Stochastic Inequalities
IMS Lecture Notes – Monograph Series
Volume 22 (1993)

FURTHER ASPECTS OF THE COMPARISON OF TWO GROUPS OF RANKED OBJECTS BY MATCHING IN PAIRS[1]

By H. A. DAVID and JINGYU LIU

Iowa State University

Suppose $\Gamma_X = (X'_{(1)}, \ldots, X'_{(n)})$ and $\Gamma_Y = (Y'_{(1)}, \ldots, Y'_{(n)})$ are two groups of stochastically ordered rv's, representing, say, the increasing strengths of the members of two chess teams. Let $\pi = (\pi_1, \ldots, \pi_n)$ be a permutation of $(1, \ldots, n)$. Then the statistic $S(\pi) = \sum_{i=1}^n I(Y'_{(i)} > X'_{(\pi_i)})$ measures the superiority of Γ_Y over Γ_X in matchings under π, where $I(y > x)$ is an indicator function. The dependence of $ES(\pi) = \sum_{i=1}^n P(Y'_{(i)} > X'_{(\pi_i)})$ on π, especially when $\pi = (1, \ldots, n)$ and when π is randomly given, has been studied in Liu and David (1992) under two different models. After a review of the main results of that paper, some new optimality results are developed. In addition, a threshold model is used to treat tied comparisons.

1. Introduction

Tournaments in which n players or teams are compared by being matched up in pairs have been studied by mathematicians and statisticians at least since Zermelo (1929). The eminent author proposed and examined a method for evaluating the strengths of contestants in a round robin chess tournament that had to be broken off before each pair of players could meet. Independently, statisticians became interested through the connection between tournaments and the method of paired comparisons. In the latter, typically, n flavors are compared by being tasted in pairs, pairwise comparison providing maximal discrimination.

The method of paired comparisons goes back to the psychometrician Thurstone (1927), other notable early contributions being Kendall and Babington Smith (1940), Bradley and Terry (1952), and Kendall (1955). The last paper is perhaps the first to stress the connection between tournaments and paired comparisons, a point pursued in David (1959), where

[1]Work supported by the U. S. Army Research Office.

AMS 1991 *subject classifications.* Primary 62G30, 62J15, 62E15, 15A39; Secondary 62P99.

Key words and phrases. Stochastic inequalities, order statistics, linear preference model, majorization, unimodality, tied comparisons, threshold model.

knock–out tournaments are also studied. Since then many authors have dealt with questions of design and analysis for round robin, knock–out, and related tournament–type situations. A review of this literature is given in David (1988).

The present paper is in the same spirit although it does not, strictly speaking, involve a tournament. We examine the standard method for comparing two chess teams, namely to match pairs of players having the same rank within their teams. Questions that arise are: Is this *ordered matching* really an optimal procedure and what are its properties? How does it compare with random matching, or with other possible matchings of the two teams? What kind of matchings are fair? Some answers are provided in Liu and David (1992) under two different probability models. After reviewing the main features of that paper, referred to as LD from here on, we present some new optimality results and deal with the complication of tied games (draws in chess).

It should be noted that, in more general terms, we are concerned with the nonparametric comparison of two groups of n objects by pairwise matching, when there is good knowledge of the within–group ranking.

Further results are given in Liu (1991).

2. The Probability Models

Let $\Gamma_X = (X'_{(1)}, \ldots, X'_{(n)})$ and $\Gamma_Y = (Y'_{(1)}, \ldots, Y'_{(n)})$ be two groups of stochastically increasing random variables (not necessarily order statistics) which represent the increasing "strengths" of the ordered objects in the two groups. Let $F_i(x)$ and $G_i(x)$ represent the continuous cdf's of $X'_{(i)}$ and $Y'_{(i)}$, respectively, $i = 1, 2, \ldots, n$. Here we assume $F_i(x) \geq F_j(x)$ for all x and any $1 \leq i < j \leq n$, i.e., $X'_{(i)}$ is stochastically smaller than $X'_{(j)}$ or $X'_{(i)} \leq_{\text{st}} X'_{(j)}$. Usually, we also assume that Γ_X and Γ_Y are independent; however, we do not assume independence within Γ_X and Γ_Y.

