ASYMPTOTIC COMPARISON OF STEP-DOWN AND STEP-UP MULTIPLE TEST PROCEDURES BASED ON EXCHANGEABLE TEST STATISTICS

BY H. FINNER AND M. ROTERS

Universität Trier

In this paper interest is focused on some theoretical investigations concerning the comparison of two popular multiple test procedures, so-called step-down and step-up procedures, in terms of their defining critical values. Such procedures can be applied, for example, to multiple comparisons with a control. In the definition of the critical values for these procedures order statistics play a central role. For $k \in \mathbb{N}_0$ fixed we consider the joint cumulative distribution function (cdf) $P(Y_{1:n} \leq c_1, \ldots, Y_{n-k:n} \leq c_{n-k})$ of the first $n-k$ order statistics and the cdf $P(Y_{n-k:n} \leq c_{n-k})$ of the $(k+1)$th largest order statistic $Y_{n-k:n}$ of n random variables $Y_1, \ldots, Y_n$ belonging to a sequence of exchangeable real-valued random variables. Our interest is focused on the asymptotic behavior of these cdf's and their interrelation if n tends to ∞. It turns out that they sometimes behave completely differently compared with the iid case treated in Finner and Roters so that positive results are only possible under additional assumptions concerning the underlying distribution. We consider different sets of assumptions which then allow analogous results for the exchangeable case. Recently, Dalal and Mallows derived a result concerning the monotonicity of a certain set of critical values in connection with the joint cdf of order statistics in the iid case. We give a counterexample for the exchangeable case underlining the difficulties occurring in this situation. As an application we consider the comparison of certain step-down and step-up procedures in multiple comparisons with a control. The results of this paper yield a more theoretical explanation of the superiority of the step-up procedure which has been observed earlier by comparing tables of critical values. As a by-product we are able to quantify the tightness of the Bonferroni inequality in connection with maximum statistics.

1. Introduction. The main purpose of this paper is some theoretical investigations concerning the comparison of two popular multiple test procedures, namely so-called step-down and step-up procedures. In general, the comparison of different multiple test procedures for a given multiple hypotheses testing problem is, as a result of the multidimensionality of the problem, extremely difficult. In some lucky cases it can be shown that a test procedure strictly dominates another one. For example, often but not always stepwise procedures are uniformly more powerful than their single-step counterparts. So it is well known that the Bonferroni–Holm step-down procedure is uni-

formly more powerful than the classical Bonferroni single-step procedure if two or more hypotheses are under consideration. It is also known that certain step-up procedures (depending on the underlying parameter configuration) are sometimes less powerful than their single-step counterparts; see, for example, Dunnett and Tamhane (1992), who derived a step-up procedure for m comparisons with a control, the so-called many-one problem.

We first give a brief description of step-down and step-up procedures in terms of p-values. Suppose we are interested in testing $m \geq 2$ hypotheses $H_1, \ldots, H_m$ and let $p_1, \ldots, p_m$ denote p-values for testing them. In the many-one situation hypotheses of interest are, for example, $H_i: \vartheta_i - \vartheta_0 \leq 0$ versus $K_i: \vartheta_i - \vartheta_0 > 0$ or the corresponding two-sided hypotheses, where ϑ_i denotes the effect of the ith treatment for $i = 1, \ldots, m$ and ϑ_0 denotes a placebo or standard treatment effect. A step-down procedure based on p-values uses a set of critical values $\alpha_1^{\mathrm{SD}} \geq \cdots \geq \alpha_m^{\mathrm{SD}}$, $\alpha_i^{\mathrm{SD}} \in (0, 1)$, and works as follows. Let $p_{1:m} \leq \cdots \leq p_{m:m}$ denote the ordered p-values and denote the corresponding hypotheses by $H_{(1)}, \ldots, H_{(m)}$. Then a hypothesis $H_{(i)}$ is rejected if and only if $p_{j:m} \leq \alpha_{m-j+1}^{\mathrm{SD}}$ for all $j \leq i$, otherwise it cannot be rejected; that is, the step-down procedure starts with the most significant p-value by comparing it with the smallest α-value and so on. Formally, this procedure can be described as follows. Let $m_1 = \max\{i: p_{j:m} \leq \alpha_{m-j+1}^{\mathrm{SD}}$ for all $j = 1, \ldots, i\}$. Then the step-down procedure rejects $H_{(1)}, \ldots, H_{(m_1)}$. The step-up procedure is also based on a set of critical values, say $\alpha_1^{\mathrm{SU}} \geq \cdots \geq \alpha_m^{\mathrm{SU}}$, $\alpha_i^{\mathrm{SU}} \in (0, 1)$. In contrast to the step-down method the inherent characteristic of the step-up method is to start by comparing the least significant p-value with the largest α-value and so on. More precisely, in short terms, the step-up procedure rejects $H_{(1)}, \ldots, H_{(m_2)}$, where $m_2 = \max\{i: p_{i:m} \leq \alpha_{m-i+1}^{\mathrm{SU}}\}$.

A widely accepted concept in multiple comparisons is the control of a pre-specified multiple level $\alpha \in (0, 1)$, often called the familywise error rate (FWE); that is, the probability of rejecting any true null hypothesis should be bounded by α. For this purpose the critical values have to be chosen appropriately. In any case, the Bonferroni–Holm procedure with $\alpha_i^{\mathrm{SD}} = \alpha/i$ controls the FWE α no matter whether the p-values are independent or not. For independent p-values one can choose the improved values $\alpha_i^{\mathrm{SD}} = 1 - (1 - \alpha)^{1/i}$. The situation for the step-up method is more complicated. In case of independent p-values Hochberg's (1988) choice $\alpha_i^{\mathrm{SU}} = \alpha/i$, that is, the same values as in the original Bonferroni–Holm procedure for possibly dependent p-values, leads to a procedure controlling the FWE α. However, the independence of the p-values allows us to use the values of Rom (1990), which are better for $m \geq 3$ and are recursively defined by $\alpha_1^{\mathrm{SU}} = \alpha$ and

$$\alpha_i^{\mathrm{SU}} = \left[\sum_{j=1}^{i-1} \alpha^j - \sum_{j=1}^{i-2} \binom{i}{j} (\alpha_{j+1}^{\mathrm{SU}})^{i-j} \right] \bigg/ i, \qquad i = 2, \ldots, m.$$

In case the p-values are dependent there exist examples where the step-up procedure with Hochberg's and hence with Rom's critical values does not control the FWE α. On the other hand, some simulations have shown that

Hochberg's critical values often lead to a conservative procedure, that is, with FWE less than α. Recently, Sarkar and Chang (1997) proved that Hochberg's step-up procedure is indeed conservative in certain situations where the test statistics are positively dependent.

