THE BASIC THEOREMS OF ANALYSIS OF VARIANCE
WITH SPECIAL REFERENCE TO FIELD EXPERIMENTS.

by

A.A. RAYNER, M.A.

THESIS SUBMITTED FOR DEGREE OF Ph.D.,
EDINBURGH UNIVERSITY.

1947.

Introduction.

The statistical procedure of analysis of variance was invented by R.A. Fisher during his stay as statistician at Rothamsted Experimental Station. His first, more or less tentative, discussion of the theory was set forth in a paper published in 1923 (11), and this was quickly followed up by the more assured and much more complete exposition in his book "Statistical Methods for Research Workers" (12), which revolutionised previous ideas on the principles of scientific experiment. Little additional work was published on the subject until 1933, but since then many workers, among whom may be mentioned M.S. Bartlett, W.G. Cochran, J. Wishart, and above all F. Yates, the present chief statistician at Rothamsted, have developed the theory on the lines laid down by Fisher. Impetus was given to this development especially by the publications of Yates and of Fisher himself (13), in which the new methods of factorial design, confounding, and covariance introduced at Rothamsted were first made more generally known. Fisher's theories met with spasmodic opposition from statisticians such as "Student", Neyman, and others, but have triumphed over all opposition and today are the basis of almost all scientific experimental work amenable to statistical treatment.

Nevertheless one would look in vain throughout the literature for any rigorous and at the same time reasonably simple mathematical treatment of the theory of analysis of variance. Fisher's own exposition is for the most part seemingly intuitive, being designed for the non-mathematical reader, as are for the most part the papers of Yates. Modern text-books such as Snedecor's "Statistical Methods" (28) present the methods without the theory behind them and appeal to the intuition of the reader. Where proofs are attempted, vital points are usually glossed over or assumed, as being beyond the scope of an elementary book. Among the very few British mathematical papers on analysis of variance are those of Irwin (15,16), but his treatment is complicated and unwieldy. Cochran (6) realised the advantages of matrix notation in a subject of this sort, and many of his theorems are equivalent to the lemmas of this thesis, but Cochran left the application of his method undeveloped.

The present thesis constitutes an attempt to put forward a progressive mathematical theory of analysis of variance as applied to the various situations met with in agricultural research in particular, but the applications are, of course, perfectly general. Matrix notation has been used throughout to simplify a subject which would otherwise prove rather unwieldy for mathematical treatment. The basic theories are those of Fisher, Yates, etc., and are now so generally accepted as to require no special references. Acknowledgment by reference is therefore made only in the case of specific points where this has seemed necessary.

Lemmas

The following lemmas will be required in the mathematical discussion of analysis of variance. No explicit proof has been given if the result is a familiar one from statistical text-books.

Lemma 1.

Any variates x, y, z,...... are independent if their moment generating function (m.g.f.) $G(\alpha, \beta, \gamma, \ldots)$ is resolvable into factors $G_1(\alpha)$, $G_2(\beta)$, $G_3(\gamma)$, ...; or, equivalently, if the multivariate probability differential $\phi(x,y,z,\ldots)dxdydz\ldots\ldots$ is resolvable into $\phi_1(x)\phi_2(y)\phi_3(z)\ldots dxdydz\ldots\ldots\ldots$

This follows from the law of compounding the generating functions associated with independent events.

Lemma 2.

If x, y, z,......(not necessarily independent of each other) are variates which are all independent of some other variate u, then any function $f(x,y,z,\ldots.)$ of x, y, z,.... is also independent of u.

Proof: Since x, y, z,... are all independent of u, the differential element of probability of x, y, z,..u is, by Lemma 1, of the form $\phi_1(x,y,z,\ldots.)\ \phi_2(u)$ $dxdydz\cdots du$. Thus, the joint m.g.f. of $f(x,y,z,\ldots)$

and u is $\iiint\ldots \exp\left[\alpha f(x,y,z,\ldots) + \beta u\right] \phi_1(x,y,z\ldots)$ $\phi_2(u)\,dxdydz\ldots du$ which factorises into

$$\iiint\ldots \exp\left[\alpha f(x,y,z,\ldots)\right] \phi_1(x,y,z,\ldots)\, dxdydz\ldots \int \exp(\beta u)$$

$\phi_2(u)\,du$, i.e. the product of two separate m.g.f.'s.

Hence, by Lemma 1, $f(x,y,z,\ldots)$ is independent of u.

Lemma 3.

If $u(x,y,z,\ldots)$ and $v(x,y,z,\ldots)$ are two functions of n variates x, y, z,...., then u and v are independent if their joint m.g.f. $G(\alpha,\beta) \equiv \int \exp(\alpha u + \beta v)\, \phi(\underset{\sim}{x})\, d\underset{\sim}{x}$ [where $\int \phi(\underset{\sim}{x})\, d\underset{\sim}{x}$ represents in matrix notation $\iiint\ldots \phi(x,y,z,\ldots)\,dxdydz\ldots$] is factorisable into $G(\alpha,o)\,G(o,\beta)$.

Likewise they are independent if their compound probability density $\phi(u,v)$ is equal to $\phi_1(u)\,\phi_2(v)$ for all values of x, y, z,....

Lemma 4.

Uncorrelated normal variates are independent.

Proof: The multivariate normal m.g.f. is $\exp(\frac{1}{2}\alpha' V \alpha)$, where α is the vector $\{\alpha_1\ \alpha_2 \ldots \alpha_n\}$ and V is the variance matrix of the variates. If the variates are standardised and uncorrelated, V = I (the unit matrix) and the m.g.f. = $\exp(\frac{1}{2}\alpha'\alpha) = \exp(\frac{1}{2}\alpha_1^2 + \frac{1}{2}\alpha_2^2 + \cdots + \frac{1}{2}\alpha_n^2)$ $= \exp(\frac{1}{2}\alpha_1^2)\exp(\frac{1}{2}\alpha_2^2)\ldots\exp(\frac{1}{2}\alpha_n^2)$, or the product of the m.g.f.'s of the separate variates. Hence, by Lemma 1, the variates are independent.

The converse, independent variates are uncorrelated, is of course, true for <u>any</u> variates.

Lemma 5.

If the set of variates x_i in a vector $\underset{\sim}{x}$ has variance matrix V, and a new set y_i or $\underset{\sim}{y}$ is formed by the linear transformation $y = Hx$ (H being in general a rectangular matrix, not necessarily square), then the variance matrix of the y_i is HVH'.

Proof:

$$V = \text{The mean value of } xx', \text{ i.e. of } \begin{bmatrix} x_1^2 & x_1x_2 & x_1x_3 & \cdots \\ x_2x_1 & x_2^2 & x_2x_3 & \cdots \text{etc} \\ x_3x_1 & x_3x_2 & x_3^2 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

Transform to the new variates $y=Hx$. Then the variance matrix of the $\underset{\sim}{y}$ is the mean value of yy'

$$= \text{Mean } (Hxx'H') = HVH'.$$

Corollary If the linear combinations $h_1x_1 + h_2x_2 + \ldots + h_nx_n$ (i.e. $h'x$) and $k_1x_1 + k_2x_2 + \ldots + k_nx_n$ (i.e. $k'x$) are uncorrelated(the x_j being independent), then $h'k \quad 0$, i.e. h and k are orthogonal.

Proof:

Here $H = \begin{bmatrix} h' \\ \hline k' \end{bmatrix}$, $H' = \left[h \,\vdots\, k \right]$, and $V = I$ if we standardise the variates.

$$\therefore HVH' = HH' = \begin{bmatrix} h'h & h'k \\ k'h & k'k \end{bmatrix}$$

But since the new variates are uncorrelated $h'k=k'h=0$. Hence, if the x_j are independent <u>normal</u> variates, $h'x$ and $k'x$ are also normal variates, and being uncorrelated are (by Lemma 4) also independent. Thus the condition $h'k = 0$ is necessary and sufficient for the statistical independence of the new normal variates $h'x$ and $k'x$.

<u>Lemma 6</u>

If $x_1, y_1, z_1, \ldots\ldots$ (not necessarily independent of each other) are variates which are all independent of $x_2, y_2, z_2, \ldots\ldots$(also not necessarily independent of each other), then any function $f_1(x_1, y_1, z_1, \quad)$ of $x_1, y_1, z_1, \ldots$ is independent of any function of $f_2(x_2, y_2, z_2, \ldots)$ of $x_2, y_2, z_2, \ldots\ldots$

Proof:

Since $x_1, y_1, z_1, \ldots$are all independent of $x_2, y_2, z_2, \ldots$, the differential element of probability of $x_1, y_1, z_1, \ldots, x_2, y_2, z_2, \ldots$ is of the form

$$\phi_1(x_1, y_1, z_1, \ldots) \, \phi_2(x_2, y_2, z_2, \ldots) \, dx_1 dy_1 dz_1 \ldots dx_2 dy_2 dz_2 \ldots$$

(by Lemma 1)

Thus the joint m.g.f. of f_1 and f_2 is

$$\iint \exp(\alpha f_1 + \beta f_2) \, \phi_1(\underset{\sim}{x}_1) \, \phi_2(\underset{\sim}{x}_2) \, d\underset{\sim}{x}_1 \, d\underset{\sim}{x}_2$$

Factorisable into

$$\int \exp(\alpha f_1) \, \phi_1(\underset{\sim}{x}_1) d\underset{\sim}{x}_1 \int \exp(\beta f_2) \, \phi_2(\underset{\sim}{x}_2) \, d\underset{\sim}{x}_2$$

i.e. the product of the two separate m.g.f.'s.

Hence f_1 and f_2 are statistically independent.

Lemma 7.

If $f_1(x_1, y_1, z_1, \ldots\ldots)$ and $f_2(x_2, y_2, z_2, \ldots)$ are independent functions of two sets of variates of which the variates of one set are all independent of the variates of the other set (but not necessarily of those in the same set), and if $f_1 + f_2$, the sum of these functions, is independent of a function of a third set of variates $f_3(x_3, y_3, z_3, \ldots\ldots)$ where $x_3, y_3, z_3, \ldots\ldots$ are not necessarily independent of each other, then each set of variates is independent of the variates of the other two sets.

Proof:

Since $f_1 + f_2$ is independent of f_3, their joint m.g.f. must be of the form

$$\iint \exp\left[\alpha(f_1 + f_2)\right] \phi(\underline{x}_1, \underline{x}_2)\, d\underline{x}_1\, d\underline{x}_2 \int \exp(\beta f_3)\phi_3(\underline{x}_3)\, d\underline{x}_3.$$

But, by Lemma 1, since the $\underline{x}_1$ are independent of the $\underline{x}_2$, this must be of the form

$$\int \exp(\alpha f_1)\phi_1(\underline{x}_1)\, d\underline{x}_1 \int \exp(\alpha f_2)\phi_2(\underline{x}_2)\, d\underline{x}_2 \int \exp(\beta f_3)\phi_3(\underline{x}_3)\, d\underline{x}_3$$

i.e. the differential element of probability of $x_1, y_1, z_1, \ldots\ldots, x_2, y_2, z_2, \ldots\ldots x_3, y_3, z_3, \ldots$ is of the form

$\phi_1(\underline{x}_1)\phi_2(\underline{x}_2)\phi_3(\underline{x}_3)\, d\underline{x}_1\, d\underline{x}_2\, d\underline{x}_3$

$\therefore$ By Lemma 1, the three sets of variates are independent of each other.

The extension to any number of sets of variates is evident.

Corollary 1 By Lemma 6, it follows that f_1, f_2, and f_3, are all independent.

Corollary 2 By Lemma 6, any other functions $f_4(x_1, y_1, z_1, \ldots)$, $f_5(x_2, y_2, z_2, \ldots)$, $f_6(x_3, y_3, z_3, \ldots)$, are also independent.

Lemma 8.

If k sets of $n_1, n_2, \ldots n_k$ observations with respective means M_j and mean square deviations s_j^2 $(j = 1, 2, \ldots k)$ are pooled in an aggregate of $n (= \sum_{j=1}^{k} n_j)$ observations with mean M and mean square deviation s^2, then $ns^2 = \sum_j n_j (s_j^2 + c_j^2)$, where $c_j = M - M_j$. This follows from the fact that the mean square deviation of the j^{th} set about M is $s_j^2 + c_j^2$.

Lemma 9.

If $v = \frac{1}{2} x' Q x$ is a quadratic form in n independent normal variates x_j, all with mean at the origin and equal variance σ^2, and if v has gamma-type probability distribution, then the number of degrees of freedom of v is equal to n', the rank of Q.

Proof:

The matrix Q, being symmetrical, is reducible to diagonal canonical form Λ by means of the transformation $H'QH$, where H is orthogonal. If we introduce new variates $y = H'x$, i.e. $x = Hy$, then $v = \frac{1}{2} y' H'QHy = \frac{1}{2} y'\Lambda y$. The y's are all normally distributed with mean at the origin, and if V is the variance matrix of the x_j , then the variance matrix of the y_j is (by Lemma 5) $H'VH$. But $V = \sigma^2 I$ and $H' H = I$. $\therefore$ the y_j also have variance σ^2 and are independent. Since the y_j are independent normal variates with the same mean and variance , $v(= \frac{1}{2} y'\Lambda y = \frac{1}{2} \sum_j \lambda_j y_j^2$) will have gamma-type distribution if the non-zero values of λ , the latent roots of Q , are all equal, this condition being both necessary and sufficient. $\therefore$ A necessary and sufficient condition that v should have gamma-type distribution is that the n' non-zero latent roots of Q should be equal.

It follows that, if v has gamma-type distribution, then, since it can be reduced to a "sum of squares" orthogonally $(\frac{1}{2} \sum \lambda_j y_j^2$), where all non-zero values of λ , being equal, are either all positive or all negative, $x'Qx$ must be a definite form, either positive definite or negative definite. Also, if ϑ = the equal non-zero latent roots of Q , the characteristic equation must be $\lambda^{n-n'}(\lambda-\vartheta)^{n'} = 0$, and by the Cayley - Hamilton theorem $Q^{n-n'} (Q-\vartheta)^{n'} = 0$, which is equivalent to $Q(Q-\vartheta) = 0$ or $Q^2 = \theta Q$. Hence Q must $= \theta M$, where M is idempotent, for $Q^2 = \theta^2 M^2 = \theta^2 M = \theta Q$.

Now $w = \frac{y}{\sigma}$ is distributed according to $dp = \frac{1}{\Gamma(\frac{1}{2})} \exp(-\frac{1}{2} w^2)\, dw$, and therefore $z = \frac{1}{2} w^2 = \frac{1}{2}\frac{y^2}{\sigma^2}$ is distributed according to $dp = \frac{1}{\Gamma(\frac{1}{2})} z \quad e^{-z} dz$. We have that $v = \frac{1}{2}\theta \sum y_j^2$, where there are n' terms. $\therefore u = \frac{v}{\theta\sigma^2}$ is a combination of n' gamma-type distributions, $z_1 + z_2 + \ldots + z_{n'}$, and its distribution is therefore given by $dp = \frac{1}{\Gamma(\frac{1}{2}n')} u^{\frac{1}{2}n'-1} e^{-u} du$. By comparison with the standard gamma-type distribution, it is seen that the number of degrees of freedom of u (and hence of v, i.e. of $\frac{1}{2} x'Qx$) is n', the rank of Q.

Corollary 1 If Q is idempotent, the number of degrees of freedom of $x'Qx$ is equal to the trace of Q, for then the non-zero latent roots of Q are all equal to 1 and tr. Q = sum of latent roots = n' = rank of Q = number of degrees of freedom.

Corollary 2 If Q is idempotent, an estimate of σ^2, the variance of the x_j, is given by the " mean square" of $x'Qx$ i.e. $x'Qx$ divided by its number of degrees of freedom.

Proof: The mean value of $x'Qx$ is $\sigma^2(\text{tr.}Q)$, since the mean value of all product terms is zero (the x_j being uncorrelated), and the diagonal terms give

$$q_{11}\sigma^2 + q_{22}\sigma^2 + \ldots\ldots + q_{nn}\sigma^2 = \sigma^2(\text{tr.}Q).$$

Hence the mean square of $x'Qx$ is an unbiassed estimate of σ^2 (byCor .1).

N.B. If $x'Qx$ is reduced to canonical form $v = y'\Lambda y$, then, since there are n-n' zero latent roots of Q and (if $x'Qx$ has gamma-type probability distribution) the n' non-zero latent roots are all equal, v is equal to $\theta \sum_j y_j^2$, where there are n' terms. Thus, if $\{x\}$ represents a point in n-dimensional "sample-space" , then, since $x'Qx$ has been shown to represent a multiple of the distance of a point from the origin in n'-dimensional space, it is evident that n-n' dimensions have been lost. This is the statistical equivalent of n-n' linear equations of constraint in dynamics. The rank of the matrix of a quadratic form therefore corresponds exactly to the definition of degrees of freedom given by Fisher (12, 13).

Lemma 10

If A and B are matrices such that the rows of A are orthogonal to the rows of B, and if $A'A$ and $C'C$ (where $C'C = A'A + B'B$) are idempotent, then $B'B$ is also idempotent.

Proof: $(C'C)^2 = (A'A + B'B)^2$

$$= (A'A)^2 + (B'B)^2 + A'AB'B + B'BA'A .$$

But $AB' = B'A = 0$, and $(A'A)^2 = A'A$,

$$\therefore (C'C)^2 = A'A + (B'B)^2$$

$$= C'C \text{ , since } C'C \text{ is idempotent,}$$

$$= A'A + B'B$$

Hence $(B'B)^2 = B'B$, so that $B'B$ is idempotent

Corollary If $y'C'Cy = y'A'Ay + y'B'By$ (where the $\{y\}$ are independent normal variates with variance σ^2) , and if $y'C'Cy$ and $y'A'Ay$ have gamma-type probability distribution with c and a degrees of freedom respectively ($C'C$ and $A'A$ being idempotent and A and B such that $A'B = B'A = 0$), then $y'B'By$ has gamma-type distribution with c-a degrees of freedom and its mean square is an estimate of σ^2 which is independent of that derived from $y'A'Ay$.

Proof: By Lemma 10, $B'B$ is idempotent. $\therefore$ by Lemma 9 Cor.1, $y'B'By$ has gamma-type distribution with degrees of freedom equal to tr. $(B'B)$. But tr.$(C'C)$ = tr.$(A'A)$ + tr.$(B'B)$. Hence the number of degrees of freedom of $y'B'By$ = tr.$(C'C)$ - tr. $(A'A)$ = c-a, and, by Lemma 9, Cor. 2, its mean square yields an unbiassed estimate of σ^2. To show that this estimate is independent of that from $y'A'Ay$, we have that $A'B = B'A = 0$, so that , by Lemma 5 (Cor.) , all the linear combinations Ay are independent of all the linear combinations By . Hence, by Lemma 6, their sums of squares, $y'A'Ay$ and $y'B'By$ are independent.

Section I.

The Hypothesis of Uniformity.

Let us consider a matrix of (say) crop yields

$$Y = [y_{ij}] = \begin{bmatrix} y_{11} & y_{12} & \cdots\cdots & y_{1n} \\ y_{21} & y_{22} & \cdots\cdots & y_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ y_{m1} & y_{m2} & \cdots\cdots & y_{mn} \end{bmatrix},$$

the elements of which are independent normal variates with the same mean μ and the same variance σ^2. Let $\bar{y}$ be the general mean of these variates, y_{10}, y_{20}, ... y_{m0} the row-means, and $y_{01}, y_{02}, \ldots y_{0k}$ **the column-means.**

We have, by Lemma 8,

$$\sum_i\sum_j (y_{ij} - \bar{\bar{y}})^2 = n \sum_i (y_{i0} - \bar{y})^2 + \sum_i\sum_j (y_{ij} - y_{i0})^2 , \quad \ldots..(1,1)$$

i.e. the sum of all mn squared deviations from the general mean is equal to n times the sum of squared deviations of row-means plus a residual sum of squares representing squared deviations from respective row-means.

A deviation of a row-mean from the general mean, e.g. $y_{10} - \bar{y}$, may be represented in vector notation as a'y, where y (the column vector of yields) is

$$\{ y_{11} \;\; y_{12} \ldots y_{1n} \,\vdots\, y_{21} \;\; y_{22} \ldots y_{2n} \,\vdots\, \ldots\ldots \,\vdots\, y_{m1} \;\; y_{m2} \ldots y_{mn} \} ,$$

and $a' = \frac{1}{mn} [m-1 \;\; m-1 \ldots m-1 \,\vdots\, -1 \;\; -1 \ldots -1 \,\vdots\, \ldots\ldots \,\vdots\, -1 \;\; -1 \ldots -1].$

Both the column vector y and the row vector a' are partitioned into m sub-vectors of n elements each.

A deviation of a variate from its row-mean, e.g. $y_{11} - y_{10}$, may similarly be represented in vector notation as $b'y$, where y is the vector of yields and b' is

$$^{1}/_{n} \Big[n-1 \quad -1 \quad -1 \ldots\ldots -1 \;\Big|\; \cdot \; \cdot \ldots\ldots \cdot \;\Big|\; \text{all other subvectors null} \Big]$$, b' being partitioned similarly to a'.

Now $a'b = \frac{1}{mn^2}\Big[(m-1)(n-1) - (n-1)(m-1)\Big] = 0$,

so that a and b are orthogonal, and hence, by Lemma 5 (Cor.), $a'y$ and $b'y$ are independent linear forms. In the same way it may be proved that the deviation of each y_{ij} from its respective row-mean is independent of the deviation of each row-mean from $\bar{y}$. Hence if we write (1,1) as

$$y'A'Ay = y'B'By + y'C'Cy, \qquad \ldots\ldots(1,2)$$

the rows of B must be orthogonal to the rows of C. Also $C'C$ is the "direct sum" of the m matrices each equal to $I_n - M_n$ thus:-

$$\begin{bmatrix} I_n - M_n & & & \\ & I_n - M_n & & \\ & & \ddots & \\ & & & I_n - M_n \end{bmatrix},$$

where I_n is the unit matrix of order n, and M_n is the matrix $\frac{1}{n}\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$ of order $n \times n$. $C'C$ is therefore idempotent. Moreover $A'A$ is equal to $I - M$, where I is the unit matrix of order mn , and $M = \frac{1}{mn} \begin{bmatrix} 1 & 1 & \ldots\ldots & 1 \\ 1 & 1 & \ldots\ldots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \ldots\ldots & 1 \end{bmatrix}$ of order $mn \times mn$, so that $A'A$ is also idempotent.

The mean value of $y'A'Ay$ is, by a well-known result, $(mn-1)\sigma^2$. To find the mean value of $\sum_i \sum_j (y_{ij} - y_{io})^2$, we have that the variance of $y_{11} - y_{10} = b'y$ is $b'b$ (by Lemma 5) $= \frac{1}{n^2}\left[(n-1)^2 + (n-1)\right] = \frac{n-1}{n}$, or (unstandardised) $= \frac{1}{n}(n-1)\sigma^2$. Thus the mean value of the mn squared residuals is equal to $m\,(n-1)\sigma^2$. But $A'A$ and $C'C$ are both idempotent matrices, so that, by Lemma 9, Cor. 2, the quadratic forms $y'A'Ay$ and $y'C'Cy$ have gamma-type distribution with $mn-1$ and $m(n-1)$ degrees of freedom respectively. Hence by Lemma 10 (Cor.), $y'B'By$ also has gamma-type distribution, the degrees of freedom being $mn-1 - m\,(n-1) = m-1$, and its mean square gives an estimate of σ^2 which is independent of the estimate furnished by the mean square of $y'C'Cy$. These two estimates of σ^2 may therefore be tested (either by the F- test or Fisher's z - test) to ascertain whether they are consistent with having been derived from the same normal population.

A matrix of yields $[y_{ij}]$, such that all elements are single samples of normal variates having the same mean and variance, satisfies "the hypothesis of uniformity". In experimental work (not necessarily agricultural) the row suffixes 1,2,...... m may correspond to different treatments e.g. different varieties of a cereal, different fertilisers, different rates of application of the same fertiliser, etc., and it is desired to test whether significant treatment differences are revealed by the experiment. If each treatment is repeated n times (n replications), the conditions being presumed constant for each replication, we have a matrix of order m x n, as above. The "null hypothesis" (that there are no differences between treatments), which in this case is identical with the uniformity hypothesis, may then be tested by comparing the estimate of variance derived from deviations of row-means from the general mean with that derived from individual deviations from respective row-means. Should the former estimate prove significantly the greater, the inference will be that the row-means cannot be considered as having been formed from normal variates with mean μ and variance σ^2. Retaining the hypotheses of normality and equal variance, we must conclude that the means of the variates differ from row to row, a conclusion which agrees with the a priori conditions of different treatments being allotted to different rows.

The above results may be illustrated by

constructing an Analysis of Variance table thus:-

Analysis of Variance.

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares
Between treatment means	$m - 1$	$n\sum_i (y_{i0} - \bar{y})^2$	S_1^2
Residuals	$m(n - 1)$	$\sum_i\sum_j (y_{ij} - y_{i0})^2$	S_2^2
Total	$mn - 1$	$\sum_i\sum_j (y_{ij} - \bar{y})^2$	S_3^2

If the hypothesis of uniformity holds, s_1^2 and s_2^2 are two estimates of σ^2 which do not differ significantly. The total sum of squares $(m-1)s_1^2 + m(n-1)s_2^2$ may then be considered equal to $(mn-1)s_3^2$, yielding a mean square which, being based on the greatest number of degrees of freedom, is the best estimate of σ^2. The hypothesis is equivalent to the assumption that each variate y_{ij} is equal to $\mu + \xi_{ij}$, where ξ_{ij} is a random normal variate with mean at the origin and variance σ^2.

If s_1^2 is significantly greater than s_2^2, the sample of yields can no longer be regarded as homogeneous.

We therefore proceed to the alternative hypothesis, that the variates have different means from row to row, estimated from the sample by y_{i0} $(i=1,2, \ldots m)$. The hypothesis is now that each sample value y_{ij} is equal to $y_{i0} + x'_{ij} = \bar{y} + (y_{i0} - \bar{y}) + x'_{ij}$, where x'_{ij} is the sample value of ξ'_{ij}, the new random variate, and $\sum_i\sum_j x'_{ij} = 0$.

Or, equivalently, each variate y_{ij} is equal to $\mu + \rho_i + \xi'_{ij}$

where $\mu + \rho_i$ is the population mean of the i^{th} treatment, ρ_i being itself a variate with mean at the origin.

No !

The mean square s_3^2 is now meaningless, but s_2^2 is still an estimate of the variance due to random experimental errors i.e. the variance of the ξ'_{ij} . The residual sum of squares is therefore usually called the "error sum of squares", and s_2^2 "the error mean square". s_2 is the standard error per plot of the experiment, or simply the "standard error of the experiment".

The differential effect of the treatments having thus been established, it is now possible to compare treatment-means by means of "Student's" t-test, using s_2 as the estimated standard error of a single yield. This is legitimate since each $y_{io} - \bar{y}$ is independent of each $y_{ij} - y_{io}$, and hence by Lemma 6, $y_{io} - \bar{y} - (y_{jo} - \bar{y}) = y_{io} - y_{jo}$, the difference of any two row- (i.e. treatment-) means, is independent of the estimate of error variance. It will be shown later (Section 9) that the necessity for establishing that s_3^2 is significantly greater than s_2^2 before comparing two treatment-means by the t-test disappears, provided that the particular comparison to be made was one determined beforehand. It is not permissible, for example, to select the highest and lowest treatments after the experiment is completed and declare them to be significantly different as the result of a t-test unless treatments as a whole are significant, i.e. s_1^2 is significantly greater than s_2^2, though of course, it is always possible to make such a comparison by means of the t-test with an estimate of variance derived from the yields of those treatments alone.

Section 2.

Conditions of uniformity in Agricultural Experimentation

It is the peculiarity of agricultural experiments that the conditions for the testing of each treatment can never be exactly the same, nor the same for replications of any one treatment. The chief reason for this lies in soil heterogeneity, the nature of which has been studied by many investigators. A strip of land divided into Plots cannot by any means be considered to have constant fertility from plot to plot. If we consider the matrix of yields $[y_{ij}]$ as representing the yields of certain fixed plots on a field under the same treatment, the infinite hypothetical population of yields under identical conditions represented by the variate y_{11}, for example, will probably not have the same mean as the similar variate y_{12}, nor can the yields of adjoining plots be regarded as independent variates, since, generally speaking, the factors which determine high or low yield and influence the actual sample values of the random variates ξ_{ij} accordingly, are likely to be similar for adjacent plots.

On the other hand the assumptions of normal distribution and constant variance are not rejected. Of these assumptions, that of normal distribution need cause no concern in ordinary crop experiments, since it is a matter of common experience in experimental work involving repetitions under identical conditions. Non-normal distributions, e.g. the Poisson, are, however, common in experiments which involve, for instance, counts of insects. The second assumption, that of constant variance, is also fundamental to the theory of analysis of variance. It is not unreasonable to presume constant variance when the treatments are similar, but cases frequently arise when the variance bears some relationship to the mean. If the two above conditions are not adequately fulfilled, recourse must usually be made to some functional transformation of the variate, though Eden and Yates (8) have demonstrated that the z-test could be safely applied to one actual case of non-normal data.

Some light is thrown on the above matters by uniformity trials, whereby a field is sown with a certain crop and receives uniform treatment, but for harvest purposes is subdivided into small equal plots, the yields of which are separately recorded. It will be noticed that this is not the same thing as our "hypothesis of uniformity" as it stands at present, since, no matter how uniform in fertility the field may be, the plot-variates cannot be considered to have equal means. Mercer and Hall (17) found that the sample of plot-yields thus obtained from a fairly uniform field showed good agreement with the hypothesis of normal distribution, and this has been the experience of many other workers. Where the distribution has been found to be non-normal, the reason probably lies in the fact that the population of yields is heterogeneous i.e. that the field shows a significant departure from uniformity. This suggests that the components of yield due to differential plot fertility, which we may call the "plot-fertility-indices" with respect to a certain crop, have normal distribution, of which the chosen field is a sample. Were it not for the fact of variation of external conditions, it would be possible to imagine an almost infinite normal population of such indices, but owing to the heterogeneous nature of such a population its standard deviation would be large. The necessity for constancy of external conditions leads us to consider the population of indices from a comparatively small area such as a single farm, where the standard deviation will be much smaller, since, for example, the soil-type will remain the same over the area. For experimental purposes, however, this error will still be much too large, and so we take as a sample of plots those from a single field. Uniformity trials show that on a fairly uniform field the combination of plot-fertility-index plus the random variate ξ_{ij} (of Section 1) does indeed produce a finite sample (though not a random one) of some hypothetical normal population. It is necessary therefore, in the theory of agricultural experimentation to postulate a new random variate made up of two independent components, that due to random error pure and simple, and that due to soil heterogeneity. Of course, since the fertility map of a field does not in general change suddenly from point to point, the component of the random variate due to soil heterogeneity will also be correlated between adjacent plots, so that the random variates are not independent.

This new hypothesis, that the plot yields are normal, but not independent, variates with the same mean and variance, will be illustrated by the results of two uniformity trials with wheat, one due to Mercer and Hall (17) and one due to Christidis (5). If Y is the matrix of plot yields with rows and columns corresponding to actual rows and columns in the field, and if M is the matrix with every element equal to $\bar{y}$ (the estimate of the mean), we may form the matrix Y-M. Then (Y-M)(Y-M)' has as its diagonal and non-diagonal elements sums of squares of rows and sums of products between rows, respectively, from which estimates of row-variances and of correlation-coefficients between rows may be obtained. Similar results may be obtained for columns from the matrix $(Y-M)'(Y-M)$. As a simple example let us consider an artificial matrix of yields constructed as follows:- Sequences of ten random digits were examined and the number of digits less than five in each sequence noted. The distribution is that of the symmetrical binomial $(\frac{1}{2}t+\frac{1}{2})^{10}$, nearly normal, with mean 5 and variance 2·5. The matrix Y-M of order 8x8 was constructed, using the true mean 5 instead of the sample estimate. In this case each variate is, of course, independent.

$$
\begin{array}{ll}
Y-M = & \left[\begin{array}{rrrrrrrr}
1 & -3 & 1 & -1 & 0 & 0 & -1 & -1 \\
0 & 2 & -3 & -1 & 0 & -1 & -1 & 2 \\
-1 & 0 & -1 & 1 & 1 & -1 & 1 & -1 \\
-3 & 0 & 0 & -2 & -2 & 1 & 2 & -4 \\
2 & -1 & 1 & 0 & 0 & 0 & -1 & -1 \\
-1 & 0 & -1 & -1 & 0 & 3 & 2 & 0 \\
-1 & -1 & -3 & 2 & 1 & 2 & 0 & 0 \\
1 & 0 & 2 & -1 & -1 & -4 & -1 & -3
\end{array}\right]
\begin{array}{r}
\text{Row Totals} \\ -4 \\ -2 \\ -1 \\ -8 \\ 0 \\ 2 \\ 0 \\ \underline{-7}
\end{array} \\
\text{Column Totals} & \begin{array}{rrrrrrrrr} -2 & -3 & -4 & -3 & -1 & 0 & 1 & -8 & \quad -20 \end{array}
\end{array}
$$

The greatest row or column total is -8 compared with a standard error of $2\sqrt{5}$ (4·5), and the grand total is -20 compared with a standard error of $4\sqrt{10}$ (12·6).

$$(Y-M)(Y-M)' = \begin{bmatrix} (14) & -10 & -3 & 1 & 8 & -3 & -3 & 8 \\ -10 & (20) & 0 & -9 & -6 & -1 & 3 & -6 \\ -3 & 0 & (7) & 4 & -3 & 0 & 5 & 1 \\ 1 & -9 & 4 & (38) & -4 & 12 & -1 & 7 \\ 8 & -6 & -3 & -4 & (8) & -5 & -4 & 8 \\ -3 & -1 & 0 & 12 & -5 & (16) & 8 & -16 \\ -3 & 3 & 5 & -1 & -4 & 8 & (20) & -18 \\ 8 & -6 & 1 & 7 & 8 & -16 & -18 & (33) \end{bmatrix}$$

The following matrix presents row-variances in the diagonal and inter-row correlation-coefficients off the diagonal, negative values being printed in red.

$$\begin{bmatrix} (1.71) & .72 & .39 & .16 & .82 & .15 & .19 & .25 \\ & (2.79) & .02 & .45 & .48 & .03 & .15 & .34 \\ & & (0.98) & .21 & .41 & .02 & .43 & .01 \\ & & & (4.29) & .26 & .66 & .04 & .00 \\ & & & & (1.14) & .45 & .32 & .55 \\ & & & & & (2.21) & .45 & .78 \\ & & & & & & (2.86) & .78 \\ & & & & & & & (3.84) \end{bmatrix}$$

The mean row-variance is 2·48, the greatest deviation from 2·5 being 1·79 compared with the theoretical standard error of 1·3. Three values of r are significant at the 5% point, whereas one would expect only one or two in a random sample of this size.

Matrices of row-variances and inter-row correlation-coefficients are now presented for some actual uniformity trials.

1. Uniformity trial on wheat, Mercer and Hall (17) 500 plots, each $\frac{1}{500}$ acre, in 20 rows and 25 columns. Yields of grain in lbs.

Table (2,1) shows the matrix of row variances and inter-row correlation-coefficients for this trial. Since the matrix is symmetrical, only elements above the diagonal have been entered. The mean row-variance is 0·208, yielding an estimate of plot standard error equal to 0·46, which compares exactly to this degree of accuracy with Mercer and Hall's figure obtained from all plots. Values of r attaining the 5% level of significance (0·40) and the 1% level (0·51) are indicated by single and double underlining respectively.

As before the negative values are in red.

Table (2,2) summarises the information concerning the inter-row values of r .

Only the most cursory examination is necessary to establish the high degree of correlation existing between rows in this example. In any case where the number of values of r significant at the 5% level does not exceed one in twenty, the fact of significant positive correlation is easily established by testing the hypothesis that positive and negative values of r

(.29)	.65	.50	.28	.50	.30	.39	.26	.13	.10	.28	.14	.23	.12	.27	.17	.33	.27	.47	.56
	(.13)	.48	.42	.59	.23	.46	.20	.14	.22	.48	.12	.23	.24	.40	.50	.43	.49	.57	.54
		(.17)	.24	.33	.03	.00	.06	.18	.19	.15	.17	.15	.24	.21	.33	.19	.23	.39	.40
			(.20)	.48	.03	.45	.29	.15	.32	.31	.51	.30	.19	.26	.16	.08	.06	.06	.10
				(.17)	.46	.51	.53	.40	.42	.18	.12	.23	.36	.40	.35	.38	.16	.38	.53
					(.14)	.49	.56	.44	.66	.47	.07	.16	.10	.03	.01	.36	.21	.27	.34
						(.31)	.60	.52	.54	.48	.27	.12	.11	.23	.09	.25	.00	.10	.27
							(.20)	.70	.87	.49	.44	.29	.29	.08	.17	.24	.10	.07	.19
								(.29)	.54	.22	.28	.24	.28	.02	.16	.12	.18	.01	.03
									(.18)	.64	.14	.42	.46	.21	.12	.13	.02	.22	.31
										(.18)	.32	.47	.39	.31	.16	.28	.31	.47	.38
											(.20)	.40	.40	.19	.33	.21	.08	.02	.10
												(.20)	.66	.28	.40	.43	.33	.37	.10
													(.18)	.51	.63	.43	.55	.53	.39
														(.13)	.42	.48	.62	.66	.64
															(.25)	.66	.54	.48	.36
																(.16)	.41	.66	.58
																	(.24)	.74	.48
																		(.33)	.80
																			(.22)

TABLE (2,1)

Distance apart of rows	Number of cases	No. of cases significant at 5% (All positive)	No. of cases significant at 1% (All positive)	Range of values	Number of negative values
Adjacent	19	17	10	0.24 to 0.80	-
1 row	18	13	7	0.03 to 0.87	-
2 rows	17	12	5	0.03 to 0.62	-
3 rows	16	9	3	0.00 to 0.66	-
4 rows	15	5	2	0.15 to 0.64	-
5 rows	14	-	-	-0.07 to 0.39	2
6 rows	13	-	-	-0.02 to 0.31	1
7 rows	12	2	1	-0.10 to 0.51	2
8 rows	11	1	-	-0.18 to 0.48	1
9 rows	10	1	-	-0.10 to 0.40	3
10 rows	9	-	-	0.00 to 0.36	-
11 rows	8	-	-	-0.21 to 0.38	1
12 rows	7	1	-	0.08 to 0.40	-
13 rows	6	1	-	0.06 to 0.50	-
14 rows	5	2	1	-0.06 to 0.53	1
15 rows	4	1	-	0.10 to 0.49	-
16 rows	3	2	1	0.27 to 0.57	-
17 rows	2	2	1	0.47 to 0.54	-
18 rows	1	1	1	0.56	-

TABLE (2,2).

are equally likely.
For example, as few as two negative values in twelve would occur only once in fifty cases on this hypothesis. It is especially noteworthy that such large values of r should be recorded for rows 16 and 17 rows apart.

Tables (2,3) and (2,4) present similar results for column-variances and inter-column correlation-coefficients.

The mean column-variance is 0·15, which gives an estimate of plot standard error equal to 0·39. The reason for the discrepancy between this value and that found from all plots will appear in the sequel.
The significance levels of the correlation-coefficient are 0·44 (5%) and 0·56 (1%).

Without making any exact statistical tests, it is evident that there is some positive correlation between adjacent columns and between columns one column apart. For columns further apart than this the results do not contravene a hypothesis of no correlation. About half of these values of r are negative, and the ranges are fairly evenly disposed about zero. There are 22 values of r significant at the 5% level, almost double the expected number out of a total of 253, if the samples were random. This, however, is not so, and in addition the number of positive significant values (12) is balanced by 10 negative ones.

(.29)	.22	.11	.23	.22	.12	.20	.10	.13	.10	.08	.23	.17	.14	.17	.30	.41	.54	.29	.33	.18	.32	.17	.38	.11
	(.12)	.30	.05	.17	.24	.24	.54	.67	.34	.24	.31	.05	.19	.01	.45	.38	.41	.06	.15	.30	.20	.38	.45	.12
		(.07)	.16	.15	.18	.09	.13	.01	.37	.07	.33	.10	.50	.25	.25	.03	.12	.16	.15	.15	.04	.18	.24	.35
			(.11)	.29	.33	.19	.05	.30	.05	.23	.49	.14	.59	.17	.17	.07	.05	.29	.11	.16	.08	.03	.50	.44
				(.18)	.25	.01	.07	.04	.03	.28	.37	.04	.02	.18	.42	.41	.31	.15	.11	.02	.38	.34	.10	.16
					(.11)	.28	.26	.32	.17	.50	.20	.20	.00	.16	.14	.07	.19	.27	.29	.35	.30	.11	.28	.29
						(.08)	.21	.10	.06	.10	.04	.48	.11	.05	.45	.40	.33	.66	.50	.18	.07	.23	.12	.23
							(.14)	.27	.14	.29	.28	.24	.02	.06	.47	.46	.34	.05	.06	.31	.13	.34	.34	.36
								(.11)	.31	.08	.25	.13	.12	.26	.10	.03	.22	.18	.07	.08	.14	.28	.10	.09
									(.13)	.03	.20	.04	.07	.09	.29	.33	.43	.08	.12	.05	.24	.33	.19	.16
										(.14)	.49	.47	.07	.05	.54	.24	.34	.07	.08	.00	.34	.11	.34	.01
											(.23)	.21	.49	.53	.70	.41	.18	.11	.08	.04	.02	.11	.40	.04
												(.13)	.10	.28	.02	.05	.03	.37	.28	.03	.20	.35	.25	.13
													(.21)	.19	.14	.18	.06	.13	.28	.03	.44	.17	.42	.36
														(.12)	.36	.00	.12	.03	.31	.07	.10	.01	.00	.06
															(.28)	.78	.59	.33	.26	.34	.07	.13	.54	.21
																(.21)	.77	.42	.21	.40	.11	.11	.40	.33
																	(.14)	.43	.13	.25	.02	.06	.42	.23
																		(.14)	.32	.33	.09	.29	.15	.28
																			(.12)	.48	.39	.14	.01	.12
																				(.11,	.47	.39	.07	.27
																					(.12)	.63	.06	.17
																						(.15)	.16	.14
																							(.15)	.44
																								(.16)

TABLE (2,3).