Let $\pi = (\pi_1, \ldots, \pi_n)$ be a permutation of $\pi_0 = (1, \ldots, n)$. Then for each given π, we can define a matching by comparing $X'_{(\pi_i)}$ and $Y'_{(i)}$, $i = 1, 2, \ldots, n$. So we will simply speak of a matching π. Correspondingly, we suppose that in a particular matching hypothetical realizations $x'_{(\pi_i)}$ of $X'_{(\pi_i)}$, and $y'_{(i)}$ of $Y'_{(i)}$, $i = 1, \ldots, n$, are compared. While $x'_{(\pi_i)}$ or $y'_{(i)}$, cannot be observed, we can make the (usually subjective) judgment whether $y'_{(i)} > x'_{(\pi_i)}$. For the moment, we assume a clear decision; the possibility of a tie is considered in Section 5. Thus we prefer the Y–group, Γ_Y, to the X–group,

Γ_X, in this particular matching if

$$\sum_{i=1}^{n} I(y'_{(i)} > x'_{(\pi_i)}) > \frac{1}{2}\, n$$

where

$$I(y > x) = \begin{cases} 0 & \text{if } y \leq x \\ 1 & \text{if } y > x \end{cases}.$$

Write

$$S(\pi) = \sum_{i=1}^{n} I(Y'_{(i)} > X'_{(\pi_i)}).$$

Then $S(\pi)$ denotes the random number of preferences for objects in Γ_Y. We regard Γ_Y as *superior* to Γ_X under matching π if the expectation

$$E[S(\pi)] = \sum_{i=1}^{n} P(Y'_{(i)} > X'_{(\pi_i)}) > \frac{1}{2}\, n.$$

It is clear that some matchings may favor one of the groups. Both ordered and random matching are clearly *fair*, i.e., $E[S(\pi)] = \frac{1}{2}\, n$, when Γ_X and Γ_Y have the same distribution. Other matchings are not necessarily fair (see LD). Let V_1 and V_2 be the values of $E[S(\pi)]$ under ordered and random matching, respectively.

Our first model is the *order statistics model.* In this model we assume that $X'_{(i)}$ and $Y'_{(i)}$ have the same *marginal* distributions as $X_{(i)}$ and $Y_{(i)}$, the i^{th} order statistics in random samples of size n from F and G, respectively. We use $X'_{(i)}$ rather than $X_{(i)}$ since we generally want to permit $P(X'_{(i)} > X'_{(j)}) > 0$ for $i < j$. The joint distribution of the $X'_{(i)}$ may, in fact, have any dependence structure, including independence.

However, our measure of superiority of Γ_Y over Γ_X, viz. $E[S(\pi)]$, depends only on the marginal distributions of $X'_{(i)}$ and $Y'_{(j)}$. We will therefore replace $X'_{(i)}$ and $Y'_{(j)}$ by the order statistics $X_{(i)}$ and $Y_{(j)}$ from here on in discussions of the order statistics model. For an ordered matching we have

$$V_1 = E[S(\pi^0)] = \sum_{i=1}^{n} P(Y_{(i)} > X_{(i)}).$$

Note that under random matching we simply have $V_2 = nP(Y > X)$, with $X \sim F$ and $Y \sim G$.

The question of whether ordered matching in the order statistics model is more effective than random matching may now be reduced to the question of whether $V_1 \geq V_2$ if $X \leq_{\text{st}} Y$. The answer is yes under certain conditions. We deal with this and related issues in Section 3.

The second model is the *linear preference model.* In this model, we assume that

$$(1) \qquad X'_{(i)} \sim F(x - \lambda_{(i)}) \quad \text{and} \quad Y'_{(i)} \sim F(x - \mu_{(i)}), \quad i = 1, \ldots, n$$

where $F(x)$ is a continuous distribution function and $\lambda_{(1)} \leq \cdots \leq \lambda_{(n)}$ and $\mu_{(1)} \leq \cdots \leq \mu_{(n)}$. The model is based on the linear model much used in the method of paired comparisons (e.g., David, 1988, p. 7). At times, we will assume that $F(x)$ is in the class of unimodal distribution functions, that contains almost all the common useful distribution functions.

