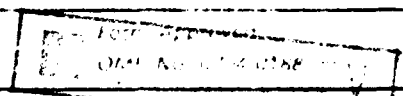


AD-A239 858 TION PAGE



AD-A239 149

2. AUTHOR(S) December 1990 3. REPORT TYPE AND DATES COVERED ~~TECHNICAL~~/DISSERTATION

4. TITLE AND SUBTITLE
A Mathematical Model for a Paired Comparison Experiment on a Continuum of Response

5. FUNDING NUMBERS
①

6. AUTHOR(S)
Kaye A. de Ruiz

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
AFIT Student Attending: University of California-Riverside

8. PERFORMING ORGANIZATION REPORT NUMBER
AFIT/CI/CIA-90-041D

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFIT/CI
Wright-Patterson AFB OH 45433-6583

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION//AVAILABILITY STATEMENT
Approved for Public Release IAW 190-1
Distributed Unlimited
ERNEST A. HAYGOOD, 1st Lt, USAF
Executive Officer

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
OTIC
RECEIVED
DEC 13 1991
G
91-09303

14. SUBJECT TERMS

15. NUMBER OF PAGES
107

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT

18. SECURITY CLASSIFICATION OF THIS PAGE

19. SECURITY CLASSIFICATION OF ABSTRACT

20. LIMITATION OF ABSTRACT

UNIVERSITY OF CALIFORNIA
RIVERSIDE

A Mathematical Model for a Paired Comparison
Experiment on a Continuum of Response

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Applied Statistics

by

Kaye A. de Ruiz

December 1990

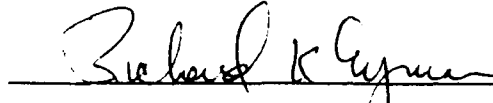
| | |
|---------------|---|
| Accession No. | |
| NTIS G | |
| DIC | |
| Unrestricted | |
| Justification | |
| By | |
| Distribution | |
| Approval | |
| Dist | 3 |
| A-1 | |

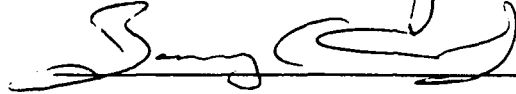


Dissertation Committee:

- Professor Robert J. Beaver, Chairperson
- Professor Barry C. Arnold
- Professor Richard K. Eyman

The Dissertation of Kaye A. de Ruiz is approved:







Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

While earning my doctoral degree, I have received much support from all of the students and staff in the statistics department. I would like to give special thanks to my major professor, Dr. Robert J. Beaver for his unceasing patience and constant help and friendship throughout the research and completion of this dissertation. My thanks also go to Dr. Barry C. Arnold and Dr. Richard K. Eyman for their assistance with the original problem development and with the professional completion of this paper. In addition, I wish to thank Peggy Franklin for her help in long-distance coordination with various university departments and Dale Anderson and Jim Bentley for their assistance in the computer programming aspects of the research. Finally, I would like to thank my parents, Glenn Abbott and Mary Lea Abbott for their support throughout all my college years.

This dissertation is dedicated to my spouse,

Samuel

and to our four sons,

Francisco Xavier

Casildo Glenn Katec

Esteban Mino Zaragoza

and

Qitos Maximilliano

for their infinite love and patient understanding.

ABSTRACT OF THE DISSERTATION

A Mathematical Model for A Paired Comparison
Experiment on a Continuum of Response


by

Kaye A. de Ruiz

Doctor of Philosophy, Graduate Program in Applied
Statistics


University of California, Riverside, December 1990

Professor Robert J. Beaver, Chairperson




This research uses a Bradley-Terry paired comparison model, extending the possible outcomes from a simple binary response to any point along a continuum. An integrated mean squared error approach is used as a basis for parameter estimation based on an underlying logistic model and for estimating a function that maps a response from the real line onto the interval $[0,1]$. Evaluation of the parameter estimates is performed using an underlying logistic sampling distribution with a logistic mapping function. The asymptotic

→ OVER



behavior of the estimates involves the use of linear combinations of order statistics, also called L-statistics. Using this approach it is shown that the parameter estimates are consistent but slightly biased. A simulation study with randomly generated values from a standard logistic distribution were used for response values in a paired comparison experiment. Samples for 3, 4, and 5-treatments with 5, 10, 25, and at times 50, replications for each pairing were considered. The values for the parameter estimates for these samples are presented along with the graphs of the estimated inverse mapping function. Density smoothing techniques using a kernel estimator were employed to further explore the shape of the derivative of the inverse mapping function. This continuous model was then reformulated into a Likert scale with five possible responses from any comparison. An ad hoc method and weighted least squares were used to determine the discriminating values for the Likert scale.



ABSTRACT OF THE DISSERTATION

A Mathematical Model for A Paired Comparison
Experiment on a Continuum of Response

by

Kaye A. de Ruiz

Doctor of Philosophy, Graduate Program in Applied
Statistics

University of California, Riverside, December 1990

Professor Robert J. Beaver, Chairperson

This research uses a Bradley-Terry paired comparison model, extending the possible outcomes from a simple binary response to any point along a continuum. An integrated mean squared error approach is used as a basis for parameter estimation based on an underlying logistic model and for estimating a function that maps a response from the real line onto the interval $[0,1]$. Evaluation of the parameter estimates is performed using an underlying logistic sampling distribution with a logistic mapping function. The asymptotic

behavior of the estimates involves the use of linear combinations of order statistics, also called L-statistics. Using this approach it is shown that the parameter estimates are consistent but slightly biased. A simulation study with randomly generated values from a standard logistic distribution were used for response values in a paired comparison experiment. Samples for 3, 4, and 5 treatments with 5, 10, 25, and at times 50, replications for each pairing were considered. The values for the parameter estimates for these samples are presented along with the graphs of the estimated inverse mapping function. Density smoothing techniques using a kernel estimator were employed to further explore the shape of the derivative of the inverse mapping function. This continuous model was then reformulated into a Likert scale with five possible responses from any comparison. An ad hoc method and weighted least squares were used to determine the discriminating values for the Likert scale.

TABLE OF CONTENTS

| | | Page |
|-----------|---|------|
| Chapter 1 | Introduction and Background | 1 |
| Chapter 2 | Basic Model | 13 |
| | 2.1 Definition and Notation | 13 |
| | 2.2 Estimation of Parameters (without order) | 18 |
| | 2.3 Basic Model with order parameter | 20 |
| Chapter 3 | Use of Logistic Distribution in Basic Model | 24 |
| | 3.1 Truncated Distribution Approach | |
| | 3.1.1 Case I: All Ordered Pairs | 25 |
| | 3.1.2 Case II: Non Ordered Pairs | 32 |
| | 3.2 Limited Integral Approach | 37 |
| | 3.3 Model with an Order Parameter | 40 |
| Chapter 4 | Properties of the Minimum Mean Square Error Parameter Estimates Involving Order Statistics | 43 |
| | 4.1 Properties of μ_i using Inverse Logistic Function | 43 |
| | 4.2 Application Using a Logistic Mapping Function | 55 |
| | 4.3 Properties of γ_{ij} using Logistic Inverse | 64 |
| Chapter 5 | Simulation Study Using Parameter Estimates | 68 |
| | 5.1 General Approach | 68 |
| | 5.2 Results on $\hat{\mu}_i$ | 69 |
| | 5.3 Results on $\hat{\phi}^{-1}$ and density estimation procedures for $\frac{d}{du}\phi^{-1}(u)$ | 73 |
| Chapter 6 | Application to the Likert Scale | 88 |
| | 6.1 Introduction | 88 |
| | 6.2 Ad Hoc Determination of θ_i 's | 90 |
| | 6.3 Weighted Least Squares Determination of θ_i 's | 95 |

LIST OF FIGURES

| | | Page |
|-------------|---|------|
| Figure 5.1 | Phi Inverse with 3 Treatments and 25 Replications | 74 |
| Figure 5.2 | Phi Inverse with 4 Treatments and 25 Replications | 75 |
| Figure 5.3 | Phi Inverse with 5 Treatments and 25 Replications | 75 |
| Figure 5.4 | Normal Density Smoothing, Bandwidth=0.1, 10 Replications | 77 |
| Figure 5.5 | Normal Density Smoothing, Bandwidth=0.2, 10 Replications | 77 |
| Figure 5.6 | Normal Density Smoothing, Bandwidth=0.3, 10 Replications | 78 |
| Figure 5.7 | Normal Density Smoothing, Bandwidth =0.4, 10 Replications | 78 |
| Figure 5.8 | Normal Density Smoothing, Bandwidth=0.5, 10 Replications | 79 |
| Figure 5.9 | Normal Density Smoothing, Bandwidth=1.0, 10 Replications | 79 |
| Figure 5.10 | Normal Density Smoothing, Bandwidth=0.1, 25 Replications | 80 |
| Figure 5.11 | Normal Density Smoothing, Bandwidth =0.2, 25 Replications | 80 |
| Figure 5.12 | Normal Density Smoothing, Bandwidth' =0.3, 25 Replications | 81 |
| Figure 5.13 | Normal Density Smoothing, Bandwidth=0.4, 25 Replications | 81 |

| | | |
|-------------|--|----|
| Figure 5.14 | Normal Density Smoothing, Bandwidth=0.5, 25 Replications | 82 |
| Figure 5.15 | Normal Density Smoothing, Bandwidth =1.0, 25 Replications | 82 |
| Figure 5.16 | Normal Density Smoothing, Bandwidth=0.1, 50 Replications | 83 |
| Figure 5.17 | Normal Density Smoothing, Bandwidth=0.2, 50 Replications | 83 |
| Figure 5.18 | Normal Density Smoothing, Bandwidth=0.3, 50 Replications | 84 |
| Figure 5.19 | Normal Density Smoothing, Bandwidth =0.4, 50 Replications | 84 |
| Figure 5.20 | Normal Density Smoothing, Bandwidth=0.5, 50 Replications | 85 |
| Figure 5.21 | Normal Density Smoothing, Bandwidth =1.0, 50 Replications | 85 |

LIST OF TABLES

| | | Page |
|-----------|---|------|
| Table 5.1 | Comparison of 3 Items | 70 |
| Table 5.2 | Comparison of 4 Items | 71 |
| Table 5.3 | Comparison of 5 Items | 71 |
| Table 5.4 | Comparison of 5 Items (different preset μ_i 's) | 72 |
| Table 6.1 | ϕ^{-1} and θ Estimates for Five Parameters | 93 |
| Table 6.2 | Average $\hat{\theta}_i$ Values | 94 |
| Table 6.3 | Theta Values | 104 |

CHAPTER 1

INTRODUCTION AND BACKGROUND

Paired comparison experiments arise naturally through many types of competitions, through sensory comparisons of items observed in pairs or through tournaments, or through combinations associated with the development of experimental designs. The method of paired comparisons is generally used when items can be judged only subjectively, and it is impractical to measure the responses on a ratio or interval scale to determine the preference ordering. A basic paired comparison experiment consists of treatments, $T_1 \dots T_k$, ($k \geq 2$) taken in pairs, with n_{ij} independent comparisons of T_i with T_j , $1 \leq i < j \leq k$. The worths associated with these treatments are assumed to lie along some underlying continuum of worth. A judge determines a preference for one of the two treatments in a pair, and thereby ranks the treatments or objects. This determination can be based upon the intensity of the sensory perception such as taste, feel, or sight, or upon

the closeness of the sensation to a perceived optimum level. By compiling the results of all n_{ij} selections, a final ranking of all treatments can be obtained (Bradley, 1976).

The method of paired comparisons dates back to at least 1860, when Fechner wrote of his account of judging which of two vessels felt heavier under a variety of circumstances. Although there were some objections to the methods he used in conducting his experiments, Fechner's model, which was based on a normal distribution, was very similar to the one used later by Mosteller in 1951 (David, 1969).

Thurstone (1927) however, was the first to model paired comparisons using the idea of a subjective psychological continuum whereby items could be discriminated and ranked in order of preference. Thurstone's law of comparative judgement can be applied to a range of situations such as comparisons of physical stimulus intensities (weight or handwriting specimens), qualitative comparative judgement (excellence of papers in an educational setting), or measurement

of psychological values (opinions on a set of public issues). Each judge employs a 'discriminal process' or method of distinguishing between subjects or items to determine the excellence or value of the subjects. The judges opinion may fluctuate from one replication to the next when judging an item and this leads to the formation of a normal distribution along a continuum for each item being compared. When the difference between the means of any two stimuli is large then it is easy to rank order the items. If, however, the means of two items distribution are close then it is difficult to distinguish between the two items. Using this basic structure, Thurstone presented five different cases and models which depend on assumptions concerning the variance and correlation structure among the treatments and stimuli.

Mosteller (1951) formalized Thurstone's contributions by developing a linear model for paired comparisons. Here, the intrinsic worth or utility of each treatment, T_i , is located along a continuum at a point S_i . The subject receives a

sensation in response to the treatment and can order the two treatments based on these sensations. This gave the probability of T_i being selected over T_j expressed as

$$\begin{aligned}
 P(T_i \rightarrow T_j) &= P[d_{ij} > 0] \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-(\mu_i - \mu_j)}^{\infty} e^{-y^2/2} dy,
 \end{aligned}
 \tag{1.1.1}$$

where $E(d_{ij}) = \mu_i - \mu_j$ without loss of generality, $d_{ij} \sim N(\mu_i - \mu_j, 1)$ and $1 \leq i \leq j \leq k$. This is based on the assumption that the sensations from a stimulus for a population of individuals are normally distributed.

Bradley and Terry (1952) proposed a slightly different model for paired comparisons of the k treatments. The nonnegative treatment parameters $\Pi_1, \Pi_2, \Pi_3 \dots \Pi_k$, associated with the k treatments under investigation, could be represented on an underlying continuum of worth, and identified only up to a scale parameter. Each comparison can be considered as a Bernoulli trial with

$$P(T_i \rightarrow T_j) = \Pi_{ij} = \frac{\Pi_i}{\Pi_i + \Pi_j} \quad (1.1.2)$$

where $\Pi_i \geq 0$ and $\sum \Pi_i = 1$ for $1 \leq i \leq j \leq k$.

Equivalently this can be written as

$$P(T_i \rightarrow T_j) = 1/4 \int_{-(\ln \mu_i - \ln \mu_j)}^{\infty} \text{sech}^2(y/2) dy. \quad (1.1.3)$$

When $\Pi_i = e^{-\mu_i}$ the Bradley-Terry model can be obtained by substituting the logistic density for the normal density in Thurstone's model. Parameters can be estimated by the method of maximum likelihood whereby the maximum likelihood estimates satisfy the system of equations

$$\frac{a_i}{p_i} - \sum_{j(\neq i)} \frac{n_{ij}}{(p_i + p_j)} = 0 \quad (1.1.4)$$

$i=1,2,\dots,t$, subject to the constraint $\sum \Pi_i = 1$. Here a_i is the number of selections favoring T_i , and p_i is the estimator for Π_i with $\sum p_i = 1$. Bradley and Terry solved these equations iteratively, and used the likelihood ratio principle in testing hypotheses. Conditions for the existence of solutions to (1.1.3) were given by Ford (1957). For testing the hypothesis of

equal preferences, Beaver (1974) proposed an alternative test based on a locally asymptotic most stringent criterion.

In his book on individual choice behavior, Luce (1959) established a choice axiom using basic probability principles. This axiom provides a method for determining the probability that one item is selected from a fixed set of items. He defines U as the universal set of possible alternatives from which a subject must be able to make a preference for one of the items. Further, T is a finite subset of U and must contain the chosen item. The notation $P_T(S)$ denotes the probability that the selected item lies in the subset S where S is a subset of T . $P(x,y)$ is written to represent $P_{\{x,y\}}(x)$ when $x \neq y$.

AXIOM 1. Let T be a finite subset of U such that for every $S \subset T$, P_S is defined

i) $P(x,y) \neq 0,1$ for all $x,y \in T$, then for RCSCS, $P_T(R) = P_S(R)P_T(S)$

ii) If $P(x,y) = 0$ for some $x,y \in T$ then for every $S \subset T$, $P_T(S) = P_{T-\{x\}}(S-\{x\})$.

Using the logistic model Luce goes on to apply this choice axiom to a variety of psychological areas. After comparing this axiom with Thurstone's discriminial process, Luce concludes that although the two are "logically distinct", the assumptions are "extremely similar for all practical purposes".

Following the establishment of these basic models, theorists studied various other areas. The effect of ties or a no preference judgement was considered by Rao and Kupper (1967) who extended the Bradley-Terry model to allow for the occurrence of such ties. They introduced a parameter, η , interpreted as a threshold of sensory discrimination for the judge. Whenever the estimate of the difference in response for two objects is less than η , a tie results. When $\eta=0$, the model reduces to that of Bradley-Terry. Davidson (1970) went further in studying the effect of ties by developing a parameter that depends only on the net performance of the judges and not on the performance of an individual paired comparison.

Another extension of the Bradley-Terry Model

involves the consideration of order effects. In psychological experiments the effect of order of presentation may be of more importance than the actual effects of the stimuli. An additive order effect parameter was introduced by Beaver and Gokhale (1975) to account for the effects of order in the presentation of stimuli. A common parameter δ_{ij} for the ordered pairs (i,j) and (j,i) would inflate the probability of one item being preferred, while deflating the probability of the second item being preferred. Since the order parameter varies from pair to pair, the number of order effect parameters becomes quite large, and the model awkward to analyze.

Later, based on the extended Rao-Kupper and Davidson models, a multiplicative order effect model was introduced by Davidson and Beaver (1977). Since the Bradley-Terry Model uses parameters linear on the log-scale, it is natural to consider such a multiplicative effect. They proposed a parameter γ , independent of the individual pairs. When $\gamma=1$ there is no order effect, while $\gamma>1$ inflates the worth of the second stimuli and $\gamma<1$ deflates the worth of the second

object. This changes the probabilities of preference from the original model to account for the order of presentation of the items within a pair. The probability that the first item is preferred when given (i, j) is expressed as

$$P(i \rightarrow j | i, j) = \frac{\Pi_i}{\Pi_i + \gamma \Pi_j} \quad (1.1.5)$$

while probability that the second object is preferred is given as

$$P(j \rightarrow i | i, j) = \frac{\gamma \Pi_j}{\Pi_i + \gamma \Pi_j} \quad (1.1.6)$$

for $1 \leq i \neq j \leq k$.

Finally, a model for triple comparisons was developed in detail along with the necessary asymptotic theory, by Pendergrass and Bradley (1960) and Beaver and Rao (1972). More recent studies have considered all possible pairs including comparison of identical items to help estimate order effects and tie parameters (Sirotnik and Beaver, 1984), as well as various applications to the design of paired comparison experiments. Such applications include mixture

problems (Charnet and Beaver, 1988) and factorial paired comparisons in experiments (El-Helbawy and Ahmed, 1984).

Little or no work has been performed that considers more than a simple binary response ($y=0$ or 1) on the Bradley-Terry model. A paper by Tutz (1986) does extend this model to allow for the strength of preference for a respondent to decide whether the treatments differ greatly or only slightly. Tutz uses an ordinal response paired comparison system with k response categories where the parameters $\theta_1 \dots \theta_k$ (determined uniquely) characterize the discrimination power of the response categories. Tutz claims that his generalized Bradley-Terry model is more precise than the original model. This becomes important if the value of a future response is to be predicted.

The present study extends this approach to develop a general mathematical model relating the sensory perceptions of judges via an arbitrary function to a paired comparison model of the Bradley-Terry type. This unknown function maps a

response from a continuous scale over the real numbers to a value between 0 and 1. In Chapter 2 the necessary definitions and the basic model that will be used are introduced. A minimum mean squared error criteria is used in determining estimates of the parameters in a reasonable general context. Specific development of the estimators of the parameters as they relate to the logistic distribution is explored in Chapter 3. Here we consider both a truncated logistic distribution and a limited integral approach to the problem. In each of these cases, we examine the estimates using data obtained from all ordered pairs (i,j) and (j,i) , as well as data that considers only the pairs (i,j) with $1 \leq i < j \leq k$. In Chapter 4 we continue to consider the properties of the estimators that involve linear combinations of order statistics or L-statistics. Here, we first consider a general form for the mean of the L-statistics and a bound on the variance term using the underlying logistic distribution. Then we provide an example using a logistic mapping function within the framework of the logistic distribution. The accuracy of the parameter

estimates is explored in Chapter 5, through simulation, while in Chapter 6 we apply the results of the study to measurements on the Likert scale technique.

