UNIVERSITY OF GÖTEBORG

Department of Statistics

RESEARCH REPORT 1992:3

ISSN 0349-8034

Exact Semiparametric Inference About the Within-Subject Variability in 2 x 2 Crossover Trails.

by

Olivier Guilbaud

Statistiska institutionen

Göteborgs Universitet
Viktoriagatan 13
S 411 25 Göteborg
Sweden

Exact Semiparametric Inference About the Within-Subject Variability in 2 x 2 Crossover Trials.

Olivier Guilbaud
Kabi Pharmacia
S-112 87 STOCKHOLM
SWEDEN

ABSTRACT. The comparison of primary interest in a 2 x 2 crossover trial typically concerns the effect of the treatments, say A and B, on the mean response level. This article deals with another important aspect, namely the within-subject response variability under A and B. Differences in drug formulation and/or administration may lead to considerable differences in within-subject variability, whatever is the difference in terms of mean level; and consideration of both these aspects may therefore be of considerable importance for the judgement of the treatments. It is shown that, although there are no within-subject treatment replications, it is possible to make various exact inferences about the A/B ratio of within-subject variances and about the (A - B)-difference in mean level, simultaneously and marginally. These inferences are semiparametric in that no distributional assumption is made about the between-subject variability, whereas a normality assumption is used for the within-subject variability. The inferences include tests, confidence regions, and a multiple test procedure. A power approximation is also given. The results are illustrated numerically.

KEYWORDS. Kendall's tau, simultaneous inferences, multiple test procedure, bioequivalence.

Olivier Guilbaud is Statistician, Kabi Pharmacia, S-112 87 Stockholm, Sweden; and Adjunct Professor of Biostatistics, University of Stockholm. Parts of the results reported in this article were presented at the joint meeting of the Society for Clinical Trials and the International Society for Clinical Biostatistics in Brussels, July 1991.

made concerning this ratio. These inferences are exact in that significance probabilities, test sizes and confidence coefficients are exact. Moreover, the inferences are semiparametric in that no distributional assumption is made about the "between" random effects, whereas the "within" random effects are assumed to be normally distributed. The within-subject variability considered here can in principle be thought of as being composed of several components of interest (Ekbohm and Melander 1989, 1990), but such components are of course not identifiable with a 2 x 2 crossover design.

Recently Cornell (1991) proposed a nonparametric test of the particular hypothesis $\sigma^2_A = \sigma^2_B$ based on Kendall's tau applied to within-subject sums and differences of responses in each group. The test is asymptotically distribution-free, not exact; and no corresponding nonparametric confidence region for σ^2_A/σ^2_B is available. In contrast, the various semiparametric inferences proposed in this article are exact, and they include tests of more general hypotheses than $\sigma^2_A/\sigma^2_B = 1$, as well as confidence regions for σ^2_A/σ^2_B. These inferences are also based on Kendall's tau, but more general linear combinations of the responses are considered. It is an open problem whether/how exact inferences can be made under essentially weaker distributional assumptions for the within-subject variability. For some results on related problems concerning differences between the marginal distributions of a bivariate distribution, see Kepner and Randles (1982) with references.

The assumptions, the notation, and some basic results are given in sec. 2 and 3. Various inferences are then given in sec. 4, including: exact tests and confidence regions for σ^2_A/σ^2_B, a related point estimate, and certain exact simultaneous confidence regions and multiple test procedures concerning differences both in mean level and in within-subject variability. An illustration is given in sec. 5. In sec. 6, some concluding remarks are made, and some additional results are briefly mentioned, including a power approximation. The Appendix contains some technical details.

2. ASSUMPTIONS AND NOTATION

Let Y_{ij1} and Y_{ij2} denote the response observations from period 1 and period 2, respectively, for subject $j = 1, ..., n_i$ in sequence group $i = 1, 2$; with each $n_i \geq 2$. Moreover, let $n_* = n_1 + n_2$, and let group 1 correspond to treatment sequence AB, and group 2 to BA. For convenience, no notational distinction is made between random quantities and the corresponding realizations in this article. The normal distribution with mean a and variance b is denoted N(a, b) and Φ denotes the distribution function of N(0, 1). The t-distribution with f degrees of freedom and noncentrality parameter δ (Owen 1985) is denoted T(f, δ).

It is assumed that the response vector (Y_{ij1}, Y_{ij2}) for subject $j = 1, ..., n_i$ in group $i = 1, 2$ can be represented as

$$(Y_{ij1}, Y_{ij2}) = (\mu_{i1}, \mu_{i2}) + (\xi_{ij} + \varepsilon_{ij1}, \ \xi_{ij} + \varepsilon_{ij2}) \qquad (2.1)$$

where: (a) (μ_{i1}, μ_{i2}) is a nonrandom vector reflecting fixed effects such as direct treatment effects, period effects and indirect treatment effects; (b) ξ_{ij} is a random variable reflecting the between-subject variability; and (c) $(\varepsilon_{ij1}, \varepsilon_{ij2})$ is a random vector reflecting the within-subject variability in each period. No restriction is put on the two vectors (μ_{i1}, μ_{i2}), $i = 1, 2$, in (2.1), and no assumption is made about the functional relationship between these vectors and the fixed effects that they reflect. Moreover, no assumption is made about the joint distribution of the n_* "between" variables ξ_{ij} in (2.1). In particular, they are not assumed to be independent or identically distributed.

The essential distributional assumption in this article concerns the n_* "within" vectors $(\varepsilon_{ij1}, \varepsilon_{ij2})$ in (2.1). These vectors are assumed to be mutually independent and independent of the n_* "between" variables ξ_{ij}. Moreover, with $=_d$ denoting equality in distribution, it is assumed that, for $i = 1, 2$ and $j = 1, ..., n_i$,

$$\begin{aligned}(\varepsilon_{ij1}, \varepsilon_{ij2}) &=_d (\varepsilon_A, \varepsilon_B), \qquad i = 1,\\ &=_d (\varepsilon_B, \varepsilon_A), \qquad i = 2,\end{aligned} \qquad (2.2)$$

where $\varepsilon_A \sim N(0, \sigma^2_A)$ and $\varepsilon_B \sim N(0, \sigma^2_B)$ are independent with $0 < \sigma^2_A < \infty$ and $0 < \sigma^2_B < 0$. Here the index $_A$ or $_B$ indicates the treatment in the relevant period. In terms of these within-subject variances, define

$$\theta = \sigma^2_A/\sigma^2_B \qquad \sigma^2_{AB} = \sigma^2_A + \sigma^2_B \tag{2.3}$$

Thus, the ratio θ is a measure of the relative variability within subjects under the two treatments. The problem is to make exact inferences about θ, or, equivalently, about

$$\gamma = (\sigma^2_A - \sigma^2_B)/\sigma^2_{AB} = (\theta - 1)/(\theta + 1) \tag{2.4}$$

under these assumptions. Clearly, γ is an increasing function of $\theta > 0$, and $-1 < \gamma < 1$.

