RANKING THE PLAYERS IN A ROUND ROBIN TOURNAMENT

by

H. A. David

Department of Biostatistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 718

October 1970

Ranking the Players in a Round Robin Tournament

by

H. A. David

University of North Carolina

1. Introduction

We consider primarily the simplest type of Round Robin tournament in which each of t players $A_1, A_2, \ldots, A_t$ meets every other player once, and each game results in a win for one of the players who receives 1 point, the loser scoring 0. The question which concerns us in this paper is how to convert the results of the tournament into a ranking of the players. A familiar procedure is to base the ranking on the total number of wins $a_1, a_2, \ldots, a_t$ of the players, but how are ties to be resolved? In any case, statistics other than total wins may be used and we shall review critically various methods which have been proposed. An important part of our approach is to compare methods by their performance in small tournaments when intuition may be regarded as providing stronger guidance than some apparently appealing general principle.

Although it is convenient to use the language of tournaments throughout, the ranking problem applies immediately to a paired-comparison experiment in which every object A_i $(i=1,2,\ldots,t)$ is compared once with every other object, with the results expressed as a preference for one or the other object.

Some extensions to more general situations are indicated.

2. Partial Orderings and Strong Equalities

The results of a Round Robin tournament T consisting of t players are conveniently expressed by means of a <u>tournament</u> (or dominance) <u>matrix</u>

$$\underset{\sim}{A} = \begin{pmatrix} 0 & \alpha_{12} & \cdots & \alpha_{1t} \\ \alpha_{21} & 0 & \cdots & \alpha_{2t} \\ \vdots & \vdots & & \vdots \\ \alpha_{t1} & \alpha_{t2} & \cdots & 0 \end{pmatrix} \quad (1)$$

having 0's along the principal diagonal and $\alpha_{ij} + \alpha_{ji} = 1$ for all i,j ($i,j=1,2,\ldots,t;\ i\neq j$). Throughout most of this paper we take $\alpha_{ij} = 1$ if $A_i \to A_j$ (A_i has defeated A_j) and $\alpha_{ij} = 0$ if $A_j \to A_i$, thus precluding any outcome intermediate between win and loss.

At the conclusion of the tournament the players can always be arranged into disjoint sets $T_1, T_2, \ldots, T_k$ (for some $k, k=1,2,\ldots,t$) with the following properties (cf. Kadane, 1966):

(a) Each player in T_h has defeated all players in $T_{h'}$ for all $h<h'$ ($h,h'=1,2,\ldots,k$);

(b) For any two players A_i, A_j in the same set T_h, either $A_i \to A_j$ or there exist other players $A_{i_1}, A_{i_2}, \ldots$ in T_h such that

$$A_i \to A_{i_1} \to A_{i_2} \to \ldots \to A_j. \quad (2)$$

Suppose now that we wish to rank the players on the basis of the tournament results only, i.e., ignoring any other information on the strength of the players. Then it is clear that players in T_h should rank ahead of those in $T_{h'}$ for $h<h'$.[1] In the extreme case $k=t$ we obtain a complete ordering of all the players. So much is universally agreed. However, at the other extreme $k=1$

---

[1] We are not concerned here with questions of statistical significance (see David, 1963a, p. 75).

there is no obviously best way of ranking the players or even any subset of them; for if A_i and A_j are any two players and if $A_j \to A_i$, then to make up for the direct defeat by A_j, player A_i has to his credit an indirect win over A_j in the manner of (2). It is this case of a strong tournament that we shall need to examine further. For $1<k<t$ the sets $T_1, T_2, \ldots, T_k$ provide a partial ordering in which only the rankings within sets remain in doubt. The outcomes of the games among the players within any one set clearly constitute a strong subtournament.

2.1 Strong tournaments. The smallest strong tournament arises when t=3 and

$$\text{(i)} \quad A_1 \to A_2 \to A_3 \to A_1 \quad \text{or} \quad \text{(ii)} \quad A_1 \to A_3 \to A_2 \to A_1, \tag{3}$$

the familiar circular triads of Kendall and Babington Smith (1940). In either case (i) or (ii) there is no justification for preferring any one player and we must declare them all equal. Since such a verdict is presumably unanimous, we shall call this strong equality among A_1, A_2, A_3. Nevertheless this is in some respects an uneasy equality. For suppose that one of the players is to be selected to participate in a future tournament. Obviously (in the absence of other information) one of A_1, A_2, A_3 should be chosen at random. But now it turns out that A_3 is unavailable. Then we are likely to prefer A_1 over A_2 for outcome (i) above and A_2 over A_1 for (ii). Some may wish to toss, but no one, I believe, will prefer A_2 over A_1 in case (i).

