the following six examples we consider a random vector

$$[X_1, \cdots, X_p] \tag{55}$$

as in section 12. The letters ξ_i, σ_{ij}, $\overline{X}_i$, v_{ij}, U_i, U_{ij} have the same meaning as in that section. Besides, we write

$$\sigma_{ijkl} = E(U_iU_jU_kU_l), \qquad v_{ijkl} = \frac{1}{N}\sum_{r=1}^{N} (X_{ir}-\overline{X}_i)(X_{jr}-\overline{X}_j)(X_{kr}-\overline{X}_k)(X_{lr}-\overline{X}_l).$$

The relations

$$\sigma_{ijkl} = \sigma_{ij}\sigma_{kl} + \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk} , \qquad i, j, k, l = 1, \cdots, p , \tag{56}$$

which hold true if the distribution of (55) is normal, will be called the normal moment relations. In the first two examples the conditions that the fourth moments are finite and that the X_i do not satisfy any quadratic relation with unit probability may be removed.

23. *Example 1: To test the hypothesis* H_1.[4]

$$H_1\colon \quad \xi_i = 0 , \qquad i = 1, \cdots, p .$$

[4] The hypothesis $\xi_i = \xi_i^\circ$, $(i = 1, \ldots, p)$, may be reduced to this by using $X_i - \xi_i^\circ$ instead of X_i.

We have

$$D = -\begin{vmatrix} M & \bar{x}' \\ \bar{x} & 0 \end{vmatrix} : |M|,$$

where $M = [\sigma_{ij}]$ and $\bar{x} = [\overline{X}_1, \cdots \overline{X}_p]$. In order to studentize D it is natural to employ v_{ij} in the place of σ_{ij}. Then

$$(57)\qquad ND_1 = -N\begin{vmatrix} V & \bar{x}' \\ \bar{x} & 0 \end{vmatrix} : |V|, \qquad V = [v_{ij}].$$

The expression (57) is Hotelling's T^2-statistic (see [**3**] and [**6**]) except for a factor depending on N. Its limiting distribution when H_1 is true is the χ^2 distribution with p degrees of freedom, valid for arbitrary parent distribution.

24. *Example 2: To test the hypothesis* H_2.

$$H_2: \quad \xi_i = \xi, \qquad i = 1, \cdots, p.$$

Here

$$D = -\begin{vmatrix} M & j' & \bar{x}' \\ j & 0 & 0 \\ \bar{x} & 0 & 0 \end{vmatrix} : \begin{vmatrix} M & j' \\ j & 0 \end{vmatrix}, \qquad j = [1, 1, \cdots, 1],$$

$$(58)\qquad ND_1 = -N\begin{vmatrix} V & j' & \bar{x}' \\ j & 0 & 0 \\ \bar{x} & 0 & 0 \end{vmatrix} : \begin{vmatrix} V & j' \\ j & 0 \end{vmatrix}.$$

The statistic (58) has been studied elsewhere (see [**6**]). Its limiting distribution when H_2 is true is the χ^2 distribution with $p - 1$ degrees of freedom, valid for arbitrary parent distribution.

25. *Example 3: To test the hypothesis* H_3.

$$H_3: \quad \sigma_{11} = \cdots = \sigma_{pp}.$$

Here

$$D = -\begin{vmatrix} \gamma_{11} & \cdots & \gamma_{1p} & 1 & v_{11} \\ \hline \gamma_{p1} & \cdots & \gamma_{pp} & 1 & v_{pp} \\ 1 & \cdots & 1 & 0 & 0 \\ v_{11} & \cdots & v_{pp} & 0 & 0 \end{vmatrix} : \begin{vmatrix} \gamma_{11} & \cdots & \gamma_{1p} & 1 \\ \hline \gamma_{p1} & \cdots & \gamma_{pp} & 1 \\ 1 & \cdots & 1 & 0 \end{vmatrix},$$

where

$$\gamma_{ij} = \sigma_{iijj} - \sigma_{ii}\sigma_{jj}.$$

If no further knowledge is assumed about the parent distribution, we may use $v_{iijj} - v_{ii}v_{jj}$ in place of γ_{ij} for studentization. If the normal moment relations are assumed, then

$$\gamma_{ij} = 2\sigma_{ij}^2.$$

We may studentize D by using v_{ij} for σ_{ij}, $(i \neq j)$, and $\bar{v} = \frac{1}{p}(v_{11} + \cdots + v_{pp})$ for each σ_{ii}. Then

$$
(59) \qquad ND_1 = -N \begin{vmatrix} 2\bar{v}^2 & 2v_{12}^2 & \cdots & 2v_{1p}^2 & 1 & v_{11} \\ 2v_{21}^2 & 2\bar{v}^2 & \cdots & 2v_{2p} & 1 & v_{22} \\ - & - & - & - & - & - \\ 2v_{p1}^2 & 2v_{p2}^2 & \cdots & 2\bar{v}^2 & 1 & v_{pp} \\ 1 & 1 & \cdots & 1 & 0 & 0 \\ v_{11} & v_{22} & \cdots & v_{pp} & 0 & 0 \end{vmatrix} : \begin{vmatrix} 2\bar{v}^2 & 2v_{12}^2 & \cdots & 2v_{1p}^2 & 1 \\ 2v_{21}^2 & 2\bar{v}^2 & \cdots & 2v_{2p}^2 & 1 \\ - & - & - & - & - \\ 2v_{p1}^2 & 2v_{p2}^2 & \cdots & 2\bar{v}^2 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{vmatrix} .
$$

The limiting distribution of (59) when H_3 is true is the χ^2 distribution with $p - 1$ degrees of freedom. If $p = 2$, (59) reduces to

$$
\frac{N(v_{11} - v_{22})^2}{(v_{11} + v_{22})^2 - 4v_{12}^2},
$$

which is the test function obtained by C. T. Hsu [5].

