Bradley-Terry models for paired comparisons incorporating judge variability

Jan William Jozef van Schaik
Iowa State University
INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or “target” for pages apparently lacking from the document photographed is “Missing Page(s)”. If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.

2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.

3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of “sectioning” the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.

4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.

5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.
van Schaik, Jan William Jozef

BRADLEY-TERRY MODELS FOR PAIRED COMPARISONS INCORPORATING JUDGE VARIABILITY

Iowa State University

University Microfilms International 300 N. Zeeb Road, Ann Arbor, MI 48106

Ph.D. 1985
PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark \(\checkmark\).

1. Glossy photographs or pages ______
2. Colored illustrations, paper or print ______
3. Photographs with dark background ______
4. Illustrations are poor copy ______
5. Pages with black marks, not original copy ______
6. Print shows through as there is text on both sides of page ______
7. Indistinct, broken or small print on several pages ______
8. Print exceeds margin requirements ______
9. Tightly bound copy with print lost in spine ______
10. Computer printout pages with indistinct print \(\checkmark\)
11. Page(s) ________ lacking when material received, and not available from school or author.
12. Page(s) ________ seem to be missing in numbering only as text follows.
13. Two pages numbered ________. Text follows.
14. Curling and wrinkled pages ______
15. Dissertation contains pages with print at a slant, filmed as received ________
16. Other___________________________________________________________________

________________________________________

University Microfilms International
Bradley-Terry models for paired comparisons 
incorporating judge variability

by

Jan William Jozef van Schaik

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

Major: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1985
TABLE OF CONTENTS

1. INTRODUCTION 1
   1.1 Choice Models 1
   1.2 Ranking Models for Paired Comparisons 4
   1.3 Extension of Paired Comparisons Models 15
   1.4 Notation 20

2. PARAMETER ESTIMATION AND HYPOTHESIS TESTING 24
   2.1 Introduction 24
   2.2 Maximum Likelihood Estimation for the Product Multinomial Distribution 26
   2.3 Two Types of Wald Statistics 33

3. JUDGE EFFECT MODELS BASED ON RANDOM CHOICE PROBABILITIES 44
   3.1 Introduction 44
   3.2 The Beta-Bernoulli Model 45
   3.3 The Beta-Binomial Model 49
   3.4 Parameter Estimation for the Beta-Binomial Model 54
   3.5 Hypothesis Testing for the Beta-Binomial Model 58
   3.6 A Reparameterized Beta-Binomial Model 67

4. A JUDGE EFFECT MODEL BASED ON RANDOM SCALE VALUES 71
   4.1 Introduction 71
   4.2 The Dirichlet-Binomial Model 71
   4.3 Parameter Estimation for the Dirichlet-Binomial Model 94
   4.4 Hypothesis Testing for the Dirichlet-Binomial Model 101
1. INTRODUCTION

1.1 Choice Models

The problem of modeling choice is very broad and complex and much work has been done by psychologists, economists and statisticians. There are two very different goals, however, behind the various modeling attempts. One is to set up a model that describes choice behavior. This area of research has been almost the exclusive realm of psychologists. Economists and statisticians, on the other hand, tend to develop choice models for the purpose of arriving at a ranking of the items to be compared. A good example of the two different modeling aims is the yearly dilemma of determining the best collegiate football team. Psychologists would be more interested in how someone came to select a team as best, whereas statisticians and economists would be interested in which team is best.

Descriptive choice models attempt to describe choice behavior and are characterized by an underlying choice strategy. This strategy specifies the relationships between the various choice situations in terms of the choice probabilities. For example, one method of selecting an item from a large choice set would be to first select a subset of items from the choice set, then choose a subset from that subset, etc., until in the end a single item is chosen. A good example of a situation where this choice strategy is applicable is at a restaurant when selecting an entree from the menu. This choice strategy requires that the choice probabilities satisfy $P_A(x) = P_A(B)P_B(x)$ for $x \in B \subseteq A$, where $P_A(x)$
denotes the probability of selecting item x from the choice set A.
Depending on whether the choice subsets are obtained through a process of selection or elimination, such choice schemes are called acceptance or discard mechanisms.

The fundamental criterion for comparing various descriptive models is how well choice behavior is described. One measure of a model's descriptive behavior is the plausibility of the conditions the choice probabilities must satisfy. For example, strong stochastic transitivity \( p_{ij} > \frac{1}{2} \) and \( p_{jk} > \frac{1}{2} \) implies \( p_{ik} > \max(p_{ij}, p_{jk}) \), where \( p_{ij} \) is abbreviated notation for \( P_A(i), A = \{i,j\} \) is a condition against which there is much empirical evidence (see, e.g., Luce (1959, §1.D.3)). Questions about the validity of a model may result if the choice probabilities must satisfy this condition as a consequence of the model.

Another consideration is the versatility of the descriptive model. In particular, can the model not only describe how items are chosen, but also provide a preference ranking of the items? One way to obtain the preference rankings is to consider the ranking probabilities as derived from the model. Tversky (1972), for example, used this strategy to describe how to obtain a ranking of the choice items using his model. Another approach would be to use some form of transitivity that is a consequence of the model, such as the strong stochastic transitivity of the previous paragraph, to produce a ranking. Still another method would be to associate with each item a parameter, included in the model, and rank the items according to their parametric values as specified by the model.
This last approach is more in line with ranking models than descriptive models. Ranking models are primarily concerned with providing a ranking of the choice items. To do this, most ranking models, like descriptive models, attempt to mathematically describe choice. However, unlike descriptive models, ranking models do not describe arbitrary choice situations, but rather a particular type of choice situation. Most ranking models work with the simplest situation possible, paired comparison experiments in which ties are not permitted.

In the next section, ranking models will be presented in more detail. Two ranking models for paired comparison experiments will be presented in particular, and examined in terms of their underlying assumptions. Methods of parameter estimation for both models, and their relationships with one another and other models will be described. Extensions of these paired comparison models is the topic of Section 3. Attention is focused on extensions that take into account the possibility of correlation between comparisons. The 3-component model described by Bock (1958), which includes a judge effect term, is examined in detail. In Chapters 3 and 4, other judge effect models are proposed. These models are compared in the later chapters with respect to how well they solve the problem of correlated comparisons. The final section of this chapter briefly explains the notation that is used throughout the remainder of this dissertation.

The first description of a paired comparison experiment that appears in the literature is that of Fechner (1860). He performed an experiment
in which the judge had to select which was the heavier of two objects based on which one felt heavier. Fechner was more interested in describing the variability of an object's weight as perceived by a judge than in developing a ranking scheme of the objects based on the judge's observations. Zermelo (1929), on the other hand, did investigate the problem of ranking. In particular, he looked at the problem of ranking participants in a chess tournament in which not every player had met every other player. His solution to the problem was rediscovered by Bradley and Terry (1952), whose approach will be described in the next section.

Much has been written about choice modeling in general and paired comparisons in particular. Two very complete overviews of the work done in paired comparisons are David (1963) and Bradley (1976). Both works consider the problem from the ranking point of view. Luce and Suppes (1965) and Luce (1977) give good summaries of the work done from a descriptive model viewpoint.

1.2 Ranking Models For Paired Comparisons

In the literature, there is a wide range of ranking methods and models for paired comparison experiments. Much of the early work can be attributed to Kendall and Babington Smith (1940). They used combinational methods to enumerate all possible outcomes of various statistics and develop the distribution theory associated with these statistics. The results were used to perform tests concerning the items (see David (1963, Chapters 2 and 3)). Guttman (1946) developed a ranking system similar to
discriminant analysis and Scheffé (1952) applied analysis of variance methods to experiments in which a preference scale (e.g., a 7-point scale with points ranging from strongly dislike to strongly like) was used, to come up with a ranking scheme. Much work has also been done in tournament problems using various scoring techniques to obtain rankings (see, e.g., Moon and Pullman (1970)).

In this section, ranking models, as opposed to ranking methods, will be presented. These are essentially two types of models, as described by Block and Marschak (1960): constant utility models and random utility models. Both model types have a function, called the utility function, which associates with each item a value called the utility. The utilities are used to impose constraints on the choice probabilities. For constant utility models, the utility function assigns a fixed utility to each item, whereas for random utility models, an item's utility is a random variable. Because some constant utility models can be viewed as random utility models, and vice versa, as Block and Marschak (1960) showed, the distinction between constant and random utility models is definitional rather than practical.

A good example of a constant utility model is the model described by Bradley and Terry (1952). They were interested in obtaining a ranking of a set of k items compared by a group of N judges, with each judge performing all \( \binom{k}{2} \) possible paired comparisons. They associated utility \( \pi_i \), which they called a scale value, with the i-th item and postulated that \( p_{ij} = \pi_i (\pi_i + \pi_j)^{-1} \) (this condition will be referred to as the
 Bradley-Terry condition). For this model, all comparisons performed by a judge are assumed to be done independently so that the probability of an observed series of comparisons for a judge is given by

\[
\prod_{i<j} \left[ \frac{\pi_i}{\pi_i + \pi_j} \right]^{X(i,j)} \left[ \frac{\pi_j}{\pi_i + \pi_j} \right]^{X(j,i)} = \prod_{i} \frac{x_i}{\pi_i} / \prod_{i<j} (\pi_i + \pi_j) ,
\]  

(1.2-1)

where \( x_{i,j} \) is 0 or 1 depending on whether item \( j \) or \( i \) is preferred and \( x_i \) is the total number of times item \( i \) is preferred by a judge (i.e., \( x_i = \sum_j x_{i,j} \)), where the symbol \( \sum_j \) is used to indicate the summation over all \( k-1 \) items \( j \neq i \).

Dykstra (1960) extended the model to the case where the \((i,j)\) comparison is performed \( N_{(i,j)} \) times, \( N_{(i,j)} > 0 \) (in the Bradley-Terry model \( N_{(i,j)} = N \)). The kernel of the likelihood function is then

\[
\prod_{i} \frac{a_i}{\pi_i} / \prod_{i<j} (\pi_i + \pi_j)^{N_{(i,j)}},
\]  

(1.2-2)

where \( a_i \) is the total number of times item \( i \) is preferred in the entire experiment, i.e., \( a_i = \sum_{t=1}^{N} x_i^t \), and \( x_i^t \) is the value of \( x_i \), defined above, for judge \( t \).

Bradley and Terry proposed using maximum likelihood estimation to obtain estimators of the scale values \( \pi_i \). The maximum likelihood equations, using the extended model, are
\[ \frac{a_i}{\pi_i} = \sum_j \frac{N(i,j)}{\pi_i + \pi_j} \quad j = 1, \ldots, k \]

and

\[ \sum_i \pi_i = 1 . \]

The restriction \( \sum \pi_i = 1 \) is introduced because \( \pi = (\pi_1, \ldots, \pi_k) \) is unique up to a multiplicative constant. Ford (1957) described the same model as Dykstra but was interested in the convergence properties of an iterative procedure used to solve the maximum likelihood equations. He showed that the procedure converges if for every possible partition of the set of all items into two subsets, some item in the first subset is preferred to some item in the second subset and vice versa. This condition is equivalent to stating that:

1. every item is compared to every other item either directly \( (N(i,j) > 0) \) or indirectly \( (N(i,k_1) > 0, N(k_1,k_2) > 0, \ldots, N(k_s,j) > 0) \), and

2. there is no item that is preferred over every other item or that is never preferred.

The first condition prevents the possibility of having separate groups of items with no basis for comparing items from different groups. The second condition excludes the possibilities that \( \pi_i = 1 \) or 0 for some item \( i \), which is indicative of item \( i \) being in a class all of its own relative to the other items.
Bradley and Terry also proposed tests of hypotheses for the scale values based on the likelihood ratio statistic $A$. For small sized experiments (few items, $N_{(i,j)}$ small) the exact distribution of $A$ was tabulated by Bradley (1954), whereas, for large sized experiments, the result that $-2\ln A$ is asymptotically distributed as a chi-squared random variable is used to obtain significance levels. A log-likelihood ratio test statistic can also be used to test the goodness-of-fit of the Bradley-Terry condition. Bradley (1955) investigated the large sample properties of the maximum likelihood estimator of $\pi$ (see Davidson and Bradley (1970) also) and using the asymptotic normality of the maximum likelihood estimator, he was able to come up with tests for contrasts involving the elements of $\pi$.

The Bradley-Terry condition has an intuitive appeal that is further enhanced by the work of Luce (1959). Luce was interested in describing choice behavior and proposed a descriptive model which is summarized in what has been called the choice axiom. This choice axiom is as follows:

\textbf{Choice axiom} Let $T$ be a finite subset of $U$ (the universal set of all items) such that, for every $S \subseteq T$, $P_T(S)$ is defined.

(i) If $p_{xy} \neq 0,1$ for all $x,y \in T$, then for $R \subseteq S \subseteq T$

$$P_T(R) = P_S(R)P_T(S).$$

(ii) If $p_{xy} = 0$ for some $x,y \in T$, then for every $S \subseteq T$

$$P_T(S) = P_{T-\{x\}}(S-\{x\}).$$

The first postulate states that the choice probability can be determined using conditioning if there is no pairwise perfect discrimination. This
is nothing other than an acceptance mechanism approach to choice restricted to the situation of imperfect discrimination. Luce notes that the postulate is more likely to be true in simple choice situations (cases where the items are easily distinguished) than in complex situations. The second postulate deals with the case of perfect discrimination. Luce noted that if the first postulate was assumed to hold for both perfect and imperfect discrimination, then the following situation results: if $x$ is always preferred to $y$ and $y$ is sometimes preferred to $z$, then $x$ is always preferred to $z$ (i.e., $p_{xy} = 1, p_{yz} > 0$ implies $p_{xz} = 1$).

It is not hard to find realistic situations in which this type of stochastic transitivity does not hold (see Luce (1959, §1.D.2), for two examples). To avoid such inconsistencies, Luce introduced the second postulate.

Luce has proven a variety of probabilistic consequences from his model. The most important of these is the implication of the existence of a positive ratio scale $\pi$ such that

$$P_A(x) = \pi_x/\sum \pi_y,$$

for $x \in A$. For the choice situation where the set $A$ contains two items, this is the Bradley-Terry condition.

Most of the models proposed in the literature as ranking models for paired comparisons experiments have been random utility models. An example of a random utility model is any model which satisfies

$$p_{ij} = P(W_i > W_j),$$

where $W_i$ and $W_j$ are the random variables describing
the utilities associated with items i and j. Fecher (1860) presented what is essentially a random utility model with normal random variables as utilities. Thurstone (1959) proposed the same model in 1927 but was more explicit in his presentation. He introduced the idea that an item's stimulus ($S_i$) as perceived by a judge (e.g., the taste of a pudding) is not constant, but rather, random. He proposed using a normal random variable with mean $S_i$ and variance $\sigma_i^2$ to model the perceived stimulus of item i. If $X_i$ is the stimulus of item i as perceived by a judge, then Thurstone's approach is given by $X_i = S_i + \epsilon_i$, where $\epsilon_i$ is $N(0,\sigma_i^2)$ and models the perception deviations associated with item i. Thurstone went on to assume that a judge must prefer the item with the largest perceived stimulus, so that

$$P_{ij} = P(X_i > X_j) = \Phi\left(\frac{S_i - S_j}{\sigma_i^2 + \sigma_j^2 + 2\rho_{ij}\sigma_i\sigma_j}\right)$$

(1.2-4)

where $\Phi$ is the cdf for the standard normal distribution and $\rho_{ij}$ is the correlation between the perceived stimulus for items i and j. Thurstone, using this basic model, defined five cases. The case I model assumes that $\rho_{ij} = \rho$ for all i,j, so that the model is not overparameterized and parameter estimates can be obtained. The case II model is exactly the same in terms of the parametric model, but assumes that the normal random variable associated with an item is the same for all judges. This permits the use of the same model for identical comparisons made by different judges, and so results of different judges can be pooled
together to facilitate in parameter estimation (this assumption also holds for case III, IV and V models). The case III model assumes \( \rho_{ij} = 0 \) for all \( i, j \), and the case IV model further assumes that \( \sigma_i = \sigma_j + d \), where \( d \) is small. This last assumption is used to simplify parameter estimation since it permits \( (\sigma_i^2 + \sigma_j^2)^{\frac{1}{2}} \) to be rewritten as \( \sqrt{2} (\sigma_i + \sigma_j) \). The case V model assumes \( \rho_{ij} = 0 \) for all \( i, j \), and \( \sigma_i = \sigma \) for all \( i \). This model, also known as the Thurstone model, is the most widely used of the five models.

Thurstone proposed to estimate the various model parameters by simply calculating \( \Phi^{-1}(\hat{\theta}_{ij}) \), where \( \hat{\theta}_{ij} \) is the observed value of \( \theta_{ij} \), and solving \( \Phi \) model equations. For example, for the Thurstone model the equations to be solved are

\[
\Phi^{-1}(\hat{\theta}_{ij}) = \frac{S_i - S_j}{\sqrt{2} \sigma}.
\]

(1.2-5)

Thurstone noted that the item stimulus and variance parameters are estimable up to a location and scale factor. Consequently, he took \( S_i = 0 \) and \( \sigma_i^2 = 1 \), for an arbitrary item \( i \), before solving the model equations. Using the alternative restriction \( \Sigma_i S_i = 0 \), rather than \( S_i = 0 \), to fix the location, the solution to the equations (1.2-5) can be found by summing those equations containing a common parameter. The solution is given by \( \hat{S}_i = k^{-1} \Sigma_j \Phi(\hat{\theta}_{ij}) \). Mosteller (1951a) noted that the assumption \( \rho_{ij} = 0 \) for all \( i, j \), is more restrictive than it needs to be for the case V model in terms of parameter estimation. He proposed a model, which is known as the Thurstone-Mosteller model, in which he relaxed the assumption to \( \rho_{ij} = \rho \) for all \( i, j \). In a later
paper (Mosteller (1951b)), he proposed a goodness-of-fit test for the Thurstone-Mosteller model based on the difference between the observed and predicted preference proportions for each comparison. Sadaswan (1982) proposed an improvement of Thurstone's estimation technique. MacKay and Chaiy (1982) and Gibson (1953) have worked on the problem of estimating the parameters for the Thurstone case III and case IV models.

The basic idea of Thurstone to use a random variable to model perceived item stimulus has been used in many other models. Eisler (1965) describes a model in which the perceived item stimulus random variable has a variance that is a function of the mean. Bradley (1953) noted that if the difference distribution of $X_i$ and $X_j$ is taken to be a logistic distribution rather than a normal distribution, as in Thurstone's models, then the Bradley-Terry condition results, i.e.,

$$\Pr_{ij} = \int_{-(\ln \pi_i - \ln \pi_j)}^{\infty} e^{-x(1 + e^{-x})} \frac{-2}{e^{-x} + e^{-x}} \, dx = \frac{\pi_i}{\pi_i + \pi_j}.$$  

This is an example of the equivalence between constant utility models and random utility models. Note that $\ln \pi_i$ is a location parameter value associated with item $i$ which corresponds to $S_i$ in Thurstone's model. The normal and logistic distributions are very similar, and early researchers found that choice probabilities predicted by the Bradley-Terry model and the Thurstone model are nearly the same.
Yellot (1977) investigated the question of uniqueness of the choice probabilities, as specified by the Bradley-Terry and Thurstone models, in terms of the distribution of the random variable modeling the perceived item stimulus. He showed (using Cramer's result that if $X_1$ and $X_2$ are independent and $X_1 + X_2$ is normally distributed, then $X_1$ and $X_2$ are normally distributed) that the Thurstone model is unique, that is, the Thurstone choice probabilities as specified by (1.2-4) are satisfied if and only if the perceived item stimulus random variables are normally distributed. The uniqueness of the Bradley-Terry model is not quite as strict. Yellot showed that the Bradley-Terry condition holds if and only if the difference distribution of the two stimuli random variables is logistic. He was able to show that a variety of distributions exists for which the difference between two random variables has a logistic distribution. The most common is the double exponential distribution, also known as the extreme value distribution.

This distribution had been mentioned earlier in conjunction with the Bradley-Terry model. Lehmann (1953), while investigating the power of rank tests, showed that the Bradley-Terry condition holds if the $X_i$ are independently distributed as $F(x; \eta_i)$. This result was used by Davidson (1969) for the case that $X_i$ is distributed as a double exponential. In particular, if $F(x; \eta_i) = \exp(-\exp(-(x-\eta_i)))$, then the Bradley-Terry condition follows since $F(x; \eta_i) = (F(x,0))^{\eta_i}$. From this (as well as from Yellot), it follows that if it can be assumed that an item's perceived stimulus can be modeled using the double exponential distribution given above, then the choice probabilities satisfy the Bradley-
Terry condition.

Justification for modeling the perceived item stimulus as randomly distributed with a normal or double exponential distribution was given by Thompson and Singh (1967). They noted that perception depended on the number of sensory receptors stimulated and the frequency of stimulation. The sensory signals sent to the brain can be processed in a variety of ways to come up with an overall sensation perception. Thompson and Singh looked at three different processing mechanisms:

1. the sum (or average) of the sensory signals,
2. the median (or some other quantile) of the sensory signals, and
3. the maximum sensory signal.

In their paper, they showed that as the number of receptors stimulated and/or the frequency of stimulation becomes very large, mechanisms 1 and 2 result in the overall perceived sensation being approximately distributed as a normal random variable. For the third mechanism, the perceived item stimulus unfortunately has a limiting distribution that depends on the distribution of the input sensory signals. If the input sensory signals have a distribution that is a member of the exponential family, then the limiting distribution is the double exponential distribution. However, Thompson and Singh did show that the choice probabilities satisfy the Bradley-Terry condition regardless of the distribution of the input sensory signals.

Audley (1960) proposed a model along similar lines. He defined a preference as occurring when K successive stimuli have been perceived from
a single item. He associated with each item a parameter that determined
the probability of a perceived stimulus from that item in a small interval
of time \((t, t + \Delta t)\) and further assumed that the stimulus probabilities are
independent of one another. For the case of \(K = 1\), the choice proba-
bilities for pairwise comparisons satisfy the Bradley-Terry condition.

The underlying thread of the paired comparisons ranking models that
have been presented is the Bradley-Terry condition. It is a simple and
intuitive condition which can be derived from a variety of situations.
Although the Thurstone model can never exactly satisfy the Bradley-Terry
condition, as was proven by Yellot, it is a close approximation to it.
As such, the Bradley-Terry model can be viewed as the basic paired
comparisons model. In the next section, various extensions of the model
are presented.

1.3 Extension of Paired Comparisons Models

Several extensions of the basic paired comparisons model have been
proposed. Most of these will be mentioned, but only those models which
in some way take into account that comparisons made by the same judge
may be correlated will be discussed.

As alluded to previously, some work in ranking models has been con-
cerned with choice situations other than paired comparisons. Pendergrass
and Bradley (1960) have considered various models for the triple choice
situation along the lines of the Bradley-Terry model. Luce's generaliza-
tion of the Bradley-Terry condition has been used for the triple compari-
son situation as well as other multiple choice situations. Audley's
model, although presented by Audley for paired comparison experiments, is also applicable to multiple choice situations. Similarly, Thurstone's approach can be applied to the multiple choice situation as noted by Luce (1959) (see also Yellot (1977)).

One obvious extension of paired comparisons models is to permit ties. Rao and Kupper (1967) and Davidson (1970) have introduced such models along the lines of the Bradley-Terry model. Glenn and David (1960) have proposed a model permitting ties which is a modification of the Thurstone-Mosteller model. Models which handle the special situation in which the items are factorial treatment combinations have been proposed by Abelson and Bradley (1954) and El-Helbawy (1974). El-Helbawy's model is rather clumsy and Bradley and El-Helbawy (1976) proposed a simpler model. Imrey et al. (1976) proposed a ranking method (not a true model) using logits which can be used in experiments with factorial treatment combinations. This method can also be used for multivariate paired comparisons experiments. For this problem, an extension of the Bradley-Terry model has been proposed by Davidson and Bradley (1969, 1970).

One type of model that has received very little attention is the nonindependent utility model. Both Thurstone's model and the Bradley-Terry model (viewed as a random utility model with the perceived item stimuli being distributed, for example, as double exponential random variables) assume the perceived item stimulus random variables are independently distributed. The Thurston-Mosteller model, on the other hand, assumes that these random variables are correlated. However, as
Hosteller (1951a) noted (and Yellot (1977) formally proved), with respect to parameter estimation the Thurstone and Thurstone-Mosteller models are equivalent. Bradley (1965) proposed a model, in the form of a Lehmann model, which introduces a correlation between the perceived item stimulus random variables. Bradley's model is

$$P(X_i < u, X_j < v) = \frac{s}{\sum p_{\alpha} K(u)^{-\alpha} K(v)^{-\alpha}}$$

where \( p_\alpha > 0 \), \( \sum p_\alpha = 1 \), \( c_\alpha > 0 \), \( K \) is any distribution function, and \( s \) is an arbitrary positive integer. He showed that this model satisfies the Bradley-Terry condition, so that, by Yellot, \( X_i - X_j \) has a logistic distribution (Davidson (1969) also noted this). Tversky (1972) (see also Corbin and Marley (1974)) showed that the descriptive model that he proposed for choice behavior is equivalent to a nonindependent random utility model, however he did not specify the form of the model.

Bock (1958) (see also Bock and Jones (1968, §6.7)) introduced a model which included factors modeling the variability of and correlation among comparisons made by a judge. His model for the perceived stimulus of item \( i \) by judge \( h \) in the \( t \)-th comparison is

$$X_{hit} = S_i + \nu_{hi} + \epsilon_{hit},$$

where \( \nu_{hi} \) and \( \epsilon_{hit} \) are random variables associated with the item's stimulus for a randomly selected judge and the deviation in perceived sensation for a given judge, respectively. The first random variable models the variability among judges in perception of an item's stimulus, e.g., some judges will find Coke sweeter than others. The second random variable models the variability among repeated perceptions of an item's
stimulus made by a single judge, e.g., a judge's perception of the
sweetness of Coke will depend on the state of the judge's taste buds at
the moment of tasting.

For a randomly selected judge, Bock assumed the random vector
$v_{hi} = (v_{h1}, \ldots, v_{hk})$ to be normally distributed with mean zero, variance
$\sigma^2$ and correlation $\rho$, and that the $\varepsilon_{hit}$ are iid normal random variables
with mean zero and variance $\delta^2$. Therefore, the difference distribution
$Y_{hij} = X_{hit} - X_{htj}$ (the subscript $t$ is dropped since the items specify
the comparison) has a normal distribution with mean $S_i - S_j$ and variance
$2\sigma^2(1-\rho) + 2\delta^2$. Like the Thurstone-Mosteller model, Bock's model
specifies that the perceived item stimulus random variables are correlated.
Bock's model goes beyond the Thurstone-Mosteller model because it specifies
that $\text{Cov}(Y_{hij}, Y_{hik}) = \sigma^2(1-\rho)$ and $\text{Cov}(Y_{hij}, Y_{hkl}) = 0$, whereas, the variance
covariance matrix for the vector $Y' = (Y_{h12}, \ldots, Y_{hij}, \ldots, Y_{h(k-1)k})$ is
$2\sigma^2(1-\rho)I$ for the Thurstone-Mosteller model.

Bock's model allows for an improvement in the estimation of $S_i$.
Weighted least squares can now be used to estimate the $S_i$ rather than
simple least squares, which is equivalent to the method used for the
Thurstone-Mosteller model (see Mosteller (1951a)). In a further effort
to obtain better estimators, Bock did not work with the raw proportions,
as did Mosteller, but rather with transformed proportion (the transformation
he used was the arcsin square root transformation which he chose
because of its variance stabilizing properties). The transformed
variables all have variances that are approximately equal to $N^{-1}$ where
N is the number of judges, and have a common correlation \( \rho \). This correlation is unknown and Bock, rather than estimating it and using asymptotic methods to derive tests, derived bounds for \( \rho(0 < \rho < 1/3) \) which were used to determine worst case type tests. Bock apparently was more interested in showing how correlations affect goodness-of-fit tests and tests of hypotheses for the \( S_i \), than in developing appropriate tests. He was able to show that ignoring the correlation (i.e., assuming \( \rho = 0 \)) had as a consequence that the goodness-of-fit results would indicate a better fitting model than is actually the case. Mosteller and others had noted that the Thurstone-Mosteller model seemed to fit data too well but did not provide an explanation. For tests of hypothesis, ignoring the correlation results in tests which tend to find significant differences in item stimulus values \( S_i \) too easily.

In Chapters 3 and 4, judge effect models more closely related to the Bradley-Terry model are presented and examined. The objective is to provide a better description of paired comparison experiments than that which can be attained with the Bradley-Terry model. The approach of incorporating a judge related random variable to achieve a better model appears to be promising in light of the results of Bock. In Chapter 3, the judge effect is taken into account through the choice probabilities, whereas, in Chapter 4, the judge variability is modeled in terms of variation in the scale values associated with the items.
1.4 Notation

A summary of the notation used throughout the remainder of this dissertation is given in this section. In Chapter 2, underlying the discussion is a product multinomial setting with T groups and C subgroups within each group. The notation used in this chapter is as follows:

- $n_t$ \( \text{the number of events in group } t \),
- $n$ \( \text{the } T \times 1 \text{ vector } (n_1, \ldots, n_T)' \),
- $n_t$ \( \text{the } C \times 1 \text{ vector the elements of which all equal } n_t \),
- $\tilde{n}$ \( \text{the } T \times C \times 1 \text{ vector } (n_1, \ldots, n_C)' \),
- $x_i^{(t)}$ \( \text{the number of occurrences of an event in the } i\text{-th subgroup of the } t\text{-th group} \),
- $\bar{x}_i^{(t)}$ \( \text{the } C \times 1 \text{ vector } (x_1^{(t)}, \ldots, x_C^{(t)})' \),
- $\tilde{x}$ \( \text{the } T \times C \times 1 \text{ vector } (\bar{x}^{(1)}', \ldots, \bar{x}^{(T)}')' \),
- $p_i^{(t)}$ \( \text{the probability of an event occurring in the } i\text{-th subgroup of the } t\text{-th group} \),
- $p^{(t)}$ \( \text{the } C \times 1 \text{ vector } (p_1^{(t)}, \ldots, p_C^{(t)})' \),
- $\hat{p}$ \( \text{the } T \times C \times 1 \text{ vector } (\hat{p}^{(1)}', \ldots, \hat{p}^{(T)}')' \),
- $\theta$ \( \text{a parameter vector} \),
- $\Theta^s$ \( \text{the } s \text{ dimensional parameter space} \),
- $f_s$ \( \text{a } T \times C \times 1 \text{ vector function such that } f_s(\theta) = \hat{p} \),
- $N$ \( \text{the number of observation vectors } X \).
In the chapters that follow Chapter 2, the discussion is focused on paired comparison experiments. A paired comparison experiment refers to an experiment in which a judge compares two items at a time and must indicate a preference for one of the items. Such an experiment can be viewed in terms of a product multinomial distribution in which the comparisons correspond to the groups and the preferred item of a comparison corresponds to the subgroup of a group. Because it will be important to identify the items involved in a comparison, reference will be made to the \((i,j)\)-th comparison rather than the \(t\)-th comparison. Further, the first subscript in the subscript pair will indicate the preferred item so that \((i,j)\) will refer to the subgroup for which the \(i\)-th item is preferred from the group consisting of the possible outcomes \{\((i,j), (j,i)\)\}. The notation used in chapters after Chapter 2 is as follows:

- \(k\) the number of items compared in the experiment,
- \(n_{(i,j)}\) the number of times items \(i\) and \(j\) are compared against one another,
- \(n\) the \(\binom{k}{2}\times 1\) vector \((n_{(1,2)}, n_{(1,3)}, \ldots, n_{(k-1,k)})'\),
- \(n_C\) the \(2\binom{k}{2}\times 1\) vector \((n_{(1,2)}, n_{(2,1)}, n_{(1,3)}, \ldots, n_{(k-1,k)}, n_{(k,k-1)})'\),
- \(X_{(i,j)}\) the number of times item \(i\) is preferred to item \(j\),
- \(X\) the \(\binom{k}{2}\times 1\) vector \((X_{(1,2)}, X_{(1,3)}, \ldots, X_{(k-1,k)})'\),
- \(X_C\) the \(2\binom{k}{2}\times 1\) vector \((X_{(1,2)}, X_{(2,1)}, X_{(1,3)}, \ldots, X_{(k-1,k)}, X_{(k,k-1)})'\).
the probability of item \( i \) being preferred to item \( j \) when items \( i \) and \( j \) are compared,

\[ P(i,j) \]

the \((k)\times 1\) vector \( (P(1,2), P(1,3), \ldots, P(k-1,k))' \),

\[ \pi_i \]

the scale value corresponding to item \( i \),

\[ \pi \]

the vector of distinct scale values,

\[ \Pi^S \]

the \( s \) dimensional parameter space satisfying

\[ \Pi^S = \{ \pi | \pi = (\pi_1, \ldots, \pi_s)', \pi_i > 0, i = 1, \ldots, s \} \],

\[ m_q \]

the number of items that have scale value equal to \( \pi_q \),

\[ m \]

the \( s \times 1 \) vector \((m_1, \ldots, m_s)' \),

\[ f_s \]

the \((2(k)\times 1)\) vector function of the \( s \times 1 \) parameter vector

\[ \pi \]

the \((i,j)\)-th element of which is \( \pi (\pi_q + \pi_r)^{-1} \), where \( \pi_q \) and \( \pi_r \) are the distinct scale values associated with items \( i \) and \( j \), respectively,

\[ N \]

the number of judges.

Summations in Chapters 3 and 4 will be abbreviated as follows:

\[ \sum_{i=1}^{k} x_i = \sum_{i=1}^{k} x_{i}', \sum_{j=1}^{k} x_{ij} = \sum_{j=1}^{k} x_{ij} \text{ and } \sum_{j\neq i}^{k} x_{ij} = \sum_{i=1}^{k} \sum_{j=i+1}^{k-1} x_{ij}. \]

Product notation will be abbreviated similarly.

A specific matrix notation is used throughout this dissertation.

The only difference between its usage in Chapter 2 and the subsequent chapters is that in Chapter 2 the generic symbols \( T \), \( \bar{n}, X \) and \( \Theta \) are used,
which are replaced in the subsequent chapters by the more specific quantities $2\binom{k}{2}$, $n_C$, $X_C$ and $\pi$, respectively. The following is a description of the notation used in Chapter 2 (which can be easily adjusted to describe the notation more specific used in the following chapters):

- $D(a)$ a diagonal matrix with the elements of $a$ down the diagonal,
- $\nu_0$ the TCxl vector equal to $D^{-1}(\tilde{\eta})f_0(\theta_0)$,
- $A_s$ the TCxs matrix equal to $D^{-1}(\nu_0) \frac{\partial f_0(\theta)}{\partial \theta}_{\theta=\theta_0}$,
- $P_s$ the TCxTC matrix equal to $A_s(A_sA_s)^{-1}A_s'$,
- $\Sigma$ the variance-covariance matrix of $X$,
- $A^-$ the generalized inverse of the matrix $A$. 
2. PARAMETER ESTIMATION AND HYPOTHESIS TESTING

2.1 Introduction

There exists many methods with which parameters can be estimated. Some criteria for selecting a parameter estimation method include the convenience of numerical evaluations, the efficiency of the method, and the existence of a suitable distribution theory. This last criterion is important when tests of hypotheses concerning the parameters are needed. One aspect of the popularity of maximum likelihood estimators (MLEs) is that under fairly general conditions they are asymptotically distributed as normal random variables, which provides convenient and efficient large sample tests of hypotheses. However, MLEs may not always be available, as in the case when the underlying distribution is too complex to evaluate the MLEs numerically. In such situations, it is necessary to use other types of estimators with tractable distributional properties although such estimators may not be quite as efficient as MLEs.

In this chapter, attention will be focused on testing hypotheses of the form

$$H: \mathbf{p} = f(\theta), \quad \theta \in \Theta,$$

where \( \mathbf{p} \) is a \( TC \times 1 \) vector of \( T \) sets of probabilities with elements \( p_{i}^{(t)} \) which satisfy \( p_{i}^{(t)} > 0, \sum_{i=1}^{C} p_{i}^{(t)} = 1 \ (t = 1, \ldots, T; \ i = 1, \ldots, C) \), and \( \theta \) is a \( k \times 1 \) vector of parameters lying in the \( k \) dimensional parameter space \( \Theta \), with \( k \leq T(C-1) \). Underlying this testing problem are observations of
a $T 	imes 1$ random vector $\mathbf{X}$ with elements that are observed frequencies for the corresponding events, where an event is classified as belonging to a particular group, subgroup combination ($T$ groups, $C$ subgroups/group).

For example, in a paired comparisons experiment a judge's observation vector records the number of times item $i$ was preferred to item $j$ when items $i$ and $j$ are compared, for all possible pairs of items. In this example the pairs of items correspond to the groups and the two subgroups per group correspond to the items of the pair. Much work has been done in the field of contingency table analysis for the case where $\mathbf{X}$ is assumed to have a multinomial or product multinomial distribution. More recent work in this field has focused on the situation where only $E(\mathbf{X})$ and $\text{Var}(\mathbf{X})$ are known. The approach used to test hypotheses in such situations has been to use a Wald statistic in which $\hat{\theta}$ is estimated by the MLE assuming a multinomial or product multinomial distribution.

Such an estimator of $\hat{\theta}$ is known as a pseudo-MLE rather than the MLE when the estimator is not based on the true distribution of $\mathbf{X}$. However, as long as the pseudo-MLE is a consistent estimator of $\hat{\theta}$ with a limiting normal distribution, it can be used to estimate $\hat{\theta}$ in the test statistic. A drawback in using a pseudo-MLE rather than, for example, the true MLE, is that the pseudo-MLE may be inefficient as compared to estimators based on the true distribution of $\mathbf{X}$ (such as the true MLE). On the other hand, pseudo-MLEs are general enough to permit tests of hypotheses to be constructed in a variety of situations, particularly those where the underlying distribution is unknown or difficult to work with. This
approach of using pseudo-MLEs to perform tests of hypotheses has been used by Brier (1980) and has recently been described in detail by Rao and Scott (1984).

In the next section, a linear approximation for the MLE for $\theta$ based on the product multinomial distribution will be derived. This estimator is used as a pseudo-MLE for $\theta$ for distributions more complex than the product multinomial distribution. In the following section, two test statistics which utilize the pseudo-MLEs are proposed and their distributions derived. These test statistics are used in Chapters 3 and 4 to test hypotheses about the items in paired comparison experiments and to test the appropriateness of the judge effect models proposed in those chapters.

2.2 Maximum Likelihood Estimation for the Product Multinomial Distribution

In this section, a closed form approximation for the MLE for the product multinomial distribution is presented. It is used as the pseudo-MLE in the test statistics described in the next section. The approximation is an extension of Theorem 14.8-3 in Bishop, Fienberg and Holland (1975) and its derivation is similar to that as presented by Cox (1984) for the multinomial distribution.

First, the product multinomial distribution is formally defined.

**Definition 2.1** Let $X^{(t)}$ be a $C \times 1$ vector of random frequencies for which $X_{i}^{(t)} > 0$ and $\sum_{i=1}^{C} X_{i}^{(t)} = n_t$, and let $X' = (X^{(1)}', \ldots, X^{(T)}')$. Define $P^{(t)}$ and $P$ similarly, but with $p_{i}^{(t)} > 0$ and $\sum_{i=1}^{C} p_{i}^{(t)} = 1,$
(t = 1, ..., T; i = 1, ..., C). \( \mathbf{X} \) is said to follow a product multinomial distribution if the distribution function of \( \mathbf{X} \) is given by

\[
h(\mathbf{x}; \mathbf{p}, \mathbf{n}) = \prod_{t=1}^{T} \mathbf{n}_t! \prod_{i=1}^{C} \left( \frac{x_i(t)}{p_i(t)} \right)^{x_i(t)},
\]

where \( \mathbf{n}' = (n_1, \ldots, n_T) \). This distribution will be denoted by \( \text{PM}(\mathbf{p}, \mathbf{n}, T) \).

In order to derive an approximation for the MLE for the product multinomial distribution two results are needed. The first of these is the Implicit Function Theorem, which is stated in the following lemma.

**Lemma 2.1** Let \( F: \mathbb{R}^{a+b} \rightarrow \mathbb{R}^b \) be continuously differentiable in an open set \( U \subset \mathbb{R}^{a+b} \) containing the point \((\mathbf{x}_a^*, \mathbf{x}_b^*)\) such that \( F(\mathbf{x}_a^*, \mathbf{x}_b^*) = 0 \), and let the matrix of first partial derivatives, \( \frac{\partial F}{\partial \mathbf{x}_b} \), evaluated at \((\mathbf{x}_a^*, \mathbf{x}_b^*)\), be nonsingular. Then, there exists a neighborhood \( U_0 \) of \( \mathbf{x}_a^* \) in \( \mathbb{R}^a \) and a unique, continuously differentiable function \( g: U_0 \rightarrow \mathbb{R}^b \) such that \( g(\mathbf{x}_a^*) = \mathbf{x}_b^* \) and \( F(\mathbf{x}_a^*, g(\mathbf{x}_a^*)) = 0 \), for all \( \mathbf{x}_a \in U_0 \).

