

ANALYSIS OF TWO-WAY DESIGNS

By J. ROY and K. R. SHAH

Indian Statistical Institute

SUMMARY. Methods are given for analysis of general two-way designs with recovery of information from row and column-contrasts, using additivity of plot and treatment effects and the physical fact of randomisation. Column-balanced designs are discussed in particular with a numerical illustration.

1. INTRODUCTION

Designs for two-way elimination of heterogeneity have been considered by various authors: Bose and Kishen (1939), Fisher (1936), Rao (1943, 1946), Yates (1937), and Youden (1937). A general method of analysis under the so-called fixed effects Normal model was given by Shrikhande (1951). The purpose of this paper is to validate Shrikhande's (1951) results and to derive methods of combined estimation after recovery of information from row and column-contrasts using only the assumption of additivity of plot and treatment effects, and the physical fact of randomisation.

To estimate treatment effects, three sets of contrasts namely, row-contrasts, column-contrasts and interaction-contrasts are introduced and their distribution induced by the randomisation considered. These contrasts are all uncorrelated. But since the contrasts in different sets have different variances, it is not possible to combine them effectively unless the ratios of these variances are known. Therefore, best linear unbiased estimates are obtained first from interaction-contrasts only, in Section 3.1. The equations for estimation turn out to be the same as those obtained by Shrikhande under the Normal model.

The equations for combined estimation after recovery of information from row- and column-contrasts are given in Section 3.2. Methods for estimating the variance ratios that are required in this problem are given in Sections 4.2 and 4.3. The analysis of variance is shown in Section 4.2. Conditions under which a two-way design compares favourably with the corresponding one-way designs are examined in Section 5 and the relative efficiency-factors worked out.

Though the analysis requires rather heavy computations in the general case, it is shown in Section 6 that if the columns of the design, ignoring rows, form a balanced incomplete block design the analysis is much simpler. The special case where columns are balanced and rows partially balanced is discussed in full and a numerical example is given in Section 8. All derivations are given in Section 7.

2. PRELIMINARIES

2.1. *The additive model.* Suppose there are mn plots or experimental units (eu's) on which a comparative trial involving v treatments is to be carried out. The eu's are arranged in a $m \times n$ two-way classification, so that each eu is determined by a pair of co-ordinates (t, u) $t = 1, 2, \dots, m; u = 1, 2, \dots, n$. With the (t, u) -th eu is associated a number x_{tu} to be called the plot effect and we assume that if the k -th treatment is applied on the (t, u) -th eu, the 'yield' would be $x_{tu} + \theta_k$ where the parameter θ_k is to be regarded as the effect of the k -th treatment; $k = 1, 2, \dots, v$. This is the so-called *additive* or *no-interaction* model. The purpose of the experiment is to compare the θ_k 's.

We now define

$$\begin{aligned} \mu &= \Sigma x_{tu} / mn, \text{ the general mean,} \\ \sigma_1^2 &= \Sigma \rho_t^2 / (m-1), \text{ the between-row variance,} \\ \sigma_2^2 &= \Sigma \gamma_u^2 / (n-1), \text{ the between-column variance,} \\ \sigma_3^2 &= \Sigma \Sigma \eta_{tu}^2 / [(m-1)(n-1)], \text{ the interaction variance,} \end{aligned} \quad \dots (2.1)$$

where $\rho_t = \frac{1}{n} \Sigma_u x_{tu} - \mu$, $\gamma_u = \frac{1}{m} \Sigma_t x_{tu} - \mu$ and $\eta_{tu} = x_{tu} - \rho_t - \gamma_u - \mu$. We shall write

$$w_i = \sigma_0^2 / \sigma_i^2 \quad i = 1, 2 \quad \dots (2.2)$$

for the ratios of the variances.

2.2. *The design.* The treatments are allocated to the eu's in the following manner. First a two-way design, that is, an arrangement of the v treatments in m rows and n columns is taken. The design is thus completely characterized by the numbers ϵ_{ijk} , $i = 1, 2, \dots, m; j = 1, 2, \dots, n; k = 1, 2, \dots, v$ where $\epsilon_{ijk} = 1$ if the k -th treatment occurs in the intersection of the i -th row and the j -th column of the design and $\epsilon_{ijk} = 0$, otherwise. The k -th treatment thus occurs in $m_{ki} = \sum_{j=1}^n \epsilon_{ijk}$ positions in the i -th row and in $n_{kj} = \sum_{i=1}^m \epsilon_{ijk}$ positions in the j -th column. We shall restrict ourselves to *equi-replicate* designs, that is to those designs, where each treatment occurs altogether in r positions. Thus $\sum_i m_{ki} = \sum_j n_{kj} = r$ and of course, $\sum_k m_{ki} = n$, $\sum_k n_{kj} = m$. We shall call $M = ((m_{ki}))$ and $N = ((n_{kj}))$ the row incidence-matrix and the column incidence-matrix respectively. The rows and the columns of the design are then allotted to the two ways of classification of the eu's independently and at random.

2.3. *Consequences of randomisation.* Let us denote the yield of the eu corresponding to the i -th row and the j -th column of the design by y_{ij} . The randomisation procedure ensures that

$$E(y_{ij}) = \mu + \sum_{k=1}^v \epsilon_{ijk} \theta_k. \quad \dots (2.3)$$

ANALYSIS OF TWO-WAY DESIGNS

and that

$$\text{cov}(y_{ij}, y_{i'j'}) = \left(\delta_{ii'} - \frac{1}{m}\right)\sigma_1^2 + \left(\delta_{jj'} - \frac{1}{n}\right)\sigma_2^2 + \left(\delta_{ii'} - \frac{1}{m}\right)\left(\delta_{jj'} - \frac{1}{n}\right)\sigma_3^2 \dots \quad (2.4)$$

where $\delta_{ii'}$ is the Kronecker symbol, $\delta_{ii'} = 1(0)$ if $i = i'$ ($i \neq i'$).