26. *Example 4: To test the hypothesis,* H_4, *that* $X_1, \cdots, X_p$ *are independent and homoscedastic.*

As a consequence of H_4 we have

$$
(60) \qquad \sigma_{ii} = \eta, \qquad \sigma_{ij} = 0, \qquad i \neq j; \qquad i, j = 1, \cdots, p.
$$

Then

$$
(61) \qquad D = -\begin{vmatrix} \gamma_{11} & & & & & & & 0 & 1 & v_{11} \\ & \ddots & & & & & & & \vdots & \vdots \\ & & \gamma_{pp} & & & & & & 1 & v_{pp} \\ & & & \sigma_{11}\sigma_{22} & & & & & 0 & v_{12} \\ & & & & \sigma_{11}\sigma_{33} & & & & 0 & v_{13} \\ & & & & & \ddots & & & \vdots & \vdots \\ 0 & & & & & & & \sigma_{p-1,p-1}\sigma_{pp} & 0 & v_{p-1,p} \\ 1 & \cdots & 1 & 0 & 0 & \cdots & & 0 & 0 & 0 \\ v_{11} & \cdots & v_{pp} & v_{12} & v_{13} & \cdots & & v_{p-1,p} & 0 & 0 \end{vmatrix} : \begin{vmatrix} \gamma_{11} & & & & & & & 0 & 1 \\ & \ddots & & & & & & & \vdots \\ & & \gamma_{pp} & & & & & & 1 \\ & & & \sigma_{11}\sigma_{22} & & & & & 0 \\ & & & & \sigma_{11}\sigma_{33} & & & & 0 \\ & & & & & \ddots & & & 0 \\ & & & & & & & & \vdots \\ 0 & & & & & & & \sigma_{p-1,p-1}\sigma_{pp} & 0 \\ 1 & \cdots & 1 & 0 & 0 & \cdots & & 0 & 0 \end{vmatrix}
$$

because under the hypothesis H_4 the dispersion matrix of the vector

$$
[U_{11} - \sigma_{11}, \cdots, U_{pp} - \sigma_{pp}, U_{12}, U_{13}, \cdots, U_{p-1,p}]
$$

is the diagonal matrix figuring in (61).

We may studentize D by using v_{iiii} and v_{ii} in place of σ_{iiii} and σ_{ii}, assuming nothing further about the parent distribution. If the normal moment relations are assumed, then, taking into account (60), we have

$$\sigma_{ii} = \eta \,, \qquad \gamma_{ii} = 2\eta^2 .$$

Hence

$$D = \frac{1}{\eta^2}\left(\frac{1}{2}\sum_{i=1}^{p} (v_{ii} - \bar{v})^2 + \sum_{i<j} v_{ij}^2\right).$$

In order to studentize D we may replace η by $\bar{v}$. Then

$$(62) \qquad ND_1 = \frac{N}{\bar{v}^2}\left(\frac{1}{2}\sum_{i} (v_{ii} - \bar{v})^2 + \sum_{i<j} v_{ij}^2\right) = \frac{N}{2}\left(\frac{1}{\bar{v}^2}\sum_{i,j=1}^{p} v_{ij}^2 - p\right).$$

When H_4 is true, the limiting distribution of (62) is the χ^2 distribution with $\frac{1}{2}(p+2)(p-1)$ degrees of freedom.

27. *Example 5: Given that* p = 4, *to test the hypothesis that the three tetrad differences are zero.*

This is equivalent to

$$H_5: \quad \sigma_{12}\sigma_{34} = \sigma_{13}\sigma_{24} = \sigma_{14}\sigma_{23} = \theta .$$

We have

$$(63) \qquad D = -\begin{vmatrix} \varphi_{11} & \varphi_{12} & \varphi_{13} & 1 & Y_1 \\ \varphi_{21} & \varphi_{22} & \varphi_{23} & 1 & Y_2 \\ \varphi_{31} & \varphi_{32} & \varphi_{33} & 1 & Y_3 \\ 1 & 1 & 1 & 0 & 0 \\ Y_1 & Y_2 & Y_3 & 0 & 0 \end{vmatrix} : \begin{vmatrix} \varphi_{11} & \varphi_{12} & \varphi_{13} & 1 \\ \varphi_{21} & \varphi_{22} & \varphi_{23} & 1 \\ \varphi_{31} & \varphi_{32} & \varphi_{33} & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix},$$

where

$$Y_1 = v_{12}v_{34}, \qquad Y_2 = v_{13}v_{24}\,, \qquad Y_3 = v_{14}v_{23}\,.$$

In the expansions of the Y_i the coefficients of $N^{-\frac{1}{2}}$ are

$$\sigma_{34}Z_{12} + \sigma_{12}Z_{34}\,, \qquad \sigma_{24}Z_{13} + \sigma_{13}Z_{24}\,, \qquad \sigma_{23}Z_{14} + \sigma_{14}Z_{23}\,,$$

where Z_{ij} is the normalized sample mean of $U_{ij} - \sigma_{ij}$. Hence $[\varphi_{ij}]$ is the dispersion matrix of the three variables

$$\sigma_{34}U_{12} + \sigma_{12}U_{34} - 2\sigma_{12}\sigma_{34}\,, \qquad \sigma_{24}U_{13} + \sigma_{13}U_{24} - 2\sigma_{13}\sigma_{24}\,,$$

$$\sigma_{23}U_{14} + \sigma_{14}U_{23} - 2\sigma_{14}\sigma_{23}\,.$$

$$\begin{aligned}
\varphi_{11} &= \sigma_{34}^2\sigma_{1122} + \sigma_{12}^2\sigma_{3344} + 2\sigma_{12}\sigma_{34}\sigma_{1234} - 4\sigma_{12}^2\sigma_{34}^2, \\
\varphi_{22} &= \sigma_{24}^2\sigma_{1133} + \sigma_{13}^2\sigma_{2244} + 2\sigma_{13}\sigma_{24}\sigma_{1324} - 4\sigma_{13}^2\sigma_{24}^2, \\
\varphi_{33} &= \sigma_{23}^2\sigma_{1144} + \sigma_{14}^2\sigma_{2233} + 2\sigma_{14}\sigma_{23}\sigma_{1423} - 4\sigma_{14}^2\sigma_{23}^2, \\
\varphi_{12} &= \sigma_{12}\sigma_{13}\sigma_{2344} + \sigma_{12}\sigma_{24}\sigma_{1433} + \sigma_{13}\sigma_{34}\sigma_{1422} + \sigma_{24}\sigma_{34}\sigma_{2311} - 4\sigma_{12}\sigma_{13}\sigma_{24}\sigma_{34}, \\
\varphi_{13} &= \sigma_{12}\sigma_{14}\sigma_{2433} + \sigma_{12}\sigma_{23}\sigma_{1344} + \sigma_{23}\sigma_{34}\sigma_{2411} + \sigma_{14}\sigma_{34}\sigma_{1322} - 4\sigma_{12}\sigma_{14}\sigma_{23}\sigma_{34}, \\
\varphi_{23} &= \sigma_{13}\sigma_{14}\sigma_{3422} + \sigma_{23}\sigma_{24}\sigma_{3411} + \sigma_{13}\sigma_{23}\sigma_{1244} + \sigma_{14}\sigma_{24}\sigma_{1233} - 4\sigma_{13}\sigma_{14}\sigma_{23}\sigma_{24}.
\end{aligned}$$