It is easy to see that when $\mu_{(i)} = \lambda_{(i)}$, $i = 1, \ldots, n$, and both Γ_X and Γ_Y are groups of independent random variables, $S(\pi^0)$ has a Binomial $\left(n, \frac{1}{2}\right)$ distribution. In general, there is no closed form for the distribution of $S(\pi)$.

In this model, we are still interested in comparing V_1 and V_2, as well as V_1 and $E[S(\pi)]$. Under certain conditions, we get similar results to those in the order statistics model. However, there are significant differences between the two models.

3. Order Statistics Model

We consider the case $G(x) = F(x - \mu)$, where $\mu \geq 0$, and write $p_{ij} = P(Y'_{(i)} > X'_{(j)}) = P(Y_{(i)} > X_{(j)})$. The following basic result is obtained in LD.

THEOREM 3.1 *Let* $(X_{(1)}, \ldots, X_{(n)})$ *and* $(Y_{(1)}, \ldots, Y_{(n)})$ *be the order statistics in two independent random samples from populations with continuous cdf's* $F(x)$ *and* $F(x - \mu)$, *respectively, where* $\mu \geq 0$. *Then for any* $1 \leq i, j \leq n$, *we have*

$$p_{ii} + p_{jj} \geq p_{ij} + p_{ji}.$$

Under the conditions of the theorem it is then shown that $V_1 \geq V_2$, meaning that ordered matching has at least as much power to identify the stronger group Γ_Y as does random matching.

We now need two definitions

DEFINITION 3.1 π is said to be a *simple matching* (or permutation) if it can be obtained from π^0 by interchanging pairs of the components of π^0, with no component involved in more than one interchange.

DEFINITION 3.2 $\pi = (\pi_1, \ldots, \pi_n)$ is said to be a *symmetric matching* (or permutation) if $\pi_{n-i+1} = n - \pi_i + 1$ for $i = 1, \ldots, n$.

For example, (3,4,1,2) is a simple matching and (2,4,1,3) is a symmetric matching.

It is shown in LD that simple and symmetric matchings, as well as combinations thereof, are fair.

Moreover, the following result is obtained.

THEOREM 3.2 *Under the conditions of* Theorem 3.1

(a)
$$\sum_{i=1}^{n} p_{ii} \le \sum_{i=1}^{n} p_{i,\pi_i}$$

for any simple permutation π.

(b) *If $F(x)$ is the cdf of a symmetric rv, then* (a) *holds for any symmetric permutation π.*

4. Linear Preference Model

The specific assumptions for this model have been given in (1). If X and Y are iid with cdf $F(x)$, then $X'_{(i)} \overset{d}{=} X + \lambda_{(i)}$ and $Y'_{(i)} \overset{d}{=} Y + \mu_{(i)}$. Also $X - Y$ is symmetrically distributed about zero, with cdf $U(x)$, say.

We have
$$P(Y'_{(i)} > X'_{(i)}) = U(\mu_{(i)} - \lambda_{(i)})$$
and
$$E[S(\pi)] = \sum_{i=1}^{n} P(Y'_{(i)} > X'_{(\pi_i)}) = \sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\pi_i)}).$$

It is easy to see that ordered, random, and simple matching is still fair, i.e., $E[S(\pi)] = \frac{n}{2}$ if $\lambda_{(i)} = \mu_{(i)}$, $i = 1, \ldots, n$. Symmetric matching is also fair for any symmetric spacing, i.e., $\lambda_{(i)} = \mu_{(i)}$ and $\lambda_{(i)} + \lambda_{(n-i+1)} =$ constant, $i = 1, \ldots, n$.