2.4. *A linear transformation.* Since the y_{ij} 's are correlated, it is convenient to make a linear transformation and obtain uncorrelated random variables. For this purpose, we use the following definitions. A linear function of the form $l = \sum \sum l_{ij} y_{ij}$ is said to be a *contrast* if $\sum \sum l_{ij} = 0$. A contrast l is said to belong to rows, or simply called a *row-contrast* if $l_{i1} = l_{i2} = \dots = l_{in}$ holds for $i = 1, 2, \dots, m$. Similarly, a contrast l is said to be a *column-contrast* if $l_{1j} = l_{2j} = \dots = l_{mj}$ holds for $j = 1, 2, \dots, n$. A contrast l is said to belong to interaction or simply called an *interaction-contrast* if $\sum_i l_{ij} = 0$ for $j = 1, 2, \dots, n$ and $\sum_j l_{ij} = 0$ for $i = 1, 2, \dots, m$. A contrast l is said to be *normalised* if $\sum \sum l_{ij}^2 = 1$. Two contrasts l and $l' = \sum \sum l'_{ij} y_{ij}$ are said to be *orthogonal* if $\sum \sum l_{ij} l'_{ij} = 0$ holds.

If then we make a linear transformation from y_{ij} 's to (i) $G^* = G/\sqrt{mn}$, where $G = \sum \sum y_{ij}$ is the grand total, (ii) a set of $(m-1)$ mutually orthogonal normalised row-contrasts (iii) a set $(n-1)$ mutually orthogonal normalised column-contrasts, and (iv) a set of $(m-1)(n-1)$ mutually orthogonal normalised interaction-contrasts, it can then be shown as in Section (7.1) that the transformation is orthogonal and that these transformed variables are uncorrelated; the variance of any normalised row-contrast being $m\sigma_1^2$, the variance of any normalised column-contrast being $m\sigma_2^2$ and that of any normalised interaction-contrast being σ_3^2 . Since the expectation of each contrast is a linear function of the θ_k 's, the method of least-squares can be used for purposes of estimation.

2.5. *Notation.* We shall write R_i for the total yield of the i -th row, C_j for that of the j -th column, and T_k for that of the k -th treatment; thus

$$R_i = \sum_{j=1}^n y_{ij}, C_j = \sum_{i=1}^m y_{ij} \text{ and } T_k = \sum_{i=1}^m \sum_{j=1}^n \epsilon_{ijk} y_{ij}.$$

We shall use the matrix notations: $\mathbf{R} = (R_1, R_2, \dots, R_m)$, $\mathbf{C} = (C_1, C_2, \dots, C_n)$, $\mathbf{T} = (T_1, T_2, \dots, T_p)$ and $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$. Also, a matrix of the form $p \times q$ with all elements unity will be denoted by \mathbf{E}_{pq} .

If \mathbf{A} is a positive semi-definite matrix of form $a \times a$ and rank b , it has b positive latent roots, say α_i , $i = 1, 2, \dots, b$. Let ξ_i of the form $1 \times n$ be a latent vector of \mathbf{A} corresponding to the latent root α_i , $i = 1, 2, \dots, b$ such that $\xi_i \xi_j' = \delta_{ij}$. Then the matrix $\mathbf{A}^* = \sum_{i=1}^b \frac{1}{\alpha_i} \xi_i \xi_i'$ will be called a *pseudo-inverse* of \mathbf{A} , in the sense of Rao (1955).

3. ESTIMATION OF TREATMENT EFFECTS

3.1. *Estimation from interaction-contrasts.* Since row-contrasts, column-contrasts and interaction-contrasts have different variances, it is not convenient to use them simultaneously for estimation of treatment effects in an efficient way unless the relative magnitudes of these variances are known. We shall, therefore, consider

first the problem of estimation from interaction-contrasts only. Recovery of information provided by row-contrasts and column-contrasts will be taken up in Section 3.2.

As we have pointed out in Section 2.4, any set of $(m-1) \times (n-1)$ mutually orthogonal normalised interaction-contrasts are mutually uncorrelated and each of them has the same variance σ_0^2 . Also, the expectation of each is a linear function of the θ_i 's. Consequently, the method of least-squares can be used to derive linear unbiased estimators with minimum variance ('best' estimators) of linear functions of treatment effects. As will be shown later in Section 7.2, the method of least-squares gives the equations:

$$\theta K = Q \quad \dots (3.1)$$

where the elements of

$$Q = T - \frac{1}{n} R M' - \frac{1}{m} C N' + \frac{rG}{mn} E_{10} \quad \dots (3.2)$$

are called the *adjusted yields* of the treatments and

$$K = r I - \frac{1}{n} M M' - \frac{1}{m} N N' + \frac{r^2}{mn} E_{00} \quad \dots (3.3)$$

will be called the *coefficient-matrix* of the two-way design.

Since $KE_{01} = 0$, $\text{rank}(K) \leq v-1$. A two-way design will be said to be *doubly connected* if its coefficient matrix K is of rank $(v-1)$. In whatever follows, we shall assume that the two-way design is doubly connected.

It is well known from the theory of least-squares (see Rao, 1952) that any linear parametric function of the form $\theta = \sum_{k=1}^v l_k \theta_k$ with $\sum_{k=1}^v l_k = 0$ admits linear unbiased estimators, and amongst them the one with minimum variance is $T = \sum l_k t_k$ where $t = (t_1, t_2, \dots, t_v)$ is any solution of (3.1). To obtain the variance of T express it in the alternative form $T = \sum_{k=1}^v m_k Q_k$ and then $V(T) = (\sum_{k=1}^v l_k m_k) \sigma_0^2$.

It may be noted that the equations (3.1) are the same as obtained by Shrikhande (1951) from the so-called 'Normal' model. The present approach demonstrates the robustness of this procedure and is aesthetically more satisfying to the authors.

