*Reprinted from*
BIOMETRICS
THE BIOMETRIC SOCIETY, Vol. 13, No. 4, December 1957

# PARTITION OF EXPERIMENTAL VECTORS CONNECTED WITH MULTINOMIAL DISTRIBUTIONS

L. C. A. CORSTEN
*Institute for Research on Varieties of Field Crops (I.V.R.O.)*
*Wageningen, The Netherlands*

## 1. *Introduction*

The investigation of variates in contingency tables (which need not be orthogonal) often gives rise to a partition of the experimental result in components each illuminating one aspect of the problem in question. This is here expressed in terms of vectors. Compare Kuiper [9].

Our presentation is perhaps more transparent than previous papers dealing with similar subjects: Fisher [3], Fog [4], Irwin [6], Lancaster [10], [12], [13]. Moreover our method is general; hence it can be applied also to more intricate cases. Our results can be applied in genetics, which in fact motivated our research. This is why we give an introductory section (Sec. 2) on genetics.

## 2. *Genetics*

We consider externally perceptible properties of individuals (plants, animals), like the colour (green or yellow) of seeds, or of eyes (brown or blue) of men. Each property to be considered is determined by genes of only one locus, dominant $A$ and recessive $a$, such that the only possible distinction with respect to that property is dominant ($AA$ or $Aa$) or recessive ($aa$).

We assume that the choice of a paternal gamete by a zygote is stochastically independent of the choice of a maternal gamete. Moreover we assume equal viability for every combination of gametes.

Then it follows that crossing individuals $Aa$ with each other will produce zygotes having the dominant and the recessive form of the property with probabilities $\frac{3}{4}$ and $\frac{1}{4}$ respectively.

We consider a second property and locus with genes $B$ and $b$. If this locus is on a chromosome different from that carrying $A$ and $a$, the dominance or recessivity will be independent for the two considered properties.

Crossing such individuals $Aa/Bb$ with each other yields zygotes with probabilities as mentioned in the following table:

$$\begin{array}{c|cc} & A & a \\ \hline B & \frac{9}{16} & \frac{3}{16} \\[1em] b & \frac{3}{16} & \frac{1}{16} \end{array} \tag{1}$$

where the capitals and lower case denote the dominant and recessive phenotypes respectively. The column and row totals represent the chances of the dominant and the recessive forms of the first and the second property respectively.

If both loci occur on the same chromosome and crossing-over does not take place, the gametes of an individual $Aa/Bb$ are only $AB$ and $ab$ with equal probabilities. Crossing such individuals with each other produces zygotes for which the following table of probabilities appears:

$$\begin{array}{c|cc} & A & a \\ \hline B & \frac{3}{4} & 0 \\ b & 0 & \frac{1}{4} \end{array} \tag{2}$$

In a similar way crossing individuals $Aa/bB$ with gametes $Ab$ and $aB$ yields:

$$\begin{array}{c|cc} & A & a \\ \hline B & \frac{1}{2} & \frac{1}{4} \\ b & \frac{1}{4} & 0 \end{array} \tag{2$'$}$$

In both (2) and (2′) the column and row totals are the same as in (1) and they have the same meaning.

If both loci occur on the same chromosome and crossing-over takes place, the gametes of an individual $Aa/Bb$ (in the coupling phase) are $AB$, $ab$, $aB$ and $Ab$ with probabilities, say, $\frac{1}{2} - \frac{1}{2}W$, $\frac{1}{2} - \frac{1}{2}W$, $\frac{1}{2}W$, $\frac{1}{2}W$ respectively (with $0 \leq W \leq \frac{1}{2}$). Crossing such individuals with each other yields zygotes to which the following table of probabilities corresponds:

$$\begin{array}{c|cc} & A & a \\ \hline B & \frac{1}{4}(W^2 - 2W + 3) & \frac{1}{4}(-W^2 + 2W) \\ b & \frac{1}{4}(-W^2 + 2W) & \frac{1}{4}(W^2 - 2W + 1) \end{array} \tag{3}$$

In a similar way individuals $Aa/bB$ (in the repulsion phase) where crossing-over takes place will have gametes $AB$, $ab$, $aB$, and $Ab$ with probabilities $\frac{1}{2}W$, $\frac{1}{2}W$, $\frac{1}{2} - \frac{1}{2}W$, and $\frac{1}{2} - \frac{1}{2}W$ respectively. The following table of probabilities for zygotes yielded by such gametes corresponds to this situation:

$$\begin{array}{c|cc} & A & a \\ \hline B & \frac{1}{4}(W^2 + 2) & \frac{1}{4}(1 - W^2) \\ b & \frac{1}{4}(1 - W^2) & \frac{1}{4}W^2 \end{array} \qquad (3')$$

In both cases (3) and (3′), the column and row totals are the same as in (1), (2), and (2′). If $W = 0$ (crossing-over does not occur), (3) reduces to (2), and (3′) reduces to (2′): complete linkage. If $W = \frac{1}{2}$ both (3) and (3′) reduce to (1): the two properties are completely independent.

Increasing linkage (i.e. decreasing $W$) appears with respect to tables like (1) as an increase of the numbers in one diagonal and a decrease of the numbers in the other diagonal, row and column totals remaining constant; the contribution from linkage to each cell of the table has the same absolute value.

### 3. *Definition of components of a two by two table*

We consider a scheme of probabilities like (1), (a scheme with probabilities $p_1$ and $q_1$ for rows and $p_2$ and $q_2$ for columns and independence of these probabilities). This scheme (vector) is represented by

$$\begin{bmatrix} p_1p_2 & p_1q_2 \\ q_1p_2 & q_1q_2 \end{bmatrix}, \quad \text{where} \quad p_i + q_i = 1 \qquad (i = 1, 2)$$

and it is called *the basis of the one-dimensional space of levels.*

It is our purpose to compare a scheme of counts arranged in a two by two table:

$$x = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \quad \text{where} \quad \sum_{i=1}^{4} x_i = n,$$

or rather of the relative frequencies $x_i/n$, with such a *level.* The experimental result may suggest that the independence is satisfied, that also the proportion $p_2 : q_2$ is not a bad description, but that the proportion $p_1 : q_1$ is wrong.

Then the following scheme may be more likely:

$$\begin{bmatrix} (p_1 + \alpha)p_2 & (p_1 + \alpha)q_2 \\ (q_1 - \alpha)p_2 & (q_1 - \alpha)q_2 \end{bmatrix}$$

wherein $\alpha$ is a suitably chosen real number. Also in this scheme the probabilities for the columns are $p_2$ and $q_2$ respectively and there is independence between row and column probabilities. The scheme is equal to

$$\begin{bmatrix} p_1p_2 & p_1q_2 \\ q_1p_2 & q_1q_2 \end{bmatrix} + \alpha \begin{bmatrix} p_2 & q_2 \\ -p_2 & -q_2 \end{bmatrix}.$$

We call

$$\alpha \begin{bmatrix} p_2 & q_2 \\ -p_2 & -q_2 \end{bmatrix}$$

a *row effect* and

$$\begin{bmatrix} p_2 & q_2 \\ -p_2 & -q \end{bmatrix}$$

the *basis of the one-dimensional space of row effects.*
Analogously,

$$\begin{bmatrix} p_1 & -p_1 \\ q_1 & -q_1 \end{bmatrix}$$

is called the *basis of the one-dimensional space of column effects.*

It may also happen that the independence is satisfied, but that both proportions $p_1 : q_1$ and $p_2 : q_2$ are wrong, and that at first instance the following scheme seems to be more plausible:

$$\begin{bmatrix} (p_1 + \alpha)(p_2 + \beta) & (p_1 + \alpha)(q_2 - \beta) \\ (q_1 - \alpha)(p_2 + \beta) & (q_1 - \alpha)(q_2 - \beta) \end{bmatrix}. \tag{4}$$

The chances for the rows are $(p_1 + \alpha)$ and $(q_1 - \alpha)$ respectively, those for columns $(p_2 + \beta)$ and $(q_2 - \beta)$ respectively, the independence being maintained. But if, in accordance with the customary practice in the analysis of variance, we postulate additivity of row and column effects, the following vector appears:

$$\begin{aligned} &\begin{bmatrix} p_1p_2 & p_1q_2 \\ q_1p_2 & q_1q_2 \end{bmatrix} + \alpha \begin{bmatrix} p_2 & q_2 \\ -p_2 & -q_2 \end{bmatrix} + \beta \begin{bmatrix} p_1 & -p_1 \\ q_1 & -q_1 \end{bmatrix} \\ &\qquad = \begin{bmatrix} p_1p_2 + \alpha p_2 + \beta p_1 & p_1q_2 + \alpha q_2 - \beta p_1 \\ q_1p_2 - \alpha p_2 + \beta q_1 & q_1q_2 - \alpha q_2 - \beta q_1 \end{bmatrix}. \end{aligned} \tag{5}$$

In this scheme row and column totals are as desired, but the vector (4) is equal to (5) plus

$$\begin{bmatrix} \alpha\beta & -\alpha\beta \\ -\alpha\beta & \alpha\beta \end{bmatrix} = \alpha\beta\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

The remark in Sec. 2—last paragraph—may be expressed as follows: Table (3) can be written as the sum of (1) and

$$\frac{3 - 8W - 4W^2}{16}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

and Table (3′) as the sum of (1) and

$$\frac{4W^2 - 1}{16}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

So it is plausible to consider a vector

$$\gamma\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

as representative of linkage, or of disturbance of independence, or, in terms of the analysis of variance, as interaction. Therefore we call the vector

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

the *basis of the one-dimensional space of interactions*. We remark that in contrast with row and column effects, this basis does not depend on the form of the *level*.

In (4) and (5) we saw that, if both row and column effects are present, and moreover additive, a disturbance of independence with respect to the scheme (4) considered as *level* occurs. The disturbance is

$$\alpha\beta\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

If $\alpha$ and $\beta$ are small, this *interaction* is very small and even absent if $\alpha$ and/or $\beta$ are zero. Then this *interaction* is negligible in comparison with possible interaction from linkage.

