ON THE COMBINED ANALYSIS OF A GROUP OF SPLIT-PLOT DESIGNS WITH HETEROGENEOUS ERROR VARIANCES

K.C. Bhuyan

SUMMARY

For a group of split-plot experiments with the same set of treatments, it is assumed that the error variance is constant for a particular experiment but it varies from experiment to experiment. Assuming error variances to be unknown and assuming that the individual analysis has already been performed, the combined estimate of contrast of whole-plot treatment parameters, sub-plot treatment parameters and the whole-plot X sub-plot interaction parameters are provided. The tests of the above contrasts are also suggested. The estimation and test are based on the work of James (1951, 1954). In real data analysis, it is observed that the suggested estimates and tests are in conformity with the estimates and tests observed in individual analysis of the experiments.

I. INTRODUCTION

In agricultural as well as industrial experiment the result obtained at a single place in a single season by a single experimeter, however, accurate it may be, can supply only limited information. So in the field of agricultural research, the experiment is repeated at a number of places over a number of seasons. This lead to the problem of analysis of groups of experiments. If the experiments are conducted in different locations under different agro-climatic conditions, there may arise cases of heterogeneous error variances. But on many occassions the experimenters fail to derive full information on treatment contrasts, as adequate solutions are not available for analysis of such groups of experiments.

Cochran (1937, 1954) was apparently the first author to discuss the combined analysis of groups of experiments with heterogeneous error variances. He suggested the weighted combination of the estimates of treatment parameters from individual experiments. But weighted combination is not suitable in all experimental conditions and it does not always yield unbiased estimate. The effects of heterogeneity of error variances are most serious when the analysis of variance technique is to be used as a method of statistical inference on treatment parameters. However, the difficulties obviate if the error variances are known. But in practice the experimenters are not aware of the error variances and use of estimated error variances as weights during analysis may vitiate the tests of significance.

In this paper, assuming error variances to be unknown, the combined estimates of treatment parameters of a group of split-plot experiments are obtained and the corresponding test is provided. It is assumed that the error variance of a particular experiment is constant but it varies from experiment to experiment. Thus for the estimation of treatment parameters usual method of least squares may be applied. Assuming the individual analysis has already been performed, the combined estimate and test of treatment parameters are provided. The estimation and test is based on the work of James (1951, 1954).

2. ESTIMATION OF TREATMENT EFFECTS FROM INDIVIDUAL EXPERIMENT

Let a split-plot experiment be repeated over p places. The model assumed for h-th experiment ($h = 1, 2, . . , p$) is

$$Y_{hijl} = \mu + \alpha_{hi} + \beta_{hj} + \gamma_{h1} + (\beta\gamma)_{hj1} + e_{hijl} \qquad (2.1)$$

where μ = general mean, α_{hi} = *i-th replication* effect at h-th place

β_{hj} = j-th whole-plot treatment effect at h-th place

γ_{h1} = 1-th sub-plot treatment effect at h-th place

$(\beta\gamma)_{hjl}$ = interaction j-th whole-plot treatment with 1-th sub-plot treatment at h-th place

e_{hijl} = random error

$i = 1, 2, . . . , r; j = 1, 2, . . . , q; 1 = 1, 2, . . . , v$

The quantities μ, α_{hi}, β_j, γ_1 and $(\beta\gamma)_{j1}$ are assumed to be fixed unknown parameters. The error terms e_{hij1} are normally distributed with

$$E(e_{hijl}) = 0$$

$$E(e_{hijl}, e_{hi'j'l'}) = \sigma_h^2 \text{ if } i = i', j = j' \text{ and } l = l'$$
$$= P_h \sigma_h^2 \; if \; i = i', j = j' \text{ and } l = l'$$
$$= 0 \text{ otherwise.}$$

Let the first and second kind of error variances for h-th experiment be

$$S_{1h}^2 = [\, 1 + (\nu - 1)\, p_h \,]\, \sigma_h^2$$

and $S_{2i}^2 = (1 - p_h)\sigma_h^2$

respectively. The restrictions on the model (2.1) are

$$\sum_i \alpha_{hi} = \Sigma\beta_j = \Sigma\gamma_1 = \sum_j (\beta\gamma)_{j1} = \sum_1 (\beta\gamma)_{j1} = 0 \quad .$$

