Comparing two groups of ranked objects by pairwise matching

Jingyu Liu
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/rtd
Part of the Statistics and Probability Commons

Recommended Citation
Liu, Jingyu, "Comparing two groups of ranked objects by pairwise matching" (1992). Retrospective Theses and Dissertations. 10329.
http://lib.dr.iastate.edu/rtd/10329

This Dissertation is brought to you for free and open access by Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Comparing two groups of ranked objects by pairwise matching

Liu, Jingyu, Ph.D.
Iowa State University, 1992
Comparing two groups of ranked objects
by pairwise matching

by

Jingyu Liu

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
Major: Statistics

Approved:
Signature was redacted for privacy.

In Charge of Major Work
Signature was redacted for privacy.

For the Major Department
Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa
1992

Copyright © Jingyu Liu, 1992. All rights reserved.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ............................................ v

1. INTRODUCTION .................................................. 1

2. ORDER STATISTICS MODEL ..................................... 7
   2.1 Introduction .................................................. 7
   2.2 Ordered Matching and Random Matching ..................... 8
   2.3 Fair Matching under the Order Statistics Model ............. 13
   2.4 Results on $P(Y_{(i)} > X_{(j)})$ .......................... 21
   2.5 Some Asymptotic Results ................................... 32
   2.6 Matching with Ties Permitted for the Order Statistics Model... 41

3. LINEAR PREFERENCE MODEL .................................... 48
   3.1 Introduction .................................................. 48
   3.2 Preliminary Results ......................................... 49
   3.3 Ordered Matching ............................................ 52
   3.4 Fair Matching under the Linear Preference Model ............ 63
   3.5 Some Optimality Results .................................... 69
   3.6 Matching with Ties Permitted for the Linear Preference Model... 74

BIBLIOGRAPHY ...................................................... 77
LIST OF TABLES

Table 2.1: Values of $E[S(\pi)]$ for all permutations $\pi$ of $(1,2,3,4)$ .... 14
Table 2.2: Values of $p_i$ for the uniform distribution ............... 25
Table 2.3: Values of $p_i$ for the Beta distribution with $p = q$ ......... 27
Table 2.4: Values of $p_i$ for the normal distribution ................. 28
Table 2.5: Values of $p_i$ for the Gamma distribution ................. 29
Table 2.6: Values of $P_i^T$ for the normal distribution 1 .............. 46
Table 2.7: Values of $P_i^T$ for the normal distribution 2 .............. 47
ACKNOWLEDGEMENTS

I am greatly indebted to Dr. H. A. David, my major professor, for his guidance, support, and encouragement during the course of this work. I feel very fortunate to have such a respected advisor. My sincere thanks and appreciation are also due to all the professors who taught me during my graduate studies at ISU.

Finally, I would like to thank my wife, Shanlin, for her support and understanding during the years of my graduate studies and for her taking care of our lovely son, Chang-Chang, in those years.
1. INTRODUCTION

This dissertation deals with the nonparametric comparison of two groups of $n$ objects when only two objects can be compared at a time. Well-known examples of this situation are comparisons of two chess or tennis teams, where it is common practice to match the best player of the first team with the best player of the second team, down to matching the two weakest players. Such a matching has considerable intuitive appeal and its properties will be studied under two different models.

It is important to realize a key element inherent in the foregoing situation, namely that a great deal is usually known about the relative strengths of the players within a team. In fact, whenever a within-group ordering of $n$ objects is available the above matching by rank is appealing. But is this ordered matching really an optimal procedure and what are its properties? How does it compare with a random matching, or with other possible matchings of the two groups of objects?

We now introduce our probability models. Let $\Gamma_X = (X'_{(1)}, X'_{(2)}, \ldots, X'_{(n)})$ and $\Gamma_Y = (Y'_{(1)}, Y'_{(2)}, \ldots, Y'_{(n)})$ be two groups of stochastically ordered random variables (not necessarily the order statistics) which represent the increasing "strengths" of the ordered objects in the two groups, respectively. Correspondingly, we suppose that in a particular ordered matching hypothetical realizations $x'_{(i)}$ of $X_{(i)}$ and $y'_{(i)}$ of $Y_{(i)}$, $i = 1, \ldots, n$, are compared. While we
cannot observe $x'_{(i)}$ or $y'_{(i)}$, we can make the usually subjective judgment whether $y'_{(i)} > x'_{(i)}$ or $y'_{(i)} < x'_{(i)}$. Then we prefer the Y-group, $\Gamma_Y$, to the X-group, $\Gamma_X$, in this particular matching if

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

where

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

For the present, we ignore ties, and assume that the random variables are absolutely continuous. We regard $\Gamma_Y$ as superior to $\Gamma_X$ under ordered matching if

$$ES \equiv E \sum_{i=1}^{n} I(Y'_{(i)} > X'_{(i)})$$

$$= \sum_{i=1}^{n} P(Y'_{(i)} > X'_{(i)}) \geq \frac{1}{2} n.$$

Other matchings may be obtained by pairing $Y'_{(i)}$ with $X'_{(\pi_i)}$, $i = 1, \ldots, n$, where $\pi = (\pi_1, \ldots, \pi_n)$ is a permutation of $(1, \ldots, n)$. Correspondingly, we will speak of a matching $\pi$ and write

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

(1.1)

to denote the random number of preferences for objects in $\Gamma_Y$. The expected value of $S(\pi)$ for fixed $\pi$ is given by

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

(1.2)

Therefore, for an ordered matching ($\pi = \pi^o = (1, \ldots, n)$), we have

$$E[S(\pi^o)] = \sum_{i=1}^{n} P(Y'_{(i)} > X'_{(i)}) = V_1 \text{ (say)}.$$
The expected value of $S(\pi)$ under random matching is given by

$$V_2 = \frac{1}{n!} \sum_{(\pi_1, \ldots, \pi_n)} \sum_{i=1}^{n} P(Y'_i > X'_i(\pi_i)),$$

where the first sum is over all possible permutations $(\pi_1, \ldots, \pi_n)$. Simplifying the above summation, we have

$$V_2 = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} P(Y'_i > X'_j).$$

**Definition 1.1** A matching $\pi$ is said to be fair if $E[S(\pi)] = \frac{1}{2}n$ when $\Gamma_X \sim \Gamma_Y$.

It is clear that ordered matching is fair. It can be shown that random matching is fair also. Other matchings are not necessarily fair and we shall examine this aspect further in later chapters.

We use $F_i(x)$ and $G_i(x)$ to represent the cdf's of $X'_i$ and $Y'_i$, and $f_i(x)$ and $g_i(x)$ to represent the pdf's of $X'_i$ and $Y'_i$, respectively, $i = 1, \ldots, n$. Here we assume $F_i(x) \geq F_j(x)$ for all $x$ and for any $1 \leq i < j \leq n$, i.e., $X'_i$ is stochastically smaller than $X'_j$ or $X'_i \leq_{st} X'_j$. Therefore, we have $X'_1 \leq_{st} X'_2 \leq_{st} \ldots \leq_{st} X'_n$. Similarly, $Y'_1 \leq_{st} Y'_2 \leq_{st} \ldots \leq_{st} Y'_n$. Usually, we also assume that $\Gamma_X$ and $\Gamma_Y$ are independent; however, we do not assume independence within $\Gamma_X$ and $\Gamma_Y$.

(a) **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 two random samples of size $n$ from $F$ and $G$, respectively. Then the pdf of $X'_i$, $f_i(x)$, is given by

$$f_i(x) = \frac{n!}{(i-1)!(n-i)!} [F(x)]^{i-1} [1 - F(x)]^{n-i} f(x).$$

(1.5)
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 $X'_{(i)}$'s may, in fact, have any dependence structure, including independence.

In particular, if $\Gamma_X = (X_{(1)}, \ldots, X_{(n)})$ and $\Gamma_Y = (Y_{(1)}, \ldots, Y_{(n)})$, then $S(\pi)$ becomes Galton's rank-order statistic (e.g., Hodges, 1955). We mention this only to make it clear that we are not considering this case which does not permit $X'_{(i)} > X'_{(j)}$ for $i < j$.

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

$$V_1 = E[S(\pi^\circ)] = \sum_{i=1}^n P(Y_i > X_i).$$

Note that random matching merely returns us to the unordered case for which we have

$$V_2 = \sum_{i=1}^n P(Y_i > X_i) = nP(Y > X),$$

where $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ are random samples from $F$ and $G$ respectively, with $X \sim F$ and $Y \sim G$.

The question of whether ordered matching is more effective than random matching may now be reduced to the question of whether $V_1 \geq V_2$ if $X \leq_{st} Y$. The answer is yes under certain conditions. We deal with this and related issues in Chapter 2. In fact, a much stronger result will be proved, namely that if $X \leq_{st} Y$, then $V_1 \geq E[S(\pi)]$ for any simple matching and any symmetrical matching, where simple
and symmetrical matchings are both fair matchings defined in Chapter 2. In addition, the properties of \( p_{i,j} = P(Y(i) > X(j)) \) are of interest in themselves, particularly their relation to \( p = P(Y > X) \), their dependence on \( G \) (for given \( F \) and \( i = j \)), and on \( i \) (for given \( F, G, \) and \( i = j \)). Also of interest is the limiting behavior of \( p_{i,j} \) as \( n \to \infty \).

(b) Linear Preference Model

In this model, we assume that \( X^i(i) \sim F(x - \lambda(i)) \) and \( Y^i(i) \sim F(x - \mu(i)) \), \( i = 1, \ldots, n \), where \( F(x) \) is a distribution function and \( \lambda(1) \leq \lambda(2) \leq \ldots \leq \lambda(n) \) and \( \mu(1) \leq \mu(2) \leq \ldots \leq \mu(n) \) are ordered real numbers. The model is based on the linear model much used in the method of paired comparison (e.g., David, 1988, p.7). At times, we will assume that \( F(x) \) is a unimodal distribution function. The definition of a unimodal distribution will be given in Chapter 3. We will see that the class of unimodal distributions 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 \( X \) and \( Y \) are groups of independent random variables; \( S(\pi^0) \) has a Binomial\((n, \frac{1}{2})\) distribution. In general, there is no closed form for the distribution of \( S(\pi) \), and we are not going to investigate this issue here.

In this model, we are still interested in comparing \( V_1 \) and \( V_2 \), as well as \( V_1 \) and \( E[S(\pi)] \). The following questions arise: Under what conditions are \( n \sum_{i=1}^{n} \mu(i) \geq n \sum_{i=1}^{n} \lambda(i) \)? The answer is yes when the sample size \( n = 2 \). When \( n > 2 \), the situation becomes more complex. We will study this and related issues in Chapter 3.
It is noted that for given \( \mu(i) \) and \( \lambda(i) \) \( (i = 1, \ldots, n) \), there is always a permutation \( \pi' \) such that \( E[S(\pi')] \geq E[S(\pi)] \) for all \( \pi \). We will give a sufficient condition for such \( \pi' \). We will also discuss some rearrangement properties of \( E[S(\pi)] \).

In the discussions of (a) and (b), we have ignored ties. However ties are possible in practice. As in a chess game, it is possible that some comparison ends in a tie. Usually, a tie is caused when the performances of the two objects are too close to tell the difference; it is not necessarily caused by two objects having exactly the same “strength”.

We now introduce an indicator function \( I(u, v; \tau) \) which is defined as follows:

\[
I(u, v; \tau) = \begin{cases} 
0 & \text{if } u - v < -\tau \\
\frac{1}{2} & \text{if } |u - v| \leq \tau \\
1 & \text{if } u - v > \tau,
\end{cases}
\]

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

For any permutation \( \pi = (\pi_1, \ldots, \pi_n) \), we define

\[
S_\tau(\pi) = \sum_{i=1}^{n} I(Y'_{(i)}, X'_{(\pi_i)}; \tau).
\]

Then \( S_\tau(\pi) \) is a random variable which measures the performance of \( \Gamma_Y \) with respect to \( \Gamma_X \) under the matching \( \pi \), with ties permitted. Correspondingly, the expectation of \( S_\tau(\pi) \) is given by

\[
E[S_\tau(\pi)] = \sum_{i=1}^{n} P(Y'_{(i)} > X'_{(\pi_i)} + \tau) + \frac{1}{2} \sum_{i=1}^{n} P(|Y'_{(i)} - X'_{(\pi_i)}| < \tau).
\]

It can be shown that when \( \tau \) is “small”, we get the same results in both models as when ties are ignored. However, when \( \tau \) is “large”, this is no longer necessarily the case. We will discuss these issues towards the ends of Chapters 2 and 3.
2. ORDER STATISTICS MODEL

2.1 Introduction

Let X and Y be independent rv's with respective cdf's $F(x)$ and $G(x)$ and pdf's $f(x)$ and $g(x)$. Let $(X(1), X(2), \ldots, X(n))$ and $(Y(1), Y(2), \ldots, Y(n))$ be the order statistics of random samples $(X_1, X_2, \ldots, X_n)$ and $(Y_1, Y_2, \ldots, Y_n)$ from $F(x)$ and $G(x)$ respectively. Consider the two groups of stochastically ordered rv's $\Gamma_X = (X'(1), X'(2), \ldots, X'(n))$ and $\Gamma_Y = (Y'(1), Y'(2), \ldots, Y'(n))$. In this chapter we assume that $X'(i)$ and $Y'(i)$ have the same marginal distributions as $X(i)$ and $Y(i)$, respectively. Note that the cdf of the i-th order statistics $X(i)$ is given by

$$F_i(x) = \sum_{k=i}^{n} \binom{n}{k} [F(x)]^k [1 - F(x)]^{n-k}$$

$$= \int_{-\infty}^{x} \frac{n!}{(i-1)!(n-i)!} [F(t)]^{i-1} [1 - F(t)]^{n-i} f(t) dt.$$ 

Hence the pdf of $X(i)$, $f_i(x)$, is given by (1.5) in Chapter 1. We can get similar expressions for the cdf and pdf of $Y(i)$, i.e., $G_i(x)$ and $g_i(x)$.

As noted in Chapter 1, for any permutation $\pi = (\pi_1, \ldots, \pi_n)$, or a matching $\pi$, $S(\pi) = \sum_{i=1}^{n} I(Y'_{\pi(i)} > X'_{(i)})$ does not necessarily have the same distribution as
\[ \sum_{i=1}^{n} I(Y(i) > X(\pi_i)). \] However, their expectations are the same, i.e.,

\[ E[S(\pi)] = \sum_{i=1}^{n} P(Y(i) > X(\pi_i)). \tag{2.1} \]

Here \( S(\pi) \) is the number of times that the \( Y \)'s in \( \Gamma_Y \) are greater than the \( X \)'s in \( \Gamma_X \) under the matching \( \pi \).

Under the order statistics model, the expectation of \( S(\pi) \) under ordered matching, i.e., \( \pi = \pi^O = (1, 2, \ldots, n) \) is given by

\[ V_1 = E[S(\pi^O)] = \sum_{i=1}^{n} P(Y(i) > X(i)) \tag{2.2} \]

and the expectation of \( S(\pi) \) under random matching is given by

\[ V_2 = \sum_{i=1}^{n} P(Y_i > X_i) = nP(Y > X). \tag{2.3} \]

The main question concerning us is how \( V_1 \) is related to \( V_2 \) and \( E[S(\pi)] \). In addition, we will investigate the properties of \( p_{ij} = P(Y(i) > X(j)) \) and some special matchings.

### 2.2 Ordered Matching and Random Matching

In this section, we will discuss the relationship between ordered matching and random matching for two groups of ordered random variables

\[ \Gamma_X = (X_1, X_2, \ldots, X_n) \text{ and } \Gamma_Y = (Y_1, Y_2, \ldots, Y_n). \]

It is noted that there are \( n! \) possible matchings for \( \Gamma_X \) and \( \Gamma_Y \). We say a matching \( \pi \) is random if the matching is picked at random from the \( n! \) possible
matchings. Therefore, the expectation of $S(\pi)$ under random matching is given by

$$V_2 = \frac{1}{n!} \sum_{(\pi_1, \ldots, \pi_n)} \sum_{i=1}^{n} P(Y(i) > X(\pi_i)).$$

where the first sum is over all possible permutations $(\pi_1, \ldots, \pi_n)$. The above summation is the sum of probabilities $P(Y(i) > X(j))$, and for any fixed $\pi_i = j$, there are $(n - 1)!$ possible permutations which contain $P(Y(i) > X(j))$, so that,

$$\sum_{(\pi_1, \ldots, \pi_n)} \sum_{i=1}^{n} P(Y(i) > X(\pi_i)) = \sum_{i=1}^{n} \sum_{j=1}^{n} (n - 1)! P(Y(i) > X(j)).$$

Also

$$P(Y(i) > X(j)) = \int_{-\infty}^{\infty} P(X(j) < x)g_i(x) \, dx$$

$$= \int_{-\infty}^{\infty} \sum_{k=j}^{n} \binom{n}{k} [F(x)]^{k-1}[1 - F(x)]^{n-k} \frac{n!}{(i - 1)!(n - i)!} [G(x)]^{i-1}[1 - G(x)]^{n-i} g(x) \, dx.$$ 

Note that

$$\sum_{k=j}^{n} \binom{n}{k} p^k(1 - p)^{n-k} = \frac{1}{B(j, n - j + 1)} \int_0^p t^{j-1}(1 - t)^{n-j} \, dt$$

for $0 \leq p \leq 1$. With $C_{n,l} = \frac{n!}{(l - 1)!(n-l)!}$, we can write

$$P(Y(i) > X(j)) =$$

$$C_{n,i} C_{n,j} \int_{-\infty}^{\infty} \int_0^{F(x)} t^{j-1}(1 - t)^{n-j} [G(x)]^{i-1}[1 - G(x)]^{n-i} g(x) \, dt \, dx$$

$$= C_{n,i} C_{n,j} \int_0^1 \int_0^{F[G^{-1}(u)]} t^{j-1}(1 - t)^{n-j} u^{i-1}(1 - u)^{n-i} \, dt \, du.$$ 

Also since

$$\sum_{i=1}^{n} \sum_{j=1}^{n} P(Y(i) > X(j)) = n^2 P(Y > X)$$

we can write $V_2$ as (2.3).

We first give a theorem that will be used frequently.
Theorem 2.2.1 If $X \sim F(x)$ and $Y \sim F(x-\mu)$ where $\mu \geq 0$ and $X$ is an absolutely continuous rv. Then for $i < j$, we have

$$P(Y(i) > X(i)) + P(Y(j) > X(j)) \geq P(Y(i) > X(j)) + P(Y(j) > X(i)). \quad (2.6)$$

Proof. Consider the 4-dimensional function

$$D(x(i), x(j), y(i), y(j)) = I(y(i) > x(i)) + I(y(j) > x(j)) - I(y(i) > x(j)) - I(y(j) > x(i)).$$

on $A = \{x(i) < x(j), y(i) < y(j)\}$. Then ignoring possible ties, $A$ can be partitioned into the following six parts:

(a) $x(i) < x(j) < y(i) < y(j)$,
(b) $x(i) < y(i) < x(j) < y(j)$,
(c) $x(i) < y(i) < y(j) < x(j)$,
(d) $y(i) < y(j) < x(i) < x(j)$,
(e) $y(i) < x(i) < y(j) < x(j)$,
(f) $y(i) < x(i) < x(j) < y(j)$.

It is easy to check that $D(x(i), x(j), y(i), y(j))$ equals $-1$ for case (e), $1$ for case (b) and $0$ otherwise. Therefore, since $Y \overset{d}{=} X + \mu$ and $\mu \geq 0$, we have

$$ED(X(i), X(j), Y(i), Y(j)) = P(X(i) < Y(i) < X(j) < Y(j)) - P(Y(i) < X(i) < Y(j) < X(j)) \geq 0.$$
Note that $ED(X_{(i)}, X_{(j)}, Y_{(i)}, Y_{(j)})$ is the LHS of (2.6) minus its RHS. It follows that (2.6) holds. □

Under some additional conditions a different proof of Theorem 2.2.1 is given in Appendix A.