CHAPTER 2

DEVELOPMENT OF THE BASIC MODEL

2.1 Definitions and Notation

In the development of a model that generalizes the 0-1 binary response of the Bradley-Terry Model to one lying in the interval $(0,1)$, the definitions and notation which follow are convenient.

Definition 1: Let $U_1 \dots U_k$ be random errors associated with a subject's perception to the k items. The variables $U_1 \dots U_k$ are assumed to be independently and identically distributed, centered at 0 with cumulative distribution function (cdf) $G(\cdot)$ which may be known or unknown. The difference $U_i - U_j$ has cdf F , centered at zero.

Definition 2: The random variable

$V_{ij} = (U_i - U_j) + (\mu_i - \mu_j)$, which is the perceived difference in the worth of items i and j , has location parameter $(\mu_i - \mu_j)$ and cdf F centered at $\mu_i - \mu_j$. The unknown parameters μ_i and μ_j

are used to define the relative worths of the i th and j th items.

Definition 3:

The response elicited is $Y_{ij} = \phi(V_{ij})$ where ϕ is a monotonic, increasing, continuous function that maps the real numbers onto the interval $[0,1]$.

Throughout this study we will use ordered pairs so that the pair (i,j) is different than the pair (j,i) with the first member of the pair being presented first to a judge, $1 \leq i < j \leq k$. The function $F_{ij}(y)$ will denote the cdf of the random variable Y_{ij} , the judges response to the ordered pair (i,j) while $F_{ji}(y)$ denotes the cdf of the random variable Y_{ji} , comparison of the pair (j,i) .

The estimators of the function ϕ and the unknown parameters $\mu_1 \dots \mu_k$, are determined using a minimum mean squared error criterion. We begin by observing that the cdf of Y_{ij} is related to the cdf of V_{ij}

$$\begin{aligned}
F_{ij}(y) &= P(Y_{ij} < y) = P(\phi(V_{ij}) < y) \\
&= P(V_{ij} < \phi^{-1}(y)) \\
&= F[\phi^{-1}(y) - (\mu_i - \mu_j)] \quad (2.1.1)
\end{aligned}$$

where F is the distribution function of $d_{ij} = U_i - U_j$. Let $F_{ij}^{(n)}(y)$ be the sample estimate of $F_{ij}(y)$, which is equal to $F(\phi^{-1}(y) - (\mu_i - \mu_j))$ the true cdf of the random variable Y_{ij} . For ease of presentation, we will let $n_{ij} = n$ for all pairs (i, j) so that there are an equal number of comparisons for each pair. The estimates of ϕ and $\underline{\mu}' = (\mu_1, \mu_2, \dots, \mu_k)$ are those which minimize the difference between $F_{ij}(y)$ and $F_{ij}^{(n)}$, which is equivalent to minimizing $F(\phi^{-1}(y) - (\mu_i - \mu_j))$ and $F^{-1}(F_{ij}^{(n)}(y))$ for $1 \leq i < j \leq k$. However, we will minimize the difference between $F^{-1}(F(\phi^{-1}(y) - \mu_i - \mu_j))$ and $F^{-1}F_{ij}^{(n)}$, or simplified to $(\phi^{-1}(y) - (\mu_i - \mu_j))$ and $F^{-1}F_{ij}^{(n)}$. If we define $f_{ij}(y) = F^{-1}(F_{ij}^{(n)}(y))$, then the criterion to be minimized is given by

$$Q = \sum_{i \neq j}^k \int_0^1 [\phi^{-1}(y) - (\mu_i - \mu_j) - f_{ij}(y)]^2 dy \quad (2.1.2)$$

By expanding Q we have

$$\begin{aligned}
Q &= k(k-1) \int_0^1 [\phi^{-1}(y)]^2 dy + \sum_{i \neq j}^k (\mu_i - \mu_j)^2 + \\
&\sum_{i \neq j}^k \int_0^1 f_{ij}^2(y) dy - 2 \sum_{i \neq j}^k (\mu_i - \mu_j) \int_0^1 \phi^{-1}(y) dy - \\
&2 \sum_{i \neq j}^k \int_0^1 \phi^{-1}(y) f_{ij}(y) dy + 2 \sum_{i \neq j}^k (\mu_i - \mu_j) \int_0^1 f_{ij}(y) dy
\end{aligned}$$

where $\sum_{i \neq j}^k$ is interpreted as the double sum over both i and j excluding terms for which $i=j$.

Minimizing Q involves taking derivatives with respect to the unknown parameters, setting these derivatives equal to 0. Solving the resulting equations provides the parameter estimates which follow:

$$\begin{aligned}
\frac{dQ}{d\mu_i} &= 2 \sum_{j(\neq i)}^k (\mu_i - \mu_j) - 2 \sum_{j(\neq i)}^k (\mu_j - \mu_i) \\
&+ 2 \sum_{j(\neq i)}^k \int_0^1 (f_{ij}(y) - f_{ji}(y)) dy.
\end{aligned}$$

Then,

$$\begin{aligned}
0 &= \sum_{j(\neq i)}^k (\hat{\mu}_i - \mu_j) - \sum_{j(\neq i)}^k (\hat{\mu}_j - \mu_i) \\
&+ \sum_{j(\neq i)}^k \int_0^1 (f_{ij}(y) - f_{ji}(y)) dy
\end{aligned}$$

for $i=1,2,\dots,k$. Using the constraint $\sum_{j=1}^k \mu_j = 0$,

this equation can be written as

$$0 = 2k\mu_i + \sum_{j(\neq i)}^k \int_0^1 (f_{ij}(y) - f_{ji}(y)) dy$$

and

$$\hat{\mu}_i = \frac{1}{2k} \sum_{j(\neq i)}^k \int_0^1 [f_{ji}(y) - f_{ij}(y)] dy, \quad (2.1.3)$$

$i=1,2,\dots,k.$

In solving for $\hat{\phi}^{-1}(y)$ we rewrite Q as

$$Q = \int_0^1 \sum_{i \neq j}^k \{ \phi^{-1}(y) - f_{ij}(y) - \delta_{ij} \}^2 dy \quad \text{which is}$$

minimized whenever the summation is minimized.

This follows from the fact that $\sum_j \sum_j (x_{ij} - a)^2$ is minimized for $a = \bar{x}$. Therefore we consider

$$\begin{aligned} \phi^{-1}(y) &= \sum_{i \neq j}^k \frac{f_{ij}(y)}{k(k-1)} + \sum_{i \neq j}^k \frac{(\mu_i - \mu_j)}{k(k-1)} \\ &= \sum_{i \neq j}^k \frac{f_{ij}(y)}{k(k-1)} \end{aligned} \quad (2.1.4)$$

Throughout the study the assumption is made that in the experiment in which we compare item i with item j , $P(j \rightarrow i) = 1 - P(i \rightarrow j)$. This in

turn leads to the relationship which will prove useful in future results

$$F^{-1}(F_{ij}(y)) = F^{-1}(1 - F_{ji}(1-y)). \quad (2.1.5)$$

For each pair (i,j) we have n responses on [0,1]. By using the order statistics and the associated values of $F_{ij}^{(n)}(y)$, the integral in 2.1.3 simply becomes the area under a step function with steps of $1/n$ at the values of the order statistics Y_1^{ij} , Y_2^{ij} , ... Y_n^{ij} . This leads to

$$\begin{aligned} \int_0^1 F^{-1}(F_n(y)) dy = & F^{-1}(0) [Y_{1:n}^{ij} - 0] + F^{-1}(1/n) [Y_{2:n}^{ij} - Y_{1:n}^{ij}] \\ & + F^{-1}(2/n) [Y_{n:3}^{ij} - Y_{n:2}^{ij}] + \dots + \\ & F^{-1}(n-1/n) [Y_{n:n}^{ij} - Y_{n-1:n}^{ij}] + F^{-1}(n/n) [1 - Y_{n:n}^{ij}]. \end{aligned} \quad (2.1.6)$$

2.2 Estimation of Parameters (without order)

Using the form given in the previous section, the estimators now appear as

$$\begin{aligned} \hat{\mu}_i = \frac{1}{2k} \sum_{j(\neq i)}^k \{ & F^{-1}(0) [Y_{1:n}^{ji} - 0] + \\ & F^{-1}(1/n) [Y_{2:n}^{ji} - Y_{1:n}^{ji}] + \dots + \\ & F^{-1}(\frac{n-1}{n}) [Y_{n:n}^{ji} - Y_{n-1:n}^{ji}] + F^{-1}(1) [1 - Y_{n:n}^{ji}] \} \end{aligned}$$

$$-F^{-1}(0) [Y_{1:n}^{ji} - 0] - F^{-1}(\frac{1}{n}) [Y_{2:n}^{ji} - Y_{1:n}^{ji}] - \dots -$$

$$F^{-1}(\frac{n-1}{n}) [Y_{n:n}^{ij} - Y_{n-1:n}^{ij}] - F^{-1}(1) [1 - Y_{n:n}^{ij}] \}$$

and

$$\hat{\phi}^{-1}(y) = \frac{1}{k(k-1)} \sum_{j (\neq i)}^k [F^{-1}(F_{ij}^{(n)}(y)) + F^{-1}(F_{ji}^{(n)}(y))]. \quad (2.2.1.)$$

The estimates of μ_1, \dots, μ_k and $\phi(\cdot)$ are based on the assumption that both the pair (i, j) and the pair (j, i) are sampled and both sets of data are available. If, however, the sampling is done only on (i, j) the model is taken to be reflective. In other words, we will employ the reasoning that $\phi^{-1}(u) = -\phi^{-1}(1-u)$ and therefore that $\phi^{-1}(y_{ji}) = \phi^{-1}(1-y_{ij})$. Recall from 2.1.5 that, in addition, $F^{-1}(F_{ji}(y)) = F^{-1}(1 - F_{ij}(1-y))$.

For this reflective case then, where the data is used twice (once for y_{ij} and once for $y_{ji} = 1 - y_{ij}$) the parameter estimates become:

$$\hat{\mu}_i = \frac{1}{2k} \sum_{(j \neq i)}^k \{ F^{-1}(0) [1 - Y_{n:n}^{ij}] + F^{-1}(\frac{1}{n}) [1 - Y_{n-1:n}^{ij}]$$

$$- (1 - Y_{n:n}^{ij}) \} + \dots + F^{-1}(\frac{n-1}{n}) [1 - Y_{1:n}^{ij} - (1 - Y_{2:n}^{ij})] +$$

$$F^{-1}(1) [1 - (1 - Y_{1:n}^{ij})] - F^{-1}(0) [Y_{1:n}^{ij} - 0] - \dots$$

$$- F^{-1}(\frac{n-1}{n}) [Y_{n:n}^{ij} - Y_{n-1:n}^{ij}] - F_{ij}^{-1}(1) [1 - Y_{n:n}^{ij}] \}$$

$$\begin{aligned}
&= \frac{1}{2k} \sum_{j(\neq i)}^k \{F^{-1}(0) - F^{-1}(1) + \\
&\quad Y_{1:n}^{ij} [F^{-1}(1) - F^{-1}(\frac{n-1}{n}) - F^{-1}(0) + F^{-1}(\frac{1}{n})] \\
&\quad Y_{2:n}^{ij} [F^{-1}(\frac{n-1}{n}) - F^{-1}(\frac{n-2}{n}) + F^{-1}(\frac{2}{n}) - F^{-1}(\frac{1}{n})] + \\
&\quad \dots + Y_{n:n}^{ij} [F^{-1}(\frac{1}{n}) - F^{-1}(0) + F^{-1}(1) - F^{-1}(\frac{n-1}{n})]\}.
\end{aligned}$$

$$\hat{\phi}^{-1}(y) = \frac{1}{k(k-1)} \sum_{i < j}^k F^{-1}(F_{ij}^{(n)}(y)) \tag{2.2.2}$$

2.3 Basic model with order parameter

With an order effect (γ_{ij}) added to the model, (see David and Beaver), we return to the minimum mean square error criterion to determine the parameter estimates for μ_i , $\phi^{-1}(y)$ and γ_{ij} .

$$Q = \sum_{i \neq j}^k \int_0^1 \{ \phi^{-1}(y) - (\mu_i - \mu_j + \gamma_{ij}) - F^{-1}(F_{ij}^{(n)}(y)) \}^2 dy \tag{2.3.1}$$

Expanding (2.3.1) leads to

$$\begin{aligned}
Q &= k(k-1) \int_0^1 [\phi^{-1}(y)]^2 dy + \sum_{j \neq i}^k (\mu_i - \mu_j + \gamma_{ij})^2 + \\
&\quad \sum_{i \neq j}^k f_{ij}^2 dy - 2 \sum_{i \neq j}^k (\mu_i - \mu_j + \gamma_{ij}) \int_0^1 \phi^{-1}(y) dy \\
&\quad - 2 \sum_{i \neq j}^k \int_0^1 \phi^{-1}(y) f_{ij}(y) dy + 2 \sum_{i \neq j}^k (\mu_i - \mu_j + \gamma_{ij}) \int_0^1 f_{ij}(y) dy.
\end{aligned}$$

Again, $\sum_{j \neq i}$ represents the double sum over both i and j , while $\sum_{j(\neq i)}$ represents the sum over j alone. Since the parameters μ_1, \dots, μ_k can be estimated only up to a scale parameter, we also use the condition again that $\sum_{j=1}^k \mu_j = 0$. For this study we also use the condition that $\gamma_{ij} = -\gamma_{ji}$. In minimizing Q , we find

$$\begin{aligned} \frac{dQ}{d\mu_i} = & 2 \sum_{j(\neq i)}^k (\mu_i - \mu_j + \gamma_{ij}) - 2 \sum_{j(\neq i)}^k (\mu_j - \mu_i - \gamma_{ij}) \\ & + 2 \sum_{j(\neq i)}^k \int_0^1 (f_{ji}(y) - f_{ij}(y)) dy. \end{aligned}$$

which, when set equal to zero leads to the estimate for μ_i given as:

$$\hat{\mu}_i = \frac{1}{2k} \left[-2\hat{\gamma}_{ij} + \sum_{j(\neq i)}^k \int_0^1 (f_{ji}(y) - f_{ij}(y)) dy \right].$$

$i=1, 2, \dots, k.$

(2.3.2)

The estimate for $\phi^{-1}(y)$ is identical to that of the basic model (2.1.4) when we assume that $\gamma_{ij} = -\gamma_{ji}$ because of the cancellation of these terms from the equation.

The estimate for γ_{ij} is found as shown the solution to the equation

$$\frac{dQ}{d\gamma_{ij}} = 2(\mu_i - \mu_j + \gamma_{ij}) - 2(\mu_j - \mu_i - \gamma_{ij}) +$$

$$2 \left[\int_0^1 f_{ij}(y) dy - \int_0^1 f_{ji}(y) dy \right] = 0,$$

from which we obtain,

$$\hat{\gamma}_{ij} = -\hat{\delta}_{ij} + \frac{1}{2} \int_0^1 (f_{ji}(y) dy - f_{ij}(y) dy)$$

(2.3.3)

with $\delta_{ij} = \mu_i - \mu_j$, $1 \leq i, j \leq k$. For the design in which all ordered pairs are used, the estimates for μ_i and γ_{ij} can be written as:

$$\hat{\mu}_i = \frac{1}{2k} \left(-2\hat{\gamma}_{ij} + \sum_{j(\neq i)}^k \left\{ F^{-1}(0) [Y_{1:n}^{ji} - 0] + \right.$$

$$F^{-1}\left(\frac{1}{n}\right) [Y_{2:n}^{ji} - Y_{1:n}^{ji}] + \dots +$$

$$F^{-1}(1) [1 - Y_{n:n}^{ji}] - F^{-1}(0) [Y_{1:n}^{ij} - 0]$$

$$\left. - \dots - F^{-1}(1) [1 - Y_{n:n}^{ij}] \right\} \right)$$

(2.3.4)

$$\hat{\gamma}_{ij} = -\hat{\delta}_{ij} + \frac{1}{2} \left(\sum_{j(\neq i)}^k \left\{ F^{-1}(0) [Y_{1:n}^{ji} - 0] + \right.$$

$$F^{-1}\left(\frac{1}{n}\right) [Y_{2:n}^{ji} - Y_{1:n}^{ji}] + \dots + F^{-1}(1) [1 - Y_{n:n}^{ji}]$$

$$\left. - F^{-1}(0) [Y_{1:n}^{ij} - 0] - \dots - F^{-1}(1) [1 - Y_{n:n}^{ij}] \right\} \right)$$

(2.3.5)

In the case of distinct pairs, the estimators in 2.3.4 and 2.3.5 can be simplified somewhat (see 2.2.2). The properties of these estimators will be addressed in Chapter 3, as they pertain to specific distributions.

CHAPTER 3

USE OF LOGISTIC DISTRIBUTION IN BASIC MODEL

3.1 Truncated Distribution Approach

The logistic distribution is commonly used in the study of paired comparison experiments. In its simplest representation, the logistic distribution is given as

$$F(x) = \frac{1}{1+e^{-x}} \quad \text{for } x \in \mathcal{R} \quad (3.1.1)$$

with density

$$\begin{aligned} f(x) &= \frac{e^{-x}}{(1+e^{-x})^2} \\ &= F(x)(1-F(x)) \end{aligned} \quad (3.1.2)$$

In order to obtain finite values for all terms in the parameter estimates given in equations 2.3.1 and 2.2.2, it is necessary to limit the range over which the logistic random variable is defined. One method involves truncating the distribution; a second method, discussed in Section 3.2, involves limiting the range of integration which appears in the estimates. Without some type of truncation or

censoring, the inverse mapping will involve $F^{-1}(0)=-\infty$ and $F^{-1}(1)=+\infty$, and hence, give rise to a divergent integral. The symmetrically truncated standard logistic distribution takes the form

$$F(x) = \frac{\frac{1}{1+e^{-x}} - \frac{1}{1+e^{\alpha}}}{\frac{1}{1+e^{-\alpha}} - \frac{1}{1+e^{\alpha}}} \quad -\alpha \leq x \leq \alpha < \infty$$

(3.1.3)

where α is the upper truncation point and $-\alpha$ is the lower truncation point. We assume that α is a known parameter. If α is not known, for all practical purposes one can take α to be a large number, say 10, so that the interval $(-\alpha, \alpha)$ contains more than 99% of the distribution. The truncation occurs for the variable v_{ij} so that the domain $-\infty < v_{ij} < +\infty$, is replaced by $-\alpha < v_{ij} < \alpha$ which then implies that the range of $U_i - U_j$ is given by $-\alpha - \delta_{ij} < U_i - U_j < \alpha - \delta_{ij}$.