Distance apart of columns	Number of cases	No. of cases significant at 5% level		No. of cases significant at 1% level		Range of values	Number of negative values
		Positive	Negative	Positive	Negative		
Adjacent	24	7	-	3	-	-(.29) to .78	6
1 column	23	3	-	1	-	-(.26) to .59	7
2 columns	22	-	1	-	-	-(.53) to .33	8
3 columns	21	-	1	-	1	-(.70) to .40	11
4 columns	20	1	1	-	-	-(.54) to .50	5
5 columns	19	2	-	-	-	-(.42) to .54	10
6 columns	18	-	1	-	1	-(.67) to .37	6
7 columns	17	1	3	-	-	-(.54) to .49	11
8 columns	16	1	1	-	-	-(.46) to .45	6
9 columns	15	1	-	1	-	-(.35) to .59	7
10 columns	14	1	-	-	-	-(.42) to .50	5
11 columns	13	1	-	1	-	-(.41) to .66	7
12 columns	12	1	-	-	-	-(.31) to .50	7
13 columns	11	-	1	-	-	-(.45) to .29	6
14 columns	10	-	-	-	-	-(.38) to .35	5
15 columns	9	-	-	-	-	-(.41) to .34	5
16 columns	8	-	1	-	-	-(.54) to .38	4
17 columns	7	-	-	-	-	-(.29) to .34	4
18 columns	6	-	-	-	-	-(.30) to .33	3
19 columns	5	1	-	-	-	-(.20) to .50	3
20 columns	4	1	-	-	-	-(.38) to .44	1
21 columns	3	1	-	-	-	.17 to .45	-
22 columns	2	-	-	-	-	-(.12) to .38	1
23 columns	1	-	-	-	-	0.11	-

TABLE (2,4)

The following tables of analysis of variance are relevant.

(a)

Source	D.F.	Sums of Squares.	Mean Squares.	F.
Rows	19	6·0939	0·321	1·6
Residuals	480	98·5783	0·205	
Total	499	104·6722†	0·210	

† Figure obtained by calculating back from Mercer and Hall's estimate of variance.

(b)

Source	D.F.	Sums of squares	Mean Squares	F
Columns	24	33·5956	1·400	9·3**
Residuals	475	71·0766	0·150	
Total	499	104·6722	0·210	

The value of F in Table (b) is highly significant, and that in Table (a) is almost significant at the 5% level.

Thus, neither row nor column-means can be regarded as derived from a single homogeneous normal population. In fact, the conditions of uniformity do not hold, and it is only by chance that the plot data are so well fitted by a normal curve.

Anticipating the results of Section 3, we may combine the above two tables into a single table of analysis of variance :-

Source	D.F.	Sums of Squares	Mean Squares	F
Rows	19	6·0939	0·321	2.2*
Columns	24	33·5956	1·400	9.8**
Residuals	456	64·9827	0·143	
Total	499	104·6722	0·210	

In seeking to explain the very striking inter-row correlations and the highly significant mean square for columns, it is pertinent to enquire, as did Christidis (5) in a different connexion, whether the drilling was done along the columns. This information is, however, not available from the original paper, yet the explanation is clearly that the main changes of fertility occur in a direction parallel to the rows. Possibly the ploughing of the field may have always been done parallel to the columns.

Wishart and Sanders (21) obtained yields for plots $\frac{1}{50}$ acre in area from Mercer and Hall's uniformity trial data by combining the yields of ten adjacent small plots (five along the rows and two across), thus obtaining plots of the size recommended for experimental purposes. It is of interest to obtain the corresponding results for these data.

(a) Matrix of row variances and inter-row correlation-coefficients

$$
\begin{bmatrix}
(5.29) & 0.61 & 0.45 & 0.67 & 0.28 & 0.66 & 0.51 & 0.72 & 0.68 & 0.74 \\
 & (2.53) & 0.67 & 0.51 & 0.71 & 0.52 & 0.62 & 0.66 & 0.26 & 0.26 \\
 & & (2.21) & 0.65 & \underline{0.89} & \underline{0.89} & 0.71 & 0.58 & 0.20 & 0.16 \\
 & & & (7.19) & 0.71 & 0.82 & 0.14 & 0.21 & 0.01 & 0.06 \\
 & & & & (4.93) & 0.77 & 0.40 & 0.27 & -0.20 & -0.22 \\
 & & & & & (3.20) & 0.52 & 0.49 & 0.25 & 0.26 \\
 & & & & & & (5.00) & \underline{0.94} & 0.73 & 0.67 \\
 & & & & & & & (12.25) & 0.86 & 0.84 \\
 & & & & & & & & (10.15) & \underline{0.99} \\
 & & & & & & & & & (14.39)
\end{bmatrix}
$$

The **significance** values of r are 0·88 (5%) and 0·96 (1%). The mean row-variance is 6·71, compared with 6·26, the variance obtained from all plots.

(b) Matrix of column-variances and inter-column correlation-coefficients.

$$
\begin{bmatrix}
(2.68) & -0.14 & 0.42 & -0.59 & 0.42 \\
 & (0.88) & 0.57 & 0.23 & 0.27 \\
 & & (3.59) & -0.29 & 0.18 \\
 & & & (9.12) & 0.22 \\
 & & & & (2.52)
\end{bmatrix}
$$

The 5% level of significance for r is at 0·63. The mean column-variance is 3·76. It is not to be expected that there would be much evidence of positive inter-column correlation even between adjacent columns, since for the original small plots positive inter-column correlation extends only as far as columns separated by a single column.

Analyses of Variance.

Number	Source	D.F.	Sums of Squares	Mean Squares	F.
1	Rows	9	38.37	4.26	
	Residuals	40	268.52	6.71	
2	Columns	4	137.89	34.47	9.17**
	Residuals	45	169.00	3.76	
3	Rows	9	38.37	4.26	1.17
	Columns	4	137.89	34.47	9.50**
	Residuals	36	130.63	2.63	
	Total	49	306.89	6.26	

2. Uniformity trial on wheat by B.G. Christidis (5):- 288 plots each 8 ins: x $7\frac{1}{2}$ ft: in 24 rows and 12 columns. Yield of grain in grams.

The matrix of row-variances and inter-row correlation-coefficients is presented in Table (2,5). The mean row-variance is 180.0, whereas the variance calculated from all plots is 195.1. The significant values of the correlation-coefficient are 0.58(5%) and 0.71(1%).

The similar matrix for columns is set out in table (2,6). The mean column-variance is 130.1, compared with the exact figure of 130.18 derived by analysis of variance and with the estimate of variance derived from all plots of 195.1. The significant values of r are 0.40(5%) and 0.52(1%).

(116)	.59	.72	.77	.40	.48	.67	.66	.78	.22	.49	.20	.51	.44	.30	.39	.49	.01	.57	.25	.32	.46	.40	.38
	(167)	.71	.78	.58	.77	.65	.67	.73	.21	.44	.24	.11	.36	.04	.07	.60	.29	.36	.23	.22	.18	.25	.34
		(209)	.80	.71	.80	.75	.76	.68	.43	.19	.16	.11	.11	.04	.03	.43	.12	.48	.32	.12	.13	.38	.13
			(202)	.62	.78	.73	.83	.88	.54	.78	.07	.23	.37	.26	.33	.61	.14	.36	.49	.42	.46	.50	.37
				(153)	.60	.40	.60	.60	.52	.14	.10	.40	.25	.43	.27	.52	.14	.54	.37	.40	.01	.51	.17
					(138)	.85	.70	.63	.49	.37	.23	.11	.12	.04	.12	.36	.26	.38	.36	.17	.10	.37	.20
						(115)	.67	.64	.26	.36	.15	.08	.31	.04	.03	.17	.22	.40	.21	.01	.17	.35	.26
							(174)	.78	.65	.54	.21	.22	.33	.23	.36	.73	.33	.44	.41	.48	.58	.79	.61
								(106)	.59	.50	.21	.53	.60	.55	.56	.72	.33	.55	.40	.53	.46	.49	.51
									(107)	.32	.20	.19	.19	.45	.48	.58	.35	.34	.52	.66	.44	.65	.43
										(284)	.11	.24	.48	.23	.51	.52	.45	.16	.38	.71	.81	.59	.71
											(404)	.74	.45	.69	.64	.04	.08	.34	.05	.51	.09	.05	.24
												(196)	.77	.88	.82	.42	.07	.62	.15	.58	.34	.41	.50
													(157)	.79	.76	.50	.39	.54	.38	.58	.56	.44	.68
														(282)	.85	.47	.15	.58	.43	.66	.40	.48	.46
															(267)	.57	.31	.35	.30	.82	.67	.58	.73
																(186)	.36	.58	.56	.73	.68	.66	.64
																	(68)	.04	.08	.42	.37	.30	.71
																		(172)	.51	.44	.32	.53	.34
																			(118)	.48	.60	.54	.18
																				(135)	.71	.72	.76
																					(132)	.72	.76
																						(160)	.74
																							(272)

TABLE (2,5).

(131)	.31	.41	.41	.15	.10	.27	.23	.45	.14	.10	.10
	(89)	.51	.61	.33	.12	.07	.14	.16	.23	.29	.09
		(92)	.48	.31	.20	.23	.36	.22	.24	.08	.17
			(125)	.54	.27	.14	.49	.06	.13	.37	.26
				(134)	.13	.08	.49	.30	.07	.60	.28
					(61)	.25	.11	.10	.29	.36	.11
						(97)	.02	.39	.58	.18	.29
							(273)	.08	.14	.43	.29
								(163)	.63	.29	.40
									(227)	.32	.42
										(75)	.39
											(94)

TABLE (2,6)

Analysis of Variance Tables

Number	Source	D.F.	Sums of Squares	Mean Squares	F
1	Rows	23	8,495.20	369.36	2.05 **
	Residuals	264	47,489.33	179.88	
2	Columns	11	20,055.20	1823.20	14.01 **
	Residuals	276	35,929.33	130.18	
3	Rows	23	8,495.20	369.36	3.41 **
	Columns	11	20,055.20	1823.20	16.81 **
	Residuals	253	27,434.13	108.44	
	Total	287	55,984.53	195.07	

The results derived from Christidis' data are rather similar to those from Mercer and Hall's. However, in this case the yields, considered as a single sample, show a significant departure from normality in respect of kurtosis. Once again there is very strong positive inter-row correlation, but there is no longer such an equality of positive and negative inter-column correlation-coefficients. In this trial the information is available that the drilling was done along the rows, and thus a possible explanation disappears. However, it is clear from the analyses of variance that once again the major changes of fertility are paralled to the rows, though there is a larger component than before parallel to the columns, thus accounting for the significant value of F for rows and for the excess of positive inter-column correlation-coefficients even for columns many columns apart.

It is noteworthy that had Christidis, in seeking to prove the superiority of long, narrow plots in respect of lowness of standard error, happened to have placed his long plots along the columns, he would have got different results, in fact, results similar to those from Mercer and Hall's trial, where long plots along the columns had a negligible effect in reducing the standard error. Long, narrow plots can be superior only if they happen to lie along the line of major fertility change, a point noticed by Day (7) and Smith (19). In cereal experiments, where it is convenient to have plots of only one drill-width, there is the possibility of additional error due to drilling variations.

Section 3.

The Principle of Randomisation

Fisher (12), (13) solved the problem of the non-independence of the plot variates by the stipulation of the Principle of Randomisation. Suppose that a set of n treatments is to be tested. If the n plots for each treatment are allotted entirely at random by some process of randomisation, then it becomes possible, as before, under the hypothesis that treatments have no differential effects, to regard the y_{ij} (the yield of the jth plot of the ith treatment) as independent normal variates with the same mean and variance. This is the same "hypothesis of uniformity" as that originally formulated, except that, in considering each variate y_{ij} equal to $\mu + \xi_{ij}$, the random variate ξ_{ij} now contains a component due to soil heterogeneity. The results of §1 will therefore hold good.

Such a type of experiment, however, would be rare in agriculture owing to its lack of precision due to high standard error. Moreover, the "fertility - map" of a field can never be exactly known, for even a previous uniformity trial can give only an approximate idea of that. Hence, unless the field happens to be fairly uniform, the possibility exists that the chosen field will not, owing to its heterogeneity, constitute a normal sample of plot-fertility-indices. The experimentalist overcomes these difficulties by local control. The field is divided into a number of small areas called "blocks", each containing a single replication of all the treatments under consideration, arranged in a different random order in each block. We now consider the theory of such an experimental design.

Section 4.

Randomised Blocks

Let us consider a matrix of yields $[y_{ij}]$ of order $m \times n$, in which the rows represent different treatments and the columns different blocks, and let us examine the matrix under the hypothesis of complete uniformity throughout the experiment i.e. that the y_{ij} are normal independent variates with the same mean (μ) and the same variance (σ^2).

We have from (1,1)

$$\sum_i\sum_j(y_{ij} - \bar{y})^2 = n\sum_i(y_{io} - \bar{y})^2 + \sum_i\sum_j(y_{ij} - y_{io})^2$$

Consider the term $\sum_i\sum_j(y_{ij} - y_{io})^2$ as the sum of squares of mn variates arranged in a matrix thus:-

$$\begin{bmatrix} y_{11} - y_{10} & y_{12} - y_{10} & \ldots\ldots & y_{1n} - y_{10} \\ y_{21} - y_{20} & y_{22} - y_{20} & \ldots\ldots & y_{2n} - y_{20} \\ \vdots & \vdots & & \vdots \\ y_{m1} - y_{mo} & y_{m2} - y_{mo} & \ldots\ldots & y_{mn} - y_{mo} \end{bmatrix}$$

The general mean of these variates $= \frac{1}{mn}(\sum_i\sum_j y_{ij} - n\sum_i y_{io}) = 0$

The mean of the jth column is $y_{oj} - \frac{1}{m}\sum_i y_{io} = y_{oj} - \bar{y}$.

Hence, applying Lemma 8 to $\sum_i\sum_j(y_{ij} - y_{io})^2$ with respect to column-means, we have:-

$$\sum_i\sum_j (y_{ij} - y_{io})^2 = m\sum_j (y_{oj} - \bar{y})^2 + \sum_i\sum_j (y_{ij} - y_{io} - y_{oj} + \bar{y})^2, \quad \ldots\ldots(4,1)$$

and therefore

$$\sum_i\sum_j (y_{ij} - \bar{y})^2 = n\sum_i (y_{io} - \bar{y})^2 + m\sum_j (y_{oj} - \bar{y})^2$$
$$+ \sum_i\sum_j (y_{ij} - y_{io} - y_{oj} + \bar{y})^2 \quad \ldots\ldots(4,2)$$

To show that the two sums of squares on the R.H.S. of (4,1) are independent, let us write, for example, the residual $y_{12} - y_{1o} - y_{o2} + \bar{y}$, or $y_{12} - \frac{1}{n}\sum_j y_{1j} - \frac{1}{m}\sum_i y_{i2} + \frac{1}{mn}\sum_i\sum_j y_{ij}$, as

$\frac{1}{mn}\left[1-m \;\; (m-1)(n-1) \;\; 1-m \ldots 1-m \,\vdots\, 1 \;\; 1-n \;\; 1 \ldots 1 \,\vdots\, 1 \;\; 1-n \;\; 1 \ldots 1 \,\vdots\, \text{etc.}\right] y = d'y$, where y is the same vector as in § 1. A deviation of a column mean from the general mean, e.g., $y_{o1} - \bar{y}$, may be written as $c'y$,

$$= \frac{1}{mn}\left[n-1 \; -1 \; -1 \ldots -1 \,\vdots\, n-1 \; -1 \; -1 \ldots -1 \,\vdots\, \text{etc.}\right] y$$

Now $c'd = \frac{1}{m^2n^2}\{(n-1)(1-m) - (m-1)(n-1) - (n-2)(1-m) + (m-1)(n-1+n-1-\overline{n-2})\} = 0$, so that c and d are orthogonal, and hence, by Lemma 5 (Cor), $c'y$ and $d'y$ are independent linear forms. The same may be proved of any column-mean deviation and of any residual. Hence, if we write (4,1) as,

$$y'C'Cy = y'D'Dy + y'E'Ey, \quad \ldots\ldots\ldots(4,3)$$

the rows of D must be orthogonal to the rows of E. The matrix $C'C$ has already been proved idempotent (§ 1) and $D'D = mH'H$, where

$H = \frac{1}{m}\left[I_n - M_n \,\vdots\, I_n - M_n \,\vdots \ldots m \text{ sub-matrices}\right]$, I_n and

M_n being as before, so that

$$D'D = \frac{1}{m}\left[\begin{array}{c|c|c|c} I_n - M_n & I_n - M_n & \cdots\cdots & I_n - M_n \\ \hline I_n - M_n & I_n - M_n & \cdots\cdots & I_n - M_n \\ \hline \vdots & \vdots & \vdots & \vdots \\ \hline I_n - M_n & I_n - M_n & \cdots\cdots & I_n - M_n \end{array}\right]$$

of order mn x mn, which is clearly idempotent. The mean value of y'D'Dy is found from the fact that the variance of $y_{01} - \bar{y} = c'y$ is c'c (by Lemma 5) $= \frac{1}{m^2n^2}\left[m(n-1)^2 + m(n-1)\right] = \frac{n-1}{mn}$. The required mean value is therefore $(n-1)\sigma^2$ (unstandardised). It follows from Lemma 9, Cor.2, that y'D'Dy has gamma-type distribution with n-1 degrees of freedom, and since we have already proved that y'C'Cy has gamma-type distribution with m(n-1) degrees of freedom, we have, in consequence of Lemma 10 (Cor.), that y'E'Ey also has gamma-type distribution and has (m-1)(n-1) degrees of freedom. In addition, its mean square is an estimate of σ^2 which is independent of the estimate derived from y'D'Dy, and also (by Lemma 7, Cor. 1) of that from y'B'By or $n\sum_i (y_{i0} - \bar{y})^2$. Thus all three component sums of squares on the right-hand side of (4,2) have independent gamma-type distribution with m-1, n-1, and (m-1)(n-1) degrees of freedom respectively, so that their mean squares may be tested in the usual manner.

The fore-going results are summarised in an analysis of variance table, thus:-

Analysis of Variance

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares
Between treatment-means.	m-1	$n\sum_i(y_{io} - \bar{y})^2$	s_1^2
Between block-means.	n-1	$m\sum_j(y_{oj} - \bar{y})^2$	s_2^2
Residuals	(m-1)(n-1)	$\sum_i\sum_j(y_{ij} - y_{io} - y_{oj} + \bar{y})^2$	s_3^2
Total	mn-1	$\sum_i\sum_j(y_{ij} - \bar{y})^2$	----

It will be desired to compare s_1^2, the estimate of variance derived from treatment-means, with s_3^2, that from residuals. It is easily seen that s_3^2 will provide an estimate of random variance even when neither block-means nor treatment-means can be regarded as derived from the same population, within the limits of sampling error. In that case the yield of the $(i,j)^{th}$ plot (y_{ij}) may be considered equal to $\bar{y} + (y_{io} - y) + (y_{oj} - y) + x''_{ij} = y_{io} + y_{oj} - \bar{y} + x''_{ij}$ (where x''_{ij} is a random component), since y_{io} and y_{oj} are the sample estimates of the mean of the i^{th} row and j^{th}

column respectively. Hence $y_{ij} - y_{io} - y_{oj} + \bar{y} = x''_{ij}$.

It would also be possible to compare s_2^2 with s_3^2 in order to see if the randomised-block layout has been significantly effective in removing the effects of soil heterogeneity. However, the significance of s_2^2 is really not in question, for in the designing of the experiment we have in effect assumed that block-means would be different. Therefore, failing the significance of s_2^2 , there is no justification for pooling the sums of squares for blocks and residuals into a combined estimate of error variance, even if s_2^2 should happen to be less than s_3^2 . Thus, the hypothesis of complete uniformity was in reality not the correct one. Each variate y_{ij} is equal to $\mu + \beta_j + \xi'_{ij}$, where β_j is the mean of the j^{th} block, estimated from the sample by $y_{oj} - \bar{y}$, so that each sample value y_{ij} is equal to $\bar{y} + (y_{oj} - \bar{y}) + x'_{ij}$. β_j is a normal independent variate with mean at the origin and variance σ_B^2, but constant for all variates relating to a given block. The mean square for blocks is now an estimate of $\sigma'^2 + m\sigma_B^2$, where σ'^2 is the variance of ξ'_{ij} i.e. the random variance. The orthogonality of the design of the experiment ensures that the other two mean squares continue to be (on a null hypothesis) independent estimates of the random variance, for the proof of their independence is unaffected by the fact that the y_{ij} may have different means, and since $\sum_j b_j$ (the sum of the sample values of β_j) = 0, $\sum_i (y_{io} - \bar{y})^2$ and $\sum_i \sum_j (y_{ij} - y_{io} - y_{oj} + \bar{y})^2$ involve only deviations due to the random component x'_{ij}.

The effectiveness of the randomised-block design is clear. In general $s_2^2 > s_3^2$, which means that $\sigma'^2 < \sigma^2$, i.e. that the precision of the experiment has been improved. Moreover, since the area of a single block is more likely to be of uniform fertility than the total area of all the experimental plots, the plot-fertility-indices of a single block are more likely to correspond with the theoretical requirements of a sample from a population distributed normally about the mean for the particular block. Thus, for an efficient experiment there is a limit to the size of block, and therefore to the number of treatments which may be tested in any one experiment, though this disadvantage may be overcome by such a device as confounding.

Section 5.

Latin-Square Design

In the randomised-block design it was seen that the effects of soil heterogeneity could be partially eliminated and the precision of the experiment increased by the division of the experimental area into blocks. The Latin-square design enables the effects of soil heterogeneity to be eliminated in two directions at right angles, and in general still further increases the precision of the experiment. Consider the field divided into n^2 plots by means of n rows and n columns, the n treatments under test being assigned by a process of randomisation so that each treatment occurs once in each row and once in each column. There are thus n replications. The process of the random allocation of treatments [described by Wishart and Sanders (21)] ensures that the variate elements of the matrix of yields under the uniformity hypothesis may be considered independent.

Let the matrix of yields of a Latin square of order n be $[y_{ijk}]$, where i refers to row, j to column, and k to treatment, and let the means of rows, columns, and treatments be respectively y_{i00}, y_{0j0}, y_{00k}. The general mean is $\bar{y}$, and the y_{ijk} are, under the hypothesis of uniformity, independent normal variates, each with mean μ and variance σ^2. Disregarding treatment suffixes, we have by (4,1)

$$\sum_i\sum_j (y_{ijk} - \bar{y})^2 = n\sum_i (y_{ioo} - \bar{y})^2 + n\sum_j (y_{ojo} - \bar{y})^2 + \sum_i\sum_j (y_{ijk} - y_{ioo} - y_{ojo} + \bar{y})^2 \quad \ldots\ldots\ldots\ldots (5,1)$$

The residual sum of squares may now be further subdivided; for consider the residuals

$(y_{ijk} - y_{ioo} - y_{ojo} + \bar{y})$ arranged in rows according to the suffix k. The mean of all these n^2 residuals is given by $\frac{1}{n^2}(\sum_i\sum_j y_{ijk} - n\sum_i y_{ioo} - n\sum_j y_{ojo} + n^2 \bar{y}) = 0$.

The mean of row 1 $= y_{oo1} - \frac{1}{n}(\sum_i y_{ioo} + \sum_j y_{ojo} - n\bar{y}) = y_{oo1} - \bar{y}$.

Hence, applying Lemma 8, we obtain

$$\sum_i\sum_j (y_{ijk} - y_{ioo} - y_{ojo} + \bar{y})^2 = n\sum_k (y_{ook} - \bar{y})^2 + \sum_i\sum_j (y_{ijk} - y_{ioo} - y_{ojo} - y_{ook} + 2\bar{y}) \quad \ldots\ldots\ldots\ldots (5,2)$$

Combining (5,1) and 5,2), we have

$$\sum_i\sum_j (y_{ijk} - \bar{y})^2 = n\sum_i (y_{ioo} - \bar{y})^2 + n\sum_j (y_{ojo} - \bar{y})^2 + n\sum_k (y_{ook} - \bar{y})^2 + \sum_i\sum_j (y_{ijk} - y_{ioo} - y_{ojo} - y_{ook} + 2\bar{y})^2 \quad \ldots\ldots\ldots\ldots (5,3)$$

In order to exemplify a typical residual, let us suppose that treatment 1 occurs in the first three rows and columns as y_{13}, y_{21}, and y_{32}, the treatment suffix being omitted, because in general the suffix k of y_{ijk} may belong to any of the n treatments according to the particular randomisation of the experiment. The residual, for example,

$$y_{211} - y_{200} - y_{010} - y_{001} + 2y$$

$$= y_{21} - \tfrac{1}{n}(y_{21} + y_{22} + \ldots + y_{2n}) - \tfrac{1}{n}(y_{11} + y_{21} + \ldots + y_{n1}) - \tfrac{1}{n}(y_{13} + y_{21} + y_{32} + \ldots) + \tfrac{2}{n^2}\sum_i\sum_j y_{ij}$$

which in vector notation may be written as

$$\tfrac{1}{n^2}\Big[2-n \;\; 2 \;\; 2-n \;\; 2 \ldots 2 \;\vdots\; n^2-3n+2 \;\; 2-n \;\; 2-n \ldots 2-n \;\vdots\; 2-n \;\; 2-n \;\; 2 \;\; 2 \ldots 2 \;\vdots\; \text{etc.}\Big]\, y, \text{ or } e'y,$$

where y is the column vector

$$\{y_{11} \;\; y_{12} \ldots y_{1n} \;\vdots\; y_{21} \;\; y_{22} \ldots y_{2n} \;\vdots\; \ldots \;\vdots\; y_{n1} \;\; y_{n2} \ldots y_{nn}\}$$, and

e' is similarly partitioned.

A deviation of a treatment-mean from the general mean e.g. $y_{001} - \bar{y}$, may be written as

$$\tfrac{1}{n^2}\Big[-1 \;\; -1 \;\; n-1 \;\; -1 \ldots -1 \;\vdots\; n-1 \;\; -1 \;\; -1 \ldots -1 \;\vdots\; -1 \;\; n-1 \;\; -1 \ldots -1 \;\vdots\; \text{etc.}\Big]\, y$$

or f'y.

Since f'e

$$= \tfrac{1}{n^4}\Big\{-(2-n)(n-1) - 2(n-2)(n-1) + (n-1)(n^2-3n+2) + (2-n)(n-1)^2 + (2-n)(n-1)\Big\} = 0,$$ e and f are orthogonal, and if we express (5, 2) in matrix notation as

$$y'E'Ey = y'F'Fy + y'G'Gy \ldots\ldots\ldots\ldots\ldots (5,4)$$

and (5,3) similarly as

$$y'A'Ay = y'B'By + y'D'Dy + y'F'Fy + y'G'Gy, \ldots\ldots (5,5)$$

it follows as before that the rows of F are orthogonal to the rows of G.

It was proved in the previous section that E'E is an idempotent matrix. F'F may also be proved idempotent, since this matrix is equal to nL'L, where

$L = \frac{1}{n}\left[K_1 - M_n \,\vdots\, K_2 - M_n \,\vdots\, \ldots\ldots \,\vdots\, K_n - M_n\right]$, the K_j being matrices of order n x n with one element in each row and column unity, all other elements zero, and such that, if $k'_{j(\ell}$ is the ℓ^{th} row of K_j , $k'_{i(\ell}\, k_{j(\ell} = 0$. Hence

$$F'F = \frac{1}{n}\left[\begin{array}{c|c|c|c} I_n - M_n & K_{12} - M_n & \ldots\ldots & K_{1n} - M_n \\ \hline K_{21} - M_n & I_n - M_n & \ldots\ldots & K_{2n} - M_n \\ \hline K_{31} - M_n & K_{32} - M_n & \ldots\ldots & K_{2n} - M_n \\ \hline \vdots & \vdots & \vdots & \vdots \\ \hline K_{n1} - M_n & K_{n2} - M_n & \ldots\ldots & I_{nn} - M_n \end{array}\right]$$

where $K_{ij} = K_i' K_j$, and $K_{ij} = K'_{ji}$ $(i \neq j)$. F'F is now seen to be idempotent, since $K_{ij} K_{ji} = I_n$ and $K_{ij}\ K_{jk} = K'_i\ K_j\ K'_j\ K_k = K'_i\ K_k = K_{ik}$ (since $K_j\ K'_j = I_n$).

To find the mean value of y'F'Fy or $n\sum_k (y_{00k} - \bar{y})^2$, we have that for $k=1$ the variance of f'y is given by $f'f = \frac{1}{n^4}\left\{(n-1)n + (n-1)^2 n\right\} = \frac{n-1}{n^2}$, or unstandardised, $\frac{n-1}{n^2}\sigma^2$. The mean value of y'F'Fy is therefore $(n-1)\sigma^2$, and , by Lemma 9, Cor. 2, it may be deduced that y'F'Fy has gamma-type distribution with **n-1** degrees of freedom. Hence, since it has already been proved that $y'E'Ey$ is similarly distributed with $(n-1)^2$ degrees of freedom (§4), Lemma 10 (Cor.) may be applied to show that $y'G'Gy$ also has gamma-type distribution, its degrees of freedom being $(n-1)^2-(n-1)$ or $(n-1)(n-2)$. Its mean square is an estimate of variance which is independent of similar estimates derived not only from $y'F'Fy$, but also, by Lemma 7, Cor.1) from all the other quadratic forms of (5,5). Thus all the component sums of squares on the right-hand side of (5,3) or (5,5) have independent gamma-type distribution with n-1, n-1, n-1, and (n-1)(n-2) degrees of freedom respectively, so that their mean squares may be compared by the usual tests for compatibility with the uniformity hypothesis.

We have the following table of analysis of variance:-

Variation	D.F.	Sums of Squares	Mean Squares
Between rows	$n-1$	$n\sum_i (y_{ioo}-\bar{y})^2$	s_1^2
Between columns	$n-1$	$n\sum_j (y_{ojo}-\bar{y})^2$	s_2^2
Between treatments	$n-1$	$n\sum_k (y_{ook}-\bar{y})^2$	s_3^2
Residuals	$(n-1)(n-2)$	$\sum_i\sum_j (y_{ijk}-y_{ioo}-y_{ojo}-y_{ook}+2\bar{y})^2$	s_4^2
Total	n^2-1	$\sum_i\sum_j (y_{ijk}-\bar{y})^2$	

The desired comparison of mean squares will be between s_3^2, the estimate of variance obtained from treatment-means, and s_4^2, the estimate obtained from residuals.

The latter will still be an estimate of the random variance even when row-means, column-means, and treatment-means can no longer be considered identical within the limits of random sampling, for then the yield of the $(i,j)^{th}$ plot with treatment k (say) may be regarded as equal to

$$\bar{y} + (y_{ioo} - \bar{y}) + (y_{ojo} - \bar{y}) + (y_{ook} - \bar{y}) + x''_{ij}$$

$$= y_{ioo} + y_{ojo} + y_{ook} - 2\bar{y} + x''_{ij},$$

where x''_{ij} is a random component. Hence the residual

$y_{ijk} - y_{ioo} - y_{ojo} - y_{ook} + 2y$ is equal to x''_{ij}.

The significance of s_1^2 and of s_2^2 may also be tested to ascertain the efficiency of the row and column arrangement in removing the effects of soil heterogeneity. Yet, as in the case of the blocks mean-square of a randomised-blocks experiment, their significance is not in question, for the design of the experiment really presumes that the row and column means will be different. Thus the uniformity hypothesis is not the correct one, the hypothesis of the Latin square being that each variate y_{ij} is equal to $\mu + \rho_i + \gamma_j + \xi'_{ij}$, where ρ_i is the mean of the i^{th} row (estimated from the sample by $y_{ioo} - \bar{y}$) and γ_j is the mean of the j^{th} column (estimated from the sample by $y_{ojo} - \bar{y}$). Both ρ_i and γ_j are normal independent variates with mean at the origin, their variances being σ_R^2 and σ_c^2 respectively. ρ_i is constant for all y_{ij} relating to the i^{th} row, and γ_j is constant for all y_{ij} relating to the j^{th} column. The mean squares for rows and columns are now estimates of $\sigma'^2 + n\sigma_R^2$ and of $\sigma'^2 + n\sigma_c^2$

respectively, where σ'^2 is the variance of ξ_{ij}, i.e. the random variance. Once again the orthogonality of the design ensures that under a null hypothesis (i.e. that treatments have no differential effect) s_3^2 and s_4^2 continue to be estimates of the random variance. In the first place the proof of their independence is unaffected by the fact that the y_{ij} may have different means. Also $\sum_i r_i = \sum_j c_j = 0$ (the sample values of ρ_i and γ_j respectively), and therefore $\sum_k (y_{ook} - \bar{y})^2$ and $\sum_i \sum_j (y_{ijk} - y_{ioo} - y_{ojo} - y_{ook} + 2\bar{y})^2$ still involve only deviations due to the random variate.

If, as usually happens, both s_1^2 and s_2^2 are greater than s_4^2, it is apparent that the precision of the experiment has been increased as compared with the corresponding randomised-blocks design, though this is slightly offset by the loss of $(n-1)^2 - (n-1)(n-2) = n-1$ degrees of freedom for error. The laying down of a Latin-square experiment, too, frequently involves practical difficulties. The design could not, for example, be easily adapted to a cereal experiment where sowing and manuring were done by drills. There is also a limitation to the number of treatments that can be tested. If the number is less than four, the error variance is not based on a sufficient number of degrees of freedom; and if the number is greater that about eight, the rows and columns become too long, with a consequent impairment of efficiency.

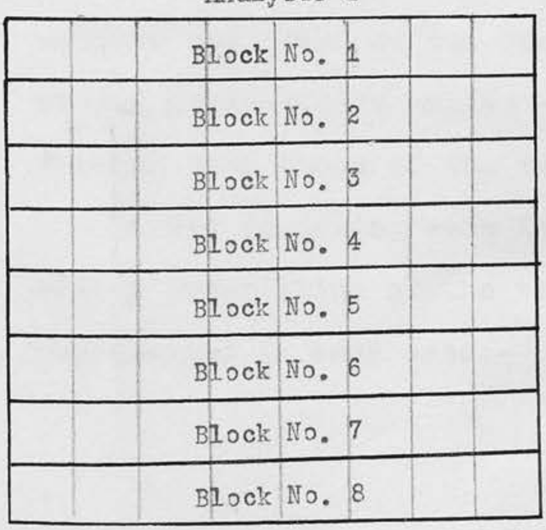
Section 6.

The Comparative Efficiency of Randomised Blocks and the Latin Square.

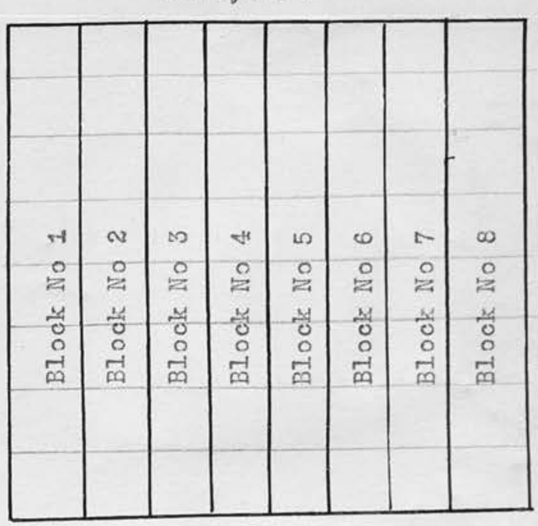
It has been seen (§2) that a set of plot-fertility-indices from a reasonably uniform field may be considered as a homogeneous normal sample. Yet it was recognised that in practice, as revealed by uniformity trials, such conditions of uniformity must rarely exist when the area of the field is large compared with the area of the ultimate plots. The reason for this is the existence of "fertility-gradients", or systematic changes of fertility, the origins of which must lie in such causes as the varying chemical constitution of the soil, the previous husbandry of the field, etc. It has also been seen (§§4,5) how by means of local control (randomised blocks, the Latin square) the total area of the experiment may be subdivided into smaller areas within which the plot-fertility-indices are more likely to approach the theoretical requirements of a homogeneous normal sample. This raises the question of the relative efficiency of the various experimental designs (different-shaped blocks, the Latin square) in eliminating the systematic effects of soil heterogeneity and in laying bare the residual variance associated with pure random error. Actual data from uniformity trials can be of little assistance in the latter respect, since it is impossible to separate the

the component of residual variance due to soil fertility and that due to random error. This suggests the use of models. Twenty-three "fertility-grids" were therefore prepared, representing in an ideal manner various possible types of fertility-gradients in 8 x 8 squares. On each of these were superimposed two "random-grids" chosen at random from a set of four, prepared and analysed so that the residual variances were known. The 46 sets of artificial data thus obtained were analysed in five different ways:- (1) with rows treated as blocks,(2) with columns treated as blocks,(3) with adjacent half-rows treated as blocks,(4) with adjacent half-columns treated as blocks,(5) as an 8 x 8 Latin square. Each block thus consisted of eight "plots", the blocks being allocated as in the diagrams below:-

Analysis 1



Analysis 2



Analysis 3

Block No 1	Block No 5
Block No 2	Block No 6
Block No 3	Block No 7
Block No 4	Block No 8

Analysis 4

Block No 1	Block No 3	Block No 5	Block No 7
Block No 2	Block No 4	Block No 6	Block No 8

The percentage efficiency of each type of analysis was estimated in the usual manner from the error mean square and the variance of the particular random-grid used (as calculated from the total sum of squares). This has meant that occasionally, owing to the vagaries of sampling, an efficiency greater than 100% has been recorded. No account was taken of the fact that the variance-estimate of the Latin-square analysis is based on fewer degrees of freedom than those of the randomised-blocks analyses.

The following were the selected "fertility-grids", with a description of the type of fertility-gradient represented in each case:-

I

0	0	10	0	0	0	10	0
0	0	10	0	0	0	10	0
0	0	10	0	0	0	10	0
0	0	10	0	0	0	10	0
0	0	10	0	0	0	10	0
0	0	10	0	0	0	10	0
0	0	10	0	0	0	10	0
0	0	10	0	0	0	10	0

Two ridges of fertility parallel to the columns of the field.

II

0	5	10	8	5	8	10	5
0	5	10	8	5	8	10	5
0	5	10	8	5	8	10	5
0	5	10	8	5	8	10	5
0	5	10	8	5	8	10	5
0	5	10	8	5	8	10	5
0	5	10	8	5	8	10	5
0	5	10	8	5	8	10	5

Same as I but with ridges less sudden

III

0	5	10	5	0	0	0	0
0	0	5	10	5	0	0	0
0	0	0	5	10	5	0	0
5	0	0	0	5	10	5	0
10	5	0	0	0	5	10	5
5	10	5	0	0	0	5	10
0	5	10	5	0	0	0	5
0	0	5	10	5	0	0	0

Two ridges diagonally across the field

IV

5	8	10	8	5	2	0	0
2	5	8	10	8	5	2	0
0	2	5	8	10	8	5	2
0	0	2	5	8	10	8	5
0	0	0	2	5	8	10	8
0	0	0	0	2	5	8	10
0	0	0	0	0	2	5	8
0	0	0	0	0	0	2	5

A single diagonal ridge, not so sudden as in III

V

12	12	12	12	12	12	12	12
10	10	10	10	10	10	10	10
8	8	8	8	8	8	8	8
6	6	6	6	6	6	6	6
4	4	4	4	4	4	4	4
2	2	2	2	2	2	2	2
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

A gradual, regular drift down the field. For combinatorial purposes a down-drift may be designated V(a), and a cross-drift V(b).