The inferences about γ proposed in this article are based on the within-subject sums and crossover (A - B)-differences,

$$Y_{ij+} = Y_{ij1} + Y_{ij2}, \tag{2.5}$$

$$\begin{aligned} Y_{ij-} &= Y_{ij1} - Y_{ij2}, \qquad i = 1, \\ &= Y_{ij2} - Y_{ij1}, \qquad i = 2; \end{aligned} \tag{2.6}$$

which are basic also for inferences about fixed effects. For instance, the ordinary t-based inferences (Jones and Kenward 1989, chap. 2) about additive fixed period effects and direct treatment effects are based on the differences (2.6) through

$$\overline{Y}_{i-} = \Sigma_j Y_{ij-}/n_i, \qquad s^2_{i-} = \Sigma_j (Y_{ij-} - \overline{Y}_{i-})^2/(n_i - 1), \tag{2.7}$$

$i = 1, 2$; with $s^2_{1\cdot}$ and $s^2_{2\cdot}$ usually combined into a pooled estimate of $\mathrm{Var}(Y_{ij\cdot})$. Note that $\mathrm{Var}(Y_{ij\cdot})$ equals σ^2_{AB} in (2.3), and that these ordinary inferences about fixed effects do <u>not</u> require that θ in (2.3) is equal to 1 as is commonly assumed.

It is convenient to introduce the function

$$
\begin{aligned}
G(c) &= 0 && \text{if } c \le -1, \\
&= (1+c)/(1-c) && \text{if } -1 < c < 1, \\
&= \infty && \text{if } c \ge 1
\end{aligned}
\tag{2.8}
$$

which, for $-1 < c < 1$, equals the inverse of the function $\gamma = \gamma(\theta)$ given by (2.4).

3. Basic Results

3.1 Linear Combinations of Response Observations

For any given real c, let $Y_{ij[c]}$ be defined in terms of (2.5) - (2.6) by

$$Y_{ij[c]} = Y_{ij+} - c\,Y_{ij-}. \tag{3.1}$$

The association between $Y_{ij[c]}$ and $Y_{ij\cdot}$ depends on the value of c, and the idea is to use this dependence on c to make inferences about γ. A key result in this context is that the vector $(Y_{ij\cdot}, Y_{ij[c]})$ can be expressed in terms of quantities in (2.1) - (2.4) as

$$
\begin{aligned}
(Y_{ij\cdot}, Y_{ij[c]}) &= (\mu_{i\cdot}, \mu_{i[c]}) + \sigma_{AB}[\,U_{ij}, (\gamma - c)U_{ij} + (1-\gamma^2)^{1/2} Z_{ij}] \\
&\quad + (0, 2\,\xi_{ij}),
\end{aligned}
\tag{3.2}
$$

$j = 1, \ldots, n_i$, $i = 1, 2$. Here $\mu_{i\cdot}$ and $\mu_{i[c]}$ are constants defined formally as $Y_{ij\cdot}$ and $Y_{ij[c]}$ through (2.5), (2.6) and (3.1) with Y_{ij1} and Y_{ij2} replaced by μ_{i1} and μ_{i2} in (2.1), respectively; whereas the vector (U_{ij}, Z_{ij}) equals $(\varepsilon_{ij1}/\sigma_A, \varepsilon_{ij2}/\sigma_B)\mathbf{M}$ for $i = 1$, and equals $(\varepsilon_{ij2}/\sigma_A, \varepsilon_{ij1}/\sigma_B)\mathbf{M}$ for $i = 2$, where **M** is the 2 x 2 orthogonal matrix with elements $(m_{11}, m_{12}, m_{21}, m_{22})$ equal to $(1+\theta)^{-1/2}(\theta^{1/2}, 1, -1, \theta^{1/2})$. This implies that, in the right hand side of (3.2), the $2n_*$ variables U_{ij} and Z_{ij} are: (a) each $N(0, 1)$ distributed; (b) mutually independent; and (c) independent of the n_* "between" variables ξ_{ij}.

It is now evident from (3.2) that the n_* variables Y_{ij-} are mutually independent, with $Y_{ij-} \sim N(\mu_{i-}, \sigma^2_{AB})$; and that, if $c = \gamma$, these n_* variables are independent of the n_* variables $Y_{ij[c]}$. It is also evident that, if $c \neq \gamma$, the Y_{ij-}'s are not independent of the $Y_{ij[c]}$'s. Actually it follows from (3.2) that if $c < \gamma$ ($c > \gamma$) then, marginally, each $Y_{ij[c]}$ is positively (negatively) regression dependent on Y_{ij-}; that is (Lehmann 1966), the conditional distribution of $Y_{ij[c]}$ given $Y_{ij-} = y$ increases (decreases) stochastically as y increases. The joint behavior of the n_* variables $Y_{ij[c]}$ depends of course on the joint behavior of the "between" ξ_{ij}'s, about which no assumption has been made.

3.2 Kendall tau Statistics

For $i = 1, 2$, let $T_i(c)$ denote Kendall's tau statistic (Kendall and Gibbons 1990) based on the n_i vectors $(Y_{ij-}, Y_{ij[c]})$ for a given c. That is, $T_i(c) = S_i(c)/N_i$, where

$$S_i(c) = \Sigma\Sigma_{r<s} \operatorname{sign}[(Y_{is-} - Y_{ir-})(Y_{is[c]} - Y_{ir[c]})] \qquad (3.3)$$

and $N_i = n_i(n_i - 1)/2$, $i = 1, 2$. In (3.3), the double sum is over the N_i distinct pairs (r, s) of indices satisfying $1 \leq r < s \leq n_i$, and sign[x] equals -1, 0, 1 as $x < 0$, $x = 0$, $x > 0$, respectively.

For $i = 1, 2$ and $1 \leq r < s \leq n_i$,

$$C_{irs} = (Y_{is+} - Y_{ir+})/(Y_{is-} - Y_{ir-}) \qquad (3.4)$$

equals the slope of the straight line connecting the two points (Y_{ir-}, Y_{ir+}) and (Y_{is-}, Y_{is+}) in group i. There are thus $N_i = n_i(n_i - 1)/2$ slopes (3.4) in group i, and it is easily shown using (3.1) that the right hand side of (3.3) equals that of

$$S_i(c) = \Sigma\Sigma_{r<s} \operatorname{sign}[C_{irs} - c] \qquad (3.5)$$

with probability one. The $N_1 + N_2$ slopes (3.4) from the two groups are basic for the exact confidence regions and the point estimate considered in sec. 4.2 and 4.3.

For $i = 1, 2$, the tau statistic $T_i(c) = S_i(c)/N_i$ has a discrete distribution concentrated on the set

$$Q_i = \{t = s/N_i;\; s = -N_i, -N_i + 2, \ldots, N_i - 2, N_i\} \tag{3.6}$$

of $N_i + 1$ values in [-1, 1], and the random vector $(T_1(c), T_2(c))$ can assume any given value (t_1, t_2) in $Q_1 \times Q_2$ with positive probability. The distribution of $(T_1(c), T_2(c))$ depends only on n_1 and n_2 if $c = \gamma$; whereas if $c \neq \gamma$, it depends also on c, γ, σ_{AB} and on the joint distribution of the n_* "between" ξ_{ij}'s.

The distribution of $(T_1(\gamma), T_2(\gamma))$ is particulary simple because: (a) $T_1(\gamma)$ and $T_2(\gamma)$ are independent; and (b) for $i = 1, 2$, $T_i(\gamma)$ is distributed according to the ordinary Kendall's tau null distribution which is completely determined by the group size n_i. The distributional results (a) and (b) just mentioned hold conditionally, given the (untied) ranks of the $Y_{ij[c]}$'s in (3.3) within each group, which is actually why (a) and (b) hold unconditionally. The $T_i(\gamma)$ distribution is symmetric about zero with variance

$$v_i = (2n_i + 5)/(9N_i); \tag{3.7}$$

and $T_i(\gamma)/v_i^{1/2} \rightarrow_d N(0, 1)$ as $n_i \rightarrow \infty$. Further details about this null distribution of Kendall´s tau are given in Kendall and Gibbons (1990).

