On A Monotonicity Problem in Step-Down Multiple Test Procedures

H. FINNER*

We consider a monotonicity problem concerning the critical values in stepwise multiple test procedures for comparing k parameters. This problem will be solved for a large class of distributional settings by means of an inequality for cumulative distribution functions of test statistics satisfying a simple monotonicity condition.

KEY WORDS: Adjusted significance levels; Hölder's inequality; Multiple level of significance; Natural procedure; Pairwise testing; Range inequality; Step-down procedure.

1. INTRODUCTION

A question that often occurs in a natural manner in stepwise multiple comparisons procedures is the monotonicity of the critical values defining these procedures. In a recent paper (Finner 1990), this problem was solved for multiple range tests by means of an inequality for the distribution function of the range statistic. Here we consider the monotonicity problem in a more general way with respect to pairwise testing problems and step-down procedures.

Let $X_1, \ldots, X_k$ be independently distributed random variables with corresponding probability measures $P^{X_i} \in \mathcal{P}$ $(i = 1, \ldots, k)$, where $\mathcal{P}$ is a prespecified class of probability measures. Then we define a function $\mu : \mathcal{P} \to \Theta$ such that $\mu_i = \mu(P^{X_i})$ $(i = 1, \ldots, k)$ denote the parameters of interest. Now we want to test all $k(k-1)/2$ hypotheses $H_0^{ij} : \mu_i = \mu_j$ versus $H_1^{ij} : \mu_i \neq \mu_j$ $(1 \leq i < j \leq k)$ by means of a step-down test procedure with multiple level α $(\alpha \in (0, 1))$. The classic solution of this problem is based on the so-called set of homogeneity hypotheses; that is,

$$H_0^Q = \bigcap_{i,j \in Q, i \neq j} H_0^{ij}, \qquad Q \subseteq I_k = \{1, \ldots, k\}, \qquad |Q| \geq 2.$$

Let $T_Q = T_Q(X_i : i \in Q)$ be real-valued test statistics depending on X_i, $i \in Q$. For the sake of simplicity, we first assume that the cumulative distribution function of T_Q is given by $F_q(z) = P_\mu(T_Q \leq z)$ for all $\mu \in H_0^Q$ where $q = |Q|$; that is, in this case the distribution of T_Q given $\mu \in H_0^Q$ depends only on the size of Q.

The most popular choice of significance levels for a step-down test procedure defined in the next paragraph is given by $\alpha_k = \alpha_{k-1} = \alpha$, $\alpha_q = 1 - (1 - \alpha)^{q/k}$, $q = 2, \ldots, k - 2$. Another choice of adjusted significance levels may be found in Lehmann and Shaffer (1979). The corresponding critical values are defined by

$$c_q(\alpha_q) = \inf\{c \in \mathbb{R} : F_q(c) \geq 1 - \alpha_q\}.$$

The step-down multiple test procedure for the extended set of hypotheses H_0^Q, $Q \subseteq I_k$, $|Q| \geq 2$ is defined as follows. Reject H_0^Q if $T_P \geq c_p(\alpha_p)$ for all P with $Q \subseteq P$. Hence a hypothesis $H_0^{ij} : \mu_i = \mu_j$ is rejected if $T_P \geq c_p(\alpha_p)$ for all P with $\{i,j\} \subseteq P$. This procedure with α_p defined as earlier keeps the multiple level (of significance) α; that is, the probability of erroneously rejecting a true null hypothesis is bounded by α (cf. Finner 1990; Lehmann and Shaffer 1977, 1979).

Because $H_0^P \subseteq H_0^Q$ often implies $T_P \geq T_Q$, the requirement $c_p(\alpha_p) \geq c_q(\alpha_q)$ seems quite natural. Examples of statistics with the property $T_P \geq T_Q$ for $H_0^P \subseteq H_0^Q$ are the range statistic $\max_{i,j \in Q} |X_i - X_j|$, the one-sided range statistic $\max_{i,j \in Q, i<j}(X_i - X_j)$, the chi-squared-type statistic $\sum_{i \in Q} (X_i - \bar{X}_\cdot^Q)^2$ with $\bar{X}_\cdot^Q = \sum_{i \in Q} X_i / q$, the mean deviation statistic $\sum_{i \in Q} |X_i - \bar{X}_\cdot^Q|$, and also the k-sample rank statistic and the k-sample sign statistic.