VI

14	13	12	11	10	9	8	7
13	12	11	10	9	8	7	6
12	11	10	9	8	7	6	5
11	10	9	8	7	6	5	4
10	9	8	7	6	5	4	3
9	8	7	6	5	4	3	2
8	7	6	5	4	3	2	1
7	6	5	4	3	2	1	0

A gradual, regular drift diagonally across the field

VII

12	12	22	12	12	12	22	12
10	10	20	10	10	10	20	10
8	8	18	8	8	8	18	8
6	6	16	6	6	6	16	6
4	4	14	4	4	4	14	4
2	2	12	2	2	2	12	2
0	0	10	0	0	0	10	0
0	0	10	0	0	0	10	0

I combined with V(a)

VIII

12	10	18	6	4	2	10	0
12	10	18	6	4	2	10	0
12	10	18	6	4	2	10	0
12	10	18	6	4	2	10	0
12	10	18	6	4	2	10	0
12	10	18	6	4	2	10	0
12	10	18	6	4	2	10	0
12	10	18	6	4	2	10	0

I combined with V(b)

IX

14	13	22	11	10	9	18	7
13	12	21	10	9	8	17	6
12	11	20	9	8	7	16	5
11	10	19	8	7	6	15	4
10	9	18	7	6	5	14	3
9	8	17	6	5	4	13	2
8	7	16	5	4	3	12	1
7	6	15	4	3	2	11	0

I combined with VI

X

12	17	22	20	17	20	22	17
10	15	20	18	15	18	20	15
8	13	18	16	13	16	18	13
6	11	16	14	11	14	16	11
4	9	14	12	9	12	14	9
2	7	12	10	7	10	12	7
0	5	10	8	5	8	10	5
0	5	10	8	5	8	10	5

II combined with V(a)*

XI

14	18	22	19	15	17	18	12
13	17	21	18	14	16	17	11
12	16	20	17	13	15	16	10
11	15	19	16	12	14	15	9
10	14	18	15	11	13	14	8
9	13	17	14	10	12	13	7
8	12	16	13	9	11	12	6
7	11	15	12	8	10	11	5

II combined with VI

XII

12	17	22	17	12	12	12	12
10	10	15	20	15	10	10	10
8	8	8	13	18	13	8	8
11	6	6	6	11	16	11	6
14	9	4	4	4	9	14	9
7	12	7	2	2	2	7	12
0	5	10	5	0	0	0	5
0	0	5	10	5	0	0	0

III combined with V(a)

* The combination of II with V(b) is omitted as not being materially different from II itself

XIII

12	15	18	11	4	2	0	0
12	10	13	16	9	2	0	0
12	10	8	11	14	7	0	0
17	10	8	6	9	12	5	0
22	15	8	6	4	7	10	5
17	20	13	6	4	2	5	10
12	15	18	11	4	2	0	5
12	10	13	16	9	2	0	0

III combined with V(b)

XIV

14	18	22	16	10	9	8	7
13	12	16	20	14	8	7	6
12	11	10	14	18	12	6	5
16	10	9	8	12	16	10	4
20	14	8	7	6	10	14	8
14	18	12	6	5	4	8	12
8	12	16	10	4	3	2	6
7	6	10	14	8	2	1	0

III combined with VI

XV

17	20	22	20	17	14	12	12
12	15	18	20	18	15	12	10
8	10	13	16	18	16	13	10
6	6	8	11	14	16	14	11
4	4	4	6	9	12	14	12
2	2	2	2	4	7	10	12
0	0	0	0	0	2	5	8
0	0	0	0	0	0	2	5

IV combined with V(a)

XVI

17	18	18	14	9	4	0	0
14	15	16	16	12	7	2	0
12	12	13	14	14	10	5	2
12	10	10	11	12	12	8	5
12	10	8	8	9	10	10	8
12	10	8	6	6	7	8	10
12	10	8	6	4	4	5	8
12	10	8	6	4	2	2	5

IV combined with V(b)

XVII

19	21	22	19	15	11	8	7
15	17	19	20	17	13	9	6
12	13	15	17	18	15	11	7
11	10	11	13	15	16	13	9
10	9	8	9	11	13	14	11
9	8	7	6	7	9	11	12
8	7	6	5	4	5	7	9
7	6	5	4	3	2	3	5

IV combined with VI

XVIII

14	12	10	8	6	4	2	0
12	10	8	6	6	4	2	2
10	8	6	6	6	4	4	4
8	6	4	6	6	6	6	6
6	4	4	6	6	6	6	8
4	2	4	4	6	6	8	10
2	2	2	4	6	8	10	12
0	2	4	6	8	10	12	14

An extreme case quoted by Wishart (20) as being unsuitable for a Latin square.

XIX

10	8	6	4	6	8	10	11
12	6	7	6	8	9	10	12
10	8	7	6	6	7	9	10
8	7	6	5	4	6	8	9
6	6	6	4	0	4	6	8
4	2	0	4	6	6	5	7
6	4	4	6	12	8	4	5
5	5	6	8	8	4	0	3

XX

2	2	4	6	6	7	5	4
0	3	5	7	6	9	6	3
6	6	7	8	9	12	6	0
12	9	8	10	12	6	4	2
10	9	8	7	6	4	3	3
8	10	8	6	0	2	3	4
10	8	6	4	0	2	3	4
12	10	8	6	4	3	4	3

In grids XIX, XX, and XXI an attempt has been made to simulate what uniformity trials show to be a common situation, i.e. contours of equal fertility level surrounding high and low points of fertility. In these grids points of high (12), medium (6), and low (0) fertility were allotted at random, the remaining plots being given indices so as to produce smooth changes from one plot to another

XXI

9	10	11	10	10	8	5	3
10	11	12	11	9	6	3	0
9	10	10	8	7	7	6	3
9	8	8	6	6	6	5	6
7	6	8	12	8	4	5	7
5	3	6	8	4	0	4	8
3	0	3	6	5	4	8	12
4	3	4	5	4	5	7	10

XXII

0	0	3	6	10	6	3	0
3	0	3	6	10	6	3	0
6	3	3	6	10	6	3	3
10	6	3	6	10	6	3	6
6	10	6	6	10	6	6	10
3	6	10	6	10	6	10	6
6	6	6	10	10	10	6	6
10	10	10	10	10	10	10	10

Fan-shaped ridges

XXIII

1	1	2	2	3	2	1	2
3	2	3	3	4	3	2	3
5	4	4	4	5	4	3	4
6	6	5	5	6	5	5	6
5	7	7	6	7	6	6	7
4	6	8	7	8	7	8	6
5	6	7	9	9	9	7	6
6	7	8	9	10	9	8	7

Fan-shaped fertility-gradient.

It is, of course, recognised that the above grids represent fertility-gradients of a very ideal type indeed, unlikely to be exactly realised in practice, but it is nevertheless interesting to see with what degree of efficiency the different types of analysis eliminate their effects. No attempt has been made to represent ridges or gradients crossing the field at angles other than 0°, 45°, or 90° This would be difficult within the limits of an 8x8 square, and would disturb the simplicity of the scheme.

The following are the four "random-grids", two of which, chosen at random, were superimposed on each of the fertility-grids. The sample values were obtained in the same manner as those in Section 2, page 25, so that the theoretical variance is 2·5.

									Totals.
(1)	-1	0	1	-1	-1	-1	-2	3	-2
	1	2	1	-2	2	2	-1	-2	3
	-2	0	2	-1	-3	-3	-1	-2	-10
	1	-3	1	0	0	-1	-2	2	-2
	2	-1	-4	3	1	-1	1	-1	0
	2	-1	-2	0	-1	2	-1	-1	-2
	0	0	2	1	-3	-1	0	0	-1
	3	-1	-1	-2	-1	0	0	0	-2
Totals:	6	-4	0	-2	-6	-3	-6	-1	-16

Analyses of Variance (Random-grid 1)

Number	Source	D.F.	Sums of Squares	Mean Squares
1	Rows	7	11.75	1.68
	Residuals	56	158.25	2.83
2	Columns	7	13.25	1.89
	Residuals	56	156.75	2.80
3	Rows	7	11.75	1.68
	Columns	7	13.25	1.89
	Residuals	49	145.00	2.96
	Total	63	170.00	2.70

									Totals
(2)	-1	-3	-1	0	0	-1	-3	-1	-10
	0	-1	-2	0	-1	2	-1	-1	- 4
	0	0	3	1	-3	-1	0	0	0
	3	-1	-1	-3	-1	0	0	0	- 3
	2	0	-1	-1	0	0	1	0	1
	1	-1	2	1	0	1	-1	1	4
	-2	-1	0	0	-1	2	0	-1	- 3
	-1	0	-1	0	-2	-2	0	-1	- 7
Totals:	2	-7	-1	-2	-8	1	-4	-3	-22

Analyses of Variance (Random-grid 2)

Number	Source	D.F.	Sums of Squares	Mean Squares
1	Rows	7	17.44	2.49
	Residuals	56	89.00	1.59
2	Columns	7	10.94	1.56
	Residuals	56	95.50	1.71
3	Rows	7	17.44	2.49
	Columns	7	10.94	1.56
	Residuals	49	78.06	1.59
	Total	63	106.44	1.69

									Totals
(3)	0	-2	1	3	-1	0	0	-2	-1
	1	-1	0	1	-2	2	3	0	4
	0	-4	-1	0	-2	-1	-2	2	-8
	2	3	0	-1	0	0	0	-3	1
	1	0	0	1	0	1	0	0	3
	2	-4	1	-3	0	0	1	-2	-5
	2	0	-1	-3	2	3	-1	1	3
	1	1	1	-1	0	-2	2	2	4
Totals:	9	-7	1	-3	-3	3	3	-2	1

Analyses of Variance (Random-grid 3)

Number	Source	D.F.	Sums of Squares	Mean Squares
1	Rows	7	17.61	2.52
	Residuals	56	157.37	2.81
2	Columns	7	21.36	3.05
	Residuals	56	153.62	2.74
3	Rows	7	17.61	2.52
	Columns	7	21.36	3.05
	Residuals	49	136.01	2.78
	Total	63	174.98	2.78

									Totals
(4)	-1	2	0	-1	0	-2	0	-4	-6
	1	-1	2	3	-2	-1	0	-2	0
	2	-2	1	-1	-1	0	2	0	1
	1	0	-2	-1	-3	1	-3	1	-6
	-1	-2	1	0	-1	0	0	-2	-5
	-2	0	-2	1	1	3	-2	1	0
	1	1	1	0	0	1	0	0	4
	0	-2	0	1	3	-2	1	1	2
Totals:	1	-4	1	2	-3	0	-2	-5	-10

Analyses of Variance

Number	Source	D.F.	Sums of Squares	Mean Squares
1	Rows	7	13.19	1.88
	Residuals	56	135.25	2.42
2	Columns	7	5.94	0.85
	Residuals	56	142.50	2.54
3	Rows	7	13.19	1.88
	Columns	7	5.94	0.85
	Residuals	49	129.31	2.64
	Total	63	148.44	2.36

The results of the analyses of the combined grids are tabulated below:-

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares				
			Analysis 1	Anal:2.	Anal:3	Anal:4	Anal:5
I	Blocks (rows)	1	1.68		1.86		1.68
		4	1.88		3.42		1.88
	Blocks (columns)	1		167.61		56.93	167.61
		4		176.56		61.21	176.56
	Error	1	23.54	2.80	23.52	16.63	2.96
		4	24.38	2.54	24.19	16.96	2.64
	Percentage Efficiency	1	11.5%	96.4%	11.5%	16.2%	91.2%
		4	9.7%	92.9%	9.8%	13.9%	89.4%
	Mean Percent: Effic:		10.6%	94.6%	10.6%	15.0%	90.3%
			52.6%		12.8%		

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares Analysis 1	Anal:2	Anal:3	Anal:4	Anal:5
II	Blocks (rows)	3	2.52		5.52		2.52
		4	1.88		5.21		1.88
	Blocks (columns)	3		84.52		51.19	84.52
		4		92.63		58.85	92.63
	Error	3	12.99	2.74	12.62	6.91	2.78
		4	13.89	2.54	13.47	6.77	2.64
	Percentage Efficiency	3	21.4%	101.5%	22.0%	40.2%	100.0%
		4	17.0%	92.9%	17.5%	34.9%	89.4%
			19.2%	97.2%	18.8%	37.6%	
	Mean Percent: Effic:		58.2%		28.2%		94.7%
III	Blocks (rows)	1	8.29		26.50		8.29
		4	5.99		33.96		5.99
	Blocks (columns)	1		7.25		30.32	7.25
		4		8.35		23.35	8.35
	Error	1	16.55	16.68	14.28	13.80	17.88
		4	18.46	18.17	14.97	16.29	19.91
	Percentage Efficiency	1	16.3%	16.2%	18.9%	19.6%	15.2%
		4	12.8%	13.0%	15.8%	14.5%	11.9%
			14.6%	14.6%	17.4%	17.0%	
	Mean Percent: Effic:		14.6%		17.2%		13.6%

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares: Analysis 1	Anal:2	Anal:3	Anal:4	Anal:5
IV	Blocks (rows)	1	20.18		62.82		20.18
		2	23.60		66.35		23.60
	Blocks (columns)	1		13.29		66.89	13.29
		2		18.92		66.71	18.92
	Error	1	15.15	16.01	9.82	9.31	15.42
		2	13.38	13.96	8.04	7.99	12.59
	Percentage Efficiency	1	17.8%	16.9%	27.5%	29.0%	17.5%
		2	12.6%	12.1%	21.0%	21.2%	13.4%
	Mean Percent: Effic:		15.2%	14.5%	24.2%	25.1%	
			14.8%		24.6%		15.4%
V	Blocks (rows)	3	157.02		151.12		157.02
		4	146.60		142.42		146.60
	Blocks (columns)	3		3.05		120.52	3.05
		4		0.85		118.35	0.85
	Error	3	2.81	22.06	3.55	7.37	2.78
		4	2.42	20.63	2.94	5.95	2.64
	Percentage Efficiency	3	98.9%	12.6%	78.3%	37.7%	100.0%
		4	97.5%	11.4%	79.7%	39.7%	89.4%
	Mean Percent: Effic:		98.2%	12.0%	79.0%	38.7%	
			54.1%		58.8%		94.7%

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares Analysis 1	Anal: 2	Anal: 3	Anal: 4	Anal: 5
VI	Blocks (rows)	2	44.49		82.35		44.49
		4	39.31		81.13		39.31
	Blocks (columns)	2		52.42		78.78	52.42
		4		54.56		83.21	54.56
	Error	2	7.95	6.96	3.21	3.66	1.59
		4	9.13	7.22	3.90	3.64	2.64
	Percentage Efficiency	2	21.3%	24.3%	52.6%	46.2%	106.3%
		4	25.8%	32.7%	60.5%	64.8%	89.4%
	Mean Percent: Effic:		23.6%	28.5%	56.6%	55.5%	
			26.0%		56.0%		97.8%
VII	Blocks (rows)	1	163.96		160.43		163.96
		3	157.02		151.12		157.02
	Blocks (columns)	1		167.61		179.07	167.61
		3		185.19		175.52	185.19
	Error	1	23.54	23.08	23.98	21.65	2.96
		3	25.56	22.06	26.31	23.27	2.78
	Percentage Efficiency	1	11.5%	11.7%	11.3%	12.5%	91.2%
		3	10.9%	12.6%	10.6%	11.9%	100.0%
	Mean Percent: Effic:		11.2%	12.2%	11.0%	12.2%	
			11.7%		11.6%		95.6%

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares Analysis 1	Anal:2	Anal:3	Anal:4	Anal:5
VIII	Blocks (rows)	3	2.52		129.27		2.52
		4	1.88		142.71		1.88
	Blocks (columns)	3		296.84		134.52	296.84
		4		292.71		147.35	292.71
	Error	3	39.53	2.74	23.69	23.03	2.78
		4	38.90	2.54	21.29	20.71	2.64
	Percentage Efficiency	3	7.0%	101.5%	11.7%	12.1%	100.0%
		4	6.1%	92.9%	11.1%	11.4%	89.4%
	Mean Percent: Effic:		6.6%	97.2%	11.4%	11.8%	
			51.9%		11.6%		94.7%
IX	Blocks (rows)	1	48.82		93.86		48.82
		3	44.66		78.62		44.66
	Blocks (columns)	1		203.04		99.79	203.04
		3		213.34		88.30	213.34
	Error	1	27.97	8.69	22.34	21.60	2.96
		3	28.88	7.80	24.64	23.43	2.53
	Percentage Efficiency	1	9.7%	31.1%	12.1%	12.5%	91.2%
		3	9.6%	35.6%	11.3%	11.9%	109.9%
	Mean Percent: Effic:		9.6%	33.4%	11.7%	12.2%	
			21.5%		12.0%		100.6%

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares Analysis 1	Anal:2	Anal:3	Anal:4	Anal:5
X	Blocks (rows)	1	163.97		161.14		163.97
		4	146.60		144.21		146.60
	Blocks (columns)	1		75.75		170.07	75.75
		4		92.63		174.56	92.63
	Error	1	12.06	23.08	12.41	11.29	2.96
		4	13.89	20.63	14.19	10.39	2.64
	Percentage Efficiency	1	22.4%	11.7%	21.8%	23.9%	91.2%
		4	17.0%	11.4%	16.6%	22.7%	89.4%
			19.7%	11.6%	19.2%	23.3%	
	Mean Percent: Effic:		15.6%		21.2%		90.3%
XI	Blocks (rows)	2	44.49		61.99		44.49
		4	39.31		60.06		39.31
	Blocks (columns)	2		80.92		75.42	80.92
		4		88.06		82.28	88.06
	Error	2	11.51	6.96	9.32	7.64	1.59
		4	13.32	7.23	10.72	7.95	2.64
	Percentage Efficiency	2	14.7%	24.3%	18.1%	22.1%	106.3%
		4	17.7%	32.6%	22.0%	29.7%	89.4%
			16.2%	28.4%	20.0%	25.9%	
	Mean Percent: Effic:		22.3%		23.0%		97.8%

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares Analysis 1	Anal:2	Anal:3	Anal:4	Anal:5
XII	Blocks (rows)	1	144.86		160.79		144.86
		3	136.12		145.94		136.12
	Blocks (columns)	1		7.25		120.32	7.25
		3		5.37		112.30	5.37
	Error	1	16.55	33.75	14.56	19.62	17.88
		3	15.94	32.28	14.71	18.91	17.44
	Percentage Efficiency	1	16.3%	8.0%	18.5%	13.8%	15.1%
		3	17.4%	8.6%	18.9%	14.7%	15.9%
			16.8%	8.3%	18.7%	14.2%	
	Mean Percent: Effic:		12.6%		16.4%		15.5%
XIII	Blocks (rows)	2	13.38		194.85		13.38
		3	7.34		180.52		7.34
	Blocks (columns)	2		196.14		210.06	196.14
		3		194.16		203.44	194.16
	Error	2	40.40	17.55	17.71	15.81	18.15
		3	39.53	16.18	17.89	15.02	17.44
	Percentage Efficiency	2	4.2%	9.6%	9.5%	10.7%	9.3%
		3	7.0%	17.2%	15.5%	18.5%	15.9%
			5.6%	13.4%	12.5%	14.6%	
	Mean Percent: Effic:		9.5%		13.6%		12.6%

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares: Analysis 1	Anal:2	Anal:3	Anal:4	Anal:5
XIV	Blocks (rows)	1	44.00		124.21		44.00
		4	31.99		117.38		31.99
	Blocks (columns)	1		76.96		113.18	76.96
		4		73.49		98.21	73.49
	Error	1	25.27	21.15	15.24	16.62	17.88
		4	26.61	21.42	15.93	18.33	19.91
	Percentage Efficiency	1	10.7%	12.8%	17.7%	16.2%	15.1%
		4	8.9%	11.0%	14.8%	12.9%	11.9%
	Mean Percent: Effic:		9.8%	11.9%	16.2%	14.6%	
			10.8%		15.4%		13.5%
XV	Blocks (rows)	2	272.31		307.63		272.31
		3	270.96		315.18		270.96
	Blocks (columns)	2		18.92		263.85	18.92
		3		19.53		275.00	19.53
	Error	2	13.38	45.05	8.96	14.44	12.59
		3	14.86	46.29	9.33	14.35	14.19
	Percentage Efficiency	2	12.6%	3.8%	18.9%	11.7%	13.4%
		3	18.7%	6.0%	29.8%	19.4%	19.6%
	Mean Percent: Effic:		15.6%	4.9%	24.4%	15.6%	
			10.2%		20.0%		16.5%

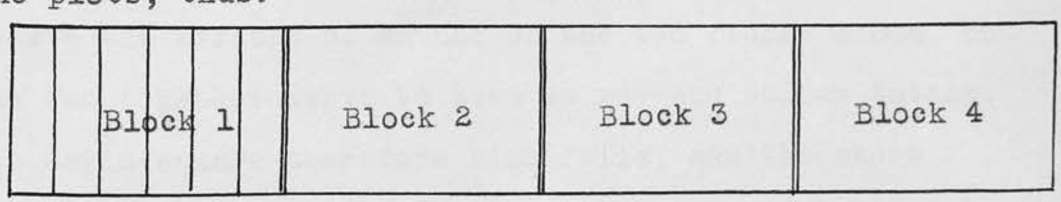
Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares				
			Analysis 1	Anal:2	Anal:3	Anal:4	Anal:5
XVI	Blocks	2	23.60		119.92		23.60
	(rows)	3	17.32		115.03		17.32
	Blocks	2		81.35		127.13	81.35
	(columns)	3		81.18		130.57	81.18
		2	21.18	13.96	9.14	8.24	12.59
	Error	3	22.56	14.58	10.35	8.41	14.19
	Percentage	2	8.0%	12.1%	18.5%	20.5%	13.4%
	Efficiency	3	12.3%	19.1%	26.9%	33.1%	19.6%
			10.2%	15.6%	22.7%	26.8%	
	Mean Percent: Effic:		12.9%		24.8%		16.5%
XVII	Blocks	1	123.32		165.11		123.32
	(rows)	3	115.46		156.11		115.46
	Blocks	1		15.57		145.18	15.57
	(columns)	3		14.53		141.78	14.53
		1	15.44	28.91	10.21	12.71	15.42
	Error	3	14.23	26.86	9.15	10.94	14.19
	Percentage	1	17.5%	9.3%	26.4%	21.2%	17.5%
	Efficiency	3	19.5%	10.4%	30.5%	25.4%	19.6%
			18.5%	9.8%	28.4%	23.3%	
	Mean Percent: Effic:		14.6%		25.8%		18.6%

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares Analysis 1	Anal: 2	Anal: 3	Anal: 4	Anal: 5
XVIII	Blocks	2	0.46		61.39		0.46
	(rows)	4	3.82		86.64		3.82
	Blocks	2		5.25		66.32	5.25
	(columns)	4		2.71		78.93	2.71
		2	13.37	12.77	5.75	5.13	14.53
	Error	4	16.20	16.34	5.75	6.81	18.13
	Percentage	2	12.6%	13.2%	29.4%	32.9%	11.6%
	Efficiency	4	14.6%	14.4%	40.3%	34.7%	13.0%
			13.6%	13.8%	34.8%	33.8%	
	Mean Percent: Effic:		13.7%		34.3%		12.3%
XIX	Blocks	1	22.39		22.82		22.39
	(rows)	4	21.87		18.55		21.87
	Blocks	1		13.28		24.78	13.28
	(columns)	4		7.05		27.80	7.05
		1	6.86	8.00	6.81	6.56	5.95
	Error	4	8.35	10.20	8.76	7.60	8.53
	Percentage	1	39.4%	33.8%	39.6%	41.2%	45.4%
	Efficiency	4	28.3%	23.1%	27.0%	31.1%	27.7%
			33.8%	28.4%	33.3%	36.2%	
	Mean Percent: Effic:		31.1%		34.8%		36.6%

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares: Analysis 1	Anal:2	Anal:3	Anal:4	Anal:5
XX	Blocks (rows)	3	12.53		33.96		12.53
		4	10.46		35.93		10.46
	Blocks (columns)	3		24.93		47.53	24.93
		4		26.14		41.39	26.14
	Error	3	13.08	11.53	10.40	8.71	11.39
		4	10.46	8.50	7.28	6.59	8.22
	Percentage Efficiency	3	21.3%	24.1%	26.7%	31.9%	24.4%
		4	22.8%	27.8%	32.4%	35.8%	28.7%
	Mean Percent: Effic		22.0%	26.0%	29.6%	33.8%	
			24.0%		31.7%		26.6%
XXI	Blocks (rows)	1	15.50		40.14		15.50
		3	16.00		44.21		16.00
	Blocks (columns)	1		15.04		46.46	15.04
		3		11.03		44.32	11.03
	Error	1	12.24	12.29	9.16	8.37	11.84
		3	10.53	11.15	7.00	6.99	10.47
	Percentage Efficiency	1	22.1%	22.0%	29.5%	32.3%	22.8%
		3	26.4%	24.9%	39.7%	39.8%	26.6%
	Mean Percent: Effic		24.2%	23.4%	34.6%	36.0%	
			23.8%		35.3%		24.7%

Fertility Grid No.	Source of Variation	Random Grid No.	Mean Squares Analysis 1	Anal:2	Anal:3	Anal:4	Anal:5
XXII	Blocks (rows)	1	39.53		36.21		39.53
		4	45.61		44.57		45.61
	Blocks (columns)	1		18.53		47.85	18.53
		4		23.32		53.32	23.32
	Error	1	7.58	10.21	8.00	6.54	6.02
		4	8.54	11.33	8.67	7.58	6.43
	Percentage Efficiency	1	35.6%	26.4%	33.8%	41.3%	44.9%
		4	27.6%	20.8%	27.2%	31.1%	36.7%
			31.6%	23.6%	30.5%	36.2%	
	Mean Percent: Effic:		27.6%		33.4%		40.8%
XXIII	Blocks (rows)	2	49.44		48.02		49.44
		4	48.75		49.21		48.75
	Blocks (columns)	2		2.91		39.30	2.91
		4		4.03		44.43	4.03
	Error	2	2.84	8.65	3.02	4.10	2.83
		4	3.53	9.12	3.48	4.07	3.46
	Percentage Efficiency	2	59.5%	19.5%	56.0%	41.2%	59.7%
		4	66.9%	25.9%	67.8%	58.0%	68.2%
			63.2%	22.7%	61.9%	49.6%	
	Mean Percent: Effic		43.0%		55.8%		64.0%

The average percentage efficiencies over the 23 different grids are: Long blocks, 25·1% ; Short blocks, 26·9%; Latin square, 51·5%, and while no great significance is attachable to such mean percentages, the generally superior efficiency of the Latin-square design is manifest. Clearly the Latin square can be inferior to a blocks design which uses either the rows or columns as blocks only when either the row or column mean square (or both) is less than the error mean square, and then not by very much. The means of the smaller and larger percentages for the long blocks are 14·4% and 35·8% respectively, and for the short blocks 24·1% and 29·8%. These figures give some indication of the greater reliability of the more compact blocks, for it must be remembered that the experimenter usually knows little or nothing about the fertility-gradients of the field and at best can only guess. Similar considerations suggest that, if long narrow plots are used, it would be advisable to place the blocks in a line perpendicular to the length of the plots, thus:-



In this manner, should the plots prove to have been placed transverse to the main fertility-slope, then at least some of the effects of soil heterogeneity will have been removed, but presumably less efficiently than if the main fertility-slope had been parallel to the length of the plots, Actually, the above type of layout is standard practice for simple experiments requiring cultivation or drilling.

As regards the individual fertility-grids, the following observations may be made:-

I and II : As expected, the long blocks provide the most and least favourable analyses. The wider blocks are not sufficiently sensitive to deal efficiently with such sudden fluctuations within the width of a single plot, even when the ridges are parallel to their longer sides ; but their efficiency improves when the ridges are less sudden.

III and IV : Long blocks could have no effect on a uniform ridge running diagonally across the field from corner to corner. In Grid III they would partially remove the effects of either of the two ridges alone, but the two together serve to even up row and column totals. The Latin square therefore also fails, and the short blocks are little better. In Grid IV there is a single but less sudden ridge which is best dealt with by the short blocks.

V : The results are similar to II , except that the short blocks show up better, as they usually will when it is a question of fertility-slope and not ridges.

VI : As in IV , but the efficiencies are higher and the Latin square especially good considering the moderate efficiencies of both types of long block.

VII : No type of block can cope with this type of simultaneous variation at right angles, but the Latin square again registers a high efficiency.

VIII : In their most favourable case the long blocks are highly efficient, but only the Latin square is independent of pre-knowledge of the grid. The short blocks fail to eliminate the ridge effects.

IX : As for VIII , except that the long blocks are less efficient for their most favourable case.

X and XI : Very similar to VII and IX respectively, but the short blocks show improvement because the ridges are less sudden.

XII , XIII , and XIV : As in III , no type of analysis is better than moderately efficient.

XV, XVI, and XVII : Similar to XII, XIII and XIV respectively, but the greater efficiency of the short blocks in dealing with the less sudden ridge is noticeable.

XVIII : As expected, long blocks and the Latin square all fail, but the short blocks are fairly efficient.

XIX, XX, and XXI : All types of analysis give moderate results, the Latin square being only slightly better than the long blocks. The short blocks are best, as is not surprising when it is considered how these grids were composed.

XXII and XXIII : The efficiencies for these two grids are fairly representative of the all-over trend.

Section 7.

A Multiple - Factor Experiment.

Suppose that we have p replications in randomised blocks of mn treatment-combinations, the latter consisting of m treatments of one type (let us say different varieties) and n treatments of a second type (say different fertilisers) in all possible combinations. Let the yields y_{ijk} (where i denotes variety , j fertiliser, and k replication) be arranged in a matrix thus :-

$$\left[\begin{array}{cccc|cccc|c|cccc}
y_{111} & y_{121} & \cdots & y_{1n1} & y_{211} & y_{221} & \cdots & y_{2n1} & \cdots & y_{m11} & y_{m21} & \cdots & y_{mn1} \\
y_{112} & y_{122} & \cdots & y_{1n2} & y_{212} & y_{222} & \cdots & y_{2n2} & \cdots & y_{m12} & y_{m22} & \cdots & y_{mn2} \\
\vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\
y_{11p} & y_{12p} & \cdots & y_{1np} & y_{21p} & y_{22p} & \cdots & y_{2np} & \cdots & y_{m1p} & y_{m2p} & \cdots & y_{mnp}
\end{array}\right]$$

Let us first consider the p replications of the mn treatment-combinations under a hypothesis of complete uniformity. By (4,2) we have

$$\sum_i\sum_j\sum_k(y_{ijk}-\bar{y})^2 = p\sum_i\sum_j(y_{ij0}-\bar{y})^2 + mn\sum_k(y_{00k}-\bar{y})^2 + \sum\sum\sum_{ijk}(y_{ijk}-y_{ij0}-y_{00k}+\bar{y})^2, \quad \ldots (7,1)$$

the three component sums of squares on the right-hand side having independent gamma-type distribution with mn-1, p-1, and (p-1)(mn-1) degrees of freedom respectively. Now consider the mn column-means of the above matrix. They are independent, normal variates with variance $\frac{\sigma^2}{p}$ (where σ^2 is the variance of the y_{ijk}), and may be regarded as being arranged in a matrix of m rows and n columns. Hence we have also by (4,2)

$$\sum_i\sum_j (y_{ij0}-\bar{y})^2 = n\sum_i (y_{i00}-\bar{y})^2 + m\sum_j (y_{0j0}-\bar{y})^2 + \sum_i\sum_j (y_{ij0}-y_{i00}-y_{0j0}+\bar{y})^2, \quad \ldots.(7,2)$$

and combining (7,1) and (7,2)

$$\sum_i\sum_j\sum_k (y_{ijk}-\bar{y})^2 = np\sum_i (y_{i00}-\bar{y})^2 + mp\sum_j (y_{0j0}-\bar{y})^2 + p\sum_i\sum_j (y_{ij0}-y_{i00}-y_{0j0}+\bar{y})^2$$
$$+ mn\sum_k (y_{00k}-\bar{y})^2 + \sum_i\sum_j\sum_k (y_{ijk}-y_{ij0}-y_{00k}+\bar{y})^2. \quad \ldots.(7,3)$$

If we write (7,2) in matrix notation as

$$\bar{y}'G\bar{y} = \bar{y}'H\bar{y} + \bar{y}'J\bar{y} + \bar{y}'K\bar{y}\ , \quad \ldots.(7,4)$$

where $\{\bar{y}\}$ is the vector $\{y_{110}\ y_{120} \ldots y_{1n0} \mid y_{210}\ y_{220} \ldots y_{2n0} \mid \ldots \mid y_{m10}\ y_{m20} \ldots y_{mn0}\}$ we know from Section 4 that each of the matrices G,H,J, and K is idempotent. We may also write (7,3) as

$$y'Ay = y'By + y'Cy + y'Dy + y'Ey + y'Fy \quad \ldots.(7,5)$$

where y is the vector $\{y_{111}\ y_{112} \ldots y_{11p} \mid y_{121}\ y_{122} \ldots y_{12p} \mid \ldots \mid y_{mn1}\ y_{mn2} \ldots y_{mnp}\}$ i.e. it is the vector y of previous sections, each element of which is now further partitioned into p elements, so that there are now mn subvectors of the type $\{y_{ij1}\ y_{ij2} \ldots y_{ijp}\}$. Now if, for example, K is the symmetrical matrix

$$\begin{bmatrix} k_{11} & k_{12} & k_{13} & \ldots \\ k_{12} & k_{22} & k_{23} & \ldots \\ k_{13} & k_{23} & k_{33} & \ldots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

of order $mn \times mn$, then the quadratic form $\bar{y}'K\bar{y}$, when referred to the vector y, must be

$$y' \cdot \frac{1}{p^2}\left[\begin{array}{cccc|cccc|c}
k_{11} & k_{11} & \cdots & k_{11} & k_{12} & k_{12} & \cdots & k_{12} & \\
k_{11} & k_{11} & \cdots & k_{11} & k_{12} & k_{12} & \cdots & k_{12} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
k_{11} & k_{11} & \cdots & k_{11} & k_{12} & k_{12} & \cdots & k_{12} & \\
\hline
k_{12} & k_{12} & \cdots & k_{12} & k_{22} & k_{22} & \cdots & k_{22} & \\
k_{12} & k_{12} & \cdots & k_{12} & k_{22} & k_{22} & \cdots & k_{22} & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \\
k_{12} & k_{12} & \cdots & k_{12} & k_{22} & k_{22} & \cdots & k_{22} & \\
\hline
 & \vdots & & & & \vdots & & &
\end{array}\right] y,$$

where each submatrix is of order pxp and has all elements the same. But $y'Dy = p\bar{y}'K\bar{y}$, so that D must be idempotent, since K is idempotent. Similarly the matrices B and C may also be proved idempotent.

From the results of Section 4 we know that the mean values of $\bar{y}'H\bar{y}$, $\bar{y}'J\bar{y}$ and $\bar{y}'K\bar{y}$ are $(m-1)\frac{\sigma^2}{p}$, $(n-1)\frac{\sigma^2}{p}$, and $(m-1)(n-1)\frac{\sigma^2}{p}$ respectively, so that the mean values of $y'By$, $y'Cy$ amd $y'Dy$ are $(m-1)\sigma^2$, $(n-1)\sigma^2$, and $(m-1)(n-1)\sigma^2$. Since B, C, and D are idempotent, it follows from Lemma 9, Cor. 2 that $y'By$, $y'Cy$, amd $y'Dy$ have gamma-type distribution with m-1, n-1, and (m-1)(n-1) degrees of freedom respectively. It is obvious that the independence of the quadratic forms on the right-hand side of (7,4) is not affected by referring them to the vector y. Hence, applying Lemma 7, Cor.1 to (7,1) and (7,2), we see that all the quadratic forms on the right-hand side of (7,5) are independent. Their mean squares must therefore be independent estimates of the variance σ^2.

An initial analysis of variance may be made as in §4, thus:-

Analysis of Variance.

Variation due to	D.F.	Sums of Squares	Mean Squares
Blocks	p-1	$mn\sum_{k}(y_{ook}-\bar{y})^2$	s_1^2
Treatment-combinations	mn-1	$p\sum_{i}\sum_{j}(y_{ijo}-\bar{y})^2$	s_2^2
Error	(p-1)(mn-1)	$\sum_{i}\sum_{j}\sum_{k}(y_{ijk}-y_{ijo}-y_{ook}+\bar{y})^2$	s_3^2
Total	mnp-1	$\sum_{i}\sum_{j}\sum_{k}(y_{ijk}-\bar{y})^2$	

Reverting now from the hypothesis of uniformity to the hypothesis of the randomised-blocks design (§4, p41), s_1^2 is an estimate of $\sigma'^2 + mn\,\sigma_B^2$, but on a null hypothesis s_2^2 and s_3^2 continue to be estimates of the random variance (now σ'^2). At this stage, therefore, a test could be made to see if treatments as a whole diverged significantly from the null hypothesis. But this is unnecessary, nor is it the object of the experiment. The sums of squares and degrees of freedom are now further subdivided as in (7,3).

Analysis of Variance.

Variation due to	D.F.	Sums of Squares	Mean Squares
Blocks	p-1	$mn\sum_k (y_{ook}-\bar{y})^2$	s_1^2
Treatment-combinations			
Varieties	m-1	$np\sum_i (y_{ioo}-\bar{y})^2$	s_4^2
Fertilisers	n-1	$mp\sum_j (y_{ojo}-\bar{y})^2$	s_5^2
Residuals ("Interaction")	(m-1)(n-1)	$p\sum_i\sum_j (y_{ijo}-y_{ioo}-y_{ojo}+\bar{y})^2$	s_6^2
Total	mn-1	$p\sum_i\sum_j (y_{ijo}-\bar{y})^2$	s_2^2
Error	(p-1)(mn-1)	$\sum_i\sum_j\sum_k (y_{ijk}-y_{ijo}-y_{ook}+\bar{y})^2$	s_3^2
Total	mnp-1	$\sum_i\sum_j\sum_k (y_{ijk}-\bar{y})^2$	

On a null hypothesis s_4^2, s_5^2, and s_6^2 are all independent estimates of the random variance and hence may be tested against s_3^2. By the tests with s_4^2 and s_5^2 the significance of variety and fertiliser means respectively is examined. The residual sum of squares under treatment-combinations is the sum of squares due to interaction between varieties and fertilisers. The interaction, AB, between two sets of treatments, A and B, is a measure of the variation in (yield-) response to treatments A when combined with the different treatments of set B, the exact measure adopted being a matter of definition. To show that the residual sum of squares for treatment-combinations does indeed provide a measure of the interaction between varieties and fertilisers, let us define the response to, or "effect" of the i^{th} treatment of set A as $y_{i00}-\bar{y}$, or what is the same thing, $\sum_j (y_{ij0}-y_{0j0})$. It is clear that a measure of the variation of the response to the i^{th} treatment over the different treatments of set B is given by the sum of squares $\sum_j (y_{ij0}-y_{i00}-y_{0j0}+\bar{y})^2$, so that for all treatments of set A we have the sum of squares $\sum_i\sum_j (y_{ij0}-y_{i00}-y_{0j0}+\bar{y})^2$ as a measure of the interaction AB, and this is identical with the sum of squared residuals for treatment-combinations. Also, since this expression is symmetrical in i and j, it follows that interaction is a symmetrical relationship, interaction between fertilisers and varieties being the same as interaction between varieties and fertilisers. If there is no interaction, the residuals $(y_{ij0}-y_{i00}-y_{0j0}+\bar{y})$ will be normally distributed with variance equal to $\frac{1}{r}(m-1)(n-1)$ times the random variance of the experiment, this being so regardless of the significance of either set of treatments. The usual test will therefore determine whether the estimate of variance, s_6^2, derived from the interaction mean square is significantly different from that of the error mean square. If so, the interaction is said to be significant, and it becomes necessary to study individual elements of the "interaction-table" given by the matrix $[y_{ij0}]$. It is noteworthy that the residual sum of squares of a randomised-block experiment is really the interaction between treatments and blocks.

Section 8.

The Split-Plot Experiment.

Any type of experimental design in agriculture may have an additional type of treatment comparison appended merely by subdividing each plot (now known as a "whole-plot") into a number of sub-plots equal to the number of sub-plot treatments (unless there is confounding), which are as usual randomised within each whole-plot. As an example, let us consider a randomised-block experiment of m varieties replicated n times, each whole-plot of which is further subdived into p fertiliser treatments. Let the yields be $[y_{ijk}]$, where i represents variety , j block, and k fertiliser, the matrix $[y_{ijk}]$ being of order mn x p. By Lemma 8, we have

$$\sum_i\sum_j\sum_k(y_{ijk}-\bar{y})^2=\sum_i\sum_j\sum_k(y_{ijk}-y_{ij0})^2+p\sum_i\sum_j(y_{ij0}-\bar{y})^2, \quad \ldots (8,1)$$

and under a hypothesis of complete uniformity throughout the experiment we have by the results of §1 the following equation of mean values.

$$(mnp-1)\sigma^2 = mn(p-1)\sigma^2 + (mn-1)\sigma^2, \qquad \ldots (8,2)$$

where σ^2 is the variance of a single sub-plot. We may therefore perform an initial analysis of variance, thus:-

Analysis of Variance.

Variation	D.F.	Sums of Squares	Mean Squares
Between whole-plots	mn-1	$p\sum_i\sum_j(y_{ijo}-\bar{y})^2$	s_1^2
Within whole- plots	mn(p-1)	$\sum_i\sum_j\sum_k(y_{ijk}-y_{ijo})^2$	s_2^2
Total	mnp-1	$\sum_i\sum_j\sum_k(y_{ijk}-\bar{y})^2$	

However, the mp treatment-combinations are not randomised over the whole of each block owing to the restriction imposed by the design of the experiment. This restriction, that all combinations of a particular variety with the different fertilisers should occur in a single whole-plot, is the same with respect to the sub-plot treatments as the blocks restriction in a randomised-blocks design. We must therefore replace the hypothesis of complete uniformity with the hypothesis (similar to that of §4, P41) that the whole-plots have, a priori, different means, even if treatments are still assumed to have no differential effect. Combining this null hypothesis with the hypothesis of the randomised-block design with respect to the whole plots, we have that each variate y_{ijk} is equal to $\mu+\beta_j+\xi_{ijo}+\xi_{ijk}$, where β_j is a normal variate (constant over the j^{th} block) with mean at the origin and variance σ_B^2, ξ_{ijo} is the random whole-plot variate (normal, constant for all sub-plots in the $(i,j)^{th}$ whole-plot, with mean at the origin and variance σ_p^2), and ξ_{ijk} is the random sub-plot variate (normal, with mean at the origin and variance σ_S^2). Hence, in the above table s_2^2 is an estimate of σ_S^2.