For this model we now assume that $F(x)$ is a unimodal cdf, that is, there exists x_0 such that $F(x)$ is convex on $(-\infty, x_0)$ and concave on (x_0, ∞). With X, Y iid unimodal, $X - Y$ is also unimodal (e.g., Dharmadhikari and Joag–dev (1988, p. 15)). This is needed for the proof in LD of the following result.

THEOREM 4.1 *Let $(X'_{(1)}, X'_{(2)})$ and $(Y'_{(1)}, Y'_{(2)})$ be independent with $X'_{(i)} \sim F(x - \lambda_{(i)})$ and $Y'_{(i)} \sim F(x - \mu_{(i)})$, $i = 1, 2$, where $\lambda_{(1)} \le \lambda_{(2)}$, $\mu_{(1)} \le \mu_{(2)}$, and $F(x)$ is an absolutely continuous unimodal distribution. If $\mu_{(1)} + \mu_{(2)} \ge \lambda_{(1)} + \lambda_{(2)}$, then we have*

$$U(\mu_{(1)} - \lambda_{(1)}) + U(\mu_{(2)} - \lambda_{(2)}) \ge U(\mu_{(1)} - \lambda_{(2)}) + U(\mu_{(2)} - \lambda_{(1)}) \tag{2}$$

Note that if either $\mu_{(1)} = \mu_{(2)}$ or $\lambda_{(1)} = \lambda_{(2)}$, then equality holds in (2).

For any fixed $\lambda_{(1)} \leq \cdots \leq \lambda_{(n)}$ and $\mu_{(1)} \leq \cdots \leq \mu_{(n)}$, let $\tilde{\pi} = (\tilde{\pi}_1, \ldots, \tilde{\pi}_n)$ be a permutation for which

$$E[S(\tilde{\pi})] = \max_{\pi} \sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\pi_i)}).$$

Theorem 4.1 tells us that if $\sum_{i=1}^{n} \mu_{(i)} \geq \sum_{i=1}^{n} \lambda_{(i)}$, then $\pi^0 = \tilde{\pi}$ for $n = 2$. However, in contrast to the order statistics model, $\pi^0 = \tilde{\pi}$ no longer necessarily holds for $n > 2$, even within the class of simple matchings.

In general, $\sum_{i=1}^{n} \mu_{(i)} \geq \sum_{i=1}^{n} \lambda_{(i)}$ does not imply

$$\sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(i)}) \geq \frac{n}{2}. \tag{3}$$

The following sufficient condition is established in LD.

THEOREM 4.2 *Let* $\{a_1, \ldots, a_m\} = \{\mu_{(i)} - \lambda_{(i)}, \text{ s.t. } \mu_{(i)} - \lambda_{(i)} \geq 0,\ i = 1, \ldots, n\}$ *and* $\{b_1, \ldots, b_{n-m}\} = \{\lambda_{(i)} - \mu_{(i)}, \text{ s.t. } \mu_{(i)} - \lambda_{(i)} < 0,\ i = 1, \ldots, n\}$. *If* $m \geq [\frac{n}{2}]$ *and*

$$\sum_{i=1}^{k} a_{(i)} \geq \sum_{i=1}^{k} b_{(i)} \qquad (k = 1, \ldots, n-m), \tag{4}$$

then (3) *holds.*

It is also shown in LD that if $\mu_{(i)} + \mu_{(j)} \geq \lambda_{(i)} + \lambda_{(j)}$ for all $1 \leq i < j \leq n$, then ordered matching is superior to random and simple matching.

4.1. *Some New Optimality Results*

The following preliminaries are needed (e.g., Marshall and Olkin (1979)).

DEFINITION 4.1 Write $x_{[i]} = x_{(n+1-i)}$, $i = 1, \ldots, n$. Then for any $x, y \in R^n$, x is majorized by y ($x \prec y$) if

(a) $\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]} \qquad k = 1, \ldots, n-1$

(b) $\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}$.

LEMMA 4.1 (Hardy, Littlewood, and Pólya (1952)). *The inequality* $\sum_{i=1}^{n} g(x_i) \leq \sum_{i=1}^{n} g(y_i)$ *holds for all continuous convex functions* $g : \mathbb{R} \to \mathbb{R}$ *if and only if* $x \prec y$.