**Proof:** See Apostol (1957, Theorem 7-6). QED

The second result required is an extension of Birch's closed form approximation for the MLE for the multinomial distribution (see Birch (1964, Lemma 4)).
Lemma 2.2 Let \( \tilde{I}^{TC} \) be the set of TCx1 vectors \( \tilde{p} = (p_1, \ldots, p_T) \) where \( p_t \) is a Cx1 vector satisfying \( p_{i,t} \geq 0 \) and \( C_{i=1}^{T} p_{i,t} = 1 \), \( t = 1, \ldots, T; i = 1, \ldots, C \). Let \( I^{TC} \) be the interior of \( \tilde{I}^{TC} \) and let \( p \in I^{TC} \). Further, let \( p \) satisfy \( p = f(\theta) \) for some \( \theta \in \Theta \), where \( \Theta \) is an open set contained in \( R^k \) \((k \leq T(C-1))\) and define the function \( L^* (\tilde{p}, \theta) : I^{TC} \times \Theta \rightarrow R^k \) to have as its \( u \)-th element

\[
\sum_{t=1}^{T} \sum_{i=1}^{C} w_t \frac{p_{i,t}}{f_{i,t}(\theta)} \frac{\partial f_{i,t}(\theta)}{\partial \theta_u},
\]

where \( f_{i,t}(\theta) = p_{i,t} \) and \( w_t \) is a known weight. If \( f \) has continuous second derivatives in a neighborhood \( \Theta_{0} \subset \Theta \) of \( \theta_{0} \) and the TCxk matrix \( \frac{\partial f}{\partial \theta} \), evaluated at \( \theta_{0} \), has full rank \( k \), then there exists a neighborhood \( I_{0}^{TC} \) of \( \tilde{p}_{0} = f(\theta_{0}) \) and a continuously differentiable function \( \Theta(p) : I_{0}^{TC} \rightarrow R^k \) such that \( L^* (\tilde{p}, \Theta(p)) = 0 \), for all \( \tilde{p} \in I_{0}^{TC} \). Furthermore, the function \( \Theta(p) \) is given by

\[
\Theta(p) = \Theta_{0} + (A' A)^{-1} A' D^{-1} (v_0) (p-p_0) + o(|| p-p_0 ||),
\]

where \( A = D^{-1} (v_0) \frac{\partial f}{\partial \theta} \bigg|_{\theta=\theta_0} \), and \( D(v_0) \) is a diagonal matrix with the elements of \( v_0 \) on the diagonal and \( v_0 \) is a TCx1 vector for which the \(((t-1)C+i)\)-th element is \( p_{0,i}/w_t \).
Proof: Note that because $\sum_{i=1}^{C} f_i(t)(\theta) = 1$ \((t = 1, \ldots, T)\), then
\[
C \sum_{i=1}^{C} \frac{\partial f_i(t)(\theta)}{\partial \theta_u} = 0 \quad (u = 1, \ldots, k) \quad (2.2-3)
\]
and
\[
C \sum_{i=1}^{C} \frac{\partial^2 f_i(t)(\theta)}{\partial \theta_u \partial \theta_r} = 0 \quad (u,r = 1, \ldots, k). \quad (2.2-4)
\]

From (2.2-3) and the fact that $p_0 = f(\theta_0)$, it follows that $L^*(p_0, \theta_0) = 0$. Also note that the \((u,r)\)-th element of $\frac{\partial L^*(p, \theta)}{\partial \theta}$ is given by
\[
\sum_{t=1}^{T} \sum_{i=1}^{C} w_t \frac{C}{f_i(t)(\theta)} \left( \frac{\partial f_i(t)(\theta)}{\partial \theta_u} \right) \frac{\partial f_i(t)(\theta)}{\partial \theta_r} + \sum_{t=1}^{T} \sum_{i=1}^{C} w_t \frac{C}{f_i(t)(\theta)} \frac{\partial^2 f_i(t)(\theta)}{\partial \theta_u \partial \theta_r}.
\]

The last term of $\frac{\partial L^*(p, \theta)}{\partial \theta}$ evaluated at \((p_0, \theta_0)\) is zero by (2.2-4), and using the fact that $p_0 = f(\theta_0)$,
\[
\frac{\partial L^*(p, \theta)}{\partial \theta} \bigg|_{(p, \theta) = (p_0, \theta_0)} = - A A \quad (2.2-5)
\]

Since by assumption $\frac{\partial f(\theta)}{\partial \theta} \bigg|_{\theta = \theta_0}$ has full rank, it follows that $\frac{\partial L^*(p, \theta)}{\partial \theta}$ evaluated at \((p_0, \theta_0)\) has full rank.
Taking $L^*$, $p$, $p_0$, $\theta$ and $\theta_0$ to be $F$, $x_a^*$, $x_b^*$, respectively, in Lemma 2.1, it follows that there exists a neighborhood $I_{TC}$ of $p_0$ and a unique, continuously differentiable function $\theta(*)$ such that 

$$\theta(p_0) = \theta_0 \quad \text{and} \quad L^*(p, \theta(p)) = 0, \quad \text{for all} \quad p \in I_{TC}.$$

Now, use a Taylor series expansion of $\theta(p)$ about $p_0$ to obtain

$$\theta(p) = \theta_0 + \frac{\partial \theta(p)}{\partial p}
\left|_{p=p_0} (p-p_0) + o(||p-p_0||) \right. \quad \text{(2.2-6)}$$

Because $L^*(p, \theta(p)) = 0$, for all $p \in I_{TC}$, it follows that

$$\frac{\partial L^*(p, \theta(p))}{\partial p}
\left|_{p=p_0} \frac{\partial \theta(p)}{\partial \theta} \left|_{p=p_0} \frac{\partial \theta(p)}{\partial p} \right|_{p=p_0} = 0, \quad \text{(2.2-7)}$$

which implies that

$$\frac{\partial \theta(p)}{\partial p}
\left|_{p=p_0} \frac{\partial L^*(p, \theta(p))}{\partial \theta} \left|_{p=p_0} \frac{\partial \theta(p)}{\partial p} \right|_{p=p_0} = - \left( \frac{\partial L^*(p, \theta(p))}{\partial \theta} \left|_{p=p_0} \frac{\partial \theta(p)}{\partial p} \right|_{p=p_0} \right)^{-1} \frac{\partial L^*(p, \theta(p))}{\partial p}
\left|_{p=p_0} \right. \quad \text{(2.2-7)}$$

It is easy to see that the $(u, (t-1)c+1)$-th element of $\partial L^*(p, \theta)/\partial p$ is

$$v_{t}(f_i(t)(\theta))^{-1} \left( \frac{\partial f_i(t)(\theta)}{\partial \theta} / \partial u \right) \quad \text{so that} \quad \frac{\partial L^*(p, \theta(p))}{\partial p}
\left|_{p=p_0} = A \cdot D^{-1}(v_0) \right. \quad \text{(2.2-7)}$$
Substituting this and (2.2-5) into (2.2-7), it follows from (2.2-6) that

\[ \tilde{\theta}(p) = \theta_0 + (A'A)^{-1}A'D^{-1}(\nu_0)(p-P_0) + o(||p-P_0||). \]  

This last lemma is the basis for obtaining a closed form approximation to the solution for the maximum likelihood equations for the product multinomial distribution. The kernel of the log likelihood for PM(p, n, T) is

\[ \sum_{t=1}^{T} \sum_{i=1}^{C} x_i(t) \log p_i(t). \]

If \( p_i(t) \) is a continuously differentiable function \( \hat{f}_i(t) \) of some parameter \( \tilde{\theta} \) for all \( i, t \), then the maximum likelihood equations for \( \tilde{\theta} \) are of the form

\[ 0 = \sum_{t=1}^{T} n_t \sum_{i=1}^{C} \frac{\hat{p}_i(t)}{p_i(t)} \frac{\partial p_i(t)}{\partial \tilde{\theta}_u}, \quad u = 1, 2, \ldots, k \]

where \( \hat{p}_i(t) = x_i(t)/n_t \). This last formula is the function \( L^*(p, \tilde{\theta}) \) in (2.2-1) with \( \hat{p}_i(t) \), \( f_i(t) \) and \( \omega_t \) corresponding to \( \hat{p}_i(t) \), \( p_i(t) \) and \( n_t \), respectively. However, because \( \hat{p}_i(t) = p_i(t) \) in a stochastic sense and not in a strict sense, the function \( \theta(p) \) corresponding to (2.2-2) does not hold in the strict sense but rather stochastically. More precisely, because \( \hat{p} = p_0 + \hat{0}_p(n^{-k}) \), where \( n = \min(n_1, \ldots, n_T) \), when \( X \) is distributed as PM(p_0, n, T), it follows (using a standard argument as presented, for example, in Bishop, Fienberg and Holland (1975, p.511)) that

\[ \theta(\hat{p}) = \theta_0 + (A'A)^{-1}A'D^{-1}(\nu_0)(\hat{p}-P_0) + o_p(n^{-k}). \]  

(2.2-8)
Note that for the special case in which all the elements of \( n \) are equal (i.e., the total number of observations in each group is the same), the above representation for the MLE becomes
\[
\hat{\theta}(\hat{\theta}) = \theta_0 + (A' A)^{-1} A' D^{-1/2}(\hat{p}_0 - \hat{\theta}_0)(\hat{p}_0 - \hat{\theta}_0) + o_p(n^{-1/2}),
\]
where \( A = D^{1/2}(\hat{p}_0) \frac{\partial f}{\partial \theta} \bigg|_{\theta = \theta_0}. \)

In the next section, the MLE for \( \theta \), as given in (2.2-8), is used as the estimator for \( \theta \) in the test statistics presented. Because the underlying distribution is not necessarily the product multinomial distribution, this estimator is referred to as a pseudo-MLE rather than the MLE. This result is summarized in the following definition.

**Definition 2.2** Let \( p \) be a TCx1 probability vector such that \( \bar{p}' = (p_1', \ldots, p_T') \), where \( p_t' = (p_1(t), \ldots, p_C(t)) \) and \( \sum_{i=1}^{C} p_i(t) = 1 \) \((t = 1, \ldots, T)\), and let \( f \) be a continuously differentiable TCx1 vector function of the kx1 parameter \( \theta \) such that \( p = f(\theta) \). Further, let \( \hat{n}' = (\hat{n}_1', \ldots, \hat{n}_T') \), where \( \hat{n}_i \) is a Cx1 vector in which the elements all equal \( \hat{n}_i \), and denote by \( D(\hat{n}) \) a diagonal matrix with the elements of \( \hat{n} \) on the diagonal. Define \( \hat{p}_0 = f(\hat{\theta}_0) \) and \( \hat{X}^{(N)} = N^{-1} \sum_{i=1}^{N} \hat{X}^{(i)} \), where \( \hat{X}^{(1)}, \ldots, \hat{X}^{(N)} \) are TCx1 vectors of observed frequencies. The pseudo-MLE for \( \theta \) based on a PM(p, n, T) distribution is given by
\[
\hat{\theta}(\hat{p}^{(N)}) = \theta_0 + (A' A)^{-1} A' D^{-1/2}(\hat{p}^{(N)} - \hat{\theta}_0)(\hat{p}^{(N)} - \hat{\theta}_0) + o_p(N^{-1/2}),
\]
where \( y_0 = D^{-1}(\theta) \theta_0 \), \( A = D^{-1}(\theta_0) \left[ \frac{\partial f}{\partial \theta} \right]_{\theta=\theta_0} \) (differentiation is performed with respect to the \( k \) parameters of \( \theta \)) and \( \hat{\theta}^{(N)} = D^{-1}(\hat{\theta}) \hat{\theta}^{(N)} \).

### 2.3 Two Types of Wald Statistics

A standard procedure for constructing test statistics is Wald's method. In its simplest form, this method uses an estimator \( \hat{\theta}_N \) of \( \theta_0 \) such that \( \sqrt{N}(\hat{\theta}_N - \theta_0) \overset{\mathcal{D}}{\rightarrow} N(0, I) \), where \( I \) is of full rank, and a consistent estimator \( \hat{\Sigma}_N \) of \( \Sigma \) to construct the test statistic

\[
N(\hat{\theta}_N - \theta_0)^T \hat{\Sigma}_N^{-1} (\hat{\theta}_N - \theta_0)
\]

for the hypothesis \( H_0: \theta = \theta_0 \). This test statistic is asymptotically distributed as a chi-square random variable with degrees of freedom equal to the rank of \( E \). In this section, the two test statistics proposed are variations of the basic Wald statistic. They will be used to test hypotheses of the form

\[
H_0: p = f_s(\theta_s), \quad \theta_s \in \Theta^s
\]

versus

\[
H_A: p = f_r(\theta_r), \quad \theta_r \in \Theta^r,
\]

where \( p \) is a TCx1 probability vector whose elements \( p_i^{(t)} \) satisfy \( p_i^{(t)} \geq 0 \), \( \sum_{t=1}^{C} p_i^{(t)} = 1 \) (\( t = 1, \ldots, T \); \( i = 1, \ldots, C \)), \( f_s \) and \( f_r \) are TCx1 vector functions of the sx1 and rx1 parameter vectors \( \theta_s \) and \( \theta_r \), respectively, and \( \Theta^s \) and \( \Theta^r \) are s and r (s<r) dimensional parameter spaces, respectively, \( (\Theta^s \subset \Theta^r) \). The notation \( \theta_s \) is fairly cumbersome. The dimension of the parameter vector, though, can be
determined from the context in which it is used, and so in the remainder of this chapter, reference will be made to just $\theta$. Because the underlying distribution of the random vector is not specified, but only $E(X)$ and $\text{Var}(X)$ are assumed known, $\theta$ is estimated using the pseudo-MLE assuming a product multinomial distribution, as given by Definition 2.2. This estimator, or any function of this estimator, will now be referred to as a pseudo-MLE.

In order to derive the distribution of the test statistics that will be proposed, two lemmas are needed. The first lemma is concerned with the distribution of $\bar{X}(N) = N^{-1} \sum_{i=1}^{N} X^{(i)}$.

**Lemma 2.3** Let $X^{(1)}, \ldots, X^{(N)}$ be i.i.d. with $E(X^{(i)}) = D(\tilde{n})p + \tilde{\mu}^{(N)}/\sqrt{N}$ and $\text{Var}(X^{(i)}) = \tilde{\Sigma}^{(N)}$, where $D(\tilde{n})$ is a diagonal matrix with the elements of $\tilde{n}$ on the diagonal, $\tilde{n} = (n_1, \ldots, n_T)$, $n_1$ is a Cx1 vector with all elements equal to $n_1$, $\tilde{\mu}^{(N)}$ is a TCx1 vector such that $\tilde{\mu}^{(N)} \to \mu$ and $\tilde{\Sigma}^{(N)}$ is a TCxTC matrix such that $\tilde{\Sigma}^{(N)} \to \Sigma$. Then, as $N \to \infty$,

$$\sqrt{N}(\bar{X}^{(N)} - D(\tilde{n})p) \xrightarrow{d} N(\tilde{\mu}, \tilde{\Sigma}) .$$

**Proof:** This will be shown using characteristic functions. By the fact that the $X^{(i)}$ are i.i.d.,

$$E(e^{it(X^{(N)}-D(\tilde{n})p)/\sqrt{N}}) = [E(e^{it(X^{(1)}-D(\tilde{n})p)/\sqrt{N}})]^N .$$

Using a Taylor series expansion,
\[
\begin{align*}
E(e^{it' (X^{(i)} - D(\tilde{u}) p)/\sqrt{N}}) &= 1 + it' E((X^{(i)} - D(\tilde{u}) p)(X^{(i)} - D(\tilde{u}) p)' )/N + o(t' t/N) \\
&= 1 + it' E((X^{(i)} - D(\tilde{u}) p)(X^{(i)} - D(\tilde{u}) p)' )/N + o(t' t/N) \\
&= 1 + it' \mu^{(N)}/N - it' \xi^{(N)} \mu^{(N)}/N + o(t' t/N) \\
&= 1 + it' \mu^{(N)}/N - it' \xi^{(N)} \mu^{(N)}/N + o(t' t/N).
\end{align*}
\]

Because \(((1 + o(N^{-1})) \rightarrow e^1 \text{ as } N \rightarrow \infty\), where \(\lambda_1 \rightarrow \lambda_1 \) and \(\lambda_2 \rightarrow \lambda_2\) (see Bishop, Fienberg and Holland (1975, §14.2.2)), then

\[
\begin{align*}
E(e^{it' (X^{(i)} - D(\tilde{u}) p)/\sqrt{N}}) \rightarrow e^{it' \mu - it' \xi^t t},
\end{align*}
\]

which is the characteristic function for the normal distribution with mean vector \(\mu\) and covariance matrix \(\xi\). \(\text{QED}\)

From this lemma it easily follows that for \(\hat{\theta}^{(N)} = D^{-1}(\tilde{u}) X^{(N)},\) \(\sqrt{N}(\hat{\theta}^{(N)} - p)\) is asymptotically distributed as \(N(D^{-1}(\tilde{u}) \mu, D^{-1}(\tilde{u}) \xi D^{-1}(\tilde{u})).\)

The next lemma is due to Moore (1977) and provides the asymptotic distribution of a generalized Wald statistic.

**Lemma 2.4** Let \(\tau_N\) be a sequence of estimators of \(\tau_N = \tau_0 + \mu^{(N)}/\sqrt{N}\)

with \(\mu^{(N)} \rightarrow \mu\) and \(\sqrt{N}(\tau_N - \tau_0) \xrightarrow{L} N(\mu, \xi),\) where the rank of \(\xi\) is \(r\).

If \(\hat{\tau}_N\) is a consistent estimator of \(\tau^*,\) a generalized inverse of \(\xi,\) and
if \( y \in C(\$) \), the column space of \( \hat{\$} \), then

\[
N(\zeta_N - \zeta_0, \\zeta_N - \zeta_0) \overset{L}{\to} \chi^2_r(\mu, \mu'),
\]

where \( \chi^2_r(c) \) denotes a chi-square random variable with \( r \) degrees of freedom and noncentrality parameter \( c \).

**Proof:** See Moore (1977, Theorem 2.b.) QED

Before presenting the test statistics, some notation that is used in the theorems and their proofs is defined. Denote by \( s \) the TCxs matrix \( D'(\theta) = 0 \), where \( V^2 = D'(\theta)P \), \( P = f_s(\theta_0), \theta_0 \in \Theta^s \)

and the differentiation is with respect to the \( s \) dimensions of \( \Theta^s \). Let

\[
P_s = A^{\prime}_s (A A)^{-1} A^{\prime}_s
\]

and define

\[
M_{r,s} = D'(\theta_0)(P - P_s)D'(\theta_0)D^{-1}(\theta_0), \quad M_{r,s} = M_{r,s} \mu
\]

and

\[
\hat{\$}_{r,s} = M_{r,s} \hat{\$}_{r,s}.
\]

One approach to the present hypothesis testing problem is to use the generalized Wald statistic

\[
W^2(f_s(\hat{\theta}), f_r(\hat{\theta}), \hat{\$}_{r,s}) = (f_r(\hat{\theta}) - f_s(\hat{\theta}))' \hat{\$}_{r,s} (f_r(\hat{\theta}) - f_s(\hat{\theta})), (2.3-1)
\]

where \( f_s(\hat{\theta}) \) and \( f_r(\hat{\theta}) \) are the pseudo-MLEs for \( \$ \) under \( H_0 \) and \( H_A \).
respectively, based on the pseudo-MLEs \( \hat{\theta} \) and \( \hat{\theta} \), respectively, for \( \theta \), and \( \mathbb{T}_{r,s} \) is a generalized inverse of \( \mathbb{T}_{r,s} \). The distribution of this statistic is chi-square, as shown by the following theorem.

**Theorem 2.1** Let \( X^{(1)}, \ldots, X^{(N)} \) be i.i.d. with \( D(n)p^{(N)} = E(X^{(i)}) = D(n)p_0 + \mu^{(N)}/\sqrt{N} \) and \( \text{Var}(X^{(i)}) = \mathbb{J}^{(N)} \), where \( \mu^{(N)} \rightarrow \mu \) and \( \mathbb{J}^{(N)} \rightarrow \mathbb{J} \) as \( N \rightarrow \infty \). Let \( f^r \) and \( f^s \) be continuously differentiable functions, and \( f^r(\hat{\theta}^{(N)}) \) and \( f^r(\hat{\theta}^{(N)}) \) be the pseudo-MLEs for \( p^{(N)} \) (using the pseudo-MLEs \( \hat{\theta}^{(N)} \) and \( \hat{\theta}^{(N)} \) for \( \theta \)) under \( H_0: p^{(N)} = p_0 = f^s(\theta_0), \theta_0 \in \Theta^s \), and \( H_A: p^{(N)} = f^r(\theta), \theta \in \Theta^r (s < r < T(C-1), \Theta^s \subset \Theta^r) \), respectively. Let \( \tilde{\theta}_0 \in \Theta^r \) be such that \( f^s(\tilde{\theta}_0) = f^r(\tilde{\theta}_0) \). Then

\[
NW^2(f^s(\hat{\theta}^{(N)}), f^r(\hat{\theta}^{(N)}), \mathbb{T}_{r,s}) \xrightarrow{L} \chi^2_t(\mu_{r,s}^-, \mathbb{T}_{r,s}^-),
\]

where \( t = \text{rank}(\mathbb{T}_{r,s}) \).

**Proof:** From Definition 2.2, the pseudo-MLE for \( \theta \) under \( H_0 \) is

\[
\hat{\theta}^{(N)} = \theta(p^{(N)}) = \hat{\theta}_0 + (A_sA_s)^{-1}A_sD^{-1/2}(v_0)(p^{(N)} - p_0) + o_p(n^{-1/2}).
\]

Because \( f^s \) is differentiable, the Taylor series expansion of \( f^s \) about \( \theta_0 \) is given by

\[
f^s(\theta) = f^s(\theta_0) + D^s(-v_0)A_s(\theta - \theta_0) + o(||\theta - \theta_0||).
\]

Therefore,
\[ f_s(\hat{\theta}^{(N)}) = f_s(\theta_0) + D^k(\nu_0)P_sD^{-k}(\nu_0)(\hat{\theta}^{(N)} - \theta_0) + o_p(N^{-1}) \]

(again, using the argument presented in Bishop, Fienberg and Holland (1975) to justify creating a stochastic expression from the analytic Taylor series expansion). Similarly, under \( H_A \)

\[ f_r(\hat{\theta}^{(N)}) = f_r(\tilde{\theta}_0) + D^k(\nu_0)P_rD^{-k}(\nu_0)(\hat{\theta}^{(N)} - \theta_0) + o_p(N^{-1}) \]

Therefore

\[ f_r(\hat{\theta}^{(N)}) - f_s(\hat{\theta}^{(N)}) = D^k(\nu_0)(P_r - P_s)D^{-k}(\nu_0)(\hat{\theta}^{(N)} - \theta_0) + o_p(N^{-1}) \]

Using Lemma 2.3 it follows that

\[ \sqrt{N} (f_r(\hat{\theta}^{(N)}) - f_s(\hat{\theta}^{(N)}) \overset{\text{d}}{\rightarrow} N(\mu_{r,s}, \tau_{r,s}) \]

and noting that \( \mu_{r,s} \in \mathcal{C}(t_{r,s}) \), it follows by Lemma 2.4 that

\[ NW^2(f_r(\hat{\theta}^{(N)}), f_s(\hat{\theta}^{(N)}), \tau_{r,s}, \tau_{r,s}) \overset{\text{d}}{\rightarrow} \chi^2_{t} (\mu_{r,s}, \tau_{r,s}) \quad \text{QED} \]

Theorem 2.1 permits the generalized Wald statistic \( W^2 \) to be used to test hypotheses about the parameter \( \theta \). The assumption underlying such tests is that the parametric function \( f(\theta) \) accurately describes the probability vector \( p \). A goodness of fit test to check the validity of this assumption follows as a corollary to Theorem 2.1.

**Corollary 2.1** Let \( X^{(1)}, \ldots, X^{(N)} \) be i.i.d. such that \( D(\bar{n})p^{(N)} = E(X^{(1)}) = D(\bar{n})p_0 + \mu^{(N)} \sqrt{N} \) and \( \text{Var}(X^{(1)}) = \tau^{(N)} \), where \( \mu^{(N)} \rightarrow \mu \) and \( \tau^{(N)} \rightarrow \tau \) as \( N \rightarrow \infty \). Let \( f_r \) be a continuously differentiable function and let \( f_r(\hat{\theta}^{(N)}) \) be a pseudo-MLE for \( p^{(N)} \) (using the pseudo-MLE \( \hat{\theta}^{(N)} \) for \( \theta \))
under $H_0$: $P^{(N)} = P = f_r(\theta_0)$, $\theta_0 \in \Theta^r$ (r$\leq$T(C-1)). Let the alternative hypothesis be $H_A$: $P^{(N)}$ unrestricted. Then

$$NW^2(f_r(\hat{\theta}^{(N)}), \hat{P}^{(N)}, \hat{\mu}_{GOF}) \xrightarrow{I} \chi^2_{t} (\mu_{GOF}^2 \hat{\sigma}_{GOF}^2)$$

where $\mu_{GOF} = M^\frac{1}{2}$, $\hat{\mu}_{GOF} = M^\frac{1}{2}M$, $M = D^\frac{1}{2}(t_0) (I_{TC} - P_r) D^{-\frac{1}{2}}(t_0) D^{-1}(t_0)$, $I_{TC}$ is a $T \times C$ dimensional identity matrix and $t = \text{rank}(I_{GOF})$.

**Proof:** Under $H_A$, $\hat{P}^{(N)}$ is a consistent estimator of $P^{(N)}$. The proof of the corollary is then identical to the proof of Theorem 2.1 once it is noted that

$$\hat{P}^{(N)} = f_r(\hat{\theta}_0) + D^\frac{1}{2}(t_0) I_{TC} D^{-\frac{1}{2}}(t_0) (\hat{P}^{(N)} - P_0). \quad QED$$

Another approach to hypothesis testing is to use a form of the Pearson chi-square statistic or the natural logarithm of the likelihood ratio statistic. The following definition presents the general form of these two statistics.

**Definition 2.3** Let $a, b$ be $T \times 1$ vectors such that $a = (a^{(1)}, \ldots, a^{(T)})$, where $a^{(t)}$ is a $C \times 1$ vector satisfying $a^{(t)}_i \geq 0$, $\sum_{i=1}^{C} a^{(t)}_i = n_t$ ($t = 1, \ldots, T; i = 1, \ldots, C$), and $b$ is similarly defined. Then,

$$X^2(a, b) = \sum_{t=1}^{T} \sum_{i=1}^{C} \frac{(a^{(t)}_i - b^{(t)}_i)^2}{b^{(t)}_i}$$

and

$$

(2.3-2)$$
\[ G^2(a,b) = 2 \sum_{t=1}^{T} \sum_{i=1}^{C} a_i^{(t)} \log \left( \frac{a_i^{(t)}}{b_i^{(t)}} \right) . \] (2.3-3)

If \( a \) is taken to be a random vector distributed as \( PM(p,n,T) \), then the asymptotic distribution of \( X^2 \) and \( G^2 \) is known to be chi-square with degrees of freedom depending on the number of independent parameters which must be estimated to determine the \( b_i^{(t)} \) values. This is not the case, however, when the underlying distribution for \( a \) is something else. Nevertheless, for the case that the random vectors \( a^{(N)} \) and \( b^{(N)} \) get arbitrarily close together, both \( X^2 \) and \( G^2 \) have the same asymptotic distribution. This result is summarized in the following lemma.

**Lemma 2.5** Let \( \hat{p}_1^{(N)} \) and \( \hat{p}_2^{(N)} \) be estimators of \( p \) such that \( \hat{p}_1^{(N)} = p + O(N^{-1}) \) and \( \hat{p}_2^{(N)} = p + O(N^{-1}) \). Then \( N X^2(\hat{p}_1^{(N)}, \hat{p}_2^{(N)}) \) and \( N G^2(\hat{p}_1^{(N)}, \hat{p}_2^{(N)}) \) have the same asymptotic distribution, which is denoted by \( N X^2(\hat{p}_1^{(N)}, \hat{p}_2^{(N)}) \sim N G^2(\hat{p}_1^{(N)}, \hat{p}_2^{(N)}) \).

These two statistics can be used in the present problem by substituting \( f(\hat{a}) \) and \( f(s) \) for \( a \) and \( b \), respectively, in Definition 2.3. The distribution of these two statistics is a weighted sum of chi-squares as the following theorem shows.

**Theorem 2.2** Let \( X^{(1)}, \ldots, X^{(N)} \) be i.i.d. with \( D(\hat{a}) p^{(N)} = E(X^{(1)}) = D(\hat{a}) p_0 + \mu^{(N)}/\sqrt{N} \) and \( \text{Var}(X^{(1)}) = \gamma^{(N)} \), where \( \mu^{(N)} \rightarrow \mu \) and \( \gamma^{(N)} \rightarrow \gamma \).
as $N \to \infty$. Let $f_s$ and $f_r$ be continuously differentiable functions, and $f_s(\hat{\theta}(N))$ and $f_r(\hat{\theta}(N))$ the pseudo-MLEs for $p(N)$ (using the pseudo-MLEs $\hat{\theta}(N)$ and $\hat{\theta}(N)$ for $\theta$) under $H_0$: $p(N) = p_0 = f_s(\theta_0)$, $\theta_0 \in \Theta$, and $H_{\Lambda}$: $p(N) = f_r(\theta)$, $\theta \in \Theta^r (s < r < T(c-1), \Theta^s \subset \Theta^r)$, respectively. Let $\tilde{\theta}_0 \in \Theta^r$ be such that $f_s(\theta_0) = f_r(\tilde{\theta}_0)$. Then, $N X^2(f_r(\hat{\theta}(N)), f_s(\hat{\theta}(N))) \overset{d}{\to} N G^2(f_r(\tilde{\theta}_0), f_s(\tilde{\theta}_0))$ and

$$N X^2(f_r(\hat{\theta}(N)), f_s(\hat{\theta}(N))) \overset{L}{\to} \sum_{i=1}^{t} \tilde{e}_{ii} \chi^2_i(\hat{\delta}_i^2) + \sum_{i=t+1}^{TC} \delta_i^2,$$

where $\hat{\delta}_i$ is the $i$-th element of $E^{1/2} \Gamma^{-1} D^{-1/2}(\tilde{E}_0)^{1/2} \Gamma^{-1} E^{1/2}$, $\tilde{e}_{ii}$ is $i$-th element of $\tilde{E}$, $\tilde{E} = \begin{pmatrix} E & 0 \\ 0 & I_{TC-t} \end{pmatrix}$, $E$ is a diagonal matrix with diagonal elements equal to the nonzero eigenvalues of $D^{-1/2}(\tilde{E}_0)^{1/2} \Gamma^{-1} E^{1/2}$, $\Gamma$ is the matrix of orthogonal eigenvectors corresponding to the elements of $E$, and $t = \text{rank}(\Gamma_{r,s})$.

**Proof:** Using the same argument as given in the proof of Theorem 2.1, it follows that $\sqrt{N} (f_r(\hat{\theta}(N)) - f_s(\hat{\theta}(N))) \overset{L}{\to} N(\mu_{r,s}, \Sigma_{r,s})$. This implies that $f_r(\hat{\theta}(N)) - f_s(\hat{\theta}(N)) = o_p(N^{-1})$, so that by Lemma 2.5,

$$N X^2(f_r(\hat{\theta}(N)), f_s(\hat{\theta}(N))) \overset{d}{\to} N G^2(f_r(\tilde{\theta}_0), f_s(\tilde{\theta}_0)).$$

To find the limiting distribution of $N X^2(f_r(\hat{\theta}(N)), f_s(\hat{\theta}(N)))$, note that

$$N X^2(f_r(\hat{\theta}(N)), f_s(\hat{\theta}(N))) =

N(f_r(\hat{\theta}(N)) - f_s(\hat{\theta}(N)))' D^{-1}(\tilde{E}_0) (f_r(\hat{\theta}(N)) - f_s(\hat{\theta}(N))) + o_p(1).$$
Because $D^{-1}(P_0)_{1,s} D^{-1}(P_0)$ is a symmetric matrix, an orthogonal matrix $\Gamma$ can be found such that $\Gamma^T D^{-1}(P_0)_{1,s} D^{-1}(P_0) \Gamma = \left( \begin{array}{cc} E & 0 \\ 0 & 0 \end{array} \right)$, where $E$ is a $t$ dimensional diagonal matrix for which the diagonal elements are the nonzero eigenvalues of $D^{-1}(P_0)_{1,s} D^{-1}(P_0)$. Defining $E = \begin{pmatrix} E & 0 \\ 0 & I_{TC-t} \end{pmatrix}$ and $z = \sqrt{N} E^{-1/2} \Gamma^T D^{-1}(P_0) (f\hat(\theta(N)) - f_\theta(\theta(N)))$, the Pearson statistic can be written as follows:

$$N X^2 (f\hat(\theta(N)), f_\theta(\theta(N))) = z^T E z + o_p(1)$$

$$= \sum_{i=1}^{t} \tilde{e}_{ii} \tilde{\sigma}_i^2 + \sum_{i=t+1}^{TC} \tilde{e}_{ii} \tilde{\sigma}_i^2 + o_p(1).$$

Because $z \xrightarrow{L} N(E^{-1/2} \Gamma^T D^{-1}(P_0) \mu_{1,s}, \begin{pmatrix} I_t & 0 \\ 0 & 0 \end{pmatrix})$, it follows that

$$N X^2 (f\hat(\theta(N)), f_\theta(\theta(N))) \xrightarrow{L} \sum_{i=1}^{t} \tilde{e}_{ii} \chi_1^2 ( \tilde{\sigma}_i^2 ) + \sum_{i=t+1}^{TC} \delta_i^2,$$

where $\delta_i$ is the $i$-th element of $E^{-1/2} \Gamma^T D^{-1}(P_0) \mu_{1,s}$. QED

Theorem 2.2 has as a corollary a goodness of fit test for $p = f(\theta)$. The proof of the corollary is analogous to the proof of the corollary to Theorem 2.1 and so the corollary will only be stated.

Corollary 2.2 Let $X^{(1)}, \ldots, X^{(N)}$ be i.i.d. such that $D(\tilde{\mu}(N)) = E(X^{(1)}) = D(\mu(\theta)) + \mu(\theta)/\sqrt{N}$ and $\text{Var}(X^{(1)}) = \mu(\theta)$, where $\mu(\theta) \xrightarrow{P} \mu$ and $\hat{\theta}(N) \xrightarrow{P} \theta$ as $N \rightarrow \infty$. Let $f_{\mu}$ be a continuously differentiable function and let $f_{\hat{\theta}(N)}$ be a pseudo-MLE for $p(N)$ (using the pseudo-MLE $\hat{\theta}(N)$ for $\theta$) under
\[ H_0: \mathbf{p}^{(N)} = \mathbf{p}_0 = f_r(\theta_0), \theta_0 \in \Theta^c \ (r \leq T(C-1)) \]. Let the alternative hypothesis be \( H_A: \mathbf{p}^{(N)} \) unrestricted. Then, \( \text{N}X^2(\hat{p}^{(N)}), f_r(\hat{\theta}^{(N)}) \) d ~ \( \text{NG}^2(\hat{p}^{(N)}), f_r(\hat{\theta}^{(N)}) \) and

\[
\text{N}X^2(\hat{p}^{(N)}), f_r(\hat{\theta}^{(N)}) \xrightarrow{d} \sum_{i=1}^{t} \tilde{e}_{ii}^2 \chi^2_1(\delta_i^2) + \sum_{i=t+1}^{TC} \delta_i^2,
\]

where \( \delta_i \) is the \( i \)-th element of \( \tilde{\mathbf{e}}^{-\frac{1}{2}} \Gamma \tilde{D}^{-\frac{1}{2}}(\mathbf{p}_0)^\dagger \mathbf{GOF}, \tilde{e}_{ii} \) is the \( i \)-th element of \( \tilde{\mathbf{e}}, \tilde{\mathbf{e}} = \begin{bmatrix} \mathbf{E} & 0 \\ 0 & \mathbf{I}_{TC-t} \end{bmatrix} \), \( \mathbf{E} \) is a diagonal matrix with diagonal elements equal to the nonzero eigenvalues of \( \tilde{\mathbf{e}}^{-\frac{1}{2}} \Gamma \tilde{D}^{-\frac{1}{2}}(\mathbf{p}_0)^\dagger \mathbf{GOF} \), \( \Gamma \) is a matrix of orthogonal eigenvectors corresponding to the elements of \( \mathbf{E}, \mathbf{GOF} \) and \( \dagger \mathbf{GOF} \) are as defined in the corollary to Theorem 2.1 and \( t = \text{rank}(\dagger \mathbf{GOF}) \).

Theorem 2.2 and its corollary provide an alternative to Theorem 2.1 and its corollary. In order to be able to use these test statistics to test the null hypothesis it is necessary to be able to evaluate a mixture of central chi-squares. To perform large sample power calculations for near alternatives, the critical values associated with a mixture of non-central chi-squares are required. Johnson and Kotz (1970, Chapter 29) provide approximations that can be used in the calculations of the critical values of both the central and non-central chi-squares mixtures. The test statistics, though, will be used in Chapter 6 as a means of comparing the various judge effect models for paired comparison experiments.
3. JUDGE EFFECT MODELS BASED ON RANDOM CHOICE PROBABILITIES

3.1 Introduction

The Bradley-Terry model is a deterministic model in a sense because it specifies that for the comparison of two items q and r, with associated scale values $\pi_q$ and $\pi_r$ respectively, the choice probability satisfies $p(q,r) = \pi_q (\pi_q + \pi_r)^{-1}$. This choice probability, however, may vary from judge to judge. Bradley and Terry (1952) noted this and proposed using a different set of scale values for each judge. The problem with such an approach is that no overall ranking of the items is determined by the model, but rather, a ranking is obtained for each judge. Another approach, which permits judge variability to be included in the model and yet retains the ability to obtain an overall ranking, would be to allow the choice probability to vary among judges in such a way that $E(p(q,r)) = \pi_q (\pi_q + \pi_r)^{-1}$, that is, the Bradley-Terry condition holds on average.

Another feature of the Bradley-Terry model is that judgements made by the same judge for the same comparison are assumed independent of one another. Although in certain situations this may be a valid assumption to make (e.g., a blind taste test in which the identity of the items are concealed from the judges), in other situations this may not be the case.

In this chapter, two judge effect models will be introduced. First, a model for which $E(p(q,r)) = \pi_q (\pi_q + \pi_r)^{-1}$ holds and all comparisons are assumed independent of one another will be described. Next, a second model in which the assumption of independence is relaxed while the
condition \( E(p(q,r)) = \pi_q (\pi_q + \pi_r)^{-1} \) continues to hold will be presented. For both of these models, estimators for the \( \pi_i \)'s are obtained. Also, test of hypotheses for the scale values and goodness of fit tests for the models are given.

3.2 The Beta-Bernoulli Model

In this section, the Bradley-Terry model will be extended so that the choice probabilities may vary among judges in such a way that \( E(p(q,r)) = \pi_q (\pi_q + \pi_r)^{-1} \). All comparisons made under this model, however, will be assumed to be independent as is the case with the Bradley-Terry model.

Consider an experiment in which each judge performs each possible comparison once. Because of the independence assumption, attention can be restricted to a single comparison. Under the model, the choice random variable \( X(q,r) \) corresponding to the comparison \((q,r)\) can take on values 1 or 0 depending on whether or not item \( q \) is preferred to item \( r \), with probabilities \( p(q,r) \) and \( 1-p(q,r) \), respectively. Further, and this is where the model differs from the Bradley-Terry model, the choice probability is considered to be randomly distributed as a Beta random variable. The parameters of the Beta distribution are taken to be \( c^*\pi_q \) and \( c^*\pi_r \), where \( \pi_q \) and \( \pi_r \) are the scale values associated with items \( q \) and \( r \), respectively, so that \( E(p(q,r)) = \pi_q (\pi_q + \pi_r)^{-1} \). The following definition formalizes this model.
Definition 3.1 A \( \binom{k}{2} \times 1 \) choice vector \( X \) is said to follow the Beta-Bernoulli model if the choice random variables \( X_{(q,r)} \) are independent of each other and the joint distribution for \( X_{(q,r)} \) and the corresponding choice probability \( p_{(q,r)} \) corresponds to a density function of the form

\[
g(x_{(q,r)}, p_{(q,r)}; \pi_q, \pi_r, c^*) = (B(c^* \pi_q, c^* \pi_r))^{-1} x_{(q,r)}^{c^* \pi_q - 1} (1-p_{(q,r)})^{c^* \pi_r - 1} p_{(q,r)}^{c^* \pi_q - 1} (1-p_{(q,r)}) \]

where \( x_{(q,r)} \in \{0,1\} \); \( x_{(q,r)} = 1-x_{(r,q)} \); \( 0 \leq p_{(q,r)} \leq 1 \),

where \( x_{(q,r)} \) is 1 or 0 depending on whether item \( q \) is preferred or not preferred, respectively, to item \( r \), \( \pi_q \) and \( \pi_r \) are the scale values associated with items \( q \) and \( r \), respectively, satisfying \( \pi_q, \pi_r > 0 \) and \( \sum_{i=1}^{k} \pi_i = 1 \), \( c^* \) is a judge variability parameter satisfying \( c^* > 0 \), and \( B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b) \).

That the parameter \( c^* \) can be viewed as a measure of variability among judges can be justified by noting that the variance of the Beta distribution, associated with the choice probability \( p_{(q,r)} \), is

\[
\pi_q \pi_r (\pi_q + \pi_r)^{-2} (c^* \pi_q + c^* \pi_r + 1)^{-1}
\]

As \( c^* \to \infty \), the variability among judges of the choice probability \( p_{(q,r)} \) goes to zero and the Beta-Bernoulli model tends to the Bradley-Terry model. On the other hand, as \( c^* \to 0 \) the variance of the Beta distribution approaches its
supremum, i.e., as \( c^* \to 0 \) judge variability increases.

The marginal distribution of \( X_{(q,r)} \) can be found by integrating out the choice probability \( p_{(q,r)} \) as follows:

\[
m(X_{(q,r)}; \pi_q, \pi_r) = \int_0^1 g(x_{(q,r)}, p_{(q,r)}; \pi_q, \pi_r, c^*) dp_{(q,r)}
\]

\[
= \frac{B(x_{(q,r)} + c^* \pi_q, 1 - x_{(q,r)} + c^* \pi_r) / B(c^* \pi_q, c^* \pi_r)}{\pi_q \pi_r (\pi_q + \pi_r)^{-1}}
\]

\[
= \frac{X(q,r)}{\pi_q \pi_r (\pi_q + \pi_r)^{-1}},
\]  (3.2-1)

where \( x_{(r,q)} = 1 - x_{(q,r)} \). Denote by \( h(x; \pi_1, \ldots, \pi_k) \) the marginal density function of the choice vector \( \bar{X} \). Because the choice random variables are assumed to be independent of one another, the marginal density function is given by

\[
h(x; \pi_1, \ldots, \pi_k) = \prod_{i,j} x_{(i,j)}^{x_{(i,j)}} / \prod_{i,j} (\pi_i + \pi_j),
\]

where \( x_{(i,j)} = \sum_j x_{(i,j)} \). This is exactly the distribution for the choice vector \( \bar{X} \) as specified by the Bradley-Terry model (see Chapter 1, Equation 1.2-1). Therefore, the MLEs and the tests of hypotheses for the \( \pi_i \)'s, as well as the goodness-of-fit test, based on the marginal distribution for \( \bar{X} \) are the same as for the Bradley-Terry model.

The Beta-Bernoulli model is easily extended to include experiments in which some judges do not perform all comparisons. In this case let

\( x_{(q,r)} = x_{(r,q)} = 0 \) (so that \( m(X_{(q,r)}; \pi_q, \pi_r) = 1 \)) if a judge does not perform the \((q,r)\)-th comparison. Then, the choice vector for this judge has a marginal distribution given by
\[ h(x; \pi_1, \ldots, \pi_k) = \prod_{i} \frac{x_i}{\pi_i} / \prod_{i<j}^\pi (\pi_i + \pi_j), \]

where \( \prod_{i<j}^\pi \) denotes the product for all comparisons made by the judge. 

For a sample of \( N \) judges in which the \((q,r)\)-th comparison is performed \( N(q,r) \) times \((N(q,r) \leq N)\), the marginal likelihood function is

\[ N \prod_{s=1}^{N} h(x^{(s)}; \pi_1, \ldots, \pi_k) = \prod_{i} a_i / \prod_{i<j} (\pi_i + \pi_j)^{(i,j)}, \]

where \( a_i = \sum_{s=1}^{N} x_i^{(s)} \) and \( x_i^{(s)} \) denotes the value of \( x_i \), defined above, for judge \( s \). This is just the kernel of the likelihood function for the Bradley-Terry model extended to the case in which each comparison is not necessarily performed an equal number of times (see Chapter 1, Equation 1.2-3).

The relationship between the Beta-Bernoulli model and the Bradley-Terry model is interesting. The implication is that the Beta-Bernoulli model cannot be viewed as an effective judge effect model. In fact, the Beta-Bernoulli model is not a true judge effect model because it does not actually incorporate a judge variability term. The parameter \( c^* \), the judge variability parameter, drops out of the marginal distribution (see 3.2-1) and is not estimable from the marginal likelihood. The Beta-Bernoulli model therefore differs from the Bradley-Terry model only in that it eases the Bradley-Terry condition for individual judges.

This easing of the Bradley-Terry condition has no real effect on the Bradley-Terry model, as has been seen in this section. Because in this presentation the aim is to obtain a model that more accurately
describes paired comparison experiments than the Bradley-Terry model, attention will next be focused on relaxing the independence assumption.

3.3 The Beta-Binomial Model

In the previous section, a model was presented for which the Bradley-Terry condition was satisfied on average and all comparisons were assumed independent of one another. Under this assumption, it is not possible to estimate the between judges variability. In this section, an extension of the Beta-Bernoulli model will be presented in which the independence assumption is relaxed in such a way that this judge variability can be estimated.

In paired comparison experiments, it is possible that a judge will perform the same comparison more than once. One way to handle such experiments would be to assume that the multiple judgements of the same pair of items by the same judge are independent of one another. Therefore, a judge who performs a particular comparison n times can be viewed as generating n independent choice vectors, and the Beta-Bernoulli model of the previous section can be applied. Under this model, each time a comparison is made the corresponding choice probability is determined.

Therefore, for n repetitions of the same comparison by a particular judge, the corresponding choice probability is determined n times. An alternative model, which may be more realistic, would require the choice probability to remain fixed for all judgements of the same comparison by a particular
judge. Under this alternative model, judgements of identical comparisons are no longer unconditionally independent of one another, but, conditional on the judge, the comparisons are independent.

This alternative model can be viewed as an extension to the Beta-Bernoulli model in which a binomial distribution is used to model the conditional distribution of a choice random variable for each judge rather than a Bernoulli distribution. The choice random variables corresponding to different comparisons are assumed to be independent of one another, and a Beta distribution, with the appropriate scale values, \( c_q^* \) and \( c_r^* \), as parameters, is used to model the variation in the choice probabilities among the judges. This is formally presented in the following definition.

**Definition 3.2** A \( (\binom{k}{2}) \times 1 \) choice vector \( \mathbf{X} \) is said to follow the Beta-Binomial model if the choice random variables \( X_{q,r} \) are independent of each other and the joint distribution for \( X_{q,r} \) and the corresponding choice probability \( p_{q,r} \) corresponds to a density function of the form

\[
g(x(q,r); p(q,r); n(q,r), c_q^*, c_r^*, c_\star) = \\
(B(c_q^*, c_r^*, c_\star))^{-1} \frac{n(q,r)}{x(q,r)} \cdot p(q,r)^{x(q,r)} (1-p(q,r))^{n(q,r)-x(q,r)}
\]

\( x(q,r) \in \{0, 1, \ldots, n(q,r)\}; x(r,q) = n(q,r) - x(q,r); 0 \leq p(q,r) \leq 1 \),

where \( x(q,r) \) is the number of times item q is preferred to item r,
\( n(q,r) \) is the number of times items q and r are compared to one another, 
\( \pi_q \) and \( \pi_r \) are the scale values associated with items q and r, respectively, satisfying \( \pi_q, \pi_r > 0 \) and \( \sum_{i=1}^{k} \pi_i = 1 \), \( c^* \) is a judge variability parameter satisfying \( c^* > 0 \), and \( B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b) \).

This model will be denoted by \( BB(n, \pi_1, \ldots, \pi_k, c^*) \). The parameter \( c^* \) measures among judge variability as can be seen by the argument given in the previous section. The marginal distribution of the choice random variable \( X_{(q,r)} \) can be determined by integrating out the choice probability \( p(q,r) \). The resulting density function is

\[
m(x_{(q,r)}; n(q,r), \pi_q, \pi_r, c^*) = \\
\frac{\Gamma(x_{(q,r)}+c^* \pi_q) \Gamma(x_{(r,q)}+c^* \pi_r) \Gamma(x_{(q,r)}+c^* \pi_r)}{\Gamma(n_{(q,r)}+c^* \pi_q + c^* \pi_r) \Gamma(c^* \pi_q) \Gamma(c^* \pi_r)} .
\]

Note that for \( n_{(q,r)} = 1 \), this reduces to the Beta-Bernoulli marginal as given in Equation (3.2-1). The Beta-Binomial, like the Beta-Bernoulli, can easily be extended to include situations where not every comparison is performed by each judge. For such situations let \( x_{(q,r)} = x_{(r,q)} = 0 \) when \( n_{(q,r)} = 0 \). The resulting contributions to the joint density function and the marginal density function are identically equal to 1.