It was mentioned by Hochberg (1988) that his step-up procedure is more powerful than the original Bonferroni–Holm procedure. This can easily be seen by comparing m_1 and m_2 defined before. Whenever a step-down and a step-up procedure use the same set of critical values, we always have $m_2 \geq m_1$; that is, the step-up procedure rejects all the hypotheses rejected by the step-down procedure, possibly more. This is obviously the case for Hochberg's and the classical Bonferroni–Holm procedures. However, this is a lame comparison since the step-down procedure with α/i always controls the FWE α while the corresponding step-up procedure may fail to do so. If we restrict attention to step-down and step-up procedures both controlling the FWE α, their comparison becomes more intricate. This can already be seen when independent p-values are at hand. Using the critical values $\alpha_i^{\mathrm{SD}} = 1 - (1 - \alpha)^{1/i}$ in the step-down procedure and Rom's values in the step-up procedure, one has $\alpha_1^{\mathrm{SD}} = \alpha_1^{\mathrm{SU}} = \alpha$ and $\alpha_i^{\mathrm{SD}} > \alpha_i^{\mathrm{SU}}$ for $i \geq 2$. As a consequence, for $m \geq 2$, the number of hypotheses rejected by the step-up method can be less than the corresponding number of the step-down procedure. On the other hand, numerical investigations for $i \geq 2$ show that, for example, for $\alpha \leq 0.2$ the differences $\alpha_i^{\mathrm{SD}} - \alpha_i^{\mathrm{SU}}$ are small and decrease rapidly to 0 when i increases. Moreover, the more informative relative differences $(\alpha_i^{\mathrm{SD}} - \alpha_i^{\mathrm{SU}})/\alpha_i^{\mathrm{SU}}$ also decrease to 0. Hence, it may be argued that the aforementioned reasoning concerning the superiority of step-up over step-down in case of equal critical values will remain valid approximately when the latter differ only slightly.

It is near at hand to ask for theoretical reasons why the difference between step-down and step-up critical values tends to 0 if m increases. The same question arises when the test statistics are dependent. Since the construction of a step-up procedure is difficult for arbitrary dependence structures, we restrict attention to the case where the test statistics are exchangeable under the corresponding null hypotheses. Such a situation appears, for example, in the aforementioned many-one problem considered in Dunnett and Tamhane (1992). A look at tables of critical values shows that step-down and step-up critical values differ only in the second decimal place for small values of i and that the difference again tends to zero if i increases. Step-up procedures which are not covered by the investigations of the present paper but show the same behavior of the critical values when compared with their step-down counterparts are Welsch's (1977) step-up procedure for all pairwise comparisons and a step-up subset selection procedure developed in Finner and Giani (1996).

The main topic of this paper will be a theoretical explanation of the closeness of the critical points for the step-down and step-up procedures. Clearly, asymptotic results do not justify the claim that step-up is better than step-down, but such investigations may help to explain the reason for the observed behavior for moderately large m and give a better understanding of the pro-

cedures. A crucial role in our investigations is played by the joint cdf of order statistics and the cdf of the largest order statistic which are both used for the determination of critical values.

Both stepwise procedures can be viewed as short-cut versions of a so-called closed multiple test procedure. The closure principle of Marcus, Peritz and Gabriel (1976) requires a (nonrandomized) level-α test φ_J (say) for each intersection hypothesis $H_J = \bigcap_{j \in J} H_j$, $\varnothing \neq J \subseteq \{1, \ldots, m\}$. We call a test φ_J a *local test*, and its actually achieved level, that is, $\sup_{\vartheta \in H_J} E_\vartheta \varphi_J$, will be referred to as the *actual level* of φ_J. Then a hypothesis H_J is finally rejected if and only if $\varphi_M = 1$ for all M with $H_M \subseteq H_J$. In the following we suppose that $H_J \neq H_M$ for $J \neq M$. In general, one is mainly interested in the results for $H_j = H_{\{j\}}$, $j = 1, \ldots, m$. Formally, H_j is rejected if all $H_J \subseteq H_j$ are rejected by their local level-α tests. However, an appropriate choice of the level-α tests φ_J finally leads to the step-down and step-up procedure, respectively, as described before. In the following it will be more convenient to work with real-valued test statistics T_i tending to larger values on the alternative of H_i instead of p-values p_i. Since we are interested in asymptotic results, we restrict attention to situations in which an exchangeable sequence of real-valued random variables $(Y_n)_{n \in \mathbb{N}}$ exists such that for each $m \in \mathbb{N}$ and $\varnothing \neq J \subseteq \{1, \ldots, m\}$ the vector $(T_j : j \in J)$ has the same distribution under H_J as $(Y_j : j \in J)$.

Now setting $\varphi_J = 1$ if and only if $\max_{j \in J} T_j > c_{|J|}^{\mathrm{SD}}$ and determining c_j^{SD} as the smallest value c_j satisfying $P(\max_{1 \leq i \leq j} Y_i \leq c_j) \geq 1 - \alpha$ for all $j = 1, \ldots, m$, we arrive at the step-down procedure discussed before, but now described in terms of test statistics instead of p-values. As a consequence, the notation changes slightly. Denoting by $T_{1:m} \leq \cdots \leq T_{m:m}$ the order statistics of $T_1, \ldots, T_m$ and by $H_{(j)}$ the hypothesis corresponding to $T_{m-j+1:m}$, we obtain with $m_1 = \max\{i : T_{m-j+1:m} > c_{m-j+1}^{\mathrm{SD}}$ for all $j = 1, \ldots, i\}$ that the step-down procedure rejects $H_{(1)}, \ldots, H_{(m_1)}$. The step-up procedure is obtained by setting $\varphi_J = 0$ if and only if $T_{j:J} \leq c_j^{\mathrm{SU}}$ for all $j = 1, \ldots, |J|$, where $T_{1:J} \leq \cdots \leq T_{|J|:J}$ denote the order statistics of T_j, $j \in J$. The critical values c_j^{SU} are now determined step by step, beginning with c_1^{SU}, as the smallest value c_j satisfying $P(Y_{1:j} \leq c_1^{\mathrm{SU}}, \ldots, Y_{j-1:j} \leq c_{j-1}^{\mathrm{SU}}, Y_{j:j} \leq c_j) \geq 1 - \alpha$ for $j = 1, \ldots, m$. The step-up procedure rejects $H_{(1)}, \ldots, H_{(m_2)}$, where $m_2 = \max\{i : T_{m-i+1:m} > c_{m-i+1}^{\mathrm{SU}}\}$.

We point out that the monotonicity of the critical values is essential for the validity of the step-down and step-up algorithms. In the step-down case the critical values defined before are always monotone. In the step-up case this is not clear in general. Of course, it is desirable to choose the critical values such that $P(Y_{j:j} \leq c_j^{\mathrm{SD}}) = 1 - \alpha$ and $P(Y_{1:j} \leq c_1^{\mathrm{SU}}, \ldots, Y_{j:j} \leq c_j^{\mathrm{SU}}) = 1 - \alpha$ for all $j = 1, \ldots, m$. This choice is always possible when the Y_i's are continuously distributed. However, the important question remains whether the step-up critical values are monotone. Dalal and Mallows (1992) proved the monotonicity in case the Y_i's are iid with continuous cdf. As a consequence, Rom's step-up critical values are monotone. They satisfy the equations

$P(Y_{1:m} \leq 1 - \alpha_1^{\mathrm{SU}}, \ldots, Y_{m:m} \leq 1 - \alpha_m^{\mathrm{SU}}) = 1 - \alpha$ for all $m \in \mathbb{N}$, where the Y_i's are iid uniformly distributed on the unit interval.