Then by the least squares procedure the estimates of the parameters in the model are obtained as follows:

$$\mu = M = \bar{Y}_{h\cdot\,\cdot\,\cdot}, \; \hat{\alpha}_{hi} = a_{hi} = \bar{Y}_{hi\cdot\,\cdot} - \bar{Y}_{h\cdot\,\cdot\,\cdot}$$

$$\hat{\beta}_{hj} = b_{hj} = \bar{Y}_{h.j.} - \bar{Y}_{h\ldots}, \; \hat{\gamma}_{h1} = t_{h1} = \bar{Y}_{h..1} - \bar{Y}_{h\ldots}$$

$$(\hat{\beta\gamma})_{hj1} = (bt)_{hj1} = \bar{Y}_{h.j1} - \bar{Y}_{h.j.} - \bar{Y}_{h..1} + \bar{Y}_{h\ldots}$$

3. ESTIMATION AND TEST WHEN THE ERROR VARIANCES ARE UNKNOWN

The main object of this analysis is to test the following hypothesis

i) Ho : $\beta_1 = \beta_2 = \ldots = \beta_q$ (3.1)

ii) Ho : $\gamma_1 = \gamma_2 = \ldots = \gamma_\nu$ (3.2)

iii) Ho : $(\beta\gamma)_{11} = (\beta\gamma)_{12} = \ldots = (\beta\gamma)_{q\nu}$ (3.3) and to estimate the following vectors of contrasts:

$$\underline{\beta} = \begin{bmatrix} 1 & -1 & 0 & 0 & \ldots & 0 \\ 1 & 0 & -1 & 0 & \ldots & 0 \\ - & - & - & - & & - \\ 1 & 0 & 0 & - & \ldots & 1 \end{bmatrix}_{(q-1)xq} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \\ \beta_q \end{bmatrix} \tag{3.4}$$

$$\underline{\gamma} = \begin{bmatrix} 1 & -1 & 0 & 0 & \ldots & 0 \\ 1 & -0 & -1 & 0 & \ldots & 0 \\ - & - & - & - & & - \\ 1 & 0 & 0 & & & 1 \end{bmatrix}_{(\nu-1)x\nu} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \\ \gamma_\nu \end{bmatrix} \tag{3.5}$$

and a similar vector of contrasts of the interaction term $(\beta\gamma)_{j1}$. Here β_j ($j = 1, 2, \ldots, q$) and γ_1 ($1 = 1, 2, \ldots, \nu$) are the effect of j-th whole-plot treatment and 1-th sub-plot treatment respectively for the experiment as a whole.

Let $\underline{T}_{11}, \underline{T}_{12}, \ldots, \underline{T}_{1p}$ be p vectors of estimates of the (q-1) contrasts of the q whole-plot treatment effects from first, second, . . . pth experiment respectively such that

$$\underline{T}_{1h} = \begin{bmatrix} 1 & -1 & 0 & \ldots & 0 \\ 1 & 0 & -1 & \ldots & 0 \\ - & - & - & & - \\ 1 & 0 & 0 & & -1 \end{bmatrix}_{(q-1)xq} \begin{bmatrix} b_{h1} \\ b_{h2} \\ \\ b_{hq} \end{bmatrix} \tag{3.6}$$

where $h = 1, 2, \ldots, p$. Let $D_{11}, D_{12}, \ldots, D_{1p}$ be variance matrices of $\underline{T}_{11}, \underline{T}_{12}, \ldots, \underline{T}_{1p}$ respectively. The estimates of these variance matrices are given such that

$$\hat{D}_{1h} = \begin{bmatrix} \frac{2}{vr} & \frac{1}{vr} & \cdots & \frac{1}{vr} \\ \frac{1}{vr} & \frac{2}{vr} & \cdots & \frac{1}{vr} \\ - & - & & - \\ \frac{1}{vr} & \frac{1}{vr} & \cdots & \frac{2}{vr} \end{bmatrix}_{(q-1)\text{x}(q-1)} s^2_{1h} \qquad (3.7)$$

Let $\underline{T}_{21}, \underline{T}_{22} \quad \ldots \underline{T}_{2p}$ be p vectors of estimates of the (ν-1) contrasts of ν sub-plot treatment effects from first, second . . . . pth experiment respectively such that

$$\underline{T}_{2h} = \begin{bmatrix} 1 & -1 & 0 & 0 & .. & 0 \\ 1 & 0 & -0 & 0 & .. & 0 \\ - & - & - & - & & - \\ 1 & 0 & 0 & 0 & .. & -1 \end{bmatrix}_{(\nu-1)\text{ x }\nu} \begin{bmatrix} t_{h1} \\ t_{h2} \\ - \\ t_{h\nu} \end{bmatrix} \qquad (3.8)$$