We may now further subdivide the sum of squares between whole-plots in accordance with the ordinary randomised-block analysis of §4. By (4,2) we have

$$p\sum_i\sum_j(y_{ijo}-\bar{y})^2 = np\sum_i(y_{ioo}-\bar{y})^2 + mp\sum_j(y_{ojo}-\bar{y})^2 + p\sum_i\sum_j(y_{ijo}-y_{ioo}-y_{ojo}+\bar{y})^2, \quad \ldots (8,3)$$

the degrees of freedom of the components on the right-hand side being, respectively, m-1, n-1, and (m-1)(n-1). The final component, when divided by its degrees of freedom, yields a mean square (s_5^2, below) which is an estimate of $\frac{1}{p}\sigma_1^2$, where $\sigma_1^2 (= p\sigma_S^2 + p^2\sigma_P^2)$ is the whole-plot random variance, i.e. s_5^2 is an estimate of $\sigma_S^2 + p\sigma_P^2$. The table is as follows:-

Analysis of Variance

Variation	D.F.	Sums of Squares	Mean Squares
Between whole-plots			
Blocks	n-1	$mp\sum_j(y_{ojo}-\bar{y})^2$	s_3^2
Varieties	m-1	$np\sum_i(y_{ioo}-\bar{y})^2$	s_4^2
Error(1)	(m-1)(n-1)	$p\sum_i\sum_j(y_{ijo}-y_{ioo}-y_{ojo}+\bar{y})^2$	s_5^2
Within whole-plots	mn(p-1)	$\sum_i\sum_j\sum_k(y_{ijk}-y_{ijo})^2$	s_2^2
Total	mnp-1	$\sum_i\sum_j\sum_k(y_{ijk}-\bar{y})^2$	

The mean square for blocks (s_3^2) is an estimate of $\sigma_S^2 + p\sigma_P^2 + m\sigma_B^2$

The varieties mean square (s_4^2) may be tested as usual against that for error (1).

The sum of squares for deviations from whole-plot means, $\sum_i\sum_j\sum_K (y_{ijK} - y_{ijo})^2$, may also be further subdivided, for we may regard the deviations as a matrix of p rows and mn columns, thus:-

$$\begin{bmatrix} y_{111} - y_{110} & y_{121} - y_{120} & y_{131} - y_{130} & \cdots\cdots & y_{ij1} - y_{ijo} & \cdots\cdots & y_{mn1} - y_{mno} \\ y_{112} - y_{110} & y_{122} - y_{120} & y_{132} - y_{130} & \cdots\cdots & y_{ij2} - y_{ijo} & \cdots\cdots & y_{mn2} - y_{mno} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{11p} - y_{110} & y_{12p} - y_{120} & y_{13p} - y_{130} & \cdots\cdots & y_{ijp} - y_{ijo} & \cdots\cdots & y_{mnp} - y_{mno} \end{bmatrix}$$

The general mean is $\frac{1}{mnp}(\sum_i\sum_j\sum_K y_{ijK} - p\sum_i\sum_j y_{ijo}) = 0$, and the mean of row k is $y_{ooK} - \frac{1}{mn}\sum_i\sum_j y_{ijo} = y_{ooK} - \bar{y}$. Hence, by Lemma 8, we have,

$$\sum_i\sum_j\sum_K (y_{ijK} - y_{ijo})^2 = mn\sum_K (y_{ooK} - \bar{y})^2 + \sum_i\sum_j\sum_K (y_{ijK} - y_{ijo} - y_{ooK} + \bar{y})^2. \qquad \ldots\ldots (8,4)$$

A further subdivision may now be made of the final sum of squares on the right-hand side of (8,4). Consider the residuals $y_{ijK} - y_{ijo} - y_{ooK} + \bar{y}$ arranged in rows according to the suffix j. The general mean of such a matrix is $\frac{1}{mnp}(\sum_i\sum_j\sum_K y_{ijK} - p\sum_i\sum_j y_{ijo} - mn\sum_K y_{ooK} + mnp\,\bar{y}) = 0$. The general column-mean is $y_{ioK} - y_{ioo} - y_{ooK} + \bar{y}$. Hence, by Lemma 8,

$$\sum_i\sum_j\sum_K (y_{ijK} - y_{ijo} - y_{ooK} + \bar{y})^2 = n\sum_i\sum_K (y_{ioK} - y_{ioo} - y_{ooK} + \bar{y})^2 + \sum_i\sum_j\sum_K (y_{ijK} - y_{ijo} - y_{ioK} + y_{ioo})^2 \ldots (8,5)$$

and combining (8,4) and (8,5) we arrive at

$$\sum_i\sum_j\sum_K (y_{ijK} - y_{ijo})^2 = mn\sum_K (y_{ooK} - \bar{y})^2 + n\sum_i\sum_K (y_{ioK} - y_{ioo} - y_{ooK} + \bar{y})^2$$
$$+ \sum_i\sum_j\sum_K (y_{ijK} - y_{ijo} - y_{ioK} + y_{ioo})^2 \qquad \ldots\ldots (8,6)$$

The first sum of squares on the right-hand side will be recognised as that for fertiliser-means and the second as that for interaction between varieties and fertilisers as in Section 7, Page 84.

The next step is to examine the independence of the three component sums of squares of (8,6). A deviation of a fertiliser-mean from the general mean, e.g. $y_{001} - \bar{y}$, may be written as $a'y$, where y is the same vector as in (7,5) and $a' = \frac{1}{mnp}\left[p-1\ -1\ -1 \ldots -1 \,\vdots\, p-1\ -1\ -1 \ldots -1 \,\vdots\, \text{etc., mn sub-vectors in all}\right]$. A residual of type $y_{i0k} - y_{i00} - y_{00k} + \bar{y}$, e.g. $y_{102} - y_{100} - y_{002} + \bar{y}$, may be written as $b'y = \frac{1}{mnp}\left[1-m\ \ mp-m-p+1\ \ 1-m \ldots 1-m \,\vdots\, \text{and } n-1 \text{ similar subvectors} \,\vdots\, 1\ \ 1-p\ \ 1 \ldots 1 \,\vdots\, \text{in all } (m-1)n \text{ similar subvectors}\right] y$. A residual of type $y_{ijk} - y_{ij0} - y_{i0k} + y_{i00}$, where we will as an example take i=1, j=2, k=3, may be written as $c'y$, where $c' = \frac{1}{mnp}\left[m\ m\ \ m-mp\ \ m \ldots m \,\vdots\, m-mn\ \ m-mn\ \ m+mnp-mn-mp\ \ m-mn \ldots m-mn \,\vdots\, m\ m\ \ m-mp\ \ m \ldots m \,\vdots\, \text{and similarly up to the } n^{th} \text{ subvector, all the rest null}\right]$.

$$a'b = \frac{1}{m^2n^2p^2}\left[n(1-m)(p-1) - n(m-1)(p-1) - n(1-m)(p-2) + n(p-1)(m-1) - n(1-p)(m-1) - n(p-2)(m-1)\right] = 0.$$ $\therefore$ a and b are orthogonal.

$$a'c = \frac{1}{m^2n^2p^2}\left[m(p-1)(n-1) - m(p-2)(n-1) - m(1-p)(n-1) + m(1-n)(p-1) - m(1-n)(p-2) - m(n-1)(p-1)\right] = 0.$$ $\therefore$ a and c are orthogonal.

$$b'c = \frac{1}{m^2n^2p^2}\left[m(1-m)(n-1)(p-2) + m(m-1)(p-1)(n-1) + m(1-p)(1-m)(n-1) + m(1-n)(1-m)(p-2) + (m-1)(p-1)(1-n)\,m + m(n-1)(p-1)(1-m)\right] = 0$$

$\therefore$ b and c are orthogonal. Similar orthogonality may be proved for all values of i, j, and k, so that from Lemmas 5 and 6 we deduce the independence of the three component sums of squares of (8, 6).

In order to find their mean values, we have

$a'a = \frac{1}{m^2n^2p^2}\left[(p-1)^2 mn - (p-1)mn\right] = \frac{1}{mnp}(p-1)$.

$\therefore$ The mean value of $mn\sum_{k}(y_{ook}-\bar{y})^2$ is $p-1$, or, unstandardised $(p-1)\sigma_s^2$, since σ_s^2 is the random variance of a single sub-plot yield.

Also $b'b = \frac{1}{m^2n^2p^2}\left[(1-m)^2(p-1)n + (m-1)^2(p-1)^2 n + (p-1)(m-1)n + (1-p)^2(m-1)n\right] = \frac{1}{mnp}(m-1)(p-1)$. $\therefore$ The mean value of $n\sum_{i}\sum_{k}(y_{iok}-y_{ioo}-y_{ook}+\bar{y})^2$ is $(m-1)(p-1) = (m-1)(p-1)\sigma_s^2$ (unstandardised). and $c'c = \frac{1}{m^2n^2p^2}\left[m^2(p-1)(n-1) + m^2(1-p)^2(n-1) + m^2(1-n)^2(p-1) + m^2(p-1)^2(n-1)^2\right] = \frac{1}{np}(p-1)(n-1)$, hence the mean value of the residual sum of squares is $m(n-1)(p-1)$ or $m(n-1)(p-1)\sigma_s^2$ (unstandardised). We thus have the following equation of mean values.

$$mn(p-1)\sigma_s^2 = (p-1)\sigma_s^2 + (m-1)(p-1)\sigma_s^2 + m(n-1)(p-1)\sigma_s^2 \qquad \ldots(8,7)$$

If we now write (8,6) in the form

$$y'Ay = y'By + y'Cy + y'Dy, \qquad \ldots.(8,8)$$

it is evident that the matrices B and C may be proved idempotent in the same manner as B, C, and D of (7,5), using a vector $\left\{y_{101}\, y_{102} \cdot\cdot y_{10p} \,\vdots\, y_{201}\, y_{202} \cdots y_{20p} \,\vdots\, \cdots \,\vdots\, y_{m01}\, y_{m02} \cdot\cdot y_{m0p}\right\}$

Also $D = K'K$ where K is "the direct sum" of m sub-matrices each equal to

$$L = M_{np} - \begin{bmatrix} M_p & & \\ & M_p & \\ & & \ddots & \\ & & & M_p \end{bmatrix} - \frac{1}{n}\begin{bmatrix} I_p & I_p & \ldots & I_p \\ I_p & I_p & \ldots & I_p \\ \vdots & \vdots & & \vdots \\ I_p & I_p & \ldots & I_p \end{bmatrix} + I_{np}$$

L is of order np x np, and M_{np} is the matrix of order np x np with all elements $\frac{1}{np}$, I_p is the unit matrix of order p x p, etc. /

$$L'L = M_{np} + \begin{bmatrix} M_p & & & \\ & M_p & & \\ & & \ddots & \\ & & & M_p \end{bmatrix} + \frac{1}{n}\begin{bmatrix} I_p & I_p & \cdots & I_p \\ I_p & I_p & \cdots & I_p \\ \vdots & \vdots & & \vdots \\ I_p & I_p & \cdots & I_p \end{bmatrix} + I_{np} - 2M_{np} - 2M_{np}$$

$$+ 2M_{np} + 2M_{np} - 2\begin{bmatrix} M_p & & & \\ & M_p & & \\ & & \ddots & \\ & & & M_p \end{bmatrix} - \frac{2}{n}\begin{bmatrix} I_p & I_p & \cdots & I_p \\ I_p & I_p & \cdots & I_p \\ \vdots & \vdots & & \vdots \\ I_p & I_p & \cdots & I_p \end{bmatrix}$$

$$= L$$

Hence $K'K = K$, and since K is symmetric it is also idempotent.

It follows that on a null hypothesis the three quadratic forms on the right-hand side of (8,8) yield independent estimates of σ_s^2 with degrees of freedom equal to the respective coefficients of (8,7). It is also apparent that the residual quadratic form, $y'Dy$ will yield an estimate of σ_s^2 even if the null hypothesis can no longer be regarded as valid. The following is the complete table of analysis of variance:-

Variation	D.F.	Sums of Squares	Mean Squares
Between whole-plots			
Blocks	n-1	$mp\sum_j (y_{ojo}-\bar{y})^2$	s_3^2
Varieties	m-1	$np\sum_i (y_{ioo}-\bar{y})^2$	s_4^2
Error(1)	$\underline{(m-1)(n-1)}$	$\underline{p\sum_i\sum_j (y_{ijo}-y_{ioo}-y_{ojo}+\bar{y})^2}$	s_5^2
Total	mn-1	$p\sum_i\sum_j (y_{ijo}-\bar{y})^2$	s_1^2
Within whole-plots			
Fertilisers	p-1	$mn\sum_k (y_{ook}-\bar{y})^2$	s_6^2
Interaction	(m-1)(p-1)	$n\sum_i\sum_k (y_{iok}-y_{ioo}-y_{ook}+\bar{y})^2$	s_7^2
Error(2)	$\underline{m(n-1)(p-1)}$	$\underline{\sum_i\sum_j\sum_k (y_{ijk}-y_{ijo}-y_{iok}+y_{ioo})^2}$	s_8^2
Total	mn(p-1)	$\sum_i\sum_j\sum_k (y_{ijk}-y_{ijo})^2$	s_2^2
Grand Total	mnp-1	$\sum_i\sum_j\sum_k (y_{ijk}-\bar{y})^2$	

The varieties mean square (s_4^2) may be tested against s_5^2, s_6^2 and s_7^2 against s_8^2. It may be desired to test s_5^2 against s_8^2 in order to see if there is any significant component of whole-plot variance. This is valid, since s_5^2 is independent of s_2^2 and hence, by Lemma 7, Cor.1, of s_8^2. In general s_5^2 will be greater than s_8^2 , but sometimes it will be less, indicating that the estimate of variance of whole-plot means is less than would have been expected from ordinary random-sampling of the population of sub-plot random variates. This can arise from accidents of sampling, or it may mean that there is competition between the sub-plots within each whole-plot.

Should the variety and fertiliser effects prove to be significant, it will be necessary to compare the varietal means (y_{ioo}) among themselves, and likewise the fertiliser means (y_{ook}). Also should the interaction between varieties and fertilisers prove significant, we will wish to compare means of the individual treatment-combinations (y_{iok}). We therefore proceed to allot standard errors for the various types of comparison. The interaction table is as follows:-

Fertilisers / Varieties	F_1	F_2	...	F_k	...	F_p	Varietal Means
V_1	y_{101}	y_{102}	...	y_{10k}	...	y_{10p}	y_{100}
V_2	y_{201}	y_{202}	...	y_{20k}	...	y_{20p}	y_{200}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
V_i	y_{i01}	y_{i02}	...	y_{i0k}	...	y_{i0p}	y_{i00}
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
V_m	y_{m01}	y_{m02}	...	y_{m0k}	...	y_{m0p}	y_{m00}
Fertiliser Means	y_{001}	y_{002}	...	y_{00k}	...	y_{00p}	—

The standard error for comparing varietal means (y_{i00}) is $\frac{s_5}{\sqrt{np}}$, since s_5^2 is an estimate of p times the error variance of a whole-plot mean. For comparing fertiliser means it is $\frac{s_8}{\sqrt{mn}}$. For the comparison of single means in the same row of the interaction-table the standard error is $\frac{s_8}{\sqrt{n}}$, and the same standard error (adjusted to suit the numbers in the groups) applies to groups of means in the same row (e.g. $y_{101}+y_{102}$ compared with $y_{105}+y_{106}$) or to groups in different columns but comprising the same rows (e.g. $y_{101}+y_{201}+y_{301}$, compared with $y_{103}+y_{203}+y_{303}$). All the above types of comparison may be made by the t-test with the appropriate number of degrees of freedom, but any other type of comparison involves the component of error variance due to whole-plots, namely σ_p^2, and, as has been shown by Nair (18) the exact test for such a comparison is the Fisher-Behrens test, named by Sukhatme (20), who has tabulated significance levels, the d-test. Tables of "d" are also reproduced in Fisher and Yates (14). The sample estimate of σ_p^2 is $\frac{s_5^2 - s_8^2}{p}$ for a single whole-plot, or $\frac{s_5^2 - s_8^2}{np}$ for a whole-plot mean. Thus for comparing single means not in the same row (e.g. y_{101} and y_{201} or y and y) the standard error is $\sqrt{\frac{s_8^2}{n} + \frac{s_5^2 - s_8^2}{np}} = \sqrt{\frac{1}{np}(p-1)\ s_8^2 + s_5^2}$, and for comparing groups of means from two different rows (e.g. the mean of $y_{101} + y_{102} + y_{103}$ with the mean of $y_{201} + y_{202} + y_{203}$) the standard error is $\sqrt{\frac{s_8^2}{nq} + \frac{s_5^2 - s_8^2}{np}}$, where q is the number of means in the group. The whole-plot component in this formula remains constant whatever the value of q, provided only single rows are involved. If the t-test is used with such a combined estimate of standard error in order to make an approximate test, the number of degrees of freedom will be that of s_5^2 i.e. $(m-1)(n-1)$. If $s_5^2 \leqq s_8^2$, the question of a component of error due to whole-plots does not arise.

Section 9.

Orthogonal Subdivision of Sets of Degrees of Freedom

Let $\{y\}$ be a vector of treatment-means. Since, in the calculation of the treatment sum of squares in the analysis of variance, adjustment is made for the number of replications of the experiment, we may for simplicity and without loss of generality consider each element, y, of this vector as being the yield of a single replication. $\{y\}$ may be divided into sub-vectors according to the number of factors in the experiment and the respective levels of each factor. For example, in a three-factor experiment with p, q, and r levels, we may arrange $\{y\}$ as pq subvectors of r elements each.

Any subdivision Ay, where A is an orthogonal matrix of order n x n (n being the total number of treatment-combinations) and the elements of its first row are all unity but normalised by division by $\sqrt{n}$, subdivides the n-1 degrees of freedom into n-1 separate orthogonal degrees of freedom by means of the linear functions of the last n-1 rows of A. Such a subdivision is called a complete orthogonal set. If the variates $\{y\}$ are standardised, the variance matrix of the yields is I, so that the variance matrix of the transformed variates Ay is, by Lemma 5, also I. This proves that each linear function is statistically independent of all the others, and hence in the case of more than one replication their squares are, by Lemma 7, Cor. 1, independent of all the other sums of squares in the analysis. It also proves that each of the squares of the linear functions has a mean value equal to the intrinsic or random variance of the experiment. Moreover, the rank of the matrix of the quadratic form corresponding to any such square is one, since the matrix is of the form $\ell\ell'$, where ℓ' is a row vector. It follows that each of the squares of the normalised linear functions has gamma-type distribution with one degree of freedom and may be tested for significance against the error mean square of the experiment, or, since only one degree of freedom is involved, the t-test may be applied directly to the normalised linear function itself.

Since $A \equiv [a_{ij}]$ is orthogonal, $\sum_j a_{ij}a_{kj} = 0$ $(i \neq k)$, and since each element of row 1 is $\frac{1}{\sqrt{n}}$, $\sum_j a_{ij} = 0$ $(i \neq 1)$. Also $\sum_j a_{ij}^2 = 1$. In other words, in addition to the condition $\sum_j a_{ij}a_{kj} = 0$ $(i \neq k)$, the coefficients of each linear function must sum to zero and must be normalised. Algebraically, the sum of the squares corresponding to each individual degree of freedom of a complete orthogonal set must always equal the total sum of squares for the treatments under consideration. This is easily proved. The column-vector of a complete orthogonal set is Ay (including row 1, though this is not one of the set).

Now $y'A'Ay = \sum_j (\sum_i a_{ij}\, y)^2$

$= \text{The sum of squares of the orthogonal set} + (\Sigma y)^2/n$

But $y'A'Ay = y'y$ (since A is orthogonal) $= \sum y^2$.

Hence the **sum** of squares of the orthogonal set

$= \Sigma y^2 - (\Sigma y)^2/n = \Sigma y^2 - \bar{y}\Sigma y = \Sigma (y-\bar{y})^2$

= treatment sum of squares.

It may be observed that since $|A| = 1 \neq 0$, the n-1 linear functions are linearly as well as statistically independent, but that if an n^{th} linear function were chosen according to the same conditions, it would not be independent of the others (Aitken, 1) and would in fact merely repeat one of the functions already chosen. On the other hand an orthogonal set need not be complete. Suppose A_p were a matrix of order pxn ($p<n$), the first row of which must be as before a normalised vector of unit elements (representing the degree of freedom taken up in fixing the general mean). Then, if A_p possesses the orthogonal property $A_p A_p' = I$, the treatment sum of squares is subdivided into p-1 single degrees of freedom, each independent of the error sum of squares, the proof being similar to that above for the complete orthogonal set. If we denote by A_{p-1} the matrix A_p without its first row, then $A'_{p-1}A_{p-1}$, the matrix of the quadratic form comprising

the sum of squares of the p-1 linear functions, has the same rank as A_{p-1}, namely p-1. $A'_{p-1} A_{p-1}$ is also idempotent, since $A'_{p-1} A_{p-1} A'_{p-1} A_{p-1} = A'_{p-1} A_{p-1}$ by the orthogonal property. Hence, since the quadratic form has p-1 degrees of freedom, both the trace and rank of its matrix, $A'_{p-1} A_{p-1}$, must be p-1. Also the matrix of the quadratic form comprising the total treatment sum of squares is known by past results to be idempotent with both trace and rank equal to n-1. If A is the matrix of the complete orthogonal set corresponding to A_p and $A = \begin{bmatrix} A_p \\ \hdashline A_{n-p} \end{bmatrix}$, then, since $A'A = A'_p A_p + A'_{n-p} A_{n-p}$, it follows that $A'_{n-p} A_{n-p}$ must be the matrix of the residual quadratic form. But A_{p-1} and A_{n-p} are such that $A'_{p-1} A_{n-p} = A'_{n-p} A_{p-1} = 0$, so that it follows by Lemma 10 (Cor.) that the residual sum of squares has n-p degrees of freedom and its mean square yields an estimate of variance independent of each of those of the orthogonal set and hence, by Lemma 7, Cor. 1, of all the other mean squares of the analysis.

There is an infinite number of complete orthogonal sets for any given set of degrees of freedom, but any such subdivision of the treatment sum of squares should conform to a predetermined plan of analysis consistent with the design of the experiment, or it will not be statistically useful. Indeed, provided this condition is satisfied a set of degrees of freedom may be subdivided non-orthogonally, for any normalised row-vector when applied to the vector of treatment-means, provided that it is independent of the error sum of squares of the experiment, gives rise to a t-test or to an F-test with one degree of freedom. As en example suppose that in a simple randomised-blocks experiment we wish to compare treatment 1 with both treatment 2 and treatment 3. The vector corresponding to the comparison of treatments 1 and 2 is, after the manner of §4, and ignoring the normalising factor,

$\{1\ 1 \ldots\ 1 \mid -1\ -1 \ldots -1 \mid \text{remaining sub-vectors null}\} \equiv d,$

and that for the comparison of treatments 1 and 3 is

$\{1\ 1 \ldots\ 1 \mid 0\ 0 \ldots\ 0 \mid -1\ -1 \ldots -1 \mid \text{the rest null}\} \equiv e.$

Both d and e (and, in fact, the vector of any linear function of the treatment means) may easily be proved independent of the vector of any residual, but they are not orthogonal to one another.

Section 10.

Interactions in General

Let us consider an experiment with three factors A,B,C at m, n, and p levels respectively. As in the previous section we regard the treatment means $\{y_{ijk}\}$, where i, j, k represent levels of A, B, C respectively, as being the result of a single replication.

By (7,3) we have

$$\sum_i\sum_j\sum_k(y_{ijk}-\bar{y})^2 = np\sum_i(y_{ioo}-\bar{y})^2 + mp\sum_j(y_{ojo}-\bar{y})^2 + p\sum_i\sum_j(y_{ijo}-y_{ioo}-y_{ojo}+\bar{y})^2$$
$$+ mn\sum_k(y_{ook}-\bar{y})^2 + \sum_i\sum_j\sum_k(y_{ijk}-y_{ijo}-y_{ook}+\bar{y})^2 \qquad \ldots\ldots(10,1$$

The sums of squares on the right-hand side represent in order the "main effect" of A, main effect of B, interaction AB, main effect of C, and residuals. Their degrees of freedom are m-1, n-1, (m-1)(n-1), p-1, and (p-1)(mn-1) respectively. Exactly as in §8, Page 88 , we now arrange the mnp residuals of (10,1) in a matrix of order n x mp with rows according to the suffix j, and deduce that

$$\sum_i\sum_j\sum_k(y_{ijk}-y_{ijo}-y_{ook}+\bar{y})^2 = n\sum_i\sum_k(y_{iok}-y_{ioo}-y_{ook}+\bar{y})^2+\sum_i\sum_j\sum_k(y_{ijk}-y_{ijo}-y_{iok}+y_{ioo})^2, \quad \ldots(10,2$$

where the sums of squares on the right-hand side have been seen to correspond to the interaction AC and residuals with (m-1)(p-1) and m(n-1)(p-1) degrees of freedom respectively.

Now consider the residuals $y_{ijk}-y_{ijo}-y_{iok}+y_{ioo}$ arranged in a matrix of order m x np with rows corresponding to the suffix i, thus:-

$$\begin{bmatrix} y_{111}-y_{110}-y_{101}+y_{100} & y_{112}-y_{110}-y_{102}+y_{100} & \cdots y_{1jk}-y_{1j0}-y_{10k}+y_{100} & \cdots y_{1np}-y_{1n0}-y_{10p}+y_{100} \\ y_{211}-y_{210}-y_{201}+y_{200} & y_{212}-y_{210}-y_{202}+y_{200} & \cdots y_{2jk}-y_{2j0}-y_{20k}+y_{200} & \cdots y_{2np}-y_{2n0}-y_{20p}+y_{200} \\ \vdots & \vdots & \vdots & \vdots \\ y_{m11}-y_{m10}-y_{m01}+y_{m00} & y_{m12}-y_{m10}-y_{m02}+y_{m00} & \cdots y_{mjk}-y_{mj0}-y_{m0k}+y_{m00} & \cdots y_{mnp}-y_{mn0}-y_{m0p}+y_{m00} \end{bmatrix}$$

The general mean of this array $= \frac{1}{mnp}\left(\sum_i\sum_j\sum_k y_{ijk} - p\sum_i\sum_j y_{ij0} - n\sum_i\sum_k y_{i0k} + np\sum_i y_{i00}\right) = 0$. The general column mean $= y_{0jk} - y_{0j0} - y_{00k} + \bar{y}$, so that an application of Lemma 8 gives

$$\sum_i\sum_j\sum_k(y_{ijk}-y_{ij0}-y_{i0k}+\bar{y})^2 = m\sum_j\sum_k(y_{0jk}-y_{0j0}-y_{00k}+\bar{y})^2 + \sum_i\sum_j\sum_k(y_{ijk}-y_{ij0}-y_{i0k}-y_{0jk}+y_{i00}+y_{0j0}+y_{00k}-\bar{y})^2, \quad \ldots\ldots(10,\ 3)$$

where the first sum of squares on the right-hand side is seen to be that due to the interaction BC with $(n-1)(p-1)$ degrees of freedom, as may be proved by the same method as that used for the sums of squares due to the interactions AB and AC.

A residual of type $y_{ijk}-y_{ij0}-y_{i0k}-y_{0jk}+y_{i00}+y_{0j0}+y_{00k}-y$ (taking as an example $i=1$, $j=2$, $k=3$) may be written as $a'y$ where $a' = \frac{1}{mnp}\Big[m-1\ \ m-1\ \ m+p-mp-1\ \ m-1 \ldots m-1 \,\vdots\, (m+n-mn-1)$ $(m+n-mn-1)\ \ (m+n+p+mnp-mn-np-mp-1)(m+n-mn-1) \ldots (m\ n-mn-1) \,\vdots$ then repeating the first sub-vector up to the n^{th} subvector $\vdots\vdots$ $-1\ -1\ p-1\ -1 \ldots -1 \,\vdots\, n-1\ \ n-1\ \ n+p-np-1\ \ n-1 \ldots n-1 \,\vdots\, -1\ -1\ p-1\ -1 \ldots \vdots$ and so on up to the $2n^{th}$ subvector, the remainder repeating the $(n+1)^{th}$ to the $2n^{th}\Big]$

A residual of type $y_{0jk}-y_{0j0}-y_{00k}+\bar{y}$ (e.g. $y_{011}-y_{010}-y_{001}+\bar{y}$) may be written as $b'y$ where

$b' = \frac{1}{mnp}\Big[(n-1)(p-1)\ \ 1-n\ 1-n\ \ldots\ 1-n \,\vdots\, 1-p\ 1\ 1\ldots 1 \,\vdots\, 1-p\ 1\ 1\ldots 1 \,\vdots$ till n^{th} subvector, then repeating the 1st to the n^{th} up to the $mn^{th}\Big]$

Now $m^2n^2p^2 a'b = (m-1)(n-1)(p-1) - (m-1)(n-1)(p-2) + (m-1)(n-1)(p-1) + (m-1)(n-1)(p-1) - (m-1)(n-1)(p-2) + (m-1)(n-1)(p-1) - (m-1)(p-1)(n-2) + (m-1)(n-2)(p-2) - (m-1)(p-1)(n-2) - (m-1)(n-1)(p-1) + (m-1)(n-1)(p-2) - (m-1)(n-1)(p-1) - (m-1)(n-1)(p-1) + (m-1)(n-1)(p-2) - (m-1)(n-1)(p-1) + (m-1)(p-1)(n-2) - (m-1)(n-2)(p-2) + (m-1)(p-1)(n-2) = 0$

$\therefore$ a and b are orthogonal, and this may similarly be proved for any residual $a'y$ and any residual $b'y$. Hence, writing (10,3) as

$$y'A'Ay = y'B'By + y'C'Cy\ , \qquad \ldots\ldots (10,4)$$

we deduce that the rows of B are orthogonal to the rows of C. The residual sum of squares $\sum_i\sum_j\sum_k(y_{ijk} - y_{ijo} - y_{iok} - y_{ojk} + y_{ioo} + y_{ojo} + y_{ook} - \bar{y})^2$ is thus independent of $\sum_j\sum_k(y_{ojk} - y_{ojo} - y_{ook} + \bar{y})^2$, and by Lemma 7, Cor.1, of all other sums of squares in the analysis. Also, by Lemma 10 (Cor.), the residual sum of squares has $m(n-1)(p-1) - (n-1)(p-1) = (m-1)(n-1)(p-1)$ degrees of freedom and the residual mean square is on a null hypothesis an estimate of the intrinsic variance of the experiment.

In Section 7 the interaction of two factors A and B was defined as a measure of the variation in the response to A at the different levels of B, this measure being provided by the sum of squares $\sum_i\sum_j(y_{ijo} - y_{ioo} - y_{ojo} + \bar{y})^2$. Thus, a measure of the interaction AB at any given levels i and j is provided by the linear expression $y_{ijo} - y_{ioo} - y_{ojo} + \bar{y}$.

A second order interaction is defined as a measure of the variation in the first order interaction AB at the different levels of C. Since $y_{ijo}-y_{ioo}-y_{ojo}+\bar{y} = \sum_k (y_{ijk}-y_{iok}-y_{ojk}+y_{ook})$, such a measure is given by the sum of squares $\sum_i\sum_j\sum_k \left[y_{ijk}-y_{iok}-y_{ojk} + y_{ook} - (y_{ijo}-y_{ioo}-y_{ojo}+\bar{y})\right]^2$ or $\sum_i\sum_j\sum_k (y_{ijk}-y_{ijo}-y_{iok}-y_{ojk}+y_{ioo}+y_{ojo}+y_{ook}-\bar{y})^2$, which is the sum of squared residuals of (10,3). This sum of squares is symmetrical with respect to i, j, and k, so that the interaction ABC may equally be defined as a measure of the variation in the interaction AC at the different levels of B, or as a measure of the variation in the interaction BC at the different levels of A. The process may be continued by the addition of a fourth factor D, when the residual sum of squares for treatments will be the third order interaction ABCD, and so on.

Combining (10,1),(10,2), and (10,3) we have the algebraic relationship

$$\sum_i\sum_j\sum_k (y_{ijk}-\bar{y})^2 = np\sum_i (y_{ioo}-\bar{y})^2 + mp\sum_j (y_{ojo}-\bar{y})^2 + mn\sum_k (y_{ook}-\bar{y})^2 + p\sum_i\sum_j (y_{ijo}-y_{ioo} -y_{ojo}+\bar{y})^2 + n\sum_i\sum_k (y_{iok}-y_{ioo}-y_{ook}+\bar{y})^2 + m\sum_j\sum_k (y_{ojk}-y_{ojo}-y_{ook}+\bar{y})^2 + \sum_i\sum_j\sum_k (y_{ijk}-y_{ijo}-y_{iok} -y_{ojk}+y_{ioo}+y_{ojo}+y_{ook}-\bar{y})^2, \quad \ldots\ldots\ldots\ldots (10,5)$$

where each component sum of squares has been shown on a null hypothesis to have gamma-type distribution with the appropriate number of degrees of freedom from the table below, each mean square yielding an independent estimate of the error variance of the experiment. The treatment sum of squares has therefore been subdivided as follows:-

Analysis of Variance

Variation due to	D.F	Sums of Squares
Main effect of A	m-1	$np \sum_i (y_{ioo} - \bar{y})^2$
Main effect of B	n-1	$mp \sum_j (y_{ojo} - \bar{y})^2$
Main effect of C	p-1	$mn \sum_k (y_{ook} - \bar{y})^2$
Interaction AB	(m-1)(n-1)	$p \sum_i \sum_j (y_{ijo} - y_{ioo} - y_{ojo} + \bar{y})^2$
Interaction BC	(n-1)(p-1)	$m \sum_j \sum_k (y_{ojk} - y_{ojo} - y_{ook} + \bar{y})^2$
Interaction CA	(m-1)(p-1)	$n \sum_i \sum_k (y_{iok} - y_{ioo} - y_{ook} + \bar{y})^2$
Interaction ABC	(m-1)(n-1)(p-1)	$\sum_i \sum_j \sum_k (y_{ijk} - y_{ijo} - y_{iok} - y_{ojk} + y_{ioo} + y_{ojo} + y_{ook} - \bar{y})^2$
Total	mnp-1	$\sum_i \sum_j \sum_k (y_{ijk} - \bar{y})^2$

The effect of the 1st level of B, i.e. $y_{o1o} - \bar{y}$, may be written as $c'y$ where

$c' = \frac{1}{mnp}\Big[$ n-1 n-1 ... n-1 ¦ -1 -1 ...-1 ¦ -1 -1 ... -1 ¦ till the n^{th} sub-vector, then repeating 1st to $n^{th}\Big]$

The effect of the 1st level of C, i.e. $y_{oo1} - \bar{y}$, may similarly be written as $d'y$ where

$d' = \frac{1}{mnp}\Big[$ p-1 -1 -1 ...-1 ¦ p-1 -1 -1 ...-1 ¦ etc., mn subvectors in all$\Big]$.

Now the vector b of the residual $b'y$ above, which measures the interaction B C at the 1st levels of both B and C, is seen to have as elements (except for a common factor) the products of corresponding elements of c and d, which are the vectors determining the effects of the first levels of B and C.

This property is a general one and is specially exemplified in the case of factors at only two levels (Section 12). To see how it arises, let us consider the treatment means y_{ij} of an experiment with all combinations of two factors A and B at m and n levels respectively, the suffix i representing the i^{th} level of A and the suffix j the j^{th} level of B. Let $a'\{y_{i0}\}$ be any linear function of the mean yields for the levels of factor A. It need not necessarily belong to an orthogonal subdivision of the level-means of A, but if it is to be of any use statistically it must be independent both of the error mean square of the experiment and of the vector of unit elements which constitutes the correction for the mean. Since in the case of a residual vector the sum of its elements corresponding to any particular value of i or of j must be zero (or else it would not be independent of all the level-mean deviations), any vector whatsoever when applied to $\{y_{i0}\}$ will determine a linear function independent of the error mean square, but to satisfy the second requirement $\sum_i a_i$ must be zero. Written in terms of the vector $y = \{y_{11}\, y_{12} \cdots y_{1n} \mid y_{21}\, y_{22} \cdots y_{2n} \mid \cdots \mid y_{m1}\, y_{m2} \cdots y_{mn}\}$, $a'\{y_{i0}\}$ becomes

$$\frac{1}{n}\left[a_1\, a_1 \ldots a_1 \mid a_2\, a_2 \ldots a_2 \mid \ldots \mid a_m\, a_m \ldots a_m\right] y \quad \ldots. \quad (10,6)$$

At any particular level of B, i.e. for any particular value of j , say $j=1$, the value of the linear function is

$$\frac{1}{n}\left[na_1 \cdot \ldots \cdot \,\vdots\, na_2 \cdot \ldots \cdot \,\vdots\, \ldots \,\vdots\, na_m \cdot \ldots \cdot \right] y. \quad \ldots (10,7)$$

The deviation of this value from the mean value is

$$\frac{1}{n}\left[(n-1)a_1 \;\; -a_1 \ldots -a_1 \,\vdots\, (n-1)a_2 - a_2 \ldots -a_2 \,\vdots\, \ldots \,\vdots\, (n-1)a_m - a_m \ldots -a_m \right] y, \quad \ldots (10,8)$$

and the sum of the squares of such deviations is the sum of squares for the interaction between the given linear function and the factor B. The deviation of the mean of the 1st level of B from the general mean is given by

$$\frac{1}{mn}\left[n-1 \; -1 \ldots -1 \,\vdots\, n-1 \; -1 \ldots -1 \,\vdots\, \ldots \,\vdots\, n-1 \; -1 \ldots -1 \right] y \ldots (10,9)$$

It is evident from the fact that $\sum_i a_i = 0$ that the vectors of (10,6),(10,8), and (10,9) are orthogonal to one another, so that in general the sum of squares for the interaction of two effects is independent of the sums of squares for the effects themselves. It is also evident that the elements of the vector of (10,8) are, except for a common factor, the products of corresponding elements of the effects vectors of (10,6) and (10,9).

By a slight extension of definition we may now define the interaction of two linear functions associated with different classifications of the variates. Let us, for example, as above consider all combinations y_{ij} of two factors A and B at m and n levels respectively, and let $a' \{y_{i0}\}$ be any linear function of the level-means of factor A and $b' \{y_{0j}\}$ any linear function of the level-means of factor B, subject to $\sum_i a_i = \sum_j b_j = 0$. Expressed in terms of the vector $y = \{y_{ij}\}$, these linear functions are

$$\frac{1}{n}\left[a_1 a_1 \ldots a_1 \,\vdots\, a_2 a_2 \ldots a_2 \,\vdots \ldots \vdots\, a_m a_m \ldots a_m\right] y \quad \ldots(10,10)$$

and $$\frac{1}{m}\left[b_1 b_2 \ldots b_n \,\vdots\, b_1 b_2 \ldots b_n \,\vdots \ldots \vdots\, b_1 b_2 \ldots b_n\right] y, \quad \ldots(10,11)$$

and it is obvious that they are independent. Their interaction is defined in accordance with the previous paragraph as $k\sum_i\sum_j a_i b_j y_{ij}$, i.e. the transforming vector has as elements, apart from a common factor k, the products of corresponding elements of the effects vectors. The interaction is therefore

$$k\left[a_1 b_1 \;\; a_1 b_2 \ldots a_1 b_n \,\vdots\, a_2 b_1 \;\; a_2 b_2 \ldots a_2 b_n \,\vdots \ldots \vdots\, a_m b_1 \;\; a_m b_2 \ldots a_m b_n\right] y \quad \ldots(10,12)$$

and its vector is clearly orthogonal to those of (10,10) and (10,11). As for the value of k, it will depend on the actual definition of the linear function taken to measure the interaction, as will be seen in Section 12. For instance, in satisfying the condition that the effects and their interaction should be independent, the three appropriate linear functions, if normalised, also satisfy the conditions that they should belong to the same orthogonal set. Hence k could be taken as $\overline{\sqrt{\sum_i\sum_j (a_i b_j)^2}}^{\,+}$, and the vector of (10,12) would be normalised. Higher order interactions may be similarly defined, all vectors being further partitioned to correspond to the additional classification.

In the special case when all the elements of the effects vectors (not normalised) are ± 1 , as in designs with factors at two levels only, the operation of multiplying corresponding elements of A and B to get AB is reversible in that, if we apply it to A and AB, we shall get B. In other words the two vectors so treated need not be associated exclusively with different factors or classifications as in the general case. The main effects and interactions of a 2^n factorial design thus constitute a finite group with identity I, since $A^2 = B^2 = \ldots\ldots = I$ (Finney, 9)

Section 11.

Analysis of Variance and Least Squares.

Yates (24) pointed out that the process of analysis of variance, as applied to a set of orthogonal data such as those obtained from the regular experimental designs, is equivalent to fitting constants representing the effects of rows, columns, treatments, etc. (according to the particular design in question) by the method of least squares. An explicit proof of this, using matrix notation, is given below.

Let $y = [y_{ijk}]$, where the y's are independent normal variates with variance σ^2, be the matrix of yields for an n x n Latin square, which we may take as the most general of the elementary designs, and let us consider fitting to the data by least squares constants representing the mean, rows, columns, and treatments - namely, $x_{000}; x_{100}, x_{200}, \ldots, x_{n00}; x_{010}, x_{020}, \ldots, x_{0n0}; x_{001}, x_{002}, \ldots x_{00n}$, subject to the conditions $\sum_i x_{i00} = \sum_j x_{0j0} = \sum_k x_{00k} = 0$, which are necessary to ensure that row, column, and treatment totals show only the effects due to the particular row, column, or treatment concerned, and also that the general mean of the yields is an unbiassed estimate of the population mean.

The observational equations are $Ax = y$, where x is the vector $\{x_{000}\ x_{100}\ x_{200} \ldots x_{n00}\ x_{010}\ x_{020} \ldots x_{0n0}\ x_{001}\ x_{002} \ldots x_{00n}\}$,

y is the vector $\{y_{11k}\ y_{12k} \dots y_{1nk} \mid y_{21k}\ y_{22k} \dots y_{2nk} \mid \dots \mid y_{n1k}\ y_{n2k} \dots y_{nnk}\}$ (the suffix k representing the particular treatment allotted by randomisation to each plot), and A is of the form

$$\left[\begin{array}{c|c|c|c} J & J_1 & I & K_1 \\ \hline J & J_2 & I & K_2 \\ \hline \vdots & \vdots & \vdots & \vdots \\ \hline J & J_n & I & K_n \end{array}\right]$$

where J is the column vector $\{1\ 1 \dots 1\}$ with n elements, J_k is a matrix of order n x n with all elements in the k^{th} column unity and all other elements zero, I is the unit matrix of order n, and the K_j are matrices of order n x n with one element of each row and column unity, all other elements zero, and such that, if $k'_{i(\ell}$ is the ℓ^{th} row of K_i, $k'_{i(\ell}\ k_{j(\ell} = 0$, i.e. corresponding rows are orthogonal.