With these remarks in mind it will be our problem to determine the coefficients in the following partition:

$$\begin{bmatrix} x_1/n & x_2/n \\ x_3/n & x_4/n \end{bmatrix} = \mu \begin{bmatrix} p_1p_2 & p_1q_2 \\ q_1p_2 & q_1q_2 \end{bmatrix} + \alpha \begin{bmatrix} p_2 & q_2 \\ -p_2 & -q_2 \end{bmatrix} + \beta \begin{bmatrix} p_1 & -p_1 \\ q_1 & -q_1 \end{bmatrix} + \gamma \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{6}$$

where $\mu$, $\alpha$, $\beta$, and $\gamma$ are unknown.

## 4. *The Partition*

The conventional solution of $\mu$, $\alpha$, $\beta$, and $\gamma$ is not simple, because the basis vectors on the righthand side of (6) are in general non-orthogonal. Orthogonality, i.e. ordinary inner products being zero, implies that the components can be obtained from orthogonal projections. Orthogonality can be obtained with a suitable choice of the following equivalent partition:

$$\begin{bmatrix} A\dfrac{x_1}{n} & B\dfrac{x_2}{n} \\ C\dfrac{x_3}{n} & D\dfrac{x_4}{n} \end{bmatrix} = \mu \begin{bmatrix} Ap_1p_2 & Bp_1q_2 \\ Cq_1p_2 & Dq_1q_2 \end{bmatrix} + \alpha \begin{bmatrix} Ap_2 & Bq_2 \\ -Cp_2 & -Dq_2 \end{bmatrix} + \beta \begin{bmatrix} Ap_1 & -Bp_1 \\ Cq_1 & -Dq_1 \end{bmatrix} + \gamma \begin{bmatrix} A & -B \\ -C & D \end{bmatrix}. \tag{7}$$

The four components are orthogonal, if the 6 inner products, of which two turn out to be identical, vanish:

$$\begin{aligned} A^2p_1p_2^2 + B^2p_1q_2^2 - C^2q_1p_2^2 - D^2q_1q_2^2 &= 0 \\ A^2p_1^2p_2 - B^2p_1^2q_2 + C^2q_1^2p_2 - D^2q_1^2q_2 &= 0 \\ A^2p_1p_2 - B^2p_1q_2 - C^2q_1p_2 + D^2q_1q_2 &= 0 \\ A^2p_2 - B^2q_2 + C^2p_2 - D^2q_2 &= 0 \\ A^2p_1 + B^2p_1 - C^2q_1 - D^2q_1 &= 0. \end{aligned}$$

The third equation minus $p_1$ times the fourth equation, and the third equation minus $p_2$ times the fifth equation, yield $C^2 = (q_2/p_2)\,D^2$ and $B^2 = (q_1/p_1)D^2$ respectively. Substitution in the third equation yields: $A^2 = (q_1q_2/p_1p_2)D^2$. If we choose $D^2 = 1/q_1q_2$, we obtain :

$$A^2 = \frac{1}{p_1p_2}, \qquad B^2 = \frac{1}{p_1q_2}, \quad \text{and} \quad C^2 = \frac{1}{q_1p_2},$$

and these values happen to satisfy *all* equations.

Substituting the values for $A$, $B$, $C$, $D$ in (7) we get:

$$\begin{bmatrix} \dfrac{x_1}{n\sqrt{p_1p_2}} & \dfrac{x_2}{n\sqrt{p_1q_2}} \\ \dfrac{x_3}{n\sqrt{q_1p_2}} & \dfrac{x_4}{n\sqrt{q_1q_2}} \end{bmatrix} = \mu\begin{bmatrix} \sqrt{p_1p_2} & \sqrt{p_1q_2} \\ \sqrt{q_1p_2} & \sqrt{q_1q_2} \end{bmatrix}$$

$$+\alpha\begin{bmatrix} \sqrt{\dfrac{p_2}{p_1}} & \sqrt{\dfrac{q_2}{p_1}} \\ -\sqrt{\dfrac{p_2}{q_1}} & -\sqrt{\dfrac{q_2}{q_1}} \end{bmatrix} + \beta\begin{bmatrix} \sqrt{\dfrac{p_1}{p_2}} & -\sqrt{\dfrac{p_1}{q_2}} \\ \sqrt{\dfrac{q_1}{p_2}} & -\sqrt{\dfrac{q_1}{q_2}} \end{bmatrix} + \gamma\begin{bmatrix} \dfrac{1}{\sqrt{p_1p_2}} & -\dfrac{1}{\sqrt{p_1q_2}} \\ -\dfrac{1}{\sqrt{q_1p_2}} & \dfrac{1}{\sqrt{q_1q_2}} \end{bmatrix}$$

or putting: $p_1p_2 = \pi_1$, $p_1q_2 = \pi_2$, $q_1p_2 = \pi_3$, and $q_1q_2 = \pi_4$, we obtain:

$$\begin{aligned} \begin{bmatrix} \dfrac{x_1}{n\sqrt{\pi_1}} & \dfrac{x_2}{n\sqrt{\pi_2}} \\ \dfrac{x_3}{n\sqrt{\pi_3}} & \dfrac{x_4}{n\sqrt{\pi_4}} \end{bmatrix} &= \mu\begin{bmatrix} \sqrt{\pi_1} & \sqrt{\pi_2} \\ \sqrt{\pi_3} & \sqrt{\pi_4} \end{bmatrix} + \frac{\alpha}{\sqrt{p_1q_1}}\begin{bmatrix} \sqrt{\pi_3} & \sqrt{\pi_4} \\ -\sqrt{\pi_1} & -\sqrt{\pi_2} \end{bmatrix} \\ &+ \frac{\beta}{\sqrt{p_2q_2}}\begin{bmatrix} \sqrt{\pi_2} & -\sqrt{\pi_1} \\ \sqrt{\pi_4} & -\sqrt{\pi_3} \end{bmatrix} + \frac{\gamma}{\sqrt{p_1q_1p_2q_2}}\begin{bmatrix} \sqrt{\pi_4} & -\sqrt{\pi_3} \\ -\sqrt{\pi_2} & \sqrt{\pi_1} \end{bmatrix}. \end{aligned} \tag{8}$$

The four new vectors (in brackets) are perpendicular to each other, indeed ($\pi_1\pi_4 = \pi_2\pi_3$). They are moreover unit vectors. The required coefficients are thus simply the inner products of the vector [on the left hand side of (8)] to be projected, and the vector on which it is projected. For example, $\mu$ is obtained by taking the inner products of both sides of (8) with the unit vector of which $\mu$ is the coefficient. In this way we obtain:

$$\mu = \frac{1}{n}(x_1 + x_2 + x_3 + x_4) = 1,$$

$$\begin{aligned} \frac{\alpha}{\sqrt{p_1q_1}} &= \frac{1}{n}\left[x_1\sqrt{\frac{\pi_3}{\pi_1}} + x_2\sqrt{\frac{\pi_4}{\pi_2}} - x_3\sqrt{\frac{\pi_1}{\pi_3}} - x_4\sqrt{\frac{\pi_2}{\pi_4}}\right] \\ &= \frac{1}{n}\left[x_1\sqrt{\frac{q_1}{p_1}} + x_2\sqrt{\frac{q_1}{p_1}} - x_3\sqrt{\frac{p_1}{q_1}} - x_4\sqrt{\frac{p_1}{q_1}}\right] \\ &= \frac{1}{n\sqrt{p_1q_1}}[(x_1 + x_2)q_1 - (x_3 + x_4)p_1], \end{aligned}$$

or

$$\alpha = \frac{(x_1 + x_2)q_1 - (x_3 + x_4)p_1}{n}.$$

In a similar way:

$$\beta = \frac{(x_1 + x_3)q_2 - (x_2 + x_4)p_2}{n}$$

and

$$\gamma = \frac{x_1 q_1 q_2 - x_2 q_1 p_2 - x_3 p_1 q_2 + x_4 p_1 p_2}{n}.$$

Substituting these values in (8), subtracting the first term on the right from both sides and multiplying both sides by $\sqrt{n}$, we obtain:

$$X = \begin{bmatrix} \dfrac{x_1 - n\pi_1}{\sqrt{n\pi_1}} & \dfrac{x_2 - n\pi_2}{\sqrt{n\pi_2}} \\[2ex] \dfrac{x_3 - n\pi_3}{\sqrt{n\pi_3}} & \dfrac{x_4 - n\pi_4}{\sqrt{n\pi_4}} \end{bmatrix}$$

$$= \frac{[(x_1 + x_2)q_1 - (x_3 + x_4)p_1]}{\sqrt{np_1q_1}} \begin{bmatrix} \sqrt{\pi_3} & \sqrt{\pi_4} \\ -\sqrt{\pi_1} & -\sqrt{\pi_2} \end{bmatrix} \tag{9}$$

$$+ \frac{[(x_1 + x_3)q_2 - (x_2 + x_4)p_2]}{\sqrt{np_2q_2}} \begin{bmatrix} \sqrt{\pi_2} & -\sqrt{\pi_1} \\ \sqrt{\pi_4} & -\sqrt{\pi_3} \end{bmatrix}$$

$$+ \frac{(x_1 q_1 q_2 - x_2 q_1 p_2 - x_3 p_1 q_2 + x_4 p_1 p_2)}{\sqrt{np_1q_1p_2q_2}} \begin{bmatrix} \sqrt{\pi_4} & -\sqrt{\pi_3} \\ -\sqrt{\pi_2} & \sqrt{\pi_1} \end{bmatrix}.$$

## 5. *Statistical considerations*

$x$ has a multinomial probability distribution:

$$P(x) = \frac{n!}{x_1! x_2! \cdots x_k!} \pi_1^{x_1} \pi_2^{x_2} \cdots \pi_k^{x_k},$$

where

$$\sum_{i=1}^{k} \pi_i = 1 \quad \text{and} \quad \sum_{i=1}^{k} x_i = n,$$

or

$$P(x) = \frac{\dfrac{(n\pi_1)^{x_1} e^{-n\pi_1}}{x_1!} \times \dfrac{(n\pi_2)^{x_2} e^{-n\pi_2}}{x_2!} \times \cdots \times \dfrac{(n\pi_k)^{x_k} e^{-n\pi_k}}{x_k!}}{\dfrac{n^{\sum x_i} e^{-n}}{(\sum x_i)!}}. \tag{10}$$

In Sec. 4 we had the case $k = 4$ and $E(x_i) = n\pi_i$.