The variance matrices of $\underline{T}_{21}, \underline{T}_{22}, \ldots, \underline{T}_{2p}$ may be denoted as $D_{21}, D_{22}, \ldots, D_{2p}$ respectively. The estimates of these variance matrices are obtained such that

$$\hat{D}_{2h} = \begin{bmatrix} \frac{2}{qr} & \frac{1}{qr} & \cdots & \frac{1}{qr} \\ \frac{1}{qr} & \frac{2}{qr} & \cdots & \frac{1}{qr} \\ - & - & & - \\ \frac{1}{qr} & \frac{1}{qr} & \cdots & \frac{2}{qr} \end{bmatrix}_{(\nu-1)\text{ x }(\nu-1)} s^2_{2h} \qquad (3.9)$$

The vector of estimates T_{3h} of the (q-1) (ν-1) contrasts of $q\nu$ interaction effects from h-th experiment and estimates of its variance matrix $\hat{D}_{3h}$ can similarly be defined. Here s_{1h}^2 and s_{2h}^2 are respectively the estimates of first and second kind of error variances from h-th experiment. These estimates are obtained as usual.

Now for combined of contrasts and test of hypothesis concerning the contrasts the following theorem may be enunciated.

Theorem: Let

$$\underline{T}_1 = \begin{bmatrix} \hat{B}_{11} \\ \hat{B}_{12} \\ \\ \hat{B}_{1\ m-1} \end{bmatrix}; \quad \underline{T}_2 = \begin{bmatrix} \hat{B}_{21} \\ \hat{B}_{22} \\ \\ \hat{B}_{2\ m-1} \end{bmatrix} \ldots, \quad \underline{T}_p = \begin{bmatrix} \hat{B}_{p1} \\ \hat{B}_{p2} \\ \\ \hat{B}_{p\ m-1} \end{bmatrix}$$

are p vectors of estimates of $(m\text{-}1)$ contrasts of the m treatment effects from first, second, ... , p-th experiment respectively. The expected values of T_h ($h = 1, 2, \ldots, p$) are the linear functions of the contrasts of $\delta_1, \delta_2 \ \ldots, \delta_m$ and that they are independently distributed around these expected values in $(m\text{-}1)$ variate normal distribution whose variance matrices are $D_1, D_2, \ldots, D_p$ respectively. In matrix notation

$$D(\underline{T}) = B\ \underline{B} \tag{3.10}$$

where T is the column vector of $T_1, T_2, \ldots, T_p$ and B is the $p(\nu-1)$ x $(\nu-1)$ identity matrices. Here.

$$B = \begin{bmatrix} B_1 \\ B_2 \\ \cdot \\ B_{m-1} \end{bmatrix} = \begin{bmatrix} 1_{11} & 1_{12} & \ldots & 1_{1m} \\ 1_{21} & 1_{22} & \ldots & 1_{2m} \\ - & - & & - \\ 1_{m-11} & 1_{m-12} & & 1_{m-1m} \end{bmatrix}_{(m-1)\mathrm{x}m} \begin{bmatrix} \delta_1 \\ \delta_2 \\ - \\ \delta_m \end{bmatrix}$$

and $$= L\delta \qquad (3.11)$$

$$\underline{T}_h = \begin{bmatrix} \hat{B}_{h1} \\ \hat{B}_{h2} \\ \cdot \\ \hat{B}_{h\ m-1} \end{bmatrix} = \begin{bmatrix} l_{11} & l_{12} & \dots & l_{1m} \\ l_{21} & l_{22} & \dots & l_{2m} \\ - & - & & - \\ l_{m-11} & l_{m-12} & \dots & l_{m-1\ m} \end{bmatrix} \begin{bmatrix} d_{h1} \\ d_{h2} \\ - \\ d_{hm} \end{bmatrix}$$

$$= L\,\underline{d} \qquad (3.12)$$

where δ_k $(k = 1, 2, \dots m)$ is the k-th treatment effect and d_{hk} $(h = 1, 2, \dots p)$ is the estimate of k-th treatment effect at h-th place. The rank of the matrix L is $(m$-$1)$. Let $\hat{D}_1, \hat{D}_2, \dots, \hat{D}_p$ be the unbiased estimates of $D_1, D_2, \dots D_p$ and distributed independently of $\underline{T}_h$ and of each other in Wishart forms with $f_1, f_2, \dots, f_p$ degrees of freedom respectively. Then under the null hypothesis

