A RANKING EXPERIMENT WITH PAIRED COMPARISONS
AND A FACTORIAL DESIGN

by

Robert M. Abelson

Thesis submitted to the Graduate Faculty of the
Virginia Polytechnic Institute
in candidacy for the degree of

MASTER OF SCIENCE
in
STATISTICS

APPROVED: APPROVED:

Director of Graduate Studies Head of Department

Dean of the School of Applied Science and Business Administration Major Professor

September, 1952

Blacksburg, Virginia

TABLE OF CONTENTS

		page
I.	INTRODUCTION	4
	1.1 The Method of Paired Comparisons	4
	1.2 The Objectives of This Paper	5
II.	THE MATHEMATICAL MODEL	6
	2.1 The Bradley-Terry Model	6
	2.2 The Factorial Model	7
	2.3 The Objectives Defined in Terms of the Model	10
III.	THE LIKELIHOOD FUNCTION	12
	3.1 The Bradley-Terry Likelihood Function	12
	3.2 The Likelihood for the General Factorial	13
IV.	MAXIMUM LIKELIHOOD ESTIMATES	15
	4.1 Properties of Maximum Likelihood Estimates	15
	4.2 Estimation in the General Case	16
	4.3 The Factorial Case	17
	4.4 The Special Case of the 2 x 2 Factorial	19
V.	LIKELIHOOD RATIO TESTS	26
	5.1 General Properties of Likelihood Ratio Tests	26

			page
	5.2	The Test Statistics for the 2 x 2 Factorial	28
	5.3	Relationships Between the Various Tests	34
VI.	AN ILLUSTRATIVE EXAMPLE		37
VII.	SUMMARY AND DISCUSSION		40
	7.1	Summary	40
	7.2	Discussion	40
VIII.	ACKNOWLEDGEMENTS		42
IX.	BIBLIOGRAPHY		43
X.	VITA		45

I. INTRODUCTION

1.1 The Method of Paired Comparisons

The experimental design in which pairs of treatments are compared in a balanced system over all $\frac{1}{2}t(t-1)$ pairs of t treatments is known as the method of paired comparisons. The method is old and various methods of analysing such experiments have been proposed. The most recent procedure developed has been presented by Bradley and Terry[1] [2] . This method has the advantage of comparative simplicity both in the formulation of the model and in application to numerical data. The model is essentially a generalisation of binomial-type trials. Initial experimental work [3] indicates that the model adequately describes the experimental situations encountered in most of such experiments.

The Bradley-Terry method affords a means of estimating treatment ratings from ranked data and procedures for testing null hypotheses on the absence of treatment differences and on group consistency. Maximum likelihood estimators, which have many of the properties of good estimators, are used to obtain treatment ratings. Likelihood ratio tests are used for tests of the various hypotheses formulated.

---

1. Numbers in square brackets refer to the bibliography.

1.2 The Objectives of This Paper

The purpose of the present paper is to investigate the means of adapting the Bradley-Terry method to experiments in which the treatments are factorial combinations of various factors taken at specified levels such as concentrations. Our objective will be the same as that of the analysis of variance for a normal deviate; that is, we wish to make tests of overall treatment equality, on the effect of each factor individually, and on the various interactions between the factors.

Investigation shows that more work must be done before the method is easily applicable to treatments arranged in a factorial array of arbitrary size. However, a procedure is developed in detail for the 2 x 2 factorial case where four treatment combinations are developed using each of two factors available at two levels. In this case it is shown that in addition to the tests of treatment equality presented by Bradley and Terry, tests can be made on each factor and on the interaction between the two factors, in a manner analogous to that of the tests made in the analysis of variance.

II. THE MATHEMATICAL MODEL

2.1 The Bradley-Terry Model

Bradley and Terry considered the problem of t treatments $T_1, \ldots, T_t$ in an experiment involving paired comparisons. They assumed that these treatments have true ratings (or preferences), $\pi_1, \ldots, \pi_t$ on some particular subjective continuum. They required that every $\pi_i \geq 0$ and that $\sum_i \pi_i = 1$. It was further assumed that when treatment T_i appears with treatment T_j in a block, where each block contains two treatments, the probability that treatment T_i obtains top rating (or a rank of 1) is $\pi_i / (\pi_i + \pi_j)$.

In their notation, r_{ijk} designates the rank of the i^{th} treatment in the k^{th} repetition of the block in which treatment T_i appears with treatment T_j. $r_{ijk} = 3 - r_{jik}$, for the treatment judged the better receives a score of one, the other treatment receives a score of two and ties are excluded. Estimates of π_i are denoted by p_i, and n denotes the number of repetitions of the design where a repetition is defined as a single set of judgements on all pairs of treatments.

It is within the framework of the Bradley-Terry model that we wish to consider the introduction of treatments in factorial array.

2.2 The Factorial Model

Let us now consider the case where we wish to study several different factors in a single experiment. If we have factors A, B, ..., D with a, b, ..., d levels respectively, we have an experiment with $t = ab \ldots d$ different treatments. For example, we might have an experiment involving a preservative chemical which we will test at three levels, A_1, A_2 and A_3, and two different cooking temperatures, B_1 and B_2. We will then have six treatment combinations A_1B_1, A_1B_2, A_2B_1, A_2B_2, A_3B_1, A_3B_2.

In the general factorial case, we assume that the levels of each factor have true rating parameters whose values are $\alpha_1, \ldots, \alpha_a, \beta_1, \ldots, \beta_b, \ldots, \delta_1, \ldots, \delta_d$ respectively, and that $\alpha_i, \beta_j, \ldots, \delta_l \geq 0$ for all $1 \leq i \leq a$, $1 \leq j \leq b$, ..., $1 \leq l \leq d$ and $\sum_{i=1}^{a} \alpha_i = 1$, $\sum_{j=1}^{b} \beta_j = 1, \ldots, \sum_{l=1}^{d} \delta_l = 1$. Furthermore, we will hypothesize that the treatment combination rating is

$$\pi_{ij\ldots l} = \alpha_i \beta_j \ldots \delta_l. \tag{1}$$

This hypothesis may be intuitively justified if we return to the Bradley-Terry model and compare it with the model presented by Thurstone [10]. In Thurstone's model, if S_i and S_j are the true treatment ratings of treatments i and j respectively on an additive subjective continuum and if P_{ij} is the probability that T_i be rated above T_j, then*

$$P_{ij} = \frac{1}{\sqrt{2\pi}} \int_{-(S_i - S_j)}^{\infty} e^{-y^2/2}\, dy. \tag{2}$$

If we redefine [1]

$$P_{ij} = 1/4 \int_{-(\log_e \pi_i - \log_e \pi_j)}^{\infty} \mathrm{Sech}^2\, y/2\, dy \tag{3}$$

then clearly $P_{ij} = \pi_i/(\pi_i + \pi_j)$.

