COMBINED INTRABLOCK AND INTERBLOCK ESTIMATES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

MANAVAZHI VIJAYA KRISHNA MENON B.A. (HONOURS), M.Sc.

*****

The Ohio State University
1959

Approved by

[Signature]
Adviser

Department of Mathematics
INTRODUCTION

The method known as the recovery of interblock information in the estimation of any "treatment" in an experimental design was originated by Yates (1939, 1940) for lattices and incomplete balanced block designs. By combining the estimate, called the intrablock estimate, obtained by minimizing the sum of squares of the differences of the observations from their means, with the interblock estimate which is obtained by minimizing the sum of squares of the differences of the block sums (i.e., the sums of observations in each of the blocks) from their means, an estimate is obtained which has lesser variance than the variance of either the intrablock or the interblock estimate. A general method of analysis applicable to any incomplete block design with equal block sizes was given in 1947 by C. R. Rao.¹

This paper presents a method of reducing the observations, assumed to be normally distributed, in any incomplete block design with equal block size and having uniquely soluble intrablock equations, to a set of independent linear combinations of the observations.

This set will be such that any linear combination of the observations which is orthogonal to the mean can be represented as a linear combination of the members of the set. This reduction is then employed to prove various limit and other theorems. Only those designs are considered which have equal block size.

The first chapter derives this reduction under the assumption that the intrablock normal equations are uniquely soluble. If $\hat{S}_\alpha$ is a linear combination of "treatments," $\hat{S}_\alpha^*$ the estimate so designated in Chapter One, and $S_\alpha^*$ the estimate obtained from $\hat{S}_\alpha$ by replacing the ratio $\theta$ of the variances of the interblock and intrablock estimates by a consistent estimate $\theta^*$, then it is proved that the ratio of the lengths of the confidence intervals for $\hat{S}_\alpha$ and $S_\alpha^*$ approaches in probability the limit 1.

Also, if $P(|\hat{S}_\alpha - S_\alpha| \leq \eta_s^\alpha) = \beta$, then the limit of $P(|\hat{S}_\alpha^* - S_\alpha^*| \leq \eta_s^\alpha) \beta$ where $\eta_s^\alpha$ is obtained from $\eta_s^\alpha$ by replacing $\theta$ by $\theta^*$. Further it is shown that $\frac{\nu}{\nu^*} \rightarrow 0$ in probability as $\theta^* \rightarrow \theta$, where $\nu_\alpha$ is the effect of "treatment" or "variety" $\alpha$. This method of reducing the observations to a set of independent quantities provides us with a systematic way of estimating $S_\alpha$ and $\theta$ and facilitates the proofs of the limit theorems mentioned above, but is not advanced as a computational method.
In the next chapter the method presented by Rao, hereinafter called Rao's method, is shown to be equivalent to our method of estimation, and both are shown to be the same as the maximum likelihood method. This equivalence with the maximum likelihood method does not appear to have been noted so far. A proof is given of the fact that our method gives the best unbiased linear estimate. Such an estimate is unique. Rao shows that the expression to be minimized is \[ \omega \sum \{ O_i - f_i(t) \}^2 + k\omega' \sum \{ O'_i - f'_i(t) \}^2, \]
where the \( O_i \) are linear combinations of the observations contained in a block, orthogonal to one another and to the block sum, the \( O'_i \) are linear combinations of the block averages, \( f_i(t) = E(O_i) \), \( f'_i(t) = E(O'_i) \), and \( \omega \) and \( k\omega' \) are the reciprocals of the variances of \( O_i \) and \( O'_i \) respectively. The \( O_i \) and \( O'_j \) are mutually independent. From his derivation it does not follow that one can use \[ \omega \sum \{ y_{ij} - E(y_{ij}) \}^2 + \frac{\omega'}{k} \sum \{ B_j - E(B_j) \}^2 \]
for minimizing, as is usually done, since the \( y_{ij} \)'s and the \( B_j \)'s are no longer mutually independent. That this can be done is brought out in this chapter.

---

In Chapter III, it is shown that, assuming the intra- and inter-block normal equations are uniquely soluble, the estimate of \( \hat{v}_a \) computed by our method is the same as the best combined linear estimate of the intrablock estimate \( \hat{v}_a \) and the interblock estimate \( \hat{v}_a \), for all \( a \), if and only if the incomplete block design is balanced.

In the fourth and last chapter an estimate \( \theta^* \) of \( \theta \) is proposed which is such that \( E(S_{\theta^*}) = S_\theta \), and an expression given for the variance of \( S_{\theta^*} \) by using the procedure indicated by Paul Meier.\(^3\)

**Symbols and Notations**

The following notations will uniformly be used:

**Symbols Connected with the Design**

\( y_{ij} \): the observation in the \( j^{th} \) block connected with the \( i^{th} \) variety.

\( b_j \): the effect (random) of the \( j^{th} \) block.

\( v_i \): the effect of the \( i^{th} \) variety.

\( \mu \): the general mean.

\( \epsilon_{ij} \): the random error.

\( \sigma^2 \): the variance of \( \epsilon_{ij} \), for all \( i, j \).

\( \sigma_{b_j}^2 \): the variance of \( b_j \), for all \( j \).

\( \theta \): \( \frac{\sigma^2 + k \sigma^2_{b_j}}{\sigma^2_{b_j}} \)

---

\(^3\) Paul Meier, "Variance of a Weighted Mean," Biometrika, IX (1959).
k: the number of plots per block.
b: the number of blocks.
v: the number of varieties.
\( r_i \): the number of replications of the \( i^{th} \) variety.
\( \lambda_{ij} = \lambda_{ji} \): the number of times the \( i^{th} \) and \( j^{th} \) varieties occur together in a block.

\( B_j \): the sum of the observations in the \( j^{th} \) block.
\( \bar{B}_j = \frac{1}{k} B_j \): average of the block sum.
\( \bar{y} \): the arithmetic mean of the observations.

Symbols for Estimates

'\(^\star\)': denotes an intra block estimate.
'\(^*\)': denotes an inter block estimate.
'\(^\wedge\)': denotes the best unbiased estimate when the variances are known.
'\(^\dagger\)': denotes the estimate obtained by replacing \( \theta \) by \( \theta^* \).

Matrix Symbols

All vectors are column vectors.

\((a^{ij})\): the inverse of the matrix \( A = (a_{ij}) \).
\( \text{diag} (\lambda_i) \)
\( \text{diag} (\lambda) \): a diagonal matrix.
(v): the column vector \[ \begin{bmatrix} v_1 \\ \vdots \\ \vdots \\ v_n \end{bmatrix} \]

\(\prime\): the transpose of a matrix.

**Summation Symbol**

\[ \sum_{\alpha} x_{\alpha} = \sum_{\alpha} x_{\alpha} \]:
the summation of \( x_{\alpha} \) extending over the domain of definition of \( x_{\alpha} \).
CHAPTER I

Let $y_{ij}$ which denotes the observation associated with the $i^{th}$ variety in the $j^{th}$ block of an incomplete block design be given as:

$$y_{ij} = v_i + b_j + \mu + \epsilon_{ij},$$ (cf. symbols and notations, page 4)

where each $b_j$ is distributed $N(0, \sigma^2)$, each $\epsilon_{ij}$ is distributed $N(0, \sigma^2)$ and the $b_j$ and the $\epsilon_{ij}$ are mutually independent.

Assuming that the intrablock normal equations are uniquely soluble for the varieties, it will now be shown that the total number $N$ of the observations $y_{ij}$ can be reduced to a set of $N-1$ independent quantities such that any linear combination of the observations which is orthogonal to the mean $\bar{y}$, can be represented as a linear form in these $N-1$ quantities and vice versa.

Whereas the $N-1$ members of the set are mutually independent, it will be noted that the $y_{ij}$ themselves are not, for two observations taken from the same block have covariance $\sigma^2$.

We start with a partially confounded factorial design having $v$ treatment combinations. Let $S_1, \ldots, S_{v-1}$ be $v-1$ normalized linear forms in these treatment combinations which are mutually orthogonal and are also orthogonal to the mean of the treatment combinations. In any
replication we assume that $b-1$ of the $S_i$ are confounded with the blocks. Since the mean is always confounded, we see that the remaining $S_i$ are orthogonal to the blocks.\(^1\)

Now let $L_\alpha, \alpha=1, \ldots, v-1$ denote the same linear form in the associated observations of any replication, as $S_\alpha$ is in the treatment combinations. By an associated observation, we mean the observation to which a particular treatment combination is applied. Let $L_\alpha^{(n)}, \ldots, L_\alpha^{(m)}$ be the observed values of $L_\alpha$ in the replications where $S_\alpha$ is unconfounded, and $M_\alpha^{(n)}, \ldots, M_\alpha^{(m)}$ the observed values in those replications where $S_\alpha$ is confounded. Thus $n_\alpha = 0$ would imply that $S_\alpha$ is confounded in all replications and similarly $m_\alpha = 0$ would mean that there are no replications in which $S_\alpha$ is confounded. We note that $n_\alpha + m_\alpha$ is the total number of replications, and if $n = \sum_i n_\alpha, m = \sum_i m_\alpha$, then $n + m = (v-1)$ (the number of replications).

If $n_\alpha \neq 0$, let $(a_{ij})$ denote an orthogonal matrix

of order $n_\alpha$ whose first row is $\frac{-1}{2} n_\alpha, \ldots, \frac{-1}{2} n_\alpha$. $(b_{ij})$ is a matrix formed similarly when $m_\alpha \neq 0$.

Let $x_{\alpha j} = \sum b_{ij}^\alpha L_{ij}$ and $y_{\alpha j} = \sum b_{ij} M_{ij}$. Then
\[
(x_{\alpha j}^{(i)}) = \frac{L_{\alpha j}^{(i)}}{\sqrt{m_\alpha}}, E(x_{\alpha j}^{(i)}) = \sqrt{m_\alpha} S_{\alpha i}.
\[
(y_{\alpha j}^{(i)}) = \frac{M_{\alpha j}^{(i)}}{\sqrt{m_\alpha}}, E(y_{\alpha j}^{(i)}) = \sqrt{m_\alpha} S_{\alpha i}.
\]
\[
E(x_{\alpha j}^{(i)}) = 0, (j \neq 1); E(y_{\alpha j}^{(i)}) = 0, (j \neq 1).
\]

Let $L_{\alpha i}$ be $\sum a_{ij}^{(i)} x_j$ where $x_j^{(i)}$ is an observation in the $i^{th}$ block, $j = 1, \ldots, k$; $i = 1, \ldots, b$. If $L_{\alpha i}$ is orthogonal to the blocks, $\sum b_{ij} = 0$ for each $i$.

\[
\text{variance of } L_{\alpha i} = \sum \text{(variance of } \sum b_{ij} x_j^{(i)})
= \sum \sigma^2(\sum b_{ij}^{(i)})^2 + \sum \sigma^2(\sum b_{ij}^{(i)})^2 = \sigma^2.
\]

Next if $M_{\alpha i} = \sum b_{ij} x_j$, since $M_{\alpha i}$ is confounded with the blocks, $b_{ij} = b_j$, say.

\[
1 = \sum b_{ij} = k \sum b_{ij}, \text{ or } \sum b_{ij}^2 = \frac{1}{k}.
\]

Thus, variance of $M_{\alpha i} = \text{variance of } \sum b_{ij} B_i = (\sum b_{ij}^2) k(\sigma^2 + k\sigma^2) = \sigma^2 + k\sigma^2 = \sigma^2$.