$$H_o: \; B_1 = B_2 = \dots = B_{m-1} = 0 \qquad (3.13)$$

to the order $1/f_h$

$$P_r \left[\Sigma (\underline{T}_h - \underline{T})' \, \omega_h \, (\underline{T}_h - \underline{T}) \leqslant 2h(a) \right] = \alpha \qquad (3.14)$$

where $\omega_h = \hat{D}_h^{-1}$

$$2h(a) = \chi^2 (A + B\chi^2)$$

$$A = 1 + \frac{1}{2(m-1)^2} \sum_h \frac{1}{f_h} \, tr\,(1 - \omega^{-1}\omega_h)^2$$

$$B = \frac{1}{(m-1)^2\,[(m-)^2 + 2]} \left[\Sigma \frac{1}{f_h} \, tr\,(1-\omega^{-1}\omega_h)^2 \right.$$

$$\left. + \frac{1}{2} \, \Sigma \frac{1}{f_h} \, (tr\,(1-\omega^{-1}\omega_h)^2 \right]$$

$$\omega = \sum_h \omega_h$$

$$\underline{T} = \omega^{-1} \Sigma \omega_h \underline{T}_h \tag{3.15}$$

χ^2 is the $\alpha\%$ point of χ^2 -variate with $(m-1)^2$ *d.f.*

This theorem can easily be proved on the basis of the work of James (1954). In practice the statistic

$$\Sigma T_h' (\omega_h T_h) - \underline{T}_h' (\omega \underline{T}) \tag{3.16}$$

is to be compared with $\chi^2 (A + B \chi^2)$. James also discussed that the statistic (3.16) is exactly distributed as χ^2 provided f_h's are large. The hypothesis (3.13) implies that all (m-1) contrasts are insignificant. But it sometimes require to test any of the contrast as significant. For this a statistic similar to that given in (3.16) is to be one element, L is a row vector of m elements such that the sum of m elements are zero. Then the statistic

$$\Sigma \omega_h T_h^2 - \frac{(\Sigma \omega_h T_h)^2}{\omega} \tag{3.17}$$

is distributed as χ^2 with (m-1) d.f., provided f_h are large. For f_h not large enough, James (1951) suggested to compare the statistic (3.17) with

$$2h(a) = \chi^2 \left[1 + \frac{3\chi^2 + (p+1)}{2(p^2-1)} \Sigma \frac{1}{f_h} (1 - \frac{\omega_h}{\omega})^2 \right] \tag{3.18}$$

where χ^2 is the $\alpha\%$ point of χ^2 -variate with (m-1) d.f. The statistic (3.15) is to be used as combined estimate of B.

For estimating $\underline{\beta}$ and $\underline{\gamma}$ the statistics to be computed are $\underline{T}_1 = [\omega_1^{-1} \Sigma \omega_{1h} T_{1h}]$ and $\underline{T}_2 = [\omega_2^{-1} \Sigma \omega_{2h} \underline{T}_{2h}]$ respectively. For estimating the vector of contrasts of the interaction effects, the statistics to be computed is $T_3 = [\omega_3^{-1} \Sigma \omega_{3h} T_{3h}]$. Here $\omega_{1h} = \hat{D}_{1h}^{-1}$, $\omega_{2h} = \hat{D}_{2h}^{-1}$, $\omega_{3h} = \hat{D}_{3h}^{-1}$, $\omega_1 = \Sigma \omega_{1h}$, $\omega_2 = \Sigma \omega_{2h}$ and $\omega_3 = \Sigma \omega_{3h}$. To test the hypothesis (3.1), (3.2) and (3.3) the following test statistics are to computed:

$$\Sigma\ \underline{T}'_{1h}\ (\omega_{1h}\underline{T}_{1h})\quad -\quad \underline{T}'_1\ (\omega_1\ \underline{T}_1) \tag{3.19}$$

$$\Sigma\ \underline{T}'_{2h}\ (\omega_{2h}\ \underline{T}_{2h})\quad -\quad \underline{T}'_2\ (\omega_2\ \underline{T}_2) \tag{3.20}$$

and

$$\Sigma\ \underline{T}'_{3h}\ (\omega_{3h}\ \underline{T}_{3h})\quad -\quad \underline{T}'_3\ (\omega_3\ \underline{T}_3) \tag{3.21}$$

These statistics are distributed with $(q\text{-}1)^2$, $(\nu\text{-}1)^2$ and $[(q\text{-}1)\ (\nu\text{-}1)]^2$ degrees of freedom respectively.

