TREATMENT CONTRASTS IN PAIRED COMPARISONS:
LARGE-SAMPLE RESULTS, APPLICATIONS
AND SOME OPTIMAL DESIGNS

by

Abdalla T. El-Helbawy and Ralph A. Bradley

FSU Technical Report No. M368
ONR Technical Report No. 102

December, 1976
The Florida State University
Department of Statistics
Tallahassee, Florida 32306

1Research supported at the Florida State University by the Army, Navy and Air Force through ONR Contract N00014-76-C-0608. Reproduction in whole or in part is permitted for any purpose of the United States Government.
TREATMENT CONTRASTS IN PAIRED COMPARISONS:
LARGE-SAMPLE RESULTS, APPLICATIONS AND SOME OPTIMAL DESIGNS

Abdalla T. El-Helbawy
University of Cairo
Cairo, United Arab Republic

Ralph A. Bradley
Florida State University
Tallahassee, Florida, U.S.A.

SUMMARY

Treatment contrasts in paired comparison experiments are considered and associated large-sample results developed. Likelihood estimators of treatment parameters are shown to be consistent when obtained under the assumption that specified orthonormal treatment contrasts are null. Under the same assumption, the distribution functions of treatment parameter estimators, their logarithms, and orthonormal contrasts among them, estimators of treatment contrasts orthonormal also to those assumed null, when suitably scaled and centered are shown to be normal with zero mean vectors and variance-covariance matrices shown in the limit for large sample sizes. The distribution functions of likelihood-ratio statistics are shown to be chi-square in the limit with large sample sizes, central under null hypotheses and non-central under alternative hypotheses specified by a class of local alternatives with non-centrality parameters given.

Much of the paper is devoted to applications of the theoretical results. Section 4 contains five applications, all illustrated with $2^3$-factorial experiments. The first four have results on optimal designs for the use of factorial treatment combinations in paired comparisons. The final one is a numerical one on a taste preference experiment for coffee with brew strength, roast color, and brand being three factors, each at two levels.

---

1Research supported at the Florida State University by the Army, Navy and Air Force through ONR Contract N00014-76-C-0608. Reproduction in whole or in part is permitted for any purpose of the United States Government.
A final section relates the results given with earlier work by Bradley. The model used here differs from that proposed earlier only in a scale-determining constraint on treatment parameters but some care is necessary in comparisons of results. An appendix contains some matrix results used in the paper and of potential use in other statistical research.

The methods of this paper, and the one on basic methodology that preceded it, provide much new flexibility to the use of paired comparisons, particularly with factorial treatment combinations. In addition, the way seems to be open for the development of a more general theory of optimal design for factorials in paired comparisons.
1. INTRODUCTION

Treatment contrasts in paired comparisons have been considered in two earlier papers in this series. Bradley and El-Helbawy (1976) provided the basic methodology based on the model of Bradley and Terry (1952) and illustrated it with reference to factorial treatment combinations. El-Helbawy and Bradley (1977) examined the properties of an iterative scheme for solution of likelihood equations. Large-sample properties of estimators and tests are considered now, along with applications and some aspects of optimal experimental design.

The paired comparisons experiment has $t$ treatments, $T_1, \ldots, T_t$, with $n_{ij}$ judgments or comparisons of $T_i$ and $T_j$, $n_{ij} \geq 0$, $i \neq j$, $n_{ii} = 0$, $n_{ji} = n_{ij}$, $i, j = 1, \ldots, t$. A parameter $\pi_i$ is associated with $T_i$, $\pi_i > 0$, $i = 1, \ldots, t$, such that the probability of selection of $T_i$ when compared with $T_j$ is

$$\text{pr}(T_i > T_j) = \frac{\pi_i}{\pi_i + \pi_j}, \quad i \neq j.$$  \hspace{1cm} (1.1)

Since (1.1) is not dependent on parameter scale, a convenient, scale-determining constraint,

$$\sum_i \gamma_i = 0,$$  \hspace{1cm} (1.2)

where

$$\gamma_i = \log \pi_i, \quad i = 1, \ldots, t,$$  \hspace{1cm} (1.3)

is used.

Treatment contrasts are specified, linear, orthonormal contrasts on the $\gamma_i$.

Let $B_m$, $B_n$ and

$$B_{m+n} = \begin{pmatrix} B_m \\ B_n \end{pmatrix}$$  \hspace{1cm} (1.4)
be respectively \( m \times t \), \( n \times t \), and \((m+n) \times t\) matrices with zero-sum, orthonormal rows, \(0 \leq m + n \leq t - 1\). Let \( \pi \) and \( \gamma(\pi) \) be column vectors with typical elements \( \pi_i \) and \( \gamma_i \) while \( 0_m \) and \( 1_t \) are respectively column vectors of \( m \) zero and \( t \) unit elements. Treatment contrasts of interest have the form \( B_{m+n}\gamma(\pi) \). We shall assume that

\[
\begin{bmatrix}
1_t \\
B_m
\end{bmatrix} \gamma(\pi) = 0_{m+1} \quad (1.5)
\]

in development below. Two main results are obtained: (i) the large sample distribution of \( \sqrt{N} [\gamma(p) - \gamma(\pi)] \) when (1.5) holds and where \( p \) is the likelihood estimator of \( \pi \) and

\[
N = \sum_{i<j} n_{ij}, \quad (1.6)
\]

and (ii) the large-sample distribution of the test statistic \( -2 \log \lambda_N(H_0, H_a) \) on the assumption that (1.5) holds where the null hypothesis is

\[
H_0: B_n\gamma(\pi) = 0_n, \quad (1.7)
\]

the alternative hypothesis is

\[
H_a: B_n\gamma(\pi) \neq 0_n, \quad (1.8)
\]

and \( \lambda_N(H_0, H_a) \) is the likelihood ratio.

The likelihood function is

\[
L(\pi) = \prod_i a_i^{\pi_i} (\pi_i + \pi_j)^{n_{ij}} \quad (1.9)
\]

on the assumption of independence of selections where \( a_i \) is the total number of selections of \( T_i \) in the entire experiment, \( \sum_i a_i = N \). The likelihood equations
for maximization of $L$ subject to (1.5) have been shown to reduce to

$$
\sum_j \left[ a_j - \sum_k n_{jk} p_j / (p_j + p_k) \right] D_{ij} = 0, \ i = 1, \ldots, t,
$$

(1.10)

$$
\sum_i \gamma_i (p) = 0,
$$

(1.11)

and

$$
B_m \gamma (p) = 0_m
$$

(1.12)

where $D_{ij}$ is the $(i, j)$-element of

$$
D = I_t - B'B_m
$$

(1.13)

$I_t$ being the $t$-dimensional identity matrix. Subject to an assumption of Ford (1957) El-Helbawy and Bradley (1977) have shown that an iterative scheme for solution of (1.10), (1.11) and (1.12) converges, that the solution is unique, and that a unique maximum of $L$ results.

Two assumptions are required for the limit theory developed below.

Assumption 1: In every partition of the indices, 1, ..., $t$, into two non-empty subsets $S_1$ and $S_2$, there exists $i \in S_1$ and $j \in S_2$ such that

$$
\lambda_{ij} > 0
$$

where we define

$$
\lambda_{ij} = \lim_{N \to \infty} \frac{n_{ij}}{N}, \ i \neq j, \ i, j = 1, \ldots, t.
$$

(1.14)

Assumption 2: The true parameter point $\pi$ has positive elements.

The first assumption essentially requires correctness of the various parts of the paired comparisons experiment. The second assumption eliminates inclusion of disparate treatments in the experiment, treatments that are never selected.
2. ASYMPTOTIC PROPERTIES OF LIKELIHOOD ESTIMATORS

Bradley and Gart (1962) consider the asymptotic properties of maximum likelihood estimators when sampling from associated populations. Let

\[ f_{ij}(x_{ij}, \pi) = \pi_i^{x_{ij}} \pi_j^{x_{ji}} / (\pi_i + \pi_j), \quad i < j, \quad i, j = 1, \ldots, t, \quad (2.1) \]

where \( x_{ij} = 1, 0 \) as \( T_i \) is selected over \( T_j \) or otherwise and \( x_{ij} + x_{ji} = 1 \).