The normal equations are $A'Ax = A'y$, where

$$A'A = \left[\begin{array}{c|c|c|c} n^2 & n\ n \dots n & n\ n \dots n & n\ n \dots n \\ \hline \begin{matrix} n \\ n \\ \vdots \\ n \end{matrix} & nI & M & M \\ \hline \begin{matrix} n \\ n \\ \vdots \\ n \end{matrix} & M & nI & M \\ \hline \begin{matrix} n \\ n \\ \vdots \\ n \end{matrix} & M & M & nI \end{array}\right],$$

a matrix of order $(3n+1) \times (3n+1)$. The matrix M, being of order $n \times n$ with all elements unity, may, owing to the linear constraints $\sum_i x_{ioo} = \sum_j x_{ojo} = \sum_k x_{ook} = 0$, be replaced by a null matrix in every case. Similarly the vectors $[n \; n \; \ldots \; n]$ of the first row of A'A may be replaced by null vectors. The vector A'y is seen to be $\{G \; R_1 \; R_2 \ldots R_n \; C_1 \; C_2 \ldots C_n \; T_1 \; T_2 \ldots T_n\}$, where G is the grand total of yields, R_i = total of yields in the i^{th} row, C_j = total for the j^{th} column, and T_k = total for the k^{th} treatment.

The above set of 3n+1 equations are orthogonal in that the equation for x_{ooo} may be solved independently of the other constants which may in turn be found independently of one another. The solutions are $x_{ooo} = \bar{y}$, $x_{ioo} = y_{ioo} - \bar{y}$, $x_{ojo} = y_{ojo} - \bar{y}$, $x_{ook} = y_{ook} - \bar{y}$.

The residual sum of squares

$$= \sum_i \sum_j [(y_{ijk} - \bar{y}) - (y_{ioo} - \bar{y}) - (y_{ojo} - \bar{y}) - (y_{ook} - \bar{y})]^2$$

$$= \sum_i \sum_j (y_{ijk} - y_{ioo} - y_{ojo} - y_{ook} + 2\bar{y})^2,$$

as in the analysis of variance of the Latin square (P44). But the residual sum of squares also = $(y-Ax)'(y-Ax)$ $= y'y - x'A'y$. In ordinary notation this is $\sum_i \sum_j y_{ijk}^2 - x_{ooo} G - \sum_i x_{ioo} R_i - \sum_j x_{ojo} C_j - \sum_k x_{ook} T_k$. The term $x_{ooo} G$ is the "correction for mean" of analysis of variance, and the remaining terms give by reason of the orthogonality mentioned above the reduction from the sum of squared residuals due to fitting each constant. The sum of squares for rows, for example,

$$\begin{aligned} &= \sum x_{ioo} R_i \\ &= \sum_i \left[(y_{ioo} - \bar{y})(n y_{ioo}) \right] \\ &= n\left[\sum_i y_{ioo}^2 - \bar{y} \sum_i y_{ioo} \right] \\ &= n\left(\sum_i y_{ioo}^2 - 2\bar{y} \sum_i y_{ioo} + n\bar{y}^2 \right) \\ &= n \sum_i (y_{ioo} - \bar{y})^2. \end{aligned}$$

Similarly, that for columns $= n \sum_j (y_{ojo} - \bar{y})^2$, and that for treatments $= n \sum_k (y_{ook} - \bar{y})^2$

As for degrees of freedom, the number for the residual sum of squares is from the theory of least squares (Aitken, 3) equal to

(No. of observations) $-$ (no. of constants fitted) $+$ (no. of linear restraints)

$= n^2 - (3n+1) + 3 = (n-1)(n-2).$

The number of degrees of freedom for rows, columns, and treatments is n-1 in each case, as may be seen from the simple consideration that there is one linear restraint on each set of constants, or a proof based on the traces of idempotent matrices may be given, using Lemma 9, Corl.

In fitting the constants by least squares we have assumed that the $[y_{ijk}]$ are independent normal variates each with the same variance σ^2, but not necessarily with the same mean, and that each variate is made up of a set of independent components thus: $y_{ijk} = x_{ooo} + x_{ioo} + x_{ojo} + x_{ook} + x_{ijk}$, where the x_{ooo}, x_{ioo}, x_{ojo}, x_{ook}

represent, respectively, mean, row, column, and treatment effects, and x_{ijk} is a random normal variate with mean at the origin and variance σ^2. But these assumptions are exactly those of the analysis of variance of a Latin square (Section 5,P.48), and we have seen above how the ordinary process of analysis of variance corresponds exactly to the fitting of constants by least squares in respect of the isolation of the components of variance and degrees of freedom deriving from the various experimental controls. The two processes are thus identical, but the method of fitting constants may still be used when an experiment, either by accident or design, lacks orthogonality, so that the ordinary procedure of analysis of variance is unavailable or needs modification (Yates,24). It is also seen from the theory of least squares that the residual sum of squares is a minimum. This is the basis of many formulae for estimating the yields of missing plots, etc.

Section 12.

Factorial Experiments at Two Levels Only.

The statistical analysis of experiments with factors at only two levels lends itself admirably to algebraic treatment. The definitions and notations used are, except where indicated, those of Yates (26).

Matrix Representation of Main Effects and Interactions.

Let us first consider one factor only, say nitrogen, at two levels n and (1), where (1) represents the plots receiving control applications of nitrogen. Without ambiguity the yields corresponding to these treatments may be represented by the same symbols, and since in the calculation of the treatment sum of squares adjustment is made for the number of replications of the experiment, we need only consider the case of a single replication and no generality will be lost. We have the symbolic relationship

$$\begin{bmatrix} I \\ N \end{bmatrix} = \begin{bmatrix} n+1 \\ n-1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ n \end{bmatrix}, \quad \ldots. \; (12,1)$$

where $n+1$ symbolises the total of yields (represented on the left-hand side by I) and $n-1$ the superiority of nitrogen over control, i.e. the "main effect" of n (represented on the left-hand side by N). Similarly, for any other single factor, e.g. potash, we have the symbolic relationship

$$\begin{bmatrix} I \\ K \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ k \end{bmatrix} \quad \ldots\ldots \quad (12,2)$$

Proceeding now to two factors, n and k, we have the following treatment combinations :- (1), n, k, nk.
The main effects and interaction (omitting Yates' conventional factor) are given by $N = (n-1) + (nk-k)$, $K = (k-1) + (nk-n)$, $NK = (nk-k)-(n-1)$. We therefore have the following symbolic equations :-

$$I = (n+1)(k+1) \qquad \text{(total effect)}$$

$$N = (n-1)(k+1)$$

$$K = (n+1)(k-1)$$

$$NK = (n-1)(k-1)$$

or
$$\{I \quad N \quad K \quad NK\} = \begin{bmatrix} (n+1)(k+1) \\ (n-1)(k+1) \\ (n+1)(k-1) \\ (n-1)(k-1) \end{bmatrix} \quad \ldots (12,3)$$

But the vector on the right-hand side is the vector formed from $\begin{bmatrix} k+1 \\ k-1 \end{bmatrix}$ and $\begin{bmatrix} n+1 \\ n-1 \end{bmatrix}$ by making ordered binary products, and is therefore the "direct-product" of the two simple vectors, a process denoted by the symbol "x" thus:-
$$\begin{bmatrix} k+1 \\ k-1 \end{bmatrix} \times \begin{bmatrix} n+1 \\ n-1 \end{bmatrix} = \begin{bmatrix} (n+1)(k+1) \\ (n-1)(k+1) \\ (n+1)(k-1) \\ (n-1)(k-1) \end{bmatrix}$$

It therefore becomes necessary to consider the algebra of direct-product vectors and matrices.

Direct-Product Vectors and Matrices. Let us consider

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix}.$$

The vector of products $\{x_1 s_1 \;\; x_1 s_2 \;\; x_2 s_1 \;\; x_2 s_2\}$ is transformed into the corresponding vector of products $\{y_1 t_1 \;\; y_1 t_2 \;\; y_2 t_1 \;\; y_2 t_2\}$ according to the relationship

$$\begin{bmatrix} y_1 t_1 \\ y_1 t_2 \\ y_2 t_1 \\ y_2 t_2 \end{bmatrix} = \begin{bmatrix} a_{11} b_{11} & a_{11} b_{12} & a_{12} b_{11} & a_{12} b_{12} \\ a_{11} b_{21} & a_{11} b_{22} & a_{12} b_{21} & a_{12} b_{22} \\ a_{21} b_{11} & a_{21} b_{12} & a_{22} b_{11} & a_{22} b_{12} \\ a_{21} b_{21} & a_{21} b_{22} & a_{22} b_{21} & a_{22} b_{22} \end{bmatrix} \begin{bmatrix} x_1 s_1 \\ x_1 s_2 \\ x_2 s_1 \\ x_2 s_2 \end{bmatrix}$$

The vector on the left-hand side is defined as the direct-product vector of y and t and may be written $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \times \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$, or more simply $(y \times t)$. Similarly, the vector on the right-hand side is the direct-product vector of x and s, or $(x \times s)$. If $y = Ax$ and $t = Bs$, then the direct-product matrix $(A \times B)$ is defined by the relationship $(y \times t) = (A \times B)(x \times s)$. In general, if $y = \{y_1 \;\; y_2 \; \ldots \; y_m\} = Ax$, where $x = \{x_1 \;\; x_2 \; \ldots \; x_n\}$ and A is in general rectangular and of order $(m \times n)$, and if $t = \{t_1 \;\; t_2 \; \ldots \; t_p\} = Bs$, where $s = \{s_1 \;\; s_2 \; \ldots s_q\}$ and B is in general rectangular and of order $p \times q$, then the direct-product vectors of x and s and of y and t are

$$(x \times s) = \{x_1 s_1 \;\; x_1 s_2 \,..\, x_1 s_q \;\; x_2 s_1 \;\; x_2 s_2 \,...\, x_2 s_q \,....\quad x_n s_q\}$$
$$(y \times t) = \{y_1 t_1 \;\; y_1 t_2 \,..\, y_1 t_p \;\; y_2 t_1 \;\; y_2 t_2 \,...\, y_2 t_p \,....\quad y_m t_p\},$$

and the direct-product matrix $(A \times B)$ of A and B is defined by $(y \times t) = (Ax \times Bs) = (A \times B)(x \times s)$ .... (12,4)

$(A \times B)$ is of order $mp \times nq$, since $(y \times t)$ is of order $mp \times 1$ and $(x \times s)$ is of order $nq \times 1$. If $A = [a_{ij}]$, then $(A \times B)$ can be formed as follows:-

$$(A \times B) = \begin{bmatrix} a_{11} B & a_{12} B & \dots & a_{1n} B \\ a_{21} B & a_{22} B & \dots & a_{2n} B \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} B & a_{m2} B & \dots & a_{mn} B \end{bmatrix}$$

Hence $(I \times A) = \begin{bmatrix} A & & \\ & A & \\ & & \ddots A \end{bmatrix}$, where I is of order $n \times n$

and A recurs n times down the diagonal; and $(I \times I)$, where the I's are of order $n \times n$ and $m \times m$ respectively, is I of order $mn \times mn$. Also $(A \times 1) = A$, for, putting $B = 1$ in (12,4), we have

$$(A \times 1)(x \times s) = (Ax \times s) = A(x \times s) \text{ , s being scalar.}$$

Multiplication Theorem The fundamental multiplicative law for general direct-product matrices is $(A \times B)(C \times D) = (AC \times BD)$. For by definition $(A \times B)(x \times s) = (Ax \times Bs)$, where x and s are arbitrary conformable vectors. Hence

$$\begin{aligned} (A \times B)(C \times D)(x \times s) &= (A \times B)(Cx \times Ds) && \text{(by definition)} \\ &= (ACx \times BDs) && \text{(again by definition)} \\ &= (AC \times BD)(x \times s) && \text{(also by definition)} . \end{aligned}$$

$\therefore (A \times B)(C \times D) = (AC \times BD)$, since x and s are arbitrary vectors.

Transposition of a Direct-Product Matrix. What is $(A \times B)'$?

$$(A \times B) = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix}.$$

$$(A \times B)' = \begin{bmatrix} a_{11}B' & a_{21}B' & \dots & a_{m1}B' \\ a_{12}B' & a_{22}B' & \dots & a_{m2}B' \\ \vdots & \vdots & & \vdots \\ a_{1n}B' & a_{2n}B' & \dots & a_{mn}B' \end{bmatrix} = (A' \times B').$$

First Theorem on Orthogonality. If M and N are both orthogonal matrices, then the direct-product matrix $(M \times N)$ is orthogonal. For

$$\begin{aligned} (M \times N)'(M \times N) &= (M' \times N')(M \times N) \\ &= (M'M \times N'N) \quad \text{(by the multiplication theorem)} \\ &= (I \times I) = I. \end{aligned}$$

"Direct-Square" of a Matrix. If we have a direct-product matrix$(A \times B)$ and put $B = A$, we obtain $(A \times A)$, the"direct-square" of A, which we may write as $A^{\{2\}}$.

For example, if $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, (though in general A is rectangular),

$$A^{\{2\}} = \begin{bmatrix} a_{11}^2 & a_{11}a_{12} & a_{12}a_{11} & a_{12}^2 \\ a_{11}a_{21} & a_{11}a_{22} & a_{12}a_{21} & a_{12}a_{22} \\ a_{21}a_{11} & a_{21}a_{12} & a_{22}a_{11} & a_{22}a_{12} \\ a_{21}^2 & a_{21}a_{22} & a_{22}a_{21} & a_{22}^2 \end{bmatrix},$$

a matrix of order 4×4. The process may be repeated to obtain the "direct-cube" of A, $A^{\{3\}}$, a matrix of order $2^3 \times 2^3$, and so on.

Second Theorem on Orthogonality. If M is an orthogonal matrix, then $M^{\{s\}}$ is also orthogonal. This follows immediately from the previous theorem on orthogonality, for if we put $N = M$, we have that $M^{\{2\}}$ is orthogonal, and the proof follows by induction.

Interaction Transformations. We have, by (12,1), symbolically $\begin{bmatrix} I \\ N \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ n \end{bmatrix}$. Suppose now that we normalise each of the row vectors of the matrix on the right-hand side and write

$$\begin{bmatrix} I \\ N \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ n \end{bmatrix} = M \begin{bmatrix} 1 \\ n \end{bmatrix}.$$

The matrix M is orthogonal so that the linear expressions for I and N constitute a complete orthogonal set. Similarly, we may write

$$\begin{bmatrix} I \\ K \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ k \end{bmatrix} = M \begin{bmatrix} 1 \\ k \end{bmatrix}$$

for any other single factor, potash.

Now, for a 2×2 experiment with n and k we have by (12,3), ignoring normalising factors,

$$\{I \quad N \quad K \quad NK\} = \begin{bmatrix} (n+1)(k+1) \\ (n-1)(k+1) \\ (n+1)(k-1) \\ (n-1)(k-1) \end{bmatrix} = \begin{bmatrix} k+1 \\ k-1 \end{bmatrix} \times \begin{bmatrix} n+1 \\ n-1 \end{bmatrix} = \{I \quad K\} \times \{I \quad N\}.$$

It is thus apparent that the following symbolic operations hold good :- $I^{\{2\}} = I$, $(I \times N) = N$, $(I \times K) = K$, $(N \times K) = NK$, where the symbols (representing vectors) on the left-hand side of each equation refer to single factors, those on the right-hand side to the 2-factor experiment.

Introducing normalising factors, we have

$$
\begin{aligned}
\{I \quad N \quad K \quad NK\} &= \{I \quad K\} \times \{I \quad N\} \\
&= M\{1 \quad k\} \times M\{1 \quad n\} \\
&= M^{\{2\}}\left(\{1 \quad k\} \times \{1 \quad n\}\right) \quad \text{(by definition of multiplication)} \\
&= M^{\{2\}}\{1 \quad n \quad k \quad nk\}
\end{aligned}
$$

The "interaction matrix" transforming yields into main effects and interactions is therefore $M^{\{2\}}$ for a 2×2 experiment. By the second theorem on orthogonality above, we know that M is orthogonal. It is, in fact,

$$
\frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}.
$$

Thus the main effects and interactions constitute a complete orthogonal set and may be tested for significance in the manner explained in Section 9.

If a third factor, say d, is now introduced, main effects and interactions are defined by the products of the preceding $(n+1)(k+1)$, $(n-1)(k+1)$, $(n+1)(k-1)$, $(n-1)(k-1)$ with the factors $(d+1)$ and $(d-1)$ in turn, subject to normalisation as before. But this is forming a direct-product vector with the additional vector $M\ \{1\ \ d\}$, so that we have, symbolically,

$$
\begin{aligned}
&\{I\ \ N\ \ K\ \ NK\ \ D\ \ ND\ \ KD\ \ NKD\} \\
&= \{I\ \ D\} \times \{I\ \ N\ \ K\ \ NK\} \\
&= M\{1\ \ d\} \times M^{\{2\}}\{1\ \ n\ \ k\ \ nk\} \\
&= M^{\{3\}}\left(\{1\ \ d\} \times \{1\ \ n\ \ k\ \ nk\}\right) \\
&= M^{\{3\}}\{1\ \ n\ \ k\ \ nk\ \ d\ \ nd\ \ kd\ \ nkd\}
\end{aligned}
$$

The interaction matrix for three factors at two levels each is therefore $M^{\{3\}}$, which is orthogonal, so that once again the degrees of freedom for treatments have been orthogonally subdivided into single degrees of freedom. Written in full the transformation of yields into main effects and interactions for a $2 \times 2 \times 2$ experiment is

$$
\begin{bmatrix} I \\ N \\ K \\ NK \\ D \\ ND \\ KD \\ NKD \end{bmatrix}
= \frac{1}{\sqrt{8}}
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
-1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\
-1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 \\
1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\
-1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \\
1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\
1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
-1 & 1 & 1 & -1 & 1 & -1 & -1 & 1
\end{bmatrix}
\begin{bmatrix} 1 \\ n \\ k \\ nk \\ d \\ nd \\ kd \\ nkd \end{bmatrix}
$$

By the introduction of further factors it may be proved that in an s- factor experiment with each factor at two levels, the interaction matrix is $M^{\{s\}}$, which is orthogonal, showing that there is one degree of freedom for each main effect and interaction, each of which is independent of the others and of all other sums of squares in the analysis of variance. This property and its corollary, that each main effect and interaction is expressible as a linear function of the yields, the square of which appears in the analysis of variance, is confined to experiments with two factors only.

As shown in Section 10, the following symbolic equations are true (subject to normalisation), using the definition that, if $A = l'y$ and $B = m'y$, then $AB = \sum_i l_i m_i y_i$:- $I^2 = I$, $IN = N$, $INK = NK$, $N^2 = K^2 = D^2 = I$, $D.ND = ND^2 = NI = N$, $KD.NKD = D$, etc. Hence I is known as the identity of the "effects group".

Since the linear expressions corresponding to each main effect and interaction are all independent of one another, it follows, by Lemma 6, that the mean square for each main effect or interaction is independent of the mean square obtained by pooling the degrees of freedom and squares corresponding to any of the other members of the orthogonal set. By the additive law of gamma-type variates such a pooled sum of squares has on the null hypothesis gamma-type distribution with degrees of freedom equal to the sum of the individual degrees of freedom, and its mean square is an unbiassed estimate of σ^2, the random variance. This is the justification, in experiments with a single or even fractional replication, for combining the degrees of freedom and squares corresponding to high order interactions (normally expected to be null) into an estimate of experimental error against which other mean squares may be tested. The particular degrees of freedom to be used as the estimate of error must be chosen beforehand from the interactions of the second of higher order which, on the basis of past experience or by reason merely of their high order, are predicted to be negligible. Naturally, should any of the chosen interactions prove to be appreciable, it is not permissible to remove them from the estimate of error. The result must be deemed a chance one, though it may be noted for reference in respect to future work. Also, should a main effect or interaction not among those chosen as estimate of error prove to be not significant, it is not a valid procedure to combine such a degree (or degrees) of freedom into a new pooled estimate of error, for such an estimate would be biassed. The following proof is this is adapted from a proof due to M.H. Quenouille.

Suppose that we have two variance estimates, u and v, of σ^2 with n_1 and n_2 degrees of freedom respectively. Let αv be the level of significance such that, if $u < \alpha v$, it is proposed to form a new variance estimate $\frac{n_1 u + n_2 v}{n_1 + n_2}$. For simplicity we take the special case of $n_1 = 2$. Then the mean value of u for $u < \alpha v$

$$= \frac{\int_0^\infty \int_0^{\alpha v} v^{\frac{1}{2}n_2 - 1} u \exp -\left(\frac{n_2 v}{2\sigma^2} + \frac{u}{\sigma^2}\right) du\,dv}{\int_0^\infty \int_0^{\alpha v} v^{\frac{1}{2}n_2 - 1} \exp -\left(\frac{n_2 v}{2\sigma^2} + \frac{u}{\sigma^2}\right) du\,dv}$$

$$= \frac{\int_0^\infty v^{\frac{1}{2}n_2 - 1} \exp\left(-\frac{n_2 v}{2\sigma^2}\right)\left\{\sigma^4 - \sigma^4 \exp(-\alpha v) - \sigma^2 \alpha v \exp(-\alpha v)\right\} dv}{\int_0^\infty v^{\frac{1}{2}n_2 - 1} \exp\left(-\frac{n_2 v}{2\sigma^2}\right)\left\{\sigma^2 - \sigma^2 \exp(-\alpha v)\right\} dv}$$

$$= \frac{\sigma^4 \Gamma(\frac{n_2}{2}) . \left(\frac{2\sigma^2}{n_2}\right)^{\frac{n_2}{2}} - \sigma^4 \Gamma(\frac{n_2}{2}) . \left(\frac{2\sigma^2}{n_2 + 2\alpha\sigma^2}\right)^{\frac{n_2}{2}} - \alpha\sigma^2 \Gamma(\frac{n_2}{2}+1) . \left(\frac{2\sigma^2}{n_2 + 2\alpha\sigma^2}\right)^{\frac{n_2}{2}+1}}{\sigma^2 \Gamma(\frac{n_2}{2}) . \left(\frac{2\sigma^2}{n_2}\right)^{\frac{n_2}{2}} - \sigma^2 \Gamma(\frac{n_2}{2}) . \left(\frac{2\sigma^2}{n_2 + 2\alpha\sigma^2}\right)^{\frac{n_2}{2}}}$$

$$= \sigma^2 \left[1 - \frac{\alpha n_2}{2\sigma^2} \left(\frac{2\sigma^2}{n_2 + 2\alpha\sigma^2}\right) \left\{ \left(\frac{n_2 + 2\alpha\sigma^2}{n_2}\right)^{\frac{n_2}{2}} - 1 \right\}^{-1} \right]$$

$$= \sigma^2 \left[1 - \left(\frac{n_2 + 2\alpha\sigma^2}{\alpha n_2}\right)^{-1} \left\{ \left(\frac{n_2 + 2\alpha\sigma^2}{n_2}\right)^{\frac{n_2}{2}} - 1 \right\}^{-1} \right]$$

Hence the mean value of $\frac{n_2 v + 2u}{n_2 + 2}$ for $u < \alpha v$ will be too low by an amount $\frac{2\sigma^2}{n_2+2}\left(\frac{1}{\alpha} + \frac{2\sigma^2}{n_2}\right)^{-1}\left[\left(1 + \frac{2\alpha\sigma^2}{n_2}\right)^{\frac{n_2}{2}} - 1\right]^{-1}$ which $\rightarrow 0$ as $n_2 \rightarrow \infty$.

In 1935 Yates (25) first published the theory of factorial experiments, though they had actually been in use at Rothamsted Experimental Station for several years. Then, and subsequently, many criticisms on various grounds were levelled at such experiments, most of them being effectively countered by Yates (25,26). One such criticism, by Wishart (23), may be mentioned here.

Wishart contended that, owing to the comparative unreliability of estimates of variance based on only one degree of freedom, it was dangerous to accept mean squares which were significent at the 5% level as being evidence of some real effect when they were very likely only chance effects. Thus, the treatment sum of squares should be examined for significance first (just as in an ordinary randomised-block experiment), and only if significance were established would it be legitimate to test individual degrees of freedom, such tests being, after all, only individual t-tests of two particular treatments. These contentions are easily refuted. If the 5% level of significance is used, only one in twenty of the effects and interactions should be significant by pure chance, whether based on one or several degrees of freedom. Realising that by adopting the 5% level he may be in error once in every twenty times, the statistician does not assert that any effects which reach this significance-level are necessarily genuine. He waits for confirmatory evidence . Then, factorial design is such that it is unnecessary to prove the significance of the treatments sum of squares as a whole before examination of individual degrees of freedom. For example, the main effect of any factor at two levels (say n) is measured by comparing half the plots of the experiment with the other half. Thus, so far as the main effect of n is concerned, the design is equivalent to a simple experiment with two treatments only - the two levels of n - , and in such a case it is obviously legitimate to use the t-test before the F-test, since the two tests are exactly equivalent.

In this section, maintaining consistency with Section 9, the linear responses of main effects or interactions have been defined by certain normalised vectors. Such a definition is not in accord with that of Yates (26), whose definition of the effect of a linear combination of the yields, $\ell' y$, is given by $\frac{\Sigma \ell y}{\lambda \Sigma \ell^2}$, where by convention $\lambda = \frac{1}{2}$ for factors at two levels only and $\lambda = 1$ in all other cases.

The definition of a linear response as a normalised vector of yields has many theoretical advantages. Such vectors may be assembled into a complete or incomplete orthogonal set, and the significance or non-significance of the responses is immediately apparent by comparison with $t \times$(S.E. of the experiment), where t has the value corresponding to the number of degrees of freedom of the error mean square and to the level of significance required. Also their squares may be entered directly in the analysis of variance table without any division. (In the case of more than one replication, the square root of the number of replications is incorporated in the normalising factor $\frac{1}{\sqrt{\Sigma \ell^2}}$ - in the denominator if the vector y is of treatment totals, in the numerator if the vector is that of treatment means). The normalised definition is also completely practical in working with a single experiment or with a set of experiments all of the same design. In point of fact, in these circumstances no normalising factor is required at all, provided adjustment to the squares is made in the analysis of variance by the proper divisor (the square of the normalising factor) and provided the standard error for the t-test is also suitably adjusted. Thus, the original definitions given by Yates (25) for the main effects and interactions of factorial experiments at two levels only were sums and differences of yields of treatment-combinations without the conventional factors he introduced before the paper was actually published.

However, if linear responses are to be made comparable for experiments with different designs, they must be reduced to a per plot basis, otherwise, for example, the interaction between two factors n and k of a three-factor experiment will, other things being equal, be less than the interaction of these same factors in a five-factor experiment. With the normalised vector definition it would be approximately half. It was to correct this anomaly that Yates altered his original definitions. For similar reasons, therefore, we now introduce a practical measure of the response on a per plot basis of a linear combination of yields, $\ell'y$, - namely $\frac{\Sigma \ell y}{\Sigma \ell^2}$. This definition is the same as that of Yates except for his conventional factor λ , i.e. the two definitions are the same except for factors at only two levels. There is, however, the additional condition which must now be imposed on the elements of the vector ℓ, that they must be integral or zero and have no factor in common.

The new definition presents other advantages besides facilitating comparisons between different experiments. For example, we shall see in the next section how, if we fit a multiple polynomial

$$f(x,y,z,\ldots) \equiv a_{000\ldots} + a_{100\ldots}x + a_{010\ldots}y + a_{001\ldots}z + \ldots$$
$$+ a_{110\ldots}xy + a_{101\ldots}xz + a_{011\ldots}yz + \ldots + a_{111\ldots}xyz + \ldots$$

by least squares to the yields of an experiment with factors x,y,z,... at two levels only, the coefficients are the linear responses, as now defined, of the main effects and interactions, e.g. $a_{011\ldots} \equiv$ the interaction YZ.

Of course, the items in the "sums of squares" column in the analysis of variance are no longer the squares of the linear responses, but the adjustments are easily made, as are those required to the standard error for application of the t-test. It is clear that the results obtained with the normalised definition still hold, since the factor $\frac{1}{\Sigma \ell^2}$ is detachable as required. Where it is theoretically preferable the normalised definition will still be used, but, if so, the fact will be specifically mentioned.

Yates' adoption of the conventional factor $\lambda = \frac{1}{2}$ for factors at two levels is more difficult to justify. It is true that for main effects his definition gives the mean response, but it does not give the mean interaction, which would require the factor $\lambda = \frac{1}{2^{r+1}}$, where r is the order of the interaction concerned. The complications brought about by varying λ within a single experiment probably caused Yates to define λ as $\frac{1}{2}$ for all main effects and interactions, but why introduce the factor λ at all? There seems to be no particular advantage (other than the very slight one already mentioned) in having any factor additional to $\frac{1}{\Sigma \ell^2}$, the necessity for which in certain circumstances has been shown. The effect is merely to introduce additional complication. For example, the formulae for deriving the yields of the various treatment-combinations from the main effects and interactions lose simplicity. For a 2 x 2 x 2 experiment in n,k, and d

$$nk = \text{Mean} + \tfrac{1}{2}(-N+K-NK+D-ND+KD-NKD) \quad \text{(Yates' definition)}$$

$$= \tfrac{1}{\sqrt{8}}(I-N+K-NK+D-ND+KD-NKD) \quad \text{(normalised definition)}$$

$$= \text{Mean} -N+K-NK+D-ND+KD-NKD \quad \text{(response per plot definition)}$$

The signs in each case come from the transpose of the matrix $M^{\{3\}}$ of P.122. The improved simplicity of the last formula is apparent.

Section 13.

Orthogonal Polynomials and Factorial Experiments

It has been seen (§11) that analysis of variance is equivalent to fitting by least squares to the yield data certain constants, including one for every treatment. If the treatments consist of a single factor at equally-spaced intervals (levels), we may fit a curve of regression to the treatment-constants by least squares and orthogonal polynomials. In agricultural experiments, where the factors are conveniently arranged at equally-spaced levels, we are chiefly interested in the linear and quadratic effects of the factor, and these may be determined by fitting a quadratic polynomial of type $a_0 + a_1 p_1(x) + a_2 p_2(x) = u$, where the functions $p(x)$ are orthogonal. Suppose that we have a factor at five equally-spaced levels and assign metric values -2,-1, 0,1,2 to the variate x , thus:-

x	-2	-1	0	1	2
u	u_{-2}	u_{-1}	u_0	u_1	u_2,

where u_{-2}, u_{-1}, etc. are the treatment constants fitted by least squares. Since these constants are deviations of treatment means from the general mean, the u's may be taken as treatment means or treatment totals (the required adjustments in either case being easily made), thus allowing the change of origin from the general mean to zero to be absorbed in the constant term, a_0. As previously, we shall in this section take the u's as being the treatment yields of a single replication.

With the metric values adopted it is apparent that we may take $p_1(x)$ as x and $p_2(x)$ as $x^2 - \alpha$, for these functions can be made orthogonal be assigning a suitable value for α. The observational equations are Pa=u, where the rows of P are $\left[1 \;\; x_i \;\; x_i^2 - \alpha\right]$. In our example $\alpha = 2$, and the equations are

$$\begin{bmatrix} 1 & -2 & 2 \\ 1 & -1 & -1 \\ 1 & 0 & -2 \\ 1 & 1 & -1 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = u$$

Because of the orthogonal relations, P'P is a diagonal matrix. In general, $P'P = \mathrm{diag}\left[n \;\; \Sigma p_1^2(x) \;\; \Sigma p_2^2(x)\right]$, where n = the number of levels of x. The normal equations are $P'Pa = P'u$, where $P'u = \left\{\Sigma u \;\; \Sigma u p_1(x) \;\; \Sigma u p_2(x)\right\}$, so that the a_j are determined independently as $a_j = \dfrac{\Sigma u p_j(x)}{\Sigma p_j^2(x)}$. Thus in the present example

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{5}(1 & 1 & 1 & 1 & 1) \\ \frac{1}{10}(-2 & -1 & 0 & 1 & 2) \\ \frac{1}{14}(2 & -1 & -2 & -1 & 2) \end{bmatrix} u$$

At the same time the residual sum of squares, since the u's are independent and have the same intrinsic variance, is equal to $(u-Pa)'(u-Pa) = u'u - a'P'Pa$. In ordinary notation this is equal to $\sum u^2 - \sum_j a_j^2 \sum_x p_j^2(x)$, and because of the orthogonality the reductions from the treatment sum of squares due to the linear and quadratic effects are $a_1^2 \sum p_1^2(x)$ and $a_2^2 \sum p_2^2(x)$. The reduction due to the fitting of the constant term is $\frac{(\sum u)^2}{n}$, which is the correction for the general mean.

It is evident that P'u, with the vectors normalised, is an incomplete orthogonal set subdividing the degrees of freedom for treatments. Each individual degree of freedom may therefore be tested either by the F-test (comparing its mean square, $a_j^2 \sum_x p_j^2(x)$, with the error mean square), or by applying the t-test directly to the linear function. The significance of the linear and quadratic effects is thus rigidly tested. It will also be noticed that a_1 and a_2 are the linear responses per plot (as defined in the last section) of the linear and quadratic effects respectively.

The orthogonal polynomial values as obtained above are particular cases of the Tchebycheff polynomials. If it is desired to ascertain the cubic, quartic, quintic, etc. effects of x, the appropriate orthogonal polynomial values may be found in Fisher and Yates' tables (14). The new terms may be fitted without alterations to any previous terms and their contributions to variance are purely additive.

If the design consists of two factors at equally-spaced levels of each, we may fit a bi-variate polynomial by least squares and orthogonal polynomials. Allotting metric values to x and y as before so that $\sum x = \sum y = 0$, it follows that the functions x and y are orthogonal to one another since $\sum xy = \sum x \sum y = 0$. Also we see at once that the function xy is orthogonal to both x and y. Hence for the fitting of the bivariate polynomial $f(x,y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy$, a_{11} must be by the results of §10 (P108) the interaction of the linear effects of the two factors. In particular, if x and y are at two levels only, a_{11} is the interaction XY. As an example, in a 5 x 4 experiment we have the scheme

y ↓ \ x →	-2	-1	0	1	2
$-\frac{3}{2}$	u_{11}	u_{21}	u_{31}	u_{41}	u_{51}
$-\frac{1}{2}$	u_{12}	u_{22}	u_{32}	u_{42}	u_{52}
$\frac{1}{2}$	u_{13}	u_{23}	u_{33}	u_{43}	u_{53}
$\frac{3}{2}$	u_{14}	u_{24}	u_{34}	u_{44}	u_{54}

so that

$$a_{00} = \tfrac{1}{20}\left[1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \quad 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \quad 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \quad 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1\right] u$$

$$a_{10} = \tfrac{1}{40}\left[-2 \; -1 \;\; 0 \;\; 1 \;\; 2 \quad -2 \; -1 \;\; 0 \;\; 1 \;\; 2 \quad -2 \; -1 \;\; 0 \;\; 1 \;\; 2 \quad -2 \; -1 \;\; 0 \;\; 1 \;\; 2\right] u$$

$$a_{01} = \tfrac{1}{100}\left[-3 \; -3 \; -3 \; -3 \; -3 \quad -1 \; -1 \; -1 \; -1 \; -1 \quad 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \quad 3 \;\; 3 \;\; 3 \;\; 3 \;\; 3\right] u$$

$$a_{11} = \tfrac{1}{200}\left[6 \;\; 3 \;\; 0 \; -3 \; -6 \quad 2 \;\; 1 \;\; 0 \; -1 \; -2 \quad -2 \; -1 \;\; 0 \;\; 1 \;\; 2 \quad -6 \; -3 \;\; 0 \;\; 3 \;\; 6\right] u$$

Where $u = \{ u_{11}\ u_{21}\ u_{31}\ u_{41}\ u_{51}\ u_{12}\ u_{22} \ldots u_{52} \ldots u_{14}\ u_{24}\ u_{34}\ u_{44}\ u_{54} \}$

It has already been seen how α can be fixed so as to make the function $x^2-\alpha$ orthogonal with x and 1, and therefore with all members of the orthogonal set $P'u$. Similarly, $y^2-\beta$ may be made orthogonal by fixing β. If this is done, it follows that the functions $(x^2-\alpha)y$, $x(y^2-\beta)$, $(x^2-\alpha)(y^2-\beta)$ are all orthogonal with the functions 1, x, y, xy, $x^2-\alpha$, $y^2-\beta$, and to each other. Hence, if we fit by least squares the bivariate polynomial

$$f(x,y) = a_{00} + a_{10}x + a_{20}(x^2-\alpha) + a_{01}y + a_{11}xy + a_{21}(x^2-\alpha)y + a_{02}(y^2-\beta) + a_{12}x(y^2-\beta) + a_{22}(x^2-\alpha)(y^2-\beta),$$

a_{21}, a_{12}, a_{22} represent respectively the interactions between quadratic effect of x and linear effect of y, between linear effect of x and quadratic effect of y, and between the two quadratic effects. If the factors are A and B, these interactions may be denoted by A"B', A'B", and A"B", the single and double dashes representing linear and quadratic effects respectively. In a 3 x 3 design the above subdivision of treatment yields would give a complete orthogonal set, but with either factor at more that three levels the subdivision is incomplete and there will be a residual treatment sum of squares. However, as in the case of a single factor, additional terms may be added representing cubic etc. effects and their interactions, the required values of the orthogonal polynomials being available in Fisher and Yates (14). As before the significance of each a_{ij} may be tested against the error mean square.

The fitting of a multivariate polynomial $f(x,y,z,\dots)$ to the treatment yields of a design with more than two factors is a simple extension of the above process.

The matrix P' of the equations $a = P'\mu$, with its vectors normalised so that $P'P = I$, is what we have called in discussing factors at two levels (§12, P. 121) an interaction matrix. Such matrices were constructed by means of forming direct-products, and an extension of this method may be used to construct the interaction matrix for factors at more than two levels. For example, consider our 5 x 4 design. Taking the first factor A alone and using the normalised definition, the interaction transformation is:-

$$\begin{bmatrix} I \\ A' \\ A'' \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{10}} & -\frac{1}{\sqrt{10}} & 0 & \frac{1}{\sqrt{10}} & \frac{2}{\sqrt{10}} \\ \frac{2}{\sqrt{14}} & -\frac{1}{\sqrt{14}} & -\frac{2}{\sqrt{14}} & -\frac{1}{\sqrt{14}} & \frac{2}{\sqrt{14}} \end{bmatrix} \begin{bmatrix} u_{10} \\ u_{20} \\ u_{30} \\ u_{40} \\ u_{50} \end{bmatrix}$$

where the interaction matrix (M, say) is orthogonal, but, of course, with only a single factor there are no interactions, only linear and quadratic effects. For the second factor B alone the transformation is

$$\begin{bmatrix} I \\ B' \\ B'' \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ -\frac{3}{\sqrt{20}} & -\frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{3}{\sqrt{20}} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} u_{01} \\ u_{02} \\ u_{03} \\ u_{04} \end{bmatrix}$$

with orthogonal interaction matrix N (say). These two transformations may be combined by direct-multiplication with the convention that $u_{i0}\,u_{0j} = u_{ij}$. We have

$$\begin{bmatrix} I \\ A' \\ A'' \\ B' \\ A'B' \\ A''B' \\ B'' \\ A'B'' \\ A''B'' \end{bmatrix} = \begin{bmatrix}
\frac{1}{2\sqrt{5}}(& 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1) \\
\frac{1}{2\sqrt{10}}(& -2 & -1 & 0 & 1 & 2 & -2 & -1 & 0 & 1 & 2 & -2 & -1 & 0 & 1 & 2 & -2 & -1 & 0 & 1 & 2) \\
\frac{1}{2\sqrt{14}}(& 2 & -1 & -2 & -1 & 2 & 2 & -1 & -2 & -1 & 1 & 2 & -1 & -2 & -1 & 2 & 2 & -1 & -2 & -1 & 2) \\
\frac{1}{10}(& -3 & -3 & -3 & -3 & -3 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 & 3) \\
\frac{1}{10\sqrt{2}}(& 6 & 3 & 0 & -3 & -6 & 2 & 1 & 0 & -1 & -2 & -2 & -1 & 0 & 1 & 2 & -6 & -3 & 0 & 3 & 6) \\
\frac{1}{2\sqrt{70}}(& -6 & 3 & 6 & 3 & -6 & -2 & 1 & 2 & 1 & -2 & 2 & -1 & -2 & -1 & 2 & 6 & -3 & -6 & -3 & 6) \\
\frac{1}{2\sqrt{5}}(& 1 & 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & 1 & 1) \\
\frac{1}{10\sqrt{2}}(& -2 & -1 & 0 & 1 & 2 & 2 & 1 & 0 & -1 & -2 & 2 & 1 & 0 & -1 & -2 & -2 & -1 & 0 & 1 & 2) \\
\frac{1}{2\sqrt{14}}(& 2 & -1 & -2 & -1 & 2 & -2 & 1 & 2 & 1 & -2 & -2 & 1 & 2 & 1 & -2 & 2 & -1 & -2 & -1 & 2)
\end{bmatrix} \begin{bmatrix} u_{11} \\ u_{21} \\ u_{31} \\ u_{41} \\ u_{51} \\ u_{12} \\ u_{22} \\ u_{32} \\ u_{42} \\ u_{52} \\ u_{13} \\ u_{23} \\ u_{33} \\ u_{43} \\ u_{53} \\ u_{14} \\ u_{24} \\ u_{34} \\ u_{44} \\ u_{54} \end{bmatrix}$$

By the first theorem on the orthogonality of direct-product matrices (§12, P.119), we know that $(N \times M)$ is orthogonal, so that the above matrix is orthogonal and each vector is normalised. It is apparent that the vector of an interaction (say $A''B'$) is (apart from normalising factors) the result of multiplying corresponding elements of the effects vectors (A'' and B'), thus agreeing with previous results.