3.1.1 CASE I: ALL ORDERED PAIRS

Case I is used to denote the paired comparison experiment involving all ordered pairs so that data is collected on all possible combinations of treatments (i, j) and (j, i) ,

$1 \leq i, j \leq k$. This sampling design allows for estimating order effects that might be present in the data. Case II, in which data is collected using only distinct pairs (i, j) $1 \leq i < j \leq k$, and then used again as data for the pair (j, i) by a simple reflection technique, is discussed in Section 3.1.2. The inverse function, $F^{-1}(\cdot)$, is needed to estimate the parameters in the basic model given by equations 2.2.1 and 2.2.2. To solve for this inverse function we observe that $V_{ij} = \phi^{-1}(Y_{ij})$. Let $p = F(y)$ and solve for y as follows:

$$p = \frac{\frac{1}{1+e^{-y}} - \frac{1}{1+e^{\alpha}}}{\frac{1}{1+e^{-\alpha}} - \frac{1}{1+e^{\alpha}}}$$

$$\frac{p(e^{\alpha} - e^{-\alpha})}{(1+e^{\alpha})(1+e^{-\alpha})} = \frac{1}{1+e^{-y}} - \frac{1}{1+e^{\alpha}}$$

$$\frac{p(e^{\alpha} - e^{-\alpha}) + 1 + e^{-\alpha}}{(1+e^{\alpha})(1+e^{-\alpha})} = \frac{1}{1+e^{-y}}$$

$$1 + e^{-y} = \frac{(1+e^{\alpha})(1+e^{-\alpha})}{p(e^{\alpha} - e^{-\alpha}) + (1+e^{-\alpha})}$$

$$e^{-y} = \frac{(1+e^{-\alpha})e^{\alpha} - p(e^{\alpha} - e^{-\alpha})}{p(e^{-\alpha} - e^{\alpha}) + 1 + e^{-\alpha}}$$

$$y = \ln \frac{p(e^{-\alpha} - e^{\alpha}) + 1 + e^{-\alpha}}{(1+e^{-\alpha})e^{\alpha} - p(e^{\alpha} - e^{-\alpha})} \quad (3.1.4).$$

and

$$F^{-1}(p) = \ln \frac{(1+e^{-\alpha})e^{\alpha} - (1-p)(e^{\alpha} - e^{-\alpha})}{(1-p)(e^{\alpha} - e^{-\alpha}) + 1 + e^{-\alpha}} \quad (3.1.5).$$

With the general form for $\hat{\mu}_i$ developed in equation (2.2.1) combined with the solution in equation 3.1.6, using the logistic distribution we have:

$$\begin{aligned} 2k\hat{\mu}_i = & \sum_{j(\neq i)}^k \left\{ F^{-1}(1) - F^{-1}(1) + \right. \\ & Y_{1:n}^{ji} \left(-\alpha - \left(\ln \frac{n(1+e^{-\alpha})e^{\alpha} - (n-1)(e^{\alpha} - e^{-\alpha})}{(n-1)(e^{\alpha} - e^{-\alpha}) + n(1+e^{-\alpha})} + \right) \right) \\ & + Y_{2:n}^{ji} \left(\ln \frac{n(1+e^{-\alpha})e^{\alpha} - e^{\alpha} + e^{-\alpha}}{e^{\alpha} - e^{-\alpha} + n(1+e^{-\alpha})} \cdot \frac{(n-2)(e^{\alpha} - e^{-\alpha}) + n(1+e^{-\alpha})}{n(1+e^{-\alpha})e^{\alpha} - (n-2)(e^{\alpha} - e^{-\alpha})} \right) \\ & + \dots + Y_{n:n}^{ji} \left(\ln \frac{n(1+e^{-\alpha})e^{\alpha} - e^{\alpha} + e^{-\alpha}}{e^{\alpha} - e^{-\alpha} + n(1+e^{-\alpha})} + \alpha \right) \\ & + Y_{1:n}^{ij} \left(\alpha + \ln \frac{e^{\alpha} - e^{-\alpha} + n(1+e^{\alpha})}{n(1+e^{-\alpha})e^{\alpha} - e^{\alpha} + e^{-\alpha}} \right) \\ & + Y_{2:n}^{ij} \left(\ln \frac{2(e^{\alpha} - e^{-\alpha}) + n(1+e^{-\alpha})}{n(1+e^{-\alpha}) - 2(e^{\alpha} - e^{-\alpha})} \cdot \frac{n(1+e^{-\alpha})e^{\alpha} - e^{\alpha} + e^{-\alpha}}{e^{\alpha} - e^{-\alpha} + n(1+e^{-\alpha})} \right) \\ & \left. + \dots + Y_{n:n}^{ij} \left(\alpha + \ln \frac{n(1+e^{-\alpha})e^{\alpha} - (n-1)(e^{\alpha} - e^{-\alpha})}{(n-1)(e^{\alpha} - e^{-\alpha}) + n(1+e^{-\alpha})} \right) \right\} \quad (3.1.6). \end{aligned}$$

With $F^{-1}(1) = \alpha$ the above can be expressed in a summarized form as follows:

$$2k\hat{\mu}_i = \sum_{j(\neq i)}^k \left\{ \sum_{m=1}^n (Y_{m:n}^{ji} c_{n-m} + Y_{m:n}^{ij} c_m) \right\} \quad (3.1.7)$$

$i=1,2,\dots,k$, where c_1, c_2, \dots, c_n , are the appropriate coefficients as given in 3.1.7 or more precisely

$$c_m = \ln \left(\frac{m(e^\alpha - e^{-\alpha}) + n(1+e^{-\alpha})}{n(1+e^{-\alpha})e^\alpha - m(e^\alpha - e^{-\alpha})} \cdot \frac{n(1+e^{-\alpha}) - (m-1)(e^\alpha - e^{-\alpha})}{(m-1)(e^\alpha - e^{-\alpha}) + n(1+e^{-\alpha})} \right) \quad (3.1.8)$$

Using the estimator given in equation 2.1.3, we can investigate the $\lim_{n \rightarrow \infty} \hat{\mu}_i$ by examining the behavior of $\int_0^1 f_{ij}(y) dy = \int_0^1 F^{-1}(F_{ij}^{(n)}(y)) dy$. Now, $F_{ij}^{(n)}(y) \rightarrow F_{ij}(y)$ uniformly in y , so that for

$$F_{ij}(y) = F(\phi^{-1}(y) - (\mu_i - \mu_j)) \quad \text{we have}$$

$$\int_0^1 F^{-1}(F_{ij}^{(n)}(y)) dy \rightarrow \int_0^1 (\phi^{-1}(y) - (\mu_i - \mu_j)) dy.$$

Since $\phi(y)$ is monotonic, increasing, and skew-symmetric at $y=\frac{1}{2}$, that is $\phi^{-1}(y) = -\phi^{-1}(1-y)$.

Hence,

$$\int_0^1 (\phi^{-1}(y) - (\mu_i - \mu_j)) dy = -(\mu_i - \mu_j).$$

This leads us to

$$\begin{aligned}
\lim_{n \rightarrow \infty} \hat{\mu}_i &= \frac{1}{2k} \sum_{j(\neq i)}^k \{f_{ji}(y) - f_{ij}(y) dy\} \\
&= \frac{1}{2k} \sum_{j=1}^k (-(\mu_j - \mu_i) + (\mu_i - \mu_j)) \\
&= \frac{1}{2k} 2k\mu_i \\
&= \mu_i
\end{aligned}$$

We can therefore conclude that, asymptotically, $\hat{\mu}_i$ is a consistent estimator of μ_i .

By considering the expectation of $\hat{\mu}_i$ we can determine if this estimate is unbiased. After taking the expectations of both sides of equation 3.1.8 we have:

$$E(\hat{\mu}_i) = \frac{1}{2k} \sum_{j(\neq i)}^k \sum_{m=2}^n \{c_{n-m} E(Y_{m:n}^{ji}) + c_m E(Y_{m:n}^{ij})\} \tag{3.1.9}$$

Since the expectations of order statistics are approximated by quantiles whereby $E(Y_{m:n}^{(n)}) \approx F_{ij}^{-1}(\frac{r}{n+1})$, we can solve for these quantiles with $F(x)=p$, as shown below:

$$p = \frac{\frac{1}{1+e^{-x}} - \frac{1}{1+e^{\alpha}}}{\frac{1}{1+e^{-\alpha}} - \frac{1}{1+e^{\alpha}}}$$

$$x = \ln \frac{p(e^\alpha - e^{-\alpha}) + 1 + e^{-\alpha}}{(1 + e^{-\alpha})e^\alpha - p(e^\alpha - e^{-\alpha})}$$

which then gives, for the value of the r th order statistic:

$$E(Y_{r:n}^{ij}) \approx \phi \left\{ \ln \left(\frac{\frac{r}{n+1}(e^\alpha - e^{-\alpha}) + 1 + e^{-\alpha}}{(1+e)e^\alpha - \frac{r}{n+1}(e^\alpha - e^{-\alpha})} \right) + (\mu_i - \mu_j) \right\} \quad (3.1.10).$$

This approximation depends upon ϕ which is unspecified so we defer the discussion of the expectation until we consider linear combinations of order statistics in Chapter 4.

The estimate of $\phi^{-1}(y)$ for the truncated logistic function in Case I appears in quite a different form. Recall from Chapter 2 (2.2.1), that

$$k(k-1)\hat{\phi}^{-1}(y) = \sum_{i < j=1}^k \left\{ F^{-1}(F_{ij}^{(n)}(y)) - F^{-1}(F_{ji}^{(n)}(y)) \right\} \quad (3.1.11).$$

The sample cumulative distribution functions are given by

$$F_{ij}^{(n)}(y) = \frac{1}{n} \sum_{i=1}^n I(Y_{ij}, y)$$

and

$$F_{ji}^{(n)}(y) = \left(\frac{1}{n}\right) \sum_{i=1}^n I(Y_{ji}, y) \quad (3.1.12)$$

where the indicator function is defined as

$$\begin{aligned} I(a, b) &= 1 \quad \text{if } a < b \\ &= 0 \quad \text{if } a \geq b. \end{aligned}$$

Using this definition, we have $\hat{\phi}^{-1}(y)$ defined as follows:

$$\begin{aligned} k(k-1)\hat{\phi}^{-1}(y) &= \sum_{i(\neq j)} \left\{ \ln \frac{F_{ij}^{(n)}(y)(e^\alpha - e^{-\alpha}) + 1 + e^{-\alpha}}{(1 + e^{-\alpha})e^\alpha - F_{ij}^{(n)}(y)(e^\alpha - e^{-\alpha})} \right\} \\ &\quad + \sum_{j(\neq i)} \left\{ \ln \frac{(1 + e^{-\alpha})e^\alpha - (1 - F_{ji}^{(n)}(y))(e^\alpha - e^{-\alpha})}{(1 - F_{ji}^{(n)}(y))(e^\alpha - e^{-\alpha}) + 1 + e^{-\alpha}} \right\} \end{aligned} \quad (3.1.13).$$

$$\begin{aligned} \text{Since } F_{ji}(y) &= 1 - F(\phi^{-1}(1-y) - (\mu_i - \mu_j)) \\ &= 1 - F_{ij}(1-y) \end{aligned}$$

and

$$F_{ij}(y) = F(\phi^{-1}(y) - (\mu_i - \mu_j))$$

where $F(x) = \frac{1}{1 + e^{-x}}$, we can consider the consistency of $\hat{\phi}^{-1}(y)$ as $n \rightarrow \infty$. Using $F_{ij}^{(n)}(y)$ to estimate $F_{ji}^{(n)}(y)$ we have from equation 2.1.5 that

$$F_{ji}^{(n)}(y) = 1 - F_{ij}^{(n)}(1-y).$$

$$\begin{aligned} \text{As } n \rightarrow \infty, \text{ then } F^{-1}(F_{ij}^{(n)}(y)) &\rightarrow F^{-1}(F_{ij}(y)) \\ &= F^{-1}(F(\phi^{-1}(y) - \delta_{ij})) \\ &= \phi^{-1}(y) - \delta_{ij}. \end{aligned}$$

Similarly, as $n \rightarrow \infty$

$$\begin{aligned} F^{-1}(F_{ij}^{(n)}(y)) &= F^{-1}(1 - F_{ij}^{(n)}(1-y)) \\ &\rightarrow F^{-1}(1 - F_{ij}(1-y)) \\ &= F^{-1}\left\{1 - [F(\phi^{-1}(1-y) - \delta_{ij})]\right\} \\ &= F^{-1}\left\{1 - \frac{1}{1 + e^{-(\phi^{-1}(1-y) - \delta_{ij})}}\right\} \\ &= F^{-1}\left\{\frac{1}{e^{\phi^{-1}(1-y) - \delta_{ij}} + 1}\right\} \\ &= F^{-1}\left\{\frac{1}{1 + e^{-(\phi^{-1}(y) - \delta_{ji})}}\right\} \\ &= F^{-1}(F(\phi^{-1}(y) - \delta_{ji})) \\ &= \phi^{-1}(y) - \delta_{ji}. \end{aligned}$$

Substituting these into 3.1.11 gives, as $n \rightarrow \infty$,

$k(k-1)\hat{\phi}^{-1}(y) \rightarrow \frac{k(k-1)}{2} 2\phi^{-1}(y)$ which implies that

$$\hat{\phi}^{-1}(y) \rightarrow \phi^{-1}(y) \text{ for fixed } y.$$

Hence,

$$\lim_{n \rightarrow \infty} \hat{\phi}^{-1}(y) = \frac{k(k-1)\phi^{-1}}{k(k-1)} = \phi^{-1}(y) \quad \forall y, \quad (3.1.14)$$

showing that $\hat{\phi}^{-1}(y)$ is a consistent estimator.

3.1.2 CASE II: NON-ORDERED PAIRS

At times it may be inconvenient or impossible to collect the ordered pairs (i,j) and (j,i) . This gives rise to the situation where we only consider the pairs (i,j) with $1 \leq i < j \leq k$. This alters the minimum mean squared error equation as follows:

$$Q = \sum_{1 \leq i < j \leq k} \left\{ \int_0^1 \left(\phi^{-1}(y) - (\mu_i - \mu_j) - F^{-1}(F_{ij}^{(n)}(y)) \right)^2 dy \right\} \quad (3.1.15).$$

This, in turn, leads to different solutions for the parameter estimators. First, consider the estimator of $\phi^{-1}(y)$ which minimizes:

$$Q = \sum_{i < j}^k \sum_{j}^k \int_0^1 (\phi^{-1}(y) - (\mu_i - \mu_j) - F^{-1}(F_{ij}^{(n)}(y)))^2 dy$$

Whereby Q is minimized when

$$\sum_{i < j}^k \sum_{j}^k \left\{ \phi^{-1}(y) - (\mu_i - \mu_j) - F^{-1}(F_{ij}^{(n)}(y)) \right\}^2$$

is minimized.

As in 2.1.4, we obtain the following:

$$\frac{k(k-1)}{2} \hat{\phi}^{-1}(y) = \sum_{i < j}^k \sum_{j}^k (\mu_i - \mu_j) + \sum_{i < j}^k \sum_{j}^k F^{-1}(F_{ij}^{(n)}(y))$$

$$\hat{\phi}^{-1}(y) = \frac{2}{k(k-1)} \left\{ \sum_{i=1}^k (k-i) \mu_i - \sum_{i=1}^k (i-1) \mu_i + \sum_{i < j}^k \sum_{j}^k F^{-1}(F_{ij}^{(n)}(y)) \right\}$$

$$\hat{\phi}^{-1}(y) = \frac{2}{k(k-1)} \left\{ \sum_{i=1}^k (k-2i+1) \mu_i + \sum_{i < j}^k \sum_{j}^k F^{-1}(1 - F_{ij}^{(n)}(1-y)) \right\}$$

(3.1.16).

Solving for $\hat{\mu}_i$ leads to the equation

$$\frac{dQ}{d\mu_i} = 2 \sum_{j=1}^{i-1} (\mu_j - \mu_i) (-1) + \sum_{j=i+1}^k (\mu_i - \mu_j)$$

$$\begin{aligned}
& + 2(i-1) \int_0^1 \hat{\phi}^{-1}(y) dy - 2(k-i) \int_0^1 \hat{\phi}^{-1}(y) dy \\
& - 2 \sum_{j=1}^{i-1} \int_0^1 F^{-1}(F_{ji}^{(n)}(y)) dy + 2 \sum_{j=i+1}^k \int_0^1 F^{-1}(F_{ij}^{(n)}(y)) dy = 0
\end{aligned}$$

or

$$\begin{aligned}
0 = & - \sum_{j=1}^{i-1} \hat{\mu}_j + (i-1)\hat{\mu}_i + (k-i)\hat{\mu}_i \\
& - (2i-k-1) \int_0^1 \hat{\phi}^{-1}(y) dy - \sum_{j=1}^{i-1} \int_0^1 F^{-1}(F_{ji}^{(n)}(y)) dy \\
& + \sum_{j=i+1}^k \int_0^1 F^{-1}(F_{ij}^{(n)}(y)) dy
\end{aligned}$$

which leads to the solution

$$\begin{aligned}
\hat{\mu}_i = & \frac{1}{k} (k+1-2i) \hat{\phi}^{-1}(y) + \sum_{j=1}^{i-1} F^{-1}(F_{ji}^{(n)}(y)) - \\
& \sum_{j=i+1}^k F^{-1}(F_{ij}^{(n)}(y)) \quad (3.1.17)
\end{aligned}$$

Using the estimate for μ_i given in 3.1.17 in equation 3.1.16 results in the estimate for $\hat{\phi}^{-1}$ which follows:

$$\begin{aligned}
\hat{\phi}^{-1}(y) = & \frac{2}{k(k-1)} \left\{ \sum_{i=1}^k (k-2i+1) \left(\frac{1}{k} \right) \left((k-2i+1) \hat{\phi}^{-1}(y) + \right. \right. \\
& \left. \left. \sum_{j=1}^{i-1} F^{-1}(F_{ji}^{(n)}(y)) - \sum_{j=i+1}^k F^{-1}(F_{ij}^{(n)}(y)) \right) \right\} +
\end{aligned}$$

$$\left. \sum_{i < j}^k \sum_{j}^k F^{-1}(F_{ij}^{(n)}(y)) \right\} \quad (3.1.18).$$

After considerable simplification the following estimator for $\hat{\phi}^{-1}$ can be obtained

$$\begin{aligned} \hat{\phi}^{-1}(y) = & \frac{6}{(5k+2)(k-1)} \left(\sum_{i=1}^k \left\{ \frac{k-2i+1}{k} \sum_{j=1}^{i-1} F^{-1}(F_{ji}^{(n)}(y)) - \right. \right. \\ & \left. \left. \frac{k-2i+1}{k} \sum_{j=1}^{i-1} F^{-1}(F_{ij}^{(n)}(y)) \right\} + \sum_{i < j}^k \sum_{j}^k F^{-1}(F_{ij}^{(n)}(y)) \right) \end{aligned} \quad (3.1.19)$$

Equations 3.1.16 and 3.1.19 will involve an iterative procedure in order to obtain actual values for $\hat{\mu}_i$ and $\hat{\phi}^{-1}$. Sampling both the pairs (i,j) and (j,i) provided a precise estimate without double use of the data, it is preferred to this latter situation. The remainder of the research is concerned with the case in which data is collected using both ordered pairs (i,j) and (j,i) .

3.2 Limited Integral Approach

An alternative method for handling the logistic distribution in order to obtain finite expressions for all terms of the parameter estimates, is by limiting the range of the integral in equation 2.1.3. Recall that as mentioned in section 3.1, the inverse mapping involves $F^{-1}(0)=-\infty$ and $F^{-1}(1)=\infty$. Instead of assuming all values in the interval $[0,1]$, we use values in the interval $[\epsilon,1-\epsilon]$, where $\epsilon>0$ and some small number. This technique bounds the distribution away from the endpoints that lead to $+\infty$ and $-\infty$.