With no further knowledge on the parent distribution we can only studentize D by means of the fourth moments. If the normal moment relations are assumed, then

$$\begin{aligned}
\varphi_{11} &= \sigma_{11}\sigma_{22}\sigma_{34}^2 + \sigma_{33}\sigma_{44}\sigma_{12}^2 + 2a\sigma_{12}\sigma_{34}, \\
\varphi_{22} &= \sigma_{11}\sigma_{33}\sigma_{24}^2 + \sigma_{22}\sigma_{44}\sigma_{13}^2 + 2a\sigma_{13}\sigma_{24}, \\
\varphi_{33} &= \sigma_{11}\sigma_{44}\sigma_{23}^2 + \sigma_{22}\sigma_{33}\sigma_{14}^2 + 2a\sigma_{14}\sigma_{33}, \\
\varphi_{12} &= b + 4\sigma_{12}\sigma_{13}\sigma_{24}\sigma_{34}, \\
\varphi_{13} &= b + 4\sigma_{12}\sigma_{14}\sigma_{23}\sigma_{34}, \\
\varphi_{23} &= b + 4\sigma_{13}\sigma_{14}\sigma_{23}\sigma_{24}.
\end{aligned}$$

Where

$$a = \sigma_{12}\sigma_{34}+\sigma_{13}\sigma_{24}+\sigma_{14}\sigma_{23}, \quad b = \sigma_{11}\sigma_{23}\sigma_{24}\sigma_{34}+\sigma_{22}\sigma_{13}\sigma_{14}\sigma_{34}+\sigma_{33}\sigma_{12}\sigma_{14}\sigma_{24}+\sigma_{44}\sigma_{12}\sigma_{13}\sigma_{23}.$$

If H_5 is true, then

$$\begin{aligned}
\varphi_{11} &= \sigma_{11}\sigma_{22}\sigma_{34}^2 + \sigma_{33}\sigma_{44}\sigma_{12}^2 + 6\theta^2, \\
\varphi_{22} &= \sigma_{11}\sigma_{33}\sigma_{24}^2 + \sigma_{22}\sigma_{44}\sigma_{13}^2 + 6\theta^2, \\
\varphi_{33} &= \sigma_{11}\sigma_{44}\sigma_{23}^2 + \sigma_{22}\sigma_{33}\sigma_{14}^2 + 6\theta^2, \\
\varphi_{12} &= \varphi_{13} = \varphi_{23} = b + 4\theta^2.
\end{aligned}$$

Substituting in (63) we get by an easy computation

$$D = \frac{c_1(Y_2 - Y_3)^2 + c_2(Y_3 - Y_1)^2 + c_3(Y_1 - Y_2)^2}{c_2c_3 + c_3c_1 + c_1c_2},$$

where

$$\begin{aligned}
c_1 &= \sigma_{11}\sigma_{22}\sigma_{34}^2 + \sigma_{33}\sigma_{44}\sigma_{12}^2 - b + 2\theta^2, \\
c_2 &= \sigma_{11}\sigma_{33}\sigma_{24}^2 + \sigma_{22}\sigma_{44}\sigma_{13}^2 - b + 2\theta^2, \\
c_3 &= \sigma_{11}\sigma_{44}\sigma_{23}^2 + \sigma_{22}\sigma_{33}\sigma_{14}^2 - b + 2\theta^2.
\end{aligned}$$

If now we replace σ_{ij} by v_{ij} and θ by $\frac{1}{3}(Y_1 + Y_2 + Y_3)$ for studentization, we obtain after an easy reduction

$$(64) \quad ND_1 = N\frac{d_1(r_{13}r_{24}-r_{14}r_{23})^2 + d_2(r_{14}r_{23}-r_{12}r_{34})^2 + d_3(r_{12}r_{34}-r_{13}r_{24})^2}{d_2d_3 + d_3d_1 + d_1d_2},$$

where the r_{ij} are the correlation coefficients of the sample and

$$\begin{aligned}
d_1 &= r_{12}^2 + r_{34}^2 + g, \\
d_2 &= r_{13}^2 + r_{24}^2 + g, \\
d_3 &= r_{14}^2 + r_{23}^2 + g.
\end{aligned}$$

$$g = \tfrac{2}{9}(r_{12}r_{34}+r_{13}r_{24}+r_{14}r_{23})^2 - (r_{12}r_{13}r_{23}+r_{12}r_{14}r_{24}+r_{13}r_{14}r_{34}+r_{23}r_{24}r_{34}).$$

When H_5 is true, the limiting distribution of (64) is the χ^2 distribution with two degrees of freedom.

The following example is taken from a paper by D. N. Lawley: a sample of size N yields the correlational matrix

$$\begin{bmatrix} 1 & \cdot 4 & \cdot 4 & \cdot 2 \\ \cdot 4 & 1 & \cdot 7 & \cdot 3 \\ \cdot 4 & \cdot 7 & 1 & \cdot 3 \\ \cdot 2 & \cdot 3 & \cdot 3 & 1 \end{bmatrix}.$$

The expression (64) has the value $0.001085N$. If the hypothesis H_5 is true and N is large, the probability that (64) may exceed this is approximately $e^{-0.000543N}$, which will be significantly small only when N is several thousand. In his paper Lawley proposed another test criterion whose limiting distribution under the hypothesis H_5 is also the χ^2 distribution with two degrees of freedom and whose value for this example is $0.00113N$.

