Jour. Ind. Soc. Ag. Statistics
Vol. XXXVII, No. 3 (1985), pp. 244-251

PAIRED AND TRIAD COMPARISONS IN SENSORY EVALUATION

S. C. RAI
Indian Agricultural Statistics Research Institute, New Delhi-110012

(Received : January, 1982)

SUMMARY

Methods of analysis of sensory evaluation experiments in paired and triad comparison designs have been discussed. The procedures permit test of hypothesis of equality of treatment ratings or preferences. We assume, in the null hypothesis that treatment ratings are equal where as the alternative hypothesis does not make any assumption regarding the equality of treatment preference. The procedure is very general and it can be used with simple calculations. It can be used to wide range of observations as it does not make any assumption regarding normality of data. The procedure is illustrated through numerical examples.

Keywords : Sensory evaluation; ranks; Chi-square; Ordinal scale.

Introduction

Sensory evaluation is assuming increasing importance in our day to day life. Generally the experiments for taste-testing or other sensory evaluations are conducted in paired or triad designs, where observations are recorded mostly in ordinal scale. Appropriate statistical methods for analysing such experiments are essential.

The analysis of data recorded in ordinal scale has received considerable attention in statistical and psychological fields. In psychology, emphasis is given to the problem of scaling and in statistics, effort is made on testing and developing various models of analysis. Bradley and Terry [1], Pendergrass and Bradley [4] have developed models for analysing ranking data in paired and triad comparisons respectively. Sadasivan and Rai [8] and Rai [5] have proposed models for fractional paired and triad comparisons respectively. Rai [6] and Win and Rai [10] have given models for triad comparison for analysing rank data. Gupta and Rai [3] developed a model of rank analysis in paired comparisons. Rai [7] has given a very general model of rank analysis in incomplete block designs.

The models mentioned above require the solutions of normal equations and some times it becomes very difficult to obtain the solution easily. Durbin [2] has proposed a model for rank analysis in incomplete block designs. The test statistic developed by Durbin follows Chi-square distribution approximately. In the present paper, models are developed for analysing rank data in paired and triad comparisons on the lines of Durbin.

2. Paired Comparisons

In paired comparison model, the performance of t treatments are compaired in pairs. There will be $\binom{t}{2}$ pairs. Both the treatments of a pair are presented one after the other to the judge for quality evaluation. The judge will allot the rank 1 to the treatment of superior quality and 2 to the other treatment. The entire process may be repeated a number of times for all treatments. The test statistic for testing the null hypothesis

$$H_0 : T_i = T_j \text{ for all } i \text{ and } j,\ i \neq j$$

against

$$H_1 : T_i \neq T_j \text{ for at least one } i \text{ and } j,$$

is given by

$$T = \frac{4}{nt} \sum_{j=1}^{t} R_j^2 - 9n(t-1)^2 \tag{1}$$

where t = number of treatments

n = number of repetitions

R_j = sum of the ranks of jth treatment from all pairs and repetitions.

The distribution of the test statistic T given at (1) is found under the assumption that each arrangement of 2 ranks (1 or 2) within a pair is equally likely because of no difference between treatments under the null hypothesis. There are 2 equally likely ways of arranging the ranks in each pair and there are $t(t-1)/2$ pairs in a repetition. Hence each arrangement of ranks over the entire array of $t(t-1)/2$ pairs is equally likely and has probability $(\frac{1}{2})^{t(t-1)/2}$. The test statistic may be calculated and its probability may be obtained.

The distribution of test statistic can be approximated to chi-square distribution with $(t-1)$ d.f. if the number of repetitions n is large. The number of times a pair is presented to a judge for evaluation is called repetition. Sum of Ranks R_j of the jth treatment, $(j = 1, 2, \ldots, t)$ is approximately normally distributed according to central limit theorem for large n. The random variable

$$[R_j - E(R_j)]/\sqrt{\text{var}(R_j)} \tag{2}$$

has approximately a standard normal distribution and hence statistic

$$T^1 = \sum_{j=1}^{t} [R_j - E(R_j)]^2/\text{var}(R_j) \tag{3}$$

is approximately distributed as χ^2 with t d.f., if R_j's $(j = 1, 2, \ldots, t)$ are independent. But sum of ranks of all the treatments are not independent because their sum is fixed as

$$\sum_{j=1}^{t} R_j = 3\,n\,t\,(t-1)/2 \tag{4}$$

Hence the knowledge of $(t-1)$ of the R_j enables us to state the value of the remaining R_j. Durbin [2] has shown that multiplication of T^1 by $(t-1)/t$ results in a statistic which is approximately distributed as chi-square with $(t-1)$ degrees of freedom. Therefore the statistic

$$T = \frac{t-1}{t}\,T^1 = \frac{t-1}{t}\sum_{j=1}^{t}\frac{[R_j - E(R_j)]^2}{\text{var}(R_j)} \tag{5}$$

is distributed as χ^2-distribution with $(t-1)$ d.f. Now the values of $E(R_j)$ and $\text{Var}(R_j)$ are to be found out in order to transform (5) into the usual form given in (1).