4. *EXAMPLE*

The data come from a group of 6 experiments on paddy conducted in Haringhata Teaching Farm, Kalyani in 1972 and 1974. The object of the experiments was to study the response of Nitrogen to different High Yielding Varieties of paddy. The experiments were conducted in split-plot design of 8 whole-plots per block and 5 subplots per whole-plot. The whole-plot treatments were the 8 different varieties of paddy and the sub-plot treatments comprised of 5 different levels of application viz., 0, 50, 100, 150 and 200 kg/ha Nitrogen as A/S. All the experimental conditions were same for the six experiments and each experiment was repeated three times.

From the individual analysis of the experiments the error mean sum of squares were observed as follows:

Error Mean Sum of Squares

First kind (s_{1h}^2)	Second Kind (s_{2h}^2)
0.034293	0.002472
0.027778	0.011372
0.502693	0.175073
0.759071	0.295578
0.163964	0.057776
0.220836	0.059994

Bartlett's (1937) χ^2 -test using different values of s_{1h}^2 was performed and first kind of error variances were observed as heterogeneous.

The second kind of error variances were also observed as heterogeneous. Applying the statistic (3.15) the estimates of the vector of contrasts of the effects of paddy and Nitrogen were observed as follows:

$$\underline{T}_1 = \begin{bmatrix} -0.45531 \\ -0.62068 \\ -0.35885 \\ -0.23230 \\ -0.13410 \\ -0.07905 \\ -0.27090 \end{bmatrix} \quad \text{and } \underline{T}_2 = \begin{bmatrix} -.25529 \\ -0.17460 \\ -0.0223155 \\ -0.13445 \end{bmatrix}$$

The statistic (3.19) was computed as 873.78 and 2h(a) for this statistics was observed as 93.76 with 49 degrees of freedom. For the statistic (3.20) the value observed was 898.22 as against 2h(a) = 29.29 with 16 degrees of freedom. The value of the statistic (3.21) was computed as 1077.85 against 2h(a) = 876.45 with 764 degrees of freedom.

From the combined analysis it may be concluded that each of paddy, Nitrogen and their interaction were significantly different. From the individual analysis it was observed that in all the experiments the effects of Nitrogen were significant. In five out of six experiments the paddy differences were found significant and only in three experiments the paddy and Nitrogen interaction differed significantly. Thus the findings by the present method is in conformity with the findings by individual analysis.

ACKNOWLEDGEMENT

The author is grateful to Dr. G.M. Saha of Indian Statistical Institute, Calcutta and Dr. T.K. Gupta of B.C.K.V. of India for their valuable suggestions in preparing this paper.

The second kind of error variances were also observed as heterogeneous. Applying the statistic (3.15) the estimates of the vector of contrasts of the effects of paddy and Nitrogen were observed as follows:

$$\underline{T}_1 = \begin{bmatrix} -0.45531 \\ -0.62068 \\ -0.35885 \\ -0.23230 \\ -0.13410 \\ -0.07905 \\ -0.27090 \end{bmatrix} \qquad \text{and } \underline{T}_2 = \begin{bmatrix} -.25529 \\ -0.17460 \\ -0.0223155 \\ -0.13445 \end{bmatrix}$$

The statistic (3.19) was computed as 873.78 and 2h(a) for this statistics was observed as 93.76 with 49 degrees of freedom. For the statistic (3.20) the value observed was 898.22 as against 2h(a) = 29.29 with 16 degrees of freedom. The value of the statistic (3.21) was computed as 1077.85 against 2h(a) = 876.45 with 764 degrees of freedom.

From the combined analysis it may be concluded that each of paddy, Nitrogen and their interaction were significantly different. From the individual analysis it was observed that in all the experiments the effects of Nitrogen were significant. In five out of six experiments the paddy differences were found significant and only in three experiments the paddy and Nitrogen interaction differed significantly. Thus the findings by the present method is in conformity with the findings by individual analysis.

ACKNOWLEDGEMENT

The author is grateful to Dr. G.M. Saha of Indian Statistical Institute, Calcutta and Dr. T.K. Gupta of B.C.K.V. of India for their valuable suggestions in preparing this paper.