28. *Example 6: To test the hypothesis,* H_6, *that the first* s *and the last* t, $(s+t=p)$, *of the variables* X_i *in* (55) *are independent.*

Under this hypothesis we have

$$(65) \qquad \sigma_{ij} = 0, \qquad i = 1, \cdots, s; \qquad j = s+1, \cdots, s+t.$$

Let the dispersion matrices of $[X_1, \cdots, X_s]$ and $[X_{s+1}, \cdots, X_{s+t}]$ be respectively M_1 and M_2, and let the matrix $V = [v_{ij}]$ be partitioned:

$$V = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix},$$

where V_{11} has s rows and columns, V_{22} has t rows and columns.

On the basis of (65) we construct

$$D = -\begin{vmatrix} \Phi_{11} & \Phi_{12} & \cdots & \Phi_{1s} & v'_1 \\ \Phi_{21} & \Phi_{22} & \cdots & \Phi_{2s} & v'_2 \\ - & - & - & - & - \\ \Phi_{s1} & \Phi_{s2} & \cdots & \Phi_{ss} & v'_s \\ v_1 & v_2 & \cdots & v_s & 0 \end{vmatrix} : \begin{vmatrix} \Phi_{11} & \cdots & \Phi_{1s} \\ - & - & - \\ \Phi_{s1} & \cdots & \Phi_{ss} \end{vmatrix},$$

where

$$v_i = [v_{i,s+1}, \cdots, v_{i,s+t}], \qquad i = 1, \cdots, s,$$

and Φ_{ij} is the covariance matrix of the vectors $[U_{i,s+1} - \sigma_{i,s+1}, \cdots, U_{i,s+t} - \sigma_{i,s+t}]$ and $[U_{j,s+1} - \sigma_{j,s+1}, \cdots, U_{j,s+t} - \sigma_{j,s+t}]$. If H_6 is true, then

$$\Phi_{ij} = \sigma_{ij} M_2, \qquad i, j = 1, \cdots, s.$$

Hence

$$\begin{bmatrix} \Phi_{11} \cdots \Phi_{1s} \\ \hline \Phi_{s1} \cdots \Phi_{ss} \end{bmatrix}^{-1} = \begin{bmatrix} a_{11}M_2^{-1} \cdots a_{1s}M_2^{-1} \\ \hline a_{s1}M_2^{-1} \cdots a_{ss}M_2^{-1} \end{bmatrix},$$

where the a_{ij} are the elements of M_1^{-1}. Then

$$D = [v_1, \cdots, v_s] \begin{bmatrix} \Phi_{11} \cdots \Phi_{1s} \\ \hline \Phi_{s1} \cdots \Phi_{ss} \end{bmatrix}^{-1} \begin{bmatrix} v'_1 \\ \cdot \\ \cdot \\ \cdot \\ v'_s \end{bmatrix}$$

$$= \sum_{i,j=1}^{s} a_{ij} v_i M_2^{-1} v'_j = tr\, M_1^{-1} V_{12} M_2^{-1} V'_{12}\,;$$

in other words, D is equal to the sum of the roots of the equation

$$|V_{12} M_2^{-1} V'_{12} - \lambda M_1| = 0\,.$$

In order to studentize D we have merely to substitute v_{ij} for σ_{ij}. Then

$$D_1 = tr V_{11}^{-1} V_{12} V_{22}^{-1} V'_{12}\,,$$

or the sum of the roots of the equation

$$|V_{12} V_{22}^{-1} V'_{12} - \lambda V_{11}| = 0\,.$$

Hence D_1 is the sum of the canonical correlation coefficients of Hotelling [4]. The limiting distribution of ND_1 when H_6 is true is the χ^2 distribution with st degrees of freedom.

Let us consider the particular case of two sets of events. Let $E_1, \cdots, E_{s+1}$ and $E'_1, \cdots, E'_{t+1}$ be two sets of events each forming a complete disjunction, and let the letters p_{ij}, p_i, p'_j, n_{ij}, n_i, n'_j represent the same quantities as in section 9. If X_i, $(i = 1 \cdots, s)$, is one or zero according as E_i happens or does not happen, and if X_{s+j} $(j = 1, \cdots, t)$, is one or zero according as E'_j happens or does not happen, then H_6 becomes the following hypothesis:

$$p_{ij} = p_i p'_j\,, \qquad i = 1, \cdots, s+1; j = 1, \cdots, t+1\,.$$

Set

$$d_{ij} = n_{ij} - \frac{n_i n'_j}{N}\,, \qquad i = 1, \cdots, s+1; j = 1, \cdots, t+1\,,$$

$$D = \begin{bmatrix} d_{11} \cdots d_{1t} \\ \hline d_{s1} \cdots d_{st} \end{bmatrix}, \qquad \Delta_1 = \frac{1}{N}\begin{bmatrix} n_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & n_s \end{bmatrix}, \qquad \Delta_2 = \frac{1}{N}\begin{bmatrix} n'_1 & & & 0 \\ & \cdot & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & n'_t \end{bmatrix},$$

$$a = \frac{1}{N}[n_1, \cdots, n_s], \qquad b = \frac{1}{N}[n'_1, \cdots, n'_t].$$

Then clearly

$$V_{11} = \Delta_1 - a'a, \qquad V_{22} = \Delta_2 - b'b, \qquad V_{12} = \frac{1}{N}D.$$

Hence

$$V_{11}^{-1} = \Delta_1^{-1} + \frac{N}{n_{s+1}} j'_s j_s, \qquad V_{22}^{-1} = \Delta_2^{-1} + \frac{N}{n_{t+1}'} j'_t j_t,$$

where

$$j_q = [1, 1, \cdots, 1], \qquad (q \text{ elements}).$$

Hence

$$ND_1 = NtrV_{11}^{-1}V_{12}V_{22}^{-1}V'_{12} = \frac{1}{N} tr\left(\Delta_1^{-1} + \frac{N}{n_{s+1}} j'_s j_s\right) D \left(\Delta_2^{-1} + \frac{N}{n'_{t+1}} j'_t j_t\right) D'$$

$$= \frac{1}{N}\left(tr\Delta_1^{-1}D\Delta_2^{-1}D' + \frac{N}{n_{s+1}} j_s D \Delta_2^{-1} D' j'_s + \frac{N}{n'_{t+1}} j_t D' \Delta_1^{-1} D j'_t \right.$$
$$\left. + \frac{N^2}{n_{s+1}n'_{t+1}} [j_s D j'_t]^2\right)$$

$$= N\left(\sum_{i=1}^{s}\sum_{j=1}^{t} \frac{d_{ij}^2}{n_i n'_j} + \frac{1}{n_{s+1}} \sum_{j=1}^{t} \frac{d^2_{s+1,j}}{n'_j} + \frac{1}{n'_{t+1}} \sum_{i=1}^{s} \frac{d^2_{i,s+1}}{n_i} + \frac{d^2_{s+1,t+1}}{n_{s+1}n'_{t+1}}\right)$$

$$= N \sum_{i=1}^{s+1}\sum_{j=1}^{t+1} \frac{d_{ij}^2}{n_i n'_j} = \textit{mean square contingency.}$$