3.2. *Recovery of information from row-contrasts and column-contrasts.* In the previous section, we had simply thrown away the row-contrasts and the column-contrasts. If the ratios $w_i = \sigma_0^2/\sigma_i^2$, $i = 1, 2$ are known, the method of weighted least-squares can be applied on all the three sets of contrasts simultaneously. If the weight for the normalised interaction-contrasts is taken as unity, the weight for normalised row-contrasts will be w_1/n and that for normalised column-contrasts will be w_2/m . As will be shown in Section 7.2, the method of weighted least-squares now gives the equations:

$$\theta \bar{K} = \bar{Q} \quad \dots (3.4)$$

ANALYSIS OF TWO-WAY DESIGNS

where

$$\bar{Q} = T - \frac{1}{n + \Delta_1} R M' - \frac{1}{m + \Delta_2} C N' + \frac{rG}{mn} \left(1 - \frac{\Delta_1}{n + \Delta_1} - \frac{\Delta_2}{m + \Delta_2} \right) E_{1r} \dots (3.5)$$

$$\bar{K} = rI - \frac{1}{n + \Delta_1} M M' - \frac{1}{m + \Delta_2} N N' + \frac{r^2}{mn} \left(1 - \frac{\Delta_1}{n + \Delta_1} - \frac{\Delta_2}{m + \Delta_2} \right) E_{rr} \dots (3.6)$$

and
$$\Delta_1 = \frac{nr_1}{n - r_1}, \Delta_2 = \frac{mr_2}{m - r_2} \dots (3.7)$$

Then the best estimator of $\theta = \sum_{k=1}^v l_k \theta_k$, where $\sum_{k=1}^v l_k = 0$, is given by $\bar{T} = \sum_{k=1}^v l_k \bar{t}_k$ where $l = (l_1, l_2, \dots, l_k)$ is any solution of (3.4). The variance of \bar{T} is most easily obtained by writing it in the form $\sum_{k=1}^v \bar{m}_k \bar{Q}_k$ and then $V(\bar{T}) = (\sum_{k=1}^v l_k \bar{m}_k) \sigma_{\bar{Q}}$.

Generally, however, the parameters Δ_1 and Δ_2 would be unknown and estimates D_1 and D_2 for them may have to be substituted. In this case, of course, the 'bestness' of \bar{T} as an estimator of θ would no longer hold, but conceivably if D_1, D_2 are at all good estimators of Δ_1, Δ_2 ; T might yet be better than T . Methods for estimating the Δ 's are given in Section 4.2.

4. ANALYSIS OF VARIANCE

4.1. *Analysis of variance of interaction-contrasts.* To estimate $\sigma_{\bar{Q}}^2$ and to carry out an omnibus test of significance of treatment differences, the analysis of variance is to be done as shown in the following table.

TABLE 4.1. ANALYSIS OF VARIANCE

source	degrees of freedom	sum of squares
(1)	(2)	(3)
rows (unadjusted)	$m - 1$	$SS_R^* = \frac{1}{n} \sum_{i=1}^m R_i^2 - \frac{G^2}{mn}$
columns (unadjusted)	$n - 1$	$SS_C^* = \frac{1}{m} \sum_{j=1}^n C_j^2 - \frac{G^2}{mn}$
treatments (adjusted for rows and columns)	$v - 1$	$SS_T = \sum_{k=1}^v Q_k l_k$
error	$r = (m-1)(n-1) - (v-1)$	$SS_e = SS_T - SS_T$
interaction	$(m-1)(n-1)$	$SS_I = SS_T - SS_R^* - SS_C^*$
total	$mn - 1$	$SS_T = \sum_{i=1}^m \sum_{j=1}^n Y_{ij}^2 - \frac{G^2}{mn}$

An unbiased estimator of σ_0^2 is provided by the error mean square $MS_0 = SS_0/v$. To examine the significance of treatment differences, one may use the customary ratio of mean squares MS_t/MS_0 where $MS_t = SS_t/(v-1)$ is the treatment mean square. The sampling distribution of this statistic under randomisation, when the treatment effects are identical, is usually approximated by the Snedecor F -distribution with $(v-1)$ and v degrees of freedom. The accuracy of this approximation is under investigation.

4.2. *Estimation of weights for use in recovery of information.* An estimate of $\Delta_1 = n\sigma_0^2/(n\sigma_1^2 - \sigma_0^2)$ can be obtained as in the case of incomplete block designs by considering the expectation of SS_R , the adjusted row sum of squares. To compute SS_R one has to carry out another analysis of variance ignoring rows. Let t_i be any solution of the following equations in θ

$$\theta K_1 = Q_1 \quad \dots (4.1)$$

$$\text{where} \quad Q_1 = T - \frac{1}{m} CN' \text{ and } K_1 = rI - \frac{1}{m} NN'. \quad \dots (4.2)$$

$$\text{Then} \quad SS_R = SS_R^* + SS_{tr} - Q_1 t_i. \quad \dots (4.3)$$

We shall show in Section 7.3 that writing $MS_R = SS_R/(m-1)$ for the adjusted row mean square, the expectation of MS_R is given by

$$E(MS_R) = [1 + (n-a_1)/\Delta_1] \sigma_0^2 \quad \dots (4.4)$$

$$\text{where} \quad a_1 = tr K_1^+ M M' / (m-1) \quad \dots (4.5)$$

in which tr denotes the trace of a matrix and K_1^+ is a pseudo-inverse of the matrix K_1 .

Consequently, we can take

$$D_1 = \frac{(n-a_1)MS_0}{MS_R - MS_0} \quad \dots (4.6)$$

as an estimator of Δ_1 in the sense that the ratio of the expectations of the numerator and the denominator of D_1 is equal to Δ_1 . An estimator D_2 of Δ_2 can be obtained in like manner.

4.3. *A positive-definite estimator of σ_1^2 .* If (4.4) is used for estimation of σ_1^2 , the estimate might on occasions turn out to be negative. We propose here an alternative procedure which is applicable to certain types of designs and has the advantage of providing a positive-definite estimator of σ_1^2 .

