

A NOTE ON TIES IN STANDARD PAIRS IN FRACTIONAL PAIRED COMPARISONS

BY

S.C. RAI

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

(Received : May, 1979)

1. INTRODUCTION

In paired-comparison taste-testing experiments, judges are allotting ranks to each member of the pair on the basis of taste quality of the treatments. A judge is tasting the samples one after the other in a pair and express his preference for one of the samples tasted. Some times it becomes extremely difficult to prefer one treatment over the other. The inability of a judge in such cases may be due to the following reasons.

1. The perception of the judge is not sharp enough to detect the treatment difference.
2. The treatments do not differ in the quality judged.

Most of the paired comparison models do not permit a judge for declaring a tie. Without going into details of merits or demerits of the models, it can be said that the procedures which do not allow the possibility of ties, are not making the full use of the information contained in the non-preferenced class. Rao and Kupper (1967) have proposed modifications in the Bradley-Terry model for applying it to tied observations. In the present paper, we are extending Rao-Kupper model when only a fraction of the pairs are considered and ties are permitted. The model is also valid for testing n 'varieties' against a standard.

2. MATHEMATICAL MODEL

Sadasivan and Rai (1973) have described a model for analysis of experiments in standard comparison pairs without permitting ties,

The model can be specified as follows :

- (i) The t treatments have true treatment ratings $\pi_1, \pi_2, \dots, \pi_t$ on a subjective continuum such that $\pi_i \geq 0$ ($i=1, 2, \dots, t$) and

$$\sum_{i=1}^t \pi_i = 1$$

- (ii) When a particular treatment (say) T_1 is compared with T_i in paired comparisons, the probability that T_1 gets top ranking (i.e. T_1 is preferred to T_i) is

$$\pi_{1i} = \frac{\pi_1}{\pi_1 + \pi_i} \quad \dots(2.1)$$

Here we will modify this model to permit judges for declaring ties also. Following the arguments given by Bradely (1953) and Rao and Kupper (1967), we can write (2.1) as

$$\pi_{1.i} = \frac{\pi_1}{\pi_1 + \theta \pi_i}, \quad \dots(2.2)$$

$$\pi_{i.1} = \frac{\pi_i}{\theta \pi_1 + \pi_i} \quad \dots(2.3)$$

and

$$\pi_{0.1i} = \frac{\pi_1 \pi_i (\theta^2 - 1)}{(\pi_1 + \theta \pi_i)(\theta \pi_1 + \pi_i)} \quad \dots(2.4)$$

where $\pi_{1.i}$ is the probability that T_1 is preferred over T_i , $\pi_{i.1}$ is the probability that T_i is preferred over T_1 and $\pi_{0.1i}$ is the probability of the tie.

3. ESTIMATION OF π_i 'S AND θ

We shall obtain the maximum likelihood estimators of π_i 's and θ and also consider their limiting joint distribution. In the paired comparisons only those pairs are considered where T_1 occurs. The number of such pairs for t treatments will be $(t-1)$. The judge is asked to rank the treatments in pairs and give rank 1 to preferred treatment and zero to non-preferred. If the judge is not able to distinguish between the treatments of a pair, he will declare 'tie' and allot the rank of zero to each member of the pair. Let $r_{0.1i_k}$ be the number of tie when T_1 is compared to T_i in the k -th replication. Similarly $r_{1.i_k}$ ($r_{i.1_k}$) is 1 if T_1 (T_i) is preferred and 0 otherwise. Also let $b_{1i_k} = r_{0.1i_k} + r_{1.i_k}$ and $b_{i1_k} = r_{0.1i_k} + r_{i.1_k}$

Therefore, $b_{1t} = \sum_k b_{1tk}$ and

$$b_{i1} = \sum_k b_{i1k} \text{ (for } i=2, 3, \dots, t)$$

and

$$b_1 = \sum_{i=2}^t b_{1i}$$

Now the probability of the observed results in the k -th repetition of the pair (T_1, T_i) is given by

$$(\theta^2 - 1)r_{0.1tk} \left(\frac{\pi_1}{\pi_1 + \theta\pi_i} \right) b_{1tk} \left(\frac{\pi_i}{\theta\pi_1 + \pi_i} \right) b_{i1k} \dots (3.1)$$

because if T_1 is preferred to T_i then $r_{1.ik} = 1$ and $b_{1tk} = 1$ and $r_{i.1k} = 0$ and $b_{i1k} = 0$ and as there is no tie, the value of $r_{0.1tk} = 0$. The value of probability given by (3.1) becomes $\frac{\pi_1}{\pi_1 + \theta\pi_i}$ which is the same as given (2.2). If T_i is preferred to T_1 , then (3.1) reduces to (2.3) and if there is tie then (3.1) becomes the same as given in (2.4). Assuming the probability independence between each pair and also between replications, we can obtain the likelihood function of the observed outcome as given below:

$$L(\pi_1, \pi_2, \dots, \pi_t; \theta) = (\theta^2 - 1) \sum_{i=2}^t \sum_{k=1}^r r_{0.1tk} \frac{\pi_1}{\pi_1 + \theta\pi_i} b_{1tk} \frac{\pi_i}{\theta\pi_1 + \pi_i} b_{i1k} \dots (3.2)$$

we will obtain the maximum likelihood estimates of π_i 's and θ by maximising $\log L$ given in (3.2) subject to the condition that

$$\sum_{i=1}^t \pi_i = 1. \text{ The normal equations will be}$$

$$\left. \begin{aligned} \frac{b_1}{\pi_1} - \sum_{i=2}^t \frac{b_{1i}}{\pi_1 + \theta\pi_i} - \theta \sum_{i=2}^t \frac{b_{i1}}{\theta\pi_1 + \pi_i} &= 0 \\ \frac{b_{i1}}{\pi_1} - \frac{\theta b_{1i}}{\pi_1 + \theta\pi_i} - \frac{b_{i1}}{\theta\pi_1 + \pi_i} &= 0 \text{ for } i=2, 3, \dots, t \\ \text{and } \frac{2\theta \sum_{k=1}^r \sum_{i=2}^t r_{0.1tk}}{\theta^2 - 1} - \sum_{i=2}^t \frac{b_{1i}\pi_i}{\pi_1 + \theta\pi_i} - \pi_i \sum_{i=2}^t \frac{b_{i1}}{\theta\pi_1 + \pi_i} &= 0 \end{aligned} \right\} (3.3)$$

To solve this system of equations, it is convenient to write (3.3) in the following form.

$$\left. \begin{aligned}
 \pi_1 &= b_1 \left[\sum_{i=2}^t \frac{b_{1i}}{\pi_1 + \theta \pi_i} + \theta \sum_{i=2}^t \frac{b_{i1}}{\theta \pi_1 + \pi_i} \right]^{-1} \\
 \pi_i &= b_{i1} \left[\frac{\theta b_{1i}}{\pi_1 + \theta \pi_i} + \frac{b_{i1}}{\theta \pi_1 + \pi_i} \right]^{-1} \quad \text{for } i=2, 3, \dots, t \\
 \theta &= \frac{(\theta^2 - 1)}{2 \sum_{k=1}^r \sum_{i=2}^t r_{0,ik}} \left[\sum_{i=2}^t \frac{b_{1i} \pi_i}{\pi_1 + \theta \pi_i} + \pi_1 \sum_{i=2}^t \frac{b_{i1}}{\pi_1 \theta + \pi_i} \right]
 \end{aligned} \right\} (3.4)$$

The solution of these equations will give the values of π_i and θ . The equations can be solved by using the iterative procedure. The procedure should continue until the agreement between $p_i^{(j-1)}$ and $(p)_i^{(j)}$ and between $\theta^{(j-1)}$ and $\theta^{(j)}$ is sufficiently close where $p_i^{(j-1)}$ and $p_i^{(j)}$ are the estimates of π_i obtained at $(j-1)th$ and $j-th$ iterations respectively and $\theta^{(j-1)}$ and $\theta^{(j)}$ are the estimates of θ at $(j-1)th$ and $j-th$ iterations respectively. The first trial value of θ may be taken as $\frac{2r(t-1)}{n} - 1$ where n is the number of times there was no ties

and the π_i 's may be taken in proportion to $\sum_k r_{i,ik}$. In most of the cases, these initial values give good approximation.

4. TEST OF HYPOTHESIS

Here we shall develop a method of testing equality of true treatment ratings. We shall test the null hypothesis ;

$$H_0 : \pi_i = \frac{1}{t}, \quad i=1, 2, \dots, t$$

against the alternative

$$H_1 : \pi_i \neq \pi_j \text{ for some } i \neq j$$

A large sample test based on the likelihood ratio

$$\lambda = \frac{\max_{\theta} L \left(\frac{1}{t}, \frac{1}{t}, \dots, \frac{1}{t}; \theta \right)}{L(p_1, p_2, \dots, p_t; \hat{\theta})} \quad \dots(4.1)$$

is obtained by using $\chi^2 = -2 \log \lambda$ as the test statistic. Under H_0 , χ^2 is distributed approximately like chi-square with $(t-1)$ degrees of freedom for large samples.