We have noted a variety of situations in which ordered matching is most effective in detecting the superior group. Now we obtain a sufficient condition for a matching $\tilde{\pi} = (\tilde{\pi}_1, \ldots, \tilde{\pi}_n)$ to be optimal, i.e., $\tilde{\pi}$ to be such that

$$\sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\tilde{\pi}_i)}) \geq \sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\pi_i)}) \tag{5}$$

for any permutation π.

Let

$$\{x_1, \ldots, x_p\} = \{\mu_{(i)} - \lambda_{(\tilde{\pi}_i)} \text{ s.t. } \mu_{(i)} - \lambda_{(\tilde{\pi}_i)} > 0 \quad i = 1, 2, \ldots, n\},$$

$$\{y_1, \ldots, y_q\} = \{\lambda_{(\tilde{\pi}_i)} - \mu_{(i)} \text{ s.t. } \lambda_{(\tilde{\pi}_i)} - \mu_{(i)} > 0 \quad i = 1, 2, \ldots, n\}.$$

Similarly, let

$$\{x'_1, \ldots, x'_{p'}\} = \{\mu_{(i)} - \lambda_{(\pi_i)} \text{ s.t. } \mu_{(i)} - \lambda_{(\pi_i)} > 0 \quad i = 1, 2, \ldots, n\},$$

$$\{y'_1, \ldots, y'_{q'}\} = \{\lambda_{(\pi_i)} - \mu_{(i)} \text{ s.t. } \lambda_{(\pi_i)} - \mu_{(i)} > 0 \quad i = 1, 2, \ldots, n\}.$$

Then, since $U(-x) = 1 - U(x)$, we have

$$\begin{aligned}\sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\tilde{\pi}_i)}) &= \sum_{j=1}^{p} U(x_j) + q - \sum_{j=1}^{q} U(y_j) + (n - p - q)U(0) \\ &= \sum_{j=1}^{p} U(x_j) - \sum_{j=1}^{q} U(y_j) + (n - p + q)U(0)\end{aligned}$$

and

$$\sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\pi_i)}) = \sum_{j=1}^{p'} U(x'_{(j)}) - \sum_{j=1}^{q'} U(y'_{(j)}) + (n - p' + q')U(0).$$

Therefore (5) is equivalent to

$$\begin{aligned}&\sum_{j=1}^{p} U(x_j) + \sum_{j=1}^{q'} U(y_{j'}) + (n - p + q)U(0) \\ \geq &\sum_{j=1}^{p'} U(x_{j'}) + \sum_{j=1}^{q} U(y_j) + (n - p' + q')U(0).\end{aligned}$$

Write

$$a_i = x_i \quad (i = 1, \ldots, p), \qquad a_{p+i} = y'_i \quad (i = 1, \ldots, q'),$$

$$a_{p+q'+i} = 0 \quad (i = 1, \ldots, n - p + q),$$

$$b_i = x'_i \qquad (i = 1, \dots, p'), \qquad b_{p'+i} = y'_i \qquad (i = 1, \dots, q),$$
$$b_{p'+q+i} = 0 \qquad (i = 1, \dots, n - p' + q').$$

By

$$\sum_{i=1}^{n} (\mu_{(i)} - \lambda_{(\tilde{\pi}_i)}) = \sum_{i=1}^{n} (\mu_{(i)} - \lambda_{(\pi_i)}),$$

we have $\sum_{i=1}^{n+q+q'} a_i = \sum_{i=1}^{n+q+q'} b_i$. Also, we can write (5) as

$$\sum_{i=1}^{n+q+q'} U(a_i) \geq \sum_{i=1}^{n+q+q'} U(b_i). \tag{6}$$

Now $-U(x)$ is a convex function on $[0, +\infty)$. Therefore, by Lemma 4.1, we see that (6) holds if $\sum_{i=1}^{k} a_{[i]} \leq \sum_{i=1}^{k} b_{[i]}$ for $k = 1, 2, \dots, n + q + q' - 1$. Summarizing the above argument, we have the following result.

