The Annals of Statistics
1976, Vol. 4, No. 3, 651-654

A NOTE ON PAIRED COMPARISON RANKINGS

BY JAGBIR SINGH

Temple University

If m objects $x_1, x_2, \cdots, x_m$ are compared pairwise, then let s_{ij} denote the number of times x_i beats x_j in n_{ij} independent comparisons. In a ranking, if x_i precedes x_j then one may require that the probability of x_i beating x_j be at least $\frac{1}{2}$. Such a ranking is called weak stochastic ranking. Let $I(R)$ be the set of all pairs (i, j) such that x_i precedes x_j in the ranking R in spite of the paired comparison outcomes resulting in $s_{ij} < s_{ji}$. A statistic $D(R) = \sum_{I(R)} (s_{ij} - s_{ji})^2/n_{ij}$ is derived and proposed for estimating a weak stochastic ranking. Since $D(R)$ is seen to be the sum of a random number of asymptotically distributed chi-square variates, a ranking is called minimum chi-square weak stochastic if $D(R) \leqq D(R_t)$, for $t = 1, 2, \cdots, m!$ It is proved that minimum chi-square rankings share at least two properties with the maximum likelihood rankings. That is, every minimum chi-square ranking is Hamiltonian ranking and when in particular $n_{ij} = 1$, every minimum chi-square ranking minimizes the violations of observed outcomes. Moreover, the branch and bound algorithm can be used for estimating the minimum chi-square rankings.

1. Introduction. In a paired comparison experiment, the elements of a set X are compared two at a time. Thus, if ties are not allowed, a comparison between x_i and x_j will result in either "x_i beats x_j" or "x_j beats x_i." A pair may be compared more than once. One considers these comparisons as constituting a sample from the collection of all possible comparisons and thinks of $\pi_{ij} = P(x_i \text{ beats } x_j)$ as being population parameters with $\pi_{ij} + \pi_{ji} = 1$. Let $\pi = (\pi_{12}, \pi_{13}, \cdots, \pi_{m-1,m})$ denote a typical point of the parameter space Ω. An arrangement $R = (x_{i_1}, x_{i_2}, \cdots, x_{i_m})$ of the elements of X is called a *weak stochastic ranking* if $\pi_{ij} \geqq \frac{1}{2}$ whenever x_i precedes x_j in R. Thompson and Remage (1964) studied the problem of estimating a *maximum likelihood* weak stochastic ranking based on paired comparison samples. Determining the maximum likelihood weak stochastic ranking R, or m.l. ranking R in short, was an optimization problem, since the solution involved maximizing the likelihood function

$$L(\pi) = \prod_{i<j} \binom{n_{ij}}{s_{ij}} \pi_{ij}^{s_{ij}} \pi_{ji}^{s_{ji}} ,$$

subject to the restriction that $\pi_{ij} \geqq \frac{1}{2}$ whenever x_i precedes x_j in R. In the likelihood function, s_{ij} is the number of times x_i beats x_j in n_{ij} independent comparisons. Let $\Omega(R) = \{\pi : \pi_{ij} \geqq \frac{1}{2} \text{ whenever } x_i \text{ precedes } x_j \text{ in the ranking } R\}$. Hence, R is a m.l. ranking if for any other ranking R_t, $t = 1, 2, \cdots, m!$

$$\sup_{\pi \in \Omega(R)} L(\pi) \geqq \sup_{\pi \in \Omega(R_t)} L(\pi) . \tag{1.1}$$

Received February 1975.

AMS 1970 *subject classifications.* Primary 62J15; Secondary 62G05.

Key words and phrases. Paired comparison rankings, maximum likelihood and minimum chi-square rankings, branch and bound algorithm.

The unrestricted maximum likelihood estimate of π_{ij} is $\hat{\pi}_{ij} = s_{ij}/n_{ij}$. Let $\hat{\pi} = (\hat{\pi}_{12}, \hat{\pi}_{13}, \cdots, \hat{\pi}_{m-1,m})$ denote the maximum likelihood estimate of π. The following is essentially a result of Thompson and Remage.

THEOREM. *The maximum of* $L(\pi)$ *over* $\Omega(R)$ *is* $L(\tilde{\pi})$, *where* $\tilde{\pi} = (\tilde{\pi}_{12}, \tilde{\pi}_{13}, \cdots, \tilde{\pi}_{m-1,m})$ *is such that*

$$\begin{aligned} \tilde{\pi}_{ij} &= \hat{\pi}_{ij} \quad \text{if } x_i \text{ precedes } x_j \text{ in } R \text{ and } s_{ij} > s_{ji} \\ &= \tfrac{1}{2} \quad \text{if } x_i \text{ precedes } x_j \text{ in } R \text{ and } s_{ij} \leqq s_{ji}\,. \end{aligned}$$

