USE OF RANKS IN GROUPS OF EXPERIMENTS

BY

S.C. RAI AND P.P. RAO

I.A.S.R.I., New Delhi-12

(Received : December, 1978)

1. INTRODUCTION

In large-scale experimental programmes it is necessary to repeat the trial of a set of treatments at a number of places and in a number of years. The aim of such repetetions is to study the susceptibility of the treatment effects to place and climatic variations. More generally the aim is to find out treatments suitable for particular tracts in which case the trials are carried out simultaneously at a number of locations situated in the region. For drawing valid conclusions in regard to the suitability of treatment effects, it becomes necessary to make 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 the interaction present;

(*ii*) the error variances homogeneous and the interaction absent;

(*iii*) the error variances heterogeneous and interaction present and

(*iv*) the error variances heterogeneous and interaction absent.

The combined analysis can be done by using analysis of variance technique under the situations (*i*) to (*iii*) mentioned above. The technique of analysis of variance is not valid when the errors are and Treatment × year interaction is absent. It is reported that in heterogeneous agricultural experimentation about 30% of the trials fall in this cotegory (Rao [8]). As a way out for overcoming the situation of errors being heterogeneous and interaction absent, transformation of data into a suitable scale was suggested. However, this does not offer complete and satisfactory solution as it is very difficult to find out the right type of transformation for a given set of data. It is therefore necessary to find an alternative to the method of analysis of variance to draw fairly accurate inferences regarding treatments. In this paper we have proposed the method of analysis of groups of experiments by ranking the individual observations of different treatments in each replication. This method is quite useful particularly for the trials when the error variances are not homogeneous and analysis of variance technique is not valid.

Friedman [2] had proposed a test which is useful when the measurement of the variable is in at least ordinal scale. It tests whether K related samples could probably have come from the same population with respect to mean ranks. Kruskal-wallis [6] whether had proposed a test where all the observations from different samples are considered together for the purpose of ranking. The test deals with the totals of such ranks for individual samples for testing whether these samples came from the same population or not. However it does not consider the situation as arising in the case of a Randomised Block design in agriculture or animal experiments in which all treatments are blocked in a single replicate and the number of blocks or replications form the size of the sample for each treatment. In the present paper, we are investigating on the results obtained from the experiments conducted at different places or during different years.

2. The Procedure of Analysis

The procedure involves first ranking the observations in each block (replication) of the individual experiment. If t treatments are compared in a block, the individual observation is ranked by giving rank 1 to the highest value, 2 to the next lower and so on. The smallest value of the observation will be given rank t. Ranking is done afresh for each block and it will have variate value $1, 2, \ldots, t$. On the hypothesis that there is no significant difference between the treatments the difference in the values in each block for different treatments will arise solely from sampling fluctuations. The rank entered for a particular treatment would then be a matter of chance. In repeated observations of various blocks, each of the numbers 1 to t would appear with equal frequency.

The set of ranks for each treatment would represent a random sample from the discontinuous rectangular distribution of $1, 2, \ldots, t$ and the rank totals for various treatments would be same under the hypothesis of equality of treatments effects. If this hypothesis is false, then the rank totals would vary from one treatment to another.

Let there be t treatments each replicated r times in a particular trial. This trial is repeated over years (or places). If the character under study is independent of the block, the set of ranks r_{ijk} being the value of j-th treatment in the i-th replication of k-th experiment for each treatment would represent a random sample of rp items from a discontinuous rectangular universe 1, 2, . . ., t. The means and variance of this universe are obtained as $(t+1)/2$ and $(t^2-1)/12$ respectively. The next step in the procedure is to obtain the mean rank $\bar{R_j}\,[=(1/rp)\sum_i \sum_k r_{ijk}]$ of each treatment. These means are all estimates of the same rectangular universe. Now as it is well known that the sampling distribution of means of samples drawn from a rectangular universe approaches normality very rapidly as given by Hilda [3], the sampling distribution of the means of the ranks will be approximately normal. The sampling distribution of the mean ranks $\bar{R_j}$ will have the mean value $\bar{R}$ which is equal to $(t+1)/2$ and the variance σ^2 which is equal to $(t^2-1)/12\, rp$. Since the true mean and true standard deviation of the chance universe are known, the hypothesis that the means of the ranks of various treatments come from a single homogeneous normal population can be tested by the statistic

$$K=\sum_{j=1}^{t}(\bar{R_j}-\bar{R})^2/\sigma^2 \qquad \ldots(1)$$

By putting the values of $\bar{R}$ and σ^2 and taking $R_j = rp\,\bar{R_j}$ where R_j is the sum of ranks of the j–th treatment, we get the value of K as

$$K=\frac{12}{rp(t^2-1)}\sum_{j=1}^{t}R_j^2-\frac{3rpt(t+1)}{(t-1)} \qquad \ldots(2)$$