THEOREM 4.3 *A sufficient condition for $\tilde{\pi}$ such that* (5) *holds is*

$$\sum_{i=1}^{k} a_{[i]} \leq \sum_{i=1}^{k} b_{[i]} \quad \textit{for } k = 1, \dots, n + q + q' - 1.$$

If there exist $\mu_{(i)}$ and $\mu_{(j)}$ such that $\mu_{(i)} + \mu_{(j)} > \lambda_{(\pi_i)} + \lambda_{(\pi_j)}$, where $i < j$ and $\pi_i > \pi_j$, then by Theorem 4.1, we can increase $\sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\pi_i)})$ by interchanging $\lambda_{(\pi_i)}$ and $\lambda_{(\pi_j)}$. Also, if there exist $\mu_{(i)}$ and $\mu_{(j)}$ such that $\mu_{(i)} + \mu_{(j)} < \lambda_{(\pi_i)} + \lambda_{(\pi_j)}$, where $i < j$ and $\pi_i < \pi_j$, we can increase $\sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\pi_i)})$ by interchanging $\lambda_{(\pi_i)}$ and $\lambda_{(\pi_j)}$. Therefore, Theorem 3.1 gives us a way to increase $\sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\pi_i)})$. We can also see that if $\sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\tilde{\pi}_i)}) > \sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\pi_i)})$ for any other permutation π, then for any $i < j$, either $\mu_{(i)} + \mu_{(j)} > \lambda_{(\tilde{\pi}_i)} + \lambda_{(\tilde{\pi}_j)}$ with $\tilde{\pi}_i < \tilde{\pi}_j$ or $\mu_{(i)} + \mu_{(j)} < \lambda_{(\tilde{\pi}_i)} + \lambda_{(\tilde{\pi}_j)}$ with $\tilde{\pi}_i > \tilde{\pi}_j$.

We therefore have the following Lemma.

LEMMA 4.2 *Suppose there exists a unique permutation $\tilde{\pi}$ such that for any $i < j$, if $\mu_{(i)} + \mu_{(j)} > \lambda_{(\tilde{\pi}_i)} + \lambda_{(\tilde{\pi}_j)}$, then $\tilde{\pi}_i < \tilde{\pi}_j$, and if $\mu_{(i)} + \mu_{(j)} < \lambda_{(\tilde{\pi}_i)} + \lambda_{(\tilde{\pi}_j)}$, then $\tilde{\pi}_i > \tilde{\pi}_j$. For such a $\tilde{\pi}$ it follows that* (5) *holds for any permutation π.*

5. The Treatment of Ties

In practice, some pairwise comparisons may result in ties. A tie occurs when the performances x and y of two objects being compared are too close to tell the difference. Accordingly, we introduce the indicator function

$$I(y > x; \tau) = \begin{cases} 0 & \text{if } y - x < -\tau \\ 1/2 & \text{if } |y - x| \leq \tau \\ 1 & \text{if } y - x > \tau \end{cases}$$

where $\tau (\geq 0)$ is called a *threshold parameter* (Glenn and David (1960)).

5.1. *Order Statistics Model*

For any permutation $\pi = (\pi_1, \dots, \pi_n)$, we now define

$$S_\tau(\pi) = \sum_{i=1}^{n} I(Y_{(i)} > X_{(\pi_i)}; \tau)$$

as a measure of the performance of Γ_Y relative to Γ_X under the matching π, with ties permitted. Then

$$\begin{aligned} (7) \qquad ES_\tau(\pi) &= \sum_{i=1}^{n} [P(Y_{(i)} > X_{(\pi_i)} + \tau) + \frac{1}{2} P(|Y_{(i)} - X_{(\pi_i)}| \leq \tau)] \\ &= \frac{1}{2} \sum_{i=1}^{n} [P(Y_{(i)} > X_{(\pi_i)} + \tau) + P(Y_{(i)} > X_{(\pi_i)} - \tau)]. \end{aligned}$$