Then \( f_{ij} \) is the probability function of the \((i,j)\)-th of \( 2^t \) associated populations and \( L(\pi) \) in (1.9) may be written

\[ L(\pi) = \prod_{i<j} \prod_{a=1}^{n_{ij}} f_{ij}(x_{ija}, \pi), \quad (2.2) \]

where \( x_{ija} \) is the value of \( x_{ij} \) in the \( a \)-th comparison of \( T_i \) and \( T_j \).

Transformation of the parameter space simplifies the study. In view of Assumption 2, there exists a positive real number \( \delta \) such that \( \pi_a \geq \delta 1_t \) where \( \pi_a \) is the true parameter vector. Let

\[ \omega_a = \{ \pi \geq \delta 1_t \text{ where } (1.5) \text{ holds} \} \quad (2.3) \]

and \( \pi_a \in \omega_a \). Let

\[ \theta = \begin{bmatrix} 0^\theta \\ 1^\theta \end{bmatrix}, \quad 0^\theta = \begin{bmatrix} 1^\theta \\ \gamma(\pi) \end{bmatrix}, \quad 1^\theta = \bar{B}_m \gamma(\pi), \text{ and } \bar{B} = \begin{bmatrix} 1^\theta \\ \gamma(\pi) \\ \bar{B}_m \end{bmatrix} \quad (2.4) \]
where $B$ is the $t \times t$ orthonormal matrix with typical element $B_{ij}$ completed by a $(t - m - 1) \times t$ matrix $\bar{B}_{-m}$. The transformed parameter space is

$$\omega_a^T = \{ \theta: \theta_j = 0_{t-m+1}, \sum_{j=m+2}^t B_{ij} \theta_j \geq \log \delta, i = 1, \ldots, t \}. \quad (2.5)$$

A one-to-one correspondence between $\omega_a^T$ and $\omega_a^T$ is given by (2.4):

$$\gamma(\pi) = \bar{B}'_{-m} \theta. \quad (2.6)$$

The method of Bradley and Gart is used with $\bar{\theta}$ as the basic parameter vector. Regularity conditions (i) and (ii) of that method are verified easily. Condition (iii) requires more detailed consideration; we need to show that the $(t - m - 1)$-square matrix

$$C(\pi) = [C_{rq}(\pi), r, q = m+2, \ldots, t], \quad (2.7)$$

is positive definite for each $\theta \in \omega_a^T$, where

$$C(\pi) = \bar{B}_{-m} \Lambda(\pi) \bar{B}'_{-m} \quad (2.8)$$

and $\Lambda(\pi)$ is the $t$-square matrix with elements

$$\Lambda_{ii}(\pi) = \pi_i \sum_j \lambda_{ij} \pi_j / (\pi_i + \pi_j)^2, \quad i = 1, \ldots, t, \quad (2.9)$$

$$\Lambda_{ij}(\pi) = -\lambda_{ij} \pi_i \pi_j / (\pi_i + \pi_j)^2, \quad i \neq j, i, j = 1, \ldots, t.$$
To verify (2.8) and (2.9), it is necessary to note that

\[
C_{r|q}(\pi) = \sum_{i<j} \lambda_{ij} E_{\theta} \left( \frac{\partial \log f_{ij}}{\partial \theta_r} \frac{\partial \log f_{ij}}{\partial \theta_q} \right)
\]

\[
= \sum_{k} \sum_{\ell} \left\{ \sum_{i<j} \lambda_{ij} E_{\pi} \left( \frac{\partial \log f_{ij}}{\partial \gamma_k} \frac{\partial \log f_{ij}}{\partial \gamma_{\ell}} \right) \right\} B_{rk} B_{q\ell}
\]

and compute the necessary expectations. Appendix Theorem A1 and Assumptions 1 and 2 show that \( A(\pi) \) has rank \((t-1)\). Observe also that the matrix \( H_{ij} \) with \((k,\ell)\)-element taken as the expectation in (2.10) is positive semi-definite and hence \( A(\pi) \) is positive semi-definite since it is a linear function of the \( H_{ij} \), the coefficients being non-negative and not all zero. Then Appendix Theorem A2 is sufficient to establish that \( C(\pi) \) is positive definite.

Conditions for the application of Theorems 1(i) and 1(iv) of Bradley and Gart have been established. It follows from Theorem 1(i) that the equations,

\[
\frac{\partial \log L(\pi)}{\partial \theta_r} = 0, \quad r = m+2, \ldots, t,
\]

have a solution \( \hat{\theta} \) which is a consistent estimator of \( \theta^\pi \), where

\[
\theta^\pi = \begin{bmatrix} 0_{m+1} \\ -m+1 \\ \theta^\alpha \end{bmatrix}
\]

is the point in \( \omega^T \) corresponding to \( \pi_a \) in \( \omega_a \). But equations (2.11) are a transform of equations (1.10), (1.11) and (1.12) which have been shown by El-Helbawy and Bradley (1977) to have a unique solution \( p \) maximizing \( L(\pi) \) on \( \omega_a \). Thus \( \hat{\theta} \) is unique and consistent for \( \theta^\alpha \). Further,

\[
\gamma(p) = \frac{\partial L}{\partial \theta}, \quad \hat{\theta}
\]

(2.12)
and

\[ \hat{p} = \exp \hat{B}' \hat{m} \hat{\theta}. \]  

(2.13)

We summarize application of Theorem 1(i) as follows:

**Theorem 1:** Given Assumptions 1 and 2 and (1.5), \( \hat{p} \), the unique solution of equations (1.10), (1.11) and (1.12), is a consistent estimator of \( \pi_a \) and \( \gamma(p) \) is a consistent estimator of \( \gamma(\pi_a), \pi_a \in \omega_a. \)

From reference Theorem 1(iv), \( \sqrt{N} \left( \hat{\theta} - \theta^0 \right) \) has a distribution function which in the limit is \( (t-m-1) \)-variate normal with zero mean vector and dispersion matrix,

\[ \sum(\pi_a) = [\zeta(\pi_a)]^{-1}, \]  

(2.14)

where \( \zeta(\pi) \) is defined in (2.8) and (2.9). From this and (2.12), we have

**Theorem 2:** Given the conditions of Theorem 1, \( \sqrt{N} \left[ \gamma(p) - \gamma(\pi_a) \right] \) has a limiting distribution function that is singular, \( t \)-variate normal in a space of \( (t-m-1) \) dimensions with zero mean vector and dispersion matrix,

\[ \sum^{-1}(\pi_a) = \overline{B}' \sum(\pi_a) \overline{B}. \]  

(2.15)

**Theorem 3:** \( \sum_1(\pi_a) \) is the \( t \)-square, principal minor of the \( (t+m+1) \)-square matrix,

\[
\begin{bmatrix}
\Lambda(\pi_a) & \frac{1}{\sqrt{t}} \cdot 1 & B'_{-m} \\
\frac{1}{\sqrt{t}} \cdot 1 & t & 0 \\
B_{-m} & 0 & \quad \left[ -1 \right]
\end{bmatrix}^{-1}
\]  

(2.16)
Theorem 3 follows from Appendix Theorem A3. Application of Theorem (iii) of Rao (1965, p. 322) and use of Theorem 2 yields the result that \( \sqrt{N} \left( \mathbf{p} - \pi_a \right) \) has a limiting distribution function that is singular, t-variate normal in a space of \((t-m-1)\) dimensions with zero mean vector and dispersion matrix,

\[
\sum_{2}^{(\pi_a)} = d(\pi_a) \sum_{1}^{(\pi_a)} d(\pi_a), \tag{2.17}
\]

where \( d(\pi_a) \) is the \( t \times t \) diagonal matrix with the elements of \( \pi_a \) as diagonal elements. It is a corollary of Theorem 2 that \( \sqrt{N} \mathbf{B}_n [\gamma(p) - \gamma(\pi_a)] \) has a limiting distribution function that is non-singular, \( n \)-variate normal in a space of \( n \) dimensions with zero mean vector and dispersion matrix, \( \mathbf{B}_n \Sigma_{1}(\pi_a)\mathbf{B}_n' \). If \( \mathbf{B}_m \) has \( \mathbf{B}_n \) for its first \( n \) rows, then the dispersion matrix is the \( n \)-square first principal minor of \( \Sigma(\pi_a) \).