Further Cov ($L_{\alpha i} M_{\alpha i}$) = 0

\[
\text{L}_{\alpha i} \text{ and } M_{\alpha i} \text{ are mutually independent.}
\]

Noting that $(a_{ij})$ and $(b_{ij})$ are orthogonal matrices we have, then,
\[
(x_{\alpha j}^{(i)}) = \sigma^2, \text{Var}(y_{\alpha j}^{(i)}) = \sigma^2, \text{and the } x_{\alpha j}^{(i)} \text{ and the } y_{\alpha j}^{(i)} \text{ are mutually independent.}
\]

Now certain of the $S_i$, say $S_{r+1}, \ldots, S_{v-1}$ may be known to be 0. We arrange the pairs $(\alpha, j)$ for $\alpha > r$, 
and for \( \alpha \leq r, j \geq 2 \) in some arbitrary order, and according to this order put \( x_\alpha = z^\beta \) and \( y_\alpha = t^\gamma \). For \( \alpha \leq r, j = 1 \) we put
\[
\begin{align*}
x_\alpha^{(i)} & \text{ for } n_\alpha \neq 0 \\
0 & \text{ for } n_\alpha = 0
\end{align*}
\]
\[
\begin{align*}
y_\alpha^{(j)} & \text{ for } m_\alpha \neq 0 \\
0 & \text{ for } m_\alpha = 0
\end{align*}
\]
We see then, using (1.1) and (1.2), that \( E(x_\alpha) = \sqrt{n_\alpha} \), \( E(y_\alpha) = \sqrt{m_\alpha} \), \( E(z^\beta) = 0 \), \( \text{Var}(x_\alpha) = \sigma^2 \), \( \text{Var}(y_\alpha) = \sigma^2 \), \( \text{Var}(z^\beta) = \sigma^2 \).

The following set of observations are therefore obtained:
\[
\begin{align*}
x_\alpha = c_\alpha S_\alpha + \varepsilon_\alpha, \alpha = 1, \ldots, r, \\
y_\alpha = d_\alpha S_\alpha + \eta_\alpha, \alpha = 1, \ldots, r, \\
z_\alpha = \varepsilon_\alpha, \alpha = r + 1, \ldots, n, \\
t_\alpha = \eta_\alpha, \alpha = r + 1, \ldots, m,
\end{align*}
\]
where \( c_\alpha = \sqrt{n_\alpha} \), \( d_\alpha = \sqrt{m_\alpha} \); \( \varepsilon_\alpha = 0 \) if \( c_\alpha = 0 \) and \( \eta_\alpha = 0 \) if \( d_\alpha = 0 \); and the non-zero \( \varepsilon \)'s and \( \eta \)'s are jointly normally and independently distributed with mean zero, the \( \varepsilon \)'s with variance \( \sigma^2 \) and the \( \eta \)'s with variance \( \sigma^2 \).

Moreover \( c_\alpha^2 + d_\alpha^2 > 0 \). The inequality \( \sigma^2 > \sigma^2 \) holds, but will not be used in this paper.

The same reduction to (1.3) is obtained in the case of quasi-factorial designs. The \( S_\alpha \) are then the quasi-factors and certain linear combinations of the quasi-factors give the varietal effects.
We proceed next to show that the system \((1.3)\) is obtained in the case of any incomplete block design with unique intrablock estimates.

From the quadratic form \(Q = \sum (v_{ij} - v_i - b_j - \mu)^2\) under the assumption \(\Sigma r_i v_i = 0, \Sigma b_j = 0,\) the first hypothesis \(v_i = 0,\ i = 1, \ldots, v,\) and the second hypothesis \(b_j = 0\) for all \(j,) we obtain the identity
\[
\sum y_{ij}^2 = \tilde{Q}_a + Q_r - Q_a + \frac{1}{k} \sum B_j^2,
\]
where \(\tilde{Q}_a\) is the minimum of \(Q\) under the assumption, \(Q_r\) is the minimum under the first hypothesis, and \(Q_r - Q_a = \tilde{Q}_r - \tilde{Q}_a.\)

Now \(Q_r - Q_a\) is a non-negative definite quadratic form in the \(\tilde{v}_i\)'s and \(\Sigma r_i v_i = 0.\) Hence there exists an orthogonal matrix \(O\) of order \(v\) having \(\frac{r_i}{v} - \frac{r_v}{v^2} (\Sigma r_i^2)^{-1}\) for its last row, such that, on applying the transformation \(O(\tilde{v}) = (K),\) \(Q_r - Q_a\) becomes \(\sum \lambda_i \tilde{k}_i^2 = \sum \tilde{\lambda}_i \tilde{k}_i^2,\) since \(\tilde{k}_v = \sum r_i \tilde{v}_i = 0.\) Putting \(\tilde{\lambda}_i \tilde{k}_i = \tilde{M}_i\) we obtain \(Q_r - Q_a = \sum \tilde{M}_i^2\). Hence
\[
(1.4) \quad \sum y_{ij}^2 = \tilde{Q}_a + \sum \tilde{M}_i^2 + \frac{1}{k} \sum B_j^2.
\]

We now prove that \(\tilde{Q}_a,\) the \(\tilde{M}_i\) and \(B_j\) are mutually independent.

In \((1.4)\) we have the rank equation \(N = (n - v - b + 1) + (v - 1) + (b).\) Hence, writing \(\tilde{Q}_a = \Sigma P_i^2,\) say, we see

---

by Cochran's theorem that the $P_i$, the $\bar{M}_i$, and the $B_j$ are mutually orthogonal linear forms in the observations.

From this it is easily verified that covariance $(P_i, B_j) = 0$ and covariance $(\bar{M}_i, B_j) = 0$, which shows that $B_j$ is independent of the $P_i$ and the $\bar{M}_i$. In order to show now that the $P_i$ and the $\bar{M}_i$ are mutually independent, we observe that $\bar{Q}_a$ and $\bar{Q}_r - \bar{Q}_a$ are independent of the $b_j$. We can thus assume that $b_j$ is fixed for all $j$. This, in turn, means that we can regard the $y_{ij}$ as independent random variables with variance $\sigma^2$, and by Cochran's theorem the $P_i$ and $\bar{M}_i$ are mutually independent with variance $\sigma^2$. Let $(\bar{v}) = A(\bar{M})$. Here $A$ is a $v \times (v - 1)$ matrix of rank $v - 1$. If $E(\bar{M}) = (M)$, then $(v) = E(\bar{v}) = A(M)$. We substitute for the $v_i$'s in $E(B_j) = \sum v_i + k\mu$, in terms of the $M_i$'s by means of the above equation, and minimize $\sum (B_j - E(B_j))^2$ in order to obtain the interblock estimates. Since the interblock estimate of $\mu$ is $\bar{v}$, the equations for estimating the $M_i$ are given by $A_j'(\Theta) = A_j' A_j(\bar{M})$ where $(\Theta) = (L) = (B - k\bar{v})$, and $E(\Theta) = A_j(M)$. Let $T$ be an orthogonal matrix such that $T'A_jA_jT = D = \text{diag}(d_1, \ldots, d_s, 0, \ldots, 0)$.

Then $T'A_j'(\Theta) = T'A_jA_j(\bar{M})$. Putting $\bar{M} = T\bar{M}$, we obtain (1.5) $T'A_j'(\Theta) = D\bar{M}$.

Since the sum of squares $\sum (B_j - E(B_j))^2$ must have

---

a minimum, the last equation must be consistent. But the last \( v - 1 - s \) elements of the column vector on the right hand side are 0. Therefore, the last \( v - 1 - s \) rows of \( T^t A \) must vanish also. From (1.5) we get the interblock estimates \( \bar{N}_1, \ldots, \bar{N}_s \) of the \( N_i = E(\bar{N}_i) \). Expressing \( E(B_j) \) in terms of the \( N_i, i = 1, \ldots, s \), and applying the sequence of hypothesis theorem\(^4\) to \( Q = \sum \{ B_j - E(B_j) \}^2 \), we obtain

\[
(1.6) \quad \sum B_j^2 = \bar{Q}_a + \sum \bar{Q}_r - Q_a + bk'y^2,
\]

where \( \bar{Q}_a \) is the minimum of \( Q \) under the assumption, \( \bar{Q}_r \) is the minimum under the hypothesis \( N_i = 0, i = 1, \ldots, s \) and \( \bar{Q}_r - Q_a = \bar{Q}_r - \bar{Q}_a \) is a non-negative definite quadratic form in the \( \bar{N}_i \)'s. Thus from (1.4)

\[
\sum y_{ij}^2 = \bar{Q}_a + \sum \bar{Q}_r - Q_a + \bar{Q}_a + bk'y^2
\]

or, \( \sum (y_{ij} - \bar{y})^2 = \bar{Q}_a + \sum \bar{Q}_r - Q_a + \bar{Q}_a \).

Applying the orthogonal transformation \( T\bar{N} = \bar{N} \), we can write the above equation as

\[
\sum (y_{ij} - \bar{y})^2 = \bar{Q}_a + \sum \bar{Q}_r - Q_a + \bar{Q}_a.
\]

Yet another orthogonal transformation enables us to write

\[
(1.7) \quad \sum (y_{ij} - y)^2 = Q_a + \sum \bar{H}_i^2 + \bar{Q}_r - Q_a + \bar{Q}_a.
\]

Since from (1.6), \( \frac{1}{k} \sum B_j^2 = \sum (B_j)^2 = \frac{1}{k} \bar{Q}_a + \frac{1}{k} \sum \bar{H}_i + bk'y^2 \), and the \( B_j \) are independent quantities with variance \( \sigma' \), we see by applying Cochran's theorem, that \( Q_a \) and the \( \bar{H}_i \)

are mutually independent and the variance of $\sqrt{e_i}H_i$ is $\sigma^{'2}$, for $i = 1,\ldots,s$. Furthermore if $\frac{1}{K}Q_a$ is written as a sum of squares, then each term will have mean 0 and variance $\sigma^{'2}$.

Finally, if the equations

(1.8) $v_j = \sum_{\alpha=1}^{\nu} c_{\alpha}^\nu H_{\alpha}^\nu$, $j = 1,\ldots,\nu$, express the $v_j$ in terms of the $H_{\alpha}$, we let $\sum_{j}^{\nu} c_{j\alpha}^\nu = c_{\alpha}^\nu$ and define $S_{\alpha}$ (for reasons explained in the next paragraph) to be $\frac{H_{\alpha}^\nu}{c_{\alpha}^\nu}$. Also let $\sqrt{c_{\alpha}^\nu} = d_{\alpha}^\nu$. Then we have

(1.9) $\tilde{H}_{\alpha} = c_{\alpha}^\nu S_{\alpha}$ and $\sqrt{c_{\alpha}^\nu} H_{\alpha} = \sqrt{c_{\alpha}^\nu} c_{\alpha}^\nu S_{\alpha} = d_{\alpha}^\nu S_{\alpha}$.

Further,

(1.10) $\sigma^{'2}_{\alpha} = \frac{\sigma^2}{c_{\alpha}^\nu}$ and $\sigma^{'2}_{\alpha} = \frac{\sigma^2}{d_{\alpha}^\nu}$.

Writing, in (1.7), $\tilde{H}_i^\nu$ as $x_i$, $\sqrt{e_i}H_i$ as $y_i$, $\tilde{Q}_a$ as $\Sigma z_i^2$ and $\frac{1}{K}Q_a$ as $\Sigma t_i^2$, we see that $x_i$, $y_i$, $z_i$, and $t_i$, give us the reduction (1.3).