3.3 A Combined tau Statistic

It is then natural to combine the tau statistics from the two groups into a single statistic for inferences about γ. The combination $T_*(c)$ considered here is a weighted sum with weights inversely proportional to the variances v_i; that is

$$T_*(c) = \Sigma_i\, w_i\, T_i(c), \tag{3.8}$$

where $w_i = v_i^{-1}/(v_1^{-1} + v_2^{-1})$ and $T_i(c) = S_i(c)/N_i$, with v_i given by (3.7) and $S_i(c)$ given by (3.3) or (3.5), $i = 1, 2$. Such a weighted sum of tau statistics was considered by Korn (1984) and Taylor (1987) to combine information from independent blocks.

The $T_*(c)$ distribution is concentrated on the set $Q_* = Q_*(n_1, n_2)$ of values in [-1, 1] defined in terms of (3.6) by

$$Q_* = \{q = w_1 t_1 + w_2 t_2;\ (t_1, t_2) \in Q_1 \times Q_2\}, \tag{3.9}$$

and $T_*(c)$ can assume any given value $q \in Q_*$ with positive probability. This set Q_* may consist of relatively few values. For instance, in the important balanced case with $n_1 = n_2 = n$ and $N_1 = N_2 = N = n(n - 1)/2$, each weight w_i equals 1/2, and $Q_*(n, n)$ then consists of the $2N + 1$ values $q = s/N$, $s = -N, -N + 1, \ldots, N - 1, N$. The $T_*(c)$ distribution depends only on n_1 and n_2 if $c = \gamma$; whereas if $c \neq \gamma$, it depends also on c, γ, σ_{AB} and on the joint distribution of the n_* "between" ξ_{ij}'s about which no assumption has been made. An important property of the $T_*(c)$-distribution is however evident from (3.5) and (3.8), namely that $T_*(c)$ strictly decreases stochastically as c increases. This monotonicity result implies in particular that, if $c < \gamma$ ($c > \gamma$), then $T_*(c)$ is strictly larger (smaller) stochastically than if $c = \gamma$, which is important for the exact inferences considered in sec. 4.

In the sequel, the distribution of $T_*(\gamma)$ is denoted $K_0(n_1, n_2)$, and $T_0(n_1, n_2)$ denotes a random variable distributed according to $K_0(n_1, n_2)$. Of course, the distributions $K_0(n_1, n_2)$ and $K_0(n_2, n_1)$ are equal. It is also convenient to introduce the set $P_0 = P_0(n_1, n_2)$ of upper tail probabilities of $K_0(n_1, n_2)$ defined by

$$P_0 = \{p = \Pr[T_0(n_1, n_2) \geq q];\ q \in Q_*\}. \tag{3.10}$$

The distribution $K_0(n_1, n_2)$ is symmetric about zero with variance

$$v_* = (\Sigma_i v_i^{-1})^{-1}, \tag{3.11}$$

where v_i is given by (3.7); and $T_0(n_1, n_2)/v_*^{1/2} \rightarrow_d N(0, 1)$ as $n_* \rightarrow \infty$.

Exact upper tail probabilities $p \in P_0(n_1, n_2)$ are given in Table 1 for the balanced case with $n_1 = n_2 \leq 10$, and in Table 2 for the unbalanced case with $n_2 - 2 \leq n_1 < n_2 \leq 10$. The algorithm described in Kendall and Gibbons (1990, sec. 5.4) was used for the distribution of each tau component of $T_*(\gamma)$ to compute these upper tail probabilities. Corresponding lower tail probabilities are obtained using the symmetry about zero of $K_0(n_1, n_2)$. A direct application of the normal approximation $N(0, v_*)$ for $K_0(n_1, n_2)$ leads to the approximation $1 - \Phi(q/v_*^{1/2})$ for $\Pr[T_*(n_1, n_2) \geq q]$, $q \in Q_*$. In the balanced case with $n_1 = n_2 = n$ and $N_1 = N_2 = N = n(n-1)/2$, it is preferable to use a continuity correction; which leads to the approximation $1 - \Phi((q - (2N)^{-1})/v_*^{1/2})$ for $\Pr[T_*(n, n) \geq q]$, $q \in Q_*(n, n)$. This normal approximation with continuity correction is excellent even for small group sizes.

4. Exact inferences

The inferences proposed in this section are exact in the validity sense: significance probabilities, test sizes and confidence coefficients are exact. This holds without any distributional assumption about the "between" variables ξ_{ij} in (2.1), which is rather appealing. Moreover, it can be verified that the proposed tests, confidence regions and point estimate have the desirable invariance property of leading to equivalent results, given observed data, if the labels A and B for the two compared treatments are exchanged. An illustration is given in sec. 5.

4.1 Exact Tests

The tests about γ considered here are specified by a null hypothesis, an alternative hypothesis, and a rejection event, respectively, of one of the following three forms,

$$H_{0\gamma}^{(-)}: \gamma \leq \gamma_0, \quad H_{1\gamma}^{(+)}: \gamma > \gamma_0, \quad R_T^{(+)} = [T_*(\gamma_0) \geq q_p], \tag{4.1}$$

$$H_{0\gamma}^{(+)}:\ \gamma \geq \gamma_0, \quad H_{1\gamma}^{(-)}:\ \gamma < \gamma_0, \quad R_T^{(-)} = [T_*(\gamma_0) \leq -q_p], \tag{4.2}$$

$$H_{0\gamma}:\ \gamma = \gamma_0, \quad H_{1\gamma}:\ \gamma \neq \gamma_0, \quad R_T = [\,|T_*(\gamma_0)| \geq q_p], \tag{4.3}$$

Here $\gamma_0 \in (-1, 1)$ is a given value, the test statistic $T_*(\gamma_0)$ is given by (3.8) with $c = \gamma_0$, and the critical value q_p equals the value $q \in Q_*$ in (3.9) that satisfies $Pr[T_0(n_1, n_2) \geq q] = p$ for a given upper tail probability $p < 1/2$ in (3.10). The distribution $K_0(n_1, n_2)$ of $T_0(n_1, n_2)$ required in this context is described in sec. 3.3 and tabulated in Table 1 and Table 2. The test (4.1) consists in rejecting $H_{0\gamma}^{(-)}$ in favor of $H_{1\gamma}^{(+)}$ if the event $R_T^{(+)}$ occurs, and to accept $H_{0\gamma}^{(-)}$ otherwise (Lehmann 1986); and similarly for (4.2) and (4.3).

From the results in sec. 3.3 it follows that the power of the test (4.1), as well as that of the test (4.2), is $> p$ under the alternative hypothesis, and is $\leq p$ under the null hypothesis, with equality in this inequality if and only if $\gamma = \gamma_0$. In particular, each of these two one-sided tests has size equal to p, that is the maximum power under the null hypothesis is equal to p; and these two tests are strictly unbiased. The two-sided test (4.3) has size 2p, because R_T equals the union of the two disjoint rejection events in (4.1) and (4.2).

To illustrate the use of Table 1 and the accuracy of the normal approximation, suppose that the two-sided test (4.3) is to be used, and that the problem is to determine the critical value q_p that corresponds to the size 2p that is the closest possible to 0.05 in the balanced case with $n_1 = n_2 = 5$ and $N_1 = N_2 = 10$. According to Table 1, the upper tail probability $p \in P_0(5, 5)$ that is the closest to 0.025 is $p = 0.0268$, which is attained with $q = 6/10 \in Q_*(5, 5)$ in (3.9). This leads to the critical value $q_p = 6/10$ in (4.3), and the size of this exact test is then $2p = 0.054$. If instead the normal approximation with continuity correction is used, the problem is to find the value closest to 0.025 that can be assumed by $1 - \Phi((s/10 - 1/20)/v_*^{1/2})$, where $s \in [-10, 10]$ is an integer and $v_* = 1/12$. This value is 0.0284, which is attained with $s = 6$. The normal approximation thus leads to the critical value 6/10 and the (approximate) size 0.057; that is to the same critical value and almost the same size as when Table 1 is used.