**Theorem 2.2.2** If $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ are two random samples from populations with cdf's $F(x)$ and $F(x - \mu)$ respectively, where $\mu \geq 0$, then

\[
(a) \sum_{i=1}^{k} \sum_{j=1}^{k} P(Y_i > X_j) \leq k \sum_{i=1}^{k} P(Y_i > X_i) \quad k = 1, \ldots, n, \\
(b) \sum_{i=1}^{n} P(Y_i > X_i) \geq \sum_{i=1}^{n} P(Y_i > X_i) = n P(Y > X).
\]

**Proof.** (a) Write $pij = P(Y_i > X_j)$. By Theorem 2.2.1, we have $pii + pjj \geq pij + pji$ for any $1 \leq i, j \leq n$. Now, for any $k \leq n$, it follows that

\[
2k \sum_{i=1}^{k} p_{ii} = \sum_{i=1}^{k} \sum_{j=1}^{k} (p_{ii} + p_{jj}) \geq \sum_{i=1}^{k} \sum_{j=1}^{k} (p_{ij} + p_{ji}) = 2 \sum_{i=1}^{k} \sum_{j=1}^{k} p_{ij},
\]

i.e., (a) holds.

(b) Note that $\sum_{i=1}^{n} \sum_{j=1}^{n} P(Y_i > X_j) = n^2 P(Y > X)$. Then from (a) by taking $k = n$, (b) follows immediately. □

Part (b) of Theorem 2.2.2 shows that ordered matching has more power to identify the stronger group than random matching.

**Lemma 2.2.1** Let $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ be two independent random samples from populations with cdf's $F(x)$ and $F(x - \mu)$, respectively, where $F'(x) = f(x)$.
is symmetric about 0. Then
\[
\sum_{i=1}^{n} P(Y(i) > X(i))
\]
\[
= \frac{n}{2} + \sum_{i=1}^{n} \int_{0}^{+\infty} [P(X(i) < x + \mu) - P(X(i) < x - \mu)] f_i(x) \, dx
\]
\[
= \frac{n}{2} + \sum_{i=1}^{n} \int_{0}^{+\infty} P(x - \mu < X(i) < x + \mu) f_i(x) \, dx.
\]

Proof. We have
\[
p_i = \int_{-\infty}^{+\infty} P(X(i) < x + \mu) f_i(x) \, dx
\]
\[
= \int_{-\infty}^{0} P(X(i) < x + \mu) f_i(x) \, dx + \int_{0}^{+\infty} P(X(i) < x + \mu) f_i(x) \, dx
\]
\[
= \int_{0}^{+\infty} P(X(i) < -x + \mu) f_i(-x) \, dx + \int_{0}^{+\infty} P(X(i) < x + \mu) f_i(x) \, dx.
\]
Since \(P(X(i) < -x + \mu) = P(X(n-i+1) > x - \mu) = 1 - P(X(n-i+1) < x - \mu)\) and \(f_i(-x) = f_{n-i+1}(x)\), it follows that
\[
\sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \int_{0}^{+\infty} f_{n-i+1}(x) \, dx - \sum_{i=1}^{n} \int_{0}^{+\infty} P(X(n-i+1) < x - \mu) f_{n-i+1}(x) \, dx
\]
\[
+ \sum_{i=1}^{n} \int_{0}^{+\infty} P(X(i) < x + \mu) f_i(x) \, dx.
\]
Also from \(\sum_{i=1}^{n} f_{n-i+1}(x) \, dx = nf(x)\), we have
\[
\sum_{i=1}^{n} p_i = \frac{n}{2} + \sum_{i=1}^{n} \int_{0}^{+\infty} [P(X(i) < x + \mu) - P(X(i) < x - \mu)] f_i(x) \, dx
\]
\[
= \frac{n}{2} + \sum_{i=1}^{n} \int_{0}^{+\infty} [P(x - \mu < X(i) < x + \mu)] f_i(x) \, dx. \quad \Box
\]
It is easy to see that when $\mu = 0$, we have $\sum_{i=1}^{n} p_i = \frac{n}{2}$. Therefore, for $\mu > 0$,

$$\sum_{i=1}^{n} \int_{0}^{+\infty} [P(x - \mu < X(i) < x + \mu)] f_i(x) \, dx$$

is the increment of $\sum_{i=1}^{n} p_i$.

### 2.3 Fair Matching under the Order Statistics Model

We have given the definition of a fair matching in Chapter 1. In this section we distinguish two classes of fair matching and then investigate the relationship between ordered matching and such fair matchings. We first introduce the following lemma.

**Lemma 2.3.1 (Gastwirth, 1968)** If $(X_1, X_2, \ldots, X_n)$ and $(Y_1, Y_2, \ldots, Y_n)$ are two iid random samples, then

$$P(Y(i) > X(j)) = \sum_{k=j}^{i+j-1} \frac{\binom{n}{k}\binom{2n}{i+j-1-k}}{\binom{2n}{i+j-1}}. \quad (2.7)$$

The proof of the above lemma is based on the argument that the event $\{Y(i) > X(j)\}$ is equivalent to the event that at least $j$ X's appear in the first $i + j - 1$ positions of a sequence with $n$ X's and $n$ Y's.

Throughout this section, we will assume that $(X_1, X_2, \ldots, X_n)$ and $(Y_1, Y_2, \ldots, Y_n)$ are two iid random samples except when otherwise indicated.

Before we go further, let us look at all possible matchings for the case of sample size $n = 4$. By Lemma 2.3.1, we have

$$P(Y(1) > X(2)) = \binom{4}{5}/\binom{8}{2} = \frac{3}{14},$$

$$P(Y(1) > X(3)) = \binom{4}{3}/\binom{8}{3} = \frac{1}{14},$$

$$P(Y(1) > X(4)) = \binom{4}{4}/\binom{8}{4} = \frac{1}{14},$$
Therefore, using the above results for each matching $\pi = (\pi_1, \pi_2, \pi_3, \pi_4)$, we can find the value of $E[S(\pi)]$ as in the Table 2.1.

Table 2.1: Values of $E[S(\pi)]$ for all permutations $\pi$ of $(1, 2, 3, 4)$

<table>
<thead>
<tr>
<th>No.</th>
<th>$(\pi_1, \pi_2, \pi_3, \pi_4)$</th>
<th>$E[S(\pi)]$</th>
<th>No.</th>
<th>$(\pi_1, \pi_2, \pi_3, \pi_4)$</th>
<th>$E[S(\pi)]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(1, 2, 3, 4)$</td>
<td>2</td>
<td>13</td>
<td>$(3, 1, 2, 4)$</td>
<td>$\frac{148}{70}$</td>
</tr>
<tr>
<td>2</td>
<td>$(1, 2, 4, 3)$</td>
<td>2</td>
<td>14</td>
<td>$(3, 1, 4, 2)$</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>$(1, 3, 2, 4)$</td>
<td>$\frac{132}{70}$</td>
<td>15</td>
<td>$(3, 2, 1, 4)$</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>$(1, 3, 4, 2)$</td>
<td>$\frac{148}{70}$</td>
<td>16</td>
<td>$(3, 2, 4, 1)$</td>
<td>$\frac{124}{70}$</td>
</tr>
<tr>
<td>5</td>
<td>$(1, 4, 2, 3)$</td>
<td>$\frac{148}{70}$</td>
<td>17</td>
<td>$(3, 4, 1, 2)$</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>$(1, 4, 3, 2)$</td>
<td>2</td>
<td>18</td>
<td>$(3, 4, 2, 1)$</td>
<td>$\frac{132}{70}$</td>
</tr>
<tr>
<td>7</td>
<td>$(2, 1, 3, 4)$</td>
<td>2</td>
<td>19</td>
<td>$(4, 1, 2, 3)$</td>
<td>$\frac{164}{70}$</td>
</tr>
<tr>
<td>8</td>
<td>$(2, 1, 4, 3)$</td>
<td>2</td>
<td>20</td>
<td>$(4, 1, 3, 2)$</td>
<td>$\frac{156}{70}$</td>
</tr>
<tr>
<td>9</td>
<td>$(2, 3, 1, 4)$</td>
<td>$\frac{132}{70}$</td>
<td>21</td>
<td>$(4, 2, 1, 3)$</td>
<td>$\frac{156}{70}$</td>
</tr>
<tr>
<td>10</td>
<td>$(2, 3, 4, 1)$</td>
<td>$\frac{116}{70}$</td>
<td>22</td>
<td>$(4, 2, 3, 1)$</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>$(2, 4, 1, 3)$</td>
<td>2</td>
<td>23</td>
<td>$(4, 3, 1, 2)$</td>
<td>$\frac{148}{70}$</td>
</tr>
<tr>
<td>12</td>
<td>$(2, 4, 3, 1)$</td>
<td>$\frac{124}{70}$</td>
<td>24</td>
<td>$(4, 3, 2, 1)$</td>
<td>2</td>
</tr>
</tbody>
</table>

By examining the Table 2.1, we find that it contains two types of fair matchings. The first type of matching is called *simple matching* which is defined as follows.

Let $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ be a permutation of $(1, 2, \ldots, n)$ and write $\pi^0 = (1, 2, \ldots, n)$. 

\[
P(Y(2) > X(3)) = \binom{4}{4} \frac{1}{4} \binom{4}{4} \frac{1}{8} = \frac{17}{70},
\]
\[
P(Y(2) > X(4)) = \binom{4}{4} \frac{1}{4} \binom{4}{1} \frac{1}{5} = \frac{1}{14},
\]
\[
P(Y(3) > X(4)) = \binom{4}{4} \frac{1}{4} \binom{4}{8} = \frac{3}{14}.
\]
Definition 2.1 $\pi$ is said to be a simple matching (or permutation) if it can be obtained from $\pi^0$ by interchanging pairs of the components of $\pi^0$ with no component involved in more than one interchange.

For example, when $n = 4$, $\pi = (3, 4, 1, 2)$ is a simple matching since it is obtained by interchanging pair $\{1, 3\}$ and pair $\{2, 4\}$ of $(1, 2, 3, 4)$. However, the matching $\pi = (2, 3, 1, 4)$ is not simple.

Lemma 2.3.2 All simple matchings are fair and the total number of simple matchings is given by $\sum_{i=0}^{[\frac{n}{2}]} (2i - 1)!! \left(\begin{array}{c} n \\ 2i \end{array}\right)$, where $(2i - 1)!! = (2i - 1)(2i - 3)\ldots1$.

Proof. Since $P(Y_i > X(j)) + P(Y_j > X(i)) = 1 - P(Y_i < X(j)) + P(Y_j > X(i)) = 1 = P(Y_i > X(i)) + P(Y_j > X(j))$, we see that interchanging a pair from an ordered matching does not affect the sum of the two probabilities. Therefore, for any simple matching $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$, by its definition, we have
\[
\sum_{i=1}^{n} P(Y_i > X(\pi_i)) = \sum_{i=1}^{n} P(Y_i > X(i)) = \frac{n}{2},
\]
i.e., $\pi$ is fair.

Now let us find the total number of simple matchings. Note that the simple matchings consist of those with $0, 1, \ldots, \left[\frac{n}{2}\right]$ interchanges of pairs. The number of simple matchings with $i$ pair exchanges is
\[
\frac{1}{i!} \binom{n}{2} \binom{n-2}{2} \binom{n-4}{2} \cdots \binom{n-2(i-1)}{2}.
\]
Therefore, the total number is given by
\[
1 + \sum_{i=1}^{\left[\frac{n}{2}\right]} \frac{1}{i!} \prod_{j=1}^{i} \binom{n-2(j-1)}{2}.
\]
\[ 1 + \sum_{i=1}^{\left\lfloor \frac{n}{2} \right\rfloor} \frac{n!}{i!2^i(n-2i)!} = \sum_{i=0}^{\left\lfloor \frac{n}{2} \right\rfloor} (2i-1)!! \binom{n}{2i}. \]

**Definition 2.2** \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \) is said to be a symmetric matching (or permutation) if \( \pi_{n-i+1} = n - \pi_i + 1 \) for \( i = 1, 2, \ldots, n \).

For example, the matchings No.11 and No.14 in Table(2.1) are symmetric.

We can show that symmetric matchings are fair also. Before we do this, we need the following lemma.

**Lemma 2.3.3** If \((X_1, X_2, \ldots, X_n)\) and \((Y_1, Y_2, \ldots, Y_n)\) are two iid random samples, then for any \(1 \leq i, j \leq n\), we have

\[ P(Y_{(j)} > X_{(i)}) = P(Y_{(n-i+1)} > X_{(n-j+1)}). \quad (2.8) \]

**Proof.** Since \( P(Y_{(j)} > X_{(i)}) \) does not depend on the common distribution of \( X \)'s and \( Y \)'s, we can take this to be symmetric about 0. Then

\[ P(Y_{(j)} > X_{(i)}) = P(-X_{(i)} > -Y_{(j)}) = P(X_{(n-i+1)} > Y_{(n-j+1)}). \]

The result follows since \( X_{(i)} \overset{d}{=} Y_{(i)}, i = 1, 2, \ldots, n. \)

The above lemma can also be shown directly by using Lemma 2.3.1.

**Lemma 2.3.4** All symmetric matchings are fair and the total number of symmetric matchings is given by

(a) \( n(n-2)(n-4) \ldots 4 \cdot 2 \) if \( n \) is even.

(b) \( (n-1)(n-3) \ldots 3 \cdot 1 \) if \( n \) is odd.
Proof. We first show that symmetric matchings are fair. Since \( P(Y(i) > X(j)) = 1 - P(X(j) > Y(i)) = 1 - P(Y(j) > X(i)) \), we have, by Lemma (2.3.3).

\[
P(Y(i) > X(j)) + P(Y(n-i+1) > X(n-j+1)) = 1.
\] (2.9)

Now let \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \) be any symmetric permutation. Replacing \( j \) by \( \pi_i \) in (2.9), we have

\[
P(Y(i) > X(\pi_i)) + P(Y(n-i+1) > X(n-\pi_i+1)) = 1.
\]

Since \( \pi_{n-i+1} = n - \pi_i + 1 \), it follows that

\[
P(Y(i) > X(\pi_i)) + P(Y(n-i+1) > X(n-\pi_i+1)) = 1.
\]

Therefore, we have

\[
\sum_{i=1}^{n} [P(Y(i) > X(\pi_i)) + P(Y(n-i+1) > X(\pi_{n-i+1}))] = n,
\]

so that

\[
2 \sum_{i=1}^{n} P(Y(i) > X(\pi_i)) = n.
\]

Now let us prove the second part of the lemma. Let \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \) be any symmetric permutation. Then if \( n = 2m \), we have \( n \) choices for \( \pi_1 \). Once \( \pi_1 \) is fixed, \( \pi_n \) is fixed also. Therefore, we have \( n - 2 \) choices for \( \pi_2 \). Similarly, once \( \pi_2 \) is fixed, \( \pi_{n-1} \) is also fixed. Therefore, we have \( n - 4 \) choices for \( \pi_3 \). Continuing this argument, we have the total number of symmetric matchings (a).

When \( n = 2m - 1 \), we have \( \pi_m = m \). Therefore, we have \( n - 1 \) choices for \( \pi_1 \).

By the same argument as above, it follows that the total number is given by (b). \( \square \)

Note:
\[
\{ \text{simple matchings} \} \cap \{ \text{symmetric matchings} \} \neq \phi.
\]
For example, if \( n = 4 \), then \((4,2,3,1)\) is both a simple and a symmetric matching (or permutation).
There exist matchings which are simple but not symmetric, e.g., \((2,1,3,4)\). Also there exist matchings which are symmetric but not simple, e.g., \((2,4,1,3)\).

(2) There exist fair matchings which are neither simple nor symmetric. For example, let \( n = 8 \), the following matching is neither simple nor symmetric:

\[
(2, 1, 4, 6, 3, 5, 7, 8).
\]

In the above matching, the middle part is symmetric and the remainder is simple. Therefore, using the same arguments used in proving that simple and symmetric matchings are fair, we see that it is fair.

In general, we have the following result:

**Lemma 2.3.5** Any matching that is a combination of simple and symmetric matchings is fair.

**Proof.** Separate the matching into two parts. The first part is the simple matching part, and the second part is the symmetric matching part. Then the rest of the proof is same as before. \( \square \)

We might ask if all fair matchings consist of simple matchings, symmetric matchings, and combinations of simple and symmetric matchings? The answer probably is yes; however, this has not been proved yet.

**Theorem 2.3.1** Let \((X_1, X_2, \ldots, X_n)\) and \((Y_1, Y_2, \ldots, Y_n)\) be two independent random samples from populations with cdf's \(F(x)\) and \(F(x-\mu)\), respectively, where \(\mu \geq 0\).
Then

\[ (a) \sum_{i=1}^{n} P(Y(i) > X(i)) \geq \sum_{i=1}^{n} P(Y(i) > X(\pi_i)) \quad (2.10) \]

for any simple permutation \((\pi_1, \pi_2, \ldots, \pi_n)\).

(b) If \(F(x)\) is the cdf of a symmetric rv, then (2.10) holds for any symmetric permutation \((\pi_1, \pi_2, \ldots, \pi_n)\).

Proof. (a) \((\pi_1, \pi_2, \ldots, \pi_n)\) is obtained by permuting the components of \(\pi^0 = (1, 2, \ldots, n)\) and once two components, say \(i\) and \(j\) \((i < j)\) are interchanged, then neither \(i\) nor \(j\) can be interchanged with any other components of \(\pi^0\). Therefore, by Theorem 2.2.1, (a) follows.

(b) For simplicity, assume \(X \sim F(x)\) is symmetric about 0. Then, since \(\pi_{n-i+1} = n - \pi_i + 1\), we have

\[
\sum_{i=1}^{n} P(Y(i) > X(\pi_i)) \\
= \frac{1}{2} \left[ \sum_{i=1}^{n} P(Y(i) > X(\pi_i)) + \sum_{i=1}^{n} P(Y(n-i+1) > X(\pi_{n-i+1})) \right] \\
= \frac{1}{2} \sum_{i=1}^{n} [P(Y(i) > X(\pi_i)) + P(Y(n-i+1) > X(n-\pi_i+1))].
\]

Since \(X(n-\pi_i+1) \overset{d}{=} -X(\pi_i)\), and \((Y-\mu)(n-i+1) \overset{d}{=} -(Y-\mu)(i)\), i.e., \(Y(n-i+1) \overset{d}{=} 2\mu - Y(i)\), it follows that

\[
P(Y(n-i+1) > X(n-\pi_i+1)) = P(2\mu - Y(i) > -X(\pi_i)) \\
= P(X(\pi_i) + \mu > Y(i) - \mu) = P((X + \mu(\pi_i)) > (Y - \mu)(i)) \\
= P(Y(\pi_i) > X(i)).
\]
Therefore, by Theorem 2.2.1, we have

\[ P(Y(i) > X(\pi_i)) + P(Y(i) > X(n - \pi_i + 1)) = P(Y(i) > X(\pi_i)) + P(Y(\pi_i) > X(i)) \leq P(Y(i) > X(i)) + P(Y(\pi_i) > X(\pi_i)). \]

Hence, (b) follows immediately. \( \square \)

Under the conditions of the above theorem, do we have that (2.10) \( V_1 \geq E[S(\pi)] \) in the notation used before) holds for any permutation \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \)? In general, the answer is no. For example, let \( F(x) = \Phi(x) \), i.e., the standard normal distribution, and \( \mu = 0.5 \). For \( n = 4 \), we can check that

\[ V_1 < P(Y(1) > X(4)) + \sum_{i=2}^{4} P(Y(i) > X(i-1)). \]

However, if \( \mu \) is "sufficiently large", (2.10) might hold for any permutation \( \pi \). Therefore, we make the following conjecture:

**Conjecture** If \((X_1, X_2, \ldots, X_n)\) and \((Y_1, Y_2, \ldots, Y_n)\) are two independent random samples from populations with cdf's \( F(x) \) and \( F(x - \mu) \), respectively, then there exists \( \mu_0 > 0 \) such that when \( \mu > \mu_0 \), \( V_1 \geq E[S(\pi)] \) for any matching (or permutation) \( \pi \).