Let $R(X_{ijk})$ denote the rank of T_j when it is compared with T_i, $(i \neq j = 1, 2, \ldots, t)$ in the kth repetition $(k = 1, 2, \ldots, n)$. The sum of ranks R_j is the sum of independent random variables $R(X_{ijk})$ and is given by

$$R_j = \sum_{k=1}^{n} \sum_{i \neq j=1}^{t} R(X_{ijk}) \tag{6}$$

Therefore

$$E(R_j) = \sum_{k=1}^{n} \sum_{i \neq j=1}^{t} E[R(X_{ijk})] = \frac{3n}{2}(t-1) \tag{7}$$

and

$$\text{Var}(R_j) = \sum_{k=1}^{n} \sum_{i \neq j=1}^{t} \text{Var}[R(X_{ijk})]$$
$$= \frac{n}{4}(t-1) \tag{8}$$

Substituting the values of $E(R_j)$ and Var (R_j) to (5) we have

$$T = \frac{t-1}{t} \sum_{j=1}^{t} \frac{\left[R_j - \frac{3n}{2}(t-1)\right]^2}{\frac{n}{4}(t-1)} \tag{9}$$

After simplification and substituting the value ΣR_j, (9) can be shown as equivalent to (1). Hence the test statistic T given at (1) follows a χ^2-distribution with $(t-1)$ d.f. for large number of repetitions.

If the number of judges is more than one as in the case of Consumer Surveys, the value of T would be computed for each judge. Let the number of judges be m and the value of T for pth judge, $(p = 1, 2, \ldots, m)$ is T_p. Then for testing over all preference, the test statistic would be $\sum_{p=1}^{m} T_p$ which is distributed as χ^2 with $m(t-1)$ d.f. because each T_p is independently distributed as χ^2 with $(t-1)$ d.f.

3. Triad Comparisons

In triad comparisons, the performance of treatments is compared by taking three at a time. All the $\binom{t}{3}$ triplets are presented to a judge in a random order and the judge is required to rank each member of a triplet on the basis of character under study. Rank 1 is allotted to the best member of the triplet, 2 to the middle one and 3 to the last preferred member. Each triplet is repeated n times.

The test statistic for testing the significance of equality of treatment effects, can be obtained. Consider the null hypothesis

$$H_0 : T_i = T_j \text{ for all } i \text{ and } j,\ i \neq j$$

against

$$H_1 : T_i \neq T_j \text{ for at least one } i \text{ and } j.$$

The test statistic is given by

$$S = \frac{3}{n\,t(t-2)} \sum_{j=1}^{t} R_j^2 - 3\,n(t-1)^2(t-2) \tag{10}$$

where t = number of treatments
n = number of repetition
R_j = Sum of ranks of the jth treatment
$(j = 1, 2, \ldots, t)$

Probabilities of obtaining the value of S under the null hypothesis may be worked out by the procedure described in earlier section for Paired Comparisons for small n and the hypothesis of equality of treatment effects may by tested. Here the distribution of the test statistic S when n is large can be found out.

For large n, R_j is approximately normal and the random variable

$$\frac{R_j - E(R_j)}{\sqrt{\text{Var}\,(R_j)}}$$

has approximately a standard normal distribution. Therefore the statistic

$$\frac{t-1}{t} \sum_{j=1}^{t} \frac{[R_j - E(R_j)]^2}{\text{Var}(R_j)} \tag{11}$$

is distributed as χ^2 with $(t - 1)$ degrees of freedom, (Durbin [2]). Here all R_j's are not independent and they are restricted by the relation

$$\sum_{j=1}^{t} R_j = nt\,(t-1)\,(t-2) \tag{12}$$

Let $R(X_{ijkp})$ be the rank of T_j when it is compared with T_i and T_k, $(i \neq k = 1, 2, \ldots, t)$ in the pth repetition. Then R_j is the sum of independent random variables $R(Xijkp)$ and is given by

$$R_j = \sum_{p=1}^{n} \sum_{i=1}^{t} \sum_{k=1}^{t} R(X_{ijk}p) \tag{13}$$

$$(i \neq k \neq j)$$

The values of $E(R_j)$ and Var (R_j) may be obtained as

$$E(R_j) = \sum_{p=1}^{n} \sum_{i=1}^{t} \sum_{k=1}^{t} E[R(X_{ijk}p)],\ (i \neq k) = n(t-1)\,(t-2) \tag{14}$$

$$\text{Var}\,(R_j) = \sum_{p=1}^{n} \sum_{i=1}^{t} \sum_{k=1}^{t} \text{Var}\,[R(X_{ijk}p)],\ (i \neq k)$$

$$= \frac{n}{3}(t-1)\,(t-2) \tag{15}$$

Substituting the values of $E(R_j)$ and Var (R_j) in (11), we have the expression as

$$\frac{t-1}{t}\sum_{j=1}^{t}\frac{[R_j - n(t-1)(t-2)]^2}{\frac{n}{3}(t-1)(t-2)} \tag{16}$$

Expression (16) can be shown to be the same as given for the test statistic S in (10).