Since the squared hyperbolic secant density is very similar to the normal density, values of $\log_e \pi_i$ correspond closely to values of S_i on the continuum postulated by Thurstone, where estimates of S_i are assumed to be normally distributed. Now in the analysis of variance of

---

*This summary of the model of Thurstone is most clearly presented by Mosteller [8].

factorial experiments a linear additive model is postulated for the treatment-factor effects and their interactions. Therefore it would seem that the analogous postulation here is that $\log_e \pi_{ij\ldots l} = \log_e \alpha_i + \log_e \beta_j + \ldots + \log_e \delta_l$ or that (1) is the logical formulation of the factorial model within the general model for paired comparisons.

Following the notation of the previous section, we will let $r_{ij\ldots l;pq\ldots r;k}$ designate the rank of the treatment combination $T_{ij\ldots l}$ in the k^{th} repetition of the block in which treatment $T_{ij\ldots l}$ appears with treatment $T_{pq\ldots r}$. As in the previous case, the scoring method requires that $r_{ij\ldots l;pq\ldots r;k} = 3 - r_{pq\ldots r;ij\ldots l;k}$.

In the present paper we shall consider in detail the two-factor two-levels-per-factor treatment combinations with factors A and B and levels designated by A_1, A_2, B_1, B_2, which are assumed to have true rating-parameters α_1, α_2, β_1, β_2, respectively under the restrictions that $\alpha_i, \beta_j \geq 0$, $i, j = 1, 2$, and $\alpha_1 + \alpha_2 = 1$, $\beta_1 + \beta_2 = 1$. In this case we have four treatment combinations, denoted by T_{11}, T_{12}, T_{21}, and T_{22}, which have ratings π_{11}, π_{12}, π_{21}, and π_{22} in the general model. We hypothesize that

$\pi_{ij} = \alpha_i \beta_j$. Estimates of α_i and β_j will be denoted by a_i and b_j respectively, and as before, n will denote the number of repetitions of the design.

2.3 The Objectives Defined in Terms of the Model

We are now able to specify those test comparisons that we wish to make. We shall use H_o to indicate the null or more restrictive hypothesis and H_a to indicate the class of alternatives allowed.

TEST I:

(4) H_o: $\pi_{ij} = 1/4$ (i, j = 1, 2),

(5) H_a: $\pi_{ij} \neq \pi_{pq}$ for some i, j, p, q.

This is a test of overall treatment equality against the general alternative.

TEST II A:

(6) H_o: $\alpha_1 = \alpha_2 = \frac{1}{2}$; $\beta_1, \beta_2 \geq 0$; $\beta_1 + \beta_2 = 1$,

(7) H_a: $\alpha_1 \neq \alpha_2$; $\beta_1, \beta_2 \geq 0$; $\beta_1 + \beta_2 = 1$.

This may be regarded as a test on A-effects independent of the existence of a B-effect.

TEST II B:

(8) H_o: $\beta_1 = \beta_2 = \frac{1}{2}$; $\alpha_1, \alpha_2 \geq 0$; $\alpha_1 + \alpha_2 = 1$,

(9) H_a: $\beta_1 \neq \beta_2$, $\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 = 1$.

This test on B-effects corresponds to Test II A.

TEST III:

(10) H_o: $\pi_{ij} = \alpha_i\beta_j$ $(i, j = 1, 2)$,

(11) H_a: $\pi_{ij} \neq \alpha_i\beta_j$ for some i, j.

This is a test of interaction, and as in the analysis of variance, the hypothesis of an additive model, namely H_o: $\log_e\pi_{ij} = \log_e\alpha_i + \log_e\beta_j$, is made against the general alternative hypothesis.

In addition, for reasons which will become apparent in Chapter 5, we also wish to consider the following tests.

TEST IV:

(12) H_o: $\alpha_1 = \alpha_2$; $\beta_1 = \beta_2$,

(13) H_a: Either $\alpha_1 \neq \alpha_2$ or $\beta_1 \neq \beta_2$ or both.

TEST V:

(12) H_o: $\alpha_1 = \alpha_2$; $\beta_1 = \beta_2$,

(14) H_a: $\alpha_1 = \alpha_2$; $\beta_1, \beta_2 \geq 0$; $\beta_1 + \beta_2 = 1$.

and

TEST VI:

(12) H_o: $\alpha_1 = \alpha_2$; $\beta_1 = \beta_2$,

(15) H_a: $\beta_1 = \beta_2$; $\alpha_1, \alpha_2 \geq 0$; $\alpha_1 + \alpha_2 = 1$.

III. THE LIKELIHOOD FUNCTION

3.1 The Bradley-Terry Likelihood Function

Consider the probability of the observed rankings in the k^{th} repetition of the block in which treatments T_i and T_j are compared. The probability of obtaining the observed result can be written [2]

$$\left(\frac{\pi_i}{\pi_i + \pi_j}\right)^{2-r_{ijk}} \left(\frac{\pi_j}{\pi_i + \pi_j}\right)^{2-r_{jik}} = \frac{\pi_i^{2-r_{ijk}}\, \pi_j^{2-r_{jik}}}{\pi_i + \pi_j},$$

for if T_i obtains top ranking, $r_{ijk} = 1$ and $r_{jik} = 2$ and the expression becomes $\pi_i / (\pi_i + \pi_j)$; on the other hand, if T_j obtains top ranking, $r_{jik} = 1$ and $r_{ijk} = 2$ and the probability is $\pi_j / (\pi_i + \pi_j)$. By multiplying similar expressions for all comparisons within a repetition and for all n repetitions, we obtain the general form of the likelihood function, which is simply the probability of the observed sample regarded as a function of the parameters. If L denotes the likelihood function, then

$$L = \prod_i (\pi_i^{2n(t-1) - \sum_{j \neq i} \sum_k r_{ijk}} \prod_{i<j} (\pi_i + \pi_j)^{-n} \tag{16}$$

$$(i, j = 1, \ldots, t)\ (k = 1, \ldots, n).$$

3.2 The Likelihood for the General Factorial Model

For the case where each treatment, $T_{ij\ldots l}$, is the factorial combination of A_i, B_j, ..., D_l, this becomes