In Chapter 3, i.e., for the linear preference model, we will prove a similar result as the above conjecture.

**Lemma 2.3.6** Let \( X \sim F(x) \) and \( Y \sim G(x) \), where both \( X \) and \( Y \) are symmetric rv's with the same point of symmetry. Let \((X_1, X_2, \ldots, X_n)\) and \((Y_1, Y_2, \ldots, Y_n)\) be two independent random samples from \( F(x) \) and \( G(x) \), respectively. Then, for any
symmetric permutation \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \), we have

\[
\sum_{i=1}^{n} P(Y(i) > X(\pi_i)) = \frac{n}{2}.
\]

**Proof.** Since

\[
\sum_{i=1}^{n} P(Y(i) > X(\pi_i)) = \frac{1}{2} \left[ \sum_{i=1}^{n} P(Y(i) > X(\pi_i)) + \sum_{i=1}^{n} P(Y(n-i+1) > X(\pi_{n-i+1})) \right]
\]

and

\[
P(Y(n-i+1) > X(\pi_{n-i+1})) = P(-Y(i) > -X(\pi_i)) = 1 - P(Y(i) > X(\pi_i)),
\]

the result follows immediately. •

### 2.4 Results on \( P(Y(i) > X(j)) \)

In this section, we will investigate the properties of \( p_{ij} = P(Y(i) > X(j)) \) under different population distributions. When \( i = j \), we simply write \( p_i = p_{ii} \). We first look at \( p_i \). By (2.5), \( p_i \) can be written as

\[
p_i = C_{n,i}^2 \int_{-\infty}^{\infty} \int_{0}^{F(x)} t^{i-1}(1 - t)^{n-i} \left[ G(x) \right]^{i-1} \left[ 1 - G(x) \right]^{n-i} g(x) dt dx
\]

\[
= C_{n,i}^2 \int_{0}^{1} \int_{0}^{F[G^{-1}(u)]} t^{i-1}(1 - t)^{n-i} u^{i-1} (1 - u)^{n-i} du dx.
\]

**Lemma 2.4.1**
(a) \( p_i \geq \frac{1}{2} \) according as \( Y \geq_{st} X \).

(b) If \( Y^* >_{st} Y \), then \( p_i^* \equiv P(Y^*_i > X(i)) > p_i \).

**Proof.** (a) Since \( p_i = \frac{1}{2} \) for \( Y \overset{d}{=} X \), by (2.11), we have \( p_i > \frac{1}{2} \) iff \( F[G^{-1}(u)] > u \), i.e., iff \( G^{-1}(u) > F^{-1}(u) \) or \( G(x) < F(x) \). Similarly, \( p_i < \frac{1}{2} \) iff \( G(x) > F(x) \).

(b) Since \( G^{*-1}(u) > G^{-1}(u) \), we have \( F[G^{*-1}(u)] > F[G^{-1}(u)] \), which proves the result. \( \Box \)

**Lemma 2.4.2** If \( X \sim F(x) \), \( Y \sim F(x - \mu) \) and \( X \) and \( Y \) are independent, then \( Y(i) - X(i) \) is symmetric about \( \mu \).

**Proof.** Since \( Y \overset{d}{=} X^* + \mu \), where \( X^* \) and \( X \) are iid, we have \( Y(i) \overset{d}{=} X^*_i + \mu \), where \( X^*_i \) and \( X(i) \) are iid. Therefore \( Y(i) - X(i) \overset{d}{=} X^*_i - X(i) + \mu \). Note that \( X^*_i - X(i) \) is symmetric about 0. It follows that \( Y(i) - X(i) \) is symmetric about \( \mu \). \( \Box \)

Under the condition of the above lemma, we can write \( p_i \) as \( p_i = P(X^*_i - X(i) > -\mu) \). Therefore, we have the following corollary.

**Corollary** If \( X \sim F(x) \), \( Y \sim F(x - \mu) \) and \( X \) and \( Y \) are independent, then \( p_i \geq \frac{1}{2} \) according as \( \mu \geq 0 \).

**Lemma 2.4.3** If \( X \sim F(x) \) and \( Y \sim F(x - \mu) \), where \( X \) is symmetric and \( X \) and \( Y \) are independent, then \( Y(i) - X(j) \) and \( Y(n-j+1) - X(n-i+1) \) have the same distribution.
Proof. For simplicity, assume $X$ is symmetric about 0. Since $X_{(j)} \overset{d}{=} -X_{(n-j+1)}$, $Y_{(i)} \overset{d}{=} 2\mu - Y_{(n-i+1)}$, and $X$ and $Y$ are independent, we have $Y_{(i)} - X_{(j)} \overset{d}{=} 2\mu - Y_{(n-i+1)} + X_{(n-j+1)} \overset{d}{=} (X_{(n-j+1)} + \mu) - (Y_{(n-i+1)} - \mu) \overset{d}{=} Y_{(n-j+1)} - X_{(n-i+1)}$. \qed

By the above lemma, we immediately have the following result.

Corollary Under the condition of Lemma 2.4.3, we have

$$P_{ij} = P_{n-j+1,n-i+1} \quad \text{for } 1 \leq i, j \leq n. \quad (2.12)$$

We can also show (2.12) directly from (2.4).

Lemma 2.4.4 If $X \sim F(x)$ and $Y \sim G(x)$, where $X$ and $Y$ are both symmetric rv's with the same point of symmetry, then

$$p_i = 1 - p_{n-i+1} \quad i = 1, \ldots, n.$$

Proof. WLG. assume $X$ is symmetric about 0. Then we have

$$X_{(i)} \overset{d}{=} -X_{(n-i+1)} \quad \text{and} \quad Y_{(i)} \overset{d}{=} -Y_{(n-i+1)}.$$

Therefore,

$$p_i = P(Y_{(i)} > X_{(i)})$$

$$= P(-Y_{(n-i+1)} > -X_{(n-i+1)})$$

$$= 1 - P(Y_{(n-i+1)} > X_{(n-i+1)})$$

$$= 1 - p_{n-i+1}. \quad \square$$

We now discuss some properties of the $p_i$'s under certain distribution assumptions. We first consider the case when $X$ is a symmetric rv with cdf $F(x)$ and pdf
\( f(x) \), and \( Y \overset{d}{=} X + \mu \), where \( X \) and \( Y \) are independent and \( \mu \geq 0 \). Without loss of generality, we assume \( f(x) \) is symmetric about 0.

For \( n = 2 \), by the corollary of Lemma 2.4.3 and Theorem 2.2.2 (b), we have

\[
p_i \geq p = P(Y > X).
\]

(2.13)

Actually, \( p_i - p \) can be expressed as follows

\[
p_i - p = \int_0^{+\infty} \left\{ \left[ F(x - \mu) - F^2(x - \mu) \right] - \left[ F(x + \mu) - F^2(x + \mu) \right] \right\} [2F(x) - 1]dF(x).
\]

By discussing the properties of \( h(t) = t - t^2 \) on \([0, 1]\), we can show (2.13) also.

The question raised here is: Does (2.13) hold for all \( n \)? We first look at some examples.

Example 2.4.1: If \( f(x) = \text{unif}[-1, 1] \), then for \( n = 3 \), (2.13) holds. We can see this as follows. Since

\[
f(x) = \begin{cases} \frac{1}{2} & \text{if } x \in [-1, 1] \\
0 & \text{otherwise}
\end{cases}
\]

and

\[
F(x) = \begin{cases} 0 & x < -1 \\
\frac{1}{2} (\mu + 1) & x \in [-1, 1] \\
1 & x > 1,
\end{cases}
\]

we have \( p_1 = p_2 = p_3 = p = 1 \) for \( \mu \geq 2 \), i.e., (2.13) holds for \( \mu \geq 2 \). For \( 0 \leq \mu \leq 2 \), by the corollary of Lemma 2.4.3, we have \( p_1 = p_3 \). Therefore, by straightforward computations, it follows that

\[
p_1 - p = p_3 - p = \frac{\mu}{288} (\mu - 2)^2 (\mu^3 - 4\mu^2 - 12\mu + 128) \geq 0
\]
and

\[ p_2 - p = \frac{\mu(\mu - 2)^2}{2^6 5} (-\mu^3 - 4\mu^2 - 18\mu + 8) \geq 0. \]

Note that

\[ p_1 - p_2 = p_3 - p_2 = (p_3 - p) - (p_2 - p) \]
\[ = \frac{\mu(\mu - 2)^2}{2^8 5} (3\mu^3 + 12\mu^2 - 84\mu + 96) \geq 0. \]

We also have \( p_2 = \min\{p_i\} \).

Let \( X \sim Unif[-1, 1] \), \( Y \overset{d}{=} X + \mu \), \( U \sim Unif[0, 1] \), and \( V \overset{d}{=} U + \lambda \). By \( X \overset{d}{=} 2U - 1 \), we have \( P(Y > X) = P(V > U) \) and \( P(Y(i) > X(i)) = P(V(i) > U(i)) \) for \( \lambda = \frac{\mu}{2} \). Therefore, we can use the standard uniform distribution to compute the \( p_i \)'s.

Table 2.2 gives numerical results for \( n = 10 \) and \( \mu = 0.4 \). The results satisfy (2.13) and also suggests that

\[ p_{i+1} - p_i \overset{\succ}{=} 0 \text{ if } i \overset{<}{=} \frac{n}{2}. \]  

(2.14)

Table 2.2: Values of \( p_i \) for the uniform distribution

<table>
<thead>
<tr>
<th>index ( i )</th>
<th>value of ( p_i )</th>
<th>index ( i )</th>
<th>value of ( p_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.95259</td>
<td>6</td>
<td>0.83280</td>
</tr>
<tr>
<td>2</td>
<td>0.90125</td>
<td>7</td>
<td>0.84282</td>
</tr>
<tr>
<td>3</td>
<td>0.86463</td>
<td>8</td>
<td>0.86463</td>
</tr>
<tr>
<td>4</td>
<td>0.84282</td>
<td>9</td>
<td>0.90125</td>
</tr>
<tr>
<td>5</td>
<td>0.83280</td>
<td>10</td>
<td>0.95259</td>
</tr>
</tbody>
</table>

\[ \sum_{i=1}^{10} p_i = 8.78815 , \quad P(Y > X) = 0.68000 \]
For the uniform distribution case, (2.13) seems to be true for all $n$. Intuitively, we can explain this as follows:

Let us compare $P(Y(i) - X(i) > 0)$ and $P(Y - X > 0)$. Since $Var(U(i)) = \frac{i(n+1-i)}{(n+1)^2(n+2)}$, it follows that $Var(Y(i) - X(i)) = 2Var(X(i)) = \frac{8i(n+1-i)}{(n+1)^2(n+2)}$. On the other hand, $Var(Y - X) = \frac{2}{3}$. Since $Var(Y(i) - X(i))$ attains its maximum value at $i = \frac{n+1}{2}$ and at this point, $Var(Y(i) - X(i)) = \frac{2}{n+1}$ which is less than $Var(Y - X)$ for $n > 2$ and equals $Var(Y - X)$ for $n = 2$. Therefore, we have

$$Var(Y(i) - X(i)) < Var(Y - X) \text{ for } n > 2.$$ 

Since both $Y(i) - X(i)$ and $Y - X$ are symmetric about $\mu$, and have unimodal distributions, therefore, we may expect that

$$P(Y(i) - X(i) > 0) \geq P(Y - X > 0) \text{ for } \mu > 0,$$

i.e., (2.13) holds for $\mu > 0$.

In general, for any two rv's $Z_1$ and $Z_2$, $Var(Z_1) < Var(Z_2)$ and $Z_1$ and $Z_2$ are both symmetric about, say, $\mu > 0$, do not imply that $P(Z_1 > 0) \geq P(Z_2 > 0)$. However, if the pdf's of $Z_1$ and $Z_2$ are both unimodal, then it is more likely that $P(Z_1 > 0) \geq P(Z_2 > 0)$ is true.

Now let us look at some other numerical results for the symmetric case.

(a) Beta distribution with parameter $p = q$.

Let $X \sim \text{Beta}(x; p, q)$ and $Y \overset{d}{=} X + \mu$. For $n = 10$, we have the numerical results in Table 2.3.
Table 2.3: Values of $p_i$ for the Beta distribution with $p = q$

<table>
<thead>
<tr>
<th>index $i$</th>
<th>value of $p_i$</th>
<th>index $i$</th>
<th>value of $p_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.94226</td>
<td>1</td>
<td>0.95049</td>
</tr>
<tr>
<td>2</td>
<td>0.93020</td>
<td>2</td>
<td>0.95527</td>
</tr>
<tr>
<td>3</td>
<td>0.92481</td>
<td>3</td>
<td>0.95791</td>
</tr>
<tr>
<td>4</td>
<td>0.92201</td>
<td>4</td>
<td>0.95929</td>
</tr>
<tr>
<td>5</td>
<td>0.92078</td>
<td>5</td>
<td>0.95990</td>
</tr>
<tr>
<td>6</td>
<td>0.92078</td>
<td>6</td>
<td>0.95990</td>
</tr>
<tr>
<td>7</td>
<td>0.92201</td>
<td>7</td>
<td>0.95929</td>
</tr>
<tr>
<td>8</td>
<td>0.92481</td>
<td>8</td>
<td>0.95791</td>
</tr>
<tr>
<td>9</td>
<td>0.93020</td>
<td>9</td>
<td>0.95527</td>
</tr>
<tr>
<td>10</td>
<td>0.94226</td>
<td>10</td>
<td>0.95049</td>
</tr>
</tbody>
</table>

$\sum_{i=1}^{10} p_i = 9.28011$  
$P(Y > X) = 0.72639$

From Table 2.3, we can see that (2.13) is satisfied. However, (2.14) is not satisfied for the parameter $p = 3$. When $p = 3$, the $p_i$’s satisfy

$$p_{i+1} - p_i \geq 0 \quad \text{if} \quad i \leq \frac{n}{2}. \quad (2.15)$$

(b) Normal distribution.

Consider $X \sim \text{Normal}(0, 1)$ and $Y \sim X + \mu$. Table 2.4 gives the values of $p_i$’s for different $\mu$.

For the normal distribution case, it seems that (2.13) and (2.15) are always true. Moreover, it is possible that (2.13) holds for any symmetric distribution. However, this is not proved yet.

Now let us consider some asymmetric cases.
Table 2.4: Values of $p_i$ for the normal distribution

<table>
<thead>
<tr>
<th>$n = 10$</th>
<th>$\mu = 0.5$</th>
<th>$\mu = 1.0$</th>
<th>$\mu = 1.5$</th>
<th>$\mu = 2.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>index $i$</td>
<td>value of $p_i$</td>
<td>value of $p_i$</td>
<td>value of $p_i$</td>
<td>value of $p_i$</td>
</tr>
<tr>
<td>1</td>
<td>0.73042</td>
<td>0.88833</td>
<td>0.96431</td>
<td>0.99103</td>
</tr>
<tr>
<td>2</td>
<td>0.77879</td>
<td>0.93689</td>
<td>0.98861</td>
<td>0.99868</td>
</tr>
<tr>
<td>3</td>
<td>0.80164</td>
<td>0.95454</td>
<td>0.99422</td>
<td>0.99959</td>
</tr>
<tr>
<td>4</td>
<td>0.81353</td>
<td>0.96238</td>
<td>0.99611</td>
<td>0.99980</td>
</tr>
<tr>
<td>5</td>
<td>0.81875</td>
<td>0.96555</td>
<td>0.99677</td>
<td>0.99985</td>
</tr>
<tr>
<td>6</td>
<td>0.81875</td>
<td>0.96555</td>
<td>0.99677</td>
<td>0.99985</td>
</tr>
<tr>
<td>7</td>
<td>0.81353</td>
<td>0.96238</td>
<td>0.99611</td>
<td>0.99980</td>
</tr>
<tr>
<td>8</td>
<td>0.80164</td>
<td>0.95454</td>
<td>0.99422</td>
<td>0.99959</td>
</tr>
<tr>
<td>9</td>
<td>0.77879</td>
<td>0.93689</td>
<td>0.98861</td>
<td>0.99868</td>
</tr>
<tr>
<td>10</td>
<td>0.73042</td>
<td>0.88833</td>
<td>0.96431</td>
<td>0.99103</td>
</tr>
</tbody>
</table>

$\sum_{i=1}^{10} p_i = 7.88627$  
$P(Y > X) = 0.63816$  

(a) Lehmann alternative i.e., $G(x) = F_k(x)$. In this case, we have

$$p = P(Y > X) = \int_{-\infty}^{\infty} P(X < x)g(x)dx = \frac{k}{k+1},$$

$$p_i = 1 - C_{n,i}^{2} \int_{-\infty}^{\infty} \int_{0}^{F_k(x)} t^{i-1}(1-t)^{n-i}[F(x)]^{i-1}[1-F(x)]^{n-i}f(x)dtdF(x)$$

$$= 1 - C_{n,i}^{2} \int_{0}^{1} \int_{0}^{s^{k}} t^{i-1}(1-t)^{n-i}s^{i-1}(1-s)^{n-i}dtds.$$  

It is easy to see that for fixed $n$, $p_i$ is an increasing function of $k$. Also we can show that

$$pn = \frac{k}{k+1} = p \text{ for all } n.$$  

In general, since

$$(1-t)^{n-i} = \sum_{m=0}^{n-i} \binom{n-i}{m} (-1)^{m}t^{m},$$
we can write \( p_i \) as follows

\[
p_i = 1 - i^2 \binom{n}{i}^2 \sum_{m=0}^{n-i} \frac{(-1)^m \binom{n-i}{m}}{(m+i)[k(m+i)+i]} \left( \frac{n+k(m+i)}{n-i} \right)^{-1}.
\]

In extensive numerical work, we found

\[ p_i > p \quad \text{for} \quad i = 1, \ldots, n - 1. \]

Also, there exists an \( i_o \) such that \( p_{i+1} - p_i > 0 \) for \( i < i_o \) and \( p_{i+1} - p_i < 0 \) for \( i > i_o \).

(b) Gamma distribution.

For \( X \sim \text{Gamma}(x; \alpha) \) and \( Y \overset{d}{=} X + \mu \), we have the values of \( p_i \) in Table 2.5.

<table>
<thead>
<tr>
<th>( n = 10, \mu = 0.5, \alpha = 3.0 )</th>
<th>( n = 10, \mu = 2.0, \alpha = 3.0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( i )</td>
<td>value of ( p_i )</td>
</tr>
<tr>
<td>1</td>
<td>0.79602</td>
</tr>
<tr>
<td>2</td>
<td>0.77698</td>
</tr>
<tr>
<td>3</td>
<td>0.75989</td>
</tr>
<tr>
<td>4</td>
<td>0.74263</td>
</tr>
<tr>
<td>5</td>
<td>0.72466</td>
</tr>
<tr>
<td>6</td>
<td>0.70552</td>
</tr>
<tr>
<td>7</td>
<td>0.68452</td>
</tr>
<tr>
<td>8</td>
<td>0.66058</td>
</tr>
<tr>
<td>9</td>
<td>0.63151</td>
</tr>
<tr>
<td>10</td>
<td>0.59137</td>
</tr>
</tbody>
</table>

| \( \sum_{i=1}^{10} p_i = 7.07368 \) | \( \sum_{i=1}^{10} p_i = 9.59490 \) |
| \( P(Y > X) = 0.59249 \) | \( P(Y > X) = 0.81391 \) |

From Table 2.5, we can see that for \( \alpha = 0.5 \), we have \( p_{10} < p \). As a matter of fact, for the exponential distribution case \( (\alpha = 1) \), \( p_n < p \) for all \( n \). We will show
this later. This result tells us that for asymmetric distributions, (2.13) does not necessarily hold. For the Gamma distribution, it seems that

\[ p_{i+1} - p_i < 0 \quad \text{for } i = 1, \ldots, n - 1. \tag{2.16} \]

(c) Beta distribution with parameter \( p \neq q \).