We notice that $T_P \geq T_Q$ always implies $c_p(\alpha) \geq c_q(\alpha)$ for all $\alpha \in (0, 1)$. Unfortunately, the choice of the adjusted significance levels often results in $c_{k-1} < c_{k-2}$; that is, a violation of the monotonicity requirement. A similar effect can be observed in the minimax approach of Lehmann and Shaffer (1979) (see ex. 3.1 in Finner 1990), where it is possible that the monotonicity is violated for many $p \in \{2, \ldots, k\}$. If the monotonicity requirement is not satisfied, this may indicate that there is something wrong with the choice of the adjusted significance levels. But if we define a new set of critical values by $c'_q = \max_{2 \leq p \leq q} c_p$, $q = 2, \ldots, k$, we obtain $c'_2 \leq \cdots \leq c'_k$, and the resulting test procedure with modified adjusted significance levels given by $\alpha'_q = 1 - F_q(c'_q)$, $q = 2, \ldots, k$, leads to the same decisions concerning the pair hypotheses H_0^{ij} as does the original procedure whenever $T_Q \leq T_P$ for all $Q \subseteq P$. This may be seen as follows. Clearly, because $c'_q \geq c_q$ for all $q = 2, \ldots, k$, the modified procedure accepts every hypothesis that is accepted by the original procedure. To prove the converse, assume that H_0^{ij} is accepted by the modified procedure. Then there exists a Q with $\{i, j\} \subseteq Q$, $q = |Q|$, and $T_Q < c'_q$. If $c'_q = c_q$, the original procedure also accepts H_0^{ij}. If $c'_q > c_q$, then there exists a $p < q$ with $c'_q = c_p$ and $c_p = c'_p$. Let $\{i, j\} \subseteq P \subset Q$ with $p = |P|$. Then $T_P \leq T_Q < c'_q = c_p$, thus H_0^{ij} is accepted by the original procedure.

A further advantage of monotonic critical values and monotonic test statistics is that a pair hypothesis H_0^{ij} can always be rejected if $T_{ij} > c_q$ for one $q \in \{2, \ldots, k\}$ and if, in addition, $T_P > c_p$ for all P with $\{i, j\} \subseteq P$, $p = |P| > q$. This property can be used to avoid testing of all $2^k - k - 1$ homogeneity hypotheses (see, for instance, the example in Sec. 4).

It is obvious from these considerations that it makes no sense to look for an optimal choice of adjusted significance

* H. Finner is Wissenschaftlicher Assistent, Fachbereich IV, Mathematik/Statistik, 54286 Trier, Germany. The author thanks the referees for their helpful comments.

Journal of the American Statistical Association
September 1993, Vol. 88, No. 423, Theory and Methods

levels without respect to the monotonicity requirement if interest is focused only on the pair hypotheses $H_0^{ij} : \mu_i = \mu_j$. On the other hand, a step-down multiple test procedure—or, more generally, a closed test procedure—does not require monotonicity. For example, the modified procedure in the preceding paragraph based on the critical values c'_q may lead to a loss in power relative to the original procedure if one is interested in all homogeneity hypotheses H_0^Q, $Q \subseteq I_k$, $|Q| \geq 2$. But monotonic critical values and monotonic test statistics often allow simplified and fast algorithms for the computation of the decisions for the pair hypotheses H_0^{ij}; see, for example, the algorithm based on the range statistic (cf. Finner 1990; Lehmann and Shaffer 1977). Finally, a violation of the monotonicity property seems to contradict the philosophy of a step-down procedure.