As will be shown in Section 7.2, the least-squares equations for estimation of θ_i 's from row-contrasts are

$$\theta \left(MM' - \frac{r^2}{m} E_{rr} \right) = \left(RM' - \frac{rQ}{m} E_{is} \right). \quad \dots (4.7)$$

ANALYSIS OF TWO-WAY DESIGNS

If t^* is any solution of the above equations in θ and if

$$p = \text{rank} \left(M M' - \frac{r^2}{m} E_{mn} \right) < m-1 \quad \dots (4.8)$$

the residual sum of squares in the analysis of variance of normalised row-contrasts can be used to estimate σ_1^2 . Thus,

$$s_1^2 = \frac{\left(\sum_{i=1}^m R_i^2 - (r^2/m) \right) - R M' t^*}{n^2(m-1-p)} \quad \dots (4.9)$$

is a positive-definite unbiased estimator of σ_1^2 . The corresponding estimator of Δ_1 is $D_1^* = n M S_0 / (n s_1^2 - M S_0)$.

5. EFFICIENCY

It is known [Kempthorne (1956), Roy (1957)] that the average variance of interaction estimators of differences of the type $\theta_k - \theta_{k'}$, is $2\sigma_0^2/h(K)$ where K is the coefficient-matrix of the design and $h(K)$ denotes the harmonic mean of the positive latent roots of K . If instead of the two-way design, a one-way design using columns as blocks were used, the average variance of intra-block estimates would then be $2 \left[\left(1 - \frac{1}{n} \right) \sigma_0^2 + \sigma_1^2 \right] / h(K_1)$ where $K_1 = rI - \frac{1}{m} N N'$. As a measure of the efficiency of the two-way design in comparison with the one-way (column) design, we propose the ratio of the reciprocals of these average variances. This turns out to be

$$E = e\phi \quad \dots (5.1)$$

where $e = h(K)/h(K_1)$ will be called the *efficiency-factor* of the two-way design relative to the one-way design using columns as blocks and $\phi = 1 + \frac{1}{\Delta_1}$.

It will be shown in Section 7.4 that the relative efficiency-factor $e \leq 1$. Consequently, the two-way design is effective only if $\phi > 1/e$. This parameter ϕ can be estimated by substituting the estimate of Δ_1 as obtained in Section 4.

6. TWO-WAY DESIGNS WITH COLUMN BALANCE

A two-way design will be said to have column balance if each treatment occurs in a column at most once, and any pair of treatments occur together in the same number, say λ , of columns; or in other words if the columns of the design regarded as blocks form a Balanced Incomplete Block Design. A column balanced design is said to be a Youden Square if the row incidence-matrix $M = E_{mn}$ and an extended Youden Square if $M = pE_{mn}$ where p is a positive integer, $p \geq 2$.

Shrikhande (1951) claims that all known column-balanced designs can be arranged in rows in such a way that (i) a partially balanced association scheme with two associate classes can be imposed on the treatments and (ii) the m_{21} 's satisfy:

$$\sum_i m_{21} m_{2'1} = \begin{cases} \mu_0 & \text{if } k = k' \\ \mu_u & \text{if } k \neq k' \text{ are } u\text{-th associates; } u = 1, 2. \end{cases} \quad \dots (6.1)$$

For a definition of a partially balanced association scheme, the reader is referred to Bose and Shimamoto (1952).

If $\mu_1 \neq \mu_2$, the designs satisfying (6.1) are said to belong to the class Y_1 (Shrikhande, 1951).

The analysis of designs belonging to the class Y_1 is particularly easy, being similar to that of partially balanced incomplete block designs with two associate classes. The analysis based on only the interaction-contrasts under the 'Normal' model is given by Shrikhande (1951); here we give the complete analysis including recovery of information from row-contrasts and column-contrasts.

For the parameters of the partially balanced association scheme, we shall use, the standard notations $p_{jk}^i, n_i; i, j, k = 1, 2$. Let now,

$$\begin{aligned} a &= \frac{\lambda v}{m} - \frac{1}{n} (\mu_0 - \mu_2) \\ b &= \frac{1}{n} (\mu_1 - \mu_2) \quad \dots (6.2) \\ c &= p_{11}^1 - p_{11}^2 \\ d &= n_1 - p_{11}^2 \end{aligned}$$

Then a solution of the equations (3.1) for a design of the class Y_1 turns out to be

$$t_k = [A Q_k + b S_1(Q_k)]/D \quad \dots (6.3)$$

where S_1 denotes summation over first associates and

$$\begin{aligned} A &= a + bc \\ D &= aA - b^2d. \end{aligned} \quad \dots (6.4)$$

Since the variances of estimates of treatment differences are given by

$$V(t_k - t_{k'}) = \begin{cases} [2(A-b)/D] \sigma_0^2, & \text{if } k, k' \text{ are first associates} \\ [2A/D] \sigma_0^2, & \text{otherwise} \end{cases}$$

the relative efficiency factor of this design turns out to be

$$e = \frac{m(v-1)D}{\lambda v[(v-1)\bar{a} - n_1 \bar{b}]} \quad \dots (6.5)$$

For combined estimation, a solution of the equations (3.4) is

$$i_k = [\bar{A} \bar{Q}_k + \bar{b} S_1(\bar{Q}_k)]/\bar{D} \quad \dots (6.6)$$

where

$$\begin{aligned} \bar{A} &= \bar{a} + \bar{b}c \\ \bar{D} &= \bar{a}\bar{A} - \bar{b}^2d \end{aligned} \quad \dots (6.7)$$

ANALYSIS OF TWO-WAY DESIGNS

and
$$\bar{a} = \frac{\lambda v + r \Delta_2}{m + \Delta_2} - \frac{1}{n + \Delta_1} (\mu_0 - \mu_2) \quad \dots (6.8)$$

$$\bar{b} = \frac{1}{n + \Delta_1} (\mu_1 - \mu_2).$$

To estimate Δ_1 we need a solution of (4.1) which in this case turns out to be

$$t_{1k} = \frac{m}{\lambda r} Q_{1k}$$

where Q_{1k} is the k -th element of Q_1 defined by (4.2). An estimate of Δ_1 is obtained by putting

$$a_1 = \frac{m\mu_0 - r^2}{\lambda(m-1)} \quad \dots (6.9)$$

in the expression (4.6) for D_1 .