The marginal density of the choice vector, denoted by \( h(x; n, \pi_1, \ldots, \pi_k, c^*) \), can be obtained due to the independence of the choice random variables for a given judge. Its value is

\[
h(x; n, \pi_1, \ldots, \pi_k, c^*) = \prod_{q<r} m(x_{(q,r)}; n(q,r), \pi_q, \pi_r, c^*) . \tag{3.3-1}
\]
The mean vector and variance-covariance matrix of \( X \) can easily be calculated from the definition of the Beta-Binomial model using conditional expectations. Using this technique, the mean turns out to be

\[
E(X) = E(E(X|p)) \\
= E(D_n p) \\
= D(D_n) \pi^*_n, \tag{3.3-2}
\]

where \( D_n \) is a diagonal matrix with the elements of \( n \) on the diagonal and \( \pi^*_n \) is a \((\binom{k}{2})\times 1\) vector in which the \((i,j)\)-th element is \( \pi_i (\pi_i + \pi_j)^{-1} \).

Because the comparisons are independent, conditional on the judge, \( E(X'X|p) \) is a diagonal matrix. Therefore, it follows that

\[
E(X_{(q,r)} X_{(s,t)}) = E(X_{(q,r)} E(X_{(s,t)})) \text{ and}
\]

\[
E(X^2_{(q,r)}) = E(E(X^2_{(q,r)}|p))
\]

\[
= E(n_{(q,r)} p_{(q,r)} (1-p_{(q,r)}) + n^2_{(q,r)} p^2_{(q,r)})
\]

\[
= \frac{n_{(q,r)} \pi^*_q}{\pi q} \left\{ 1 + \frac{(n_{(q,r)} - 1)(c \pi_q + 1)}{(c \pi_q + c \pi r + 1)} \right\}. \tag{3.3-3}
\]

From (3.3-2) and (3.3-3) it follows that

\[
\text{Var}(X_{(q,r)}) = \frac{n_{(q,r)} \pi^*_q}{\pi q} \left\{ 1 + \frac{(n_{(q,r)} - 1)(c \pi_q + 1)}{(c \pi_q + c \pi r + 1)} - \frac{n_{(q,r)} \pi^*_q}{\pi q} \right\}
\]

and

\[
\text{Cov}(X_{(q,r)}, X_{(s,t)}) = 0 \text{ (for all pairs } (q,r) \neq (s,t)) . \tag{3.3-4}
\]
The variance formulas given in 3.3-4 provide an alternative way to interpret the Beta-Binomial model. From Equation 3.3-4, as $c^* \rightarrow \infty$, \[ \text{Var}(X_{(q,r)}) = n_{(q,r)} \pi_q \pi_r (\pi_q + \pi_r)^{-2}, \] and as $c^* \rightarrow 0$, \[ \text{Var}(X_{(q,r)}) = n_{(q,r)}^2 \pi_q \pi_r (\pi_q + \pi_r)^{-2}. \] Denoting by $X(i)_{(q,r)}$ the result for the $i$-th comparison of items $q$ and $r$ ($i = 1, 2, \ldots, n_{(q,r)}$), then, \[ \text{Var}(X_{(q,r)}) = \text{Var}(\sum_{i=1}^{n_{(q,r)}} X(i)_{(q,r)}) \]
\[ = n_{(q,r)} \sum_{i=1}^{n_{(q,r)}} \text{Var}(X(i)_{(q,r)}) + 2 \sum_{i<j} \text{Cov}(X(i)_{(q,r)}, X(j)_{(q,r)}) \]
\[ = \frac{n_{(q,r)} \pi_q \pi_r}{(\pi_q + \pi_r)^2} + \frac{n_{(q,r)} (n(q,r)-1) \rho \pi_q \pi_r}{(\pi_q + \pi_r)^2} \]
\[ = \frac{n_{(q,r)} \pi_q \pi_r}{(\pi_q + \pi_r)^2} \left[ 1 + (n(q,r)-1)\rho \right], \]
where $\rho$ is the unconditional correlation between $X(i)_{(q,r)}$ and $X(j)_{(q,r)}$, $i \neq j$. The case $\rho=0$ corresponds to the case that $c^* \rightarrow \infty$ so that for the Bradley-Terry model (which becomes equivalent to the Beta-Binomial model as $c^* \rightarrow \infty$) there is no correlation between comparisons made on the same pair of items. In this case all judges have the same choice probabilities. This result can also be derived by recalling that for the Bradley-Terry model all comparisons are assumed independent of one another. More interesting is the case where $c^* \rightarrow 0$, which corresponds to $\rho=1$. As judge variability increases, the comparisons performed by a judge become perfectly correlated, which is a seeming contradiction.
However, the judge variability for the Beta-Binomial model is a measure of the variability of the choice probabilities distributed as Beta random variables. As the variability among the choice probabilities increases the choice probabilities tend to have more extreme values (i.e., the density function for the Beta distribution becomes more U shaped). The variation among judges is maximized when the choice probabilities are at the extremes, with \( \pi_q \) ones and \( \pi_r \) zeros. Judges with a choice probability value of 1 for the \((q,r)\) pair will always prefer item \( q \) to item \( r \). Judges with choice probability value 0 for the \((q,r)\) pair will always prefer item \( r \) to item \( q \). Consequently, the correlation between the results of two comparisons of items \( q \) and \( r \) performed by the same judge are perfectly correlated, i.e., \( \rho=1 \) for the Beta-Binomial model as \( c^* \to 0 \). In view of the correspondence between the Beta-Binomial with \( c^* \to 0 \) and \( c^* \to \infty \), with models with \( \rho=1 \) and \( \rho=0 \), respectively, the Beta-Binomial model can be viewed as an intra-class correlation model where a class consists of those comparisons of a pair of items performed by a judge. Since the same \( c^* \) parameter is used for the Beta distribution of preferences probabilities for each pair of items, this phenomenon occurs simultaneously for all pairs of items as \( c^* \to 0 \).

Since the \( \pi_i \) values need not be identical, the Beta distribution for the preference probabilities will be centered at different values for different pairs of items. Consequently, the value of \( \rho \) can vary with the pair of items when \( c^* > 0 \), but \( \rho \to 1 \) for every pair of items as \( c^* \to 0 \). Similarly, \( \rho \to 0 \) for every pair of items as \( c^* \to \infty \).
In the next section, various methods of estimation for the parameters $c^*$ and $\pi_1$ are discussed. In the subsequent section, various methods for testing hypotheses concerning the parameters and for testing the goodness of fit of the Beta-Binomial model will be presented. First, some notation will be defined that will be used throughout the next two sections and the correspondence with the notation in Chapter 2 will be indicated.

For the problem of a paired comparison experiment with $k$ items, there are $T = \binom{k}{2}$ groups and $C = 2$ subgroups/group. In the calculations that follow, the complete $2(\frac{k}{2})x1$ choice vectors are required, i.e., each choice vector must contain the elements $X_{(q,r)}$ and $X_{(r,q)}$ (note that $X_{(r,q)} = n_{(q,r)} - X_{(q,r)}$). The complete choice vector will be denoted by $X_C$ to distinguish it from the choice vector $X$ used throughout this chapter. Similarly, $n_C$ will be used to denote a $2(\frac{k}{2})x1$ vector that contains both $n_{(q,r)}$ and $n_{(r,q)}$ (note that $n_{(q,r)} = n_{(r,q)}$). This vector $n_C$ corresponds to the vector $\tilde{n}$ of Chapter 2. The parameter $\theta$ of the previous chapter will correspond to the parameter $\bar{\pi}$, where the elements of $\bar{\pi}$ are the distinct scale values of the Beta-Binomial model, and $\Theta^S$ will now be denoted by $\Pi^S = \{\pi|\pi = (\pi_1, ..., \pi_s), \pi^T = 1, \pi_i > 0, i = 1, ..., s\}$, where $\pi$ is a $s x 1$ vector with the $i$-th element equal to the number of items associated with $\pi_i$. In general, no subscript will be associated with the parameter $\bar{\pi}$ to indicate its dimension as this will be clear from the context the parameter is used in. The function $f_{-s}(\pi)$ will denote a $2(\frac{k}{2})x1$ vector such that $\pi \in \Pi^S$ and whose elements are of the form $\pi_q (\pi_q + \pi_r)^{-1}$. In particular, if $\pi \in \Pi^S$ and the items $i$ and $j$ are
associated with the scale values \( \pi_q \) and \( \pi_r \), respectively, then the 
(i,j)-th element of \( f_S(\pi) \) is \( \pi_q (\pi_q + \pi_r)^{-1} \). For the case that \( \pi \in \Pi^k \), 
the (i,j)-th element of \( f_k(\pi) \) is \( \pi_i (\pi_i + \pi_j)^{-1} \). Note that the function 
\( f_S \) is continuously differentiable on the parameter space \( \Pi^S \) (1 \( \leq \) s \( \leq \) k).

3.4 Parameter Estimation for the Beta-Binomial Model

In this section, three different methods of parameter estimation 
are used to obtain estimates of the scale values and the judge variability 
parameter. The first estimation technique is maximum likelihood estimation. The MLEs for \( c^* \) and the vector of scale values \( \pi \) are obtained by 
maximizing the likelihood function

\[
\prod_{i=1}^{N} h(x^{(i)}; \pi_1, \pi_2, ..., \pi_k, c^*)
\]

where \( h(x; \pi_1, \pi_2, ..., \pi_k, c^*) \) is given in (3.3-1), \( x^{(i)} \) is the choice vector 
associated with judge \( i \), and \( N \) is the number of judges, with respect to 
\( c^* \) and the elements of \( \pi \in \Pi^S \) (1 \( \leq \) s \( \leq \) k). A closed form solution for the 
MLEs does not exist, but it is possible to obtain the MLEs using 
numerical maximization techniques. To this end there are several com­
puter packages available (see Chapter 5 for more details).

A second method of parameter estimation is the method of moments. 
The scale values \( \pi_i \) can be estimated, using this technique, from the 
first moments of \( X_{c_i} \). From (3.3-2), \( E(X_{(i,j)}) = \pi_{(i,j)} (\pi_i + \pi_j)^{-1} \), so 
that the method of moments equations are
\[ \bar{x}_{(q,r)} = N^{-1} \sum_{i=1}^{N} x_{(i)}(q,r) = \frac{n_{(q,r)}}{\pi^q + \pi^r}, \quad (q,r = 1, \ldots, k), \] (3.4-1)

where \( x_{(q,r)} \) is the number of times judge \( i \) prefers item \( q \) to item \( r \).

Consider the problem of estimating \( \pi \in \pi^s \) (1 ≤ s ≤ k). Summing over the \( m_q(k-1) \) equations corresponding to the \( m_q(k-1) \) different comparisons involving \( \pi_q \), where \( m_q \) is the number of items that have \( \pi_q \) as their scale value, the above equations reduce to

\[ \sum_{(i,j)} \bar{x}_{(i,j)} = \pi_q \sum_{(i,j)} \frac{n_{(i,j)}}{\pi^q + \pi_j}, \quad (q = 1, \ldots, s), \]

where \( \sum_{(i,j)} \) denotes the summation over the \( m_q(k-1) \) comparisons \( (i,j) \) involving an item \( i \) which has scale value \( \pi_q \). Note that for \( m_q = 1 \) (q = 1, ..., k) these are the MLE equations for the Bradley-Terry model (see Chapter 1, Equation (1.2-3)). These equations can be solved iteratively with methods available for the Bradley-Terry model. However, if \( \bar{x}_{(i,j)} > 0 \) for all comparisons \( (i,j) \) associated with \( \pi_q \) then a closed form solution exists for \( \pi_q \). Rewrite Equation (3.4-1) as

\[ \frac{n_{(q,r)}}{\bar{x}_{(q,r)}} = \frac{\pi^q + \pi^r}{\pi^q}, \quad (q,r = 1, \ldots, k). \]

Summing over the \( m_q(k-1) \) equations associated with \( \pi_q \) and using the fact that \( \sum_{i} \pi_i = 1 \), where \( \pi = (\pi_1, \ldots, \pi_s) \), or equivalently, \( \sum_{i} \pi_i = 1 \),

\[ \sum_{(i,j)} \frac{n_{(i,j)}}{\bar{x}_{(i,j)}} = \sum_{(i,j)} \frac{\pi^q + \pi_j}{\pi^q} = \frac{m_q}{\pi^q} ((k-2)\pi^q + 1). \]
Therefore, $\pi_q$ can be estimated by

$$\hat{\pi}_q = \left\{ \frac{-1}{m} \sum_{q} \frac{n(i,j)}{x(i,j)} + 2^{-k} \right\}^{-1}. \quad (3.4-2)$$

The problem of estimating $c^*$ by the method of moments involves the second order moment of $X^c$. From (3.3-3)

$$E(X^2(q,r)) = \frac{n(q,r)\pi_q}{\pi_q + \pi_r} \left[ 1 + \frac{(n(q,r)-1)(c^*\pi_q + 1)}{c^*\pi_q + c^*\pi_r + 1} \right]. \quad (3.4-3)$$

Note that $c^*$ can only be estimated from $E(X^2(q,r))$ if $n(q,r) > 1$. If $n(q,r) = 1$ for all comparisons, then $c^*$ is not estimable and the Beta-Binomial reduces to the Beta-Bernoulli model. The method of moments equation associated with the $(q,r)$-th comparison is

$$\frac{-2}{x(q,r)} = N^{-1} \sum_{i=1}^{N} (x(q,i))^2 = \frac{n(q,r)\pi_q}{\pi_q + \pi_r} \left[ 1 + \frac{(n(q,r)-1)(c^*\pi_q + 1)}{c^*\pi_q + c^*\pi_r + 1} \right].$$

Substituting estimates $\hat{\pi}_q$ and $\hat{\pi}_r$ for $\pi_q$ and $\pi_r$ and letting

$$\hat{x}(q,r) = n(q,r)\hat{\pi}_q\hat{\pi}_r^{-1},$$

the above equation can be rewritten as

$$\frac{(\hat{x}(q,r)^{-1})}{(n(q,r)-1)} = \frac{(c^*\pi_q + 1)}{(c^*\pi_q + c^*\pi_r + 1)}.$$
Because up to \( \binom{k}{2} \) such estimators can be calculated to estimate \( c^* \), the average of these estimators will be used to estimate \( c^* \), that is,

\[
\hat{c}^* = d^{-1} \sum_{(i,j)} \hat{c}^*(i,j),
\]

(3.4-4)

where \( \sum_{(i,j)}^* \) is the summation over all comparisons for which \( n_{(i,j)} > 1 \) and \( d \) is the number of such comparisons. Alternatively, the \( c^*(i,j) \) could be weighted inversely proportional to \( n^{-1}_{(i,j)} \) resulting in the estimator

\[
\hat{c}^* = d_w^{-1} \sum_{(i,j)} n_{(i,j)}\hat{c}^*(i,j),
\]

(3.4-5)

where \( d_w = \sum_{(i,j)} n_{(i,j)} \).

The final estimation procedure considered is the pseudo-MLE approach of Chapter 2. This approach can be used to obtain estimators for the \( \pi_i \)'s, but not an estimator for \( c^* \) (the reason is that the pseudo-MLE approach is based on the Bradley-Terry condition, i.e., \( p(q,r) = \pi_q (\pi_q + \pi_r)^{-1} \), which does not involve the parameter \( c^* \)). The pseudo-MLE's for the \( \pi_i \)'s can be substituted into (3.4-4) or (3.4-5) to obtain an estimator for \( c^* \).

Consider the problem of estimating \( \pi \in \Pi^S \) \((1 \leq s \leq k)\). Because \( f_s \) is continuously differentiable everywhere on the parameter space, any value \( \pi_0 \in \Pi^S \) may be used as a basis for the pseudo-MLE. A convenient choice would be to take \( \pi_0 = s^{-1} l \), where \( l \) is the sxl vector of ones. The pseudo-MLE of \( \pi \) is then given by

\[
\hat{\pi}(N) = \pi_0 + (A A^{-1} A^{-1} A^{-1} (\Sigma_0^{-1} + \Sigma_0^{-1} \beta(N) - \Sigma_0^{-1} \beta_0),
\]
where $v_0 = D^{-1}(n_C)P_0$, $A = D^{-1}(v_0) \left. \frac{\partial f_s(\pi)}{\partial \pi} \right|_{\pi = \pi_0}$, $P(N) = N^{-1} \sum_{i=1}^{N} D^{-1}(n_C)x(i)$, $D(a)$ denotes a diagonal matrix with the elements of $a$ down the diagonal and $P_0 = f_s(\pi_0)$. From Lemma 2.3, it follows that $P(N)$ is a consistent estimator of $D^{-1}(n_C)E(\pi)$, and if $D^{-1}(n_C)E(\pi) = f_s(\pi)$, then it follows that $\hat{\pi}(N)$ is a consistent estimator of $\pi$. This estimator can be improved upon iteratively by using $\hat{\pi}(N)$ for $\pi_0$ (after $\hat{\pi}(N)$ has been rescaled so that it satisfies $\sum \pi = 1$) and then recalculating $\hat{\pi}(N)$.

3.5 Hypothesis Testing for the Beta-Binomial Model

Once estimators for the scale values and the judge variability parameter have been calculated, tests of hypotheses concerning the scale values and the judge variability parameter can be performed. In this section such tests, as well as a goodness-of-fit test for the Beta-Binomial model, are presented.

One approach to hypothesis testing is based on MLEs and the likelihood ratio statistic. Define

$$
\Lambda = \frac{L(x^{(1)}, \ldots, x^{(N)}; \hat{\pi}_1, \ldots, \hat{\pi}_k, c^*)}{L(x^{(1)}, \ldots, x^{(N)}; \hat{\pi}_1, \ldots, \hat{\pi}_k, c^*)},
$$

where

$$
L(x^{(1)}, \ldots, x^{(N)}; \hat{\pi}_1, \ldots, \hat{\pi}_k, c^*) = \prod_{i=1}^{N} h(x^{(i)}; \hat{\pi}_1, \ldots, \hat{\pi}_k, c^*),
$$

$h(x^{(i)}; \hat{\pi}_1, \ldots, \hat{\pi}_k, c^*)$ is given in (3.3-1), $x^{(i)}$ is the choice vector associated with judge $i$, $(\hat{\pi}_1, c^*)$ and $(\hat{\pi}_1, c^*)$ are the MLEs for $(\pi_1, c^*)$.
under $H_0$ and $H_A$, respectively, and $N$ is the number of judges. Tests of hypotheses about the $\pi_i$'s and $c^*$ can be performed using the likelihood ratio statistic $\Lambda$. The statistic $-2\ln\Lambda$ has an asymptotic chi-square distribution under the null hypothesis, which can be shown by appealing to the following result found in Wilks (1962, p. 419).

**Lemma 3.1** Let $\Omega_k$ be a $k$ dimensional parameter space and let $\Omega_{k0}$ be an open $k$ dimensional interval in $\Omega_k$ that contains the true parameter $\theta_0 = (\theta_{10}, \ldots, \theta_{k0})'$. Let $w_k$ be a $k'$ dimensional subspace of $\Omega_{k0}$ ($k' < k$) and postulate $H_0: \theta \in w_k$, and $H_A: \theta \in \Omega_{k0} - w_k$. Let $(x_1, \ldots, x_N)$ be a sample from $F(x;\theta)$, $\theta \in \Omega_{k0}$ and let $F(x;\theta)$ be such that for all $\theta \in \Omega_{k0}$ and for all $i,j$

1. $\int \frac{\partial^2 \ln F(x;\theta)}{\partial \theta_i \partial \theta_j} \frac{\partial \ln F(x;\theta)}{\partial \theta_j} dF(x;\theta) + \int \frac{\partial^2 \ln F(x;\theta)}{\partial \theta_i \partial \theta_j} dF(x;\theta)$

and

2. $\int \frac{\partial^2 \ln F(x;\theta)}{\partial \theta_i \partial \theta_j} dF(x;\theta) < \infty$.

Then, $-2\ln\Lambda$, where $\Lambda$ is the likelihood ratio test for $H_0$ and $H_A$, is asymptotically distributed as a chi-square random variable, with $k-k'$ degrees of freedom.

For the Beta-Binomial model, $dF(x;\theta) = h(x;n,\pi_1,\ldots,\pi_k,c^*)$ in equation (3.3-1) and $\Omega_k$ is the $(k+1)$ dimensional space defined by $[0,1]x\ldots x[0,1]x\mathbb{R}$. Let $H(x;n,\pi_1,\ldots,\pi_k,c^*) = \ln h(x;n,\pi_1,\ldots,\pi_k,c^*)$. Then,
\[ H(x;n,\pi, c^*) = \sum_{i<j} \ln \left( \frac{n(i,j)}{x(i,j)} \right) + \ln \Gamma \left( x(i,j) + c^* \pi_i \right) + \ln \Gamma \left( x(j,i) + c^* \pi_j \right) \]

\[ + \ln \Gamma \left( c^* \pi_i + c^* \pi_j \right) - \ln \Gamma \left( n(i,j) + c^* \pi_i + c^* \pi_j \right) \]

\[ - \ln \Gamma \left( c^* \pi_i \right) - \ln \Gamma \left( c^* \pi_j \right) \}

\[ = \sum_{i<j} \ln \left( \frac{n(i,j)}{x(i,j)} \right) + \sum_{i,j} \left\{ \ln \Gamma \left( x(i,j) + c^* \pi_i \right) - \ln \Gamma \left( c^* \pi_i \right) \right\} \]

\[ + \sum_{i<j} \left\{ \ln \Gamma \left( c^* \pi_i + c^* \pi_j \right) - \ln \Gamma \left( n(i,j) + c^* \pi_i + c^* \pi_j \right) \right\} . \]

Noting that \( \frac{\partial \ln \Gamma(x)}{\partial x} = -\gamma - x^{-1} + \sum_{s=1}^{\infty} (s^{-1} - (x+s)^{-1}) \), where \( \gamma \) is Euler's constant (see Artin (1964, p.17)), it follows that

\[ \frac{\partial H(x;n,\pi, c^*)}{\partial \pi_i} = \sum_{j} c^* \left\{ [-\gamma - \frac{1}{x(i,j) + c^* \pi_i}] + \sum_{s=1}^{\infty} \left( \frac{1}{s} - \frac{1}{x(i,j) + c^* \pi_i + s} \right) \right\} \]

\[ - \left[ -\gamma - \frac{1}{c^* \pi_i} + \sum_{s=1}^{\infty} \left( \frac{1}{s} - \frac{1}{c^* \pi_i + s} \right) \right] \}

\[ + \sum_{j} c^* \left\{ [-\gamma - \frac{1}{c^* \pi_i + c^* \pi_j}] + \sum_{s=1}^{\infty} \left( \frac{1}{s} - \frac{1}{c^* \pi_i + c^* \pi_j + s} \right) \right\} \]

\[ - \left[ -\gamma - \frac{1}{n(i,j) + c^* \pi_i + c^* \pi_j} + \sum_{s=1}^{\infty} \left( \frac{1}{s} - \frac{1}{n(i,j) + c^* \pi_i + c^* \pi_j + s} \right) \right] \} \]
\[
= \sum_j c_{i,j} \left( \frac{1}{c^{\pi_{i,j}}} - \frac{1}{c^{\pi_{i,j} + x_{i,j}}} + \frac{1}{n(i,j) + c^{\pi_{i,j} + x_{i,j}}} \right) \\
- \frac{1}{c^{\pi_{i,j} + x_{i,j}}}
\]

\[
+ \sum_j \sum_{s=1}^{\infty} \frac{1}{c^{\pi_{i,j} + x_{i,j} + s}} - \frac{1}{c^{\pi_{i,j} + x_{i,j} + s}} + \frac{1}{n(i,j) + c^{\pi_{i,j} + x_{i,j} + s}} \\
- \frac{1}{c^{\pi_{i,j} + x_{i,j} + s}}
\]

\[
= \sum_j c_{i,j} \left( \frac{1}{c^{\pi_{i,j}}} - \frac{1}{c^{\pi_{i,j} + x_{i,j}}} + \frac{1}{n(i,j) + c^{\pi_{i,j} + x_{i,j}}} \right)
\]

Similarly it follows that

\[
\frac{\partial H(x; n, \pi, c^*)}{\partial c} = \sum_i \sum_j \frac{1}{c^{\pi_{i,j}}} - \frac{1}{c^{\pi_{i,j} + x_{i,j}}} + \frac{1}{n(i,j) + c^{\pi_{i,j} + x_{i,j} + s}} \\
- \frac{1}{c^{\pi_{i,j} + x_{i,j} + s}}
\]

From \(\frac{\partial^2 H(x; n, \pi, c^*)}{\partial c^2}\) and \(\frac{\partial^2 H(x; n, \pi, c^*)}{\partial c^2}\) it is easily seen that

\[
\frac{\partial^2 H(x; n, \pi, c^*)}{\partial c^2} = \sum_j (c^*)^2 \left( \frac{1}{c^{\pi_{i,j} + x_{i,j}}} - \frac{1}{c^{\pi_{i,j} + x_{i,j} + s}} \right)
\]

\[
\frac{\partial^2 H(x; n, \pi, c^*)}{\partial c \partial \pi} = \sum_j (c^*)^2 \left( \frac{1}{c^{\pi_{i,j} + x_{i,j}}} - \frac{1}{c^{\pi_{i,j} + x_{i,j} + s}} \right)
\]

\[
\frac{\partial^2 H(x; n, \pi, c^*)}{\partial \pi_i \partial \pi_j} = (c^*)^2 \sum_s (c^*)^{\pi_{i,j} + x_{i,j} + s}
\]
Because $c^* > 0$ and $\pi_i > 0$ ($i = 1, \ldots, k$) by definition of the Beta-Binomial model, it is possible to determine $\Omega_{k0}$ such that the first and second partial derivatives of $H(x;n,\pi,c^*)$, as given above, are bounded by finite functions. Therefore, conditions 1 and 2 of Lemma 3.1 are satisfied (for condition 1 use the Lebesque Dominated Convergence Theorem and the equality $(\partial^2 \log f(x;\theta))/\partial \theta_1 \partial \theta_1 = (\partial f(x;\theta))/\partial \theta_1)$), and it follows that $-2\ln\Lambda$ is asymptotically distributed as a chi-square random variable for the Beta-Binomial model.

Using the likelihood ratio approach, tests of hypotheses concerning the $\pi_i$'s can be performed. For example, for $H_0: \pi \in \Pi^S$ versus $H_A: \pi \in \Pi^R$ ($s < r$), the test statistic

$$-2\ln(L(x^{(1)}, \ldots, x^{(N)}; n, \hat{\pi}_1, \ldots, \hat{\pi}_k, c^*)/L(x^{(1)}, \ldots, x^{(N)}; n, \hat{\pi}_1, \ldots, \hat{\pi}_k, c^*))$$

has an asymptotic chi-square distribution with $r-s$ degrees of freedom. Another hypothesis that can be tested using the likelihood ratio is that of no judge variability, i.e., $H_0: c^* = \infty$ versus $H_A: c^*$ unrestricted. The likelihood ratio, though, cannot be used to perform a goodness of fit
test for the Beta-Binomial model. On the other hand, the test statistics of the previous chapter can be used to perform such a test. These test statistics, as presented in Theorems 2.1 and 2.2, can also be used as alternatives to likelihood ratio tests to perform tests of hypotheses concerning the scale values.

In considering the asymptotic local power of these test statistics, it must be verified that there exists a sequence \( \mu^{(N)} \rightarrow \mu \) and a sequence \( \pi^{(N)} \in \Pi^r \) such that for \( \pi_s \in \Pi^s \) (\( s < r \))

\[
D(n_C)f_s(\pi_s) + \frac{\mu^{(N)}}{\sqrt{N}} = D(n_C)f_r(\pi_r) \tag{3.5-1}
\]

for each \( N \), where \( D(n_C) \) is a diagonal matrix with the elements of \( n_C \) on the diagonal. It is easy to show that such sequences exist. Define \( \alpha \) to be a \( k \times 1 \) vector whose elements are the elements of \( \pi_s \) in such a way that \( f_s(\pi_s) = f_k(\alpha) \). Define \( \alpha^{(N)} = \alpha + \varepsilon/\sqrt{N} \) where \( \varepsilon \) is such that \( \alpha^{(N)} \) consists of \( r \) distinct elements. Then, the \((i,j)\)-th element of

\[
D(n_C)(f_k(\alpha^{(N)}) - f_k(\alpha))
\]

is

\[
n(i,j) \left( \frac{\alpha^{(N)}}{\alpha_i + a_j} - \frac{\alpha_i^{(N)}}{\alpha_i^{(N)} + a_j} \right) = n(i,j) \frac{(\varepsilon_i a_j - \varepsilon_j a_i)(\alpha_i + a_j)^{-1}(\alpha_i^{(N)} + a_j)\alpha_i^{(N)}}{(\alpha_i + a_j)(\alpha_i^{(N)} + a_j)}.
\]

Define \( \mu^{(N)}(i,j) = n(i,j)(\varepsilon_i a_j - \varepsilon_j a_i)(\alpha_i + a_j)^{-1}(\alpha_i^{(N)} + a_j)^{-1} \) and \( \pi_r^{(N)} \) to be the vector whose elements are the \( r \) distinct elements of \( \alpha^{(N)} \). Because \( \mu^{(N)} \rightarrow \mu \), where \( \mu(i,j) = n(i,j)(\varepsilon_i a_j - \varepsilon_j a_i)(\alpha_i + a_j)^{-2} \), and \( \pi_r^{(N)} \in \Pi^r \) for each \( N \), it has been shown that there exist sequences \( \mu^{(N)} \rightarrow \mu \) and \( \pi_r^{(N)} \in \Pi^r \) that satisfy (3.5-1). Consequently, Theorems 2.1 and 2.2 can be used for the Beta-Binomial model.
The results of these theorems are summarized in Theorem 3.1. Recall that $u_{r,s} = M_{r,s}u$ and $t_{r,s} = M_{r,s}t$, where $M_{r,s} = D^{\frac{1}{2}}(v_0)(P_r - P_s) \times D^{-\frac{1}{2}}(v_0)D^{-1}(n_C)$, $P_s = A_s(A_s^tA_s)^{-1}A_s^t$, $A_s = D^{-\frac{1}{2}}(v_0)\frac{\partial f_s(\pi)}{\partial \pi}$, $\pi = \pi_0$.

$v_0 = D^{-1}(n_C)P_0$, $P_0 = f_s(\pi_0)$ and $\pi$ is the variance-covariance matrix of $X^C$, the elements of which are given by (3.3-4). Also, $\delta_i$ is used to denote the $i$-th element of $E^{-\frac{1}{2}}\Gamma D^{-\frac{1}{2}}(P_0)u_{r,s}$ and $\tilde{e}_{ii}$ the $i$-th diagonal element of $\tilde{E}$, where $\tilde{E} = \begin{pmatrix} E & 0 \\ 0 & I \end{pmatrix}$, $E$ is a diagonal matrix with diagonal elements equal to the nonzero eigenvalues of $D^{-\frac{1}{2}}(P_0) \tilde{f}_r,sD^{-\frac{1}{2}}(P_0)$, $\Gamma$ is the matrix of orthogonal eigenvectors corresponding to the elements of $E$, $I_K$ is the $K$ dimensional identity matrix and $t = \text{rank}(\tilde{f}_r,s)$.

**Theorem 3.1** Let $X^{(1)}, \ldots, X^{(N)}$ be i.i.d. $BB(n, \pi_1^{(N)}, \ldots, \pi_k^{(N)}, c)$ where the scale values $\pi_i^{(N)}$ are such that $D(n_C)\pi^{(N)} = E(X^{(1)}) = D(n_C)f_s(\pi_0) + \mu^{(N)}/\sqrt{N}$ and $\mu^{(N)} \to \mu$ as $N \to \infty$. Let $f_s(\hat{\pi}^{(N)})$ and $f_r(\hat{\pi}^{(N)})$ be the pseudo-MLEs for $\pi^{(N)}$ (using the pseudo-MLEs $\hat{\pi}_r$ and $\hat{\pi}_s$ for $\pi$) under $H_0: \pi^{(N)} = f_s(\pi_0)$, $\pi_0 \in \Pi^s$ and $H_1: \pi^{(N)} = f_r(\pi)$, $\pi \in \Pi^r$ ($s < r < k$), respectively. Then,

$$\begin{align*}
\text{NW}^2(\hat{\pi}^{(N)}), &\rightarrow \chi_r^2(\tilde{u}_{r,s} \tilde{t}_{r,s}), \\
\text{NX}^2(\hat{\pi}^{(N)}), &\rightarrow \text{NG}^2(\hat{\pi}^{(N)}), \\
\text{NX}^2(\hat{\pi}^{(N)}), &\rightarrow \chi_{s-1}^2(\tilde{\alpha}_{i-s}^2), \\
\text{NX}^2(\hat{\pi}^{(N)}), &\rightarrow \sum_{i=1}^{2} \frac{\delta_i^2}{\delta_i^2} + \frac{2}{2}.
\end{align*}$$
where $W^2$, $X^2$ and $G^2$ are as defined in (2.3-1), (2.3-2) and (2.3-3), respectively.

**Proof:** These results are a direct consequence of Theorems 2.1 and 2.2. The only thing that needs to be verified is that the rank of $\mathcal{P}_{r,s}$ equals $r-s$. This is done as follows.

If $f_s(\pi)$ has as its $(i,j)$-th element $q_{(q+\pi)}^{-1}$, then the $(i,j)$-th row of $\mathcal{P}_{r,s}$ consists of all zeros except the $q$-th column which contains $q_{(q+\pi)}^{-2}$ and the $r$-th column which contains $q_{(q+\pi)}^{-2}$. It is easy to see that the rank of $\mathcal{P}_{r,s}$ is $s$, and so the rank of $\mathcal{P}$ is $s$.

Furthermore, it is easily seen that $\mathcal{C}(\mathcal{P}_{r,s}) \subset \mathcal{C}(\mathcal{P})$, where $\mathcal{C}(X)$ represents the column space of $X$. Therefore, it follows that the rank of $\mathcal{P}_{r,s}$, and that of $\mathcal{P}_{r,s}$, equals $r-s$. Since the variance covariance matrix $\mathcal{P}$ has rank equal to $\binom{k}{2}$ and $\mathcal{C}(\mathcal{P}_{r,s}) \subset \mathcal{C}(\mathcal{P})$, it follows that $\text{rank}(\mathcal{P}_{r,s}) = r-s$. QED

Theorem 3.1 permits $W^2$, $X^2$ and $G^2$ to be used to test hypotheses about the scale values $\pi_i$, but not about the judge variability parameter $\pi$. A corollary to Theorem 3.1, similar to the corollaries to Theorems 2.1 and 2.2, provides a goodness of fit test for the Beta-Binomial model.

**Corollary 3.1** Let $X^{(1)}, \ldots, X^{(N)}$ be i.i.d. such that $E[X^{(1)}] = D[n_{\pi}]\hat{P}(N) = \hat{E}(X^{(1)}) = D[n_{\pi}]f_k(\pi) + \mu^{(N)}\sqrt{N}$ and $\text{Var}(X^{(1)}) = \mathcal{P}(N)$, where $\mu^{(N)} \rightarrow \mu$ and $\mathcal{P}(N) \rightarrow \mathcal{P}$, the variance covariance matrix of the Beta-Binomial model, as $N \rightarrow \infty$. Let $f_k(\hat{\pi}(N))$ be the pseudo-MLE for $P(N)$ (using the pseudo-MLE $\hat{\pi}(N)$ for $\pi$) under
H_0: \mathbf{P}^{(N)} = f_k(\mathbf{\pi}_0), \mathbf{\pi}_0 \in \mathbb{R}^k. Let the alternative hypothesis be H_A: \mathbf{P}^{(N)} is unrestricted. Then,

\[ \text{NW}^2(f_k(\hat{\mathbf{\pi}}^{(N)}), \mathbf{P}^{(N)}), \hat{\mathbf{f}}_{\text{GOF}}) \Rightarrow \chi^2_{(2)^{-k}}(\mathbf{\mu}_{\text{GOF}}, \hat{\mathbf{f}}_{\text{GOF}}, \mathbf{\mu}_{\text{GOF}}), \]

\[ \text{NX}^2(\hat{\mathbf{P}}^{(N)}, f_k(\hat{\mathbf{\pi}}^{(N)})) \Rightarrow \text{NG}^2(\hat{\mathbf{P}}^{(N)}, f_k(\hat{\mathbf{\pi}}^{(N)})) \]

and

\[ \text{NX}^2(\hat{\mathbf{P}}^{(N)}, f_k(\hat{\mathbf{\pi}}^{(N)})) \Rightarrow \sum_{i=1}^{(k)-k} \tilde{e}_{ii} \chi_1^2(\tilde{\delta}_i^2) + \sum_{i=(2)^{-k}}^{2(2)} \delta_i^2, \]

where \( \mathbf{\mu}_{\text{GOF}} = M^\dagger, \hat{\mathbf{f}}_{\text{GOF}} = M^\dagger M^\prime, M = D^{1/2}(\mathbf{\pi}_0)(I_{2(k)} - \mathbf{P}_k)D^{-1/2}(\mathbf{\pi}_0)D^{-1}(\mathbf{\pi}_0), \)

\( \delta_i \) is the \( i \)-th element of \( \tilde{\mathbf{E}}^{-1}_{2(k)}D^{-1}(\mathbf{\pi}_0)\mathbf{\mu}_{\text{GOF}}, \)

\( \tilde{\mathbf{E}} = \begin{pmatrix} E & 0 \\ 0 & I \end{pmatrix}, E \) is a diagonal matrix with diagonal elements equal to the nonzero eigenvalues of \( D^{-1/2}(\mathbf{\pi}_0)\hat{\mathbf{f}}_{\text{GOF}}D^{-1/2}(\mathbf{\pi}_0) \) and \( \Gamma \) is the matrix of orthogonal eigenvectors corresponding to the eigenvalues in \( E \).

\textbf{Proof:} This follows from the corollaries to Theorems 2.1 and 2.2 once it is noted that \( \mathcal{C}(\mathbf{P}_k) \subset \mathcal{C}(\hat{\mathbf{f}}_{\text{GOF}}) \mathbf{GOF} \) so that the rank of \( \hat{\mathbf{f}}_{\text{GOF}} \) is \( (k)_{2}^{-k} \). \( \text{QED} \)

For both Theorem 3.1 and its corollary, the true covariance matrices \( \hat{f}_{r,s} \) and \( \hat{f}_{\text{GOF}} \) are assumed known. A consistent estimator for these matrices can be used by Lemma 2.4 without altering the results of Theorem 3.1 and its corollary. In Chapter 5, these matrices will be estimated by substituting estimators for \( \mathbf{\pi}_1, \ldots, \mathbf{\pi}_k \) and \( c^* \), as given in
Section 3.4. Note that these estimators are all consistent, so that the estimates of $\hat{r}_{rs}$ and $\hat{r}_{GOF}$, based on these estimators, are themselves consistent.

3.6 A Reparameterized Beta-Binomial Model

Another form of the Beta-Binomial model was proposed by Lancaster and Quade (1983). Their model differs from the Beta-Binomial proposed in Section 3.3 with respect in the way in which the model is parameterized. Specifically, Lancaster and Quade modeled the choice random variables using a binomial distribution and assumed that the choice random variables were all independent of one another. Furthermore, they assumed that the choice probabilities were independently distributed as Beta random variables. However, rather than use scale values as the parameters for the Beta distribution, they used $\alpha(q,r)$ and $\beta(q,r)$ as the parameters associated with $p(q,r)$. The joint distribution of the choice vector then is

$$h(x;n,\alpha,\beta) = \prod_{q<r} \frac{\Gamma(x(q,r) + \alpha(q,r)) \Gamma(x(r,q) + \beta(q,r)) \Gamma(\alpha(q,r) + \beta(q,r))}{\Gamma(\alpha(q,r)) \Gamma(\beta(q,r))}. $$

In order to reduce the number of parameters from $\binom{k}{2}$ to $k+1$, Lancaster and Quade introduced two restrictions on the model. The first condition they imposed was that the means of the Beta distribution should satisfy the Bradley-Terry condition, i.e., $E(p_{q,r}) = \alpha(q,r)^{x(q,r)} + \beta(q,r)^{x(q,r)}$ where $\sum_{q=1}^{k+1} \pi_q = 1$. This is the same condition
that both the Beta-Bernoulli and the Beta-Binomial models satisfy. The
other condition they imposed is that the variance of the choice
probabilities should be proportional to the binomial variance, i.e.,
\[
\text{Var}(p_{q,r}) = \lambda E(p_{q,r})(1-E(p_{q,r})).
\]
This condition is equivalent to
\[
\frac{\alpha(q,r)\beta(q,r)}{(\alpha(q,r)+\beta(q,r))^2(\alpha(q,r)+\beta(q,r)+1)} = \lambda \frac{\alpha(q,r)\beta(q,r)}{(\alpha(q,r)+\beta(q,r))^2},
\]
which reduces to \(\alpha(q,r) + \beta(q,r) = v\), where \(v = \lambda^{-1} - 1\). Their justi-
fication for this restriction is that it is reasonable to expect the
variance to be small when \(p_{q,r}\) is close to 0 or 1. For the Beta-
Binomial model, as well as the Beta-Bernoulli model, \(\text{Var}(p_{q,r}) = \pi_q \hat{\pi}_r (\pi_q + \pi_r)^{-2}(c^*_q \pi_q + c^*_r \pi_r + 1) \neq \lambda \pi_q \hat{\pi}_r (\pi_q + \pi_r)^{-2} = \lambda E(p_{q,r})(1-E(p_{q,r})),\)
that is, the Lancaster and Quade restriction on the variance of the
choice probabilities does not hold. It is with respect to this
parameterization that the Lancaster and Quade and the Beta-Binomial
differ from one another. As a result, for Lancaster and Quade's model
the parameters associated with the Beta distribution for \(p_{q,r}\) are
\(\pi_q \hat{\pi}_r (\pi_q + \pi_r)^{-1}\) and \(\pi_q \hat{\pi}_r (\pi_q + \pi_r)^{-1}\) whereas for the Beta-Binomial models
these parameters are \(c^*_q \pi_q\) and \(c^*_r \pi_r\).

Using these restrictions, Lancaster and Quade reparameterized
the model in terms of \(\pi_i (i=1,\ldots,k)\) and \(v\) and obtained
Note that this is just a reparameterization of the marginal distribution
for the Beta-Binomial model as given in (3.3-1). The parameter \( v \) is a
between judge variability parameter like \( c^* \) for the Beta-Binomial model,
and as \( v \to \infty \) this model tends to the Bradley-Terry model, as does the
Beta-Binomial model when \( c^* \to \infty \). The scale values \( \pi_i \) for Lancaster
and Quade's model similarly correspond to the scale values of the Beta-
Binomial. Lancaster and Quade obtained MLEs for the \( \pi_i \) and \( v \) by
numerical techniques and performed tests of hypotheses about the scale
values and the judge variability parameter using the likelihood ratio
test.

In this chapter, the Beta-Binomial model was introduced as a
possible judge effect model for paired comparison experiments. It was
seen that for the case \( n_{(q,r)} = 1 \) for all comparisons, so that the Beta-
Bernoulli model holds, the model is no longer a true judge effect model
because by among judge variability cannot be estimated. With respect
to estimation and hypothesis testing of the scale values, the Beta-
Bernoulli model and the Bradley-Terry model are equivalent. The Beta-
Binomial model is an improvement over the Bradley-Terry model in that
it incorporates a measure of judge variability. The Beta-Binomial

\[
h(x;\eta, \pi_1, \ldots, \pi_k, v) =
\]

\[
\prod_{q < r} \binom{n(q,r)}{x(q,r)} \frac{\Gamma(x(q,r) + \nu \pi r(q,r)) \Gamma(x(q,r) + \nu \pi r(q,r)) \Gamma(\pi r(q,r))}{\Gamma(n(q,r) + \nu \pi r(q,r)) \Gamma(v \pi r(q,r)) \Gamma(\nu \pi r(q,r))}
\]

Note that this is just a reparameterization of the marginal distribution
for the Beta-Binomial model as given in (3.3-1). The parameter \( v \) is a
between judge variability parameter like \( c^* \) for the Beta-Binomial model,
and as \( v \to \infty \) this model tends to the Bradley-Terry model, as does the
Beta-Binomial model when \( c^* \to \infty \). The scale values \( \pi_i \) for Lancaster
and Quade's model similarly correspond to the scale values of the Beta-
Binomial. Lancaster and Quade obtained MLEs for the \( \pi_i \) and \( v \) by
numerical techniques and performed tests of hypotheses about the scale
values and the judge variability parameter using the likelihood ratio
test.