The main purpose of this article is to show that the "natural" procedure (cf. Finner 1990) based on the adjusted significance levels

$$\begin{aligned} \alpha_q &= \alpha, \qquad q = k \\ &= \min\{1 - F_{k-1}(c_{k-1}(\alpha)),\ 1 - (1-\alpha)^{(k-1)/k}\}, \\ &\qquad\qquad\qquad\qquad\qquad\qquad q = k-1 \\ &= 1 - (1-\alpha)^{q/k}, \qquad 2 \leq q \leq k-2 \qquad (1) \end{aligned}$$

yields the monotonicity of the corresponding critical values for all test statistics with the property $T_Q \leq T_P$ for $Q \subseteq P$. The adjustment for $q = k - 1$ in (1) is necessary, because there are examples where $c_{k-1}(\alpha) > c_{k-2}(1 - (1-\alpha)^{(k-2)/k})$, whereas no problems occur by choosing $\alpha_{k-1} = 1 - (1 - \alpha)^{(k-1)/k}$.

The solution given in Section 2 is based on an inequality for various cumulative distribution functions (cdf's) obtained in Finner (1992), which is of the same type as the range inequality.

The case where the cdf of T_Q depends not only on the size q but also on the special set Q itself will be considered separately in Section 3. It will be shown that the corresponding critical values $c_Q(\alpha_q)$, $\varnothing \neq Q \subseteq I_k$, $|Q| = q \geq 2$ do not always satisfy $c_Q(\alpha_q) \leq c_P(\alpha_p)$ for all $Q \subset P$ although the monotonicity property $T_Q \leq T_P$ is satisfied. But here we obtain the weaker property that for every $Q \subseteq I_k$, $|Q| \geq 3$, there exists at least one $j \in Q$ such that $c_{Q\setminus\{j\}}(\alpha_{q-1}) \leq c_Q(\alpha_q)$, where $\alpha_q = 1 - (1-\alpha)^{q/k}$, $q = 2, \ldots, k$. Finally, the practical advantage of monotonicity properties is illustrated by reanalyzing a data set of an unbalanced one-way layout with a step-down F-test procedure.

2. SOLUTION OF THE MONOTONICITY PROBLEM

Theorem 2.1. Let X_i, $i \in \mathbb{N}$, be independent identically distributed random variables with values in Ω and let $T_Q : \Omega^q \to \mathbb{R}$ be real-valued Borel-measurable statistics with cumulative distribution function $F_q(z) = P(T_Q \leq z)$ for all $\varnothing \neq Q \subset \mathbb{N}$, $q = |Q|$, $z \in \mathbb{R}$. If $T_P(\omega_i : i \in P) \leq T_Q(\omega_i : i \in Q)$ for all $(\omega_i : i \in Q) \in \Omega^q$, $\varnothing \neq P \subset Q \subset \mathbb{N}$, $p = |P|$, $q = |Q|$, then

$$F_p(z)^{1/p} \geq F_q(z)^{1/q} \quad \text{for all} \quad z \in \mathbb{R}. \qquad (2)$$

The proof of Theorem 2.1 may be found in Finner (1992). We point out that inequality (2) is a special case of a generalization of Hölder's inequality for n functions. Similarly, as in Finner (1990), we now obtain the following corollary.

Corollary 2.1. Let $k \in \mathbb{N}$, $k \geq 2$, $\alpha \in (0, 1)$, $\alpha_q = 1 - (1-\alpha)^{q/k}$, $q = 1, \ldots, k$, and $c_q(\alpha_q)$ as defined earlier. Then, under the assumptions of Theorem 2.1, we have $c_k(\alpha_k) \geq \cdots \geq c_1(\alpha_1)$.

All that remains to be done is to verify the monotonicity property of the underlying statistics T_Q in various situations. The range statistic and the one-sided range statistic trivially possess this property. To prove the monotonicity property for the chi-squared type statistic and the mean deviation statistic, let $P = \{1, \ldots, p\}$ and $Q = \{1, \ldots, p+1\}$. Then it may easily be seen that

$$\sum_{j \in Q} (x_j - \bar{x}^Q_{\cdot})^2 = \sum_{j \in P} (x_j - \bar{x}^P_{\cdot})^2 + (p+1)(x_{p+1} - \bar{x}^Q_{\cdot})^2/p$$

and

$$\begin{aligned} \sum_{j \in Q} |x_j - \bar{x}^Q_{\cdot}| &= \sum_{j \in P} |x_j - \bar{x}^Q_{\cdot}| \\ &\quad + \sum_{j \in P} |x_{p+1} - \bar{x}^P_{\cdot}|/(p+1) \\ &\geq \sum_{j \in P} |x_j - \bar{x}^Q_{\cdot} + (x_{p+1} - \bar{x}^P_{\cdot})/(p+1)| \\ &= \sum_{j \in P} |x_j - \bar{x}^P_{\cdot}|; \end{aligned}$$

hence the chi-squared type statistic and the mean deviation statistic satisfy the desired monotonicity property.