The expression for χ^2 is given below :

$$\begin{aligned}
 I^2 = 2 & \left[r(t-1) \log \frac{r(t-1)}{n} + n \log 2 + \sum_{i=2}^t \sum_{k=1}^r \right. \\
 & \left. r_{0,ik} \log \left\{ \frac{n(\hat{\theta}^2 - 1)}{\sum_{i=2}^t \sum_{k=1}^r r_{0,ik}} \right\} \right. \\
 & + b_{1i} \log p_1 + \sum_{i=2}^t b_{1i} \log p_i - b_{1i} \sum_{i=2}^t \log (p_1 + \hat{\theta} p_i) \\
 & \left. - \sum_{i=2}^t b_{1i} \log (\hat{\theta} p_1 + p_i) \right] \quad \dots(4.2)
 \end{aligned}$$

A test can be developed to test the adequacy of the model. The chi-square test of goodness of fit can be used for testing the appropriateness of the model. The procedures involved, required the computation of expected cell frequencies which can be obtained by using the estimators p_1, \dots, p_t given in section 3. χ^2 is obtained at $(t-2)$ d.f. by using sums of the terms of the form $(O-E)^2/E$ where O and E are respectively, the observed and expected frequencies in each cell.

The test for testing the threshold parameter $\theta = \theta_0$ and the experimental model for combination of experiments may be developed along the lines of Rao and Kupper (1967).

5. NUMERICAL ILLUSTRATION

In this section the use of the proposed model will be illustrated by application to experimental data. A taste-testing experiment using four brands of Bread was conducted in paired comparisons by taking fractional pairs (Zurmati and Rai 1977). For the purpose of illustration, we will use the data from that experiment. The data based on 30 replications are presented in the following table using the notation defined in Section 3.

TABLE 1
Data of Numerical example

Pairs (l, i)	$r_{1,i}$	r_{0-1t}	r_{i-1}	b_{1t}	b_{i1}
1, 2	17	5	8	22	13
1, 3	21	4	5	25	9
1, 4	14	6	10	20	16

After obtaining a good set of starting values of θ and p_i as explained in section 3, we will obtain the estimates of the parameters. Four iterations were required to get the precise estimates. The initial and final estimates of the parameters are given in the Table 2. below :

TABLE 2
Estimates of parameters

	θ	p_1	p_2	p_3	p_4
Initial value	1.40	0.4208	0.1983	0.1240	0.2479
Final value	1.25	0.3802	0.2116	0.1244	0.2838

A test for $H_0 : \pi_i = \frac{1}{t}$; $i=1, 2, \dots, t$ gave the value of $\chi^2=126.69$ which is significant at 3 d.f. for 0.01 probability level. This indicates that there is significant difference in the preference of various brands of bread. A test of goodness of fit of the model was carried out as outlined in section 4. The value of χ^2 was found to be 5.15 at 2 degrees of freedom which indicates an adequate fit for the data.

SUMMARY AND CONCLUSIONS

A model for analysis of paired comparisons when only fractional pairs are considered have been developed. The model accommodates presence of ties in ranking which is quite natural. The model given by Sadasivan and Rai (1973) does not make provision for ties. This model is modified by introducing an additional parameter called 'Threshold' parameter in the model. The threshold

parameter permits ties in the model. The method has been developed to estimate the parameters of the model. A test for testing the goodness of fit of the proposed model is also given. Some of the procedures developed in the paper, have been explained through a numerical example.

REFERENCES

- [1] Bradley, R.A. (1953) : 'Some Statistical Methods in Taste Testing and Quality Evaluation'. *Biometrics* 9, 22-38.
- [2] Bradley, R.A. (1954) : 'Incomplete Block Rank Analysis'. On appropriateness of the model for a method of paired comparisons. *Biometrics*, 10, 375-90.
- [3] Bradley, R.A. and Terry, M.B. (1952) : 'Rank analysis of Incomplete Block designs-I. The method of paired comparisons. *Biometrika*, 39, 324-45.
- [4] David, H.A. (1963) : 'The method of paired comparisons. Charles Griffin and Co.
- [5] Dykstra, O. Jr. (1960) : 'Rank analysis of Incomplete Block Design'. A method of paired comparisons; A method of paired comparisons employing unequal repetitions in pairs. *Biometrics*, 16; 176-88.
- [6] Glenn, W.A. and David, H.A. (1960) : 'Ties in paired—Comparisons experiments using a modified Thurstone—Mosteller Model, *Biometrics*, 16; 86-109.
- [7] Rai, S.C. (1971) : 'On model for rank analysis in fractional paired comparisons. *Jour. Ind. Soc. Agril. Stat.* 23; 61-68.
- [8] Rao, P.V. and Kupper, L.L. (1967) : 'Ties in paired—comparison experiments : A generalization of Bradley Terry model. *J.A.S.A.* 62, 194-204.
- [9] Sadasivan, G. and Rai, S.C. (1973) : 'A Bradley—Terry model for standard comparison pairs, *Sankhya*, 35; 25-34.
- [10] Zurmati, M.R. and Rai, S.C. (1977) : 'On Taste Quality of different brand of bread. *Jour. Ind. Soc. Agril. Stat.* 142-43.