MATH. SCAND. 1 (1953)

CONTINGENCY TABLES AND APPROXIMATE χ^2-DISTRIBUTIONS

DAVID FOG

Within various domains of mathematical statistics we are concerned with random variables which are approximately χ^2-distributed; for instance in the treatment of contingency tables. It is well known that a variable which is exactly χ^2-distributed with n degrees of freedom, may be split up into a sum of n squares of independent variables z_i, each of which is normally distributed with zero mean and unit variance. In the following we shall show, for a number of cases, that a variable, approximately χ^2-distributed with n degrees of freedom, can be split up in approximately the same way, i.e. into a sum of squares of n variables having approximately the same properties as the z's above. The investigation includes not only two-dimensional contingency tables, but similar tables of an arbitrary number of dimensions, and leads to a decomposition into a sum of squares of uncorrelated variables, having zero mean and variances which only depend on the number N of experiments and tend to 1 for $N \to \infty$. The exact values of the variances are found as functions of N.

1. Multinomial distribution without secondary conditions. As an introduction to the contingency tables we shall start with the multinomial distribution. Consider an experiment, which may give q different results, denoted by the integers 1, 2, ..., q. The corresponding probabilities $p_1, p_2, \ldots, p_q$ are assumed to be known. We consider a series of N independent experiments, N being given. The results Nos. 1, 2, ..., q are assumed to occur $n_1, n_2, \ldots, n_q$ times respectively, so that

$$\sum_1^q n_i = N . \tag{1.1}$$

It is well known that the expression

Received March 20, 1953.

$$\sum_{1}^{q} \frac{(n_i - \nu_i)^2}{\nu_i}, \tag{1.2}$$

$\nu_i = E(n_i)$ being the mean of n_i, is approximately χ^2-distributed with $q-1$ degrees of freedom. Below we shall give a decomposition of (1.2) into a sum of $q-1$ squares with certain specified properties.

As the distribution of the set $(n_1, n_2, \ldots, n_q)$ is an ordinary multinominal distribution, the means, variances, and covariances of the n's are determined by

$$\nu_i = E(n_i) = Np_i, \ \mathrm{var}(n_i) = Np_i(1-p_i), \ \mathrm{cov}(n_i, n_j) = -Np_ip_j, \ (i \neq j). \tag{1.3}$$

Introducing the new variables

$$y_i = \frac{n_i - \nu_i}{\nu_i^{1/2}} \tag{1.4}$$

we get from (1.3)

$$E(y_i) = 0, \ \mathrm{var}(y_i) = 1 - p_i, \ \mathrm{cov}(y_i, y_j) = -(p_ip_j)^{1/2}, \quad (i \neq j), \tag{1.5}$$

and from (1.1)

$$\sum_{1}^{q} p_i^{1/2} y_i = 0. \tag{1.6}$$

Consider an orthogonal substitution of order q

$$\begin{cases} z_m = \sum_{i=1}^{q} h_{mi} y_i, & m = 1, 2, \ldots, q-1, \\ z_q = \sum_{i=1}^{q} p_i^{1/2} y_i, \end{cases} \tag{1.7}$$

the q^{th} linear form being that of (1.6), the others needing no specification. This substitution is applied to the y's introduced in (1.4); from (1.6) follows $z_q = 0$, and owing to the orthogonality of the transformation we get

$$\sum_{1}^{q} y_i^2 = \sum_{1}^{q-1} z_m^2. \tag{1.8}$$

For each variable $z_1, z_2, \ldots, z_{q-1}$ the mean $E(z_m) = 0$. Further we find

$$\mathrm{cov}(y_1, z_m) = h_{m1}(1-p_1) - \sum_{i=2}^{q} h_{mi}(p_1p_i)^{1/2} = h_{m1} - p_1^{1/2} \sum_{i=1}^{q} h_{mi} p_i^{1/2} = h_{m1}$$

or, after replacement of a subscript,

$$\mathrm{cov}(y_i, z_m) = h_{mi}. \tag{1.9}$$

Consequently

$$\text{var}(z_m) = \sum_i h_{mi}^2 = 1 \tag{1.10}$$

and for $m \neq m_1$

$$\text{cov}(z_m, z_{m_1}) = \sum_i h_{mi} h_{m_1 i} = 0 \,. \tag{1.11}$$

So we have obtained the desired result:

The sum of squares

$$\sum_1^q y_i^2 = \sum_1^q \frac{(n_i - \nu_i)^2}{\nu_i}$$

may be written as a sum of squares of $q-1$ *normalized and mutually orthogonal linear forms* $z_1, z_2, \ldots, z_{q-1}$ *given in* (1.7). *These are uncorrelated, with zero mean and unit variance.*

This result is mentioned in a paper by Irwin [1].