<u>Tables of Orthogonal Polynomials</u> Tables of orthogonal polynomial values for some simple factorial designs will now be appended. Only linear and quadratic effects will be tabulated, since cubic or higher effects are rarely required. In the notation used a_{121}, for example, would represent the coefficient of xy^2z in our fitted polynomial $f(x,y,z)$, i.e. it would be the interaction $A'B''C'$ on a response per plot basis.

3 x 2

$$a_{10} = \tfrac{1}{4}\begin{bmatrix}-1 & 0 & 1 & -1 & 0 & 1\end{bmatrix} = A'$$
$$a_{20} = \tfrac{1}{12}\begin{bmatrix}1 & -2 & 1 & 1 & -2 & 1\end{bmatrix} = A''$$
$$a_{01} = \tfrac{1}{6}\begin{bmatrix}-1 & -1 & -1 & 1 & 1 & 1\end{bmatrix} = B$$
$$a_{11} = \tfrac{1}{4}\begin{bmatrix}1 & 0 & -1 & -1 & 0 & 1\end{bmatrix} = A'B$$
$$a_{21} = \tfrac{1}{12}\begin{bmatrix}-1 & 2 & -1 & 1 & -2 & 1\end{bmatrix} = A''B$$

4 x 2

$$a_{10} = \tfrac{1}{40}\begin{bmatrix}-3 & -1 & 1 & 3 & -3 & -1 & 1 & 3\end{bmatrix} = A'$$
$$a_{20} = \tfrac{1}{8}\begin{bmatrix}1 & -1 & -1 & 1 & 1 & -1 & -1 & 1\end{bmatrix} = A''$$
$$a_{01} = \tfrac{1}{8}\begin{bmatrix}-1 & -1 & -1 & -1 & 1 & 1 & 1 & 1\end{bmatrix} = B$$
$$a_{11} = \tfrac{1}{40}\begin{bmatrix}3 & 1 & -1 & -3 & -3 & -1 & 1 & 3\end{bmatrix} = A'B$$
$$a_{21} = \tfrac{1}{8}\begin{bmatrix}-1 & 1 & 1 & -1 & 1 & -1 & -1 & 1\end{bmatrix} = A''B$$

5 x 2

$$a_{10} = \frac{1}{20}\left[-2 \;\; -1 \;\; 0 \;\; 1 \;\; 2 \qquad -2 \;\; -1 \;\; 0 \;\; 1 \;\; 2\right] = A'$$

$$a_{20} = \frac{1}{28}\left[2 \;\; -1 \;\; -2 \;\; -1 \;\; 2 \qquad 2 \;\; -1 \;\; -2 \;\; -1 \;\; 2\right] = A''$$

$$a_{01} = \frac{1}{10}\left[-1 \;\; -1 \;\; -1 \;\; -1 \;\; -1 \qquad 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1\right] = B$$

$$a_{11} = \frac{1}{20}\left[2 \;\; 1 \;\; 0 \;\; -1 \;\; -2 \qquad -2 \;\; -1 \;\; 0 \;\; 1 \;\; 2\right] = A'B$$

$$a_{21} = \frac{1}{28}\left[-2 \;\; 1 \;\; 2 \;\; 1 \;\; -2 \qquad 2 \;\; -1 \;\; -2 \;\; -1 \;\; 2\right] = A''B$$

3 x 3

$$a_{10} = \frac{1}{6}\left[-1 \;\; 0 \;\; 1 \qquad -1 \;\; 0 \;\; 1 \qquad -1 \;\; 0 \;\; 1\right] = A'$$

$$a_{20} = \frac{1}{18}\left[1 \;\; -2 \;\; 1 \qquad 1 \;\; -2 \;\; 1 \qquad 1 \;\; -2 \;\; 1\right] = A''$$

$$a_{01} = \frac{1}{6}\left[-1 \;\; -1 \;\; -1 \qquad 0 \;\; 0 \;\; 0 \qquad 1 \;\; 1 \;\; 1\right] = B'$$

$$a_{11} = \frac{1}{4}\left[1 \;\; 0 \;\; -1 \qquad 0 \;\; 0 \;\; 0 \qquad -1 \;\; 0 \;\; 1\right] = A'B'$$

$$a_{21} = \frac{1}{12}\left[-1 \;\; 2 \;\; -1 \qquad 0 \;\; 0 \;\; 0 \qquad 1 \;\; -2 \;\; 1\right] = A''B'$$

$$a_{02} = \frac{1}{18}\left[1 \;\; 1 \;\; 1 \qquad -2 \;\; -2 \;\; -2 \qquad 1 \;\; 1 \;\; 1\right] = B''$$

$$a_{12} = \frac{1}{12}\left[-1 \;\; 0 \;\; 1 \qquad 2 \;\; 0 \;\; -2 \qquad -1 \;\; 0 \;\; 1\right] = A'B''$$

$$a_{22} = \frac{1}{36}\left[1 \;\; -2 \;\; 1 \qquad -2 \;\; 4 \;\; -2 \qquad 1 \;\; -2 \;\; 1\right] = A''B''$$

4 x 3

$$a_{10} = \frac{1}{60}\left[-3 \;\; -1 \;\; 1 \;\; 3 \qquad -3 \;\; -1 \;\; 1 \;\; 3 \qquad -3 \;\; -1 \;\; 1 \;\; 3\right] = A'$$

$$a_{20} = \frac{1}{12}\left[1 \;\; -1 \;\; -1 \;\; 1 \qquad 1 \;\; -1 \;\; -1 \;\; 1 \qquad 1 \;\; -1 \;\; -1 \;\; 1\right] = A''$$

$$a_{01} = \frac{1}{8}\left[-1 \;\; -1 \;\; -1 \;\; -1 \qquad 0 \;\; 0 \;\; 0 \;\; 0 \qquad 1 \;\; 1 \;\; 1 \;\; 1\right] = B'$$

$$a_{11} = \frac{1}{40}\left[3 \;\; 1 \;\; -1 \;\; -3 \qquad 0 \;\; 0 \;\; 0 \;\; 0 \qquad -3 \;\; -1 \;\; 1 \;\; 3\right] = A'B'$$

$$a_{21} = \frac{1}{8}\left[-1 \;\; 1 \;\; 1 \;\; -1 \qquad 0 \;\; 0 \;\; 0 \;\; 0 \qquad 1 \;\; -1 \;\; -1 \;\; 1\right] = A''B'$$

$$a_{02} = \frac{1}{24}\left[1 \;\; 1 \;\; 1 \;\; 1 \qquad -2 \;\; -2 \;\; -2 \;\; -2 \qquad 1 \;\; 1 \;\; 1 \;\; 1\right] = B''$$

$$a_{12} = \frac{1}{120}\left[-3 \;\; -1 \;\; 1 \;\; 3 \qquad 6 \;\; 2 \;\; -2 \;\; -6 \qquad -3 \;\; -1 \;\; 1 \;\; 3\right] = A'B''$$

$$a_{22} = \frac{1}{24}\left[1 \;\; -1 \;\; -1 \;\; 1 \qquad -2 \;\; 2 \;\; 2 \;\; -2 \qquad 1 \;\; -1 \;\; -1 \;\; 1\right] = A''B''$$

5 x 3

$$a_{10}=\frac{1}{30}\left[-2\ -1\ \ 0\ \ 1\ \ 2 \qquad -2\ -1\ \ 0\ \ 1\ \ 2 \qquad -2\ -1\ \ 0\ \ 1\ \ 2\right]=A'$$

$$a_{20}=\frac{1}{42}\left[2\ -1\ -2\ -1\ \ 2 \qquad 2\ -1\ -2\ -1\ \ 2 \qquad 2\ -1\ -2\ -1\ \ 2\right]=A''$$

$$a_{01}=\frac{1}{10}\left[-1\ -1\ -1\ -1\ -1 \qquad 0\ \ 0\ \ 0\ \ 0\ \ 0 \qquad 1\ \ 1\ \ 1\ \ 1\ \ 1\right]=B'$$

$$a_{11}=\frac{1}{20}\left[2\ \ 1\ \ 0\ -1\ -2 \qquad 0\ \ 0\ \ 0\ \ 0\ \ 0 \qquad -2\ -1\ \ 0\ \ 1\ \ 2\right]=A'B'$$

$$a_{21}=\frac{1}{28}\left[-2\ \ 1\ \ 2\ \ 1\ -2 \qquad 0\ \ 0\ \ 0\ \ 0\ \ 0 \qquad 2\ -1\ -2\ -1\ \ 2\right]=A''B'$$

$$a_{02}=\frac{1}{30}\left[1\ \ 1\ \ 1\ \ 1\ \ 1 \qquad -2\ -2\ -2\ -2\ -2 \qquad 1\ \ 1\ \ 1\ \ 1\ \ 1\right]=B''$$

$$a_{12}=\frac{1}{60}\left[-2\ -1\ \ 0\ \ 1\ \ 2 \qquad 4\ \ 2\ \ 0\ -2\ -4 \qquad -2\ -1\ \ 0\ \ 1\ \ 2\right]=A'B''$$

$$a_{22}=\frac{1}{84}\left[2\ -1\ -2\ -1\ \ 2 \qquad -4\ \ 2\ \ 4\ \ 2\ -4 \qquad 2\ -1\ -2\ -1\ \ 2\right]=A''B''$$

4 x 4

$$a_{10}=\frac{1}{80}\left[-3\ -1\ \ 1\ \ 3 \qquad -3\ -1\ \ 1\ \ 3 \qquad -3\ -1\ \ 1\ \ 3 \qquad -3\ -1\ \ 1\ \ 3\right]=A'$$

$$a_{20}=\frac{1}{16}\left[1\ -1\ -1\ \ 1 \qquad 1\ -1\ -1\ \ 1 \qquad 1\ -1\ -1\ \ 1 \qquad 1\ -1\ -1\ \ 1\right]=A''$$

$$a_{01}=\frac{1}{80}\left[-3\ -3\ -3\ -3 \qquad -1\ -1\ -1\ -1 \qquad 1\ \ 1\ \ 1\ \ 1 \qquad 3\ \ 3\ \ 3\ \ 3\right]=B'$$

$$a_{11}=\frac{1}{400}\left[9\ \ 3\ -3\ -9 \qquad 3\ \ 1\ -1\ -3 \qquad -3\ -1\ \ 1\ \ 3 \qquad -9\ -3\ \ 3\ \ 9\right]=A'B'$$

$$a_{21}=\frac{1}{80}\left[-3\ \ 3\ \ 3\ -3 \qquad -1\ \ 1\ \ 1\ -1 \qquad 1\ -1\ -1\ \ 1 \qquad 3\ -3\ -3\ \ 3\right]=A''B'$$

$$a_{02}=\frac{1}{16}\left[1\ \ 1\ \ 1\ \ 1 \qquad -1\ -1\ -1\ -1 \qquad -1\ -1\ -1\ -1 \qquad 1\ \ 1\ \ 1\ \ 1\right]=B''$$

$$a_{12}=\frac{1}{80}\left[-3\ -1\ \ 1\ \ 3 \qquad 3\ \ 1\ -1\ -3 \qquad 3\ \ 1\ -1\ -3 \qquad -3\ -1\ \ 1\ \ 3\right]=A'B''$$

$$a_{22}=\frac{1}{16}\left[1\ -1\ -1\ \ 1 \qquad -1\ \ 1\ \ 1\ -1 \qquad -1\ \ 1\ \ 1\ -1 \qquad 1\ -1\ -1\ \ 1\right]=A''B''$$

5 x 4

$$a_{10}=\frac{1}{40}\left[-2\ -1\ \ 0\ \ 1\ \ 2 \qquad -2\ -1\ \ 0\ \ 1\ \ 2 \qquad -2\ -1\ \ 0\ \ 1\ \ 2 \qquad -2\ -1\ \ 0\ \ 1\ \ 2\right]=A'$$

$$a_{20}=\frac{1}{56}\left[2\ -1\ -2\ -1\ \ 2 \qquad 2\ -1\ -2\ -1\ \ 2 \qquad 2\ -1\ -2\ -1\ \ 2 \qquad 2\ -1\ -2\ -1\ \ 2\right]=A''$$

$$a_{01}=\frac{1}{100}\left[-3\ -3\ -3\ -3\ -3 \qquad -1\ -1\ -1\ -1\ -1 \qquad 1\ \ 1\ \ 1\ \ 1\ \ 1 \qquad 3\ \ 3\ \ 3\ \ 3\ \ 3\right]=B'$$

$$a_{11}=\frac{1}{200}\left[6\ \ 3\ \ 0\ -3\ -6 \qquad 2\ \ 1\ \ 0\ -1\ -2 \qquad -2\ -1\ \ 0\ \ 1\ \ 2 \qquad -6\ -3\ \ 0\ \ 3\ \ 6\right]=A'B'$$

$$a_{21}=\frac{1}{280}\left[-6\ \ 3\ \ 6\ \ 3\ -6 \qquad -2\ \ 1\ \ 2\ \ 1\ -2 \qquad 2\ -1\ -2\ -1\ \ 2 \qquad 6\ -3\ -6\ -3\ \ 6\right]=A''B'$$

$$a_{02}=\frac{1}{20}\left[1\ \ 1\ \ 1\ \ 1\ \ 1 \qquad -1\ -1\ -1\ -1\ -1 \qquad -1\ -1\ -1\ -1\ -1 \qquad 1\ \ 1\ \ 1\ \ 1\ \ 1\right]=B''$$

$$a_{12}=\frac{1}{40}\left[-2\ -1\ \ 0\ \ 1\ \ 2 \qquad 2\ \ 1\ \ 0\ -1\ -2 \qquad 2\ \ 1\ \ 0\ -1\ -2 \qquad -2\ -1\ \ 0\ \ 1\ \ 2\right]=A'B''$$

$$a_{22}=\frac{1}{56}\left[2\ -1\ -2\ -1\ \ 2 \qquad -2\ \ 1\ \ 2\ \ 1\ -2 \qquad -2\ \ 1\ \ 2\ \ 1\ -2 \qquad 2\ -1\ -2\ -1\ \ 2\right]=A''B''$$

5 x 5

$$a_{10} = \frac{1}{50}\left[-2\ -1\ \ 0\ \ 1\ \ 2\quad -2\ -1\ \ 0\ \ 1\ \ 2\quad -2\ -1\ \ 0\ \ 1\ \ 2\quad -2\ -1\ \ 0\ \ 1\ \ 2\quad -2\ -1\ \ 0\ \ 1\ \ 2\right] = A'$$

$$a_{20} = \frac{1}{70}\left[2\ -1\ -2\ -1\ \ 2\quad 2\ -1\ -2\ -1\ \ 2\quad 2\ -1\ -2\ -1\ \ 2\quad 2\ -1\ -2\ -1\ \ 2\quad 2\ -1\ -2\ -1\ \ 2\right] = A''$$

$$a_{01} = \frac{1}{50}\left[-2\ -2\ -2\ -2\ -2\quad -1\ -1\ -1\ -1\ -1\quad 0\ \ 0\ \ 0\ \ 0\ \ 0\quad 1\ \ 1\ \ 1\ \ 1\ \ 1\quad 2\ \ 2\ \ 2\ \ 2\ \ 2\right] = B'$$

$$a_{11} = \frac{1}{100}\left[4\ \ 2\ \ 0\ -2\ -4\quad 2\ \ 1\ \ 0\ -1\ -2\quad 0\ \ 0\ \ 0\ \ 0\ \ 0\quad -2\ -1\ \ 0\ \ 1\ \ 2\quad -4\ -2\ \ 0\ \ 2\ \ 4\right] = A'B'$$

$$a_{21} = \frac{1}{140}\left[-4\ \ 2\ \ 4\ \ 2\ -4\quad -2\ \ 1\ \ 2\ \ 1\ -2\quad 0\ \ 0\ \ 0\ \ 0\ \ 0\quad 2\ -1\ -2\ -1\ \ 2\quad 4\ -2\ -4\ -2\ \ 4\right] = A''B'$$

$$a_{02} = \frac{1}{70}\left[2\ \ 2\ \ 2\ \ 2\ \ 2\quad -1\ -1\ -1\ -1\ -1\quad -2\ -2\ -2\ -2\ -2\quad -1\ -1\ -1\ -1\ -1\quad 2\ \ 2\ \ 2\ \ 2\ \ 2\right] = B''$$

$$a_{12} = \frac{1}{140}\left[-4\ -2\ \ 0\ \ 2\ \ 4\quad 2\ \ 1\ \ 0\ -1\ -2\quad 4\ \ 2\ \ 0\ -2\ -4\quad 2\ \ 1\ \ 0\ -1\ -2\quad -4\ -2\ \ 0\ \ 2\ \ 4\right] = A'B''$$

$$a_{22} = \frac{1}{196}\left[4\ -2\ -4\ -2\ \ 4\quad -2\ \ 1\ \ 2\ \ 1\ -2\quad -4\ \ 2\ \ 4\ \ 2\ -4\quad -2\ \ 1\ \ 2\ \ 1\ -2\quad 4\ -2\ -4\ -2\ \ 4\right] = A''B''$$

In the following three-factor tables the vectors are applicable to a vector of treatment yields of the form $\{u_{111}\ \ u_{211}\ldots u_{m11}\ \ u_{121}\ \ u_{221}\ldots u_{m21}\ldots u_{1n1}\ \ u_{2n1}\ldots u_{mn1}\ldots u_{11p}\ \ u_{21p}\ldots u_{m1p}\ldots\ u_{1np}\,u_{2np}\ldots u_{mnp}\}$

<u>3 x 2 x 2</u>

$$a_{100}=\tfrac{1}{8}\left[-1\ \ 0\ \ 1\quad -1\ \ 0\ \ 1\quad -1\ \ 0\ \ 1\quad -1\ \ 0\ \ 1\right]=A'$$

$$a_{200}=\tfrac{1}{24}\left[1\ -2\ \ 1\quad 1\ -2\ \ 1\quad 1\ -2\ \ 1\quad 1\ -2\ \ 1\right]=A''$$

$$a_{010}=\tfrac{1}{12}\left[-1\ -1\ -1\quad 1\ \ 1\ \ 1\quad -1\ -1\ -1\quad 1\ \ 1\ \ 1\right]=B$$

$$a_{110}=\tfrac{1}{8}\left[1\ \ 0\ -1\quad -1\ \ 0\ \ 1\quad 1\ \ 0\ -1\quad -1\ \ 0\ \ 1\right]=A'B$$

$$a_{210}=\tfrac{1}{24}\left[-1\ \ 2\ -1\quad 1\ -2\ \ 1\quad -1\ \ 2\ -1\quad 1\ -2\ \ 1\right]=A''B$$

$$a_{001}=\tfrac{1}{12}\left[-1\ -1\ -1\quad -1\ -1\ -1\quad 1\ \ 1\ \ 1\quad 1\ \ 1\ \ 1\right]=C$$

$$a_{101}=\tfrac{1}{8}\left[1\ \ 0\ -1\quad 1\ \ 0\ -1\quad -1\ \ 0\ \ 1\quad -1\ \ 0\ \ 1\right]=A'C$$

$$a_{201}=\tfrac{1}{24}\left[-1\ \ 2\ -1\quad -1\ \ 2\ -1\quad 1\ -2\ \ 1\quad 1\ -2\ \ 1\right]=A''C$$

$$a_{011}=\tfrac{1}{12}\left[1\ \ 1\ \ 1\quad -1\ -1\ -1\quad -1\ -1\ -1\quad 1\ \ 1\ \ 1\right]=B\,C$$

$$a_{111}=\tfrac{1}{8}\left[-1\ \ 0\ \ 1\quad 1\ \ 0\ -1\quad 1\ \ 0\ -1\quad -1\ \ 0\ \ 1\right]=A'BC$$

$$a_{211}=\tfrac{1}{24}\left[1\ -2\ \ 1\quad -1\ \ 2\ -1\quad -1\ \ 2\ -1\quad 1\ -2\ \ 1\right]=A''BC$$

<u>4 x 2 x 2</u>

$$a_{100}=\tfrac{1}{80}\left[-3\ -1\ \ 1\ \ 3\quad -3\ -1\ \ 1\ \ 3\quad -3\ -1\ \ 1\ \ 3\quad -3\ -1\ \ 1\ \ 3\right]=A'$$

$$a_{200}=\tfrac{1}{16}\left[1\ -1\ -1\ \ 1\quad 1\ -1\ -1\ \ 1\quad 1\ -1\ -1\ \ 1\quad 1\ -1\ -1\ \ 1\right]=A''$$

$$a_{010}=\tfrac{1}{16}\left[-1\ -1\ -1\ -1\quad 1\ \ 1\ \ 1\ \ 1\quad -1\ -1\ -1\ -1\quad 1\ \ 1\ \ 1\ \ 1\right]=B$$

$$a_{110}=\tfrac{1}{80}\left[3\ \ 1\ -1\ -3\quad -3\ -1\ \ 1\ \ 3\quad 3\ \ 1\ -1\ -3\quad -3\ -1\ \ 1\ \ 3\right]=A'B$$

$$a_{210}=\tfrac{1}{16}\left[-1\ \ 1\ \ 1\ -1\quad 1\ -1\ -1\ \ 1\quad -1\ \ 1\ \ 1\ -1\quad 1\ -1\ -1\ \ 1\right]=A''B$$

$$a_{001}=\tfrac{1}{16}\left[-1\ -1\ -1\ -1\quad -1\ -1\ -1\ -1\quad 1\ \ 1\ \ 1\ \ 1\quad 1\ \ 1\ \ 1\ \ 1\right]=C$$

$$a_{101}=\tfrac{1}{80}\left[3\ \ 1\ -1\ -3\quad 3\ \ 1\ -1\ -3\quad -3\ -1\ \ 1\ \ 3\quad -3\ -1\ \ 1\ \ 3\right]=A'C$$

$$a_{201}=\tfrac{1}{16}\left[-1\ \ 1\ \ 1\ -1\quad -1\ \ 1\ \ 1\ -1\quad 1\ -1\ -1\ \ 1\quad 1\ -1\ -1\ \ 1\right]=A''C$$

$$a_{011}=\tfrac{1}{16}\left[1\ \ 1\ \ 1\ \ 1\quad -1\ -1\ -1\ -1\quad -1\ -1\ -1\ -1\quad 1\ \ 1\ \ 1\ \ 1\right]=BC$$

$$a_{111}=\tfrac{1}{80}\left[-3\ -1\ \ 1\ \ 3\quad 3\ \ 1\ -1\ -3\quad 3\ \ 1\ -1\ -3\quad -3\ -1\ \ 1\ \ 3\right]=A'BC$$

$$a_{211}=\tfrac{1}{16}\left[1\ -1\ -1\ \ 1\quad -1\ \ 1\ \ 1\ -1\quad -1\ \ 1\ \ 1\ -1\quad 1\ -1\ -1\ \ 1\right]=A''BC$$

5 x 2 x 2

$$a_{100} = \frac{1}{40}\left[-2\ -1\ \ 0\ \ 1\ \ 2 \quad -2\ -1\ \ 0\ \ 1\ \ 2 \quad -2\ -1\ \ 0\ \ 1\ \ 2 \quad -2\ -1\ \ 0\ \ 1\ \ 2\right] = A'$$

$$a_{200} = \frac{1}{56}\left[2\ -1\ -2\ -1\ \ 2 \quad 2\ -1\ -2\ -1\ \ 2 \quad 2\ -1\ -2\ -1\ \ 2 \quad 2\ -1\ -2\ -1\ \ 2\right] = A''$$

$$a_{010} = \frac{1}{20}\left[-1\ -1\ -1\ -1\ -1 \quad 1\ \ 1\ \ 1\ \ 1\ \ 1 \quad -1\ -1\ -1\ -1\ -1 \quad 1\ \ 1\ \ 1\ \ 1\ \ 1\right] = B$$

$$a_{110} = \frac{1}{40}\left[2\ \ 1\ \ 0\ -1\ -2 \quad -2\ -1\ \ 0\ \ 1\ \ 2 \quad 2\ \ 1\ \ 0\ -1\ -2 \quad -2\ -1\ \ 0\ \ 1\ \ 2\right] = A'B$$

$$a_{210} = \frac{1}{56}\left[-2\ \ 1\ \ 2\ \ 1\ -2 \quad 2\ -1\ -2\ -1\ \ 2 \quad -2\ \ 1\ \ 2\ \ 1\ -2 \quad 2\ -1\ -2\ -1\ \ 2\right] = A''B$$

$$a_{001} = \frac{1}{20}\left[-1\ -1\ -1\ -1\ -1 \quad -1\ -1\ -1\ -1\ -1 \quad 1\ \ 1\ \ 1\ \ 1\ \ 1 \quad 1\ \ 1\ \ 1\ \ 1\ \ 1\right] = C$$

$$a_{101} = \frac{1}{40}\left[2\ \ 1\ \ 0\ -1\ -2 \quad 2\ \ 1\ \ 0\ -1\ -2 \quad -2\ -1\ \ 0\ \ 1\ \ 2 \quad -2\ -1\ \ 0\ \ 1\ \ 2\right] = A'C$$

$$a_{201} = \frac{1}{56}\left[-2\ \ 1\ \ 2\ \ 1\ -2 \quad -2\ \ 1\ \ 2\ \ 1\ -2 \quad 2\ -1\ -2\ -1\ \ 2 \quad 2\ -1\ -2\ -1\ \ 2\right] = A''C$$

$$a_{011} = \frac{1}{20}\left[1\ \ 1\ \ 1\ \ 1\ \ 1 \quad -1\ -1\ -1\ -1\ -1 \quad -1\ -1\ -1\ -1\ -1 \quad 1\ \ 1\ \ 1\ \ 1\ \ 1\right] = BC$$

$$a_{111} = \frac{1}{40}\left[-2\ -1\ \ 0\ \ 1\ \ 2 \quad 2\ \ 1\ \ 0\ -1\ -2 \quad 2\ \ 1\ \ 0\ -1\ -2 \quad -2\ -1\ \ 0\ \ 1\ \ 2\right] = A'BC$$

$$a_{211} = \frac{1}{56}\left[2\ -1\ -2\ -1\ \ 2 \quad -2\ \ 1\ \ 2\ \ 1\ -2 \quad -2\ \ 1\ \ 2\ \ 1\ -2 \quad 2\ -1\ -2\ -1\ \ 2\right] = A''BC$$

3 x 3 x 2

a_{100} =	$\frac{1}{12}$	-1	0	1	-1	0	1	-1	0	1	-1	0	1	-1	0	1	-1	0	1	= A'
a_{200} =	$\frac{1}{36}$	1	-2	1	1	-2	1	1	-2	1	1	-2	1	1	-2	1	1	-2	1	= A"
a_{010} =	$\frac{1}{12}$	-1	-1	-1	0	0	0	1	1	1	-1	-1	-1	0	0	0	1	1	1	= B'
a_{110} =	$\frac{1}{8}$	1	0	-1	0	0	0	-1	0	1	1	0	-1	0	0	0	-1	0	1	= A'B'
a_{210} =	$\frac{1}{24}$	-1	2	-1	0	0	0	1	-2	1	-1	2	-1	0	0	0	1	-2	1	= A"B'
a_{020} =	$\frac{1}{36}$	1	1	1	-2	-2	-2	1	1	1	1	1	1	-2	-2	-2	1	1	1	B"
a_{120} =	$\frac{1}{24}$	-1	0	1	2	0	-2	-1	0	1	-1	0	1	2	0	-2	-1	0	1	= A'B"
a_{220} =	$\frac{1}{72}$	1	-2	1	-2	4	-2	1	-2	1	1	-2	1	-2	4	-2	1	-2	1	= A"B"
a_{001} =	$\frac{1}{18}$	-1	-1	-1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	1	1	1	= C
a_{101} =	$\frac{1}{12}$	1	0	-1	1	0	-1	1	0	-1	-1	0	1	-1	0	1	-1	0	1	= A'C
a_{201} =	$\frac{1}{36}$	-1	2	-1	-1	2	-1	-1	2	-1	1	-2	1	1	-2	1	1	-2	1	= A"C
a_{011} =	$\frac{1}{12}$	1	1	1	0	0	0	-1	-1	-1	-1	-1	-1	0	0	0	1	1	1	= B'C
a_{111} =	$\frac{1}{8}$	-1	0	1	0	0	0	1	0	-1	1	0	-1	0	0	0	-1	0	1	= A'B'C
a_{211} =	$\frac{1}{24}$	1	-2	1	0	0	0	-1	2	-1	-1	2	-1	0	0	0	1	-2	1	= A"B'C
a_{021} =	$\frac{1}{36}$	-1	-1	-1	2	2	2	-1	-1	-1	1	1	1	-2	-2	-2	1	1	1	= B"C
a_{121} =	$\frac{1}{24}$	1	0	-1	-2	0	2	1	0	-1	-1	0	1	2	0	-2	-1	0	1	= A'B"C
a_{221} =	$\frac{1}{72}$	-1	2	-1	2	-4	2	-1	2	-1	1	-2	1	-2	4	-2	1	-2	1	= A"B"C

3 x 3 x 3

$$a_{100} = \tfrac{1}{18}\left[-1\ 0\ 1 \quad -1\ 0\ 1 \quad -1\ 0\ 1 \quad -1\ 0\ 1 \quad -1\ 0\ 1 \quad -1\ 0\ 1 \quad -1\ 0\ 1 \quad -1\ 0\ 1 \quad -1\ 0\ 1\right] = A'$$

$$a_{200} = \tfrac{1}{54}\left[1\ {-2}\ 1 \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1\right] = A''$$

$$a_{010} = \tfrac{1}{18}\left[-1\ {-1}\ {-1} \quad 0\ 0\ 0 \quad 1\ 1\ 1 \quad -1\ {-1}\ {-1} \quad 0\ 0\ 0 \quad 1\ 1\ 1 \quad -1\ {-1}\ {-1} \quad 0\ 0\ 0 \quad 1\ 1\ 1\right] = B'$$

$$a_{110} = \tfrac{1}{12}\left[1\ 0\ {-1} \quad 0\ 0\ 0 \quad -1\ 0\ 1 \quad 1\ 0\ {-1} \quad 0\ 0\ 0 \quad -1\ 0\ 1 \quad 1\ 0\ {-1} \quad 0\ 0\ 0 \quad -1\ 0\ 1\right] = A'B'$$

$$a_{210} = \tfrac{1}{36}\left[-1\ 2\ {-1} \quad 0\ 0\ 0 \quad 1\ {-2}\ 1 \quad -1\ 2\ {-1} \quad 0\ 0\ 0 \quad 1\ {-2}\ 1 \quad -1\ 2\ {-1} \quad 0\ 0\ 0 \quad 1\ {-2}\ 1\right] = A''B'$$

$$a_{020} = \tfrac{1}{24}\left[1\ 1\ 1 \quad -2\ {-2}\ {-2} \quad 1\ 1\ 1 \quad 1\ 1\ 1 \quad -2\ {-2}\ {-2} \quad 1\ 1\ 1 \quad 1\ 1\ 1 \quad -2\ {-2}\ {-2} \quad 1\ 1\ 1\right] = B''$$

$$a_{120} = \tfrac{1}{36}\left[-1\ 0\ 1 \quad 2\ 0\ {-2} \quad -1\ 0\ 1 \quad -1\ 0\ 1 \quad 2\ 0\ {-2} \quad -1\ 0\ 1 \quad -1\ 0\ 1 \quad 2\ 0\ {-2} \quad -1\ 0\ 1\right] = A'B''$$

$$a_{220} = \tfrac{1}{108}\left[1\ {-2}\ 1 \quad -2\ 4\ {-2} \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1 \quad -2\ 4\ {-2} \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1 \quad -2\ 4\ {-2} \quad 1\ {-2}\ 1\right] = A''B''$$

$$a_{001} = \tfrac{1}{18}\left[-1\ {-1}\ {-1} \quad -1\ {-1}\ {-1} \quad -1\ {-1}\ {-1} \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 1\ 1\ 1 \quad 1\ 1\ 1 \quad 1\ 1\ 1\right] = C'$$

$$a_{101} = \tfrac{1}{12}\left[1\ 0\ {-1} \quad 1\ 0\ {-1} \quad 1\ 0\ {-1} \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad -1\ 0\ 1 \quad -1\ 0\ 1 \quad -1\ 0\ 1\right] = A'C'$$

$$a_{201} = \tfrac{1}{36}\left[-1\ 2\ {-1} \quad -1\ 2\ {-1} \quad -1\ 2\ {-1} \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1 \quad 1\ {-2}\ 1\right] = A''C'$$

$$a_{011} = \tfrac{1}{12}\left[1\ 1\ 1 \quad 0\ 0\ 0 \quad -1\ {-1}\ {-1} \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad -1\ {-1}\ {-1} \quad 0\ 0\ 0 \quad 1\ 1\ 1\right] = B'C'$$

$$a_{111} = \tfrac{1}{8}\left[-1\ 0\ 1 \quad 0\ 0\ 0 \quad 1\ 0\ {-1} \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 1\ 0\ {-1} \quad 0\ 0\ 0 \quad -1\ 0\ 1\right] = A'B'C'$$

$$a_{211} = \tfrac{1}{24}\left[1\ {-2}\ 1 \quad 0\ 0\ 0 \quad -1\ 2\ {-1} \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad -1\ 2\ {-1} \quad 0\ 0\ 0 \quad 1\ {-2}\ 1\right] = A''B'C'$$

$$a_{021} = \tfrac{1}{36}\left[-1\ {-1}\ {-1} \quad 2\ 2\ 2 \quad -1\ {-1}\ {-1} \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 1\ 1\ 1 \quad -2\ {-2}\ {-2} \quad 1\ 1\ 1\right] = B''C'$$

$$a_{121} = \tfrac{1}{24}\left[1\ 0\ {-1} \quad -2\ 0\ 2 \quad 1\ 0\ {-1} \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad 0\ 0\ 0 \quad -1\ 0\ 1 \quad 2\ 0\ {-2} \quad -1\ 0\ 1\right] = A'B''C'$$

continued on next page.

3 x 3 x 3 contd.

$$a_{221} = \frac{1}{72}\left[-1\ \ 2\ {-1} \quad 2\ {-4}\ \ 2 \quad -1\ \ 2\ {-1} \quad 0\ \ 0\ \ 0 \quad 0\ \ 0\ \ 0 \quad 0\ \ 0\ \ 0 \quad 1\ {-2}\ \ 1 \quad -2\ \ 4\ {-2} \quad 1\ {-2}\ \ 1\right] = A''B''C'$$

$$a_{002} = \frac{1}{54}\left[1\ \ 1\ \ 1 \quad 1\ \ 1\ \ 1 \quad 1\ \ 1\ \ 1 \quad -2\ {-2}\ {-2} \quad -2\ {-2}\ {-2} \quad -2\ {-2}\ {-2} \quad 1\ \ 1\ \ 1 \quad 1\ \ 1\ \ 1 \quad 1\ \ 1\ \ 1\right] = C''$$

$$a_{102} = \frac{1}{36}\left[-1\ \ 0\ \ 1 \quad -1\ \ 0\ \ 1 \quad -1\ \ 0\ \ 1 \quad 2\ \ 0\ {-2} \quad 2\ \ 0\ {-2} \quad 2\ \ 0\ {-2} \quad -1\ \ 0\ \ 1 \quad -1\ \ 0\ \ 1 \quad -1\ \ 0\ \ 1\right] = A'C''$$

$$a_{202} = \frac{1}{108}\left[1\ {-2}\ \ 1 \quad 1\ {-2}\ \ 1 \quad 1\ {-2}\ \ 1 \quad -2\ \ 4\ {-2} \quad -2\ \ 4\ {-2} \quad -2\ \ 4\ {-2} \quad 1\ {-2}\ \ 1 \quad 1\ {-2}\ \ 1 \quad 1\ {-2}\ \ 1\right] = A''C''$$

$$a_{012} = \frac{1}{36}\left[-1\ {-1}\ {-1} \quad 0\ \ 0\ \ 0 \quad 1\ \ 1\ \ 1 \quad 2\ \ 2\ \ 2 \quad 0\ \ 0\ \ 0 \quad -2\ {-2}\ {-2} \quad -1\ {-1}\ {-1} \quad 0\ \ 0\ \ 0 \quad 1\ \ 1\ \ 1\right] = B'C''$$

$$a_{112} = \frac{1}{24}\left[1\ \ 0\ {-1} \quad 0\ \ 0\ \ 0 \quad -1\ \ 0\ \ 1 \quad -2\ \ 0\ \ 2 \quad 0\ \ 0\ \ 0 \quad 2\ \ 0\ {-2} \quad 1\ \ 0\ {-1} \quad 0\ \ 0\ \ 0 \quad -1\ \ 0\ \ 1\right] = A'B'C''$$

$$a_{212} = \frac{1}{72}\left[-1\ \ 2\ {-1} \quad 0\ \ 0\ \ 0 \quad 1\ {-2}\ \ 1 \quad 2\ {-4}\ \ 2 \quad 0\ \ 0\ \ 0 \quad -2\ \ 4\ {-2} \quad -1\ \ 2\ {-1} \quad 0\ \ 0\ \ 0 \quad 1\ {-2}\ \ 1\right] = A''B'C''$$

$$a_{022} = \frac{1}{108}\left[1\ \ 1\ \ 1 \quad -2\ {-2}\ {-2} \quad 1\ \ 1\ \ 1 \quad -2\ {-2}\ {-2} \quad 4\ \ 4\ \ 4 \quad -2\ {-2}\ {-2} \quad 1\ \ 1\ \ 1 \quad -2\ {-2}\ {-2} \quad 1\ \ 1\ \ 1\right] = B''C''$$

$$a_{122} = \frac{1}{72}\left[-1\ \ 0\ \ 1 \quad 2\ \ 0\ {-2} \quad -1\ \ 0\ \ 1 \quad 2\ \ 0\ {-2} \quad -4\ \ 0\ \ 4 \quad 2\ \ 0\ {-2} \quad -1\ \ 0\ \ 1 \quad 2\ \ 0\ {-2} \quad -1\ \ 0\ \ 1\right] = A'B''C''$$

$$a_{222} = \frac{1}{216}\left[1\ {-2}\ \ 1 \quad -2\ \ 4\ {-2} \quad 1\ {-2}\ \ 1 \quad -2\ \ 4\ {-2} \quad 4\ {-8}\ \ 4 \quad -2\ \ 4\ {-2} \quad 1\ {-2}\ \ 1 \quad -2\ \ 4\ {-2} \quad 1\ {-2}\ \ 1\right] = A''B''C''$$

Section 14.

Graduation of Bivariate Data by Orthogonal Polynomials and Least Squares.

The discussion of the previous section suggests that a general method of graduation by means of orthogonal polynomials for a set of m x n data, independent and of equal weight, for $x = 0,1,2,\ldots m-1$, $y = 0,1,2,\ldots n-1$, might be investigated. We therefore extend to the bivariate case Aitken's method (2) of graduation of a set of univariate data. Were it not for the fact that the theoretical and practical work becomes unduly unwieldy, there seems no reason why the multivariate case should not be similarly treated, and indeed those cases of most common occurrence in agriculture, namely three or four variates (factors) graduated by orthogonal polynomials up to degree 2 in each variate should not present undue difficulty. It was seen in the preceding section that the orthogonal polynomials determining e.g. a_{21} were obtained by multiplication of corresponding polynomials in the sets determining a_{20} and a_{01}. In the same way we find that in the present case the bivariate orthogonal polynomials are obtained by similar multiplications of the Tchebycheff polynomials of the univariate case.

In the following exposition we shall need to make use of the calculus of finite differences for functions of two variables. The various formulae required are set out below: Δ_x, Σ_x, E_x denote operations with respect to x alone, Δ_y, Σ_y, E_y operations with respect to y alone. In dealing with the product function uv, the subscript 1 refers to operations on u alone and the subscript 2 to operations on v alone.

<u>Differencing</u> $\Delta_x\Delta_y u_{x,y} = (E_x-1)(E_y-1)u_{x,y}$

$$= (E_xE_y - E_x - E_y + 1)u_{x,y}$$

$$= u_{x+1,y+1} - u_{x+1,y} - u_{x,y+1} + u_{x,y}.$$

<u>Summation</u> $\sum_{x=1}^{m}\sum_{y=1}^{n}\phi(x,y)$, where $\phi(x,y) = \Delta_x\Delta_y f(x,y)$,

$$= \sum_{x=1}^{m}\left[\phi(x,1)+\phi(x,2)+\ldots+\phi(x,n)\right]$$

$$= \sum_{x=1}^{m}(1+E_y+E_y^2+\ldots+E_y^{n-1})\,\phi(x,1)$$

$$= \sum_{x=1}^{m}\Delta_x\frac{1-E_y^n}{1-E_y}(E_y-1)\,f(x,1) \quad \text{since } \phi(x,1)=\Delta_x\Delta_y f(x,1)$$

$$= (E_x^m-1)(E_y^n-1)\,f(1,1)$$

$$= f(m+1,n+1) - f(m+1,1) - f(1,n+1) + f(1,1)$$

<u>Indefinite Summation</u> If $\phi(x,y)=\Delta_x\Delta_y f(x,y)$, then $\Sigma_x\Sigma_y(x,y) = f(x,y)$ where upper and lower limits to both x and y may be introduced as above.

