JOURNAL OF COMBINATORIAL THEORY (A) 13, 162–169 (1972)

# Some Inconsistencies in Judging Problems

ROGER R. DAVIDSON AND ROBERT E. ODEH

*University of Victoria, Victoria, British Columbia, Canada*

*Communicated by Editor-in-Chief*

Received March 23, 1971

Suppose that $I$ individuals are ordered on the basis of the sums of ranks assigned independently by $J$ judges, and that there is a unique winner. If the winner is deleted and the ranks assigned to the remaining individuals are adjusted, then an ordering of the reduced set is obtained. This ordering is said to be consistent with the original ordering if the relative positions of all remaining individuals are unchanged.

An examination is made of conditions under which an individual, who was beaten by the winner and at least one other person in the original ranking, emerges as the winner in the reduced ranking. Such an inconsistency is called an interchange.

## 1. INTRODUCTION

In the classical problem of multiple rankings, one of the objectives is to obtain a consensus ranking of a group of individuals on the basis of the rankings assigned by a panel of judges. The problem of multiple rankings was first discussed by Friedman [1] and has been treated by Kendall in his book "Rank Correlation Methods" [4]. Alternative methods of obtaining a consensus ranking have been suggested by Kendall [4] and Kemeny and Snell [3].

Suppose that each of $J$ judges ranks a group of $I$ individuals in increasing order of preference, with his $k$-th choice receiving the rank $(I - k + 1)$, $k = 1,..., I$.[1] Let $A_j = (r_{1j} ,..., r_{Ij})$ be the ranking given by judge $j$, where $r_{ij}$ is the rank assigned to individual $i$ by judge $j$, $1 \leqslant i \leqslant I$, $1 \leqslant j \leqslant J$. Then Kendall [4] suggests that the consensus ranking be that based on the rank sums

$$r_i = \sum_{j=1}^{J} r_{ij}, \qquad 1 \leqslant i \leqslant I,$$

[1] This ranking system, which is symmetric to that used by Kendall [4], leads to simplifications in the computational aspects of the problem to be discussed.

which when reordered give

$$r_{(1)} \leqslant r_{(2)} \leqslant \cdots \leqslant r_{(I-1)} \leqslant r_{(I)},$$

the winner being the individual who receives the rank sum $r_{(I)}$. This ranking is optimal in the sense that it minimizes $\sum_{i=1}^{I} (r_i - iJ)^2$, the sum of squares of differences between the actual rank sums and what they would be if all $J$ rankings were identical (cf. [4, pp. 114–5]).

The methods of obtaining a consensus ranking which have been suggested by Kemeny and Snell [3] are based on a measure of distance between rankings. Suppose that in assessing the pair of individuals $(i, l)$, judge $j$ assigns a score

$$a_{(i,l)j} \quad \begin{cases} 1, & \text{if } i \text{ is preferred to } l, \\ -1, & \text{if } l \text{ is preferred to } i, \\ 0, & \text{if there is no preference.} \end{cases}$$

Then a measure of the distance between the rankings $A_j$ and $A_k$ is given by

$$d(A_j, A_k) = \frac{1}{4} \sum_{i=1}^{I} \sum_{l=1}^{I} | a_{(i,l)j} - a_{(i,l)k} |.$$

The measure $d(\cdot, \cdot)$ is the unique measure of distance which satisfies axioms 1, 2, and 3 given by Kemeny and Snell [3] and for which the minimum positive distance is $\frac{1}{2}$. (Note that the distance $d(\cdot, \cdot)$ is one-half of the distance used in [3] in that the latter assumes a minimum positive distance of 1).

A second measure of distance is

$$s(A_j, A_k) = \left[ \sum_{i=1}^{I} (r_{ij} - r_{ik})^2 \right]^{1/2}.$$

This measure fails to satisfy only one of the axioms required of $d(\cdot, \cdot)$, namely, the sufficient condition for equality in the triangle inequality (cf. [3, Axiom 1.3]). It should be noted that the distance $d(\cdot, \cdot)$ is a function of the distance $s(\cdot, \cdot)$, but that the correspondence is not one-to-one. The distance $s(\cdot, \cdot)$ serves as the basis for the Spearman rank correlation coefficient, namely,

$$\rho(A_j, A_k) = 1 - 6 \frac{s^2(A_j, A_k) + T_j + T_k}{(I^3 - I)},$$

where $T_j$ is a correction for ties in the ranking $A_j$. If a ranking contains a set of $t$ tied ranks then the sum of the ranks remains the same but the sum of squares of the ranks is reduced by $(t^3 - t)/12$. Thus the correction

$$T_j = (1/12) \sum_{t \in A_j} (t^3 - t),$$

where $\sum_{t \in A_j}$ denotes summation over all blocks of ties in the ranking $A_j$.