The same probability (10) will be obtained if we formally assert that the $x_i$ have independent Poisson distributions with $\lambda_i = n\pi_i$ and probabilities $\lambda_i^{x_i} e^{-\lambda_i}/x_i!$ under the condition that their sum is equal to $\sum x_i$ ; for the sum of Poisson distributed variables has also a Poisson distribution with parameter $\sum \lambda_i$ (in this case equal to $\sum n\pi_i = n$), and its probability occurs in the denominator of $P(x)$, as is proper in the case of a conditional probability.

The same formal assertion implies that $X_i = (x_i - n\pi_i)/\sqrt{n\pi_i}$ has mean zero and unit variance. If further $n\pi_i$ is sufficiently large (e.g. $> 9$), the distribution of $X_i$ will be approximated by the normal (0, 1) distribution.

The vector $X$ that can be represented as a point in a $k$-dimensional space $K$, has thus approximately a probability density

$$Ce^{-X_1^2/2} \cdot e^{-X_2^2/2} \cdots = Ce^{-X^2/2}$$

however, with the restriction

$$\sum_{i=1}^{k} x_i = n, \quad \text{or} \quad \sum_{i=1}^{k} X_i \sqrt{n\pi_i} + \sum_{i=1}^{k} n\pi_i = n, \quad \text{or} \quad \sum_{i=1}^{k} X_i \sqrt{\pi_i} = 0.$$

In other words $X$ is situated in the $(k-1)$ dimensional subspace of $K$, perpendicular to the vector $(\sqrt{\pi_1}, \sqrt{\pi_2}, \cdots, \sqrt{\pi_k})$.

Thus we find the (known) result that

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

has approximately a $\chi^2$-distribution with $(k-1)$ dimensions (or degrees of freedom).

In Sec. 4 we succeeded in splitting up $X$ (situated in a 3-dimensional space) in three perpendicular components each of which has a special meaning. From the foregoing it follows that, if $E(x_i) = n\pi_i$, the square of every component vector—which is equal to the square of the coefficient belonging to it in (9)—has approximately a $\chi^2$-distribution with one dimension.

In the light of our definitions in Sec. 3, we can even say that the square of the length of the first component has a one-dimensional $\chi^2$-distribution, if only row effect is absent irrespective of whether the other effects are present or not. The same is true for the second component if only column effect is absent, and the same for the third if there are no interaction and not too large simultaneous main effects. In other words, each of these components affords us a specific test criterion for the null hypothesis that there is no row effect, or no column effect, or no interaction which can be tested independent of the validity of the other two hypotheses to a certain extent. The last restriction concerns the simultaneous occurrence of considerable row and column effects which influences the component for interaction. The numerical value of each of these three statistics is thus obtained by splitting up the known test criterion $X^2$ for goodness of fit in three terms, the first of which is equal to

$$\frac{[(x_1 + x_2)q_1 - (x_3 + x_4)p_1]^2}{np_1q_1},$$

the second equal to

$$\frac{[(x_1 + x_3)q_2 - (x_2 + x_4)p_2]^2}{np_2q_2},$$

and the third is the rest:

$$\frac{[x_1q_1q_2 - x_2q_1q_2 - x_3p_1q_2 + x_4p_1p_2]^2}{np_1q_1p_2q_2}.$$

Just as is the case in tests with $X^2$, the critical region in the one-dimensional $\chi^2$-distributions will be a one-sided upper critical region.

*Example* (Fisher [3]). Counting of seedlings of self-fertilized maize, which was heterozygous $Aa/bB$ (i.e. in repulsion phase) with respect to two properties, viz. starchy versus sugary and green versus white, gave the following results:

| | starchy | sugary |
|---|---|---|
| green | 1997 | 904 |
| white | 906 | 32 |

If the properties are independent, then the vector *level* will be equal to

$$\begin{bmatrix} \frac{9}{16} & \frac{3}{16} \\ \frac{3}{16} & \frac{1}{16} \end{bmatrix} \quad \text{with} \quad p_1 = p_2 = \tfrac{3}{4}.$$

The ordinary $\chi^2$-test criterion for these probabilities will be found to be equal to $12.21 + 47.13 + 180.19 = 287.69$ which is significant, e.g. at the 1% level.

In order to investigate the origin of this discrepancy from our expectation (compare case (3′) in Sec. 2), we calculate the squares of the coefficients in (9): corresponding with row effect

$$\frac{[\frac{1}{4}(x_1 + x_2) - \frac{3}{4}(x_3 + x_4)]^2}{n\frac{1}{4}\cdot\frac{3}{4}} = \frac{[x_1 + x_2 - 3(x_3 + x_4)]^2}{3n} = 0.65,$$

with column effect

$$\frac{[x_1 + x_3 - 3(x_2 + x_4)]^2}{3n} = 0.78,$$

with interaction

$$\frac{(x_1 - 3x_2 - 3x_3 + 9x_4)^2}{9n} = 286.27,$$

which are the same as those obtained by Fisher. The sum of these squares is indeed equal to the value of $X^2$. Further, we see that the deviation from our expectation is practically exclusively due to interaction or linkage.

In many cases we have no theoretical indications about the chances for rows and columns. Then (and in the case that row and column effects appear to be present and a further investigation of dependence is wished), the unknown chances are estimated according to the maximum likelihood method assuming independence between row and column chances. These estimates are proportional to the marginal totals and imply that $X^2$ has a one-dimensional $\chi^2$-distribution.

A more elementary way of approaching this problem and its consequence for $X^2$ and for the partition of $X^2$ runs as follows. Consider the set of all possible values of $x/n$ in a $2 \times 2$ table with the independent chances $p_1$ and $p_2$ unknown and with fixed $n$. In the subset where the marginal totals are fixed, the conditional probability distribution of $x/n$ will be:

$$\frac{\dfrac{n!}{x_1!x_2!x_3!x_4!}(p_1p_2)^{x_1}(p_1q_2)^{x_2}(q_1p_2)^{x_3}(q_1q_2)^{x_4}}{\dfrac{n!}{(x_1 + x_2)!(x_3 + x_4)!}p_1^{x_1+x_2}q_1^{x_3+x_4}\dfrac{n!}{(x_1 + x_3)!(x_2 + x_4)!}p_2^{x_1+x_3}q_2^{x_2+x_4}}$$

which appears to be independent of the unknown $p_i$. We may therefore assume any $p_i$ to be true in the multinomial distribution from which the conditional distribution in the chosen subset can be obtained. It is thus permissible to assume that in the original multinomial distribution the $p_i$ are equal to the chosen marginal totals. We saw that the unconditional distribution of $X$ *with the assumed* $p_i$ *as parameters* could be approximated by the three-dimensional normal distribution. By imposing the conditions of the considered subset to this $X$, i.e. that the marginal totals of $x/n$ should be equal to the assumed row and column probabilities, $X$ will be limited to a one-dimensional space. In the subset the conditional distribution of such a $X^2$ will be that of a one-dimensional $\chi^2$. Because this conditional distribution does not depend on the marginal totals defining the subset, it is valid in general. Further, it follows that by assuming the $p_i$ equal to the corresponding observed frequencies, the two components for row and column effects in the partition (9) of such a $X^2$ will vanish, and that by this very choice the square of the component for interaction has a one-dimensional $\chi^2$-distribution.

In this particular case the component for interaction can be reduced by substituting the marginal totals for $p_i$ and $q_i$ to

$$\frac{n(x_1x_4 - x_2x_3)^2}{(x_1 + x_2)(x_3 + x_4)(x_1 + x_3)(x_2 + x_4)}$$

or using: $x_1x_4 - x_2x_3 = x_1(x_1 + x_2 + x_3 + x_4) - (x_1 + x_2)\,(x_1 + x_3)$ to

$$\frac{\left[x_1 - \dfrac{(x_1 + x_2)(x_1 + x_3)}{n}\right]^2}{n\left(\dfrac{x_1 + x_2}{n}\right)\left(\dfrac{x_1 + x_3}{n}\right)\left(\dfrac{x_3 + x_4}{n}\right)\left(\dfrac{x_2 + x_4}{n}\right)}$$

which will be useful later on.

*Example*: If we take the same numbers as in the previous example and we suppose nothing to be known about probabilities, the estimates of expected numbers are

$$\begin{bmatrix} 2193.69 & 707.30 \\ 709.30 & 228.69 \end{bmatrix}$$

and $X^2 = 17.64 + 54.70 + 54.55 + 169.17 = 296.06$, which is much more than the 0.1% point of the one-dimensional $\chi^2$-distribution (10.827). So there is interaction or lack of independence. The fact that the latter $X^2$ turns out to be larger than the former should not surprise us, as maximum likelihood estimators do not in general lead to a minimal value of $\chi^2$. Therefore it is possible that in this example the values $p_1 = p_2 = \frac{3}{4}$ yield a lower value of the goodness of fit criterion $X^2$ than the values of $p_1$ and $p_2$, estimated by the maximum likelihood method, do.

6. $2^n$ *tables*

It is not difficult to generalize the treatment of $2 \times 2$ tables to that of $2^n$ tables. We indicate the generalization by presenting the case $n = 3$. In that case the basis vector for the space of *levels* is defined to be a three-way table with given probabilities for the completely independent rows, columns, and layers ($p_i + q_i = 1$). The probabilities for each of the eight cells are then given by Fig. 1. The probability for row 1 (back face) is $p_1$, for row 2 (front face) $q_1$, for column 1 (left side) $p_2$, for column 2 (right side) $q_2$, for layer 1 (bottom) $p_3$, and for layer 2 (upper face) $q_3$.

We call a vector *row effect* if the sum of this vector and the vector *level* looks like Fig. 2. The probability for back space is $p_1 + \alpha$, for front face $(q_1 - \alpha)$, for left side $p_2$, for right side $q_2$, for bottom $p_3$, for upper face $q_3$, all these being independent.

We choose as a basis vector for *row effects, column effects*, and *layer effects* the vectors represented by Figs. 3, 4, and 5 respectively.

We call a vector *row × column interaction*, if the sum of this vector and the vector *level* looks like Fig. 6. The probabilities in the six faces are equal to those in the vector *level* by itself, so that there is no *main effect* in this sum. The chances in the vertical edges are:

$$\begin{bmatrix} p_1p_2 + \delta & p_1q_2 - \delta \\ q_1p_2 - \delta & q_1q_2 + \delta \end{bmatrix}$$

so that independence between rows and columns is disturbed. Further, the probabilities for layers are independent of those for the other classifications.

