SANKHYĀ

THE INDIAN JOURNAL OF STATISTICS

Edited by: P. C. MAHALANOBIS

VOL. 17, PART 2 AUGUST 1956

ON THE RECOVERY OF INTER BLOCK INFORMATION IN VARIETAL TRIALS

By C. RADHAKRISHNA RAO

Indian Statistical Institute, Calcutta

1. INTRODUCTION

The object of this note is to make further comments on the method of combined intra and inter block analysis of experiments suggested by the author earlier in 1947, derive explicit expressions in the case of linked block (LB) designs and suggest a new method more suitable to designs which admit easy estimation of block parameters as in the case of LB, lattice and similar designs (Roy and Laha, 1950; Ramakrishnan, 1956, appearing in this issue).

2. THE Q METHOD

2.1. *The general analysis.* In the intra block analysis of varietal trials the normal equation for the estimation of varietal differences can be written, assuming equal replication for all varieties,

$$Q_i = \frac{r(k-1)}{k} t_i - \frac{\lambda_{i1}}{k} t_1 \cdots - \frac{\lambda_{iv}}{k} t_v \qquad \ldots \quad (2.1.1)$$

$$i = 1, \ldots, v$$

together with the consistent equation

$$0 = t_1 + \cdots + t_v$$

where Q_i = the total yield for the i-th variety minus the sum of block means in which it occurs,

λ_{ij} = the number of blocks in which the i-th and j-th varieties occur together,

r = the number of replications,

and k = the block size.

The corresponding equations for the combined intra and inter block analysis is shown to be (Rao, 1947, equation 3.12 with P changed to $Q(c)$ where c stands for combined)

$$Q_i(c) = \frac{R(k-1)}{k} t_i - \frac{\Lambda_{i1}}{k} t_1 - \cdots - \frac{\Lambda_{iv}}{k} t_v \qquad \ldots \quad (2.1.2)$$

together with the consistent condition

$$0 = t_1 + \cdots + t_v$$

where

$$R = r[w + w'/(k-1)],$$

$$\Lambda_{ij} = \lambda_{ij}(w - w'),$$

$$Q(c) = wQ_i + w'Q'_i \qquad \ldots \quad (2.1.3)$$

Q'_i = sum of means of blocks in which the i-th variety occurs minus r times the grand mean

w = the reciprocal of the estimated intra block variance and w' that of inter block variance (Rao, 1947, equations 4.1, 4.2).

It is seen that the equality of any two λ_{ij} imply the equality of the corresponding Λ_{ij}. Since the equations (2.1.1) and (2.1.2) are similar, it was noted in Rao (1947) that the solutions of the former as functions of Q, r and the distinct λ_{ij} provide solutions for the latter by writing $Q(c)$, R, Λ_{ij} for Q, r, λ_{ij}. The same is true of the expressions for the variances. It is necessary for the application of this method that r and λ_{ij} are treated as parameters and the solutions are obtained as their functions without recognizing the relationships such as

$$r(k-1) = \sum_i \lambda_{ij} \qquad \ldots \quad (2.1.4)$$

or any other relationship among the λ_{ij} values, since such relationships may not be true in terms of R and Λ_{ij}. If a relationship such as (2.1.4) is explicitely used then R and Λ_{ij} should be defined as

$$R = r[w + w'(v-k)/v(k-1)]$$

$$\Lambda_{ij} = \lambda_{ij}(w - w') + \frac{rk}{v} w'. \qquad \ldots \quad (2.1.5)$$

With these new definitions it is immaterial whether r is considered as an independent parameter or not in solving the intra block equations.

It is, however, possible to solve the intra block equations by not recognizing any relationship among r and λ_{ij}, in which case the expressions for the combined case can be obtained by changes given in (2.1.3). A certain amount of care may be necessary involving the actual examination of the method of solution instead of depending on published formulae. It will be seen that the method of solving is similar for the equations (2.1.1) and (2.1.2) thus establishing the correspondence between the solutions. Let us consider a few examples.