This statistic is distributed as χ^2 with $(t-1)$ d.f. for large rp as it is the sum of squares of standardised normal variates. Now if K if significantly greater than what might reasonably have been expected from chance, we can conclude that the mean ranks averaged over years or places differ significantly and there is significant difference in the treatment effects. The χ^2—value representing the treatment X year (or place) interaction may be obtained by the expression

$$K'=\frac{12}{r(t^2-1)}\left[\sum_{k=1}^{p}\sum_{j=1}^{t}\left(\sum_{i=1}^{r}r_{ijk}\right)^2-\frac{1}{p}\sum_{j=1}^{t}R_j^2\right] \ldots(3)$$

Which is distributed as χ^2 with $(t-1)\ (p-1)$ d.f. The significance of this value indicates the presence of interaction of treatments with years (or places).

Comparison between years (or places) or between blocks within years is not done in the method of ranks analysis since the totals of all ranks in any experiment is the same for all blocks. The treatment means can be compared with the help of the rank means $\bar{R}_j$. The C.D. for comparison of any two treatment rank means is given by (at 5% level)

$$\sqrt{\frac{(t^2-1)}{6rp}} \times 1.96 \qquad .. (4)$$

For comparing the treatment main effects, the sample size for each treatment is rp which is generally sufficiently large even for few replications per experiment. In order to test the interaction 'treatment X year (or places)', it is necessary that the number of replications per-experiment should be moderately large. Friedman [2] had shown that for $t=4$ or more, even four replications are sufficient for the distribution of Statistic k to follow χ^2—distribution. It is therefore, recommended that this method may be applied in the cases where the number of replications per experiment is four or more.

3. ILLUSTRATION

As an illustration of the method described above the data of an agricultural field experiment conducted at three Agricultural Research Stations in Gujrat during 1963 are taken. Five different organic manures ($T_1, \ldots, T_5$) were applied and their effects were studied on the yield of paddy. Experiments were conducted in randomised block design with six replications at each centre. The error variances were heterogeneous and treatment X place interaction was absent and hence no conclusion could be drawn by using analysis of variance technique. For application of this procedure the observations in each block are ranked for different treatments and their sums of ranks alongwith the values of K are presented below for each centre.

Sums of ranks and the value of K

Centre No.	*Sums of ranks* (R_j)					K
	T_1	T_2	T_3	T_4	T_5	
I	23	12	13	19	23	9.33
II	16	10	24	15	25	13.50
III	25	17	16	13	19	6.67
Over all the places	64	39	53	47	67	15.11

Now K for over all the the places is distributed as χ^2 with 4 degeees of freedom. This is significant at 0.01 probability level indicating that there is significant difference in the treatment effects.

The K statistic for Treatment X years interaction is obtained as 29.50-15.11=14.39 This K statistic is distributed as χ^2 with 8 d.f. This is not significant at 5 percent level. So it may be inferred that the treatment differences are consistent with places.

Now we will attempt to find out the best treatment. Mean of the treatment ranks is normally distributed with standard deviation

$$\sqrt{(t^2-1)/12\,rp}.$$

For comparing two treatments we have to find out the standard error of difference of the treatment means which is given by

$$\sqrt{(t^2-1)/6\,rp}.$$

In the present example the standard error of difference between two treatment means is 0.4714. Arrange the means of the treatment ranks in ascending order of their magnitude and compare their difference with the C.D. = 0.923 The treatment means are presented below :

T_2	T_4	T_3	T_1	T_5
2.167	2.611	2.944	3.555	3.722

It is observed that treatments 2 and 4 are significantly superior to treatments 1 and 5 at 5 percent level of significance. Therefore treatments 2 and 4 may be recommended for adoption.

4. Efficiency

It is evident that the procedure described above does not utilise all the information furnished by the data since it uses only ranks and does not use the quantitative value of the observation. It is desirable to obtain some notion about the amount of information lost in the above procedure in situations where the analysis of variance provides the proper test. When $t=2$, this procedure is equivalent to the binomial series test of significance of a mean difference. Cochran [1] had already established that the binomial series test of a mean difference has an efficiency of 63.7 percent. It follows that this method when $t=2$ has an efficiency of 63.7%. This provides a measure of the minimum efficiency of the procedure beeause when $t=2$, a classification in terms of greater or smaller is substituted for the exact quantitative measurement; as t increases a more and more finely sub-divided scale is substituted for the exact measurements. Therefore it seems reasonable that the loss in information through using ranks decreases as t increases.