Given an observed value $T_*(\gamma_0)$, the significance probability, that is the prob-value, corresponding to the tests (4.1) - (4.3) is equal to $\Pr[T_0(n_1, n_2) \geq T_*(\gamma_0)]$, $\Pr[T_0(n_1, n_2) \leq T_*(\gamma_0)]$, and $\Pr[\,|T_0(n_1, n_2)| \geq |T_*(\gamma_0)|]$, respectively; where as in sec. 3.3, $T_0(n_1, n_2)$ has distribution $K_0(n_1, n_2)$. It is evident how such prob-values can be calculated, either exactly using Table 1 and Table 2, or approximately using the normal approximation.

In case the null and alternative hypotheses of interest are formulated in terms of θ, these hypotheses should of course be reformulated into equivalent hypotheses in terms of γ using (2.4) before the tests just described are applied.

4.2 Exact Confidence Regions

Let $C_{(1)} \leq C_{(2)} \leq \cdots$ denote the $N_1 + N_2$ order statistics corresponding to the slopes (3.4) from the two groups. It is then clear from (3.5) and (3.8) that, as c varies from $-\infty$ to ∞, the value of $T_*(c)$ changes from 1 to -1 in a series of steps downward which occur at the ordered slopes. More precisely, with $\#\{C_{irs} < c\}$ and $\#\{C_{irs} \leq c\}$ denoting the number of slopes <u>in group i</u> that are $< c$ and $\leq c$ respectively, $T_*(c)$ in (3.8) can be represented as

$$T_*(c) \;=\; 1 - \Sigma_i\, (w_i/N_i)\, [\#\{C_{irs} < c\} + \#\{C_{irs} \leq c\}]; \qquad (4.4)$$

whereas the limits from the left and from the right of c, $T_*(c - 0)$ and $T_*(c + 0)$, can be represented by the right hand side of (4.4) with $[+]$ replaced by $2\#\{C_{irs} < c\}$ and $2\#\{C_{irs} \leq c\}$ respectively. The $N_1 + N_2$ order statistics $C_{(k)}$ are distinct with probability one, and the jump downward, $T_*(c - 0) - T_*(c + 0)$, of the step function (4.4) at each c equal to a slope from group i then equals $2w_i/N_i$.

For any given $p \in P_0(n_1, n_2)$ in (3.10) with $p < 1/2$, let $q_p \in Q_*$ be defined as in (4.1) - (4.3), and let $L = L(p)$ and $U = U(p)$ be the integer-valued random variables

$$L = \min\{k;\ T_*(C_{(k)} + 0) < q_p\}, \tag{4.5}$$

$$U = \max\{k;\ T_*(C_{(k)} - 0) > -q_p\}. \tag{4.6}$$

These variables satisfy $L \le U$ with probability one and, for any given $c \in (-\infty, \infty)$, the following relations among events hold,

$$[c \ge C_{(L)}] \supset [T_*(c) < q_p] \supset [c > C_{(L)}], \tag{4.7}$$

$$[c \le C_{(U)}] \supset [T_*(c) > -q_p] \supset [c < C_{(U)}]. \tag{4.8}$$

The three events in (4.7), as well as the three events in (4.8), have equal probability, because the events $[C_{(L)} = c]$ and $[C_{(U)} = c]$ have zero probability .

Now, the relations (4.7) and (4.8) hold in particular with $c = \gamma$, in which case the two middle events each have probability 1 - p by the definition of q_p. It follows that with $c = \gamma$, each of the six events in (4.7) and (4.8) has probability 1 - p, which implies that the events $[C_{(L)} \le \gamma]$, $[\gamma \le C_{(U)}]$, and $[C_{(L)} \le \gamma \le C_{(U)}]$ have probability 1 - p, 1 - p, and 1 - 2p, respectively. Thus, given observed data, the confidence coefficient associated with the confidence regions $[C_{(L)}, \infty)$, $(-\infty, C_{(U)}]$, and $[C_{(L)}, C_{(U)}]$ for γ is 1 - p, 1 - p, and 1 - 2p, respectively.

In the balanced case with $n_1 = n_2 = n$ and $N_1 = N_2 = N = n(n-1)/2$, the jump $T_*(c - 0) - T_*(c + 0)$ equals $1/N$ at each of the ordered slopes $C_{(1)} < C_{(2)} < \cdots$, and $q_p \in Q_*(n, n)$ in (3.9) is of the form $q_p = s_p/N$ for some integer $s_p \in [-N, N]$. Therefore, in this balanced case, $T_*(C_{(k)} + 0) = (N - k)/N$ and $T_*(C_{(k)} - 0) = (N - k + 1)/N$ for $k = 1, 2, ..., 2N$, and the integers L and U defined by (4.5) and (4.6) with $q_p = s_p/N \in Q_*(n, n)$ are given by $L = N - s_p + 1$ and $U = N + s_p$. Thus, $C_{(L)}$ and $C_{(U)}$ are symmetrically placed in the sequence of ordered slopes if $n_1 = n_2$, whereas this is not necessarily the case if $n_1 \ne n_2$.

Continuing the illustration with $n_1 = n_2 = 5$ and $N = 10$ in sec. 4.1, suppose that a confidence interval for γ is to be determined so that the associated confidence coefficient 1 - 2p is as close to 0.95 as possible. Then, as shown previously, Table

1 leads to the exact value $p = 0.0268$ and $q_p = s_p/N = 6/10 \in Q_*(5, 5)$, whereas the normal approximation with continuity correction leads to the approximate value $p = 0.0284$ and the same $q_p = 6/10$. Thus, according to both Table 1 and the normal approximation, q_p in (4.5) and (4.6) should be equal to $s_p/N = 6/10$, which implies $L = N - s_p + 1 = 5$ and $U = N + s_p = 16$. The confidence interval for γ then equals $[C_{(5)}, C_{(16)}]$. The exact confidence coefficient associated with this interval is $1 - 2 \times 0.0268 = 0.946$, whereas the normal approximation gives almost the same value, $1 - 2 \times 0.0284 = 0.943$.

From (2.4) it is known apriori that $\gamma \in (-1, 1)$. Nevertheless, the endpoints $C_{(L)}$ and/or $C_{(U)}$ of the confidence regions for γ described may not belong to this open interval (-1, 1), although the confidence coefficient associated with such a region is exact and the ordinary coverage interpretation in repeated sampling holds. Such disturbing phenomena are not unusual when confidence regions for ratios of variances in random-effects models are considered, see for instance Lehmann (1986, p. 421).

Confidence regions in terms of θ are obtained through the transformation (2.8). That is, the confidence regions $[C_{(L)}, \infty)$, $(-\infty, C_{(U)}]$ and $[C_{(L)}, C_{(U)}]$ for γ are transformed into the regions $[G(C_{(L)}), \infty]$, $[0, G(C_{(U)}]$ and $[G(C_{(L)}),\ G(C_{(U)})]$ for θ with the same exact confidence coefficient as the original γ-regions. If $C_{(L)}$ and/or $C_{(U)}$ do not belong to (-1, 1), then of course the corresponding transformed bounds for θ are zero or infinite.