No positive results are available to date for the more general exchangeable case. On the contrary, at the end of Section 3 we give a somewhat artificial example where the monotonicity property is violated. This is regrettable in so far as the assumption of the existence of monotone critical values plays an important role in our theoretical results.

The asymptotic behavior of the critical values for the step-down and step-up methods is closely related to the following question: If either

$$P(Y_{1:n} \leq c_1, \ldots, Y_{n:n} \leq c_n) \quad \text{or} \quad P(Y_{n:n} \leq c_n)$$

converges to $\beta \in [0, 1]$ for n tending to ∞, what will the other one do? Will it have the same limit?

It turns out that an answer to this question is much more intricate than in the special case of an iid sequence $(Y_n)_{n \in \mathbb{N}}$, which was treated in Finner and Roters (1994) and will be summarized briefly in Section 2. There as well as in this paper we consider more generally the relationship between $P(Y_{1:n} \leq c_1, \ldots, Y_{n-k:n} \leq c_{n-k})$ and $P(Y_{n-k:n} \leq c_{n-k})$ for a fixed $k \in \mathbb{N}_0$ and n tending to ∞. The number $k \in \mathbb{N}_0$ as well as large values of $n \in \mathbb{N}$ are of some importance in an application concerning software proofreading in Dalal and Mallows (1992). By the way, the complication induced by the introduction of k does not affect the complexity of the proofs of our theorems. The main results concerning the relationship between

(1.1) $$\lim_{n \to \infty} P(Y_{n-k:n} \leq c_{n-k}) = \beta$$

and

(1.2) $$\lim_{n \to \infty} P(Y_{1:n} \leq c_1, \ldots, Y_{n-k:n} \leq c_{n-k}) = \beta$$

in the exchangeable case will be presented in Section 3. In the following we often refer to (1.1) and (1.2) as the $(n-k)$-*limit property* and *up-to-$(n-k)$-limit property*, respectively. The first theorem in Section 3 gives sufficient conditions for the full equivalence of (1.1) and (1.2). In the following two theorems, which are proved in the Appendix, we impose various assumptions allowing us to establish the validity of (1.1) by assuming (1.2) and vice versa. An example shows that the monotonicity of the sequence $(c_n)_{n \in \mathbb{N}}$ cannot be dispensed with in general to prove that (1.2) implies (1.1).

In Section 4 these results are applied to multiple comparisons with a control under various distributional assumptions. We give conditions for the similar limiting behavior of the critical values and briefly discuss the asymptotic behavior of the actual levels of the local tests in the step-down and step-up procedures.

Moreover, in Section 5 we study the conservativeness of the corresponding procedures based on a Bonferroni adjustment, especially Hochberg's step-up procedure, which turns out to be rather conservative in certain situations.

Finally, our results enable us to quantify the tightness of the Bonferroni inequality in various situations.

2. Asymptotic results in the iid case. To state the result for the iid case, let $k \in \mathbb{N}_0$, $g_k(c) = \exp(-c) \sum_{j=0}^{k} c^j / j!$ for $c \in [0, \infty)$, $g_k(\infty) = 0$ and $F_n^k(x_1, \ldots, x_{n-k}) = P(X_{1:n} \leq x_1, \ldots, X_{n-k:n} \leq x_{n-k})$ for all $(x_1, \ldots, x_{n-k}) \in (-\infty, \infty]^{n-k}$, $n > k$. Then we have the following:

THEOREM 2.1 [Finner and Roters (1994)]. *Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of real-valued iid random variables with cdf F, and let $k \in \mathbb{N}_0$, $c \in [0, \infty]$ and $x_n \in (-\infty, \infty]$ such that $\alpha_n = 1 - F(x_n) < 1$ for all $n \in \mathbb{N}$.*

(a)

$$\lim_{n \to \infty} n\alpha_n = c \tag{2.1}$$

implies

$$\lim_{n \to \infty} F_n^k(x_1, \ldots, x_{n-k}) = g_k(c). \tag{2.2}$$

(b) *If $c < \infty$ or $(x_n)_{n \in \mathbb{N}}$ is nondecreasing, then* (2.2) *implies* (2.1).
(c) $\limsup_{n \to \infty} n\alpha_n = c$ *iff* $\liminf_{n \to \infty} F_n^k(x_1, \ldots, x_{n-k}) = g_k(c)$.
(d) (i) $\liminf_{n \to \infty} n\alpha_n = c$ *implies* $\limsup_{n \to \infty} F_n^k(x_1, \ldots, x_{n-k}) \leq g_k(c)$.
   (ii) *If, in addition, $(x_n)_{n \in \mathbb{N}}$ is a nondecreasing sequence, then*

$$\liminf_{n \to \infty} n\alpha_n = c \quad iff \quad \limsup_{n \to \infty} F_n^k(x_1, \ldots, x_{n-k}) = g_k(c).$$

The bridging between Theorem 2.1 and the limit behavior of $P(X_{n-k:n} \leq x_{n-k})$ is given by the well-known relationship [cf., e.g., Leadbetter, Lindgren and Rootzén (1983), page 32]

$$\lim_{n \to \infty} n\alpha_n = c \quad \text{iff} \quad \lim_{n \to \infty} P(X_{n-k:n} \leq x_{n-k}) = g_k(c),$$

where $c \in [0, \infty]$, or, more generally,

$$\liminf_{n \to \infty} n\alpha_n = c \quad \text{iff} \quad \limsup_{n \to \infty} P(X_{n-k:n} \leq x_{n-k}) = g_k(c),$$

$$\limsup_{n \to \infty} n\alpha_n = c \quad \text{iff} \quad \liminf_{n \to \infty} P(X_{n-k:n} \leq x_{n-k}) = g_k(c).$$

Theorem 2.1 will be the basis for the proofs of the results concerning the exchangeable case in Section 3.

As mentioned before, in case of independent p-values the critical values $\alpha_i^{\mathrm{SD}} = 1 - (1 - \alpha)^{1/i}$ for the step-down and Rom's critical values α_i^{SU} (say) for the step-up procedure satisfy the equations $P(X_{m:m} \leq 1 - \alpha_m^{\mathrm{SD}}) = 1 - \alpha$ and $P(X_{1:m} \leq 1 - \alpha_1^{\mathrm{SU}}, \ldots, X_{m:m} \leq 1 - \alpha_m^{\mathrm{SU}}) = 1 - \alpha$ for all $m \in \mathbb{N}$, where the X_i's are now assumed to be iid uniformly distributed on the unit interval. A comparison of the step-down and step-up procedures has already

been carried out by Finner, Hayter and Roters (1993). Obviously, Theorem 2.1 yields $\lim_{n \to \infty} n\alpha_n^{\mathrm{SD}} = \lim_{n \to \infty} n\alpha_n^{\mathrm{SU}} = -\ln(1 - \alpha)$; that is, α_n^{SD} and α_n^{SU} are asymptotically equal for large n. This proves the numerical findings concerning the critical values of the step-down and step-up procedures described before. Moreover, plugging the step-down critical values into the step-up procedure or vice versa the step-up critical values into the step-down procedure, we obtain asymptotically the same actual levels for the local tests, namely α.