At next we consider the k-sample rank statistic (Dwass 1960; Steel 1960) and the k-sample sign statistic (Nemenyi 1963), as defined in Miller (1980). Let $X_i = (X_{i1}, \ldots, X_{in})$, $i = 1, \ldots, k$ and assume that the X_{ij}, $j = 1, \ldots, n$, $i = 1, \ldots, k$ are independent identically distributed with values in $\mathbb{R}$. Let $R_{i1:i'}, \ldots, R_{in:i'}$ be the ranks of the observations $x_{i1}, \ldots, x_{in}$ with respect to the combined sample $x_{i1}, \ldots, x_{in}, x_{i'1}, \ldots, x_{i'n}$ of size $2n$. The rank sum for sample i with respect to sample i' is given by $R_{ii'} = \sum_{j=1}^{n} R_{ij:i'}$, and the q-sample rank statistic for testing H_0^Q, $Q \subseteq I_k$, $|Q| = q \geq 2$ is defined as

$$T_q = \max_{i,i' \in Q, i \neq i'} R_{ii'}.$$

The problem of ties must be solved such that T_Q and T_P are identically distributed for $|P| = |Q|$. Because $T_P(x_i : i \in P) \leq T_Q(x_i : i \in Q)$ for all $P \subseteq Q$ and $(x_i : i \in Q) \in \Omega^{nq}$, Theorem 2.1 can be applied.

In the case of the k-sample sign statistic, we consider a slightly modified version of the procedure in Miller (1980). Let $D^+_{ii':j} = 1$ if $x_{ij} - x_{i'j} > 0$ and $D^+_{ii'} = 0$ if otherwise. Set $S^+_{ii'} = \sum_{j=1}^{n} D^+_{ii':j}$ and define the statistics $T_Q = \max_{i,i' \in Q, i \neq i'} S^+_{ii'}$. Obviously, $T_P(x_i : i \in P) \leq T_Q(x_i : i \in Q)$ for all $P \subseteq Q$ and $(x_i : i \in Q) \in \Omega^{nq}$. Hence the corresponding distribution functions satisfy (2).

Unfortunately, Theorem 2.1 is no longer valid if the X_i are dependent; for example, if the X_i are studentized random variables given by $X_i = Z_i/S$, where $Z_1, \ldots, Z_k$, S are assumed to be independent. In this case the monotonicity of the critical values of the natural procedure may be violated. An example concerning the studentized range statistic was given by Finner (1990). But it seems that the F distribution makes an exception. Let $T_p = \sum_{i=1}^{p} X_i^2$, $T_\nu = \sum_{i=1}^{\nu} Y_i^2$, where T_p and T_ν are independently distributed chi-squared variates with p and ν degrees of freedom. Let $G_{p,\nu}$ denote the cdf of T_p/T_ν and let $f_{p,\nu,\alpha}$ denote the critical values satisfying $P(T_p/T_\nu \le f_{p,\nu,\alpha}) = 1 - \alpha$ (i.e., $f_{p,\nu,\alpha} = pF_{p,\nu,\alpha}/\nu$, where $F_{p,\nu,\alpha}$ denote the critical values of the F distribution with p and ν degrees of freedom). Then a numerical check indicates the validity of the inequality $G_{p,\nu}(z)^{1/p} \ge G_{p+1,\nu}(z)^{1/(p+1)}$ for all $\nu \ge 2$ (with equality for $\nu = 2$ and reversed inequality for $\nu = 1$), $z > 0$, and $p \in \mathbb{N}$. But the correctness of this inequality would imply the monotonicity of the critical values $\tilde{c}_p(\alpha_p) = f_{p,\nu,\alpha_p}$ as well as of $c_p(\alpha_p) = f_{p-1,\nu,\alpha_p}$, where $\alpha_p = 1 - (1 - \alpha)^{p/k}$, $p = (1), 2, \ldots, k$. In pairwise testing situations with corresponding distributional assumptions, the latter set of critical values with the modification $\alpha_{k-1} = \alpha$ is widely used (see also the example in Sec. 4). Hence it would be of interest to know whether or not this inequality holds for all $\nu \ge 2$.