Are there other instances when we may reasonably speak of strong equality? The answer is yes if t is odd. For it is then always possible to find a tournament outcome which is made up of cycles. If t is prime there are $\frac{1}{2}(t-1)$ cycles, all of length t. For example, if t=7, writing the players as 0,1,...,6 for short, we have the 3 cycles

$$0 \to 1 \to 2 \to 3 \to 4 \to 5 \to 6 \to 0,$$
$$0 \to 2 \to 4 \to 6 \to 1 \to 3 \to 5 \to 0,$$
$$0 \to 3 \to 6 \to 2 \to 5 \to 1 \to 4 \to 0.$$

To generate the 2nd and 3rd cycles we have simply kept on adding, respectively, 2 and 3 mod 7 (cf. David, 1963b). When t is not prime, we may proceed in a similar manner, i.e., by repeated addition of $r=2,3,\ldots,\frac{1}{2}(t-1)$; however, if r divides t the corresponding cycle will be replaced by r cycles, of length t/r, starting with $0,1,\ldots,r-1$. Thus for $t=15$ the 3rd cycle is

$$0 \to 3 \to 6 \to 9 \to 12 \to 0, \quad 1 \to 4 \to 7 \to 10 \to 13 \to 1,$$
$$2 \to 5 \to 8 \to 11 \to 14 \to 2.$$

When the complete tournament results can be resolved into cycles, there will again, I think, be general agreement that the players are equal. One may perhaps wish to add that some (those occurring in the same cycle) are more equal than others.

The process fails when t is even because the $\frac{1}{2}t$ outcomes $0 \to \frac{1}{2}t, 1 \to \frac{1}{2}t+1, \ldots, \frac{1}{2}t-1 \to t-1$ are not part of any cycle.

Strong equality of an odd number of players is also possible <u>within</u> a larger tournament of any size. For example, if (3) holds and if each of A_1, A_2, A_3 has the same record against each of the remaining players $A_4, A_5, \ldots, A_t$, then the strong equality of A_1, A_2, A_3 is clearly preserved. The smallest strong tournament of this type is that for $t=5$ with A_4 defeating and A_5 losing to each of A_1, A_2, A_3, but with $A_5 \to A_4$. A convenient listing of all non-isomorphic tournaments for $t \leq 6$ is given by Moon (1968, pp. 91-5). For $t=3,4,5,6$ the list includes respectively, 1, 1, 6, 35 strong tournaments.

3. Row-Sum Scores

Once all partial orderings, with possible equalities, have been effected in the manner of section 2, we are faced with the more difficult problem of how to rank the players within the resultant strong subtournaments. Some readers may wonder why we do not simply rank the t players on the basis of their number of wins. This familiar method gives A_i the score $a_i = \sum_{\substack{j=1 \\ j \neq i}}^{t} \alpha_{ij}$, more fully called the row-sum score, since a_i is the sum of the ith row of the matrix $\underset{\sim}{A}$ in (1). It is easy to verify that this method is in complete accord with section 2, as any reasonable method must be. There is indeed little wrong with the method in the case of a balanced tournament such as the Round Robin, as we shall see in some detail. However, it is by no means the only reasonable method and it inevitably produces some tied rankings unless the original tournament T is transitive. In the latter case the row-vector of scores, arranged in descending order of magnitude is $(t-1, t-2, \ldots, 1, 0)$. In a strong tournament $a_i = t-1$ is impossible, as is $a_i = 0$; hence there remain only $t-2$ possible scores for the t players, showing that there must be at least two tied pairs or a tied triple of players. Corresponding remarks apply to strong subtournaments.

Other methods have therefore been proposed, probably with the primary aim of providing tie-breaking mechanisms. The most important of these were originated by Kendall and Wei (Kendall, 1955), Brunk (1960), and Slater (1961). Of course, these procedures will leave strong equalities intact. Also they may do more than break ties and place a player ahead of one with a larger number of wins.

