RANK ANALYSIS OF GROUPS OF SPLIT-PLOT EXPERIMENTS

S.C. RAI AND P.P. RAO
I.A.S.R.I., New Delhi-110012
(Received : November, 1983)

SUMMARY

A method which does not involve some of the assumptions needed in the case of analysis of variance has been developed. In this new method ranks of treatments are considered instead of their actual values and suitable tests are developed for comparing the means of ranks of different treatments in the case of factorial experiments laid out in split-plot designs.

INTRODUCTION :

In agricultural experimental programmes, it is necessary and also desirable to repeat the trials for a set of treatments at a number of places or during different years. The main aim of such repetitions is to study the susceptibility of treatment effects to places and climatic variations so that the proper recommendation may be made regarding the utility of treatments for various tracts. For drawing valid and appropriate conclusions regarding the suitability of treatment effects, it becomes necessary to perform the joint statistical analysis of the data by combining the results of individual trials. The results may be classified as belonging to one of the following four types :

(*i*) The error variances homogeneous and interaction present.

(*ii*) The error variances homogeneous and interaction absent.

(*iii*) The error variances heterogeneous and interaction present.

(*iv*) The error variances heterogenous and interaction absent.

The usual technique of analysis of variance is not valid when the error variances are heterogenos and interaction absent.

Split-plot design is quite useful and popular in agricultural experimentation because it permits the treatments which require larger plots and it gives higher precision over randomised block design for sub-plot treatments. The analysis of groups of split-plot designs poses certain statistical problems. The combined results may be classified into one of 32 groups. Viz. 4 with main-plot errors, 4 with sub-plot errors and 2 with presence or absence of interaction between main-plot and sub-plot treatments ($4 \times 4 \times 2$).

The analysis of most of the groups of split-plot experiments is not possible because error variances (main-plot error, or sub-plot error or both) are heterogeneous. It is also possible that various effects may not follow additive model and thereby invalidate the use of analysis of variance technique. In the present paper, we are developing a method of analysis by utilising the ranks of observations obtained from split-plot experiments and drawing valid conclusons about treatment effects. The method enables us to pool the results of similar experiments conducted over years at the same place or at different locations during the same year without any statistical handicap.

2. The Procedure of Analysis :

Rai [4] has developed a method for analysing ordered observations in complete block designs. Rai and Rao [5] have evolved a general method of analysing the data of groups of experiments conducted in randomised block designs at various places or during different years. Koch [3] has suggested non-parametric method for analysing complex split-plot experiments. Here we will outline a procedure for analysing groups of experiments conducted in split-plot designs at various places or over different times without assuming normality.

2.1 Combined analysis of Main-plot treatments :

Let us assume the following :—

Number of main-plot treatments $=\alpha$

Number of sub-plot treatments $=\beta$

Number of replications $=r$

Number of years or places $=p$

The procedure involves first summing the observations of all sub-plots of a main-plot in each block of the individual experiment and then rank them within each block. The observations are to be ranked by giving rank 1 to the lowest value, 2 to next higher and so on. The highest value of the observation in a block will have the rank value α. Ranking is to be done afresh for each block and it will have the variate values 1, 2, ..., α. On the hypothesis that there is no significant difference between the main-plot treatments, the difference in the ranks of various treatments will arise solely from sampling fluctuations. The rank obtained by a particular observation would then be a matter of chance. The trials in split-plot designs are repeated at p places or years. The set of ranks r_{ijk} being the rank of jth main-plot treatment in the ith replication of k-th trial (place or year) for each treatment would represent a random sample of rp items from a discontinuous rectangular universe 1, 2: ..., α. The mean and variance of this universe are obtained as $\frac{1}{2}(\alpha+1)$ and $(1/12)(\alpha^2-1)$ respectively.

Now the next step in the analysis is to obtain the mean rank

$$\bar{R}_j=\left(\frac{1}{rp}\sum_{i=1}^{r}\sum_{k=1}^{p} r_{ijk}\right) \qquad \ldots(1)$$

for j-th main-plot treatment. $(j=1, 2, \ldots, \alpha)$.