2. Minimum chi-square rankings. Define $I(R) = \{(i, j) : s_{ij} > s_{ji}$, and x_j precedes x_i in the ranking $R\}$. Notice that for $(i, j) \in I(R)$, $\tilde{\pi}_{ij} = \frac{1}{2}$, that is, an observed outcome between x_i and x_j is being violated by ranking x_j ahead of x_i in the ranking R. We notice that the logarithm of the likelihood ratio, $\lambda(R) = \sup_{\Omega(R)} L(\pi)/L(\hat{\pi})$, can be written as

$$\begin{aligned} (2.1)\qquad \ln \lambda(R) &= \textstyle\sum_{I(R)} n_{ij}[\ln \frac{1}{2} - \hat{\pi}_{ij} \ln \hat{\pi}_{ij} - \hat{\pi}_{ji} \ln \hat{\pi}_{ji}] \\ &= -\textstyle\sum_{I(R)} n_{ij}[\hat{\pi}_{ij} \ln \{1 + (\hat{\pi}_{ij} - \hat{\pi}_{ji})\} + \hat{\pi}_{ji} \ln \{1 + (\hat{\pi}_{ji} - \hat{\pi}_{ij})\}]\,. \end{aligned}$$

Since $-1 < (\hat{\pi}_{ij} - \hat{\pi}_{ji}) = -(\hat{\pi}_{ji} - \hat{\pi}_{ij}) < 1$, using Taylor's series expansion of the logarithmic functions, simplifying and ignoring terms of higher orders, we approximate

$$-2 \ln \lambda(R) \approx \textstyle\sum_{I(R)} n_{ij}(\hat{\pi}_{ij} - \hat{\pi}_{ji})^2 = D(R)\,, \quad \text{say.}$$

The statistic $D(R)$ can also be written as

$$(2.2)\qquad D(R) = \textstyle\sum_{I(R)} (s_{ij} - s_{ji})^2/n_{ij} = 4 \sum_{I(R)} (s_{ij} - n_{ij}/2)^2/n_{ij}\,.$$

When n_{ij} is the same for all pairs then we have the even simpler statistic $D(R) = \sum_{I(R)} (s_{ij} - s_{ji})^2$.

If R is a m.l. ranking, then $-2 \ln \lambda(R) \leqq -2 \ln \lambda(R_t)$ for any ranking R_t. Instead of using the method of maximum likelihood, we propose to use $D(R)$ to estimate a weak stochastic ranking. From (2.2), $D(R)$ is seen to be the sum of asymptotically and independently distributed chi-square variables, each with one degree of freedom. We notice, however, that the number of elements in the set $I(R)$ is random; and, therefore, the distribution of $D(R)$ is not really chi-square. Regardless, a ranking R is to be called minimum chi-square weak stochastic ranking, or m.c. ranking, if $D(R) \leqq D(R_t)$ for any other ranking R_t. We now prove two interesting properties of the m.l. rankings retained by the m.c. rankings.

PROPERTY 1. *Every* m.c. *ranking is a Hamiltonian ranking in the sense that, if* x_j *is the immediate predecessor of* x_i *in a* m.c. *ranking, then* $s_{ij} \leqq s_{ji}$.

PROOF. Suppose x_j is the immediate predecessor of x_i in a m.c. ranking R but $s_{ij} > s_{ji}$. By definition of $I(R)$, $(i, j) \in I(R)$. Let R_t be the ranking obtained from R by interchanging x_i and x_j. Note that neither $(i, j) \in I(R_t)$, nor $(j, i) \in I(R_t)$. Consider any other subscript pair $(k, l) \in I(R_t)$. It follows that x_k precedes x_l in the ranking R_t, and $s_{kl} < s_{lk}$. Since x_j is the immediate predecessor of x_i in R, and R_t differs from R by the interchange of x_i and x_j only, we notice that $(k, l) \in I(R)$. Hence $I(R_t) \subset I(R)$. Thus $D(R_t) < D(R)$. This is a contradiction of the hypothesis that R is a m.c. ranking.

PROPERTY 2. *When each pair is compared exactly once, then every* m.c. *ranking minimizes the number of violations of observed paired comparison outcomes.*

PROOF. If $n_{ij} = 1$, then $\hat{\pi}_{ij}$ is either zero or one. In either case, it follows from (2.1) that $-2 \ln \lambda(R)$ is a constant multiple of the number of elements in the set $I(R)$. On the other hand, $D(R)$ equals the number of elements in the set $I(R)$. Hence, when $n_{ij} = 1$, the m.l. rankings and the m.c. rankings are the same. Since the m.l. rankings minimize the violations of observed outcomes when $n_{ij} = 1$ for all pairs, the same is true for the m.c. rankings.