Estimation of dispersion matrices may be required. This is done through estimation of \( \Lambda(\pi_a) \) by \( \Lambda(p) \).

3. LIKELIHOOD RATIO TESTS

The general test situation is that described by (1.7) and (1.8) given assumption (1.5). Likelihood estimators \( p_0 \) and \( p \) are obtained as described for \( p \) in Section 1; when \( p_0 \) is obtained, \( \mathbf{B}_m \) is replaced by \( \mathbf{B}_{m+n} \). Asymptotic properties of \( p_0 \) under (1.5) and (1.7) are thus available also from use of Section 2.

The likelihood ratio statistics is

\[-2 \log \lambda_N(H_0, H_a) = 2[\log L(p) - \log L(p_0)]. \tag{3.1}\]
The results of Davidson and Lever (1970) for associated populations are used to obtain the asymptotic distribution function of \(-2 \log \lambda_N(H_0, H_a)\) under \(H_0\) and \(H_a\) with local alternatives defined below.

The class of local alternatives is defined as follows: Let \(\{\pi^N\}\) be a sequence of values of \(\pi\) such that each \(\pi^N\) satisfies (1.5) with

\[
\mathbb{E}_{\pi^{N}}(\delta_i^N) = N^{-1/2} \delta_i^N, \tag{3.2}
\]

\[
\lim_{N \to \infty} \delta_i^N = \delta_i, \quad i = 1, \ldots, n, \tag{3.3}
\]

where \(\delta_i^N\) is a column vector with \(i\)-th element \(\delta_i^N\), and

\[
\mathbb{E}_{\pi^{m+n}}(\xi^{N}) = \mathbb{E}_{\pi^{m+n}}(\xi^0) \tag{3.4}
\]

where \(\pi^0\) is a value of \(\pi\) under \(H_0\) satisfying (1.5) and (1.7).

Theorem 4 of Davidson and Lever is used. Let \(F_n(\lambda^2, z)\) be the distribution function of the non-central, chi-square variate with \(n\) degrees of freedom and non-centrality parameter \(\lambda^2\).

**Theorem 4:** Under the sequence of local alternatives \(\{\pi^N\}\),

\[
\lim_{N \to \infty} \Pr \{-2 \log \lambda_N(H_0, H_a) \leq z\} = F_n(\lambda^2, z), \tag{3.5}
\]

where

\[
\lambda^2 = \delta^0 \cdot \mathbb{E}_0^{-1} \delta, \tag{3.6}
\]

\(\xi_0\) is the \(n\)-square principal minor of \(\mathbb{E}_0^{-1}\), \(\mathbb{E}_n\) provides the first \(n\) rows of \(\mathbb{E}_m\),

and

\[
\mathbb{E}_0 = \mathbb{E}_m \Delta(\pi_0) \mathbb{E}_m', \tag{3.7}
\]
the form of (2.8). When $H_0$ is true, $\delta = 0$, $n^2 = 0$, and the limit distribution is the central, chi-square distribution with $n$ degrees of freedom.

Theorem 4 is established through verification of the regularity conditions of Davidson and Lever. The troublesome one is the verification of the positive definiteness of $\zeta(\mu)$ in (2.8) and hence also of $\zeta_0$ in (3.7); this has been done in Section 2.

The likelihood ratio statistic in (3.1) may be partitioned. Retain assumption (1.5) and consider the nested sequence of hypotheses, $H_{a0}, H_{a1}, \ldots, H_{aj}, J \geq 2$, such that $H_{a0} = H_a$, $H_{aj} = H_0$, nested in the sense that, if $H_{aj}$ consists of $h_j$ distinct equations selected from the set (1.7), $h_j > h_{j-1}$, the $h_{j-1}$ equations for $H_{a, j-1}$ are included in those for $H_{aj}, j = 1, \ldots, J, h_0 = 0, h_J = n$. It is easily seen that

$$-2 \log \lambda_n(H_0, H_a) = \sum_{j=1}^{J} -2 \log \lambda_n(H_{aj}, H_{aj-1}). \quad (3.8)$$

Theorem 4 may be applied to each term in the sum in (3.8). The results of Good (1967) establish that the asymptotically chi-square variates, the $j$-th of which has $(h_j - h_{j-1})$ degrees of freedom, are also asymptotically independent.

Bradley and El-Helbawy (1976) in their concluding remarks suggested that (1.5) may not be constraining in the test of $H_0$ versus $H_a$. The remark was made because of the similarity of analyses in their Tables 3 and 4 of their example, a similarity that simplifies interpretation. There seems to be no basis for this to be true generally but it is possible that in many examples the likelihood surface is such that (3.1) is little affected by (1.5). The simplest situation occurs when the constrained maxima of $L(\mu)$ under $H_0$ and $H_a$ are the unconstrained maxima.
4. APPLICATIONS AND OPTIMAL DESIGN CONSIDERATIONS

(i). Bradley and El-Helbawy (1976) considered specifically the application of the general consideration of treatment contrasts to mixed factorials, reparameterized the general model to introduce factorial parameters, and illustrated the methodology through use of a $2^3$-factorial example. El-Helbawy and Bradley (1976) gave additional details on factorials. We consider here a $2^3$-factorial and the large-sample results of this paper.

Take $t = 8$, replace $T_i$ by $T_a$ and $\pi_i$ by $\pi_a$ is (1.1) where $\alpha = (\alpha_1, \alpha_2, \alpha_3)$, $\alpha_s = 0, 1$, $s = 1, 2, 3$, and let $\alpha_s$ represent the level of factor $s$. The factorial parameters are introduced [See (4.4), Bradley and El-Helbawy (1976)] when we set

$$
\pi_a = \pi_{a_1}^{1} \pi_{a_2}^{2} \pi_{a_3}^{3} \pi_{a_1 a_2}^{12} \pi_{a_1 a_3}^{13} \pi_{a_2 a_3}^{23} \pi_{a_1 a_2 a_3}^{123}
$$

(4.1)

for all $a$. From (1.3),

$$
\gamma_a = \log \pi_a = \sum_{s=1}^{3} \log \pi_{a_s}^{s} + \sum_{s<s'} \log \pi_{a_s a_{s'}}^{s s'} + \log \pi_{a_1 a_2 a_3}^{123}
$$

(4.2)

$$
= \sum_{s=1}^{3} \gamma_{a_s}^{s} + \sum_{s<s'} \gamma_{a_s a_{s'}}^{s s'} + \gamma_{a_1 a_2 a_3}^{123}
$$

the final form defining the factorial parameters $\gamma_{a_s}^{s}$, $\gamma_{a_s a_{s'}}^{s s'}$, $\gamma_{a_1 a_2 a_3}^{123}$. The factorial parameters in (4.1) and (4.2) are subject to constraints, the usual ones of the analysis of variance when (4.2) is considered. The constraints are
\[
\frac{1}{s} \sum_{s=0}^{3} \gamma_{ss} = 0, \quad s = 1, 2, 3, \tag{4.3}
\]

\[
\frac{1}{s} \sum_{s=0}^{3} \gamma_{ss'} = 0, \quad s < s', \quad s, s' = 1, 2, 3, \tag{4.4}
\]

\[
\frac{1}{3} \gamma_{123} = \frac{1}{3} \gamma_{123} = \frac{1}{3} \gamma_{123} = 0. \tag{4.5}
\]

Consider the null hypothesis of no two-factor interaction between the first two factors. This may be done by setting \(\gamma_{12} = 0\) and (4.4) requires that \(\gamma_{12} = 0, \quad \alpha_1, \alpha_2 = 0, 1\). Alternatively, postulate that

\[
\gamma_{12} - \gamma_{01} - \gamma_{10} + \gamma_{11} = 0. \tag{4.6}
\]

In the general test situation of this paper, \(B_{n=1}\) in (1.5) does not exist,

\[
B_{n=1} = \frac{1}{\sqrt{8}} \quad (1, 1, -1, -1, -1, -1, 1, 1) \tag{4.7}
\]

in (1.7) when we take the 8 elements of \(\gamma\) in ascending order of the triplet \((\alpha_1, \alpha_2, \alpha_3)\), and \(B_{n=1}\) is the first row of \(B_m\) which now consists of the 7 standard factorial treatment contrasts. With the definition of (4.7) in (1.7), (4.6) results.