We have from (1.8),

$$\sigma^{'2}_{v_j} = \sum_{\alpha=1}^{\nu} c_{j\alpha}^\nu \sigma^{'2}_{\alpha} = \sigma^2(\sum_{\alpha=1}^{\nu} c_{j\alpha}^\nu)$, $j = 1,\ldots,\nu$.

Thus if the variances of the $\tilde{v}_j$ are small, then all the $c_{j\alpha}$ are small. The $c_{\alpha}$ are therefore large and from (1.10) this implies that the $S_{\alpha}$ have small variances. Conversely if the variances of the $S_{\alpha}$ are small then so are those of the $\tilde{v}_\alpha$. The $\tilde{v}_\alpha$ and the $S_{\alpha}$ may therefore be said to be comparable.
We note for future reference

\begin{equation}
\begin{aligned}
Q_r - Q_a &= \sum_i \tilde{H}_i^2 = \sum_i e_i^2 \tilde{S}_i^2 = \frac{\sum_i \tilde{S}_i^2}{\tilde{S}_i^2}.
\end{aligned}
\end{equation}

\begin{equation}
\frac{1}{k} \left( Q_r - Q_a \right) = \sum_i \left( \tilde{e}_i \tilde{H}_i^2 \right)^2 = \sum_i \left( d_i \tilde{S}_i \right)^2 = \frac{\tilde{S}_i^2}{\tilde{S}_i^2}.
\end{equation}

We remark that for an incomplete balanced block design, we have corresponding to (1.7), the equation,

\[ \Sigma (y_{ij} - \bar{y})^2 = \tilde{Q}_a + \lambda \Sigma \tilde{v}_i^2 + \mu \Sigma \tilde{v}_i^2 + \tilde{Q}_a. \]

This can be reduced by an orthogonal transformation of the form \( O(v) = (S) \), where the last row of \( O \) consists of the same element \( v^T \), to

\begin{equation}
\begin{aligned}
\tilde{Q}_a + \lambda \Sigma \tilde{S}_i^2 + \mu \Sigma \tilde{S}_i^2 + \tilde{Q}_a &= \Sigma \tilde{z}_i^2 + \lambda \Sigma \tilde{S}_i^2 + \mu \Sigma \tilde{S}_i^2 + \Sigma \tilde{t}_i^2.
\end{aligned}
\end{equation}

Here \( x_i = \sqrt{\lambda} \tilde{S}_i \) and \( y_i = \sqrt{\mu} \tilde{S}_i \) and the \( z_i \) and the \( t_i \) are as previously defined.

We have thus shown that the equations (1.3) arise in partially confounded factorial designs as well as in those incomplete block designs with equal block size which have unique intrablock estimates for the varieties.

The likelihood function \( L \) of the joint distribution of the \( x_i \), the \( y_i \), the \( z_i \), and the \( t_i \) is

\begin{equation}
\begin{aligned}
L = \text{constant} - \frac{m}{2} \log \theta - N \log \sigma - \frac{1}{2} \sigma^2 \theta \left\{ \right.
&\theta \Sigma z_i^2 + \theta \Sigma (x_{a} - c_{a} S_{a})^2 + \Sigma t_i^2 + \Sigma (y_{a} - d_{a} S_{a})^2 \left. \right\}.
\end{aligned}
\end{equation}

If \( \theta \) is known we obtain from this the maximum likelihood estimates
\[
\begin{align*}
\sigma^2 & = \frac{1}{n\theta}\left\{\sum x^2 + \theta \sum (x_\alpha - c_\alpha \hat{S}_\alpha)^2 + \sum t^2 + \sum (y_\alpha - d_\alpha \hat{S}_\alpha)^2\right\}, \\
\hat{S}_\alpha & = \frac{\theta c_\alpha x_\alpha + d_\alpha y_\alpha}{\theta c_\alpha^2 + d_\alpha^2}.
\end{align*}
\]

Since \( \hat{S}_\alpha = x_\alpha/c_\alpha \), \( \hat{S}_\alpha = y_\alpha/d_\alpha \), \( \hat{S}_\alpha \) can be written as
\[
\frac{\theta c_\alpha \hat{S}_\alpha + d_\alpha ^2 \hat{S}_\alpha}{\theta c_\alpha^2 + d_\alpha^2}.
\]

Usually however \( \theta \) is not known and has to be estimated.

Applying the maximum likelihood method to (1.13) under the assumption that \( \theta \) is not known, one would get the equations (1.14) with \( \theta \) replaced by its maximum likelihood estimate \( \hat{\theta} \), and the additional equation
\[
(1.15) \quad \hat{\theta} = \frac{1}{m\hat{\sigma}^2} \left( \sum t^2 + \sum (y_\alpha - d_\alpha \hat{S}_\alpha)^2 \right).
\]
Together with (1.14) this gives
\[
(1.16) \quad \hat{\sigma}^2 = \frac{1}{n}\left\{\sum x^2 + \sum (x_\alpha - c_\alpha \hat{S}_\alpha)^2\right\}.
\]
\[
\left\{\sum t^2 + \sum (y_\alpha - d_\alpha \hat{S}_\alpha)^2\right\}/m
\]
\[
\left\{\sum x^2 + \sum (x_\alpha - c_\alpha \hat{S}_\alpha)^2\right\}/m
\]

The solution of (1.16) and (1.14) for \( \hat{\theta} \) leads to a cubic equation in \( \hat{\theta} \) which appears to be too complicated to be useful. We therefore proceed in a different way. We take (1.16) as a basis for an estimate of \( \theta \) and shall assume that \( c_\alpha \) is not zero, \( \alpha = 1, \ldots, r \).

We replace in (1.16) \( \hat{S}_\alpha \) by the intrablock estimate \( \tilde{S}_\alpha \). The numerator and denominator of (1.16) then become biased estimates of \( \sigma^2 \) and \( \hat{\sigma}^2 \) respectively. The bias in the denominator is corrected by replacing \( n \) by \( n - r \).
To correct the bias in the numerator we compute

\[ E \sum (y_{\alpha} - d_{\alpha} \tilde{S}_{\alpha})^2 = E \sum (y_{\alpha} - d_{\alpha} S_{\alpha})^2 + E \sum d_{\alpha}^2 (S_{\alpha} - \tilde{S}_{\alpha})^2 \]

\[ = s \sigma^2 + \sigma^2 \sum \frac{d_{\alpha}^2}{c_{\alpha}} \], where \( s \) is the number of non-zero \( y_{\alpha} \)'s.

In this way we are led to the estimate

(1.17)

\[ \theta^* = \frac{n - r}{m} \left\{ \sum t_{\alpha}^2 + \sum (y_{\alpha} - d_{\alpha} \tilde{S}_{\alpha})^2 - (\sum \frac{d_{\alpha}^2}{c_{\alpha}^2} \sum z_{\alpha}^2 / n - r) \right\} / \sum z_{\beta}^2. \]

We now assume that we have a sequence of designs for estimating the \( S_{\alpha} \), for which \( n \to \infty \), \( m \to \infty \). Under these conditions \( \theta^* \) is a consistent estimate of \( \theta \). We obtain an estimate \( S^* \) of \( S_{\alpha} \) by replacing in the second equation of (1.14) the parameter \( \theta \) by \( \theta^* \). Thus

(1.18) \[ S^* = \frac{\theta^* c_{\alpha}^2 \tilde{S}_{\alpha} + d_{\alpha}^2 \tilde{S}_{\alpha}}{\theta^* c_{\alpha}^2 + d_{\alpha}^2}. \]

The estimate \( S^* \) is the one customarily used. It seems worth while to note its relation to the maximum likelihood estimates of \( S_{\alpha} \) for the case when \( \theta \) is known, and if \( \theta \) is replaced by a consistent estimate \( \theta^* \). If we replace in (1.18) \( \theta^* \) by \( \theta \) we obtain the maximum likelihood estimate \( \hat{S}_{\alpha} \) for the case that \( \theta \) is known and if \( \theta^* \) is replaced by \( \hat{\theta} \) from equation (1.16) one obtains the maximum likelihood estimate of \( S_{\alpha} \) for the case that \( \theta \) is not known. The following equation (1.19) shows, moreover, that for large enough values of \( n \) and \( m \) the estimate \( S^* \) is for all practical purposes, as good as the estimate \( \hat{S}_{\alpha} \).
In the following we shall use the notation proposed in (5) and shall repeatedly use the fact that all rules applicable to ordinary limit and order relations are also applicable to stochastic limit and order relations.  

We shall prove

**Theorem 1.1.** \( \text{plim} \frac{\hat{S}_\alpha - S^*}{\sigma_{\hat{S}_\alpha}} = 0. \)

**Proof:** Putting \( c^2_\alpha = n_\alpha, \ d^2_\alpha = m_\alpha, \) we have

\[
\hat{S}_\alpha - S^* = \frac{n_\alpha m_\alpha (\bar{S}_\alpha - \bar{S}_\alpha)(-\theta^* + \theta)}{(\theta n_\alpha + m_\alpha)(\theta n_\alpha + m_\alpha)}.
\]

Furthermore,

\[
\sigma^2 = \frac{\theta}{\theta n_\alpha + m_\alpha} \sigma^2, \text{ and } \text{Var}(\hat{S}_\alpha - \bar{S}_\alpha) = \frac{\theta n_\alpha + m_\alpha}{n_\alpha m_\alpha} \sigma^2.
\]

Hence,

\[
\sqrt{\frac{m_\alpha n_\alpha}{n_\alpha + m_\alpha}} (\bar{S}_\alpha - \bar{S}_\alpha) = o_p(1).
\]

Furthermore,

\[
\sqrt{\frac{m_\alpha n_\alpha}{\theta n_\alpha + m_\alpha}} \leq \min\left(\frac{\sqrt{n_\alpha}}{\theta} \ , \ \sqrt{\frac{n_\alpha}{m_\alpha}}\right) \leq \max(1, \frac{1}{\theta^*}) = o_p(1),
\]

and since \( \theta - \theta^* = o_p(1), \) we have

\[
\frac{\hat{S}_\alpha - S^*}{\sigma_{\hat{S}_\alpha}} = o_p(1) o_p(1) = o_p(1),
\]

which is just another way of writing \( \text{plim} \frac{\hat{S}_\alpha - S^*}{\sigma_{\hat{S}_\alpha}} = 0. \)

---

Let $\sigma^*$ be any consistent estimate of $\sigma$ (for instance its intrablock estimate) and form $\tilde{\sigma}^* = \sqrt{\frac{\theta^*}{\theta^*_n + m_\alpha}} \sigma^*$, by replacing $\theta$ by $\theta^*$ and $\sigma$ by $\sigma^*$, in $\sigma^*$. Then we have

$\lim \frac{\tilde{\sigma}^*}{\sigma^*} = \lim \sqrt{\frac{\theta^*}{\theta^*_n + m_\alpha}} \tilde{\sigma}^* = 1.$

We determine $\eta$ so that

$P( |S_{a} - S_{a}^*| \leq \eta \sigma^*_{a}) = \beta.$

Consider the two confidence intervals

$|S_{a} - S_{a}| \leq \eta \sigma^*_{a}$

and

$|S_{a}^* - S_{a}| \leq \eta \bar{\sigma}^*.$

We deduce immediately from theorem 1.1 and (1.19)

**Theorem 1.2.** The ratio of the difference between the end points of the confidence intervals (1.21) and (1.22) to the length of the confidence interval (1.21) approaches in probability the limit 0.