4. The Method of Kendall and Wei

We illustrate a slightly simpler version of this method (Moon, 1968, p. 44) on the following tournament matrix

$$\underset{\sim}{A} = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix} . \tag{4}$$

The first column of Table 1 gives the row-sum scores $\underset{\sim}{a}=\underset{\sim}{A}\underset{\sim}{1}$, where $\underset{\sim}{1}$ is a column of t 1's. We note a triple tie. The starting point of the Kendall-Wei method is to obtain a second score vector $\underset{\sim}{a}^{(2)}$ by assigning to each player the total number of games won by all the opponents defeated by him.[2] For example, $a_1^{(2)} = 3+1 = 4$, and in general $\underset{\sim}{a}^{(2)} = \underset{\sim}{A}^2\underset{\sim}{1}$. This gives the second column, with A_2 now behind A_1 and A_5 who are still tied. The idea is to give more credit to a player for defeating a strong (i.e., high-scoring) opponent than for a win over a weak opponent. Kendall (1955) writes, "this is as far as one would wish to go on practical grounds, perhaps" but then investigates continuation of this process of re-allocation which clearly corresponds to repeated powering of the matrix $\underset{\sim}{A}$. The score vector $\underset{\sim}{a}^{(3)} = \underset{\sim}{A}^3\underset{\sim}{1}$ is given in the next column, but produces no change in the ranking. Actually, Kendall powered not $\underset{\sim}{A}$ but the matrix $\underset{\sim}{A} + \frac{1}{2}\underset{\sim}{I}$ obtained by giving each player half a point for tieing with himself. The scores $\underset{\sim}{b}^{(2)} = (\underset{\sim}{A} + \frac{1}{2}\underset{\sim}{I})^2\underset{\sim}{1} = \underset{\sim}{a}^{(2)} + \underset{\sim}{a} + \frac{1}{4}\underset{\sim}{1}$ appearing in the next column give the same ranking in this example. Now if this process is continued indefinitely the rankings will settle down.

In fact, since (Thompson, 1958) for $t>3$ the matrix $\underset{\sim}{A}$ of any strong tournament is primitive (i.e., $\underset{\sim}{A}^n$ has all its elements positive from a certain finite integer $n=n_o$ on) it is known from Frobenius theory (e.g., Brauer, 1961) that

---

[2] This has long been a tie-breaking method (the Sonneborn-Berger system) used in chess tournaments.

$$\lim_{n\to\infty} \left(\frac{A}{\lambda}\right)^n \underset{\sim}{1} = \underset{\sim}{s},$$

where λ is the unique positive characteristic root of $\underset{\sim}{A}$ with the largest absolute value and $\underset{\sim}{s}$ is a vector of positive terms, the column eigenvector satisfying

$$\underset{\sim}{A}\underset{\sim}{s} = \lambda \underset{\sim}{s}. \tag{5}$$

Here $\underset{\sim}{s}$ is determined only up to a constant multiplier. Replacing $\underset{\sim}{A}$ by $\underset{\sim}{A} + \frac{1}{2}\underset{\sim}{I}$ increases λ by $\frac{1}{2}$ but leaves $\underset{\sim}{s}$ unchanged. By the Kendall-Wei method we mean the ranking of the players according to the components of $\underset{\sim}{s}$.

Table 1. Various score vectors for the tournament with matrix $\underset{\sim}{A}$ of (4)

	$\underset{\sim}{a}$	$\underset{\sim}{a}^{(2)}$	$\underset{\sim}{a}^{(3)}$	$\underset{\sim}{b}^{(2)}$	$\underset{\sim}{s}$	$\underset{\sim}{s}^*$	$\underset{\sim}{v}$
A_1	2	4	7	$6\frac{1}{4}$	.4623	.4623	11
A_2	2	3	6	$5\frac{1}{4}$	.3880	.3880	7
A_3	3	5	9	$8\frac{1}{4}$	.5990	.2514	19
A_4	1	2	4	$3\frac{1}{4}$	.2514	.5990	3
A_5	2	4	7	$6\frac{1}{4}$	.4623	.4623	9