In this chapter, the Beta-Binomial model was introduced as a
possible judge effect model for paired comparison experiments. It was
seen that for the case \( n_{(q,r)} = 1 \) for all comparisons, so that the Beta-
Bernoulli model holds, the model is no longer a true judge effect model
because by among judge variability cannot be estimated. With respect
to estimation and hypothesis testing of the scale values, the Beta-
Bernoulli model and the Bradley-Terry model are equivalent. The Beta-
Binomial model is an improvement over the Bradley-Terry model in that
it incorporates a measure of judge variability. The Beta-Binomial

\[
h(x;\eta, \pi_1, \ldots, \pi_k, v) =
\]

\[
\prod_{q < r} \binom{n(q,r)}{x(q,r)} \frac{\Gamma(x(q,r) + \nu \pi r(q,r)) \Gamma(x(q,r) + \nu \pi r(q,r)) \Gamma(\pi r(q,r))}{\Gamma(n(q,r) + \nu \pi r(q,r)) \Gamma(v \pi r(q,r)) \Gamma(\nu \pi r(q,r))}
\]
models, however, ignores any possible correlation between different comparisons by the same judge. In the next chapter, a model is proposed that introduces a correlation between various comparisons by a given judge.
4. A JUDGE EFFECT MODEL BASED ON RANDOM SCALE VALUES

4.1 Introduction

A limitation of the models presented in the previous chapter is that different comparisons made by the same judge are assumed independent. A more reasonable model would allow for correlations among comparisons by the same judge. In this chapter, such a model is developed. First, the model is defined and briefly compared with the models presented in the previous chapter. Next, the problem of parameter estimation is discussed and two test statistics for hypothesis testing are introduced.

4.2 The Dirichlet-Binomial Model

In the previous chapter, a judge effect was included in the models by assuming that the choice probabilities were independently distributed random variables. Such an assumption implies that the choice probabilities, and the choice random variables, cannot be correlated with one another. Another approach to introducing a judge effect is to view the vector of scale values \((\pi_1, \ldots, \pi_K)\) as a random vector. Associated with each randomly chosen judge is a vector \(v\) representing the judge's scale values, and the corresponding choice probabilities can be determined from these scale values using the Bradley-Terry condition.

Conditional on a particular judge, the choice random variables \(X_{(q,r)}\) are assumed to be independently distributed as

\[
\text{Binomial} \left( n_{(q,r)} \cdot v^+ q \left( \sum_r v^+_r \right)^{-1} \right)
\]

random variables. The vector of scale
values in turn is assumed to have a Dirichlet distribution with parameters \( \pi_1, \ldots, \pi_k \), the true scale values, and \( c^* \), a judge variability parameter. This is formalized in the following definition.

**Definition 4.1** A \((k \choose 2)\)x1 choice vector \( X \) is said to follow a Dirichlet-Binomial model if the density function for the joint distribution of \( X \) and \( V \), the kx1 vector of random scale values, is given by

\[
g(X, V; \pi_1, \ldots, \pi_k, c^*) = \frac{\Gamma(c^*)}{\prod_i c_i^{\pi_i - 1}} \prod_i v_i^{\pi_i - 1} \prod_{q < r} x(q, r)^{v_q - v_r} x(r, q)^{v_r - v_q} \cdot \frac{x(q, r)}{x(q, r)}^{v_q - v_r} \cdot \frac{x(r, q)}{x(r, q)}^{v_r - v_q},
\]

where \( x(q, r) \in \{0, \ldots, n(q, r)\} \); \( x(q, r) = n(q, r) - x(r, q) \); \( v_i > 0 \); \( \sum_{i=1}^k v_i = 1 \),

where \( x(q, r) \) is the number of times item \( q \) is preferred to item \( r \), \( n(q, r) \) is the number of times items \( q \) and \( r \) are compared to one another, \( \pi_i \) is the true scale value associated with item \( i \) satisfying \( \pi_i > 0 \) and \( \sum_{i=1}^k \pi_i = 1 \), and \( c^* \) is a judge variability parameter satisfying \( c^* > 0 \).

The model will be denoted by DB \((n, \pi_1, \ldots, \pi_k, c^*)\). The parameter \( c^* \) is a measure of among judge variability much like \( c^* \) in Chapter 3. Under the Dirichlet distribution, the random vector \( V \) has mean vector \((\pi_1, \ldots, \pi_k)\)' and each element of the variance-covariance matrix is proportional to \((c^* + 1)^{-1}\). Therefore, as \( c^* \to \infty \), the vector of random
scale values \( V \) becomes less variable and more tightly centered around the true vector of scale values \((\pi_1, \ldots, \pi_k)^t\) and the Dirichlet-Binomial model tends to the Bradley-Terry model.

In order to compare the Dirichlet-Binomial model with the models of the previous chapter, the values of \( E(X) \) and \( \text{Var}(X) \) will be calculated when the choice vector \( X \) is distributed as \( \text{DB}(n, \pi_1, \ldots, \pi_k, c^*) \). The following lemmas are needed to calculate the values.

**Lemma 4.1**
\[
\int_0^u x^{a-1} (u-x)^{b-1} \, dx = u^{a+b-1} \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad (a, b > 0).
\]

**Proof:** See Gradshteyn and Ryzik (1980, p.284). \( \text{QED} \)

**Definition 4.2**
\[(a, n) = a(a+1)\ldots(a+n-1), \quad (a \in \mathbb{R}; \ n \in \{1, 2, \ldots\}).\]

**Lemma 4.2**
\[
\Gamma(z+m) = (z,m) \Gamma(z), \quad (z \in \mathbb{R}; \ m \in \{1, 2, \ldots\}).
\]

**Proof:** \[
\Gamma(z) = \frac{\Gamma(z+1)}{z} = \frac{\Gamma(z+2)}{z(z+1)} = \ldots = \frac{\Gamma(z+m)}{(z,m)}. \quad \text{QED}
\]

**Lemma 4.3**
\[(a, m+n) = (a, m)(a+m, n), \quad (a \in \mathbb{R}; \ n, m \in \{1, 2, \ldots\}).\]

**Proof:**
\[
(a, m+n) = [a(a+1)\ldots(a+m-1)] \times [(a+m)(a+m+1)\ldots(a+m+n-1)]
= (a, m)(a+m, n). \quad \text{QED}
\]
Lemma 4.4
\[ \sum_{n=0}^{\infty} \frac{(a,n)(b,n)}{(c,n)n!} = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}, \quad (a,b \in \mathbb{R}, \, c \in \mathbb{R}-\{0,-1,-2,\ldots\}), \]
holds if either
(i) \( c-a-b > 0 \),
or (ii) \(-a \in \{0,1,2,\ldots\} \),
or (iii) \(-b \in \{0,1,2,\ldots\} \).

Proof: See Carlson (1977, p.242, Corollary 8.3-4). QED

Lemma 4.5
\[ \int_{0}^{1} \int_{0}^{1-u_{1}} \int_{0}^{1-u_{1}-u_{2}} [B(b)]^{-1} \frac{b_{1}-1}{u_{1}} \frac{b_{2}-1}{u_{2}} \frac{b_{3}-1}{u_{3}} \frac{b_{4}-1}{(u_{1}+u_{2})(u_{1}+u_{3})} \, du_{3} \, du_{2} \, du_{1} \]
holds if
(i) \( b_{1}+b_{2}+b_{3} > 2 \)
and (ii) \( b_{1}+b_{2} > 1 \),
where \( b = (b_{1},b_{2},b_{3},b_{4}) \), \( b_{1},b_{2},b_{3},b_{4} > 0 \).

Proof:
\[ \int_{0}^{1} \int_{0}^{1-u_{1}} \int_{0}^{1-u_{1}-u_{2}} [B(b)]^{-1} \frac{b_{1}-1}{u_{1}} \frac{b_{2}-1}{u_{2}} \frac{b_{3}-1}{u_{3}} \frac{b_{4}-1}{(u_{1}+u_{2})(u_{1}+u_{3})} \, du_{3} \, du_{2} \, du_{1} \]
\[
\begin{align*}
&= \int_{0}^{\infty} \int_{0}^{\infty} \int_{0}^{\infty} \int_{0}^{\infty} [B(b)]^{-1} u_1 b_1 - 1 u_2 b_2 - 1 u_3 b_3 - 1 (1-u_1-u_2-u_3) b_4 - 1 \\
&\quad \times \int_{0}^{\infty} \int_{0}^{\infty} -(u_1+u_2)s -(u_1+u_3)t \\
&\quad \times e^{-u_1(s+t)-u_2s-u_3t} du_3 du_2 du_1 ds dt \\
&\quad \times (by \ Tonelli's \ Theorem) \\
&= \int_{0}^{\infty} \int_{0}^{\infty} \int_{0}^{\infty} \int_{0}^{\infty} [B(b)]^{-1} u_1 b_1 - 1 u_2 b_2 - 1 u_3 b_3 - 1 (1-u_1-u_2-u_3) b_4 - 1 \\
&\quad \times e^{u_2 t+u_3 s+(1-u_1-u_2-u_3)(s+t)} \\
&\quad \times du_3 du_2 du_1 ds dt \\
&= \int_{0}^{\infty} \int_{0}^{\infty} \int_{0}^{\infty} \int_{0}^{\infty} [B(b)]^{-1} u_1 b_1 - 1 u_2 b_2 - 1 u_3 b_3 - 1 (1-u_1-u_2-u_3) b_4 - 1 \\
&\quad \times \left[ \sum_{N=0}^{\infty} \frac{1}{N!} (u_2 t+u_3 s+(1-u_1-u_2-u_3)(s+t))^N \right] \\
&\quad \times du_3 du_2 du_1 ds dt
\end{align*}
\]
\[
= \sum_{N=0}^{\infty} \int_0^\infty \int_0^\infty e^{-(s+t)} \int_0^1 \int_0^{1-u_1} \int_0^{1-u_1-u_2} \frac{(B(b))^{-1}}{N!} u_1^{-1} u_2^{-1} u_3^{-1} \\
\times (1-u_1-u_2-u_3)^b (u_2 t + u_3 s + (1-u_1-u_2-u_3)(s+t))^N du_3 du_2 du_1 dsdt \\
\text{(by the Monotone Convergence Theorem)}
\]

\[
= \sum_{N=0}^{\infty} \int_0^\infty \int_0^\infty e^{-(s+t)} \int_0^1 \int_0^{1-u_1} \int_0^{1-u_1-u_2} \frac{(B(b))^{-1}}{N!} u_1^{-1} u_2^{-1} u_3^{-1} \\
\times (1-u_1-u_2-u_3)^b (u_2 t + u_3 s + (1-u_1-u_2-u_3)(s+t))^3 du_3 du_2 du_1 dsdt \\
\text{(by the Multinomial Theorem)}
\]

\[
= \sum_{N=0}^{\infty} \int_0^\infty \int_0^\infty e^{-(s+t)} \int_0^1 \int_0^{1-u_1} \int_0^{1-u_1-u_2} \frac{(B(b))^{-1}}{N!} u_1^{-1} u_2^{-1} u_3^{-1} \\
\times (1-u_1-u_2-u_3)^b (u_2 t + u_3 s + (1-u_1-u_2-u_3)(s+t))^3 du_3 du_2 du_1 dsdt \\
\text{(by the Multinomial Theorem)}
\]
\[
\begin{aligned}
&= \sum_{N=0}^{\infty} \int_{0}^{\infty} \int_{0}^{\infty} e^{-(s+t)} \sum_{m_1, m_2, m_3} \frac{[B(b)]^{-1}}{m_1! m_2! m_3!} t_{m_1 m_2 m_3}^{s+t} \times B(b_1, b_2 + m_1, b_3 + m_2, b_4 + m_3) ds dt \\
&\quad \text{(by repeated application of Lemma 4.1)} \\
&= \sum_{N=0}^{\infty} \sum_{m_1, m_2, m_3} \frac{(b_1, m_1) (b_2, m_2) (b_3, m_3)}{m_1! m_2! m_3!} \int_{0}^{\infty} \int_{0}^{\infty} e^{-(s+t)} t_{m_1 m_2 m_3}^{s+t} ds dt \\
&\quad \text{(by repeated application of Lemma 4.2)} \\
&= \sum_{N=0}^{\infty} \sum_{m_1, m_2, m_3} \frac{(b_2, m_1) (b_3, m_2) (b_4, m_3)}{m_1! m_2! m_3!} \int_{0}^{1} \int_{0}^{1} e^{-x (ur)} u_{m_1 m_2 m_3}^{(1-u)r} r dr du \\
&\quad \text{(using the transformation } r = s+t, u = t(s+t)^{-1}) \\
&= \sum_{N=0}^{\infty} \sum_{m_1, m_2, m_3} \frac{(b_2, m_1) (b_3, m_2) (b_4, m_3)}{m_1! m_2! m_3!} \int_{0}^{1} e^{-r N+1} dr \int_{0}^{1} u_{m_1 m_2 m_3}^{(1-u)} du \\
&= \sum_{N=0}^{\infty} \sum_{m_1, m_2, m_3} \frac{(b_2, m_1) (b_3, m_2) (b_4, m_3)}{m_1! m_2! m_3!} \Gamma(N+2) B(m_1 + 1, m_2 + 1) \\
&\quad \text{with } m_1 + m_2 + m_3 = N
\end{aligned}
\]
\[
\sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \sum_{m_3=0}^{\infty} \frac{(2, m_1+2, m_2+2, m_3)(b_2, m_1)(b_3, m_2)}{(c, m_1+2, m_2+2, m_3)(2, m_1+2, m_2)} \\
\sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \frac{(2, m_1+2, m_2)(b_2, m_1)(b_3, m_2)}{(c, m_1+2, m_2)(2, m_1+2, m_2)} \sum_{m_3=0}^{\infty} \frac{(2, m_1+2, m_3)(b_4, m_3)}{(c, m_1+2, m_2+2, m_3)} \\
\text{(by application of Lemma 4.3)}
\]

\[
\sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \frac{(b_2, m_1)(b_3, m_2)}{(c, m_1+2, m_2)} \frac{\Gamma(c+b_1+2)}{\Gamma(c-2)\Gamma(c+b_4+1+2)} \\
\text{(by application of Lemma 4.4, using the assumption that } b_1+b_2+b_3 > 2) \\
\sum_{m_1=0}^{\infty} \sum_{m_2=0}^{\infty} \frac{(b_2, m_1)(b_3, m_2)}{(c, m_1+2, m_2)} \frac{\Gamma(c+b_1+2)(c+b_2+2)}{\Gamma(c-2)\Gamma(c+b_4+1+2)} \\
\text{(by application of Lemma 4.2)}
\]

\[
\frac{(c-1)(c-2)}{(c-b_4+1)(c-b_4+2)} \sum_{m_1=0}^{\infty} \frac{(b_2, m_1)}{(c-b_4+1, m_1)} \sum_{m_2=0}^{\infty} \frac{(b_3, m_2)}{(c-b_4+1, m_2)} \frac{(1, m_2)}{m_2} \\
\text{(by application of Lemma 4.3)}
\]

\[
\frac{(c-1)(c-2)}{(c-b_4+1)(c-b_4+2)} \sum_{m_1=0}^{\infty} \frac{(b_2, m_1)}{(c-b_4+1, m_1)} \frac{\Gamma(c+b_1+2)\Gamma(c-b_2-b_4-l+m_1)}{\Gamma(c-b_3-b_4+1+m_1)} \\
\text{(by application of Lemma 4.3)}
\]
by application of Lemma 4.4, using the assumption that $b_1 + b_2 > 1$)

$$
\frac{(c-1)(c-2)}{(c-b_4-2)} \sum_{m_1=0}^{\infty} \frac{(b_2, m_1)}{(c-b_4-1, m_1)} \frac{1}{c-b_3-b_4-1+m_1}.
$$

QED

The unconditional mean for the Dirichlet-Binomial model, $E(X)$, is now derived.

**Theorem 4.1** Let the choice vector $X$ be distributed as $DB(n, \pi_1, \ldots, \pi_k, c^*)$. Then, $E(X) = D(n)\pi^*$, where $\pi^*$ is a $(k \times 1)$ vector in which the $(q,r)$-th element is $\pi_q (\pi_q + \pi_r)^{-1}$ and $D(n)$ is a diagonal matrix with the elements of $n$ on the diagonal.

**Proof:** First, note that the joint distribution of $X|v$ is just the joint distribution of $\binom{k}{q}$ independent binomial random variables.

Then, the $(q,r)$-th element of the vector $E(X|v)$ is $n(q,r)\frac{v_q}{v_q + v_r}$. Using the identity $E(X) = E(E(X|v))$, the $(q,r)$-th element of $E(X)$ can be calculated as

$$
E\left(\frac{n(q,r)\frac{v_q}{v_q + v_r}}{\frac{v_q}{v_q + v_r}}\right) = n(q,r) \sum_{i=1}^{\infty} \frac{\Gamma(c^*)}{\Gamma(c^* i)} \prod_{i=1}^{c^*} \pi_i^{-1} dv.
$$
\[
\frac{n(q,r)\Gamma(c^*)}{\prod (c^*_{\pi_i})} \int \cdots \int (v_q + v_r)^{-1} \left( \prod_{\substack{i=1, \ i \neq q}} v_i \right)^{k-1} c^*_{\pi_i-1} c^* q (1-\Sigma v_i) \, dv_{k-1} \cdots dv_1
\]

\[
= \frac{n(q,r)\Gamma(c^*)}{\Gamma(c^*_{\pi_q})\Gamma(c^*_{\pi_r})\Gamma(c^*(1-\pi_q-\pi_r))} \int_0^{1-v_q} \int_0^{1-v_r} (v_q + v_r)^{-1} c^*_{\pi_q} c^*_{\pi_r}^{-1} \\
\times (1-v_q - v_r) \, dv_q \, dv_r
\]

(by repeated application of Lemma 4.1 and the fact that \(\Sigma \pi_i = 1\))

\[
= \frac{n(q,r)\Gamma(c^*)}{\Gamma(c^*_{\pi_q})\Gamma(c^*_{\pi_r})\Gamma(c^*(1-\pi_q-\pi_r))} \int_0^{\pi_q} \int_0^{\pi_q} c^*_{\pi_q} c^*_{\pi_q}^{-1} \\
\times (1-\pi_q - \pi_q) \, dx_q \, dx_r
\]

(using the transformation \(x_q = v_q, x_r = v_q + v_r\))

\[
= \frac{n(q,r)\Gamma(c^*)c^*_{\pi_q}}{(c^*_{\pi_q} + \pi_r + 1)\Gamma(c^*(1-\pi_q-\pi_r))} \int_0^{\pi_q} \int_0^{\pi_q} c^*_{\pi_q} c^*_{\pi_q}^{-1} \pi_r (1-\pi_r) \, dx_r
\]

(by Lemma 4.1)
To obtain the variance-covariance matrix under the Dirichlet-Binomial model it is necessary to calculate the matrix of expected values of the squares and cross-products for the elements of the choice vector $X$. The calculations for the expected cross-products consists of two separate computations, one for the cross-product between comparisons with no item in common and one for the cross-products between comparisons with one item in common. This latter case is further broken down into three subcases depending on whether the item in common is preferred in both, one, or none of the comparisons.

\[ E(XX') = M, \text{ a } \binom{k}{2} \times \binom{k}{2} \text{ matrix with elements} \]

\[
(1) \quad m(q,r)(q,r) = \frac{n(q,r)^{\pi q}}{\pi + \pi r} + \frac{n(q,r)(n(q,r)-1)^{\pi q}}{q + \pi r}(c^* \pi q + 1),
\]

\[
(2) \quad m(q,r)(s,t) = \frac{n(q,r)^{\pi q}}{\pi + \pi r}(s^* \pi r)
\]

\[ (q,r = 1,2,...,k; q \neq r), \]

\[ (q,r,s,t = 1,2,...,k; q \neq s, q \neq t, r \neq s, r \neq t), \]

\[ \text{QED} \]
Proof: Note that the joint distribution of $X|V$ is just the joint distribution of $\binom{k}{2}$ independent binomial random variables. Therefore, the diagonal elements of $E(XX' | V)$ are the second order moments of binomial distributions, and the off-diagonal elements are the products of two binomial means. Specifically, the $( (q,r)(q,r))$ diagonal element of $E(XX' | V)$ is

$$
\sum_{n=0}^{\infty} \frac{c^*_{(\pi+r)+1,n}}{c (\pi+r+s)+1+n}
$$

and the off-diagonal elements are the products of two binomial means.
and the \(((q,r)(s,t))\) off diagonal element of \(E(XX' | Y)\) is

\[
E \left( \frac{n(q,r)^n(s,t)^q \cdot s}{(v_q + v_r)(v_s + v_t)} \right).
\]

Then, \(E(XX')\) is calculated using the identity \(E(XX') = E(E(XX' | Y))\), as follows:

**Case 1** Using 4.2-1, \(E(XX')\) becomes

\[
E(\frac{n(q,r)^n(s,t)^q \cdot s}{(v_q + v_r)(v_s + v_t)}) + \frac{n(q,r)^n(q,r)^{-1} v^2}{(v_q + v_r)^2}.
\]

From Theorem 4.1, \(E(\frac{n(q,r)^n(s,t)^q}{v_q + v_r}) = \frac{n(q,r)^n}{(v_q + v_r)^2}\). Now

\[
E \left( \frac{v^2_{q \cdot r}}{(v_q + v_r)^2} \right) = \int \frac{v^2_{q \cdot r}}{(v_q + v_r)^2} \frac{\Gamma(c^*)}{\prod \Gamma(c^* \cdot i)} \prod \frac{c^* \cdot i - 1}{v_i} dv_i
\]

\[
= \frac{\Gamma(c^*)}{\prod \Gamma(c^* \cdot i)} \int \prod \frac{v_{q \cdot r}}{(v_q + v_r)^2} \left( \sum_{i=1}^{k-1} v_i \right)^{-2} \prod_{i=1}^{k-1} \frac{c^* \cdot i - 1}{v_i} \prod_{i \neq q} \frac{c^* \cdot i - 1}{v_i} \prod_{i=1}^{k-1} \frac{c^* \cdot i - 1}{v_i}
\]

\[
\times \frac{v_{q \cdot r}}{(1 - \sum_{i=1}^{k-1} v_i)^k} \left( 1 - \sum_{i=1}^{k-1} v_i \right) dv_{k-1} \ldots dv_1
\]

\[
= \frac{\Gamma(c^*)}{\prod \Gamma(c^* \cdot q) \Gamma(c^* \cdot r) \Gamma(c^* \cdot (1 - \pi^\cdot q \cdot r))} \int_0^{1-v_q} \left( \frac{1}{v_q + v_r} \right)^{-2} \frac{v_{q \cdot r}}{(1 - v_q - v_r)} \left( 1 - v_q - v_r \right) dv_r dv_q
\]
(by repeated application of Lemma 4.1, and the fact that $\Sigma_1 \pi_1 = 1$)

$$\frac{\Gamma(c^*)}{\Gamma(c^*\pi_q)\Gamma(c^*\pi_r)\Gamma(c^*(1-\pi_q-\pi_r))} \int_0^1 \int_0^x (x q q^{-2} c^\pi q + 1 \frac{c^\pi q - 1}{x q q} (x q q - x q)) \times (1 - x q) \, dx q \, dx r$$

(using the transformation $x q = v q, x r = v q + v r$)

$$\frac{\Gamma(c^*)(c^*\pi + 1)c^\pi q}{\Gamma(c^*(\pi q + \pi r) + 2)\Gamma(c^*(1-\pi q - \pi r))} \int_0^1 (c^*(1-\pi q) + 1) c^*(1-\pi q - \pi r) + 1 \, dx r$$

(by Lemma 4.1)

$$\frac{\pi q (c^* \pi + 1)}{(\pi q + \pi r)(c^* (\pi q + \pi r) + 1)}$$

Therefore, $m(q, r)(q, r) = \frac{n(q, r)(q, r) \pi q + n(q, r)(q, r) - 1}{\pi q + \pi r} \frac{\pi q (c^* \pi + 1)}{(\pi q + \pi r)(c^* (\pi q + \pi r) + 1)}$.

Case 2 Using 4.2-2, the calculations for the $(q, r, (s, t))$ cross-product term of $E(XX')$ reduces to calculating $E(\frac{n(q, r)(s, t)v q v s}{(v q + v r)(v s + v t)}).$

Now
\[
E\left(\frac{v_q v_s}{(v_q + v_r)(v_s + v_t)}\right) = \sum_{i \neq q, s} \frac{\Gamma(c^*_i)}{\prod_{i} \Gamma(c^*_{\pi_i})} \prod_{i} v_i^{c^*_{\pi_i}-1} d\nu
\]

\[
= \frac{\Gamma(c^*)}{\prod_{i} \Gamma(c^*_{\pi_i})} \int \ldots \int_{v_i < 1} (v_q + v_r)^{-1}(v_s + v_t)^{-1} \prod_{i=1}^{k-1} c^*_{\pi_i}^{-1} \\
\times (1 - \sum_{i=1}^{k-1} v_i)^{-1} v_q v_s d\nu_k \ldots d\nu_1
\]

\[
= K \int_{0}^{1-v_q} \int_{0}^{1-v_r} \int_{0}^{1-v_q-v_r} (v_q + v_r)^{-1}(v_s + v_t)^{-1} c^*_{\pi_q} c^*_{\pi_r}^{-1} \\
\times c^*_{\pi_t} c^*_{\pi_t-1} (1-v_q-v_r-v_s-v_t) d\nu_t d\nu_s d\nu_r d\nu_q,
\]

(4.3-3)

where \( K = \frac{\Gamma(c^*)}{\Gamma(c^*_{\pi_q})\Gamma(c^*_{\pi_r})\Gamma(c^*_{\pi_t})} \frac{\Gamma(c^*_{\pi_t})}{\Gamma(c^*(1-\pi_r-\pi_s-\pi_t))} \frac{\Gamma(c^*_{\pi_t})}{\Gamma(c^*(1-\pi_q-\pi_r-\pi_s-\pi_t))} \),

(by repeated application of Lemma 4.1, and the fact that \( \sum_{i=1}^{k-1} v_i = 1 \)).

Looking at the innermost double integral,

\[
\int_{0}^{1-v_q} \int_{0}^{1-v_r} (v_q + v_r)^{-1} c^*_{\pi_q} c^*_{\pi_t-1} \\
\times (1-v_q-v_r-v_s-v_t) d\nu_t d\nu_s
\]

\[
= c^*(1-\pi_q-\pi_r-\pi_s-\pi_t)^{-1} d\nu_t d\nu_s
\]
In the proof of Theorem 4.1, it has been shown that $E\left[\frac{v_q}{v_q + v_r}\right] = \frac{\pi_q}{\pi_q + \pi_r}$, where $v_q, v_r$ have a joint distribution given by a Dirichlet distribution. Noting that the above double integral is proportional to such an expectation, it is easily seen to be equal to

$$
(1-v_q-v_r) \frac{c^*(1-\pi_q - \pi_r - \pi_s - \pi_t)}{(1-v_q-v_r)} \int_0^{1-x_s} \int_0^{1-x_t} (x_q^++x_t^-)^{-1} x_s^+ x_t^- \left(1 - \frac{v_q + v_r}{1-v_q-v_r}\right) dv_t dv_s
gains
(1-v_q-v_r) \int_0^{1-x_s} \int_0^{1-x_t} (x_q^++x_t^-)^{-1} x_s^+ x_t^- \left(1 - \frac{v_q + v_r}{1-v_q-v_r}\right)
\times (1-x_q^--x_t^+) \left(1 - \frac{v_q + v_r}{1-v_q-v_r}\right) \frac{c^*(1-\pi_q - \pi_r - \pi_s - \pi_t)}{(1-v_q-v_r)} \Gamma(c^*(1-\pi_q - \pi_r - \pi_s - \pi_t)) \frac{\pi_s}{\pi_q + \pi_r} \cdot
$$

Substituting this into (4.3-3), $E\left\{\frac{v_q v_s}{(v_q + v_r)(v_s + v_r)}\right\}$ becomes
\[
\frac{\pi_s}{\pi_s + \pi_t} \frac{\Gamma(c^*)}{\Gamma(c^*_{q})\Gamma(c^*_{r})\Gamma(c^*_{(1-\pi_q-\pi_r)})} \int_0^1 \int_0^{1-v_q} (v_{q+r})^{-1} v^*_q v^*_r \left(1-v^*_q-v^*_r\right) \left(c^*_{\pi_q} c^*_{\pi_r}\right)^{-1} \\
\times (1-v^*_q-v^*_r) \text{d}v^*_r \text{d}v^*_q.
\]

However, this is nothing other than \(\pi_s / (\pi_s + \pi_r) E\left[n_{q}/(v_q + v_r)\right]\), where \(v_q\) and \(v_r\) are jointly distributed with a Dirichlet distribution. It follows that

\[
E\left[n_{q} n_{s}(s,t)\pi_q \pi_s \right] = \frac{\pi_q \pi_s}{(\pi_q + \pi_r)(\pi_s + \pi_t)}.
\]

Therefore, \(m(q,r)(s,t) = \frac{n(q,r,s)n_{q,s}(s,t)\pi_q \pi_s}{(\pi_q + \pi_r)(\pi_s + \pi_t)}\).

**Case 3** Using (4.3-2), \(E(XX)\) becomes \(E\left[n_{q,s} h(q,r,s) / (v_{q+r})(v_{q+s})\right]\) where

\(h(q,r,s)\) equals \(v^2_q, v_q v_s, \) or \(v_r v_s, v_r v_s, \) depending on whether the item in common (item \(q\)) is preferred in both, one, none of the comparisons, respectively.

In general,

\[
E\left[\frac{h(q,r,s)}{(v_{q+r})(v_{q+s})}\right] = \sum_{i=1}^{\epsilon} \frac{h(q,r,s)}{(v_{q+r})(v_{q+s})} \frac{\Gamma(c^*)}{\Gamma(c^*_{\pi_i})} \prod_{i} c^*_{\pi_i^{-1}} \text{d}v^*_i
\]
\[
\frac{\Gamma(c^*)}{\prod_{i} \Gamma(c^{*}_{\pi_i})} \left( \sum_{i=1}^{k-1} \frac{h(q,r,s)}{(v^q + v^r)(v^q + v^s)} \prod_{i=1}^{k-1} v_i \right) \left( 1 - \sum_{i=1}^{k-1} v_i \right)^{c^{*}_{\pi_k}-1}
\]

\[
= \frac{\Gamma(c^*)}{\Gamma(c^{*}_{\pi}) \Gamma(c^{*}_{\pi_r}) \Gamma(c^{*}_{\pi_s}) \Gamma(c^{*}_{1-\pi_q - \pi_r - \pi_s})} \left( \begin{array}{ccc}
1 & 1-v_q & 1-v_q - v_r \\
0 & 0 & 0
\end{array} \right)
\]

\[
\times \frac{h(q,r,s)}{(v^q + v^r)(v^q + v^s)} \left( \begin{array}{cccc}
1 & 1-v_q & 1-v_q - v_r & 1-v_q - v_r - v_s \\
1 & 1-v_q & 1-v_q & 1-v_q - v_r - v_s \\
v_q & v_r & v_s & (1-v_q - v_r - v_s)
\end{array} \right)
\]

\[
dv_k \cdots dv_1
\]

(by repeated application of Lemma 4.1 and the fact that \(\sum_{i=1}^{k} \pi_i = 1\)).

a) In this subcase the item in common is preferred in both comparisons so that \(h(q,r,s) = v^2_q\). Now, apply Lemma 4.5 to (4.3-4) with

\[
b_1 = c^{*}_{\pi_q +2}, b_2 = c^{*}_{\pi_r}, b_3 = c^{*}_{\pi_s} \text{ and } b_4 = c^{*}_{1-\pi_q - \pi_r - \pi_s}.
\]

The two conditions necessary for Lemma 4.5 are satisfied since \(c^{*}_{\pi_i} > 0\) \((i = 1,2,\ldots,k)\) by assumption. Therefore, (4.3-4) equals

\[
\frac{\Gamma(c^*) \Gamma(c^{*}_{\pi_q +2})}{\Gamma(c^{*}_{\pi}) \Gamma(c^{*}_{\pi + \pi_r} + \pi_s)} \sum_{n=0}^{\infty} \frac{(c^{*}_{\pi_r + \pi_s} + n)_{\pi_r + \pi_s + 1,n}}{(c^{*}_{\pi_q + \pi_r + \pi_s})_{\pi_q + \pi_r + \pi_s + 1,n} c^{*}_{\pi_q + \pi_r}}.
\]

Then,
\begin{align*}
\mathbb{m}(q,r)(q,s) &= \mathbb{E}\left[ \frac{n(q,r)n(q,s)v_r^2}{(v_q+v_r)(v_q+v_s)} \right] = \frac{n(q,r)n(q,s)\pi^q_q (c^*q+1)}{\pi^{q+\pi_{r+s}}_{r+s}} \\
& \quad \times \sum_{n=0}^{\infty} \frac{(c^*r,r,n)}{(c^*(q+r+s)+1,n)} \frac{1}{c^*(q+r+s)+1+n}.
\end{align*}

b) In this subcase the item in common is preferred in one of the comparisons. Consequently, \( h(q,r,s) = v_q v_r \) or \( v_q v_s \). Without loss of generality assume \( h(q,r,s) = v_q v_r \) and apply Lemma 4.5 with \( b_1 = c^*q+1 \), \( b_2 = c^*r+1 \), \( b_3 = c^*s \) and \( b_4 = c^*(1-\pi_{q+r+s}) \). The lemma can be applied since the assumption that \( c^*_i > 0 \) for \( i = 1, 2, \ldots, k \) guarantees that the two conditions hold. Then, (4.3-4) becomes

\begin{align*}
\frac{\Gamma(c^*)\Gamma(c^*+1)\Gamma(c^*+1)}{\Gamma(c^*_q)\Gamma(c^*_r)\Gamma(c^*_s)} \frac{c^*(c^*+1)}{c^*(q+r+s)} \sum_{n=0}^{\infty} \frac{(c^*q+1,n)}{(c^*(q+r+s)+1,n)} \\
& \quad \times \frac{1}{c^*(q+r+s)+1+n}.
\end{align*}

Therefore,

\begin{align*}
\mathbb{m}(r,q)(q,s) &= \mathbb{E}\left[ \frac{n(q,r)n(q,s)v_r^2}{(v_q+v_r)(v_q+v_s)} \right] \\
& = \frac{n(q,r)n(q,s)c^*q r \pi}{\pi^{q+\pi_{r+s}}_{r+s}} \sum_{n=0}^{\infty} \frac{(c^*q+1,n)}{(c^*(q+r+s)+1,n)} \\
& \quad \times \frac{1}{c^*(q+r+s)+1+n}.
\end{align*}
c) The item in common is not preferred in both comparisons so that

\[ h(q,r,s) = v_r v_s. \]

Because \( c^{\pi_i} > 0 \) (i = 1, 2, ..., k) by assumption, Lemma 4.5 can be applied to (4.3-4), where \( b_1 = c^{\pi_q}, b_2 = c^{\pi_r+1}, \]

\( b_3 = c^{\pi_s+1} \) and \( b_4 = c^{(1-\pi_q - \pi_r - \pi_s)}. \) Then, (4.3-4) reduces to

\[
\frac{\Gamma(c^*) \Gamma(c^{\pi+1}) \Gamma(c^{\pi+2})}{\Gamma(c^{\pi_q}) \Gamma(c^{\pi_s}) \Gamma(c+2)} \sum_{n=0}^{\infty} \frac{(c^{\pi+1, n})}{(c^{\pi_q+\pi_r+\pi_s}+1, n)}
\]

so that

\[
m(r,q)(s,q) = E \left\{ \frac{n(q,r)n(q,s)v_r v_s}{(v_q v_r)(v_s v_t)} \right\}
\]

\[
= \frac{n(q,r)n(q,s)c^{\pi q \pi r \pi s}}{\pi_q^{\pi q} \pi_r^{\pi r} \pi_s^{\pi s}} \sum_{n=0}^{\infty} \frac{(c^{\pi+1, n})}{(c^{\pi_q+\pi_r+\pi_s}+1, n)}
\]

\[
\times \frac{1}{c^{(\pi_q+\pi_r+\pi_s)+1, n}}. \quad \text{QED}
\]

**Corollary 4.1** Let the choice vector \( X \) be distributed as DB(\( \pi_1, ..., \pi_k, c^* \)). Then, the variance-covariance matrix \( \Sigma \) associated with \( X \) has the following elements:
(1) $\sigma(q,r)(q,r) = \frac{n(q,r)^n(q,r)-1\pi q (c^\pi +1)}{(\pi^\pi r)(c^\pi q^\pi q^\pi +1)} + \frac{n(q,r)^q q}{q^q q^q +1} (1 - \frac{n(q,r)^q q}{q^q q^q +1})$, $(q,r = 1,2,\ldots,k; q\neq r)$,

\[
\frac{n(q,r)^q q}{q^q q^q +1} (1 - \frac{n(q,r)^q q}{q^q q^q +1})
\]

(2) $\sigma(q,r)(s,t) = 0$, $(q,r,s,t = 1,2,\ldots,k; q\neq s, q\neq t; r\neq s; r\neq t)$,

(3) a) $\sigma(q,r)(q,s) = n(q,r)^n(q,s) \left\{ \frac{\pi q (c^\pi q^\pi +1)}{\pi^\pi q^\pi q^\pi +1} \right\} \sum_{n=0}^{\infty} \frac{(c^\pi q^\pi q^\pi n)}{(c^\pi q^\pi q^\pi +1,n)}$

\[
\times \frac{1}{c^\pi (q^\pi q^\pi +1+n)} - \frac{q^2}{(q^\pi q^\pi +1,n)}
\]

$(q,r,s = 1,2,\ldots,k; q\neq r; q\neq s; r\neq s)$,

b) $\sigma(r,q)(q,s) = n(q,r)^n(q,s) \left\{ \frac{c^\pi q^\pi r^\pi}{q^\pi q^\pi +1} \right\} \sum_{n=0}^{\infty} \frac{(c^\pi q^\pi q^\pi +1,n)}{(c^\pi q^\pi q^\pi +1,n)}$

\[
\times \frac{1}{c^\pi (q^\pi q^\pi +1+n)} - \frac{q^\pi q^\pi}{(q^\pi q^\pi +1,n)}
\]

$(q,r,s = 1,2,\ldots,k; q\neq r; q\neq s; r\neq s)$,

c) $\sigma(r,q)(s,q) = n(q,r)^n(q,s) \left\{ \frac{c^\pi q^\pi r^\pi}{q^\pi q^\pi +1} \right\} \sum_{n=0}^{\infty} \frac{(c^\pi q^\pi q^\pi +1,n)}{(c^\pi q^\pi q^\pi +1,n)}$
Proof: The results follow easily from Theorems 4.1 and 4.2 using

\[
\text{Var}(X) = E(XX^T) - (EX)(EX)^T.
\]

Theorem 4.1 and the corollary to Theorem 4.2 give the values of

\[
E(X) \text{ and } \text{Var}(X) \text{ when } X \text{ is distributed as } DB(n, \pi_1, \ldots, \pi_k, c^*).
\]

For the Beta-Binomial model of Chapter 3, the \((q,r)\)-th element of \(E(X)\) is

\[
n(q,r) \pi_q (\pi_r + n)^{-1}
\]

(see Equation 3.3-2). This is also the mean for the \((q,r)\)-th choice random variable under the Dirichlet-Binomial model. The variances of the choice random variables for the Beta-Binomial and the Dirichlet-Binomial models are also equal. However, the covariance structure for the models in Chapter 3 differs from that for the Dirichlet-Binomial model. For the models of Chapter 3, the off-diagonal elements are equal to zero, but this is not the case for the Dirichlet-Binomial model. This is because in Chapter 3 different comparisons are assumed independent of one another and for the Dirichlet-Binomial model this assumption is not made. It turns out that for the Dirichlet-Binomial model comparisons with an object in common are correlated with one another, but comparisons with no object in common are uncorrelated. This is reasonable because knowledge of how items
a and b relatively compare to one another provides no information about how items c and d relatively compare to one another, but should provide some information about the relative rankings of items a and c for example.

In the next sections, parameter estimation and hypothesis testing under the Dirichlet-Binomial model will be discussed. First, some notation is defined that is used throughout these sections and the correspondence with the notation in Chapter 2 will be discussed.

For the problem of a paired comparison experiment with k items there are $T = \binom{k}{2}$ groups and $C=2$ subgroups within each group. In the calculations that follow, the complete $2 \binom{k}{2} \times 1$ choice vectors are used, i.e., each choice vector contains the elements $x_{(r,q)}$ and $x_{(q,r)}$ (note that $x_{(r,q)} = n_{(q,r)} - x_{(q,r)}$). The complete choice vector will be denoted by $X_C$ to distinguish it from the choice vector $X$.

Similarly, $n_C$ will be used to denote a $2 \binom{k}{2} \times 1$ vector that contains both $n_{(r,q)}$ and $n_{(q,r)}$ (note that $n_{(q,r)} = n_{(r,q)}$). The vector $n_C$ corresponds to the vector $\tilde{n}$ of Chapter 2. The parameter $\theta$ of the previous chapter corresponds to the parameter $\pi$, where the elements of $\pi$ are the distinct scale values of the Dirichlet-Binomial model, and $\Theta^S$ is now denoted by $\Pi^S = \{ \pi | \pi = (\pi_1, ..., \pi_s), \sum_1^{s} \pi_i = 1, \pi_i > 0, i = 1, ..., s \}$, where $\pi$ is a $s \times 1$ vector with the $i$-th element equal to the number of items associated with $\pi_i$. In general, no subscript is associated with the parameter $\pi$ to indicate its dimension since this will be clear from the context the parameter is used in. The function $I_{-1}(\pi)$ denotes a $2 \binom{k}{2} \times 1$ vector with elements of the form $\pi_q (\pi_q + \pi_r)^{-1}$. 
where \( \pi \in \Pi^s \). In particular, if \( \pi \in \Pi^s \) and the item \( i \) and \( j \) have associated with them the scale values \( \pi_q \) and \( \pi_r \), respectively, then the \((i,j)\)-th element of \( f_{s}^r(\pi) \) is \( \pi_q (\pi_q + \pi_r)^{-1} \). For the case that \( \pi \in \Pi^k \), the \((i,j)\)-th element of \( f_{k}^r(\pi) \) is \( \pi_i (\pi_i + \pi_j)^{-1} \). Note that the function \( f_{s}^r(\pi) \) is continuously differentiable on the parameter space \( \Pi^s \) (\( 1 \leq s \leq k \)).

4.3 Parameter Estimation for the Dirichlet-Binomial Model

A common method of parameter estimation is maximum likelihood estimation. Unfortunately, the Dirichlet-Binomial model does not easily lend itself to this approach. The first step towards maximum likelihood estimation would be to obtain the marginal distribution of \( X \), which has a density function of the form (see Definition 4.1)

\[
\begin{align*}
  m(x; \pi_1, \ldots, \pi_k, c^*) &= \left\{ \frac{\Gamma(c^*)}{\Pi_i \Gamma(c_{pi_i})} \left( \prod_{i} \pi_i \right)^{c^*_i - 1} \right\} \\
  &\times \prod_{q < r} \frac{\binom{n(q,r)}{x(q,r)} x(q,r)}{\binom{n(q,r)}{x(q,r)}} \frac{n(q,r)}{(v_q + v_r)^{n(q,r)}} \, dv.
\end{align*}
\]

Because no closed form formula exists for the marginal density function, maximizing the likelihood function would require the use of numerical integration. Such calculations are not only quite expensive and have to be performed for each observed value of \( X \), but for multiple integration
no reliable software exists. Therefore, maximum likelihood estimation is not feasible for the Dirichlet-Binomial model.

Another common estimation technique is the method of moments. Using this approach, the parameter $\pi_i$ ($i=1,...,k$) can be estimated from the first moment of $X_c$. By Theorem 4.1, the $(q,r)$-th element of $E(X_c)$ is $n(q,r)^{-1} \left( \pi_q + \pi_r \right)^{-1}$ and the method of moments equations are

$$
\bar{x}(q,r) = N^{-1} \sum_{i=1}^{N} x(i) = \frac{n(q,r) \pi_q}{\pi_q + \pi_r}, \quad (q,r = 1,2,...,k),
$$

where $x(i)$ is the number of times judge $i$ prefers item $q$ to item $r$ and $N$ is the number of judges. Note that these equations are exactly the same as the method of moments equation for the scale values for the Beta-Binomial model, as given in Equation (3.4-1). Therefore, the method of moments estimators are as in Equation (3.4-2), i.e.,

$$
\hat{\pi}_q = \left[ m_q^{-1} \sum (i,j) n(i,j) \bar{x}(i,j) \right]^{-1}, \quad (q=1,...,s),
$$

where $m_q$ is the number of items that have $\pi_q$ as their scale value and $\Sigma (i,j)$ denotes the summation over the $m_q(k-1)$ comparisons $(i,j)$ involving an item $i$ which has associated with it the scale value $\pi_q$.

The problem of estimating $c^*$ using the method of moments approach is much more complex. It involves the use of the expected squares
and cross-products of which are given by Theorem 4.2. The second order moments are the easiest to work with to estimate $c^*$. Letting 

$$\frac{-2}{X(q,r)} = N^{-1} \sum_{i=1}^{N} [X(i)_{(q,r)}]^2,$$

the method of moments equation associated with the $(q,r)$-th comparison is

$$\frac{-2}{X(q,r)} = \frac{n_{(q,r)}}{n^q + n^r} \left[ 1 + \frac{(n_{(q,r)}-1)(c^*_q + 1)}{c^*_r + 1} \right].$$

This is the same as the method of moments equation for $c^*$ for the Beta-Binomial model as given in Equation (3.4-3). Therefore, the method of moments estimators for $c^*$ based on $x_{(q,r)}$ is

$$\hat{c}^*(q,r) = (1 - B(q,r))/B(q,r) + (B(q,r)-1)\hat{n}_{q,r},$$

where $\hat{n}_q$ and $\hat{n}_r$ are estimators for $n_q$ and $n_r$, respectively, and $B(q,r) = (X^2_{(q,r)}/\hat{X}_{(q,r)}) - (n_{(q,r)}-1)$ and $\hat{X}_{(q,r)} = n_{(q,r)} \hat{\pi}_{q} \hat{\pi}_{r} - 1$.