Thus if $m$ and $n$ are sufficiently large we may for all practical purposes replace the exact confidence interval (1.21) by the interval (1.22). We prove next the

**Theorem 1.3.** Limit $P(|S_{a}^* - S_{a}| \leq \eta \sigma^*) = \beta$, as $n, m, \rightarrow \infty$ and $\theta^* \rightarrow \theta$ in probability.

**Proof:** We denote $\hat{S}_{a} - S_{a}$ by $A$, $S_{a}^* - S_{a}$ by $B$, and $S_{a}^* - \hat{S}_{a}$ by $C$. Let $\epsilon$ and $\epsilon_1$ be arbitrary positive constants such that $\epsilon < \eta$. Since $B = A + C$ and $\eta \sigma^* = (\eta - \epsilon) \sigma^* + \epsilon \bar{\sigma}^*$, we have
(1.23) \( P(|B| \leq \eta \sigma^*) \geq P(|A| \leq (\eta - \epsilon) \sigma^*, |C| \leq \epsilon \sigma^*) \).

(1.24) Now \( P(|C| \geq \epsilon \sigma^*) + P(|A| \leq (\eta - \epsilon) \sigma^*, |C| \leq \epsilon \sigma^*) \)
\( \geq P(|A| \leq (\eta - \epsilon) \sigma^*, |C| \geq \epsilon \sigma^*) + P(|A| \leq (\eta - \epsilon) \sigma^*, |C| \leq \epsilon \sigma^*) \)
\( = P(|A| \leq (\eta - \epsilon) \sigma^*). \)

But since from theorem 1.1 and (1.19), \( \text{plim}_{\sigma^*} = 0 \),
we have for sufficiently large \( n \) and \( m \), \( P(|C| \geq \epsilon \sigma^*) \leq \epsilon. \)

By using this in (1.24), we obtain
\( P(|A| \leq (\eta - \epsilon) \sigma^*, |C| \leq \epsilon \sigma^*) \geq P(|A| \leq (\eta - \epsilon) \sigma^*) - \epsilon. \)

Together with (1.23) this gives
(1.25) \( P(|B| \leq \eta \sigma^*) \geq P(|A| \leq (\eta - \epsilon) \sigma^*) - \epsilon. \)

Now suppose that \( \sigma^* \leq \sigma^*_a \). Then since \( A \) is normally distributed with mean 0,
\( P(|A| \leq (\eta - \epsilon) \sigma^*) \)
\( = P(|A| \leq (\eta - \epsilon) \sigma^*_a) - P((\eta - \epsilon) \sigma^*_a < |A| \leq (\eta - \epsilon) \sigma^*) \)
\( \geq P(|A| \leq (\eta - \epsilon) \sigma^*_a) - P(|A| \leq (\eta - \epsilon)(\sigma^*_a - \sigma^*) \)
\( = P(|A| \leq (\eta - \epsilon) \sigma^*_a) - P(|A| \leq (\eta - \epsilon)(\sigma^*_a - \sigma^*), (\sigma^*_a - \sigma^*) \leq \epsilon_2) \)
\( - P(|A| \leq (\eta - \epsilon)(\sigma^*_a - \sigma^*), (\sigma^*_a - \sigma^*) > \epsilon_2), \)

where \( \epsilon_2 \) is an arbitrary positive constant,
\( \geq P(|A| \leq (\eta - \epsilon) \sigma^*_a) - P(|A| \leq (\eta - \epsilon) \epsilon_2) - P(\sigma^*_a - \sigma^* > \epsilon_2). \)

By choosing \( \epsilon_2 \) sufficiently small, the second term on the right can be made arbitrarily small, say less than \( \epsilon. \)

By choosing \( n \) and \( m \) sufficiently large the third term becomes \( \epsilon. \)

Therefore we obtain from (1.25),
(1.26) \( P(|B| \leq \eta \sigma^*) \geq P(|A| \leq (\eta - \epsilon) \sigma^*_a) - 3 \epsilon. \)
and this inequality is evidently true if \( \sigma* \geq \sigma^* \).

The right side
\[
= P( |A| \leq \eta \sigma^*_a ) - P( (\eta - \epsilon) \sigma^*_a < |A| \leq \eta \sigma^*_a ) - 3 \epsilon.
\]

By choosing \( \epsilon \) sufficiently close to 0 we see that
\[
P( (\eta - \epsilon) \sigma^*_a < |A| \leq \eta \sigma^*_a )
\]
can be made \( \leq \epsilon \).

Thus we have from (1.26),
\[
(1.27) \quad P( |B| \leq \eta \bar{\sigma}^*) \geq P( |A| \leq \eta \sigma^*_a ) - 4 \epsilon.
\]

Similarly, denoting \( \hat{S}_a - S^* \) by \( D \) and observing that \( B + D = A \), we have
\[
P( |A| \leq (\eta + \epsilon) \bar{\sigma}^*) \geq P( |B| \leq \eta \bar{\sigma}^* \& |D| \leq \epsilon \bar{\sigma}^* )
\]
\[
\geq P( |B| \leq \eta \bar{\sigma}^*) - \epsilon, \text{ for sufficiently large } n \text{ and } m.
\]

By proper choice of \( \epsilon \), we can make the left side of this inequality \( \leq \beta + \epsilon \), so that for sufficiently large \( n \) and \( m \),
\[
(1.28) \quad \beta + 2 \epsilon \geq P( |B| \leq \eta \bar{\sigma}^* ).
\]

Since \( \epsilon \) is arbitrary, we have the result, from (1.27) and (1.28).

Finally, we prove the

**Theorem 1.4.** \( \lim_{n \to \infty} \frac{V_i - V_i^*}{\sigma_{\mu_i}} = 0 \), as \( \theta* \to \theta \) in probability.

**Proof:** Let \( V_i \) be equal to \( \frac{\sum_i c_{ia} S_a}{\sqrt{\lambda_i}} \), \( i = 1, \ldots, v_0 \).

Then \( \hat{V}_i = \sum_a c_{ia} \hat{S}_a \) and \( V_i^* = \sum_a c_{ia} S^* \). Therefore,
\[
(1.29) \quad \frac{V_i - V_i^*}{\sigma_{\mu_i}} = \sum_a c_{ia} (\hat{S}_a - S_a) \frac{\hat{S}_a}{\sigma_{\mu_a}} \frac{\sigma_{\mu_a}}{\sigma_{\mu_i}}.
\]
Consider one of the terms on the right, say,
\[ c_{ia} \frac{S_a - S^*_a}{\sigma^*_a} \]

From theorem 1.1 we have \( \lim_{n \to \infty} \frac{S_a}{\sigma_a} = 0 \).

Now \( \frac{2 \sigma^2_a}{\sigma^2} = \frac{c_{ia}^2 \sigma^2}{\sum c_{ia}^2 \sigma^2} \leq 1 \).

\[ \therefore \lim_{n \to \infty} c_{ia} \frac{S_a}{\sigma} \frac{S_a - S^*_a}{\sigma_a} = 0. \]

Thus each of the \( v - 1 \) terms on the right hand side of (1.29) converges to 0 in probability. The theorem is therefore proved.
CHAPTER II

As a preliminary to proving the equivalence of the maximum likelihood method, the method of Chapter I and Rao's method of getting a combined estimate, we derive the joint probability density function $P(y_{ij})$ of the $y_{ij}$.

The joint probability density function of the $y_{ij}$, given the $b_j$ is

$$
\frac{1}{(\sqrt{2\pi}\sigma)^b} \exp\left(-\frac{\sum(y_{ij} - v_i - b_j - \mu)^2}{2\sigma^2}\right)
$$

and the joint probability density function of the $b_j$ is

$$
(2.1) \frac{1}{(\sqrt{2\pi}\sigma^*)^b} \exp\left(-\frac{\sum b_j^2}{2\sigma^*^2}\right).
$$

Therefore

$$
(2.2) P(y_{ij}) = \frac{1}{(\sqrt{2\pi}\sigma)^b} \frac{1}{(\sqrt{2\pi}\sigma^*)^b} \int e^{-\frac{\sum(y_{ij} - v_i - b_j - \mu)^2}{2\sigma^2} - \frac{\sum b_j^2}{2\sigma^*^2} db}
$$

where $db$ denotes $db_1 db_2 \ldots \ldots db_b$.

Omitting the factor $-\frac{1}{2}$, the expression in the exponent becomes

$$
\sum \frac{(y_{ij} - v_i - \mu)}{\sigma^2}^2 + k \frac{\sum b_j^2}{\sigma^2} + \sum \frac{b_j^2}{\sigma^*^2} + 2 \sum b_j \sum \frac{(y_{ij} - v_i - \mu)}{\sigma^2}.
$$

Let $\sigma^2 = A$ and $\sigma^*^2 = B$. 
Then the expression given above becomes

\[
(2.3) \begin{cases}
\frac{1}{A} \sum (y_{ij} - v_i - \mu)^2 &+ \frac{A + kB}{AB} \sum b_j^2 - 2 \sum b_j.
\end{cases}
\]

\[
= \frac{1}{A} \sum (y_{ij} - v_i - \mu)^2 - \frac{B}{A(A + kB)} \sum \left( \frac{1}{A} \sum (y_{ij} - v_i - \mu)^2 \right)^2 + \\
+ \frac{A + kB}{AB} \sum \left\{ b_j - \frac{B}{A + kB} \sum (y_{ij} - v_i - \mu) \right\}^2.
\]

Now since (2.1) is the probability density function of the $b_j$, we see that

\[
\int_{-\infty}^{\infty} \frac{1}{(\sqrt{2\pi})^b} e^{-\frac{b_j^2}{2b}} \, db = 1.
\]

From this we obtain the identity

\[
\int_{-\infty}^{\infty} e^{-\frac{b_j^2}{2b}} \, db = (\sqrt{2\pi})^b \] and it is easily deduced that

\[
\int_{-\infty}^{\infty} e^{-\frac{c(b_j - d)^2}{2b}} \, db = \left(\frac{\sqrt{2\pi}}{\sqrt{c}}\right)^b.
\]
Using this, we have from (2.2) and (2.3),
\[
P(y_{ij}) = \frac{1}{(\sqrt{2\pi A})^{bk}(\sqrt{2\pi B})^{b}} \left( \frac{2\pi B}{A + kB} \right)^b e^{-\frac{1}{2}X},
\]
where
\[
1 = \frac{\beta(j^2A)k^j}{\alpha + kB}
\]
\[
(2.4) \quad X = \sum \frac{(y_{ij} - v_i - \mu)^2}{A} - \frac{B}{A(A + kB)} \sum (B_j - \sum v_i - k\mu)^2
\]
is the quantity to be minimized in estimating the \(v_i\)'s and \(\mu\) by the maximum likelihood method. We prove now

**Theorem 2.1.** The maximum likelihood method and the method of Chapter I for estimating the \(S_\alpha\)'s are equivalent.

**Proof:** We know that
\[
\sum y_{ij}^2 = \tilde{Q}_a + \tilde{Q}_r - Q_a + \sum \frac{B_j^2}{k},
\]
where \(\tilde{Q}_a\) is independent of the \(v_i\), the \(b_j\), and \(\mu\); and \(\tilde{Q}_r - Q_a\) is independent of the \(b_j\) and \(\mu\), and is a quadratic form in the \(S_\alpha\). Therefore,
\[
(2.5) \quad \sum (y_{ij} - v_i - \mu)^2 = \tilde{Q}_a + \tilde{Q}_r - Q_a + \frac{1}{k} \sum (B_j - \sum v_i - k\mu)^2,
\]
where now \(\tilde{Q}_r - Q_a\) is the same quadratic form in the \(S_\alpha\) as \(\tilde{Q}_r - Q_a\) is in the \(S_\alpha\).