The limited integral approach is quite similar to the development in Section 3.1, in which we developed specific parameter estimators for $\phi^{-1}(y)$ and μ_i in the logistic distribution. However, this development is slightly simplified. The parameter estimates are again based on those given in 2.2.1. Here, we will consider the situation whereby the ordered pairs (i,j) and (j,i) are sampled. For $0<p<1$, the inverse of the function, $F(y)=p$, is given by:

$$F(y) = \frac{1}{1+e^{-y}}$$

or

$$y = F^{-1}(p) = \ln \frac{p}{1-p} \quad (3.2.1)$$

This gives then, for the parameter estimates

$$\hat{\mu}_i = \frac{1}{2k} \sum_{j(\neq i)} [f_{j i}(y) - f_{i j}(y)] dy \quad (3.2.2)$$

where $f_{ij}(y)$ is the same as in Section 3.1. A brief development of the estimator $\hat{\mu}_i$, $i=1,2,\dots,k$, which follows is almost identical in procedure to the one given in the Section 3.1 in which infinite values for $F^{-1}(0)$ and $F^{-1}(1)$ were avoided by using a truncated logistic distribution.

$$\begin{aligned} \hat{\mu}_i = \frac{1}{2k} \sum_{j(\neq i)}^k \left\{ F^{-1}(\epsilon) [Y_{1:n}^{j i} - \epsilon] + F^{-1}\left(\frac{1}{n}\right) [Y_{2:n}^{j i} - Y_{1:n}^{j i}] + \right. \\ \left. \dots + F^{-1}(1-\epsilon) [(1-\epsilon) - Y_{n:n}^{j i}] - F^{-1}(\epsilon) [Y_{1:n}^{i j} - \epsilon] - \dots \right. \\ \left. - F^{-1}(1-\epsilon) [(1-\epsilon) - Y_{n:n}^{i j}] \right\} \\ \hat{\mu}_i = \frac{1}{2k} \sum_{j(\neq i)}^k \left\{ (-\epsilon) \ln \frac{\epsilon}{1-\epsilon} + Y_{1:n}^{j i} \left[\ln \frac{\epsilon(n-1)}{1-\epsilon} \right] + \right. \end{aligned}$$

$$\begin{aligned}
& Y_{2:n}^{ji} \left[\ln \frac{n-2}{2(n-1)} \right] + \dots + Y_{n-1:n}^{ji} \left[\ln \frac{n-2}{2(n-1)} \right] + \\
& Y_{n:n}^{ji} \left[\ln \frac{\epsilon(n-1)}{1-\epsilon} \right] + (1-\epsilon) \left(\ln \frac{1-\epsilon}{\epsilon} \right) + \epsilon \ln \frac{\epsilon}{1-\epsilon} + \\
& Y_{1:n}^{ij} \left[\frac{1-\epsilon}{(n-1)\epsilon} \right] + Y_{2:n}^{ij} \left[\ln \frac{2(n-1)}{n-2} \right] + \dots + \\
& \left. Y_{n-1:n}^{ij} \left[\ln \frac{2(n-1)}{n-2} \right] + Y_{n:n}^{ij} \left[\ln \frac{1-\epsilon}{(n-1)\epsilon} \right] - (1-\epsilon) \left(\ln \frac{1-\epsilon}{\epsilon} \right) \right\}
\end{aligned}$$

After combining like terms the estimator for μ_i is:

$$\begin{aligned}
\hat{\mu}_i &= \frac{1}{2k} \sum_{j(\neq i)}^k \left\{ \left(\ln \frac{\epsilon(n-1)}{\epsilon} \right) (Y_{1:n}^{ji} + Y_{n:n}^{ji}) \right. \\
&+ \left(\ln \frac{n-2}{2(n-1)} \right) (Y_{2:n}^{ji} + Y_{n-1:n}^{ji}) + \dots + \left(\ln \frac{1-\epsilon}{(n-1)\epsilon} \right) \times \\
&\left. (Y_{1:n}^{ij} + Y_{n:n}^{ij}) + \left(\ln \frac{2(n-1)}{n-2} \right) (Y_{2:n}^{ij} + Y_{n-1:n}^{ij}) + \dots \right\} \\
\hat{\mu}_i &= \frac{1}{2k} \sum_{j(\neq i)}^k \left\{ \left(\ln \frac{\epsilon(n-1)}{\epsilon} \right) (Y_{1:n}^{ji} + Y_{n:n}^{ji} - Y_{1:n}^{ij} - Y_{n:n}^{ij}) \right. \\
&+ \sum_{m=2}^{n-1} \left(Y_{m:n}^{ji} + Y_{n-m+1:n}^{ji} - Y_{m:n}^{ij} - Y_{n-m+1:n}^{ij} \right) \times \\
&\quad \left. \left(\ln \frac{(m-1)(n-m)}{m(n-m+1)} \right) \right\} \tag{3.2.3}.
\end{aligned}$$

The properties of $\hat{\mu}_i$ will be explored in Chapter 4, which deals with the properties of the estimators that are linear combinations of order

statistics.

3.3 Model with an Order Parameter

In Chapter 2, the basic model was extended to include an order parameter, γ_{ij} , and the general estimate for this parameter was given in equation 2.3.5. We can make this estimate more specific using the logistic distribution as we have done with μ_i and $\phi^{-1}(y)$. Here we will only consider the estimate as it appears using the limited integral approach of Section 3.2. Recall that in the general form

$$\hat{\gamma}_{ij} = -\hat{\delta}_{ij} + \frac{1}{2} \int_0^1 (F^{-1}(F_{ji}(y)) - F^{-1}(F_{ij}(y))) dy$$

and as we observe in Section 3.1,

$$\lim_{n \rightarrow \infty} \hat{\gamma}_{ij} = -\hat{\delta}_{ij} + \frac{1}{2} \int_0^1 \left\{ \phi^{-1}(y_{ji}) - (\delta_{ji} + \gamma_{ji}) - [\phi^{-1}(y_{ij}) - \delta_{ij} + \gamma_{ij}] \right\} dy$$

(3.3.1)

This equation gives the estimate for γ_{ij} , in 2.3.3,

in which γ_{ij} is determined for a specific (i,j) pair.

Since $\phi^{-1}(y)$ is skew symmetric at $y=\frac{1}{2}$ then 3.3.1 becomes

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{\gamma}_{ij} &= -\hat{\delta}_{ij} + \frac{1}{2} \left\{ -(\delta_{ji} + \gamma_{ji}) - (-\delta_{ij} + \gamma_{ji}) \right\} \\ &= -\delta_{ij} + \delta_{ij} + \gamma_{ij} \\ &= \gamma_{ij}. \end{aligned}$$

Using the development from $\hat{\mu}_i$ given in the discussion preceding 3.2.3, we have the following estimate for γ_{ij} , as it appears for the logistic distribution, however, the range of integration is limited from ϵ to $1-\epsilon$:

$$\begin{aligned} \hat{\gamma}_{ij} &= \hat{\delta}_{ij} + \frac{1}{2} \left\{ \left(\ln \frac{\epsilon(n-1)}{\epsilon} \right) (Y_{1:n}^{ji} + Y_{n:n}^{ji} - Y_{1:n}^{ij} - Y_{n:n}^{ij}) \right. \\ &\quad \left. + \left(\ln \frac{n-2}{(n-1)2} \right) (Y_{2:n}^{ji} + Y_{n-1:n}^{ji} - Y_{2:n}^{ij} - Y_{n-1:n}^{ij}) + \dots \right. \\ &\quad \left. + \sum_{m=3}^{n-2} \left(\ln \frac{(n-m)(m-1)}{(n-m+1)m} \right) (Y_{m:n}^{ji} + Y_{n-m+1:n}^{ji} - Y_{m:n}^{ij} - Y_{n-m+1:n}^{ij}) \right\} \end{aligned} \tag{3.3.2}$$

for $0 < \epsilon < 1$.

The properties of $\hat{\gamma}_{ij}$ will be similar to those of $\hat{\mu}_i$, $i=1,2,\dots,k$, and will be discussed in Chapter 4.

CHAPTER 4
PROPERTIES OF THE MINIMUM MEAN SQUARE ERROR
PARAMETER ESTIMATES INVOLVING ORDER STATISTICS

4.1 Properties of $\hat{\mu}_i$ using the inverse logistic function.

In the previous chapter we presented estimates for $\hat{\mu}_i$ under different sampling designs and different functional assumptions. In this chapter we consider the estimates for μ_i , $i=1,2,\dots,k$, using the limited integral approach. After developing a general form for the estimate, we apply this by taking $\phi(y)$ to be the logistic cdf which made the computation slightly more simple.

One approach to assessing the properties of estimates given in 3.2.3 and 3.3.2, is to use the theory of L-statistics, that is, linear combinations of order statistics. This approach will be explored for the estimators $\hat{\mu}_i$ and $\hat{\gamma}_{ij}$, $1 \leq i, j \leq k$, in Section 4.3. There are several approaches to establishing the asymptotic theory

associated with linear combinations of order statistics such as those given by Shorak (1972), Chernoff, Gastwirth and Johns (1967), Boos (1979), or Stigler (1974). Shorak's approach places more stringent restrictions on the distribution function and less stringent requirements on the coefficients usually referred to as weight functions in the literature. However, the approach of Shorak along with that of Chernoff, Gastwirth and Johns, and that of Boos are difficult to apply. We prefer the approach presented by Stigler (1974), who places more restrictions on the weight function.

Stigler considers the statistics T_n and S_n given by (in Stigler's notation)

$$S_n = \frac{1}{n} \sum_{i=1}^n J\left(\frac{i}{n+1}\right) X_{i:n} \quad \text{and}$$

$$T_n = \frac{1}{n} \sum_{i=2}^{n-1} J\left(\frac{i}{n+1}\right) X_{i:n} \quad (4.1.1).$$

in which $J(u)$ is a weight function defined on $[0,1]$ and $X_{i:n}$ is the i th ordered observation in a sample of size n . T_n is the expression for the linear combination of the order statistics of

equation 3.2.3, excluding the endpoints. Stigler proposes that S_n and T_n are asymptotically equivalent when the weight function $J(u)$ satisfies certain conditions. For example, $J(u)$ must be bounded and continuous almost everywhere with respect to F^{-1} . It is assumed that the variance of the parent distribution is finite. We give without proof the following theorems due to Stigler (1974), which provide us with the means to show that S_n and therefore T_n are asymptotically normal.

Theorem 1. Assume that $E|X_1|^2 < \infty$, and that $J(u)$ is bounded and continuous a.e. F^{-1} . Then

$$\lim_{n \rightarrow \infty} n\sigma^2(\hat{T}_n) = \sigma^2(J, F)$$

and

$$\lim_{n \rightarrow \infty} n\sigma^2(S_n) = \sigma^2(J, F),$$

where

$$\sigma^2(J, F) =$$

$$2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} J(F(x))J(F(y))F(x)(1-F(y))dx dy$$

(4.1.2)

Theorem 2. Assume that $E(X_1^2) < \infty$, and that $J(u)$ is bounded and continuous a.e. F^{-1} .

Then $\sigma^2(J,F) > 0$ implies

$$L\left(\frac{S_n - E(S_n)}{\sigma(S_n)}\right) \rightarrow N(0,1) \quad \text{as } n \rightarrow \infty.$$

$\sigma^2(J,F)$ is given by 4.1.2. ~~HN~~

Theorem 3. Assume that $E|X_i| < \infty$, and that $J(u)$ is bounded and continuous a.e. with respect to Lebesgue measure. Then as $n \rightarrow \infty$, $E(S_n) \rightarrow \mu(J,F)$, where

$$\begin{aligned} \mu(J,F) &= \int_0^{\infty} \int_{0F(x)}^1 J(u) du \, dx - \int_{-\infty}^0 \int_0^{F(x)} J(u) du \, dx \\ &= \int_0^1 J(u) F^{-1}(u) du \end{aligned} \quad (4.1.3)$$

Theorem 4.

Assume that $\int [F(x)(1-F(x))]^{\frac{1}{2}} dx < \infty$ and that $J(u)$ is bounded and satisfies Holder condition with $\alpha > \frac{1}{2}$ (except possibly at a finite number of points of F^{-1} measure zero). Then

$$n^{\frac{1}{2}}(E(S_n) - \mu(J,F)) \rightarrow 0,$$

where $\mu(J,F)$ is given by 4.1.3.

Theorem 5. Assume that for some $\epsilon > 0$,

$\lim_{x \rightarrow \infty} x^{\epsilon} [1-F(x)+F(-x)] = 0$, and that $J(u)$ is bounded and continuous a.e. F^{-1} . If in

addition, $J(u)=0$ for $0 < u < \alpha$ and $1 - \alpha < u < 1$, then $n\sigma^2(S_n) \rightarrow \sigma^2(J, F)$ (given by 4.1.2), and if $\sigma^2(J, F) > 0$,

$$L \left(\frac{S_n - E(S_n)}{\sigma(S_n)} \right) \rightarrow N(0, 1) \quad \text{as } n \rightarrow \infty.$$

(4.1.4)

Furthermore, the assumptions that $E|X_i| < \infty$

and that $\int [F(x)(1-F(x))]^{\frac{1}{2}} dx$ can be dropped from Theorems 3 and 4 and the conclusions will still hold.

In order to show the asymptotic normality of T_n using Theorem 5, the underlying cumulative distribution function (cdf), $F(x)$, as well as the weight function $J(u)$, must satisfy the following conditions:

- i) $J(u)=0$ for $0 < u < \alpha$ and $1 - \alpha < u < 1$
- ii) $\lim_{x \rightarrow \infty} x^\epsilon [1 - F(x) + F(-x)] = 0$, for some $\epsilon > 0$
- iii) $J(u)$ is bounded and continuous
a.e.

To determine which, if any, of the above conditions are satisfied, we first need to find $J(u)$. We proceed to develop the mean and

variance terms in general, given $J(u)$ does satisfy the conditions. Recalling the basic equations for $\hat{\mu}_i$ in 3.2.3, we observe that there are really two linear combinations of order statistics to consider, one involving Y^{ij} and the second involving Y^{ji} . This is true for any distribution and mapping function satisfying the definitions provided in Chapter 2. This gives then,

$$\hat{\mu}_i = \frac{1}{2k} \sum_{j(\neq i)}^k \{S_{ji}^n - S_{ij}^n\} \quad \text{where}$$

$$S_{ij}^n = \sum_{r=1}^n F^{-1}\left(\frac{r}{n}\right) \{Y_{r+1:n}^{ij} - Y_{r:n}^{ij}\}$$

$$= \sum_{r=1}^n Y_{r:n}^{ij} \left\{ F^{-1}\left(\frac{r-1}{n}\right) - F^{-1}\left(\frac{r}{n}\right) \right\}$$

(4.1.5)

This enables us to find a general form for $J(u)$ as follows:

$$S_{ij}^n = - \frac{1}{n} \sum_{r=1}^n \left\{ \frac{F^{-1}\left(\frac{r}{n}\right) - F^{-1}\left(\frac{r-1}{n}\right)}{\frac{1}{n}} \right\} Y_{r:n}^{ij}$$

so that

$$J(u) = \lim_{n \rightarrow \infty} \left\{ \frac{F^{-1}(u+\Delta u) - F^{-1}(u)}{\Delta u} \right\}$$

$$= \frac{d}{du} F^{-1}(u) \quad (4.1.6,$$

for $u = \frac{r}{h}$ and $\Delta u = \frac{1}{h}$.

Using 4.1.6, we can determine the general form for $\mu(J, F_{ij})$ as shown:

$$\begin{aligned} \mu(J, F_{ij}) &= \int_0^1 J(u) F_{ij}^{-1}(u) du \\ &= \int_0^1 \frac{d}{du} F^{-1}(u) F_{ij}^{-1}(u) du. \end{aligned}$$

Recalling equation 2.1.1, we have then

$$F_{ij}(u) = F(\phi^{-1}(u) - \delta_{ij}) = z$$

$$F^{-1}(z) = \phi^{-1}(u) - \delta_{ij}$$

and

$$\phi(F^{-1}(z) + \delta_{ij}) = u.$$

Therefore,

$$F_{ij}^{-1}(u) = \phi(F^{-1}(u) + \delta_{ij}).$$

By letting $G(u) = F^{-1}(u) = x$, then

$$dx = G'(u) du = \frac{d}{du} F^{-1}(u).$$

This results in

$$\begin{aligned}
 \mu(J, F_{ij}) &= \int_0^1 G'(u) F_{ij}^{-1}(u) du \\
 &= \int_0^1 G'(u) \phi(G(u) + \delta_{ij}) du \\
 &= \int_{-\infty}^{\infty} \phi(x + \delta_{ij}) dx
 \end{aligned}$$

and incorporating the limited integral approach gives

$$\begin{aligned}
 \mu(J, F_{ij}) &= \int_{-\alpha + \delta_{ij}}^{\alpha + \delta_{ij}} \phi(x + \delta_{ij}) dx \\
 &\quad \text{for } \delta_{ij} > 0 \quad (4.1.7).
 \end{aligned}$$

To center the function around zero instead of $\frac{1}{2}$, we let $\psi(z) = \phi(z) - \frac{1}{2}$. Then equation 4.1.7 becomes

$$\begin{aligned}
 \mu(J, F_{ij}) &= \int_{-\alpha + \delta_{ij}}^{\alpha + \delta_{ij}} \left\{ \phi(x) - \frac{1}{2} + \frac{1}{2} \right\} dx \\
 &= \int_{-\alpha + \delta_{ij}}^{\alpha + \delta_{ij}} \psi(x) dx + \frac{1}{2} (\alpha + \delta_{ij} - (-\alpha + \delta_{ij})) \\
 &= \alpha + \int_{-\alpha + \delta_{ij}}^{\alpha + \delta_{ij}} \psi(x) dx.
 \end{aligned}$$

By observing the symmetric properties that exist for $\psi(x)$ we rewrite the equation as follows:

$$\mu(J, F_{ij}) = \alpha + \int_{-\alpha+\delta_{ij}}^{\alpha-\delta_{ij}} \psi(x) dx + \int_{\alpha-\delta_{ij}}^{\alpha+\delta_{ij}} \psi(x) dx$$

which reduces to

$$\begin{aligned} \mu(J, F_{ij}) &= \alpha + \int_{\alpha-\delta_{ij}}^{\alpha+\delta_{ij}} \left(\phi(x) - \frac{1}{2} \right) dx \\ &= \alpha - \delta_{ij} + \int_{\alpha-\delta_{ij}}^{\alpha+\delta_{ij}} \phi(x) dx \end{aligned} \quad (4.1.8)$$

Similarly, we can determine the mean for the Y^{ji} portion of the equation, giving

$$\mu(J, F_{ji}) = \alpha + \delta_{ij} - \int_{\alpha+\delta_{ji}}^{\alpha-\delta_{ji}} \phi(x) dx \quad (4.1.9)$$

To get the overall mean we combine the two which results in

$$\begin{aligned} \mu(J, F) &= 2\delta_{ij} - \int_{\alpha+\delta_{ji}}^{\alpha-\delta_{ji}} \phi(x) dx - \int_{\alpha-\delta_{ij}}^{\alpha+\delta_{ij}} \phi(x) dx \\ &= 2\delta_{ij} + B(\alpha, \delta_{ij}) \end{aligned} \quad (4.1.10)$$

where $B(\alpha, \delta_{ij})$ is the bias term given by the integral portion of the equation for $\mu(J, F)$.

The bias in 4.1.10 is small when α is large. In this case the value of $\phi(x)$ is close to 1 which results in the mean reducing to the following:

$$\begin{aligned} \mu(J, F) &\approx 2\delta_{ij} - \{\alpha - \delta_{ji} - (\alpha + \delta_{ji})\} - \{\alpha + \delta_{ij} - (\alpha - \delta_{ij})\} \\ &= 2\delta_{ij} - 4\delta_{ij} \\ &= -2\delta_{ij}. \end{aligned} \tag{4.1.11}$$

Using equation 4.1.5 and the fact that as $n \rightarrow \infty$ $E(S_n) \rightarrow \mu(J, F)$ we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{\mu}_i &= -\frac{1}{2k} \sum_{j(\neq i)}^k \mu(J, F) \\ &= \frac{1}{2k} \sum_{j(\neq i)}^k -2\delta_{ij} + \bar{B}(\alpha, \mu_i) \end{aligned} \tag{4.1.12}$$

$$\text{where } \bar{B}(\alpha, \mu_i) = \frac{1}{2k} \sum_{j(\neq i)}^k B(\alpha, \delta_{ij}).$$

Finally then we have,

$$\lim_{n \rightarrow \infty} \hat{\mu}_i \approx \mu_i$$

when α is sufficiently large.

We have then, considering that there will always be at least some small amount of bias present

$$\lim_{n \rightarrow \infty} E[\hat{\mu}_i] = \mu_i + B^*(\alpha, \delta_{ij}) \quad (4.1.13)$$

where $B^*(\alpha, \delta_{ij}) = \mu_i - \lim_{n \rightarrow \infty} \hat{\mu}_i$ for all α .