The limit distributions are specified by Theorem 4. But \(\chi^2\) in (3.6) depends on \(\pi_0\) and the comparison proportions \(\lambda_{ij}\). We take \(\pi_0 = 1/8\) consistent with \(H_0\) and the concept that any other effects present are of the same order of magnitude relative to \(N\) as the two-factor interaction under test. The experiment is assumed to be as balanced as possible but permitting optimality considerations;
in this example, we take \( \lambda_{ij} = a \) or \( b \) respectively as \( T_i \) and \( T_j \) represent factorial treatment combinations with factor levels \( \alpha_1 \) and \( \alpha_2 \) such that
\[
\alpha_1^{+\alpha_2} \quad \text{does or does not have the same sign for the two treatments, } 12a+16b = 1.
\]
From (2.7) and (3.7), \( \Lambda(I_8) \) has diagonal elements, \( b + 3a/4 \), and \((i, j)\)-elements, \( i \neq j \), \( -a/4 \) or \(-b/4\) again dependent on agreement or disagreement of signs of \( \alpha_1^{+\alpha_2} \) for \( T_i \) and \( T_j \). Then,
\[
C_0 = \begin{bmatrix}
2b & 0' \\
0 & -6
\end{bmatrix}
\]
and \( \Sigma_0^{-1} = 2b \). The limit distributions for the test of interaction between factors 1 and 2 are chi-square with 1 degree of freedom, central under \( H_0 \) and non-central under \( H_a \) with non-centrality parameter \( \lambda^2 = 2b\delta^2 \), where the sequence \( \{n_i\} \) is such that (3.2) yields \( \delta = \sqrt{8N} \gamma_{00}^{12} \). Note that asymptotic power is maximized when \( \lambda^2 \) and hence \( b \) is taken as large as possible, \( b = 1/16 \) and \( a = 0 \). If the experiment is designed only to test for the specified two-factor interaction, all paired comparisons should be made between treatments with differing signs of \( \alpha_1^{+\alpha_2} \) determined from their factorial formulations. This conclusion is in agreement with intuition and constitutes a beginning towards optimal design considerations in planning factorial experiments in paired comparisons.

(ii). Tests for the two-factor interaction of subsection 4(i) are available also when other specified factorial effects are assumed to be null. The results and conclusions of 4(i) still apply. For example, suppose that the three-factor and other two-factor interactions are assumed to be null. Then \( B_m \) in (1.5) is
\[
\mathbf{B}_{m=3} = \frac{1}{\sqrt{8}} \begin{bmatrix}
1 & -1 & 1 & -1 & 1 & -1 & 1 \\
1 & -1 & -1 & 1 & 1 & -1 & 1 \\
1 & -1 & -1 & 1 & -1 & 1 & -1 \\
\end{bmatrix}
\]

and
\[
\mathbf{\tilde{B}}_m = \frac{1}{\sqrt{8}} \begin{bmatrix}
1 & 1 & -1 & -1 & -1 & 1 & 1 \\
1 & 1 & 1 & 1 & -1 & -1 & -1 \\
1 & 1 & -1 & -1 & 1 & 1 & -1 \\
1 & -1 & 1 & -1 & 1 & -1 & 1 \\
\end{bmatrix}
\]

with \( B_{n=1} \) as in (4.7) and in the first row of \( \mathbf{\tilde{B}}_m \). It follows that
\[
\mathbf{C}_0 = \begin{bmatrix}
2b & 0 \\
0 & \frac{1}{2} \\
\end{bmatrix}
\text{ and } \mathbf{\Sigma}_0^{-1} = 2b.
\]

Returning to the remark of Bradley and El-Helbawy (1976) discussed at the end of Section 3, we note that, while the statistics for the tests of subsections 4(i) and 4(ii) may differ, their limit distributions do not under either \( H_0 \) or \( H_a \).

(iii). Let us consider a somewhat more complex optimal design problem than that of 4(i). Suppose that the \( 2^3 \)-factorial experiment is planned to detect possible interactions involving factor one represented by \( \gamma_{00}^{12}, \gamma_{00}^{13} \) and \( \gamma_{000}^{123} \) from (4.2) after use of (4.3)-(4.5). Now
\[
\bar{B}_{n=3} = \frac{1}{\sqrt{6}} \begin{bmatrix}
1 & 1 & -1 & -1 & -1 & 1 & 1 \\
1 & -1 & 1 & -1 & -1 & 1 & 1 \\
1 & -1 & -1 & 1 & -1 & 1 & -1 \\
\end{bmatrix}, \quad (4.8)
\]

\(\bar{B}_m\) does not exist (unless other contrasts are assumed null), and \(\bar{B}_m\) consists of the 7 factorial contrasts possible with \(\bar{B}_{n=3}\) in (4.8) being the first three rows of \(\bar{B}_m\).

The comparison proportions \(\lambda_{ij}\) must be specified with as much symmetry as possible maintained but with optimal design consideration possible. We define comparison proportions as follows: 
- \(\lambda_{ij} = a\) for comparisons of treatment combinations yielding information on \(\gamma_{00}\) alone or \(\gamma_{00}\) alone,
- \(\lambda_{ij} = b\) for comparisons yielding information on both \(\gamma_{00}\) and \(\gamma_{00}\) but not on \(\gamma_{000}\),
- \(\lambda_{ij} = c\) for comparisons yielding information only on the pair \(\gamma_{00}\) and \(\gamma_{00}\) or the pair \(\gamma_{000}\),
- \(\lambda_{ij} = d\) for comparisons yielding information on \(\gamma_{00}, \gamma_{00}\) and \(\gamma_{000}\),
- \(\lambda_{ij} = e\) for comparisons yielding information only on \(\gamma_{000}\).

8\(a\) + 4\(b\) + 8\(c\) + 4\(d\) + 4\(e\) = 1. With the association of factorial treatment combinations to the treatments by the ascending order of \((a_1, a_2, a_3)\), we have the symmetric matrix,

\[
[\lambda_{ij}] = \begin{bmatrix}
0 & c & c & b & d & a & a & e \\
0 & b & c & a & d & e & a \\
0 & c & a & e & d & a \\
0 & e & a & a & d & 0 & c & 0 & c \\
0 & b & c & 0 & b & c \\
\end{bmatrix}.
\]

(symmetric)
From direct use of (2.8) with \( n_0 = 1 \), it follows that \( C_0 \) is a diagonal matrix and the 3x3 principal minor of \( C_0 \) is

\[
\Sigma_0^{-1} = \frac{1}{2} \begin{bmatrix}
(a+b+c+d) & 0 & 0 \\
0 & (a+b+c+d) & 0 \\
0 & 0 & (2c+d+e)
\end{bmatrix}
\]  

(4.9)

The application is finished when we compute \( \lambda^2 \) in Theorem 4: 

\[
\delta' = \sqrt{8N} \begin{bmatrix}
\gamma_{22}^{12} & \gamma_{23}^{13} & \gamma_{24}^{123}
\end{bmatrix}
\text{from (3.2)}
\]

\[
\lambda^2 = 4N[(a+b+c+d)\left(\gamma_{00}^{12}\right)^2 + (\gamma_{00}^{13})^2 + (2c+d+e)(\gamma_{00}^{123})^2].
\]

If the experiment is to be designed for detection of \( \gamma_{00}^{12}, \gamma_{00}^{13} \) and \( \gamma_{00}^{123} \), with equal emphasis on each, the optimal design would maximize \( (2a+2b+4c+5d+e) \) subject to the constraints that \( (8a+4b+8c+4d+4e) = 1 \) and each parameter is non-negative. The maximum of \( \lambda^2 \) occurs when \( d = 1/4, a = b = c = e = 0 \), from elementary programming considerations. Thus all comparisons used should yield information on all three interactions of interest and all four of these treatment pairings should be used to avoid confounding; the comparisons are \( T_{000} \) vs \( T_{100} \), \( T_{001} \) vs \( T_{101} \), \( T_{010} \) vs \( T_{110} \), and \( T_{100} \) vs \( T_{111} \). If some other factorial effects or treatment contrasts are assumed null, specifying \( B_m \), the conditional test of interactions involving factor one may be used; then, since \( C_0 \) remains diagonal, has the same 3x3 first principal minor, and is only reduced in dimensionality, \( \Sigma_0^{-1} \) remains as in (4.9) and the same optimality results are obtained.
(iv). In principle, Theorem 2 may be used to consider optimal designs for specified treatment contrasts in paired comparisons. Optimality criteria would be applied to \( \Sigma_1(\pi_a) \) in (2.15). But, in practice, difficulty enters since \( \Sigma_1(\pi_a) \) does depend on the unknown parameter vector \( \pi_a \) and some simplifying argument is required.