If we work with test statistics instead of p-values, the question is whether the differences $c_n^{\mathrm{SU}} - c_n^{\mathrm{SD}}$ converge to 0. Assuming that the test statistics are iid with continuous cdf F under the corresponding null hypotheses, the critical values are given by $c_n^{\mathrm{SD}} = F^{-1}(1 - \alpha_n^{\mathrm{SD}})$ and $c_n^{\mathrm{SU}} = F^{-1}(1 - \alpha_n^{\mathrm{SU}})$. Generally, the behavior of $c_n^{\mathrm{SU}} - c_n^{\mathrm{SD}}$ for large n is heavily dependent on the tail behavior of F, and, for cdf's with polynomially decreasing tails such as the Cauchy distribution, information about the speed of convergence of $\alpha_n^{\mathrm{SD}} - \alpha_n^{\mathrm{SU}}$ is important. In this connection we could show by rather cumbersome calculations that

$$\lim_{n \to \infty} n^3(\alpha_n^{\mathrm{SD}} - \alpha_n^{\mathrm{SU}}) = c^2/2 \quad \text{for } \alpha \in (0, 1/2),$$

where $c = -\ln(1 - \alpha)$ [cf. Finner and Roters (1997)]. Moreover, we have for $\alpha \in [1/2, 1)$ that $0 \le \liminf_{n \to \infty} n^3(\alpha_n^{\mathrm{SD}} - \alpha_n^{\mathrm{SU}}) \le c^2/2 \le \limsup_{n \to \infty} n^3(\alpha_n^{\mathrm{SD}} - \alpha_n^{\mathrm{SU}}) \le \exp(c)c^2/2$, and $\lim_{n \to \infty} n^3(\alpha_n^{\mathrm{SD}} - \alpha_n^{\mathrm{SU}}) = c^2/2$ in case the limit exists. Altogether, if F is the cdf of a normal, exponential or Cauchy distribution, it can indeed be shown that $\lim_{n \to \infty} (c_n^{\mathrm{SU}} - c_n^{\mathrm{SD}}) = 0$.

3. Asymptotic results in the exchangeable case. For abbreviation we refer to the following setup as model $\mathscr{M}$. Let $(X_n)_{n \in \mathbb{N}}$ be a sequence of iid random variables with values in $(\mathscr{X}, \mathscr{B})$ and let Z be a random variable with values in $(\mathscr{Y}, \mathscr{C})$ such that $(X_n)_{n \in \mathbb{N}}$ and Z are stochastically independent. Moreover, let g denote a real-valued measurable function on $\mathscr{X} \times \mathscr{Y}$, and let without loss of generality a sequence of exchangeable random variables $(Y_n)_{n \in \mathbb{N}}$ be given as $Y_n = g(X_n, Z)$, $n \in \mathbb{N}$ [cf., e.g., Tong (1990), page 111]. For each $z \in \mathscr{Y}$ we write $Y_n^z = g(X_n, z)$ and $G(x|z) = P(g(X_n, z) \le x)$, which is the conditional cdf of Y_n given $Z = z$. Moreover, we use the abbreviation $\overline{G}(x|z) = 1 - G(x|z)$.

We first give a simple example showing that in case of exchangeable random variables both the $(n-k)$-limit and the up-to-$(n-k)$-limit may exist but have different values. For this purpose let $(X_n)_{n \in \mathbb{N}_0}$ be a sequence of iid random variables with pdf $f(x) = \exp(-x)I_{(0, \infty)}(x)$, $x \in \mathbb{R}$, $Y_n = X_n - X_0$ for $n \in \mathbb{N}$, and define $(c_n)_{n \in \mathbb{N}}$ by $c_1 = -\ln 2$ and $c_n = \ln n$ for $n \ge 2$. Then obviously $\lim_{n \to \infty} n P(X_1 > c_n + x) = \lim_{n \to \infty} n \exp(-\ln n - x) = \exp(-x)$ for all $x \in \mathbb{R}$. Hence, Theorem 2.1 entails

$$\lim_{n \to \infty} P\big(Y_{1:n} \le c_1, \ldots, Y_{n-k:n} \le c_{n-k}\big)$$

$$= \lim_{n \to \infty} \int P\big(X_{1:n} \le c_1 + x, \ldots, X_{n-k:n} \le c_{n-k} + x\big) \, dP^{X_0}(x)$$

$$= \int_{\ln 2}^{\infty} g_k(\exp(-x)) \exp(-x)\, dx$$

$$< \int_{0}^{\infty} g_k(\exp(-x)) \exp(-x)\, dx$$

$$= \lim_{n \to \infty} \int P(X_{n-k:n} \le c_{n-k} + x)\, dP^{X_0}(x)$$

$$= \lim_{n \to \infty} P(Y_{n-k:n} \le c_{n-k}),$$

which reveals that for obtaining equality in this inequality the condition $P^{X_0}(\bigcap_{n \in \mathbb{N}}\{x \in \mathbb{R}\colon P(X_1 > c_n + x) < 1\}) = 1$ is indispensable.

In view of this example one of the main objectives will be to elaborate suitable conditions which allow us to prove the equivalence of the $(n-k)$-limit property (1.1) and the up-to-$(n-k)$-limit property (1.2). The first theorem of this section provides sufficient conditions which yield the full equivalence of (1.1) and (1.2). Its proof is nearly trivial but the verification of the first condition can be very hard.

THEOREM 3.1. *Given model $\mathscr{M}$, let $c_n \in (-\infty, \infty]$, $n \in \mathbb{N}$, and suppose that the sets*

$$A = \left\{ z \in \mathscr{Z}\colon \lim_{n \to \infty} n\overline{G}(c_n|z) \text{ exists in } [0, \infty] \right\}$$

and

$$B = \bigcap_{n \in \mathbb{N}} \left\{ z \in \mathscr{Z}\colon \overline{G}(c_n|z) < 1 \right\}$$

both have P^Z-probability 1, that is, $P(Z \in A \cap B) = 1$. Then for all $k \in \mathbb{N}_0$ the following limits exist and it holds

$$\lim_{n \to \infty} P(Y_{n-k:n} \le c_{n-k}) = \lim_{n \to \infty} P(Y_{1:n} \le c_1, \ldots, Y_{n-k:n} \le c_{n-k}) = E g_k(r(Z)),$$

where r is measurable with $r(z) = \lim_{n \to \infty} n\overline{G}(c_n|z)$ for all $z \in A$.