<u>Indefinite Summation of a Product</u>

$$\Delta_x\Delta_y(u_{x,y}v_{x,y}) = (E_{x1}E_{x2}-1)(E_{y1}E_{y2}-1)u_{x,y}v_{x,y}$$

$$= E_{x1}\Delta_{x2}E_{y1}\Delta_{y2}(1+E_{x1}^{-1}\Delta_{x1}\Delta_{x2}^{-1})(1+E_{y1}^{-1}\Delta_{y1}\Delta_{y2}^{-1})\,u_{x,y}v_{x,y}$$

In the following exposition we shall need to make use of the calculus of finite differences for functions of two variables. The various formulae required are set out below : Δ_x, Σ_x, E_x denote operations with respect to x alone, Δ_y, Σ_y, E_y operations with respect to y alone. In dealing with the product function uv, the subscript 1 refers to operations on u alone and the subscript 2 to operations on v alone.

<u>Differencing</u> $\Delta_x\Delta_y u_{x,y} = (E_x-1)(E_y-1)u_{x,y}$

$$= (E_xE_y - E_x - E_y + 1)u_{x,y}$$

$$= u_{x+1,y+1} - u_{x+1,y} - u_{x,y+1} + u_{x,y}.$$

<u>Summation</u> $\sum_{x=1}^{m}\sum_{y=1}^{n}\phi(x,y)$, where $\phi(x,y) = \Delta_x\Delta_y f(x,y)$,

$$= \sum_{x=1}^{m}\left[\phi(x,1)+\phi(x,2)+\ldots+\phi(x,n)\right]$$

$$= \sum_{x=1}^{m}(1+E_y+E_y^2+\ldots+E_y^{n-1})\,\phi(x,1)$$

$$= \sum_{x=1}^{m}\Delta_x\frac{1-E_y^n}{1-E_y}(E_y-1)\,f(x,1) \quad \text{since } \phi(x,1) = \Delta_x\Delta_y f(x,1)$$

$$= (E_x^m-1)(E_y^n-1)\,f(1,1)$$

$$= f(m+1,n+1) - f(m+1,1) - f(1,n+1) + f(1,1)$$

<u>Indefinite Summation</u> If $\phi(x,y) = \Delta_x\Delta_y f(x,y)$, then $\sum_x\sum_y(x,y) = f(x,y)$ where upper and lower limits to both x and y may be introduced as above.

<u>Indefinite Summation of a Product</u>

$$\Delta_x\Delta_y(u_{x,y}v_{x,y}) = (E_{x_1}E_{x_2}-1)(E_{y_1}E_{y_2}-1)\,u_{x,y}v_{x,y}$$

$$= E_{x_1}\Delta_{x_2}E_{y_1}\Delta_{y_2}(1+E_{x_1}^{-1}\Delta_{x_1}\Delta_{x_2}^{-1})(1+E_{y_1}^{-1}\Delta_{y_1}\Delta_{y_2}^{-1})\,u_{x,y}v_{x,y}$$

$$\therefore \sum_x \sum_y u_{x,y} v_{x,y} = \Delta_x^{-1} \Delta_y^{-1} u_{x,y} v_{x,y}$$

$$= E_{x_1}^{-1} \Delta_{x_2}^{-1} E_{y_1}^{-1} \Delta_{y_2}^{-1} \left(1 - E_{x_1}^{-1} \Delta_{x_1} \Delta_{x_2}^{-1} + E_{x_1}^{-2} \Delta_{x_1}^{2} \Delta_{x_2}^{-2} - \ldots\right)\left(1 - E_{y_1}^{-1} \Delta_{y_1} \Delta_{y_2}^{-1} + E_{y_1}^{-2} \Delta_{y_1}^{2} \Delta_{y_2}^{-2} - \ldots\right) u_{x,y} v_{x,y}$$

$$= \left(E_{x_1}^{-1} \Delta_{x_2}^{-1} E_{y_1}^{-1} \Delta_{y_2}^{-1} - E_{x_1}^{-2} \Delta_{x_1} \Delta_{x_2}^{-2} E_{y_1}^{-1} E_{y_2}^{-1} - E_{x_1}^{-1} \Delta_{x_2}^{-1} E_{y_1}^{-2} \Delta_{y_1} \Delta_{y_2}^{-2} + E_{x_1}^{-3} \Delta_{x_1}^{2} \Delta_{x_2}^{-3} E_{y_1}^{-1} \Delta_{y_2}^{-1}\right.$$

$$\left. + E_{x_1}^{-2} \Delta_{x_1} \Delta_{x_2}^{-2} E_{y_1}^{-2} \Delta_{y_1} \Delta_{y_2}^{-2} - E_{x_1}^{-1} \Delta_{x_2}^{-1} E_{y_1}^{-3} \Delta_{y_1}^{2} \Delta_{y_2}^{-3} + \ldots\right) u_{x,y} v_{x,y}$$

$$= u_{x-1,y-1} \sum_x \sum_y v_{x,y} - \Delta_x u_{x-2,y-1} \sum_x^2 \sum_y v_{x,y} - \Delta_y u_{x-1,y-2} \sum_x \sum_y^2 v_{x,y}$$

$$+ \Delta_x^2 u_{x-3,y-1} \sum_x^3 \sum_y v_{x,y} + \Delta_x \Delta_y u_{x-2,y-2} \sum_x^2 \sum_y^2 v_{x,y} + \Delta_y^2 u_{x-1,y-3} \sum_x \sum_y^3 v_{x,y} - \ldots$$

<u>Advancing Difference Formula.</u>

$$u_{x,y} = (1 + \Delta_x)^x \, (1 + \Delta_y)^y u_{0,0}$$

$$= (1 + x\Delta_x + x_{(2)}\Delta_x^2 + \ldots)(1 + y\Delta_y + y_{(2)}\Delta_y^2 + \ldots) u_{0,0}$$

$$= (1 + x\Delta_x + y\Delta_y + x_{(2)}\Delta_x^2 + xy\Delta_x\Delta_y + y_{(2)}\Delta_y^2 + \ldots) u_{0,0}$$

<u>Reduced Descending Factorials.</u> $\quad x_{(r)} y_{(s)} = \dfrac{x^{(r)} y^{(s)}}{\underline{|r}\,\underline{|s}}$

$$\Delta_x \Delta_y \left[x_{(r)} y_{(s)}\right] = \Delta_x \left[x_{(r)} y_{(s-1)}\right] = x_{(r-1)} y_{(s-1)}$$

$$\sum_x \sum_y x_{(r)} y_{(s)} = x_{(r+1)} y_{(s+1)}$$

Let the mn data corresponding to $x = 0,1,2,\ldots m-1$, $y = 0,1,2,\ldots n-1$ be represented by the vector

$$u = \{u_{0,0}\ u_{0,1}\ u_{0,2}\ldots u_{0,n-1}\ \ u_{1,0}\ u_{1,1}\ \ldots u_{1,n-1}\ldots u_{m-1,0}\ u_{m-1,1}\ldots u_{m-1,n-1}\}$$

The problem is to fit a polynomial $U(x,y : m,n)$ of degree (ℓ,k), where $\ell < m$, $k < n$, by the principle of least squares.

Let us take U in the form $a_{0,0} + a_{0,1}T_{0,1}(x,y)$ $+ a_{0,2}T_{0,2}(x,y) + \ldots + a_{0,k}T_{0,k}(x,y) + a_{1,0}T_{1,0}(x,y) + a_{1,1}T_{1,1}(x,y) + \ldots$ $+ a_{1,k}T_{1,k}(x,y) + \ldots + a_{\ell,k}T_{\ell,k}(x,y)$, where, for example, $T_{\ell,k}(x,y)$ is a polynomial in x of degree not $> \ell$ and in y of degree not $> k$. Also we impose on the T-polynomials the orthogonal conditions $\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} T_{p,q}(x,y)\ T_{r,s}(x,y) = 0$ for $r \neq p$ or $s \neq q$, $\neq 0$ when $r = p$ and $s = q$ simultaneously.

The observational equations are $u = Ta$, where the rows of T are $[1\ \ T_{0,1}\ T_{0,2}\ldots T_{0,k}\ T_{1,0}\ T_{1,1}\ldots T_{1,k}\ldots T_{\ell,k}]$ for $x=0$, $y=0$; $x=0$, $y=1$; $x=0$, $y=2$;... $x=0$, $y=n-1$; $x=1$, $y=0$; $x=1$, $y=1$;... $x=1$, $y=n-1$;.... $x=m-1$, $y=n-1$, and

$$a = \{a_{0,0}\ a_{0,1}\ a_{0,2}\ldots a_{0,k}\ a_{1,0}\ a_{1,1}\ldots a_{1,k}\ldots a_{\ell k}\}.$$

$$T'T = \text{diag.}\left[\ mn\ \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} T^2_{0,1}(x,y)\ \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} T^2_{0,2}(x,y)\ldots\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} T^2_{0,k}(x,y)\ldots\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} T^2_{\ell,k}(x,y)\right]$$

Hence the normal equations $T'Ta = T'u$ give the $a_{i,j}$ independently of each other as

$$a_{i,j} = \frac{\sum_x\sum_y u_{x,y}\,T_{i,j}(x,y)}{\sum_x\sum_y T^2_{i,j}(x,y)} \quad \ldots\ldots(14,1)$$

Since the u's are independent and of equal weight, the residual sum of squares $= (u-Ta)'(u-Ta) = u'u - a'T'Ta$

$$= \sum_i\sum_j u^2_{i,j} - \sum_i\sum_j\left[a^2_{i,j}\sum_x\sum_y T^2_{i,j}(x,y)\right]$$

$$= \sum_i\sum_j u^2_{i,j} - \sum_i\sum_j\left[a_{i,j} \times \text{Numerator on R.H.S. of } (14,1)\right],$$

the sum of squared residuals being reduced by a single term for each $a_{i,j}$.

<u>Derivation of the T-Polynomials</u> $T_{p,q}(x,y)$, a polynomial of degree p in x, q in y, may be expressed as

$$c_{p,q}P_{p,q}(x,y) + c_{p,q-1}P_{p,q-1}(x,y) + c_{p-1,q}P_{p-1,q}(x,y) + \ldots + c_{0,0}P_{0,0}(x,y)$$

where $P_{p,q}(x,y)$ is an arbitrary polynomial of degree p in x, q in y.

$$\therefore \sum_{x=0}^{m-1}\sum_{y=0}^{n-1} T_{p,q}(x,y)\; T_{r,s}(x,y) = \sum_x\sum_y \left\{T_{r,s}(x,y)\left[c_{p,q}P_{p,q}(x,y) + c_{p,q-1}P_{p,q-1}(x,y) + \ldots + c_{0,0}P_{0,0}(x,y)\right]\right\}.$$

Hence the orthogonal conditions are equivalent to

$$\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} P_{p,q}(x,y)\quad T_{r,s}(x,y) = 0$$

for all values of $p \leqq r$, $q \leqq s$, but not $p=r$, $q=s$ simultaneously. If now in the formula for summation of a product we put $u_{x,y} = (x+p)_{(p)}\,(y+q)_{(q)}$, then since $(x+p-a)_{(p-a+1)} = 0$ when $x=0$ for $a=1,2,\ldots p$ and $(y+q-b)_{(q-b+1)} = 0$ when $y=0$ for $b=1,2,\ldots q$, we have

$$\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} u_{x,y}v_{x,y} = u_{m-1,n-1}\sum_0^{m-1}\sum_0^{n-1} v_{x,y} - \Delta_x u_{m-2,n-1}\left(\sum_0^{m-1}\right)^2\sum_0^{n-1} v_{x,y}$$
$$- \Delta_y u_{m-1,n-2}\sum_0^{m-1}\left(\sum_0^{n-1}\right)^2 v_{x,y} + \ldots\ldots$$

$\therefore$ The orthogonal conditions are

$$(m+p-1)_{(p)}(n+q-1)_{(q)}\sum\sum T_{r,s}(x,y) - (m+p-2)_{(p-1)}(n+q-1)_{(q)}\sum^2\sum T_{r,s}(x,y)$$
$$-(m+p-1)_{(p)}(n+q-2)_{(q-1)}\sum\sum^2 T_{r,s}(x,y) + (m+p-3)_{(p-2)}(n+q-1)_{(q)}\sum^3\sum T_{r,s}(x,y)$$
$$+(m+p-2)_{(p-1)}(n+q-2)_{(q-1)}\sum^2\sum^2 T_{r,s}(x,y) + (m+p-1)_{(p)}(n+q-3)_{(q-2)}\sum\sum^3 T_{r,s}(x,y)$$
$$-\ldots\ldots\ldots = 0$$

, for all values of $p \leqq r$, $q \leqq s$, but not

$p=r$, $q=s$ simultaneously. Taking these equations in order from the lowest ($p=0$, $q=0$), we derive in succession
$\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} T_{r,s}(x,y) = 0 = \sum^2\sum T_{r,s}(x,y) = \sum\sum^2 T_{r,s}(x,y) = \sum^3\sum T_{r,s}(x,y)$
$= \sum^2\sum^2 T_{r,s}(x,y) = \sum\sum^3 T_{r,s}(x,y) = \ldots = \sum^{r+1}\sum^{s-1} T_{r,s}(x,y) = \sum^{r-1}\sum^{s+1} T_{r,s}(x,y)$
$= \sum^r\sum^s T_{r,s}(x,y)$ where r and s are both positive. Putting $\sum_x^r\sum_y^s T_{r,s}(x,y) \equiv G_{2r,2s}(x,y)$, a polynomial of degree $(2r,2s)$, we have that $G_{2r,2s}(x,y)$ and its differences with respect to x, $\Delta_x G_{2r,2s}(x,y), \Delta_x^2 G_{2r,2s}(x,y), \ldots \Delta_x^{r-1} G_{2r,2s}(x,y)$ vanish at $x=0$ and $x=m$, and hence that $G_{2r,2s}(x,y)$ contains the factors $x(x-1)(x-2)\ldots(x-r+1)(x-m)(x-m-1)(x-m-2)\ldots(x-m-r+1)$. Similarly $G_{2r,2s}(x,y)$ and its differences with respect to y, $\Delta_y G_{2r,2s}(x,y), \Delta_y^2 G_{2r,2s}(x,y), \ldots \Delta_y^{s-1} G_{2r,2s}(x,y)$ vanish at $y=0$ and $y=n$, and hence $G_{2r,2s}(x,y)$ contains the factors $y(y-1)\ldots(y-s+1)(y-n)(y-n-1)\ldots(y-n-s+1)$. $\therefore$ Except for a numerical factor, $G_{2r,2s}(x,y)$ is the product of these $(2r+2s)$ factors. Also $\Delta_x^r\Delta_y^s G_{2r,2s}(x,y) = T_{r,s}(x,y)$, $\therefore T_{r,s}(x,y) = \Delta_x^r\Delta_y^s x_{(r)}(x-m)_{(r)}\, y_{(s)}(y-n)_{(s)}$, the numerical factor being chosen as $\frac{1}{(\lfloor r \lfloor s)^2}$ because then $x_{(r)}$, $(x-m)_{(r)}$, $y_{(s)}$, and $(y-n)_{(s)}$ are all integers, x and y being integral.

$\therefore T_{r,s}(x,y)$ is an integer.

To obtain $T_{r,s}(x,y)$ we first express $(x-m)_{(r)}(y-n)_{(s)}$ as a Gregory-Newton series thus:-

$$(x-m)_{(r)}(y-n)_{(s)} = (x-r)_{(r)}(y-s)_{(s)} - (m-r)(x-r)_{(r-1)}(y-s)_{(s)}$$
$$-(n-s)(x-r)_{(r)}(y-s)_{(s-1)} + (m-r+1)_{(2)}(x-r)_{(r-2)}(y-s)_{(s)}$$
$$+(m-r)(n-s)(x-r)_{(r-1)}(y-s)_{(s-1)} + (n-s+1)_{(2)}(x-r)_{(r)}(y-s)_{(s-2)} - \ldots$$
$$+(-1)^{r+s}(m-1)_{(r)}(n-1)_{(s)}.$$

We then have $x_{(r)}(x-m)_{(r)}\,y_{(s)}(y-n)_{(s)} = (2r)_{(r)}(2s)_{(s)}\,x_{(2r)}\,y_{(2s)}$
$$-(m-r)(2r-1)_{(r)}(2s)_{(s)}\,x_{(2r-1)}\,y_{(2s)} - (n-s)(2r)_{(r)}(2s-1)_{(s)}\,x_{(2r)}\,y_{(2s-1)}$$
$$+(m-r+1)_{(2)}(2r-2)_{(r)}(2s)_{(s)}\,x_{(2r-2)}\,y_{(2s)} + (m-r)(n-s)(2r-1)_{(r)}(2s-1)_{(s)}\,x_{(2r-1)}\,y_{(2s-1)}$$
$$+(n-s+1)_{(2)}(2r)_{(r)}(2s-2)_{(s)}\,x_{(2r)}\,y_{(2s-2)} + \ldots + (-1)^{r+s}(m-1)_{(r)}(n-1)_{(s)}\,x_{(r)}\,y_{(s)},$$

using the identities $x_{(r)}(x-r)_{(r)} = (2r)_{(r)}\,x_{(2r)}$ and $x_{(r)}(x-r)_{(r-1)} = (2r-1)_{(r)}\,x_{(2r-1)}$. Hence, reducing the suffixes of x and y by r and s respectively we have

$$T_{r,s}(x,y) = \Delta_x^r\Delta_y^s\left[x_{(r)}(x-m)_{(r)}\,y_{(s)}(y-n)_{(s)}\right]$$
$$= (2r)_{(r)}(2s)_{(s)}\,x_{(r)}\,y_{(s)} - (m-r)(2r-1)_{(r)}(2s)_{(s)}\,x_{(r-1)}\,y_{(s)} - \ldots\ldots$$
$$+(-1)^{r+s}(m-1)_{(r)}(n-1)_{(s)}, \qquad \ldots.(14,2)$$

the general term, the $(h,k)^{th}$, being

$$(-1)^{h+k}(2r-h)_{(r)}(2s-k)_{(s)}\,x_{(r-h)}\,y_{(s-k)}\,(m-r+h-1)_{(h)}(n-s+k-1)_{(k)}.$$

But (14,2) is a Gregory-Newton advancing-difference expansion written in reverse, the term corresponding to that containing $\Delta_x^p\Delta_y^q T_{r,s}(o,o)$ being the $(r-p,s-q)^{th}$, i.e. $\Delta_x^p\Delta_y^q\,T_{r,s}(o,o) = (-1)^{r+s-p-q}(r+p)_{(r)}(s+q)_{(s)}$

$$(m-p-1)_{(r-p)}(n-q-1)_{(s-q)} \qquad \ldots.(14,3)$$

Now (14,3) is simply the product of two univariate polynomial values $\Delta^p T_r(o)$ and $\Delta^q T_s(o)$, so that tables of terminal values and differences for the bivariate case may be constructed from the univariate case by appropriate multiplications. For example, let us take the case $m = 7$ $r = 3$; $n = 6$, $s = 3$. We have the following univariate tables in which (as we shall see later) the usual cancellations should be made:-

$m = 7$

$r =$	0	1	2	3
T_r	1	-3	5	-1
ΔT_r		1	-5	2
$\Delta^2 T_r$			2	-2
$\Delta^3 T_r$				1
ΣT_r^2	7	28	84	6

$n = 6$

$s =$	0	1	2	3
T_s	1	-5	5	-5
ΔT_s		2	-6	12
$\Delta^2 T_s$			3	-15
$\Delta^3 T_s$				10
ΣT_s^2	6	70	84	180

from which the corresponding bivariate table may be constructed:-

(r,s)	(0,0)	(0,1)	(0,2)	(0,3)	(1,0)	(1,1)	(1,2)	(1,3)	(2,0)	(2,1)	(2,2)	(2,3)	(3,0)	(3,1)	(3,2)	(3,3)
$T_{r,s}$	1	-5	5	-5	-3	15	-15	15	5	-25	25	-25	-1	5	-5	5
$\Delta_x T_{r,s}$					1	-5	5	-5	-5	25	-25	25	2	-10	10	-10
$\Delta_y T_{r,s}$		2	-6	12		-6	18	-36		10	-30	60		-2	6	-12
$\Delta_x^2 T_{r,s}$									2	-10	10	-10	-2	10	-10	10
$\Delta_x\Delta_y T_{r,s}$						2	-6	12		-10	30	-60		4	-12	24
$\Delta_y^2 T_{r,s}$			3	-15			-9	45			15	-75			-3	15
$\Delta_x^3 T_{r,s}$													1	-5	5	-5
$\Delta_x^2\Delta_y T_{r,s}$										4	-12	24		-4	12	-24
$\Delta_x\Delta_y^2 T_{r,s}$							3	-15			-15	75			6	-30
$\Delta_y^3 T_{r,s}$				10				-30				50				-10
$\Delta_x^3\Delta_y T_{r,s}$														2	6	12
$\Delta_x^2\Delta_y^2 T_{r,s}$											6	-30			-6	30
$\Delta_x\Delta_y^3 T_{r,s}$								10				-50				20
$\Delta_x^3\Delta_y^2 T_{r,s}$															3	-15
$\Delta_x^2\Delta_y^3 T_{r,s}$												20				-20
$\Delta_x^3\Delta_y^3 T_{r,s}$																10
$\sum\sum T_{r,s}^2$	42	490	588	1260	168	1960	2352	5040	504	5880	7056	15120	36	420	504	1080

The values of $\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} T^2_{r,s}(x,y)$ in the above table are found as follows. In virtue of the orthogonal relations, we may replace $T_{r,s}(x,y)$ by its term of highest degree and write $\sum_x\sum_y T^2_{r,s}(x,y) = \sum_x\sum_y (2r)_{(r)}(2s)_{(s)}\, x_{(r)}\, y_{(s)}\, T_{r,s}(x,y)$. Summing this product by parts, using the fact that $\sum\sum T_{r,s}(x,y) = \sum^2\sum T_{r,s}(x,y)$ $= \sum\sum^2 T_{r,s}(x,y) = \ldots\ldots = \sum^{r+1}\sum^{s-1} T_{r,s}(x,y) = \sum^r\sum^s T_{r,s}(x,y)$ $= \sum^{r+1}\sum^{s-1} T_{r,s}(x,y) = 0$, we obtain $\sum_x\sum_y T_{r,s}(x,y)$ $= (-1)^{r+s}\frac{\underline{|2r}}{(\underline{|r})^2}\frac{\underline{|2s}}{(\underline{|s})^2}\sum_x^{r+1}\sum_y^{s+1} T_{r,s}(x,y)$
$=(-1)^{r+s}\frac{\underline{|2r}\,\underline{|2s}}{(\underline{|r}\,\underline{|s})^2}\sum_x\sum_y x_{(r)}(x-m)_{(r)}\,y_{(s)}(y-n)_{(s)}$ (since $\sum_x^{r+1}\sum_y^{s+1}\Delta_x^r\Delta_y^s = \sum_x\sum_y$).

Applying now r summations by parts with respect to x alone according to the formula $\sum_{x=0}^{m-1} v_x\,\Delta_x u_x = u_x v_x\Big]_0^m - \sum_0^{m-1} u_{x+1}\,\Delta_x v_x$, and noting that at each step one or other of the factors in the partial integrate vanishes at $x=0$ or $x=m$, we get

$$\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} T^2_{r,s}(x,y) = (-1)^{r+s}\frac{\underline{|2r}\,\underline{|2s}}{(\underline{|r}\,\underline{|s})^2}\sum_{y=0}^{n-1}\left[(-1)^r\frac{m(m^2-1^2)(m^2-2^2)\ldots(m^2-r^2)}{\underline{|2r+1}}\,y_{(s)}(y-n)_{(s)}\right].$$

A further similar application of s summations by parts with respect to y alone brings us to the result

$$\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} T^2_{r,s}(x,y) = \frac{m(m^2-1^2)(m^2-2^2)\ldots(m^2-r^2)}{(2r+1)(\underline{|r})^2}\,\frac{n(n^2-1^2)(n^2-2^2)\ldots(n^2-s^2)}{(2s+1)(\underline{|s})^2}$$

which is seen to be the product $\sum_{x=0}^{m-1} T^2_r(x)\,.\,\sum_{x=0}^{n-1} T^2_r(x)$, so that again the univariate tables may be used, this time to obtain the values of $\sum\sum T^2_{r,s}(x,y)$ in the table above. The usual cancellations of the univariate tables may therefore be made, and, in fact are essential if the $a_{r,s}$ are to be the same as the effects and interactions as already defined.

The $(r,s)^{th}$ reduced bivariate factorial moment of the data is $\frac{m_{(r,s)}}{\underline{|r}\,\underline{|s}} = \sum_x\sum_y x_{(r)}\, y_{(s)}\, u_{x,y}$. For evaluating $a_{r,s}$ we have $a_{r,s}\sum_x\sum_y T^2_{r,s}(x,y) = \sum_x\sum_y T_{r,s}(x,y).u_{x,y}$

But $T_{r,s}(x,y) = T_{r,s}(o,o) + x\Delta_x T_{r,s}(o,o) + y\Delta_y T_{r,s}(o,o)$
$+ x_{(2)}\Delta^2_x T_{r,s}(o,o) + xy\Delta_x\Delta_y T_{r,s}(o,o) + y_{(2)}\Delta^2_y T_{r,s}(o,o) + \ldots$

$\therefore a_{r,s}\sum_x\sum_y T^2_{r,s}(x,y) = T_{r,s}(o,o)\sum\sum u_{x,y} + \Delta_x T_{r,s}(o,o)\sum\sum x u_{x,y}$
$+ \Delta_y T_{r,s}(o,o)\sum\sum y u_{x,y} + \Delta^2_x T_{r,s}(o,o)\sum\sum x_{(2)} u_{x,y} + \ldots.$
$= T_{r,s}(o,o)\, m_{(o,o)} + \Delta_x T_{r,s}(o,o)\, m_{(1,o)} + \Delta_y T_{r,s}(o,o)\, m_{(o,1)}$
$+ \Delta^2_x T_{r,s}(o,o)\frac{m_{(2,o)}}{\underline{|2}} + \ldots..$

Hence $a_{r,s}$ is found by combining the reduced factorial moments of $u_{x,y}$ with the appropriate entries in the column (r,s) of the table of terminal values and differences for the special values of m and n, and then dividing by $\sum\sum T^2_{r,s}$. The reduced bivariate factorial moments are obtained by repeated summation on the values of y for each value of x and combining the results with a table of values of $x_{(r)}$ for $x = o,1,2,\ldots m-1$ and $r = o, 1,2,\ldots.$.

The terminal graduated values (z) and differences are

$z_{oo} = a_{o,o} + a_{1,o}\, T_{1,o}(o,o) + a_{o,1}\, T_{o,1}(o,o) + a_{2,o}\, T_{2,o}(o,o)$
$+ a_{1,1}\, T_{1,1}(o,o) + a_{o,2}\, T_{o,2}(o,o) + a_{3,o}T_{3,o}(o,o) + \ldots.$

$\Delta_x z_{o,o} = a_{1,o}\Delta_x T_{1,o}(o,o) + a_{2,o}\Delta_x T_{2,o}(o,o) + a_{1,1}\Delta_x T_{1,1}(o,o)$
$+ a_{3,o}\Delta_x T_{3,o}(o,o) + \ldots..$

$$\Delta_y z_{0,0} = a_{0,1}\Delta_y T_{0,1}(0,0) + a_{1,1}\Delta_y T_{1,1}(0,0) + a_{0,2}\Delta_y T_{0,2}(0,0) + \ldots$$
$$\Delta_x^2 z_{0,0} = a_{2,0}\Delta_x^2 T_{2,0}(0,0) + a_{3,0}\Delta_x^2 T_{3,0}(0,0) + \ldots\ldots$$
$$\Delta_x\Delta_y z_{0,0} = a_{1,1}\Delta_x\Delta_y T_{1,1}(0,0) + \ldots\ldots$$
$$\Delta_y^2 z_{0,0} = a_{0,2}\Delta_y^2 T_{0,2}(0,0) + \ldots$$

and so on, so that the table can be used a second time, using the entries in rows as multipliers in order to find the terminal graduated values and differences, from which all the graduated values may be found.

A similar check to that in the univariate case could be made by calculating the remote terminal values and differences, $z_{m-1,n-1}$, $\Delta_x z_{m-2,n-1}$, $\Delta_y z_{m-1,n-2}$, etc. Since $T_{r,s}(x,y) = T_r(x).T_s(y) = (-1)^r\, T_r(m-x-1).(-1)^s\, T_s(n-y-1)$ [Aitken(2)] which $= (-1)^{r+s}\, T_{r,s}(m-x-1,n-y-1)$, we have that

$$z_{m-1,n-1} = a_{0,0} - a_{1,0}T_{1,0}(0,0) - a_{0,1}T_{0,1}(0,0) + a_{2,0}T_{2,0}(0,0) + \ldots\ldots$$

Also, $\Delta_x^p\Delta_y^q T_{r,s}(x,y) = \Delta_x^p T_r(x).\Delta_y^q T_s(y)$

$$= (-1)^{p+r}\,\Delta_x^p T_r(m-x-p-1).\,(-1)^{q+s}\,\Delta_y^q T_s(n-y-q-1)$$
$$= (-1)^{p+q+r+s}\,\Delta_x^p\Delta_y^q T_{r,s}(m-x-p-1,n-y-q-1),$$

so that
$$\Delta_x z_{m-2,n-1} = -a_{1,0}\Delta_x T_{1,0}(m-2,n-1) + a_{2,0}\Delta_x T_{2,0}(m-2,n-1)$$
$$+ a_{1,1}\Delta_x T_{1,1}(m-2,n-1) - a_{3,0}\Delta_x T_{3,0}(m-2,n-1) - \ldots$$
$$= a_{1,0}\Delta_x T_{1,0}(0,0) - a_{2,0}\Delta_x T_{2,0}(0,0) - a_{1,1}\Delta_x T_{1,1}(0,0) + a_{3,0}\Delta_x T_{3,0}(0,0) + \ldots$$

and so on for the other differences. This shows that by reading all entries in the bivariate tables of orthogonal polynomials with positive sign, we could calculate $z_{m-1,n-1}$, $\Delta_x z_{m-2,n-1}$, $\Delta_y z_{m-1,n-2}$, etc. However, since each $z_{i,j}$ is in any case calculated independently, such a check is unnecessary.

Example of the Method. Graduation of the following bivariate data by means of a polynomial of degree (3,3)

x =	0	1	2	3	4	5	6	Totals.
y = 0	9	13	11	7	5	3	3	51
1	10	10	17	18	16	5	2	78
2	14	20	27	26	20	13	14	134
3	10	10	24	32	24	30	13	143
4	5	13	18	16	24	29	22	127
5	3	6	19	12	20	22	19	101
Totals	51	72	116	111	109	102	73	634

Calculation of Reduced Bivariate Factorial Moments:-

x = 0	Σ	Σ^2	Σ^3	Σ^4	x = 1	Σ	Σ^2	Σ^3	Σ^4
9	51				13	72			
10	42	103			10	59	162		
14	32	61	104		20	49	103	188	
10	18	29	43	60	10	29	54	85	122
5	8	11	14	17	13	19	25	31	37
3	3	3	3	3	6	6	6	6	6

x = 2.	Σ	Σ^2	Σ^3	Σ^4
11	116			
17	105	310		
27	88	205	397	
24	61	117	192	286
18	37	56	75	94
19	19	19	19	19

x = 3	Σ	Σ^2	Σ^3	Σ^4
7	111			
18	104	290		
26	86	186	338	
32	60	100	152	216
16	28	40	52	64
12	12	12	12	12

x = 4.	Σ	Σ^2	Σ^3	Σ^4
5	109			
16	104	324		
20	88	220	436	
24	68	132	216	320
24	44	64	84	104
20	20	20	20	20

x = 5	Σ	Σ^2	Σ^3	Σ^4
3	102			
5	99	347		
13	94	248	497	
30	81	154	249	366
29	51	73	95	117
22	22	22	22	22

x = 6.	Σ	Σ^2	Σ^3	Σ^4
3	73			
2	70	252		
14	68	182	375	
13	54	114	193	291
22	41	60	79	98
19	19	19	19	19

These summations, may be combined into the following table:-

$x =$	0	1	2	3	4	5	6
Σ	51	72	116	111	109	102	73
Σ^2	103	162	310	290	324	347	252
Σ^3	104	188	397	338	436	497	375
Σ^4	60	122	286	216	320	366	291

Values of $x_{(r)}$ are as follows:-

$x =$	0	1	2	3	4	5	6
$r = 0$	1	1	1	1	1	1	1
1		1	2	3	4	5	6
2			1	3	6	10	15
3				1	4	10	20

Suitable combination of the above two tables gives the reduced bivariate factorial moments thus:-

$$\frac{m_{(0,0)}}{\lfloor 0 \, \lfloor 0} = 634, \quad \frac{m_{(1,0)}}{\lfloor 1 \, \lfloor 0} = 2021, \quad \frac{m_{(2,0)}}{\lfloor 2 \, \lfloor 0} = 3218, \quad \frac{m_{(3,0)}}{\lfloor 3 \, \lfloor 0} = 3027$$

$$\frac{m_{(0,1)}}{\lfloor 0 \, \lfloor 1} = 1788, \quad \frac{m_{(1,1)}}{\lfloor 1 \, \lfloor 1} = 6195, \quad \frac{m_{(2,1)}}{\lfloor 2 \, \lfloor 1} = 10374, \quad \frac{m_{(3,1)}}{\lfloor 3 \, \lfloor 1} = 10096$$

$$\frac{m_{(0,2)}}{\lfloor 0 \, \lfloor 2} = 2335, \quad \frac{m_{(1,2)}}{\lfloor 1 \, \lfloor 2} = 8475, \quad \frac{m_{(2,2)}}{\lfloor 2 \, \lfloor 2} = 14622, \quad \frac{m_{(3,2)}}{\lfloor 3 \, \lfloor 2} = 14552,$$

$$\frac{m_{(0,3)}}{\lfloor 0 \, \lfloor 3} = 1661, \quad \frac{m_{(1,3)}}{\lfloor 1 \, \lfloor 3} = 6198, \quad \frac{m_{(2,3)}}{\lfloor 2 \, \lfloor 3} = 10879, \quad \frac{m_{(3,3)}}{\lfloor 3 \, \lfloor 3} = 10976$$

The computation is arranged round the table of orthogonal polynomial values as on the next page.

$a_{r,s}$ =	15.095	0.829	-0.940	-0.100	0.708	0.544	0.008	-0.104	-0.990	-0.013	0.122	-0.025	-0.028	-0.226	0.127	0.083
$a_{r,s}\sum\sum T^2_{r,s}$ =	634	406	-553	-129	119	1067	19	-523	-499	-79	859	-379	-1	-95	64	90
r,s =	0,0	0,1	0,2	0,3	1,0	1,1	1,2	1,3	2,0	2,1	2,2	2,3	3,0	3,1	3,2	3,3
Moments																
634	1	-5	5	-5	-3	15	-15	15	5	-25	25	-25	-1	5	-5	5
2021					1	-5	5	-5	-5	25	-25	25	2	-10	10	-10
1788		2	-6	12		-6	18	-36		10	-30	60		-2	6	-12
3218									2	-10	10	-10	-2	10	-10	10
6195						2	-6	12		-10	30	-60		4	-12	24
2335			3	-15			-9	45			15	-75			-3	15
3027													1	-5	5	-5
10374										4	-12	24		-4	12	-24
8475							3	-15			-15	75			6	-30
1661				10				-30				50				-10
10096														2	-6	12
14622											6	-30			-6	30
6198								10				-50				20
14552															3	-15
10879												20				-20
10976																10
$\sum\sum T^2_{r,s}$ =	42	490	588	1260	168	1960	2352	5040	504	5880	7056	15120	36	420	504	1080

$z_{0,0} = 8.834$
$\Delta_x z_{0,0} = 2.142$
$\Delta_y z_{0,0} = 1.650$
$\Delta_x^2 z_{0,0} = -3.024$
$\Delta_x\Delta_y z_{0,0} = 4.646$
$\Delta_y^2 z_{0,0} = -1.503$
$\Delta_x^3 z_{0,0} = 1.322$
$\Delta_x^2\Delta_y z_{0,0} = -1.680$
$\Delta_x\Delta_y^2 z_{0,0} = -3.849$
$\Delta_y^3 z_{0,0} = 0.040$
$\Delta_x^3\Delta_y z_{0,0} = -0.218$
$\Delta_x^2\Delta_y^2 z_{0,0} = 3.210$
$\Delta_x\Delta_y^3 z_{0,0} = 1.870$
$\Delta_x^3\Delta_y^2 z_{0,0} = -0.864$
$\Delta_x^2\Delta_y^3 z_{0,0} = -2.164$
$\Delta_x^3\Delta_y^3 z_{0,0} = 0.830$

From the calculated values of $z_{0,0}$, $\Delta_x z_{0,0}$, $\Delta_y z_{0,0}$, $\Delta_x^2 z_{0,0}$, ... etc. we may obtain the following graduated values:-

x =	0	1	2	3	4	5	6	Totals.
y = 0	8.834	10.976	10.094	7.510	4.546	2.524	2.766	47.3
1	10.484	17.272	19.356	17.840	13.828	8.424	2.732	89.9
2	10.631	18.216	22.627	23.886	22.015	17.036	8.971	123.3
3	9.315	15.718	21.527	25.648	26.987	24.450	16.943	140.6
4	6.576	11.688	17.676	23.126	26.624	26.756	22.108	134.6
5	2.454	8.036	12.694	16.320	18.806	20.044	19.926	98.3
Totals	48.3	81.9	104.0	114.3	112.8	99.2	93.5	634.0

A check on $z_{6,5}$, derived alternatively as described on P. 159 gave 19.926, agreeing exactly with the result of the above table. The grand total of the graduated values is the same as that of the ungraduated values.

The residual sum of squares may be calculated from the following table of $z_{ij} - u_{ij}$, in which signs are ignored.

x =	0	1	2	3	4	5	6
y = 0	0.166	2.024	0.906	0.510	0.454	0.476	0.234
1	0.484	7.272	2.356	0.160	2.172	3.424	0.732
2	3.369	1.784	4.373	2.114	2.015	4.036	5.029
3	0.685	5.718	2.473	6.352	2.987	5.550	3.943
4	1.576	1.312	0.324	7.126	2.624	2.244	0.108
5	0.546	2.036	6.306	4.320	1.194	1.956	0.926

$$\sum_i \sum_j (z_{ij} - u_{ij})^2 = 436.0$$

The residual sum of squares may also be computed as follows and each $a_{r,s}$ tested for significance.

Degree of curve fitted $=(r,s)$	$a_{r,s}$ × Numerator of (14,1)	Residuals from 12,242, the total S.S.	D.F. $=42-(r+1)(s+1)$	Mean Squares	F
0,0	9570.23	2671.77	41	65.17	-
0,1	336.57	2335.20	40	58.38	5.8*
0,2	519.82	1815.38	39	46.55	11.2**
0,3	12.90	1802.48	38	47.43	0.3
1,0	84.25	2587.52	40	64.69	1.3
1,1	580.45	1670.50	38	43.96	13.2**
1,2	0.15	1150.53	36	31.96	0.0
1,3	54.39	1083.24	34	31.86	1.7
2,0	494.01	1841.19	39	47.21	10.5**
2,1	1.03	1175.46	36	32.65	0.3
2,2	104.80	550.69	33	16.69	6.3*
2,3	9.48	473.92	30	15.80	0.6
3,0	0.03	1841.16	38	48.45	0.0
3,1	21.47	1153.96	34	33.94	0.6
3,2	8.13	521.06	30	17.37	0.5
3,3	7.47	436.82	26	16.80	0.4

The above table assumes that, for example, the fitting of $a_{1,0}$ and $a_{0,1}$ involves also the fitting of $a_{1,1}$, so that it would not be correct to test $a_{1,0}$ against the residual left by subtraction of the sums of squares due to $a_{0,1}$ and $a_{1,0}$ using 27 degrees of freedom. As expected, since the data are approximately a normal bivariate sample, the constants which prove significant are chiefly those associated with second degree terms. The residual sum of squares agrees reasonably well with that found otherwise.

The above tests of significance are, of course, only approximate and in default of more exact knowledge about the true variance of the variates. In actual experimental work the tests will be made against the error mean square. As regards degrees of freedom, if the curve fitted is of degree (r,s), there are $(r+1)(s+1)$ constants and this is the rank of the matrix T. The vector of residuals is $u-Ta$ or $\left[I-T(T'T)^{-1}T'\right]u$, where the matrix $I-T(T'T)^{-1}T'$ is symmetric, idempotent, and hence of rank $mn-(r+1)(s+1)$. The sum of squared residuals must therefore have $mn-(r+1)(s+1)$ degrees of freedom, and its mean square is on a null hypothesis an estimate of variance which is independent of the estimates derived from the linear combinations $\{a_{r,s}\} = (T'T)^{-1}T'u$.

For the sake of comparison we will now graduate the same set of data by the methods of the previous section.