The equation given for the variance term (equation 4.1.2), involves so many variable terms that we were not able to give a precise form for the variance in the general case. Instead we can bound the variance which allows us to make the variance term as small as we want simply by increasing the sample size. We can modify 4.1.2 as follows:

$$\begin{aligned} \sigma^2(J, F_{ij}) &= \int_{\epsilon < u < v < 1-\epsilon} \int J(u) u(1-v) dF_{ij}^{-1}(u) J(v) dF_{ij}^{-1}(v) \\ &\leq \int_{\epsilon < u < v < 1-\epsilon} \int J(u) dF_{ij}^{-1}(u) J(v) dF_{ij}^{-1}(v) \\ &\leq \frac{1}{2} \int_{\epsilon}^{1-\epsilon} \int_{\epsilon}^{1-\epsilon} J(u) \left\{ \phi'(F^{-1}(z) + \delta_{ij}) \right\} dF^{-1}(z) dz. \end{aligned}$$

The weight function, $J(u)$, will be finite as long as the density $f(u)$, corresponding to $F(u)$, is not zero over $(-\alpha, \alpha)$. Since we know that if $G(u)=z$ then $g(u) du=dz$ which implies

$$du = \frac{dz}{g(u)} = \frac{dz}{g(F^{-1}(z))}.$$

If $\frac{1}{x}$ is bounded then we know x is bounded so $dF^{-1}(z)$ is bounded and we can conclude that the variance terms is finite. From Theorem 1 then, we have that

$$n\sigma^2(S_n) \rightarrow \sigma^2(J, F)$$

then

$$\sigma^2(S_n) \rightarrow \frac{\sigma^2(J, F)}{n}$$

where $\sigma^2(J, F) = \left(\frac{1}{2k}\right)^2 \sum_{j(\neq i)}^k \left\{ \sigma^2(J, F_{ij}) + \sigma^2(J, F_{ji}) \right\}.$

4.2 Application Using a Logistic Mapping Function

Using the general form for $\mu(J,F)$ and the bounded variance we can consider more specific mean and variance terms using a logistic mapping distribution. The common coefficient for the terms of equation 3.2.3 is

$$c_m = \ln \left(\frac{m(n-m+1)}{(m-1)(n-m)} \right) \quad (4.2.1)$$

By letting $m=nx$ we have

$$c_m = \ln \left\{ \frac{nx(n-nx+1)}{(nx-1)(n-nx)} \right\}.$$

From 4.1.1, we see that in order to obtain $J(x)$, we must take n times the coefficient, resulting in

$$\begin{aligned} n \left[\ln \frac{nx(n-nx+1)}{(nx-1)(n-nx)} \right] &= n \left[\ln \frac{x(n-nx+1)}{(nx-1)(1-x)} \right] \\ &= n \left[\ln \frac{x}{1-x} + \ln \frac{n-nx+1}{nx-1} \right] \\ &= n \left[\ln \frac{x}{1-x} - \ln \frac{x-\frac{1}{n}}{1-(x-\frac{1}{n})} \right] \end{aligned}$$

Now let $f(x) = \ln \frac{x}{1-x}$ and $h = \frac{1}{n}$, then

$$n \left[\ln \frac{x}{1-x} - \ln \frac{x-\frac{1}{n}}{1-(x-\frac{1}{n})} \right] = \frac{1}{h} [f(x) - f(x-h)]$$

→ $f'(x)$ as $h \rightarrow 0$.

Consider,

$$\begin{aligned} f'(x) &= \frac{d}{dx} \left(\ln \frac{x}{1-x} \right) \\ &= \frac{1-x}{x} \left[\frac{1}{1-x} + \frac{x}{(1-x)^2} \right] \\ &= \frac{1}{x} + \frac{1}{1-x} \end{aligned}$$

Hence,

$$J(x) = \frac{1}{x(1-x)} \quad (4.2.2).$$

Therefore,

$$\begin{aligned} J(u) &= \frac{1}{u(1-u)} \quad \text{for } u \in (\epsilon, 1-\epsilon) \\ &= 0 \quad \text{for } u \in [0, \epsilon), (1-\epsilon, 1] \end{aligned}$$

is the weight function estimating μ_i , $i=1,2,\dots,k$.

With this weight function, $J(u)$, we systematically proceed to determine if the remaining three conditions concerning $F(x)$ and $J(u)$ are satisfied

- i) Since the weight function ranges only from $m=2$ to $n-1$, $J(u)$ is defined to be zero for

either of the endpoints as required by this condition.

ii) For this condition we look at the tail areas of the distribution of the random variable Y for which $0 \leq Y \leq 1$. The tail area condition for Y can be expressed as

$$\lim_{y \rightarrow 1} y^\epsilon [1 - F_{ij}(y) - F_{ij}(1-y)] = 0$$

for which

$$F_{ij}(y) = F(\phi^{-1}(y) - (\mu_i - \mu_j))$$

where F is a central logistic distribution. Therefore, for fixed δ_{ij}

$$\lim_{y \rightarrow 1} y^\epsilon \left\{ 1 - \frac{1}{1 + e^{-(\ln \frac{y}{1-y} - \delta_{ij})}} + \frac{1}{1 + e^{-(\ln \frac{1-y}{y} - \delta_{ij})}} \right\}$$

$$= \lim_{y \rightarrow 1} y^\epsilon \left\{ 1 - \frac{1}{1 + \frac{1-y}{y} e^{-\delta_{ij}}} + \frac{1}{1 + \frac{y}{1-y} e^{-\delta_{ij}}} \right\}$$

$$= 1 [1 - 1 + 0]$$

$$= 0, \quad \text{and this condition is met.}$$

iii) In finding the original weight function, it was shown that the derivative of $F(x)$, given by $F(x)(1-F(x))$ was bounded by 1 and hence, $J(u)$ is also bounded and continuous a.e. satisfying this final condition.

Knowing that the conditions for Stigler's results hold, we now proceed to find the asymptotic mean, $\mu(J,F)$, and variance, $\sigma^2(J,F)$. Since there are two types of linear combinations of order statistics in the equation for $\hat{\mu}_i$, one for Y^{ij} and one for Y^{ji} we compute their means separately, then combine them since the observations are independent. After appropriately substituting into equation 4.1.4 we find:

$$\begin{aligned}
 \mu(J, F_{ij}) &= \int_{\epsilon}^{1-\epsilon} J(u) F_{ij}^{-1}(u) du \\
 &= \int_{\epsilon}^{1-\epsilon} \frac{1}{u(1-u)} \left[\frac{u}{e^{-\delta_{ij}} + u(1-e^{-\delta_{ij}})} \right] du. \\
 &= \frac{1}{1 - e^{-\delta_{ij}} + e^{-\delta_{ij}}} \ln \left(\frac{u + (1-u)e^{-\delta_{ij}}}{1-u} \right) \Big|_{\epsilon}^{1-\epsilon} \\
 &= \ln \left(\frac{u + (1-u)e^{-\delta_{ij}}}{1-u} \right) \Big|_{\epsilon}^{1-\epsilon}
 \end{aligned}$$

Similarly, we find

$$\mu(J, F_{ji}) = \ln \left(\frac{u + (1-u)e^{-\delta_{ji}}}{1-u} \right) \Big|_{\epsilon}^{1-\epsilon}$$

Combining these give

$$\begin{aligned} \mu(J, F) &= \mu(J, F_{ji}) - \mu(J, F_{ij}) \\ &= \ln \left\{ \frac{u + (1-u)e^{-\delta_{ji}}}{1-u} \cdot \frac{1-u}{u + (1-u)e^{-\delta_{ij}}} \right\} \Big|_{\epsilon}^{1-\epsilon} \\ &= \ln \left\{ \frac{1-\epsilon + \epsilon e^{-\delta_{ji}}}{1-\epsilon + \epsilon e^{-\delta_{ij}}} \right\} - \ln \left\{ \frac{\epsilon + (1-\epsilon)e^{-\delta_{ji}}}{\epsilon + (1-\epsilon)e^{-\delta_{ij}}} \right\} \\ &= \ln \left\{ \frac{2(1-\epsilon)\epsilon + (1-\epsilon)^2 e^{-\delta_{ij}} + \epsilon^2 e^{-\delta_{ji}}}{2(1-\epsilon)\epsilon + \epsilon^2 e^{-\delta_{ij}} + (1-\epsilon)^2 e^{-\delta_{ij}}} \right\} \end{aligned}$$

By taking the limit as $\epsilon \rightarrow 0$, we have

$$\begin{aligned} &\lim_{\epsilon \rightarrow 0} (\mu(J, F_{ji}) - \mu(J, F_{ij})) \\ &= \ln \left(\frac{1-0}{1-0} \cdot \frac{e^{-\delta_{ij}}}{e^{-\delta_{ji}}} \right) \\ &= \ln (e^{-\delta_{ij}} + e^{\delta_{ji}}) \end{aligned}$$

$$= -(2\mu_i - 2\mu_j) \quad (4.2.3)$$

Summing over all j , with $\delta_{ij} = \mu_i - \mu_j$, gives:

$$\begin{aligned} \frac{1}{2k} \sum_{j(\neq i)}^k 2(\mu_i - \mu_j) \\ = -(2k\mu_i + 0). \end{aligned}$$

$$\text{This leads us to } \lim_{\epsilon \rightarrow 0} E(\hat{\mu}_i) = -\mu_i. \quad (4.2.4)$$

The minus sign is correct because we need to rewrite the original equation for $\hat{\mu}_i$ with a negative sign factored out of the summation, as shown in section 4.1 with S_n . This is a finite analogue of the integral given in 4.1.4, so that

$$E(\hat{\mu}_i) = \frac{1}{2k} \sum_{j(\neq i)}^k \left\{ -\mu(J, F_{ji}) - (-\mu(J, F_{ij})) \right\}$$

which does gives us the $\lim_{\epsilon \rightarrow 0} E(\hat{\mu}_i) = \mu_i$.

To obtain the variance term, we return to equation 4.1.2, which, with the appropriate substitutions, is given as follows:

$$\sigma^2(J, F) = 2 \int_{0 < u < v < 1} J(u) J(v) u(1-v) dF^{-1}(u) dF^{-1}(v)$$

with $u=F(x)$ and $v=F(y)$. We evaluate the integral as shown below:

$$\begin{aligned} &= 2 \int_{\epsilon < u < v < 1-\epsilon} \frac{1}{u(1-u)} \cdot \frac{1}{v(1-v)} u(1-v) \frac{1}{u(1-u)} \cdot \frac{1}{v(1-v)} du dv \\ &= 2 \int_{\epsilon < u < v < 1-\epsilon} \frac{1}{v^2(1-v)u(1-u)^2} du dv \\ &= 2 \int_{\epsilon}^{1-\epsilon} \frac{1}{v^2(1-v)} \left(\int_{\epsilon}^v \frac{1}{u(1-u)^2} du \right) dv \\ &= 2 \int_{\epsilon}^{1-\epsilon} \frac{1}{v^2(1-v)} \left(\frac{1}{1-u} - \ln \frac{1-u}{u} \right) \Big|_{\epsilon}^v dv \quad (4.2.5) \end{aligned}$$

Expanding and combining like terms then results in

$$\begin{aligned} \sigma^2(J, F) = 2 \left\{ \int_{\epsilon}^{1-\epsilon} \frac{1}{v^2(1-v)^2} dv + \int_{\epsilon}^{1-\epsilon} \frac{1}{v^2(1-v)} \ln \frac{v}{1-v} dv \right. \\ \left. + \left(\left(\ln \frac{1-\epsilon}{\epsilon} \right) - \frac{1}{1-\epsilon} \right) \left(\int_{\epsilon}^{1-\epsilon} \frac{1}{v^2(1-v)} dv \right) \right\} \end{aligned} \quad (4.2.6)$$

By using the appropriate integral tables 4.2.6 becomes

$$\sigma^2 = 2 \left\{ \left(-\frac{1-2v}{v(1-v)} - 2 \ln \frac{1-v}{v} \right) \frac{1-\epsilon}{\epsilon} + \left(\ln \frac{1-\epsilon}{\epsilon} - \frac{1}{1-\epsilon} \right) \left(-\frac{1}{v} - \ln \frac{1-v}{v} \right) \frac{1-\epsilon}{\epsilon} \right. \\ \left. + \int_{\frac{\epsilon}{1-\epsilon}}^{\frac{1-\epsilon}{\epsilon}} \frac{1}{\left(\frac{x}{1+x} \right)^2 \left(1 - \frac{x}{1+x} \right)} (\ln x) \frac{1}{(1+x)^2} dx \right\}$$

where the last integral expression arises when we substitute x for $\frac{v}{1-v}$. After algebraic simplification, this last integral can be separated into parts for integration as follows:

$$\int_{\frac{\epsilon}{1-\epsilon}}^{\frac{1-\epsilon}{\epsilon}} \frac{1+x}{x^2} (\ln x) dx = \int_{\frac{\epsilon}{1-\epsilon}}^{\frac{1-\epsilon}{\epsilon}} \frac{1}{x^2} (\ln x) dx + \int_{\frac{\epsilon}{1-\epsilon}}^{\frac{1-\epsilon}{\epsilon}} \frac{1}{x} (\ln x) dx \\ = \frac{1}{x} \left(-\ln x - 1 \right) \Big|_{\frac{\epsilon}{1-\epsilon}}^{\frac{1-\epsilon}{\epsilon}} + \frac{1}{2} \left((\ln x)^2 \right) \Big|_{\frac{\epsilon}{1-\epsilon}}^{\frac{1-\epsilon}{\epsilon}}.$$

(4.2.7)

Therefore

$$\sigma^2(J, F) = 2 \left\{ \frac{2(1-\epsilon)-1}{(1-\epsilon)\epsilon} - 2 \ln \frac{\epsilon}{1-\epsilon} + \frac{1-2\epsilon}{(1-\epsilon)\epsilon} \right. \\ + 2 \ln \frac{1-\epsilon}{\epsilon} + \left(\ln \frac{1-\epsilon}{\epsilon} - \frac{1}{1-\epsilon} \right) \left(-\frac{1}{1-\epsilon} - \ln \frac{\epsilon}{1-\epsilon} + \frac{1}{\epsilon} + \ln \frac{1-\epsilon}{\epsilon} \right) \\ \left. + \frac{\epsilon}{1-\epsilon} \left(\ln \frac{\epsilon}{1-\epsilon} \right) - \frac{\epsilon}{1-\epsilon} + \frac{1-\epsilon}{\epsilon} \ln \frac{\epsilon}{1-\epsilon} + \frac{1-\epsilon}{\epsilon} \right\}$$

After combining terms we arrive at the expression for the variance term:

$$\sigma^2(J, F) = 2 \left\{ \frac{6\epsilon^2 - 7\epsilon + 2}{\epsilon(1-\epsilon)^2} + \frac{4\epsilon}{1-\epsilon} \ln \frac{\epsilon}{1-\epsilon} + 2 \left(\ln \frac{1-\epsilon}{\epsilon} \right)^2 \right\}. \quad (4.2.8)$$

Recalling that $\lim_{n \rightarrow \infty} n\sigma^2(\hat{T}_n) = \sigma^2(J, F)$ gives us the asymptotic variance of \hat{T}_n as

$$\sigma^2(T_n) = \frac{2}{n} \left\{ \frac{6\epsilon^2 - 7\epsilon + 2}{\epsilon(1-\epsilon)^2} + \frac{4\epsilon}{1-\epsilon} \ln \frac{\epsilon}{1-\epsilon} + 2 \left(\ln \frac{1-\epsilon}{\epsilon} \right)^2 \right\} \quad (4.2.9)$$

This is reasonable since the variance depends on the range of the sample (as determined by ϵ) and, of course, on the sample size.

Finally, then we can conclude that

$$\frac{E(\hat{\mu}_i) - \mu_i}{\sigma(T_n)} \xrightarrow{d} N(0, 1) \quad \text{as } n \rightarrow \infty \quad (4.2.13)$$

where $E(\hat{\mu}_i)$ is given in 4.1.10.

4.3 Properties of $\hat{\gamma}_{ij}$ Using Logistic Inverse

When an order effect exists, the estimator for the order parameter, γ_{ij} , also involves order statistics as shown in equation 3.3.2. The coefficients of all of the terms from $m=2$ to $m=n-1$ inclusive, are identical to those of $\hat{\mu}_i$ as discussed in Section 4.2. Therefore, we can draw similar conclusions to those found in Section 4.2. Again, the weight function $J(u) = \frac{1}{u(1-u)}$ and the mean and variance for T_n , the L-statistic, will be those given in 4.1.4 and 4.1.2. An added complexity arises, however, when considering $\hat{\gamma}_{ij}$ since $\hat{\gamma}_{ij}$ is a function of the ordered pair (i,j) and (j,i) only.

Using the discussion in Section 4.2, on the conditions for Stigler's asymptotic distribution of L-Statistics, we can proceed with the asymptotic distribution of $\hat{\gamma}_{ij}$. Using the development preceding equation 4.1.5 we rewrite equation 3.3.2 as

$$\hat{\gamma}_{ij} = \delta_{ij} + \frac{1}{2} \{ S_{ji}^n - S_{ij}^n \} \quad (4.3.1)$$

where S_{ij}^n is defined in equation 4.1.5. From 4.1.10, we found that prior to summing over all $j(\neq i)$, $\mu(J, F) = 2\delta_{ij} + B(\alpha, \delta_{ij})$.

Recall that $f_{ji}(y) = F^{-1}(F_{ji}^{(n)}(y))$
and, as $n \rightarrow \infty$, $f_{ji}(y) \rightarrow F^{-1}(F(\phi^{-1}(y) - (\delta_{ji} + \gamma_{ji})))$;

$$= \phi^{-1}(y) - (\delta_{ji} + \gamma_{ij})$$

and

$$f_{ij}(y) = F^{-1}(F_{ij}^{(n)}(y)) \rightarrow F^{-1}(F(\phi^{-1}(y) - (\delta_{ij} + \gamma_{ij}))).$$

$$= \phi^{-1}(y) - (\delta_{ij} + \gamma_{ij}).$$

By proceeding similarly to the development for $\hat{\mu}_i$ in Section 4.1, we obtain

$$\mu(J, F_{ij}) = \alpha - (\delta_{ij} + \gamma_{ij}) + \int_{\alpha + \delta_{ji} + \gamma_{ji}}^{\alpha - (\delta_{ji} + \gamma_{ji})} \phi(x) dx$$

$$\mu(J, F_{ji}) = \alpha + (\delta_{ij} + \gamma_{ij}) - \int_{\alpha - (\delta_{ij} + \gamma_{ij})}^{\alpha + \delta_{ij} + \gamma_{ij}} \phi(x) dx.$$

Combining these results gives

$$\begin{aligned} \mu(J,F) = & 2(\delta_{ij} + \gamma_{ij}) - \int_{\alpha + \delta_{ji} + \gamma_{ji}}^{\alpha - (\delta_{ji} + \gamma_{ji})} \phi(x) dx \\ & - \int_{\alpha - (\delta_{ij} + \gamma_{ij})}^{\alpha + \delta_{ij} + \gamma_{ij}} \phi(x) dx. \end{aligned} \quad (4.3.2)$$

Again recall that $E(S_n) \rightarrow \mu(J,F)$ as $n \rightarrow \infty$ so that

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{\gamma}_{ij} = & -\delta_{ij} + \gamma_{ij} + \frac{1}{2} \left\{ 2(\delta_{ji} + \gamma_{ji}) - \int_{\alpha + \delta_{ji} + \gamma_{ji}}^{\alpha - (\delta_{ji} + \gamma_{ji})} \phi(x) dx \right. \\ & \left. - \int_{\alpha - (\delta_{ij} + \gamma_{ij})}^{\alpha + \delta_{ij} + \gamma_{ij}} \phi(x) dx \right\} \\ \approx & \gamma_{ij} \end{aligned} \quad (4.3.3)$$

for large α , as discussed in section 4.1.

Therefore, since there is always some amount of bias, $\hat{\gamma}_{ij}$ is a biased estimator of γ_{ij} and

$$\lim_{n \rightarrow \infty} E(\hat{\gamma}_{ij}) = \gamma_{ij} + B(\alpha, \delta_{ij}, \gamma_{ij}) \quad (4.3.4)$$

where $B(\alpha, \delta_{ij}, \gamma_{ij}) = \lim_{n \rightarrow \infty} \hat{\gamma}_{ij} - \gamma_{ij}$.

From section 4.1, we know that the variance term is bounded for the L-statistics we are investigating. The asymptotic variance for

the linear combination of order statistics, $\sigma^2(T_n)$ is given in equation 4.2.9. This leads to the variance of $\hat{\gamma}_{ij}$ as follows:

$$\begin{aligned}\sigma^2(\hat{\gamma}_{ij}) &= \sigma^2(T_n) \\ &= \frac{2}{n} \left\{ \frac{6\epsilon^2 - 7\epsilon + 2}{\epsilon(1-\epsilon)^2} + \frac{4\epsilon}{1-\epsilon} \ln \frac{\epsilon}{1-\epsilon} + 2 \left(\ln \frac{1-\epsilon}{\epsilon} \right)^2 \right\}\end{aligned}$$

(4.3.4).

Recall that the bias term becomes small for large n , so that we can make the variance as small as necessary. We conclude then that $\hat{\gamma}_{ij}$ is a consistent estimator of γ_{ij} .