We consider \( X \sim \text{Beta}(x; p, q) \) and \( Y \overset{d}{=} X + \mu \). In this case, numerical results suggest that (2.16) holds for \( p < q \). It is easy to show that if we have \( X' \sim \text{Beta}(x; q, p) \) and \( Y' \overset{d}{=} X' + \mu \), then

\[ p_i = p_{n-i+1} = P(Y_{n-i+1} > X'_{n-i+1}) \quad \text{for } i = 1, \ldots, n. \]

In the above, we discussed the properties of \( p_i \). From the numerical results, we can see that \( p_{i+1} - p_i \) has some trends. The following lemma gives a simple expression for \( p_{i+1} - p_i \).

**Lemma 2.4.5** If \( X \) and \( Y \) are independent rv's with cdf's \( F(x) \) and \( G(x) \), and pdf's \( f(x) \) and \( g(x) \), respectively, then

\[
\begin{align*}
p_{i+1} - p_i &= \binom{n}{i}^2 \int_{-\infty}^{\infty} G^i(x)[1 - F(x)]^{n-i} d\{F^i(x)[1 - G(x)]^{n-i}\} \\
&= P(Y(i) < X(i+1) < Y(i+1)) - P(X(i) < Y(i+1) < X(i+1)).
\end{align*}
\tag{2.17}
\]

**Proof.** We have

\[
\begin{align*}
p_i &= P(Y(i) > X(i)) = P(\text{at least } i \text{ X's } < Y(i)) \\
&= P(\text{exactly } i \text{ X's } < Y(i)) + P(\text{at least } (i + 1) \text{ X's } < Y(i)) \\
&= P(\text{exactly } i \text{ X's } < Y(i)) + P(Y(i) > X(i+1)).
\end{align*}
\]
Now

\[ p_{i+1} - P(X_{(i+1)} < Y(i)) = P(\text{at most } i \ Y's < X_{(i+1)}) - P(\text{at most } (i - 1) \ Y's < X_{(i+1)}) \]

\[ = P(\text{exactly } i \ Y's < X_{(i+1)}) \]

\[ = \int_{-\infty}^{\infty} P(\text{exactly } i \ Y's < x) f_{i+1}(x) dx \]

\[ = \int_{-\infty}^{\infty} \binom{n}{i} G^i(x)[1 - G(x)]^{n-i} f_{i+1}(x) dx. \]

Also

\[ P(\text{exactly } i \ X's < Y(i)) = \int_{-\infty}^{\infty} \binom{n}{i} F^i(x)[1 - F(x)]^{n-i} g(x) dx. \]

So, by the above and some rearrangement, we have

\[ p_{i+1} - p_i = P(\text{exactly } i \ Y's < X_{(i+1)}) - P(\text{exactly } i \ X's < Y(i)) \]

\[ = \binom{n}{i}^2 \int_{-\infty}^{\infty} G^i(x)[1 - F(x)]^{n-i} d\{F^i(x)[1 - G(x)]^{n-i}\}. \]

Noting that

\[ P(\text{exactly } i \ Y's < X_{(i+1)}) = P(Y(i) < X_{(i+1)} < Y(i+1)) \]

and

\[ P(\text{exactly } i \ X's < Y(i)) = P(X(i) < Y(i+1) < X(i+1)), \]

we obtain the second stated result. \( \Box \)
2.5 Some Asymptotic Results

In the last section, we discussed some properties of \( p'_{ij} \)'s. Among these \( p'_{ij} \)'s, \( p_1 = P(Y_1 > X_1) \) and \( p_n = P(Y_n > X_n) \) have some special interest. In this section, we will investigate the limit behaviors of \( p_1 \) and \( p_n \) as well as \( p_{n_i}, n_j \) for \( n_i = [n\alpha] + 1 \), and \( n_j = [n\beta] + 1 \), where \( 0 < \alpha, \beta < 1 \). We begin with known results.

Lemma 2.5.1 If there exist \( a_n > 0 \) and \( b_n \) such that

\[
a_n X(n) + b_n \xrightarrow{d} Z,
\]

then the cdf of \( Z \) must be of one of the following three types:

\[
\begin{align*}
\Lambda_1(x) &= 0 & x \leq 0 \\
&= \exp(-x^{-\alpha}) & x > 0, \alpha > 0, \\
\Lambda_2(x) &= \exp[-(-x)^\alpha] & x \leq 0, \alpha > 0 \\
&= 1 & x > 0, \\
\Lambda_3(x) &= \exp(-e^{-x}) & -\infty < x < \infty.
\end{align*}
\]

Lemma 2.5.2 If the distribution of \((U_n, V_n)\) converges weakly to \( T(u)H(v) \), where \( T(u) \) and \( H(v) \) are continuous distribution functions, then

\[
\lim_{n \to \infty} P(U_n \pm V_n < \mu) = \int_{-\infty}^{\infty} H(\mu \mp u)dT(u). \tag{2.19}
\]

The proofs of the above two lemmas can be found in Galambos 1987, p76 and p130.

Based on Lemma 2.5.1 and 2.5.2, we have the following result:
Lemma 2.5.3 Let \((X_1, \ldots, X_n)\) and \((Y_1, \ldots, Y_n)\) be two independent random samples with population cdf's \(F(x)\) and \(F(x - \mu)\), respectively, where \(\mu > 0\). Assume there exist \(a_n > 0\) and \(b_n\) such that

\[
an_nX_{(n)} + b_n \overset{\text{d}}{\rightarrow} Z \sim \Lambda(x), \quad \text{and} \quad \lim_{n \to \infty} a_n = c.
\]

Then

\[
\lim_{n \to \infty} P(Y_{(n)} > X_{(n)}) = \int_{-\infty}^{\infty} \Lambda(c\mu + x) d\Lambda(x),
\]

where \(\Lambda(x)\) can only be one of \(\Lambda_1(x), \Lambda_2(x)\) and \(\Lambda_3(x)\) given in Lemma 2.5.1.

Proof. Since \(Y_{(n)} \overset{\text{d}}{=} X^*_{(n)} + \mu\), where \(X^*_{(n)}\) and \(X_{(n)}\) are iid, we have

\[
an_nX_{(n)} + b_n \overset{\text{d}}{=} a_nX^*_{(n)} + b_n \overset{\text{d}}{=} Z.
\]

Also by \(a_n > 0\), it follows that

\[
P(Y_{(n)} > X_{(n)}) = P((a_nX_{(n)} + b_n) - (a_nX^*_{(n)} + b_n) < a_n\mu).
\]

Now, if \(c < \infty\), then for any \(\varepsilon > 0\), there exists \(N > 0\), such that when \(n > N\), we have \(c - \varepsilon < a_n < c + \varepsilon\). Therefore, it follows that, for \(n > N\)

\[
P((a_nX_{(n)} + b_n) - (a_nX^*_{(n)} + b_n) < (c - \varepsilon)\mu)
\]

\[
\leq P(Y_{(n)} > X_{(n)})
\]

\[
\leq P((a_nX_{(n)} + b_n) - (a_nX^*_{(n)} + b_n) \leq (c + \varepsilon)\mu).
\]

Let \(n \to \infty\), by Lemma 2.5.2, we have

\[
\int_{-\infty}^{\infty} \Lambda[(c - \varepsilon)\mu + x] d\Lambda(x) \leq \lim_{n \to \infty} P(Y_{(n)} > X_{(n)})
\]

\[
\leq \int_{-\infty}^{\infty} \Lambda[(c + \varepsilon)\mu + x] d\Lambda(x).
\]
Let $\varepsilon \to 0$. Since $\Lambda(x)$ is a distribution function, our stated result follows immediately by the monotone convergence theorem. If $c = +\infty$, then for any $m > 0$, there exists $N > 0$ such that when $n > N$, we have $a_n > m$. Therefore,

$$1 \geq P(Y(n) > X(n)) \geq P((a_n X(n) + b_n) - (a_n X^*_n + b_n) < m).$$

Again, by Lemma 2.5.2 and the monotone convergence theorem, we can show that

$$\lim_{n \to \infty} P(Y(n) > X(n)) = 1.$$

This completes the proof. $\Box$

Now let us consider the limit of $p_{n_i,n_j}$. Let $n_k = [n\lambda] + 1$, where $0 < \lambda < 1$, and $\xi_\lambda$ be population quantile corresponding to $X(n_k)$. It is well known that if $0 < f(\xi_\lambda) < \infty$, then

$$\sqrt{n}(X(n_k) - \xi_\lambda) \xrightarrow{d} N(0, \frac{\lambda(1-\lambda)}{f^2(\xi_\lambda)})$$

(see David, 1981, p255). Therefore, we have the following result:

**Lemma 2.5.4** Let $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ be two independent random samples with population cdf's $F(x)$ and $F(x - \mu)$, respectively. Assume $f(x) = F'(x)$ exists. Let $n_i = [n\alpha] + 1$ and $n_j = [n\beta] + 1$, where $0 < \alpha < 1$ and $0 < \beta < 1$. Then, if $0 < f(\xi_\alpha) < \infty$ and $0 < f(\xi_\beta) < \infty$, we have

$$\lim_{n \to \infty} P(Y(n_i) > X(n_j)) = \begin{cases} 
0 & \text{if } \xi_\alpha - \xi_\beta + \mu < 0 \\
\frac{1}{2} & \text{if } \xi_\alpha - \xi_\beta + \mu = 0 \\
1 & \text{if } \xi_\alpha - \xi_\beta + \mu > 0,
\end{cases}$$

**Proof.** By (2.20) and Lemma 2.5.2, using the same proof of Lemma 2.5.3 for $a_n = \sqrt{n}$, our stated result follows. $\Box$
It is easily shown that when \( F(x) = 1 - e^{-x} \), \( X(n) = -\log n \) has the limit distribution \( \Lambda_3 = e^{-e^{-x}} \) (e.g., David 1981, p263). Therefore, if \( G(x) = F(x - \mu) \), we have

\[
\lim_{m \to \infty} P(Y(n) > X(n)) = \lim_{m \to \infty} P((X(n) - \log n) - (X^*_n - \log n) < \mu)
\]
\[
= \int_{-\infty}^{\infty} e^{-(x+\mu)} e^{-x} dx
\]
\[
= \int_{-\infty}^{\infty} (e^{-x}) e^{-\mu} e^{-x} dx
\]
\[
= \frac{1}{1 + e^{-\mu}}.
\]

The above result tells us if \( X(n) \) and \( X^*_n \) are iid, the largest order statistics from an exponential distribution, then

\[
P(X(n) - X^*_n \leq x) \to \frac{1}{1 + e^{-x}},
\]

i.e., \( X(n) - X^*_n \) converges in distribution to a standard logistic distribution.

Now we consider the exponential case for finite \( n \). Note that

\[
p_n = P(Y(n) > X(n))
\]
\[
= n \int_{-\infty}^{\infty} F^n(x)G^{n-1}(x)g(x)dx
\]
\[
= n \int_{0}^{\infty} (1 - e^{-x-\mu})^n (1 - e^{-x})^{n-1} e^{-x} dx
\]
\[
= n \int_{0}^{1} (1 - ct)^n (1 - t)^{n-1} dt,
\]

where \( c = e^{-\mu} \). Letting \( s = 1 - t \), and writing \( c_1 = 1 - c \), we have

\[
p_n = n \int_{0}^{1} (c_1 + cs)^n s^{n-1} ds
\]
\[ f(n) = \sum_{k=0}^{n} (-1)^k \frac{(n!)^2}{(n-k)!(n+k)!} e^k. \]

Therefore,

\[ p_n - p_{n-1} = \sum_{k=1}^{n-1} (-1)^k \frac{[(n-1)!!]^2}{(n-k)!(n+k)!} k^2 e^k + (-1)^n \frac{(n!)^2}{(2n)!} e^n. \]

From the above, we can see that when \( c \) is small, especially when \( c < \frac{1}{4} \) i.e., \( \mu > 2 \log 2 \), we have \( p_n - p_{n-1} \leq 0 \) for all \( \mu \geq 0 \). This possibly holds for all \( \mu \geq 0 \). However, a general proof has not been obtained yet.

Now let us compare \( p_n \) with \( p = P(Y > X) \). Note that

\[ p = \int_{-\infty}^{\infty} P(X < y)g(y)dy = \int_{\mu}^{\infty} (1 - e^{-y})e^\mu - y dy = 1 - \frac{1}{2} e^{-\mu}. \]
For \( n = 2 \), we have

\[
p_2 = 1 - \frac{2}{3} e^{-\mu} + \frac{1}{6} e^{-2\mu}.
\]

Therefore,

\[
p_2 - p = -\frac{1}{6} e^{-\mu} + \frac{1}{6} e^{-2\mu} < 0,
\]
i.e., \( p_2 < p \). Since \( p_n \) is monotone decreasing for large \( \mu \), we have \( p_n < p \) for large \( \mu \).

Let us consider the limit for the smallest order statistic. Here we temporarily use \( X_{1:n} \) to denote the smallest order statistic for sample size \( n \). Then

\[
p_1(n) = \Pr(Y_{1:n} > X_{1:n})
\]

\[
= \int_{-\infty}^{\infty} P(X_{1:n} > x - \mu) f_1(x) dx
\]

\[
= n \left( \int_0^\mu e^{-nx} dx + \int_{\mu}^\infty e^{-n(x-\mu)} e^{-nx} dx \right)
\]

\[
= 1 - \frac{1}{2} e^{-n\mu}.
\]

Therefore, \( n \to \infty \), \( p_1(n) = 1 \).

We now find a recurrence relation for \( p_n \). We have

\[
p_n = n \int_0^1 (c_1 + ct)^n t^{n-1} dt
\]

\[
= \frac{n}{(n+1)c} \int_0^1 t^{n-1} d(c_1 + ct)^{n+1}
\]

\[
= \frac{n}{(n+1)c} \left[ t^{n-1} (c_1 + ct)^{n+1} \right]_0^1 - (n-1) \int_0^1 t^{n-2} (c_1 + ct)^{n+1} dt
\]

\[
= \frac{n}{(n+1)c} [1 - (n-1) \int_0^1 t^{n-2} (c_1 + ct)^{n-1} (c_1^2 + 2c_1 ct + c^2 t^2) dt]
\]

\[
= \frac{n}{(n+1)c} [1 - c_1^2 p_{n-1} - 2c_1 c(n-1) \int_0^1 t^{n-1} (c_1 + ct)^{n-1} dt -
\]
\[-(n - 1)c^2 \int_0^1 t^n(c_1 + ct)^n \, dt.\]

Also

\[
\int_0^1 t^{n-1}(c_1 + ct)^{n-1} \, dt = \frac{1}{nc} \left( \int_0^1 t^{n-1} \, dt \right) \int_0^1 t^{n-1}(c_1 + ct)^n \, dt
\]

\[
= \frac{1}{nc} \left[ 1 - (n - 1) \int_0^1 t^{n-2}(c_1 + ct)^n \, dt \right]
\]

\[
= \frac{1}{nc} \left[ 1 - (n - 1) \int_0^1 t^{n-2}(c_1 + ct)^{n-1} \, dt \right]
\]

\[
= \frac{1}{nc} \left[ 1 - c_1p_{n-1} - (n - 1)c \int_0^1 t^{n-1}(c_1 + ct)^{n-1} \, dt \right].
\]

Writing

\[x = \int_0^1 t^{n-1}(c_1 + ct)^{n-1} \, dt,
\]

we have the equation

\[x = \frac{1}{nc} \left[ 1 - c_1p_{n-1} - (n - 1)c x \right].\]

Therefore,

\[x = \frac{1 - c_1p_{n-1}}{(2n - 1)c}.\]

Also, since

\[
\int_0^1 t^n(c_1 + ct)^n \, dt = \frac{1}{nc} \left( \int_0^1 \, dt \right) \int_0^1 t^n d(c_1 + ct)^n
\]

\[
= \frac{1}{nc} \left[ 1 - n \int_0^1 t^{n-1}(c_1 + ct)^{n} \, dt \right] = \frac{1}{nc} (1 - pn),
\]
it follows that

\[ p_n = \frac{n}{(n+1)c} \left[ 1 - c_1^2 p_{n-1} - \frac{2(n-1)c_1}{2n-1} (1 - c_1 p_{n-1}) - \frac{(n-1)c}{n} (1 - pn) \right], \]

i.e.,

\[ 2cp_n = \frac{n + (n-1)c}{2n-1} - \frac{nc_1^2}{2n-1} p_{n-1} \]

or

\[ p_n = \frac{n + (n-1)c}{2c(2n-1)} - \frac{nc_1^2}{2c(2n-1)} p_{n-1}. \]

Let \( n \to \infty \) and write \( L = \lim_{n \to \infty} p_n = \lim_{n \to \infty} p_{n-1} \). Then

\[ \frac{(1 + c)^2}{2} L = \frac{1 + c}{2} \quad \text{i.e.,} \quad L = \frac{1}{1 + c} = \frac{1}{1 + e^{-\mu}}. \]

Again, we get the same result as in Lemma 2.5.3.

For the exponential distribution case, we can interpret \( p_1(n) = P(Y_1;n > X_1;n) \)

and \( p_n \) as follows:

In a life test, suppose we have two identical systems. Each of them has \( n \) iid

components. System 1 begins to work first, and after time \( \mu \) system 2 begins to work.

Then \( p_1(n) \) and \( p_n \) are the probabilities that system 2 lasts longer than system 1

corresponding to series and parallel systems respectively. When \( n \) is large, we have

\( p_1(n) \approx 1 \) and \( p_n \approx \frac{1}{1 + e^{-\mu}}. \)

Let \( F(x) = \frac{1}{1 + e^{-kx}} \) \((k > 0)\), i.e., the logistic distribution , and \( G(x) = F(x-\mu) \),

where \( \mu > 0 \). Then writing \( a = e^{-k\mu} \) and \( a' = 1 - a \), we have

\[ p_n = -n \int_{-\infty}^{\infty} (1 + e^{-k(x+\mu)})^{-n}(1 + e^{-kx})-(n+1)de^{-kx} \]

\[ = n \int_{0}^{\infty} (1 + at)^{-n}(1 + t)-(n+1)dt \]
\[
\begin{align*}
= n \int_1^\infty (a' + at)^{-n} t^{-(n+1)} dt \\
= -\frac{n}{(n-1)a} \int_1^\infty t^{-(n+1)} d(a' + at)^{-(n-1)} \\
= \frac{n}{(n-1)a} [1 - (n+1) \int_1^\infty t^{-(n+2)} (a' + at)(a' + at)^{-n} dt] \\
= \frac{n}{(n-1)a} [1 - (n+1)a' \int_1^\infty t^{-(n+2)} (a' + at)^{-n} dt - \frac{n+1}{n} cpn].
\end{align*}
\]

Since
\[
\int_1^\infty t^{-(n+2)} (a' + at)^{-n} dt \\
= \int_1^\infty t^{-(n+2)} (a' + at)^{-(n+1)} (a' + at) dt \\
= a' \int_1^\infty t^{-(n+2)} (a' + at)^{-(n+1)} dt \\
+ a \int_1^\infty t^{-(n+1)} (a' + at)^{-n} dt \\
= \frac{a'}{n+1} pn + \frac{1}{n} \int_1^\infty t^{-(n+1)} d(a' + at)^{-n} \\
= \frac{a'}{n+1} pn + \frac{1}{n} + \frac{n+1}{n} \int_1^\infty t^{-(n+2)} (a' + at)^{-n} dt,
\]

we have
\[
\int_1^\infty t^{-(n+2)} (a' + at)^{-n} dt = \frac{2}{2n+1} \left( \frac{a'}{n+1} pn + 1 \right).
\]

Hence,
\[
pn = \frac{n}{(n+1)a} [1 - \frac{n(n+1)a'}{2n+1} \left( \frac{a'}{n+1} pn + 1 \right) - \frac{(n+1)a}{n} pn],
\]
\[ p_n = \frac{1}{2a} \left( 1 - \frac{na^2}{2n + 1} p_{n+1} - \frac{n + 1}{2n + 1} a \right). \]

Let \( n \to \infty \) and write \( L = \lim_{n \to \infty} p_n \). Then

\[ L = \frac{1}{2a} \left( 1 - \frac{1}{2} a^2 L - \frac{1}{2} \right) \]

so that

\[ L = \frac{1}{1 + a} = \frac{1}{1 + e^{-k\mu}}. \]

If we choose \( F(x) = e^{-e^{-x}} \), i.e., the extreme value distribution, and \( G(x) = F(x - \mu) \), where \( \mu > 0 \), then it is well known that \( p_n = \frac{1}{1 + e^{-\mu}} \) for all \( n \).