DEFINITION. Let $\{A_1, ..., A_J\}$ be the rankings assigned to a group of $I$ individuals by $J$ judges.

(i) A ranking $B$ is a *mean consensus ranking* if for every ranking $C$,

$$\sum_{j=1}^{J} d^2(A_j, B) \leqslant \sum_{j=1}^{J} d^2(A_j, C).$$

(ii) A ranking $B$ is a *median consensus ranking* if for every ranking $C$,

$$\sum_{j=1}^{J} d(A_j, B) \leqslant \sum_{j=1}^{J} d(A_j, C).$$

(iii) A ranking $B$ is a *rank sum consensus ranking* if for every ranking $C$,

$$\sum_{j=1}^{J} \rho(A_j, B) \geqslant \sum_{j=1}^{J} \rho(A_j, C).$$

The definitions of the mean and median consensus rankings are given by Kemeny and Snell [3], who illustrate the procedure for obtaining them. In practice, however, these two types of consensus rankings are difficult to obtain. On the other hand, it follows from Kendall [4, pp. 114–5] that the ranking determined by the rank sums $(r_{(1)}, ..., r_{(I)})$ is a rank sum consensus ranking.

In the remainder of this study attention is restricted to consensus rankings based on rank sums. Furthermore, it will be assumed that each judge gives a ranking of the individuals which is free of ties. The individuals will be numbered according to their rank sums, that is, individual $k$ has rank sum $r_{(k)}$. Note that the rank sum consensus ranking will contain ties whenever two or more individuals have the same rank sum. However, attention will be restricted to consensus rankings where the winner is unique, i.e., $r_{(I)} > r_{(I-1)}$.

Now suppose that the winner is deleted from the set of individuals under consideration and that the ranks assigned by the judges are adjusted as follows:

$$r_{ij}^* = \begin{cases} r_{ij}, & \text{if } r_{ij} < r_{Ij}, \\ r_{ij} - 1, & \text{if } r_{ij} > r_{Ij}. \end{cases}$$

This is equivalent to behaving as if the winner was not in the set of individuals under consideration in the first instance.

The ordering of the reduced set is said to be *consistent* with the original ordering if

$$r_{(i)} < r_{(j)} \Rightarrow r_i^* < r_j^* \quad \text{and} \quad r_{(i)} = r_{(j)} \Rightarrow r_i^* = r_j^*$$

for $i \neq j$, $i, j = 1, \ldots, I-1$, where

$$r_i^* = \sum_{j=1}^{J} r_{ij}^*, \qquad 1 \leqslant i \leqslant I-1.$$

In this study, the following type of inconsistency is considered:

$$r_{(k)} < r_{(I-1)} \quad \text{but} \quad r_k^* > r_{I-1}^* \qquad \text{for some } k, \qquad 1 \leqslant k \leqslant I-2.$$

This type of inconsistency arises when one or more individuals, who were beaten by the winner and at least one other person in the original consensus ranking, emerge as the winner or joint winners in the reduced ranking. Such an inconsistency is referred to as an *interchange* of winner in the reduced set.

To illustrate the occurrence of an interchange consider the following ranking configuration in the case $I = 4$, $J = 5$:

| Individuals | Judges 1 | 2 | 3 | 4 | 5 | $r_{(i)}$ | $r_i^*$ |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 1 | 1 | 3 | 8 | 7 |
| 2 | 3 | 3 | 3 | 2 | 1 | 12 | 12 |
| 3 | 1 | 2 | 2 | 4 | 4 | 13 | 11 |
| 4 | 4 | 4 | 4 | 3 | 2 | 17 | — |

It is noted that the individual who was third in the original ranking becomes the winner in the reduced ranking.

The extent to which the reduced ranking can differ from the original ranking upon deletion of the winner is illustrated by the following ranking configuration for $I = 5$ and $J = 6$:

| Individuals | Judges 1 | 2 | 3 | 4 | 5 | 6 | $r_{(i)}$ | $r_i^*$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 3 | 3 | 4 | 4 | 16 | 16 |
| 2 | 4 | 5 | 1 | 2 | 2 | 3 | 17 | 15 |
| 3 | 5 | 4 | 2 | 1 | 3 | 2 | 17 | 15 |
| 4 | 3 | 3 | 5 | 5 | 1 | 1 | 18 | 14 |
| 5 | 2 | 2 | 4 | 4 | 5 | 5 | 22 | — |

The individual who ranked lowest in the original ranking emerges as the winner in the reduced ranking and the individual who ranked second in the original ranking becomes last in the reduced set.