2.2. *The balanced incomplete block.* The intra block equations are

$$Q_i = \frac{r(k-1)}{k} t_i - \frac{\lambda}{k} \sum_{j \neq i} t_j, \qquad i = 1, \ldots, v \qquad \ldots \quad (2.2.1)$$

$$0 = t_1 + \cdots + t_v.$$

Eliminating $\sum_{j \neq i} t_j$ by using the last equation

$$Q_i = \frac{r(k-1)+\lambda}{k} t_i$$

from which
$$t_i = \frac{kQ_i}{r(k-1)+\lambda}$$

and
$$V(t_i - t_j) = V \frac{k(Q_i - Q_j)}{r(k-1)+\lambda} = \frac{2k}{r(k-1)+\lambda} \sigma^2.$$

It is clear that, if instead of the equations (2.2.1) we had

$$Q(c) = \frac{R(k-1)}{k} t_i - k \sum_{j \neq i} t_j, \qquad i = 1, \ldots, v \qquad \ldots \quad (2.2.2)$$

$$0 = t_1 + \cdots + t_v.$$

We obtain the same expressions for estimates and variances with Q, r, λ changed to $Q(c)$, R, Λ. In this special case the transformation is

$$\Lambda = \lambda(w - w')$$

$$R = r[w + w'/(k-1)].$$

For the variance, σ^2 has to be dropped since it is already taken into account in the above transformation.

2.3. *Partially balanced incomplete block.* Let us consider a partially balanced design with two classes of associates (Bose and Nair, 1939; Nair and Rao, 1942). The Q equations are

$$Q_i = \frac{r(k-1)}{k} t_i - \frac{\lambda_1}{k} \Sigma_{1i} t_j - \frac{\lambda_2}{k} \Sigma_{2i} t_j, \; i = 1, \ldots, v \qquad \ldots \quad (2.3.1)$$

$$0 = t_1 + \ldots + t_v$$

where Σ_{1i} and Σ_{2i} indicate respectively the summations over the first and second associates of the i-th variety.

To solve these equations we follow the method followed by Bose and Nair (1939). By summing over the i-th equation and its first associates after eliminating $\Sigma_{2i} t_j$ by using the last equation we find

$$\Sigma_{1i} Q_j = A_{22} t_i + B_{22} \Sigma_{1i} t_j$$

while the equation (2.3.1) can be written

$$Q_i = A_{12} t_i + B_{12} \Sigma_{1i} t_j$$

where

$$k A_{12} = r(k-1) + \lambda_2$$

$$k B_{12} = \lambda_2 - \lambda_1$$

$$k A_{22} = (\lambda_2 - \lambda_1) p^2_{12}$$

$$k B_{22} = r(k-1) + \lambda_2 + (\lambda_2 - \lambda_1)(p^1_{11} - p^2_{11}).$$

Eliminating $\Sigma_{1i} t_j$ and solving for t_i we get

$$t_i = [Q_i B_{22} + B_{12} \Sigma_{1i} Q_j] \div \Delta$$

$$= [(B_{22} + B_{12}) Q_i + B_{12} \Sigma_{2i} Q_j] \div \Delta$$

where

$$\Delta = A_{12} B_{22} - A_{22} B_{12}.$$

The second formula is convenient if the number of second associates is smaller in number. The variance of $t_i - t_r$ is

$$2(B_{22} + B_{12}) \sigma^2 / \Delta$$

if t_i and t_r are first associates and

$$2 B_{22} \sigma^2 / \Delta$$

if they are second associates.

In this method the solution for varietal differences and the expressions for the variances would have been the same if $Q(c)$, R, Λ_1, Λ_2 as defined in (2.1.3) were used instead of Q, r, λ_1, λ_2. Hence the expressions for the combined analysis can be obtained from the above by changing Q, r, λ to $Q(c)$, R, Λ. In the expressions for the variance, σ^2 should be dropped.

The analysis is similar for designs involving more associates. Special cases such as lattice designs, triangular designs etc. may be considered in a similar way and explicit expressions obtained or the general expression derived above may be used with special values of λ_1, λ_2 and p^i_{jk}. The general expressions for designs with three associate classes are given in Rao (1947) in such way as to provide combined estimates by changing Q, r, λ to $Q(c)$, R, Λ.

2.4. *Linked blocks.* In linked block (LB) designs introduced by Youden (1951) there are v varieties each replicated r times, b blocks of k plots and any two blocks have μ varieties in common. Roy and Laha (1956) provided a very simple method of analysing such designs using only the linked block property. This method will be explained in § 3 (the P method) where an alternative approach to intra and inter block analysis is considered.