Hence S is distributed like a χ^2 distribution with $(t-1)$ degrees of freedom for large value of n.

4. Numerical Illustrations

In order to illustrate the procedures developed in the paper, two numerical examples, one for paired comparisons and the other for triad comparisons are included.

Example: I

Four brands of mango juice were compared in paired comparison for their taste quality. (Win [10]). Brands were as follows :

T_1 = Sun-Sip mango juice
T_2 = Mohun's mango juice
T_3 = Noga mango juice
T_4 = Dipy's mango juice

There were 6 pairs namely $T_1 T_2$, $T_1 T_3$, $T_1 T_4$, $T_2 T_3$, $T_2 T_4$ and $T_3 T_4$. Each brand of mango juice of a pair was presented to a judge for tasting and giving ranks. Rank 1 was given to the mango juice which was found better and rank 2 was allotted to the other brand. In this way each pair was presented 40 times. Table 1 gives the frequency of a brand of mango juice getting first or second rank alongwith the sum of ranks.

TABLE 1—PREFERENCE MATRIX AND SUM OF RANKS

Treatments	*Number of times ranked as*		*Sum of ranks*
	First	*Second*	*(R_j)*
T_1	81	39	159
T_2	58	62	182
T_3	60	60	180
T_4	41	79	199

For testing the null hypothesis regarding the equality of treatment effects,

$H_0 : T_i = T_j$ for all i and j, ($i \neq j = 1, 2, 3, 4$)

against

$H_1 : T_i \neq T_j$ for at least one i and j. Calculate the value of T given in (1) by using R_j values from the above table.

$T = 20.1$

Since T is distributed as χ^2 with 3 d.f., it is found that calculated value of T is significant at 1% level, indicating the rejection of the null hypothesis. This indicates that treatments differ significantly from one another.

Example : 2

Chapatis prepared from different wheat varieties were compared in a taste-testing experiment conducted in the Cereal Laboratory of I.A.R.I., New Delhi in the year 1973. The varieties were as follows :

T_1 = Sharbati Sonara
T_2 = Sonalika
T_3 = K-65
T_4 = C-306
T_5 = K-68

The design followed was triad comparison and each triplet was repeated 50 times. The judge was to allot rank 1 to the best treatment, rank 2 to the second best and rank 3 to the last preferred treatment in a triplet. The data on frequency distribution of different ratings are given in Table 2.

TABLE 2—FREQUENCY DISTRIBUTION ON DIFFERENT RATINGS AND SUM OF RANKS

Number of times the treatment is allotted			*Treatments*		
	T_1	T_2	T_3	T_4	T_5
Rank 1	120	90	105	80	105
Rank 2	105	88	115	90	102
Rank 3	75	122	80	130	93
Sum of Ranks R_j	555	632	575	650	588

For testing the significance of null hypothesis regarding equality of treatment effects, we calculate the value of test-statistic S given in (10) by using the values of sum of ranks R_j. The value of S is obtained as 25.2 which is distributed as χ^2 with 4 d.f. This value is significant at 1% level which indicates that treatments differ significantly from one another.

REFERENCES

[1] Bradley, R. A. and Terry, M. E. (1952) : Rank analysis of incomplete block designs I—The method of paired comparisons, *Biometrika*, **39** : 324-345.

[2] Durbin, J. (1951) : Incomplete blocks in ranking experiments, *British Jl. Psycho.* (Stat. Section), **4** : 85-90.

[3] Gupta, S. C. and Rai, S. C. (1980) : Rank analysis in paired comparison designs. *Jl. Ind. Soc. Agr. Stat.*, **32** : 87-98.

[4] Pendergrass, R. C. and Bradley, R. A. (1960) : Ranking in triple comparisons. *In* I. Olkin *et al.* (eds.), *Contribution to Prob. and Stat.* Stanford University Press, Stanford, California : 331-351.

[5] Rai, S. C. (1971) : *Ranking in fractional triad comparisons. Jl. Ind. Soc. Agr. Stat.*, **23** : 52-61.

[6] Rai, S. C. (1976) : A model for Rank analysis in Triad Comparisons. *Jl. Ind. Soc. Agr. Stat.*, **28** : 89-98.

[7] Rai, S. C. (1981) : *Rank Analysis in Incomplete Block Designs.* I.A.S.R.I. publication.

[8) Sadasivan, G. and Rai, S. C. (1973) : A Bradley-Terry model for Standard Comparison pairs. *Sankhya*, **35** : 25-34.

[9] Win, Kyi (1976) : Analysis of Experiments involving Rankings in Triad Comparisons. M.Sc. (Agril. Stat.) *Thesis*, I.A.R.I., New Delhi.

[10] Win, Kyi and Rai, S. C. (1979) : Analysis of experiments involving rankings in triad comparisons. *Jl. Ind. Soc. Agr. Stat.*, **31** : 97-110.