In Table 1 $\underset{\sim}{s}$ has been normalized ($\underset{\sim}{s}'\underset{\sim}{s}=1$). The rankings given by $\underset{\sim}{s}$ are the same as in the preceding 3 columns. (In other cases it may take a little longer for the process to settle down.) An incidental feature in this example is that the tie between A_1 and A_5 remains unbroken by any of the rankings in spite of the fact that the equality of A_1 and A_5 is not strong (in the sense of section 2). We omit a formal proof which can be based on induction on n. The present example is the smallest where matrix-powering methods fail to break an equality which is not strong. However, all ties are broken by the method of section 4.2 leading to equation (8) and to the score vector $\underset{\sim}{v}$ given in the last column of Table 1.

The smallest example for which $\underset{\sim}{s}$ reverses the order of two players, as determined by $\underset{\sim}{a}$, is given by

$$\underset{\sim}{A} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad (6)$$

Here $\underset{\sim}{s}' = (.6382, .5400, .3415, .2159, .3712)$ so that A_5 is ranked ahead of A_3. This is a dubious improvement over ranking by $\underset{\sim}{a}$. It is instructive to look in some detail at the even simpler case t=4.

4.1 <u>The strong tournament of 4 players</u>. Ignoring the transitive tournament for t=4, we can without loss of generality take $A_1 \to A_2 \to A_3 \to A_1$. Now if A_4 wins or loses all his games against A_1, A_2, A_3 the equality of this trio is, of course, unimpaired. The remaining outcomes, leading to a strong tournament, are of two kinds

(i) $A_1 \to A_4$, $A_2 \to A_4$, $A_4 \to A_3$,

(ii) $A_4 \to A_1$, $A_4 \to A_2$, $A_3 \to A_4$,

and yield, respectively,

$$\underset{\sim}{s}_1' = (.6256, .5516, .4484, .3213),$$
$$\underset{\sim}{s}_2' = (.3213, .4484, .6256, .5516).$$

We see that the equality of A_1 and A_2 is now broken although they have the same record against A_4; moreover, A_1 ranks ahead of A_2 in (i) and behind A_2 in (ii). Since there is only one distinct type of strong tournament for t=4, the elements of $\underset{\sim}{s}_2'$ are simply a permutation of those in $\underset{\sim}{s}_1'$. The particular permutation and its inverse are, respectively,

$$P = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix} \quad \text{and} \quad P^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix},$$

the second row of P^{-1} giving the ranks of the elements in $\underset{\sim}{s}_2'$. Similarly, an interchange of wins and losses in all six games including (i) corresponds to the permutations

$$Q = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix} = Q^{-1},$$

thus resulting in the ranking (3, 4, 1, 2).

This last example illustrates a feature of the Kendall-Wei method applying for any t: interchange of wins and losses does not necessarily reverse a ranking.

Thus the Kendall-Wei method has some disconcerting consequences. Similar remarks hold for other methods based on powering the tournament matrix. The idea of giving more credit to a player for defeating a strong (i.e., high-scoring) opponent than a weak one retains usefulness in breaking ties among top-scorers (see the Appendix) but I do not find it acceptable for arriving at a complete ranking.

4.2 Variants. An interesting variant of the Kendall-Wei method has been proposed by Ramanujacharyulu (1964). In addition to the nth scores $\underset{\sim}{a}^{(n)} = \underset{\sim}{A}^n \underset{\sim}{1}$, which he calls the 'iterated power of order n', this author suggests scores $\underset{\sim}{a}^{*(n)} = (\underset{\sim}{A}')^n \underset{\sim}{1}$, the 'iterated weakness of order n'. We see that $\underset{\sim}{a}^{*(n)}$ is the result of n-1 re-allocations of losses rather than wins. The strongest player is now the one suffering the fewest iterated losses, i.e., having the lowest score. As $n \to \infty$ we have, with the same λ as in (5),

$$\underset{\sim}{A}'\underset{\sim}{s}^* = \lambda \underset{\sim}{s}^*,$$

or

$$\underset{\sim}{s}^{*\prime}\underset{\sim}{A} = \lambda \underset{\sim}{s}^{*\prime},$$

i.e., $\underset{\sim}{s}^{*\prime}$ is simply the row-eigenvector of $\underset{\sim}{A}$ corresponding to the principal characteristic root λ. Of course, $\underset{\sim}{A}'$ may be obtained from $\underset{\sim}{A}$ by interchange of wins and losses. We have already pointed out that this does not necessarily lead to a reversal of rankings made in the manner of section 4.1, so that $\underset{\sim}{s}^*$ does not necessarily give the same rankings as $\underset{\sim}{s}$.