These means are all estimates of the same rectangular universe. The sampling distribution of the means of ranks will be approximately normal even for a moderate number of observations, Hilda [2]. The sampling distribution of mean rank will be normal with mean

$$\bar{R}=\frac{1}{2}(\alpha+1) \text{ and variance } \sigma^2=\frac{1}{12\,rp}(\alpha^2-1).$$

The hypothesis that means of the ranks of various treatments come from a single homogenous normal population can be tested by the statistic

$$M_1=\sum(\bar{R}_j-\bar{R})^2/\sigma^2$$

The value of M_1 can be put as

$$M_1=\frac{12}{rp\,(\alpha^2-1)}\sum_{j=1}^{\alpha} R_j^2-\frac{3\,rp\alpha\,(\alpha+1)}{(\alpha-1)} \qquad \ldots(2)$$

where R_j is the sum of ranks of j-th treatment. It may be seen that M_1 is distributed as χ^2 with $(\alpha-1)$ *d.f.* for large rp as it is the sum of squares of standardised normal variates. If M_1 is significantly greater than might reasonably have been expected from chance, it may be concluded that mean ranks differ significantly and there is significant difference in the main-plot treatment effects. The C.D. at 5 percent level of significance for comparison of any two treatment rank means is given by

$$\sqrt{(\alpha^2-1)/6\,rp} \times 1.96 \qquad \ldots(3)$$

For testing the presence of interaction between main-plot treatment effects × place (year), the following test statistic may be used

$$M_1^1 = \frac{12}{r(\alpha^2-1)} \left[\sum_{k=1}^{p} \sum_{j=1}^{\alpha} \left(\sum_{i=1}^{r} r_{ijk} \right)^2 - \frac{1}{p} \sum_{j=1}^{\alpha} R_j^2 \right] \qquad \ldots(4)$$

M_1^1 is distributed as χ^2 with $(\alpha-1)(p-1)$ *d f.* The significance of the value of M_1^1 indicates the presence of interaction of main-plot treatments with place (year).

2.2. Combined analysis of Sub-plot treatments :

For pooled analysis of sub-plot treatments, ranking of sub-plot observations will be done within each main-plot of a block. Rank the Sub-plot observations by giving rank 1 to the lowest value, 2 to next higher and so on for each main-plot of a block. Let r_{ijsk} be the rank of s-th sub-plot treatment in the j-th main-plot of i-th replication of k-th experiment. The r_{ijsk} would represent a random sample of $r\ \alpha\ p$ items from a discontinuous rectangular universe 1, 2, ..., β. The mean rank for each sub-plot will be given by

$$\overline{R}_s = \frac{1}{r\,p\,\alpha} \sum_{i=1}^{r} \sum_{j=1}^{\alpha} \sum_{k=1}^{p} r_{ijsk} \qquad \ldots(5)$$

The sampling distribution of $\overline{R}_s$ would be normal with mean $\frac{1}{2}(\beta+1)$ and variance$=\dfrac{1}{12\,r\,p\,\alpha}(\beta^2-1)$.

The hypothesis that the means of ranks of various sub-plot treatments come from a single homogeneous normal population can be tested by the statistic

$$M_2 = \frac{12}{r\,p\,\alpha\,(\beta^2-1)} \sum_{s=1}^{\beta} R_s^2 - \frac{3\,r\,p\,\alpha\,\beta\,(\beta+1)}{\beta-1} \qquad \ldots(6)$$

where $R_s = rp\alpha\,\bar{R}_s$.