For convenience in proving some results, we introduce matrix notation and express the equations in terms of matrices. Let N be the incidence matrix with b rows and v columns representing the blocks and varieties. The (i, j)-th element is unity if the i-th block contains the j-th variety and zero otherwise. It is easy to see for LB designs

$$NN' = \begin{pmatrix} k & \mu & \dots & \mu \\ . & . & \dots & . \\ \mu & \mu & \dots & k \end{pmatrix} = (k-\mu)\, I + \mu U$$

where U is a matrix with all elements unity and I is the unit matrix. Also

$$N'N = \begin{pmatrix} r & \lambda_{12} & \dots & \lambda_{1v} \\ . & . & \dots & . \\ \lambda_{v1} & \lambda_{v2} & \dots & r \end{pmatrix}$$

where λ_{ij} is the number of blocks in which the i-th and j-th varieties occur together. The intra-block Q equations are

$$\underset{\sim}{Q} = \left(r\, I - \frac{1}{k}\, N'N\right) \underset{\sim}{t} \qquad \dots \quad (2.4.1)$$

where $\underset{\sim}{Q}$ is the column vector $(Q_1, \dots, Q_v)$ and $\underset{\sim}{t}$, the column vector $(t_1, \dots, t_v)$.

Multiplying both sides by $N'N$ and simplifying

$$N'N\underset{\sim}{Q} = \left(rN'N - \frac{1}{k}\, N'NN'N\right)_{\sim}$$

$$= \left(rN'N - \frac{k-\mu}{k} N'N - \frac{\mu}{k}\, N'UN\right) \underset{\sim}{t}$$

$$= \left(r - \frac{k-\mu}{k}\right) N'N \underset{\sim}{t} \qquad \dots \quad (2.4.2)$$

since $N'UN\underset{\sim}{t} = 0$.

Eliminating $N'N\underset{\sim}{t}$ from (1.4.1) and (2.4.2)

$$[r(k-1)+\mu]\underset{\sim}{Q} + N'N\underset{\sim}{Q} = r[r(k-1)+\mu]\underset{\sim}{t}$$

$$\frac{\underset{\sim}{Q}}{r} + \frac{N'N\underset{\sim}{Q}}{r[r(k-1)+\mu]} = \underset{\sim}{t}\, . \qquad \dots \quad (2.4.3)$$

Comparing elements on both sides

$$t_i = \frac{Q_i}{r} + \frac{rQ_i + \Sigma\lambda_{is}Q_s}{r[r(k-1)+\mu]}$$

$$= Q_i\left(\frac{1}{r} + \frac{1}{\mu b}\right) + \frac{\Sigma\lambda_{is}Q_s}{\mu br},$$

since $r(k-1) = \mu(b-1)$.

The variance of $t_i - t_j$ is

$$2\left\{\left(\frac{1}{r} + \frac{1}{\mu b}\right) - \frac{\lambda_{ij}}{\mu br}\right\}\sigma^2$$

which depends only on λ_{ij} besides the parameters r, b, μ as shown by Roy and Laha (1956).

From the above analysis it is not clear how the combined intra and inter block estimates can be obtained. We shall write down the combined equations and follow the above procedure. Using matrix notation the combined equations

$$Q_i(c) = \frac{R(k-1)}{k} t_i - \frac{\Lambda_{i1}}{k} t_1 \cdots - \frac{\Lambda_{iv}}{k} t_v$$

of (2.1.2) can be written

$$\underset{\sim}{Q}(c) = \left(rw\, I - \frac{w-w'}{k} N'N\right)\underset{\sim}{t}. \qquad \ldots \quad (2.4.4)$$

We follow the same procedure as above. Multiplying both sides by $N'N$ and simplifying we get

$$N'N\underset{\sim}{Q}(c) = \left[rw\, N'N - \frac{(w-w')(k-\mu)}{k} N'N\right]\underset{\sim}{t}$$

$$= \left[rw - \frac{w-w'}{k}(k-\mu)\right] N'N\,\underset{\sim}{t} = g(w-w')N'N\,\underset{\sim}{t}$$

where $g = [krw-(k-r)(w-w')]/(w-w')$. ... (2.4.5)

Eliminating $N'NT$ from (2.4.4) and (2.4.5) as before

$$g\underset{\sim}{Q}(c) + N'N\underset{\sim}{Q}(c) = grw\,\underset{\sim}{t}. \qquad \ldots \quad (2.4.6)$$

Comparing the elements in the vectors on both sides

$$t_i = \frac{(g+r)Q_i(c) + \Sigma\lambda_{ij}Q_j(c)}{grw}.$$

The variance of $t_i - t_j$ is

$$2\,\frac{g+r-\lambda_{ij}}{grw}. \qquad \dots \quad (2.4.7)$$

which depends only on λ_{ij} besides other parameters common to all pairs (i, j).