29. *Extension of example* 6 *to several sets of variables.*—If the hypothesis is that

$$(X_1, \cdots, X_{s_1}), (X_{s_1+1}, \cdots, X_{s_1+s_2}), \cdots, (X_{s_1+ \cdots + s_{\kappa-1}+1}, \cdots, X_{s_1+\cdots+s_\kappa})$$

are mutually independent vectors, then our method of construction gives the test function $D_1 = \Sigma D_1(ij)$, $(i, j = 1, \cdots, \kappa; i < j)$, where $D_1(ij)$ is the D_1 in example 6 for the ith and jth vectors. The details of the construction are omitted. The limiting distribution of ND_1 when the hypothesis is true is the χ^2 distribution with $\Sigma s_i s_j$, $(i, j = 1, \cdots, \kappa; i < j)$, degrees of freedom.

30. *The case of* k *samples, the hypothesis* $\mathbf{H}'$, *and the test function* $N\Delta_1$.—In this section we consider again the k random vectors (1); the notation used here is the same as in section 1. Let $f_\lambda(x_1, \cdots, x_m)$, $(\lambda = 1, \cdots, l; l < m)$, be l functions defined in the whole m-dimensional space and possessing continuous third derivatives of every kind in each of the neighborhoods of the points $(\mu_{1a}, \cdots, \mu_{ma})$, $(a = 1, \cdots, k)$. It is assumed that the matrices

$$F_a = \begin{bmatrix} f_{1a}{}^{(1)} \cdots f_{1a}{}^{(m)} \\ \text{— — — —} \\ f_{la}{}^{(1)} \cdots f_{la}{}^{(m)} \end{bmatrix}, \qquad a = 1, \cdots, k,$$

where

$$f_{\lambda a}{}^{(i)} = \frac{\partial}{\partial \mu_{ia}} f_\lambda(\mu_{1a}, \cdots, \mu_{ma}),$$

are of rank l. Let

$$\theta_{\lambda a} = f_\lambda(\mu_{1a}, \cdots, \mu_{ma}), \qquad \lambda = 1, \cdots, l; a = 1, \cdots, k.$$

We call the hypothesis $\mathbf{H}'$ the following hypothesis:

$$\mathbf{H}': \ \theta_{\lambda 1} = \theta_{\lambda 2} = \cdots = \theta_{\lambda k}, \qquad \lambda = 1, \cdots, l.$$

Let

$$Y_{\lambda a} = f_\lambda(\bar{U}_{1a}, \cdots, \bar{U}_{ma}), \qquad y_a = [Y_{1a}, \cdots, Y_{la}],$$

$$\Delta = -\begin{vmatrix} \frac{1}{g_1}\Phi_1 & & O & I & y'_1 \\ & \cdot & & \cdot & \cdot \\ & & \cdot & \cdot & \cdot \\ O & & \frac{1}{g_k}\Phi_k & I & y'_k \\ I & \cdots & I & O & 0' \\ y_1 & \cdots & y_k & 0 & 0 \end{vmatrix} : \begin{vmatrix} \frac{1}{g_1}\Phi_1 & & O & I \\ & \cdot & & \cdot \\ & & \cdot & \cdot \\ O & & \frac{1}{g_k}\Phi_k & I \\ I & \cdots & I & O \end{vmatrix},$$

where the g_a are defined in (9) and $\Phi_a = F_a[\eta_{ij}]F'_a$. Φ_a is the dispersion matrix of $[R_{1a}, \cdots, R_{la}]$, where $R_{\lambda a}$ is the coefficient of $N^{-1/2}$ in the expansion of $Y_{\lambda a}$ and is non-singular under our assumptions. If the hypothesis $\mathbf{H}'$ is true, the limiting distribution of $N\Delta$ (as the sample sizes become infinite in the manner specified in section 3) is the χ^2 distribution with $l(k - 1)$ degrees of freedom. This proposition is a consequence of theorem 2 and lemma 4. Its proof is similar to that for the limiting distribution of ND as set forth in section 20, and is omitted.

Now, if we write

$$y = [y_1, \cdots, y_k]\,,$$

$$\Phi = \left|\begin{array}{cccc} \frac{1}{g_1}\Phi_1 & & O & I \\ & \cdot & & \cdot \\ & & \cdot & \cdot \\ & & \cdot & \cdot \\ O & & \frac{1}{g_k}\Phi_k & I \\ I & \cdots & I & O \end{array}\right|,$$

then

$$\Delta = [y, 0]\, \Phi^{-1}\, [y, 0]^{-1}$$

$$= y\left(\begin{bmatrix} g_1\Phi_1^{-1} & & O \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ O & & g_k\Phi_k^{-1} \end{bmatrix} - \begin{bmatrix} g_1\Phi_1^{-1} \\ \cdot \\ \cdot \\ \cdot \\ g_k\Phi_k^{-1} \end{bmatrix} (g_1\Phi_1^{-1} + \cdots + g_k\Phi_k^{-1})^{-1} [g_1\Phi_1^{-1}, \cdots, g_k\Phi_k^{-1}]\right) y'.$$

Hence

$$(66) \quad N\Delta = \sum_a N_a y_a \Phi_a^{-1} y'_a - \left(\sum_a N_a y_a \Phi_a^{-1}\right)\left(\sum_a N_a \Phi_a^{-1}\right)^{-1}\left(\sum_a N_a \Phi_a^{-1} y'_a\right).$$