$$(17) \quad L = \prod_i \prod_j \cdots \prod_l \pi_{ij\ldots l}^{\,2n(ab\ldots d-1) - \sum_k \sum_{ij\ldots l \neq pq\ldots r} r_{ij\ldots l;pq\ldots r;k}}$$

$$\prod_{ij\ldots l \neq pq\ldots r} (\pi_{ij\ldots l} + \pi_{pq\ldots r})^{-n/2}$$

$$i,\ p = 1, \ldots, a \quad j,q = 1,\ldots,b \quad l,r = 1,\ldots,d$$

$$k = 1, \ldots, n$$

where the inequality $ij\ldots l \neq pq\ldots r$ requires that $ij\ldots l$ and $pq\ldots r$ be two distinct points in a factor space, wherein there are $ab\ldots d$ permissible points in a hyper-rectangular grid in the positive quadrant of a space with axes corresponding to the factor-types. Since this condition allows each pair to be counted twice, the exponent in the second product in (17) is $-n/2$ rather than $-n$ as in (16). Under the hypothesis of no interaction, viz., that $\pi_{ij\ldots l} = \alpha_i \beta_j \ldots \delta_l$,

$$(18) \quad L = \prod_i \prod_j \cdots \prod_l (\alpha_i \beta_j \ldots \delta_l)^{\,2n(ab\ldots d-1) - \sum_k \sum_{ij\ldots l \neq pq\ldots r} r_{ij\ldots l;pq\ldots r;k}}$$

$$\prod_{ij\ldots l \neq pq\ldots r} (\alpha_i \beta_j \ldots \delta_l + \alpha_p \beta_q \ldots \delta_r)^{-n/2}$$

In the case of the 2 x 2 factorial there are six comparisons to be made and the second product in (18) becomes

$$[(\alpha_1\beta_1 + \alpha_1\beta_2)(\alpha_1\beta_1 + \alpha_2\beta_1)(\alpha_1\beta_1 + \alpha_2\beta_2)(\alpha_1\beta_2 + \alpha_2\beta_1)$$

$$(\alpha_1\beta_2 + \alpha_2\beta_2)(\alpha_2\beta_1 + \alpha_2\beta_2)]^{-n} = \alpha_1^{-n}\alpha_2^{-n}\beta_1^{-n}\beta_2^{-n}\ [(\alpha_1\beta_1 + \alpha_2\beta_2)$$

$$(\alpha_1\beta_2 + \alpha_2\beta_1)]^{-n}\quad (\alpha_1 + \alpha_2)(\beta_1 + \beta_2)]^{-2n},$$

from which we see that (18) reduces to

$$\text{(19)}\quad L = \alpha_1^{c_1}\ \alpha_2^{c_2}\ \beta_1^{d_1}\ \beta_2^{d_2}\ [(\alpha_1\beta_1 + \alpha_2\beta_2)(\alpha_1\beta_2 + \alpha_2\beta_1)]^{-n}$$

$$(\alpha_1 + \alpha_1)(\beta_1 + \beta_2)]^{-2n},$$

where c_i is defined as

$$\text{(20)}\qquad c_i = 11n - \sum_k \sum_j \sum_p \sum_q r_{ijpqk} \quad (pq \neq ij)$$

and d_j is defined as

$$\text{(21)}\qquad d_j = 11n - \sum_k \sum_i \sum_p \sum_q r_{ijpqk}. \quad (pq \neq ij)$$

IV. MAXIMUM LIKELIHOOD ESTIMATES

4.1 Properties of Maximum Likelihood Estimators

Since the parameters of the various forms of the likelihood function (the π's, α's, β's, etc.) are unknown, we must obtain estimates of them. We will do this by a general method known as the method of maximum likelihood.

It can be shown that maximum likelihood estimators have many of the properties we desire in "good" estimators, that is, they are sufficient, if sufficient estimators exist, efficient and consistent. In general, however, they are not unbiassed [7].

The method is based on the following principle: If $\prod_{i=1}^{n} f(x_i, \theta_1, \ldots, \theta_k)$ is the probability of a random sample $x_1, \ldots, x_n$ from a population with parameters $\theta_1, \ldots, \theta_k$ then the maximum likelihood estimate of each θ_h, $(h=1,\ldots,k)$ is $\hat{\theta}_h$, if it exists, such that $\prod_{i=1}^{n} f(x_i, \hat{\theta}_1, \ldots, \hat{\theta}_k) \geq \prod_{i=1}^{n} f(x_i, \theta_1', \ldots, \theta_k')$ where θ_h' is any other possible value of θ_h. Therefore, in the usual case where the likelihood function has a maximum with slope zero, the procedure to be followed is clear. If the parameters are not linearly related, it is only necessary to take derivatives of the likelihood function with respect to each of the parameters, set these derivatives equal to zero and solve the resulting set of equations for the estimates. For the type of situation with which we are dealing, where linear restrictions exist, the method of Lagrange multipliers will be used for the restraints on the parameters. Before differentiating, however, there is one modification to be made. Since likelihood functions are products of functions of the parameters, it is considerably more convenient to maximize the logarithm of the likelihood function, which, of course, has its maximum at the same point as that of the function itself.

4.2 Estimation in the General Case

In the situation described by Bradley and Terry, where

(16) $$L = \prod_i \pi_i^{2n(t-1) - \sum_{j \neq i} \sum_k r_{ijk}} \prod_{i<j} (\pi_i + \pi_j)^{-n}$$

and

$$\Sigma \pi_i = 1,$$

let

(22) $$F = \log L - \lambda(\Sigma \pi_i - 1),$$

where λ is the Lagrange multiplier. Then

$$\frac{\partial F}{\partial \lambda} = \Sigma\pi_i - 1$$

and

$$\frac{\partial F}{\partial \pi_i} = \frac{2n(t-1) - \sum\limits_{j\neq i} \sum\limits_{k} r_{ijk}}{\pi_i} - \sum_{j\neq i} \frac{n}{\pi_i + \pi_j} - \lambda.$$

By setting these derivatives equal to zero and solving for the estimates of the π's, we see that $\lambda = 0$ and thus the estimating equations are

$$(23) \qquad \frac{2n(t-1) - \sum\limits_{j\neq i} \sum\limits_{k} r_{ijk}}{p_i} - \sum_{j\neq i} \frac{n}{p_i + p_j} = 0$$

and

$$(24) \qquad \Sigma p_i = 1.$$

Values of p_i are available in tables for three treatments for n up to 10 [2] , for four treatments for n up to eight [2,9], and for five treatments for n up to four [9]. For larger values of t and n the equations can be solved by an iterative procedure that has been described in [2].