An examination of conditions under which interchanges can take place is given in the next section. Although attention has been restricted to the study of interchanges, a similar approach can be used in studying other types of inconsistencies.

## 2. Some Necessary Conditions for an Interchange

A number of conditions which must be satisfied in order that an interchange be possible will now be presented. Let $\alpha(k, l)$ denote the number of judges who rank individual $k$ over individual $I$, but rank individual $I$ over individual $l$. Note that

$$\alpha(k, l) - \alpha(l, k) = (r_{(k)} - r_k^*) - (r_{(l)} - r_l^*). \tag{1}$$

If $D_{kl} = r_{(k)} - r_{(l)}$ is the difference between the rank sums of individuals $k$ and $l$, then for $k > l$ the following inequality holds:

$$D_{kl} \geqslant 2\alpha(k, l) + (1 - I)\,\alpha(l, k) + (2 - I)(J - \alpha(k, l) - \alpha(l, k)). \tag{2}$$

This follows in that individual $k$ must receive, from $\alpha(k, l)$ judges, ranks exceeding those of individual $l$ by at least 2; from $\alpha(l, k)$ judges, ranks less than those of individual $l$ by no more than $(I - 1)$; and from the remaining judges, ranks less than those of individual $l$ by no more than $(I - 2)$.

Consider the occurrence of an interchange, that is, for some

$$k = 1, 2, \ldots, I - 2,\ D_{(I-1)k} = D > 0 \qquad \text{and} \qquad r_{I-1}^* < r_k^*.$$

In this situation individual $k$ overcomes a deficit $D$ and ends up with a rank sum in the reduced set which exceeds that of the person who was second in the original ranking. From (1) it follows that an interchange can occur only if

$$\alpha(I - 1, k) - \alpha(k, I - 1) \geqslant D + 1. \tag{3}$$

Now $0 \leqslant \alpha(k, l) \leqslant J - 1$ for $k, l = 1, \ldots, I - 1$, so that the occurrence of an interchange requires $J \geqslant D + 2$. By combining (2) and (3) one obtains

$$I \geqslant 2 + \frac{D + \alpha(k, I - 1) + 2}{J - \alpha(k, I - 1) - D - 1}\,. \tag{4}$$

The bound in (4) is a minimum when $\alpha(k, I-1) = 0$. Thus we have the following:

THEOREM. *A necessary condition for there to be an interchange in the ranking based on rank sums when I individuals are ranked by J judges is that for a deficit* $D > 0$, *J is at least* $D + 2$ *and*

$$I \geqslant [3 + (D+1)/(J-D-1)], \qquad (5)$$

*where* $[\cdot]$ *denotes the greatest integer function.*

It should be noted that the bound given by (5) is not a sufficient condition for the occurrence of an interchange. For example, when $I = 3$ and $J = 6$, (5) is satisfied when the deficit $D = 1$. However, as it will now be shown, no interchange is possible when $I = 3$ and $J = 6$. Now $r_1^* + r_2^* = 18$ so that $r_2^* - r_1^*$ is even. By (5) an interchange can only occur when $D = 1$, in which case, from (1), $\alpha(2, 1) - \alpha(1, 2)$ is odd. It then follows from (3) that the occurrence of an interchange requires $\alpha(2, 1) - \alpha(1, 2) \geqslant 3$. But this requires $r_{(3)} \leqslant r_{(2)}$, contrary to the assumption that $r_{(3)} > r_{(2)}$. Thus no interchange can take place.

From the theorem above it follows that no interchange is possible unless $I, J \geqslant 3$ and $I + J \geqslant 8$. In addition, as has been noted, no interchange is possible in the case $I = 3$, $J = 6$. Observing that interchanges can occur for $(I, J) \in \{(5, 3); (4, 4); (3, 5); (4, 6)\}$, we now show that interchanges are possible for all $(I, J)$ such that $I + J \geqslant 8$, $I, J \geqslant 3$ except $(3, 6)$.

Given a ranking configuration that leads to an interchange for $(3, J)$ one obtains a ranking configuration that leads to an interchange for $(3, J+2)$ and $(3, J+3)$ as follows:

(i) Let two additional judges assign ranks 1 and 3 to individual 1, ranks 3 and 1 to individual 2, and ranks 2 and 2 to individual 3. Then the augmented rank sum $r_{(i)}^{+} = r_{(i)} + 4$, $i = 1, 2, 3$ while $\alpha(2, 1) - \alpha(1, 2)$ is unchanged.

