A NOTE ON THE TUKEY-KRAMER PROCEDURE FOR FAIRWISE COMPARISONS
OF CORRELATED MEANS

Lawrence D. Brown†

Department of Statistics
Hebrew University, Jerusalem, Israel,
and Department of Mathematics
Cornell University, Ithaca, New York

1. MAIN RESULT

Let \( X \) have a multivariate normal distribution on \( \mathbb{R}^k \) with mean \( \mu \) and
covariance matrix \( \sigma^2Z \). Suppose \( \mu \) and \( \sigma^2 \) are unknown and \( \Sigma \) is known. Let
\( (Z^2/\sigma^2) \) be an independent chi-square random variable with \( v \) degrees of freedom.

Introduce \( q_{k,v}(a) \), the upper \( a \)th quantile of the Studentized range. By
definition \( q_{k,v}(a) \) is the value for which

\[
P(\max_{i,j} |X_i - X_j| \geq q_{k,v}(a)) = a
\]

when \( u = 0 \), \( Z = 1 \).

The purpose of this note is to prove

Theorem: Let \( k = 3 \). Let \( q = q_{k,v}(a) \) and

\[
\frac{Z}{\sqrt{\frac{2}{3}}} = (a_{11} - 2a_{13} + a_{33})/2
\]

Then

\[
Pr(\max_{i,j} |X_i - X_j| \geq q_{i,j}Z \text{ for some } i,j \leq a
\]

The following trivial corollary contains the statistical motivation for (3).

† Research supported in part by N.S.F. grant B200031.
Corollary: Let \( k = 3 \). The simultaneous confidence intervals
\[
\mu_i - \mu_j \in [X_i - X_j \pm \omega_{ij}q2] \quad (1 \leq i < j \leq k)
\]
have confidence level at least \( 1 - \alpha \).

2. DISCUSSION

One class of problems covered by the theorem is one-way analysis of variance. Here one observes,
\[
Y_{in} = \mu + \tau_i + \varepsilon_{in}, \quad n = 1, \ldots, N_i, \; i = 1, \ldots, k
\]
with \( \varepsilon_{in} \) independent normal \( (0, \sigma^2) \) variables and one sets
\[
X_i = Y_{i..} = N_i^{-1} \sum_{n=1}^{N_i} Y_{in}. \quad \text{In the balanced case one has } N_i = N. \text{ For this case Tukey (1953) proposed what has come to be called the T-procedure for simultaneously setting confidence intervals for all the contrasts } \tau_i - \tau_j. \text{ This is the special procedure defined by the intervals in (4) in the special case } \omega_{ij} = \omega = N^{-1/2}. \text{ The constant } q_k(\alpha) \text{ defined by (1) is devised precisely for this classic case, and has been extensively tabulated.}

Tukey (1953) and Kramer (1956) proposed the confidence intervals in (4) as the natural extension of the T-procedure for cases where the \( N_i \) are not necessarily equal. Here \( \Sigma \) is diagonal with diagonal elements \( N_i^{-1} \) so that
\[
\omega^2_{ij} = \left( N_i^{-1} + N_j^{-1} \right) / 2.
\]
These proposals were based on a conjecture that (3) is valid, or nearly valid. When \( k = 3 \) Kurtz (1956) proved that (3) is valid in this situation. Kurtz' proof is the foundation for the proof of our main theorem, which of course includes also the case of non-diagonal \( \Sigma \).

Dunnett (1980) provided extensive numerical evidence of the validity of (4) when the \( \omega_{ij} \) are given by (6). Brown (1979) then proved the validity of (4) in this case when \( k = 4, 5 \). Recently Heyter (1984) has proved the validity of (4) for all values of \( k \) when the \( \omega_{ij} \) are given by (6).

In analysis of covariance one observes
\[
Y_{in} = \mu + \tau_i + b_i \beta + e_{in}
\]
Tukey-Kramer Procedure

with \( \mu, \tau, \epsilon_i \) as before, \( b_x \) known constants and \( \theta \) an unknown constant. Then the least squares estimates \( \hat{x}_1 = \hat{\theta} + \hat{\tau}_1 \) of \( \mu + \tau_i \) have covariance matrix \( \sigma^2 Z \). \( Z \) is not, in general, a diagonal matrix. For such situations Tukey (1953) and Kramer (1957) again proposed the multiple comparison procedure described by (4) for comparisons involving all contrasts \( \tau_i - \tau_j \). Our main theorem validates this proposal when \( k = 3 \).