3. INDEPENDENT NON-IDENTICALLY DISTRIBUTED TEST STATISTICS

In this section we consider the problems occurring in the case where the distribution of T_Q under H_0^Q depends not only on the size of Q but also on the special set Q itself. For example, if the X_i are given by $X_i = (X_{i1}, \ldots, X_{in_i})$, $i = 1, \ldots, k$ with different sample sizes n_i, or if the X_i have different variances, then the distribution of $T_Q(X_i : i \in Q)$ depends on the special set Q. But in this case the statistics T_Q often satisfy the monotonicity property $T_P \ge T_Q$ if $H_0^P \subseteq H_0^Q$. This is obviously true for maximum statistics of the type $T_Q = \max_{\{i,j\}\subseteq Q, i\ne j} T_{\{i,j\}}$. Furthermore, as in Section 2 we easily obtain that the chi-squared statistic $\sum_{i\in Q} n_i(\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}^Q)^2$ and the mean deviation statistic $\sum_{i\in Q} n_i|\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot}^Q|$ possess this property, where $\bar{X}_{i\cdot} = \sum_{j=1}^{n_i} X_{ij}/n_i$ and $\bar{X}_{\cdot\cdot}^Q = \sum_{i\in Q}\sum_{j=1}^{n_i} X_{ij}/\sum_{r\in Q} n_r$. Under the assumption that the cdf is now given by $F_Q(z) = P_\mu(T_Q \le z)$ for all $\mu \in H_0^Q$, the critical values are defined by

$$c_Q(\alpha_q) = \inf\{c \in \mathbb{R} : F_Q(c) \ge 1 - \alpha_q\},$$

with $Q \subseteq I_k$, $q = |Q| \ge 2$.

We first show that a similar result as Corollary 2.1 cannot hold in general. For example, let X_i be normally distributed with mean 0 and variance $1/n_i$, $n_i \in \mathbb{N}$, $i = 1, \ldots, q$ and let $Q = \{1, \ldots, q\}$, $q \ge 2$. Because $\lim_{n_q\to\infty} P(\max_{1\le i\le q} |X_i| \le c) = P(\max_{1\le i\le q-1} |X_i| \le c)$ for all $c > 0$, it is obvious that there exist examples where $c_Q(\alpha_q) < c_{Q\setminus\{q\}}(\alpha_{q-1})$ for large values of n_q and α_q defined as in Corollary 2.1. Although this result may be disappointing, it is not astonishing, because the adjusted significance levels do not take into account unequal sample sizes. On the other hand, it seems impossible to find a universal choice of significance levels depending on the imbalance concerning the underlying distributions such that the desired monotonicity property is satisfied.

Now although the critical values do not satisfy the monotonicity property $c_Q(\alpha_q) \ge c_P(\alpha_p)$ for all $P \subset Q$ in general, a natural (minimal) requirement should be that for every set Q, $|Q| \ge 3$, there exists at least one $j \in Q$ such that $c_Q(\alpha_q) \ge c_{Q\setminus\{j\}}(\alpha_{q-1})$. A slight generalization of Theorem 2.1 and Corollary 2.1 shows that this property is satisfied for $\alpha_q = 1 - (1 - \alpha)^{q/k}$, $q = 2, \ldots, k$ and under the assumption that the X_i are independently distributed. Again this result is an immediate consequence of the generalized Hölder inequality in Finner (1992).