Let us consider an application similar to that of 4(iii) with the 2\(^3\) factorial. Let \( B_m \) consist of the 4 orthonormal rows for factorial contrasts orthogonal to those for \( \gamma_{00}, \gamma_{01}, \) and \( \gamma_{000} \) and let (1.5) apply. Then \( B_{n=3} \) of application 4(iii) equals \( B_m \). From Theorem 2, \( \sqrt{N} \left[ y(p) - y(\pi_a) \right] \) has a limiting distribution function that is singular, 8-variate normal in a space of 3 dimensions with zero mean vector and dispersion matrix, \( \Sigma_1(\pi_a) = B_m' \Sigma(\pi_a) B_m = B_{n=3}' \Sigma(\pi_a) B_{n=3} \). As a consequence, \( \sqrt{N} B_{n=3} \left[ y(p) - y(\pi_a) \right] = \sqrt{N} \left[ \gamma_{12}^{12} (p) - \gamma_{00}^{12} (\pi_a), \gamma_{13}^{13} (p) - \gamma_{00}^{13} (\pi_a), \gamma_{000}^{123} (p) - \gamma_{000}^{123} (\pi_a) \right] \) has the trivariate normal limiting distribution with zero mean vector and dispersion matrix \( B_{n=3}' \Sigma_1(\pi_a) B_{n=3} = \Sigma(\pi_a) \), the matrix of interest in this example.

Let us make the simplifying assumption that \( \pi_a = \lambda_8 \) and impose the comparison proportions \( \lambda_{ij} \) of application 4(iii). Then \( C(\lambda_8) = B_m' \Lambda(\lambda_8) B_m = B_{n=3}' \Lambda(\lambda_8) B_{n=3} \) from (2.8), \( \Sigma(\lambda_8) = C^{-1}(\lambda_8) \) from (2.14), and \( C(\lambda_8) \) is given by (4.9). Then

\[
\Sigma(\lambda_8) = \frac{1}{2} \begin{bmatrix}
1/(a+b+c+d) & 0 & 0 \\
0 & 1/(a+b+c+d) & 0 \\
0 & 0 & 1/(2c+d+e)
\end{bmatrix}.
\]  

(4.10)

In this application, optimality results depend on (4.10).
The paired comparisons design with \( a = b = c = e = 0, d = 1/4 \) is \( A_-, D_-, \) and \( E \)-optimal minimizing \( \text{tr} \Sigma(\Gamma_8), |\Sigma(\Gamma_8)| \), and the larger of \( 1/(a+b+c+d) \) and \( 1/(2c+d+e) \) respectively for estimation of \( \delta_{00}, \delta_{10}, \delta_{20} \) and \( \delta_{00}, \delta_{01}, \delta_{02} \) given (1.5) as applied and \( \tau_a = 1/8 \). The optimal design is the same one obtained in application 4(iii). Let \( A = a+b+c+d \) and \( B = 2c+d+e, 8a+4b+8c+4d+4e = 1, a,b,c,d,e \geq 0. \) All three optimality criteria are met if comparison proportions are found that maximize both \( A \) and \( B \) simultaneously. Note that \( B = \frac{1}{4} - 2a - b. \) Take any fixed \( a \) and \( b \) meeting the conditions; then \( B \) is fixed and \( A \) is maximized when \( e = 0. \) Let \( c = \frac{q}{2} \left( \frac{1}{4} - 2a - b \right) \) and \( d = (1-q) \left( \frac{1}{4} - 2a - b \right), 0 \leq q \leq 1. \) \( B \) remains fixed, \( A \) is maximized when \( q = 0, \) and \( c = 0. \) Now let \( d \) be fixed, \( a = \frac{c^*}{2} \left( \frac{1}{4} - d \right) \) and \( b = (1-q^*) \left( \frac{1}{4} - d \right), 0 \leq q^* \leq 1. \) Then \( B = d \) and \( A = d + (1-q^*) \left( \frac{1}{4} - d \right) \) maximized when \( q^* = 0, a = 0. \) Finally, \( b = \frac{1}{4} - d, A = 1/4, B = d \) and \( B \) is maximized for maximum \( d, d = 1/4. \)

Note: Applications 4(i)-(iv) seem to be the first results available in optimal design in paired comparisons. Littell and Boyett (1976) did compare two possible designs for an \( R \times C \) factorial in paired comparisons on the basis of asymptotic relative efficiencies. The way may now be open for a more general theory of optimal design for paired comparisons relative to specified treatment contrasts.

(v). This section on applications is concluded with numerical examples illustrating use of results in Section 2. The data of Bradley and El-Helbawy (1976) are used: their example gave the basic analysis for a \( 2^3 \)-factorial on coffee preferences where the three factors were brew strength, roast color and coffee brand. Twenty-six preference judgments were obtained on each of the 28 possible treatment comparisons. Two situations from the reference paper are considered: (a) No treatment contrasts assumed null (the conditions of line 1 of the reference Table 3), \( B_m \) does not exist and (b) factor 3 does not interact (the conditions of line 7 of the reference Table 3),
\[ B_{m=3} = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 \end{bmatrix} \]  \quad (4.11)

A data summary is given in Table 1 along with parameter estimates from (1.10) - (1.13) under situations (a) and (b) taken from the reference paper. Estimated variance-covariance calculations are shown separately now for the two situations.

Table 1
Coffee Test Preference Data and Parameter Estimates

<table>
<thead>
<tr>
<th>Treatments</th>
<th>( T_1 ) = ( T_{000} )</th>
<th>( T_2 ) = ( T_{001} )</th>
<th>( T_3 ) = ( T_{010} )</th>
<th>( T_4 ) = ( T_{011} )</th>
<th>( T_5 ) = ( T_{100} )</th>
<th>( T_6 ) = ( T_{101} )</th>
<th>( T_7 ) = ( T_{110} )</th>
<th>( T_8 ) = ( T_{111} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Preferences, ( a_i )</td>
<td>114</td>
<td>92</td>
<td>104</td>
<td>85</td>
<td>77</td>
<td>102</td>
<td>71</td>
<td>83</td>
</tr>
<tr>
<td>(a) Estimates ( p_i )</td>
<td>1.5739</td>
<td>1.0206</td>
<td>1.2915</td>
<td>0.8897</td>
<td>0.7592</td>
<td>1.2408</td>
<td>0.6728</td>
<td>0.8548</td>
</tr>
<tr>
<td>(b) Estimates ( p_i )</td>
<td>1.2724</td>
<td>1.2500</td>
<td>1.0806</td>
<td>1.0616</td>
<td>0.9805</td>
<td>0.9633</td>
<td>0.7685</td>
<td>0.7550</td>
</tr>
</tbody>
</table>