If \( F(x) \) is the standard normal distribution and \( G(x) = F(x - \mu) \), then it can be shown that \( a_n = \left( \frac{1}{2} \log n \right)^2 \) (e.g., David 1981, p. 264). Therefore, by Lemma 2.5.3, we have \( \lim_{n \to \infty} p_n = 1 \).

### 2.6 Matching with Ties Permitted for the Order Statistics Model

As we mentioned before, in practice it is possible that some comparison ends in a tie. A tie usually is caused when the performances of the two objects are too close to tell the difference. In this case, in order to compare \( \Gamma_X \) with \( \Gamma_Y \), we introduce the following indicator function

\[
I(\mu, \nu; \tau) = \begin{cases} 
0 & \text{if } \mu - \nu < -\tau \\
\frac{1}{2} & \text{if } |\mu - \nu| \leq \tau \\
1 & \text{if } \mu - \nu > \tau,
\end{cases}
\]
where \( \tau \) is called a threshold parameter (Glenn and David, 1960). We assume \( \tau \geq 0 \).

Then for any permutation \( \pi = (\pi_1, \ldots, \pi_n) \), we define

\[
S_{\tau}(\pi) = \sum_{i=1}^{n} I(Y_{(i)}, X_{(\pi_i)}; \tau)
\]

which measures the performance of \( \Gamma_Y \) with respect to \( \Gamma_X \) under the matching \( \pi \) with ties permitted. Again, we assume that \( \Gamma_X \) and \( \Gamma_Y \) are independent and the assumptions on \( \Gamma_X \) and \( \Gamma_Y \) are the same as before. Therefore, the expectation of \( S_{\tau}(\pi) \) is given by

\[
ES_{\tau}(\pi) = \sum_{i=1}^{n} P(Y_{(i)} > X_{(\pi_i)} + \tau) + \frac{1}{2} \sum_{i=1}^{n} P(|Y_{(i)} - X_{(\pi_i)}| \leq \tau)
\]

\[
= \frac{1}{2} \left( \sum_{i=1}^{n} P(Y_{(i)} > X_{(\pi_i)} + \tau) + \sum_{i=1}^{n} P(Y_{(i)} > X_{(\pi_i)} - \tau) \right). \quad (2.21)
\]

For \( \pi^o = (1, 2, \ldots, n) \), we write

\[
V_1^T = ES_{\tau}(\pi^o)
\]

which is \( ES_{\tau}(\pi) \) under ordered matching. Let \( V_2^T \) be the expectation of \( S_{\tau}(\pi) \) under random matching, then using the same arguments as before, we have

\[
V_2^T = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} E[I(Y_{(i)}, X_{(j)}; \tau)]
\]

\[
= \frac{1}{n} \left[ \sum_{i=1}^{n} \sum_{j=1}^{n} P(Y_{(i)} > X_{(j)} + \tau) \right]

\[
+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} P(|Y_{(i)} - X_{(j)}| \leq \tau)
\]

\[
= \frac{n}{2} [P(Y > X + \tau) + P(Y > X - \tau)],
\]
where $X$ and $Y$ are independent with cdf's $F(x)$ and $F(x - \mu)$ respectively.

The questions concerning us here are how $V_1^T$ compare to $V_2^T$ and how $V_1^T$ compares to $V_1$.

**Lemma 2.6.1** Let $X \sim F(x)$, where $F(x)$ is an absolutely continuous unimodal distribution function. If $Y \overset{d}{=} X + \mu$ with $\mu \geq 0$, and $X$ and $Y$ are independent, then for any $\tau \geq 0$, we have

$$P(Y > X) \geq \frac{1}{2}[P(Y > X + \tau) + P(Y > X - \tau)]. \quad (2.22)$$

**Proof.** Let $U(x)$ be the cdf of $X_1 - X_2$, where $X_1$ and $X_2$ are iid with cdf $F(x)$. Then $X_1 - X_2$ is symmetric about 0 and $U(x)$ is a unimodal distribution function also (Dharmadhikari and Joag-dev, 1988,p15). Now we can write (2.22) as

$$U(\mu) \geq \frac{1}{2}[U(\mu + \tau) + U(\mu - \tau)]. \quad (2.23)$$

If $\tau \leq \mu$, then since $U(x)$ is concave on $[0, +\infty)$, (2.23) follows immediately, by the definition of a concave function. Now assume $\tau > \mu$. Consider the function

$$H(x) = \begin{cases} \frac{U(\mu - \tau) - U(0)}{\mu - \tau}x + U(0) & x \leq 0 \\ \frac{U(\mu + \tau) - U(0)}{\mu + \tau}x + U(0) & x > 0. \end{cases}$$

Then based on the fact that $U(x)$ is unimodal and $U'(x)$ is symmetric about 0, it follows that $H(x)$ is a concave function. Therefore, we have

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

Since

$$H(\mu) = \frac{\tau}{\mu + \tau}U(0) + \frac{\mu}{\mu + \tau}U(\mu + \tau)$$
and $U(x)$ is concave on $[0, \infty)$, we have

$$H(\mu) \leq U\left[\frac{\tau}{\mu + \tau} \cdot 0 + \frac{\mu}{\mu + \tau}(\mu + \tau)\right] = U(\mu). \quad \square$$

**Corollary** Under the conditions of Lemma 2.6.1 and if $X_{(i)}$ has a unimodal distribution ($i = 1, 2, \ldots, n$), then

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

The proof of above corollary is directly from Lemma 2.6.1. By this corollary, we immediately have $V_1 \geq V_1^T$.

In the corollary, we need $X_{(i)}$ to be unimodal. Alam (1972) shows that if the density function $f(x)$ of $X$ satisfies the condition that $\frac{1}{f(x)}$ is convex, then its order statistics are unimodal. This condition is satisfied by Normal, Gamma, Cauchy, Laplace, Logistic, Uniform, etc.

**Theorem 2.6.1** Let $X$ and $Y$ be independent absolutely continuous rv's with cdf $F(x)$ and $F(x - \mu)$, respectively, where $\mu \geq 0$.

(a) If $0 \leq \tau \leq \mu$, then

$$V_1^T \geq V_2^T$$

and

$$V_1^T \geq ES_\tau(\pi) \quad (2.24)$$

for any simple permutation $\pi$. If $X$ is a symmetric rv, then (2.24) holds for any symmetric permutation also.

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

$$ES_\tau(\pi) = \frac{n}{2} \quad (2.25)$$
for any simple permutation $\pi$. If $X$ is a symmetric rv, then (2.25) holds for any symmetric permutation $\pi$ also.

Proof. (a) Note that

$$V_1^\tau = \frac{1}{2} \left[ \sum_{i=1}^{n} P(Y(i) - \tau > X(i)) + \sum_{i=1}^{n} P(Y(i) + \tau > X(i)) \right]$$

and $Y(i) - \tau$ $(i = 1, 2, \ldots, n)$ are the order statistic from the population with cdf $F[x - (\mu - \tau)]$, where $\mu - \tau \geq 0$. Then by Theorem 2.2.2. (b), we have

$$\sum_{i=1}^{n} P(Y(i) - \tau > X(i)) \geq n P(Y > X + \tau)$$

and

$$\sum_{i=1}^{n} P(Y(i) + \tau > X(i)) \geq n P(Y > X - \tau).$$

Therefore, we have $V_1^\tau \geq V_2^\tau$. Using a similar argument and Theorem 2.3.1, we can show the remainings of (a).

(b) If $\mu = 0$, by the definition of simple permutation and the fact that

$$P(Y(i) > X(\pi_i) + \tau) + P(Y(i) > X(\pi_i) - \tau) = 1,$$

it follows that (2.25) holds for any simple permutation $\pi$. If $X$ is symmetric, for simplicity assume $X$ is symmetric about 0, then for any symmetric permutation $\pi = (\pi_1, \ldots, \pi_n)$, we have $Y(i) = -Y(n-i+1)$ and $X(\pi_i) = -X(\pi_{n-i+1}) = X(\pi_{n-i+1})$. Therefore,

$$\sum_{i=1}^{n} P(Y(i) > X(\pi_i) + \tau) = \sum_{i=1}^{n} P(-Y(n-i+1) > X(\pi_{n-i+1}) + \tau) = \sum_{i=1}^{n} P(Y(i) < X(\pi_i) - \tau).$$
Hence (2.25) follows immediately. □

In Theorem 2.6.1, we have shown that $V_1^T \geq V_2^T$ for $0 \leq \tau \leq \mu$. However, if $\tau > \mu$, then $V_1^T \geq V_2^T$ does not necessarily hold.

Now, write

$$P_i^T = EI(Y_{(i)}, X_{(i)}; \tau) = \frac{1}{2}[P(Y_{(i)} > X_{(i)} + \tau) + P(Y_{(i)} > X_{(i)} - \tau)].$$

Then Table 2.6 and 2.7 give the numerical values of $P_i^T$ for different $\tau$ and $\mu$ with $F(x) = \Phi(x)$, i.e., the standard normal distribution.

**Table 2.6: Values of $P_i^T$ for the normal distribution**

<table>
<thead>
<tr>
<th>$\mu = 0.6$, $\tau = 0.9$</th>
<th>$\mu = 0.8$, $\tau = 0.9$</th>
<th>$\mu = 1.0$, $\tau = 0.9$</th>
</tr>
</thead>
<tbody>
<tr>
<td>index $i$</td>
<td>value of $P_i^T$</td>
<td>index $i$</td>
</tr>
<tr>
<td>1</td>
<td>0.66016</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.65544</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0.64983</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>0.64626</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0.64456</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>0.64456</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>0.64626</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>0.64983</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>0.65544</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>0.66016</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$V_1^T = 6.51252$</th>
<th>$V_1^T = 7.14881$</th>
<th>$V_1^T = 7.80500$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_2^T = 6.35790$</td>
<td>$V_2^T = 6.78574$</td>
<td>$V_2^T = 7.19316$</td>
</tr>
</tbody>
</table>

From Table 2.7, we can see that $V_1^T < V_2^T$ for $\mu = 0.6$ and $\tau = 1.0$. We can also see that

$$P_i^T = P_{n-i+1}^T \quad \text{for} \quad i = 1, 2, \ldots, n. \quad (2.26)$$
Table 2.7: Values of $P_i^T$ for the normal distribution

<table>
<thead>
<tr>
<th>$\mu = 0.6$, $\tau = 1.0$</th>
<th>$\mu = 0.8$, $\tau = 1.0$</th>
<th>$\mu = 1.0$, $\tau = 1.0$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>index $i$</strong></td>
<td><strong>value of $P_i^T$</strong></td>
<td><strong>index $i$</strong></td>
</tr>
<tr>
<td>1</td>
<td>0.64192</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.63083</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0.62263</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>0.61784</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0.61564</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>0.61564</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>0.61784</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>0.62263</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>0.63083</td>
<td>9</td>
</tr>
<tr>
<td>10</td>
<td>0.64193</td>
<td>10</td>
</tr>
</tbody>
</table>

$V_1^T = 6.25772$, $V_1^T = 6.84579$, $V_1^T = 7.48895$

$V_2^T = 6.29850$, $V_2^T = 6.71111$, $V_2^T = 7.10675$

In general, if $F(x)$ is the cdf of a symmetric rv, then we can show that (2.26) holds.
3. LINEAR PREFERENCE MODEL

3.1 Introduction

Again, let $\Gamma_X = (X^i_{(1)}, X^i_{(2)}, \ldots, X^i_{(n)})$ and $\Gamma_Y = (Y^i_{(1)}, Y^i_{(2)}, \ldots, Y^i_{(n)})$ be two groups of stochastically ordered random variables which may represent the increasing "strengths" of the members of two chess teams or two tennis teams, etc. Here we assume that $X^i_{(i)} \sim F(x - \lambda(i))$ and $Y^i_{(i)} \sim F(x - \mu(i))$, $i = 1, \ldots, n$, where $F(x)$ is a distribution function, $\lambda(1) \leq \lambda(2) \leq \ldots \leq \lambda(n)$ and $\mu(1) \leq \mu(2) \leq \ldots \leq \mu(n)$ are ordered real numbers. We also assume that $\Gamma_X$ and $\Gamma_Y$ are independent. However, we do not need to assume independence within each group.

Let $X$ and $Y$ be iid with cdf $F(x)$. Then we have $X^i_{(i)} \overset{d}{=} X + \lambda(i)$ and $Y^i_{(i)} \overset{d}{=} Y + \mu(i)$. Also $X - Y$ has a symmetric distribution with mean zero. If $U(x)$ is the cdf of $X - Y$, we can write the preference probability of $Y^i_{(i)}$ over $X^i_{(i)}$ as

$$P(Y^i_{(i)} > X^i_{(i)}) = U(\mu(i) - \lambda(i)).$$  \hspace{1cm} (3.1)

This representation of the preference probability is based on the linear model much used in the method of paired comparisons (e.g., David, 1988, p.7). As indicated in Chapter 1, for any $\pi = (\pi_1, \ldots, \pi_n)$, or a matching $\pi$, $S(\pi) = \sum_{i=1}^{n} I(Y^i_{(i)} > X^i_{(\pi_i)})$ is the number of preferences of the $Y$'s in $\Gamma_Y$ over the $X$'s in $\Gamma_X$ under the matching.
The expectation of $S(\pi)$ can be written as

$$E[S(\pi)] = \sum_{i=1}^{n} P(Y_{(i)} > X_{(\pi_i)}) = \sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(\pi_i)}).$$  \hspace{1cm} (3.2)

In this case, the expectation of $S(\pi)$ under ordered matching, i.e., $\pi = \pi^0 = (1, 2, \ldots, n)$ is given by

$$V_1 = E[S(\pi^0)] = \sum_{i=1}^{n} U(\mu_{(i)} - \lambda_{(i)}).$$ \hspace{1cm} (3.3)

and the expectation of $S(\pi)$ under random matching is given by

$$V_2 = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} U(\mu_{(i)} - \lambda_{(j)}).$$ \hspace{1cm} (3.4)

In this chapter we will assume that $F(x)$ is a unimodal distribution function and discuss the properties of $E[S(\pi)]$ under ordered matching, fair matching, and some general situations. We will also discuss some rearrangement properties of $E[S(\pi)]$.

### 3.2 Preliminary Results

In this section, we state some known definitions, mainly from Marshall and Oklin (1979), and results needed later. Only references to the proofs of these results will be given here.

**Definition 3.1** A distribution function is said to be unimodal if there exists $x_0$ such that it is convex on $(-\infty, x_0)$ and concave on $(x_0, +\infty)$.

It is easy to see that if $F(x)$ is absolutely continuous, then $F(x)$ being unimodal is equivalent to that its density function $f(x) = F'(x)$ is nondecreasing on $(-\infty, x_0)$ and nonincreasing on $(x_0, +\infty)$. 
By the above definition, we can see that almost all the common distribution functions are unimodal.

**Lemma 3.2.1** If $X_1$ and $X_2$ are iid unimodal, then $X_1 - X_2$ is unimodal.

For the proof, see Dharmadhikari and Joag-dev, 1988, p15.

**Definition 3.2** For any $u, v \in R^n$, $u \leq^b v$ means that $v$ can be reached from $u$ by successive interchanges of the components of $u$, each of which corrects an inversion of the natural (i.e., nondecreasing) order.

For example, if $u = (x(2), x(1), x(3), \ldots, x(n))$ and $v = (x(1), x(2), \ldots, x(n))$, then $u \leq^b v$ since $v$ can be obtained from $u$ by interchanging the components $x(2)$ and $x(1)$ of $u$ which corrects an inversion of the natural order.

Let $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ be a permutation of $(1, 2, \ldots, n)$. And for any $x \in R^n$ write $\pi(x) = (x\pi_1, x\pi_2, \ldots, x\pi_n)$. Also write $x \uparrow = (x(1), x(2), \ldots, x(n))$ and $x \downarrow = (x(n), x(n-1), \ldots, x(1)) = (x[1], x[2], \ldots, x[n])$.

**Definition 3.3** For any $x, y, u$ and $v \in R^n$, $(x, y) \leq^a (u, v)$ means that $u$ is a permutation of $x$ and $v$ is a permutation of $y$, and there exist permutations $\pi(1)$ and $\pi(2)$ such that $\pi(1)(x) = x \uparrow$, $\pi(2)(u) = u \uparrow$ and $\pi(1)(y) \leq^b \pi(2)(v)$.

For example, let

$x = (4, 2, 6, 7)$, \quad $u = (2, 4, 7, 6)$,

$y = (3, 1, 9, 7)$, \quad $v = (1, 3, 9, 7)$.

Then, for $\pi(1) = (2, 1, 3, 4)$ and $\pi(2) = (1, 2, 4, 3)$, we have $\pi(1)(x) = x \uparrow$ and $\pi(2)(u) = u \uparrow$. Also, since

$\pi(1)(y) = (1, 3, 9, 7) \leq^b \pi(2)(v) = (1, 3, 7, 9)$. 

we have

\[(x, y) \leq^a (u, v)\]

Clearly, \((x \uparrow, y) \leq^a (x \uparrow, v)\) if \(y \leq^b v\).

**Definition 3.4** A function \(g\) of two vector arguments is called an arrangement increasing (AI) function if \(g(x, y) \leq g(u, v)\) when \((x, y) \leq^a (u, v)\).

Note. If \(g\) is AI, then \(g(x \downarrow, y \uparrow) = g(x \uparrow, y \downarrow) \leq g(x, y) \leq g(x \downarrow, y \uparrow) = g(x \downarrow, y \downarrow)\).

**Lemma 3.2.2** \(g(u, v) = \sum_{i=1}^{n} \phi(u_i, v_i)\) is AI if and only if \(\frac{\partial^2}{\partial r \partial s} \phi(r, s) \geq 0\) provided the derivative exists.

For the proof, see Marshall and Olkin 1979, p150 and Hollander, Proschan and Sethuraman (1977).

**Definition 3.5** For any \(x, y \in \mathbb{R}^n\), \(x \prec y\) if

\begin{enumerate}
\item[(a)] \(\sum_{i=1}^{k} x[i] \leq \sum_{i=1}^{k} y[i] \quad k = 1, 2, \ldots, n - 1\).
\item[(b)] \(\sum_{i=1}^{n} x[i] = \sum_{i=1}^{n} y[i]\).
\end{enumerate}

When \(x \prec y\), \(x\) is said to be majorized by \(y\). Also

\(x \prec_w y\) if \(\sum_{i=1}^{k} x[i] \leq \sum_{i=1}^{k} y[i] \quad k = 1, 2, \ldots, n\).

\(x \prec^w y\) if \(\sum_{i=1}^{k} x(i) \geq \sum_{i=1}^{k} y(i) \quad k = 1, 2, \ldots, n\).

We say \(x\) is weakly submajorized by \(y\) if \(x \prec_w y\) and \(x\) is weakly supermajorized by \(y\) if \(x \prec^w y\).
Lemma 3.2.3 (Hardy, Littlewood, and Pólya, 1952) The inequality
\[ \sum_{i=1}^{n} g(x_i) \leq \sum_{i=1}^{n} g(y_i) \]
for all continuous convex functions \( g : \mathbb{R} \rightarrow \mathbb{R} \) if and only if \( x \leq y \).