PROOF. From Theorem 2.1 we obtain, by conditioning on $Z = z$, $z \in A \cap B$,

$$\lim_{n \to \infty} P(Y_{n-k:n} \le c_{n-k}) = \int_{A \cap B} \lim_{n \to \infty} P(Y_{n-k:n}^z \le c_{n-k})\, dP^Z(z)$$

$$= \int_{A \cap B} g_k(r(z))\, dP^Z(z) = E g_k(r(Z))$$

$$= \int_{A \cap B} \lim_{n \to \infty} P(Y_{1:n}^z \le c_1, \ldots, Y_{n-k:n}^z \le c_{n-k})\, dP^Z(z)$$

$$= \lim_{n \to \infty} P(Y_{1:n} \le c_1, \ldots, Y_{n-k:n} \le c_{n-k}). \qquad \square$$

The next two theorems in this section, the proofs of which are given in the Appendix, work according to the following scheme: Either the $(n-k)$-limit property (1.1) or the up-to-$(n-k)$-limit property (1.2) together with some

further suitable requirements is assumed to be fulfilled which then allows the verification of $P(Z \in A \cap B) = 1$ in Theorem 3.1. A key requirement will be the existence of $\lim_{x \to \infty} \overline{G}(x|z)/\overline{G}(x|z_0)$ for all z, $z_0 \in \mathscr{Z}$ and the structure of this limit as a function of z or of (z, z_0). We start with the case where these limits are almost surely positive and finite.

THEOREM 3.2. *Given model $\mathscr{M}$, let $c_n \in (-\infty, \infty]$, $n \in \mathbb{N}$, $k \in \mathbb{N}_0$, $\beta \in [0, 1]$, $P(Z \in B) = 1$, and suppose there exists a $z_0 \in \mathscr{Z}$ such that the following conditions are fulfilled*:

(3.1)
$$\forall\, x \in \mathbb{R}\colon \overline{G}(x|z_0) > 0,$$
$$\forall\, z \in \mathscr{Z}\colon \exists\, t(z) := \lim_{x \to \infty} \overline{G}(x|z)\big/\overline{G}(x|z_0) \quad and \quad P\big(t(Z) \in (0, \infty)\big) = 1.$$

Then we obtain that:

(a) *the $(n - k)$-limit property* (1.1) *implies the up-to-$(n - k)$-limit property* (1.2).
  (b) (i) *If $\beta > 0$, or*,
      (ii) *if $\beta = 0$, $(c_n)_{n \in \mathbb{N}}$ is nondecreasing and*

(3.2) $P\big(\overline{G}(x|z_0) \geq \overline{G}(x|Z) \text{ for eventually all large } x \in \mathbb{R}\big) > 0,$

then (1.2) *implies* (1.1).

It is worth noting that under the assumptions of Theorem 3.2 we have the full equivalence of the $(n - k)$-limit property (1.1) and the up-to-$(n - k)$-limit property (1.2) for $\beta > 0$; that is, the main statements are the same as in Theorem 2.1. This changes in case $\lim_{x \to \infty} \overline{G}(x|z)/\overline{G}(x|z_0) \in \{0, \infty\}$ for almost all z, z_0 with $z \neq z_0$. The next theorem covers this situation. It turns out that it is only possible to prove that the up-to-$(n - k)$-limit property (1.2) implies the $(n - k)$-limit property (1.1) if the sequence $(c_n)_{n \in \mathbb{N}}$ is nondecreasing. Otherwise there exist counterexamples, one of which will be sketched after Theorem 3.3. To obtain a positive result at all, additional assumptions seem to be unavoidable. The main property utilized will be a certain ordering structure on $\mathscr{Z}$. Because of the rather general distributions P^Z dealt with in the theorem, a certain continuity assumption will be made.

THEOREM 3.3. *Given model $\mathscr{M}$, let $c_n \in (-\infty, \infty]$, $n \in \mathbb{N}$, $k \in \mathbb{N}_0$, $\beta \in [0, 1]$, $P(Z \in B) = 1$, and suppose the following conditions are fulfilled*:

$$\forall\, z \in \mathscr{Z}\colon \forall\, x \in \mathbb{R}\colon \overline{G}(x|z) > 0,$$
$$\forall\, z, z_0 \in \mathscr{Z}\colon \exists\, t(z|z_0) := \lim_{x \to \infty} \overline{G}(x|z)\big/\overline{G}(x|z_0),$$
$$\forall\, z_0 \in \mathscr{Z}\colon P(t(Z|z_0) \in \{0, \infty\}) = 1,$$
$$\forall\, z_0, z_1 \in \mathscr{Z}\colon t(\cdot|z_0) \leq t(\cdot|z_1) \;\; [P^Z] \quad or \quad t(\cdot|z_0) \geq t(\cdot|z_1) \;\; [P^Z],$$
$$\forall\, z_0 \in \mathscr{Z}\colon P\big(t(\cdot|Z) = t(\cdot|z_0) \, [P^Z]\big) = 0.$$

In this case we have the following statements:

(a) *The $(n-k)$-limit property* (1.1) *implies the up-to-$(n-k)$-limit property* (1.2).

(b) *If $(c_n)_{n \in \mathbb{N}}$ is nondecreasing, then* (1.2) *implies* (1.1).

We briefly sketch an example showing that the monotonicity requirement in the second part of the last theorem cannot be dispensed with in order to conclude the $(n-k)$-limit property (1.1) from the up-to-$(n-k)$-limit property (1.2). Let $(X_n)_{n \in \mathbb{N}_0}$ be a sequence of iid standard normal random variables, $\beta \in (0, 1)$, $k \in \mathbb{N}_0$. Define $c_n = u_{1/n} - u_\beta$ for $n \in \mathbb{N}$ even, and $c_n = n$ for $n \in \mathbb{N}$ odd, where Φ is the cdf of the standard normal distribution, and u_γ denotes the $(1-\gamma)$-quantile, $\gamma \in (0, 1)$, of the standard normal distribution. Then one can show that for $Y_n = X_n - X_0$, $n \in \mathbb{N}$, the up-to-$(n-k)$-limit property (1.2) holds, whereas the $(n-k)$-limit (1.1) does not exist. This example also reveals that the behavior of the joint cdf of order statistics for exchangeable random variables is sometimes completely different from that for iid random variables in the following sense. Suppose we have $\lim_{n \to \infty} P(X_{1:n} \leq x_1, \ldots, X_{n:n} \leq x_n) = \beta \in (0, 1)$ for iid random variables X_i, $i \in \mathbb{N}$. Then $\lim_{n \to \infty} P(X_{1:n} \leq x_1, \ldots, X_{n-k:n} \leq x_{n-k}) = g_k(-\ln \beta) > \beta$ for $k \in \mathbb{N}$. In contrast to this, the sequence of exchangeable random variables defined before satisfies $\lim_{n \to \infty} P(Y_{1:n} \leq c_1, \ldots, Y_{n-k:n} \leq c_{n-k}) = \beta$ for all $k \in \mathbb{N}_0$. This observation is important with regard to numerical computations because it shows that one has to be very careful in the construction of a sequence of critical values satisfying $P(Y_{1:n} \leq c_1, \ldots, Y_{n:n} \leq c_n) = \beta$ for all $n \in \mathbb{N}$ since for large n upward deviations from the true value of c_n in $P(Y_{1:n} \leq c_1, \ldots, Y_{n:n} \leq c_n)$ may have no significant effect on the resulting probability. In view of the last example and part (b) of Theorem 3.3, the monotonicity of the sequence $(c_n)_{n \in \mathbb{N}}$ is without doubt a desirable property. It is more than regrettable that in the important case where the critical values are chosen such that $P(Y_{1:j} \leq c_1^{\mathrm{SU}}, \ldots, Y_{j:j} \leq c_j^{\mathrm{SU}}) = 1 - \alpha$ for all $j = 1, \ldots, m$, their monotonicity cannot be guaranteed in general. The following disillusioning counterexample dashes the hope of the full generalization of Dalal and Mallows' (1992) result to the exchangeable case.