4.3 A Point Estimator

Following Hodges and Lehmann (1963), the point estimate of γ considered here equals a value c that makes $T_*(c)$ as close as possible to the mean zero of the symmetric distribution $K_0(n_1, n_2)$ of $T_*(\gamma)$. Let L_0 and U_0 be the integer-valued random variables defined by the right-hand side of (4.5) and (4.6), respectively, with q_p replaced by zero. Moreover, define the estimator

$$\hat{\gamma} = (C_{(U_0)} + C_{(L_0)})/2 \qquad (4.9)$$

of γ. The two order statistics in (4.9) satisfy $C_{(U_0)} \leq C_{(L_0)}$. In the balanced case with $n_1 = n_2$, $\hat{\gamma}$ equals the median of the $N_1 + N_2$ slopes (3.4); whereas if $n_1 \neq n_2$ then this is not necessarily the case.

The estimator (4.9) is exactly or approximately median unbiased in the sense that the probabilities of $\hat{\gamma}$ overestimating and underestimating γ satisfy

$$(1 - \varepsilon_0)/2 \leq \Pr[\gamma < \hat{\gamma}] \leq (1 + \varepsilon_0)/2, \tag{4.10}$$

$$(1 - \varepsilon_0)/2 \leq \Pr[\hat{\gamma} < \gamma] \leq (1 + \varepsilon_0)/2, \tag{4.11}$$

where $\varepsilon_0 = \Pr[T_0(n_1, n_2) = 0]$ is the probability mass of $K_0(n_1, n_2)$ at the point zero; see the Appendix. Thus, if ε_0 is equal to or close to zero, then $\hat{\gamma}$ is exactly or approximately median unbiased. Actually, $\hat{\gamma}$ may be exactly or approximately median unbiased even if $\varepsilon_0 > 0$ is large, but this then depends on the joint distribution of the slopes (3.4) and, in particular, on the joint distribution of the n_* "between" ξ_{ij}'s in (2.1) about which no assumption has been made. For large n_*, ε_0 is close to zero, if not equal to zero, because $K_0(n_1, n_2)$ tends to a normal distribution as $n_* \to \infty$.

As is the case for the confidence regions for γ considered in sec. 4.2, $\hat{\gamma}$ may not belong to the open interval (-1, 1). The corresponding point estimator $G(\hat{\gamma})$ of θ in terms of (2.8) is exactly or approximately median unbiased in the sense that (4.10) and (4.11) hold with θ and $G(\hat{\gamma})$ substituted for γ and $\hat{\gamma}$, respectively.

4.4 Simultaneous Inferences

The comparison of primary interest in a 2 x 2 crossover trial typically concerns the effect of the treatments on the level of the response variable. A crucial assumption usually made (Jones and Kenward 1989 chap. 2) in this context is that the period effects and the direct treatment effects are fixed and additive, with no other disturbing fixed effects being present, that is no carry-over or aliased effects. In terms of (2.1) this assumption implies

$$\mu_{i1} - \mu_{i2} = \pi + \Delta, \quad i = 1,$$
$$= \pi - \Delta, \quad i = 2; \tag{4.12}$$

where the constant Δ equals the (A - B)-difference of the fixed direct treatment effects, and the constant π equals the (period 1 - period 2)-difference of the fixed period effects, with Δ typically being of primary interest. In addition to the assumptions made in sec. 2, it is assumed throughout the subsequent part of this article that (4.12) holds, so that inferences about Δ and (Δ, γ) can also be considered. The additional assumption that (4.12) holds does not influence marginal inferences about γ based on (3.8) because, as can be seen from (3.2) and (3.3), (3.8) is not influenced by the μ_{ij}'s in (2.1).

Exact inferences about Δ can be made (Jones and Kenward 1989 chap. 2) using essentially a two-sample t-statistic based on (2.7) as follows. Let $\hat{\Delta} = (\bar{Y}_{1\cdot} + \bar{Y}_{2\cdot})/2$, $s^2_{\cdot} = [(n_1 - 1) s^2_{1\cdot} + (n_2 - 1) s^2_{2\cdot}]/(n_* - 2)$, $s^2_* = s^2_{\cdot} [1/n_1 + 1/n_2]/4$ and $\sigma^2_* = \sigma^2_{AB}[1/n_1 + 1/n_2]/4$. Moreover, for any given $d \in (-\infty, \infty)$, let

$$t(d) = [\hat{\Delta} - d]/s_*, \tag{4.13}$$

which is $T(n_* - 2, \delta)$ distributed with noncentrality parameter $\delta = (\Delta - d)/\sigma_*$. The basic tests about Δ, similar to (4.1) - (4.3), are of one of the following three forms

$$H_{0\Delta}^{(-)}: \Delta \le \Delta_0, \quad H_{1\Delta}^{(+)}: \Delta > \Delta_0, \quad R_t^{(+)} = [t(\Delta_0) \ge t_{p,v}], \tag{4.14}$$

$$H_{0\Delta}^{(+)}: \Delta \ge \Delta_0, \quad H_{1\Delta}^{(-)}: \Delta < \Delta_0, \quad R_t^{(-)} = [t(\Delta_0) \le -t_{p,v}], \tag{4.15}$$

$$H_{0\Delta}: \Delta = \Delta_0, \quad H_{1\Delta}: \Delta \ne \Delta_0, \quad R_t = [\, |t(\Delta_0)| \ge t_{p,v}]. \tag{4.16}$$

Here $t(\Delta_0)$ is given by (4.13) with d equal to the value Δ_0, and $t_{p,v}$ denotes the upper p point ($0 < p < 1/2$) of the central t-distribution $T(\nu, 0)$ with $\nu = n_* - 2$ degrees of freedom. The size of the tests (4.14) - (4.16) equals p, p and 2p, respectively; and (Lehmann 1986) these three tests are strictly unbiased. The confidence coefficient associated with the corresponding confidence regions $[\hat{\Delta} - t_{p,v}\, s_*, \infty)$, $(-\infty, \hat{\Delta} + t_{p,v}\, s_*]$ and $[\hat{\Delta} - t_{p,v}\, s_*, \hat{\Delta} + t_{p,v}\, s_*]$ is equal to 1 - p, 1 - p and 1 - 2p, respectively.

It is then of interest to know how these exact inferences about Δ are related to the exact inferences about γ considered in sec. 4.1 and 4.2, and in particular whether it is possible to make exact simultaneous inferences. It turns out that such inferences can easily be made as follows.

Define a "rectangular" confidence region for (Δ, γ) as the direct product of two marginal regions: (a) a given Δ-region of the form $[\hat{\Delta} - t_{p,v}\, s_*, \infty)$, $(-\infty, \hat{\Delta} + t_{p,v}\, s_*]$, or $[\hat{\Delta} - t_{p,v} s_*, \hat{\Delta} + t_{p,v} s_*]$, as just described; and (b) a given γ-region of the form $[C_{(L)}, \infty)$, $(-\infty, C_{(U)}]$, or $[C_{(L)}, C_{(U)}]$, as described in sec 4.2; with the two given associated marginal confidence coefficients possibly distinct. The exact confidence coefficient associated with such a "rectangular" confidence region for (Δ, γ) is simply equal to the product of the two marginal confidence coefficients; that is, <u>as if</u> the two random regions (a) and (b) used in the direct product were independent. The confidence coefficient of the "rectangular" confidence region can thus be made equal to a desired value by choosing the two marginal confidence coefficients appropriately.