If $k = 2$, we write

$$y_1 + y_2 = s\,, \qquad y_1 - y_2 = d\,,$$

so that

$$y_1 = \tfrac{1}{2}(s + d)\,, \quad y_2 = \tfrac{1}{2}(s - d)\,,$$

and substitute in (66). Direct computation shows that the result is independent of s and is equal to

$$\tfrac{1}{4} d \{N_1\Phi_1^{-1} + N_2\Phi_2^{-1} - (N_1\Phi_1^{-1} - N_2\Phi_2^{-1})\,(N_1\Phi_1^{-1} + N_2\Phi_2^{-1})^{-1}\,(N_1\Phi_1^{-1} - N_2\Phi_2^{-1})\} d'\,.$$

But

$$N_1\Phi_1^{-1} + N_2\Phi_2^{-1} - (N_1\Phi_1^{-1} - N_2\Phi_2^{-1})\,(N_1\Phi_1^{-1} + N_2\Phi_2^{-1})^{-1}\,(N_1\Phi_1^{-1} - N_2\Phi_2^{-1})$$

$$= N_1\Phi_1^{-1} + N_2\Phi_2^{-1} - (\overline{N_1\Phi_1^{-1} + N_2\Phi_2^{-1}} - 2N_2\Phi_2^{-1})(N_1\Phi_1^{-1} + N_2\Phi_2^{-1})^{-1} (2N_1\Phi_1^{-1} - \overline{N_1\Phi_1^{-1} + N_2\Phi_2^{-1}}$$

$$= 4N_1N_2\Phi_2^{-1}\,(N_1\Phi_1^{-1} + N_2\Phi_2^{-1})^{-1}\,\Phi_1^{-1} = 4\left(\frac{1}{N_1}\Phi_1 + \frac{1}{N_2}\Phi_2\right)^{-1}.$$

Hence (66) reduces to

$$N\Delta = (y_1 - y_2)\left(\frac{\Phi_1}{N_1} + \frac{\Phi_2}{N_2}\right)^{-1}(y'_1 - y'_2)\,, \tag{67}$$

a result which is to be expected.

If $\Phi_a = \Phi$ for all a, then (66) and (67) reduce to

$$N\Delta = \sum_a N_a y_a \Phi^{-1} y'_a - \frac{1}{N}\left(\sum_a N_a y_a\right)\Phi^{-1}\left(\sum_a N_a y'_a\right), \tag{68}$$

$$N\Delta = \left(\frac{1}{N_1} + \frac{1}{N_2}\right)(y_1 - y_2)\Phi^{-1}(y'_1 - y'_2)\,, \quad \text{if} \quad k = 2\,. \tag{69}$$

When the unknown populational constants involved in the Φ_a in (68) are replaced by appropriate functions of sample means, we have a studentized statistic whose limiting distribution is the same as that of $N\Delta$ when the hypothesis $\mathbf{H}'$ is true. This statistic we shall call $N\Delta_1$ and propose to use as a test function for $\mathbf{H}'$ when the samples are large. The actual test consists in computing $N\Delta_1$, referring to the χ^2 distribution with $l(k - 1)$ degrees of freedom, and rejecting $\mathbf{H}'$ if $N\Delta_1$ is significantly large. The power of the test tends in general to unity as its limit, a fact which may be deduced in the same manner as done in section 22.

In the four examples which follow we consider k random vectors

$$[X_{1a}, \cdots, X_{pa}]\,, \qquad a = 1, \cdots, k\,, \tag{70}$$

each having the properties of (30) described in section 12. In example 1′ only the finiteness of the second moments and the non-singularity of the dispersion matrices need be assumed. The meanings of the symbols ξ_{ia}, σ_{ija}, σ_{ijkla}, $\overline{X}_{ia}$, v_{ija}, v_{ijkla} are self-evident.

31. *Example* 1′: *To test the hypothesis* H_1':

$$H'_1: \quad \xi_{i1} = \xi_{i2} = \cdots = \xi_{ik}\,, \qquad i = 1, \cdots, p\,.$$

Here

$$\Phi_a = M_a = [\sigma_{ija}]\,.$$

Using (66) and (67) we have

$$N\Delta = \sum_a N_a \bar{x}_a M_a^{-1} x_a^{-1} - \left(\sum_a N_a \bar{x}_a M_a^{-1}\right)\left(\sum_a N_a M_a^{-1}\right)^{-1}\left(\sum_a N_a M_a^{-1} x_a^{-1}\right),$$

where

$$\bar{x}_a = [\overline{X}_{1a}, \cdots, \overline{X}_{pa}]\,,$$

and

$$N\Delta = (\bar{x}_1 - \bar{x}_2)\left(\frac{1}{N_1}M_1 + \frac{1}{N_2}M_2\right)^{-1}(\bar{x}'_1 - \bar{x}'_2)\,, \quad \text{if } k = 2\,.$$

If no further knowledge is assumed about the parent distributions, we may studentize $N\Delta$ by employing $V_a = [v_{ija}]$ for M_a. Then

$$(71)\begin{cases} N\Delta_1 = \sum_a N_a\bar{x}_a V_a^{-1}\bar{x}'_a - \left(\sum_a N_a\bar{x}_a V_a^{-1}\right)\left(\sum_a N_a V_a^{-1}\right)^{-1}\left(\sum_a N_a V_a^{-1}\bar{x}'_a\right), \\ N\Delta_1 = (\bar{x}_1 - \bar{x}_2)\left(\frac{1}{N_1}V_1 + \frac{1}{N_2}V_2\right)^{-1}(\bar{x}'_1 - \bar{x}'_2), \quad \text{if } k = 2. \end{cases}$$

The limiting distribution of (71) when H'_1 is true is the χ^2 distribution with $p(k-1)$ degrees of freedom.