Similarly to estimate Δ_2 we need MS_O the adjusted column mean square given by

$$(n-1) MS_O = SS_O^* + SS_n - Q_2 t_2 \quad (6.10)$$

t_2 being any solution for θ in

$$\theta K_2 = Q_2 \quad \dots (6.11)$$

where
$$Q_2 = T - \frac{1}{n} RM' \quad \text{and} \quad K_2 = rI - \frac{1}{n} MM'. \quad \dots (6.12)$$

In this case $t_2 = (t_{21}, \dots, t_{22})$ is given by

$$t_{2k} = [A' Q_{2k} + bS(Q_{2k})] / D' \quad \dots (6.13)$$

where $A' = A + (r-\lambda)/m$ and $D' = [a + (r-\lambda)/m] A' - b^2 d$. An estimate of Δ_2 is given by

$$D_2 = \frac{(m-a_1)MS_O}{MS_O - MS_0} \quad \dots (6.14)$$

where
$$a_2 = \frac{(r-\lambda)[(v-1)A' - \pi_1 b]}{(m-1)D'}. \quad \dots (6.15)$$

The derivation of (6.9) and (6.15) is given in Section 7.5.

7. DERIVATION OF RESULTS

In this section we prove some of the results stated earlier.

7.1. Some lemmas.

Lemma 1: If $l_{ij}^{\alpha\beta}$, $i, \alpha = 1, 2, \dots, m; j, \beta = 1, 2, \dots, n$ are real numbers chosen to satisfy: $(n) \sum_{ij} l_{ij}^{\alpha\beta} l_{ij}^{\alpha'\beta'} = \delta_{\alpha\alpha'} \delta_{\beta\beta'}$, where δ is the Kronecker symbol,

(b) $I_{ij}^{(\alpha\alpha)} = 1/\sqrt{mn}$, (c) $I_{ij}^{(\alpha\alpha')} = I_{ij}^{(\alpha\alpha)}$, $\alpha = 1, 2, \dots, m-1$ and (d) $I_{ij}^{(\beta\beta)} = I_{ij}^{(\beta\beta)}$, $\beta = 1, 2, \dots, n-1$ then we have:

$$\sum_{\alpha=1}^{m-1} I_{ij}^{(\alpha\alpha)} I_{i'j'}^{(\alpha\alpha')} = \frac{1}{n} \left(\delta_{ii'} - \frac{1}{m} \right)$$

$$\sum_{\beta=1}^{n-1} I_{ij}^{(\beta\beta)} I_{i'j'}^{(\beta\beta')} = \frac{1}{m} \left(\delta_{jj'} - \frac{1}{n} \right) \quad \dots (7.1)$$

$$\sum_{\alpha=1}^{m-1} \sum_{\beta=1}^{n-1} I_{ij}^{(\alpha\beta)} I_{i'j'}^{(\alpha\beta')} = \left(\delta_{ii'} - \frac{1}{m} \right) \left(\delta_{jj'} - \frac{1}{n} \right)$$

and

$$\sum_{ij} \sum_{i'j'} I_{ij}^{(\alpha\beta)} I_{i'j'}^{(\alpha'\beta')} \delta_{ii'} = \begin{cases} n \delta_{\alpha\alpha'}, & \text{when } \beta = \beta' = n \\ 0 & \text{otherwise} \end{cases}$$

$$\sum_{ij} \sum_{i'j'} I_{ij}^{(\alpha\beta)} I_{i'j'}^{(\alpha'\beta')} \delta_{jj'} = \begin{cases} m \delta_{\beta\beta'}, & \text{when } \alpha = \alpha' = m \\ 0 & \text{otherwise} \end{cases} \quad \dots (7.2)$$

$$\sum_{ij} \sum_{i'j'} I_{ij}^{(\alpha\beta)} I_{i'j'}^{(\alpha'\beta')} \delta_{ii'} \delta_{jj'} = \delta_{\alpha\alpha'} \delta_{\beta\beta'}$$

These follow easily from the properties of orthogonal matrices.

Lemma 2: Let $Z_{\alpha\beta} = \sum_{ij} I_{ij}^{(\alpha\beta)} y_{ij}$. Then

$$E(Z_{\alpha\beta}) = \begin{cases} \frac{1}{\sqrt{mn}} (m\mu + r \sum_k \theta_k) & \text{when } \alpha = m, \beta = n \\ \sum_k a_k^{(\alpha\beta)} \theta_k & \text{otherwise} \end{cases} \quad \dots (7.3)$$

where $a_k^{(\alpha\beta)} = \sum_{ij} I_{ij}^{(\alpha\beta)} \epsilon_{ijk}$

Also, these $Z_{\alpha\beta}$'s are mutually uncorrelated and they have variances given by

$$V(z_{\alpha\beta}) = \begin{cases} 0 & \text{if } \alpha = m; \beta = n \\ n\sigma_1^2 & \text{if } \beta = n; \alpha = 1, 2, \dots, m-1 \\ m\sigma_2^2 & \text{if } \alpha = m; \beta = 1, 2, \dots, n-1 \\ \sigma_3^2 & \text{if } \alpha = 1, 2, \dots, m-1; \beta = 1, 2, \dots, n-1. \end{cases} \quad \dots (7.4)$$

These results are obtained by direct computation using the expectations and covariances of y_{ij} 's given by (2.3) and (2.4) and the properties of $I_{ij}^{(\alpha\beta)}$'s given by (7.2).