Intuitively it may be somewhat surprising that the confidence coefficient of the "rectangular" confidence region can be factorized as described, because actually the marginal random regions (a) and (b) are not independent. To show the factorization result, note first that by (4.7) and (4.8), the event of γ being covered by the random region (b) is equal with probability one to a certain event C defined in terms of $T_*(\gamma)$. Now, it can be shown using the results in sec. A.2 in the Appendix that $T_*(\gamma)$ is independent of $\hat{\Delta}$ and s_* in (4.13). Thus, the event D of Δ being covered by the random region (a) is independent of the event C, that is $\Pr[C \cap D] = \Pr[C]\Pr[D]$ holds as required.

Consider now two given tests: (i) a test about Δ of the form (4.14), (4.15) or (4.16); and (ii) a test about γ of the form (4.1), (4.2) or (4.3); with the two given marginal sizes, α_t and α_T respectively, possibly distinct. An exact multiple test procedure can be based on these two tests (i) and (ii), viewed as a family of tests, with each test either rejecting or accepting its null hypothesis H_0 according to whether its rejection event occurs or not. Thus, briefly, the four possible decisions with this family are: to reject H_0 in (i) and reject H_0 in (ii); to reject H_0 in (i) and accept H_0 in (ii); to accept H_0 in (i) and reject H_0 in (ii); or to accept H_0 in (i) and accept H_0 in (ii). The corresponding type I familywise error rate (FWE) equals the probability of making a type I error with at least one of the two tests (i) and (ii). This FWE depends on the unknown Δ and γ configuration. It is therefore desirable to be able to control this FWE in the strong sense (Hochberg and Tamhane 1987) through an upper bound that is valid for any Δ and γ. Now, define $\alpha_{tT} = 1 - (1 - \alpha_t)(1 - \alpha_T)$ in terms of the marginal sizes α_t and α_T of the tests (i) and (ii). It can be shown using the results in sec. A.2 that the FWE is $\leq \alpha_{tT}$ for any Δ and γ, with equality in this inequality for Δ and γ equal to the boundary values Δ_0 and γ_0 specifying the null hypotheses of the tests (i) and (ii). Thus this simple bound α_{tT} is the best possible upper bound that controls the type I FWE in the strong sense, and it can be made equal to a desired value by choosing the two marginal sizes α_t and α_T appropriately.

One or both of the given tests (i) and (ii) just considered may be two-sided, that is of the form (4.3) or/and (4.16). Suppose now that whenever such a two-sided test rejects its null hypothesis, a supplementary directional decision is made stating on which side of the null value the unknown parameter lies. For instance, if the rejection event R_T occurs, the supplementary decision is to decide $\gamma > \gamma_0$ or $\gamma < \gamma_0$ if the subevent $R_T^{(+)}$ or $R_T^{(-)}$, respectively, occurs in (4.1)-(4.3); and similarly for (4.14)-(4.16). In addition to the possibility of making a type I error, there is then the possibility of making a type III error (Hochberg and Tamhane 1987), that is a directional decision in the wrong direction. Now, it can be shown using the results in sec. A.2 that the probability of making a type I or/and type III error with at least one of the two tests (i) and (ii) is bounded from above by the bound α_{tT} derived previously considering errors of type I only. Thus, in this sense, natural and informative directional decisions can be made without additional cost. This appealing result parallels that in Hochberg and Tamhane (1987, theorem 2.2, p. 41) for fixed-effects linear models.

5. AN ILLUSTRATION

Koch (1972) reported the following data from a 2 x 2 crossover trial where 10 children were randomly assigned to the two sequence groups, with $n_1 = n_2 = 5$. Briefly, the two treatments A and B applied in the two periods to each child were A: to drink first 100 ml of grapefruit juice and then an elixir of Pentobarbital; and B: to drink first 100 ml of water and then an elixir of Pentobarbital; with the two treatment periods separated by one week. The amount of Pentobarbital elixir given in each period was proportional to the child's body weight, and the corresponding response observation consisted of a measurement of the amount of drug (in μg/ml) in a 10 ml sample of blood taken 15 minutes after the elixir administration. The measurements (Y_{ij1}, Y_{ij2}) from period 1 and period 2 for the children are, in group i = 1 with treatment sequence AB: (1.75, 0.55), (0.30, 1.05), (0.35, 0.63), (0.20, 1.55), (0.30, 8.20); and in group i = 2 with BA: (7.20, 0.35), (7.10, 1.55), (0.75, 0.25), (2.15, 0.35), (3.35, 1.50). In what follows only point estimates and confidence regions based on these data are considered; it should be evident how the tests and the multiple test procedure described in sec. 4.1 and 4.4 are performed.

The estimates required for the inferences about Δ in (4.12) based on (4.13) with $n_* - 2 = 8$ degrees of freedom are: $\hat{\Delta} = -2.595$ and $s_* = 0.998$. For instance, the 95 % confidence interval for Δ, with endpoints $-2.595 \pm 2.306 \times 0.998$, is equal to [-4.90, -0.29]. This gives a clear indication that $\Delta < 0$, which means that the mean response level is lower under A than under B.

Suppose now that it is also of interest to compare A and B with respect to the within-subject variability of the response measurement. The confidence regions for γ and the point estimate of γ described in sec. 4.2 and 4.3 are based on the $N_1 + N_2 = 20$ order statistics $C_{(k)}$ corresponding to the slopes (3.4) from the two groups. Since the crossover trial considered is balanced, the selection of the relevant order statistics to be used in this context is very simple: the point estimate $\hat{\gamma}$ equals the median $(C_{(10)} + C_{(11)})/2$; and, as shown in sec. 4.2, the interval $[C_{(6)}, C_{(15)}]$ is a confidence interval for γ with exact confidence coefficient 0.946. With the present data, $\hat{\gamma} = -0.993$ and $[C_{(6)}, C_{(15)}] = [-1.031. -0.667]$. This gives a clear indication that γ is much smaller than zero, which means that the within-subject response variability is much smaller under A than under B. As mentioned in sec. 4.2 and 4.3, the endpoints of the confidence interval for γ and the point estimate $\hat{\gamma}$ do not always lie within (-1, 1), although this interval is known apriori to contain γ. In the present case, $\hat{\gamma}$ is very close to the lower endpoint of (-1, 1), whereas $C_{(6)}$ actually is below it. Results in terms of $\theta = \sigma^2_A/\sigma^2_B$ are obtained by using the tranformation (2.8): the point estimate of θ equals 0.003, and the confidence interval for θ with exact confidence coefficient 0.946 equals (0, 0.200]; which indicates the relative smallness of σ^2_A compared to σ^2_B.

The two confidence intervals for Δ and γ just considered can also be viewed simultaneously. As shown in sec. 4.4, the confidence rectangle $[-4.90, -0.29] \times [-1.031, -0.667]$ for (Δ, γ) based on these two marginal intervals has an exact confidence coefficient simply equal to the product of the two marginal confidence coefficients 0.95 and 0.946, that is equal to 0.899.

Cornell (1991) also considered these data, and his asymptotically distribution-free test of the hypothesis $\sigma^2_A = \sigma^2_B$ based on Kendall's tau leads to a prob-value < 0.01. No corresponding nonparametric confidence interval for σ^2_A/σ^2_B is available.

6. CONCLUDING REMARKS

A 2 x 2 crossover trial is of course not primarily designed to compare the two treatments with respect to the within-subject response variability, but such a comparison may nevertheless provide additional information of considerable importance for the overall judgement of the two treatments. It has been shown in this article that it is possible to make various exact inferences of interest in this context, without any assumption about the joint distribution of the between-subject random effects ξ_{ij} in (2.1), which is rather appealing. It is also appealing that no assumption is required about fixed effects that may influence the mean response level, unless the inferences considered concern such fixed effects. An open problem is whether/how exact inferences can be made under essentially weaker distributional assumptions for the within-subject random effects than those used in this article.