COUNTEREXAMPLE 3.4.   Let $f = 2I_{[0, 0.15]} + 2I_{[0.65, 1]}$ and $g = 2I_{(0.15, 0.65)}$ be pdf's with respect to Lebesgue measure, and suppose that $(V_n)_{n \in \mathbb{N}}$, $(W_n)_{n \in \mathbb{N}}$ are sequences of iid random variables with $V_1 \sim f$, $W_1 \sim g$. Moreover, let $Z \sim \mathrm{Bin}(1, 1/2)$ and assume $(V_n)_{n \in \mathbb{N}}$, $(W_n)_{n \in \mathbb{N}}$ and Z are independent. If $Y_n = (1 - Z)V_n + ZW_n$, $n \in \mathbb{N}$, then $(Y_n)_{n \in \mathbb{N}}$ is a sequence of real-valued exchangeable random variables, and with $c_1 = 0.4$, $c_2 = 0.63$ as well as $\beta = 0.4$ we obtain

$$P(Y_1 \leq c_1) = P(Y_{1:2} \leq c_1, Y_{2:2} \leq c_2) = \beta$$

but

$$P(Y_{1:3} \leq c_1, Y_{2:3} \leq c_2, Y_{3:3} \leq c_2) = 0.4072 > \beta;$$

that is, the c_3 satisfying $P(Y_{1:3} \leq c_1, Y_{2:3} \leq c_2, Y_{3:3} \leq c_3) = \beta$ has to be less than c_2. $\square$

Reviewing the results of this section, one may think that the lack of a more general characterization of the limit behavior of $P(Y_{1:n} \leq c_1, \ldots, Y_{n-k:n} \leq c_{n-k})$ in terms of the limit behavior of $P(Y_{n-k:n} \leq c_{n-k})$ and vice versa is unsatisfactory, especially the need for the condition $P(Z \in A) = 1$. The question whether this condition is indispensable is difficult to answer. However, in case $P(Z \in A) < 1$, $P(Z \in B) = 1$ and $(c_n)_{n \in \mathbb{N}}$ is monotone, it seems to be possible that, for example, $\lim_{n \to \infty} P(Y_{n-k:n} \leq c_{n-k}) = \beta$ but $\liminf_{n \to \infty} P(Y_{1:n} \leq c_1, \ldots, Y_{n-k:n} \leq c_{n-k}) < \beta$, but yet we have not been successful in constructing such a situation.

4. Application in multiple comparisons with a control. As an application of the previous section we now study the asymptotic behavior of actual levels and critical values for step-down and step-up test procedures in many-one problems, that is, comparison of m parameters ϑ_i with a control parameter ϑ_0. Depending on the model (with or without nuisance parameter) and hypotheses (one- or two-sided, i.e., $H_i: \vartheta_i - \vartheta_0 \leq 0$ or $H_i: \vartheta_i - \vartheta_0 = 0$), we consider four types of test statistics, namely, $Y_i = X_i - X_0$, $Y_i = |X_i - X_0|$, $Y_i = (X_i - X_0)/S$, $Y_i = |X_i - X_0|/S$, $i = 1, \ldots, m$, where $X_0, X_1, \ldots, X_m, S$ are assumed to be independent real-valued random variables and where the distribution of S is independent of the ϑ_i's and arbitrary with support $(0, \infty)$. Moreover, $X_1, \ldots, X_m$ are assumed to be iid under $\vartheta_1 = \cdots = \vartheta_m$. The cdf of X_1 will be denoted by F, which may be the cdf of (i) a normal distribution $N(\vartheta, \sigma^2)$ with $\vartheta \in \mathbb{R}$ unknown, $\sigma^2 > 0$ known or unknown, (ii) an exponential distribution $\mathrm{Exp}(\vartheta, \sigma)$ with Lebesgue density $f(x) = (1/\sigma) \exp((x - \vartheta)/\sigma) I_{[\vartheta, \infty)}(x)$ or (iii) a Cauchy distribution $\mathrm{Cauchy}(\vartheta)$ with location parameter ϑ. In the latter case we only study the nonstudentized test statistics. In the following calculations we always assume without loss of generality $\vartheta_0 = \vartheta_1 = \cdots = \vartheta_m = 0$ and $\sigma^2 = 1$. With the notation of the previous section Z is equal to X_0 or (X_0, S). Remember that the critical values of the step-down and step-up procedures are determined by $P(Y_{j:j} \leq c_j^{\mathrm{SD}}) = 1 - \alpha$ and $P(Y_{1:j} \leq c_1^{\mathrm{SU}}, \ldots, Y_{j:j} \leq c_j^{\mathrm{SU}}) = 1 - \alpha$ for all $j = 1, \ldots, m$. Before we are able to apply the results of the last section, we have to study the behavior of $\overline{G}(x|z)/\overline{G}(x|z_0)$ for $x \to \infty$, where, as before, $G(\cdot|z)$ denotes the conditional cdf of Y_i given $Z = z$. It turns out that in all cases $t(z|z_0) = \lim_{x \to \infty} \overline{G}(x|z)/\overline{G}(x|z_0)$ exists and enjoys all the corresponding properties required either in Theorem 3.2 or in Theorem 3.3.