4.3 The Factorial Case

In the case where the treatments are factorial combinations, to find estimates for α_i, β_j, ..., δ_l we must maximizes $F = \log L - \lambda(\Sigma\alpha_i - 1) - \mu(\Sigma\beta_j - 1) - \cdots - \eta(\Sigma\delta_l - 1)$, where L is defined by (18) and $\lambda, \mu, \ldots, \eta$ are Lagrange multipliers. This is done by differentiating with respect to $\lambda, \mu, \ldots, \eta$ and also with respect to each of the ab...d parameters. For example,

(25)
$$\frac{\partial F}{\partial \alpha_i} = \frac{\sum_{j=1}^{b} \cdots \sum_{l=1}^{d} \left[2n(ab\ldots d-1) - \sum_k \sum_{\substack{pq\ldots r \\ \neq ij\ldots l}} r_{ij\ldots l;pq\ldots r;k}\right]}{\alpha_i}$$

$$-2n \sum_{j=1}^{b} \cdots \sum_{l=1}^{d} \sum_{p=1}^{a} \cdots \sum_{r=1}^{d} {}_{pq\ldots r \neq ij\ldots l} \left[\frac{\beta_j \ldots \delta_l}{(\alpha_i\beta_j\ldots\delta_l + \alpha_p\beta_q\ldots\delta_r)}\right] - \lambda .$$

(26) $$\frac{\partial F}{\partial \lambda} = \sum_{i=1}^{a} \alpha_i - 1; \quad \frac{\partial F}{\partial \mu} \quad \sum_{j=1}^{b} \beta_j - 1; \ldots; \quad \frac{\partial F}{\partial \eta} = \sum_{l=1}^{d} \delta_l - 1.$$

As in the previous section the Lagrange multipliers λ, μ and η can be shown to be zero, so that we obtain as one of the set of estimating equations,

(27) $$a_i^{-1} \sum_j \cdots \sum_l \left[2n(ab\ldots d - 1) - \sum_k \sum_{\substack{pq\ldots r \\ \neq ij\ldots l}} r_{ij\ldots l;pq\ldots r;k}\right]$$

$$-2n \sum_{\substack{j=1 \\ pq\ldots r \neq ij\ldots l}}^{b} \cdots \sum_{l=1}^{d} \sum_{p=1}^{a} \cdots \sum_{r=1}^{d} \left[\frac{b_j \ldots d_l}{a_i b_j \ldots d_l + a_p b_q \ldots d_r}\right] = 0.$$

We also obtain, as restrictions on the estimates,

$$(28) \qquad \sum_{i=1}^{a} a_i = 1, \quad \sum_{j=1}^{b} b_j = 1, \ \ldots, \quad \sum_{l=1}^{d} d_l = 1.$$

Equations similar to (27) may be obtained by differentiating with respect to each of the other parameters. This is a rather formidable set of equations which it would in general be impractical to attempt to solve without the use of a high speed calculator. In the special case with which we are concerned, however, the equations are greatly simplified and can easily be solved with the use of a desk calculator.

4.4 The Special Case of the 2 x 2 Factorial

As we have already seen in Chapter III, for the 2 x 2 factorial without interaction, the likelihood function reduces to

$$(29) \quad L = \alpha_1^{c_1} \alpha_2^{c_2} \beta_1^{d_1} \beta_2^{d_2} \left[(\alpha_1\beta_1 + \alpha_2\beta_2)(\alpha_1\beta_2 + \alpha_2\beta_1)\right]^{-n}$$
$$\left[(\alpha_1 + \alpha_2)(\beta_1 + \beta_2)\right]^{-2n}.$$

$$(30) \quad F = \log_e L - \lambda(\alpha_1 + \alpha_2 - 1) - \mu(\beta_1 + \beta_2 - 1).$$

Differentiating with respect to α_1, α_2, β_1, β_2, λ and μ, substituting the estimates a_1, a_2, b_1, b_2 for their respective parameters and setting the derivatives equal to zero, we obtain

(31) $a_1 + a_2 = 1,$

(32) $b_1 + b_2 = 1,$

(33) $$\frac{c_1}{a_1} - \frac{2n}{a_1 + a_2} - \frac{nb_1}{a_1b_1 + a_2b_2} - \frac{nb_2}{a_1b_2 + a_2b_1} - \lambda = 0,$$

(34) $$\frac{c_2}{a_2} - \frac{2n}{a_1 + a_2} - \frac{nb_2}{a_1b_1 + a_2b_2} - \frac{nb_1}{a_1b_2 + a_2b_1} - \lambda = 0,$$

(35) $$\frac{d_1}{b_1} - \frac{2n}{b_1 + b_2} - \frac{na_1}{a_1b_1 + a_2b_2} - \frac{na_2}{a_1b_2 + a_2b_1} - \mu = 0,$$

(36) $$\frac{d_2}{b_2} - \frac{2n}{b_1 + b_2} \quad \frac{na_2}{a_1b_1 + a_2b_2} - \frac{na_1}{a_1b_2 + a_2b_1} - \mu = 0,$$

where a_i is the maximum likelihood estimate of α_i and b_j is the maximum likelihood estimate of β_j. If we multiply (33) by a_1 and (34) by a_2 and add, noting that (31), $a_1 + a_2 = 1$, the result is

(37) $c_1 + c_2 - 2n - n - n - \lambda = 0.$

Since $c_1 + c_2 = 4n$, (37) becomes $\lambda = 0$. By a similar process we can show that $\mu = 0$. Now, by substituting $a_2 = 1 - a_1$ and $b_2 = 1 - b_1$ into (33) and (35) we obtain two independent equations in two unknowns,

$$(38) \quad \frac{c_1}{a_1} - 2n - \frac{nb_1}{1-a_1-b_1+2a_1b_1} - \frac{n(1-b_1)}{a_1+b_1-2a_1b_1} = 0$$

and

$$(39) \quad \frac{d_1}{b_1} - 2n - \frac{na_1}{1-a_1-b_1+2a_1b_1} - \frac{n(1-a_1)}{a_1+b_1-2a_1b_1} = 0.$$