We choose as a basis vector for *row × column interaction, row × layer interaction*, and *column × layer interaction* the vectors represented by Figs. 7, 8, and 9 respectively.

Finally, we call a vector *second-order interaction* if the sum of this vector and the vector *level* looks like Fig. 10.

The probabilities in the six faces are equal to those in the vector *level* by itself, so that *main effect* is absent in this sum. The probabilities in the twelve edges are also equal to those in the vector *level* by itself, so that there is no *first-order interaction* in the sum.

But the disturbance of independence between row and column probabilities in the bottom is different from that in the upper face. Similar remarks can be made about row and layer probabilities and about column and layer probabilities. In other words, the relation between any pair of classifications cannot be described without involving the third classification.

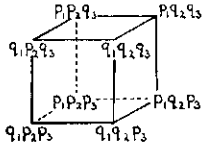


Fig. 1
Level

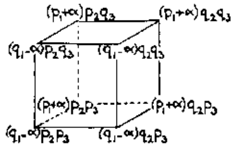


Fig. 2
Level + row effect

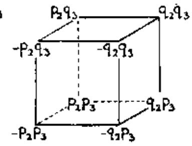


Fig. 3
Basis for row effects

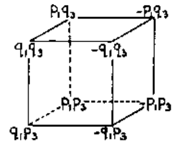


Fig. 4
Basis for column effects

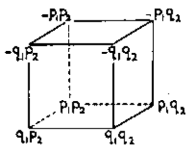


Fig. 5
Basis for layer effects

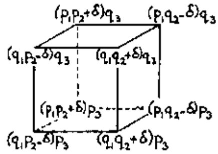


Fig. 6
Level + row × column interaction

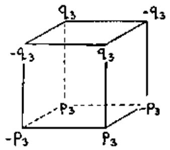


Fig. 7
Basis for row × column interaction

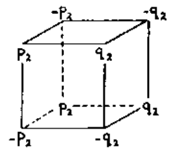


Fig. 8
Basis for row × layer interaction

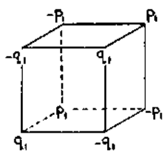


Fig. 9
Basis for column × layer interaction

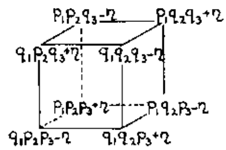


Fig. 10
Level + second order interaction

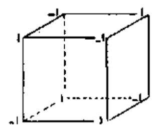

Fig. 11
Basis for second order interaction

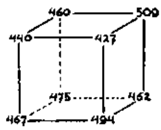


Fig. 12
Example of a $2^3$ table (*Roberts* e.a.)

We choose as a basis vector for *second-order interaction* the vector represented by Fig. 11.

About the thus defined interactions, remarks similar to that in Sec. 3 may be made. If main effects and interactions are present, we have in the place of $p_1p_2p_3$ a chance equal to

$$(p_1 + \alpha)(p_2 + \beta)(p_3 + \gamma) + \delta(p_3 + \gamma) + \epsilon(p_2 + \beta) + \zeta(p_1 + \alpha) + \eta$$
$$= p_1p_2p_3 + \alpha p_2p_3 + \beta p_1p_3 + \gamma p_1p_2 + (\alpha\beta + \delta)p_3 + (\alpha\gamma + \epsilon)p_2$$
$$+ (\beta\gamma + \zeta)p_1 + (\alpha\zeta + \beta\epsilon + \gamma\delta + \alpha\beta\gamma + \eta).$$

From this it will be seen that three main effects together introduce contributions to all of the four interactions. In that case certain first-order interactions may nullify the additional contribution to the second-order interaction. If there are two main effects only, there will be a contribution to the first-order interaction between them. If in this case one or both of the other first-order interactions are nonzero, a contribution to the second-order interaction may take place. If there is only one main effect and simultaneously an interaction between the two other classifications, there will be a contribution to the second-order interaction. If there are no main effects at all, the second-order interaction obtains no contributions from possible first-order interactions. We may conclude that disturbing contributions to interactions may be caused by main effects, but that they will be negligible if $\alpha$, $\beta$, and $\gamma$ are small. The drawback of such contributions is not great as it is not our intention to estimate the different effects, but to investigate the origin of discrepancies between expectation and experimental result.

It will be our purpose to split up a vector of experimental numbers $x_i/n$ ($i = 1, 2, \cdots, 8$; $\sum_i x_i = n$) into eight components in the directions of the defined basis vectors. To facilitate this partition, we again divide the numbers on similar places by the square root of the probability of that place in the vector *level*. Putting those probabilities equal to $\pi_i$, we obtain the following formulation of our problem:

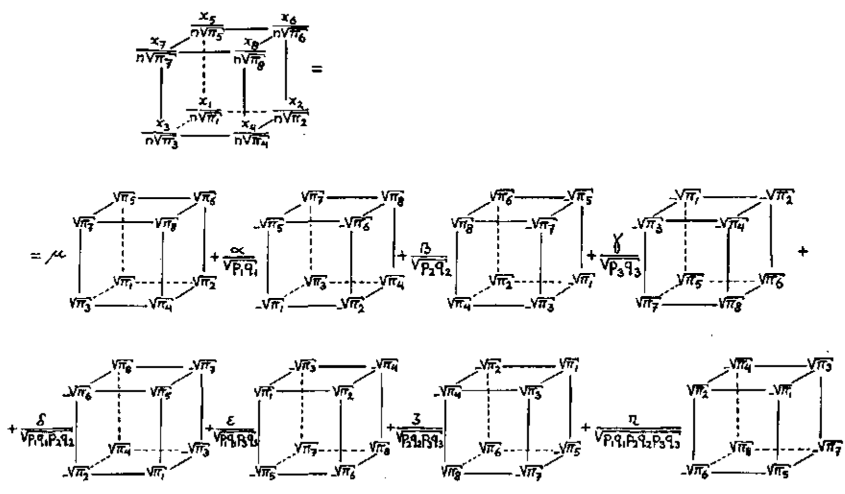

Fig. 13
Equation (11): Partition of $2^3$ table

Calculate the coefficients $\mu$, $\alpha$, $\beta$, $\cdots$ $\eta$ in the equation represented by Fig. 13 which will be called equation (11). All the "cubes" on the right-hand side of this equation are orthogonal unit vectors.

$$\left(\text{Observe that } \frac{\pi_1}{\pi_2} = \frac{\pi_3}{\pi_4} = \frac{\pi_5}{\pi_6} = \frac{\pi_7}{\pi_8} \quad \text{and} \quad \frac{\pi_1}{\pi_3} = \frac{\pi_2}{\pi_4} = \frac{\pi_5}{\pi_7} = \frac{\pi_6}{\pi_8}.\right)$$

The required coefficients are thus again the inner products of the vector on the left side to be projected and the vector on which it is projected:

$$\mu = 1,$$

$$\alpha = \frac{(x_1 + x_2 + x_5 + x_6)q_1 - (x_3 + x_4 + x_7 + x_8)p_1}{n},$$

$$\beta = \frac{(x_1 + x_3 + x_5 + x_7)q_2 - (x_2 + x_4 + x_6 + x_8)p_2}{n},$$

$$\gamma = \frac{(x_1 + x_2 + x_3 + x_4)q_3 - (x_5 + x_6 + x_7 + x_8)p_3}{n},$$

$$\delta = \frac{(x_1 + x_5)q_1q_2 - (x_2 + x_6)q_1p_2 - (x_3 + x_7)p_1p_2 + (x_4 + x_8)p_1p_2}{n},$$

$$\epsilon = \frac{(x_1 + x_2)q_1q_3 - (x_3 + x_4)p_1q_3 - (x_5 + x_6)q_1p_3 + (x_7 + x_8)p_1p_3}{n},$$

$$\zeta = \frac{(x_1 + x_3)q_2q_3 - (x_2 + x_4)p_2q_3 - (x_5 + x_7)q_2p_3 + (x_6 + x_8)p_2p_3}{n},$$

$$\eta = \frac{x_1q_1q_2q_3 - x_2q_1p_2q_3 - x_3p_1q_2q_3 + x_4p_1p_2q_3}{n}$$

$$- \frac{x_5q_1q_2p_3 - x_6q_1p_2p_3 - x_7p_1q_2p_3 + x_8p_1p_2p_3}{n}.$$

Substituting these values in equation (11), Fig. 13, subtracting the first term on the fight from both sides and multiplying on both sides by $\sqrt{n}$, we get on the left a vector $X$ consisting of numbers $(x_i - n\pi_i)/\sqrt{n\pi_i}$ the square of which is again equal to the well-known goodness of fit criterion. $X^2$ has approximately a $\chi^2$-distribution with 7 degrees of freedom, if the null hypothesis that the vector *level* contains the probabilities for each cell is true. Further, the seven components of $X$ on the right are perpendicular, and $X^2$ is equal to the sum of the seven squares of these components. Each of these squares has under the null hypothesis a $\chi^2$-distribution with one degree of freedom.

Now again every component can serve as a statistic for testing a specific hypothesis. With the first three components we test $\alpha = 0$, $\beta = 0$, and $\gamma = 0$ respectively, with the following three components, the hypotheses $\delta = 0$, $\epsilon = 0$, and $\zeta = 0$, under the condition that $\alpha\beta$, $\alpha\gamma$, and $\beta\gamma$ are negligible respectively, and with the last component, the hypothesis $\eta = 0$, under the condition that $\alpha\zeta + \beta\epsilon + \gamma\delta + \alpha\beta\gamma$ is negligible. Each of these hypotheses is tested on the assumption of independence of the three classifications.

The mentioned restrictions do not trouble us if no theoretical chances are available. Generalizing the considerations at the end of Sec. 5, we take the $p_i$ equal to marginal totals with the consequence that $\alpha = \beta = \gamma = 0$. If any product $\alpha\beta$, $\alpha\gamma$, or $\beta\gamma$ appears to be not negligible and a further investigation of interactions is required, it is recommended to take the relative $p_i$ from marginal totals.