We thus obtain the expressions for the combined analysis by following the same method of solving although it may be difficult to obtain the intra-block solutions in such a way as to provide the combined solutions by changing Q, r, λ to $Q(c)$, R and Λ. But this is clearly unnecessary unless it is simpler to do so.

3. The P method

3.1. *Intra block analysis.* Adopting matrix notation as in sub-section 2.4, we write the Q equations

$$\underset{\sim}{Q} = \left(rI - \frac{1}{k}N'N\right)\underset{\sim}{t}. \qquad \dots \quad (3.1.1)$$

Instead of solving these equations directly we may add to it b other consistent equations

$$\underset{\sim}{B} = k\underset{\sim}{b} + N\underset{\sim}{t} \qquad \dots \quad (3.1.2)$$

where $\underset{\sim}{B}$ = the column vector containing the block totals

$\underset{\sim}{b}$ = the column vector of b additional constants $b_1, b_2, \dots, b_b$ which may be referred to as block constants.

Multiplying (3.1.2) by $\frac{1}{k}N'$ and adding to (3.1.1) gives

$$\underset{\sim}{T} = r\underset{\sim}{t} + N'\underset{\sim}{b} \qquad \dots \quad (3.1.3)$$

where $\underset{\sim}{T}$ is the column vector of total yields of varieties.

Eliminating $\underset{\sim}{t}$ from the equations (3.1.2), (3.1.3)

$$\underset{\sim}{P} = \underset{\sim}{B} - \frac{1}{r}\,N\underset{\sim}{T} = \left(k - \frac{NN'}{r}\right)\underset{\sim}{b}. \qquad \dots \quad (3.1.4)$$

Writing these equations in full

$$P_j = \frac{k(r-1)}{r}\,b_j - \frac{\mu_{j1}}{r}\,b_1 - \dots - \frac{\mu_{jb}}{r}\,b_b \qquad \dots \quad (3.1.5)$$

where P_j = total yield of the j-th block minus the sum of the mean yields of varieties occurring in the j-th block.

μ_{ij} = the number of varieties common to the i-th and j-th blocks.

The equations (3.1.5) will be referred to as the P equations. It may be easier to solve these equations for the b constants which may be subject to a restriction of the type

$$b_1+\cdots+b_b = 0.$$

Having obtained these values they may be substituted in (3.1.3) to obtain $\underset{\sim}{t}$. The i-th equation in (3.1.3) is

$$T_i = r\,t_i + \Sigma_i b_s$$

where Σ_i denotes summation over the blocks in which the i-th variety occurs. Hence

$$t_i = \frac{1}{r}\,T_i - \frac{1}{r}\,\Sigma_i b_s.$$

The estimate of t_i-t_j is $\frac{1}{r}\,\{T_i-T_j-(\Sigma_i b_s-\Sigma_j b_s)\}$ and the variance of (t_i-t_j) is

$$V(t_i-t_j) = \frac{1}{r^2}\,V(T_i-T_j)+\frac{1}{r^2}\,V(\Sigma_i b_s-\Sigma_j b_s)$$

$$= \sigma^2\left\{\frac{2}{r}+\frac{C_{ij}}{r^2}\right\}$$

where $\sigma^2 C_{ij} = V(\Sigma_i b_s-\Sigma_j b_s)$.

To compute this variance the simplest way is to find the expression for $\Sigma_i b_s-\Sigma_j b_s$ in terms of P_j and use the following general result.

If $$l_1 b_1+\cdots+l_b b_b = d_1 P_1+\cdots+d_b P_b$$

then $$V(l_1 b_1+\cdots+l_b b_b) = (d_1 l_1+\cdots+d_b l_b)\sigma^2.$$

The sum of squares due to varieties can also be obtained by using the solutions of the P equations as was first shown by Yates (1939, 1940). So the entire analysis can be carried out by using the P equations (3.1.5) and the T eqations (3.1.3) involving varietal totals. This may be referred to as the P method. In the next section it is shown that the P method could be used in the intra and inter block analysis also.