For large N the variables $z_1, z_2, \ldots, z_{q-1}$ are approximately normal and, being uncorrelated, they are also approximately independent.

2. Two-dimensional contingency tables. Let u and v be two independent variables assuming the values $u_1, u_2, \ldots, u_q$, and $v_1, v_2, \ldots, v_r$, respectively, with given probabilities. As in § 1 we consider a series of N

	u_1	u_2	...	u_q	
v_1	n_{11}	n_{21}	...	n_{q1}	b_1
v_2	n_{12}	n_{22}	...	n_{q2}	b_2
...	...	...	...	...	...
v_r	n_{1r}	n_{2r}	...	n_{qr}	b_r
	a_1	a_2	...	a_q	N

experiments, N being given. The number of occurrences of the couple (u_i, v_j) is denoted n_{ij}. The variables n_{ij} shown in the accompanying table are connected in a multinomial distribution. Denoting the marginal frequencies (the frequencies of u_i and v_j) by a_i and b_j we have

$$\sum_j n_{ij} = a_i \,, \quad \sum_i n_{ij} = b_j \,. \tag{2.1}$$

As the equation $\sum_{ij} n_{ij} = N$ may be derived from the $q+r$ equations (2.1) in two different ways, only $q+r-1$ of them are linearly independent.

In the following the marginal frequencies a_i, b_j are assumed to be fixed. Under these circumstances we have (Wilks [4], p. 216)

$$\nu_{ij} = E(n_{ij}) = \frac{a_i b_j}{N} \tag{2.2}$$

and

$$\text{var}(n_{ij}) = \frac{a_i b_j}{N}\left(1+\frac{(a_i-1)(b_j-1)}{N-1}-\frac{a_i b_j}{N}\right), \tag{2.3}$$

$$\text{cov}(n_{ij}, n_{i_1 j}) = \frac{a_i a_{i_1} b_j}{N}\left(\frac{b_j-1}{N-1}-\frac{b_j}{N}\right), \tag{2.4}$$

$$\text{cov}(n_{ij}, n_{i_1 j_1}) = \frac{a_i a_{i_1} b_j b_{j_1}}{N}\left(\frac{1}{N-1}-\frac{1}{N}\right). \tag{2.5}$$

In (2.4) and (2.5), as well as in later formulas of the same kind, different subscripts like i and i_1, j and j_1 denote different numbers.

As in § 1 we introduce new variables

$$y_{ij} = \frac{n_{ij}-\nu_{ij}}{\nu_{ij}^{1/2}}. \tag{2.6}$$

The sum of the squares of these variables is known to be approximately χ^2-distributed with $(q-1)(r-1)$ degrees of freedom. In the following we shall give a transformation of this sum of squares analogous to that of (1.8).

From (2.1) we get the equations

$$\sum_j y_{ij} b_j^{1/2} = 0, \quad \sum_i y_{ij} a_i^{1/2} = 0\ , \tag{2.7}$$

among which $q+r-1$ are linearly independent. In view of the following, we note that these equations are equivalent to $q+r-1$ other equations

$$\sum_{ij} h_{mij} y_{ij} = 0\ , \tag{2.8}$$

the left sides of which are normalized and mutually orthogonal.

Immediately we have $E(y_{ij}) = 0$, and further we get from (2.3), (2.4), and (2.5)

$$\text{var}(y_{ij}) = 1+\frac{(a_i-1)(b_j-1)}{N-1}-\frac{a_i b_j}{N}, \tag{2.9}$$

$$\text{cov}(y_{ij}, y_{i_1 j}) = (a_i a_{i_1})^{1/2}\left(\frac{b_j-1}{N-1}-\frac{b_j}{N}\right), \tag{2.10}$$

$$\text{cov}(y_{ij}, y_{i_1 j_1}) = (a_i a_{i_1} b_j b_{j_1})^{1/2}\left(\frac{1}{N-1}-\frac{1}{N}\right). \tag{2.11}$$

Consider now a linear form

$$z = \sum_{ij} h_{ij} y_{ij} \tag{2.12}$$

which is orthogonal to all forms (2.7) and, accordingly, to (2.8). This requirement is equivalent to

$$\sum_i h_{ij} a_i^{1/2} = 0, \quad \sum_j h_{ij} b_j^{1/2} = 0 , \tag{2.13}$$

giving $q+r-1$ linearly independent equations between the coefficients h_{ij}.