For a given ranking R, $D(R)$ is very simple to compute. DeCani (1972) observed that the branch and bound algorithm is useful for estimating the m.l. rankings. To see that the algorithm is also useful for computing m.c. rankings, let

$$c_{ij} = (s_{ij} - s_{ji})^2/n_{ij} \qquad \text{if} \quad s_{ij} < s_{ji}$$
$$= 0 \,.$$

Now we can write

$$D(R) = \sum_{(i,j)\in R} c_{ij} \,,$$

where $(i, j) \in R$ indicates that x_i precedes x_j in R. Let $Z = \min D(R)$. Obviously $Z \geqq 0$. We are seeking that ranking which gives Z. Let $Z_{i_1 i_2 \cdots i_\gamma}$ be a lower bound on Z when x_{i_1} precedes x_{i_2}, x_{i_2} precedes x_{i_3}, and so on $x_{i_{\gamma-1}}$ precedes x_{i_γ} in the ranking. It is easy to see that

$$Z_{i_1 i_2 \cdots i_\gamma} = \sum_{j=1}^{\gamma-1} \sum_{k=j+1}^{\gamma} c_{i_j i_k} \,.$$

Thus, given any partial ranking, we can obtain lower bound on Z. We briefly outline the algorithm; for details, see deCani.

For some x_i and x_j, compute z_{ij} and z_{ji}, and choose the smaller of the two. Suppose it is z_{ij}. Then for some x_k, compute z_{ijk}, z_{ikj}, z_{kij}, and choose the smallest one. Continue this way. Thus, if m objects are to be ranked, at the $(m-1)$th stage, m lower bounds are computed based on the smallest lower bound computed at the previous stage. Call $Z_m(1)$ the smallest of the lower bounds. Ties can be resolved arbitrarily. Delete all the other $(m-1)$th stage lower bounds since they clearly do not give m.c. rankings.

The process of calculation generated a tree. The $(\gamma-1)$th stage has γ branches terminating in γ nodes. The nodes with smallest lower bounds were chosen for subsequent branching. From $(m-1)$th stage, go down the tree, eliminating nodes with lower bounds bigger than $Z_m(1)$. If a node cannot be deleted then from this node branch back up reaching a new $(m-1)$th stage and a new minimal lower bound $Z_m(2)$. Choose the smaller of $Z_m(1)$ and $Z_m(2)$, and proceed to eliminate additional nodes. Sooner or later the algorithm will terminate after finding all m.c. rankings. In using the algorithm, one can search and remove all those branches which would not lead to Hamiltonian rankings.

Consider an example from Thompson and Remage (1964). In the example, $m = 4$, $n_{ij} = 4$, $s_{12} = 3$, $s_{13} = 1$, $s_{14} = 1$, $s_{23} = 3$, $s_{24} = 3$, and $s_{34} = 3$. The c_{ij} are always very easy to compute. In this example, additional simplification occurs, since $n_{ij} = 4$ for all (i, j). Hence if we replace n_{ij} by one in the definition of c_{ij}, we find $c_{12} = 0$, $c_{21} = 4$, $c_{13} = 4$, $c_{31} = 0$, $c_{14} = 4$, $c_{41} = 0$, $c_{23} = 0$, $c_{42} = 4$, $c_{24} = 0$, $c_{42} = 4$, $c_{34} = 0$, and $c_{43} = 4$. Algorithm will proceed as shown in the following table based on a preliminary best to worst ranking x_2, x_3, x_4, x_1.

TABLE 1
Branching sequences and bound values

First sequence	Second sequence	Third sequence
1. $z_{23} = 0$		
$z_{32} = 4$		Branch up from z_{32}
2. $z_{234} = 0$		$z_{324} = 4$
$z_{243} = 4$	Branch up from z_{243}	$z_{342} = 8^*$
$z_{423} = 8^*$		$z_{432} = 12^*$
3. $z_{2341} = 4$	$z_{2431} = 8^*$	$z_{3241} = 8^*$
$z_{2314} = 8^*$	$z_{2413} = 12^*$	$z_{3214} = 12^*$
$z_{2134} = 12^*$	$z_{2143} = 16^*$	$z_{3124} = 8^*$
$z_{1234} = 8^*$	$z_{1243} = 12^*$	$z_{1324} = 12^*$

* Means deleted nodes.

The m.c. ranking is x_2, x_3, x_4, x_1, which is also the m.l. ranking.

REFERENCES

DECANI, J. S. (1972). A branch and bound algorithm for maximum likelihood paired comparison ranking. *Biometrika* **59** 131–135.

THOMPSON, W. A. and REMAGE, R. (1964). Rankings from paired comparisons. *Ann. Math. Statist.* **35** 739–747.

DEPARTMENT OF STATISTICS
TEMPLE UNIVERSITY
PHILADELPHIA, PENNSYLVANIA 19122