Since we have presented an analysis for ordered pairs, the estimate of order effects is also included since order parameters are frequently used in the study of paired comparison experiments.

CHAPTER 5
SIMULATION STUDY USING PARAMETER ESTIMATES

5.1 General Approach

This chapter uses simulated data to empirically explore the parameter estimators and determine the surface validity of the resulting estimates. A random number generator provided values of a uniform random variable x , defined over the interval $(0,1)$, which were then transformed to values of a logistic random variable V , using the transformation:

$$V_{ij} = \ln \frac{x}{1-x} \quad (5.1.1)$$

The logistic random variables are the V_{ij} 's discussed in Chapter 2 with $V_{ij} \sim L(0,1)$ with the range $(-\infty, +\infty)$.

The logistic data was generated for paired comparisons with 3, 4, or 5, items to be compared. Replications of 5, 10, and 25, at each of the item combinations were considered. Ordered sampling of both (i,j) and (j,i) was performed as

indicated by the results in Chapter 3. Each parameter estimate was then programmed separately, using the appropriate equation given in Chapter 4.

5.2 Results for $\hat{\mu}_i$

In order to estimate $\hat{\mu}_i$ with simulated data, we need to set δ_{ij} at some predetermined value based on the μ_i 's. By presetting the μ_i 's we have a basis for comparing the estimates with the original predetermined μ_i 's. The μ_i 's were selected so that the effect of different sample sizes on the estimate could be revealed with only one μ_i different (see Table 5.1), then with 2, 3, and 4 μ_i 's being different (see Tables 5.2 and 5.3). Sorting of the variables was necessary because of the order statistics contained in the estimate (3.1.8). After summing the linear combinations of order statistics over all $j \neq i$ we can determine how close our estimate is to the preset μ_i 's. Recall that the parameter can only be estimated up to a scale factor, so we do not expect the estimates to be identical to the

preestablished values but to be in similar proportion to them.

Tables 5.1 to 5.3 reflect the results of the simulation for $\hat{\mu}_i$. As expected, the estimates are closer in proportion to the true preset values of μ_i as the sample size increases. The estimated values obtained from identical preset μ_i 's are quite similar in most of the examples while the relative proportions for the

Table 5.1
Comparison of 3 Items

| Replications | Parameter | Preset Value | Estimated Value |
|--------------|-----------|--------------|-----------------|
| 5 | μ_1 | 1 | 0.354 |
| | μ_2 | 1 | 0.658 |
| | μ_3 | -2 | -1.620 |
| 10 | μ_1 | 1 | 0.290 |
| | μ_2 | 1 | 0.378 |
| | μ_3 | -2 | -1.129 |
| 25 | μ_1 | -2 | -1.129 |
| | μ_2 | 1 | 0.374 |
| | μ_3 | 1 | 0.370 |

Table 5.2
Comparison of 4 Items

| Replications | Parameter | Preset Value | Estimated Value |
|--------------|-----------|--------------|-----------------|
| 5 | μ_1 | 1 | 0.115 |
| | μ_2 | 1 | 0.513 |
| | μ_3 | 1 | 0.209 |
| | μ_4 | -3 | -1.620 |
| 10 | μ_1 | 1 | 0.387 |
| | μ_2 | 1 | 0.300 |
| | μ_3 | 1 | 0.087 |
| | μ_4 | -3 | -1.340 |
| 25 | μ_1 | 1 | 0.231 |
| | μ_2 | 1 | 0.206 |
| | μ_3 | 1 | 0.302 |
| | μ_4 | -3 | -1.080 |

Table 5.3
Comparison of 5 Items

| Replications | Parameter | Preset Value | Estimated Value |
|--------------|-----------|--------------|-----------------|
| 5 | μ_1 | 1 | 0.184 |
| | μ_2 | 1 | 0.295 |
| | μ_3 | -2 | -0.368 |
| | μ_4 | -1 | -0.195 |
| | μ_5 | 1 | 0.169 |
| 10 | μ_1 | 1 | 0.180 |
| | μ_2 | 1 | 0.200 |
| | μ_3 | -2 | -0.604 |
| | μ_4 | -1 | -0.190 |
| | μ_5 | 1 | 0.239 |
| 25 | μ_1 | 1 | 0.192 |
| | μ_2 | 1 | 0.193 |
| | μ_3 | -2 | -0.484 |
| | μ_4 | -1 | -0.197 |
| | μ_5 | 1 | 0.218 |

Table 5.4
 Comparison of 5 Items (Different preset μ 's)

| Replications | Parameter | Preset Value | Estimated Value |
|--------------|-----------|--------------|-----------------|
| 25 | μ_1 | -2 | -0.489 |
| | μ_2 | 1 | 0.193 |
| | μ_3 | 1 | 0.201 |
| | μ_4 | 1 | 0.199 |
| | μ_5 | -1 | -0.187 |
| 25 | μ_1 | -2 | 0.353 |
| | μ_2 | 3 | -0.594 |
| | μ_3 | 1 | -0.162 |
| | μ_4 | -1 | 0.165 |
| | μ_5 | -1 | 0.160 |
| 25 | μ_1 | 2 | -0.338 |
| | μ_2 | 2 | -0.300 |
| | μ_3 | -4 | 1.000 |
| | μ_4 | 1 | -0.119 |
| | μ_5 | -1 | 0.154 |

μ_i 's set at different levels are not quite as accurate. With an increase in sample size these tend closer to the correct proportions indicated by the preset values.

5.3 Results on ϕ^{-1} and Density estimation procedures for estimating $\frac{d}{du} \hat{\phi}^{-1}(u)$ (optional)

The estimate developed for ϕ^{-1} as shown in equation 2.2.1 was programmed using the basic logistic distribution for the sample data. Since this function provides a one-to-one mapping from (0,1) onto the real numbers, we would expect the sample data to reveal a logistic cdf when graphed. Since $\phi(v) = \frac{1}{1+e^{-x}}$ then $\hat{\phi}^{-1}(p)$ is the logistic cdf and $\phi'(v)$ is the logistic density. As is typical for most estimates, the larger the sample size, the smoother the curve. Each of the graphs in Figures 5.1 to 5.3 were plotted based on the data generated under the simulation discussed in Section 5.1. Since these graphs are quite close to a true logistic cdf, it would appear that the estimate developed through this research is adequate for moderate sample sizes.

Following the consideration of $\hat{\phi}^{-1}$, we further explored the data through density estimation procedures. The kernel estimator (Silverman, 1986) was used with a normal kernel to graph $\phi'(v)$ obtained from the simulated data. This

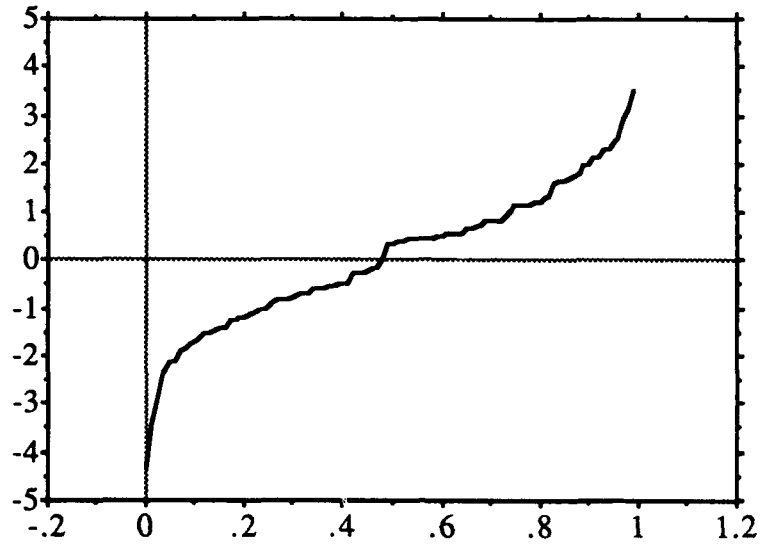


Figure 5.1: Phi Inverse with 3 Treatments and
25 Replications

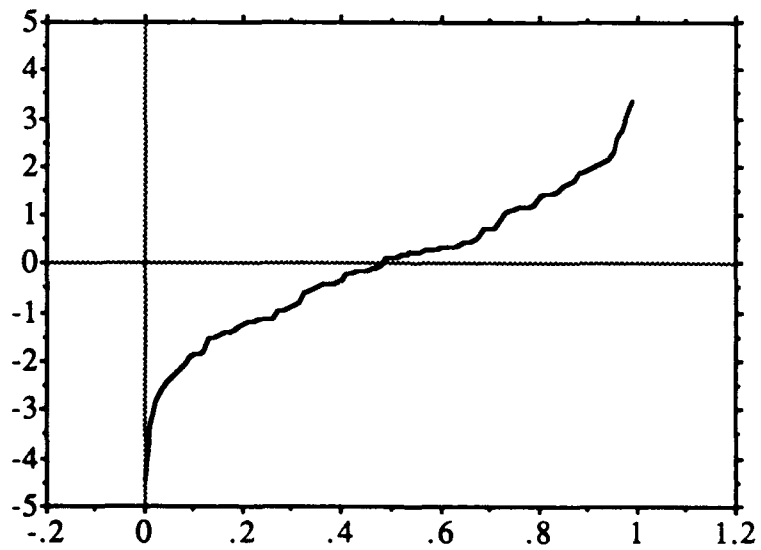


Figure 5.2: Phi Inverse with 4 Treatments and
25 Replications

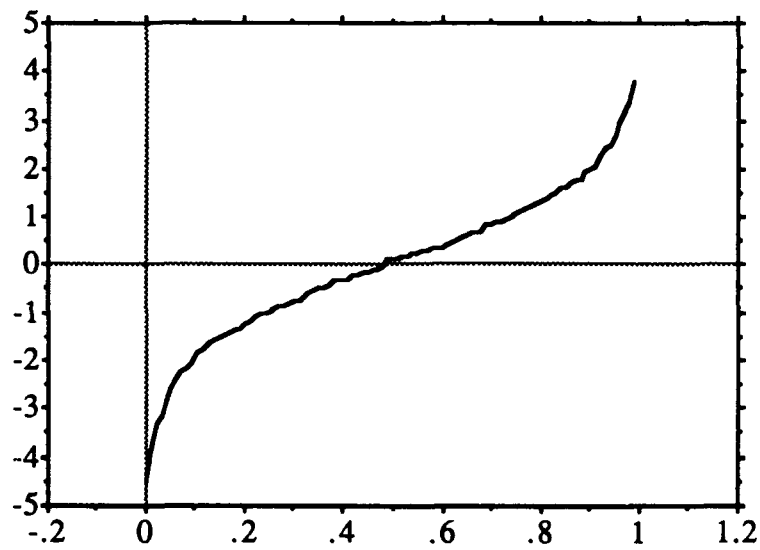


Figure 5.3: Phi Inverse with 5 Treatments,
25 Replications

estimator is defined by Silverman to be

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x-x_i}{h}\right) \quad (5.3.1)$$

where h is the window width, bandwidth, or smoothing parameter. Various bandwidths were considered as suggested by Silverman. Figures 5.4 to 5.9 show the simulated data using 4 treatments and 10 replications, Figures 5.10 to 5.16 also use 4 treatments but with 25 replications and finally, Figures 5.17 to 5.22 show 4 treatments with 50 replications. In each of these three situations with differing sample sizes, the first bandwidth is too small revealing spurious structure, while the last one or two figures in each set use too large a bandwidth obscuring the true nature of the distribution. Further, then we see that the graphs shown in Figures 5.7, 5.13 and 5.18 reveal a logistic density with slightly heavier tails than that of a true normal density function.

Silverman discusses the asymptotic properties of the kernel estimator and "global accuracy" of \hat{f} using the mean integrated square error (MISE) given by