The mean of z is zero. For a closer investigation of z we form, using (2.9)-(2.11),

$$\operatorname{cov}(y_{11}, z) = h_{11}\left(1+\frac{(a_1-1)(b_1-1)}{N-1}-\frac{a_1 b_1}{N}\right)$$

$$+\sum_{i=2}^{q} h_{i1}(a_1 a_i)^{1/2}\left(\frac{b_1-1}{N-1}-\frac{b_1}{N}\right)+\sum_{j=2}^{r} h_{1j}(b_1 b_j)^{1/2}\left(\frac{a_1-1}{N-1}-\frac{a_1}{N}\right)$$

$$+\sum_{i=2}^{q}\sum_{j=2}^{r} h_{ij}(a_1 a_i b_1 b_j)^{1/2}\left(\frac{1}{N-1}-\frac{1}{N}\right).$$

By means of (2.13) the two sums in the second line may be converted into

$$-h_{11}a_1\left(\frac{b_1-1}{N-1}-\frac{b_1}{N}\right)-h_{11}b_1\left(\frac{a_1-1}{N-1}-\frac{a_1}{N}\right),$$

while the double-sum in the third line may be reduced to

$$-\sum_{i=2}^{q} h_{i1} b_1 (a_1 a_i)^{1/2}\left(\frac{1}{N-1}-\frac{1}{N}\right) = h_{11} a_1 b_1\left(\frac{1}{N-1}-\frac{1}{N}\right).$$

As the terms with denominator N vanish we get

$$\operatorname{cov}(y_{11}, z) = h_{11}\left(1+\frac{(a_1-1)(b_1-1)-\Sigma' a_1(b_1-1)+a_1 b_1}{N-1}\right), \tag{2.14}$$

Σ' denoting the sum of the term quoted and the term obtained by interchanging the letters a and b, i.e.,

$$\Sigma' a_1(b_1-1) = a_1(b_1-1)+b_1(a_1-1) . \tag{2.15}$$

Now

$$(a_1-1)(b_1-1)-\Sigma' a_1(b_1-1)+a_1 b_1 = 1 , \tag{2.16}$$

and we get from (2.14)

$$\operatorname{cov}(y_{ij}, z) = h_{ij}\left(1+\frac{1}{N-1}\right). \tag{2.17}$$

Hence

$$\operatorname{var}(z) = \sum_{ij} h_{ij}^2 \left(1+\frac{1}{N-1}\right),$$

or, if z is assumed to be normalized,

$$\operatorname{var}(z) = 1+\frac{1}{N-1}. \tag{2.18}$$

Under the conditions (2.13) the forms (2.12) determine a linear space of $qr-(q+r-1)=(q-1)(r-1)$ dimensions. This space may be represented by $(q-1)(r-1)$ normalized, mutually orthogonal forms

$$z_m = \sum_{ij} h_{mij} y_{ij}, \qquad 1 \leqq m \leqq (q-1)(r-1)\,. \tag{2.19}$$

All these forms have the variance (2.18). Further they are uncorrelated, since for two such forms $z=\Sigma_{ij} h_{ij} y_{ij}$ and $z'=\Sigma_{ij} h'_{ij} y_{ij}$ we get by means of (2.17)

$$\operatorname{cov}(z,z') = \sum_{ij} h_{ij} h'_{ij} \left(1+\frac{1}{N-1}\right) = 0\,,$$

owing to the orthogonality.

We now construct an orthogonal substitution of order qr, adding to (2.19) the equations

$$z_m = \sum_{ij} h_{mij} y_{ij}, \quad (q-1)(r-1) < m \leqq qr\,, \tag{2.20}$$

the right sides of which are the forms of (2.8). Applying this substitution to the y's in (2.6), which satisfy (2.7)—and the equivalent system (2.8)—, we get

$$\sum_{ij} y_{ij}^2 = \sum_{m} z_m^2, \quad 1 \leqq m \leqq (q-1)(r-1)\,. \tag{2.21}$$

It has thus been proved:

The sum of qr squares

$$\sum_{ij} y_{ij}^2 = \sum_{ij} \frac{(n_{ij}-\nu_{ij})^2}{\nu_{ij}}$$

may be written as a sum of squares of $(q-1)(r-1)$ *normalized and mutually orthogonal linear forms* (2.19). *These forms are uncorrelated, and all have mean* 0 *and variance* $1+(N-1)^{-1}$.