(ii) Let three additional judges assign ranks 1, 2 and 3 to individual 1, ranks 3, 1 and 2 to individual 2, and ranks 2, 3 and 1 to individual 3. Then the augmented rank sum $r_{(i)}^{+} = r_{(i)} + 6$, $i = 1, 2, 3$ while $\alpha(2, 1) - \alpha(1, 2)$ is unchanged.

Given a ranking configuration that leads to an interchange for $(I, J)$ one obtains a ranking configuration leading to an interchange for $(I + 1, J)$ by adding an individual who is ranked lowest by all judges.

With the use of the theorem above and of standard backtracking procedures, a list of all vectors of rank sums $\mathbf{r} = (r_{(1)}, \ldots, r_{(I)})$ which lead to interchanges can be generated. For each such vector $\mathbf{r}$ one can determine the proportion of those ranking configurations which yield $\mathbf{r}$ for which an interchange results. Furthermore, the proportion of all possible ranking configurations which lead to an interchange for the case $(I, J)$ is found. A description of the counting techniques and a tabulation of the results for all $(I, J)$ such that $I + J \leqslant 9$ are available from the authors.

## 3. Interchanges and Concordance

A measure of the agreement among $J$ judges in their rankings $\{A_1, ..., A_J\}$ of $I$ individuals can be based on each of the measures of distance between rankings introduced in Section 1, namely,

$$T_s = \sum_{1 \leqslant j < k \leqslant J} \sum s^2(A_j, A_k),$$

$$T_d = \sum_{1 \leqslant j < k \leqslant J} \sum d^2(A_j, A_k).$$

Clearly, small values of $T_s$ or $T_d$ correspond to good agreement.

A widely used measure of agreement among a set of rankings is provided by Kendall's coefficient of concordance:

$$W = 12S/J^2(I^3 - I),$$

where $S = \sum_{i=1}^{I} (r_i - \frac{1}{2}J(I + 1))^2$ is the sum of squares of the deviations of the rank sums from the average rank sum. If all $J$ rankings are identical $W = 1$, and if the rankings are totally different $W = 0$. The concordance $W$, in the form given, assumes that none of the rankings $\{A_1, ..., A_J\}$ contain ties. It can be shown by adapting the presentation in [4, pp. 95–96] that

$$W = 1 - 12T_s/J^2(I^3 - I).$$

As might be expected, an interchange usually occurs in ranking configurations in which there is substantial lack of agreement among the judges. To illustrate that this is not always the case consider the following ranking configuration for $I = 5$ and $J = 4$:

| Individuals | Judges 1 | 2 | 3 | 4 | $r_{(i)}$ | $r_i^*$ |
|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 2 | 5 | 5 |
| 2 | 2 | 2 | 3 | 1 | 8 | 8 |
| 3 | 3 | 4 | 4 | 3 | 14 | 14 |
| 4 | 5 | 3 | 2 | 5 | 15 | 13 |
| 5 | 4 | 5 | 5 | 4 | 18 | — |

By deleting the winner and adjusting ranks, the individual who was third in the original ranking emerges as the winner in the reduced ranking. In this case $T_s = 46$, $S = 114$ and $W = 0.713$. Less than 1% of the ranking configurations for $I = 5$ and $J = 4$ have higher concordance $W$ (the 0.05 and 0.01 significance points of $S$ have been given by Friedman [2] for $I = 3, 4, 5, 6, 7$ and $J = 3, 4, 5, 6, 8, 10, 15, 20$).

## Acknowledgment

The authors gratefully acknowledge the support received under a University of Victoria faculty research grant, and support received under Canadian National Research Council grants NRC A7166 and NRC A5203.

The authors are indebted to Mr. C. Knight for his assistance with the computations.

## References

1. M. Friedman, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *J. Amer. Statist. Assoc.* **32** (1937), 675–699.
2. M. Friedman, A comparison of alternative tests of significance for the problem of *m* rankings, *Ann. Math. Statist.* **11** (1940), 86–92.
3. J. G. Kemeny and J. L. Snell, "Mathematical Models in the Social Sciences," Chap. II, Blaisdell, Waltham, Mass., 1962.
4. M. G. Kendall, "Rank Correlation Methods," Chaps. 6, 7, Hafner, New York, 1962.