But from (2.4) the quantity to be minimized in the maximum likelihood method is
\[
X = \sum \frac{(y_{ij} - v_i - \mu)^2}{A} - \frac{B}{A(A + kB)} \sum (B_j - \sum v_i - k\mu)^2.
\]
Substituting from (2.5), $X$ becomes

$$\frac{1}{A} \left\{ \overline{Q}_a + \overline{Q}_r - \overline{Q}_a \right\} + \left\{ \frac{A}{Ak} - \frac{B}{A(A + kB)} \right\} \overline{\left( B_j - \sum (j) \nu_i - k\mu \right)^2} =$$

$$(2.6) \quad \frac{1}{A} \left\{ \overline{Q}_a + \overline{Q}_r - \overline{Q}_a \right\} + \frac{1}{k(A + kB)} \overline{\left( B_j - \sum (j) \nu_i - k\mu \right)^2}.$$

But from (1.6), we have $B_j^2 = \overline{Q}_a + \overline{Q}_r - \overline{Q}_a + k^2b(y - \mu)^2$.

Here $\overline{Q}_a$ is independent of the $\nu_i$ and $\mu$, and $\overline{Q}_r - \overline{Q}_a$ is a quadratic form in the $\overline{S}_a$ independent of $\mu$. Therefore

$$\overline{\left( B_j - \sum (j) \nu_i - k\mu \right)^2} = \overline{Q}_a + \overline{Q}_r - \overline{Q}_a + k^2b(\overline{y} - \mu)^2,$$

where $\overline{Q}_r - \overline{Q}_a$ is the same quadratic form in the $(\overline{S}_a - \overline{S}_a)$ as $\overline{Q}_r - \overline{Q}_a$ is in the $\overline{S}_a$.

Substituting in (2.6), we obtain

$$X = \frac{1}{A} \left\{ \overline{Q}_a + \overline{Q}_r - \overline{Q}_a \right\} + \frac{1}{k(A + kB)} \left\{ \overline{Q}_a + \overline{Q}_r - \overline{Q}_a \right\} +$$

$$\frac{(\sum B_j - k\mu)^2}{kb(A + kB)}.$$

Since all terms except the last term in the expression on the right hand side of the preceding equation are independent of $\mu$, we see that the maximum likelihood estimate of $\mu$ is $\overline{y}$ and the substitution of $\overline{y}$ for $\mu$ makes the last term vanish.

Thus we have only to minimize

$$\frac{1}{A} \left\{ \overline{Q}_a + \overline{Q}_r - \overline{Q}_a \right\} + \frac{1}{k(A + kb)} \left( \overline{Q}_a + \overline{Q}_r - \overline{Q}_a \right).$$

In order to complete the proof we have merely to note
that in the notation of Chapter I,

\[ Q_a = \sum z_a^2, \quad Q_r - Q_a = \sum x_a^2, \] variance of \( x_a \) = variance of \( z_a \) = \( \sigma^2 = A; \)

\[ \frac{1}{k} Q_a = \frac{1}{k} t_a^2, \quad \frac{1}{k} Q_r - Q_a = \sum y_a^2, \] and variance of \( t_a \) = variance of \( y_a \) = \( \sigma' = A + kB. \)

As already pointed out in the Introduction, we mean by Rao's method, the method of obtaining estimates by minimizing the expression

\[ \sum \omega(y_{ij} - v_i - b_j - \mu)^2 + \sum \omega' (B_j - \sum v_i - k\mu)^2 \]

with respect to the \( v_i \), the \( b_j \) and \( \mu \) under the assumption that \( \sum r_i v_i = \sum b_j = 0 \). Here \( \omega = \frac{1}{\sigma^2} \) and \( \omega' = \frac{1}{\sigma'^2}. \)

**Theorem 2.2.** The maximum likelihood method is equivalent to Rao's method of estimation.

**Proof:** In the maximum likelihood method we have to minimize (cf. 2.4)

\[ X = \frac{1}{A} \sum (y_{ij} - v_i - b_j - \mu)^2 - \frac{B}{A(A + kB)} \sum (B_j - \sum v_i - k\mu)^2. \]

Now \( \sum (y_{ij} - v_i - b_j - \mu)^2 \) can be written as

\[ \sum (y_{ij} - v_i - b_j - \mu)^2 + k \sum b_j^2 + 2 \sum b_j \sum (y_{ij} - v_i - b_j - \mu) \]

\[ = \sum (y_{ij} - v_i - b_j - \mu)^2 + k \{ \sum b_j + \sum (y_{ij} - v_i - b_j - \mu)^2 \} \]

\[ - \frac{1}{k} \sum \left\{ \sum (y_{ij} - v_i - b_j - \mu)^2 \right\} \]
\[= \sum (y_{ij} - v_i - b_j - \mu)^2 + \frac{1}{k} \sum (B_j - \sum v_i - k\mu)^2 - \]

\[\frac{1}{k} \sum (B_j - \sum v_i - kb_j - k\mu)^2.\]

But \[\frac{\partial x}{\partial b_j} = \sum (y_{ij} - v_i - b_j - \mu) + (B_j - \sum v_i - kb_j - k\mu)\]

\[= 2(B_j - \sum v_i - kb_j - k\mu).\]

Since \[\frac{\partial x}{\partial b_j} = 0\] is one of the conditions for a minimum, we have \[B_j - \sum v_i - kb_j - k\mu = 0,\] i.e., the last term in (2.8) becomes zero. Thus from (2.7), we have merely to minimize

\[\frac{1}{A} \sum (y_{ij} - v_i - b_j - \mu)^2 + \frac{1}{Ak} \sum (B_j - \sum v_i - k\mu)^2 - \]

\[\frac{B}{A(A + kB)} \sum (B_j - \sum v_i - k\mu)^2\]

\[= \frac{1}{A} \sum (y_{ij} - v_i - b_j - \mu)^2 + \]

\[\frac{1}{k(A + kB)} \sum (B_j - \sum v_i - k\mu)^2\]

which is the expression to be minimized in Rao's method. The maximum likelihood method and Rao's method will therefore lead to the same estimates for the \(v_i\), the \(b_j\), and \(\mu\).

Let \(y\), \(i = 1, \ldots, n\), be normally distributed random variables, not necessarily independent of one another.
Let their variance-covariance matrix be \( \Lambda \).

Let \( E(y_i) = \sum t_i^* a_{ij} t_j \) for each \( i \), the \( t_j \) being certain parameters to be estimated. In matrix notation, we have \( E(y) = A(t) \).

**Definition.** A linear function \( p't = \sum p_i t_i \) of the parameters \( t_i \) is said to be *estimable* if there exists \( b'y = \sum b_i y_i \) such that \( E(b'y) = p't \).

Rao shows\(^1\) that the best unbiased estimate of \( p't \) is \( m^' \Lambda^{-1} y \) where \( m \) is a column vector satisfying the equation \( p = A^' \Lambda^{-1} A m \). The best unbiased estimate is also the same as \( p't^* \) where \( t^* \) is any solution of the equation

\[
A^' \Lambda^{-1} y = A^' \Lambda^{-1} A t^*.
\]

(2.9) These equations can be obtained by minimizing with respect to the \( t_i \), the expression

\[
\{y' - E(y')\} \Lambda^{-1} \{y - E(y)\}.
\]

Now consider an incomplete block design. Let \( O_i, i = 1, \ldots, k - 1 \), be \( k - 1 \) linear combinations of the observations in block \( j \), orthogonal to one another and to the block sum. Let \( O_i, i = 1, \ldots, b(k - 1) \) denote in some order the linear forms \( O_i, i = 1, \ldots, k - 1, j = 1, \ldots, b \). The \( O_i \) are mutually independent, and have the same variance \( \sigma^2 \).

Let \( O'_i, i = 1, \ldots, b - 1 \) be \( b - 1 \) linear combinations of the block averages \( \overline{B}_1, \ldots, \overline{B}_b \), which are mutually orthogonal, and are orthogonal to the sum \( \overline{B}_1 + \ldots + \overline{B}_b \) of the block averages. The \( O'_i \) are mutually independent, and are independent of the \( O_j \). Further the variance of \( O'_i = \sigma'^2 \) for all \( i \). We see then from (2.9) that the expression to be minimized for the best unbiased estimates of the \( t_i \)'s, is

\[
(2.10) \quad \frac{1}{n} \sum_{i} \{ O_i - E(O_i) \}^2 + \frac{1}{k} \sum_{j} \{ B_j - E(B_j) \}^2
\]

This is the expression contained in (1). The method commonly adopted is to minimize

\[
(2.11) \quad \frac{1}{n} \sum_{i} \{ y_{ij} - E(y_{ij}) \}^2 + \frac{1}{k} \sum_{j} \{ B_j - E(B_j) \}^2.
\]

(2.11) is very similar to (2.10), but certainly the use of (2.11) for minimizing cannot be justified from the fact that (2.10) leads to best unbiased estimates. It appears as though (2.11) has been used on the basis of its similarity to (2.10), and not on the basis of the true fact (cf. theorem 2.2) that it is equivalent to (2.9) which would be the only justification for its use.

We prove next

**Theorem 2.3.** The estimate \( \hat{S}_\alpha \) is the best unbiased linear estimate of \( S_\alpha \). Such an estimate is unique.

**Proof:** Any linear combination \( L_p \) of the observations
which is orthogonal to the mean can be expressed
(cf. Chapter I) as a linear combination of the $S^\alpha$, $\bar{S}^\alpha$, $z^\alpha$, and $t^\alpha$.
Thus $L_\beta = \Sigma a_\sigma \bar{S}^\alpha + \Sigma c_\sigma z^\alpha + \Sigma d_\sigma t^\alpha + \Sigma b_\sigma \bar{S}^\alpha$.
If $L_\beta$ is an unbiased estimate of $S^\beta$, then
$E(L_\beta) = \Sigma a_\sigma S^\alpha + \Sigma b_\sigma S^\alpha = S^\beta$.
But there can be no non-trivial linear relation among the $S^\alpha$.
Hence $a_\beta + b_\beta = 1$ and $a_\sigma + b_\sigma = 0$ for $\alpha \neq \beta$.
$\therefore b_\alpha = 1 - a_\alpha$ for $\alpha \neq \beta$.
Assuming that $L_\beta$ is an unbiased estimate of $S^\beta$ we have
$L_\beta = (a_\sigma \bar{S}^\sigma + b_\sigma \bar{S}^\sigma) + \Sigma a_\sigma \bar{S}^\alpha - \Sigma a_\sigma \bar{S}^\alpha + \Sigma c_\sigma z^\alpha + \Sigma d_\sigma t^\alpha$.
$\therefore \text{Var}(L_\beta) = \text{Var}(a_\sigma \bar{S}^\sigma + b_\sigma \bar{S}^\sigma) + \Sigma a_\sigma^2 (\sigma^2_{\bar{S}^\sigma} + \sigma^2_{\bar{S}^\sigma}) + \sigma^2 \Sigma c_\sigma^2 + \sigma^2 \Sigma d_\sigma^2$.
Clearly the variance of $L_\beta$ will be least if and only if
$a_\sigma = 0$ for $\alpha \neq \beta$, $c_\alpha = 0$ for all $\alpha$, $d_\alpha = 0$ for all $\alpha$ and when $a_\beta$, $b_\beta$ are so chosen - subject to the restriction
$a_\beta + b_\beta = 1$ - that the variance of $a_\beta \bar{S}^\beta + b_\beta \bar{S}^\beta$ is least.
Thus $L_\beta$ will have minimum variance if and only if
$L_\beta = a_\beta \bar{S}^\beta + b_\beta \bar{S}^\beta$, where $a_\beta + b_\beta = 1$ and $a_\beta$ and $b_\beta$ are so chosen that $\sigma^2_{L_\beta}$ is minimum. But this means that $L_\beta = \bar{S}^\beta$, completing the proof of the theorem.
CHAPTER III

We know from Chapter II that
\[
\begin{bmatrix}
S_1 \\
\vdots \\
S_{v-1} \\
0_{v-1}
\end{bmatrix} =
\begin{bmatrix}
v_1 \\
\vdots \\
v_{v-1} \\
v_v
\end{bmatrix},
\]
where \( A \) is a non-singular matrix whose \( v \)th row is \( r_1, \ldots, r_v \). Let \( (v) = \begin{bmatrix} v_1 \\ \vdots \\ v_{v-1} \end{bmatrix} \) and \( (S) = \begin{bmatrix} S_1 \\ \vdots \\ S_{v-1} \end{bmatrix} \).