If we multiply (38) by $a_1(1 - a_1 - b_1 + 2a_1b_1)(a_1+b_1-2a_1b_1)$ and (39) by $b_1(1-a_1-b_1+2a_1b_1)(a_1+b_1-2a_1b_1)$ it can be seen that (38) is cubic in a_1 and quadratic in b_1, while (39) is quadratic in a_1 and cubic in b_1. A possible method of solution would be to solve (38) for b_1 algebraically, substitute the solution in (39) and solve for a_1. The solution of b_1, however, involves a radical, and when (39) is squared in order to rationalize it, and powers of a_1 are collected, it becomes a thirteenth degree equation in a_1, which, while tedious, could be solved numerically in any particular case. However, there is a simpler method available. If we multiply (38) by a_1 and (39) by b_1, and then take the sum and difference of the resulting equations, we obtain two new equations,

$$(40)\quad c_1 + d_1 - 2n(a_1 + b_1) - n - \frac{2na_1b_1}{1-a_1-b_1+2a_1b_1} = 0$$

and

$$(41)\quad c_1 - d_1 - 2n(a_1 - b_1) - \frac{n(a_1-b_1)}{a_1+b_1-2a_1b_1} = 0.$$

We transform by setting

$$(42)\qquad v = a_1 + b_1$$

and

$$(43)\qquad u = a_1 - b_1$$

and simplifying the algebra by writing

$$(44)\qquad x = \frac{c_1 + d_1 - n}{n}$$

and

$$(45)\qquad w = \frac{c_1 - d_1}{n}.$$

We note that

$$u^2 - v^2 = 4a_1b_1$$

and

$$a_1 + b_1 - 2ab = \frac{u^2 + 2v - v^2}{2}.$$

Equations (40) and (41) become

(46) $$v = 1 \pm \sqrt{u^2 + 2 - w/(w - 2u)}$$

and

(47) $$u^2 = v^2 - 2v + 3 + \frac{x - 3}{2v - x + 1} .$$

Elimination of u between (46) and (47) results in a 5th degree degree equation in v,

(48) $$\begin{aligned} &128v^5 - (160+160x)v^4 + (210+244x-18w^2+66x^2)v^3 \\ &+ (61 - 271x - 117x^2 - 9x^3 - 9w^2 + 21w^2x)v^2 \\ &+ (4 - 54x + 120x^2 + 18x^3 - 8w^2x^2 + 6w^2x)v \\ &+ (w^2x^3 - w^2x^2 + 12x^2 - 18x^3 - 2x) = 0. \end{aligned}$$

This equation can be solved easily by Horner's method after w and x have been replaced by their values calculated from the observed sums of ranks, using equations (20), (21), (44) and (45). We can then calculate the value of u, using (46) or (47) and thence a_1 and b_1 for $a_1 = (u + v)/2$ and $b_1 = (v - u)/2$. The values of a_2 and b_2 are then obtained from (31) and (32). For larger designs the equations rapidly become much more complex.

There is one more question which we must consider in connection with the 2 x 2 factorial; namely, what are the maximum likelihood estimates of α_i under the assumption that $\beta_1 = \beta_2$ and what are the maximum likelihood estimates of β_j under the assumption that $\alpha_1 = \alpha_2$. If we let $\beta_1 = \beta_2$, the likelihood function becomes

$$L = 2^{-2n}\alpha_1^{c_1}(1 - \alpha_1)^{c_2}$$

$$\frac{\partial \log_e L}{\partial \alpha_1} = \frac{c_1}{\alpha_1} - \frac{c_2}{1-\alpha_1},$$

from which

(49) $$\frac{c_1}{\hat{\alpha}_1} - \frac{c_2}{1 - \hat{\alpha}_1} = 0,$$

where $\hat{\alpha}_1$ is the maximum likelihood estimate of α_1 under the assumption that $\beta_1 = \beta_2$. Solving for $\hat{\alpha}_1$ and noting that $c_1 + c_2 = 4n$, we find that

(50) $$\hat{\alpha}_1 = \frac{c_1}{4n}.$$

Similarly it can be shown that

(51) $$\hat{\alpha}_2 = \frac{c_2}{4n},$$

(52) $$\hat{\beta}_1 = \frac{d_1}{4n}$$

and

(53) $$\hat{\beta}_2 = \frac{d_2}{4n} ,$$

where $\hat{\beta}_j$ is the maximum likelihood estimate of β_j under under $\alpha_1 = \alpha_2$. Since under these hypotheses we have binomial situations, we should expect that the estimates of the parameters would be the same as those derived from the binomial distribution. That this is so is easily seen by noting, for example, that the number of times a treatment containing A_1 is compared with a treatment containing A_2 is equal to 4n, and that the number of times the treatments containing A_1 obtain top rating is c_1.

V. LIKELIHOOD RATIO TESTS

5.1 General Properties of Likelihood Ratio Tests

There are many different tests which may be used on the various hypotheses which we will examine, but we will confine our attention to one type of test; the likelihood ratio. It can be shown [4] that this test is uniformly most powerful if such a test exists. This will be true if the hypotheses depend on only one parameter, and if a sufficient estimator exists for the parameter. If a uniformly most powerful test does not exist, the likelihood ratio test is usually a good compromise between the possible tests which might be used [6].

The test is as follows. Given a sample whose likelihood is

$$L = \prod_{i=1}^{n} f(x_i, \theta_1, \ldots, \theta_k),$$

we wish to test the hypothesis H_o that $(\theta_1, \ldots, \theta_k)$ belongs to the subspace ω of the parameter space Ω. If the parameters are allowed to vary freely over the entire space Ω, L will have a maximum in that space which will be denoted by $L(\hat{\Omega})$. Likewise L will have a maximum in ω which will be denoted by $L(\hat{\omega})$. Of course, if ω is a point, as is often the case, $L(\hat{\omega})$ is simply the value of

L at that point. It seems quite reasonable to use the ratio of the two maxima, $\lambda = \dfrac{L(\hat{\omega})}{L(\hat{\Omega})}$, as the test statistic. λ in this case should not be confused with the same symbol used to indicate a Lagrange multiplier. Necessarily $0 \leq \lambda \leq 1$, for, since L is a product of density functions, it is always positive and since ω is a subspace of Ω, $L(\hat{\omega})$ is less than, or at most equal to $L(\hat{\Omega})$. The critical region for testing H_o will be an interval of the form $0 \leq \lambda < A$, where $0 < A < 1$, since, obviously, larger values of λ make H_o more tenable. In order to specify A, it is necessary that we know the distribution of λ under H_o. In many cases this is difficult and, in fact, if H_o is not a simple hypothesis, a unique distribution for λ may not exist, that is, the distribution may depend on the unknown values of parameters not specified by H_o, and thus we cannot determine the value of A. For large samples, however, the difficulty is resolved by the use of the following theorem.[11].