Lemma 3.2.4 (Tomić, 1949) The inequality \( \sum_{i=1}^{n} g(x_i) \leq \sum_{i=1}^{n} g(y_i) \) holds for all continuous increasing convex functions \( g \) if and only if \( x \prec_w y \). It holds for all decreasing convex functions if and only if \( x \succ_w y \).

Lemma 3.2.5 (Mitrinovic, 1970, p22) If \( \phi(x) \) is a concave function on \( I = [0, a] \), if \( x \in I \) (\( i = 1, \ldots, n \)) and \( x_1 + \ldots + x_k \in I \), then
\[ \phi(x_1) + \phi(x_2) + \ldots + \phi(x_k) \geq \phi(x_1 + x_2 + \ldots + x_k) + (k - 1)\phi(0). \]

3.3 Ordered Matching

For any fixed \( \lambda(1) \leq \lambda(2) \leq \ldots \leq \lambda(n) \) and \( \mu(1) \leq \mu(2) \leq \ldots \leq \mu(n) \), let \( \tilde{\pi} = (\pi_1, \pi_2, \ldots, \pi_n) \) be a permutation for which
\[ E[S(\tilde{\pi})] = \max_\pi \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i)). \tag{3.5} \]
One question here is under what conditions \( \tilde{\pi} = \pi^o \), i.e., \( E[S(\pi)] \) attains its maximum value under ordered matching?

We first consider the case \( n = 2 \).

Theorem 3.3.1 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 \( \lambda(1) \leq \lambda(2) \), \( \mu(1) \leq \mu(2) \), and \( F(x) \) is an absolutely continuous unimodal distribution. If \( \mu(1) + \mu(2) \geq \lambda(1) + \lambda(2) \), then we have
$$P(Y'_1 > X'_1) + P(Y'_2 > X'_2) \geq P(Y'_1 > X'_2) + P(Y'_2 > X'_1)$$

or

$$U(\mu_1 - \lambda_1) + U(\mu_2 - \lambda_2) \geq U(\mu_1 - \lambda_2) + U(\mu_2 - \lambda_1). \quad (3.6)$$

Proof. Let \( c = \mu_1 + \mu_2 - \lambda_1 - \lambda_2 \). Consider the function

\[ h(x) = U(x) + U(c-x) \quad \text{for} \quad x \geq c \geq 0. \]

By the definition of \( U(x) \) and Lemma 3.2.1, we know that \( U(x) \) is also an absolutely continuous unimodal distribution. If \( u(x) = U'(x) \), we have that \( u(x) \) is symmetric about 0 and nonincreasing on \((0, +\infty)\). Now \( h'(x) = u(x) - u(c-x) \). Since \( x - (c-x) = 2x - c \geq 0 \) for \( x \geq \frac{c}{2} \), it follows that \( h'(x) \leq 0 \) for \( x \geq \frac{c}{2} \), i.e., \( h(x) \) is a nonincreasing function for \( x \geq \frac{c}{2} \). Let \( x_1 = \max\{\mu_1 - \lambda_1, \mu_2 - \lambda_2\} \) and \( x_2 = \mu_2 - \lambda_1 \), then \( x_2 \geq x_1 \geq \frac{c}{2} \). Therefore, \( h(x_1) \geq h(x_2) \), i.e., \( (3.6) \) holds. \( \square \)

Note:

1. If \( \mu_1 + \mu_2 = \lambda_1 + \lambda_2 \), by \( \mu_1 - \lambda_1 = -(\mu_2 - \lambda_2) \), \( \mu_1 - \lambda_2 = -(\mu_2 - \lambda_1) \) and by the properties of \( U(x) \), we have

\[ U(\mu_1 - \lambda_1) + U(\mu_2 - \lambda_2) = U(\mu_1 - \lambda_2) + U(\mu_2 - \lambda_1) = 1. \]

2. If either \( \mu_1 = \mu_2 \) or \( \lambda_1 = \lambda_2 \), then equality holds in \( (3.6) \).

Corollary If \( \mu_1 \neq \mu_2 \), \( \lambda_1 \neq \lambda_2 \), and \( u(x) = U'(x) \) is strictly decreasing on
[0, +∞), then \( \mu_1 + \mu_2 \geq \lambda_1 + \lambda_2 \) if and only if (3.6) holds.

**Proof.** We need to show only that (3.6) implies \( \mu_1 + \mu_2 \geq \lambda_1 + \lambda_2 \). If (3.6) holds and \( \mu_1 + \mu_2 < \lambda_1 + \lambda_2 \), then switch \( \mu \) and \( \lambda \) in Theorem 3.2.1. Therefore, in the proof of Theorem 3.3.1, we have \( x_1 = \max\{\lambda_1 - \mu_1, \lambda_2 - \mu_2\} \) and \( x_2 = \lambda_2 - \mu_1 \). Now we have \( x_2 > x_1 \), since otherwise we could get either \( \mu_1 = \mu_2 \) or \( \lambda_1 = \lambda_2 \). Hence, by the given condition on \( u(x) \), it follows that

\[
U(\lambda_1 - \mu_1) + U(\lambda_2 - \mu_2) > U(\lambda_1 - \mu_2) + U(\lambda_2 - \mu_1),
\]

i.e.,

\[
U(\mu_1 - \lambda_1) + U(\mu_2 - \lambda_2) < U(\mu_1 - \lambda_2) + U(\mu_2 - \lambda_1)
\]

which contradicts (3.6). 

Theorem 3.2.1 tells us for \( n = 2 \), that if

\[
\sum_{i=1}^{n} \lambda_i \leq \sum_{i=1}^{n} \mu_i,
\]

then

\[
E[S(\pi^0)] = \max_{\pi} \sum_{i=1}^{n} U(\mu_i - \lambda_{\pi_i}).
\]

However, for \( n > 2 \), (3.7) no longer implies (3.8). For example, let \( F(x) = \Phi(x) \), the standard normal cdf. Then \( U(x) = \Phi(\frac{x}{\sqrt{2}}) \). Consider

\[
\mu_1 = 9\sqrt{2}, \quad \mu_2 = 10\sqrt{2}, \quad \mu_3 = 12\sqrt{2},
\]

\[
\lambda_1 = \sqrt{2}, \quad \lambda_2 = 11\sqrt{2}, \quad \lambda_3 = 13\sqrt{2}.
\]

Then (3.7) holds, and

\[
\sum_{i=1}^{3} U(\mu_i - \lambda_{\pi_i}) = \Phi(8) + \Phi(-1) + \Phi(-1) = 1.73147.
\]
but
\[
\Phi\left(\frac{\mu(1) - \lambda(3)}{\sqrt{2}}\right) + \Phi\left(\frac{\mu(2) - \lambda(1)}{\sqrt{2}}\right) + \Phi\left(\frac{\mu(3) - \lambda(2)}{\sqrt{2}}\right) \\
= \Phi(-4) + \Phi(1) + \Phi(9) \doteq 1.8431.
\]

For \( n = 3 \), a sufficient condition for (3.8) to hold is given by the following lemma.

Lemma 3.3.1  \( \text{For } n = 3, \text{ if} \)
\[
\mu(i) + \mu(j) \geq \lambda(i) + \lambda(j), \quad \text{for any } 1 \leq i < j \leq 3, \quad (3.9)
\]
and also either \( \mu(1) + \mu(2) \geq \lambda(1) + \lambda(3) \) or \( \mu(1) + \mu(3) \geq \lambda(2) + \lambda(3) \) then (3.8) holds.

Proof. For any permutation \( \pi_1, \pi_2, \pi_3 \) of \( 1, 2, 3 \), if there exists \( i \) such that \( \pi_i = i \), then by (3.9) and Theorem 3.2.1, we have
\[
\sum_{i=1}^{3} U(\mu(i) - \lambda(i)) \geq \sum_{i=1}^{3} U(\mu(i) - \lambda(\pi_i)).
\]
Therefore, we only need to show the remaining cases, i.e.,
\[
\sum_{i=1}^{3} U(\mu(i) - \lambda(i)) \geq U(\mu(1) - \lambda(3)) + U(\mu(2) - \lambda(1)) + U(\mu(3) - \lambda(2)) \quad (3.10)
\]
and
\[
\sum_{i=1}^{3} U(\mu(i) - \lambda(i)) \geq U(\mu(1) - \lambda(2)) + U(\mu(2) - \lambda(3)) + U(\mu(3) - \lambda(1)). \quad (3.11)
\]
If \( \mu(1) + \mu(2) \geq \lambda(1) + \lambda(3) \), by Theorem 3.2.1, we have
\[
U(\mu(1) - \lambda(1)) + U(\mu(2) - \lambda(3)) \geq U(\mu(1) - \lambda(3)) + U(\mu(2) - \lambda(1)).
\]
Adding $U(\mu(3) - \lambda(2))$ to both sides of this inequality, the RHS is the RHS of (3.10).

Note that

$$U(\mu(2) - \lambda(2)) + U(\mu(3) - \lambda(3)) \geq U(\mu(2) - \lambda(3)) + U(\mu(3) - \lambda(2)),$$

(3.10) follows immediately. Similarly, if $\mu(1) + \mu(3) \geq \lambda(1) + \lambda(2)$, we can show that (3.10) holds.

The proof of (3.11) is same as that of (3.10). □

Note:

(i) (3.9) does not imply (3.8). For example, $F(x) = \Phi(x)$ and

$$\mu(1) = 1.1\sqrt{2}, \quad \mu(2) = 3.1\sqrt{2}, \quad \mu(3) = 5.1\sqrt{2},$$

$$\lambda(1) = \sqrt{2}, \quad \lambda(2) = 3\sqrt{2}, \quad \lambda(3) = 5\sqrt{2}.$$

We have

$$\sum_{i=1}^{3} U(\mu(i) - \lambda(i)) = 3\Phi(0.1) = 1.6194.$$

But

$$U(\mu(1) - \lambda(3)) + U(\mu(2) - \lambda(1)) + U(\mu(3) - \lambda(2))$$

$$\Phi(-3.9) + \Phi(2.1) + \Phi(2.1) = 1.96428 > 1.6194.$$

The above example also tells us that even if $\mu(i) \geq \lambda(i)$ for $i = 1, 2, 3$, it is not necessary that (3.8) holds.

(ii) The conditions in Lemma 3.3.1 are not necessary for (3.8). For example,
\( F(x) = \Phi(x) \) and
\[
\begin{align*}
\mu(1) &= 10\sqrt{2}, \quad \mu(2) = 20\sqrt{2}, \quad \mu(3) = 30\sqrt{2}, \\
\lambda(1) &= 6\sqrt{2}, \quad \lambda(2) = 16\sqrt{2}, \quad \lambda(3) = 25\sqrt{2}.
\end{align*}
\]

**Theorem 3.3.2** For any fixed \( n \), there exists a constant \( c \) such that when \( \mu(i) - \lambda(i) \geq c, (i = 1, 2, \ldots, n) \), (3.8) holds, provided \( U'''(x) \) exists.

**Proof.** Since \( U(x) \uparrow 1 \) as \( x \to +\infty \), there exists \( c \) such that \( nU(c) \geq n - \frac{1}{2} \). Now, if \( \mu(i) - \lambda(i) \geq c \) (\( i = 1, 2, \ldots, n \)), then for any permutation \((k_1, k_2, \ldots, k_n)\), if there exists \( l \) such that \( \mu(l) - \lambda(k_l) < 0 \), then
\[
\sum_{i=1}^{n} U(\mu(i) - \lambda(k_i)) \leq \frac{1}{2} + (n - 1) \leq n - \frac{1}{2} \leq nU(c) \leq \sum_{i=1}^{n} U(\mu(i) - \lambda(i)).
\]

Therefore, we only need show that
\[
\sum_{i=1}^{n} U(\mu(i) - \lambda(i)) \geq \sum_{i=1}^{n} U(\mu(i) - \lambda(k_i)), \tag{3.12}
\]
where \((k_1, k_2, \ldots, k_n)\) is any permutation such that \( \mu(i) - \lambda(k_i) \geq 0 \) (\( i = 1, 2, \ldots, n \)).

Let \( \phi(r, s) = U(r - s) \), since \( U(x) \) is concave on \([0, +\infty)\), we have
\[
\frac{\partial^2}{\partial r \partial s} \phi(r, s) = -U'''(r - s) \geq 0
\]
for \( r \geq s \). Therefore, by Lemma 3.2.2 we have that \( \sum_{i=1}^{n} U(\mu(i) - \lambda(k_i)) \) is arrangement increasing among the permutations such that \( \mu(i) - \lambda(k_i) \geq 0 \). In particular, (3.8) holds. **Box**
Lemma 3.3.2 If \( \min_{1 \leq i \leq n} \{\mu_i\} \geq \max_{1 \leq i \leq n} \{\lambda_i\} \), i.e. \( \mu(1) \geq \lambda(n) \), then

\[ \sum_{i=1}^{n} U(\mu(i) - \lambda(k_i)) \text{ is Al, and (3.8) holds.} \]

**Proof.** Use the same argument as in the proof of Theorem 3.3.2.

Lemma 3.3.3 If

\[ \sum_{i=1}^{n} U(\mu(i) - \lambda(i)) > \sum_{i=1}^{n} U(\mu(i) - \lambda(k_i)) \]  

(3.13)

for any \((k_1, k_2, \ldots, k_n) \neq (1, 2, \ldots, n)\), then we have \( \mu(i) + \mu(j) > \lambda(i) + \lambda(j) \) for any \( 1 \leq i < j \leq n \).

**Proof.** If there exist \( l \) and \( m \) such that \( \mu(l) + \mu(m) \leq \lambda(l) + \lambda(m) \), where \( l < m \), then we have

\[ U(\mu(l) - \lambda(m)) + U(\mu(m) - \lambda(l)) \geq U(\mu(l) - \lambda(l)) + U(\mu(m) - \lambda(m)). \]

Adding the terms \( U(\mu(i) - \lambda(i)) \) (\( 1 \leq i \leq n \) and \( i \neq l, m \)) to both sides, this leads to a contradiction of (3.13). \( \square \)

Note:

(1) If (3.13) holds, then \( \mu_i \neq \mu_j \) and \( \lambda_i \neq \lambda_j \) for any \( i \neq j \). Otherwise if there exist \( l \) and \( m \) such that \( \mu(l) = \mu(m) \), where \( l < m \), we have

\[ U(\mu(l) - \lambda(l)) + U(\mu(m) - \lambda(m)) \]
\[ = U(\mu(m) - \lambda(l)) + U(\mu(l) - \lambda(m)) \]
which can lead to a contradiction of (3.13).

(2) If (3.8) holds and there exist \( \mu \)'s or \( \lambda \)'s that are equal, then we can find other permutation \((k'_1, k'_2, \ldots, k'_n)\) such that

\[
\sum_{i=1}^{n} U(\mu(i) - \lambda(i)) = \sum_{i=1}^{n} U(\mu(i) - k'_i),
\]

i.e., the ordered matching is not unique in attaining the maximum value.

**Lemma 3.3.4** For \( n = 2 \), and \( U(x) \) strictly increasing, the following are equivalent.

(1) \( \mu(1) + \mu(2) \geq \lambda(1) + \lambda(2) \).

(2) \( U(\mu(1) - \lambda(1)) + U(\mu(2) - \lambda(2)) \geq 1 \).

(3) \( U(\mu(1) - \lambda(2)) + U(\mu(2) - \lambda(1)) \geq 1 \).

**Proof.** (a) "(1) \Rightarrow (2)"

If (1) holds, we have \( \mu(1) - \lambda(1) \geq \lambda(2) - \mu(2) \). By the symmetry of \( U(x) \), we have

\[
U(\mu(1) - \lambda(1)) + U(\mu(2) - \lambda(2)) = U(\mu(1) - \lambda(1)) + 1 - U(\lambda(2) - \mu(2)) \geq 1.
\]

(b) "(2) \Rightarrow (3)"

If (2) holds, then since

\[
U(\mu(1) - \lambda(1)) + U(\mu(2) - \lambda(2)) = U(\mu(1) - \lambda(1)) - U(\lambda(2) - \mu(2)) + 1 \geq 1,
\]

we have \( \mu(1) - \lambda(1) \geq \lambda(2) - \mu(2) \), i.e., \( \mu(1) - \lambda(2) \geq \lambda(1) - \mu(2) \). Therefore,

\[
U(\mu(1) - \lambda(2)) + U(\mu(2) - \lambda(1))
\]
\[ \mu(1) - \lambda(2) = \frac{1}{1} - \frac{1}{1} \]

(c) \( \dot{3} = \hat{1} \). Since

\[ U(\mu(1) - \lambda(2)) + U(\mu(2) - \lambda(1)) = U(\mu(1) - \lambda(2)) - U(\lambda(1) - \mu(2)) + 1 \geq 1, \]

we have \( \mu(1) - \lambda(2) \geq \lambda(1) - \mu(2) \), i.e., \( \mu(1) + \mu(2) \geq \lambda(1) + \lambda(2) \). \( \square \)

In general, \( \sum_{i=1}^{n} \mu(i) \geq \sum_{i=1}^{n} \lambda(i) \) does not imply

\[ \sum_{i=1}^{n} U(\mu(i) - \lambda(i)) \geq \frac{n}{2}, \] (3.14)

i.e., \( \Gamma_Y \) is better than \( \Gamma_X \) in an ordered matching. For example, let \( F(x) = \Phi(x) \), \( n = 3 \), and

\[ \mu(1) = \sqrt{2}, \quad \mu(2) = 2\sqrt{2}, \quad \mu(3) = 20\sqrt{2}, \]

\[ \lambda(1) = 6\sqrt{2}, \quad \lambda(2) = 7\sqrt{2}, \quad \lambda(3) = 8\sqrt{2}. \]

Then \( \sum_{i=1}^{3} \mu(i) > \sum_{i=1}^{3} \lambda(i) \). But

\[ \sum_{i=1}^{3} U(\mu(i) - \lambda(k_i)) = \sum_{i=1}^{n} \Phi\left( \frac{\mu(i) - \lambda(k_i)}{\sqrt{2}} \right) < \frac{3}{2} \]

for any permutation \( (k_1, k_2, k_3) \) of \( (1, 2, 3) \).

Writing \( \sum_{i=1}^{n} (\mu(i) - \lambda(i)) = c \) and \( m = \# \) of \( \mu(i) - \lambda(i) \geq 0 \) \( (i = 1, 2, \ldots, n) \), we have the following result.

**Lemma 3.3.5** For \( n = 3 \), if \( c \geq 0 \) and \( m > 1 \), then (3.14) holds.
Proof. WLG. assume $\mu(1) - \lambda(1) \geq 0$ and $\mu(2) - \lambda(2) \geq 0$, then by Lemma 3.2.5, we have

$$U(\mu(1) - \lambda(1)) + U(\mu(2) - \lambda(2)) \geq U(\mu(1) - \lambda(1) + \mu(2) - \lambda(2)) + U(0).$$

Hence

$$\sum_{i=1}^{3} U(\mu(i) - \lambda(i)) \geq U(\mu(1) - \lambda(1) + \mu(2) - \lambda(2)) + U(0) + 1 - U(\mu(1) - \lambda(1) + \mu(2) - \lambda(2) - c).$$

Since $c \geq 0$, we have $\mu(1) - \lambda(1) + \mu(2) - \lambda(2) \geq \mu(1) - \lambda(1) + \mu(1) - \lambda(2) - c$.

Therefore, (3.14) holds.

For $m = 1$, there is no general result.

Let $\{a_1, a_2, \ldots, a_m\} = \{\mu(i) - \lambda(i), \text{s.t. } \mu(i) - \lambda(i) \geq 0, \ i = 1, 2, \ldots, n\}$ and $\{b_1, b_2, \ldots, b_{n-m}\} = \{\lambda(i) - \mu(i), \text{s.t. } \mu(i) - \lambda(i) < 0, \ i = 1, 2, \ldots, n\}$.