If $\sigma_{ija} = \sigma_{ij}$ for all i, j, and a, we write $M = [\sigma_{ij}]$ and get

$$N\Delta = \sum_a N_a\bar{x}_a M^{-1}\bar{x}'_a - N\bar{x}M^{-1}\bar{x}',$$

where $\bar{x} = \frac{1}{N}\sum_a N_a\bar{x}_a$ is the row vector whose components are the grand means $\overline{X}_1, \cdots, \overline{X}_p$. Hence, writing $[a_{ij}] = M^{-1}$, we have

$$(72)\ N\Delta = \sum_{a=1}^{k} N_a \sum_{i,j=1}^{p} a_{ij}\overline{X}_{ia}\overline{X}_{ja} - N\sum_{i,j=1}^{p} a_{ij}\overline{X}_i\overline{X}_j = \sum_{i,j} a_{ij}\left(\sum_a N_a\overline{X}_{ia}\overline{X}_{ja} - N\overline{X}_i\overline{X}_j\right)$$

$$= \sum_{i,j} a_{ij}\sum_a N_a(\overline{X}_{ia} - \overline{X}_i)(\overline{X}_{ja} - \overline{X}_j).$$

In order to studentize (72) we use

$$v_{ij} = \frac{1}{N}\sum_{a=1}^{k}\sum_{r=1}^{N_a}(X_{iar} - \overline{X}_i)(X_{jar} - \overline{X}_j)$$

in place of σ_{ij}. Setting $[a_{ij}] = [v_{ij}]^{-1}$ we have

$$(73)\qquad N\Delta_1 = \sum_{i,j} a_{ij}\sum_a N_a(\overline{X}_{ia} - \overline{X}_i)(\overline{X}_{ja} - \overline{X}_j).$$

Consider now the particular case of k sets of events. Let

$$(74)\qquad E_{1a}, E_{2a}, \cdots, E_{ma}, \qquad a = 1, \cdots, k,$$

be k sets of events, each forming a complete disjunction. Let $P(E_{ia}) = p_{ia}$. If X_{ia}, $(i = 1, \cdots, m-1;\ a = 1, \cdots, k)$, is one or zero according as E_{ia} happens or does not happen, then H'_1 becomes the hypothesis

$$(75)\qquad p_{i1} = p_{i2} = \cdots = p_{ik} = p_i, \qquad i = 1, \cdots, m.$$

Since

$$\sigma_{iia} = p_{ia}(1 - p_{ia}) , \qquad \sigma_{ija} = -p_{ia}p_{ja} , \qquad i \neq j ,$$

we have, when (75) is true,

$$\sigma_{iia} = p_i(1 - p_i) , \qquad \sigma_{ija} = -p_i p_j , \qquad i \neq j .$$

Hence (73) can be used. Now if N_a trials of experiment are made on the ath set (74), if the number of happenings of E_{ia} is n_{ia}, and if $n_i = \sum_a n_{ia}$, then

$$\bar{X}_{ia} = \frac{n_{ia}}{N_a} , \qquad \bar{X}_i = \frac{n_i}{N} ,$$

$$v_{ii} = \frac{n_i}{N}\left(1 - \frac{n_i}{N}\right) , \qquad v_{ij} = -\frac{n_i n_j}{N^2} , \qquad i \neq j; \qquad i, j = 1, \cdots , m - 1 .$$

Hence

$$a_{ii} = \frac{N}{n_i} + \frac{N}{n_m} , \qquad a_{ij} = \frac{N}{n_m} , \qquad i \neq j ,$$

(76) $$N\Delta_1 = \sum_{i=1}^{m} \sum_{a=1}^{k} \frac{N\left(n_{ia} - \frac{n_i N_a}{N}\right)^2}{n_i N_a} .$$

The limiting distribution of (76) when (75) is true is the χ^2 distribution with $(m - 1)(k - 1)$ degrees of freedom.

32. *Example 2′: To test the hypothesis* H'_2:

$$H'_2: \quad \sigma_{ija} = \sigma_{ij} , \qquad i, j = 1, \cdots , p; \; a = 1, \cdots , k .$$

Here

$$y_a = [v_{11a}, v_{12a}, \cdots , v_{22a}, v_{23a}, \cdots , v_{p-1,pa}, v_{ppa}] ,$$

and therefore Φ_a is dispersion matrix of the system

$$U_{11a} - \sigma_{11a}, U_{12a} - \sigma_{12a}, \cdots , U_{p-1,pa} - \sigma_{p-1,pa}, U_{ppa} - \sigma_{ppa} ,$$

where

(77) $$U_{ija} = (X_{ia} - \xi_{ia})(X_{ja} - \xi_{ja}) .$$

The elements of Φ_a are

(78) $$E\{(U_{ija} - \sigma_{ija})(U_{kla} - \sigma_{kla})\} = \sigma_{ijkla} - \sigma_{ija}\sigma_{kla} .$$

If nothing is assumed about the parent distributions, we may employ v_{ijkla} for studentization. If the normal moment relations (56) are assumed for each vector (70), then (78) reduces to $\sigma_{ika}\sigma_{jla} + \sigma_{ila}\sigma_{jka}$, which under the hypothesis H'_2 is equal to $\sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}$. Hence $\Phi_a = \Phi$ and the formula (68) can be used. Replacing each σ_{ij} by $\bar{v}_{ij} = \frac{1}{N}\sum_a N_a v_{ija}$ in Φ we get the studentized statistic $N\Delta_1$ whose limiting distribution when H'_2 is true is the χ^2 distribution with $\frac{1}{2}p(p+1)(k-1)$ degrees of freedom.

33. *Example 3′: To test the hypothesis* H′$_3$:

$$H'_3: \quad \rho_{ija} = \rho_{ij}, \qquad i, j = 1, \cdots, p: a = 1, \cdots, k,$$

where ρ_{ija} *is the correlation coefficient of* X_{ia} *and* X_{ja}.

Here

$$y_a = [r_{12a}, r_{13a}, \cdots, r_{23a}, \cdots, r_{p-1,pa}],$$

where the r_{ija} are the sample correlation coefficients.

Now

$$r_{ija} = \frac{\bar{U}_{ija} - \bar{U}_{ia}\bar{U}_{ja}}{(\bar{U}_{iia} - \bar{U}_{ia}^2)^{\frac{1}{2}}(\bar{U}_{jja} - \bar{U}_{ja}^2)^{\frac{1}{2}}},$$

where $U_{ia} = X_{ia} - \xi_{ia}$ and U_{ija} is defined in (77). Setting $\bar{U}_{ia} = N_a^{-\frac{1}{2}}Z_{ia}$, $\bar{U}_{ija} = \rho_{ija}(\sigma_{iia}\sigma_{jja})^{\frac{1}{2}} + N_a^{-\frac{1}{2}}Z_{ija}$ in r_{ija} and expanding in powers of $N_a^{-\frac{1}{2}}$ we obtain the following coefficients of $N_a^{-\frac{1}{2}}$:

$$\frac{Z_{ija}}{\sqrt{\sigma_{iia}\sigma_{jja}}} - \frac{1}{2}\rho_{ija}\left(\frac{Z_{iia}}{\sigma_{iia}} + \frac{Z_{jja}}{\sigma_{jja}}\right),$$

which is the normalized sample mean of

$$T_{ija} = \frac{U_{ija}}{\sqrt{\sigma_{iia}\sigma_{jja}}} - \frac{1}{2}\rho_{ija}\left(\frac{U_{iia}}{\sigma_{iia}} + \frac{U_{jja}}{\sigma_{jja}}\right).$$