For $\underset{\sim}{A}$ of (6) the scores s_i^* are given in Table 1. (Since $\underset{\sim}{A}$ is self-conjugate under interchange of wins and losses the s_i^* are permutations of the s_i). They give the ranking ($3\frac{1}{2}$, 2, 1, 5, $3\frac{1}{2}$) instead of the $\underset{\sim}{s}$-ranking ($2\frac{1}{2}$, 4, 1, 5, $3\frac{1}{2}$).

Ramanujacharyulu advocates the 'power-weakness ratio' $\underset{\sim}{r}$, where $r_i = s_i/s_i^*$. In our example this leads right back to the original ranking by $\underset{\sim}{a}$. The same applies to the difference $d_i = s_i - s_i^*$, favored by Hasse already in 1961 on the grounds that interchange of wins and losses simply reverses the sign of d_i, and hence reverses the corresponding ranking. This desirable feature of $\underset{\sim}{d}$ is, however, hardly decisive. For tournament (i) of section 4.1, for example, we have:

	$\underset{\sim}{s}$	$\underset{\sim}{s}^*$	$\underset{\sim}{r}$	$\underset{\sim}{d}$
A_1	.6256	.4484	1.3952	.1772
A_2	.5516	.3213	1.7168	.2303
A_3	.4484	.6256	.7168	-.1772
A_4	.3213	.5516	.5825	-.2303

Thus $\underset{\sim}{s}^*$, $\underset{\sim}{r}$, and $\underset{\sim}{d}$ all rank A_2 ahead of A_1. The rewards of steadiness (not losing to a weaker player) are greater than we may think reasonable.

Recently other motivations for the Kendall-Wei approach and use of equation (5) have been put forward by Daniels (1969) and Pullman and Moon (1969). These authors also suggest a number of modifications, mostly based on the concept of 'fair scores'. Thus, to take the most useful of their proposals, suppose that A_i is 'worth' V_i in the sense that any player who beats A_i wins V_i from him. One way of choosing the V_i is to equate expected gains and losses; i.e., we require

$$\sum_j \pi_{ij} V_j = V_i \sum_j \pi_{ji} \qquad i=1,2,\ldots,t, \tag{7}$$

where $\pi_{ij} = \Pr\{A_i \to A_j\}$, $i \neq j$, and $\pi_{ii} = 0$ by convention.[3]

In a Round Robin of n rounds, π_{ij} is estimated by α_{ij}/n. Correspondingly we may estimate V_i by scores v_i satisfying

$$\sum_j \alpha_{ij} v_j = v_i \sum \alpha_{ji}, \tag{8}$$

or, defining $q_{ij} = \alpha_{ij}/\sum_j \alpha_{ji}$, and $\underset{\sim}{Q} = (q_{ij})$, by

$$\underset{\sim}{Q}\underset{\sim}{v} = \underset{\sim}{v}. \tag{8'}$$

This is, in fact, the characteristic equation corresponding to $\underset{\sim}{Q}$ whose largest eigenvalue is 1 (Pullman and Moon). Daniels points out that for the Bradley-Terry model $\pi_{ij} = \pi_i/(\pi_i+\pi_j)$ $(\pi_i \geq 0, \Sigma\pi_i = 1)$ equation (7) gives (apart from a multiplicative constant) $V_i = \pi_i$. In that case, the v_i of (8), which may be obtained without iteration, are therefore simple estimates of the π_i, but the v_i

---

[3] This convention is more in line with our previous procedures than $\pi_{ii} = \frac{1}{2}$ used by Daniels.

may, of course, be used quite generally. For the 8 non-isomorphic strong tournaments existing for $t \leq 5$ (n=1), $\underset{\sim}{v}$ produces the same rankings as $\underset{\sim}{s}$ of (5) except for breaking a tie in the tournament of Table 1 (see last column, where the v_i have been taken as the smallest positive integers satisfying (8)).