ANALYSIS OF TWO-WAY DESIGNS

Lemma 3: With $Z_{\alpha\beta}$ and $a_k^{(\alpha\beta)}$'s as defined in Lemma 2, we have

$$Q_k = \sum_{\alpha=1}^{m-1} Z_{\alpha n} a_k^{(\alpha n)} = \frac{1}{n} \sum_{i=1}^n m_{ki} R_i - \frac{rQ}{mn}$$

$$Q_k^* = \sum_{\beta=1}^{n-1} Z_{m\beta} a_k^{(m\beta)} = \frac{1}{m} \sum_{j=1}^n n_{kj} C_j - \frac{rQ}{mn} \quad \dots (7.5)$$

$$Q_k = \sum_{\alpha=1}^{m-1} \sum_{\beta=1}^{n-1} Z_{\alpha\beta} a_k^{(\alpha\beta)} = T_k - \frac{1}{n} \sum_{i=1}^n m_{ki} R_i - \frac{1}{m} \sum_{j=1}^n n_{kj} C_j + \frac{rQ}{mn}$$

$$\gamma_{kk'} = \sum_{\alpha=1}^{m-1} a_k^{(\alpha n)} a_{k'}^{(\alpha n)} = \frac{1}{n} \sum_{i=1}^n m_{ki} m_{k'i} - \frac{r^2}{mn}$$

$$\gamma_{kk'}^* = \sum_{\beta=1}^{n-1} a_k^{(m\beta)} a_{k'}^{(m\beta)} = \frac{1}{m} \sum_{j=1}^n n_{kj} n_{k'j} - \frac{r^2}{mn} \quad \dots (7.6)$$

$$\gamma_{kk'} = \sum_{\alpha=1}^{m-1} \sum_{\beta=1}^{n-1} a_k^{(\alpha\beta)} a_{k'}^{(\alpha\beta)} = r\delta_{kk'} - \frac{1}{n} \sum_{i=1}^n m_{ki} m_{k'i} - \frac{1}{m} \sum_{j=1}^n n_{kj} n_{k'j} + \frac{r^2}{mn}.$$

These results are obtained by direct computation using (7.1) and (7.2).

7.2. *Derivation of least-square equations.* According to the method of weighted least-squares, we have to minimise $\sum \frac{1}{V(Z_{\alpha\beta})} [Z_{\alpha\beta} - E(Z_{\alpha\beta})]^2$, where the summation is over all values of α, β except $\alpha = m, \beta = n$. On multiplication by σ_0^2 , this reduces to minimising

$$L = \sum_{\alpha=1}^{m-1} \sum_{\beta=1}^{n-1} [Z_{\alpha\beta} - E(Z_{\alpha\beta})]^2 + \frac{rQ}{n} \sum_{\alpha=1}^{m-1} [Z_{\alpha n} - E(Z_{\alpha n})]^2 + \frac{rQ}{m} \sum_{\beta=1}^{n-1} [Z_{m\beta} - E(Z_{m\beta})]^2.$$

Equating the partial derivative of L with respect to θ_k ($k = 1, 2, \dots, v$) to zero, we get the equations:

$$\sum_{k'=1}^v \left(\gamma_{kk'} + \frac{rQ}{n} \gamma_{kk'}^* + \frac{rQ}{m} \gamma_{kk'}^* \right) \theta_{k'} = Q_k + \frac{rQ}{n} Q_k^* + \frac{rQ}{m} Q_k^* \quad \dots (7.7)$$

$$k = 1, 2, \dots, v$$

where Q_k and $\gamma_{kk'}$, etc., are given by (7.5) and (7.6). This, in matrix notation is our equation (3.4) for combined estimation. To obtain the equation (3.1) for estimation from interaction-contrasts only, set $w_1 = w_2 = 0$ in (7.7). Similarly if we want estimates from row-contrasts only, the equations would be $\sum_{k'=1}^v \gamma_{kk'}^* \theta_{k'} = Q_k^*$ which, in matrix notation, is our equation (4.7).

7.3. *Expectation of the adjusted row sum of squares.* Since the adjusted row sum of squares SS_R defined by (4.3) is invariant under the transformation $y'_{ij} = y_{ij} - \sum_{k=1}^v \epsilon_{ijk} \theta_k$, its distribution and therefore its expectation does not involve θ_k 's. Consequently in computing $E(SS_R)$ we can ignore the terms involving θ_k 's. Now, since

$$E(SS_R^*) = (n-1)n\sigma_1^2 + \text{terms in } \theta_k\text{'s}$$

$$\text{and } E(SS_{tr}) = (v-1)\sigma_0^2 + \text{terms in } \theta_k\text{'s} \quad \dots (7.8)$$

it follows from (4.3) that all that we need now to compute $E(SS_R)$ is $E(Q_1 t_1)$ where Q_1, t_1 are defined by (4.1) and (4.2). Let K_1^* be a pseudo-inverse of the matrix K_1 defined by (4.2), so that a particular solution of (4.1) is $t_1 = Q_1 K_1^*$. Hence

$$\begin{aligned} E(Q_1 t_1) &= E(Q_1 K_1^* Q_1) = E(\text{tr}(K_1^* Q_1' Q_1)) \\ &= \text{tr}(K_1^* E(Q_1' Q_1)) \\ &= \text{tr} K_1^* D(Q_1) + \text{terms in } \theta_k\text{'s} \quad \dots (7.9) \end{aligned}$$

where $D(Q_1)$ stands for the dispersion matrix of Q_1 . To compute $D(Q_1)$ we express Q_1 in the form $Q_1 = Q + \frac{1}{n} R \left(M' - \frac{rE_{nn}}{m} \right)$. Since the elements of Q are interaction-contrasts and those of $\frac{1}{n} R \left(M' - \frac{rE_{nn}}{m} \right)$ are row-contrasts, these are uncorrelated, and therefore $D(Q_1) = D(Q) + D \left[\frac{1}{n} R \left(M' - \frac{rE_{nn}}{m} \right) \right]$. Since $D(Q) = K \sigma_0^2$ and $D(R) = \left(I - \frac{E}{m} \right) n^2 \sigma_1^2$; we get on simplification

$$D(Q_1) = \left(K_1 + \frac{r^2 E_{nn}}{mn} \right) \sigma_0^2 + MM' \left(\sigma_1^2 - \frac{\sigma_0^2}{n} \right). \quad \dots (7.10)$$

Using (7.8), (7.9) and (7.10), we get finally

$$\begin{aligned} E(SS_R) &= E(SS_R^*) + E(SS_{tr}) - E(Q_1 t_1) \\ &= [n(m-1) - \text{tr} K_1^* MM'] \sigma_1^2 + (\text{tr} K_1^* MM') \sigma_0^2 / n \end{aligned}$$

from which (4.4) follows.