In the absence of any theoretical analysis we present here the results obtained by analysing the data both by the method of usual analysis of variance and by this technique. The groups of experiments where analysis of variance is a valid test have been taken for this study. A comparison of the results will, of course, offer no conclusive evidence as to the relative efficiency of the two methods, but it will at least suggest whether the loss of information in using this method is so great as to vitiate completely its usefulness. The results obtained from 12 groups of experiments collected under the project of 'National Index of Field Experiments' are presented below :

Camparison of Analysis of variance and Ranking Method

Reference	*Analysis of variance*		*Ranking Method*	
	Value of F	*Significance level*	*Value of K*	*Significance level*
Mh. 64 (243), 65 (185)	3.02	0.01	38.79	0.01
Mh. 64 (246), 65 (197)	2.37	0.05	28.12	N.S.
Mh. 61 (228), 62 (225)	5.81	0.01	21.87	0.01
Mh. 63 (294), 64 (255), 65 (223)	13.81	0.01	28.00	0.01
Mh. 62 (169), 63 (219)	5.04	0.01	22.90	0.05
Mh. 64 (233), 65 (171)	1.84	0.05	32.70	N.S.
Mh. 62 (211), 63 (272)	29.52	0.01	62 50	0.01
Mh. 61 (199), 63 (293)	6.83	0.01	13.60	0.05
Mh. 64 (188), 65 (23) V_1	5.96	0.05	9.75	0.05
Mh. 64 (188), 65 (23) V_2	4.89	0.05	12.75	0.05
Mh. 64 (188), 65 (23) V_3	17.09	0.01	15.00	0.01
Mh. 64 (188), 65 (23) V_4	6.90	0.05	12.75	0.05

It may be seen from the above table that out of 12 groups of experiments, 7 indicated the presence of treatment difference at 1 percent level of significance by the variance Ratio F test. The analysis by ranking method indicates only 5 groups of experiments where treatment differences are significant at 1 percent level of significance. The remaining 2 experiments gave significance of treatment effects at 5 percent level of significance. Treatment effects were observed to be signifficantly different from one an other at 5 percent level of signifficance for 5 groups of Experiments in case of analysis of variance F test. Out of these 5 groups of experiments 3 gave the same result by the ranking method. In case of 2 groups of experiments, the results were not significant even at 5 percent level of significance in the ranking method. The difference in these cases by the two methods is observed to be present where the results are just on the border of significance level. These results give a good idea of the powerful parameteric F test. In 8 of the 12 cases analysed, the probability levels yielded by the two tests were essentially the same.

5. Conclusions

The method described in the paper uses solely information on 'order' and makes no use of the quantitative values of the variate 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 can not validly be applied. The main step in the application of this method is to rank the individual observations and calculate the value of k from the table of ranks. The sampling distribution of this statistic approaches the χ^2 distribution as the number of sets of ranks increases. The computation involved for obtaining this statistic is very simple and less time consuming. The theoretical discussion of the efficiency of this procedure relative to the analysis of variance indicates that in situations where the latter method can validly be applied and when the number of sets of ranks is large, the maximum loss of information in this method is 36 percent. The amount of information lost appears to be greatest when there are only two ranks in each set and the loss decreases as the number of treatments increases. The application of this method and analysis of variance to the same body of the data provides further evidence as to its relative efficiency. In 8 out of 12 cases, the probabilily levels yielded by the two tests were essentially the same. For the 12 sets of groups of experiments, the ranking method has rejected H_0 in 5 cases at 1 percent level of significance while F test provides the rejection of 7 cases at this significance level.

References

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

[2] Friedman, M. (1937) : The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Amer. Stat. Assoc.*, 675-701.

[3] Hilda Frost Dunlap (1931) : An empirical determination of the distribution of means, Standard deviations and correlation coefficients drawn from rectangular populations. *Ann. Math. Stat.* 66-81.

[4] Kendall, M. G. (1948) : *Rank Correlation Methods*. London : Griffin.

[5] Kendall, M.G. and Smith, B.B. (1939) : The problem of *m*-rankings. *Ann. Math. Stat.*, 275-287.

[6] Kruskal, W.H. and Wallis W. A. (1952) : Use of ranks in one criterion variance analysis. *J. Amer. Stat. Assoc.* 583-621.

[7] Rai, S.C. (1976) : A model for rank analysis in triad comparisons. *J. Ind. Soc, Agri. Stat.* 89-98.

[8] Rao, P. P. (1975) : An investigation into the causes of and remedical measures for heterogeniety of error variances in groups of Agricultural Field Experiments, Diploma Thesis, IARS.

[9] I.A.R.S. : National Index of Agricultural Field Experiments. 1960-65,