For large $rp\alpha$, M_2 is distributed like χ^2 with $(\beta-1)$ d.f. as it is the sum of squares of standard normal variates. The significance of M_2 will indicate the rejection of hypothesis of equality of treatment effects. The C.D. at 5p ercent level of significance for comparison of two sub-plot tseatment rank means is given by

$$\sqrt{(\beta^2-1)/6\,rp\alpha} \times 1.96 \qquad \ldots(7)$$

For testing the presence of interaction between sub-plot treatments and places (years) the test statistic is

$$M_2^1 = \frac{12}{r(\beta^2-1)} \Big[\sum_{k=1}^{p} \sum_{j=1}^{\alpha} \sum_{s=1}^{\beta} \Big(\sum_{i=1}^{r} r_{ijsk} \Big)^2 - \frac{1}{p} \sum_{j=1}^{\beta} R_s^2 \Big] \qquad \ldots(8)$$

which is distributed as χ^2 with $(\beta-1)\,(p-1)$ d.f.

The significance of M_2^1 will indicate the presence of interaction between sub-plot treatments and places (years).

2.3. Combined analysis of Main-plot X Sub-plot treatment interactions :

There will be $\alpha\beta$ combinations of main-plot and sub-plot treatments. Rank each observation of a replication by giving rank 1 to the lowest value, 2 to next higher and so on. The highest value of the observation will have the rank $\alpha\beta$. Let $r_{(ms)ik}$ be the rank of the treatment combination of m-th main-plot $\times$ s-th sub-plot treatment in the ith replication of kth trial. Then this would represent a random sample of rp items from a discontinuous rectangular universe 1, 2, ..., $\alpha\beta$. The rank of means of interaction wiil be obtained as

$$\bar{R}_{ms}=\frac{1}{rp}\sum_{i=1}^{r}\sum_{k=1}^{p} r_{(ms)ik} \qquad \ldots(9)$$

for $m=1, 2, \ldots \alpha$ and $s=1, 2, \ldots, \beta$.

The sampling distribution of $\bar{R}_{ms}$ will be normal with mean $\bar{R}''$ $\frac{1}{2}(\alpha\beta+1)$ and variance $\sigma^2_{''} = \frac{1}{12rp}(\alpha^2\beta^2-1)$. The hypothesis that the means of the ranks of the various interactions of main-plot and sub-plot treatments come from a single, homogeneous normal population can be tested by the statistic

$$M_3=\frac{12}{rp\,(\alpha^2\beta^2-1)}\sum_{ms} R^2_{ms}-\frac{3rp\,\alpha\beta\,(\alpha\beta+1}{(\alpha\beta-1)} \qquad \ldots(10)$$

nhere $R_{ms}=rp\,\bar{R}_{ms}$.

This statistic is distributed as χ^2 with $(\alpha\beta-1)$ d.f. for large rp as it is the sum of squares of standardised normal variates. The C.D. at 5 percent level of significance for the comparison of any two interaction of rank means is given by

$$\sqrt{(\alpha^2\beta^2-1)/6\,rp}\times 1.96 \qquad \ldots(11)$$

For testing the presence of interaction between main-plot $\times$ Sub-plot treatments and places (years), the test statistic is given by

$$M^1_3=\frac{12}{r\,(\alpha^2\beta^2-1)}\left[\sum_{k=1}^{p}\sum_{ms}\left(\sum_{i=1}^{r} r_{(ms)ik}\right)^2 - \frac{1}{p}\sum_{ms} R^2_{ms}\right] \qquad \ldots(12)$$

which is distributed as χ^2 with $(\alpha\beta-1)$ d.f. The significance of this value will indicate the presence of interaction of main-plot $\times$ sub-plot treatments with places (years).

3. ILLUSTRATION:

As an illustration of the method described above, the data of an agricultural experiment conducted during three years at State Agriculture Farm, Kalyani, West Bengal are taken. In order to see the effect of different levels of nitrogen (N_0, N_1, N_2, and N_3) on the yields of various varieties of wheat (V_1, V_2, V_3, V_4, V_5 and V_6), split-plot experiments taking nitrogen in the main-plot and varieties in the sub-plot, were carried out during 1971, 72 and 73. There were two replications. The error variances were heterogeneous for both main-plot treatment effects (Error 1) and sub-plot treatment effects (Error 2) and hence no conclusions could be drawn from analysis of variance technique.