This estimator can be calculated only for those comparisons for which $n_{(q,r)} > 1$. If $n_{(q,r)} = 1$, then $c^*$ can only be estimated from the expected cross-product terms with the method of moments. Letting

$$\overline{X}_{(q,r),(s,t)} = N^{-1} \sum_{i=1}^{N} X(i)_{(q,r)} X(i)_{(s,t)},$$

the method of moment equations are:

$$\overline{X}_{(q,r),(q,s)} = \frac{\pi_q (c^*_q + 1)}{n^q + \pi^r + n^s} \sum_{n=0}^{\infty} \frac{(c^*_q + 1, n)}{(\pi_q + \pi^r + \pi^s + 1, n)} \frac{1}{c^*_r + 1 + n},$$

$q, r, s = 1, \ldots, k; q \neq r; q \neq s; r \neq s$.
\[
\tilde{x}_{(r,q)(q,s)} = \sum_{n=0}^{\infty} \frac{(c^* (\pi_q + \pi_r + \pi_s) + 1, n)}{(c^* (\pi_q + \pi_r + \pi_s) + 1, n)} \frac{1}{c^* (\pi_q + \pi_r + \pi_s) + 1 + n}
\]

\[(q, r, s = 1, \ldots, k; q \neq r; q \neq s; r \neq s),\]

and

\[
\tilde{x}_{(r,q)(s,q)} = \sum_{n=0}^{\infty} \frac{(c^* (\pi_q + \pi_r + \pi_s) + 1, n)}{(c^* (\pi_q + \pi_r + \pi_s) + 1, n)} \frac{1}{c^* (\pi_q + \pi_r + \pi_s) + 1 + n}
\]

\[(q, r, s = 1, \ldots, k; r \neq q; s \neq q; r \neq s).\]

Obviously, \(c^*\) cannot be explicitly determined from any one of the equations. However, numerical root finding techniques can be used to improve an initial estimate (see Kennedy and Gentle (1980, p.72)).

Such initial values can be determined by using only the first term of the infinite sums and then solving for \(c^*\).

The initial estimate for \(c^*\) based on the \((q,r)(q,s)\) expected cross-product is obtained from

\[
\tilde{x}_{(q,r)(q,s)} = \frac{\pi_q (c^* (\pi_q + 1))}{(\pi_q + \pi_r + \pi_s)(c^* (\pi_q + \pi_r + \pi_s) + 1)}.
\]

Letting \(B_{(q,r)(q,s)} = \tilde{x}_{(q,r)(q,s)}(\hat{\pi}_q + \hat{\pi}_r + \hat{\pi}_s)(\hat{\pi}_q)^{-1}\), \(c^*\) is initially estimated by

\[
c_{(q,r)(q,s)} = (B_{(q,r)(q,s)}^{-1}) / (B_{(q,r)(q,s)}(\hat{\pi}_q + \hat{\pi}_r) + \hat{\pi}_q).
\]

Similarly, the initial estimate for \(c^*\) based on the \((r,q)(q,s)\) expected cross-product is
\[ c^{\ast}(r,q)(q,s) = \frac{B(r,q)(q,s)}{\left(1 - B(r,q)(q,s)\left(\hat{\pi}^q + \hat{\pi}^r\right)\right)}, \]

where \( B(r,q)(q,s) = \frac{\pi}{\pi^q + \pi^r + \pi^s} \). To obtain an initial estimator for \( c^{\ast} \) based on the cross-product term in which the common item is not preferred in either comparison the first two terms of the infinite sum are required. The method of moments equations for this case are given by

\[
\hat{x}(r,q)(s,q) = \frac{c \pi^r \pi^s}{\pi^q + \pi^r + \pi^s} \left( \frac{1}{c \pi^r} + \frac{c \pi^r + 1}{c \left(\pi^q + \pi^r + \pi^s\right) + 1} \right),
\]

\((q,r,s = 1, \ldots, k; q\neq r; q\neq s; r\neq s)\).

Letting

\[ B^{(1)}(r,q)(s,q) = \left[ -\frac{\hat{\pi}^r \hat{s}}{\pi^q + \pi^r + \pi^s} \right], \]

the initial estimator for \( c^{\ast} \) is obtained by solving

\[
\left(c^{\ast}\right)^2 \left(\frac{\hat{\pi}^q + \hat{\pi}^r}{\pi^q + \pi^r}\right) \left(\hat{\pi}^q + \hat{\pi}^r\right) B^{(1)}(r,q)(s,q) - \hat{\pi}^r + c^{\ast} \left(\frac{\hat{\pi}^q + \hat{\pi}^r}{\pi^q + \pi^r}\right) B^{(1)}(r,q)(s,q) - 1 = 0.
\]

There is an easier method to calculate such an initial estimator by noting that
\[
\tilde{x}_{(r,q)}(s,q) = N^{-1} \sum_{i=1}^{N} x^{(i)}_{(r,q)} x^{(i)}_{(s,q)}
\]

\[
= N^{-1} \sum_{i=1}^{N} x^{(i)}_{(r,q)} (1 - x^{(i)}_{(q,s)})
\]

\[
= \tilde{x}_{(q,s)} - \tilde{x}_{(r,q)}(q,s).
\]

Using the fact that \( \tilde{x}_{(r,q)} \) approximates \( \pi_r (\pi_q + \pi_r)^{-1} \) and using the first term of the \((r,q)(q,s)\) expected cross-product, the method of moments equation associated with the \((r,q)(s,q)\) cross-product is

\[
\tilde{x}_{(r,q)}(s,q) = \frac{\pi_r}{\pi_q + \pi_r} - \frac{c^* \pi_r \pi_q}{(\pi_q + \pi_r + \pi_s)(c^* (\pi_q + \pi_r) + 1)}.
\]

Letting

\[
B^{(2)}_{(r,q)}(s,q) = \left\{ \begin{array}{cc}
\frac{\hat{\pi}_r}{\hat{\pi}_q + \hat{\pi}_r} - \tilde{x}_{(r,q)}(q,s) & \frac{\hat{\pi}_r \hat{\pi}_q}{\hat{\pi}_q + \hat{\pi}_r} \\
\frac{\hat{\pi}_q + \hat{\pi}_s}{\hat{\pi}_q + \hat{\pi}_r}
\end{array} \right.,
\]

\( c^* \) is initially estimated by

\[
\hat{c}^*(r,q)(s,q) = B^{(2)}_{(r,q)}(s,q)/(1 - B^{(2)}_{(r,q)}(s,q)(\hat{\pi}_q + \hat{\pi}_r)).
\]

From each of these initial estimates, \( \hat{c}^*(q,r)(s,t) \), an estimate of \( c^* \) can be calculated. These estimators, along with the estimators in (4.3-1) for \( c^* \) based on the second order moments, can be averaged.
together to obtain an estimator for $c^*$ based on all the data. This average can be an unweighted mean or a weighted mean with weights inversely proportional to $(n(q,r)^n(s,t))^{-1}$.

Another method of parameter estimation is to use the pseudo-MLEs of Chapter 2. This estimation technique can be used to obtain estimates of the scale values $\pi_i$. Because the technique is based on the Bradley-Terry condition and both the Dirichlet-Binomial and Beta-Binomial models satisfy this condition on average, therefore, the pseudo-MLE is the same for both models. For $\pi \in \mathbb{R}^s$, the pseudo-MLE is given by

$$\hat{\pi}(N) = \pi_0 + (A' A)^{-1} A' D^{-1}(\pi_0)(\hat{p}(N) - E_0)$$

where $\pi_0 = s^{-1/2}, 1$ is a sxl vector of ones, $p_0 = f_s(\pi_0), v_0 = D^{-1}(\mu_0)E_0$, $A = D^{-1}(\Sigma_0) \frac{\partial f_s(\pi)}{\partial \pi} \bigg|_{\pi=\pi_0}$, $\hat{p}(N) = \pi N \sum_{i=1} D^{-1}(\mu_C)X^i_C$ and $D(\mu)$ is a diagonal matrix with the elements of $\mu$ down the diagonal. From Lemma 2.3 it follows that $\hat{p}(N)$ is a consistent estimator of $D^{-1}(\mu_C)E(X)$. If $D^{-1}(\mu_C)E(X) = f_s(\pi)$, then it follows that $\hat{\pi}(N)$ is a consistent estimator of $\pi$. This estimator can be improved iteratively by using $\hat{\pi}(N)$ for $\pi_0$ (after $\hat{\pi}(N)$ has been rescaled so that it satisfies $m^T \hat{\pi}(N) = 1$, where $m = (m_1, \ldots, m_s)^T$) and then recalculating $\hat{\pi}(N)$. An estimator for $c^*$ can be obtained by using $\hat{\pi}(N)$ to estimate $\pi$ in the method of moments procedure.
4.4 Hypothesis Testing for the Dirichlet-Binomial Model

A straightforward method of testing the scale values and the judge variability parameter would be to directly test the estimates of these parameters. For example, using MLEs for \((\pi_1, \ldots, \pi_k)\) and \(c^*\), tests of hypotheses could be performed using the likelihood ratio. However, MLEs for the scale values and \(c^*\) are not available for the Dirichlet-Binomial model as was seen in the previous section.

Another way to approach the problem of hypothesis testing for the scale values is to test hypotheses of functions of the scale values using the Wald type test statistics of Chapter 2. Since for the Dirichlet-Binomial model \(E(X|\pi)\) equals the expected mean vector for the Beta-Binomial model, the argument that there exists a sequence \(\mu^{(N)} \to \mu\) and a sequence \(\pi^{(N)}_{r,s} \to \pi^r\) such that for \(\pi_{s} \in \Pi^s, (s < r)\)

\[
D(n_C)\tilde{f}_{s}(\pi_{s}) + \mu^{(N)}/\sqrt{N} = D(n_C)\tilde{f}_{r}(\pi^{(N)}_{r})
\]

for each \(N\), where \(D(s)\) is a diagonal matrix with the elements of \(s\) on the diagonal, is exactly as presented in Section 3.5. Consequently, sequences of alternative hypotheses as postulated in Theorems 2.1 and 2.2 exist so these theorems can be used to test hypotheses about the scale values for the Dirichlet-Binomial model and provide insights about local asymptotic power. Recall that \(\mu_{r,s} = M_{r,s}\mu\) and

\[
\mu'_{r,s} = M_{r,s}^{-1} M_{r,s}', \quad \text{where}
\]
\[ M_{r,s} = D^{-1}(\nu_0)(p_r - p_s)D^{-1}(\nu_0)D^{-1}(\nu_0), P_s = A_s(A_s)'^{-1}A_s', \]

\[ A_s = D^{-1}(\nu_0) \left| \frac{\partial F_s(\pi)}{\partial \pi} \right| \bigg|_{\pi = \pi_0} \]

\[ \nu_0 = D^{-1}(\nu_0)\nu_0, \quad P_0 = A_s(\pi_0) \]

\( \delta_i \) is the variance-covariance matrix of \( X^i \), the elements of which are given by Theorem 4.2. Also, \( \delta_i \) is used to denote the \( i \)-th element of \( E^{-1}\Gamma D^{-1}(p_0)^{-1}r_s \), and \( \delta_{ii} \) the \( i \)-th diagonal element of \( \tilde{E} \), where \( \tilde{E} = \begin{pmatrix} E & 0 \\ 0 & I \end{pmatrix} \). \( E \) is a diagonal matrix with diagonal elements equal to the nonzero eigenvalues of \( D^{-1}(p_0)^{-1}r_s \), \( \Gamma \) is the matrix of orthogonal eigenvectors corresponding to the elements of \( E \), \( I_K \) is the \( K \) dimensional identity matrix and \( -t = \text{rank}(r_s) \). The results of Theorem 2.1 and 2.2 can be summarized as follows.

**Theorem 4.3** Let \( X^{(1)}, \ldots, X^{(N)} \) be i.i.d. DBfn\(^{(N)}\) where the scale values \( \pi^{(N)} \) are such that \( D(\nu_0)p^{(N)} = E(X^{(i)}) = D(\nu_0)f_s(\pi_0) + \mu^{(N)}/\sqrt{N} \) and \( \mu^{(N)} \rightarrow \mu \) as \( N \rightarrow \infty \). Let \( f_s(\hat{\pi}^{(N)}) \) and \( f_r(\hat{\pi}^{(N)}) \) be the pseudo-MLEs for \( \hat{\pi}^{(N)} \) (using the pseudo-MLEs \( \hat{\pi}^{(N)} \) and \( \hat{\pi}^{(N)} \) for \( \pi \)) under \( H_0: p^{(N)} = f_s(\pi_0), \quad \pi_0 \in \Pi_0 \) and \( H_A: p^{(N)} = f_r(\pi), \quad \pi \in \Pi^r \) \((s < r < k)\), respectively. Then

\[ \text{NW}^2(f_s(\hat{\pi}^{(N)}), f_r(\hat{\pi}^{(N)}), r_s) \overset{L}{\rightarrow} \chi^2_{r-s}(\mu' - r_s \mu, r_s \mu' r_s) \]
\[ N^2(\hat{f}_r(\hat{N})), \hat{f}_s(\hat{N})), d \sim NG^2(\hat{f}_r(\hat{N})), \hat{f}_s(\hat{N})), \]

and

\[ N^2(\hat{f}_r(\hat{N})), \hat{f}_s(\hat{N})), \xrightarrow{L} \sum_{i=1}^{r-s} \hat{e}_{ii} X^2(\hat{\delta}_1^2) + \sum_{i=r-s+1}^{2(k)} \hat{\delta}_1^2, \]

where \( W, X^2 \) and \( G^2 \) are as defined in 2.3-1, 2.3-2 and 2.3-3 respectively.

**Proof:** The proof is exactly as the proof of Theorem 3.1. \( \text{QED} \)

The test statistics \( W^2, X^2 \) and \( G^2 \) can also be used to perform a goodness of fit test for the Dirichlet-Binomial model as follows.

**Corollary 4.2** Let \( X^{(1)}, \ldots, X^{(N)} \) be i.i.d. such that \( D(n^c)p^{(N)} = E(X^{(1)}) = D(n^c)f_k(n^c) + j(N)/\sqrt{N} \) and \( \text{Var}(X^{(1)}) = \hat{f}^{(N)} \), where \( j(N) \to j \) and \( \hat{f}(N) \to \hat{f} \), the variance-covariance matrix of the Dirichlet-Binomial model, as \( N \to \infty \). Let \( f_k(\pi) \) be the pseudo-MLE for \( p^{(N)} \) (using the pseudo-MLE \( \hat{\pi}^{(N)} \) for \( \pi \)) under \( H_0: p^{(N)} = f_k(\pi), \pi \in \Pi_k \).

Let the alternative hypothesis be \( H_A: p^{(N)} \) unrestricted. Then,

\[ NW^2(\hat{f}_k(\hat{N})), \hat{p}^{(N)}, \hat{\delta}_{GOF}) \xrightarrow{L} \chi^2_{(k)^{2-k} - (\hat{\mu}_{GOF}^T \hat{\delta}_{GOF} \hat{\mu}_{GOF})}, \]

\[ N^2(\hat{p}, \hat{f}_k(\hat{N})), \hat{p}^{(N)}, \hat{f}_k(\hat{N})), \xrightarrow{d} NG^2(\hat{p}^{(N)}, \hat{f}_k(\hat{N})), \]

and

\[ N^2(\hat{p}, \hat{f}_k(\hat{N})), \xrightarrow{L} \sum_{i=1}^{(k)^{2-k}} \hat{e}_{ii} X^2(\hat{\delta}_1^2) + \sum_{i=(k)^{2-k}+1}^{2(k)} \hat{\delta}_1^2, \]
where $\mathbf{u}_{\text{GOF}} = \mathbf{M} \mathbf{u}$, $\mathbf{t}_{\text{GOF}} = \mathbf{M} \mathbf{t}'$, $\mathbf{M} = D^{-\frac{1}{2}}(\mathbf{e}_0)(\mathbf{I} - \mathbf{F}_k)D^{-\frac{1}{2}}(\mathbf{e}_0)D^{-1}(\mathbf{e}_0)$, $\delta_1$

is the $i$-th element of $E^{-\frac{1}{2}} \Gamma D^{-\frac{1}{2}}(\mathbf{e}_0)\mathbf{u}_{\text{GOF}}$, $E = \begin{pmatrix} E & 0 \\ 0 & I \end{pmatrix}_{2(k-k)}$, $E$ is a diagonal matrix with diagonal elements equal to the nonzero eigenvalues of $D^{-\frac{1}{2}}(\mathbf{e}_0)\mathbf{t}_{\text{GOF}}D^{-\frac{1}{2}}(\mathbf{e}_0)$ and $\Gamma$ is the matrix of orthogonal eigenvectors corresponding to the eigenvalues in $E$.

**Proof:** The proof is exactly as the proof of the corollary to Theorem 3.1. QED

For both Theorem 4.3 and its corollary, the true covariance matrices $\mathbf{t}_{\tau, s}$ and $\mathbf{t}_{\text{GOF}}$ are assumed known. A consistent estimator for these matrices can be used by Lemma 2.4 without altering the results of Theorem 4.3 and its corollary. In Chapter 5, these matrices will be estimated by substituting estimators for $\pi_1, \ldots, \pi_k$ and $c^*$, as given in Section 4.3. Note that these estimators are all consistent, so that the estimates of $\mathbf{t}_{\tau, s}$ and $\mathbf{t}_{\text{GOF}}$, based on these estimators, are themselves consistent.
5. EXAMPLES

5.1 Introduction

In the preceding chapters, several models for pair comparison experiments were introduced and discussed. In each of the following sections, four of these models, the Thurstone-Mosteller, Bradley-Terry, Beta-Binomial and Dirichlet-Binomial models, will be fitted to a data set. The estimated parameter values for these models will be compared as well as the results of tests of hypotheses.

The programs that calculated the parameter estimates and performed the tests of hypotheses for each of these models are given in Appendices A-D. The programs were written in Fortran and make use of IMSL subroutines to evaluate the cumulative normal distribution function, perform various matrix operations and calculate the zeros of a function. In Section 3.4, mention was made of using numerical techniques to maximize the likelihood function of the Beta-Binomial with respect to its parameters. The IMSL subroutine ZXMWD (constrained minimization) was used for this purpose. This approach turned out to be quite prohibitive with respect to the number of computations required and MLEs for the scale values and judge variability parameter are not obtained for the Beta-Binomial model.

5.2 Example 1

The data in this example come from an experiment performed by Hopkins (1954). Six subjects were asked to indicate which of two test solutions tasted sweeter. Pairs of solutions were randomly chosen...
from four different solutions in such a way that each judge compared each possible pair of solutions 20 times. For this experiment, \( N = 6 \) and \( n_{(i,j)} = 20 \), for all pairs of solutions. The data collected are given in Table 1.

Table 1. Data from taste-test experiment III of Hopkins (1954).
(The rows denote preferred treatments)

<table>
<thead>
<tr>
<th></th>
<th>W</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>-</td>
<td>4</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>X</td>
<td>16</td>
<td>-</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>Y</td>
<td>4</td>
<td>3</td>
<td>-</td>
<td>6</td>
</tr>
<tr>
<td>Z</td>
<td>12</td>
<td>8</td>
<td>14</td>
<td>-</td>
</tr>
</tbody>
</table>

Judge B

<table>
<thead>
<tr>
<th></th>
<th>W</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>-</td>
<td>3</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>X</td>
<td>12</td>
<td>-</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>Y</td>
<td>11</td>
<td>12</td>
<td>-</td>
<td>15</td>
</tr>
<tr>
<td>Z</td>
<td>13</td>
<td>14</td>
<td>5</td>
<td>-</td>
</tr>
</tbody>
</table>

Judge G

<table>
<thead>
<tr>
<th></th>
<th>W</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>-</td>
<td>8</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td>X</td>
<td>12</td>
<td>-</td>
<td>6</td>
<td>13</td>
</tr>
<tr>
<td>Y</td>
<td>13</td>
<td>14</td>
<td>-</td>
<td>15</td>
</tr>
<tr>
<td>Z</td>
<td>11</td>
<td>7</td>
<td>5</td>
<td>-</td>
</tr>
</tbody>
</table>

Judge H

<table>
<thead>
<tr>
<th></th>
<th>W</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>-</td>
<td>5</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>X</td>
<td>15</td>
<td>-</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>Y</td>
<td>17</td>
<td>12</td>
<td>-</td>
<td>17</td>
</tr>
<tr>
<td>Z</td>
<td>11</td>
<td>7</td>
<td>3</td>
<td>-</td>
</tr>
</tbody>
</table>

Judge E

Various estimates of the scale values under the saturated model (i.e., no scale values equal) are given in Table 2. Note that for the Bradley-Terry, Beta-Binomial and Dirichlet-Binomial models these values are all
Table 2. Scale value estimates for Example 1

<table>
<thead>
<tr>
<th></th>
<th>$\pi_W$</th>
<th>$\pi_X$</th>
<th>$\pi_Y$</th>
<th>$\pi_Z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM$^a$</td>
<td>MM$^b$</td>
<td>-.0944</td>
<td>.0787</td>
<td>.0418</td>
</tr>
<tr>
<td>BT$^c$</td>
<td>MM</td>
<td>.2138</td>
<td>.2819</td>
<td>.2658</td>
</tr>
<tr>
<td></td>
<td>MLE$^d$</td>
<td>.2138</td>
<td>.2819</td>
<td>.2658</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.0178)</td>
<td>(.0214)</td>
<td>(.0206)</td>
</tr>
<tr>
<td>BB$^e$</td>
<td>MM</td>
<td>.2138</td>
<td>.2819</td>
<td>.2658</td>
</tr>
<tr>
<td></td>
<td>pMLE$^f$</td>
<td>.2138</td>
<td>.2819</td>
<td>.2658</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.6240)</td>
<td>(.7304)</td>
<td>(.7077)</td>
</tr>
<tr>
<td>DB$^g$</td>
<td>MM</td>
<td>.2138</td>
<td>.2819</td>
<td>.2658</td>
</tr>
<tr>
<td></td>
<td>pMLE</td>
<td>.2138</td>
<td>.2819</td>
<td>.2658</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.039)</td>
<td>(1.168)</td>
<td>(1.141)</td>
</tr>
</tbody>
</table>

$^a$Thurstone-Hosteller model (estimates for $S_1$ are given in this row).

$^b$Method of moments estimate.

$^c$Bradley-Terry model.

$^d$Maximum likelihood estimate.

$^e$Beta-Binomial model.

$^f$Pseudo-MLE.

$^g$Dirichlet-Binomial model.

Equivalent. This is expected for the method of moment estimators because the equations from which these estimators are calculated are the same for all three models. Similarly, for the Bradley-Terry model the method of moment estimator and the MLEs are equivalent because the underlying
equations used to calculate these estimators are equivalent. The scale values estimates for the Thurstone-Mosteller model, although different than the other scale value estimates, generate choice probabilities that are equivalent to at least three decimal places to the choice probabilities obtained using the scale value estimates from the other models. Therefore, in terms of ordering the taste-test solutions, the four models are for practical purposes identical.

The differences between the models can be found in the amount of variability taken into account by the models. This difference is reflected in the variance estimates of the scale value estimates. In Table 2, the standard errors of the pseudo-MLEs for the scale values for the Bradley-Terry, Beta-Binomial and Dirichlet-Binomial models are given below the estimates (the MLE and pseudo-MLE for the Bradley-Terry model are equivalent as is seen by the definition of the pseudo-MLE (Definition 2.2) and noting that the Bradley-Terry model postulates a product binomial distribution for the choice vector). The standard errors, for any given scale value, increase from the Bradley-Terry model to the Beta-Binomial model to the Dirichlet-Binomial model. An explanation for the increase from the Bradley-Terry model to the Beta-Binomial model can be found by looking at the covariance matrices of the two models. For both models all the covariances between comparisons are equal to zero. The variance of a comparison for the Beta-Binomial model is equal to its variance under the Bradley-Terry model plus an extra term equal to \( n(i,j)(n(i,j)^{-1})^{1/p(i,j)}(c^{1/p(i,j)} + 1)^{-1}(p(i,j)^{-2} \), which is
positive under the Beta-Binomial model. Therefore, the variance of any comparison, and as such the variance of a pseudo-MLE, is larger for the Beta-Binomial model than for the Bradley-Terry model. The increase in the standard errors of the pseudo-MLEs from the Beta-Binomial model to the Dirichlet-Binomial model can be explained due to the fact that for the Dirichlet-Binomial model, positive covariances are present in the covariance matrix, which is not the case for the Beta-Binomial model, and both models postulate the same variance for a comparison. This positive covariance structure is a result of the estimated values for the scale values and the judge variability parameters used in the calculation of the covariance matrix for the Dirichlet-Binomial model for this example, not some inherent feature of the Dirichlet-Binomial model. A second reason for the increase in the standard errors from the Beta-Binomial model to the Dirichlet-Binomial model is the value of the judge variability parameter estimate used in the calculations. For the Beta-Binomial model \( c^* \) was estimated as 17.273 (based on the 6 different comparison variances), whereas, for the Dirichlet-Binomial model \( c^* \) was estimated as 9.206 (based on the 6 comparison variances and the 12 nonzero comparison covariances). As the value of \( c^* \) increases, the variance of a comparison decreases (see Equation 3.3-4 or Corollary 4.1). Consequently, the variance of the pseudo-MLE of a scale value decreases. Therefore, part of the reason why the standard errors of the scale value estimate are smaller for the Beta-Binomial model than for the Dirichlet-Binomial model is because a larger
estimate of $c^*$ has been used in the calculations for the Beta-Binomial model. If the value of $c^*$ obtained for the Dirichlet-Binomial model had been used in the calculation of the Beta-Binomial models, the resulting standard errors would have been .7591, 8871, .8598 and .8091.

A goodness of fit test was performed for each model. The results are given in Table 3. From the test results it seems that all four models fit the data well with, comparatively, the Thurstone-Mosteller model being the best fit and the Dirichlet-Binomial model being the worst fit. However, this may be misleading. Recall that Bock (1958) showed that for the Thurstone-Mosteller model the goodness of fit test

Table 3. Tests of hypotheses for example 1

<table>
<thead>
<tr>
<th>Test</th>
<th>df</th>
<th>$M_X^a$</th>
<th>$-2\ln b$</th>
<th>$NW_{BB}^c$</th>
<th>$NW_{DB}^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2 vs H1</td>
<td>3</td>
<td>.9240</td>
<td>.9245</td>
<td>.3067($c^* = 17.273$)</td>
<td>1.058($c^* = 9.206$)</td>
</tr>
<tr>
<td>H3 vs H2</td>
<td>3</td>
<td>5.350</td>
<td>5.328</td>
<td>1.230($c^* = 9.435$)</td>
<td>0.5566($c^* = 7.026$)</td>
</tr>
</tbody>
</table>

$^a$ Test statistic proposed by Mosteller (1951b) for the Thurstone-Mosteller model.

$^b$ Likelihood ratio test for the Bradley-Terry model.

$^c$ Test statistic of Theorem 3.1 for the Beta-Binomial model.

$^d$ Test statistic of Theorem 4.3 for the Dirichlet-Binomial model.
statistic showed a better fit (smaller test statistic value) than was actually the case because it ignored any correlation between comparisons. Although Bock's result is only for the Thurstone-Mosteller model, it is interesting to note that the Dirichlet-Binomial model, which includes some correlations between comparisons for a judge, has a larger goodness of fit value than does the Bradley-Terry model. The Beta-Binomial model, which does not include any correlations between comparison but only a judge variability parameter, has a smaller goodness of fit value.

A test of equality of the four scale values was also performed for each model and the results are presented in Table 3. All four models fail to reject the hypothesis that at least one scale value is different. Bock showed that ignoring any correlation among comparisons in such situations results in inflated test statistic values for the Thurstone-Mosteller model. With this in mind and looking at the value of the test statistic for the Thurstone-Mosteller model in Table 3, there is no evidence of differences between the scale values for the Thurstone-Mosteller model. The same conclusion can be reached for the Bradley-Terry model although with a lesser degree of significance than with the Thurstone-Mosteller model. It is interesting to note that the value of the test statistic decreases (i.e., there is less evidence of unequal scale values) from the Bradley-Terry model to the Beta-Binomial model to the Dirichlet-Binomial model, that is, as judge variability and correlations between comparisons are incorporated into the model.
For this data set, the judge variability is rather substantial (note that solution Y was rated highest by four judges and lowest by the other two judges; similarly solutions X and W both receive highest and lowest ratings). This is reflected in the estimates of $c^*$ for the Beta-Binomial and Dirichlet-Binomial model for the various hypotheses (see Table 3). Recall that as $c^* \to \infty$, both these models tend to the Bradley-Terry models which assumes that all judges have the same choice probabilities. Models that do not incorporate some measure of the judge variability can be expected to perform worse in such situations than those that do incorporate such a measure. In this example, though, there are only six judges for the Beta-Binomial and Dirichlet-Binomial models. Consequently, results from these models should be carefully considered since the asymptotic theory for the distribution of the test statistic requires a large number of judges.

5.3 Example 2

In the second example, data from a study conducted by Dr. D. Anderson of the Department of Physical Education at Iowa State University are examined. The object of the study was to determine the relative importance of ten different objectives in physical education (see Table 4).
### Table 4. Objectives of physical education

<table>
<thead>
<tr>
<th></th>
<th>1. Organic vigor</th>
<th>2. Democratic values</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Social competency</td>
<td>Cultural appreciation</td>
</tr>
<tr>
<td>5</td>
<td>Leisure time activities</td>
<td>Self realization</td>
</tr>
<tr>
<td>7</td>
<td>Emotional stability</td>
<td>Neuromuscular skills</td>
</tr>
<tr>
<td>9</td>
<td>Spiritual and moral values</td>
<td>Mental development</td>
</tr>
</tbody>
</table>

A total of 360 of the 381 respondents completed all 45 possible comparisons, performing each comparison once (so that $N=360$ and $n(i,j) = 1$, for all comparisons). The raw data is given in Appendix E, and in Table 5 the various objective preferences are presented for the 360 judges.

### Table 5. Preferences in objectives in physical education. (The rows label preferred treatments)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>255</td>
<td>161</td>
<td>248</td>
<td>203</td>
<td>125</td>
<td>134</td>
<td>147</td>
<td>201</td>
<td>173</td>
</tr>
<tr>
<td>2</td>
<td>105</td>
<td>0</td>
<td>58</td>
<td>134</td>
<td>120</td>
<td>34</td>
<td>43</td>
<td>90</td>
<td>109</td>
<td>71</td>
</tr>
<tr>
<td>3</td>
<td>199</td>
<td>302</td>
<td>0</td>
<td>275</td>
<td>184</td>
<td>120</td>
<td>142</td>
<td>154</td>
<td>236</td>
<td>154</td>
</tr>
<tr>
<td>4</td>
<td>112</td>
<td>226</td>
<td>85</td>
<td>0</td>
<td>131</td>
<td>47</td>
<td>51</td>
<td>92</td>
<td>164</td>
<td>88</td>
</tr>
<tr>
<td>5</td>
<td>157</td>
<td>240</td>
<td>176</td>
<td>229</td>
<td>0</td>
<td>125</td>
<td>142</td>
<td>127</td>
<td>212</td>
<td>158</td>
</tr>
<tr>
<td>6</td>
<td>235</td>
<td>326</td>
<td>240</td>
<td>313</td>
<td>235</td>
<td>0</td>
<td>201</td>
<td>204</td>
<td>270</td>
<td>275</td>
</tr>
<tr>
<td>7</td>
<td>226</td>
<td>317</td>
<td>218</td>
<td>309</td>
<td>218</td>
<td>159</td>
<td>0</td>
<td>196</td>
<td>262</td>
<td>224</td>
</tr>
<tr>
<td>8</td>
<td>213</td>
<td>270</td>
<td>206</td>
<td>268</td>
<td>233</td>
<td>156</td>
<td>164</td>
<td>0</td>
<td>230</td>
<td>195</td>
</tr>
<tr>
<td>9</td>
<td>159</td>
<td>251</td>
<td>124</td>
<td>196</td>
<td>148</td>
<td>90</td>
<td>98</td>
<td>130</td>
<td>0</td>
<td>116</td>
</tr>
<tr>
<td>10</td>
<td>187</td>
<td>289</td>
<td>206</td>
<td>272</td>
<td>202</td>
<td>85</td>
<td>136</td>
<td>165</td>
<td>244</td>
<td>0</td>
</tr>
</tbody>
</table>
Because each comparison is performed only once by a judge, the Beta-Binomial model reduces to the Beta-Bernoulli model and so is indistinguishable from the Bradley-Terry model. In this section, therefore, attention will be focused on the Thurstone-Mosteller, Bradley-Terry and Dirichlet-Binomial models.

Table 6 contains estimated values for the scale values for the three models under four hypothesized parameterizations. Each hypothesis includes the condition \( \pi_{ij} = \pi_i (\pi_i + \pi_j)^{-1} \) for the preference probabilities. The first hypothesis specifies no conditions on the ten scale values for the model, the second specifies nine scale values for the model by taking \( \pi_3 = \pi_{10} \), the third hypothesis specifies eight distinct scale values for the model by taking \( \pi_1 = \pi_5 \) and \( \pi_3 = \pi_{10} \) and the fourth hypothesis specifies seven distinct scale values by taking \( \pi_1 = \pi_5 \) and \( \pi_3 = \pi_8 = \pi_{10} \). Unlike with the previous example, the pseudo-MLEs for the Dirichlet-Binomial model are not all exactly equivalent to the other estimators of the scale values for the Dirichlet-Binomial and Bradley-Terry models. The differences, however, are negligible in practical terms. Similarly, the scale value estimates under the Thurstone-Mosteller model do not generate choice probabilities that coincide as closely as they did in the previous example with the choice probabilities generated by the scale value estimates of the Bradley-Terry or Dirichlet-Binomial model. The differences, however, between the choice probabilities are less than 2%. Therefore, in practical terms, the scale value estimates for the three models generate
### Table 6. Scale value estimates for example 2

<table>
<thead>
<tr>
<th>Method</th>
<th>$\pi_1$</th>
<th>$\pi_2$</th>
<th>$\pi_3$</th>
<th>$\pi_4$</th>
<th>$\pi_5$</th>
<th>$\pi_6$</th>
<th>$\pi_7$</th>
<th>$\pi_8$</th>
<th>$\pi_9$</th>
<th>$\pi_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1&lt;sup&gt;a&lt;/sup&gt; TM&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.02167</td>
<td>-0.6829</td>
<td>0.1207</td>
<td>-0.4891</td>
<td>-0.03774</td>
<td>0.5362</td>
<td>0.3994</td>
<td>0.2304</td>
<td>-0.2241</td>
<td>0.1255</td>
</tr>
<tr>
<td>BT&lt;sup&gt;c&lt;/sup&gt; MM&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.08959</td>
<td>0.02843</td>
<td>0.1034</td>
<td>0.03971</td>
<td>0.08130</td>
<td>0.2031</td>
<td>0.1620</td>
<td>0.1269</td>
<td>0.05971</td>
<td>0.1059</td>
</tr>
<tr>
<td>MLE&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.08959</td>
<td>0.02843</td>
<td>0.1034</td>
<td>0.03971</td>
<td>0.08130</td>
<td>0.2031</td>
<td>0.1620</td>
<td>0.1269</td>
<td>0.05971</td>
<td>0.1059</td>
</tr>
<tr>
<td>(0.00304)(0.00119)(0.00346)(0.00153)(0.00279)(0.00643)(0.00523)(0.00417)(0.00214)(0.00354)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DB&lt;sup&gt;f&lt;/sup&gt; MM&lt;sup&gt;g&lt;/sup&gt;</td>
<td>0.08959</td>
<td>0.02843</td>
<td>0.1034</td>
<td>0.03971</td>
<td>0.08130</td>
<td>0.2031</td>
<td>0.1620</td>
<td>0.1269</td>
<td>0.05971</td>
<td>0.1059</td>
</tr>
<tr>
<td>pMLE&lt;sup&gt;g&lt;/sup&gt;</td>
<td>0.08959</td>
<td>0.02843</td>
<td>0.1034</td>
<td>0.03971</td>
<td>0.08130</td>
<td>0.2031</td>
<td>0.1620</td>
<td>0.1269</td>
<td>0.05971</td>
<td>0.1059</td>
</tr>
<tr>
<td>(0.00561)(0.00243)(0.00628)(0.00306)(0.00521)(0.01060)(0.00892)(0.00737)(0.00412)(0.00639)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

<sup>a</sup>Model hypothesis specifying no restriction on the scale values.

<sup>b</sup>Thurstone-Mosteller model (estimates for $S_1$ are given in these rows).

<sup>c</sup>Bradley-Terry model.

<sup>d</sup>Method of moments estimate.

<sup>e</sup>Maximum likelihood estimate.

<sup>f</sup>Dirichlet-Binomial model.

<sup>g</sup>Pseudo-MLE.
<table>
<thead>
<tr>
<th></th>
<th>$\pi_1$</th>
<th>$\pi_2$</th>
<th>$\pi_3$</th>
<th>$\pi_4$</th>
<th>$\pi_5$</th>
<th>$\pi_6$</th>
<th>$\pi_7$</th>
<th>$\pi_8$</th>
<th>$\pi_9$</th>
<th>$\pi_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H2</strong>&lt;sup&gt;h&lt;/sup&gt;</td>
<td>$0.02167$</td>
<td>$-0.6829$</td>
<td>$0.1231$</td>
<td>$-0.4891$</td>
<td>$-0.03774$</td>
<td>$0.5362$</td>
<td>$0.3994$</td>
<td>$0.2304$</td>
<td>$-0.2241$</td>
<td>$0.1231$</td>
</tr>
<tr>
<td>BT</td>
<td>$0.08960$</td>
<td>$0.02843$</td>
<td>$0.1046$</td>
<td>$0.03972$</td>
<td>$0.08130$</td>
<td>$0.2031$</td>
<td>$0.1620$</td>
<td>$0.1269$</td>
<td>$0.05971$</td>
<td>$0.1046$</td>
</tr>
<tr>
<td>MLE</td>
<td>$0.08960$</td>
<td>$0.02843$</td>
<td>$0.1046$</td>
<td>$0.03972$</td>
<td>$0.08130$</td>
<td>$0.2031$</td>
<td>$0.1620$</td>
<td>$0.1269$</td>
<td>$0.05971$</td>
<td>$0.1046$</td>
</tr>
<tr>
<td></td>
<td>$(0.00304)$</td>
<td>$(0.00119)$</td>
<td>$(0.00249)$</td>
<td>$(0.00153)$</td>
<td>$(0.00279)$</td>
<td>$(0.00647)$</td>
<td>$(0.00523)$</td>
<td>$(0.00418)$</td>
<td>$(0.00214)$</td>
<td>$(0.00249)$</td>
</tr>
<tr>
<td><strong>DB</strong></td>
<td>$0.08960$</td>
<td>$0.02843$</td>
<td>$0.1046$</td>
<td>$0.03972$</td>
<td>$0.08130$</td>
<td>$0.2031$</td>
<td>$0.1620$</td>
<td>$0.1269$</td>
<td>$0.05971$</td>
<td>$0.1046$</td>
</tr>
<tr>
<td>pMLE</td>
<td>$0.08959$</td>
<td>$0.02843$</td>
<td>$0.1046$</td>
<td>$0.03971$</td>
<td>$0.08130$</td>
<td>$0.2031$</td>
<td>$0.1620$</td>
<td>$0.1269$</td>
<td>$0.05971$</td>
<td>$0.1046$</td>
</tr>
<tr>
<td></td>
<td>$(0.00561)$</td>
<td>$(0.00243)$</td>
<td>$(0.00426)$</td>
<td>$(0.00306)$</td>
<td>$(0.00521)$</td>
<td>$(0.01060)$</td>
<td>$(0.00892)$</td>
<td>$(0.00737)$</td>
<td>$(0.00412)$</td>
<td>$(0.00426)$</td>
</tr>
<tr>
<td><strong>H3</strong>&lt;sup&gt;i&lt;/sup&gt;</td>
<td>$0.00803$</td>
<td>$-0.6829$</td>
<td>$0.1231$</td>
<td>$-0.4891$</td>
<td>$-0.00803$</td>
<td>$0.5362$</td>
<td>$0.3994$</td>
<td>$0.2304$</td>
<td>$-0.2241$</td>
<td>$0.1231$</td>
</tr>
<tr>
<td>BT</td>
<td>$0.08537$</td>
<td>$0.02844$</td>
<td>$0.1046$</td>
<td>$0.03973$</td>
<td>$0.08537$</td>
<td>$0.2031$</td>
<td>$0.1620$</td>
<td>$0.1270$</td>
<td>$0.05973$</td>
<td>$0.1046$</td>
</tr>
<tr>
<td>MLE</td>
<td>$0.08537$</td>
<td>$0.02844$</td>
<td>$0.1046$</td>
<td>$0.03973$</td>
<td>$0.08537$</td>
<td>$0.2031$</td>
<td>$0.1620$</td>
<td>$0.1270$</td>
<td>$0.05973$</td>
<td>$0.1046$</td>
</tr>
<tr>
<td></td>
<td>$(0.00214)$</td>
<td>$(0.00119)$</td>
<td>$(0.00249)$</td>
<td>$(0.00153)$</td>
<td>$(0.00214)$</td>
<td>$(0.00651)$</td>
<td>$(0.00526)$</td>
<td>$(0.00418)$</td>
<td>$(0.00214)$</td>
<td>$(0.00249)$</td>
</tr>
<tr>
<td><strong>DB</strong></td>
<td>$0.08537$</td>
<td>$0.02844$</td>
<td>$0.1046$</td>
<td>$0.03973$</td>
<td>$0.08537$</td>
<td>$0.2031$</td>
<td>$0.1620$</td>
<td>$0.1270$</td>
<td>$0.05973$</td>
<td>$0.1046$</td>
</tr>
<tr>
<td>pMLE</td>
<td>$0.08537$</td>
<td>$0.02844$</td>
<td>$0.1046$</td>
<td>$0.03973$</td>
<td>$0.08537$</td>
<td>$0.2031$</td>
<td>$0.1620$</td>
<td>$0.1270$</td>
<td>$0.05973$</td>
<td>$0.1046$</td>
</tr>
<tr>
<td></td>
<td>$(0.00372)$</td>
<td>$(0.00243)$</td>
<td>$(0.00426)$</td>
<td>$(0.00306)$</td>
<td>$(0.00372)$</td>
<td>$(0.01060)$</td>
<td>$(0.00892)$</td>
<td>$(0.00737)$</td>
<td>$(0.00412)$</td>
<td>$(0.00426)$</td>
</tr>
</tbody>
</table>

<sup>h</sup>Model hypothesis specifying restriction $\pi_3 = \pi_{10}$ on the scale values.

<sup>i</sup>Model hypothesis specifying restrictions $\pi_1 = \pi_5$, $\pi_3 = \pi_{10}$ on the scale values.
Table 6. (Continued)

<table>
<thead>
<tr>
<th></th>
<th>$\pi_1$</th>
<th>$\pi_2$</th>
<th>$\pi_3$</th>
<th>$\pi_4$</th>
<th>$\pi_5$</th>
<th>$\pi_6$</th>
<th>$\pi_7$</th>
<th>$\pi_8$</th>
<th>$\pi_9$</th>
<th>$\pi_{10}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_4^J$ TM</td>
<td>-0.0803</td>
<td>-0.6829</td>
<td>0.1589</td>
<td>-0.4891</td>
<td>-0.0803</td>
<td>0.5362</td>
<td>0.3994</td>
<td>0.1589</td>
<td>-0.2241</td>
<td>0.1589</td>
</tr>
<tr>
<td>BT TM</td>
<td>0.08552</td>
<td>0.02851</td>
<td>0.1117</td>
<td>0.03982</td>
<td>0.08552</td>
<td>0.2033</td>
<td>0.1622</td>
<td>0.1117</td>
<td>0.05985</td>
<td>0.1117</td>
</tr>
<tr>
<td>MLE</td>
<td>0.08552</td>
<td>0.02851</td>
<td>0.1117</td>
<td>0.03982</td>
<td>0.08552</td>
<td>0.2033</td>
<td>0.1622</td>
<td>0.1117</td>
<td>0.05985</td>
<td>0.1117</td>
</tr>
<tr>
<td></td>
<td>(0.00215)(0.00119)(0.00210)(0.00154)</td>
<td>(0.00215)(0.00660)(0.00530)(0.00210)(0.00214)(0.00210)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DB MM</td>
<td>0.08552</td>
<td>0.02851</td>
<td>0.1117</td>
<td>0.03982</td>
<td>0.08552</td>
<td>0.2033</td>
<td>0.1622</td>
<td>0.1117</td>
<td>0.05985</td>
<td>0.1117</td>
</tr>
<tr>
<td>pMLE</td>
<td>0.08552</td>
<td>0.02851</td>
<td>0.1117</td>
<td>0.03982</td>
<td>0.08552</td>
<td>0.2033</td>
<td>0.1622</td>
<td>0.1117</td>
<td>0.05985</td>
<td>0.1117</td>
</tr>
<tr>
<td></td>
<td>(0.00373)(0.00244)(0.00340)(0.00307)</td>
<td>(0.00373)(0.01061)(0.00893)(0.00340)(0.00413)(0.00340)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^j\text{Model hypothesis specifying restriction } \pi_1 = \pi_5, \pi_3 = \pi_8 = \pi_{10} \text{ on the scale values.}\)
the same orderings of the physical education objectives.