A more difficult situation will arise if a first-order interaction appears to be considerable or if it is expected in advance that any pair of classifications is not independent. This situation is obviously contrary to the assumption of independence of the three classifications and will be considered in Sec. 8.

*Example*: (Roberts, Dawson, and Madden [14].) Crossing mice *AaBbCc* with *aabbcc* gave numbers represented in Fig. 12, (*A* in back face, *B* in left side, and *C* in bottom). In the level $p_i$ was $\frac{1}{2}$ ($i = 1, 2, 3$).

The value of $\chi^2$ for row × layer interaction, e.g. is

$$\frac{\epsilon^2}{np_1q_1p_3q_3} = \frac{(x_1 + x_2 + x_7 + x_8 - x_3 - x_4 - x_5 - x_6)^2}{n},$$

and for second-order interaction is

$$\frac{\eta^2}{np_1q_1p_2q_2p_3q_3} = \frac{(x_1 + x_4 + x_6 + x_7 - x_2 - x_3 - x_5 - x_8)^2}{n}.$$

| | | |
|---|---|---|
| The seven values for $\chi^2$ are: | for row effect | 1.63 |
| | column effect | 0.67 |
| | layer effect | 1.03 |
| | row × column interaction | 0.13 |
| | row × layer interaction | 4.25 |
| | column × layer interaction | 0.13 |
| | second-order interaction | 2.79 |
| | total | 10.63 |

The one-dimensional $\chi^2$ has at the 5% level of significance the critical value: 3.84; the 7-dimensional $\chi^2$: 14.07. The total $\chi^2$ is not significant. Concluding that the row × layer interaction is significant, would be rash: the probability that at least one of seven one-dimensional $\chi^2$ is larger than 4.25 is about 0.24. So there is only a slight indication that linkage between *A* and *C* may exist.

It is not always possible to attach a simple meaning to a second-order interaction, but in cases like this it could be caused by the fact that one of the eight gene combinations, as a consequence of diminished viability, is much less frequent than is expected on account of main effects and first-order interactions only.

About the calculations we can remark the following: $\chi^2$ for row $\times$ column interaction can be determined as the interaction $\chi^2$ in the two by two row-column table which can be formed by adding along the four vertical edges; the formula for $\delta$ is then the same as that for $\gamma$ in a two by two table. Thus one can calculate this $\chi^2$ again from the test criterion $X^2$ for this two by two table and by subtraction of the main effects for rows and columns. The second-order interaction $\chi^2$ is found as the difference of the test criterion $X^2$ for the $2^3$ table and the sum of those for main effects and first-order interactions.

If the probabilities $p_i$ are not known, we take them such that the main effects equal zero. In our example we then obtain the value $\chi^2$ for

| | |
|---|---|
| row $\times$ column interaction | 0.12 |
| row $\times$ layer interaction | 4.34 |
| column $\times$ layer interaction | 0.14 |
| second-order interaction | 2.69 |
| total | 7.29 |

The 4-dimensional $\chi^2$ has at the 5% level of signicance a critical value: 9.49. The total $\chi^2$ is not significant. The probability that at least one of four one-dimensional $\chi^2$ is larger than 4.34 is about 0.14; so there is a slight indication that linkage between $A$ and $C$ may exist.

An example of the partition of a $2^5$ table in 31 components is given by Haldane [5].

## 7. $m \times n \times \cdots$ *tables*

In some particular cases a partition of $m \times n \times \cdots$ tables (and even of non-orthogonal tables) may have sense. We indicate the case of a $2 \times 3$ table which can be generalized easily.

An inquiry into the attitude with respect to a political proposal may be summarized in a table of experimental counts:

| | for | against | no opinion |
|---|---|---|---|
| men | $x_1$ | $x_2$ | $x_3$ |
| women | $x_4$ | $x_5$ | $x_6$ |

In this case an appropriate definition of a vector *level* will be:

$$\begin{bmatrix} p_1p_2p_3 & p_1p_2q_3 & p_1q_2 \\ q_1p_2p_3 & q_1p_2q_3 & q_1q_2 \end{bmatrix}, \quad \text{where} \quad p_i + q_i = 1, \qquad (i = 1, 2, 3).$$

The chances for rows (men and women) are $p_1$ and $q_1$ respectively, those for columns 1 and 2 together (politically interested), and for column 3 (politically not interested), $p_2$ and $q_2$ respectively, and those of columns 1 and 2, $p_2p_3$ and $p_2q_3$ (i.e. $p_3$ and $q_3$ under the condition of being in column 1 or 2). Moreover, all these probabilities are independent. Without further explanation we define as *basis vector* for *row effect*:

$$\begin{bmatrix} p_2p_3 & p_2q_3 & q_2 \\ -p_2p_3 & -p_2q_3 & -q_2 \end{bmatrix}.$$

*Column effects* will have two basis vectors, one corresponding with modifications of $p_2$ and $q_2$, independence being maintained:

$$\begin{bmatrix} p_1p_3 & p_1q_3 & -p_1 \\ q_1p_3 & q_1q_3 & -q_1 \end{bmatrix}, \tag{12}$$

and one corresponding with similar modifications of $p_3$ and $q_3$:

$$\begin{bmatrix} p_1p_2 & -p_1p_2 & 0 \\ q_1p_2 & -q_1p_2 & 0 \end{bmatrix} \text{ or rather } \begin{bmatrix} p_1 & -p_1 & 0 \\ q_1 & -q_1 & 0 \end{bmatrix}. \tag{13}$$

If two such column effects occur together, viz. $\beta$ times (12) and $\gamma$ times (13) and, if they are additive, we obtain:

$$\begin{bmatrix} p_1(p_2+\beta)\left(p_3+\frac{\gamma}{p_2}\right)-\frac{\beta\gamma p_1}{p_2} & p_1(p_2+\beta)\left(q_3-\frac{\gamma}{p_2}\right)+\frac{\beta\gamma p_1}{p_2} & p_1(q_2-\beta) \\ q_1(p_2+\beta)\left(p_3+\frac{\gamma}{p_2}\right)+\frac{\beta\gamma q_1}{p_2} & q_1(p_2+\beta)\left(q_3-\frac{\gamma}{p_2}\right)-\frac{\beta\gamma q_1}{p_2} & q_1(q_2-\beta) \end{bmatrix}.$$

This is:

$$\frac{\beta\gamma}{p_2}\begin{bmatrix} p_1 & -p_1 & 0 \\ q_1 & -q_1 & 0 \end{bmatrix}$$

less than a level with probabilities $p_1$, $p_2 + \beta$, $p_3 + \gamma/p_2$, etc. With respect to this level there is a deficit of $\beta\gamma/p_2$ times the column effect (13). As $\beta$ will be small in comparison with $p_2$, $\beta\gamma/p_2$, however, will be negligible in comparison with $\gamma$, and even will be zero, if $\beta$ or $\gamma$ is zero. Similar remarks can be made about simultaneous occurrence of additive row and column effects; there will be disturbances of independence which will be negligible in comparison with proper interactions. We choose as *basis vectors for interactions*:

$$\begin{bmatrix} p_3 & q_3 & -1 \\ -p_3 & -q_3 & 1 \end{bmatrix}$$

which represents the disturbance of independence between row probabilities, $p_1$ and $q_1$, and column probabilities, $p_2$ and $q_2$, and

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \end{bmatrix}$$

which represents a similar disturbance of independence between row probabilities and column probabilities, $p_3$ and $q_3$. It will be our purpose again to split up a vector of experimental numbers $x_i/n$ $(i = 1, 2, \cdots 6;$ $\sum x_i = n)$ in six components in the directions of the defined basis vectors. We divide again the numbers on similar places by the square root of the probability of that place in the vector *level*. Putting those probabilities equal to $\pi_i$, we obtain:

$$\begin{aligned}
\begin{bmatrix} \dfrac{x_1}{n\sqrt{\pi_1}} & \dfrac{x_2}{n\sqrt{\pi_2}} & \dfrac{x_3}{n\sqrt{\pi_3}} \\[2ex] \dfrac{x_4}{n\sqrt{\pi_4}} & \dfrac{x_5}{n\sqrt{\pi_5}} & \dfrac{x_6}{n\sqrt{\pi_6}} \end{bmatrix} &= \mu \begin{bmatrix} \sqrt{\pi_1} & \sqrt{\pi_2} & \sqrt{\pi_3} \\ \sqrt{\pi_4} & \sqrt{\pi_5} & \sqrt{\pi_6} \end{bmatrix} \\
&+ \frac{\alpha}{\sqrt{p_1 q_1}} \begin{bmatrix} \sqrt{\pi_4} & \sqrt{\pi_5} & \sqrt{\pi_6} \\ -\sqrt{\pi_1} & -\sqrt{\pi_2} & -\sqrt{\pi_3} \end{bmatrix} \\
&+ \frac{\beta}{\sqrt{p_2 q_2}} \begin{bmatrix} \sqrt{p_1 q_2 p_3} & \sqrt{p_1 q_2 q_3} & -\sqrt{p_1 p_2} \\ \sqrt{q_1 q_2 p_3} & \sqrt{q_1 q_2 q_3} & -\sqrt{q_1 p_2} \end{bmatrix} \\
&+ \frac{\gamma}{\sqrt{p_2 p_3 q_3}} \begin{bmatrix} \sqrt{p_1 q_3} & -\sqrt{p_1 p_3} & 0 \\ \sqrt{q_1 q_3} & -\sqrt{q_1 p_3} & 0 \end{bmatrix} \\
&+ \frac{\delta}{\sqrt{p_1 q_1 p_2 q_2}} \begin{bmatrix} \sqrt{q_1 q_2 p_3} & \sqrt{q_1 q_2 q_3} & -\sqrt{q_1 p_2} \\ -\sqrt{p_1 q_2 p_3} & -\sqrt{p_1 q_2 q_3} & \sqrt{p_1 p_2} \end{bmatrix} \\
&+ \frac{\epsilon}{\sqrt{p_1 q_1 p_2 p_3 q_3}} \begin{bmatrix} \sqrt{q_1 q_3} & -\sqrt{q_1 p_3} & 0 \\ -\sqrt{p_1 q_3} & \sqrt{p_1 p_3} & 0 \end{bmatrix}.
\end{aligned} \tag{14}$$