> Suppose the c. d. f.* of the population depends on parameters $\theta_1, \theta_2, \ldots, \theta_h$, and that λ is the likelihood ratio for the hypothesis
>
> $$H_o: \theta_1 = \theta_1^o, \theta_2 = \theta_2^o, \ldots, \theta_m = \theta_m^o,$$

*Cumulative distribution function

> where $m \leq h$. Then under certain regularity conditions the asymptotic distribution of $-2 \log_e \lambda$, under the assumption that H_o is true, is the χ^2 distribution with m degrees of freedom.

The proof is given in Kendall[2].

5.2 The Test Statistics for the 2 x 2 Factorial

In the present problem there are several different tests we wish to make, involving various null hypotheses on α, β, and π, and various alternative hypotheses. These have been outlined in section 2.3.

TEST I:

(4) $$H_o: \pi_{ij} = 1/4 \quad (i,j = 1, 2),$$

(5) $$H_a: \pi_{ij} \neq \pi_{pq} \text{ for some } i, j, p, q.$$

On substituting 1/4 in the likelihood function we see that under H_o,

(54) $$L(\hat{\omega}_1) = 2^{-6n}$$

and under H_a, in view of (23) and (24),

(55) $$L(\hat{\Omega}_1) = \prod_{i=1}^{2} \prod_{j=1}^{2} p_{ij}^{6n-\sum_k \sum_{pq \neq ij} r_{ijpqk}} \prod_{pq \neq ij} (p_{ij}+p_{pq})^{-2n}$$

It now follows that

$$(56)\quad -2\log_e\lambda_1 = 12n\log_e 2 + 2\sum_{i=1}^{2}\sum_{j=1}^{2}\Big(6n - \sum_{k=1}^{n}\sum_{pq\neq ij} r_{ijpqk}\Big)\log_e p_{ij}$$
$$-4n\sum_{pq\neq ij}\log_e(p_{ij}+p_{pq})$$

This is the situation discussed by Bradley and Terry, for which the exact distribution of B_1, a monotonic function of λ_1, has been tabled for those values of t and n specified earlier. Since it is impractical, however, to compute tables for some of the other tests that we wish to make, it is necessary that we use larger values of n and resort to large-sample approximations. In this case, $-2\log_e\lambda_1$ is distributed approximately as χ^2 with three degrees of freedom. The use of this approximation will lead to the announcement of too many significant results. However, for n = 10 or more, it is not too serious if we maintain a reasonable degree of suspicion toward borderline significance.

TEST II A:

(6) H_o: $\alpha_1 = \alpha_2 = \frac{1}{2}$; $\beta_1 \geq 0$; $\beta_2 \geq 0$; $\beta_1 + \beta_2 = 1$,

(7) H_a: $\alpha_1 \neq \alpha_2$; $\beta_1 \geq 0$; $\beta_2 \geq 0$; $\beta_1 + \beta_2 = 1$.

In this second test we wish to test the A effect, that is, whether the effect of A_1 is significantly different from that of A_2 or not.

For $L(\hat{\omega})$ we substitute $\frac{1}{2}$, $\frac{1}{2}$, $\hat{\beta}_1$, $\hat{\beta}_2$ for α_1, α_2, β_1, β_2 in L, where $\hat{\beta}_1$ and $\hat{\beta}_2$ are obtained by the use of (52) and (53).

Then,

(57) $$L(\hat{\omega}_{2A}) = 2^{-10n}\, n^{-4n}\, d_1^{d_1}\, d_1^{d_2}.$$

For $L(\hat{\Omega}_2)$ we substitute a_1, a_2, b_1, b_2 in L and obtain the result $L(\hat{\Omega}_2) = a_1^{c_1}\, a_2^{c_2}\, b_1^{d_1}\, b_2^{d_2}[(a_1b_1+a_2b_2)(a_1b_2+a_2b_1)]^{-n}$.

The test statistic is

(58) $$-2\log_e\lambda_{2A} = 20n\log_e 2 + 8n\log_e n + 2c_1\log_e a_1 + 2c_2\log_e a_2$$
$$- 2d_1[\log_e d_1 - \log_e b_1] - 2d_2[\log_e d_2 - \log_e b_2]$$
$$- 2n\log_e(a_1b_1 + a_2b_2)(a_1b_2 + a_2b_1),$$

which is distributed approximately as χ^2 with one degree of freedom.

TEST II B:

Another test of the same type which we can make is on the B effect. Here, we have

(8) $$H_o: \beta_1 = \beta_2 = \tfrac{1}{2};\ \alpha_1, \alpha_2 \geq 0;\ \alpha_1 + \alpha_2 = 1,$$

(9) $$H: \beta_1 \neq \beta_2;\ \alpha_1, \alpha_2 \geq 0;\ \alpha_1 + \alpha_2 = 1.$$

For $L(\hat{\omega}_{2B})$ we substitute $\hat{\alpha}_1, \hat{\alpha}_2, \frac{1}{2}, \frac{1}{2}$ for $\alpha_1, \alpha_2, \beta_1, \beta_2$ from which

(59) $$L(\hat{\omega}_{2B}) = 2^{-10n}\, n^{-4n}\, c_1^{c_1}\, c_2^{c_2}.$$

$L(\hat{\Omega})$ remains in the form $L(\hat{\Omega}_2)$. In this case the test statistic becomes,

(60) $$-2 \log_e \lambda_{2B} = 20n \log_e 2 + 8n \log_e n - 2c_1 [\log_e c_1 - \log_e a_1]$$
$$-2c_2 [\log_e c_2 - \log_e a_2] + 2d_1 \log_e b_1 + 2d_2 \log_e b_2$$
$$- 2n \log_e (a_1 b_1 + a_2 b_2)(a_1 b_2 + a_2 b_1),$$

which is also distributed approximately as χ^2 with one degree of freedom.

<u>TEST III</u>:

(10) $$H_o: \pi_{ij} = \alpha_i \beta_j,$$

(11) $$H_a: \pi_{ij} \neq \alpha_i \beta_j \text{ for some } i, j.$$

A moment's reflection will show that $L(\hat{\omega}_3) \equiv L(\hat{\Omega}_2)$ and $L(\hat{\Omega}_3) \equiv L(\hat{\Omega}_1)$ from which

$$(61)\quad -2\log_e\lambda_3 = 4n \sum_{ij\neq pq} \log_e(p_{ij} + p_{pq})$$

$$+2n\log_e (a_1b_1 + a_2b_2)(a_1b_2 + a_2b_1)$$

$$-2\sum_{i=1}^{2}\sum_{j=1}^{2}(6n - \sum_{k=1}^{n}\sum_{\substack{pq\\ \neq ij}} r_{ijpqk})\log_e p_{ij}$$

$$-2c_1\log_e a_1 - 2c_2\log_e a_2$$

$$-2d_1\log_e b_1 - 2d_2\log_e b_2.$$

This test statistic, like the two previous ones, is distributed approximately as χ^2 with one degree of freedom. Clearly this is the test of the AB interaction. On its outcome will depend the question of whether we can assign overall true ratings to A_i and B_j or whether we must specify the level of B when giving ratings to A_i and specify the level of A when giving ratings to B_j.