Hence Φ is the dispersion matrix of the system

$$T_{12a}, T_{13a}, \cdots, T_{23a}, \cdots, T_{p-1,pa}.$$

The elements of Φ_a are

$$(79)\quad E(T_{ija}T_{kla}) = \tau_{ijkla} - \tfrac{1}{2}\rho_{ija}(\tau_{iikla} + \tau_{jjkla}) - \tfrac{1}{2}\rho_{kla}(\tau_{kkija} + \tau_{llija}) + \tfrac{1}{4}\rho_{ija}\rho_{kla}(\tau_{iikka} + \tau_{iilla} + \tau_{jjkka} + \tau_{jjkka}),$$

where

$$\tau_{ijkla} = \frac{\sigma_{ijkla}}{\sqrt{\sigma_{iia}\sigma_{jja}\sigma_{kka}\sigma_{lla}}}.$$

With no further knowledge on the parent distributions we can only studentize by means of v_{ijkla}. If the normal moment relations (56) are assumed for each vector (70), then

$$\tau_{ijkl} = \rho_{ija}\rho_{kla} + \rho_{ika}\rho_{jla} + \rho_{ila}\rho_{jka} .$$

If the hypothesis H'_3 is also taken into account, (79) becomes

$$\rho_{ik}\rho_{jl} + \rho_{il}\rho_{jk} - (\rho_{ij}\rho_{ik}\rho_{il} + \rho_{ij}\rho_{jk}\rho_{jl} + \rho_{ik}\rho_{jk}\rho_{kl} + \rho_{il}\rho_{jl}\rho_{kl}) + \tfrac{1}{2}\rho_{ij}\rho_{kl}(\rho_{ik}^2 + \rho_{il}^2 + \rho_{jk}^2 + \rho_{jl}^2) .$$

Hence in this case $\Phi_a = \Phi$ and the formula (68) can be used. The studentization consists in replacing ρ_{ij} by one of the following functions:

$$\frac{1}{k}\sum_a r_{ija} , \qquad \frac{1}{N}\sum N_a r_{ija} , \qquad \sum_a N_a v_{ija} / \sum_a N_a (v_{iia} v_{jja})^{\frac{1}{2}} .$$

Then we get a test function $N\Delta_1$ whose limiting distribution when H'_3 is true is the χ^2 distribution with $\frac{1}{2}p(p-1)(k-1)$ degrees of freedom.

34. *Example* 4′: *Given that*

$$(80) \qquad \sigma_{iia} = \sigma_{ii} , \qquad i = 1, \cdots, p; \; a = 1, \cdots, k ,$$

to test the hypothesis H′$_4$:

$$\mathrm{H}'_4: \quad \sigma_{ija} = \sigma_{ij} , \qquad i \neq j; i, j = 1, \cdots, p; \; a = 1, \cdots, k .$$

Here

$$y_a = [v_{12a}, v_{13a}, \cdots, v_{23a}, \cdots, v_{p-1,pa}] .$$

Hence Φ_a is the dispersion matrix of the system

$$U_{12a} - \sigma_{12a}, U_{13a} - \sigma_{13a}, \cdots, U_{23a} - \sigma_{23a}, \cdots, U_{p-1,pa} - \sigma_{p-1,pa} .$$

Φ_a is a certain arrangement of the elements

$$(81) \qquad \sigma_{ijkla} - \sigma_{ija}\sigma_{kla} .$$

Without any further knowledge about the parent distribution we have to employ v_{ijkla} for studentization. If the normal moment relations (56) are assumed for each vector (70), then (81) becomes $\sigma_{ika}\sigma_{jla} + \sigma_{ila}\sigma_{jka} = \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}$ under the assumption (80) and the hypothesis H'_4. Hence formula (68) can be used. Replacing each σ_{ij} by $\bar{v}_{ij} = \frac{1}{N}\sum_a N_a v_{ija}$, we get a test function $N\Delta_1$ whose limiting distribution when H'_4 is true is the χ^2 distribution with $\frac{1}{2}p(p-1)(k-1)$ degrees of freedom.

A final remark

One of our assumptions on the statistic T in (7) is that the function f is defined in the whole mk-dimensional space. But we give examples in which the functions playing the role of f have less extensive domains of definition. This difficulty may be overcome by observing that we can extend the definition of the functions in question by assigning any constant value, for example zero, as the value of the functions outside their natural domains of definition. The same consideration applies to the functions in sections 19 and 20.

REFERENCES

1. CRAMÉR, H. *Random Variables and Probability Distributions.* Cambridge University Press, 1937.
2. DOOB, J. L. "The limiting distributions of certain statistics," *Annals of Math. Stat.*, vol. 6 (1935), pp. 160–170.
3. HOTELLING, H. "The Generalization of Student's Ratio," *Annals of Math. Stat.*, vol. 2 (1931), pp. 359–378.
4. ———. "Relations between two sets of variates," *Biometrika*, vol. 28 (1936), pp. 321–377.
5. HSU, C. T. "On samples from a normal bivariate population," *Annals of Math. Stat.*, vol. 11 (1940), pp. 410–426.
6. HSU, P. L. "Notes on Hotelling's generalized *T*," *Annals of Math. Stat.*, vol. 9 (1938), pp. 231–243.
7. ———. "The limiting distribution of a general class of statistics," *Science Record (Academia Sinica)*, vol. 1 (1942), pp. 37–41.
8. LAWLEY, D. N. "The estimation of factor loadings by the method of maximum likelihood," *Proc. Roy. Soc. Edinburgh*, vol. 60 (1940), pp. 64–82.
9. NEYMAN, J., and E. S. PEARSON. "On the problem of *k* samples," *Bulletin de l'Academie Polonaise des Sciences et des Lettres*, A (1931), pp. 460–481.
10. WILKS, S. S. "On the independence of *k* sets of normally distributed statistical variables," *Econometrika*, vol. 3 (1935), pp. 309–326.