We prove now the following:

**Theorem 3.1.** \((v) = C(S)\), where \( C \) is non-singular.

**Proof:** Since
\[
\begin{bmatrix}
S_1 \\
\vdots \\
S_{v-1} \\
0_{v-1}
\end{bmatrix} =
\begin{bmatrix}
v_1 \\
\vdots \\
v_{v-1} \\
v_v
\end{bmatrix},
\]
we have

\[(3.1) \begin{bmatrix} v_1 \\ \vdots \\ v_v \end{bmatrix} = A^{-1} \begin{bmatrix} S_1 \\ \vdots \\ S_{v-1} \\ 0_{v-1} \end{bmatrix}.
\]

Let the minor matrix corresponding to \( a^{vv} \) be \( C \). We see from (3.1) that \( (v) = C(S) \). We have merely to show now that \( |C| \neq 0 \). But \( a^{vv} = \frac{|C|}{|A^{-1}|} \) and \( a^{vv} = r_v \neq 0 \).

\( \therefore \) \( |C| \neq 0 \) and \( C \) is non-singular.

The next theorem follows easily from the preceding one.

**Theorem 3.2.** There exists a non-singular matrix \( C \)
with non-zero diagonal elements such that on suitably renumbering \(S_1, \ldots, S_{v-1}\), we have \((v) = C(S)\).

**Proof:** That there exists a non-singular matrix \(C_1\) such that \((v) = C_1(S)\) has already been proved above.

There exist elementary matrices \(A_1, A_2, \ldots, A_m\) such that \(C_1A_i\), \(i = 1, \ldots, m\), is the matrix obtained from \(C_1\) by interchanging two columns, say the \(p_i^{th}\) and the \(q_i^{th}\), and also such that \(C_1(A_1 \ldots A_m)\) is a matrix with non-zero diagonal elements. But if \(S\) is multiplied on the left by \(A_i^{-1}\), the effect is merely to interchange the \(p_i^{th}\) and the \(q_i^{th}\) rows of \(S\).

But \((v) = \left\{ C_1(A_1 \ldots A_m) \right\} \left\{ (A_m^{-1} \ldots A_1^{-1})S \right\}\). Setting \(C_1(A_1 \ldots A_m) = C\) and noting that \(A_m^{-1} \ldots A_1^{-1}(S)\) is a column vector whose elements are obtained by rearranging \((S)\), we have the result.

It is assumed henceforth that \((v) = C(S)\) where \(C\) is as in the preceding theorem.

We have now

\[(3.2) \quad (\forall) = C(\overline{S}), \quad (\overline{v}) = C(\overline{S}) \quad \text{and} \quad (\hat{v}) = C(\hat{S}).\]

Let

\[\frac{\sigma^2_{v_i}}{\sigma^2_{v_i} + \sigma^2_{v_i}} = \lambda_i', \quad \frac{\sigma^2_{v_i}}{\sigma^2_{v_i} + \sigma^2_{v_i}} = \mu_i',\]

\[\frac{\sigma^2_{S_i}}{\sigma^2_{S_i} + \sigma^2_{S_i}} = \tau_i \quad \text{and} \quad \frac{\sigma^2_{S_i}}{\sigma^2_{S_i} + \sigma^2_{S_i}} = \gamma_i',\]

for \(i = 1, \ldots, v-1.\)
Then,

\[(3.3) \text{diag}(\lambda) + \text{diag}(\mu) = \text{diag}(\xi) + \text{diag}(\eta) = I,\]

where \(I\) is the identity matrix.

From (1.14), we have \((\hat{S}) = \text{diag}(\xi) (\hat{S}) + \text{diag}(\eta) (\bar{S}).\)

Using (3.2) we then obtain

\[(3.4) (\hat{\nu}) = C\{\text{diag}(\xi) (\hat{S}) + \text{diag}(\eta) (\bar{S})\}.\]

If now we define \(v_i^* = \lambda_i \tilde{v}_i + \mu_i \bar{v}_i, \quad i = 1, \ldots, v-1,\)

then \(v_i^*\) is the best unbiased linear combination of \(\tilde{v}_i\)

and \(\bar{v}_i\). The column vector \((v^*)\) can be written as

\[(3.5) (v^*) = \text{diag}(\lambda) (\tilde{\nu}) + \text{diag}(\mu) (\bar{\nu}) = \text{diag}(\lambda) C(\tilde{S}) + \text{diag}(\mu) C(\bar{S}).\]

We have the definition:

An incomplete block design is called balanced if \(r_i = r,\)

\(i = 1, \ldots, v\) and \(\lambda_{ij} = \lambda\) for all \(i, j, i \neq j.\)

The main result of this chapter is the following:

**Theorem 3.3.** \((v^*) = (\hat{\nu})\) if and only if the design is balanced.

We prove first a few lemmas.

**Lemma 3.1.** \(\bar{Q}_i = Q_a = \sum \bar{v}_i \tilde{Q}_i,\) where \(\tilde{Q}_i\) is the adjusted total for the \(i^{th}\) variety, i.e., \(\tilde{Q}_i = (\text{sum of the yields of the } i^{th} \text{ variety}) - (\text{sum of the averages of the blocks containing the } i^{th} \text{ variety}) = V_i - \sum \tilde{S}_i,\)

\(\bar{Q}_a\) is the minimum of \(Q\) under the hypothesis \(v_i = 0,\) for all \(i,\) and \(\bar{Q}_r\) is the minimum of \(Q\) under the hypothesis \(v_i = 0,\) for all \(i.\)
Minimizing Q under the restrictions \( \sum_i r_i v_i = 0 \),
\( \sum_i b_j = 0 \), we have

\[
\begin{align*}
V_i &= r_i \bar{v}_i + \sum_{(j)} \bar{b}_j + r_i \bar{\mu} \\
B_j &= \sum_{(j)} \bar{v}_i + k \bar{b}_j + k \bar{\mu}.
\end{align*}
\]

Minimizing Q under the hypothesis \( v_i = 0 \) for all i gives

\[
B_j = k \bar{b}_j + k \bar{\mu}.
\]

From (3.6) and (3.7) we obtain

\[
\begin{align*}
\bar{B}_j + \bar{\mu} &= \bar{B}_j - \frac{\sum v_i}{k}, \\
\bar{b}_j + \bar{\mu} &= \bar{B}_j.
\end{align*}
\]

Now let \( \Sigma' \) denote the sum over all pairs \( i,j \) for which the \( i \)th variety occurs in the \( j \)th block. Then

\[
\begin{align*}
\bar{Q}_r - Q_a &= \Sigma' \left\{ (\bar{v}_i - \bar{v}_a)\right\}^2 \\
&= \sum' \left\{ \bar{v}_i^2 - 2 \bar{v}_i \frac{\sum (j) \bar{v}_a}{k} + \left( \frac{\sum (j) \bar{v}_a}{k} \right)^2 \right\}.
\end{align*}
\]

But \( \sum' (\bar{v}_i \sum (j) \bar{v}_a) \) becomes on summing first with respect to

\[
\begin{align*}
\Sigma' \left( \frac{\sum (j) \bar{v}_a}{k} \right)^2 &= \sum' \left( \frac{\sum (j) \bar{v}_a}{k} \right)^2.
\end{align*}
\]

Thus we obtain

\[ (3.9) \quad \overline{Q_r - Q_a} = \sum ri \bar{v}_i^2 - \sum' (\bar{v}_i \sum \frac{(j)}{k} \bar{v}_a) \]

On the other hand,

\[ \sum \bar{v}_i \bar{b}_i = \sum \bar{v}_i (v_i - \sum (i) \bar{b}_j) \]

On substituting from (3.6), this becomes

\[ \sum \bar{v}_i (ri \bar{v}_i + \sum (i) \bar{b}_j + ri \bar{p} - \frac{1}{k} \sum (j) \bar{v}_a = \sum (i) \bar{b}_j - ri \bar{p}) \]

\[ = \sum \bar{v}_i (ri \bar{v}_i - \sum (i) \frac{1}{k} \sum \bar{v}_a) \], which from (3.9) is the same as \( Q_r - Q_a \).

**Lemma 3.2.** \( \overline{Q_r - Q_a} = k \sum \bar{v}_i \bar{c}_i \), where \( \bar{c}_i \) is the minimum of

\[ Q = \sum (B_j - \sum (j) v_i - k \bar{p})^2 \] under the assumption \( \sum ri v_i = 0 \),

\( \bar{Q}_r \) is the minimum of \( Q \) under the hypothesis \( v_i = 0 \) for all \( i \) and \( \bar{c}_i = \sum (i) B_j - ri \bar{y} \).

**Proof:** The normal equations under the assumption are

\[ (3.10) \left\{ \begin{array}{l}
\sum (i) B_j = \sum (i) \sum (j) \bar{v}_i + kr_i \bar{p}, \\
\sum B_j = \sum \sum (j) \bar{v}_i + kb \bar{p} = kb \bar{p},
\end{array} \right. \]

since \( \sum (j) \bar{v}_i = \sum ri \bar{v}_i = 0. \)

Thus

\[ (3.11) \quad \bar{p} = \bar{y}. \]

Minimizing \( Q \) under the hopothesis, we have the equation

\[ \sum B_j = kb \bar{p}, \text{ i.e.,} \]

\[ (3.12) \quad \bar{p} = \bar{y}. \]
From (3.10), (3.11), and (3.12) we obtain

\[(3.13) \quad Q_r - Q_a = \sum_j (\bar{\varphi}_j \bar{v}_i)^2\]

\[= \sum_{i=1}^{\bar{v}_i^2} + \sum_{i,j=1}^{i,j} (2\lambda_{ij} \bar{v}_i \bar{v}_j).\]

On the other hand,

\[k\sum_i \bar{v}_i \bar{q}_i = k\sum_i (\sum_{i} \bar{\epsilon}_j - \bar{r}_i \bar{y}).\]

From (3.10) and (3.11), this becomes

\[k \sum_i \bar{v}_i = \sum_i (r_i \bar{v}_i + \sum_{j=1}^{\lambda_{ij} \bar{v}_i \bar{v}_j})\]

\[= \sum r_i \bar{v}_i^2 + 2 \sum_{i,j=1}^{\lambda_{ij} \bar{v}_i \bar{v}_j}\]

\[= Q_r - Q_a, \text{ from (3.13)}.\]

This proves the lemma.