TEST IV:

The fourth type of test which is of some interest is

(12) $\quad H_o$: $\alpha_1 = \alpha_2$; $\beta_1 = \beta_2$,

(13) $\quad H_a$: Either $\alpha_1 \neq \alpha_2$ or $\beta_1 \neq \beta_2$.

Here we see that $L(\hat{\omega}_4) \equiv L(\hat{\omega}_1)$ and $L(\hat{\Omega}_5) = L(\hat{\Omega}_2)$, so that

$$(62) \quad -2\log_e\lambda_4 = 12n\log_e 2 + 2c_1\log_e a_1 + 2c_2\log_e a_2 + 2d_1\log_e b_1 + 2d_2\log_e b_2 - 2n\log_e (a_1b_1 + a_2b_2)(a_1b_2 + a_2b_1) .$$

This is distributed approximately as χ^2 with two degrees of freedom.

<u>TEST V</u>:

(12) H_o: $\alpha_1 = \alpha_2$; $\beta_1 = \beta_2$,

(14) H_a: $\alpha_1 = \alpha_2$; β_1, $\beta_2 \geq 0$; $\beta_1 + \beta_2 = 1$.

In this case $L(\hat{\omega}_5) \equiv L(\hat{\omega}_4)$ and $L(\hat{\Omega}_5) = L(\hat{\omega}_{2A})$.

$$(63) \quad -2\log_e\lambda_5 = -8n\log_e 2 - 8n\log_e n + 2d_1\log_e d_1 + 2d_2\log_e d_2 .$$

<u>TEST VI</u>:

(12) H_o: $\alpha_1 = \alpha_2$; $\beta_1 = \beta_2$,

(15) H_a: $\beta_1 = \beta_2$; α_1, $\alpha_2 \geq 0$; $\alpha_1 + \alpha_2 = 1$.

Here $L(\hat{\omega}_6) \equiv L(\hat{\omega}_4)$ and $L(\hat{\Omega}_6) \equiv L(\hat{\omega}_{2B})$ so that

$$(64) \quad -2\log_e\lambda_6 = -8n\log_e 2 - 8n\log_e n + 2c_1\log_e c_1 + 2c_2\log_e c_2 .$$

These last two statistics are distributed approximately as χ^2 with one degree of freedom.

5.3 Relationships Between the Various Tests

An examination of the following tables reveals the relationships between the tests.

Table I. The A Effect

Test Statistic	Equation Number	Hypotheses	Distribution
$-2\log_e\lambda_{2A}$	(58)	Ho: $\alpha_1=\alpha_2$; $\beta_1,\beta_2\geq 0$; $\beta_1+\beta_2=1$	χ^2_1
		H_a: $\alpha_1\neq\alpha_2$; $\beta_1,\beta_2\geq 0$; $\beta_1+\beta_2=1$	
$-2\log_e\lambda_5$	(63)	H_o: $\alpha_1=\alpha_2$; $\beta_1=\beta_2$	χ^2_1
		H_a: $\alpha_1=\alpha_2$; $\beta_1,\beta_2\geq 0$; $\beta_1+\beta_2=1$	
$-2\log_e\lambda_4$	(62)	H_o: $\alpha_1=\alpha_2$; $\beta_1=\beta_2$	χ^2_2
		H_a: Either $\alpha_1\neq\alpha_2$ or $\beta_1\neq\beta_2$ or both	

Table II. The B Effect

Test Statistic	Equation Number	Hypotheses	Distribution
$-2\ \log_e \lambda_{2B}$	(60)	H_o: $\beta_1 = \beta_2$; $\alpha_1, \alpha_2 \geq 0$; $\alpha_1 + \alpha_2 = 1$ H_a: $\beta_1 \neq \beta_2$; $\alpha_1, \alpha_2 \geq 0$; $\alpha_1 + \alpha_2 = 1$	χ^2_1
$-2\ \log_e \lambda_6$	(64)	H_o: $\alpha_1 = \alpha_2$; $\beta_1 = \beta_2$ H_a: $\beta_1 = \beta_2$; $\alpha_1, \alpha_2 \geq 0$; $\alpha_1 + \alpha_2 = 1$	χ^2_1
$-2\ \log_e \lambda_4$	(62)	H_o: $\alpha_1 = \alpha_2$;, $\beta_1 = \beta_2$ H_a: Either $\alpha_1 \neq \alpha_2$ or $\beta_1 \neq \beta_2$ or both	χ^2_2

Table III. Interaction

Test Statistic	Equation Number	Hypotheses	Distribution
$-2\ \log_e \lambda_4$	(62)	H_o: $\alpha_1 = \alpha_2$; $\beta_1 = \beta_2$ H_a: Either $\alpha_1 = \alpha_2$ or $\beta_1 = \beta_2$ or both	χ^2_1
$-2\ \log_e \lambda_3$	(61)	H_o: $\pi_{ij} = \alpha_i \beta_j$ H_a: $\pi_{ij} \neq \alpha_i \beta_j$ for some i, j	χ^2_1
$-2\ \log_e \lambda_1$	(56)	H_o: $\pi_{ij} = \pi_{pq}$ H_a: $\pi_{ij} \neq \pi_{pq}$ for some i,j,p,q	χ^2_3

We see that the two degrees of freedom associated with $-2 \log_e \lambda_4$ may be partitioned in two ways, one yielding a test of the A-effect and the other the B-effect. In turn the three degrees of freedom of the general test may be partitioned into an interaction effect with one degree of freedom and $- 2 \log_e \lambda_4$ with two degrees of freedom. Tests II A and II B are not orthogonal but this is of no consequence as we do not intend to compare them with each other.

In this chapter three types of tests, the test of overall treatment equality, the tests of the A and B effects, and the test of the AB interaction, have been developed for practical use. In addition, three other tests have been developed in order to exhibit the nature of the method and these may sometimes be of supplementary interest.