The six new vectors (in brackets) on the right are orthogonal unit vectors. The coefficients are obtained again by taking inner products:

$$\mu = 1,$$

$$\alpha = \frac{1}{n}[(x_1 + x_2 + x_3)q_1 - (x_4 + x_5 + x_6)p_1],$$

$$\beta = \frac{1}{n}[(x_1 + x_2 + x_4 + x_5)q_2 - (x_3 + x_6)p_2],$$

$$\gamma = \frac{1}{n}[(x_1 + x_4)q_3 - (x_2 + x_5)p_3 ,$$

$$\delta = \frac{1}{n}[(x_1 + x_2)q_1q_2 - x_3q_1p_2 - (x_4 + x_5)p_1q_2 + x_6p_1p_2],$$

$$\epsilon = \frac{1}{n}[x_1q_1q_3 - x_2q_1p_3 - x_4p_1q_3 + x_5p_1p_3].$$

Substituting these values in equation (14), subtracting the first term on the right from both sides and multiplying on both sides by $\sqrt{n}$, we get on the left a vector $X$ consisting of numbers $(x_i - n\pi_i)/\sqrt{n\pi_i}$ , the square of which is again equal to the goodness of fit criterion. Under the null hypothesis that the vector *level* is true, the square of each term on the right has a $\chi^2$-distribution with one degree of freedom. It will be remarked that the $\chi^2$ for row effect, the first column effect, and the first interaction can be obtained as the $\chi^2$ for the similar components in a $2 \times 2$ table which is deduced from our table by amalgamating the columns 1 and 2. The $\chi^2$ for the second column effect and the second interaction are equal to

$$\frac{[(x_1 + x_4)q_3 - (x_2 + x_5)p_3]^2}{p_3q_3 \cdot np_2} \quad \text{and} \quad \frac{(x_1q_1q_3 - x_2q_1p_3 - x_4p_1q_3 + x_5p_1p_3)^2}{p_1q_1p_3q_3 \cdot np_2}$$

respectively. They will be obtained as the $\chi^2$ for column effect and interaction in the $2 \times 2$ table consisting of the first and second column [see equation (9)], with the restriction, however, that we do not use the experimental total, $x_1 + x_2 + x_4 + x_5$ , in the denominator, but the expected total $np_2$ .

If there are no theoretical values $p_i$ , then we take them such that main effects are absent, i.e. from marginal totals (these are also maximum likelihood estimations). Then the $\chi^2$ for the first interaction will be found as the ordinary $\chi^2$ test criterion for the $2 \times 2$ table obtained by amalgamating the first two columns and the $\chi^2$ for the second interaction as the $\chi^2$ *for interaction* in a $2 \times 2$ table consisting of the first two columns, but with row and column probabilities estimated from the marginal totals of the whole $2 \times 3$ table. In this case $np_2$ is of course taken equal to $x_1 + x_2 + x_4 + x_5$ .

*Example*: The inquiry mentioned in the beginning of this section may have the following result:

| | for | against | no opinion |
|---|---|---|---|
| men | 1154 | 475 | 243 |
| women | 1083 | 442 | 362 |

(These numbers are taken from *Introduction to the Theory of Statistics* by A. M. Mood, page 273, where they occur as an example of a $2 \times 3$ contingency table.) Theoretical probabilities being absent, we calculate the two-dimensional $\chi^2$ for the whole table with the aid of expected values obtained from marginal totals:

| | | |
|---|---|---|
| 1114.04 | 456.67 | 301.29 |
| 1122.96 | 460.33 | 303.71 |

and we find 26.78. The critical value of a two-dimensional $\chi^2$ is 9.21 at the 1% level. So we conclude that there is association in our numbers. In order to investigate the origin of this association, we calculate the $\chi^2$ for the first interaction, which is found as the ordinary $\chi^2$ from the amalgamated table:

| | opinion | no opinion |
|---|---|---|
| men | 1629 | 243 |
| women | 1525 | 362 |

and with expected values also obtained by amalgamation of the calculated expected values:

$$\begin{bmatrix} 1570.71 & 301.29 \\ 1583.29 & 303.71 \end{bmatrix}$$

and we find 26.77. By subtraction it follows that $\chi^2$ for the second interaction is equal to 0.01. We conclude therefore that there was a difference in interest between the sexes, but that no difference in attitude against the proposal could be detected between men and women.

In general each one-dimensional $\chi^2$ in a contingency table, where the probabilities for rows and columns are estimated from the data and where a similar partition takes place, is connected with one of the $2 \times 2$ tables which are obtained by successive amalgamations of the data. Such a $\chi^2$ is calculated as the ordinary $\chi^2$ for *interaction* (as derived in Sec. 5) in this $2 \times 2$ table of which the table of expected values is obtained by a similar amalgamation of the *complete* scheme of expected values (estimated by the maximum likelihood method). The last results where theoretical chances are unknown are the same as those of Lancaster [10] and Kimball [8]. In connection with the foregoing it may be remarked that the method, suggested by Lancaster [11] as an exact one, for calculating $\chi^2$ in a contingency table where cells with small expectations are pooled, does not seem to be correct.

When pooling of cells takes place it is not correct to say only that one or more one-dimensional $\chi^2$ reduced to zero. If, e.g. in a $3 \times 3$ table, two cells in the same row are pooled, in other words are conceived as one cell, these cells do not contribute information about the estimation of the probabilities for the columns to which these cells belonged before pooling. For the vector level can be described as:

<table>
<tr><td>$p_1p_2p_3$</td><td>$p_1p_2q_3p_4$</td><td>$p_1p_2q_3q_4$</td></tr>
<tr><td>$p_1q_2p_3$</td><td>$p_1q_2q_3p_4$</td><td>$p_1q_2q_3q_4$</td></tr>
<tr><td>$q_1p_3$</td><td colspan="2">$q_1q_3$</td></tr>
</table>

with $p_i + q_i = 1$ $(i = 1, 2, 3, 4)$.

The estimation of $p_1$ and $p_2$ takes place from row totals in the same way as in the $3 \times 3$ table. Similarly, the $p_3$ will be estimated from the totals of the first column and the total of the second and the third column together. But $p_4$ will be estimated from the proportion of the totals in what has been left of the second and the third column, namely in the first and the second row. In the expressions for several of the one-dimensional chi-squares that do not vanish by pooling, the (estimated) values of $p_4$ and $q_4$ will in general be different from the values obtained by estimation in the complete $3 \times 3$ table.

The same conclusion follows from the fact that as a consequence of such pooling in the complete partition of a $3 \times 3$ table according to main effects and interactions, the basis of column effect with respect to $p_4$ will coincide with that for interaction in the $2 \times 2$ table formed by the second and third column on the one hand, and by the sum of the first and second row and the third row on the other hand. Thus not only this interaction but also the mentioned main effect should vanish in the partition after pooling, in order that $X^2$, which measures the discrepancy with respect to a level, as represented here, contain three interaction components only.

## 8. *Further remarks about $2^3$ tables*

Several authors (Kendall [7], Simpson [15]) warn against amalgamating $2^3$ tables to $2 \times 2$ tables, even when second-order interaction happens to be absent. They demonstrate the possibility:

(a) that interactions, which in each of the amalgamated classes separately tend in the same direction, seem to be absent after amalgamation, or

(b) that the amalgamated classes separately do not show dependence between the two other properties, but that they do together. In our opinion this warning is exaggerated in many cases and danger threatens from another direction. In our view, as has been shown in Sec. 6, dependence between two classifications in a $2^3$ table will just be tested in a $2 \times 2$ table obtained by amalgamation in the $2^3$ table, irrespective of whether second-order interaction is present or not.

While second-order interaction is absent, the case (a) may be constructed by adding a level, a row $\times$ layer interaction, and a column $\times$ layer interaction. In the bottom of the table this appears as adding of a level and suitable row and column effects in a $2 \times 2$ table so that a small interaction occurs in it (see Sec. 3). In a similar way a small interaction which will have the same direction appears in the upper face. However, in the $2 \times 2$ table obtained by amalgamation of bottom and upper face, interaction will be absent.

Also, while second-order interaction is absent, the case (b) may be constructed in the following way. First we form the sum of a level with $p_3 = \frac{1}{2}$, a row $\times$ layer interaction, and a column $\times$ layer interaction. Because the interactions in the bottom and in the upper face of this sum are identical, we can add a row $\times$ column interaction to it such that both interactions in the bottom and in the upper face separately vanish. The $2 \times 2$ table obtained by amalgamation of bottom and upper face shows row $\times$ column interaction of course. These disturbing interactions, however, will be negligible in comparison with proper interactions due to dependence.

The danger to which we alluded consists in maintaining the hypothesis of independence in the model of a $2^3$ table and the corresponding partition, although a first-order interaction appears to be considerable, or independence between one or more classifications can be expected to be impossible in advance. For the discussed $\chi^2$-test and the partition in a $2^3$ table—and with this remark we proceed on what we said in Sec. 6—are only justified if the three classifications are independent.

In a case, e.g. like the example given by Simpson [15] where the result of a treatment against a disease is investigated by counting *dead* and *alive* in males and females, it is not obvious that the probability of being treated is the same for males and females.

If *one first-order interaction* must be taken in account, we have to choose a new model to test other interactions. The new model for this case is that for a $2 \times 4$ table. Let the experimental result multiplied by $n = \sum_{i=1}^{8} x_i$ be:

| | treated | | not treated | |
|---|---|---|---|---|
| | male | female | male | female |
| alive | $x_1$ | $x_2$ | $x_5$ | $x_6$ |
| dead | $x_3$ | $x_4$ | $x_7$ | $x_8$ |

After the choice of a level similar to that in the foregoing section:

$$\begin{bmatrix} p_1p_2p_3p_4 & p_1p_2q_3p_4 & p_2q_2p_4 & q_1p_4 \\ p_1p_2p_3q_4 & p_1p_2q_3q_4 & p_1q_2q_4 & q_1q_4 \end{bmatrix},$$

the basis vectors for interactions may be:

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} p_3 & q_3 & -1 & 0 \\ -p_3 & -q_3 & 1 & 0 \end{bmatrix},$$

and

$$\begin{bmatrix} p_2p_3 & p_2q_3 & q_2 & 1 \\ -p_2p_3 & -p_2q_3 & -q_2 & -1 \end{bmatrix}.$$

For convenience we put the $p_i$ as unknown for the rest of this section so that no main effects are present.