3.2. *The intra and inter block analysis.* To the Q equations (2.1.2) for the combined intra and inter block estimates

$$\underset{\sim}{Q}(c) = \left(w\,r\,I-\frac{w-w'}{k}\,N'N\right)\underset{\sim}{t} \qquad \ldots \quad (3.2.1)$$

we add the consistent equations

$$\underset{\cdot}{B}-km\underset{\cdot}{u} = k\underset{\cdot}{b}+N\underset{\cdot}{t} \qquad \ldots \quad (3.2.2)$$

where $\underset{\cdot}{u}$ is the column vector with all its elements unity and m, the grand mean.

Eliminating $N'N\underset{\cdot}{t}$ as before

$$\underset{\cdot}{Q}(c)+\frac{w-w'}{k}N'(\underset{\cdot}{B}-km\underset{\cdot}{u}) = wr\underset{\cdot}{t}+(w-w')N'\underset{\cdot}{b}$$

or writing out in full

$$u(T_i-rm) = wrt_i+(w-w')\Sigma_i b_s, \qquad i = 1, \ldots, v. \qquad \ldots \quad (3.2.3)$$

Substituting for t_i in (3.2.2) we obtain

$$w\,P_i = b_i k\left(w-\frac{w-w'}{r}\right)-\frac{w-w'}{r}\Sigma\mu_{is}b_s \qquad \ldots \quad (3.2.4)$$

where P_i is same as defined in (3.1.5). The equations (3.2.4) are solved in the same way as (3.1.5). The values of b are substituted in (3.2.3) to obtain a solution for t_i

$$wrt_i = wT_i-(w-w')\Sigma_i b_s \qquad \ldots \quad (3.2.5)$$

$$t_i-t_j = \frac{u(T_i-T_j)}{wr}-\frac{(w-w')(\Sigma_i b_s-\Sigma_j b_s)}{wr}.$$

If $\qquad \Sigma_i b_s-\Sigma_j b_s = u(c_1P_1+c_2P_2+\cdots+c_bP_b)$

then $\qquad V(t_i-t_j) = \frac{2}{rw}\left\{1+\frac{(w-w')}{r}(\Sigma_i c_s-\Sigma_j c_s)\right\}.$

As an application let us consider the analysis of the LB design. The P equations (3.2.4) for combined estimation are

$$w\,P_i = b_i k\left(w-\frac{w-w'}{r}\right)-\mu\frac{(w-w')}{r}\Sigma b_s$$

$$= b_i\left\{kw-\frac{(k-\mu)(w-w')}{r}\right\}-\mu(b_1+\cdots+b_b).$$

Setting$(b_1+\cdots+b_b = 0)$ the solutions are

$$b_i = rw\,P_i/[r\,k\,w-(k-\mu)(w-w')]$$

$$= rw\,P_i/g(w-w')$$

where g is as defined in (2.4.6)

$$g(w-w') = k\, r\, w-(k-\mu)(w-w').$$

Substituting for b_i in 3.2.3

$$t_i = \frac{T_i}{r}-\frac{1}{g}\,\Sigma_i P_s$$

omitting the constant m since only the differences of $t_1, \ldots, t_v$ are to be considered

$$(t_i-t_j) = \frac{1}{r}(T_i-T_j)-\frac{1}{g}(\Sigma_i P_s-\Sigma_j P_s)$$

$$V(t_i-t_j) = \frac{2}{w}\left\{\frac{1}{r}+\frac{1}{g}-\frac{\lambda_{ij}}{rg}\right\}$$

which is same as the expression (2.4.7) obtained by following the Q method.

REFERENCES

BOSE, R. C. AND NAIR, K. R. (1939): Partially balanced incomplete block designs. *Sankhyā*, 4, 337.

NAIR, K. R. AND RAO, C. R. (1942): A note on partially balanced incomplete block designs. *Science and Culture*, 7, 516.

RAMAKRISHNAN, C. S. (1956): On the dual of a PBIB design, and a new class of designs with two replications. *Sankhyā*, 17, 133-142.

RAO, C. R. (1947): General methods of analysis for incomplete block designs. *J. Amer. Stat. Ass.*, 42, 541.

ROY, J. AND LAHA, R. G. (1956): Classification and analysis of linked block designs. *Sankhyā*, 17 115-132.

YATES, F. (1939): The recovery of inter-block information in varietal trials arranged in three dimensional lattice. *Ann. Eugenics*, 9, 136.

——— (1940): The recovery of inter-block information in balanced incomplete block designs. *Ann. Eugenics*, 10, 317.

Paper received : February, 1956.