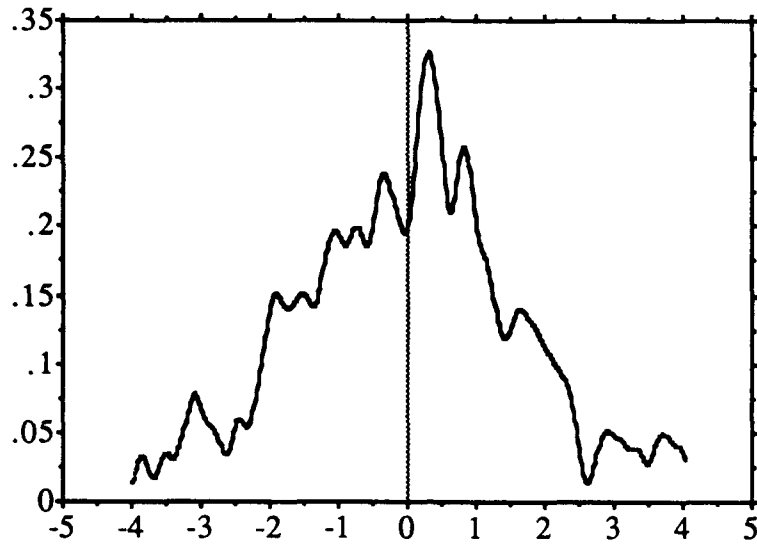


Figure 5.4: Normal Density Smoothing, Bandwidth = 0.1
10 Replications

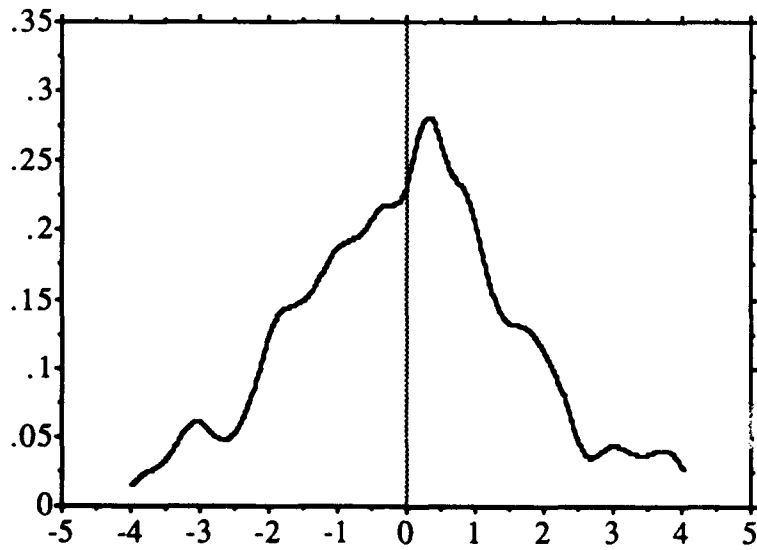


Figure 5.5: Normal Density Smoothing, Bandwidth = 0.2
10 Replications

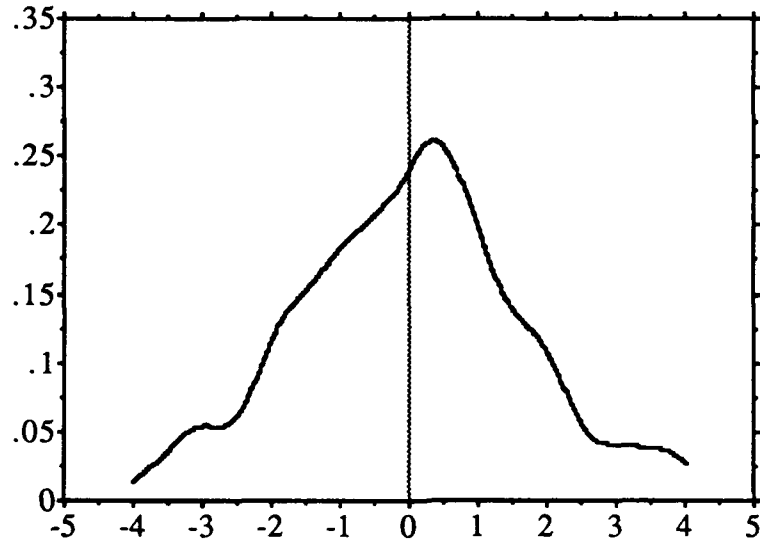


Figure 5.6: Normal Density Smoothing, Bandwidth = 0.3
10 Replications

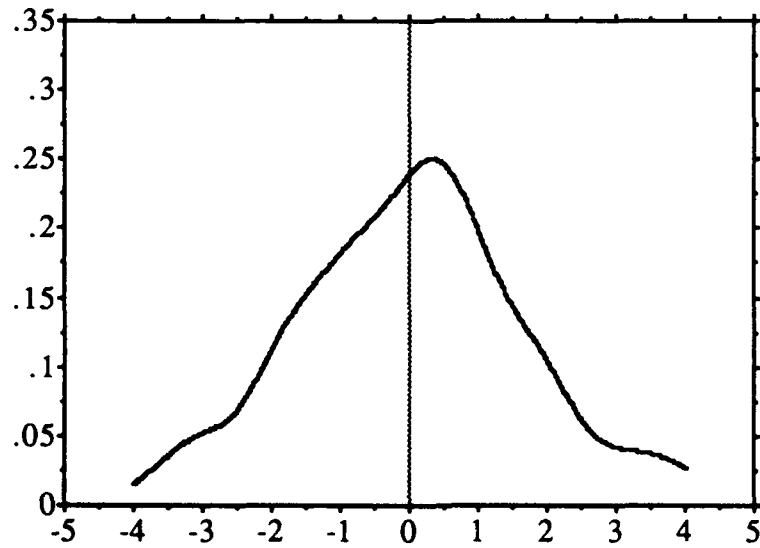


Figure 5.7: Normal Density Smoothing, Bandwidth = 0.4
10 Replications

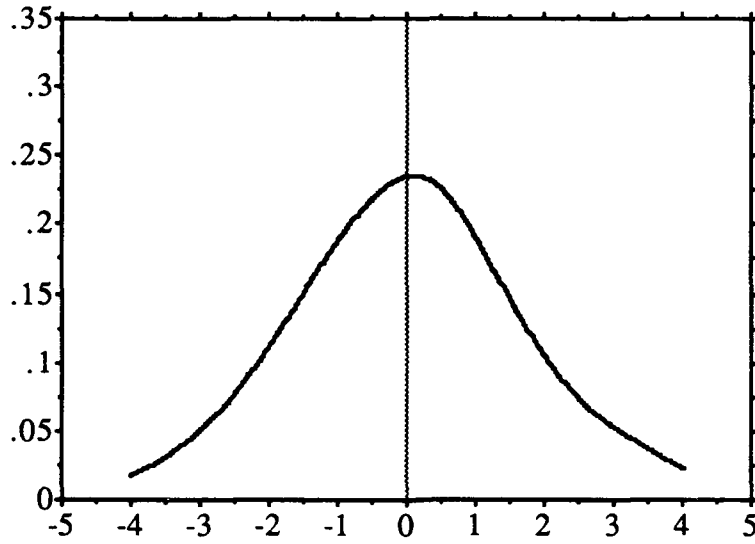


Figure 5.8: Normal Density Smoothing, Bandwidth = 0.5
10 Replications

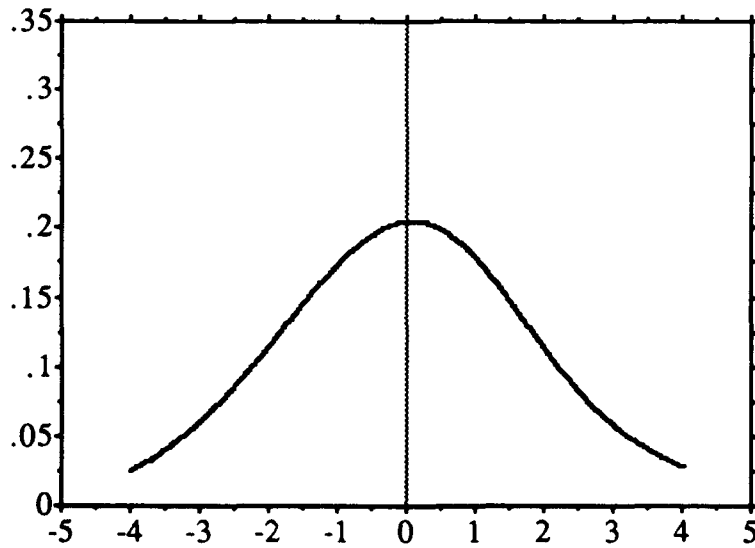


Figure 5.9: Normal Density Smoothing, Bandwidth = 1.0
10 Replications

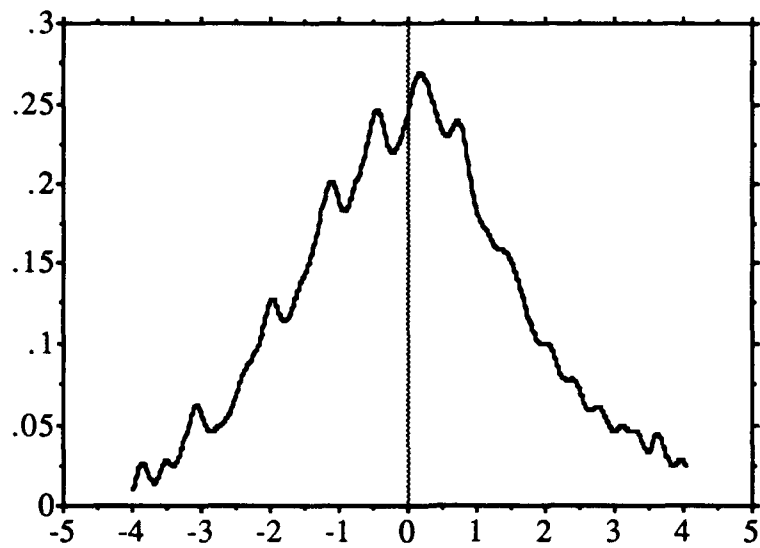


Figure 5.10: Normal Density Smoothing, Bandwidth = 0.10.
25 Replications

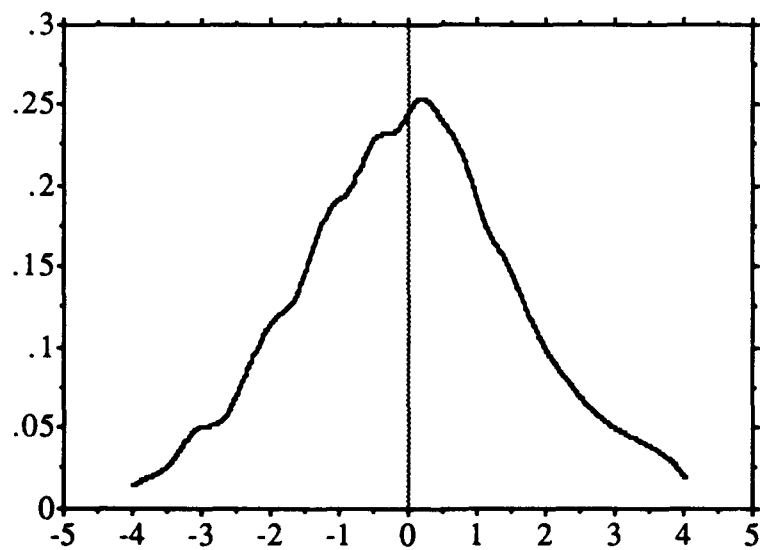


Figure 5.11 Normal Density Smoothing, Bandwidth = 0.20.
25 Replications

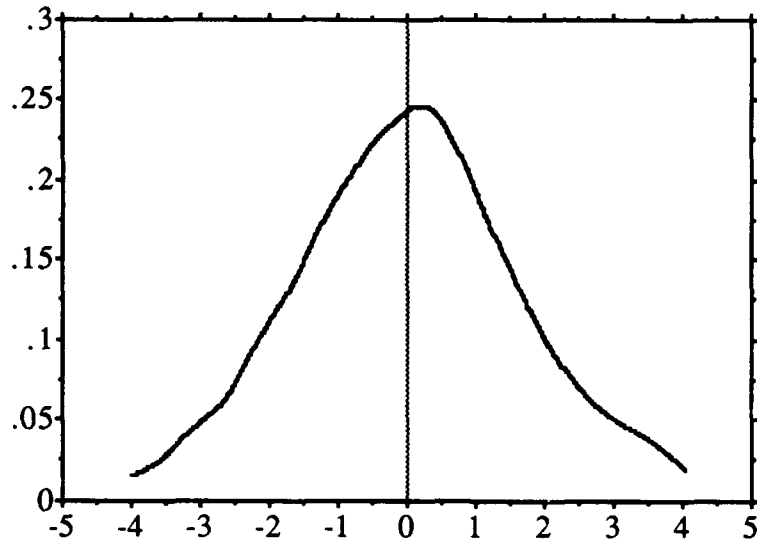


Figure 5.12: Normal Density Smoothing, Bandwidth = 0.30.
25 Replications

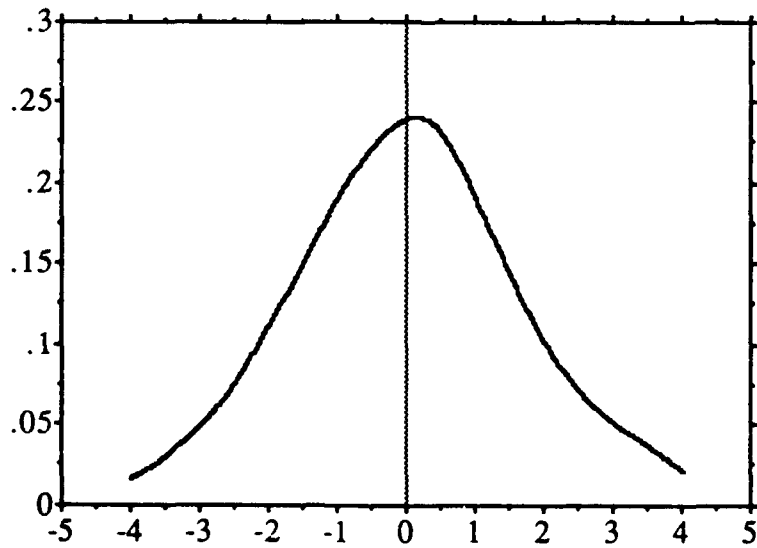


Figure 5.13: Normal Density Smoothing, Bandwidth = 0.40.
25 Replications

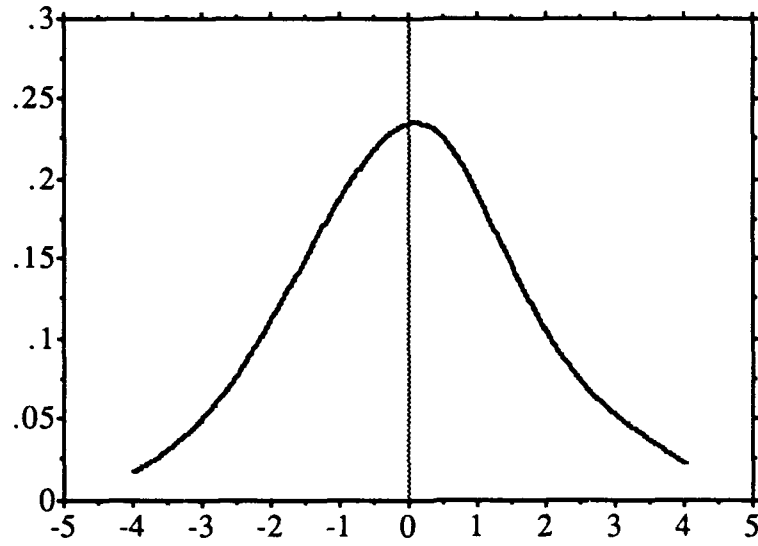


Figure 5.14: Normal Density Smoothing, Bandwidth = 0.50.
25 Replications

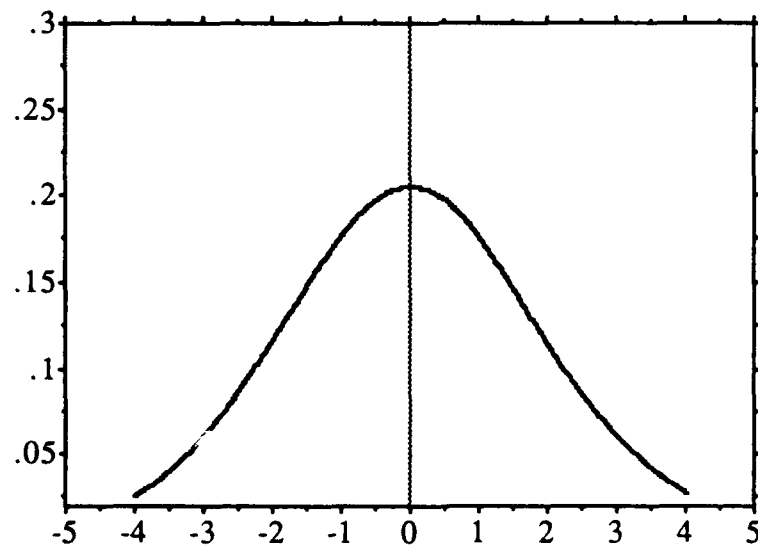


Figure 5.15: Normal Density Smoothing, Bandwidth = 1.0.
25 Replications

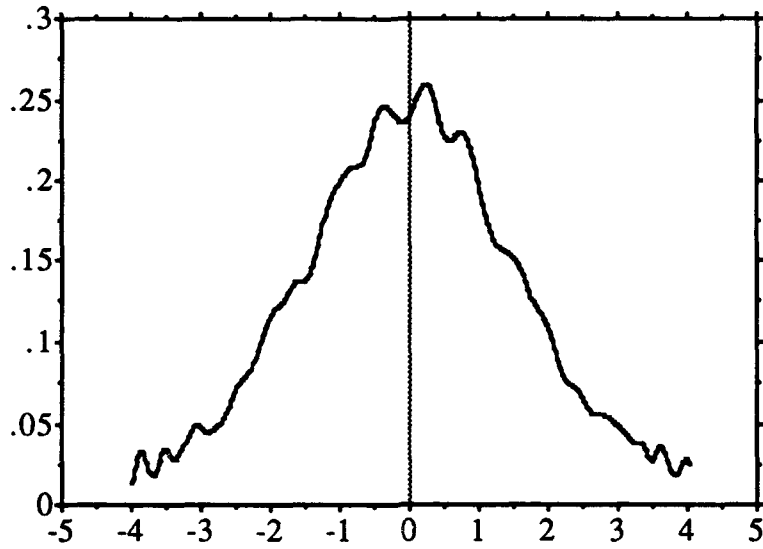


Figure 5.16: Normal Density Smoothing, Bandwidth = 0.10.
50 Replications

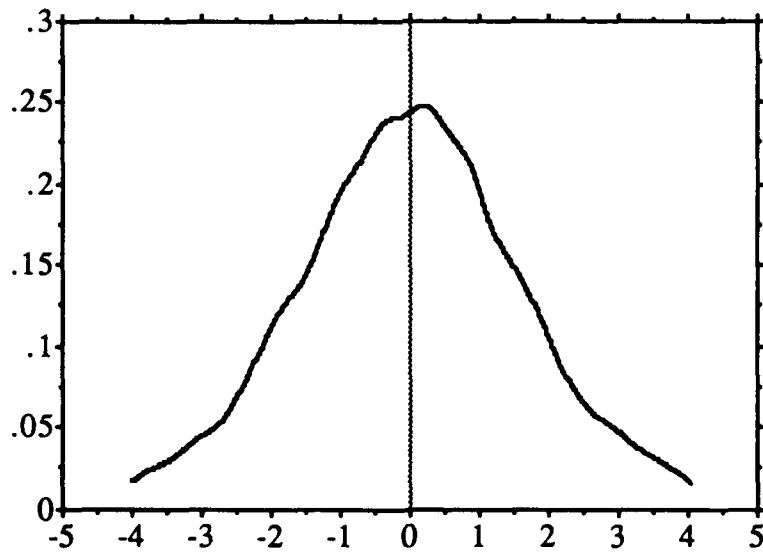


Figure 5.17: Normal Density Smoothing, Bandwidth = 0.20.
50 Replications

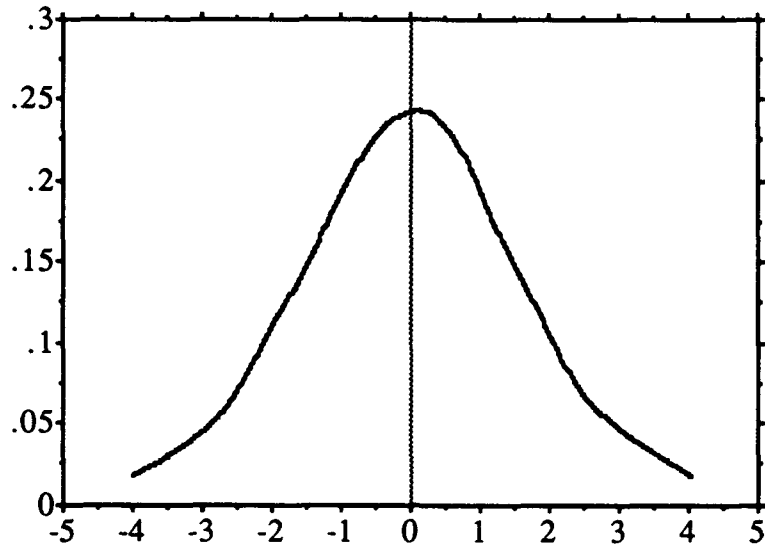


Figure 5.18: Normal Density Smoothing, Bandwidth = 0.30.
50 Replications

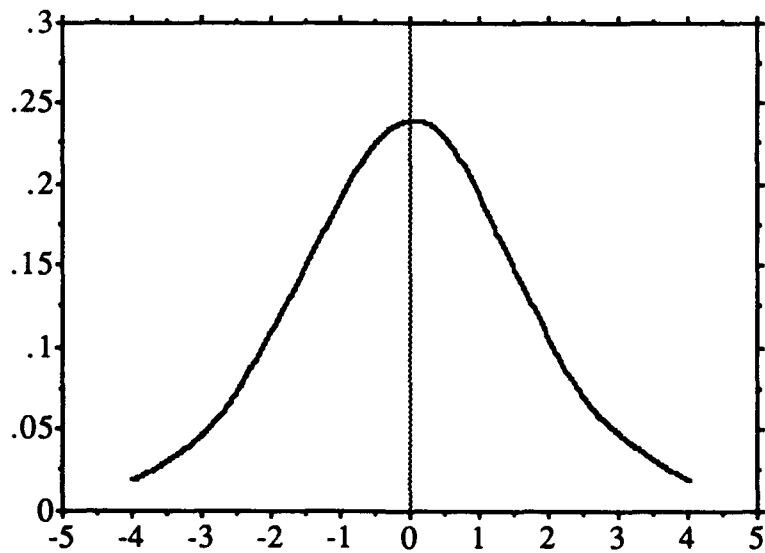


Figure 5.19: Normal Density Smoothing, Bandwidth = .40.
50 Replications

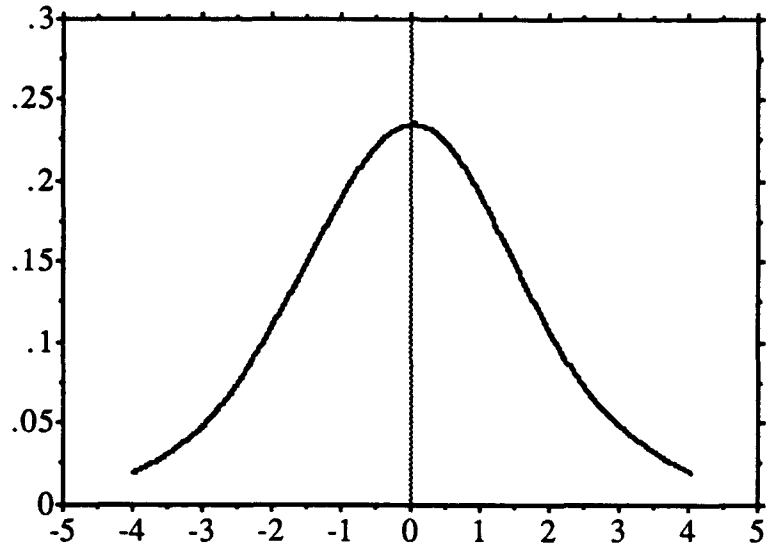


Figure 5.20: Normal Density Smoothing, Bandwidth = 0.50.
50 Replications

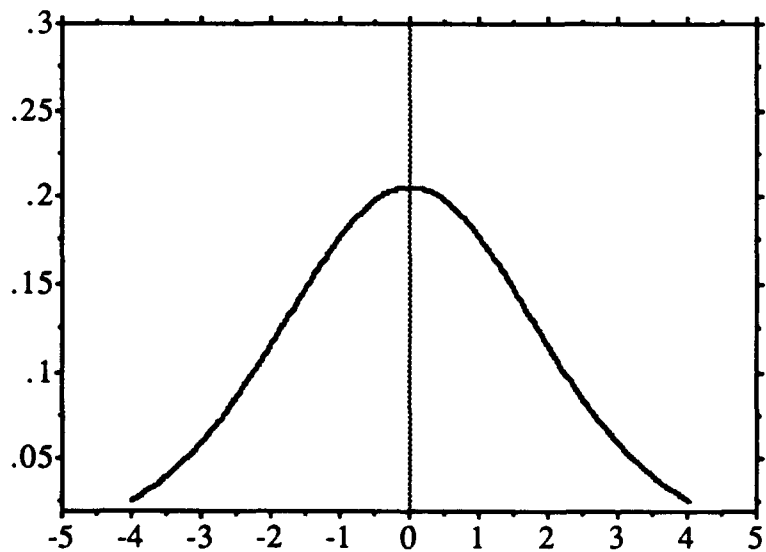


Figure 5.21: Normal Density Smoothing, Bandwidth = 1.0.
50 Replications

$$\text{MISE}(\hat{f}) = E \int \{\hat{f}(x) - f(x)\}^2 dx. \quad (5.3.2)$$

This can be further decomposed into two elements as shown by Silverman: the integrated square bias and the integrated variance.

$$\begin{aligned} \text{MISE}(\hat{f}) &= \int \text{MSE}_x(\hat{f}(x)) dx \\ &= \int \left\{ E(\hat{f}(x)) - f(x) \right\}^2 dx + \int \text{var } \hat{f}(x) dx \end{aligned} \quad (5.3.3)$$

Silverman uses this then to show that, for a given f , the bias term does not depend on the sample size directly, but depends only the the weight function, that is the kernel estimator. When the sample size is increased, the bandwidth will need to be adjusted accordingly in order to obtain the most appropriate smoothing function. This can easily be seen from the graphs shown in Figures 5.4 to 5.21 where the sample size ranges from 120 data points for Figures 5.4 to 5.9 to 600 data points for Figures 5.16 to 5.21.

Since the normal density kernel estimator satisfies the required properties of boundedness and bounded variation as well as the usual probability density conditions, we can use arguments given by Silverman (attributed to Bertrand-Retali). This leads to the statement that with probability 1,

$$\sup_x |\hat{f}(x) - f(x)| \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (5.3.4)$$

Based on this, we conclude that the kernel estimator used in the current study has asymptotic uniform consistency.

CHAPTER 6
APPLICATION OF THE MODEL TO A LIKERT SCALE

6.1 Introduction

Recently, Gerhard Tutz (1986) discussed the use of an ordered response in the Bradley-Terry type of paired comparison experiment. This extends the traditional model of preference or no preference to one where the strength of preference is also characterized. His approach uses random utility functions and develops specific parameters within the framework of a weighted least-squares method. In this chapter we will use the minimum mean squared error estimate for ϕ^{-1} obtained in Chapter 2, to develop ad hoc estimates for the θ_i 's used in the Likert Scale. A second approach to approximating the θ_i 's will employ weighted least squares.

The Likert Scale was originally developed by Rensis Likert (1932) and his associates at the Division of Program Surveys of the US Department of Agriculture. They applied sampling methods to the study of attitudes in

order to develop a technique for measuring individual's feelings. Likert was the first to apply the method of internal consistency to attitude measurement and use a summated rating system to eliminate the need for a group of judges to sort statements (Young, 1966). The Likert technique uses a series of questions concerning a particular area or problem. Each question requires an answer from one of five words such as strongly approve, approve, undecided, disapprove, or strongly disapprove. These can be replaced with strongly agree, agree, no opinion, disagree and strongly disagree or by almost always, always, frequently, occasionally, rarely, and almost never, as necessary to fit the nature of the question. A score is allotted to each of the possible choices and the total score for each subject is obtained by simply summing the score for each question.