**Lemma 3.3.6** If $m \geq \lceil \frac{n}{2} \rceil$ and

$$\sum_{i=1}^{k} a(i) \geq \sum_{i=1}^{k} b(i) \quad (k = 1, 2, \ldots, n - m),$$

then (3.14) holds.

Proof.

$$\sum_{i=1}^{n} U(\mu(i) - \lambda(i)) = \sum_{i=1}^{m} U(a(i)) + \sum_{i=1}^{n-m} U(-b(i))$$

$$= \sum_{i=1}^{m} U(a(i)) + (n - m) - \sum_{i=1}^{n-m} U(b(i))$$
\[ \geq \sum_{i=1}^{n-m} U(a(i)) - \sum_{i=1}^{n-m} U(b(i)) + \frac{1}{2}[m - (n - m)] + (n - m) \]
\[ = \sum_{i=1}^{n-m} U(a(i)) - \sum_{i=1}^{n-m} U(b(i)) + \frac{n}{2}. \]

Since \( U(x) \) is an increasing concave function on \([0, +\infty)\), we have \(-U(x)\) is a decreasing convex function on \([0, +\infty)\). By Lemma 3.2.4, we have

\[ \sum_{i=1}^{n-m} -U(a(i)) \leq \sum_{i=1}^{n-m} -U(b(i)) \]

that is

\[ \sum_{i=1}^{n-m} U(a(i)) \geq \sum_{i=1}^{n-m} U(b(i)). \]

Therefore, (3.14) follows. \( \Box \)

Now, let us consider the relationship between ordered matching and random matching. Let \( V_1 \) and \( V_2 \) be the expectations of \( S(\pi) \) under ordered matching and random matching, respectively, where the expressions of \( V_1 \) and \( V_2 \) are given by (3.3) and (3.4). Then we have the following result.

**Theorem 3.3.3** Let \( (X'_1, X'_2, \ldots, X'_n) \) and \( (Y'_1, Y'_2, \ldots, Y'_n) \) be independent with \( X'_i \sim F(x - \lambda(i)) \) and \( Y'_i \sim F(x - \mu(i)) \), \( i = 1, 2, \ldots, n \), where \( \lambda(1) \leq \lambda(2) \leq \cdots \leq \lambda(n), \mu(1) \leq \mu(2) \leq \cdots \leq \mu(n) \), and \( F(x) \) is an absolutely continuous unimodal distribution. If

\[ \mu(i) + \mu(j) \geq \lambda(i) + \lambda(j) \quad \text{for any} \quad 1 \leq i < j \leq n, \]
then we have
\[ k \sum_{i=1}^{k} P(Y_{(i)}' > X_{(i)}') \geq \sum_{i=1}^{k} \sum_{j=1}^{k} P(Y_{(i)}' > X_{(j)}') \quad k = 1, \ldots, n. \]

In particular, for \( k = n \), we have \( V_1 \geq V_2 \).

**Proof.** Using the given condition and Theorem 3.3.1, we can prove this result by following the same arguments as in the proof of Theorem 2.2.2. \( \square \)

### 3.4 Fair Matching under the Linear Preference Model

In Chapter 2, we discussed fair matchings for the order statistics model, in which simple and symmetric matchings are fair. For the linear preference model, we see that simple matchings are still fair. However, symmetric matchings are not necessarily fair in this case. To see this, let us consider \( n = 4 \) and \( F(x) = \Phi(x) \), i.e., the standard normal distribution function. Let \( \lambda_{(i)} = \mu_{(i)} \quad (i = 1, 2, 3, 4) \) and

\[ \lambda(1) = \sqrt{2}, \quad \lambda(2) = 1.2\sqrt{2}, \quad \lambda(3) = 1.7\sqrt{2}, \quad \lambda(4) = 2\sqrt{2}. \]

Note that \( \pi = (\pi_1, \pi_2, \pi_3, \pi_4) = (2, 4, 1, 3) \) is a symmetric matching. But

\[ ES(\pi) = \sum_{i=1}^{4} U(\mu_{(i)} - \lambda_{(\pi_i)}) \]

\[ = \Phi(-0.2) + \Phi(-0.8) + \Phi(0.7) + \Phi(0.3) \]

\[ = 2.0085. \]

This tells us that \( \pi \) is not fair. The following lemma shows that under certain condition, symmetric matchings are still fair.
**Lemma 3.4.1** If \( \lambda(i) = \mu(i) \) and

\[
\lambda(i+1) - \lambda(i) = \lambda(n-i+1) - \lambda(n-i)
\]

for \( i = 1, 2, \ldots, n \), then

\[
ES(\pi) = \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i)) = \frac{n}{2}
\]

for any symmetric matching \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \).

**Proof.** For any \( i \) and \( j \), by (3.15), we have

\[
\lambda(j) - \lambda(i) = \lambda(n-i+1) - \lambda(n-j+1).
\]

In order to show this, we can assume \( i < j \). Then

\[
\begin{align*}
\lambda(j) - \lambda(i) &= \lambda(j) - \lambda(j-1) + \lambda(j-1) - \lambda(j-2) + \cdots + \lambda(i+1) - \lambda(i) \\
&= \lambda(n-j+2) - \lambda(n-j+1) + \lambda(n-j+3) - \lambda(n-j+2) + \cdots \\
&\quad + \lambda(n-i+1) - \lambda(i) \\
&= \lambda(n-i+1) - \lambda(n-j+1).
\end{align*}
\]

Since

\[
ES(\pi) = \frac{1}{2} \left[ \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i)) + \sum_{i=1}^{n} U(\mu(n-i+1) - \lambda(\pi_{n-i+1})) \right]
\]

and \( \pi_{n-i+1} = n - \pi_i + 1 \), by (3.17), we have

\[
ES(\pi) = \frac{1}{2} \left[ \sum_{i=1}^{n} U(\lambda(i) - \lambda(\pi_i)) + \sum_{i=1}^{n} U(\lambda(n-i+1) - \lambda(n-\pi_i+1)) \right]
\]
The following lemma tells us that simple matchings are fair.

**Lemma 3.4.2** If \( \lambda(i) = \mu(i) \) for \( i = 1, 2, \ldots, n \), then (3.16) holds for any simple matching \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \).

**Proof.** Since \( U(\mu(i) - \lambda(j)) = U(\lambda(i) - \mu(j)) = 1 - U(\mu(j) - \lambda(i)) \), then by the definition of simple permutation, (3.16) follows. □

**Lemma 3.4.3** If \( \mu(i) + \mu(j) \geq \lambda(i) + \lambda(j) \) for all \( 1 \leq i < j \leq n \), then

\[
\begin{align*}
(a) & \quad \sum_{i=1}^{n} U(\mu(i) - \lambda(i)) \geq \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i)) \\
(b) & \quad \sum_{i=1}^{n} U(\mu(i) - \lambda(i)) \geq \frac{n}{2}.
\end{align*}
\]

**Proof.** (a) This follows easily from Theorem 3.3.1 and the definition of simple permutation.

(b) If \( n = 2m \), then by \( \mu(2k-1) + \mu(2k) \geq \lambda(2k-1) + \lambda(2k) \) \( (k = 1, 2, \ldots, m) \) and Lemma 3.3.4, we have

\[
U(\mu(2k-1) - \lambda(2k-1)) + U(\mu(2k) - \lambda(2k)) \geq 1.
\]
for \( k = 1, 2, \ldots, m \). Hence
\[
\sum_{k=1}^{m} \left[ U(\mu(2k-1) - \lambda(2k-1)) + U(\mu(2k) - \lambda(2k)) \right] \geq m = \frac{n}{2},
\]
i.e., (b) holds.

If \( n = 2m + 1 \), then there must exist \( l \) such that \( \mu(l) \geq \lambda(l) \). Hence \( U(\mu(l) - \lambda(l)) \geq \frac{1}{2} \). Therefore, using the same argument as for \( n \) even, we have
\[
\sum_{i \neq l}^{l} U(\mu(i) - \lambda(i)) \geq m,
\]
hence
\[
\sum_{i=1}^{n} U(\mu(i) - \lambda(i)) \geq m + \frac{1}{2} = \frac{n}{2}. \quad \square
\]

Note:

(1) Under the condition given in above lemma, if \( n = 2m \), i.e., \( n \) is even, then by Lemma 3.3.4 and the definition of simple matching, we have
\[
\sum_{i=1}^{n} U(\mu(i) - \lambda(\pi(i))) \geq \frac{n}{2} \quad (3.18)
\]
for any simple matching \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \). However, if \( n = \text{odd} \), then (3.18) does not necessarily hold. We can see this as follows. Let \( n = 3 \), \( F(x) = \Phi(x) \) and
\[
\lambda(1) = 4\sqrt{2}, \quad \lambda(2) = 4.1\sqrt{2}, \quad \lambda(3) = 10.2\sqrt{2},
\]
\[
\lambda(1) = \sqrt{2}, \quad \lambda(2) = 7.2\sqrt{2}, \quad \lambda(3) = 15\sqrt{2}.
\]
Then \( \mu(i) + \mu(j) \geq \lambda(i) + \lambda(j) \) for \( 1 \leq i < j \leq 3 \). But, for the simple matching \((1, \ 3, \ 2)\), we have
\[
U(\mu(1) - \lambda(1)) + U(\mu(2) - \lambda(3)) + U(\mu(3) - \lambda(2))
\]
\[ = \Phi(-3) + \Phi(-3) + \Phi(10.9) < \frac{3}{2}. \]

(2) If \( \mu(i) \geq \lambda(i) \) for \( i = 1, 2, \ldots, n \), then by Lemma 3.3.4 and the definition of simple matching, it follows that (3.18) holds for any simple matching \( \pi \).

**Lemma 3.4.4** If \( \mu(i) = \lambda(i) \) for \( i = 1, 2, \ldots, n \), then

\[
U(\mu(1) - \lambda(n)) + \sum_{i=2}^{n} U(\mu(i) - \lambda(i-1)) \geq \frac{n}{2}. \quad (3.19)
\]

**Proof.** Let \( x = \lambda(n) - \mu(1) \) and \( x_i = \mu(i) - \lambda(i-1) \geq 0 \) (\( i = 2, 3, \ldots, n \)), we have

\[
x = \sum_{i=2}^{n} x_i. \text{ Since } U(x) \text{ is concave on } [0, +\infty), \text{ it follows by Lemma 3.2.5 that}
\]

\[
\sum_{i=2}^{n} U(x_i) \geq U\left( \sum_{i=2}^{n} x_i \right) + (n - 2)U(0)
\]

\[
= U(x) + (n - 2)\frac{1}{2}.
\]

Note that

\[
U(x) = 1 - U(-x) = 1 - U(\mu(1) - \lambda(n)),
\]

(3.19) follows immediately. \( \square \)

From the above lemma, we have

\[
U(\mu(1) - \lambda(n)) + \sum_{i=2}^{n} U(\mu(i) - \lambda(i-1)) \geq \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i)) \quad (3.20)
\]

for any simple matching \( \pi = (\pi_1, \pi_2, \ldots, \pi_n) \). One might feel that (3.20) holds for any matching \( \pi \). However, this is not true. We can see this by the following example:
Let \( \lambda(i) = \mu(i) = i\sqrt{2} \) \( (i = 1, 2, \ldots, n) \), and \( F(x) = \Phi(x) \). Consider \( (\pi_1, \pi_2, \ldots, \pi_n) = (n, n - 1, 1, 2, \ldots, n - 2) \). Then

\[
\sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i))
\]

\[
= U(\mu(1) - \lambda(n)) + U(\mu(2) - \lambda(n-1)) + \sum_{i=3}^{n} U(\mu(i) - \lambda(n-2))
\]

\[
= \Phi(-n + 1) + \Phi(-n + 3) + (n - 2)\Phi(2)
\]

and

\[
U(\mu(1) - \lambda(n)) + \sum_{i=2}^{n} U(\mu(i) - \lambda(i-1)) = \Phi(-n + 1) + (n - 1)\Phi(1).
\]

Since \( \Phi(1) = 0.8413 \) and \( \Phi(2) = 0.9772 \). Taking \( n = 12 \), we have

\[
U(\mu(1) - \lambda(n)) + \sum_{i=2}^{n} U(\mu(i) - \lambda(i-1)) < \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i)).
\]

Note.

Let \( T_n \) be the total number of simple matchings, then

\[
\lim_{n \to \infty} \frac{T_n}{n!} = 0.
\]

This can be shown as follows.

Since by Lemma 2.3.2, we can write \( T_n \) as

\[
T_n = \sum_{i=0}^{[n/2]} (2i - 1)!! \binom{n}{2i},
\]

where \((2i-1)!! = (2i-1)(2i-3) \ldots 1\). Note that

\[
\sum_{i=0}^{[n/2]} \binom{n}{2i} = 2^{n-1},
\]
we have
\[ \frac{T_n}{n!} \leq \frac{(2^\left( \frac{n}{2} \right) - 1)!!2^{n-1}}{n!} \leq \frac{4}{n(n-2) \cdots 2} \leq \frac{c}{n}, \]
where \( c \) is some constant. Therefore, we have \( \frac{T_n}{n!} \to 0 \) as \( n \to \infty \).

### 3.5 Some Optimality Results

For given \( \mu_i \)'s and \( \lambda_i \)'s, it is interesting to find a permutation \( \pi_1, \pi_2, \ldots, \pi_n \) such that
\[
\sum_{i=1}^{n} U(\mu_i) - \lambda(\pi_i) \geq \sum_{i=1}^{n} U(\mu_i) - \lambda(\pi_i)
\]
for any permutation \( \pi_1, \pi_2, \ldots, \pi_n \).

We have shown that under certain conditions, \( (\pi_1, \pi_2, \ldots, \pi_n) = (1, 2, \ldots, n) \), i.e., the ordered matching, gives the maximum value. Now let us find a sufficient condition for \( \pi_1, \pi_2, \ldots, \pi_n \) such that (3.21) holds. Let
\[
\{x_j\} = \{\mu(i) - \lambda(\pi_i) | \mu(i) - \lambda(\pi_i) > 0 \} \quad (i = 1, 2, \ldots, n),
\]
\[
\{y_j\} = \{\lambda(\pi_i) - \mu(i) | \lambda(\pi_i) - \mu(i) > 0 \} \quad (i = 1, 2, \ldots, n).
\]

Similarly, let
\[
\{x'_j\} = \{\mu(i) - \lambda(\pi_i) | \mu(i) - \lambda(\pi_i) > 0 \} \quad (i = 1, 2, \ldots, n),
\]
\[
\{y'_j\} = \{\lambda(\pi_i) - \mu(i) | \lambda(\pi_i) - \mu(i) > 0 \} \quad (i = 1, 2, \ldots, n).
\]

Suppose \( \{x_j\} \) has \( p \) elements, \( \{y_j\} \) has \( q \) elements, \( \{x'_j\} \) has \( p' \) elements, and \( \{y'_j\} \) has \( q' \) elements. Then
\[
\sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i)) = \sum_{j=1}^{p} U(x_j) + q - \sum_{j=1}^{q} U(y_j) + (n - p - q)U(0)
\]
\[
= \sum_{j=1}^{p} U(x_j) - \sum_{j=1}^{q} U(y_j) + (n - p + q)U(0)
\]

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, (3.21) is equivalent to
\[
\sum_{j=1}^{p} U(x_j) + \sum_{j=1}^{q} U(y_j) + (n - p + q)U(0) \\
\geq \sum_{j=1}^{p'} U(x'_j) + \sum_{j=1}^{q} U(y_j) + (n - p' + q')U(0).
\]

Write
\[
a_i = x_i \quad (i = 1, 2, \ldots, p),
\]
\[
a_{p+i} = y_i \quad (i = 1, 2, \ldots, q'),
\]
\[
a_{p+q'+i} = 0 \quad (i = 1, 2, \ldots, n - p + q),
\]
\[
b_i = x'_i \quad (i = 1, 2, \ldots, p'),
\]
\[
b_{p'+i} = y_i \quad (i = 1, 2, \ldots, q),
\]
\[
b_{p'+q'+i} = 0 \quad (i = 1, 2, \ldots, n - p' + q').
\]

By
\[
\sum_{i=1}^{n} (\mu(i) - \lambda(\pi_i)) = \sum_{i=1}^{n} (\mu(i) - \lambda(\pi_i)),
\]
we have \( \sum_{i=1}^{n+q+q'} a_i = \sum_{i=1}^{n+q+q'} b_i \). Also, we can write (3.21) as

\[
\sum_{i=1}^{n+q+q'} U(a_i) \geq \sum_{i=1}^{n+q+q'} U(b_i).
\]

(3.22)

Now \(-U(x)\) is a convex function on \([0, +\infty)\). Therefore, by Lemma 3.2.3, we have that (3.22) holds if \( \sum_{i=1}^{k} a[i] \leq \sum_{i=1}^{k} b[i] \), for \( k = 1, 2, \ldots, n+q+q'-1 \), where \((a[1], a[2], \ldots, a[n]) = a \downarrow\). Summarizing the above argument, we have the following lemma.

**Theorem 3.5.1** A sufficient condition for \((\tilde{\pi}_1, \tilde{\pi}_2, \ldots, \tilde{\pi}_n)\) such that (3.21) holds is

\[
\sum_{i=1}^{k} a[i] \leq \sum_{i=1}^{k} b[i] \quad \text{for} \quad k = 1, 2, \ldots, n + q + q' - 1.
\]

Note:

(1) If there exist \(i\) and \(j\) such that \(\lambda(i) = \lambda(j)\) or \(\mu(i) = \mu(j)\), then the permutation \((\tilde{\pi}_1, \tilde{\pi}_2, \ldots, \tilde{\pi}_n)\) which satisfies (3.21) is not unique.

(2) For any given \(\sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i))\), 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 3.3.1, 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 \(\pi_i < \pi_j\), then we can increase \(\sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i))\) by interchanging \(\lambda(\pi_i)\) and \(\lambda(\pi_j)\). Therefore, Theorem 3.3.1 gives us a way to increase \(\sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i))\).
From (2), it is easy to see that if \( \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i)) > \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i)) \) for any other permutation \((\pi_1, \pi_2, \ldots, \pi_n)\), then for any \(i < j\), either \(\mu(i) + \mu(j) > \lambda(\pi_i) + \lambda(\pi_j)\) with \(\pi_i < \pi_j\) or \(\mu(i) + \mu(j) < \lambda(\pi_i) + \lambda(\pi_j)\) with \(\pi_i > \pi_j\).

Based on the above arguments, we have the following Lemma.

**Lemma 3.5.1** Suppose there exists a unique permutation \((\pi_1, \pi_2, \ldots, \pi_n)\) such that for any \(i < j\), if \(\mu(i) + \mu(j) > \lambda(\pi_i) + \lambda(\pi_j)\), then \(\pi_i < \pi_j\), and if \(\mu(i) + \mu(j) < \lambda(\pi_i) + \lambda(\pi_j)\), then \(\pi_i > \pi_j\). For such \((\pi_1, \pi_2, \ldots, \pi_n)\), it follows that (3.21) holds for any permutation \((\pi_1, \pi_2, \ldots, \pi_n)\).

In general the permutation \((\pi_1, \pi_2, \ldots, \pi_n)\) in Lemma 3.5.2 is not unique. Therefore, in case that the \((\pi_1, \pi_2, \ldots, \pi_n)\) is not unique, then (3.21) does not necessarily hold for such \((\pi_1, \pi_2, \ldots, \pi_n)\). For example, let \(n = 3\), and \(F(x) = \Phi(x)\). Consider

\[
\mu(1) = 1.1\sqrt{2}, \quad \mu(2) = 3.1\sqrt{2}, \quad \mu(3) = 5.1\sqrt{2},
\]
\[
\lambda(1) = \sqrt{2}, \quad \lambda(2) = 3\sqrt{2}, \quad \lambda(3) = 5\sqrt{2}.
\]

Then both permutations \((1, 2, 3)\) and \((3, 1, 2)\) satisfy the conditions in Lemma 3.5.1. However, it is easy to check that for \((\pi_1, \pi_2, \pi_3) = (1, 2, 3)\), (3.21) does not hold. But for \((\pi_1, \pi_2, \pi_3) = (3, 1, 2)\), (3.21) holds.