From the intrablock normal equations we obtain on eliminating the \(\bar{\epsilon}_j\)'s and \(\bar{p}\), the following equations for the \(\bar{v}_i\)'s:

---

\[
\begin{bmatrix}
\bar{Q}_1 \\
\vdots \\
\bar{Q}_v
\end{bmatrix} = A
\begin{bmatrix}
\bar{v}_1 \\
\vdots \\
\bar{v}_v
\end{bmatrix},
\]
where
\[
A = \begin{bmatrix}
\frac{r_1 k - 1}{k}, & -\frac{\lambda_{i1}}{k}, & \ldots, & -\frac{\lambda_{ij}}{k}, & \ldots, & -\frac{\lambda_{iv}}{k} \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
-\frac{\lambda_{vi}}{k}, & -\frac{\lambda_{iv}}{k}, & \ldots, & \frac{r_1 k - 1}{k}, & \ldots, & -\frac{\lambda_{iv}}{k} \\
-\frac{\lambda_{vi}}{k}, & -\frac{\lambda_{vi}}{k}, & \ldots, & -\frac{\lambda_{iv}}{k}, & \ldots, & \frac{r_1 k - 1}{k}
\end{bmatrix}
\]

Here \(\sum_{i=1}^{v} \bar{v}_i = \bar{1} = 0\).

If in \(A \begin{bmatrix}
\bar{v}_1 \\
\vdots \\
\bar{v}_v
\end{bmatrix}\), the last row is omitted and \(\bar{v}_v\) is replaced by \(-\frac{\sum_{i=1}^{v} r_i \bar{v}_i}{r_v}\), we obtain

\[
(3.14) \quad \begin{bmatrix}
\bar{Q}_1 \\
\vdots \\
\bar{Q}_{v-1}
\end{bmatrix} = A
\begin{bmatrix}
\bar{v}_1 \\
\vdots \\
\bar{v}_{v-1}
\end{bmatrix},
\]
where \(A\) is the \((v - 1) \times (v - 1)\) matrix such that
\[
a_{ii} = \frac{r_i (k - 1)}{k} + \frac{\lambda_{iv}}{k} \frac{r_i}{r_v}\text{ and}
\]
\[
a_{ij} = -\frac{\lambda_{ij}}{k} + \frac{\lambda_{iv}}{k} \frac{r_j}{r_v}, \text{ for } i \neq j; \ i, j = 1, \ldots, v.
\]
We now have the

**Lemma 3.3:** \( \overline{Q_r - Q_a} = (\forall) / I + \left[ \begin{array}{cccc} \frac{r_1}{r_v} & \cdots & \frac{r_1}{r_v} \\ \vdots & \ddots & \vdots \\ \frac{r_{v-1}}{r_v} & \cdots & \frac{r_{v-1}}{r_v} \end{array} \right] A(\forall), \)

where \( (\forall) = \left[ \begin{array}{c} \bar{v}_1 \\ \vdots \\ \bar{v}_{v-1} \end{array} \right]. \)

**Proof:** From lemma (3.1), we have

\[
\overline{Q_r - Q_a} = \sum_i \bar{v}_i \bar{Q}_i = \sum_i \bar{v}_i \bar{Q}_i + \bar{v}_v \bar{Q}_v = (\forall) A(\forall) + (\forall) \left( \frac{p_1 \bar{v}_1 + \cdots + p_{v-1} \bar{v}_{v-1}}{\bar{Q}_1 + \cdots + \bar{Q}_{v-1}} \right),
\]

where \( p_i = \frac{r_i}{r_v} \) for \( i = 1, \ldots, v - 1. \)

\[
= (\forall) A(\forall) + (\forall) \left[ \begin{array}{c} p_1 \\ \vdots \\ p_{v-1} \end{array} \right] \left( \begin{array}{c} 1, \ldots, 1 \end{array} \right) \left[ \begin{array}{c} \bar{Q}_1 \\ \vdots \\ \bar{Q}_{v-1} \end{array} \right] = (\forall) A(\forall) + (\forall) \left[ \begin{array}{c} p_1 \\ \vdots \\ p_{v-1} \end{array} \right] A(\forall), \text{ from (3.14).}
\]

\[
= (\forall) / I + \left[ \begin{array}{cccc} p_1 & \cdots & p_1 \\ \vdots & & \vdots \\ p_{v-1} & \cdots & p_{v-1} \end{array} \right] A(\forall),
\]

and the lemma is proved.
We obtain analogously to the intrablock equations for the \( \tilde{v}_i \) given in (3.14), the equations

\[
(Q) = \begin{bmatrix}
Q_1 \\
\vdots \\
Q_{v-1}
\end{bmatrix} = B(\tilde{v})
\]

for the interblock estimates \( \tilde{v}_i, i = 1, \ldots, v - 1, \)

where \( B \) is the matrix such that

\[
(3.15) \quad b_{ii} = \frac{r_i}{k} - \frac{\lambda_{iv} r_i}{k r_v}, \quad i = 1, \ldots, v - 1,
\]

\[
b_{ij} = \frac{\lambda_{ij}}{k} - \frac{\lambda_{iv} r_j}{k r_v}, \quad i \neq j; \quad i, j = 1, \ldots, v - 1.
\]

Also \( \sum_i r_i \tilde{v}_i = \sum_i Q_{ti} = 0. \)

The next lemma is derived from lemma (3.2) in the same way as the preceding one was derived from lemma 3.1.

**Lemma 3.4.** \( \overline{Q_r - Q_a} = k(\overline{a})' I + \begin{bmatrix}
\rho_1 \\
\vdots \\
\rho_{v-1}
\end{bmatrix} B(\overline{\tilde{v}}). \)

**Lemma 3.5.** The matrix \( P = I + \begin{bmatrix}
\rho_1 \\
\vdots \\
\rho_{v-1}
\end{bmatrix} \) occurring in lemmas (3.4) and (3.3), is non-singular.

**Proof:** \( P = \begin{bmatrix}
1 + \rho_1, \rho_1 \\
\vdots \\
\rho_{v-1}, \rho_{v-1}
\end{bmatrix}. \)
Subtracting the first column from the other columns,

\[ |P| = \begin{vmatrix} 1 + \rho_1, -1, -1, \ldots, -1 \\ \rho_2, 1, 0, \ldots, 0 \\ \vdots \\ \rho_{V-1}, 0, \ldots, 0 \end{vmatrix}, \]

the minor of the first element on the principal diagonal being the identity. Adding all the other rows to the first row we have

\[ |P| = 1 + \sum_{i=1}^{V-1} \rho_i, 0, \ldots, 0. \]

Therefore \( P \) is non-singular.

**Lemma 3.6.** An incomplete block design is balanced if and only if there exists a diagonal matrix \( D \) such that \( A = BD \).

**Proof:** The condition is necessary because if the design is balanced, then \( r_i = r \) and \( \lambda_{ij} = \lambda \), for all \( i, j, j \neq i \), and then \( A = \text{diag} \left( \frac{r(k-1)}{k} + \lambda \right) \) and \( B = \text{diag} \left( \frac{r - \lambda}{k} \right) \), so that \( D = \text{diag} \left( \frac{r(k-1)}{k} + \lambda \right) \).

We prove now that the condition is also sufficient.

Adding \( A \) and \( B \) we obtain \( A + B = \text{diag}(r_i) \).

Substitution of \( A = BD \) gives \( B = \text{diag}(r_i) (D + I)^{-1} \),
showing that $B$ is diagonal, since $D + I$ is diagonal.

Hence, $b_{ij} = 0$ for $i, j = 1, \ldots, v - 1; i \neq j$;

i.e., $\lambda_{ij} - \lambda_{iv} \frac{r_{ij}}{r_v} = 0$. (We shall assume henceforth that $i$ and $j$ vary always from 1 to $v - 1$ unless expressly stated to the contrary.)

We obtain from the preceding equation

(3.16) $\frac{\lambda_{ij}}{r_j} = \frac{\lambda_{iv}}{r_v}$

or $\lambda_{ij} = \frac{\sum_{j, j \neq i}^{i} \lambda_{ij}}{r_j} = \frac{r_i(k - 1)}{bk - r_i}$

Hence,

(3.17) $\lambda_{ij} = \frac{r_i r_j (k - 1)}{bk - r_i}$

Interchanging $i$ and $j$, we get $\lambda_{ji} = \frac{r_j r_i (k - 1)}{bk - r_j}$.

Since $\lambda_{ij} = \lambda_{ji}$, we then have $\frac{r_i r_j (k - 1)}{bk - r_i} = \frac{r_j r_i (k - 1)}{bk - r_j}$.

Assuming that $k > 1$, we obtain $r_i = r_j = r$, say.

From (3.17) we then see that $\lambda_{ij} = \frac{r^2 (k - 1)}{bk - r}$ = $\lambda$, say.

From (3.16) we now obtain $\lambda_{iv} = \frac{\lambda r_v}{r}$, so that

(3.18) $\lambda_{1v} = \lambda_{2v} = \ldots = \lambda_{v-1,v} = \lambda = \frac{\lambda}{r} r_v$.

The number of varieties other than the $i^{th}$ occurring in the $r_i = r$ blocks containing the $i^{th}$ variety

$= r(k - 1) = \sum_{j, j \neq i}^{i} \lambda_{ij} = (v - 2)\lambda + \wedge$.  

Therefore,

\[(3.19) \quad A = r(k - 1) - \lambda(v - 2).\]

Similarly, from the blocks containing the \(i\)th variety we obtain \(r_\nu(k - 1) = (v - 1)A\). Hence,

\[(3.20) \quad A = \frac{r_\nu(k - 1)}{v - 1}.\]

From (3.18) and (3.20), \(\frac{\lambda r_\nu}{r} = \frac{r_\nu(k - 1)}{v - 1}\). Hence,

\[(3.21) \quad r(k - 1) = \lambda(v - 1).\]

Substituting in (3.19), we get \(A = \lambda\).

This together with (3.18) shows that \(r = r_\nu\). The sufficiency of the condition \(A = BD\) is thus established.

We now proceed with the proof of theorem 3.3.

**Proof:** Let \((v^*) = (\hat{v})\).

From equations (3.4) and (3.5), this means

\[C \text{ diag}(\xi) (\mathcal{S}) + C \text{ diag}(\eta) (\mathcal{S}) = \text{ diag}(\lambda) C(\mathcal{S}) + \text{ diag}(\mu) C(\mathcal{S}).\]

But, since the \(\mathcal{S}_i\)'s and the \(\mathcal{S}_j\)'s are mutually independent, there can be no linear relation connecting them. Therefore,

\[(3.22) \quad C \text{ diag}(\xi) = \text{ diag}(\lambda) C\]

and \(C \text{ diag}(\eta) = \text{ diag}(\mu) C\).

Using (3.3), we see that the last equation is equivalent to (3.22). From (3.22) we obtain, \(\text{ diag}(\xi) C^{-1} = C^{-1} \text{ diag}(\lambda)\).

That is,

\[(3.23) \quad c^{ij} \xi_i = c^{ij} \lambda_j, \text{ for all } i, j, = 1, \ldots, v - 1.\]

\[\therefore \text{ if } c^{ij} \neq 0, \text{ then } \xi_i = \lambda_j.\]
Also from (3.22), \( c_{ij} \xi_j = c_{ij} \lambda_i \).
But by assumption, \( c_{11}, \ldots, c_{v-1,v-1} \), are all non-zero. Hence,

(3.24) \( \xi_i = \lambda_i \), for all \( i \).

We note that since \( \xi_i = \frac{S_i}{\sigma^2 S_i + \sigma^2 S_j} \),

(3.25) \( \xi_i = \xi_j \) if and only if \( \frac{\sigma^2 S_i}{\sigma^2 S_i} = \frac{\sigma^2 S_j}{\sigma^2 S_j} \).