4. Inconsistencies, Upsets, and Weak and Strong Orderings

Two distinct lines of approach, different from any of the foregoing, are taken by Brunk (1960) and Slater (1961). We consider these only briefly, taking the latter first. Slater points out that corresponding to any tournament outcome there is one or more ranking of the players for which the number i of inconsistencies is minimized. By an inconsistency[4] is meant a defeat of a player by one ranked below him. For example, consider the circular triad $A_1 \to A_2 \to A_3 \to A_1$. Of the 6 possible rankings, three namely $A_1A_2A_3$, $A_2A_3A_1$, $A_3A_1A_2$ have i=1, whereas $A_1A_3A_2$, $A_3A_2A_1$, $A_2A_1A_3$ have i=2. Slater therefore rules out the second set but expresses no preference among the members of the first set.[5] He proposes i as a general statistic for a test of randomness in place of Kendall's widely-used number c of circular triads. Unfortunately i is for all but very small sample sizes very much more difficult to evaluate than c (nor is it necessarily better than c for being more complicated; cf. David, 1963a, p. 34). Several interesting methods have been developed to determine i and the associated ranking(s), e.g., Remage and Thompson (1966) by dynamic programming, Phillips (1969) by ingenious elementary methods, and deCani (1969) by linear programming. With Slater's approach, as with the methods of section 4, it is possible for the resultant ranking to be in discord with any row-sum ranking. Finding this aspect unsatisfactory, Ryser (1964) and Fulkerson (1965) have devised methods leading

---

[4] Ryser (1964) uses the more evocative word 'upset'.

[5] Since each player is ranked first once, second once, and third once, strong equality is not really broken.

to rankings which minimize the number of upsets subject to keeping row-sums monotone.

The case for Slater's i has been strengthened by a probabilistic basis provided by Thompson and Remage (1964) who show that Slater's nearest adjoining order is also the maximum-likelihood weak stochastic order, i.e., the ranking obtained by maximizing the likelihood function

$$L = \prod_{i<j} \pi_{ij}^{\alpha_{ij}} (1-\pi_{ij})^{1-\alpha_{ij}} \qquad 1 \leq i < j \leq t \tag{9}$$

with respect to the π_{ij}, subject to the restriction that for any ordered triple $A_{i_1}\ A_{i_2}\ A_{i_3}$

$$\pi_{i_1 i_2} \geq \tfrac{1}{2},\ \pi_{i_2 i_3} \geq \tfrac{1}{2},\ \pi_{i_1 i_3} \geq \tfrac{1}{2}. \tag{10}$$

In contrast, Brunk (1960) maximizes L subject to the strong stochastic transitivity condition

$$\pi_{i_1 i_2} \geq \tfrac{1}{2},\ \pi_{i_2 i_3} \geq \tfrac{1}{2},\ \pi_{i_1 i_3} \geq \max\ (\pi_{i_1 i_2}, \pi_{i_2 i_3}). \tag{11}$$

It is instructive to compare the two approaches on the circular triad outcome $A_1 \rightarrow A_2 \rightarrow A_3 \rightarrow A_1$ for which (9) reduces to

$$L = \pi_{12}\pi_{23}(1-\pi_{13}).$$

For each of the 6 possible rankings Table 2 shows the estimates p_{ij} of the π_{ij} obtained when L is maximized under (10) and (11), together with corresponding likelihood L. For example, for the ranking $A_1A_2A_3$ maximization of L under (10) obviously leads to $p_{12}=1$, $p_{23}=1$, $p_{13}=\frac{1}{2}$, whereas under (11) we get $p_{12}=\frac{2}{3}$, $p_{23}=\frac{2}{3}$,

Table 2. Probability estimates p_{ij} and likelihood values L for the rankings $A_{i_1} A_{i_2} A_{i_3}$ when $A_1 \to A_2 \to A_3 \to A_1$.