The Likert Scale was compared to one developed by Thurstone (Edwards and Kenney, 1946) and was found to have higher reliability with fewer items to consider. The Likert technique is

also easier to administer and less time-consuming. This scale is used frequently in social science and educational areas, as well as in industry to measure worker satisfaction, etc. Combining the use of paired comparisons on a continuous response scale with the Likert technique could provide insight into the theory behind this widely used scale.

6.2 Ad Hoc Procedure for Applying the Model to the Likert Scale

In our original model ϕ mapped the real numbers onto $[0,1]$, but in applying ϕ to the Likert scale technique it will map the reals onto integer values. For example, for the five θ values:

$$\begin{array}{lll}
 \phi(V_{ij}) = -2 & \text{if} & \theta_0 < V_{ij} < \theta_1 \\
 \phi(V_{ij}) = -1 & \text{if} & \theta_1 < V_{ij} < \theta_2 \\
 \phi(V_{ij}) = 0 & \text{if} & \theta_2 < V_{ij} < \theta_3 \\
 \phi(V_{ij}) = 1 & \text{if} & \theta_3 < V_{ij} < \theta_4 \\
 \phi(V_{ij}) = 2 & \text{if} & \theta_4 < V_{ij} < \theta_5
 \end{array}$$

for $-\infty = \theta_0 < \theta_1 < \dots < \theta_4 < \theta_5 = \infty$

(6.2.1).

Here v_{ij} is the same logistic random variable as defined in Chapter 2 (see Definition 2). The θ_i 's could be symmetrically placed for convenience which results in $\theta_1 = -\theta_4$ and $\theta_2 = -\theta_3$. The two parameters, θ_0 and θ_5 , could be replaced by $-\infty$ and $+\infty$, respectively. This gives us only two parameters to solve for, θ_1 and θ_2 .

The original minimum mean squared error estimate for ϕ^{-1} (2.2.1) is modified to allow for the discontinuous case as follows:

$$Q = \sum_{i \neq j}^k \sum_{j}^k \left\{ \sum_{r=1}^p \left(\phi^{-1} \left(r - \frac{p+1}{2} \right) - (\mu_i - \mu_j) - F_{v_{ij}}^{-1} \left(\sum_{s=2}^r \frac{f_{ij}^{s-(p+1)/2}}{n_{ij}} \right) \right) \right\}^2$$

$$\hat{\phi}^{-1} \left(r - \frac{p+1}{2} \right) = \frac{\sum \sum F_{v_{ij}}^{-1} (\hat{q}_{r-(p+1)/2}^{ij})}{k(k-1)} \quad (6.2.2)$$

where $\hat{q}_{ij} = P(\phi(y) \leq r - \frac{p+1}{2}) = \sum_{s=1}^r \frac{f_{s-(p+1)/2}^{ij}}{n_{ij}}$, p is the total number of θ 's in the model, and r is the value of the i th θ parameter.

By using the estimate for $\phi^{-1}(y) = v_{ij}$, we can map any v_{ij} to its appropriate y value, similar to the example provided in 6.2.1. Simple estimates of the values of $\theta_1 \dots \theta_p$ are found as the

midpoints between each of the $\hat{\phi}^{-1}(y)$ values for $-\infty < y < \infty$.

Table 6.1 provides the values for ϕ^{-1} and its estimate as given in equation 6.2.2. The estimate is quite close to the value of the parameter for each of the combinations of numbers of treatments to be compared and number of replications. Using these five parameter values the midpoints are computed, thereby determining the θ_i 's. This results in the $\hat{\theta}_i$ values given in Table 6.1.

Since the θ parameters are to be symmetric with $\theta_1 = -\theta_4$ and $\theta_2 = -\theta_3$, we can average the estimates for θ_1 and $-\theta_4$, and for θ_2 and $-\theta_3$. In addition, we let $\theta_0 = -\infty$ and $\theta_5 = +\infty$. After averaging the values for each $\hat{\theta}_i$ in all of the cases displayed in Table 6.1, the values are presented in Table 6.2 (second column).

Table 6.1

ϕ^{-1} and θ Estimates for Five Parameters

| Number of Treatments | Replications | ϕ^{-1} | $\hat{\phi}^{-1}$ | i | $\hat{\theta}_i$ |
|----------------------|--------------|-------------|-------------------|---|------------------|
| 3 | 10 | -2 | -2.0515 | | |
| | | -1 | -1.0885 | 1 | -1.5700 |
| | | 0 | -0.2310 | 2 | -0.6598 |
| | | 1 | 1.0885 | 3 | 0.6285 |
| | | 2 | 1.8265 | 4 | 1.4575 |
| 4 | 10 | -2 | -2.0515 | | |
| | | -1 | -1.0517 | 1 | -1.5516 |
| | | 0 | -0.0817 | 2 | -0.5065 |
| | | 1 | 1.1334 | 3 | 0.5259 |
| | | 2 | 2.0688 | 4 | 1.150 |
| 4 | 10 | -2 | | | |
| | | -1 | -0.986 | 1 | -1.5728 |
| | | 0 | -0.027 | 2 | -0.5065 |
| | | 1 | 1.003 | 3 | 0.488 |
| | | 2 | 2.0194 | 4 | 1.150 |
| 5 | 10 | -2 | -2.0660 | | |
| | | -1 | -1.0796 | 1 | -1.5728 |
| | | 0 | -0.0778 | 2 | -0.5787 |
| | | 1 | 1.0962 | 3 | 0.5092 |
| | | 2 | 2.0194 | 4 | 1.5578 |
| 5 | 25 | -2 | -2.0639 | | |
| | | -1 | -1.0215 | 1 | -1.5427 |
| | | 0 | -0.0085 | 2 | -0.5150 |
| | | 1 | 1.0013 | 3 | 0.4960 |
| | | 2 | 2.0998 | 4 | 1.5456 |

Table 6.2
Average $\hat{\theta}_i$ Values

| $\hat{\theta}_i$ | Averaged Over All 5 Cases | Averaged Over Cases with 25 Replications | Practical θ_i |
|------------------|---------------------------|--|----------------------|
| $\hat{\theta}_1$ | -1.55 | -1.52 | -1.5 |
| $\hat{\theta}_2$ | -0.57 | -0.51 | -0.5 |
| $\hat{\theta}_3$ | 0.49 | 0.49 | 0.5 |
| $\hat{\theta}_4$ | 1.54 | 1.53 | 1.5 |

If, however, we average only the cases with 25 replications, we can see the $\hat{\theta}_i$ parameters drawing closer to a truly reflective situation, as expected. For practical purposes we could determine the θ_i 's to be those given in the final column of Table 6.2. These then would be the cut-off points to determine the appropriate y value for a given region. For example, if the comparison resulted in a $V_{ij}=-0.35$, it would be mapped to 0 on the Likert Scale, revealing no difference between the two items compared. If $V_{ij}=3.5$, then it would be mapped to 2, indicating a strong preference for the first item in the comparison. Some scaling adjustments may need to

be made since the paired comparison estimates are only accurate to a scale parameter. The example used in this discussion centers the scale around 0, for convenience. As shown in discussions in previous chapters, the values of $\delta_{ij} = \mu_i - \mu_j$ can change the relative position of the V_{ij} 's on the real line.

6.3 Weighted Least Squares Approach to Determining θ_i 's

A more theoretical approach to determining the θ_i 's involves the use of weighted least squares as described by Beaver (1977) for univariate models involving paired comparisons. In the current research we have a transformation that is linear in the log scale, which fits the model given in Beaver's article. Here the sample proportions are given by

$$P_1 = \frac{1}{1 + e^{-(\mu_i - \mu_j + \theta_1)}}$$

$$P_2 = \frac{1}{1 + e^{-(\mu_i - \mu_j + \theta_2)}} - \frac{1}{1 + e^{-(\mu_i - \mu_j + \theta_1)}}$$

$$\begin{aligned}
p_3 &= \frac{1}{1+e^{-(\mu_i - \mu_j + \theta_3)}} - \frac{1}{1+e^{-(\mu_i - \mu_j + \theta_2)}} \\
p_4 &= \frac{1}{1+e^{-(\mu_i - \mu_j + \theta_4)}} - \frac{1}{1+e^{-(\mu_i - \mu_j + \theta_3)}} \\
p_5 &= 1 - \frac{1}{1+e^{-(\mu_i - \mu_j + \theta_4)}} \quad (6.3.1).
\end{aligned}$$

By considering the following linear combinations of these proportions, the logit transformation can be used to produce a model linear in μ_i , μ_j and $\theta_1 \dots \theta_4$:

$$\begin{aligned}
p_1 &= \frac{1}{1+e^{-(\mu_i - \mu_j + \theta_1)}} \\
p_1 + p_2 &= \frac{1}{1+e^{-(\mu_i - \mu_j + \theta_2)}} \\
p_1 + p_2 + p_3 &= \frac{1}{1+e^{-(\mu_i - \mu_j + \theta_3)}} \\
p_1 + p_2 + p_3 + p_4 &= \frac{1}{1+e^{-(\mu_i - \mu_j + \theta_4)}} \quad (6.3.2).
\end{aligned}$$

Since the sum of the proportions must be 1, $\sum_{i=1}^5 p_i = 1$; hence the linear combinations in 6.3.2

represent 4 independent linear functions of the 5 multinomial probabilities. When the logit transformation $\ln \frac{p_i}{1-p_i}$ is applied to p_i , we have

$$\begin{aligned} \ln \frac{p_i}{1-p_i} &= \ln \frac{\frac{1}{e^{-(\mu_i - \mu_j + \theta_1)}}}{1 + \frac{1}{e^{-(\mu_i - \mu_j + \theta_1)}}} \\ &= \ln \frac{1}{e^{-(\mu_i - \mu_j + \theta_1)}} \\ &= \ln \left(e^{(\mu_i - \mu_j + \theta_1)} \right) \end{aligned}$$

so that

$$\ln \frac{p_i}{1-p_i} = \mu_i - \mu_j + \theta_1 \quad (6.3.3).$$

Similarly using logits, θ_2 , θ_3 and θ_4 can be expressed in a loglinear form as follows:

$$\begin{aligned} \ln \left(\frac{p_1 + p_2}{1 - (p_1 + p_2)} \right) &= \mu_i + \mu_j + \theta_2 \\ \ln \left(\frac{p_1 + p_2 + p_3}{1 - (p_1 + p_2 + p_3)} \right) &= \mu_i - \mu_j + \theta_3 \\ \ln \left(\frac{p_1 + p_2 + p_3 + p_4}{1 - (p_1 + p_2 + p_3 + p_4)} \right) &= \mu_i - \mu_j + \theta_4 \quad (6.3.4). \end{aligned}$$

Following the approach in Section 6.2, we consider the θ_i 's to be reflective with $\theta_1 = -\theta_4$ and $\theta_2 = -\theta_3$.

We again consider the pairs (i, j) $1 \leq i < j \leq t$ and (j, i) $1 \leq j < i \leq t$ with $t(t-1)$ pairs in all. The basic model given by Beaver for the loglinear situation is

$$F(\pi) = K \ln(A\pi) \quad (6.3.5).$$

where $\pi'_{ij} = (\pi_{1,ij}, \pi_{2,ij}, \dots, \pi_{5,ij})$ and

$$\pi' = (\pi'_{12}, \dots, \pi'_{t,t-1});$$

$$K = I \otimes K_0 \quad \text{and} \quad A = I \otimes A_0;$$

and \otimes denotes the right direct matrix product.

For an arbitrary pair (i, j) , with $i < j$, we have

$$A_0 \pi_{ij} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \end{bmatrix} \quad (6.3.6)$$

The above then results in

$$F_{ij}(p) = K_0 \ln(A_0 \underline{p}_{ij}) =$$

$$K_0 \begin{bmatrix} \ln p_1 - \ln (p_2 + p_3 + p_4 + p_5) \\ \ln (p_1 + p_2) - \ln (p_3 + p_4 + p_5) \\ \ln (p_1 + p_2 + p_3) - \ln (p_4 + p_5) \\ \ln (p_1 + p_2 + p_3 + p_4) - \ln (p_5) \end{bmatrix}$$

(6.3.7)

where

$$K_0 = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

This gives, for more specific pairs, the following:

$$F(\underline{p}_{12}) = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & \dots & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & \dots & 0 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & \dots & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \underline{\beta}$$

$$F(\underline{p}_{13}) = \begin{bmatrix} 1 & 0 & -1 & 0 & \dots & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & \dots & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & \dots & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & \dots & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \underline{\beta}$$

and similarly for all of the (i,j) pairs with $i < j$, up to $(t,t-1)$ as given below:

$$F(\underline{p}_{t-1,t}) = \begin{bmatrix} 0 & \dots & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & \dots & 0 & 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & \dots & 0 & 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & \dots & 0 & 1 & -1 & 0 & 0 & 0 & 1 \end{bmatrix} \underline{\beta}$$

Combining 6.3.5 thru 6.3.7 provides a design matrix in the form $F(\underline{p}) = X\underline{\beta}$ where

$$\underline{\beta} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \dots \\ \alpha_{t-1} \\ \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} \quad \text{and } \alpha_i = \mu_i - \mu_j, \text{ t is}$$

the number of
treatments being
compared.

The design matrix takes the form

$$X = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & \dots & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & -1 & 0 & \dots & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & \dots & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & \dots & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & \dots & 0 & 0 & 1 & 0 & 0 & 0 \\ & & & \dots & & & & & & \\ 0 & 0 & \dots & 1 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \dots & 1 & -1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Hence we now have a model linear in the parameters $\alpha_{1j} \dots \alpha_{t-1}$, $\theta_1 \dots \theta_4$, and the weighted least squares

technique as applied by Beaver is applicable. Without any loss of generality, we let $\mu_i - \mu_j = 0$, $1 \leq i < j \leq k$, since these parameters are location parameters. This simplification makes the computation slightly easier, although we could also solve for the locations parameter, $\hat{\mu}_i$, if these parameters were of interest at this point in the research.. In terms of the Likert scale, however, the θ_i 's are of major concern and we concentrate on this aspect of the model for the remainder of this section.

Using the data generated as explained in Section 5.1, we can determine the estimates for θ_1 , θ_2 , θ_3 , and θ_4 , using a weighted least squares analysis. Each comparison of (i,j) and (j,i) generates estimates for each of the θ_i 's from the replications at each comparison. These are then averaged over all of the pairs to obtain an estimate for θ_1 , θ_2 , θ_3 , and θ_4 . Table 6.3 shows the results of this simulation. The parameters are quite close to the original Likert Scale values with reflectivity of the parameters, that is $\theta_1 = -\theta_4$ and $\theta_3 = -\theta_2$. Of course with smaller sample sizes the

estimates are not quite as accurate and reveal less reflectivity.

Any V_{ij} in the interval $[\theta_2, \theta_3]$, would be mapped to $\frac{1}{2}$ in the y scale, implying no true difference exists between the items being compared. A V_{ij} in the interval $[\theta_3, \theta_4]$, would indicate a preference for the first of the items compared, while a V_{ij} in $[\theta_1, \theta_2]$ reveals a preference for the second item. Strong preferences are shown for the first or second item by obtaining a value of V_{ij} greater than θ_4 or less than θ_1 , respectively.

These findings support the interpretation of the Likert Scale and provide a more theoretical basis for the widespread conclusions frequently made from studies using this technique. Although simulations have revealed that the parameter estimates are valid for moderate sample sizes, it would be important to actually use these estimates in a practical study.

TABLE 6.3

| Treat-ments | Replica-tions | θ | Value |
|-------------|---------------|------------|---------|
| 4 | 10 | θ_1 | -1.5607 |
| | | θ_2 | -1.0517 |
| | | θ_3 | 1.1334 |
| | | θ_4 | 1.3327 |
| 3 | 25 | θ_1 | -2.0310 |
| | | θ_2 | -0.9892 |
| | | θ_3 | 1.0239 |
| | | θ_4 | 1.9753 |
| 4 | 25 | θ_1 | -2.0214 |
| | | θ_2 | -0.9857 |
| | | θ_3 | 1.0030 |
| | | θ_4 | 2.0270 |
| 5 | 25 | θ_1 | -2.0639 |
| | | θ_2 | -1.0215 |
| | | θ_3 | 1.0013 |
| | | θ_4 | 2.0898 |

BIBLIOGRAPHY

1. Beaver, R.J. (1974). "Locally asymptotically most stringent tests for paired comparison experiments." J Amer. Statist. Assn., 69:423-27.
2. Beaver, J.J. and Gokhale, D.V. (1975). "A model to incorporate within pair order effects in paired comparisons." Communications in Statistics, 4:923-39.
3. Beaver, R.J. and Rao, P.V. (1972). "The use of limit theorems in paired and triple comparison Model Building." Journal of Mathematical Psychology, 9:92-103.
4. Beaver, Robert J. (1977). "Weighted least squares analysis of several univariate models." J. Amer. Statist. Assn., 72:629-34.
5. Boos, Dennis D. (1979). "A differential for L-statistics." Annals of Statistics. 7:955-9.
6. Bradley, R.A. and Terry, M.E. (1952). "The rank analysis of incomplete block designs I. The method of paired comparisons", Biometrika, 39:323-45.
7. Bradley, R. A. (1976). "Science, statistics and paired comparisons", Biometrics, 32:-213-32.
8. Charnet R. and Beaver R.J. (1988). "Design and analysis of paired comparison experiments involving mixtures." J. Statist. Plann. Inf., 20:91-103.
9. Chernoff, Herman, Gastwirth, Joseph L., and Jones, M.V. (1967). "Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation." Annals of Statistics. 38:52-72.
10. David, H.A. (1969). The Method of Paired Comparisons, Second Edition, Griffin, London.
11. David, H.A. (1981). Order Statistics. Wiley Series in Probability and Statistics.

12. Davidson, R. R. (1970). "On extending the Bradley-Terry model to accomodate ties in paired comparison experiments." J Amer Statist. Assn., 65:317-28.
13. Edwards, Allen L. and Kenney Kathryn C. (1946). "A comparison of the Thrustone and Likert scale techniques of attitude scale construction." Journal of Applied Psychology 30:72-83.
14. El-Helbawy, A.T. and Ahmed E.A. (1984), "Optimal design results for 2 factorial paired comparison experiments." Commun.cation in Statistics, 13(22) 2827-2845.
15. Ford, L.R. Jr. (1957). "Solution of a ranking problem from binary comparisons." American Math. Monthly, 64(8):28-33.
16. Hogg, R.V. and Craig, A.T. (1978). Introduction to Mathematical Statistics. Fourth Edition. Macmillan Publishers.
17. Likert, Rensis. "A technique for the measurement of attitudes." Archives of Psychology Columbia University Press. No.140, 1932.
18. Luce, R.D. Individual Choice Behavior: A Theoretical Analysis. New York: Wiley, 1959.
19. Moestellar, F. (1951). "Remarks on the methods of paired comparisons: I. The least squares solution assuming equal standard deviations and equal correlations." Psychometrika, 16:3-9.
20. Pendergrass R.N. and Bradley R.A. (1960). "Ranking in Triple Comparisons." Contributions to Probability and Statistics. Stanford University Press: 331-51.
21. Rao, P.V. and Kupper, L.L. (1967). "Ties in paired-comparison experiments: A generalization of the Bradley-Terry Model." J. Amer. Statist. Assn., 62:194-204.

22. Shorak, Galen R. (1972). "Convergence of quantile and spacing processes with applications." Annals of Math. Stat. 43:1400-11.
23. Silverman, B. W. Density Estimation for Statistics and Data Analysis. London and New York: Chapman and Hall, 1986.
24. Sirotnik, B.W. and Beaver, R.J. (1984). "Paired comparison experiments involving all possible pairs." British Journal of Math. and Statist. Psychology, 37:22-33.
25. Stigler, S.M. (1974). "Linear functions of order statistics with smooth weight functions." Annals of Statistics, 2:676-693.
26. Thurstone, L.L. (1972). "A law of comparative judgement." Psychological Review, 34:273-86.
27. Tutz, G.A. (1986). "Bradley-Terry-Luce models with an ordered response." Journal of Mathematical Psychology, 30:306-16.
28. Young, Pauline V. Scientific Social Surveys and Research. New Jersey: Prentice-Hall, 1966.