For large N the forms (2.19) are approximately normal with unit variance and approximately independent.

This problem has been studied by Lancaster [2], who states the variance to be 1, which is only true asymptotically.

3. Three-dimensional contingency tables. Let u, v, and w be three independent variables assuming the values $(u_1, u_2, \ldots, u_q)$, $(v_1, v_2, \ldots, v_r)$, and $(w_1, w_2, \ldots, w_s)$, respectively, with given probabilities. The theory of § 2 may be applied to any two of these variables; however, as the independence of the variables two by two does not imply the independence of all three, it is of interest to study the whole three-dimensional table and not only the two-dimensional marginals.

In many respects the situation is here analogous to that described in § 2, and for such parts of the investigation we may confine ourselves to brief indications. However, the analogy is not perfect; at a certain point something new appears, which makes a more profound treatment of the three- and multidimensional contingency tables necessary.

As before the number of experiments is assumed to be N, and the frequency of the triple (u_i, v_j, w_k) is called n_{ijk}. Denoting the one-dimensional frequencies by a_i, b_j, c_k, respectively, we have

$$\sum_{jk} n_{ijk} = a_i, \quad \sum_{ik} n_{ijk} = b_j, \quad \sum_{ij} n_{ijk} = c_k . \tag{3.1}$$

Only $q+r+s-2$ of these $q+r+s$ equations are linearly independent.

For fixed marginal frequencies a_i, b_j, c_k we find in analogy to (2.2–3)

$$v_{ijk} = E(n_{ijk}) = \frac{a_i b_j c_k}{N^2},$$

$$\operatorname{var}(n_{ijk}) = \frac{a_i b_j c_k}{N^2}\left(1+\frac{(a_i-1)(b_j-1)(c_k-1)}{(N-1)^2}-\frac{a_i b_j c_k}{N^2}\right),$$

and corresponding expressions for the covariances (compare (2.4–5)).

Consider now the variables

$$y_{ijk} = \frac{n_{ijk}-v_{ijk}}{v_{ijk}^{1/2}}. \tag{3.2}$$

Introducing the y's in (3.1) we get the equations

$$\sum_{jk} y_{ijk}(b_j c_k)^{1/2} = 0, \quad \sum_{ik} y_{ijk}(a_i c_k)^{1/2} = 0, \quad \sum_{ij} y_{ijk}(a_i b_j)^{1/2} = 0, \tag{3.3}$$

which are equivalent to $q+r+s-2$ linearly independent equations

$$\sum_{ijk} h_{mijk} y_{ijk} = 0 , \tag{3.4}$$

the left sides of which are normalized and mutually orthogonal.

We have $E(y_{ijk}) = 0$, and corresponding to the formulas (2.9–11) we find

$$\text{(3.5)} \qquad \operatorname{var}(y_{ijk}) = 1 + \frac{(a_i-1)(b_j-1)(c_k-1)}{(N-1)^2} - \frac{a_i b_j c_k}{N^2},$$

$$\text{(3.6)} \qquad \operatorname{cov}(y_{ijk}, y_{i_1jk}) = (a_i a_{i_1})^{1/2}\left(\frac{(b_j-1)(c_k-1)}{(N-1)^2} - \frac{b_j c_k}{N^2}\right),$$

$$\text{(3.7)} \qquad \operatorname{cov}(y_{ijk}, y_{i_1j_1k}) = (a_i a_{i_1} b_j b_{j_1})^{1/2}\left(\frac{c_k-1}{(N-1)^2} - \frac{c_k}{N^2}\right),$$

$$\text{(3.8)} \qquad \operatorname{cov}(y_{ijk}, y_{i_1j_1k_1}) = (a_i a_{i_1} b_j b_{j_1} c_k c_{k_1})^{1/2}\left(\frac{1}{(N-1)^2} - \frac{1}{N^2}\right).$$

We now consider a linear form

$$\text{(3.9)} \qquad z = \sum_{ijk} h_{ijk} y_{ijk}\,,$$

the coefficients h_{ijk} satisfying the equations

$$\text{(3.10)} \qquad \sum_i h_{ijk} a_i^{1/2} = 0, \quad \sum_j h_{ijk} b_j^{1/2} = 0, \quad \sum_k h_{ijk} c_k^{1/2} = 0\,.$$

It is orthogonal to all forms (3.3), and accordingly to (3.4).