(A) If $Y_i = X_i - X_0$, then $G(x|z) = F(x + z)$ and

$$t(z|z_0) = \begin{cases} \begin{cases} \infty, \\ 1, \quad \text{for} \begin{cases} z < z_0 \\ z = z_0, \\ z > z_0 \end{cases} \quad \text{if } F \sim N(0, 1), \\ 0, \end{cases} \\ \exp(z_0 - z), \quad \text{if } F \sim \mathrm{Exp}(0, 1), \\ 1, \quad\quad\quad\quad \text{if } F \sim \mathrm{Cauchy}(0). \end{cases}$$

(B) If $Y_i = |X_i - X_0|$, then $G(x|z) = F(x + z) - F(-x + z)$ and

$$t(z|z_0) = \begin{cases} \begin{cases} \infty, \\ 1, \quad \text{for} \\ 0, \end{cases} \begin{cases} |z| > |z_0| \\ |z| = |z_0|, \\ |z| < |z_0| \end{cases} \qquad \text{if } F \sim N(0, 1), \\ \exp(z_0 - z), \quad \text{if } F \sim \text{Exp}(0, 1), \\ 1, \qquad\qquad \text{if } F \sim \text{Cauchy}(0). \end{cases}$$

(C) If $Y_i = (X_i - X_0)/S$, then $G(x|z) = F(xs + u)$ [with $z = (u, s)$] and

$$t(u, s|u_0, s_0) = \begin{cases} \begin{cases} \infty, \\ 1, \quad \text{for} \\ 0, \end{cases} \begin{cases} u < u_0, \ s = s_0 \text{ or } s < s_0 \\ u = u_0, \ s = s_0, \\ u > u_0, \ s = s_0 \text{ or } s > s_0 \end{cases} \quad \text{if } F \sim N(0, 1), \\ \begin{cases} \infty, \\ \exp(u_0 - u), \quad \text{for} \\ 0, \end{cases} \begin{cases} s < s_0 \\ s = s_0, \\ s > s_0 \end{cases} \qquad \text{if } F \sim \text{Exp}(0, 1). \end{cases}$$

(D) If $Y_i = |X_i - X_0|/S$, then $G(x|z) = F(xs + u) - F(-xs + u)$ [with $z = (u, s)$] and

$$t(u, s|u_0, s_0) = \begin{cases} \begin{cases} \infty, \\ 1, \quad \text{for} \\ 0, \end{cases} \begin{cases} |u| > |u_0|, \ s = s_0 \text{ or } s < s_0 \\ |u| < |u_0|, \ s = s_0 \text{ or } s > s_0, \quad \text{if } F \sim N(0, 1), \\ |u| = |u_0|, \ s = s_0 \end{cases} \\ \begin{cases} \infty, \\ \exp(u_0 - u), \quad \text{for} \\ 0, \end{cases} \begin{cases} s > s_0 \\ s = s_0, \\ s < s_0 \end{cases} \qquad \text{if } F \sim \text{Exp}(0, 1). \end{cases}$$

The structure of the t-functions now tells us whether Theorem 3.2 or Theorem 3.3 is applicable. In case of $t(z) \in (0, \infty)$ $[P^Z]$, we are in the situation of Theorem 3.2; in case of $t(z) \in \{0, \infty\}$ $[P^Z]$, Theorem 3.3 is applicable. We note that the ordering structure required in Theorem 3.3 is given in all cases. Hence, by Theorem 3.2, if $F \sim \text{Exp}(0, 1)$ or $F \sim \text{Cauchy}(0)$ in (A) or (B) we have the full equivalence of the $(n - k)$-limit property (1.1) and the up-to-$(n - k)$-limit property (1.2) for $\beta > 0$. In all the other cases, especially in the important case of $F \sim N(0, 1)$, (1.1) implies (1.2), but Theorem 3.3 yields that (1.2) implies (1.1) only if $(c_n)_{n \in \mathbb{N}}$ is a nondecreasing sequence, which is, in view of the unsolved monotonicity problem, a blemish without doubt. Nevertheless, the use of the step-down critical values, which are always monotone, in the step-up procedure yields actual levels nearly equal to α for the corresponding local tests φ_J when $|J|$ tends to ∞. This observation may be taken as a first hint of the validity of the conjecture that the step-down and the step-up critical values do not differ very much in the examples considered before.

As in the iid case a direct comparison of the critical values of the step-down and the step-up procedures would be desirable. But such a comparison turns out to be more intricate than in the iid case. We only state the mere results; details will be reported elsewhere. For example, in the situation of (A) with $X_i \sim$ Exp(0, 1), $i \in \mathbb{N}_0$, it can be shown by using a method similar to that in the iid case that the differences between the exact step-down and the corresponding exact step-up critical values, that is, $c_m^{\text{SD}} - c_m^{\text{SU}}$, tend to 0 for $m \to \infty$. This could also be concluded for normally distributed X_i's provided it turned out one day that the step-up critical values are monotone in this case. These considerations underline again the importance of the monotonicity question. In the Cauchy case we were not able to prove anything concerning the differences of the critical values. However, in this case the quotients $c_m^{\text{SD}}/c_m^{\text{SU}}$ tend to 1 for $m \to \infty$.

What are the practical implications of these findings? Well, the calculation of the step-up critical values in the many-one situation is cumbersome, especially if an additional nuisance parameter is at hand and the number of hypotheses increases. Then one can calculate, for example, the first four or five step-up values exactly. If the difference $c_i^{\text{SU}} - c_i^{\text{SD}}$ is negligible for $i = 4, 5$, one may use the step-down values as an approximation for the exact step-up values for $i > 5$, hoping that this approximation behaves well in view of the asymptotic results.

5. How tight is the Bonferroni inequality? Finally, we briefly shed some light on Bonferroni-type adjustments. Since the calculation of exact critical values for a step-up procedure is cumbersome, it is tempting to try the Bonferroni adjustment as proposed by Hochberg (1988). Unfortunately, it is not known whether such a step-up procedure based on exchangeable random variables controls the given multiple level α. Sometimes it is argued in the literature, mostly based on simulation studies, that such a procedure should be conservative. As noted already in Section 1, the recent work of Sarkar and Chang (1997) shows that Hochberg's procedure is conservative for certain positively dependent test statistics. Our results now allow us to study the asymptotic behavior of the actual levels of the local tests φ_J for the intersection hypotheses $H_J = \bigcap_{i \in J} H_i$ for $|J| \to \infty$. As an example we consider the situation in item (A) of the preceding section for the normal, exponential and Cauchy distributions. Using the Bonferroni levels $\alpha_j^B = \alpha/j$ leads to critical values c_j^B defined by $P(Y_1 \leq c_j^B) = 1 - \alpha_j^B$, $j \in \mathbb{N}$. The step-down procedure using these critical values is nothing but the classical Bonferroni–Holm procedure which always controls the multiple level α. The actual levels of the local tests φ_J occurring in the constructions of the step-down and the step-up procedures via the closure principle are given by $1 - P(Y_{j:j} \leq c_j^B)$ and $1 - P(Y_{1:j} \leq c_1^B, \ldots, Y_{j:j} \leq c_j^B)$, respectively, where $j = |J|$. To obtain the actual levels for $j \to \infty$, in view of Theorem 3.1 we only have to study the limit behavior of $j\overline{G}(c_j^B|z) = jP(X_1 - z > c_j^B)$ for $j \to \infty$. Now, if $X_j \sim N(0, 1)$, $j \in \mathbb{N}_0$, we obtain $c_j^B = \sqrt{2}u_{\alpha/j}$ and $\lim_{j \to \infty} jP(X_1 - z > c_j^B) = 0$ for all $z \in \mathbb{R}$; hence, $\lim_{j \to \infty} P(Y_{j:j} \leq c_j^B) = \lim_{j \to \infty} P(Y_{1:j} \leq c_1^B, \ldots, Y_{j:j} \leq c_j^B) = 1$.