(a). The matrix \( \Lambda(\pi) \) in (2.8) and (2.9) is estimated with \( \pi = p_a \) from the (a)-estimates of \( p \) in Table 1; designated as \( \Lambda(p_a) \), it is shown in the first section of Table 2. Theorem 3 and (2.16) are used to obtain the estimated variance-covariance matrix of \( \sqrt{N} [\gamma(p) - \gamma(\pi_{-a})] \) in Theorem 2, \( N = 26 \times 28 \). The dispersion matrix \( \Sigma_1(p_a) \) is the 8-square principal minor of the inverse matrix (2.16) with \( \Lambda(p_a) \) and \( \sqrt{8} \frac{1}{8} \) substituted. Here \( B_{m} \) does not exist and the inverse matrix is \( 9 \times 9 \). The desired principal minor is shown in the second section of Table 2.
The factorial parameters are estimated from $\gamma(p_a)$ through the relationship,

$$\bar{B}_m \gamma = \sqrt{8} \left( \gamma_0', \gamma_0, \gamma_0', \gamma_0, \gamma_{00}, \gamma_{00}, \gamma_{000}, \gamma_{000} \right),$$

(4.12)

after use of (4.3)-(4.5) with $\bar{B}_m$ consisting of the 7 sets of coefficients of the orthonormal factorial treatment contrasts ordered to match the order of the parameters in the right-hand side of (4.12). The estimates values are, in the same order, 0.1532, 0.1035, 0.0094, -0.0198, 0.1921, -0.0239, and 0.0390, natural logarithms being used. The vector of factorial parameters,

$$\sqrt{N} \left[ \gamma_0'(p) - \gamma_0'(\pi_a), \ldots, \gamma_{123}(p) - \gamma_{123}(\pi_a) \right],$$

order of elements matching (4.12), calculated from $\sqrt{N} \left[ \gamma(p) - \gamma(\pi_a) \right]$ of Theorem 2 through use of (4.12), since the transformation applied is linear, has a non-singular seven-variate normal limiting distribution with zero mean vector and dispersion matrix estimated by $\Sigma_2(p_a) = \frac{1}{8} \bar{B}_m \Sigma_1(p_a) \bar{B}_m^t$ evaluated in the last section of Table 2. This result may be used in various ways, for example, in the calculation of confidence regions on the factorial parameters. Note the small covariances and nearly equal variances of $\Sigma_2(p_a)$. 
<table>
<thead>
<tr>
<th>$\Delta(p_a)$</th>
<th>0.0579</th>
<th>-0.0085</th>
<th>-0.0088</th>
<th>-0.0082</th>
<th>-0.0078</th>
<th>-0.0083</th>
<th>-0.0074</th>
<th>-0.0081</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0612</td>
<td>-0.0088</td>
<td>-0.0088</td>
<td>-0.0087</td>
<td>-0.0088</td>
<td>-0.0085</td>
<td>-0.0088</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0601</td>
<td>-0.0086</td>
<td>-0.0086</td>
<td>-0.0083</td>
<td>-0.0089</td>
<td>-0.0080</td>
<td>-0.0085</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0610</td>
<td>-0.0088</td>
<td>-0.0086</td>
<td>-0.0087</td>
<td>-0.0088</td>
<td>-0.0089</td>
<td>-0.0089</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0600</td>
<td>-0.0084</td>
<td>-0.0086</td>
<td>-0.0089</td>
<td>-0.0089</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.0604</td>
<td></td>
<td></td>
<td>-0.0081</td>
<td>-0.0086</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.0587</td>
<td></td>
<td></td>
<td>-0.0088</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.0608</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Xi_1(p_a)$</td>
<td>13.2315</td>
<td>-1.8341</td>
<td>-1.7610</td>
<td>-1.8964</td>
<td>-1.9859</td>
<td>-1.7703</td>
<td>-2.0666</td>
<td>-1.9172</td>
</tr>
<tr>
<td></td>
<td>12.5094</td>
<td>-1.7738</td>
<td>-1.7573</td>
<td>-1.7882</td>
<td>-1.7660</td>
<td>-1.8271</td>
<td>-1.7630</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.7553</td>
<td>-1.8121</td>
<td>-1.8761</td>
<td>-1.7471</td>
<td>-1.9391</td>
<td>-1.8262</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.5538</td>
<td>-1.7584</td>
<td>-1.7993</td>
<td>-1.7819</td>
<td>-1.7485</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.7678</td>
<td>-1.8578</td>
<td>-1.7488</td>
<td>-1.7527</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.6696</td>
<td>-1.9170</td>
<td>-1.8120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>13.0520</td>
<td>-1.7715</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12.5912</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Xi_2(p_a)$</td>
<td>1.8350</td>
<td>0.0246</td>
<td>0.0065</td>
<td>0.0009</td>
<td>0.0494</td>
<td>0.0103</td>
<td>0.0081</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8234</td>
<td>0.0037</td>
<td>-0.0026</td>
<td>0.0321</td>
<td>0.0201</td>
<td>0.0091</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8145</td>
<td>0.0140</td>
<td>0.0011</td>
<td>0.0037</td>
<td>0.0109</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8141</td>
<td>-0.0007</td>
<td>0.0058</td>
<td>-0.0219</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.8475</td>
<td>0.0056</td>
<td>0.0108</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.8141</td>
<td>-0.0016</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1.8152</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(b) Calculations for situation (b) follow the same pattern as for (a). We now use \( \pi = p_b \) from the (b)-estimates of \( p \) in Table 1 and calculate \( \Lambda(p_b) \) for the first section of Table 3. In the second section of Table 3, we have \( \Sigma_1(p_b) \), again calculated from (2.16) with \( \Lambda(p_b), \sqrt{\delta} \), \( b \) and \( \x_m=3 \) in (4.11) substituted. Now (4.12) is replaced by

\[
\bar{\beta}_{m=3} \gamma = \sqrt{\delta} (\gamma_0^1, \gamma_0^2, \gamma_0^3, \gamma_0^{12}) \tag{4.13}
\]

with \( \bar{\beta}_{m=3} \) consisting of the 4 sets of coefficients of the orthonormal factorial treatment contrasts, orthonormal to those of \( \x_m=3 \), in the order of the parameters in the right-hand side of (4.13). The estimated values are, in the same order, 0.1503, 0.1017, 0.0089 and -0.0201. The vector of parameters,

\[
\sqrt{N} [\gamma_0^1(p) - \gamma_0^1(\pi_a), \ldots, \gamma_0^{12}(p) - \gamma_0^{12}(\pi_a)],
\]

order of elements matching (4.13), has a non-singular, four-variate normal limiting distribution with zero mean vector and dispersion matrix estimated by

\[
\Sigma_2(p_b) = \frac{1}{8} \cdot \bar{\beta}_{m=3} \Sigma_1(p_b) \bar{\beta}_{m=3}'
\]
evaluated in the last section of Table 3. The variances and covariances of \( \Sigma_2(p_b) \) are much like those of \( \Sigma_1(p_a) \).
Table 3
Matrix Calculations for Situation (b)