A decision problem related to the directional decisions described in sec. 4.4 concerns the position of γ and Δ relative to given intervals (γ'_0, γ''_0) and (Δ'_0, Δ''_0). For instance, in the context of bioequivalence trials, these intervals may consist of γ-values and of Δ-values that are considered to be essentially equivalent to zero. Following Lehmann (1957, sec. 10), decision rules can be based on intersections of "one-sided" rejection events of the form (4.1) - (4.2) with γ'_0 and γ''_0 substituted for γ_0, and of the form (4.14) - (4.15) with Δ'_0 and Δ''_0 substituted for Δ_0; with the possible decisions reflecting various degrees of inconclusiveness. A very simple decision rule of this kind is to make the decision D: $\gamma \in (\gamma'_0, \gamma''_0)$ and $\Delta \in (\Delta'_0, \Delta''_0)$ if the intersection of the following four rejection events occurs: $R_T^{(+)}(\gamma'_0)$, $R_T^{(-)}(\gamma''_0)$, $R_t^{(+)}(\Delta'_0)$, $R_t^{(-)}(\Delta''_0)$, with obvious notation; and to make no decision otherwise. The probability of deciding D if actually D is not true is then bounded from above by the maximum of the four sizes. This can easily be shown by using the fact that the probability of an intersection is at most as large as that of each of the intersected events. Thus, if each of the four one-sided tests has a size ≤ 0.05, then the probability of deciding D erroneously is ≤ 0.05. The corresponding decision rule concerning Δ only was proposed in the context of bioequivalence trials by Westlake (1981) in terms of confidence intervals and by Schuirmann (1981) in terms of two one-sided tests.

If the n_* "between" ξ_{ij}'s in (2.1) are assumed to be independent and distributed according to a common parent ξ-distribution with variance $0 \leq \sigma^2_\xi < \infty$, it is possible to derive a normal approximation for the power of the tests (4.1)-(4.3) along the lines in Konijn (1956), using (3.2) and Hoeffding's (1948) representation of Kendall's tau as a U-statistic. The asymptotic result supporting the approximation assumes that γ becomes close to γ_0 as n_* becomes large. More precisely it is assumed, with $\gamma_0 \in (-1, 1)$ and $\alpha \in (0, 1/2)$ given, that the sequence of situations considered in such that: (a) n_1 & $n_2 \rightarrow \infty$; (b) p tends to α at the rate $p = \alpha + o(n_*^{-1/2})$; (c) the ξ-distribution remains fixed; and (d) the "within" variances σ^2_A and σ^2_B vary in such a way that σ^2_{AB} in (2.3) tends to some limit $\sigma^2_\infty \in (0, \infty)$ and γ in (2.4) tends to γ_0 at the rate $\gamma = \gamma_0 + h/n_*^{1/2} + o(n_*^{-1/2})$ for some $h \in (-\infty, \infty)$. No assumption about the shape of the ξ-distribution is necessary.

Briefly, the approximation in question is $\Pr(R_T^{(+)}) \approx \Phi(\kappa - z_p)$, $\Pr(R_T^{(-)}) \approx \Phi(-\kappa - z_p)$, and $\Pr(R_T) \approx \Phi(|\kappa| - z_p) + \Phi(-|\kappa| - z_p)$; with $z_p = \Phi^{-1}(1-p)$ and $\kappa = K\, n_*^{1/2}(\gamma - \gamma_0)$ for a certain K which depends on the ξ-distribution and which satisfies $K_{min} \leq K \leq K_{max}$ with $K_{max} = 3/[\pi(1 - \gamma_0^2)^{1/2}]$ and $K_{min} = K_{max}\ \exp\{-2\sigma^2_\xi/[\sigma^2_{AB}(1-\gamma_0^2)]\}$. The bound K_{min} is of particular interest for planning purposes. If estimates s^2_+ and s^2_- of $Var(Y_{ij+})$ and $Var(Y_{ij-})$ are available, an estimate K^*_{min} of K_{min} is easily obtained by substituting the estimate max $(0, s^2_+/s^2_- - 1)$ for the unknown quantity $2\sigma^2_\xi/\sigma^2_{AB} = Var(Y_{ij+})/Var(Y_{ij-}) - 1$. This estimate K^*_{min} can then be substituted for K in the definition of κ to get an approximate lower bound for the power of (4.1)-(4.3) under the alternative hypothesis. Useful additional relations and approximations concerning Δ and γ simultaneously can be obtained through (A.3). These results could be used for planning purposes, including the determination of the number n_*. A more detailed description of these results and applications is however beyond the scope of this article.

A problem that has not been considered in this article is how ties among observations should be handled. This requires further investigations.

APPENDIX

A.1 Proof of (4.10) and (4.11)

The proof parallells that of Hodges and Lehmann (1963) for estimates of location and translation parameters. More precisely, the relations (4.7) and (4.8) hold with L_0, U_0, γ and zero substituted for L, U, c and q_p, respectively; with the two middle events (and thus all six events) each having probability $(1 - \varepsilon_0)/2$. The inequalities (4.10) and (4.11) then follow easily using the fact that $C_{(U_0)} \le \hat{\gamma} \le C_{(L_0)}$.

A.2 A Representation and Some Related Results

For $i = 1, 2$, define $\bar{U}_i = \Sigma_j U_{ij}/n_i$, $s^2_{iU} = \Sigma_j(U_{ij} - \bar{U}_i)/(n_i - 1)$ and $U^*_{ij} = (U_{ij} - \bar{U}_i)/s_{iU}$, $j = 1, ..., n_i$, in terms of the U_{ij}'s in (3.2); and let $\mathbf{U}^*_i$ denote the vector with components U^*_{ij}, $j = 1, ..., n_i$. It then follows from (3.2) that for $i = 1, 2$, $\bar{Y}_{i\cdot}$ and $s^2_{i\cdot}$ in (2.7), and $S_i(c)$ in (3.3), can be represented as

$$\bar{Y}_{i\cdot} = \mu_{i\cdot} + \sigma_{AB}\bar{U}_i, \qquad s^2_{i\cdot} = \sigma^2_{AB}\, s^2_{iU}, \tag{A.1}$$

$$S_i(c) = \Sigma\Sigma_{r<s} \operatorname{sign}[(U^*_{is} - U^*_{ir})\{(\gamma - c)s_{iU}(U^*_{is} - U^*_{ir}) + V_{irs}\}] \tag{A.2}$$

with $V_{irs} = (1 - \gamma^2)^{1/2} (Z_{is} - Z_{ir}) + 2(\xi_{is} - \xi_{ir})/\sigma_{AB}$. By well known properties of random samples from normal distributions, $\bar{U}_1$, $\bar{U}_2$, s_{1U}, s_{2U}, $\mathbf{U}^*_1$ and $\mathbf{U}^*_2$ are independent; and these random variables and vectors are independent of the V_{irs}'s in (A.2), because the U_{ij}'s are independent of the Z_{ij}'s and of the ξ_{ij}'s in (3.2).