For $\pi^0 = (1, \dots, n)$ corresponding to ordered matching, we write

$$(8) \qquad V_1^\tau = ES_\tau(\pi^0).$$

Let V_2^τ be the expectation of $S_\tau(\pi)$ under random matching. Then we have

$$\begin{aligned} V_2^\tau &= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} EI(Y_{(i)} > X_{(j)}; \tau) \\ &= \frac{1}{2n} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{n} [P(Y_{(i)} > X_{(\pi_i)} + \tau) + P(Y_{(i)} > X_{(\pi_i)} - \tau)] \right\} \end{aligned}$$

so that

$$(9) \qquad V_2^\tau = \frac{n}{2} [P(Y > X + \tau) + P(Y > X - \tau)],$$

where X and Y are independent with respective cdf's $F(x)$ and $F(x - \mu)$.

The questions that concern us here are how V_1^τ and V_2^τ compare, and also their relations to V_1 and V_2, respectively.

LEMMA 5.1 *Let $X \sim F(x)$, where $F(x)$ is an absolutely continuous unimodal cdf. If $Y \stackrel{d}{=} X + \mu$ with $\mu \geq 0$, and X and Y are independent, then for any $\tau \geq 0$,*

$$P(Y > X) \geq \frac{1}{2}\,[P(Y > X + \tau) + P(Y > X - \tau)]. \tag{10}$$

PROOF Let $U(x)$ be the cdf of $X_1 - X_2$, where X_1 and X_2 are iid with cdf $F(x)$. As noted earlier, $X_1 - X_2$ is symmetrically distributed about 0 and $U(x)$ is a unimodal cdf. Then (10) may be written

$$U(\mu) \geq \frac{1}{2}\,[U(\mu + \tau) + U(\mu - \tau)],$$

a result which is evident from the unimodality of U. □

COROLLARY *If $X_{(i)}$ has a unimodal distribution ($i = 1, \ldots, n$), then under the conditions of* Lemma 5.1

$$P(Y_{(i)} > X_{(i)}) \geq \frac{1}{2}\,[P(Y_{(i)} > X_{(i)} + \tau) + P(Y_{(i)} > X_{(i)} - \tau)].$$

We have thus shown that under the conditions stated $V_2 > V_2^\tau$ and $V_1 > V_1^\tau$. For the corollary we need $X_{(i)}$ to have a unimodal distribution. Alam (1972) shows that if the density function $f(x)$ of X satisfies the condition that $1/f(x)$ is convex, then the order statistics have unimodal distributions. This condition is satisfied by the following distributions (among others): normal, logistic, Cauchy, uniform, and the gamma and Weibull families for shape parameters ≥ 1.

THEOREM 5.1 *Let X and Y be independent absolutely continuous rv's with respective cdf's $F(x)$ and $F(x - \mu)$, where $\mu \geq 0$.*

(a) If $0 \leq \tau \leq \mu$, then $V_1^\tau \geq V_2^\tau$ and

$$V_1^\tau \geq ES_\tau(\pi) \tag{11}$$

for any simple permutation π. If X is a symmetric rv, then (11) *holds also for any symmetric permutation.*

(b) If $\mu = 0$ and $\tau \geq 0$, then

$$ES_\tau(\pi) = \frac{1}{2}\,n \tag{12}$$

for any simple permutation. If X is a symmetric rv, then (12) *holds also for any symmetric permutation.*

PROOF (a) From (7) we have

$$V_1^\tau = \frac{1}{2}\sum_{i=1}^{n}[P(Y_{(i)} - \tau > X_{(i)}) + P(Y_{(i)} + \tau > X_{(i)})].$$

Note that $Y_{(i)} \mp \tau$ $(i = 1, \ldots, n)$ is the i^{th} order statistic from a population with cdf $F[x - (\mu \mp \tau)]$, where $\mu \mp \tau \geq 0$. Then since $V_1 \geq V_2$, we have

$$\sum_{i=1}^{n} P(Y_{(i)} > X_{(i)} \pm \tau) \geq nP(Y > X \pm \tau),$$

which establishes $V_1^\tau \geq V_2^\tau$.