(i_1, i_2, i_3)			Under (10)			L	Under (11)			L
			$p_{i_1 i_2}$,	$p_{i_2 i_3}$,	$p_{i_1 i_3}$		$p_{i_1 i_2}$,	$p_{i_2 i_3}$,	$p_{i_1 i_3}$	
1	2	3	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{4}{27}$
2	3	1	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{4}{27}$
3	1	2	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{4}{27}$
1	3	2	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{4}$
3	2	1	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{4}$
2	1	3	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	1	$\frac{1}{4}$

$p_{13}=\frac{2}{3}$. The next two rows follow by cyclic interchange. The startling result, due to Brunk, is that under (11) the last three rankings give the higher value of L. As he points out, in none of these rankings is a player ranked more than one place ahead of one to whom he lost.

What conclusions are we to draw from this embarrassment of choices in making a ranking? Personally, I prefer to leave A, B, C tied although this leads to a lower value of L, viz., $\frac{1}{8}$. General principles, such as restricted maximization of L, are, of course, of interest but confidence in them is somewhat undermined by what they lead to in a situation with which we can really come to grips. Moreover, the numerical values of the estimates maximizing L under (10) and (11) are hardly realistic; for example, under (11) the ranking A_1 A_3 A_2 gives $p_{13}=\frac{1}{2}$, $p_{32}=\frac{1}{2}$, and yet $p_{12}=1$, etc. (The unrestricted maximum of L is 1, with $p_{12}=1$, $p_{23}=1$, $p_{13}=0$ which are also unrealistic values providing no guidance for ranking.)

References

Brauer, A. (1961). On the characteristic roots of power-positive matrices. Duke Math. J. **28**, 439-45.

Brunk, H. D. (1960). Mathematical models for ranking from paired comparisons. J. Amer. Statist. Ass. **55**, 503-20.

Bühlmann, H. and Huber, P. J. (1963). Pairwise comparison and ranking in tournaments. Ann. Math. Statist. **34**, 501-10.

Daniels, H. E. (1969). Round-robin tournament scores. Biometrika **56**, 295-99.

David, H. A. (1963a). The Method of Paired Comparisons. London, Griffin; New York, Hafner.

David, H. A. (1963b). The structure of cyclic paired-comparison designs. J. Aust. Math. Soc. **3**, 117-27.

DeCani, J. S. (1969). Maximum likelihood paired comparison ranking by linear programming. Biometrika **56**, 537-45.

Fulkerson, D. R. (1965). Upsets in round robin tournaments. Canad. J. Math. **17**, 957-69.

Hasse, Maria (1961). Über die Behandlung graphentheoretischer Probleme unter Verwendug der Matrizenrechnung. Wiss. Zeit. Tech. Univ. Dresden **10**, 1313-6.

Huber, P. J. (1963). Pairwise comparison and ranking: optimum properties of the row sum procedure. Ann. Math. Statist. **34**, 511-20.

Kadane, J. B. (1966). Some equivalence classes in paired comparisons. Ann. Math. Statist., **37**, 488-94.

Kendall, M. G. (1955). Further contributions to the theory of paired comparisons. Biometrics **11**, 43-62.

Kendall, M. G. and Babington Smith, B. (1940). On the method of paired comparisons. Biometrika **33**, 239-51.

Moon, J. W. (1968). Topics on Tournaments. Holt, Rinehart and Winston, New York.

Phillips, J. P. N. (1969). A further procedure for determining Slater's i and all nearest adjoining orders. Brit. J. Math. Statist. Psychol. **22**, 97-101.

Pullman, N. J. and Moon, J. W. (1969). Tournaments and handicaps. pp. 219-23 in Information Processing **68**, Amsterdam, North-Holland Publishing Co.

Ramanujacharyulu, C. (1964). Analysis of preferential experiments. Psychometrika **29**, 257-61.

Remage, R., Jr. and Thompson, W. A., Jr. (1966). Maximum-likelihood paired comparison rankings. Biometrika **53**, 143-49.

Ryser, H. J. (1964). Matrices of zeros and ones in combinatorial mathematics. pp. 103-24 in Schneider H. (Ed.) Recent Advances in Matrix Theory, Madison, Univ. Wisconsin Press.

Slater, P. (1961). Inconsistencies in a schedule of paired comparisons. Biometrika **48**, 303-12.

Thompson, G. L. (1958). Lectures on game theory, Markov chains and related topics. Sandia Corporation Monograph SCR-11.

Thompson, W. A., Jr. and Remage, R., Jr., (1964). Rankings from paired comparisons. Ann. Math. Statist. **35**, 739-47.