The mean of z is 0. As in § 2 it may be shown that

$$\text{(3.11)} \qquad \operatorname{cov}(y_{ijk}, z) = h_{ijk}\left(1 - \frac{1}{(N-1)^2}\right).$$

In the proof we make use of the identity (compare (2.16))

$$\text{(3.12)} \qquad \begin{aligned}(a_1-1)(b_1-1)(c_1-1) - \Sigma' a_1(b_1-1)(c_1-1) \\ + \Sigma' a_1 b_1 (c_1-1) - a_1 b_1 c_1 = -1\,,\end{aligned}$$

Σ' denoting as in § 2 the sum of the term quoted and the terms obtained by permutations of letters. The identity (3.12) is easily extended to ϱ dimensions, the right side then being $(-1)^\varrho$.

From (3.11) we find for a normalized z

$$\text{(3.13)} \qquad \operatorname{var}(z) = 1 - \frac{1}{(N-1)^2}.$$

Under the conditions (3.10) the forms (3.9) determine a linear space of $(q-1)(r-1)(s-1)$ dimensions; this space may be represented by $(q-1)(r-1)(s-1)$ normalized, mutually orthogonal forms

$$\text{(3.14)} \qquad z_m = \sum_{ijk} h_{mijk} y_{ijk}, \quad 1 \leqq m \leqq (q-1)(r-1)(s-1)\,,$$

all with variance (3.13). As in § 2, they are seen to be uncorrelated.

Now we have come to the point, where the discussion departs from that of section 2. In fact, we have not yet a sufficient number of forms for writing down an orthogonal substitution corresponding to that of (2.19–20). Consider, therefore, a linear form

$$z = \sum_{ijk} h_{ij} c_k^{1/2} y_{ijk} , \tag{3.15}$$

the quantities h_{ij} satisfying the equations

$$\sum_i h_{ij} a_i^{1/2} = 0, \quad \sum_j h_{ij} b_j^{1/2} = 0 . \tag{3.16}$$

This form is seen to be orthogonal to all forms (3.3)—and (3.4)—as well as to any form (3.9).

The mean of z in (3.15) is 0. By means of (3.5–8) it is found by a reduction similar to the former ones, but more tedious, that

$$\operatorname{cov}(y_{ijk}, z) = h_{ij} c_k^{1/2} \left(1+\frac{1}{N-1}\right), \tag{3.17}$$

whence for a normalized z

$$\operatorname{var}(z) = 1+\frac{1}{N-1} . \tag{3.18}$$

Further any form (3.15) is uncorrelated with any form (3.9).

Under the conditions (3.16) the forms (3.15) determine a linear space of $(q-1)(r-1)$ dimensions. It may be represented by $(q-1)(r-1)$ normalized, mutually orthogonal forms

$$z_m = \sum_{ijk} h_{mij} c_k^{1/2} y_{ijk} , \tag{3.19}$$

all with variance (3.18) and—as before (§ 1, § 2)—mutually uncorrelated.

Further

$$z = \sum_{ijk} h_{ik} b_j^{1/2} y_{ijk} \tag{3.20}$$

and

$$z = \sum_{ijk} h_{jk} a_i^{1/2} y_{ijk} , \tag{3.21}$$

with suitable conditions on the coefficients, give rise to $(q-1)(s-1)$ and $(r-1)(s-1)$ normalized, mutually orthogonal forms

$$z_m = \sum_{ijk} h_{mik} b_j^{1/2} y_{ijk} \tag{3.22}$$

and

$$z_m = \sum_{ijk} h_{mjk} a_i^{1/2} y_{ijk}\,, \tag{3.23}$$

all with variance (3.18) and mutually uncorrelated. Any form (3.22) or (3.23) is orthogonal to and uncorrelated with all forms (3.9) and (3.15), and is, further, orthogonal to the forms (3.3).

Now we are able to construct the desired orthogonal substitution of order qrs. In (3.14), (3.19), (3.22), and (3.23) we have altogether

$$\begin{aligned}(q-1)(r-1)(s-1)+(q-1)(r-1)+(q-1)(s-1)+(r-1)(s-1) \\ = qrs-(q+r+s)+2\end{aligned} \tag{3.24}$$

normalized and mutually orthogonal forms. The orthogonal substitution is formed by adding to these the $q+r+s-2$ forms

$$z_m = \sum_{ijk} h_{mijk} y_{ijk}\,,$$

constituted by the first members of the equations (3.4). Applying this substitution to the y's in (3.2) we get

$$\sum_{ijk} y_{ijk}^2 = \sum_m z_m^2\,, \quad 1 \leqq m \leqq qrs-(q+r+s)+2\,.$$