The standard errors for the MLEs for the Bradley-Terry model and the pseudo-MLEs for the Dirichlet-Binomial model are presented in the table beneath the estimates. As in the previous example, the standard errors for the Dirichlet-Binomial model are all larger than those for the Bradley-Terry model. The variance of a comparison under the Dirichlet-Binomial model is equal to that for the Bradley-Terry model because \( n_{(i,j)} = 1 \) (see Corollary 4.1), so that the increase in the standard errors cannot be attributed to an increase in the variance of the comparisons (which would be due to including the judge variability parameter). Rather, the increase in the standard errors is because the covariances of the covariance matrix of the Dirichlet-Binomial model are all nonnegative. This is due to the particular values of the scale value estimates and judge variability parameter estimate used to calculate the covariance matrix of the Dirichlet-Binomial model.

Table 7 contains the results of goodness of fit tests and tests of hypotheses concerning the scale values. The test statistics reject all three models as very poor fits of the data. This lack of fit for the Bradley-Terry model can be attributed to the fact that the Bradley-Terry condition does not predict the choice probabilities well. For the Dirichlet-Binomial model the lack of fit can be due to two possible reasons. The first, is that, as with the Bradley-Terry model, the Bradley-Terry condition does not adequately model the choice
Table 7. Tests of hypotheses for example 2

H1: unrestricted alternative

H2: $\pi_{(i,k)} = \pi_i (\pi_1 + \pi_3)^{-1}$

H3: $\pi_3 = \pi_5$

H4: $\pi_1 = \pi_5$, $\pi_3 = \pi_10$

H5: $\pi_1 = \pi_5$, $\pi_3 = \pi_8 = \pi_{10}$

<table>
<thead>
<tr>
<th>Test</th>
<th>df</th>
<th>$\chi^2$</th>
<th>$-2\ln\lambda$</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2 vs H1</td>
<td>36</td>
<td>126.1</td>
<td>128.5</td>
<td>304.6 ($\hat{c}^* = 4.277$)</td>
</tr>
<tr>
<td>H3 vs H2</td>
<td>1</td>
<td>.0249</td>
<td>.2411</td>
<td>.0725 ($\hat{c}^* = 4.247$)</td>
</tr>
<tr>
<td>H4 vs H3</td>
<td>1</td>
<td>3.862</td>
<td>3.935</td>
<td>1.119 ($\hat{c}^* = 4.409$)</td>
</tr>
<tr>
<td>H5 vs H4</td>
<td>1</td>
<td>16.63</td>
<td>20.46</td>
<td>10.05 ($\hat{c}^* = 4.217$)</td>
</tr>
</tbody>
</table>

*aTest statistic proposed by Mosteller (1951a) for the Thurstone-Mosteller model.*

*bLikelihood ratio test for the Bradley-Terry model.*

*cTest statistic of Theorem 4.3 for the Dirichlet-Binomial model.*

probabilities. Another possible reason for the lack of fit is that the covariance matrix for the Dirichlet-Binomial model is not the true covariance matrix for the underlying distribution of the choice vectors. The goodness of fit tests proposed in Corollaries 3.1 and 4.2 reply on the assumption that the limiting covariance matrix of the choice vectors equals the covariance matrix as specified by the model. A quick
comparison of the sample covariance matrix with that of the Dirichlet-Binomial model shows that this is not the case for this problem. All the covariances in the sample covariance matrix are nonzero, which is in contradiction to what is predicted by the Dirichlet-Binomial model.

Assuming that all three models fit the data well, the next point of interest would be to test for equality among the scale values in order to reduce the number of parameters in the model. From the test results given in Table 7, it follows that the restriction \( \pi_3 = \pi_{10} \) holds for the Bradley-Terry model. For the Dirichlet-Binomial model, the restriction \( \pi_1 = \pi_5 \) holds in addition to \( \pi_3 = \pi_{10} \). That the Dirichlet-Binomial model is able to reduce the number of scale values further than the Bradley-Terry model is to be expected in view of the larger standard errors for the scale value estimates for the Dirichlet-Binomial model.

The test results for the Thurstone-Mosteller model are very similarly to those for the Bradley-Terry model, as is expected. The best fitting Thurstone-Mosteller model however satisfies the restriction \( \pi_1 = \pi_5 \) and \( \pi_3 = \pi_{10} \) by recalling Bock's result, that ignoring correlation inflates the test statistic values.
6. SUMMARY AND CONCLUSIONS

In Chapters 3 and 4, the Beta-Binomial and Dirichlet-Binomial models for paired comparison experiments were introduced. Methods of parameter estimation and hypothesis testing were presented at the same time. Both models are based on the expected choice probabilities satisfying the Bradley-Terry condition. Because of this, the method of moment estimators for the scale values are equivalent for the Bradley-Terry, Beta-Binomial and Dirichlet-Binomial models, as was seen in the case of the two examples in Chapter 5. The pseudo-MLEs for the scale values are also based on the expected choice probabilities. For the Bradley-Terry model, these estimators are equivalent to the MLEs and method of moment estimators. For the Beta-Binomial and Dirichlet-Binomial models, this is not the case. However, in the examples of the previous chapter, the difference between these estimators and the method of moment estimators was extremely small. An advantage of the pseudo-MLE as compared to the method of moment estimator is that it requires fewer iterations to compute the estimate. Another advantage is that a covariance matrix for the pseudo-MLEs is calculated as the estimate is evaluated. For the method of moment estimators, a covariance matrix is not available.

Recall that Bock (1958) showed that for the Thurstone-Mosteller model ignoring correlations between comparisons results in incorrect test statistics. For tests of goodness of fit, the test statistic will be smaller than it actually is whereas for tests of the scale values the test statistics are larger than is truly the case. With this in
mind, the Beta-Binomial and Dirichlet-Binomial models were developed to include a measure of judge variability in an attempt to more accurately assess the variability of estimated scale values in paired comparison models.

Tables 8 and 9 contain the positive eigenvalues associated with the limiting distribution of the test statistic $N^2$ (see Theorem 2.1) for the Bradley-Terry, Beta-Binomial and Dirichlet-Binomial models (with 4 items) for the goodness of fit test and the test of equality of the scale values. These eigenvalues were calculated for various values of $c$ and various sets of scale values and can be used as a basis for comparing the Bradley-Terry, Beta-Binomial and Dirichlet-Binomial models.

In Table 8, two things are worth noting. First, for the Beta-Binomial model, the eigenvalues are all larger than those for the Bradley-Terry model, whereas the opposite is true for the Dirichlet-Binomial model (this explains the order of the goodness of fit statistic values for these models in Example 1 of Chapter 5). The inclusion of a judge variability term in the model increases the expected value for the goodness of fit test statistic but also including the correlations between comparisons by a judge decreases the expected value for the goodness of fit test statistic. Therefore, if no correlations are present, using the Bradley-Terry model and ignoring any effect due to judge variability will result in the Bradley-Terry model being rejected as a poor fit too often. Ignoring correlations between comparisons, when they are
Table 8. Nonzero eigenvalues for the goodness of fit test using $N \chi^2$

**H1:** unrestricted alternative

**H2:** $p(i,j) = \pi_1 (\pi_i + \pi_j)^{-1}$, $i,j = 1,2,3,4$

Scale values: $\pi_1 = .25$, $\pi_2 = .25$, $\pi_3 = .25$, $\pi_4 = .25$

<table>
<thead>
<tr>
<th>$c^*$</th>
<th>BT$^a$</th>
<th>BB$^b$</th>
<th>DB$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>.8333</td>
<td>.8333</td>
<td>.8333</td>
</tr>
<tr>
<td></td>
<td>.2690</td>
<td>.2690</td>
<td>.2690</td>
</tr>
<tr>
<td>5</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>.6429</td>
<td>.6429</td>
<td>.6429</td>
</tr>
<tr>
<td></td>
<td>.3664</td>
<td>.3664</td>
<td>.3664</td>
</tr>
<tr>
<td>10</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>.5833</td>
<td>.5833</td>
<td>.5833</td>
</tr>
<tr>
<td></td>
<td>.4187</td>
<td>.4187</td>
<td>.4187</td>
</tr>
<tr>
<td>15</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>.5588</td>
<td>.5588</td>
<td>.5588</td>
</tr>
<tr>
<td></td>
<td>.4419</td>
<td>.4419</td>
<td>.4419</td>
</tr>
<tr>
<td>20</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>.5455</td>
<td>.5455</td>
<td>.5455</td>
</tr>
<tr>
<td></td>
<td>.4549</td>
<td>.4549</td>
<td>.4549</td>
</tr>
</tbody>
</table>

Scale Values: $\pi_1 = .2138$, $\pi_2 = .2819$, $\pi_3 = .2658$, $\pi_4 = .2358$

<table>
<thead>
<tr>
<th>$c^*$</th>
<th>BT</th>
<th>BB</th>
<th>DB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>.8396</td>
<td>.8339</td>
<td>.8278</td>
</tr>
<tr>
<td></td>
<td>.2718</td>
<td>.2711</td>
<td>.2698</td>
</tr>
<tr>
<td>5</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>.6488</td>
<td>.6434</td>
<td>.6380</td>
</tr>
<tr>
<td></td>
<td>.3709</td>
<td>.3667</td>
<td>.3622</td>
</tr>
<tr>
<td>10</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>.5874</td>
<td>.5837</td>
<td>.5800</td>
</tr>
<tr>
<td></td>
<td>.4219</td>
<td>.4186</td>
<td>.4151</td>
</tr>
<tr>
<td>15</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>.5619</td>
<td>.5591</td>
<td>.5564</td>
</tr>
<tr>
<td></td>
<td>.4444</td>
<td>.4418</td>
<td>.4391</td>
</tr>
<tr>
<td>20</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>.5479</td>
<td>.5457</td>
<td>.5435</td>
</tr>
<tr>
<td></td>
<td>.4568</td>
<td>.4548</td>
<td>.4526</td>
</tr>
</tbody>
</table>

$^a$Eigenvalues for the Bradley-Terry model.

$^b$Eigenvalues for the Beta-Binomial model.

$^c$Eigenvalues for the Dirichlet-Binomial model.
Table 8. (continued)

Scale values: $\pi_1 = .4 \quad \pi_2 = .3 \quad \pi_3 = .2 \quad \pi_4 = .1$

<table>
<thead>
<tr>
<th>c*</th>
<th>BT</th>
<th>BB</th>
<th>DB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>5</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>10</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>15</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>20</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
</tbody>
</table>

Scale values: $\pi_1 = .3333 \quad \pi_2 = .3333 \quad \pi_3 = .1667 \quad \pi_4 = .1667$

<table>
<thead>
<tr>
<th>c*</th>
<th>BT</th>
<th>BB</th>
<th>DB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>5</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>10</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>15</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>20</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
</tbody>
</table>
Table 8. (continued)

| Scale values: $\pi_1 = .375$ $\pi_2 = .375$ $\pi_3 = .125$ $\pi_4 = .125$ |
|---|---|---|---|---|
| $c^*$ | BT  | BB  | DB  |
| 1  | .5  | .5  | .5  | .8619 | .8333 | .8129 | .3527 | .3063 | .3000 |
| 5  | .5  | .5  | .5  | .6769 | .6429 | .6267 | .3900 | .3774 | .3675 |
| 10 | .5  | .5  | .5  | .6088 | .5833 | .5728 | .4326 | .4234 | .4068 |
| 15 | .5  | .5  | .5  | .5787 | .5588 | .5511 | .4516 | .4445 | .4292 |
| 20 | .5  | .5  | .5  | .5617 | .5455 | .5394 | .4622 | .4566 | .4431 |

| Scale values: $\pi_1 = .6667$ $\pi_2 = .1111$ $\pi_3 = .1111$ $\pi_4 = .1111$ |
|---|---|---|---|---|
| $c^*$ | BT  | BB  | DB  |
| 1  | .5  | .5  | .5  | .9091 | .7992 | .7992 | .4357 | .4357 | .2756 |
| 5  | .5  | .5  | .5  | .7369 | .6212 | .6212 | .4259 | .4259 | .3019 |
| 10 | .5  | .5  | .5  | .6552 | .5707 | .5707 | .4465 | .4465 | .3566 |
| 15 | .5  | .5  | .5  | .6154 | .5501 | .5501 | .4588 | .4588 | .3897 |
| 20 | .5  | .5  | .5  | .5918 | .5389 | .5389 | .4666 | .4666 | .4108 |

actually present (regardless of whether a judge variability has been taken into account), results in just the opposite situation, i.e., the model will seem to fit the data better than it actually does.
The second item worth noting is that as $c^* \to \infty$ the eigenvalues for the Beta-Binomial and Dirichlet-Binomial models decrease, increase, respectively, to the eigenvalues of the Bradley-Terry model. This is expected because as $c^* \to \infty$ both the Beta-Binomial and Dirichlet-Binomial tend to the Bradley-Terry model. Therefore, the bias in the goodness of fit test due to ignoring judge variability and the correlations decreases as judge variability decreases. Looking at Table 8, it would seem that for $c^* > 5$ the differences between the Bradley-Terry, Beta-Binomial and Dirichlet-Binomial models for four items would be negligible and the use of the Beta-Binomial or Dirichlet-Binomial model instead of the Bradley-Terry model would not be warranted in view of the extra computational requirements of the Beta-Binomial and Dirichlet-Binomial models.

In Table 9, in contrast to Table 8, the eigenvalues for the Beta-Binomial model are always larger than those for the Bradley-Terry model. Therefore, ignoring any effect due to judge variability when it is actually present will increase the expected value of the test statistic, and significant differences between the scale values will be more easily found. Similarly, because the eigenvalues for the Dirichlet-Binomial model are always larger than those for the Beta-Binomial model, ignoring any correlations between the comparisons by a judge affects the sensitivity of tests of the scale values so that differences are again more easily found. For $c^* > 5$, the differences between the eigenvalues for the 3 models seem to be small enough so that the value of $N X^2$ for
Table 9. Nonzero eigenvalues for the test of equality of the scale values using $NX^2$

H1: $\pi_1 = \pi_2 = \pi_3 = \pi_4$

H2: $p_{(i,j)} = \pi_i (\pi_i + \pi_j)^{-1}$ for $i,j = 1,2,3,4$

Scale values: $\pi_1 = .25$ $\pi_2 = .25$ $\pi_3 = .25$ $\pi_4 = .25$

<table>
<thead>
<tr>
<th>c</th>
<th>BT$^a$</th>
<th>BB$^b$</th>
<th>DB$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>5</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>10</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>15</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>20</td>
<td>.5</td>
<td>.5</td>
<td>.5</td>
</tr>
</tbody>
</table>

Scale values: $\pi_1 = .2138$ $\pi_2 = .2819$ $\pi_3 = .2658$ $\pi_4 = .2358$

<table>
<thead>
<tr>
<th>c</th>
<th>BT$^a$</th>
<th>BB$^b$</th>
<th>DB$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.4993</td>
<td>.4993</td>
<td>.4984</td>
</tr>
<tr>
<td>5</td>
<td>.4993</td>
<td>.4993</td>
<td>.4984</td>
</tr>
<tr>
<td>10</td>
<td>.4993</td>
<td>.4993</td>
<td>.4984</td>
</tr>
<tr>
<td>15</td>
<td>.4993</td>
<td>.4993</td>
<td>.4984</td>
</tr>
<tr>
<td>20</td>
<td>.4993</td>
<td>.4993</td>
<td>.4984</td>
</tr>
</tbody>
</table>

$^a$Eigenvalues for the Bradley-Terry model.

$^b$Eigenvalues for the Beta-Binomial model.

$^c$Eigenvalues for the Dirichlet-Binomial model.
Table 9. (continued)

<p>| Scale values: $\pi_1 = .4$ $\pi_2 = .3$ $\pi_3 = .2$ $\pi_4 = .1$ |
|-----------------------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>$c^*$</th>
<th>BT</th>
<th>BB</th>
<th>DB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.4908</td>
<td>.4668</td>
<td>.7635</td>
</tr>
<tr>
<td>5</td>
<td>.4908</td>
<td>.4668</td>
<td>.5890</td>
</tr>
<tr>
<td>10</td>
<td>.4908</td>
<td>.4668</td>
<td>.5454</td>
</tr>
<tr>
<td>15</td>
<td>.4908</td>
<td>.4668</td>
<td>.5286</td>
</tr>
<tr>
<td>20</td>
<td>.4908</td>
<td>.4668</td>
<td>.5197</td>
</tr>
</tbody>
</table>

<p>| Scale values: $\pi_1 = .3333$ $\pi_2 = .3333$ $\pi_3 = .1667$ $\pi_4 = .1667$ |
|-----------------------------|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>$c^*$</th>
<th>BT</th>
<th>BB</th>
<th>DB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.4933</td>
<td>.4857</td>
<td>.7893</td>
</tr>
<tr>
<td>5</td>
<td>.4933</td>
<td>.4857</td>
<td>.6071</td>
</tr>
<tr>
<td>10</td>
<td>.4933</td>
<td>.4857</td>
<td>.5576</td>
</tr>
<tr>
<td>15</td>
<td>.4933</td>
<td>.4857</td>
<td>.5381</td>
</tr>
<tr>
<td>20</td>
<td>.4933</td>
<td>.4857</td>
<td>.5277</td>
</tr>
</tbody>
</table>
Table 9. (continued)

| Scale values: $\pi_1 = .375$ $\pi_2 = .375$ $\pi_3 = .125$ $\pi_4 = .125$ |
|-----------------|-----------------|-----------------|-----------------|
| **c** | **BT** | **BB** | **DB** |
| 1 | .4856 | .4856 | .4665 |
|   | .7633 | .7632 | .7331 |
|   | 1.246 | 1.246 | 1.205 |
| 5 | .4856 | .4856 | .4665 |
|   | .5879 | .5879 | .5647 |
|   | .7854 | .7854 | .7575 |
| 10 | .4856 | .4856 | .4665 |
|   | .5428 | .5428 | .5214 |
|   | .6546 | .6546 | .6305 |
| 15 | .4856 | .4856 | .4665 |
|   | .5253 | .5253 | .5046 |
|   | .6031 | .6031 | .5805 |
| 20 | .4856 | .4856 | .4665 |
|   | .5160 | .5160 | .4957 |
|   | .5756 | .5756 | .5539 |

| Scale values: $\pi_1 = .6667$ $\pi_2 = .1111$ $\pi_3 = .1111$ $\pi_4 = .1111$ |
|-----------------|-----------------|-----------------|-----------------|
| **c** | **BT** | **BB** | **DB** |
| 1 | .4895 | .4895 | .4250 |
|   | .6992 | .6992 | .6071 |
|   | 1.069 | 1.069 | .9465 |
| 5 | .4895 | .4895 | .4250 |
|   | .5533 | .5533 | .4804 |
|   | .6730 | .6730 | .5904 |
| 10 | .4895 | .4895 | .4250 |
|   | .5236 | .5236 | .4546 |
|   | .5880 | .5880 | .5138 |
| 15 | .4895 | .4895 | .4250 |
|   | .5128 | .5128 | .4452 |
|   | .5568 | .5568 | .4856 |
| 20 | .4895 | .4895 | .4250 |
|   | .5072 | .5072 | .4403 |
|   | .5406 | .5406 | .4710 |

the 3 models should not differ by very much. Therefore, for $c^* > 5$, the extra precision in the tests of significance is slight and as such the use of Beta-Binomial and Dirichlet-Binomial models cannot be justified in view of the increased computations required by the use of these models. However, this conclusion, as well as that from
Table 8, may change if more items are used. The results of Example 2 suggest that larger \(c^*\) values have a greater impact on the test statistic values when a larger number of items are involved in the experiment.

In conclusion, the Bradley-Terry model is a very useful model for many cases. However, when judge variability is substantial, as indicated by smaller values of \(c^*\), the use of a judge effect model such as the Dirichlet-Binomial or Beta-Binomial becomes justified. When correlations between comparisons are present the Dirichlet-Binomial model can possibly be used and when correlations are not present the Beta-Binomial model is appropriate. A limitation of the Dirichlet-Binomial model is that it specifies a particular type of covariance structure for the comparisons. An open research topic is an investigation into a judge effect model that specifies correlations between all comparisons, not only those with an object in common. One possible approach to the problem would be to use Lancaster and Quade's approach of using \(a_{(q,r)}^{(1)}, \ldots, a_{(q,r)}^{(k)}\) as the parameters of the Dirichlet distribution for the \((q,r)\)-th comparison and reducing the number of parameters in the model by specifying restrictions the model must satisfy.

The number of cases investigated in this dissertation were not sufficient to be able to make very specific recommendations about how much judge variability can be present without seriously affecting the large sample chi-square approximations for the usual Pearson chi-square test based on the Bradley-Terry model. The results presented indicate that the allowable judge variability is dependent on the number of items.
being compared. Another possible effect requiring further investigation is the effect of imbalance in the number of comparisons made for each comparison pair by the judges. Until more is known about the influence of these factors, one way to avoid some pitfalls of using the Bradley-Terry model is to use the Dirichlet-Binomial model and the assorted Wald statistics.
7. BIBLIOGRAPHY


8. ACKNOWLEDGEMENTS

I wish to express my thanks to Dr. K. Koehler for his many years of help with ideas and suggestions for my research. His patience with me and willingness to sit down and help me see the forest from the trees were boundless, and I am deeply grateful for that. I would also like to thank Dr. H. A. David for his constructive criticisms and insightful suggestions concerning this dissertation. I would like to thank Dr. B. C. Carlson for his help in the proof of Theorem 4.1, without whom it would not have been possible. Further, I would like to thank the Department of Statistics for the learning opportunities given to me and the Department of Entomology for the opportunity to practice what I have learned.

College for me has also been a time of enormous personal growth. I wish to thank my friend Terry Callanan who in his quiet and unassuming way was a real source of inspiration for me. I am deeply indebted to Cindy Martin for all her love and caring that have helped me grow as a more loving and caring person. She was always there to share in the laughter and fun, and in the tears and pain. Also, I wish to express my deepest thanks to the community of St. Thomas Aquinas for their witness of God's love for me. Uncountable are the times they have touched me and lifted my spirits with that love expressed through them. In particular, I wish to thank Father John Tilp for his loving concern and the many hours he spent listening to me and helping me sort through the illogic of emotions and feelings.
A great many thanks go to Mrs. Darlene Wicks for her patience and skill in typing this dissertation. Also, for her patience with me and my joking in and around the office.

I am deeply indebted to the love and moral support given me during my stay in California and later in Iowa by the California van Schaiks. And to my housemates, past and present, for their sharing with me of their lives and keeping me sane through the many trials and tribulations.

Finally, but not least, to my parents and brother in the Netherlands. I cannot express my gratitude and thanks for all the love and support you have given me in my lifetime, and especially during my college days when I was so far from home. To you, I dedicate this dissertation.
9. APPENDIX A

Listing of the computer program that calculates scale value estimates and performs tests of hypotheses for the Thurstone-Mosteller model.
INTEGER UNIT, DIM, N(10,10), X(10,10), SO, SA, HO(10), HA(10), DF
REAL P(10,10), D(10,10)
REAL*8 TO(10,10), TA(10,10), CHI2, PI, DARSIN, DFLOAT, DSQRT
LOGICAL GOF

C UNIT - DEVICE NUMBER DATA IS TO BE READ FROM.
C DIM - DIMENSION ARRAYS HAVE BEEN SET UP WITH. THIS EQUALS THE MAXIMUM
C NUMBER OF ITEMS IN AN EXPERIMENT THE PROGRAM CAN HANDLE.
C NI - NUMBER OF ITEMS IN THE EXPERIMENT.
C NJ - NUMBER OF JUDGES IN THE EXPERIMENT.
C N - SYMMETRIC ARRAY GIVING THE NUMBER OF TIMES THE FIRST DIMENSION
C X - CUMULATIVE CHOICE MATRIX CONTAINING THE NUMBER OF TIMES THE FIRST
C DIMENSION ITEM IS PREFERRED TO THE SECOND DIMENSION ITEM OVER ALL
C JUDGES.
C SO, SA - DEGREES OF FREEDOM UNDER HO, HA.
C HO, HA - VECTORS SPECIFYING THE NULL, ALTERNATIVE HYPOTHESES BY
C SPECIFYING WHICH SCALE VALUE IS ASSOCIATED WITH WHICH ITEM.
C EG., A ONE IN THE FIFTH POSITION SPECIFIES THAT ITEM FIVE
C HAS SCALE VALUE ONE ASSOCIATED WITH IT.
C P - MATRIX OF CHOICE PROBABILITIES, THE FIRST DIMENSION PREFERRED TO
C THE SECOND DIMENSION.
C D - MATRIX OF STANDARD NORMAL POINTS ASSOCIATED WITH THE MATRIX P.
C TO, TA - MATRICES OF CHOICE PROBABILITIES CALCULATED BY THE THURSTONE-
C MOSTELLER MODEL UNDER HO, HA.
C GOF - FLAG INDICATING WHETHER OR NOT A GOODNESS-OF-FIT TEST FOR THE
C MODEL IS TO BE PERFORMED.

DATA UNIT/10/, DIM/10/, NI/10/, PI/3.14159265358979D00/,
+ HO/1, 2, 3, 4, 5, 6, 7, 8, 9, 10/, HA/1, 2, 3, 4, 5, 6, 7, 8, 9, 10/
+ GOF/.TRUE./

C READ INPUT DATA
CALL RD(UNIT, DIM, NI, NJ, N, X)
WRITE(6,2000)
2000 FORMAT('INPUT DATA READ IN'/ 'ON MATRIX')
DO 1 I1=1, NI
WRITE(6,2001) (N(I1, I2), I2=1, NI)
2001 FORMAT(IX, 10(I3,2X))
1 CONTINUE
WRITE(6,2002) NJ
2002 FORMAT('NUMBER OF JUDGES: ', I3/ 'ON MATRIX')
DO 2 I1=1, NI
WRITE(6,2001) (X(I1, I2), I2=1, NI)
2 CONTINUE

C CALCULATION OF THE CHOICE PROBABILITIES AND ASSOCIATED STANDARD
C NORMAL POINTS AS SPECIFIED BY THE DATA.
WRITE(6,2003)
SA=0
DO 10 I1=1,NI
   P(I1,I1)=0.5
   D(I1,I1)=0.0
   DO 5 I2=1,I1
      IF (N(I1,I2).EQ.0) GO TO 5
      X1=FLOAT(X(I1,I2))
      IF (X1.EQ.0) X1=0.00001
      XN=FLOAT(NJ)*FLOAT(N(I1,I2))
      P(I1,I2)=X1/XN
      CALL MDNRIS(P(I1,I2),D(I1,I2),IER)
      D(I2,I1)=-D(I1,I2)
      SA=SA+1
  5 CONTINUE
WRITE(6,2004) (P(I1,I2),I2=1,I1)
2004 FORMAT(1X,10(D11.4,2X))
10 CONTINUE
WRITE(6,2005)
2005 FORMAT('OSTANDARD NORMAL POINTS ASSOCIATED WITH THE ABOVE CHOICE', + ' PROBABILITIES(MOSTELLERS X')')
DO 11 I1=1,NI
WRITE(6,2004) (D(I1,I2),I2=1,11)
11 CONTINUE

C CALCULATE THE SCALE VALUES AND CHOICE PROBABILITIES FOR THE THURSTONE-MOSTELLER MODEL UNDER HO.
WRITE(6,2006)
2006 FORMAT(//'=',20('**'),' CALCULATION OF THE SCALE VALUES FOR THE T', + ' HURSTONE-MOSTELLER MODEL UNDER HO ',20('**'))
   CALL TM(DIM,NI,SO,D,H0,T0)
C
IF (GOF) GO TO 15
C
C CALCULATE THE SCALE VALUES AND CHOICE PROBABILITIES FOR THE THURSTONE-MOSTELLER MODEL UNDER HA.
WRITE(6,2007)
   CALL TM(DIM,NI,SA,D,HA,TA)
   GO TO 16
C
C CALCULATE THE CHOICE PROBABILITIES FOR THE ALTERNATIVE GOODNESS OF FIT HYPOTHESIS.
15 WRITE(6,2008)
2008 FORMAT(//'=',20('**'),' THE ALTERNATIVE HYPOTHESIS IS THE UNRESTRICTED', + ' ICPON Hypothesis ',20('**'))
C
C CALCULATE THE TEST STATISTIC PROPOSED BY MOSTELLER(1951B).
DO 25 I2=1,IEND
   IF (GOF) TA(I1,I2)=DBLE(P(I1,I2))
   TO(I1,I2)=DSQRT(TO(I1,I2))
   TO(I1,I2)=DARSIN(TO(I1,I2))
   TO(I1,I2)=TO(I1,I2)*180.0D0/PI
   TA(I1,I2)=DSQRT(TA(I1,I2))
   TA(I1,I2)=DARSIN(TA(I1,I2))
   TA(I1,I2)=TA(I1,I2)*180.0D0/PI
   DN=DFLOAT(N(I1,I2))**DFLOAT(NJ)
   CHI2=CHI2+((TA(I1,I2)-TO(I1,I2))**2)*DN/821.0D00
20   CONTINUE
25   CONTINUE
DN=DFLOAT(N(I1,I2))**DFLOAT(NJ)
CHI2=CHI2+((TA(I1,I2)-TO(I1,I2))**2)*DN/821.0D00
2009  FORMAT(///1X,20('*')/'-FOR THE TEST OF H0 VS. HA THE CHI-SQUARE ',
  +  'STATISTIC HAS A VALUE OF ',D12.6,' WITH ',I2,' DEGREES OF',
  +  'F FREEDOM'/1')
STOP
END

SUBROUTINE RD(UNIT,DIM,NI,NJ,N,X)

INTEGER DIM,N(DIM,DIM),X(DIM,DIM),UNIT,CNT(10,10)
LOGICAL FLAG

FLAG=.TRUE.
DO 5 I1=1,NI
   READ(UNIT,1000) (N(I1,I2),I2=1,NI)
1000   FORMAT(10I2)
5   CONTINUE

DO 15 I1=1,NI
   DO 10 I2=I1,NI
      IF (N(I1,I2).NE.N(I2,I1)) GO TO 40
      X(I1,I2)=0
      X(I2,I1)=0
10   CONTINUE
15   CONTINUE
NJ=0
20 DO 30 I1=1,NI
   READ(UNIT,1001,END=35) (CNT(I1,I2),I2=1,NI)
1001 FORMAT(10I3)
   DO 25 I2=1,NI
      X(I1,I2)=X(I1,I2)+CNT(I1,I2)
25 CONTINUE
30 CONTINUE
C
   NJ=NJ+1
   DO 32 I1=1,NI
      DO 31 I2=I1,NI
         IC=CNT(I1,I2)+CNT(I2,I1)
         IF (IC.EQ.N(J,N,J)) GO TO 31
         FLAG=.FALSE.
         WRITE(6,2001) NJ
      2001 FORMAT('***ERROR*** JUDGE ',13,' HAS AN INCORRECT X MATRIX')
      31 CONTINUE
   32 CONTINUE
   GO TO 20
C
   35 IF (FLAG) RETURN
      STOP
40 FORMAT(6,2002)
2002 FORMAT('***ERROR*** INCORRECT N MATRIX')
   STOP
END
C
SUBROUTINE TM(DIM,NI,S,D,H,T)
C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C CALCULATION OF THE SCALE VALUES AND THE CHOICE PROBABILITIES FOR THE C
C THURSTONE-MOSTELLER MODEL UNDER THE SPECIFIED HYPOTHESIS. C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C
INTEGER DIM,S,M(10,2),H(DIM)
REAL D(DIM,DIM)
REAL*8 T(DIM,DIM),SUM,PI(10),P12(10),DBLE,DFLOAT
C
C DETERMINE THE NUMBER S OF DISTINCT SCALE VALUES HYPOTHESIZED. THE C
C CODES ASSOCIATED WITH THESE SCALE VALUES (H(I)) ARE STORED FOR FUTURE C
C USE INTO M(I,1).
S=1
   M(1,1)=H(1)
   DO 10 I1=2,NI
      IEND=I1-1
      DO 5 I2=1,IEND
         IF (H(I2).EQ.H(I1)) GO TO 10
      5 CONTINUE
   10 CONTINUE
C
S=S+1  
M(S,1)=H(I1)  
10 CONTINUE  
WRITE(6,2000) (H(I),I=1,NI)  
2000 FORMAT('HYPOTHESIS VECTOR',14X,10(2X,I3))  
WRITE(6,2001) S  
2001 FORMAT(' NUMBER OF DISTINCT SCALE VALUES ',I3)  
C  
C DETERMINE THE ESTIMATE FOR EACH DISTINCT SCALE VALUE. THIS IS DONE BY  
C DETERMINING IF ITEM I1 HAS THE SCALE VALUE BEING ESTIMATED ASSOCIATED  
C WITH IT, AND IF IT DOES, ADDING TOGETHER THE D(I1,I2) VALUES FOR ALL  
C ITEMS I2 THAT DO NOT HAVE THAT PARTICULAR SCALE VALUE ASSOCIATED WITH  
C THEM. THIS SUM IS CALCULATED OVER ALL ITEMS I1 THAT HAVE THE SCALE  
C VALUE, THEN DIVIDED BY THE NUMBER OF ITEMS I1 THAT HAVE THAT  
C PARTICULAR SCALE VALUE(IE, BY M(I,2)).  
DO 25 I1=1,S  
M(I,2)=0  
SUM=0.0D0  
DO 20 I1=1,NI  
IF (H(I1).NE.M(I,1)) GO TO 20  
M(I,2)=M(I,2)+1  
DO 15 I2=1,NI  
IF (H(I2).EQ.M(I,1)) GO TO 15  
SUM-SUM+DBLE(D(I1,12))  
15 CONTINUE  
20 CONTINUE  
PI(I)=SUM/DFLOAT(M(I,2))/DFLOAT(NI)  
25 CONTINUE  
C  
C USING THE ESTIMATED SCALE VALUE, DETERMINE THE CHOICE PROBABILITIES.  
DO 35 I1=1,NI  
DO 30 I2=1,NI  
P12(I2)=PI(H(I1))-PI(H(I2))  
CALL MDNORD(P12(I2),T(I1,12))  
30 CONTINUE  
WRITE(6,2006) (P12(I2),I2=1,NI)  
2006 FORMAT(1X,10(D11.4,2X))  
35 CONTINUE  
C
WRITE(6,2007)
2007 FORMAT('OCHOICE PROBABILITIES UNDER THE HYPOTHESES D MODEL(MOSTE',
+ + 'Llers P')')
DO 40 I1=1,NI
   WRITE(6,2006) (T(I1,I2),I2=1,I1)
40 CONTINUE
S=S^1
RETURN
END
10. APPENDIX B

Listing of the computer program that calculates scale value estimates and performs tests of hypotheses for the Bradley-Terry model.
INTEGER N(10,10),X(10,10),SO,SA,H0(10),HA(10),DF,UNIT,DIM
REAL*8 P(10,10),LLBT,LO,LA,LAMBDA,DFLOAT
LOGICAL GOF

C UNIT - Device number data is to be read from.
C DIM - Dimension arrays have been set up with. This equals the maximum
C number of items in an experiment the program can handle.
C NI - Number of items in the experiment.
C NJ - Number of judges in the experiment.
C N - Symmetric array giving the number of times the first dimension
C item is compared to the second dimension item by a judge.
C X - Cumulative choice matrix containing the number of times the first
C dimension item is preferred to the second dimension item over all
C judges.
C SO,SA - Degrees of freedom under H0, HA.
C H0,HA - Vectors specifying the null, alternative hypotheses by
C specifying which scale value is associated with which item.
C EG., a one in the fifth position specifies that item five
C has scale value one associated with it.
C P - Matrix of choice probabilities, the first dimension preferred to
C the second dimension.
C LO,LA - Value of the log likelihood function for the Bradley-Terry
C model under H0, HA.
C GOF - Flag indicating whether or not a goodness-of-fit test for the
C model is to be performed.

DATA UNIT/10/,DIM/10/,NI/10/,HO/1,2,3,4,5,6,7,8,9,10/,
+ HA/1,2,3,4,5,6,7,8,9,10/,GOF/.TRUE./

C READ INPUT DATA.
CALL RD(UNIT,DIM,NI,NJ,N,X)
WRITE(6,2000)
2000 FORMAT('INPUT DATA READ IN'/'ON MATRIX')
DO 1 I1=1,NI
   WRITE(6,2001) (N(I1,I2),I2=1,NI)
2001 FORMAT(1X,10(I3,2X))
1 CONTINUE
WRITE(6,2002) NJ
2002 FORMAT('NUMBER OF JUDGES: ',I3/'ON MATRIX')
DO 2 I1=1,NI
   WRITE(6,2001) (X(I1,I2),I2=1,NI)
2 CONTINUE
C
C CALCULATE THE SCALE VALUES, CHOICE PROBABILITIES AND LOG LIKELIHOOD
C FOR THE BRADLEY-TERRY MODEL UNDER H0.
WRITE(6,2003)
2003 FORMAT(//'=',20('**'),'CALCULATION OF THE SCALE VALUES FOR THE BR',
+ 'ADLEY-TERRY MODEL UNDER H',20('**'))
CALL BT(DIM,NI,SO,NJ,N,X,H0,P)
LO=LLBT(NI,DIM,P,X)

C IF (GOF) GO TO 10
C
C CALCULATE THE SCALE VALUES, CHOICE PROBABILITIES AND LOG LIKELIHOOD
C FOR THE BRADLEY-TERRY MODEL UNDER HA.
WRITE(6,2004)
2004 FORMAT('/=',20('**'),'CALCULATION OF THE SCALE VALUES FOR THE BR',
    + 'ADLEY-TERRY MODEL UNDER HA',20('**'))
CALL BT(DIM,NI,SA,NJ,N,X,HA,P)
LA=LLBT(NI,DIM,P,X)
GO TO 25
C
C CALCULATE THE CHOICE PROBABILITIES AND LOG LIKELIHOOD FOR THE
C ALTERNATIVE GOODNESS-OF-FIT HYPOTHESIS.
10 WRITE(6,2005)
2005 FORMAT('/=',20('**'),'THE CHOICE PROBABILITIES FOR THE UNRESTRICT',
    + 'ED ALTERNATIVE HYPOTHESIS',20('**'))
SA=0
DO 20 I1=1,NI
    P(I1,I1)=0.5D00
    DO 15 12=1,I1
        IF (N(I1,I2).EQ.0) GO TO 15
        P(I1,I2)=DFLOAT(X(I1,I2))/(DFLOAT(NJ)*DFLOAT(N(I1,12)))
        P(I2,I1)=1.0D00-P(I1,12)
        SA=SA+1
    15 CONTINUE
WRITE(6,2006) (P(I1 ,12),12=1 ,I1 )
2006 FORMAT(1X,10(D11.4,2X))
20 CONTINUE
LA=LLBT(NI,DIM,P,X)

C CALCULATE -2(LOG LIKELIHOOD RATIO) AND THE DEGREES OF FREEDOM.
25 LAMBDA=-2.0D00*(LO=-LA)
IF (SA.LE.SO) GO TO 30
DF=SA-SO
WRITE(6,2007) LAMBDA,DF
2007 FORMAT('/=',20('**'),'=-2*LN(LAMBDA)=',D14.6,' WITH ',12,' DEGRE',
    + 'ES OF FREEDOM'/')
STOP
30 WRITE(6,2008) SO,SA
2008 FORMAT('INCORRECT HYPOTHESIS SPECIFICATION'/10X,'HO HAS ',12,
    + 'DEGREES OF FREEDOM'/10X,'HA HAS ',12,'DEGREES OF ',
    + 'FREEDOM')
WRITE(6,2000)
STOP
END
C
SUBROUTINE RD(UNIT,DIM,NI,NJ,N,X)
THE COMPARISON MATRIX N IS INPUTTED AND CHECKED FOR CONSISTENCY. THE CHOICE VECTORS FOR EACH JUDGE ARE READ IN NEXT, CHECKED FOR CONSISTENCY AGAINST THE COMPARISON MATRIX AND ADDED TO THE CHOICE MATRIX X.

```
INTEGER DIM,N(DIM,DIM),X(DIM,DIM),UNIT,CNT(10,10)
LOGICAL FLAG

FLAG=.TRUE.
DO 5 I1=1,NI
   READ(UNIT,1000) (N(I1,I2),I2=1,NI)
1000 FORMAT(10I2)
5 CONTINUE

DO 15 I1=1,NI
   DO 10 I2=I1,NI
      IF (N(I1,I2).NE.N(I2,I1)) GO TO 40
      X(I1,I2)=0
      X(I2,I1)=0
   10 CONTINUE
15 CONTINUE

NJ=0
20 DO 30 I1=1,NI
   READ(UNIT,1001,END=35) (CNT(I1,I2),I2=1,NI)
1001 FORMAT(10I3)
   DO 25 I2=1,NI
      X(I1,I2)=X(I1,I2)+CNT(I1,I2)
25 CONTINUE
30 CONTINUE

NJ=NJ+1
32 DO 31 I1=1,NI
   DO 31 I2=I1,NI
      IC=CNT(I1,I2)+CNT(I2,I1)
      IF (IC.EQ.N(I1,I2)) GO TO 31
      FLAG=.FALSE.
      WRITE(6,2001) NJ
2001 FORMAT('***ERROR*** JUDGE ',I3,' HAS AN INCORRECT X MATRIX')
31 CONTINUE
32 CONTINUE
   GO TO 20

35 IF (FLAG) RETURN
STOP
40 WRITE(6,2002)
```
SUBROUTINE BT(DIM, NI, S, NJ, N, X, H, P)

INTEGER S, DIM, M(10, 2), H(DIM), N(DIM, DIM), X(DIM, DIM)
REAL*8 P1(10), P2(10), SCALE, A, SUM, DEV, MAX, P(DIM, DIM),
+ DFLOAT, DMAX1, DABS
LOGICAL FLAG

FLAG=.FALSE.

C DETERMINE THE NUMBER S OF DISTINCT SCALE VALUES HYPOTHESIZED. THE
C CODES ASSOCIATED WITH THESE SCALE VALUES (H(I)) ARE STORED FOR FUTURE
C USE INTO M(I,1).
M(1,1)=H(1)
S=1
DO 2 II=2, NI
  IEND=II-1
  DO 1 I2=1, IEND
    IF (H(I2).EQ.H(I1)) GO TO 2
  1 CONTINUE
  S=S+1
  M(S,1)=H(I1)
 2 CONTINUE
WRITE(6,2000) (H(I), I=1, NI)
2000 FORMAT('HYPOTHESIS VECTOR', 14X, 10(2X, 13))
WRITE(6,2001) S
2001 FORMAT('NUMBER OF DISTINCT SCALE VALUES ', I3)

C INITIALIZE THE ITERATION COUNTER(IT) AND THE VECTOR OF SCALE VALUES
C (P1). FOR THE CASE OF 1 SCALE VALUE (IE, ALL ITEMS EQUAL), SKIP OVER
C THE ITERATIVE CALCULATIONS OF THE MLE'S FOR THE SCALE VALUES.
IT=0
DO 4 I=1, S
  P1(I)=1.0D00/DFLOAT(NI)
4 CONTINUE
M(1,2)=NI
IF (S.EQ.1) GO TO 21

C ITERATE CalculATe THE MLE'S FOR THE SCALE VALUES. THIS IS DONE
C BY DETerminING If ITEM I1 HAS The SCALE VALUE BEING ESTIMATED
C ASSOCIATED WITH IT, AND IF SO, CUMULATING X(I1,I2) AND C N(I1,I2)/(PI(I1)+PI(I2)) FOR THOSE ITEMS I2 THAT DO NOT HAVE THAT C PARTICULAR SCALE VALUE ASSOCIATED WITH THEM. THESE SUMS ARE THEN C USED TO CALCULATE THE NEXT ITERATION OF MLES.

5 SCALE=0.0D00
   DO 20 I=1,S
   M(I,2)=0
   SUM=0.0D00
   A=0.0D00
   DO 15 I1=1,NI
      IF (H(I1).NE.M(I,1)) GO TO 15
      M(I,2)=M(I,2)+1
      DO 10 12=1,NI
         IF (H(I2).EQ.M(I,1)) GO TO 10
         A=A+DFLOAT(X(I1,12))
         SUM=SUM+DFLOAT(N(I1,12))/(PI(H(I1))+PI(H(I2)))
      10 CONTINUE
   15 CONTINUE
   P2(I)=A/(DFLOAT(NJ)*SUM)
   SCALE=SCALE+DFLOAT(M(I,2))*P2(I)
   20 CONTINUE

C OUTPUT INFORMATION ABOUT THE HYPOTHESIS AND THE INITIAL SCALE VALUE C PARAMETER ESTIMATES.
21 IF (FLAG) GO TO 24
   WRITE(6,2002) (M(I,1),I=1,S)
   WRITE(6,2003) (M(I,2),I=1,S)
   WRITE(6,2004) (PI(I),I=1,S)
   WRITE(6,2005) (P1(I),I=1,S)
2002 FORMAT('DISTINCT SCALE VALUE CODES',5X,10(2X,13))
2003 FORMAT('NUMBER OF ASSOCIATED ITEMS',5X,10(2X,13))
2004 FORMAT('CURATIVE CALCULATION OF SCALE VALUE ESTIMATES'/' INT', + 10(1X,D10.4))
   IF (S.EQ.1) GO TO 29
   FLAG=.TRUE.