VI. An Illustrative Example

In order to illustrate the method for 2 x 2 factorially arranged treatments in the method of paired comparisons, let us consider a hypothetical example. Suppose that in an experiment with four treatment combinations, A_1B_1, A_1B_2, A_2B_1, A_2B_2, in which all comparisons of pairs of treatments were repeated ten times, $(n = 10)$, the sums of ranks were observed to be $\Sigma r_{11} = 38$, $\Sigma r_{12} = 44$, $\Sigma r_{21} = 48$, $\Sigma r_{22} = 50$. We wish to test overall treatment differences, the effect of factor A, the effect of factor B, and the AB interaction.

To find overall treatment ratings, we can refer to Appendix A of Bradley and Terry [2]. If we divide the observed sums of ranks by two we can enter the table for $t = 4$, $n = 5$, at 19, 22, 24, 25, where we find that the estimates of the treatment ratings are $p_{11} = .47$, $p_{12} = .24$, $p_{21} = .16$, $p_{22} = .13$. These estimates can be calculated more precisely by an iterative procedure making repeated use of the formula [2]:

$$p_i^{(1)} = \left[2n(t-1) - \Sigma r_i\right]$$

$$\left[\frac{n}{p_1^{(1)} + p_i^{(0)}} + \dots + \frac{n}{p_{i-1}^{(1)} + p_i^{(0)}} + \frac{n}{p_i^{(0)} + p_{i+1}^{(0)}} + \dots + \frac{n}{p_i^{(0)} + p_t^{(0)}}\right]^{-1},$$

where the superscript in parentheses indicates the order of iteration. Thus, $p_{11} = .4732$, $p_{12} = .2421$, $p_{21} = .1580$, $p_{22} = .1267$.

To find a_i and b_j, we substitute in (20) and (21) to obtain $c_1 = 28$, $c_2 = 12$, $d_1 = 24$, $d_2 = 16$ from which, by (44) and (45), $x = 4.2$ and $w = 0.4$. Equation (48) then becomes

$$128v^5 - 832v^4 + 2396.16v^3 - 3795.2v^2$$
$$+ 3209.0368v - 1121.27232 = 0,$$

which can be solved by Horner's method, yielding $v = 1.31275$. Then, by equation (47), $u = .095133$, and, since $a_1 = (v+u)/2$ and $b_1 = (v-u)/2$, we find that $a_2 = .70394$, $a_2 = .29606$, $b_1 = .60881$, $b_2 = .39119$. Substituting these values in the test statistics (logarithms to the base e of numbers between zero and one can be found in Fisher and Yates [5], p. 84), we find that $-2 \log_e \lambda_1 = 8.61188$. Entering Table IV

(Distribution of χ^2) of Fisher and Yates, under three degrees of freedom, 8.61188 is seen to be significant at the five percent level. For the other statistics, we find that $-2 \log_e \lambda_{2A} = 6.72692$, $-2 \log_e \lambda_{2B} = 3.8434$, $-2 \log_e \lambda_3 = .27488$. Entering Table IV under one degree of freedom, we see that the A-effect is significant at the one percent level and the B-effect at the five percent level. This indicates that a difference exists between the effect of A_1 and that of A_2, and also that possibly a difference exists between the effect of B_1 and the effect of B_2. The evidence for this, however, is not very strong, as the use of the large sample approximation leads to higher levels of significance than the true values. 0.27488 is not significant at any useful level so we do not reject the hypothesis that interaction is absent.

VII. SUMMARY AND DISCUSSION

7.1 Summary

A method is presented for analysing a 2 x 2 factorial experiment in which the data consist of relative rankings in pairwise comparisons. Maximum likelihood estimates are developed for the ratings of the various levels of each factor and for the treatment combinations. Likelihood ratio tests of the most important hypotheses likely to arise are derived in detail. The large sample approximations are used. In addition, the method is presented in a manner such that tests of other hypotheses in which the experimenter might be interested can easily be derived.

The equations for the analysis of a factorial design of arbitrary size are presented. It can be seen, however, that the complexity of these equations render an attempt at their solution impractical in most cases and more work must be done if a useful method of analyzing experiments of this type is to be found.

7.2 Discussion

One promising approach to the problem might be to omit certain comparisons of the type A_iB_j with A_pB_q where $p \neq i$ and $q \neq j$, since it is these comparisons which cause the mathematical complexity. In an experiment of large size the loss of information in omitting these comparisons might not be too serious.

Another approach might be the use of incomplete blocks containing more than two treatments. The case of blocks of three has been investigated by Bradley and Terry. Extension of this work to the factorial design might be profitable.

These approaches are presented in a purely speculative manner without any real evidence that they will provide a solution to the problem. It is believed, however, that investigation in this direction will provide some insight into the behavior of data derived from experiments of this type.

VIII. ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to Professor Ralph A. Bradley, without whose many ideas and suggestions this paper could not have been written. Appreciation is also expressed to Professor M. C. K. Tweedie and to for their many helpful suggestions.

The author would also like to thank his wife, for their assistance in the final preparation of this thesis.

IX. BIBLIOGRAPHY

1. Bradley, R. A., Some Statistical Methods in Taste Testing and Quality Evaluation, Biometrics, 9, (1953)
2. Bradley, R. A. and Terry, M. E., The Rank Analysis of Incomplete Block Designs I - The Method of Paired Comparisons, Biometrika 39, (1952).
3. Bradley, R. A., Bi-Annual Report No. 5 on Statistical Methods for Sensory Difference Testing, Virginia Agricultural Experiment Station, (1952).
4. Fisher, R. A., Two New Properties of Mathematical Likelihood, Proceedings of the Royal Society, Series A, Vol. 144, (1934).
5. Fisher, R. A. and Yates, F., Statistical Tables for Biological, Agricultural and Medical Research, Third Edition, Oliver and Boyd, Edinburgh, (1948).
6. Kendall, M. G., The Advanced Theory of Statistics, Vol. II, Griffin, London, (1946).
7. Mood, A. M., Introduction to the Theory of Statistics, McGraw-Hill, New York, (1950).
8. Mosteller, F., Remarks on the Method of Paired Comparisons, Psychometrika, 16, (1951).

9. Report No. 1 Computation of Tables for the Rank Analysis of Incomplete Block Designs, Quartermaster Food and Container Institute for the Armed Forces, (1952).

10. Thurstone, L. L., Psychophysical Analysis, Amer. J. Psychol. 38, (1927).

11. Wilks, S. S., Mathematical Statistics, Princeton University Press, (1946), chapter 7.

The vita has been removed from the scanned document