7.4. *Relative efficiency factor.* To prove that the relative efficiency factor e defined in Section 5 cannot exceed unity, we need the following result in matrix theory.

Lemma: If A and B are positive-definite matrices of the same order and C = A - B is positive definite (semi-definite), then D = B^{-1} - A^{-1} is positive definite (semi-definite).

ANALYSIS OF TWO-WAY DESIGNS

Proof: We observe that if P and Q are symmetric matrices of the same order and P is positive definite, a necessary and sufficient condition for Q to be positive definite (semi-definite) is that the roots of the determinantal equation $|Q - \lambda P| = 0$ are all positive (non-negative). Now, the determinantal equation $|C - \lambda A| = 0$ is equivalent to the equation $|D - \lambda B^{-1}| = 0$. This follows by pre- and post-multiplying the former equation by A^{-1} and B^{-1} respectively. Also, C being positive definite (semi-definite) implies that the root λ are all positive (non-negative) which in turn implies that D is positive definite (semi-definite).

Next consider a $(v-1) \times v$ matrix P satisfying $PP' = I$ and $P'P = I - \frac{E}{v}$. Since our design is doubly connected it follows that $A = PK_1P'$ and $B = PKP'$ are both positive definite and $A - B = \frac{1}{n}PMM'P'$ is positive definite or semi-definite.

$$\begin{aligned} \text{Now } H(K_1) - H(K) &= H(A) - H(B) = (v-1) \left(\frac{1}{trA^{-1}} - \frac{1}{trB^{-1}} \right) \\ &= \frac{(v-1)tr(B_1^{-1} - A^{-1})}{(trA^{-1})(trB^{-1})} > 0 \end{aligned}$$

from which it follows that $e < 1$.

7.5. Estimation of weights for two-way designs with column-balance.

Since in this case $K_1 = \frac{\lambda v}{n} \left[I - \frac{E}{v} \right]$, we easily get $K_1' = \frac{m}{\lambda v} \left[I - \frac{E}{v} \right]$. Hence $K_1' MM' = \frac{m}{\lambda v} MM' - \frac{mnr}{\lambda v^2} E$ from which (6.9) follows very easily using (4.5).

Since $NN' = (r-\lambda)I + \lambda E$, we have $trK_1'NN' = (r-\lambda)trK_1'$. Now if $t = Q_1Y$ be any solution in θ of the equations (6.11), it can be shown that $trK_1't = tr(Y) - \frac{1}{v}$ (sum of all elements of Y). Since (6.13) provides a solution of the equations (6.11), (6.15) follow immediately.

8. NUMERICAL EXAMPLE

Table 8.1 gives the yields and the lay-out of a design of the Y_1 class with treatments indicated by numbers within brackets.

TABLE 8.1. YIELDS AND THE LAY-OUT AT A DESIGN OF THE Y_1 CLASS (WITH TREATMENTS INDICATED BY NUMBERS WITHIN BRACKETS)

$i \setminus j$	1	2	3	4	5	6	7	8	9	10	R_i
1	140.1 (2)	161.8 (5)	112.2 (1)	153.9 (4)	116.5 (4)	189.2 (0)	160.3 (2)	152.7 (3)	178.0 (3)	134.0 (1)	1409.6
2	102.6 (1)	129.2 (6)	89.5 (3)	97.4 (1)	103.9 (3)	142.5 (5)	138.8 (6)	106.9 (4)	133.3 (5)	87.9 (2)	1132.0
3	155.0 (8)	163.8 (3)	138.3 (6)	141.0 (5)	79.8 (2)	141.0 (4)	161.2 (4)	130.1 (1)	155.8 (2)	107.1 (5)	1283.2
C_j	398.6	456.8	340.0	392.9	300.2	473.3	460.3	305.7	467.1	329.0	4014.8

SANKHYĀ : THE INDIAN JOURNAL OF STATISTICS : SERIES A

The parameters of the design are : $m = 3, n = 10, v = 6, r = 5, \lambda = 2,$

$n_1 = 1, n_2 = 4, \mu_0 = 0, \mu_1 = 0, \mu_2 = 8, p_{11}^1 = 0, p_{11}^2 = 0.$

$$a = \frac{\lambda v}{m} - \frac{1}{n}(\mu_0 - \mu_2) = 3.0 \quad b = \frac{1}{n}(\mu_1 - \mu_2) = 0.1$$

$$c = p_{11}^1 - p_{11}^2 = 0 \quad d = n_1 - p_{11}^2 = 1$$

$$A = a + bc = 3.0 \quad D = aA - b^2d = 15.2$$

The computational details of estimation are given in Table 8.2.

TABLE 8.2. COMPUTATIONAL LAY-OUT FOR INTERACTION ESTIMATES AND COMBINED ESTIMATES

treat- ments	1st class	2nd class	T_k	$[R]_k$	$[C]_k$	mnQ_k	mnD_k	mQ_k	nQ_k	nD'_{1k}	\bar{Q}_k	$\bar{D}i_k$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	
1	3	683.2	664.6	1837.1	-940.2	-3543.4	-107.3	-814.4	-3903.3	-37.12	-145.29	
2	4	623.0	6897.6	1956.1	-1462.8	-5748.9	-84.4	-638.6	-3236.8	-50.04	-207.24	
3	1	689.0	6046.4	1039.8	1333.8	4717.8	109.9	252.6	1156.3	39.20	153.82	
4	2	890.1	6897.0	2022.4	-430.8	-1801.5	17.0	-96.6	-539.2	-13.45	-58.85	
5	6	668.3	6530.0	2190.0	-127.0	-321.7	-61.1	333.0	1730.1	-0.31	3.93	
6	5	751.4	6530.0	2129.0	1730.0	6737.7	155.2	984.0	4854.0	61.92	248.63	
			4014.8	180	30	0*	0*	0*	0*	0*	0*	
			4014.8*	12044.4*								

$$[R]_k = \sum m_{1i} R_{1i}, [C]_k = \sum n_{2j} C_j;$$

$$mnQ_k = mnT_k - m[R]_k - n[C]_k + rQ_k;$$

$$mnD_k = A.mnQ_k + bS_1 (mnQ_k);$$

$$mQ_k = mT_k - [C]_k;$$

$$nQ_k = nT_k - [R]_k; nD'_{1k} = A'nQ_{1k} + bS_1 (nQ_{1k});$$

$$\bar{Q}_k = T_k - \frac{1}{n+D_1} [R]_k - \frac{1}{m+D_2} [C]_k + \frac{r}{mn} \left[1 - \frac{D_1}{n+D_1} - \frac{D_2}{m+D_2} \right]$$

$$\bar{D}i_k = \lambda \bar{Q}_k + bS_1 (\bar{Q}_k).$$

*Denotee check.