For the factor x we have the interaction transformation

$$\begin{bmatrix} I \\ X' \\ X'' \\ X''' \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{7}}(1 & 1 & 1 & 1 & 1 & 1 & 1) \\ \frac{1}{\sqrt{28}}(-3 & -2 & -1 & 0 & 1 & 2 & 3) \\ \frac{1}{\sqrt{84}}(5 & 0 & -3 & -4 & -3 & 0 & 5) \\ \frac{1}{\sqrt{6}}(-1 & 1 & 1 & 0 & -1 & -1 & 1) \end{bmatrix} u_{io}, \quad \ldots.(14,4)$$

where the values of the orthogonal polynomials for the linear and quadratic effects are obtainable as in the last section, or may be found, together with the cubic values, in Fisher and Yates (14). For the factor y the transformation is

$$\begin{bmatrix} I \\ Y' \\ Y'' \\ Y''' \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{6}}(1 & 1 & 1 & 1 & 1 & 1) \\ \frac{1}{\sqrt{70}}(-5 & -3 & -1 & 1 & 3 & 5) \\ \frac{1}{\sqrt{84}}(5 & -1 & -4 & -4 & -1 & 5) \\ \frac{1}{\sqrt{180}}(-5 & 7 & 4 & -4 & -7 & 5) \end{bmatrix} u_{oj} \ldots(14,5)$$

Combining (14,4) and (14,5) by direct multiplication and reverting from the normalised definition to the response per plot definition, we have the transformation of Table (14,1) where u is now the vector $\{9 \; 13 \; 11 \; 7 \; 5 \; 3 \; 3 \,\vdots\, 10 \; 10 \; 17 \; 18 \; 16 \; 5 \; 2 \,\vdots\, 14 \; 20 \; 27 \; 26 \; 20 \; 13 \; 14 \,\vdots\, 10 \; 10 \; 24 \; 32 \; 24 \; 30 \; 13 \,\vdots\, 5 \; 13 \; 18 \; 16 \; 24 \; 29 \; 22 \,\vdots\, 3 \; 6 \; 19 \; 12 \; 22 \; 19\}$

The values of the a_{ij} so obtained check exactly with the previous results, as do the graduated values. The defect of this method is that the transforming matrix is likely to be unwieldy.

a_{00} =	$\frac{1}{42}$ [	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
a_{10} =	$\frac{1}{168}$ [	-3	-2	-1	0	1	2	3	-3	-2	-1	0	1	2	3	-3	-2	-1	0	1	2	3
a_{20} =	$\frac{1}{504}$ [	5	0	-3	-4	-3	0	5	5	0	-3	-4	-3	0	5	5	0	-3	-4	-3	0	5
a_{30} =	$\frac{1}{36}$ [	-1	1	1	0	-1	-1	1	-1	1	1	0	-1	-1	1	-1	1	1	0	-1	-1	1
a_{01} =	$\frac{1}{490}$ [	-5	-5	-5	-5	-5	-5	-5	-3	-3	-3	-3	-3	-3	-3	-1	-1	-1	-1	-1	-1	-1
a_{11} =	$\frac{1}{1960}$ [	15	10	5	0	-5	-10	-15	9	6	3	0	-3	-6	-9	3	2	1	0	-1	-2	-3
a_{21} =	$\frac{1}{5880}$ [	-25	0	15	20	15	0	-25	-15	0	9	12	9	0	-15	-5	0	3	4	3	0	-5
a_{31} =	$\frac{1}{420}$ [	5	-5	-5	0	5	5	-5	3	-3	-3	0	3	3	-3	1	-1	-1	0	1	1	-1
a_{02} =	$\frac{1}{588}$ [	5	5	5	5	5	5	5	-1	-1	-1	-1	-1	-1	-1	-4	-4	-4	-4	-4	-4	-4
a_{12} =	$\frac{1}{2352}$ [	-15	-10	-5	0	5	10	15	3	2	1	0	-1	-2	-3	12	8	4	0	-4	-8	-12
a_{22} =	$\frac{1}{7056}$ [	25	0	-15	-20	-15	0	25	-5	0	3	4	3	0	-5	-20	0	12	16	12	0	-20
a_{32} =	$\frac{1}{504}$ [	-5	5	5	0	-5	-5	5	1	-1	-1	0	1	1	-1	4	-4	-4	0	4	4	-4
a_{03} =	$\frac{1}{1260}$ [	-5	-5	-5	-5	-5	-5	-5	7	7	7	7	7	7	7	4	4	4	4	4	4	4
a_{13} =	$\frac{1}{5040}$ [	15	10	5	0	-5	-10	-15	-21	-14	-7	0	7	14	21	-12	-8	-4	0	4	8	12
a_{23} =	$\frac{1}{15120}$ [	-25	0	15	20	15	0	-25	35	0	-21	-28	-21	0	35	20	0	-12	-16	-12	0	20
a_{33} =	$\frac{1}{1080}$ [	5	-5	-5	0	5	5	-5	-7	7	7	0	-7	-7	7	-4	4	4	0	-4	-4	4

a_{00}	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1]	= u
a_{10}	-3	-2	-1	0	1	2	3	-3	-2	-1	0	1	2	3	-3	-2	-1	0	1	2	3]	= u
a_{20}	5	0	-3	-4	-3	0	5	5	0	-3	-4	-3	0	5	5	0	-3	-4	-3	0	5]	= u
a_{30}	-1	1	1	0	-1	-1	1	-1	1	1	0	-1	-1	1	-1	1	1	0	-1	-1	1]	= u
a_{01}	1	1	1	1	1	1	1	3	3	3	3	3	3	3	5	5	5	5	5	5	5]	= u
a_{11}	-3	-2	-1	0	1	2	3	-9	-6	-3	0	3	6	9	-15	-10	-5	0	5	10	15]	= u
a_{21}	5	0	-3	-4	-3	0	5	15	0	-9	-12	-9	0	15	25	0	-15	-20	-15	0	25]	= u
a_{31}	-1	1	1	0	-1	-1	1	-3	3	3	0	-3	-3	3	-5	5	5	0	-5	-5	5]	= u
a_{02}	-4	-4	-4	-4	-4	-4	-4	-1	-1	-1	-1	-1	-1	-1	5	5	5	5	5	5	5]	= u
a_{12}	12	8	4	0	-4	-8	-12	3	2	1	0	-1	-2	-3	-15	-10	-5	0	5	10	15]	= u
a_{22}	-20	0	12	16	12	0	-20	-5	0	3	4	3	0	-5	25	0	-15	-20	-15	0	25]	= u
a_{32}	4	-4	-4	0	4	4	-4	1	-1	-1	0	1	1	-1	-5	5	5	0	-5	-5	5]	= u
a_{03}	-4	-4	-4	-4	-4	-4	-4	-7	-7	-7	-7	-7	-7	-7	5	5	5	5	5	5	5]	= u
a_{13}	12	8	4	0	-4	-8	-12	21	14	7	0	-7	-14	-21	-15	-10	-5	0	5	10	15]	= u
a_{23}	-20	0	12	16	12	0	-20	-35	0	21	28	21	0	-35	25	0	-15	-20	-15	0	25]	= u
a_{33}	4	-4	-4	0	4	4	-4	7	-7	-7	0	7	7	-7	-5	5	5	0	-5	-5	5]	= u

(Table 14,1)

Calculation of the Optimal Levels of Factors in Combination.

These may be obvious by inspection, but in any case approximate calculated values are not difficult to obtain. As an example let us suppose that the data of P.160 represent the yields of an experiment with two factors x and y at seven and six levels respectively. From the graduated values of P.164 we may select a square of nine of them within which ranges the maximum value must evidently lie. The chosen values are

23.886	22.015	17.036
25.648	26.987	24.450
23.126	26.624	26.756,

and by a change of origin we denote them to be

$z_{-1,-1}$	$z_{0,-1}$	$z_{1,-1}$
$z_{-1,0}$	$z_{0,0}$	$z_{1,0}$
$z_{-1,1}$	$z_{0,1}$	$z_{1,1}$

By the advancing difference formula

$$z_{x,y} = (1 + x\Delta_x + y\Delta_y + x_{(2)}\Delta_x^2 + xy\Delta_x\Delta_y + y_{(2)}\Delta_y^2 + x_{(3)}\Delta_x^3 + x_{(2)}y\Delta_x^2\Delta_y + xy_{(2)}\Delta_x\Delta_y^2 + y_{(3)}\Delta_y^3 + \ldots)\, z_{0,0}$$

For a maximum value of $z_{x,y}$ we must have

$$\frac{\partial z_{x,y}}{\partial x} = \left(\Delta_x + \frac{2x-1}{2}\Delta_x^2 + y\Delta_x\Delta_y + \frac{3x^2-6x+2}{\lfloor 3}\Delta_x^3 + \frac{2x-1}{2}y\Delta_x^2\Delta_y + y_{(2)}\Delta_x\Delta_y^2 + \cdots\right) z_{0,0} = 0$$

$$\frac{\partial z_{x,y}}{\partial y} = \left(\Delta_y + \frac{2y-1}{2}\Delta_y^2 + x\Delta_x\Delta_y + \frac{3y^2-6y+2}{\lfloor 3}\Delta_y^3 + \frac{2y-1}{2}x\Delta_x\Delta_y^2 + x_{(2)}\Delta_x^2\Delta_y + \cdots\right) z_{0,0} = 0$$

Approximate equations for x and y small are therefore

$$\left.\begin{aligned} &(\Delta_x - \tfrac{1}{2}\Delta_x^2 + \tfrac{1}{3}\Delta_x^3)z_{0,0} + x(\Delta_x^2 - \Delta_x^3)z_{0,0} + y(\Delta_x\Delta_y - \tfrac{1}{2}\Delta_x^2\Delta_y - \tfrac{1}{2}\Delta_x\Delta_y^2)z_{0,0} = 0 \\ \text{and } &(\Delta_y - \tfrac{1}{2}\Delta_y^2 + \tfrac{1}{3}\Delta_y^3)z_{0,0} + x(\Delta_x\Delta_y - \tfrac{1}{2}\Delta_x\Delta_y^2 - \tfrac{1}{2}\Delta_x^2\Delta_y)z_{0,0} + y(\Delta_y^2 - \Delta_y^3)z_{0,0} = 0 \end{aligned}\right\} \quad (14,6)$$

It is necessary to express these equations in terms of differences which involve values of z evenly ranged about the origin. This is achieved by means of the substitutions

$$\Delta_x^2 z_{0,0} = \Delta_x^2 z_{-1,0} + \Delta_x^3 z_{-1,0}$$
$$\Delta_x^3 z_{0,0} = \Delta_x^3 z_{-1,0} + \Delta_x^4 z_{-1,0}$$
$$\Delta_x^4 z_{-1,0} = \Delta_x^4 z_{-2,0} + \Delta_x^5 z_{-2,0}$$
$$\Delta_x^2 \Delta_y z_{0,0} = \Delta_x^2 \Delta_y z_{-1,0} + \Delta_x^3 \Delta_y z_{-1,0}$$

and similar expressions for differences of y. The equations (14,6) then become

$$\Delta_x z_{0,0} - \tfrac{1}{2}\Delta_x^2 z_{-1,0} - \tfrac{1}{2}\Delta_x^3 z_{-1,0} + \tfrac{1}{3}\Delta_x^3 z_{-1,0} + \tfrac{1}{3}\Delta_x^4 z_{-2,0} + \tfrac{1}{3}\Delta_x^5 z_{-2,0}$$
$$+ x(\ \Delta_x^2 z_{-1,0} - \Delta_x^4 z_{-2,0} - \Delta_x^5 z_{-3,0})$$
$$+ y(\ \Delta_x\Delta_y z_{0,0} - \tfrac{1}{2}\Delta_x^2\Delta_y z_{-1,0} - \tfrac{1}{2}\Delta_x^3\Delta_y z_{-1,0} - \tfrac{1}{2}\Delta_x\Delta_y^2 z_{0,-1} - \tfrac{1}{2}\Delta_x\Delta_y^3 z_{0,-1}\ ) = 0$$

and $\Delta_y z_{0,0} - \tfrac{1}{2}\Delta_y^2 z_{0,-1} - \tfrac{1}{2}\Delta_y^3 z_{0,-1} + \tfrac{1}{3}\Delta_y^3 z_{0,-1} + \tfrac{1}{3}\Delta_y^4 z_{0,-2} + \tfrac{1}{3}\Delta_y^3 z_{0,-2}$

$$+ x(\ \Delta_x\Delta_y z_{0,0} - \tfrac{1}{2}\Delta_x\Delta_y^2 z_{0,-1} - \tfrac{1}{2}\Delta_x\Delta_y^3 z_{0,-1} - \tfrac{1}{2}\Delta_x^2\Delta_y z_{-1,0} - \tfrac{1}{2}\Delta_x^3\Delta_y z_{-1,0}\ )$$
$$+ y(\ \Delta_y^2 z_{0,-1} - \Delta_y^4 z_{0,-2} - \Delta_y^5 z_{0,-2}\ ) = 0$$

In the present example, where the fitted polynomial is predominantly of the second degree, we may ignore differences of x or y of the third and higher orders, whereupon, expressed in terms of the z's, the equations reduce to

$$\tfrac{1}{2}(z_{1,0} - z_{-1,0}) + x(z_{1,0} - 2z_{0,0} + z_{-1,0}) + y(\ \tfrac{1}{2}z_{1,0} + \tfrac{1}{2}z_{0,1} - z_{0,0} - \tfrac{1}{2}z_{-1,1} - \tfrac{1}{2}z_{1,-1}$$
$$+ \tfrac{1}{2}z_{-1,0} + \tfrac{1}{2}z_{0,-1}\ ) = 0$$

and $\tfrac{1}{2}(\ z_{0,1} - z_{0,-1}) + x(\ \tfrac{1}{2}z_{1,0} + \tfrac{1}{2}z_{0,1} - z_{0,0} - \tfrac{1}{2}z_{-1,1} - \tfrac{1}{2}z_{1,-1} + \tfrac{1}{2}z_{-1,0} + \tfrac{1}{2}z_{0,-1})$

$$+ y(z_{0,1} - 2z_{0,0} + z_{0,-1}) = 0$$

or $3.876x - 2.301y = -0.599$

and $2.301x - 5.355y = -2.304,$

whence $x = 0.14, \; y = 0.49.$

The method is easily extensible to data with more than two factors.

Section 15. Confounding

The principles of confounding have been discussed by Fisher (13), Yates (24,25,26), Barnard (4), and Finney (10). Fisher describes confounding as an artifice which "consists of increasing the number of blocks..... beyond the number of replications in the experiment so that each replication occupies two or more blocks; and, at the same time, arranging that the experimental contrasts between the different blocks within each replication shall be contrasts between unimportant interactions, the study of which the experimenter is willing to sacrifice, for the sake of increasing the precision of the remaining contrasts, in which he is specially interested". Yates' description is "a device whereby the necessity of including every combination of the treatments of a factorial design in each block (or row and column in a Latin square) is avoided.....The treatment combinations of each replication are divided into two or more groups (each group being assigned to a separate block) in such a way that the contrasts between the different groups represent high-order interactions, which.....are usually of less interest than the main effects and interactions between two factors only. Thus in any one replication the contrasts representing certain interactions are identified, or confounded, with the block differences".

Both these descriptions (they are hardly formal definitions) are more restricted than they need be. Two experimental contrasts are said to be confounded when they are identified with one another with respect to the unit plots making up the contrasts. For example, in a split-plot experiment the whole-plot treatments are confounded with whole-plots with respect to the sub-plots as unit, but may obviously be estimated from the whole-plot yields. Similarly, in a simple confounding experiment in which the treatment-combinations are divided into two blocks for each replication so that a particular contrast is totally confounded, the latter is confounded with blocks but is capable of estimation from comparisons of block-pairs. Nor is it necessary that one of the contrasts confounded should be some effect of soil heterogeneity, whether it be due to blocks, whole-plots, or rows and columns of a Latin square. In a half-replicate design, for example, every treatment contrast is confounded with some other treatment contrast, the one being called the "alias" of the other.

However, the term "confounding" is most often understood in the particular sense of an experimental device for controlling the heterogeneity of the population. Thus, in field experiments, if the number of treatment-combinations is large, the blocks of an ordinary randomised-block design become too big to exert an effective control over soil heterogeneity. Confounding is the device most commonly used to counteract this. Any treatment contrast $\ell' y$ (where $\Sigma \ell = 0$), corresponding to a single degree of freedom, may be confounded by allotting the treatment-combinations to different blocks according to whether they correspond to positive or negative elements of $\{\ell\}$. The only practical application of this is when the elements of $\{\ell\}$ (not normalised) are all ± 1, in which case, if each replication occupies two blocks, every other treatment contrast belonging to the same orthogonal set as $\ell' y$ will be unconfounded. This is evident since any other vector $\{m\}$ of the orthogonal set has its elements divided into two groups corresponding to the positive and negative elements of $\{\ell\}$, and the sums of elements within these groups, being equal must therefore each be zero. But this division into groups is the same division by which the treatment-combinations are allotted to the two different blocks. Hence in the linear function m'y, the coefficients of the components of yield due to each block variate sum to zero, and m'y is thus unaffected by block differences. The designing of confounded experiments is therefore seen to reduce in many cases to the search for suitable vectors with elements all ± 1. Suitability will usually mean that the treatment contrast to be confounded should be a high-order interaction, or at least an interaction which may be predicted (perhaps from previous experiments) to be negligible in comparison to the random variation. In some cases of partial confounding, i.e. a design where the contrasts confounded are not the same for each replication, even main effects may be confounded in order to secure a balanced design.

Confounding is especially simple for 2^n factorial designs, since every main effect and interaction is determined by a vector with all elements ± 1 (Section 12). Finney has summarised the rules governing the structure of such designs. In accordance with the notation of section 12, capital letters A,B,C,D,..... AB,AC,BC,..... ABC,ABD,..... etc. are used to denote main effects and interactions of factors a,b,c,d,....., while small letters are used for the treatment-combinations, e.g. abd is the combination of the higher levels ("presence") of a, b, d with the control levels ("absence") of the remaining factors. It was seen at the end of Section 10 (P.108) that if we regard AB as the "product" of A and B, then B is the product of A and AB. (Yates calls B the "generalised interaction" of A and AB). It is useful to define a similar symbolic product for the small letters, e.g. the product of "abc" and "bce" is defined to be "ae".

In a 2^n confounded arrangement the block containing the control treatment (1), representing the absence of all factors, is called the "principal block". This leads to the first rule governing the structure of such a design: that every treatment-combination in the principal block contains an even number of the letters occurring in any confounded interaction. This is seen to be a simple consequence of the symbolic representation of an interaction by the product $(a \pm 1)(b \pm 1)(c \pm 1)(d \pm 1)\ldots.$, the minus sign in the bracket associated with a particular factor being taken if that letter occurs in the interaction, e.g. ABC is $(a-1)(b-1)(c-1)(d+1)\ldots.$ The sign of (1) is clearly the same as that of any product of o,2,4,....letters from the brackets containing minus signs. When only a single contrast is to be confounded, this rule alone enables the principal block to be written down without actually working out the interaction matrix. Another property of the principal block, that the product of any two treatment-combinations which are members of the principal block is also a member, follows from the fact that the product will still have an even number of the letters of any confounded interaction,

Still greater control over soil heterogeneity will be given if each replication is divided into 4,8,16,..... equal blocks, when the 3,7,15,.... additional degrees of freedom, respectively, (per replication) now allotted for block differences will entail the confounding of 3,7,15,.....treatment contrasts. These degrees of freedom for blocks (within replications) may be subdivided into single degrees of freedom with all elements of the unnormalised subdividing vectors equal to ± 1 by means of the matrices $M^{\{2\}}$, $M^{\{3\}}$, $M^{\{4\}}$,....., where M is the matrix $\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$, and its direct square, direct cube, etc. are, ignoring normalising factors, the interaction matrices of Section 12. The rows of these matrices, M, $M^{\{2\}}$, $M^{\{3\}}$, $M^{\{4\}}$,...., being linearly independent, determine for designs with each replication divided into 2, 4,8,16,....blocks, respectively, all possible ways of forming block contrasts by means of an equal number of additions and subtractions. But it has been seen (Sections 10, 12) that any 2,3,4,... rows (other than the first) of $M^{\{2\}}, M^{\{3\}}, M^{\{4\}}, \ldots.$, respectively, generate all the others by means of multiplication of corresponding elements, i.e. the other rows are interactions of the given 2,3,4,... It follows that, if any set of contrasts is simultaneously confounded, all possible products ("generalised interactions") of these contrasts are also confounded. Thus, if the replication is divided into four blocks, only two treatment contrasts may be selected for confounding, for the third contrast confounded is automatically the product of the other two. Hence, one would not choose for confounding the interactions ABC and ABCD, for example, since this would mean that the main effect D would also be confounded, and in general this is not desirable. Keeping the restrictions of this paragraph in view, the statistician may easily derive the treatment-combinations for the principal block of his design by applying the two rules of the previous paragraph.

Once the principal block has been written down, another block of the design may be generated from it by multiplication by any treatment-combination not a member of the principal block. A repetition of this process will give all the blocks of the scheme. The reason for this is easy to see when we consider that, when each replication is divided into 2^p blocks, the contents of the blocks are determined by the 2^{n-p} treatment-combinations corresponding to each of the 2^p different permutations of positive and negative signs in corresponding elements of the p vectors representing the treatment contrasts selected for confounding. But any one such permutation of signs means that the 2^{n-p} treatment-combinations so determined have either an even or odd number of letters (depending on the particular sign in each vector) in common with the letters of each confounded interaction. In other words, their orthogonality (even number of letters in common) or non-orthogonality to each of the confounded interactions is fixed. Consequently, if in such a permutation the sign in one vector is changed, the new set of 2^{n-p} treatment-combinations now determined differs from the former in respect of orthogonality to the confounded interactions for only one interaction, and, as has been seen, this is brought about by multiplying by a treatment-combination which is non-orthogonal to the interaction concerned. That is, in the case of the principal block, such a multiplying treatment-combination is not a member of the principal block. But in changing one sign of a particular permutation we have obtained a new permutation which determines the contents of a second block, and so the rule is proved.

The contents of the blocks may also be obtained in the case of a 2^n experiment with 2^k blocks per replication from the symbolic relationship

$$\{\text{Column vector of confounded interactions}\} = M^{\{k\}}\{\text{Column vector of block totals}\}$$

Hence $[M^{\{k\}}]'\{\text{Vector of confounded interactions}\}$

$$= \{\text{Vector of block totals}\}$$

Substitute for the vector of confounded interactions the appropriate row vectors of the interaction matrix post-multiplied by the column vector of treatment-combinations

$$\{(1)\ \ a\ \ b\ \ ab\ \ c\ \ ac\ \ bc\ \ abc\ldots\ldots\text{etc.}\} = \{y\},$$

and we obtain a relationship giving the contents of the blocks. The matrix $M^{\{n\}}_{2^k}$, comprising the 2^k rows of the interaction matrix $M^{\{n\}}$, corresponding to the 2^k-1 confounded interactions and the row of unit elements, is of order $2^k \times 2^n$, and when premultiplied by $[M^{\{k\}}]'$ gives a product matrix of order $2^k \times 2^n$ with all elements either o or 1. Applying this product matrix to $\{y\}$, we obtain the contents of each block. The unit elements of $[M^{\{k\}}]'[M_{2^k}^{\{n\}}]$ correspond to particular permutations of signs of $M_{2^k}^{\{n\}}$, a fact noted in the previous paragraph.

The subject of confounding in designs with all factors at three levels and in designs with some factors at three levels and some at two levels has been exhaustively treated by Yates (26). With a factor at four levels (or qualities), use is made of the orthogonal subdivision into three degrees of freedom corresponding to the matrix $M^{\{2\}}$. The problem of confounding in this case thus reduces to the case of a 2^n design, provided the other factors occur at two or four levels. In such a design, if the factors have four equally-spaced levels (not qualities), use might be made of the fact that the quadratic effect of a four-level factor is represented by a vector with all elements ± 1, and therefore presents opportunities for confounding. Confounded designs have been worked out for experiments with all factors at five levels, but such applications must be comparatively rare in practice.

Section 16.

Analysis of Covariance

Part of the observed variation of a variate y (the dependent variate) may be due to its regression on a number of concomitant variates $x_1, x_2, \ldots . x_k$, the latter being known as the independent variates, though they need not be, and in general are not, statistically independent of one another. By making allowance for this regression the precision of an experiment may be greatly increased. This is done by means of a process rather inadequately named the Analysis of Covariance.

In the first place, if we have a sum of squares resolved into a number of component sums of squares in accordance with some particular experimental design, thus:- $y'A'Ay = y'B'By + y'C'Cy + \ldots . + y'N'Ny$,
then not only does the same partitioning apply to the sums of squares of all the concomitant variates, but also to the sums of products (or covariances) of any two variates, dependent or independent. For example,

$$y'A'Ax_i = y'B'Bx + y'C'Cx + \ldots . + y'N'Nx$$

illustrates the partitioning of the sums of products of y and x_i. Hence, corresponding to each component set of the partitioned degrees of freedom, we have the sums of squares for all variates and the sums of products for each pair of variates, from which it is possible to obtain in the usual manner the various partial regression and correlation coefficients for each set, and the sums of squares for y corrected for regression. Unless it is desired to make the tests of significance of analysis of variance for any of the x-variates, no restriction is placed on their probability distributions.

Let $a'y$ be any deviation of a class-mean from the general mean or any residual of the type discussed in the theory of analysis of variance, and let Ay be the column vector representing the assembly of all similar mean-deviations (or residuals), whose mean will be zero. If we let $a'y = b_1 a'x_1 + b_2 a'x_2 + \ldots\ldots + b_k a'x_k$, we have a set of observational equations which may be written as $Ay = AXb$, where b is the vector $\{b_1\, b_2 \ldots . b_k\}$ and X is the matrix

$$\begin{bmatrix} x_{11} & x_{21} \cdots & x_{k1} \\ x_{12} & x_{22} \cdots & x_{k2} \\ x_{13} & x_{23} \cdots & x_{k3} \\ \vdots & \vdots & \vdots \end{bmatrix}$$

with the same number of rows as there are elements in the vectors $\{y\}$, $\{x_1\}$, $\{x_2\}$, ... etc., and x_{ij} represents the j^{th} element of the vector x_i. A may be a symmetric matrix as in the case of an error residual, but, if it is, it will be singular, being not of full rank. The normal equations are $X'A'Ay = X'A'AXb$, so that $b = (X'A'AX)^{-1}X'A'Ay$.

The vector of residuals is $\left[I - AX(X'A'AX)^{-1}X'A'\right]Ay$, and the sum of squared residuals is

$y'A'\left[I-AX(X'A'AX)^{-1}X'A'\right]Ay$. The latter may be alternatively expressed as $y'A'Ay(1-\hat{R}^2_{y,123...})$, where $\hat{R}_{y,123...}$ is the sample estimate of the multiple correlation coefficient, for, if $\{\hat{y}\}$ is a regression estimate of $\{y\}$,

$$\hat{R}_{y,123...} = \hat{y}'A'Ay.(\hat{y}A'A\hat{y})^{-\frac{1}{2}}.(y'A'Ay)^{-\frac{1}{2}}.$$

But $\hat{y}'A'A\hat{y} = b'X'A'AXb = b'X'A'Ay = \hat{y}'A'Ay$

$$\therefore \hat{R}^2_{y,123...} = \frac{\hat{y}'A'A\hat{y}}{y'A'Ay}$$, and the result follows.

On a hypotheses of uniformity the vectors $\{b\}$ obtained from each set of the partitioned degrees of freedom would be sample estimates of a certain population vector $\{\beta\}$. However, in an analysis of variance no such assumption is made concerning the different blocks, rows, or columns used to control soil heterogeneity and the treatments may or may not have a differential effect. The object of the covariance analysis is to discover how far the significance of treatments, as tested in the analysis of variance of the variate y, is attributable to the regression of y on the independent variates. Attention is therefore confined to only two sources of variation - treatments and error.

The regression and correlation coefficients obtained from the error line of the analysis provide a measure of the association between the random variation of y and the random variation of the x's. This is known as the "error regression", and from it may be found how much of the estimated variance of y is due to regression. The treatments regression, on the other hand, may be quite different from the error regression owing to the treatments having differential effects on the variates. Cochran (6) concluded that a test comprising a comparison of the residual mean squares of the treatments and error regressions would not be suitable. He also investigated the possibility of taking the error regression out of the treatments sum of squares and testing this residual sum of squares with the residual sum of squares for error. If the treatments and error sums of squares for y are $y'A'Ay$ and $y'B'By$*, with p and q degrees of freedom respectively, the matrices $A'A$ and $B'B$ are both idempotent and such that $A'B = B'A = 0$. The residual sum of squares for error after deducting the sum of squares for regression is $y'B'\left[I-BX(X'B'BX)^{-1} X'B' \right]By$ with q-k degrees of freedom (if there are k independent variates), and, since the matrix $M'M$ of this quadratic form is idempotent, its mean square yields an estimate

---

*There is no connection between the vector b and the matrix B as there is between a and A above. The notation b for a regression coefficient is universal and is retained.

of the intrinsic variance of y. When the error regression is applied to the treatment-means of y, the vector of residuals is $Ay-AXb$, where $b=(X'B'BX)^{-1}X'B'By$, or $A\left[I-X(X'B'BX)^{-1}X'B'B\right]y$, so that the sum of squared residuals is $y'\left[I-B'BX(X'B'BX)^{-1}X'\right]A'A\left[I-X(X'B'BX)^{-1}X'B'B\right]y$, or $y'L'Ly$. Now

$$
\begin{aligned}
M'L &= \left[B'-B'BX(X'B'BX)^{-1}X'B'\right]\left[A-AX(X'B'BX)^{-1}X'B'B\right] \\
&= 0. \qquad \text{Since } B'A=0.
\end{aligned}
$$

Hence, as seen in the proof of Lemma 10 (Cor.), the two quadratic forms $y'M'My$ and $y'L'Ly$ are independent.

But
$$
\begin{aligned}
L'L = A'A &- \left[B'BX(X'B'BX)^{-1}X'A'A\right]-\left[A'AX(X'B'BX)^{-1}X'B'B\right] \\
&+\left[B'BX(X'B'BX)^{-1}X'A'AX(X'B'BX)^{-1}X'B'B\right],
\end{aligned}
$$

so that
$$
\begin{aligned}
(L'L)^2 = {} & A'A-\left[A'AX(X'B'BX)^{-1}X'B'B\right] \\
&+\left[B'BX(X'B'BX)^{-1}X'A'AX(X'B'BX)^{-1}X'B'B\right]-\left[B'BX(X'B'BX)^{-1}X'A'A\right] \\
&+\left[A'AX(X'B'BX)^{-1}X'A'A\right]-\left[A'AX(X'B'BX)^{-1}X'A'AX(X'B'BX)^{-1}X'B'B\right] \\
&-\left[B'BX(X'B'BX)^{-1}X'A'AX(X'B'BX)^{-1}X'A'A\right] \\
&+\left[B'BX(X'B'BX)^{-1}X'A'AX(X'B'BX)^{-1}X'A'AX(X'B'BX)^{-1}X'B'B\right] \\
\neq {} & L'L.
\end{aligned}
$$

The matrix $L'L$ is therefore not idempotent. Hence the mean square of the quadratic form $y'L'Ly$ would not be an estimate of the intrinsic variance of y. Cochran, in fact, showed that such a quadratic form, not having a matrix with equal non-zero latent roots, is not distributed as a gamma-type variate. The F-test therefore fails.

The correct test, first published by Bartlett (27), may be described as follows. The treatments and error lines of the table are pooled so as to give sums of squares and sums of products for "treatments + error" with $p+q$ degrees of freedom. The sum of squares due to regression is now obtained from these, leaving a residual sum of squares with $p+q-k$ degrees of freedom. From this is subtracted the residual sum of squares for error, $y'B'\left[I-BX(X'B'BX)^{-1}X'B'\right]By$ with $q-k$ degrees of freedom, so that a new mean square with p degrees of freedom is obtained for comparison by the F-test with the residual sum of squares for error. We now examine the validity of this procedure.

It is perhaps not immediately obvious that this test is suitable for the purpose on hand. Just as the residual sum of squares for error is that due to the intrinsic random variance of the dependent variate, so the residual sum of squares for treatments + error is that due to the intrinsic "random + treatment" variance of the dependent variate. Their difference is therefore a sum of squares due to the intrinsic treatments variance of y, so that a comparison with the intrinsic error variance by means of the F-test, if valid, is suggested.

If the sum of squares for treatments+error is $y'C'Cy = y'A'Ay + y'B'By$ with $p+q$ degrees of freedom, where $A'A$ and $B'B$ are both idempotent and such that $A'B = B'A = 0$, it is easily proved in the same manner as in Lemma 10 that $C'C$ is also idempotent. The sum of squared residuals for treatments + error is $y'C'\left[I - CX(X'C'CX)^{-1}X'C'\right]Cy$, while that for error is $y'B'\left[I - BX(X'B'BX)^{-1}X'B'\right]By = y'M'My$. Their difference is $y'\left[C'C - B'B - C'CX(X'C'CX)^{-1}X'C'C + B'B(X'B'BX)^{-1}X'B'B\right]y$ or $y'N'Ny$. Putting $C'C = A'A + B'B$, we have

$$N'N = A'A - \left[(A'A+B'B)X(X'C'CX)^{-1}X'(A'A+B'B)\right] + \left[B'BX(X'B'BX)^{-1}X'B'B\right].$$

Also $M'M = B'B - \left[B'BX(X'B'BX)^{-1}X'B'B\right]$,

$$\begin{aligned}\therefore M'M.N'N = & -B'BX(X'C'CX)^{-1}X'A'A - B'BX(X'C'CX)^{-1}X'B'B \\ & + B'BX(X'B'BX)^{-1}X'B'B + B'BX(X'C'CX)^{-1}X'A'A \\ & + B'BX(X'C'CX)^{-1}X'B'B - B'BX(X'B'BX)^{-1}X'B'B \\ = & \ 0,\end{aligned}$$

which is the criterion for the independence of the two quadratic forms $y'M'My$ and $y'N'Ny$ and is equivalent to $M'N = N'M = 0$. Applying now Lemma 10 (Cor.), we see that $y'N'Ny$ must have gamma-type distribution with $(p+q-k) - (q-k) = p$ degrees of freedom, and its mean square is, on a null hypothesis (that the intrinsic treatments variance is zero), an estimate of the intrinsic random variance of y independent of the estimate from the error residual. The F-test is therefore valid and exact.

The significance of the error regression may be tested by analysis of variance, thus:-

Analysis of Variance

Source of Variation.	D.F.	Sums of Squares	Mean Squares
Regression	k	$y'B'BX(X'B'BX)^{-1}X'B'By$	s_1^2
Deviations from Regression	q-k	$y'B'[I - BX(X'B'BX)^{-1}X'B']By$	s_2^2
Total	q	$y'B'By$	s_3^2

The product of the matrices of the quadratic forms for regression and for deviations from regression is zero and the matrices are both idempotent. The corresponding mean squares are therefore independent estimates of the variance of the y's, and are hence amenable to the F-test. Should the test prove significant we will wish to assume the amended hypothesis that part of the random variance of the y's is due to regression, and that the residual mean square s_2^2 is an estimate of what we have called the intrinsic variance of the y's, i.e. the random variance after allowance for regression. Owing to the fact that the b's are not independent of one another their individual contributions to the sum of squares for regression are not easily obtainable. However, each b may be tested for significance by means of the t-test, for $\{b\}=(X'B'BX)^{-1}X'B'By$, so that the variance matrix of the b's is, by Lemma 5, $(X'B'BX)^{-1}$, since the y's are independent and of equal weight and B'B is idempotent. Hence the variances of the b's are given by $d_{ii}\sigma^2$, where d_{ii} is a diagonal term of $(X'B'BX)^{-1}$, and the value of t is given by $\frac{b_i}{\sqrt{d_{ii}}\, s_2}$, s_2^2 being the estimate of variance from the residual sum of squares for error with q-k degrees of freedom. Numerator and denominator are easily seen to be independent, for $\{b\}=(X'B'BX)^{-1}X'B'By = Ry$ and $s_2^2=\frac{1}{q-k}\, y'M'My$ where $M = B-[BX(X'B'BX)^{-1}X'B'B]$, so that $RM=0$, remembering that, since B is the matrix of the error residuals, B is idempotent as well as B'B. This was proved for the error residuals of the split-plot design on Page 90, and may similarly be proved for any regular design. It arises from the fact that the matrix of the vector of error residuals of an analysis of variance (here B) is the same as the matrix of the vector of residuals after fitting constants by least squares (Section 11), and such a matrix is always idempotent.

Values of y adjusted for regression may be obtained and, if the effect of treatments has proved significant in the analysis of covariance, the adjusted treatment means may be compared. Referred to the mean as origin, an adjusted treatment mean is $a_i'y - a_i'Xb$, where b is the vector of error regression coefficients and a_i' is a row of A. The error regression coefficients are used for this adjustment because they are a measure of the regressions when block, etc. differences have been removed. The means we wish to compare are independent of block, etc. differences and for purposes of comparison we are testing a null hypothesis in respect of treatments. Hence the adjustment is made by means of the error regressions. It remains to find the standard error of the difference of two adjusted means and to prove that the t-test is valid for such a comparison. The difference between two adjusted treatment means is

$$(a_i' - a_j')\, y - (a_i' - a_j')\, Xb$$
$$= (a_i' - a_j') \left[I - X(X'B'BX)^{-1} X'B'B \right] y = Sy$$

By Lemma 5, the variance of this difference

$$= (a_i' - a_j') \left[I - X(X'B'BX)^{-1} X'B'B \right] \left[I - X(X'B'BX)^{-1} X'B'B \right]' (a_i - a_j)\sigma^2$$
$$= (a_i' - a_j') \left[I - X(X'B'BX)^{-1} X' \right] (a_i - a_j)\, \sigma^2$$
$$(\text{since } a_i'B = a_j'B = 0)$$
$$= \frac{2\sigma^2}{r} - (a_i' - a_j')\, X\, (X'B'BX)^{-1} X' (a_i - a_j)\, \sigma^2,$$

where r = number of plot-yields per treatment mean. This is equivalent to the results of Wishart (29). To show that the difference of two adjusted means is independent of the estimate of variance s_2^2, we have that

$$SM = (a_i' - a_j')\left[I - X(X'B'BX)^{-1}X'B'B\right]\left[B - BX(X'B'BX)^{-1}X'B'B\right]$$

$$= (a_i' - a_j')\left[X(X'B'BX)^{-1}X'B'B - X(X'B'BX)^{-1}X'B'B\right]$$

$$= 0 \text{ , again since } a_i'B = a_j'B = 0 \text{ and } B \text{ is idempotent.}$$

The validity of the t-test is therefore proved.

List of References.

1. Aitken, A.C.- Determinants and Matrices (Oliver and Boyd) : P.65.
2. Aitken, A.C.- On the graduation of data by the orthogonal polynomials of least squares: Proc.R.S. Edin., Vol.LIII , P. 54.
3. Aitken, A.C.- On linear approximation by least squares: Proc.R.S. Edin., Vol.LXII, Part II, No.16, P.138.
4. Barnard, M.M.- An enumeration of the confounded arrangements in the 2x2x2... factorial designs: Supp.J.R.S.S., Vol.3, No.2, P.195.
5. Christidis, B.G.- The importance of the shape of plots in field experimentation: J.Agr.Sc., Vol.XXI, P.14.
6. Cochran, W.G.- The distribution of quadratic forms in a normal system with application to the analysis of covariance: Proc. Camb.Phil.Soc.,30, P.178.
7. Day., J.W.- The relation of size, shape, and number of replications of plots to probable error in field experiments: Journ.Am.Soc.Agron.,12, P.100.
8. Eden, T. and Yates, F.- On the validity of Fisher's z-test when applied to an actual example of non-normal data: J.Agr.Sc., Vol.XXIII, P.6.
9. Finney, D.J.- The fractional replication of factorial arrangements: Ann.Eugen, Vol.12. Pt.4, P.291.

10. Finney, D.J.- Recent developments in the design of field experiments. III Fractional replication: J.Agr.Sc., Vol.XXXVI, Pt.3, P.184.

11. Fisher, R.A. and Mackenzie, W.A.- Studies in crop variation. II The manurial response of different potato varieties: J.Agr.Sc., XIII, P.311.

12. Fisher, R.A.- Statistical Methods for Research Workers (Oliver and Boyd).

13. Fisher, R.A.- Design of Experiments (Oliver and Boyd).

14. Fisher, R.A. and Yates, F.- Statistical Tables for Biological, Medical, and Agricultural Research. (Oliver and Boyd).

15. Irwin, J.O.- Mathematical theorems involved in the analysis of variance : J.R.S.S., 94, P.284.

16. Irwin, J.O.- On the independence of constituent items in the analysis of variance: Supp.J.R.S.S.,1, P.236.

17. Mercer, W.B. and Hall, A.D.- The experimental error of field trials: J.Agr.Sc., Vol.IV, P.107.

18. Nair, K.S.- The calculation of standard errors and tests of significance of different types of treatment comparisons in split-plot and strip arrangements of field experiments: Indian J.Agr.Sc., Vol.XIV, Pt.IV, P.315.

19. Smith, H.F.- An empirical law describing heterogeneity in the yields of agricultural crops: J.Agr.Sc., Vol.28, P.1.

20. Sukhatme, P.V.- On Fisher and Behrens' test of significance for the difference in means of two normal samples: Sankhya, Vol.4, P.39.

21. Wishart, J. and Sanders, H.G.- Principles and Practice of Field Experimentation(Empire Cotton Growing Corp.)

22. Wishart, J.- Field trials: their layout and statistical analysis (Imp. Bureau of Plant Breeding and Genetics).

23. Wishart,J.- Field experiments of factorial design: J.Agr.Sc., Vol.XXVIII, P.299.

24. Yates, F.- The principles of orthogonality and confounding in replicated experiments: J.Agr.Sc., Vol.XXIII, P.108.

25. Yates, F.- Complex experiments: Supp. J.R.S.S.,2,P.181.

26. Yates,F.- Design and Analysis of Factorial Experiments: Tech.Communication No.35, Imp. Bureau of Soil Science, Harpenden.

Additional References.

27. Bartlett,M.S.- The vector representation of a sample: Proc. Camb.Phil.Soc., 30, P.327.

28. Snedecor, G.W.- Statistical Methods (Collegiate Press, Ames, Iowa).

29. Wishart, J.- Tests of significance in the analysis of covariance: Supp. J.R.S.S., 3, P.79.