**Lemma 3.5.2** For any \(\mu_1, \mu_2, \ldots, \mu_n\), we have

\[
\sum_{i=2}^{n} U(\mu(i) - \mu(i-1)) \leq (n-1)U(\frac{\mu(n) - \mu(1)}{n-1}). \tag{3.23}
\]

**Proof.** Let \(x_i = \mu(i+1) - \mu(i)\) for \(i = 1, 2, \ldots, n-1, c = \mu(n) - \mu(1)\), then \(x_i \geq 0\).
and \( \sum_{i=1}^{n-1} x_i = c \). Therefore,

\[
\left( \frac{c}{n-1}, \frac{c}{n-1}, \ldots, \frac{c}{n-1} \right) \preceq (x_1, x_2, \ldots, x_{n-1}).
\]

Since \(-U(x)\) is a convex function on \([0, +\infty)\), we have by Lemma 3.2.3

\[-[U(x_1) + \ldots + U(x_{n-1})] \geq -[U\left( \frac{c}{n-1} \right) + \ldots + U\left( \frac{c}{n-1} \right)],\]

i.e., (3.23) holds. \( \Box \)

Note: By Lemma 3.4.4 and Lemma 3.5.3, we have

\[
\frac{n}{2} \leq U(\mu(1) - \mu(n)) + \sum_{i=2}^{n} U(\mu(i) - \mu(i-1))
\]

\[
\leq U(\mu(1) - \mu(n)) + (n-1)U\left( \frac{\mu(n) - \mu(1)}{n-1} \right).
\]

**Lemma 3.5.3** For any real number \( \xi \), we have

\[
\sum_{i=1}^{n} (\mu(i) - \lambda(i) - \xi)^2 \leq \sum_{i=1}^{n} (\mu(i) - \lambda(\pi_i) - \xi)^2
\]

\[
\leq \sum_{i=1}^{n} (\mu(i) - \lambda(n-i+1) - \xi)^2,
\]

where \( (\pi_1, \pi_2, \ldots, \pi_n) \) is any permutation of \( (1, 2, \ldots, n) \).

**Proof:** Since

\[
\sum_{i=1}^{n} (\mu(i) - \lambda(i) - \xi)^2
\]

\[
= \sum_{i=1}^{n} (\mu(i) - \lambda(i))^2 - 2 \sum_{i=1}^{n} (\mu(i) - \lambda(i))\xi + n\xi^2
\]
and

\[ \sum_{i=1}^{n} (\mu(i) - \lambda(\pi_i) - \xi)^2 \]

\[ = \sum_{i=1}^{n} (\mu(i) - \lambda(\pi_i))^2 - 2 \sum_{i=1}^{n} (\mu(i) - \lambda(\pi_i))\xi + n\xi^2, \]

we only need to show

\[ \sum_{i=1}^{n} (\mu(i) - \lambda(\pi_i))^2 \geq \sum_{i=1}^{n} (\mu(i) - \lambda(i))^2. \]

Since

\[ \sum_{i=1}^{n} (\mu(i) - \lambda(\pi_i))^2 - \sum_{i=1}^{n} (\mu(i) - \lambda(i))^2 \]

\[ = \sum_{i=1}^{n} (\mu^2(i) + \lambda^2(\pi_i)) - 2 \sum_{i=1}^{n} \mu(i)\lambda(\pi_i) \]

\[ - \sum_{i=1}^{n} (\mu^2(i) + \lambda^2(i)) + 2 \sum_{i=1}^{n} \mu(i)\lambda(i) \]

\[ = 2(\sum_{i=1}^{n} \mu(i)\lambda(i) - \sum_{i=1}^{n} \mu(i)\lambda(\pi_i)) \geq 0. \]

i.e., the first inequality holds. Similarly by

\[ \sum_{i=1}^{n} \mu(i)\lambda(n-i+1) \leq \sum_{i=1}^{n} \mu(i)\lambda(\pi_i), \]

we can show the other inequality. \qed

3.6 Matching with Ties Permitted for the Linear Preference Model

Now we consider the case that ties are permitted when we compare \( \Gamma_X \) and \( \Gamma_Y \) in pairs. As in Section 2.6, we introduce the indicator function \( I(\mu, \nu; \tau) \), where
Let $\tau \geq 0$. We call $\tau$ the threshold parameter as before. Therefore, for any matching or permutation $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$, our measure of superiority of $\Gamma_Y$ over $\Gamma_X$, viz. $E[\sum_{i=1}^{n} I(Y(i), X(\pi_i); \tau)]$ is given by

$$ES_\tau(\pi) = E[\sum_{i=1}^{n} I(Y(i), X(\pi_i); \tau)]$$

$$= \sum_{i=1}^{n} P(Y'_(i) > X'(\pi_i) + \tau) + \frac{1}{2} \sum_{i=1}^{n} P(\mid Y'_(i) - X'(\pi_i) \mid \leq \tau)$$

$$= \frac{1}{2} [\sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i) - \tau) + \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_i) + \tau)].$$

For $\pi^o = (1, 2, \ldots, n)$, write

$$V_1^T = ES_\tau(\pi^o).$$

Let $V_2^T$ be $ES_\tau(\pi)$ under random matching. Then, $V_2^T$ can be written as

$$V_2^T = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} E[I(Y'_(i) > X'(j); \tau)]$$

$$= \frac{1}{2n} [\sum_{i=1}^{n} \sum_{j=1}^{n} U(\mu(i) - \lambda(j) - \tau) + \sum_{i=1}^{n} \sum_{j=1}^{n} U(\mu(i) - \lambda(j) + \tau)].$$

Now we are still interested in the relationship between $V_1^T$ and $V_2^T$, and the relationship between $V_1^T$ and $ES_\tau(\pi)$.

**Lemma 3.6.1** If $\mu(i) + \mu(j) \geq \lambda(i) + \lambda(j) + 2\tau$ for any $1 \leq i, j \leq n$, then

(a) $V_1^T \geq V_2^T$.

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

**Proof.** Both (a) and (b) can be shown by using the given condition and Theorem 3.3.1.
Lemma 3.6.2  (a) If $\lambda(i) = \mu(i)$ for $i = 1, 2, \ldots, n$, then

$$ES_{\tau}(\pi) = \frac{n}{2}$$

(3.24)

for any simple matching $\pi$.

(b) If $\lambda(i) = \mu(i)$ and $\lambda(i+1) - \lambda(i) = \lambda(n-i+1) - \lambda(n-i)$ for $i = 1, 2, \ldots, n$ then (3.24) holds for any symmetric matching $\pi$.

Proof. (a) Since for any $i$ and $j$,

$$\left[U(\mu(i) - \lambda(j) - \tau) + U(\mu(i) - \lambda(j) + \tau)\right]$$

$$+ U[(\mu(j) - \lambda(i) - \tau) + U(\mu(j) - \lambda(i) + \tau)]$$

$$= 2.$$

By the definition of simple matching (3.24) follows.

(b) Since

$$ES_{\tau}(\pi) = \frac{1}{4} \left[ \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_{i}) - \tau) + \sum_{i=1}^{n} U(\mu(n-i+1) - \lambda(\pi_{n-i+1}) - \tau) \right]$$

$$+ \frac{1}{4} \sum_{i=1}^{n} U(\mu(i) - \lambda(\pi_{i}) + \tau) + \sum_{i=1}^{n} U(\mu(n-i+1) - \lambda(\pi_{n-i+1}) + \tau)$$

by $\pi_{n-i+1} = n - \pi_{i} + 1$ and the given condition, it is readily shown that (3.24) holds. $\square$


For small $\mu$, we state and prove Theorem 2.2.1 as follows:

**Theorem 2.2.1.** Let $X \sim F(x)$ and $Y \sim F(x - \mu)$ where $X$ is an absolutely continuous rv and $X$ and $Y$ are independent. If $F'(x)$ is bounded, then there exists $\mu_0 > 0$ such that for $|\mu| < \mu_0$ and $i < j$, we have

$$P(Y(i) > X(i)) + P(Y(j) > X(j)) \geq P(Y(i) > X(j)) + P(Y(j) > X(i))$$

(3.25)

according as $\mu \geq 0$

Proof. Let $h(\mu) \equiv P(Y(i) > X(i)) + P(Y(j) > X(j)) - P(Y(i) > X(j)) - P(Y(j) > X(i))$

$$= P(X(i) < X^*(i) + \mu) + P(X(j) < X^*(j) + \mu) - P(X(j) < X^*(i) + \mu) - P(X(i) < X^*(j) + \mu),$$

where $X_{(l)}$ and $X_{(l)}^*$ are iid, $l = i, j$.

Combining the first and third term, and combining the second and the last term, we have

$$h(\mu) = \int_{-\infty}^{\infty} [P(X(i) < x + \mu) - P(X(j) < x + \mu)][f_i(x) - f_j(x)] dx,$$
Consider

\[ U(x, \mu) = [P(X_i < x + \mu) - P(X_j < x + \mu)][f_i(x) - f_j(x)]. \]

Then

\[ \frac{\partial U(x, \mu)}{\partial \mu} = [f_i(x + \mu) - f_j(x + \mu)][f_i(x) - f_j(x)]. \]

Since \( f_i(x) \) and \( f_j(x) \) are density functions and bounded, then there exists \( M > 0 \) such that

\[ \left| \frac{\partial U(x, \mu)}{\partial \mu} \right| \leq M[f_i(x) + f_j(x)]. \]

Therefore, \( \int_{-\infty}^{\infty} \frac{\partial U(x, \mu)}{\partial \mu} \, dx \) is uniformly convergent. Hence \( h'(\mu) \) is continuous and

\[ h'(\mu) = \int_{-\infty}^{\infty} [f_i(x + \mu) - f_j(x + \mu)][f_i(x) - f_j(x)] \, dx. \]

Note that \( h'(0) = \int_{-\infty}^{\infty} [f_i(x) - f_j(x)]^2 \, dx > 0 \) (otherwise we would have \( f_i(x) = f_j(x) \) a.e.). So there exists \( \delta_0 > 0 \) such that \( h'(0) > \delta_0 \). For such \( \delta_0 \), by the continuity of \( h'(\mu) \), there exists \( \mu_0 > 0 \) such that when \( |\mu| \leq \mu_0 \), we have \( |h'(\mu) - h'(0)| < \delta_0 \). Therefore we have \( h'(\mu) > h'(0) - \delta_0 > 0 \). By \( h'(0) = 0 \) and \( h'(\mu) > 0 \), it follows that when \( |\mu| \leq \mu_0 \), (3.25) holds. \( \square \)
APPENDIX B: SOME MATHEMATICAL RESULTS

Here we will derive some mathematical results by using the probability $P(Y(i) > X(j))$. In section 2.2, we express $P(Y(i) > X(j))$ as (2.4) and (2.5).

(a) We first assume that $F(x) = G(x)$. Then we have

$$P(Y(i) > X(j)) = C_{n,i} C_{n,j} \int_0^1 \int_0^u t^{j-1}(1-t)^{n-j} u^{i-1}(1-u)^{n-i} dt du,$$

where $C_{n,i} l = \frac{n!}{(l-1)!(n-l)!}$.

Now, let $i = j$. Then $P(Y(i) > X(j)) = \frac{1}{2}$. Therefore,

$$\int_0^1 \int_0^u t^{j-1}(1-t)^{n-j} u^{i-1}(1-u)^{n-i} dt du = \frac{1}{2} \left[ \frac{(i-1)!(n-i)!}{n!} \right]^2$$

for $i = 1, 2, \ldots, n$.

Since

$$\int_0^u t^{j-1}(1-t)^{n-j} dt = \sum_{k=j}^{n} \binom{n}{k} u^k (1-u)^{n-k},$$

we have

$$P(Y(i) > X(j)) = C_{n,i} \sum_{k=j}^{n} \binom{n}{k} \int_0^1 u^{k+i-1}(1-u)^{2n-k-i} du$$

$$= C_{n,i} \sum_{k=j}^{n} \binom{n}{k} B(k+i, 2n-k-i+1)$$

$$= C_{n,i} \sum_{k=j}^{n} \binom{n}{k} \frac{(k+i-1)!(2n-k-i)!}{(2n)!}.$$
Since \( P(Y(i) > X(j)) = P(X(i) > Y(j)) = 1 - P(Y(j) > X(i)), \) i.e.,

\[
P(Y(i) > X(j)) + P(Y(j) > X(i)) = 1,
\]
it follows that

\[
C_{n,i} \sum_{k=j}^{n} \binom{n}{k} \frac{(k + i - 1)!(2n - k - i)!}{(2n)!} + C_{n,j} \sum_{k=i}^{n} \binom{n}{k} \frac{(k + j - 1)!(2n - k - j)!}{(2n)!} = 1 \tag{3.27}
\]

for any \( 1 \leq i, j \leq n. \)

Let \( i = j. \) By (3.27), we have

\[
\sum_{k=i}^{n} \binom{n}{k} \frac{(k + i - 1)!(2n - k - i)!}{(2n)!} = \frac{1}{2} \frac{(i - 1)!(n - i)!}{n!} \tag{3.28}
\]

for \( i = 1, 2, \ldots, n. \)

In particular, let \( i = 1. \) It follows that

\[
\sum_{k=1}^{n} \frac{(2n - k - 1)!}{(n-k)!} = \frac{(2n - 1)!}{n!} \tag{3.29}
\]

From Lemma 2.3.3., we have

\[
P(Y(j) > X(i)) = P(Y(n-i+1) > X(n-j+1)).
\]

Therefore,

\[
C_{n,j} \sum_{k=i}^{n} \binom{n}{k} \frac{(k + j - 1)!(2n - k - j)!}{(2n)!} = C_{n,i} \sum_{k=n-j+1}^{n} \binom{n}{k} \frac{(n + k - i)!(n - k + i - 1)!}{(2n)!}
\]
for any $1 \leq i, j \leq n$. Or

$$
\sum_{k=i}^{n} \binom{n}{k} \frac{(k + j - 1)!(2n - k - j)!}{j!(n-j)!} = \sum_{k=n-j+1}^{n} \binom{n}{k} \frac{(n + k - i)!(n - k + i - 1)!}{i!(n-i)!}
$$

(3.30)

for any $1 \leq i, j \leq n$.

(b) Assume $F(x) = 1 - e^{-x}$ and $G(x) = F(x - \mu)$, where $\mu \geq 0$. Then it is not difficult to find that

$$
P(Y(n) > X(n)) = \sum_{k=0}^{n} (-1)^{k} \frac{(n!)^{2}}{(n-k)!(n+k)!} e^{-k\mu}.
$$

Let $\mu = 0$. By $P(Y(n) > X(n)) = \frac{1}{2}$, it follows that

$$
\sum_{k=0}^{n} (-1)^{k} \frac{(n!)^{2}}{(n-k)!(n+k)!} = \frac{1}{2}.
$$

(3.31)

Since $Y(n) \overset{d}{=} X'(n) + \mu$, where $X'(n)$ and $X(n)$ are iid, then

$$
P(Y(n) > X(n)) = P(X'(n) + \mu > X(n)).
$$

Write $h_n(\mu) = P(X'(n) + \mu > X(n))$. Clearly, $h_n(\mu)$ is a monotone increasing function, i.e.,

$$
\sum_{k=0}^{n} (-1)^{k} \frac{(n!)^{2}}{(n-k)!(n+k)!} e^{-k\mu}
$$

is a monotone increasing function with respect to $\mu$ on $[0, \infty)$. By the results in Section 3.4, we have

$$
\lim_{n \to \infty} P(Y(n) > X(n)) = \frac{1}{1 + e^{-\mu}},
$$
i.e.,

\[
\lim_{n \to \infty} \sum_{k=0}^{n} (-1)^{k} \frac{(n!)^2}{(n-k)!(n+k)!} e^{-k\mu} = \frac{1}{1 + e^{-\mu}}. \tag{3.32}
\]

If we write \( c = e^{-\mu} \), then

\[
\sum_{k=0}^{n} (-1)^{k} \frac{(n!)^2}{(n-k)!(n+k)!} c^k
\]

is a monotone decreasing function with respect to \( c \) on \([0,1)\), and

\[
\lim_{n \to \infty} \sum_{k=0}^{n} (-1)^{k} \frac{(n!)^2}{(n-k)!(n+k)!} c^k = \frac{1}{1 + c}.
\]

(c) Assume \( F(x) = \frac{1}{1 + e^{-kx}} \), \((k > 0)\) and \( G(x) = F(x - \mu) \). Then

\[
P(Y_n > X_n) = n \int_{-\infty}^{+\infty} F^n(x)G^{n-1}(x)g(x)dx
\]

\[
= n \int_{-\infty}^{+\infty} F^n(x+\mu)F^{n-1}(x)f(x)dx
\]

\[
= n \int_{0}^{+\infty} [1 + at]^{-n}(1 + t)^{-(n+1)}dt,
\]

where \( a = e^{-k\mu} \). By partial integration, we have

\[
P(Y_n > X_n) = \sum_{i=0}^{n-1} \frac{(n+i-1)!(n-i-1)!}{[(n-1)!]^2} a^i \frac{(2n-1)!}{[(n-1)!]^2} \int_{0}^{+\infty} \frac{dt}{(1 + t)(1 + at)^{2n}}.
\]

Since \( P(Y_n > X_n) = \frac{1}{2} \), when \( \mu = 0 \) or \( a = 1 \), we have

\[
\sum_{i=1}^{n} (n+i-2)!(n-i)! = \frac{1}{2} [(n-1)!]^2 + \frac{1}{4} (2n)!.
\]
(d) Assume \( G(x) = F^k(x) \). We have

\[
P(Y_{i+1} > X_{i+1}) - P(Y_i > X_i) = \binom{n}{i} \int_0^1 (i+1-t)^{n-i-1} (1-t)^{i+1} \left[ (n-i)t^k - ik(1-t)t^{k-1} \right] dt
\]

\[
= \binom{n}{i} \sum_{m=0}^{n-i} \binom{n-i}{m} (-1)^m \left[ (n-i) \frac{\Gamma[i(k+1)+mk+1]\Gamma(n-i)}{\Gamma[i(k+1)+mk+n-i+1]} \right. \\
\left. - i k \frac{\Gamma[i(k+1)+mk]\Gamma(n-i+1)}{\Gamma[i(k+1)+mk+n-i+1]} \right]
\]

\[
= \binom{n}{i} \frac{2}{n-i} \sum_{m=0}^{n-i} \binom{n-i}{m} (-1)^m \frac{(i+1)^{m+1}}{m!(n-i-m)!} \frac{n+mk}{n+k(i+m)!}
\]

\[
= \frac{n!}{i!} \frac{2}{n-i} \sum_{m=0}^{n-i} (-1)^m \frac{i+mk}{m!(n-i-m)!} \frac{[i(k+1)+mk-1]!}{[n+k(i+m)]!}
\]

\[(i = 1, 2, \ldots, n - 1).
\]

Taking \( k = 1 \), we have \( P(Y_{i+1} > X_{i+1}) - P(Y_i > X_i) = 0 \). Therefore, it follows that

\[
\sum_{m=0}^{n-i} (-1)^m \frac{i+m}{m!(n-i-m)!} \frac{(2i+m-1)!}{n+m+i)!} = 0 \quad (3.34)
\]

for \( i = 1, 2, \ldots, n - 1 \).

Let \( i = 1 \). We have

\[
\sum_{m=0}^{n-1} (-1)^m \frac{(m+1)^2}{(n-m-1)!(n+m+1)!} = 0.
\]

Taking \( k = m + 1 \), it follows that

\[
\sum_{k=1}^{n} (-1)^k \frac{k^2}{(n-k)!(n+k)!} = 0 \quad (n > 1). \quad (3.35)
\]