Table 1. Treatment Means and Sample Sizes of Duncan's (1957) Data

Treatment	1	2	3	4	5	6	7
$\bar{X}_{i\cdot}$	680	734	743	851	873	902	945
n_i	3	2	5	5	3	2	3

Theorem 3.1. Let X_i, $i \in \mathbb{N}$ be independently distributed random variables with values in Ω_i and let $T_Q : \Omega_Q \to \mathbb{R}$, $\Omega_Q = \times_{i\in Q}\Omega_i$ be real-valued Borel-measurable statistics with cdf $F_Q(z) = P(T_Q \le z)$ for all $\emptyset \ne Q \subset \mathbb{N}$. If $T_{Q\setminus\{j\}}(\omega_i : i \in Q\setminus\{j\}) \le T_Q(\omega_i : i \in Q)$ for all $(\omega_i : i \in Q) \in \Omega_Q$, $Q \subset \mathbb{N}$, $q = |Q| \ge 2$, then for all $z \in \mathbb{R}$,

$$F_Q(z) \le \prod_{j\in Q} F_{Q\setminus\{j\}}(z)^{1/(q-1)} \le \max_{j\in Q} F_{Q\setminus\{j\}}(z)^{q/(q-1)}.$$

Corollary 3.1. Let $k \in \mathbb{N}$, $k \ge 2$, $\alpha \in (0, 1)$, $\alpha_q = 1 - (1 - \alpha)^{q/k}$, $q = 1, \ldots, k$, and $c_Q(\alpha_q) = \inf\{c \in \mathbb{R} : F_Q(c) \ge 1 - \alpha_q\}$. Then, under the assumptions of Theorem 3.1, for every Q, $Q \subseteq I_k$, $q = |Q| \ge 2$, there exists a $j \in Q$ such that $c_{Q\setminus\{j\}}(\alpha_{q-1}) \le c_Q(\alpha_q)$.

Proof. By the definition of c_Q and with Theorem 3.1, we obtain $(1 - \alpha)^{q/k} = 1 - \alpha_q \le F_Q(c_Q(\alpha_q)) \le \max_{j\in Q} F_{Q\setminus\{j\}}(c_Q(\alpha_q))^{q/(q-1)}$; hence $1 - \alpha_{q-1} \le \max_{j\in Q} F_{Q\setminus\{j\}}(c_Q(\alpha_q))$. But the last inequality implies the existence of a $j_0 \in Q$ with $1 - \alpha_{q-1} \le F_{Q\setminus\{j_0\}}(c_Q(\alpha_q))$, and now the definition of $c_{Q\setminus\{j_0\}}$ implies $c_{Q\setminus\{j_0\}}(\alpha_{q-1}) \le c_Q(\alpha_q)$.

In practice it is recommended to use the step-down algorithm based on the set of homogeneity hypotheses with adjusted significance levels $\alpha_q = 1 - (1 - \alpha)^{q/k}$, $q = 2, \ldots, k - 2$, and $\alpha_{k-1} = \alpha_k = \alpha$. If the imbalance concerning the distributions P^{X_i} is not too large, then one can hope that at least all critical values c_Q with $3 \le |Q| \le k - 2$ satisfy the monotonicity property $c_{Q\setminus\{j\}}(\alpha_{q-1}) \le c_Q(\alpha_q)$. Another possibility is the use of other conservative procedures, such as the Tukey–Kramer range test procedure in the unbalanced normal case. Finally, we note that the critical values of χ^2- or F-test procedures are invariant against unequal sample sizes.

Table 2. Adjusted Significance Levels and Critical Values for Duncan's (1957) Data; $k = 7$, $\nu = 16$, and $\alpha = .1$

q	α_q	c_q
2	.0297	30,748.634
3	.0442	41,174.748
4	.0584	49,512.687
5	.0725	56,973.847
6	.1000	60,524.268
7	.1000	70,511.229

Table 3. Test Statistics for All Pair Hypotheses H_0^{ij} for Duncan's (1957) Data

i, j	T_{ij}	i, j	T_{ij}
1, 2	3,499.2	3, 4	29,160.0
1, 3	7,441.9	3, 5	31,687.5
1, 4	54,826.9	3, 6	36,115.7
1, 5	55,873.5	3, 7	76,507.5
1, 6	59,140.8	4, 5	907.5
1, 7	105,337.5	4, 6	3,715.7
2, 3	115.7	4, 7	16,567.5
2, 4	19,555.7	5, 6	1,009.2
2, 5	23,185.2	5, 7	7,776.0
2, 6	28,224.0	6, 7	2,218.8
2, 7	53,425.2		

4. EXAMPLE

As a numerical example that may illustrate the usefulness of monotonicity properties of critical values and test statistics, we reanalyze a data set of an unbalanced one-way layout given in Duncan (1957).