If we want interaction estimates only we need proceed only up to column (7) in this table. The analysis of variance table may be prepared at this stage as shown in Table 8.3.

ANALYSIS OF TWO-WAY DESIGNS

TABLE 8.3. ANALYSIS OF VARIANCE

source	degrees of freedom	sum of squares	mean squares	F
rows (unadjusted)	2	$SS_R^* = 7059.34$		
columns (unadjusted)	9	$SS_C^* = 11753.65$		
treatments (adjusted for rows and columns)	6	$SS_T = 2204.16$	$MS_T = 440.83$	3.39
error	13	$SS_E = 1600.66$	$MS_E = 130.05$	
interaction	18	$SS_I = 3894.81$		
total	29	$SS_T = 22707.70$		

For recovery of information from row-contrasts and column-contrasts further computations may be made as shown in the columns (8) to (12) of Table 8.2. The constants required are given below.

$$A' = A + (r-\lambda)/m = 4.9, \quad D' = [a + (r-\lambda)/m]A' - b^2d = 24,$$

$$a_1 = \frac{m\mu_0 - r^2}{\lambda(m-1)} = 0.5, \quad a_2 = \frac{(r-\lambda)[(c-1)A' - n_1b]}{\lambda(n-1)D'} = 0.3389,$$

We obtain

$$Q_1 t_1' = 1402.40 \quad Q_2 t_2' = 4607.75$$

$$SS_R = SS_R^* + SS_T - Q_1 t_1' = 7861.09 \quad SS_C = SS_C^* + SS_T - Q_2 t_2' = 9349.95$$

$$D_1 = \frac{(n-a_1)MS_E}{MS_R - MS_E} = 0.3251 \quad D_2 = \frac{(m-a_2)MS_E}{MS_C - MS_E} = 0.3808$$

$$\bar{a} = \frac{\lambda v + rD_2}{m + D_2} - \frac{1}{n + D_1}(\mu_0 - \mu_2) = 4.01570; \quad \bar{b} = \frac{1}{n + D_1}(\mu_1 - \mu_2) = 0.09685$$

$$\bar{A} = \bar{a} + \bar{b}c = 4.01570, \quad D = \bar{a}\bar{A} - \bar{b}^2d = 16.11719.$$

The two sets of estimates are given below in Table 8.4.

TABLE 8.4. ESTIMATES OF TREATMENT EFFECTS

k	t_k	\hat{t}_k
1	- 7.77	- 9.01
2	-12.61	-12.65
3	10.35	9.54
4	- 4.06	- 3.65
5	- 0.71	0.24
6	11.82	13.43

Estimated variances of the estimates of treatment differences for the two sets are obtained as shown below.

$$\text{Est. } V(i_k - i_{k'}) = \begin{cases} \frac{2(A-b)}{D} \cdot MS_0 = 65.03, & \text{if } k, k' \text{ are first associates} \\ \frac{2A}{D} \cdot MS_0 = 68.74, & \text{otherwise} \end{cases}$$

$$\text{Est. } V(\bar{i}_k - \bar{i}_{k'}) = \begin{cases} \frac{2(\bar{A}-\bar{b})}{\bar{D}} MS_0 = 63.24, & \text{if } (k, k') \text{ are first associates} \\ \frac{2\bar{A}}{\bar{D}} MS_0 = 64.81, & \text{otherwise.} \end{cases}$$

This shows that recovery of information from row-contrasts and column-contrasts has not resulted in appreciable gain in precision.

Next we compare this design with the design formed by taking columns as blocks. The relative efficiency factor e turns out to be 0.97938. Since $\phi = 1 + \frac{1}{\Delta_1} = 4.07598$ we get $E = e \times \phi = 3.99$ as the efficiency of the two-way design relative to the column-design, which shows that the gain in precision is appreciable.

REFERENCES

- BOSE, R. C. and KISHEN, K. (1939): On partially balanced Youden's Squares. *Science and Culture*, 5, 136-137.
- BOSE, R. C. and SHIMAMOTO, T. (1952): Classification and analysis of partially balanced incomplete block designs with two associate classes. *J. Amer. Stat. Ass.*, 47, 151-184.
- FISHER, R. A. (1936): *The Design of Experiments*, Oliver and Boyd, Edinburgh.
- KEMPTHORNE, OSCAR (1956): The efficiency factor of an incomplete block design. *Ann. Math. Stat.*, 27, 846-849.
- RAO, C. R. (1943): Quasi-Latin squares in experimental arrangements. *Current Science*, 7, 322-323.
- (1946): Confounded factorial designs in quasi-Latin squares. *Sankhyā*, 12, 295-304.
- (1952): *Advanced Statistical Methods in Biometric Research*, John Wiley and Sons, New York.
- (1955): Analysis of dispersion for multiply classified data with unequal numbers in cells. *Sankhyā*, 15, 253-260.
- ROY, J. (1958): On the efficiency factor of block designs. *Sankhyā*, 19, 181-188.
- SHRIMANUK, S. S. (1951): Designs with two-way elimination heterogeneity. *Ann. Math. Stat.*, 22, 235-247.
- YATES, F. (1937): The design and analysis of factorial experiments. *Imp. Bur. Soil. Sci. Tech. Comm.*, 35.
- YOUDEK, W. J. (1937): Use of incomplete block replications in estimating tobacco-mosaic virus. *Contr. Boyce Thompson Inst.*, 9, 41-48.

Paper received : November, 1960.