Appendix

We formally establish here some properties of the methods of section 4 based on powering the tournament matrix.

Theorem 1. In a strong tournament with $t \geq 4$ the nth score $a_i^{(n)} = (\underset{\sim}{A}^n \underset{\sim}{1})_i$ is non-decreasing in n $(i=1,2,\ldots,t)$ and $a_i^{(n+2)} > a_i^{(n)}$ for $n=1,2,\ldots$.

Proof. $a_i^{(n)}$ is the sum of the ith row of $\underset{\sim}{A}^n$ and therefore equals the total number of paths of length n starting at A_i in the directed graph corresponding to $\underset{\sim}{A}$. But for each such path there is at least one path of length n+1, since in the graph of a strong tournament at least one path must leave each vertex. Hence $a_i^{(n+1)} \geq a_i^{(n)}$, with equality holding only when all paths of length n from A_i end up at vertices of outdegree 1. For $t \geq 4$ there are at most two such vertices, say A_1 and A_2, with $A_1 \rightarrow A_2$. Thus $a_i^{(n)} = a_i^{(n+1)} = a_i^{(n+2)}$ would imply that all these paths of length n from A_i become paths of length n+1 also ending up at A_1 or A_2. This would require all the paths of length n from A_i to end at A_1, which is impossible. It follows that $a_i^{(n+2)} > a_i^{(n)}$, which completes the proof.

We may note also that $a_i^{(n+1)} > a_i^{(n)}$ for $n > t+2$. This result can be verified directly for $t=4$ and follows for $t>4$ from the fact that all elements of $\underset{\sim}{A}^n$ are positive for $n > t+2$ (Moon, §13).

Theorem 2. In a strong tournament T with $t \geq 4$, suppose $a_1=a_2=a$ and $A_1 \rightarrow A_2$. Then the Kendall-Wei method ranks A_1 ahead of A_2 if $a=t-2$ and A_2 ahead of A_1 if $a=1$.

Proof. Take $a=t-2$ first and let A_3 be the player who defeated A_1. Then from $\underset{\sim}{A}\underset{\sim}{s} = \lambda \underset{\sim}{s}$ we have

$$s_2 + \quad + s_4 + \ldots + s_t = \lambda s_1,$$
$$s_3 + s_4 + \ldots + s_t = \lambda s_2, \tag{A1}$$

so that

$$s_2 - s_3 = \lambda(s_1 - s_2). \tag{A2}$$

Now if A_3 had lost only to A_2, then A_1, A_2, A_3 would have been strongly equal. However, since T is strong, A_3 must have lost to at least one other player. Hence $s_3 < s_2$ so that $s_1 > s_2$ by (A2) since $\lambda > 0$.

The case a=1 follows similarly.

Comments.

1. For a=1 the result $s_2 > s_1$ may be proved directly by noting that instead of (A1) we now have $s_2 = \lambda s_1$ with $\lambda > 1$, since λ lies (strictly) between the smallest and the largest row-sums of $\underset{\sim}{A}$ (Brauer, 1961).

We see also that in (A2) $s_2 - s_3 > s_1 - s_2$, which is pleasing in so far as A_3 has a lower row-sum score than the common score of A_1 and A_2.

2. Similar results hold for other matrix-powering methods of ranking. For example, for $a_1 = a_2 = 1$ and $A_2 \rightarrow A_3$ we have from $\underset{\sim}{a}^{(n)} = \underset{\sim}{A}^n \underset{\sim}{1}$ that

$$a_1^{(n+1)} = a_2^{(n)}, \; a_2^{(n+1)} = a_3^{(n)}.$$

From Theorem 1 it now follows that

$$a_1^{(n+1)} = a_2^{(n)} = a_3^{(n-1)} \leq a_3^{(n)} = a_2^{(n+1)}$$

and that the inequality is strict for one of any two consecutive values of n.

The last result makes it clear that the original Kendall-Wei powering, with scores $\underset{\sim}{b}^{(n)} = (\underset{\sim}{A} + \tfrac{1}{2}\underset{\sim}{I})^n \underset{\sim}{1}$, gives $b_1^{(n)} < b_2^{(n)}$ for n=2,3,... . Thus this method breaks such ties. However, like the modified method, it may introduce fresh ties elsewhere in the ranking.