C CALCULATE THE MAXIMUM ABSOLUTE DEVIATION BETWEEN THE ESTIMATES OF C THIS ITERATION AND THOSE OF THE PREVIOUS ITERATION. IF THE C DEVIATION IS SMALL ENOUGH OR THE NUMBER OF ITERATIONS GETS TOO LARGE, C STOP ITERATING.
24 MAX=0.0D00
   DO 25 I=1,S
      P2(I)=P2(I)/SCALE
      DEV=P2(I)-P1(I)
      DEV=DABS(DEV)
      MAX=DMAX1(DEV,M)
      P1(I)=P2(I)
   25 CONTINUE
   IT=IT+1
   IREM=MOD(IT,5)
IF (IREM.EQ.O) WRITE(6,2005) IT,(P1(I),I=1,S),MAX
2005 FORMAT(1X,I3,11(1X,D10.4))
IF ((IT.LT.500).AND.(MAX.GT.1.0D-05)) GO TO 5
IF (IREM.NE.O) WRITE(6,2005) IT,(P1(I),I=1,S),MAX

C C CALCULATE THE VARIANCE-COVARIANCE MATRIX ASSOCIATED WITH THE SCALE
C VALUE ESTIMATES.
CALL VCBT(S,DIM,NI,NJ,H,M,N,P1)
29 WRITE(6,2006)
2006 FORMAT('OCHOICE PROBABILITIES CALCULATED USING THE SCALE VALUES',
    + ESTIMATED UNDER THE HYPOTHESIZED BRADLEY-TERRY MODEL')

C C CALCULATION OF THE CHOICE PROBABILITIES USING THE SCALE VALUES
C ESTIMATED ABOVE.
DO 35 I1=1,NI
    DO 30 I2=1,I1
        P(I1,I2)=P1(H(I1))/(P1(H(I1))+P1(H(I2)))
        P(I2,I1)=P1(H(I2))/(P1(H(I1))+P1(H(I2)))
30    CONTINUE
WRITE(6,2007) (P(I1,I2),I2=1,I1)
2007 FORMAT(1X,10(D11.4,2X))
35 CONTINUE
S=S-1
RETURN
END

C SUBROUTINE VCBT(S,DIM,NI,NJ,H,M,N,PI)

C CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C CALCULATION OF FISHER'S INFORMATION MATRIX AND ITS INVERSE FOR THE C
C HYPOTHESIZED BRADLEY-TERRY MODEL.
C CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C
INTEGER S,DIM,H(DIM),M(DIM,2),N(DIM,DIM)
REAL*8 XC,XN,XD,L(10,10),PI(DIM),LI(10,10),WKAREA(150),DFLOAT

C C CALCULATE THE NEGATIVE OF FISHER'S INFORMATION MATRIX FOR THE BRADLEY-
C TERRY MODEL UNDER THE SPECIFIED HYPOTHESIS.
IS=S-1
DO 50 I1=1,IS
    DO 50 I2=1,I1
        L(I1,I2)=0.0D00
    50 CONTINUE
DO 45 J1=1,NI
    IF (H(J1).NE.M(I1,1)) GO TO 25
    DO 20 J2=1,NI
        IF (H(J2).NE.M(I2,1)) GO TO 15
        IF (J1.EQ.J2) GO TO 5
        XN=DFLOAT(N(J1,J2))
        XD=(PI(H(J1))+PI(H(J2)))**2
L(I1,I2) = L(I1,I2) - \( \frac{XN}{XD} \)

GO TO 20

DO 10 K = 1, NI
    IF (H(K) .EQ. M(I1, 1)) GO TO 10
    XN = DFL0AT(N(J1, K)) * PI(H(K))
    XD = ((PI(H(J1))) + PI(H(K)))**2 * PI(H(J1))
    L(I1,I2) = L(I1,I2) + XN/XD
  10 CONTINUE
  GO TO 20

IF (H(J2).NE.M(S, 1)) GO TO 20
    XC = DFL0AT(M(I2, 2))/DFL0AT(M(S, 2))
    XN = DFL0AT(N(J1, J2))
    XD = ((PI(H(J1))) + PI(H(J2)))**2
    L(I1,I2) = L(I1,I2) + XC*XN/XD
  20 CONTINUE
  GO TO 45

IF (H(J1).NE.M(S, 1)) GO TO 45
    DO 40 J2 = 1, NI
        IF (H(J2).NE.M(I2, 1)) GO TO 30
        XC = DFL0AT(M(I1, 2))/DFL0AT(M(S, 2))
        XN = DFL0AT(N(J1, J2))
        XD = ((PI(H(J2))) + PI(H(J1)))**2
        L(I1,I2) = L(I1,I2) + XC*XN/XD
  30 CONTINUE
  40 CONTINUE

C
C CALCULATE THE INVERSE OF THE ABOVE COMPUTED MATRIX.
  IDGT = 0
  CALL LINV2F(L, IS, DIM, LI, IDGT, WKAREA, IER)
C
C CALCULATE THE VARIANCE AND COVARIANCES ASSOCIATED WITH THE LAST SCALE
C VALUE (IT DEPENDS ON THE PREVIOUSLY COMPUTED SCALE VALUES BECAUSE THE
C SUM OF THE SCALE VALUES IS ONE). THEN OUTPUT THE VARIANCE-COVARIANCE
C MATRIX.
  55 LI(S,S) = 0.0D00
    DO 65 I1 = 1, IS
        LI(I1,S) = 0.0D00
        DO 60 I2 = 1, IS
LI(I1,I2)=LI(I1,I2)/DFLOAT(NJ)
LI(I1,S)=LI(I1,S)*DFLOAT(M(I2,2))/DFLOAT(M(S,2))*LI(I1,I2)
60 CONTINUE
LI(S,I1)=LI(I1,S)
LI(S,S)=LI(S,S)*DFLOAT(M(I1,2))/DFLOAT(M(S,2))*LI(I1,S)
65 CONTINUE
WRITE(6,2002)
2002 FORMAT('COVARIANCE-COVARIANCE MATRIX FOR THE SCALE VALUES')
DO 75 I1=1,S
WRITE(6,2003) (LI(I1,I2),I2=1,I1)
2003 FORMAT(1X,10(D14.7,2X))
75 CONTINUE
RETURN
END

REAL FUNCTION LLBT(NI,DIM,P,X)

C
C CALCULATION OF THE LOG LIKELIHOOD FOR THE BRADLEY-TERRY MODEL FOR
C THE GIVEN MATRIX OF CHOICE PROBABILITIES.
C
INTEGER DIM,X(DIM,DIM)
REAL*8 P(DIM,DIM),X1,X2,DFLOAT,DLOG

LLBT=0.0DOO
DO 10 I1=1,NI
DO 5 I2=1,I1
X1=DFLOAT(X(I1,I2))
X2=DFLOAT(X(I2,I1))
LLBT=LLBT+X1*DLOG(P(I1,I2))+X2*DLOG(P(I2,I1))
5 CONTINUE
10 CONTINUE
WRITE(6,2001) LLBT
2001 FORMAT('THE VALUE OF THE LOG-LIKELIHOOD FUNCTION FOR THE BRADLEY-
+ TERRY MODEL UNDER THE HYPOTHESIS IS ',D14.7)
RETURN
END
11. APPENDIX C

Listing of the computer program that calculates estimates for the scale values and judge variability parameter, and performs tests of hypotheses for the Beta-Binomial model.
INTEGER UNIT,DIM,DC2,S,DFO,DFA,DF,H0(10),HA(10),M(10,2),
+ N(10,10),X1(10,10),X2(10,10)
REAL*8 CO,CA,PI(10),W,SING(90),WK(180),FO(90),FA(90),VO(90),
+ PO(90,90),PA(90,90),SO(90,90),MWK(90,90),DGT
LOGICAL CESTO,CESTA,GOF
C
C UNIT = DEVICE NUMBER DATA IS TO BE READ FROM.
C DIM = DIMENSION ARRAYS HAVE BEEN SET UP WITH. THIS EQUALS THE MAXIMUM
C NUMBER OF ITEMS IN AN EXPERIMENT THE PROGRAM CAN HANDLE.
C DC2 = DIM*(DIM-1)
C NI = NUMBER OF ITEMS IN THE EXPERIMENT.
C NIC2 = NI*(NI-1)
C NJ = NUMBER OF JUDGES IN THE EXPERIMENT.
C N = SYMMETRIC ARRAY GIVING THE NUMBER OF TIMES THE FIRST DIMENSION
C ITEM IS COMPARED TO THE SECOND DIMENSION ITEM BY A JUDGE.
C X1 = CUMULATIVE CHOICE MATRIX CONTAINING THE NUMBER OF TIMES THE
C FIRST DIMENSION ITEM IS PREFERRED TO THE SECOND DIMENSION ITEM
C OVER ALL JUDGES.
C X2 = CUMULATIVE(OVER ALL JUDGES) MATRIX OF THE SQUARE OF THE ELEMENTS
C OF EACH JUDGE'S CHOICE MATRIX.
C DFO,DFA = DEGREES OF FREEDOM UNDER HO, HA.
C HO,HA = VECTORS SPECIFYING THE NULL, ALTERNATIVE HYPOTHESES BY
C SPECIFYING WHICH SCALE VALUE IS ASSOCIATED WITH WHICH ITEM.
C EG., A ONE IN THE FIFTH POSITION SPECIFIES THAT ITEM FIVE
C HAS SCALE VALUE ONE ASSOCIATED WITH IT.
C M(1,1): IDENTIFIES THE SCALE VALUE CODE ASSOCIATED WITH SCALE
C VALUE I.
C M(1,2): THE NUMBER OF ITEMS WITH SCALE VALUE I.
C S = NUMBER OF DISTINCT SCALE VALUES.
C PI = VECTOR OF SCALE VALUES.
C CO,CA = JUDGE VARIABILITY PARAMETER UNDER HO, HA.
C FO,FA = VECTOR OF CHOICE PROBABILITIES PREDICTED BY THE BETA-BINOMIAL
C MODEL UNDER HO, HA.
C PO,PA = PROJECTION MATRICES FOR THE CHOICE PROBABILITY VECTOR UNDER
C HO, HA.
C VO = VECTOR OF CHOICE PROBABILITIES FOR THE BETA-BINOMIAL MODEL UNDER
C HO, DIVIDED BY THE NUMBER OF TIMES A COMPARISON IS MADE.
C SO = VARIANCE-COVARIANCE MATRIX FOR THE COMPARISONS UNDER THE BETA-
C BINOMIAL MODEL UNDER HO.
C CESTO,CESTA = FLAG DETERMINING WHETHER OR NOT THE JUDGE VARIABILITY
C PARAMETER SHOULD BE ESTIMATED UNDER HO, HA.
C GOF = FLAG INDICATING WHETHER OR NOT A GOODNESS-OF-FIT TEST FOR THE
C MODEL IS TO BE PERFORMED.
C
DATA UNIT/10/,DIM/10/,NI/10/,CESTO/.TRUE./,CESTA/.TRUE./,
+ CO/1.0D00/,CA/1.0D00/,HO/1,2,3,4,5,6,7,8,9,10/,
+ HA/1,2,3,4,5,6,7,8,9,10/,GOF/.TRUE./
C
C IGNORE OVERFLOW AND UNDERFLOW ERROR LIMITS AND SUPPRESS WARNING
C MESSAGES.
CALL ERRSET(207,256,-1,0)
CALL ERRSET(208,0,-1,0)
NIC2=NI*(NI-1)
DC2=DIM*(DIM-1)

C READ INPUT DATA.
CALL RD(UNIT,DIM,NI,NJ,N,X1,X2)

C CALCULATE VARIOUS PARAMETER ESTIMATES FOR THE BETA-BINOMIAL MODEL
C UNDER HO.
WRITE(6,2000)
2000 FORMAT(//1X,132('*')/'OCALCULATION OF ESTIMATORS FOR THE JUDGE V',
+ 'ARIABILITY PARAMETER AND THE SCALE VALUE PARAMETERS FOR ',
+ 'THE BETA-BINOMIAL MODEL UNDER HO')

C SETUP THE M MATRIX FOR HO.
CALL SET(DIM,NI,S,HO,M,CO,CEST0)
DF0'=S'='1

C CALCULATE THE METHOD OF MOMENT ESTIMATORS FOR THE SCALE VALUES AND
C THE JUDGE VARIABILITY PARAMETER FOR THE BETA-BINOMIAL MODEL UNDER HO.
WRITE(6,2001)
2001 FORMAT(/'^',20('*'),' METHOD OF MOMENT ESTIMATION OF THE PARAMET',
+ 'ERS ',20('*'))
CALL MM(DIM,NI,NJ,S,N,X1,X2,H0,M,CEST0,PI,C0)

C INITIALIZE SCALE VALUE ESTIMATES AND CALCULATE PSEUDO-MLES FOR THE
C SCALE VALUES FOR THE BETA-BINOMIAL MODEL UNDER HO. ALSO CALCULATE
C AND OUTPUT VO AND SO.
WRITE(6,2002)
2002 FORMAT(/'^',20('*'),' PSEUDO-MAXIMUM LIKELIHOOD ESTIMATION OF TH',
+ 'E PARAMETERS ',20('*'))
DO 11=1,S
PI(I)=1.0D0/DFLOAT(NI)
11 CONTINUE
CALL PMLE(DIM,DC2,NI,NJ,S,H0,M,N,X1,PI,CO,FO,PO,VO,SO)

C IF (GOF) GO TO 5

C CALCULATE VARIOUS PARAMETER ESTIMATES FOR THE BETA-BINOMIAL MODEL
C UNDER HA.
WRITE(6,2003)
2003 FORMAT(//1X,132('*')/'OCALCULATION OF ESTIMATORS FOR THE JUDGE V',
+ 'ARIABILITY PARAMETER AND THE SCALE VALUE PARAMETERS FOR ',
+ 'THE BETA-BINOMIAL MODEL UNDER HA')

C SETUP THE M MATRIX FOR HA.
CALL SET(DIM,NI,S,HA,M,CA,CESTA)
DFA = S^1

C CALCULATE THE METHOD OF MOMENT ESTIMATORS FOR THE SCALE VALUES AND
C THE JUDGE VARIABILITY PARAMETER FOR THE BETA-BINOMIAL MODEL UNDER HA.
WRITE(6,2001)
   CALL MM(DIM,NI,NJ,S,N,X1,X2,HA,CESTA,PI,CA)
C
C INITIALIZE SCALE VALUE ESTIMATES AND CALCULATE PSEUDO-MLEs FOR THE
C SCALE VALUES FOR THE BETA-BINOMIAL MODEL UNDER HA.
WRITE(6,2002)
   DO 2 I=1,S
      PI(I)=1.0D00/DFLOAT(NI)
   2 CONTINUE
   CALL PMLE(DIM,DC2,NI,NJ,S,HA,M,N,X1,PI,CA,FA,PA,SING,MWK)
   GO TO 20
C
C CALCULATE FA,PA FOR THE ALTERNATIVE GOODNESS-OF-FIT HYPOTHESIS.
5 WRITE(6,2004)
   2004 FORMAT(/'ALTERNATIVE HYPOTHESIS IS THE UNRESTRICTED HYPOTHESIS ',
            + 'FOR A GOODNESS OF FIT TEST')
   I=1
   DO 15 J2=1,NI
       DO 15 J1=1,NI
          IF (J1.EQ.J2) GO TO 15
          FA(I)=DFLOAT(X1(J1,J2))/(DFLOAT(N(J1,J2))*DFLOAT(NJ))
          DO 10 I1=1,NIC2
             PA(I1,I)=0.0D00
             10 CONTINUE
             PA(I1,I)=1.0D00
             I=I+1
       15 CONTINUE
   DFA = NIC2/2
C
C CALCULATE THE DIFFERENCE VECTOR FA-FO AND THE DIFFERENCE PROJECTION
C MATRIX PA-PO. THIS MATRIX IS THEN PRE-MULTIPLIED BY A DIAGONAL MATRIX
C WITH THE SQUARE ROOT OF THE ELEMENTS OF VO DOWN THE DIAGONAL, POST-
C MULTIPLIED BY THE INVERSE OF A DIAGONAL MATRIX WITH THE SQUARE ROOT
C OF THE ELEMENTS OF VO DOWN THE DIAGONAL AND POST-MULTIPLIED AGAIN BY
C A DIAGONAL MATRIX WITH THE ELEMENTS OF THE COMPARISON MATRIX N DOWN
C THE DIAGONAL.
20 DF=DFA-DFO
   DO 30 I1=1,NIC2
      FA(I1)=FA(I1)-FO(I1)
   DO 25 I2=1,NIC2
      J2=(I2-1)/(NI-1)+1
      J1=I2-(J2-1)*(NI-1)
      IF (J1.GE.J2) J1=J1+1
      PA(I1,I2)=PA(I1,I2)-PO(I1,I2)
      PA(I1,I2)=DSQRT(VO(I1))*PA(I1,I2)
PA(I1,I2) = PA(I1,I2) / (DSQRT(VO(I2)) * DFLOAT(N(J1,J2)))

25 CONTINUE
30 CONTINUE

C PRE- AND POST-MULTIPLY SO BY PA-PO CALCULATED ABOVE.
CALL VMULFF(PA, SO, NIC2, NIC2, NIC2, DC2, DC2, PO, DC2, IER)
CALL VMULFP(PO, PA, NIC2, NIC2, NIC2, DC2, DC2, SO, DC2, IER)

C PERFORM AN ACCURACY CHECK FOR THE CALCULATION OF THE GENERALIZED
INVERSE, TO SEE THAT IGNORING UNDERFLOW AND OVERFLOW CONDITIONS HAS
NOT UPSET CALCULATIONS GREATLY. THIS IS DONE BY CALCULATING THE
GENERALIZED INVERSE( USING SINGULAR VALUE DECOMPOSITION) AND THEN
PRE- AND POST-MULTIPLYING THE GENERALIZED INVERSE BY THE ORIGINAL
MATRIX.
WRITE(6,5000)

5000 FORMAT(/,1X,'ACCURACY CHECK FOR THE GENERALIZED INVERSE',
               'RSE CALCULATION'//1X,'MSM MATRIX(PARTIAL)')
DO 11 I1=1,NIC2
   IF (I1.LE.12) WRITE (6,5001) (30(I1,12),12=1,6)
5001 FORMAT(1X,6(D11.4,3X))

C SAVE SO FOR FUTURE USE AND INITIALIZE PO.
DO 35 I2=1,NIC2
   PA(I1,I2) = SO(I1,I2)
   PO(I1,I2) = 0.0D0
35 CONTINUE
PO(I1,I1) = 1.0D0

40 CONTINUE

C PERFORM THE SINGULAR VALUE DECOMPOSITION.
CALL LSVDF(SO, DC2, NIC2, NIC2, PO, DC2, NIC2, SING, WK, IER)
WRITE(6,5002) (SING(I), I=1,NIC2)
5002 FORMAT(/,1X,'SINGULAR VALUES OF THE MSM MATRIX'/10(1X,D11.4,1X))
WRITE(6,5003) DF
5003 FORMAT(/,1X,'NUMBER OF NON-ZERO SINGULAR VALUES ',12)

C CALCULATE THE GENERALIZED INVERSE.
DO 55 I2=1,NIC2
   DO 45 I1=1,DF
      PO(I1,I2) = PO(I1,I2) / SING(I1)
45 CONTINUE
   DF = DF + 1
   DO 50 I1=DF,NIC2
      PO(I1,I2) = 0.0D0
50 CONTINUE
   DF = DF + 1
55 CONTINUE

C PRE- AND POST- MULTIPLY THE GENERALIZED INVERSE BY THE ORIGINAL

CALL VMULFF(SO, PO, NIC2, NIC2, DC2, DC2, MWK, DC2, IER)
WRITE(6,5004)

5004 FORMAT(/,1X,'GENERALIZED INVERSE OF MSM(PARTIAL)'
DO 65 I1=1,NIC2
   IF (I1.LE.12) WRITE (6,5001) (MWK(I1,I2), I2=1,6)
65 CONTINUE
DO 60 I2=1,NIC2
    PO(I1,I2)=MWK(I1,I2)
60    CONTINUE
65 CONTINUE
C PRE- AND POST-MULTIPLY THE GENERALIZED INVERSE.
    CALL VMULFF(PA,MWK,NIC2,NIC2,NIC2,DC2,DC2,SO,DC2,IER)
    CALL VMULFF(SO,PA,NIC2,NIC2,NIC2,DC2,DC2,MWK,DC2,IER)
    WRITE(6,5005)
5005 FORMAT('RESULT OF PRE- AND POST-MULTIPLYING THE GENERALIZED',
      ' INVERSE OF THE MSM MATRIX BY THE MSM MATRIX')
    DO 70 I1=1,12
      WRITE(6,5001) (MWK(I1,I2),12=1,6)
70    CONTINUE
    WRITE(6,5006)
5006 FORMAT('RESULT OF PRE- AND POST-MULTIPLYING THE GENERALIZED',
      ' INVERSE OF THE MSM MATRIX BY THE MSM MATRIX')
C CALCULATE THE TEST STATISTIC AND OUTPUT THE RESULTS.
    CALL VMULFM(FA,P0,NIC2,1,NIC2,DC2,DC2,W,1,IER)
    CALL VMULFF(SO,FA,1,NXC2,1,DC2,DC2,W,1,IER)
    W=W+FLOAT(NJ)*W
    WRITE(6,2005) W,DF
2005 FORMAT('VALUE OF N*W=',D14.7,' WITH ',I3,' DEGREES OF',
      ' FREEDOM')
STOP
END
C SUBROUTINE RD(UNIT,DIM,NI,NJ,N,X1,X2)
C
INTEGER DIM,N(DIM,DIM),X1(DIM,DIM),X2(DIM,DIM),
UNIT,CNT(10,10)
LOGICAL FLAG
C
FLAG=.TRUE.
DO 5 I1=1,NI
      READ(UNIT,1000) (N(I1,I2),I2=1,NI)
1000 FORMAT(I10)
5 CONTINUE
C
DO 10 I1=1,NI
  DO 10 I2=1,NI
    IF (N(I1,I2).NE.N(I2,I1)) GO TO 103
    X1(I1,I2)=0
103    CONTINUE
X1(I2,I1)=0
X2(I1,I2)=0
X2(I2,I1)=0
10 CONTINUE
C
NJ=0
20 DO 30 I1=1,NI
   READ(UNIT,1001,END=55) (CNT(I1,I2),I2=1,NI)
1001 FORMAT(IOI3)
   DO 25 I2=1,NI
      IC=CNT(I1,I2)
      X1(I1,I2)=X1(I1,I2)+IC
      X2(I1,I2)=X2(I1,I2)+(IC**2)
25 CONTINUE
30 CONTINUE
NJ=NJ+1
40 DO 35 I1=1,NI
   IC=CNT(I1,I2)+CNT(I2,I1)
   IF (IC.EQ.NI(I1,I2)) GO TO 35
   FLAG=.FALSE.
   WRITE(6,4001) NJ,I1,I2
4001 FORMAT('***ERROR*** JUDGE ',I3, ' HAS AN INCORRECT X MATRIX F', + 'OR THE ',I2,',',I2,' COMPARISON')
35 CONTINUE
40 CONTINUE
GO TO 20
C
55 WRITE(6,2000)
2000 FORMAT('INPUT DATA READ IN'/ON MATRIX')
   DO 101 I1=1,NI
      WRITE(6,2001) (N(I1,I2),I2=1,NI)
2001 FORMAT(1X,10(I3,2X))
101 CONTINUE
   WRITE(6,2002) NJ
2002 FORMAT(ONUMBER OF JUDGES: ',I3/OX MATRIX')
   DO 102 I1=1,NI
      WRITE(6,2001) (X1(I1,I2),I2=1,NI)
102 CONTINUE
   IF (FLAG) RETURN
STOP
103 WRITE(6,4002)
4002 FORMAT('***ERROR*** INCORRECT N MATRIX')
STOP
END
C
SUBROUTINE SET(DIM,NI,S,H,M,C,CEST)
SET UP THE M MATRIX TO CONTAIN THE DISTINCT SCALE VALUE CODES AND C
THE NUMBER OF ITEMS ASSOCIATED WITH EACH CODE, FOR THE HYPOTHESIS C
SPECIFIED.

INTEGER DIM,S,H(DIM),M(DIM,2)
REAL*8 C
LOGICAL CEST

S=1
M(1,1)=H(1)
M(1,2)=1
DO 10 I1=2,NI
   DO 5 I2=1,S
      IF (H(I1).NE.M(I2,1)) GO TO 5
      M(I2,2)=M(I2,2)+1
   GO TO 10
5 CONTINUE
   S=S+1
   M(S,1)=H(I1)
   M(S,2)=1
10 CONTINUE

WRITE(6,2000) (H(I),I=1,NI)
2000 FORMAT('HYPOTHESIS VECTOR',14X,10(2X,I3))
WRITE(6,2001) S
2001 FORMAT(' NUMBER OF DISTINCT SCALE VALUES ',I3)
WRITE(6,2002) (M(I,1),I=1,S)
2002 FORMAT(' DISTINCT SCALE VALUE CODES',5X,10(2X,I3))
WRITE(6,2003) (M(I,2),I=1,S)
2003 FORMAT(' NUMBER OF ASSOCIATED ITEMS',5X,10(2X,I3))
IF (CEST) GO TO 15
WRITE(6,2004) C
2004 FORMAT(' THE VALUE OF THE JUDGE VARIABILITY PARAMETER IS HYPOTHE',
     + 'SIZED TO EQUAL ',D14.7)
RETURN
15 WRITE(6,2005)
2005 FORMAT(' THE VALUE OF THE JUDGE VARIABILITY PARAMETER IS NOT SPE',
     + 'CIFIED BY THE HYPOTHESIS')
RETURN
END

SUBROUTINE MM(DIM,NI,NJ,S,N,X1,X2,H,M,CEST,PI1,CWT)

CALCULATE THE METHOD OF MOMENTS ESTIMATOR FOR THE SCALE VALUES AND C
THE JUDGE VARIABILITY PARAMETER UNDER THE SPECIFIED HYPOTHESIS. C
C

INTEGER DIM,S,N(DIM,DIM),X1(DIM,DIM),X2(DIM,DIM),
+ H(DIM),M(DIM,2),D,DW
REAL*8 SUM,A,SCALE,DEV,MAX,PI1(DIM),PI2(10),B,C,CT,CWT,CLIMIT,
+ DFLOAT,DMAX1,DABS
LOGICAL CEST

C
DATA CLIMIT/30.0D00/
COMMON /JUDGE/DXX,XN,P1,P2,P3
C
C INITIALIZE THE ITERATION COUNTER(IT) AND THE VECTOR OF SCALE VALUES
C (PI). FOR THE CASE OF 1 SCALE VALUE(IE, ALL ITEMS EQUAL), SKIP OVER
C THE ITERATIVE CALCULATIONS OF THE MLE'S FOR THE SCALE VALUES.
IT=0
DO 4 I=1,S
  PI1(I)=1.0D00/DFLOAT(NI)
4 CONTINUE
WRITE(6,2000) (PI1(I),1=1,S)
2000 FORMAT('ITERATIVE CALCULATION OF THE SCALE VALUE ESTIMATES'/
+ ' INT',10(1X,D10.4))
IF (S.EQ.1) GO TO 26
C
C ITERATIVELY CALCULATE THE METHOD OF MOMENT ESTIMATORS FOR THE SCALE
C VALUES. THIS IS DONE BY DETERMINING IF ITEM I1 HAS THE SCALE VALUE
C BEING ESTIMATED ASSOCIATED WITH IT, AND IF SO, CUMULATING X(I1,I2)
C AND N(I1,I2)/(PI(I1)+PI(I2)) FOR THOSE ITEMS OTHER THAN I1. THESE SUMS
C ARE THEN USED TO CALCULATE THE NEXT ITERATION OF METHOD OF MOMENT
C ESTIMATORS.
5 SCALE=0.0D00
DO 20 I=1,S
  SUM=0.0D00
  A=0.0D00
  DO 15 I1=1,NI
    IF (H(I1).NE.M(I1)) GO TO 15
    DO 10 I2=1,NI
      IF (I2.EQ.I1) GO TO 10
      A=A+DFLOAT(X1(I1,I2))
      SUM=SUM+DFLOAT(N(I1,I2))/(PI1(H(I1))+PI1(H(I2)))
10    CONTINUE
15  CONTINUE
  PI2(I)=A/(DFLOAT(NI)*SUM)
  SCALE=SCALE+DFLOAT(M(I,2))*PI2(I)
20 CONTINUE
C
C CALCULATE THE MAXIMUM ABSOLUTE DEVIATION BETWEEN THE ESTIMATES OF
C THIS ITERATION AND THOSE OF THE PREVIOUS ITERATION. IF THE
C DEVIATION IS SMALL ENOUGH OR THE NUMBER OF ITERATIONS GETS TOO LARGE,
C STOP ITERATING.
24 MAX=0.0D00
DO 25 I=1,S
PI2(I)=PI2(I)/SCALE
DEV=PI2(I)-PI1(I)
DEV=DABS(DEV)
MAX=DMAX1(DEV,MAX)
PI1(I)=PI2(I)
25 CONTINUE
IT=IT+1
IREM=MOD(IT,10)
IF (IREM.EQ.0) WRITE(6,2001) IT,(PI1(I),I=1,S),MAX
2001 FORMAT(1X,I3,11(1X,D10.M))
IF ((IT.LT.500).AND.(MAX.GT.1.0D-05)) GO TO 5
IF (IREM.NE.0) WRITE(6,2001) IT,(PI1(I),I=1,S),MAX
C 26 IF (.NOT.CEST) RETURN
C C CALCULATION OF THE METHOD OF MOMENT ESTIMATOR FOR THE JUDGE
C VARIABILITY PARAMETER BASED ON THE SECOND MOMENTS OF THE COMPARISONS
C UNDER THE BETA-BINOMIAL MODEL. ESTIMATES ARE TRUNCATED TO FALL BETWEEN
C ZERO AND CLIMIT. UNWEIGHTED AND WEIGHTED AVERAGES OF THE VARIOUS
C ESTIMATES FROM THE VARIOUS SAMPLE MOMENTS ARE CALCULATED AND USED AS
C THE FINAL ESTIMATES.
WRITE(6,2002)
2002 FORMAT('CALCULATION OF THE METHOD OF MOMENTS ESTIMATOR FOR THE '
+ 'JUDGE VARIABILITY PARAMETER')
ILO=0
IUP=0
D=0
DW=0
CT=O.ODOO
CWT=O.ODOO
DO 35 I1=1,NI
C=O.ODOO
DO 30 I2=1,I1
IF (N(I1,I2).LE.1) GO TO 30
IF (I1.EQ.I2) GO TO 30
B=DFLOAT(X2(I1,I2))*(PI1(H(I1))+PI1(H(I2)))
B=B/(DFLOAT(N(I1,I2))*PI1(H(I1))*DFLOAT(NJ))
B=(B-1.0D0)/(DFLOAT(N(I1,I2))-1.0D0)
C=(1.0D0-B)/(B*PI1(H(I2)))+(B-1.0D0)*PI1(H(I1)))
IF (C.T.GT.0.ODOO) GO TO 27
ILO=ILO+1
GO TO 29
27 IF (C.LT.CLIMIT) GO TO 28
IUP=IUP+1
C=CLIMIT
28 CT=CT+C
CWT=CWT+DFLOAT(N(I1,I2))*C
29 D=D+1
DW=DW+N(I1, I2)
30 CONTINUE
35 CONTINUE
CT=CT/DFLOAT(D)
CWT=CWT/DFLOAT(DW)
WRITE(6,2003) CT, CWT, D, ILO, IUP, CLIMIT
2003 FORMAT(' JUDGE VARIABILITY ESTIMATOR (BASED ON VARIANCES): ',
+ 'D14.7,' (UNWEIGHTED AVERAGE)'/51X,D14.7,
+ '(WEIGHTED AVERAGE)'' AVERAGE BASED ON',1X,I3,
+ 'VALUES OF WHICH ',I3,' WERE TRUNCATED AT 0.0000 AND ',
+ I3,' WERE TRUNCATED AT ',D10.3)
RETURN
END
C SUBROUTINE PMLE(DIM, DC2, NI, N, S, H, M, N, X, PI, C, PO, M2, VO, M4)
C CALCULATE THE PSEUDO-MLES FOR THE SCALE VALUES AND THE ASSOCIATED C
C VARIANCES AND COVARIANCES.
C INTEGER DIM, S, DC2, N(DIM, DIM), X(DIM, DIM), H(DIM), M(DIM, 2)
REAL*8 P(90), PO(DC2), VO(90), A(90, 10), XM, PX, MAX, PI(DIM), C, B(10),
+ ATA(10, 10), IATA(10, 10), WK(130), M1(90, 90), M2(DC2, DC2),
+ M3(90, 10), M4(90, 90), DFLOAT, DSQRT, DABS, DMAX1
LOGICAL STOPIT
C STOPIT=.FALSE.
IS=S=1
NJC2=NI*(NI-1)
C C ITERATIVELY CALCULATE THE PSEUDO-MLES. FOR THE CASE OF A SINGLE SCALE
C VALUE TO BE ESTIMATED, SKIP THE ITERATIVE CALCULATIONS AND CALCULATE
C THE VALUES FOR THE CHOICE PROBABILITY VECTOR, PROJECTION MATRIX AND
C RELATED VALUES.
WRITE(6,2000) (PI(I), I=1, S)
2000 FORMAT(' ITERATIVE CALCULATION OF THE SCALE VALUE ESTIMATES'/
+ ' INT', 10(1X, D10.4))
IF (S.EQ.1) GO TO 26
IT=0
5 I=1
DO 15 J2=1, NI
DO 15 J1=1, NI
IF (J1.EQ.J2) GO TO 15
P(I)=DFLOAT(X(J1, J2))/(DFLOAT(N(J1, J2))*DFLOAT(NJ))
PO(I)=PI(H(J1))/(PI(H(J1))+PI(H(J2)))
VO(I)=PO(I)/DFLOAT(N(J1, J2))
PO(I)=P(I)/PO(I)/DSQRT(VO(I))
C THE MATRIX A BEING CALCULATED IS THE MATRIX OF DERIVATIVES OF THE
DO 10 I1=1,IS
   A(I,I1)=0.0D0
   IF (H(J1).EQ.H(J2)) GO TO 10
   PX=(PI(H(J1))+PI(H(J2)))**2
   XM=DFLOAT(M(I1,2))/DFLOAT(M(S,2))
   IF (M(I1,1).EQ.H(J1)) A(I,I1)=PI(H(J2))/PX
   IF (M(I1,1).EQ.H(J2)) A(I,I1)=-PI(H(J1))/PX
   IF (M(S,1).EQ.H(J1)) A(I,I1)=A(I,I1)*XM*PI(H(J2))/PX
   IF (M(S,1).EQ.H(J2)) A(I,I1)=A(I,I1)+XM*PI(H(J1))/PX
   A(I,I1)=A(I,I1)/DSQRT(VO(I))
10 CONTINUE
   I=I+1
15 CONTINUE
   CALL VMULFM(A,P0,NIC2,IS,1,DC2,DC2,P,DC2,IER)
   CALL VTPROF(A,NIC2,IS,DC2,P0)
   CALL VCVTSF(P0,IS,ATA,DIM)
   IDGT=0
   CALL LINV2F(ATA,IS,DIM,IATA,DIM1,IER)
   CALL VMULFF(IATA,P,IS,IS,1,DIM,B,DIM,IER)
   IF (STOPIT) GO TO 25
C
C UPDATE THE SCALE VALUE ESTIMATES. IF THE MAXIMUM ABSOLUTE DEVIATION
C BETWEEN THESE NEW ESTIMATES AND THOSE OF THE PREVIOUS ITERATION IS
C SMALL ENOUGH OR THE NUMBER OF ITERATIONS IS TOO LARGE, STOP ITERATING.
PX=0.0D0
MAX=0.0D0
DO 20 I=1,IS
   IF (B(I).GT.PI(I)) GO TO 19
   B(I)=1.0D0-PI(I)
19 PI(I)=PI(I)+B(I)
   PX=PX+DFLOAT(M(I,2))*B(I)
   B(I)=DABS(B(I))
   MAX=DMAX1(MAX,B(I))
20 CONTINUE
   PI(S)=PI(S)+PX+DFLOAT(M(S,2))
   PX=DABS(PX)
   MAX=DMAX1(MAX,PX)
   IT=IT+1
   IREM=MOD(IT,10)
   IF (IREM.EQ.0) WRITE(6,2001) IT,(PI(I),I=1,S),MAX
2001 FORMAT(1X,I3,11(1X,D10.4))
   IF ((IT.GT.500).OR.(MAX.LT.1.0D-05)) STOPIT=.TRUE.
   GO TO 5
C
25 IF (IREM.NE.0) WRITE(6,2001) IT,(PI(I),I=1,S),MAX
   CALL VMULFF(A,IATA,NIC2,IS,IS,DC2,DIM,M1,DC2,IER)
C CALCULATE THE CHOICE PROBABILITY VECTOR, THE PROJECTION MATRIX AND
C THE VARIANCE-COVARIANCE MATRIX FOR THE COMPARISONS UNDER THE BETA-
C BINOMIAL MODEL.

26 J=1
  DO 45 J2=1,NI
    DO 45 J1=1,NI
      IF (J1.EQ.J2) GO TO 45
      P0(J)=PI(H(J1))/(PI(H(J2))+PI(H(J1)))
      VO(J)=PO(J)/DFLOAT(N(J1,J2))
      PX=(DFLOAT(N(J1,J2))-1.0D00)*(C*PI(H(J1))+1.0D00)
      PX=PX/(C*(PI(H(J1))+PI(H(J2)))+1.0D00)
      M4(J,J)=DFLOAT(N(J1,J2))*PO(J)*PX
      J=J+1
    CONTINUE
  CONTINUE

C CALCULATE THE VARIANCES AND COVARIANCES FOR THE SCALE VALUE ESTIMATES.

DO 50 I=1,NIC2
  DO 50 J=1,NIC2
    M2(I,J)=M4(I,J)/(DSQRT(VO(I))*DSQRT(VO(J)))
  CONTINUE

CALL VMULFF(M2,M1,NIC2,NIC2,IS,DC2,DC2,M3,DC2,IER)
CALL VMULFM(M1,M3,NIC2,IS,IS,DC2,DC2,IATA,DIM,IER)
WRITE(6,2002)
2002 FORMAT('COVARIANCE-COVARIANCE MATRIX FOR THE SCALE VALUE ESTIMATE',
    + 'RS')
  DO 55 I=1,S
    IATA(S,I)=0.0D00
  CONTINUE
  DO 65 I=1,S
    IATA(I,J)=IATA(I,J)/DFLOAT(NJ)
    IATA(S,J)=IATA(S,J)-DFLOAT(M(I,2))/DFLOAT(M(S,2))*IATA(I,J)
  CONTINUE

CONTINUE
WRITE(6,2003) (IATA(I,J),J=1,I)
2003 FORMAT(1X,10(D11.4,2X))
65 CONTINUE
   DO 70 J=1,IS
      IATA(S,S)=IATA(S,S)-DFLOAT(M(J,2))/DFLOAT(M(S,2))*IATA(S,J)
70 CONTINUE
   WRITE(6,2003) (IATA(S,I),I=1,S)
C
C CALCULATE THE PROJECTION MATRIX ASSOCIATED WITH THE HYPOTHESES
C CHOICE PROBABILITY VECTOR.
   CALL VMULFP(A,M1,NIC2,IS,NIC2,DC2,DC2,M2,DC2,IER)
   RETURN
END
12. APPENDIX D