With the mentioned example in mind we prefer to define the vector level as:

$$\begin{bmatrix} p_1p_2p_4 & p_1q_2p_4 & q_1p_3p_4 & q_1q_3p_4 \\ p_1p_2q_4 & p_1q_2q_4 & q_1p_3q_4 & q_1q_3q_4 \end{bmatrix},$$

which is the same as that for a $2^3$ table if $p_2 = p_3$. A basis for interactions may be formed by:

$$\begin{bmatrix} p_2 & q_2 & -p_3 & -q_3 \\ -p_2 & -q_2 & p_3 & q_3 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{bmatrix},$$

and (15)

$$\begin{bmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

The first of these three vectors represents dependence of death rate from treatment, the second and the third represent interaction of sex and death rate within the treated and within the not-treated individuals respectively. If one prefers to consider dependence between sex and death rate and further dependence of death rate from the treatment within the sexes separately, one should only transpose the second and third column in each of the four vectors. We proceed in the former version.

The last pair of vectors can be replaced in several ways by two other vectors, the one representing a common interaction between sex and death rate in both treated and not-treated individuals, the other representing a difference in dependence of death rate from sex between treated and not-treated individuals, and, moreover, such that a partitioning of $X^2$ in independent components corresponds to this choice. We prefer:

$$\begin{bmatrix} p_1p_2q_2 & -p_1p_2q_2 & q_1p_3q_3 & -q_1p_3q_3 \\ -p_1p_2q_2 & p_1p_2q_2 & -q_1p_3q_3 & q_1p_3q_3 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}.$$

The first reason for this preference is that the mentioned difference in dependence—which may be called second-order interaction—is represented by the same basis vector as in the $2^3$ table. A second reason will be given in the treatment of the following case. A third reason will appear later on. The component of $X^2$ corresponding to the first vector of (15) can be calculated as the test criterion in the $2 \times 2$ table obtained by neglecting sex. The computation of the other components will be treated in the following.

It is also possible that *two first-order interactions* must be taken in account. Let us assume also that an interaction between treatment and death rate exists in the discussed $2 \times 4$ table. We take as a new level:

$$\begin{bmatrix} p_1p_2p_4 & p_1q_2p_4 & q_1p_3p_5 & q_1q_3p_5 \\ p_1p_2q_4 & p_1q_2q_4 & q_1p_3q_5 & q_1q_3q_5 \end{bmatrix}$$

which is the same as that for a $2 \times 4$ table if $p_4 = p_5$ and as that for a $2^3$ table if, moreover, $p_2 = p_3$. This level corresponds to that for two separate and independent $2 \times 2$ tables. A basis for interactions may be formed by the last two vectors of the set given by (15).

This pair can be replaced again in several ways by another pair of vectors which express a common interaction between sex and death rate in both $2 \times 2$ tables, and a difference between such interactions respectively, and which admit a partition of $X^2$ in independent components. We choose:

$$\begin{bmatrix} p_1p_2q_2p_4q_4 & -p_1p_2q_2p_4q_4 & q_1p_3q_3p_5q_5 & -q_1p_3q_3p_5q_5 \\ -p_1p_2q_2p_4q_4 & p_1p_2q_2p_4q_4 & -q_1p_3q_3p_5q_5 & q_1p_3q_3p_5q_5 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}.$$

The difference in interaction is expressed again by the same vector as in the previous cases. The component of $X^2$ corresponding to the first vector of this pair is equal to:

$$\frac{(q_2q_4x_1 - p_2q_4x_2 - q_2p_4x_3 + p_2p_4x_4 + q_3q_5x_5 - p_3q_5x_6 - q_3p_5x_7 + p_3p_5x_8)^2}{n(p_1p_2q_2p_4q_4 + q_1p_3q_3p_5q_5)}.$$

According to the end of Sec. 5, this may be reduced to:

$$\frac{(x_1 - np_1p_2p_4 + x_5 - nq_1p_3p_5)^2}{np_1p_2q_2p_4q_4 + nq_1p_3q_3p_5q_5}.$$

As we know that a quantity like

$$\frac{x_1 - np_1p_2p_4}{\sqrt{np_1p_2q_2p_4q_4}}$$

(with $n$ sufficiently large) has a standard normal distribution, we may consider $x_1$ as a normally distributed variable with expectation $np_1p_2p_4$ and variance $np_1p_2q_2p_4q_4$. The component is thus the square of a normally (0, 1) distributed combination of two normally (0, 1) distributed variables of the considered kind and such that the numbers $np_1$ and $nq_1$ have some weight in this combination, indeed, but not too much. This balanced combination of statistics for testing independence in $2 \times 2$ tables, which is also recommended by van Eeden [2], is the second reason for our preference mentioned in the previous case, which is obtained by equalizing $p_5$ to $p_4$ and $q_5$ to $q_4$ respectively. It may be remarked that the chosen measure of dependence agrees with Kendall's [7] quantities $\delta$ both in the case of a $2 \times 4$ table and in that of two $2 \times 2$ tables. The common interaction is thus not obtained by amalgamating the treated and not-treated classes, but by *adding* the two values of this measure of dependence, and this as a consequence of choosing the appropriate model.

The case where *three first-order interactions* must be taken in account will be considered now. If none of the faces of the original $2^3$ table plays a special role in this case, we must have a definition of a level containing three interactions which is independent of the choice of the properties allotted to rows, columns, or layers respectively. After Bartlett [1] we choose a vector consisting of $\pi_i$ $(i = 1 \cdots 8)$ with

$$\sum_{i=1}^{8} \pi_i = 1 \quad \text{and} \quad \pi_1\pi_4\pi_6\pi_7 = \pi_2\pi_3\pi_5\pi_8 = \gamma. \tag{16}$$

We remark that this relation was also valid for the levels of the $2^3$ table, the $2 \times 4$ table, and the two $2 \times 2$ tables. Conceiving e.g. $\pi_1/\pi_2 : \pi_3/\pi_4 = \pi_1\pi_4/\pi_2\pi_3$ as a measure of interaction in the $2 \times 2$ table

$$\begin{bmatrix} \pi_1 & \pi_2 \\ \pi_3 & \pi_4 \end{bmatrix}$$

as is recommended for measuring linkage, we see that the relation (16) involves equality of interactions in every pair of faces of the original $2^3$ table. The difference between the experimental vector and the maximum likelihood estimate of the level—which cannot be obtained in that simple way (i.e. from marginal totals) as in the cases discussed till now—will be called second-order interaction again. The corresponding $X^2$ has a one-dimensional $\chi^2$-distribution because the estimation of the level implies the estimation of six parameters.

The estimates $\hat{\pi}_i$ are obtained by determination of the maximum of the likelihood function:

$$C + \sum_{i=1}^{8} x_i \log \pi_i \quad \text{under the conditions (16).}$$

Differentiation of the function:

$$\sum_{i=1}^{8} x_i \log \pi_i + \lambda(\pi_1\pi_4\pi_6\pi_7 - \pi_2\pi_3\pi_5\pi_8) - \mu\left(\sum_{i=1}^{8} \pi_i - 1\right)$$

with respect to $\pi_i$ gives the equations:

$$P_i \equiv \frac{x_i + \lambda\gamma}{\pi_i} - \mu = 0, \qquad (i = 1, 4, 6, 7)$$

and

$$P_i \equiv \frac{x_i - \lambda\gamma}{\pi_i} - \mu = 0, \qquad (i = 2, 3, 5, 8).$$

$\sum_{i=1}^{8} \pi_i P_i$ yields: $\mu = n$. Putting $\lambda\gamma = \delta$, we obtain:

$$\begin{aligned} x_1 + \delta &= n\hat{\pi}_1 , & x_5 - \delta &= n\hat{\pi}_5 , \\ x_2 - \delta &= n\hat{\pi}_2 , & x_6 + \delta &= n\hat{\pi}_6 , \\ x_3 - \delta &= n\hat{\pi}_3 , & x_7 + \delta &= n\hat{\pi}_7 , \\ x_4 + \delta &= n\hat{\pi}_4 , & x_8 - \delta &= n\hat{\pi}_8 . \end{aligned}$$

From this it will be seen that second-order interaction is a multiple of the vector

$$\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}$$

here again, namely with coefficient $\delta/n$. By substitution in (16) the following equation for $\delta$ is obtained:

$$(x_1 + \delta)(x_4 + \delta)(x_6 + \delta)(x_7 + \delta) = (x_2 - \delta)(x_3 - \delta)(x_5 - \delta)(x_8 - \delta).$$

Lancaster [12] showed that the test criterion $X^2 = (\delta^2/n) \sum_{i=1}^{8} 1/\hat{\pi}_i$ calculated after solution of this cubic equation is *asymptotically* equal to

$$\frac{\left(\frac{x_1}{\pi_1} - \frac{x_2}{\pi_2} - \frac{x_3}{\pi_3} + \frac{x_4}{\pi_4} - \frac{x_5}{\pi_5} + \frac{x_6}{\pi_6} + \frac{x_7}{\pi_7} - \frac{x_8}{\pi_8}\right)^2}{n \sum_{i=1}^{8} \frac{1}{\pi_i}}. \tag{17}$$

Now the component of $X^2$ aimed at testing second-order interaction in the discussed models with no, one, and two interactions is equal to the same expression (17) for any $n$ with the restriction only that the $\pi_i$ may stand for estimates of the true $\pi_i$ occurring in the level of the relative model which, however, converge to the true $\pi_i$ for large $n$.

We see that the test criterion for second-order interaction belonging to any of the four discussed models—on condition that the underlying hypothesis expressed by the level is true—is asymptotically the same as that belonging to the following level if we observe the order of our treatment. This order implied that every model represented a stronger assumption about the $\pi_i$ than the following models.

In other words, we may conclude that Bartlett's test of second-order interaction (admitting three first-order interactions) is asymptotically independent of whether any of these interactions is present or not; that our test of second-order interaction admitting two first-order interactions is asymptotically independent of whether any of these two interactions is present or not, but is not valid if three interactions occur in fact; that our test of second-order interaction admitting one first-order interaction is asymptotically independent of whether this interaction is present or not, but is not valid if one or two of the other true interactions is not zero; that the test of second-order interaction assuming no interactions (i.e. Lancaster's procedure) is only justified if no true interaction occurs in fact; that the value of any of the relative statistics is asymptotically equal to those admitting more interactions if only the null hypothesis belonging to the first statistic is not too narrow in the sense that interactions are supposed zero although they are present. This result, an extension of Lancaster's [12] remark, was the third reason for our preference in the choice of specific basis vectors.