It has thus been proved:

The sum of qrs squares

$$\sum_{ijk} y_{ijk}^2 = \sum_{ijk} \frac{(n_{ijk}-\nu_{ijk})^2}{\nu_{ijk}}$$

may be written as a sum of squares of $qrs-(q+r+s)+2$ *normalized and mutually orthogonal linear forms* (3.14), (3.19), (3.22), *and* (3.23). *These forms all have zero mean and are mutually uncorrelated. The* $(q-1)(r-1)(s-1)$ *forms* (3.14) *have variance* $1-(N-1)^{-2}$, *while the remaining forms have variance* $1+(N-1)^{-1}$.

As in § 2, for large N, these forms are approximately normal with unit variance and approximately independent.

4. Multidimensional contingency tables. Now it is evident how to pass to an arbitrary number of dimensions, ϱ. Let the frequencies in N experiments be $n_{i_1 i_2 \cdots i_\varrho}$ and the one-dimensional marginal frequencies $a_{i_1}, b_{i_2}, \ldots, g_{i_\varrho}$, where $1 \leqq i_1 \leqq q_1$, $1 \leqq i_2 \leqq q_2, \ldots, 1 \leqq i_\varrho \leqq q_\varrho$. The mean of $n_{i_1 i_2 \cdots i_\varrho}$ for fixed marginal distributions $a_{i_1}, b_{i_2}, \ldots, g_{i_\varrho}$ then becomes

$$\nu_{i_1 i_2 \cdots i_\varrho} = \frac{a_{i_1} b_{i_2} \cdots g_{i_\varrho}}{N^{\varrho-1}}\,, \tag{4.1}$$

and with

$$y_{i_1 i_2 \cdots i_\varrho} = \frac{n_{i_1 i_2 \cdots i_\varrho} - \nu_{i_1 i_2 \cdots i_\varrho}}{(\nu_{i_1 i_2 \cdots i_\varrho})^{1/2}} \tag{4.2}$$

the result may be formulated as follows:

The sum of $q_1 q_2 \cdots q_\varrho$ squares

$$\sum_{i_1 i_2 \cdots i_\varrho} (y_{i_1 i_2 \cdots i_\varrho})^2,$$

may be written as a sum of squares of $q_1 q_2 \cdots q_\varrho - (q_1 + q_2 + \cdots + q_\varrho) + \varrho - 1$ normalized and mutually orthogonal forms

$$z = \sum_{i_1 i_2 \cdots i_\varrho} h_{i_1 i_2 \cdots i_\varrho} y_{i_1 i_2 \cdots i_\varrho}, \tag{4.3}$$

having zero mean, zero covariances, and variances as stated below:

$$\begin{cases} \Sigma (q_i - 1)(q_j - 1) & \text{forms have variances } 1 + (N-1)^{-1}, \\ \Sigma (q_i - 1)(q_j - 1)(q_k - 1) & \quad ,, \quad ,, \quad ,, \quad 1 - (N-1)^{-2}, \\ \Sigma (q_i - 1)(q_j - 1)(q_k - 1)(q_l - 1) & \quad ,, \quad ,, \quad ,, \quad 1 + (N-1)^{-3}, \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ (q_1 - 1)(q_2 - 1) \cdots (q_\varrho - 1) & \quad ,, \quad ,, \quad 1 + (-1)^\varrho (N-1)^{-(\varrho-1)}. \end{cases} \tag{4.4}$$

The meaning of Σ in (4.4) is most easily explained by an example. For $\varrho = 3$

$$\Sigma (q_i - 1)(q_j - 1) = (q_1 - 1)(q_2 - 1) + (q_1 - 1)(q_3 - 1) + (q_2 - 1)(q_3 - 1).$$

As in the previous cases, for large N the z's in (4.3) are approximately normal with unit variance and approximately independent.

REFERENCES

1. J. O. Irwin, *A note on the subdivision of χ^2 into components*, Biometrika 36 (1949), 130-134.
2. H. O. Lancaster, *The derivation and partition of χ^2 in certain discrete distributions*, Biometrika 36 (1949), 117-129.
3. H. O. Lancaster, *The exact partition of χ^2 and its application to the problem of the pooling of small expectations*, Biometrika 37 (1950), 267-270.
4. S. S. Wilks, *Mathematical statistics*, Princeton, 1950.

THE ROYAL VETERINARY AND AGRICULTURAL COLLEGE, COPENHAGEN, DENMARK