<table>
<thead>
<tr>
<th>$\Lambda(p_b)$</th>
<th>0.0609</th>
<th>-0.0089</th>
<th>-0.0088</th>
<th>-0.0088</th>
<th>-0.0087</th>
<th>-0.0087</th>
<th>-0.0083</th>
<th>-0.0083</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.0611</td>
<td>-0.0088</td>
<td>-0.0088</td>
<td>-0.0088</td>
<td>-0.0087</td>
<td>-0.0084</td>
<td>-0.0083</td>
<td>-0.0083</td>
</tr>
<tr>
<td></td>
<td>0.0618</td>
<td>-0.0089</td>
<td>-0.0089</td>
<td>-0.0089</td>
<td>-0.0086</td>
<td>-0.0086</td>
<td>-0.0088</td>
<td>-0.0088</td>
</tr>
<tr>
<td></td>
<td>0.0618</td>
<td>-0.0089</td>
<td>-0.0089</td>
<td>-0.0087</td>
<td>-0.0086</td>
<td>-0.0088</td>
<td>-0.0088</td>
<td>-0.0088</td>
</tr>
<tr>
<td></td>
<td>0.0619</td>
<td>-0.0089</td>
<td>-0.0088</td>
<td>-0.0088</td>
<td>-0.0087</td>
<td>-0.0088</td>
<td>-0.0088</td>
<td>-0.0088</td>
</tr>
<tr>
<td></td>
<td>0.0607</td>
<td>-0.0089</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(symmetric)</td>
<td>0.0606</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Sigma_1(p_b)$</th>
<th>7.1892</th>
<th>3.6239</th>
<th>0.0225</th>
<th>-3.5428</th>
<th>0.0023</th>
<th>-3.5630</th>
<th>-0.0833</th>
<th>-3.6487</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7.1747</td>
<td>-3.5377</td>
<td>0.0131</td>
<td>-3.5549</td>
<td>-0.0041</td>
<td>-3.6330</td>
<td>-0.0821</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.0980</td>
<td>3.5378</td>
<td>0.0256</td>
<td>-3.5346</td>
<td>-0.0257</td>
<td>-3.5859</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.0937</td>
<td>-3.5315</td>
<td>0.0244</td>
<td>-3.5753</td>
<td>-0.0194</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.0883</td>
<td>3.5311</td>
<td>-0.0018</td>
<td>-3.5590</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.0901</td>
<td>-3.5514</td>
<td>0.0075</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(symmetric)</td>
<td>7.2101</td>
<td>3.6605</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.2270</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Sigma_2(p_b)$</th>
<th>1.7993</th>
<th>0.0269</th>
<th>0.0023</th>
<th>-0.0054</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.7883</td>
<td>0.0016</td>
<td>-0.0037</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(symmetric)</td>
<td>1.7790</td>
<td>-0.0003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.7797</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5. CONCLUDING REMARKS

(i). The ability to consider treatment contrasts in paired comparisons brings much additional flexibility in practice to the use of paired comparisons. A primary application is to factorial treatment combinations as seen in the examples of Section 4. The asymptotic theory developed in this paper provides the justification for analyses given by Bradley and El-Helbawy (1976) and for examination of asymptotic power and distributions of estimators. As seen from the applications of Section 4, optimal designs for factorials in paired comparisons may be considered now; it may well be possible to develop general results.

(ii) In considering treatment contrasts, we have rescaled the parameters and the reader must be warned that they differ from those of Bradley (1955) and Davidson and Bradley (1970). While this does not affect the likelihood ratio, it does affect the dispersion matrices of the asymptotic distributions of estimators. We illustrate for the two basic distributions.

Consider two sets of parameters, \( p_1, \ldots, p_t, \Sigma \log p_i = 0 \) and \( p_1^*, \ldots, p_t^*, \Sigma p_i^* = 1 \). The probabilities of (1.1) have the same form for both sets of parameters. The two sets of parameters are related: \( p_i^* = \pi_i / \Sigma \pi_i, p_i = \pi_i^* / (\Sigma \pi_i^*)^{1/t} \), \( i = 1, \ldots, t \). Corresponding estimator sets, \( p_{1i}, \ldots, p_{ti} \), and \( p_{1i}^*, \ldots, p_{ti}^* \), are subject to the corresponding conditions and relationships. We have shown above, Theorems 2 and 3, that \( \sqrt{N} \left( \log p - \log \pi \right) \) has an asymptotic dispersion matrix obtainable from (2.11) or equivalently from

\[
\begin{bmatrix}
\Lambda(p) & 1 \\
1 & 0
\end{bmatrix}^{-1}
\]
when no parameter assumptions such as (1.5) and (1.7) are applied. Suppose that we desire the similar result for $\sqrt{N} (p^*_t - \pi^*_t)$. If the dispersion matrix of $\sqrt{N} (p^*_t - \pi^*_t)$ is $\Sigma^*$ and that of $\sqrt{N} (\log p - \log \pi)$ is $\Sigma$, then $\Sigma^* = T \Sigma^* T'$ where $T = [t_{ij}, i, j = 1, \ldots, t]$ = $[\partial \log \pi_i / \partial \pi^*_j, i, j = 1, \ldots, t]$ and $t_{ii} = \left(\frac{t-1}{t}\right) / \pi_i^*$, $t_{ij} = -1/t \pi_j^*$, $i \neq j$, $i, j = 1, \ldots, t$. It may be verified through change in the differentiation in (2.10) that $\Sigma^*$ is the t-square principal minor of

$$\begin{bmatrix} T' \Lambda(\pi^*) T & 1' \\ 1 & 0 \end{bmatrix}^{-1}.$$

This follows since $\Lambda(\pi) = \Lambda(\pi^*)$. Comparison of this result with formula (2.6) of Bradley (1976) shows that the two are identical when $T' \Lambda(\pi^*) T$ is evaluated.

(iii). Bradley (1955) and Davidson and Bradley (1970) gave rather weak arguments in establishing the rank of matrices corresponding to $\Lambda(\pi)$. Appendix Theorem A1 may be used to establish rigorously the ranks of their matrices.
REFERENCES


_Amer. Math. Monthly_ 64 (October, Part II), 28-33.

Good, I. J. (1967). Discussion on "Topics in the investigation of linear relations 

Gainesville.

APPENDIX

The following three theorems have been used in the paper.

**Theorem A1:** Let $\Lambda = [\Lambda_{ij}, \ i, j = 1, \ldots, t]$ satisfy the conditions,

(i) \[ \Lambda_{ii} > 0, \ i = 1, \ldots, t, \]

(ii) \[ \Lambda_{ij} < 0, \ i \neq j, \ i, j = 1, \ldots, t, \]

(iii) \[ \sum_j \Lambda_{ij} = 0, \ i = 1, \ldots, t, \]

(iv) \[ \Lambda_{ij} = 0 \iff \Lambda_{ji} = 0, \ i \neq j, \ i, j = 1, \ldots, t, \]

and

(v) \[ \Lambda \] cannot be expressed in the form,

\[
\Lambda = \begin{bmatrix}
\Lambda_{11} & 0_{12} \\
0_{21} & \Lambda_{22}
\end{bmatrix},
\]

by rearrangement of rows and columns where $0_{12}$ and $0_{21}$ are null matrices of appropriate dimensions. Given these conditions, $\Lambda$ has rank $(t-1)$.

**Proof:** If $t = 2$, the result is trivial. If $t > 2$, it is shown through use of the Gaussian elimination procedure that $\Lambda$ has the same rank, namely $(t-1)$, as the upper triangular matrix, $\Lambda^{(t-1)} = [\Lambda_{ij}^{(t-1)}, i, j = 1, \ldots, t]$, where $\Lambda_{ii}^{(t-1)} > 0$, $i = 1, \ldots, (t-1)$, and $\Lambda_{tt}^{(t-1)} = 0$.

Since $\Lambda_{11} > 0$, let $\Lambda^{(1)} = [\Lambda_{ij}^{(1)}, i, j = 1, \ldots, t]$ where

\[ \Lambda_{ij}^{(1)} = \Lambda_{ij}, \ j = 1, \ldots, t, \]

and

\[ \Lambda_{ij}^{(1)} = \Lambda_{ij} - \frac{\Lambda_{ii} \Lambda_{ij}}{\Lambda_{11}} , \ i = 2, \ldots, t, \ j = 1, \ldots, t. \ (A1) \]
Let \( \bar{A} \) be the \((t-1)\)-square, lower principal minor of \( A^{(1)} \).