Listing of the computer program that calculates estimates for the scale values and judge variability parameter, and performs tests of hypotheses for the Dirichlet-Binomial model.
INTEGER UNIT, DIM, DC2, S, DFO, DFA, DF, HO(10), HA(10), M(10, 2),
+ N(10, 10), XI(10, 10), X2(10, 10), XX(90, 90)
REAL*8 CO, CA, PI(10), W, SING(90), WK(180), FO(90), FA(90), VO(90),
+ PO(90, 90), PA(90, 90), SO(90, 90), MK(90, 90), DGT
LOGICAL CESTO, CESTA, GOF
C UNIT = DEVICE NUMBER DATA IS TO BE READ FROM.
C DIM = DIMENSION ARRAYS HAVE BEEN SET UP WITH. THIS EQUALS THE MAXIMUM
C NUMBER OF ITEMS IN AN EXPERIMENT THE PROGRAM CAN HANDLE.
C DC2 = DIM*(DIM-1)
C NI = NUMBER OF ITEMS IN THE EXPERIMENT.
C NIC2 = NI*(NI-1)
C NJ = NUMBER OF JUDGES IN THE EXPERIMENT.
C N = SYMMETRIC ARRAY GIVING THE NUMBER OF TIMES THE FIRST DIMENSION
C ITEM IS COMPARED TO THE SECOND DIMENSION ITEM BY A JUDGE.
C X1 = CUMULATIVE CHOICE MATRIX CONTAINING THE NUMBER OF TIMES THE
C FIRST DIMENSION ITEM IS PREFERRED TO THE SECOND DIMENSION ITEM
C OVER ALL JUDGES.
C X2 = CUMULATIVE(OVER ALL JUDGES) MATRIX OF THE SQUARE OF THE ELEMENTS
C OF EACH JUDGE'S CHOICE MATRIX.
C XX = CUMULATIVE(OVER ALL JUDGES) MATRIX OF ALL POSSIBLE CROSS PRODUCTS
C BETWEEN THE ELEMENTS OF A JUDGE'S CHOICE MATRIX.
C DFO, DFA = DEGREES OF FREEDOM UNDER HO, HA.
C HO, HA = VECTORS SPECIFYING THE NULL, ALTERNATIVE HYPOTHESES BY
C SPECIFYING WHICH SCALE VALUE IS ASSOCIATED WITH WHICH ITEM.
C EG., A ONE IN THE FIFTH POSITION SPECIFIES THAT ITEM FIVE
C HAS SCALE VALUE ONE ASSOCIATED WITH IT.
C M = M(I,1): IDENTIFIES THE SCALE VALUE CODE ASSOCIATED WITH SCALE
C VALUE I.
C M(I,2): THE NUMBER OF ITEMS WITH SCALE VALUE I.
C S = NUMBER OF DISTINCT SCALE VALUES.
C PI = VECTOR OF SCALE VALUES.
C CO, CA = JUDGE VARIABILITY PARAMETER UNDER HO, HA.
C FO, FA = VECTOR OF CHOICE PROBABILITIES PREDICTED BY THE DIRICHLET-
C BINOMIAL MODEL UNDER HO, HA.
C PO, PA = PROJECTION MATRICES FOR THE CHOICE PROBABILITY VECTOR UNDER
C HO, HA.
C VO = VECTOR OF CHOICE PROBABILITIES FOR THE DIRICHLET-BINOMIAL MODEL
C UNDER HO, DIVIDED BY THE NUMBER OF TIMES A COMPARISON IS MADE.
C SO = VARIANCE-COVARIANCE MATRIX FOR THE COMPARISONS UNDER THE
C DIRICHLET-BINOMIAL MODEL UNDER HO.
C CESTO, CESTA = FLAG DETERMINING WHETHER OR NOT THE JUDGE VARIABILITY
C PARAMETER SHOULD BE ESTIMATED UNDER HO, HA.
C GOF = FLAG INDICATING WHETHER OR NOT A GOODNESS-OF-FIT TEST FOR THE
C MODEL IS TO BE PERFORMED.
C
DATA UNIT/10/, DIM/10/, NI/10/, CESTO/.TRUE./, CESTA/.TRUE./,
+ CO/1.0D00/, CA/1.0D00/, HO/1, 2, 3, 4, 5, 6, 7, 8, 9, 10/,
+ HA/1, 2, 3, 4, 5, 6, 7, 8, 9, 10/, GOF/.TRUE./
C IGNORE OVERFLOW AND UNDERFLOW ERROR LIMITS AND SUPPRESS WARNING
C MESSAGES.
   CALL ERRSET(207,256,-1,0)
   CALL ERRSET(208,0,-1,0)
   NIC2=NI*(NI-1)
   DC2=DIM*(DIM-1)
C
C READ INPUT DATA.
   CALL RD(UNIT,DIM,DC2,NI,NJ,N,X1,X2,XX)
C
C CALCULATE VARIOUS PARAMETER ESTIMATES FOR THE DIRICHLET-BINOMIAL MODEL
C UNDER HO.
   WRITE(6,2000)
2000 FORMAT(//1X,132('*')/'CALCULATION OF ESTIMATORS FOR THE JUDGE V',
      + 'ARIABILITY PARAMETER AND SCALE VALUE PARAMETERS FOR THE',
      + 'DIRICHLET-BINOMIAL MODEL UNDER HO')
C
C SETUP THE M MATRIX FOR HO.
   CALL SET(DIM,NI,S,HO,M,CO,CESTO)
   DF0=S^1
C
C CALCULATE THE METHOD OF MOMENT ESTIMATORS FOR THE SCALE VALUES AND
C THE JUDGE VARIABILITY PARAMETER FOR THE DIRICHLET-BINOMIAL MODEL
C UNDER HO.
   WRITE(6,2001)
2001 FORMAT(/'=',20('*'),' METHOD OF MOMENT ESTIMATION OF THE PARAMET',
      + 'ERS ',20('*'))
   CALL MM(DIM,DC2,NI,NJ,S,N,X1,X2,XX,H0,M,CEST0,PI,CO)
C
C INITIALIZE SCALE VALUE ESTIMATES AND CALCULATE PSEUDO-MLES FOR THE
C SCALE VALUES FOR THE DIRICHLET-BINOMIAL MODEL UNDER HO. ALSO CALCULATE
C AND OUTPUT VO AND SO.
   WRITE(6,2002)
2002 FORMAT(/'=',20('*'),' PSEUDO-MAXIMUM LIKELIHOOD ESTIMATION OF TH',
      + 'E PARAMETERS ',20('*'))
   DO 1 I=1,S
      PI(I)=1.0DOO/DFLOAT(NI)
   1 CONTINUE
   CALL PMLE(DIM,DC2,NI,NJ,S,N,X1,PI,CO,FO,PO,VO,SO)
C
   IF (GOF) GO TO 5
C
C CALCULATE VARIOUS PARAMETER ESTIMATES FOR THE DIRICHLET-BINOMIAL MODEL
C UNDER HA.
   WRITE(6,2003)
2003 FORMAT(//1X,132('*')/'CALCULATION OF ESTIMATORS FOR THE JUDGE V',
      + 'ARIABILITY PARAMETER AND SCALE VALUE PARAMETERS FOR THE',
      + 'DIRICHLET-BINOMIAL MODEL UNDER HA')
C SETUP THE M MATRIX FOR HA.
   CALL SET(DIM, NI, S, HA, M, CA, CESTA)
   DFA=S-1
C
C CALCULATE THE METHOD OF MOMENT ESTIMATORS FOR THE SCALE VALUES AND
C THE JUDGE VARIABILITY PARAMETER FOR THE DIRICHLET-BINOMIAL MODEL
C UNDER HA.
   WRITE(6,2001)
   CALL MM(DIM, DC2, NI, NJ, S, N, X1, X2, XX, HA, M, CESTA, PI, CA)
   WRITE(6,2002)
C
C INITIALIZE SCALE VALUE ESTIMATES AND CALCULATE PSEUDO-MLES FOR THE
C SCALE VALUES FOR THE DIRICHLET-BINOMIAL MODEL UNDER HA.
   DO 2 I=1,S
   PI(I)=1.0D00/DFLOAT(NI)
   2 CONTINUE
   CALL PMLE(DIM, DC2, NI, NJ, S, HA, M, N, X1, PI, CA, FA, PA, SING, MWK)
   GO TO 20
C
C CALCULATE FA, PA FOR THE ALTERNATIVE GOODNESS-OF-FIT HYPOTHESIS.
   5 WRITE(6,2004)
5004 FORMAT(/'ALTERNATIVE HYPOTHESIS IS THE UNRESTRICTED HYPOTHESIS '
      + 'FOR A GOODNESS OF FIT TEST')
   I=1
   DO 15 J2=1,NI
   DO 15 J1=1,NI
      IF (J1.EQ.J2) GO TO 15
      FA(I)=DFLOAT(X1(J1,J2))/(DFLOAT(N(J1,J2))*DFLOAT(NJ))
   DO 10 I1=1,NIC2
      PA(I,I1)=0.0D00
   10 CONTINUE
   PA(I,I)=1.0D00
   I=I+1
   15 CONTINUE
   DFA=NIC2/2
C
C CALCULATE THE DIFFERENCE VECTOR FA-FO AND THE DIFFERENCE PROJECTION
C MATRIX PA-PO. THIS MATRIX IS THEN PRE-MULTIPLIED BY A DIAGONAL MATRIX
C WITH THE SQUARE ROOT OF THE ELEMENTS OF VO DOWN THE DIAGONAL, POST-
C MULTIPLIED BY THE INVERSE OF A DIAGONAL MATRIX WITH THE SQUARE ROOT
C OF THE ELEMENTS OF VO DOWN THE DIAGONAL AND POST-MULTIPLIED AGAIN BY
C A DIAGONAL MATRIX WITH THE ELEMENTS OF THE COMPARISON MATRIX N DOWN
C THE DIAGONAL.
   20 DF=DFA-DF0
   DO 30 I1=1,NIC2
      FA(I1)=FA(I1)-FO(I1)
   DO 25 I2=1,NIC2
      J2=(I2-1)/(NI-1)+1
   30 CONTINUE
J1=I2-(J2-1)*(NI-1)
IF (J1.GE.J2) J1=J1+1
PA(I1,I2)=PA(I1,I2)-PO(I1,I2)
PA(I1,I2)=DSQRT(V0(I1)) Pa(I1,I2)
PA(I1,I2)=PA(I1,I2)/(DSQRT(V0(I2))*DFLOAT(N(J1,J2)))
25 CONTINUE
30 CONTINUE
C
C PRE- AND POST- MULTIPLY SO BY PA-PO CALCULATED ABOVE.
CALL VMULFF(PA,SO,NIC2,NIC2,NIC2,DC2,DC2,PO,DC2,IER)
CALL VMULFP(PO,PA,NIC2,NIC2,NIC2,DC2,DC2,SO,DC2,IER)
WRITE(6,5000)
C
C PERFORM AN ACCURACY CHECK FOR THE CALCULATION OF THE GENERALIZED
C INVERSE, TO SEE THAT IGNORING UNDERFLOW AND OVERFLOW CONDITIONS HAS
C NOT UPSET CALCULATIONS GREATLY. THIS IS DONE BY CALCULATING THE
C GENERALIZED INVERSE (USING SINGULAR VALUE DECOMPOSITION) AND THEN
C PRE- AND POST- MULTIPLYING THE GENERALIZED INVERSE BY THE ORIGINAL
C MATRIX.
5000 FORMAT(/,1X,132('*')/1X,'ACCURACY CHECK FOR THE GENERALIZED INVER',
    + 'SE CALCULATION'/1X,'MSM MATRIX(PARTIAL)' )
DO 40 I1=1,NIC2
  IF (II.LE.12) WRITE(6,5001) (S0(I1,12),12=1,6)
5001 FORMAT(1X,6(D11.4,3X))
C
C SAVE SO FOR FUTURE USE AND INITIALIZE PO.
DO 35 I2=1,NIC2
  PA(I1,I2)=SO(I1,I2)
  P0(I1,I2)=0.0000
35 CONTINUE
P0(I1,11)=1.0000
40 CONTINUE
C
C PERFORM THE SINGULAR VALUE DECOMPOSITION.
CALL LSVDF(SO,DC2,NIC2,NIC2,PO,DC2,NIC2,SING,WK,IER)
WRITE(6,5002) (SING(I),I=1,NIC2)
5002 FORMAT(/1X,'SINGULAR VALUES OF THE MSM MATRIX'/10(1X,D11.4,3X))
WRITE(6,5003) DF
5003 FORMAT(/1X,'NUMBER OF NON-ZERO SINGULAR VALUES ',12)
C
C CALCULATE THE GENERALIZED INVERSE.
DO 55 I2=1,NIC2
  DO 45 11=1,DF
    P0(I1,I2)=P0(I1,I2)/SING(I1)
45 CONTINUE
  DF=DF+1
  DO 50 I1=DF,NIC2
    P0(I1,I2)=0.0000
50 CONTINUE
  DF=DF-1
55 CONTINUE
CALL VMULFF(S0,PO,NIC2,NIC2,NIC2,DC2,DC2,MWK,DC2,IER)
WRITE(6,5004)
5004 FORMAT(/1X,'GENERALIZED INVERSE OF MSM(PARTIAL)')
DO 65 I1=1,NIC2
   IF (I1.LE.12) WRITE(6,5001) (MWK(I1,I2),I2=1,6)
   DO 60 I2=1,NIC2
      PO(I1,I2)=MWK(I1,I2)
   60 CONTINUE
65 CONTINUE
C PRE- AND POST-MULTIPLY THE GENERALIZED INVERSE.
CALL VMULFF(PA,MWK,NIC2,NIC2,NIC2,DC2,DC2,SO,DC2,IER)
CALL VMULFF(SO,PA,NIC2,NIC2,NIC2,DC2,DC2,MWK,DC2,IER)
WRITE(6,5005)
5005 FORMAT(/1X,'RESULT OF PRE- AND POST-MULTIPLYING THE GENERALIZED',
     + ' INVERSE OF THE MSM MATRIX BY THE MSM MATRIX')
DO 70 I1=1,12
   WRITE(6,5001) (MWK(I1,I2),I2=1,6)
70 CONTINUE
WRITE(6,5006)
5006 FORMAT(/,1X,132('*'))
C
C CALCULATE THE TEST STATISTIC AND OUTPUT THE RESULTS.
CALL VMULFM(FA,P0,NIC2,1,NIC2,DC2,DC2,S0,DC2,IER)
CALL VMULFF(S0,FA,1,NIC2,1,DC2,DC2,W,1,IER)
W=DFLOAT(NJ)*W
WRITE(6,2005) W,DF
2005 FORMAT(/20('*')/'VALUE OF N*W=',D14.7,' WITH ',13,' DEGREES OF',
     + ' FREEDOM'/*1')
STOP
END
C
SUBROUTINE RD(UNIT,DIM,DC2,NI,NJ,N,XI,X2,XX)
C
C THE COMPARISON MATRIX N IS INPUTTED AND CHECKED FOR CONSISTENCY. C
C THE CHOICE VECTORS FOR EACH JUDGE ARE READ IN NEXT, CHECKED FOR C
C CONSISTENCY AGAINST THE COMPARISON MATRIX AND ADDED TO THE CHOICE C
C MATRIX XI. THE SQUARES OF THE ELEMENTS ARE ADDED TO X2 AND THE CROSS C
C PRODUCTS OF THE ELEMENTS ARE ADDED TO XX.
C
integer DIM,DC2,N(DIM,DIM),XI(DIM,DIM),X2(DIM,DIM),XX(DC2,DC2),
   + UNIT,CNT(10,10)
logical flag
C
flag=.true.
DO 5 I1=1,NI
   READ(UNIT,1000) (N(I1,I2),I2=1,NI)
1000 FORMAT(10I0)
5 CONTINUE
C
DO 10 I1=1,NI
DO 10 I2=1,NI
IF (N(I1,I2).NE.N(I2,I1)) GO TO 103
X1(I1,I2)=0
X1(I2,I1)=0
X2(I1,I2)=0
X2(I2,I1)=0
10 CONTINUE
C
NIC2=NI*(NI+1)
DO 15 I=1,NIC2
DO 15 J=1,NIC2
XX(I,J)=0
15 CONTINUE
C
NJ=0
20 DO 30 I1=1,NI
READ(UNIT,1001,END=55) (CNT(I1,12),12=1,NI)
1001 FORMAT(10I3)
DO 25 I2=1,NI
IC=CNT(I1,I2)
X1(I1,I2)=X1(I1,I2)+IC
X2(I1,I2)=X2(I1,I2)+(IC**2)
25 CONTINUE
30 CONTINUE
C
NJ=NJ+1
DO 40 I1=1,NI
DO 35 I2=1,NI
IC=CNT(I1,I2)+CNT(I2,I1)
IF (IC.EQ.N(I1,I2)) GO TO 35
FLAG=.FALSE.
WRITE(6,4001) NJ,I1,I2
4001 FORMAT('***ERROR*** JUDGE ',',I3,' HAS AN INCORRECT X MATRIX F',+OR THE ',',I2,',',I2,' COMPARISON')
35 CONTINUE
40 CONTINUE
C
I=1
DO 50 I2=1,NI
DO 50 I1=1,NI
IF (I1.EQ.I2) GO TO 50
IF (I1.LT.I2) GO TO 49
J=1
DO 45 J2=1,NI
DO 45 J1=1,NI
IF (J1.EQ.J2) GO TO 45
IF (J1.LT.J2) GO TO 44
IF (J.LE.I) GO TO 44
   XX(I,J)=XX(I,J)+CNT(I1,I2)*CNT(J1,J2)
44  J=J+1
45  CONTINUE
49  I=I+1
50  CONTINUE
GO TO 20
C
55 WRITE(6,2000)
2000 FORMAT('INPUT DATA READ IN/ON MATRIX')
   DO 101 I1=1,NI
      WRITE(6,2001) (N(I1,I2),I2=1,NI)
2001 FORMAT(1X,10(I3,2X))
101 CONTINUE
   WRITE(6,2002) NJ
2002 FORMAT('NUMBER OF JUDGES: ',I3/0X MATRIX')
   DO 102 I1=1,NI
      WRITE(6,2001) (XI(II,12),12=1,NI)
102 CONTINUE
   IF (FLAG) RETURN
   STOP
103 WRITE(6,4002)
4002 FORMAT('***ERROR*** INCORRECT N MATRIX')
   STOP
END
C
SUBROUTINE SET(DIM,NI,S,H,M,C,CEST)
C
CC SET UP THE M MATRIX TO CONTAIN THE DISTINCT SCALE VALUE CODES AND C THE NUMBER OF ITEMS ASSOCIATED WITH EACH CODE, FOR THE HYPOTHESIS C SPECIFIED.
CC
INTEGER DIM,NI,S,H(DIM),M(DIM,2)
REAL*8 C
LOGICAL CEST
C
S=1
M(1,1)=H(1)
M(1,2)=1
DO 10 I1=2,NI
   DO 5 I2=1,S
      IF (H(I1).NE.M(I2,1)) GO TO 5
      M(I2,2)=M(I2,2)+1
      GO TO 10
5  CONTINUE
S=S+1
M(S,1)=H(I1)
M(S,2)=1
10 CONTINUE

C
WRITE(6,2000) (H(I),I=1,NI)
2000 FORMAT(/'HYPOTHESIS VECTOR',4X,10(2X,I3))
WRITE(6,2001) S
2001 FORMAT('NUMBER OF DISTINCT SCALE VALUES ',3X)
WRITE(6,2002) (M(I,1),I=1,S)
2002 FORMAT('DISTINCT SCALE VALUE CODES',5X,10(2X,I3))
WRITE(6,2003) (M(I,2),I=1,S)
2003 FORMAT('NUMBER OF ASSOCIATED ITEMS',5X,10(2X,I3))
IF (CEST) GO TO 15
WRITE(6,2004) C
2004 FORMAT('THE VALUE OF THE JUDGE VARIABILITY PARAMETER IS HYPOTHE',+
'SIZED TO EQUAL ',D14.7)
RETURN
15 WRITE(6,2005)
2005 FORMAT('THE VALUE OF THE JUDGE VARIABILITY PARAMETER IS NOT SPE',+
'IFIED BY THE HYPOTHESIS')
RETURN
END

C
SUBROUTINE MM(DIM,DC2,NI,NJ,S,N,X1,X2,XX,H,M,CEST,PI1,CWT)
C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C CALCULATE THE METHOD OF MOMENTS ESTIMATOR FOR THE SCALE VALUES AND C
C THE JUDGE VARIABILITY PARAMETER UNDER THE SPECIFIED HYPOTHESIS. C
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C
INTEGER DIM,DC2,S,N(DIM,DIM),XI(DIM,DIM),X2(DIM,DIM),XX(DC2,DC2), +
H(DIM),M(DIM,2),D,DW,DC,DWC
REAL*8 SUM,A,SCALE,DEV,MAX,PI1(DIM),PI2(10),B,C,CT,CWT, +
STEP,DXX,XN,P1,P2,P3,CR,CL,FR,FL,CM,EPS,CCT,CCWT,CLIMIT, +
FC,DFLOAT,DMAX1,DABS
LOGICAL CEST
C
DATA CLIMIT/30.0D00/
COMMON /JUDGE/DXX,XN,P1,P2,P3
EXTERNAL FC
C
C INITIALIZE THE ITERATION COUNTER(IT) AND THE VECTOR OF SCALE VALUES C
C (P ). FOR THE CASE OF 1 SCALE VALUE(IE, ALL ITEMS EQUAL), SKIP OVER C
C THE ITERATIVE CALCULATIONS OF THE MLE'S FOR THE SCALE VALUES.
IT=0
DO 4 I=1,S
PI1(I)=1.0D00/DFLOAT(NI)
4 CONTINUE
WRITE(6,2000) (PI1(I),I=1,S)
2000 FORMAT('ITERATIVE CALCULATION OF THE SCALE VALUE ESTIMATES')/
ITERATIVELY CALCULATE THE METHOD OF MOMENT ESTIMATORS FOR THE SCALE
VALUES. THIS IS DONE BY DETERMINING IF ITEM I1 HAS THE SCALE VALUE
BEING ESTIMATED ASSOCIATED WITH IT, AND IF SO, CUMULATING X(I1,I2)
AND N(I1,I2)/(PI(I1)*PI(I2)) FOR THOSE ITEMS OTHER THAN I1. THESE SUMS
ARE THEN USED TO CALCULATE THE NEXT ITERATION OF METHOD OF MOMENT
ESTIMATORS.

SCALE=0.000
DO 20 I=1,S
  SUM=0.000
  A=0.000
  DO 15 I1=1,N1
    IF (H(I1).NE.M(I,1)) GO TO 15
    DO 10 I2=1,N1
      IF (I2.EQ.I1) GO TO 10
      A=A+DFLOAT(X(I1,I2))
      SUM=SUM+DFLOAT(N(I1,I2))/(PI1(H(I1))+PI1(H(I2)))
  10 CONTINUE
  15 CONTINUE
  PI2(I)=A/(DFLOAT(N1)*SUM)
  SCALE=SCALE+DFLOAT(M(I,2))*PI2(I)
20 CONTINUE

CALCULATE THE MAXIMUM ABSOLUTE DEVIATION BETWEEN THE ESTIMATES OF
THIS ITERATION AND THOSE OF THE PREVIOUS ITERATION. IF THE
DEVIATION IS SMALL ENOUGH OR THE NUMBER OF ITERATIONS GETS TOO LARGE,
STOP ITERATING.

MAX=0.000
DO 25 I=1,S
  PI2(I)=PI2(I)/SCALE
  DEV=PI2(I)-PI1(I)
  DEV=DABS(DEV)
  MAX=DMAX1(DEV,MAX)
  PI1(I)=PI2(I)
25 CONTINUE
IT=IT+1
IREM=MOD(IT,10)
IF (IREM.EQ.0) WRITE(6,2001) IT,(PI1(I),I=1,S),MAX
2001 FORMAT(1X,I3,11(1X,D10.4))
IF ((IT.LT.500).AND.(MAX.GT.1.0D-05)) GO TO 5
IF (IREM.NE.0) WRITE(6,2001) IT,(PI1(I),I=1,S),MAX

CALCULATION OF THE METHOD OF MOMENT ESTIMATOR FOR THE JUDGE
VARIABILITY PARAMETER BASED ON THE SECOND MOMENTS OF THE COMPARISONS
UNDER THE DIRICHLET-BINOMIAL MODEL. ESTIMATES ARE TRUNCATED TO FALL
C BETWEEN ZERO AND CLIMIT. UNWEIGHTED AND WEIGHTED AVERAGES OF THE  
C VARIOUS ESTIMATES FROM THE VARIOUS SAMPLE MOMENTS ARE CALCULATED AND  
C USED AS THE FINAL ESTIMATES.

WRITE(6,2002)  
2002 FORMAT('CALCULATION OF THE METHOD OF MOMENTS ESTIMATOR FOR THE ',  
+ 'JUDGE VARIABILITY PARAMETER')

ILO=0  
IUP=0  
D=0  
DW=0  
CT=0.0D00  
CWT=0.0D00  
DC=0  
DWC=0  
CCT=0.0D00  
CCWT=0.0D00  
DO 35 I1=1,NI  
  C=0.0D00  
  DO 30 I2=1,I1  
    IF (N(I1,I2).LE.1) GO TO 30  
    IF (I1.EQ.I2) GO TO 30  
    B=DFLOAT(X2(I1,I2))*(PI1(H(I1))+PI1(H(I2)))  
    B=B/(DFLOAT(N(I1,I2))*PI1(H(I1))*DFLOAT(NJ))  
    B=(B**1.0D0)/(DFLOAT(N(I1,I2))*(1.0D0))  
    C=(1.0D00-B)/(((B**PI1(H(I2)))+(B**1.0D00)*PI1(H(I1))))  
    IF (C.GT.CLIMIT) GO TO 27  
    ILO=ILO+1  
    GO TO 29  
27 IF (C.LT.CLIMIT) GO TO 28  
  IUP=IUP+1  
28 CT=CT+C  
  CWT=CWT+DFLOAT(N(I1,I2))**C  
  CCWT=CCWT+(DFLOAT(N(I1,I2))**2.0D00)*C  
  D=D+1  
  DW=DW+N(I1,I2)  
  DWC=DWC+(N(I1,I2))**2  
30 CONTINUE  
35 CONTINUE  
  IF (D.EQ.0) GO TO 40  
  CT=CT/DFLOAT(D)  
  CWT=CWT/DFLOAT(DW)  
WRITE(6,2003) CT,CWT,D,ILO,IUP,CLIMIT  
2003 FORMAT(' JUDGE VARIABILITY ESTIMATOR (BASED ON VARIANCES): ',  
+ 'D14.7,' (UNWEIGHTED AVERAGE)'/51X,D14.7,  
+ ' (WEIGHTED AVERAGE)'/9X,1X,I3,  
+ ' AVERAGE BASED ON',1X,I3,  
+ ' VALUES OF WHICH ',I3,' WERE TRUNCATED AT 0.0D00 AND ',  
+ I3,' WERE TRUNCATED AT ',D10.3)
C CALCULATION OF THE METHOD OF MOMENT ESTIMATOR FOR THE JUDGE
C VARIABILITY PARAMETER BASED ON THE CROSS PRODUCT MOMENTS OF THE
C COMPARISONS UNDER THE DIRICHLET-BINOMIAL MODEL. ESTIMATES ARE
C TRUNCATED TO FALL BETWEEN ZERO AND CLIMIT. UNWEIGHTED AND WEIGHTED
C AVERAGES OF THE VARIOUS ESTIMATES FROM THE VARIOUS SAMPLE MOMENTS ARE
C CALCULATED AND USED AS THE FINAL ESTIMATES. ESTIMATION IS DONE BY
C FINDING TWO ESTIMATES THAT BRACKET THE TRUE ESTIMATE AND USING
C REGULA FALSI TO COME UP WITH THE TRUE ESTIMATE.

DO 200 I2=1,NI
DO 200 I1=1,NI
   IF (I1.EQ.I2) GO TO 200
   IF (I1.LT.I2) GO TO 199
   N1=N(I1,I2)
   IF (N1.LE.0) GO TO 199
   J=1
   DO 150 J2=1,NI
      DO 150 J1=1,NI
         IF (J1.EQ.J2) GO TO 150
         IF (J1.LT.J2) GO TO 199
         N2=N(J1,J2)
         IF (N2.LE.0) GO TO 149
         STEP=1.0D0
         IX=XX(I,J)
         XN=DFLOAT(N1)*DFLOAT(N2)
         IF (.NOT.((I1.EQ.I1).AND.(I2.NE.J2))) GO TO 50
            P1=PI1(H(I1))
            P2=PI1(H(I2))
            P3=PI1(H(J2))
            GO TO 80
      50   IF (.NOT.((I1.EQ.I2).AND.(I2.NE.I1))) GO TO 60
         P1=PI1(H(I1))
         P2=PI1(H(J1))
         P3=PI1(H(I2))
         IX=82*X1(I1,I2)-IX
         GO TO 80
60 IF (.NOT.((I2.EQ.J1).AND.(I1.NE.J2))) GO TO 70
  P1=PI1(H(J1))
  P2=PI1(H(I1))
  P3=PI1(H(J2))
  IX=N1*X1(J1,J2)^IX
  GO TO 80
70 IF (.NOT.((I2.EQ.J2).AND.(I1.NE.J1))) GO TO 149
  P1=PI1(H(I2))
  P2=PI1(H(I1))
  P3=PI1(H(J1))
  IX=IX-N1*N2+N1*X1(J1,J2)+N2*X1(I1,I2)

C
C FIND THE INITIAL VALUE CR AND FIND CL SUCH THAT FC(CR) AND FC(CL)
C HAVE DIFFERENT SIGNS.
80 DXX=DFLOAT(IX)/DFLOAT(NJ)
  CR=DXX*(P1+P2+P3)/(XN*P1)
  CR.=CR^1.0D00/(CR*(P1+P2)+P1)
  IF (CR.LT.STEP) CR.=STEP
  IF (CR.GT.CLIMIT) CR.=CLIMIT
  FR=FC(CR)
  IF (FR.LE.0.0D00) GO TO 90

C
C FIND CL SUCH THAT FC(CL)>0.
  CL=CR+STEP
  FL=FC(CL)

C
C IF THE FUNCTION FC IS DECREASING, STEP THE OTHER WAY.
  IF (FR.GE.FL) GO TO 85
    STEP.=STEP
    CL=CR+STEP
    FL=FC(CL)

C
C ITERATIVELY STEP UNTIL ZERO IS BRACKETED OR A BOUNDARY IS REACHED.
85 IF ((CL.LT.STEP).OR.(CL.GT.CLIMIT)) GO TO 110
  IF (FL.LT.0.0D00) GO TO 100
  CL=CL+STEP
  FL=FC(CL)
  GO TO 85
90 IF (FR.EQ.0.0D00) GO TO 110

C
C FIND CL SUCH THAT FC(CL)<0.
  CL=CR+STEP
  FL=FC(CL)

C
C IF THE FUNCTION FC IS INCREASING, STEP THE OTHER WAY.
  IF (FL.LE.FR) GO TO 95
    STEP.=STEP
    CL=CR+STEP
    FL=FC(CL)
C ITERATIVELY STEP UNTIL ZERO IS BRACKETED OR A BOUNDARY IS REACHED.
95 IF ((CL.LT.STEP).OR.(CL.GT.CLIMIT)) GO TO 110
   IF (FL.GT.0.000) GO TO 100
   CL=CL+STEP
   FL=FC(CL)
   GO TO 95
C
C IF CL>CR, SWAP
100 IF (CL.LT.CR) GO TO 105
    STEP=CR
    CR=CL
    CL=STEP
105 EPS=1.0D-10
    IDGT=8
    LOOP=500

C CALL THE IMSL ROUTINE TO PERFORM REGULA FALSI.
   CALL ZFALSE(FC,EPS,IDGT,CL,CR,CM,LOOP,IER)
   GO TO 115
110 CM=CL
   STEP=DABS(STEP)
   IF (CM.GE.STEP) GO TO 114
   ILO=ILO+1
   CM=STEP
114 IF (CM.LT.CLIMIT) GO TO 115
   IUP=IUP+1
   CM=CLIMIT
115 DC=DC+1
   DWC=DWC+N1*N2
   CCT=CCT+CM
   CCWT=CCWT+XN*CM
149 J=J+1
150 CONTINUE
199 I=I+1
200 CONTINUE
   CCT=CCT/DFLOAT(DC)
   CCWT=CCWT/DFLOAT(DWC)
   WRITE(6,2004) CCT,CCWT,DC,ILO,STEP,IUP,CLIMIT
2004 FORMAT('JUDGE VARIABILITY ESTIMATOR (BASED ON COVARIANCES): ',
       + 'D14.7,' (UNWEIGHTED AVERAGE)'/53X,D14.7,
       + ' (WEIGHTED AVERAGE)'/ AVERAGE BASED ON',1X,I3,
       + ' VALUES OF WHICH ',I3,' WERE TRUNCATED AT ',D10.3,
       + ' AND ',I3,' WERE TRUNCATED AT ',D10.3)
C
C COMBINE THE ESTIMATES BASED ON THE SECOND MOMENTS AND THE CROSS
C PRODUCT MOMENTS TO COME UP WITH THE FINAL ESTIMATE.
   IF (D.EQ.0) RETURN
   CCT=CCT*DFLOAT(DC)
CCWT = CCWT \* DFLOAT(DWC)
D = D + DC
DW = DW + DWC
CT = (CT + CCT) / DFLOAT(D)
CWT = (CWT + CCWT) / DFLOAT(DW)
WRITE(6,2005) CT, CWT, D

2005 FORMAT('OJUDGE VARIABILITY ESTIMATOR (COMBINED): ', D14.7,
+ ' (UNWEIGHTED AVERAGE)'/41X, D14.7,' (WEIGHTED AVERAGE)'/
+ ' AVERAGE BASED ON ', I3, ' VALUES')
RETURN
END

C

SUBROUTINE PMLE(DIM, DC2, NI, NJ, S, H, M, N, X, PI, C, P0, M2, VO, M4)

C

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C CALCULATE THE PSEUDO-MLES FOR THE SCALE VALUES AND THE ASSOCIATED C
C VARIANCES AND COVARIANCES.
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
C

INTEGER DIM, S, DC2, N(DIM, DIM), X(DIM, DIM), H(DIM), M(DIM, 2)
REAL*8 P(90), PO(DC2), VO(90), A(90, 10), XM, PX, MAX, PI(DIM), C, B(10),
+ ATA(10, 10), IATA(10, 10), WK(130), M1(90, 90), M2(DC2, DC2),
+ M3(90, 10), M4(90, 90), VC, DFLOAT, DSQRT, DABS, DMAX1
LOGICAL STOPIT
C

STOPIT = .FALSE.
IS = S - 1
NIC2 = NI \* (NI - 1)

C

ITERATIVELY CALCULATE THE PSEUDO-MLES. FOR THE CASE OF A SINGLE SCALE
C VALUE TO BE ESTIMATED, SKIP THE ITERATIVE CALCULATIONS AND CALCULATE
C THE VALUES FOR THE CHOICE PROBABILITY VECTOR, PROJECTION MATRIX AND
C RELATED VALUES.
WRITE(6,2000) (PI(I), I = 1, S)

2000 FORMAT('IITERATIVE CALCULATION OF THE SCALE VALUE ESTIMATES'/
+ ' INT', 10(1X, D10.4))
IF (S.EQ.1) GO TO 26
IT = 0
5 I = 1
DO 15 J2 = 1, NI
DO 15 J1 = 1, NI
IF (J1 .EQ. J2) GO TO 15
P(I) = DFLOAT(X(J1, J2)) / DFLOAT(N(J1, J2)) * DFLOAT(NJ))
PO(I) = PI(H(J1)) / (PI(H(J1)) + PI(H(J2)))
VO(I) = PO(I) / DFLOAT(N(J1, J2))
PO(I) = (P(I) - PO(I)) / DSQRT(VO(I))
C THE MATRIX A BEING CALCULATED IS THE MATRIX OF DERIVATIVES OF THE
C CHOICE PROBABILITY VECTOR WITH RESPECT TO THE HYPOTHESIZED SCALE
C VALUES. THIS IS POST-MULTIPLIED BY THE INVERSE OF A DIAGONAL MATRIX
C WITH THE SQUARE ROOTS OF THE ELEMENTS OF VO DOWN THE DIAGONAL.

DO 10 I1=1,IS
  A(I,I1)=0.0D0
  IF (H(J1).EQ.H(J2)) GO TO 10
  PX=(PI(H(J1))+PI(H(J2)))**2
  XM=DFLOAT(M(I1,2))/DFLOAT(M(S,2))
  IF (M(I1,1).EQ.H(J1)) A(I,I1)=PI(H(J2))/PX
  IF (M(I1,1).EQ.H(J2)) A(I,I1)=PI(H(J1))/PX
  IF (M(S,1).EQ.H(J1)) A(I,I1)=A(I,I1)+XM*PI(H(J2))/PX
  IF (M(S,1).EQ.H(J2)) A(I,I1)=A(I,I1)+XM*PI(H(J1))/PX
  A(I,I1)=A(I,I1)/DSQRT(VO(I))
10 CONTINUE

I=I+1
15 CONTINUE
  CALL VMULFM(A,P0,NIC2,IS,1,DC2,DC2,P,DC2,IER)
  CALL VTPROF(A,NIC2,IS,DC2,P0)
  CALL VCVTSF(P0,IS,ATA,DIM)
  IDGT=0
  CALL LINV2F(ATA,IS,DIM,lATA,IDGT,WK,IER)
  CALL VMULFF(lATA,P,IS,IS,1,DIM,DC2,B,DIM,IER)
  IF (STOPIT) GO TO 25

C UPDATE THE SCALE VALUE ESTIMATES. IF THE MAXIMUM ABSOLUTE DEVIATION
C BETWEEN THESE NEW ESTIMATES AND THOSE OF THE PREVIOUS ITERATION IS
C SMALL ENOUGH OR THE NUMBER OF ITERATIONS IS TOO LARGE, STOP ITERATING.

PX=0.0D0
MAX=0.0D0
DO 20 I=1,IS
  IF (B(I).GT.PI(I)) GO TO 19
  B(I)=1.0D0-5*PI(I)
19 PI(I)=PI(I)+B(I)
  PX=PX+DFLOAT(M(I,2))*B(I)
  B(I)=DABS(B(I))
  MAX=DMAX1(MAX,B(I))
20 CONTINUE
  PI(S)=PI(S)+PX/DFLOAT(M(S,2))
  PX=DABS(PX)
  MAX=DMAX1(MAX,PX)
  IT=IT+1
  IREM=MOD(IT,10)
  IF (IREM.EQ.0) WRITE(6,2001) IT,(PI(I),I=1,S),MAX
2001 FORMAT(1X,I3,11(1X,D10.4))
  IF ((IT.GT.500).OR.(MAX.LT.1.0D-05)) STOPIT=.TRUE.
  GO TO 5

C

25 IF (IREM.NE.0) WRITE(6,2001) IT,(PI(I),I=1,S),MAX
  CALL VMULFF(lATA,A,IS,NIC2,IS,DC2,B,DC2,DC2,IER)
C
C CALCULATE THE CHOICE PROBABILITY VECTOR, THE PROJECTION MATRIX AND
C THE VARIANCE-COVARIANCE MATRIX FOR THE COMPARISONS UNDER THE
C DIRICHLET-BINOMIAL MODEL.

26 J=1
   DO 45 J2=1,NI
   DO 45 J1=1,NI
      IF (J1.EQ.J2) GO TO 45
      N2=N(J1,J2)
      PO(J)=PI(H(J1))/(PI(H(J2))+PI(H(J1)))
      VO(J)=PO(J)/DFLOAT(N2)
      PX=(DFLOAT(N2)+1.0D00)*(C*PI(H(J1))+1.0D00)
      PX=PX/(C*(PI(H(J1))+PI(H(J2)))+1.0D00)
      PX=1.0D00+PX*DFLOAT(N2)*PO(J)*PX
      M4(J,J)=DFLOAT(N2)*PO(J)*PX
   1=1
   DO 40 12=1,NI
   DO 40 11=1,NI
      IF (I1.EQ.I2) GO TO 40
      N1=N(I1,12)
      M2(I,J)=0.0D00
      IF (I.EQ.J) GO TO 35
      IF (I.LT.J) GO TO 30
      M4(I,J)=0.0D00
      IF ((J1.EQ.I1).AND.(J2.NE.I2))
      + M4(I,J)=VC(N1,N2,C,PI(H(J1)),PI(H(I2)),PI(H(J2)))
      IF ((J1.EQ.I2).AND.(J2.NE.I1))
      + M4(I,J)=VC(N1,N2,C,PI(H(J1)),PI(H(I1)),PI(H(J2)))
      IF ((J2.EQ.I1).AND.(J1.NE.I2))
      + M4(I,J)=VC(N1,N2,C,PI(H(I1)),PI(H(J1)),PI(H(I2)))
      IF ((J2.EQ.I2).AND.(J1.NE.I1))
      + M4(I,J)=VC(N1,N2,C,PI(H(J1)),PI(H(I2)),PI(H(J2)))
      IF ((J1.EQ.I2).AND.(J2.EQ.I1)) M4(I,J)=M4(J,J)
      GO TO 35
   30 M4(I,J)=M4(J,I)
   35 I=I+1
   40 CONTINUE
   J=J+1
   45 CONTINUE
   IF (S.EQ.I) RETURN

C CALCULATE THE VARIANCES AND COVARIANCES FOR THE SCALE VALUE ESTIMATES.
   DO 50 I=1,NIC2
   DO 50 J=1,NIC2
      M2(I,J)=M4(I,J)/(DSQRT(VO(I))*DSQRT(VO(J)))
   50 CONTINUE
   CALL VMULFF(M2,M1,NIC2,NIC2,IS,DC2,DC2,M3,DC2,IER)
   CALL VMULFM(M1,M3,NIC2,IS,IS,DC2,DC2,IATA,DIM,IER)
   WRITE(6,2002)
   2002 FORMAT('COVARIANCE-COVARIANCE MATRIX FOR THE SCALE VALUE ESTIMATO',
     + 'RS')
DO 55 I=1,S
   IATA(S,I)=0.0D00
55 CONTINUE
DO 65 I=1,S
   DO J=1,S
      IATA(I,J)=IATA(I,J)/DFLOAT(NJ)
      IATA(S,J)=IATA(S,J)/DFLOAT(M(I,2))/DFLOAT(M(S,2))*IATA(I,J)
   CONTINUE
65 CONTINUE
WRITE(6,2003) (IATA(I,J),J=1,1)
2003 FORMAT(1X,10(D11.1,2X))
65 CONTINUE
DO 70 J=1,S
   IATA(S,S)=IATA(S,S)/DFLOAT(M(J,2))/DFLOAT(M(S,2))*IATA(S,J)
70 CONTINUE
WRITE(6,2003) (IATA(S,I),I=1,S)
C
C CALCULATE THE PROJECTION MATRIX ASSOCIATED WITH THE HYPOTHEZIZED
C CHOICE PROBABILITY VECTOR.
CALL VMULFP(A,M1,NIC2,IS,NIC2,DC2,DC2,M2,DC2,IER)
CALL VMULFP(A,M1,NIC2,IS,NIC2,DC2,DC2,M2,DC2,IER)
RETURN
END
C
REAL FUNCTION FC8(C)
C
REAL*8 C,XIJ,NIJ,P1,P2,P3,X1,X2,X3,DFLOAT
C
COMMON /JUDGE/ XIJ,NIJ,P1,P2,P3
C
FC=0.0D00
X1=P1*(C*P1+1.0D00)/(P1+P2+P3)
X2=1.0D00
X3=C*(P1+P2)+1.0D00
FC=X2/X3
DO 5 N=1,200
   X2=X2*(C*P2+DFLOAT(N)+1.0D00)/(C*(P1+P2+P3)+DFLOAT(N))
   X3=X3+1.0D00
   X2=X2/X3
   IF (X2.LT.1.0D-10) GO TO 10
   FC=FC+X2
5 CONTINUE
10 FC=NIJ*X1*FC-XIJ
RETURN
REAL FUNCTION VC*8(N1,N2,C,P1,P2,P3)

THIS FUNCTION CALCULATES THE COVARIANCE BETWEEN TWO COMPARISONS FOR THE DIRICHLET-BINOMIAL MODEL.

REAL*8 C,P1,P2,P3,XN,X1,X2,X3,X4,DFLOAT

VC=0.0D00
XN=DFLOAT(N1)*DFLOAT(N2)
X1=P1*(C*P1+1.0D00)/(P1+P2+P3)
X3=1.0D00
X4=C*(P1+P2)+1.0D00
VC=X3/X4

DO 5 N=1,250
X3=X3*(C*P2+DFLOAT(N)+1.0D00)/(C*(P1+P2+P3)+DFLOAT(N))
X4=X4+1.0D00
X2=X3/X4
IF (X2.LT.1.0D-15) GO TO 10
VC=VC+X2
5 CONTINUE
10 X2=P1*P1/((P1+P2)*(P1+P3))
VC=XN*(X1*VC-X2)
RETURN
END
This appendix contains a listing of the data set used in Example 2. Observations in which not every comparison was performed by a judge were deleted before any calculations were done. Data pertinent to the computations are in columns 11-55. The judge ID number is in columns 1-4 and the course number of the course the judge was taking when the experiment was performed is in columns 5-7. In column 8 is the quarter or semester when the course was taken (1-fall, 2-winter, 3-spring, 4-summer) and in columns 9-10 is the year in which the course was taken. Columns 56 through 69 contain other information about the judge.

The data in columns 11 through 55 are the results of the 45 comparisons performed by a judge. An A indicates that the first item was preferred and a B indicates that the second item was preferred. The order in which the comparison results are listed and the order of presentation of the items within each comparison are as follows:

<table>
<thead>
<tr>
<th>Column</th>
<th>1st item</th>
<th>2nd item</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>organic vigor</td>
<td>democratic values</td>
</tr>
<tr>
<td>12</td>
<td>social competency</td>
<td>spiritual and moral values</td>
</tr>
<tr>
<td>13</td>
<td>cultural appreciation</td>
<td>neuromuscular skills</td>
</tr>
<tr>
<td>14</td>
<td>leisure time activities</td>
<td>emotional stability</td>
</tr>
<tr>
<td>15</td>
<td>self-realization</td>
<td>mental development</td>
</tr>
<tr>
<td>16</td>
<td>social competency</td>
<td>democratic values</td>
</tr>
<tr>
<td>17</td>
<td>spiritual and moral values</td>
<td>cultural appreciation</td>
</tr>
<tr>
<td>Column</td>
<td>1st item</td>
<td>2nd item</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>18</td>
<td>neuromuscular skills</td>
<td>leisure time activities</td>
</tr>
<tr>
<td>19</td>
<td>emotional stability</td>
<td>self-realization</td>
</tr>
<tr>
<td>20</td>
<td>mental development</td>
<td>organic vigor</td>
</tr>
<tr>
<td>21</td>
<td>organic vigor</td>
<td>social competency</td>
</tr>
<tr>
<td>22</td>
<td>cultural appreciation</td>
<td>democratic values</td>
</tr>
<tr>
<td>23</td>
<td>leisure time activities</td>
<td>spiritual and moral values</td>
</tr>
<tr>
<td>24</td>
<td>self-realization</td>
<td>neuromuscular skills</td>
</tr>
<tr>
<td>25</td>
<td>mental development</td>
<td>emotional stability</td>
</tr>
<tr>
<td>26</td>
<td>social competency</td>
<td>cultural appreciation</td>
</tr>
<tr>
<td>27</td>
<td>democratic values</td>
<td>leisure time activities</td>
</tr>
<tr>
<td>28</td>
<td>spiritual and moral values</td>
<td>self-realization</td>
</tr>
<tr>
<td>29</td>
<td>neuromuscular skills</td>
<td>mental development</td>
</tr>
<tr>
<td>30</td>
<td>emotional stability</td>
<td>organic vigor</td>
</tr>
<tr>
<td>31</td>
<td>organic vigor</td>
<td>cultural appreciation</td>
</tr>
<tr>
<td>32</td>
<td>leisure time activities</td>
<td>social competency</td>
</tr>
<tr>
<td>33</td>
<td>self-realization</td>
<td>democratic values</td>
</tr>
<tr>
<td>34</td>
<td>mental development</td>
<td>spiritual and moral values</td>
</tr>
<tr>
<td>35</td>
<td>emotional stability</td>
<td>neuromuscular skills</td>
</tr>
<tr>
<td>36</td>
<td>cultural appreciation</td>
<td>leisure time activities</td>
</tr>
<tr>
<td>37</td>
<td>social competency</td>
<td>self-realization</td>
</tr>
<tr>
<td>38</td>
<td>democratic values</td>
<td>mental development</td>
</tr>
<tr>
<td>39</td>
<td>spiritual and moral values</td>
<td>emotional stability</td>
</tr>
<tr>
<td>40</td>
<td>neuromuscular skills</td>
<td>organic vigor</td>
</tr>
<tr>
<td>41</td>
<td>organic vigor</td>
<td>leisure time activities</td>
</tr>
<tr>
<td>Column</td>
<td>1st Item</td>
<td>2nd Item</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>42</td>
<td>self-realization</td>
<td>cultural appreciation</td>
</tr>
<tr>
<td>43</td>
<td>mental development</td>
<td>social competency</td>
</tr>
<tr>
<td>44</td>
<td>emotional stability</td>
<td>democratic values</td>
</tr>
<tr>
<td>45</td>
<td>neuromuscular skills</td>
<td>spiritual and moral values</td>
</tr>
<tr>
<td>46</td>
<td>leisure time activities</td>
<td>self-realization</td>
</tr>
<tr>
<td>47</td>
<td>cultural appreciation</td>
<td>mental development</td>
</tr>
<tr>
<td>48</td>
<td>social competency</td>
<td>emotional stability</td>
</tr>
<tr>
<td>49</td>
<td>democratic values</td>
<td>neuromuscular skills</td>
</tr>
<tr>
<td>50</td>
<td>spiritual and moral values</td>
<td>organic vigor</td>
</tr>
<tr>
<td>51</td>
<td>organic vigor</td>
<td>self-realization</td>
</tr>
<tr>
<td>52</td>
<td>mental development</td>
<td>leisure time activities</td>
</tr>
<tr>
<td>53</td>
<td>emotional stability</td>
<td>cultural appreciation</td>
</tr>
<tr>
<td>54</td>
<td>neuromuscular skills</td>
<td>social competency</td>
</tr>
<tr>
<td>55</td>
<td>spiritual and moral values</td>
<td>democratic values</td>
</tr>
</tbody>
</table>