It follows immediately from (A.1), (A.2) and the definition (3.8) of $T_*(c)$ that: (a) $T_*(\gamma)$ is independent of $(\bar{Y}_{1\cdot}, \bar{Y}_{2\cdot}, s_{1\cdot}, s_{2\cdot})$; whereas (b) for $c \neq \gamma$, $T_*(c)$ is independent of $(\bar{Y}_{1\cdot}, \bar{Y}_{2\cdot})$ but not of $(s_{1\cdot}, s_{2\cdot})$. Thus, $t(d)$ in (4.13) is independent of $T_*(c)$ if $c = \gamma$, but not necessarily otherwise. Some useful relations concerning the case $c \neq \gamma$ can be derived using the fact that $t(d)$ and $T_*(c)$ are conditionally independent given $(s_{1\cdot}, s_{2\cdot})$. By a straight forward application of the generalized Kimball (1951) inequality (Esary, Proschan and Walkup 1967), it follows from this conditional independence that, for any given $-\infty \le a < 0 < b \le \infty$ and $q \in Q_*$ in (3.9),

$$\Pr[a \le t(d) \le b, T_*(c) \le q] \sim \Pr[a \le t(d) \le b] \Pr[T_*(c) \le q], \tag{A.3}$$

where ~ denotes ≤, =, ≥, if $c < \gamma$, $c = \gamma$, $c > \gamma$, respectively. The fact that t(d) is independent of $T_*(\gamma)$ is thus contained in this more general result. Note also that each side of (A.3) is a nonincreasing function of c.

Table 1. Upper Tail Probabilities of $K_0(n_1, n_2)$ for Equally Large Groups.

	n in each group							
s	3	4	5	6	7	8	9	10
0	.6389	.5920	.5670	.5517	.5415	.5343	.5290	.5249
1	.3611	.4080	.4330	.4483	.4585	.4657	.4710	.4751
2	.1389	.2413	.3059	.3481	.3772	.3981	.4137	.4257
3	.0278	.1181	.1977	.2572	.3008	.3333	.3581	.3774
4		.0451	.1153	.1798	.2320	.2730	.3053	.3310
5		.0122	.0597	.1184	.1727	.2185	.2562	.2869
6		.0017	.0268	.0729	.1236	.1706	.2113	.2457
7			.0101	.0416	.0850	.1299	.1714	.2079
8			.0030	.0218	.0558	.0962	.1364	.1736
9			.0006	.0104	.0349	.0692	.1066	.1431
10			.0001	.0044	.0208	.0483	.0816	.1163
11				.0016	.0116	.0326	.0612	.0932
12				.0005	.0061	.0213	.0449	.0736
13				.0001	.0030	.0134	.0322	.0572
14				.0000	.0013	.0081	.0226	.0438
15				.0000	.0005	.0047	.0154	.0330
16					.0002	.0026	.0103	.0244
17					.0001	.0013	.0066	.0177
18					.0000	.0007	.0042	.0127
19					.0000	.0003	.0025	.0089
20					.0000	.0001	.0015	.0061
21					.0000	.0001	.0008	.0041

NOTE: For $n_1 = n_2 = n \leq 10$, the table gives values of $Pr[T_0(n_1, n_2) \geq s/N]$ for s = 0, 1, 2, ... ; where $N = n(n-1)/2$, and $T_0(n_1, n_2)$ has distribution $K_0(n_1, n_2)$. Corresponding lower tail probabilities are obtained by using the symmetry about zero of $K_0(n_1, n_2)$. If $c = \gamma$, then the random variable (3.8) has distribution $K_0(n_1, n_2)$.

Table 2. Upper Tail Probabilities of $K_0(n_1, n_2)$ Close to 0.10, 0.05, 0.025, 0.010, 0.005 for Unequally Large Groups.

n_1	n_2	t	P(t)
3	4	.5429	.1181
		.5810	.0764
		.7524	.0417
		.7905	.0278
		1.0000	.0069
3	5	.4323	.1208
		.4710	.0931
		.5742	.0569
		.6129	.0361
		.6645	.0333
		.7161	.0222
		.8065	.0097
		.8581	.0069
		1.0000	.0014
4	5	.3909	.1122
		.4091	.0913
		.5091	.0580
		.5273	.0424
		.6273	.0264
		.6455	.0171
		.7455	.0101
		.7637	.0059
		.8638	.0028
4	6	.3670	.1015
		.3872	.0975
		.4747	.0534
		.4815	.0482
		.5623	.0262
		.5690	.0220
		.6498	.0112
		.6566	.0084
		.7104	.0066
		.7374	.0042
5	6	.3519	.1077
		.3620	.0963
		.4380	.0572
		.4481	.0487
		.5241	.0261
		.5342	.0210
		.6000	.0129
		.6101	.0098
		.6759	.0057
		.6861	.0042
5	7	.3267	.1024
		.3307	.0962
		.4178	.0513
		.4297	.0448
		.4931	.0255
		.5050	.0216
		.5683	.0110
		.5802	.0093
		.6277	.0053
		.6396	.0042
6	7	.3053	.1007
		.3084	.0999
		.3925	.0502
		.3988	.0454
		.4517	.0274
		.4579	.0245
		.5452	.0108
		.5514	.0099
		.5981	.0054
		.6044	.0050 −
6	8	.2920	.1007
		.2946	.0989
		.3679	.0526
		.3704	.0496
		.4336	.0268
		.4362	.0245
		.5095	.0103
		.5171	.0098
		.5626	.0051
		.5702	.0047
7	8	.2724	.1027
		.2744	.0983
		.3505	.0514
		.3525	.0482
		.4203	.0256
		.4245	.0243
		.4841	.0102
		.4861	.0099
		.5416	.0052
		.5457	.0049
7	9	.2614	.1042
		.2631	.0997
		.3350	.0500 +
		.3368	.0497
		.3985	.0256
		.4002	.0241
		.4704	.0100 +
		.4721	.0092
		.5150	.0051
		.5167	.0050 −
8	9	.2514	.1027
		.2529	.0990
		.3200	.0510
		.3214	.0494
		.3800	.0255
		.3814	.0244
		.4457	.0103
		.4471	.0099
		.4971	.0051
		.5000	.0047
8	10	.2426	.1006
		.2438	.0989
		.3094	.0501
		.3106	.0499
		.3690	.0251
		.3702	.0239
		.4310	.0102
		.4322	.0100 −
		.4760	.0051
		.4772	.0048
9	10	.2341	.1018
		.2351	.0992
		.3013	.0501
		.3023	.0491
		.3550	.0257
		.3561	.0248
		.4160	.0100 +
		.4171	.0100 −
		.4584	.0052
		.4594	.0049

NOTE: For $n_2 - 2 \le n_1 < n_2 \le 10$, the (n_1, n_2)-subtable gives, for each target value p = 0.10, 0.05, 0.025, 0.010, 0.005: (a) the largest P(t)-value $\le$ p, if there is such a value; and (b) the smallest P(t)-value $\ge$ p. Here $P(t) = \Pr[T_0(n_1, n_2) \ge t]$, $t \in Q_*$, where $T_0(n_1, n_2)$ has distribution $K_0(n_1, n_2)$. Corresponding lower tail probabilities are obtained by using the symmetry about zero of $K_0(n_1, n_2)$. If $c = \gamma$, then the random variable (3.8) has distribution $K_0(n_1, n_2)$.

1991:1	Olofsson, J.	On some prediction methods for categorical data
1991:2	Jonsson, R.	On the problem of optimal inference in the simple error component model for panel data
1991:3	Holm, S. & Svennson, E.	Statistical rank methods for ordinal categorical data
1992:1	Teräsvirta, T., Tjøstheim, D. & Granger, C.W.J.	Aspects of Modelling Nonlinear Time Series
1992:2	Palaszewski, B.	A conditional stepwise test for deviating parameters