We show that the elements of \( \bar{A} \) satisfy conditions (i)-(v). From (ii), (ii) and (iii), \( A_{i1}^{(1)} \geq A_{ii} + A_{i1} \geq 0, \ i = 2, \ldots, t \). Suppose that \( A_{i1}^{(1)} = 0 \). Then \( A_{ii} = A_{i1} \), by (iii) \( A_{ij} = 0, \ j \neq 1, \ i \), and by (iv) \( A_{ji} = 0, \ j \neq 1, \ i \). In addition, from (Al), \( \Lambda_{ii} = -\Lambda_{11} > 0 \) implies that \( \Lambda_{i1}^{(1)} = -\Lambda_{11} \left(1 + \frac{A_{11}}{\Lambda_{11}}\right) = 0 \) only if \( \Lambda_{11} = -\Lambda_{11} \) and \( \Lambda_{ij} = \Lambda_{ji} = 0, \ j \neq 1, \ i \). Thus (v) is contradicted, \( \Lambda_{ii}^{(1)} > 0 \), and (i) follows for \( \bar{A} \). It is obvious from (Al) that \( \Lambda_{ij}^{(1)} < \Lambda_{ij}, \ i, j = 2, \ldots, t \), and (ii) follows for \( \bar{A} \). That (iii) follows also for \( \bar{A} \) results directly from (Al) and (iii). To prove (iv) for \( \bar{A} \), note again that \( \Lambda_{ij}^{(1)} \leq \Lambda_{ij} < 0, \ i \neq j, \ i, j = 2, \ldots, t \). If \( \Lambda_{ji}^{(1)} = 0 \), \( \Lambda_{ji} = \Lambda_{ij} = 0 \), and either \( \Lambda_{ji} \) or \( \Lambda_{ij} \) is zero with the consequence that \( \Lambda_{ij} \) or \( \Lambda_{ji} \) is zero and \( \Lambda_{ij}^{(1)} = 0 \). If \( \Lambda_{ji}^{(1)} < 0 \), either \( \Lambda_{ij} < 0 \), so that \( \Lambda_{ij} < 0 \) and \( \Lambda_{ij}^{(1)} < 0 \) or \( \Lambda_{ji} = 0 \) with \( -\Lambda_{ji} \Lambda_{11} < 0 \), and then \( -\Lambda_{ji} \Lambda_{11} < 0 \) and \( \Lambda_{ij}^{(1)} < 0 \). To prove (v), suppose that (v) is violated for \( \bar{A} \). It follows immediately that (v) is also violated for \( \bar{A} \) and a contradiction results. \( \bar{A} \) has the five required properties.

The elimination procedure is repeated on \( \bar{A} \) and its successors until a final \( 2 \times 2 \) matrix is obtained satisfying conditions (i)-(v). When the process is applied to the \( 2 \times 2 \) minor, the final element must be zero. \( \bar{A}^{(t-1)} \) has now been obtained with obvious rank \((t-1)\).

Theorem A2: Let \( \Lambda \) be a \( t \times t \) symmetric positive semi-definite matrix of rank \((t-1)\) and let \( B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \) be a \( t \times t \) non-singular matrix with first row \( B_1 \) orthogonal to the columns of \( \Lambda \). Then \( B_2 \Lambda B_2^t \) is positive definite.
Proof: It follows at once that

\[
\begin{bmatrix}
0 & 0_{t-1}' \\
0_{t-1} & B_2\Lambda B_2'
\end{bmatrix}
\]

is positive semi-definite of rank (t-1). Therefore the (t-1)-square matrix

\[
B_2\Lambda B_2'
\]

is positive semi-definite of rank (t-1) and hence is positive definite.

**Theorem A3:** Let \( A \) be as in Theorem A2 and let \( B = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \end{bmatrix} \) be a \( t \times t \) orthonormal matrix with first row \( B_1 \) orthogonal to the columns of \( A \) and \( B_2 \) and \( B_3 \) being submatrices of \( B \) with \( m \) and \( (t-m-1) \) rows respectively. Let

\( \Sigma = B_3'(B_3\Lambda B_3')^{-1} B_3. \) Then \( \Sigma \) is the \( t \)-square principal minor of

\[
\begin{bmatrix}
\Lambda & [B_1', B_2'] \\
[B_1] & 0 \\
B_2 & 0
\end{bmatrix}^{-1}.
\]

(A2)

Proof: That \((B_3\Lambda B_3')^{-1}\) exists follows from Theorem A2. The proof will be complete when we show that there exist matrices \( F \) and \( G \) of dimensions \((m+1)\times t\) and \((m+1)\times(m+1)\) respectively such that

\[
\begin{bmatrix}
\Sigma & F' \\
F & G
\end{bmatrix}
\begin{bmatrix}
\Lambda & [B_1', B_2'] \\
[B_1] & 0 \\
B_2 & 0
\end{bmatrix}
= \begin{bmatrix} I_t & 0 \\
0 & I_{m+1} \end{bmatrix}.
\]

(A3)
It is necessary to solve the equations,

\[ \Sigma \Lambda + F' \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = I_t, \]  \hspace{1cm} (A4)

\[ \Sigma [B'_1, B'_2] = 0, \]  \hspace{1cm} (A5)

\[ F \Lambda + G \begin{bmatrix} B'_1 \\ B'_2 \end{bmatrix} = 0, \]  \hspace{1cm} (A6)

and

\[ F[B'_1, B'_2] = I_{m+1} \]  \hspace{1cm} (A7)

for \( F \) and \( G \). Equation (A5) results from the definition of \( \Sigma \) and the orthogonality of the rows of \( B \). Right multiplication of (A4) by \([B'_1, B'_2]\) yields

\[ F = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} [I_t - \Lambda \Sigma]. \]  \hspace{1cm} (A8)

Clearly, (A7) is satisfied with \( F \) as in (A8). It then follows from (A6) and (A8) that

\[ G = -\begin{bmatrix} B_1 \\ B_2 \end{bmatrix} [I_t - \Lambda \Sigma] \Lambda [B'_1, B'_2]. \]
<table>
<thead>
<tr>
<th>Report Number</th>
<th>Govt. Accession No.</th>
<th>Recipient's Catalog Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONR Report No. 102</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Title (and subtitle)</th>
<th>Type of Report &amp; Period Covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment Contrasts in Paired Comparisons: Large-Sample Results, Applications and Some Optimal Designs</td>
<td>Technical Report</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Contract or Grant Number(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdalla T. El-Helbawy</td>
<td>ONR N00014-76-C-0608</td>
</tr>
<tr>
<td>Ralph A. Bradley</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performing Organization Name and Address</th>
<th>Program Element, Project, Task Area &amp; Work Unit Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Florida State University</td>
<td></td>
</tr>
<tr>
<td>Department of Statistics</td>
<td></td>
</tr>
<tr>
<td>Tallahassee, Florida 32306</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controlling Office Name and Address</th>
<th>Report Date</th>
<th>Number of Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistics and Probability Program</td>
<td>December, 1976</td>
<td>33</td>
</tr>
<tr>
<td>Office of Naval Research</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arlington, Virginia 22217</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Monitoring Agency Name &amp; Address (if different from Controlling Office)</th>
<th>Security Class (of this report)</th>
<th>Declassification/Downgrading Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unclassified</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distribution Statement (of this Report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approved for public release; distribution unlimited.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distribution Statement (of the abstract entered in Block 20, if different from report)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Supplementary Notes</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Key Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired comparisons, treatment contrasts, factorial treatments, asymptotic power, optimal design, taste testing.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment contrasts in paired comparison experiments are considered and associated large-sample results developed. Likelihood estimators of treatment parameters are shown to be consistent when obtained under the assumption that specified orthonormal treatment contrasts are null. Under the same assumption, the distribution functions of treatment parameter estimators, their logarithms, and orthonormal contrasts among them, estimators</td>
</tr>
</tbody>
</table>
of treatment contrasts orthonormal also to those assumed null, when suitably scaled and centered are shown to be normal with zero mean vectors and variance-covariance matrices shown in the limit for large sample sizes. The distribution functions of likelihood-ratio statistics are shown to be chi-square in the limit with large sample sizes, central under null hypotheses and non-central under alternative hypotheses specified by a class of local alternatives with non-centrality parameters given.

Much of the paper is devoted to applications of the theoretical results. Section 4 contains five applications, all illustrated with $2^3$-factorial experiments. The first four have results on optimal designs for the use of factorial treatment combinations in paired comparisons. The final one is a numerical one on a tastepreference experiment for coffee with brew strength, roast color, and brand being three factors, each at two levels.

A final section relates the results given with earlier work by Bradley. The model used here differs from that proposed earlier only in a scale-determining constraint on treatment parameters but some care is necessary in comparisons of results. An appendix contains some matrix results used in the paper and of potential use in other statistical research.

The methods of this paper, and the one on basic methodology that preceded it, provide much new flexibility to the use of paired comparisons, particularly with factorial treatment combinations. In addition, the way seems to be open for the development of a more general theory of optimal design for factorials in paired comparisons.