Now we wish to prove that \( A = BD \) where \( D \) is diagonal, for, by lemma 3.6, this will then ensure that the design is balanced. Since by lemma 3.5, \( D \) is non-singular, this is equivalent to proving that \( PA = PBD \), or that \( PA = \left( kPB \right) \ (D) \).

But \( PA \) is, by lemma 3.3, the matrix of the quadratic form \( Q_r - Q_a \), regarded as a quadratic form in the \( v_i \)'s.

\( kPB \) is by lemma 3.4, the matrix of the quadratic form \( Q_r - Q_a \). Thus we have to prove that

(3.26) \( (the \ matrix \ of \ Q_r - Q_a) = (the \ matrix \ of \ Q_r - Q_a) D \).

However, \( Q_r - Q_a = \frac{\sigma^2 S_i}{\sigma^2 S_i} \), from (1.11).

\[
= \sigma^2 S_i \left( \frac{1}{\sigma^2 S_i} \right) \left( S_i \right) \\
= \sigma^2 \left( v_i \right) \left( C^{-1} \right) \text{diag} \left( \frac{1}{\sigma^2 S_i} \right) \left( C^{-1} v_i \right), \text{ since } \left( v_i \right) = C(S) .
\]
Similarly \( \overline{Q_r - Q_a} = k \sum_i \frac{\sigma^2}{\sigma_{S_i}^2} S_i^2 \), from (1.11)

\[
= k \sigma^2 (\bar{v})(C^{-1}) \text{diag} \left( \frac{1}{\sigma_{S_i}^2} \right) C^{-1}(\bar{v}).
\]

Therefore from (3.26), we have to prove that

\[
(C^{-1}) \text{diag} \left( \frac{1}{\sigma_{S_i}^2} \right) C^{-1} = (C^{-1}) \text{diag} \left( \frac{1}{\sigma_{S_i}^2} \right) C^{-1} E,
\]

where \( E \) is a diagonal matrix.

Or \( \text{diag} \left( \frac{1}{\sigma_{S_i}^2} \right) C^{-1} = \text{diag} \left( \frac{1}{\sigma_{S_i}^2} \right) C^{-1} E \).

Or

\[
(3.27) \left( \frac{1}{\sigma_{S_j}^2} \right) c^{ji} = e_i \frac{1}{\sigma_{S_j}^2} c^{ji}, \text{ for all } i, j.
\]

Now if \( c^{ji} \neq 0 \), we have from (3.23) and (3.24) that

\[
k_j = \lambda_i = \bar{k}_i.
\]

This implies, using (3.25) that \( \frac{\sigma_{S_j}^2}{\sigma_{S_i}^2} = \frac{\sigma_{S_i}^2}{\sigma_{S_j}^2} = e_i \), say,

establishing (3.27) and completing the proof of the fact that if \( (v^*) = (\bar{v}) \), then the design is balanced.

Conversely, let the design be balanced.

Then \( \overline{Q_r - Q_a} = \lambda \bar{\sum_i v_i}^2 = \lambda \bar{\sum_i S_i^2} \) and

\( \overline{Q_r - Q_a} = \mu \bar{\sum_i v_i}^2 = \mu \bar{\sum_i S_i^2} \), from (1.12).

Here the \( v_i \)'s and the \( S_j \)'s are connected by the equation

\[
A \begin{bmatrix} v_1 \\ \vdots \\ v_v \end{bmatrix} = \begin{bmatrix} S_1 \\ \vdots \\ S_{v-1} \\ 0 \end{bmatrix}.
\]
The matrix $A$ is orthogonal with its last row consisting of the same element, $v^{-\frac{1}{2}}$.

\[ \begin{align*}
\vdots \quad [v_1] &= A \begin{bmatrix} S_1 \ \cdots \ \cdots \ S_{v-1} \\ v \end{bmatrix}, \\
\vdots
\end{align*} \]

and the last column of $A$ consists entirely of the same element, $v^{-\frac{1}{2}}$.

\[ \begin{align*}
\therefore \quad 
\sigma_{v_i}^2 &= \text{Var} \left( \frac{v_i}{\sum_{j=1}^{v_i} a_{ij} S_j} \right) = \frac{\sum_{j=1}^{v_i} a_{ij}^2}{\lambda} \sigma_{S_j}^2 = \frac{\sigma_{v_i}^2}{\lambda} \sigma_{S_j}^2 \\
&= \frac{\sigma^2(1 - 1)}{\lambda} = \frac{1}{\lambda} - 1, \quad i = 1, \ldots, v.
\end{align*} \]

Similarly, $\sigma_{v_i}^2 = \frac{\sigma^2(1 - 1)}{\mu}$, $i = 1, \ldots, v$.

Thus, $\frac{\sigma_{v_i}^2}{\sigma_{v_i}^2} = \frac{\lambda \sigma_{v_i}^2}{\mu \sigma_{v_i}^2} = \frac{\sigma_{S_j}^2}{\sigma_{S_j}^2}$.

Substitution in (3.4) and (3.5) yields the result.

Theorem 3.3 has been proved by Sprott\textsuperscript{3} under the assumption that all the $r_i$'s are equal. The method of proof used here is different from that of Sprott, and the result is more general in that the $r_i$'s are not assumed to be all the same. Sprott's method can be employed to yield a proof of theorem 3.3.

CHAPTER IV

We first prove that if $\frac{x}{\sigma^2}$ has a $\chi^2$ distribution with $n$ degrees of freedom, then

(4.1) \[ E\left(\frac{1}{x}\right) = \frac{1}{(n - 2) \sigma^2} \quad \text{provided } n > 2. \]

Proof: \[ E\left(\frac{1}{x}\right) = \frac{1}{n/2} \frac{n}{\sigma} \frac{1}{\Gamma(n/2)} \int_0^\infty \frac{1}{x} e^{-\frac{x}{2\sigma^2}} dx = \frac{1}{n/2} \frac{n}{\sigma} \frac{1}{\Gamma(n/2)} \int_0^\infty x e^{-\frac{x}{2\sigma^2}} dx. \]

But from the formula \[ \int_0^\infty x e^{-\lambda x} dx = \frac{\Gamma(\lambda + 1)}{\lambda}, \quad (\lambda > 0), \]

we have \[ E\left(\frac{1}{x}\right) = \frac{1}{n/2} \frac{n}{\sigma} \frac{1}{\Gamma(n/2)} \frac{\Gamma\left(\frac{n}{2} - 1\right)}{\Gamma\left(\frac{n}{2}\right)} \frac{2}{\sigma^2} \frac{1}{(n - 1) - 1} \]

\[ = \frac{1}{(n - 2) \sigma^2}. \]

Now, let $\theta^* = \frac{1}{(n - r - 2)} \left\{ \sum_{m=1}^{n-r-2} \left( \frac{y_{\beta} - (d_{\beta}/c_{\beta}) x_{\beta}}{\sum z_{\beta}^2} \right)^2 - 1/n - r - 2 \left( \frac{\sum (d_{\beta}/c_{\beta})^2}{\sum z_{\beta}^2} \right)^{1/2} \right\}$, where $\alpha$ is a fixed number between 1 and $r$. 
\[
E(\theta^*) = \frac{n-r-1}{m-1} E\left\{\sum_{p=0}^{E} (y_p - \bar{d}_p x_p)^2\right\} E\left(\frac{1}{\sum_{p}^{E} z_p^2}\right) - \frac{1}{m-1} \sum_{p=0}^{E} \left(\frac{d_p}{c_p}\right)^2.
\]

But \(\frac{1}{\sigma^2} \sum_{p}^{E} z_p^2\) has a \(\chi^2\) distribution with \(n-r\) degrees of freedom. Therefore applying (4.1), we get

\[
\frac{1}{\Sigma z_p^2} = \frac{1}{(n-r-2)\sigma^2}.
\]

Hence, \(E(\theta^*)\) becomes

\[
\frac{n-r-2}{m-1} \left\{(m-s)\sigma^2 + (s-1)\sigma^2 + \sigma^2 \sum_{p=0}^{E} \left(\frac{d_p}{c_p}\right)^2\right\} \frac{1}{(n-r-2)\sigma^2} - \frac{1}{m-1} \sum_{p=0}^{E} \left(\frac{d_p}{c_p}\right)^2
\]

\[
= \frac{\sigma^2}{\sigma^2}.
\]

Thus \(\theta^*\) is an unbiased estimate of \(\theta\).

Consider the estimate \(S^\alpha\) obtained by replacing \(\theta\) by \(\theta^*\) in the maximum likelihood estimate \(\hat{S}^\alpha = \frac{\theta^c x^\alpha + d^\alpha y^\alpha}{c^2 + d^2}\) (cf. 1.14).

\[
S^\alpha = \frac{\theta^* c^\alpha x^\alpha + d^\alpha y^\alpha}{\theta^* c^2 + d^2}.
\]

We note that \(\theta^*, x^\alpha,\) and \(y^\alpha\) are mutually independent.

\[
\text{Var}(x^\alpha | \theta^*) = \text{Var} x^\alpha = \sigma^2 \text{ and } \text{Var}(y^\alpha | \theta^*) = \text{Var} y^\alpha = \sigma^2.
\]

Thus we have

\[
\text{Var}(S^\alpha | \theta^*) = \frac{\theta^* 2c^2 \sigma^2 + d^2 \sigma^2}{\left(\theta^* c^2 + d^2\right)^2} = \frac{(\theta^2 c^2 + d^2 \theta) \sigma^2}{\left(\theta^* c^2 + d^2\right)^2}.
\]

Let \(\frac{1}{\hat{\theta}}\) denote \(\frac{\theta \sigma^2}{c^2 + d^2}\) = \(\text{Var}(S^\alpha)\).

Then

\[
\text{Var}(S^\alpha | \theta^*) - \frac{1}{\hat{\theta}} = \frac{(\theta^2 c^2 + d^2 \theta) \sigma^2}{\left(\theta^2 c^2 + d^2\right)^2} - \frac{\theta \sigma^2}{c^2 + d^2}
\]
From this we obtain

\[ (4.2) \quad \text{Var}(S^*_\alpha) = \frac{1}{\Omega} + \frac{1}{\Omega \theta} \int_{-\infty}^{\infty} \frac{c^2 d^2 (\theta - \theta^*)^2}{(\theta c^2 + d^2)^2} \, d\theta, \]

where \( f(\theta^*); \) is the probability function of \( \theta^*. \)

Further \( E(S^*_\alpha \mid \theta^*) = \frac{\theta^* c^2 S_\alpha + d^2 S_\alpha}{\theta c^2 + d^2} = S_\alpha. \)

Therefore,

\[ (4.3) \quad E(S^*_\alpha) = S_\alpha, \text{ also}. \]

From the manner in which (4.2) and (4.3) have been derived it is clear that any estimate \( \hat{\theta}^* \) could have been used, as long as it is independent of \( x_\alpha \) and \( y_\alpha \), and equations (4.2) and (4.3) would still hold good.
REFERENCES


I, Manavazhi Vijaya Krishna Menon, was born in Palghat, Kerala, India, April 25, 1928. I received my secondary school education in Burma and India, and my undergraduate and graduate training at Madras University, which granted me the Bachelor of Arts (Honours) degree in mathematics in 1948, and the Master of Science degree in mathematics in 1950. I was assistant lecturer in mathematics at University college, Mandalay, Burma, from 1951 to 1956. In September, 1956, I came to Ohio State University to work for the degree of doctor of philosophy in the field of mathematical statistics, and while completing the requirements for the degree, held the position of Assistant Instructor in mathematics.