The remaining results in (a) follow from Theorem 3.2 and (b) is easily proved. □

Numerical work for the standard normal shows that $V_1^\tau \geq V_2^\tau$ does not necessarily hold if $\tau > \mu$.

5.2. *Linear Preference Model*

Corresponding to (7) we now have

$$\begin{aligned} ES_\tau(\pi) &= \sum_{i=1}^{n}[P(Y'_{(i)} > X'_{(\pi_i)} + \tau) + \frac{1}{2}P(|Y'_{(i)} - X'_{(\pi_i)}| \leq \tau)] \\ &= \frac{1}{2}\sum_{i=1}^{n}[U(\mu_{(i)} - \lambda_{(\pi_i)} - \tau) + U(\mu_{(i)} - \lambda_{(\pi_i)} + \tau)] \end{aligned}$$

Then $V_1^\tau = ES_\tau(\pi^0)$ and

$$V_2^\tau = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n}[U(\mu_{(i)} - \lambda_{(j)} - \tau) + U(\mu_{(i)} - \lambda_{(j)} + \tau)],$$

which does not simplify further, in contrast to (9).

The next lemma follows directly from Theorem 4.1.

LEMMA 5.2 *If* $\mu_{(i)} + \mu_{(j)} \geq \lambda_{(i)} + \lambda_{(j)} + 2\tau$ *for any* $1 \leq i, j \leq n$, *then*

(a) $V_1^\tau \geq V_2^\tau$;

(b) $V_1^\tau \geq ES_\tau(\pi)$ *for any simple matching* π.

We conclude with the following easily proved result.

LEMMA 5.3 *(a) If* $\lambda_{(i)} = \mu_{(i)}$ *for* $i = 1, 2, \ldots, n$, *then*

$$ES_\tau(\pi) = \frac{1}{2}\, n \tag{13}$$

for any simple matching π.

(b) If $\lambda_{(i)} = \mu_{(i)}$ *and* $\lambda_{(i+1)} - \lambda_{(i)} = \lambda_{(n-i+1)} - \lambda_{(n-i)}$ *for* $i = 1, \ldots, n$ *then* (13) *holds for any symmetric matching* π.

REFERENCES

ALAM, K. (1972). Unimodality of the distribution of an order statistic. *Ann. Math. Statist.* **43** 2041–4.

BRADLEY, R. A. AND TERRY, M. E. (1952). The rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika* **39** 324–45.

DAVID, H. A. (1959). Tournaments and paired comparisons. *Biometrika* **46** 139–49.

DAVID, H. A. (1988). *The Method of Paired Comparisons.* 2nd ed., Charles Griffin, London; Oxford University Press, New York.

DHARMADHIKARI, S. AND JOAG–DEV, K. (1988). *Unimodality, Convexity, and Applications.* Academic Press, Boston, MA.

GLENN, W. A. AND DAVID, H. A. (1960). Ties in paired–comparison experiments using a modified Thurstone–Mosteller model. *Biometrics* **16** 86–109.

HARDY, G. H., LITTLEWOOD, J. E. AND PÓLYA, G. (1952). *Inequalities.* 2nd ed. Cambridge University Press, London and New York.

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

KENDALL, M. G. AND SMITH, B. B. (1940). On the method of paired comparisons. *Biometrika* **31** 324–45.

LIU, J. (1991). Comparing Two Groups of Ranked Objects by Pairwise Matching. Ph.D. Thesis, Iowa State University.

LIU, J. AND DAVID, H. A. (1992). Comparing two groups of ranked objects by pairwise matching. *J. Statist. Plann. Inf.* **33** (to appear).

MARSHALL, A. W. AND OLKIN, I. (1979). *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York.

THURSTONE, L. L. (1927). A law of comparative judgment. *Psychol. Rev.* **34** 273–86.

ZERMELO, E. (1929). Die Berechnung der Turnier–Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung. *Math. Zeit.* **29** 436–60.

DEPARTMENT OF STATISTICS
IOWA STATE UNIVERSITY
AMES, IA 50011