Finally, a remark proceeding from a consideration of Bartlett's [1] numerical example also discussed by Lancaster [12]. This example showed a three-way classification of numbers of root-stocks according to time of planting (*at once* and *in spring*), to length of cutting (*long* and *short*), and to success (*alive* and *dead*):

| | at once | | in spring | |
|---|---|---|---|---|
| | long | short | long | short |
| alive | 156 | 107 | 84 | 31 |
| dead | 84 | 133 | 156 | 209 |

Partition of the four-dimensional $\chi^2$ corresponding to a $2^3$ table yielded 95.58 for interaction between time of planting and success, 45.40 for interaction between length of cutting and success, 0.00 for interaction between length of cutting and time of planting, and 0.07 for second-order interaction. In connection with these large interactions, Bartlett's and Lancaster's criteria are not equivalent and they will not be expected to be equal. If we formally follow the procedure discussed in this section, a new model, assuming two first-order interactions, would be needed for a further investigation of second-order interaction. We would consider these two $2 \times 2$ tables:

| | alive | | dead | |
|---|---|---|---|---|
| | long | short | long | short |
| at once | 156 | 197 | 84 | 133 |
| in spring | 84 | 31 | 156 | 209 |

The two-dimensional $\chi^2$, 7.41 (the sum of 6.50 and 0.91) could be partitioned in 5.26 for interaction between time of planting and length of cutting, and 2.15 for second-order interaction. Bartlett's criterion was equal to 2.27, so that the two criteria do not differ much now. The difference could be ascribed to a (formal) interaction between time of planting and length of cutting.

But in our opinion the whole procedure (also Bartlett's) seems to be wrong in this special example. For equality or unequality of dependence in the two considered $2 \times 2$ tables has no practical sense and will not be an object of investigation in this case. Moreover, the fact that the number of root-stocks is equal for all treatment combinations—which led up to an interaction $\chi^2$ exactly equal to zero—suggests that these numbers were not random but fixed before the execution of the experiment. An interaction between time of planting and length of cutting must thus be excluded from the model. For that reason we have already referred to this interaction with the term *formal.*

In this and in similar cases we have to consider four independent binomial distributions defined by four chances $\pi_i$, in this example, chances of *alive* according to:

| | at once | | in spring | |
|---|---|---|---|---|
| | long | short | long | short |
| alive | $\pi_1$ | $\pi_2$ | $\pi_3$ | $\pi_4$ |
| dead | $1 - \pi_1$ | $1 - \pi_2$ | $1 - \pi_3$ | $1 - \pi_4$ |

Here interactions are to be defined again. In the particular case where the numbers in every column are equal (as in this example), say $n$, the experimental result may be partitioned as follows:

$$\frac{1}{4n}\begin{bmatrix} x_1 & x_2 & x_3 & x_4 \\ x_5 & x_6 & x_7 & x_8 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} \pi & \pi & \pi & \pi \\ 1-\pi & 1-\pi & 1-\pi & 1-\pi \end{bmatrix}$$
$$+\frac{\beta}{4}\begin{bmatrix} 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix} + \frac{\gamma}{4}\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$$
$$+\frac{\delta}{4}\begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}.$$

This partition corresponds to a partition of the three-dimensional test criterion for independence in a particular $2 \times 4$ table in three independent components. The second vector on the right represents interaction between time of planting and success, and the third vector, interaction between length of cutting and success. The sum of the first three vectors on the right is proportional to:

$$\begin{bmatrix} \pi+\beta+\gamma & \pi+\beta-\gamma & \pi-\beta+\gamma & \pi-\beta-\gamma \\ 1-\pi-\beta-\gamma & 1-\pi-\beta+\gamma & 1-\pi+\beta-\gamma & 1-\pi+\beta+\gamma \end{bmatrix},$$

i.e. a vector where $\pi_1 - \pi_2 = \pi_3 - \pi_4$, $\pi_1 - \pi_3 = \pi_2 - \pi_4$ and similar relations between $1 - \pi_i$ are valid. Such a vector where the (positive or negative) raising of the chance of *alive* or of *dead* by long cutting is the same for both times of planting, and where this raising by planting at once is the same for both lengths of cutting, seems to be a natural definition of the hypothesis *no second-order interaction* for this case. The fourth vector will represent *second-order interaction*, i.e. inequality of the mentioned raisings. The corresponding partition of $X^2$ is for this example numerically equivalent to that at the beginning of our remark about it. If the numbers in the columns are not equal, a first-order interaction, e.g. between time of planting and success, can be tested in the $2 \times 2$ table obtained by neglecting the other classification (length of cutting).

A test for second-order interaction will imply then (and also when the partition described shows considerable first-order interactions as in this example) a maximum likelihood estimation of the $\pi_i$ under the hypothesis of "no second-order interaction," i.e. $\pi_1 + \pi_4 = \pi_2 + \pi_3$. To that end we determine the maximum of the likelihood function:

$$C + \sum_{i=1}^{4} x_i \log \pi_i + \sum_{i=1}^{4} (n_i - x_i) \log (1 - \pi_i) \quad \text{under that condition.}$$

Differentiating the function:

$$\sum_{i=1}^{4} x_i \log \pi_i + \sum_{i=1}^{4} (n_i - x_i) \log (1 - \pi_i) + \lambda(\pi_1 - \pi_2 - \pi_3 + \pi_4)$$

with respect to $\pi_i$ yields four quadratic equations in $\pi_1$, $\pi_2$, $\pi_3$, and $\pi_4$ respectively. The usable solutions $(0 \leq \pi_i \leq 1)$ are:

$$\hat{\pi}_1 = \frac{\lambda - n_1 + \sqrt{(\lambda - n_1)^2 + 4\lambda x_1}}{2\lambda},$$

$$\hat{\pi}_2 = \frac{\lambda + n_2 - \sqrt{(\lambda + n_2)^2 - 4\lambda x_2}}{2\lambda},$$

$$\hat{\pi}_3 = \frac{\lambda + n_3 - \sqrt{(\lambda + n_3)^2 - 4\lambda x_3}}{2\lambda},$$

$$\hat{\pi}_4 = \frac{\lambda - n_4 + \sqrt{(\lambda - n_4)^2 + 4\lambda x_4}}{2\lambda}.$$

Substitution in the relation between the $\pi_i$ yields the following equation for $\lambda$:

$$\sqrt{(n_1 - \lambda)^2 + 4\lambda x_1} + \sqrt{(n_2 + \lambda)^2 - 4\lambda x_2} + \sqrt{(n_3 + \lambda)^2 - 4\lambda x_3} + \sqrt{(n_4 - \lambda)^2 + 4\lambda x_4} = \sum_{i=1}^{4} n_i$$

from which a solution different from the trivial solution $\lambda = 0$ is required, unless an approximate solution which may be given by

$$\frac{-\frac{x_1}{n_1} + \frac{x_2}{n_2} + \frac{x_3}{n_3} - \frac{x_4}{n_4}}{\sum_{i=1}^{4} \frac{x_i(n_i - x_i)}{n_i^3}}$$

and which can be improved by usual methods, is exactly zero. In that case the solution of $\lambda$ is zero.

In Bartlett's example this approximate solution of $\lambda$ was equal to 4.91. This could be improved to 4.92078. Solving the $\hat{\pi}_i$ with the help of this value gave the following table of expected numbers:

$$\begin{bmatrix} 157.11 & 105.78 & 82.89 & 31.56 \\ 82.89 & 134.22 & 157.11 & 208.44 \end{bmatrix}$$

The one-dimensional $\chi^2$ for second-order interaction according to our definition of no second-order interaction appeared to be equal to 0.082.

ACKNOWLEDGEMENT

This work was begun while the author was assistant at the Department of Mathematics of the Agricultural University, Wageningen. He is much indebted to Professor Dr. N. H. Kuiper whose many suggestions and sound criticisms were an indispensable help to him.

REFERENCES

1. Bartlett, M. S. Contingency table interactions. *J. Roy. Stat. Soc. Suppl. 2*, [1935], 248.
2. Eeden, C. van. Methoden voor het vergelijken, toetsen en schatten van onbekende kansen. *Statistica 7*, [1953], 141–162.
3. Fisher, R. A. *Statistical Methods for Research Workers*, 12th ed., [1954], 101–106 and 299 et sqq.
4. Fog, D. Contingency tables and approximate $\chi^2$-distributions. *Math. Scand. 1*, [1953], 93–103.
5. Haldane, J. B. S. Note on the preceding analysis of Mendelian segregations. *Biometrika 31*, [1939], 68.
6. Irwin, J. O. A note on the subdivision of $\chi^2$ into components. *Biometrika 36*, [1949], 130–134.
7. Kendall, M. G. *The Advanced Theory of Statistics*, Vol. 1, [1945], 310–317.
8. Kimball, A. W. Short-cut formulas for the exact partition of $\chi^2$ in contingency tables. *Biometrics 10*, [1954], 452–458.
9. Kuiper, N. H. Variantie-analyse. *Statistica 6*, [1952], 149–194.
10. Lancaster, H. O. The derivation and partition of $\chi^2$ in certain discrete distributions. *Biometrika 36*, [1949], 117–129.
11. Lancaster, H. O. The exact partition of $\chi^2$ and its application to the problem of the pooling of small expectations. *Biometrika 37*, [1950], 267–270.
12. Lancaster, H. O. Complex contingency tables treated by the partition of $\chi^2$. *J. Roy. Stat. Soc. B 13*, [1951], 242–249.
13. Lancaster, H. O. A reconciliation of $\chi^2$, considered from metrical and enumerative aspects. *Sankhya 13*, [1953], 1–10.
14. Roberts, E., Dawson, W. M., and Madden, M. Observed and theoretical ratios in Mendelian inheritance. *Biometrika 31*, [1939], 62.
15. Simpson, E. H. The interpretation of interaction in contingency tables. *J. Roy. Stat. Soc. B 13*, [1951], 238–241.