3.1. Grouping of Main-plot treatments :

Following the procedure explained under 2.1, we rank the main-plot treatments and present below the sum of ranks of each main-plot treatment.

Main-plot treatments	*Sum of ranks*		
	1971	*1972*	*1973*
N_0	2	2	2
N_1	4	5	4
N_2	7	7	7
N_3	7	6	7

The value of M_1 given at (2) is worked out as 19.47. This is distributed as χ^2 with 3 d.f. and it is highly significant indicating that main-plot treatment effects differ significantly from one an other. Comparing the mean ranks of various treatments with the help of C.D. which is worked out as 1.265 at 5 per cent level, we conclude that treatments N_2 and N_3 are superior to No and N_1. The performance of No is inferior to all. The value of $M^1{}_1$ given at (4) is 0.53 which is not significant indicating the absence of interaction between the main-plot treatments and years.

3.2. Grouping of Sub-plot treatments

Rank the Sub-plot treatments as explained in 2.2 and obtain the sum of ranks of each sub-plot treatments. The value of M_2 is worked out as 35.2 which is distributed as x^2 with 5 *d.f.* This value is highly significant which indicates as significant variation in the effect of sub-plot treatments. The value of C.D. by using (7) has been worked out as 0 96 at 5 per cent level of significance. The mean ranks of sub-plot treatments are presented below:

$$V_1=2.9, V_2=3.1, V_3=4.6, V_4=3.7, V_5=2.9, V_6=3.7.$$

By using C.D. we may conclude that varieties V_3, V_4 and V_6 are similar in their effects and V_3 is superior to V_1, V_2 and V_5.

For testing the presence of interaction between sub-plot treatments and years, the value of M_2^1 from (8) is worked out as 51.1 which is distributed as x^2 with 18 *d.f.* This is highly significant which indicates the presence of interaction between the sub-plot treatments and years.

3.3. Grouping of interaction of Main-plot and Sub-plot treatments :

Rank the main-plot × sub-plot treatment interaction as described in 2.3 and obtain the sum of ranks for each treatment. The value of M_3 from (10) is worked out as 114.3. This is distributed as x^2 with 23 *d.f.* which is highly significant indicating the presence of interaction between main-plot and sub-plot treatments. For testing the presence of interaction of main-plot X sub-plot tretments with years, the value of M_3^1 from (12) which is a function of sum of treatment ranks, may be computed.

Conclusions :

The method described in the paper uses information on 'ranks' and makes no use of the quantitative values of observations as such. For this reason no assumption is required to be made as to the nature of underlying universe. The method is thus applicable to a wide class of problems to which the analysis of variance technique can not validly be applied. The various problems encountered in the pooled analysis of data pertaining to split-plot experiments have been discussed. The technique of rank analysis has been suggested to overcome the situations when the error variances are heterogeneous in pooling the results from groups of split-plot experiments. Models for combined analysis of main-plot treatments, sub-plot treatments and interaction of main-plot and sub-plot treatments have been developed. The procedures for the presence of interaction between treatments and places or years have also been evolved.

REFERENCES

[1] Cochran, W.G. (1937). The efficiencies of the binomial series test of significance of a mean and of correlation coefficient. J.R.S.S. (C); 69-73.

[2] Hilda Frost Dunlop (1931). An empirical determination of the distribution of means, standard deviations and correlation coefficients drawn from rectangular populations. *Ann. Math. Stat.;* 66-81.

[3] Koch, G.G. (1970). The use of non-parametric methods in statistical analysis of a complex split-plot experiment *Biometrics;* **26**; 105-128.

[4] Rai, S.C. (1981). An analysis of ordered observations in block designs. *Jl. Ind. Soc. Agri. Stat.;* **33**; 7-14.

[5] Rai, S.C. and Rao, P.P. (1980). Use of ranks in groups of experiments. *Jl. Ind. Soc. Agri. Stat.*, **32**; 25-32.

[6] National Index of Agricultural Field Experiments. 1960-65 : *IARS.*