There are linear models other than the analysis of covariance, (7), which lead to the above canonical model with \( Z \) non-diagonal. One such model is that of an unbalanced two-way design mixed model (blocks random, treatments fixed) in which simultaneous confidence intervals are desired for all treatment contrasts.

3. PROOF OF THE THEOREM

Let \( \beta_1 = (1,-1,0)' \), \( \beta_2 = (1,0,-1)' \), \( \beta_3 = (0,1,1)' \). Assume, without loss of generality, that \( \mu = 0 \). Then the left side of (3) can be rewritten as

\[
\Pr[|\beta_i'X| \geq (\beta_i'Z\beta_i/2)^{1/2}qZ, \text{ some } i = 1,2,3].
\]

Let \( Z = C'C, Y = C^{-1}X \), and \( \gamma_i = CB_i \). Then \( Y \) is multivariate normal, mean \( \hat{\theta} \), covariance \( \sigma^2 I \). The expression (8) equals

\[
\Pr[|Y_i'Y_i| \geq (\gamma_i'\gamma_i/2)^{1/2}qZ, \text{ some } i = 1,2,3].
\]

Note that \( \{Y_1, Y_2, Y_3\} \) spans a two-dimensional subspace of \( R^3 \), since \( \beta_2 - \beta_1 = \beta_3 \) so that \( Y_2 - Y_1 = Y_3 \). We may thus assume, without loss of generality, that \( C \) has been chosen to yield

\[
(Y_1)_3 = 0, \quad i = 1,2,3.
\]

Then \( Y_3 \) plays no role in the expression (9) (i.e., \( Y_1'Y \) depends only on \( Y_1, Y_2 \)).

The conditional distribution of \( (Y_1, Y_2) \) given \( Y_1^2 + Y_2^2 = r^2 \) is uniform on the circle of radius \( r \). As Kurs (1956) noted, a simple geometric argument then shows that

\[
\Pr[|Y_i'Y| \geq t ||Y_i||, \text{ some } i = 1,2,3, Y_1^2 + Y_2^2 = r^2].
\]
is maximized (for every $t, r$) by choosing $\gamma_1, \gamma_2, \gamma_3$ to form equal angles (i.e., $\gamma_1 = (1,0,0)^t$, $\gamma_2 = (+1/2, -\sqrt{3}/2, 0)^t$, $\gamma_3 = (-1/2, -\sqrt{3}/2, 0)^t$, or any rotation of these preserving the condition (10)). In fact for this choice of $(\gamma_i)$ the probability in (11) equals $\min(\delta \cos^{-1}(t/2\pi), 1)$ since the sets $[t - \gamma_i \parallel\gamma_i\parallel, i = 1, 2, 3$ overlap only when $t \leq 1/2$, and then the probability in (11) is one.

The expression (9) results from setting $Z = q_i / 2$ and taking the expectation of (11) over the distribution of $Z$ and $R^2 = \gamma_1^2 + \gamma_2^2$. It is thus maximized by taking $\gamma_1, \gamma_2, \gamma_3$ to form equal angles in a two-dimensional plane. This is equivalent to choosing $Z = 1$ in (8). When $Z = 1$ the probability in (8) is exactly $a$, by definition of $q_i$.

4. GENERALIZATIONS: REMARKS ON THE GEOMETRY OF THE PROOF

A look at the proof shows that the manipulations leading to (11) reduce the theorem to a geometric question on the unit circle. (Take $r = 1$ without loss of generality.) When $k \geq 4$ the problem can be similarly rephrased as a geometric question on the $k-1$-dimensional unit sphere. For example, when $k = 4$ the theorem can be reduced as follows to a geometric question:

Begin by defining $\beta_1 = (1,-1,0,0)^t$, $\beta_2 = (1,0,-1,0)^t$, $\beta_3 = (1,0,0,-1)^t$, ..., $\beta_6 = (0,0,1,-1)^t$. $Z = CIC$, and $\gamma_4 = \beta_4$, as in the preceding proof. To simplify the picture, as before, note that $C$ can be chosen to yield $(\gamma_1)_k = 0$, $i = 1, \ldots, 6$; and then define $\gamma_i = (\gamma_{11}, \gamma_{12}, \gamma_{13})^t$ and define the unit vectors in $R^3$, $\gamma_i = \gamma_i / \parallel \gamma_i \parallel$, $i = 1, \ldots, 6$. Note that the set of vectors $(\gamma_i)$ is constrained by three linear relations, e.g.

$$\parallel \gamma_3 \parallel p_3 - \parallel \gamma_2 \parallel p_2 = \parallel \gamma_6 \parallel p_6$$

since $\beta_3 - \beta_2 = \beta_6$), and by the angular relations

$$p_i^2 p_j = (\beta_i^t Z \beta_j)^{1/2}(\beta_i^t Z \beta_j)^{1/2}, \quad 1 \leq i, j \leq 6. \quad (13)$$

These relations will be examined later from a geometric perspective. Note that $Z$ plays an important role in (13) and also in (12). (In fact, $(\rho_i)$ is uniquely determined, up to rotation, by $Z$ through (13).)

Let $A_3(S)$ denote the area of a set $S$ on the unit sphere in $R^3$. Then the theorem for $k = 4$ would be valid if

$$A_3\{x: \parallel x \parallel = 1, \mid p_i^2 \mid \geq t, \text{some } i = 1, \ldots, 6\} \quad (14)$$
Tukey-Kramer Procedure

is maximized by the choice \( Z = 1 \) for each \( 0 \leq t \leq 1 \). Note that each of the sets \( \{ x : \| x \| = 1, \rho_1 x \leq t \} \) or \( \{ x : \| x \| = 1, -\rho_1 x \leq t \} \) appearing in (14) is the interior of a circle on the unit sphere.

The conjecture that (14) is maximized for all \( k \) by the choice \( Z = 1 \) is sufficient to prove the theorem for \( k = 4 \), but it is not necessary, since the expression in (9) corresponds to an expectation of the area in (14) with \( t^2 \) a rescaled \( F \) random variable \( \sum_{i=1}^{2} Y_i^2 \). Indeed, the proofs of Brown (1979) and Hayter (1984) for the case \( Z \) diagonal, as in (6), proceed without establishing the truth of the conjecture.

The conditions (12), (13) can be rephrased geometrically as follows:

(i) There are 4 2-dimensional planes containing the origin.
(ii) Each plane contains 3 of the 6 vectors \( \{ \rho_i \} \).
(iii) Each \( \rho_i \) lies in 2 of these planes.

Thus, the set \( \{ \rho_i \} \) is uniquely determined as the set of unit vectors determined by the intersections of four planes through the origin. Different choices of \( Z \) lead to all possible configurations for these planes.

The assumption that \( Z \) is diagonal, as in (6), corresponds to an additional geometric condition as follows: To each \( \rho_i \) there corresponds a unique vector \( \rho_j(i) \) not lying in either of the two planes in (15) which contain \( \rho_i \). If \( Z \) is diagonal if and only if

\[
\rho_j(i) \rho_i = 0, \quad i = 1, \ldots, 6.
\]

When \( Z = 1 \) the vectors \( \{ \rho_i \} \) of course satisfy the conditions in (15) and (16). Furthermore, the three lines \( \rho_j(i) \) lying in each plane in (15) cut the unit circle of that plane into six equal segments. The convex body with vertices \( \rho_j(i), i = 1, \ldots, 6 \) is a well-known semi-regular convex body with 14 faces—six squares and eight equilateral triangles, all with edges of length 1; each square is bordered by four triangles and each triangle by three squares.

After drawing pictures and examining cardboard constructions, I (and others I have talked with) believe that, subject to (15) and (16) the area in (14) is indeed maximized by the choice \( Z = 1 \). It also seems reasonable that (14) is maximized by \( Z = 1 \) subject to (15). However, it appears to be quite difficult to establish the truth (or falsity) of these geometric conjectures.

SUMMARY

This note contains a proof that the Tukey-Kramer multiple comparison procedure for three correlated means (with known correlations) is conserva-
tive. The geometry of the proof is then discussed in the hope that a proof using it can be found for more than three means.

REFERENCES