The treatment means and sample sizes are given in Table 1. The mean square error is given as $s^2 = 5{,}395.0$ with $\nu = 16$ degrees of freedom. Assuming that the sample means are independent normally distributed with means μ_i and variance σ^2/n_i, σ^2 unknown, we use the test statistics $T_Q = \sum_{i\in Q} n_i(\bar{X}_i. - \bar{X}^Q_{\cdot\cdot})^2$ with corresponding critical values $c_q = s^2(q-1)f_{q-1,\nu,\alpha_q}$ (cf. Table 2) for testing the homogeneity hypotheses H_0^Q, $|Q| = q \geq 2$. Note that T_Q can also be written as $T_Q = \sum_{i,j\in Q, i<j} n_i n_j(\bar{X}_i. - \bar{X}_j.)^2/n^Q_\cdot = \sum_{i,j\in Q,i<j}(n_i + n_j)T_{ij}/n^Q_\cdot$, with $T_{ij} = T_{\{i,j\}}$ and $n^Q_\cdot = \sum_{i\in Q} n_i$; hence Table 3 can be used to compute T_Q. From Tables 2 and 3, we see that $T_{17}, T_{37} > c_7$; hence H_0^{17} and H_0^{37} can be rejected without further tests of homogeneity hypotheses. Furthermore, we find $T_{14}, T_{15}, T_{27} > c_4$, $T_{16} > c_5$, $T_{35}, T_{36} > c_2$, and $T_{ij} < c_2$ for the remaining pairs (i, j), which implies acceptance of the latter hypotheses. Because $T_{\{2,3,5\}} = 37{,}023.6 < c_3$ and $T_{\{2,3,6\}} = 40{,}724.0 < c_3$, H_0^{35} and H_0^{36} are also accepted. Finally, calculation of all T_Q with $|Q| \geq 5$ shows that $T_Q > c_{|Q|}$ for all Q with $|Q| \geq 5$; hence H_0^{14}, H_0^{15}, H_0^{16}, and H_0^{27} can be rejected without explicit tests of all H_0^Q, $|Q| = 3, 4$ with $\{i,j\} \subseteq Q$ for these pair hypotheses.

[*Received November 1991. Revised September 1992.*]

REFERENCES

Duncan, D. B. (1957), "Multiple Range Tests for Correlated and Heteroscedastic Means," *Biometrics,* 13, 164–176.

Dwass, M. (1960), "Some *k*-Sample Rank-Order Tests," in *Contributions to Probability and Statistics. Essays in Honor of H. Hotelling,* eds. I. Olkin, S. G. Ghurye, W. Hoeffding, W. G. Madow and H. B. Mann, Stanford, CA: Stanford University Press, pp. 198–202.

Finner, H. (1990), "Some New Inequalities for the Range Distribution With Application to the Determination of Optimum Significance Levels of Multiple Range Tests," *Journal of the American Statistical Association,* 85, 191–194.

——— (1992), "A Generalization of Hölder's Inequality and Some Probability Inequalities," *The Annals of Probability,* 20, 1893–1901.

Lehmann, E. L., and Shaffer, J. P. (1977), "On a Fundamental Theorem in Multiple Comparisons," *Journal of the American Statistical Association,* 72, 576–578.

——— (1979), "Optimum Significance Levels for Multistage Comparisons Procedures," *The Annals of Statistics,* 7, 27–45.

Miller, R. G., Jr. (1980), *Simultaneous Statistical Inference* (2nd. ed.), New York: Springer.

Nemenyi, P. (1963), "Distribution-Free Multiple Comparisons," unpublished Ph.D. thesis, Princeton University; also issued as a technical report, State University of New York, Downstate Medical Center.

Steel, R. G. D. (1960), "A Rank Sum Test for Comparing all Pairs of Treatments," *Technometrics,* 2, 197–207.