This means that the probability of rejecting H_J tends to 0 for $|J| \to \infty$ in both the step-down and the step-up procedure, provided that H_J is true; that is, the actual level of φ_J tends to 0. Moreover, the Bonferroni inequality $P(Y_{j:j} > c_j^B) = P(\cup_{i \in J}\{Y_i > c_j^B\}) \le \sum_{i \in J} P(Y_i > c_j^B) = j\alpha_j^B = \alpha$ becomes extremely conservative in this case as $j = |J|$ increases.

If $X_j \sim \text{Exp}(0, 1)$, $j \in \mathbb{N}_0$, we obtain $c_j^B = -\ln(2\alpha/j)$ for $j \ge 2\alpha$, which implies $jP(X_1 - z > c_j^B) = 2\alpha \exp(-z)$ for all $z > 0$. This yields $\lim_{j \to \infty} P(Y_{j:j} \le c_j^B) = \lim_{j \to \infty} P(Y_{1:j} \le c_1^B, \dots, Y_{j:j} \le c_j^B) = (1 - \exp(-2\alpha))/(2\alpha) > 1 - \alpha$, but the left-hand side of this inequality is only slightly greater than the right-hand side for small values of α. So the Bonferroni inequality works quite well in this case. If the X_j, $j \in \mathbb{N}_0$, are Cauchy distributed we obtain the value $\exp(-\alpha/2)$ for the left-hand side of the aforementioned inequality, that is, a considerable loss.

Hence, in all three examples treated we obtain conservative local tests φ_J for large values of $|J|$. This may explain why Hochberg's step-up procedure for dependent test statistics in simulation studies mostly turned out to be conservative. It is worth mentioning that for exchangeable normal random variables Y_j, $j \in \mathbb{N}$, with correlation $\rho > 0$ we also have $\lim_{j \to \infty} P(Y_{j:j} \le c_j^B) = \lim_{j \to \infty} P(Y_{1:j} \le c_1^B, \dots, Y_{j:j} \le c_j^B) = 1$, exactly as in the many-one situation considered before, in which $\rho = 1/2$.

For iid random variables $(X_n)_{n \in \mathbb{N}}$ with continuous cdf F the results of Section 2 yield

$$\lim_{j \to \infty} P(X_{j:j} \le c_j^B) = \lim_{j \to \infty} P(X_{1:j} \le c_1^B, \dots, X_{j:j} \le c_j^B) = g_0\left(\lim_{j \to \infty} j\alpha_j^B\right)$$

$$= g_0(\alpha) = \exp(-\alpha) > 1 - \alpha;$$

that is, for small values of α the actual levels occurring in the step-down and step-up procedures are slightly smaller than α for large values of j if Bonferroni adjusted critical values are used. It can easily be seen that in the three examples concerning the exchangeable case the loss by using Bonferroni adjusted critical values is always greater than in the independent case. This is not very surprising because exchangeable random variables are positively dependent.

APPENDIX

PROOF OF THEOREM 3.2. (a) The case $\beta = 0$ being trivial, we may assume $\beta > 0$, in which case $\lim_{n \to \infty} c_n = \infty$ by virtue of (3.1). Let $c \in [0, \infty]$ and suppose that there exists a subsequence $(n_j)_{j \in \mathbb{N}}$ of $\mathbb{N}$ with $\lim_{j \to \infty}(n_j - k)\overline{G}(c_{n_j-k}|z_0) = c$. Then the assumptions of the theorem imply

$$\lim_{j \to \infty}(n_j - k)\overline{G}(c_{n_j-k}|z) = \lim_{j \to \infty}(n_j - k)\overline{G}(c_{n_j-k}|z_0)\big[\overline{G}(c_{n_j-k}|z)/\overline{G}(c_{n_j-k}|z_0)\big]$$

$$= ct(z) \quad [P^Z].$$

Hence, with Theorem 2.1, we obtain

$$\beta = \lim_{j \to \infty} P\big(Y_{n_j-k:n_j} \le c_{n_j-k}\big)$$

$$= \int \lim_{j \to \infty} P\big(Y^z_{n_j-k:n_j} \le c_{n_j-k}\big) \, dP^Z(z)$$

$$= \int g_k\big(ct(z)\big) \, dP^Z(z).$$

The last equation determines c uniquely. Hence, $\lim_{n \to \infty}(n-k)\overline{G}(c_{n-k}|z_0) = c$, thus $\lim_{n \to \infty} n\overline{G}(c_n|z_0) = c$, and finally $\lim_{n \to \infty} n\overline{G}(c_n|z) = ct(z)$ $[P^Z]$. Thus, $P(Z \in A) = 1$, and together with $P(Z \in B) = 1$ the assertion of (a) follows from Theorem 3.1.

(b)(i) Let $c = \limsup_{n \to \infty} n\overline{G}(c_n|z_0)$ and let $(n_j)_{j \in \mathbb{N}}$ be a subsequence of $\mathbb{N}$ such that $\lim_{j \to \infty}(n_j - k)\overline{G}(c_{n_j-k}|z_0) = c$. If $c < \infty$ we obtain similarly to before that $\lim_{n \to \infty} c_n = \infty$ and

$$\limsup_{n \to \infty} n\overline{G}(c_n|z) = \lim_{j \to \infty} (n_j - k)\overline{G}(c_{n_j-k}|z) = ct(z) \quad [P^Z].$$

For $c = \infty$ the same relation holds true because the sequence

$$\big(\overline{G}(c_n|z)/\overline{G}(c_n|z_0)\big)_{n \in \mathbb{N}}$$

is bounded away from 0 P^Z-a.s. Together with Theorem 2.1 we obtain

$$g_k\big(ct(z)\big) = \liminf_{n \to \infty} P\big(Y^z_{1:n} \le c_1, \ldots, Y^z_{n-k:n} \le c_{n-k}\big)$$

$$\le \liminf_{j \to \infty} P\big(Y^z_{1:n_j} \le c_1, \ldots, Y^z_{n_j-k:n_j} \le c_{n_j-k}\big)$$

$$\le \limsup_{j \to \infty} P\big(Y^z_{1:n_j} \le c_1, \ldots, Y^z_{n_j-k:n_j} \le c_{n_j-k}\big)$$

$$\le \lim_{j \to \infty} P\big(Y^z_{n_j-k:n} \le c_{n_j-k}\big)$$

$$= g_k\big(ct(z)\big) \quad [P^Z].$$

Thus, $\lim_{j \to \infty} P\big(Y^z_{1:n_j} \le c_1, \ldots, Y^z_{n_j-k:n_j} \le c_{n_j-k}\big) = g_k(ct(z))$ $[P^Z]$. This implies

$$\beta = \lim_{j \to \infty} P\big(Y_{1:n_j} \le c_1, \ldots, Y_{n_j-k:n_j} \le c_{n_j-k}\big)$$

$$= \int \lim_{j \to \infty} P\big(Y^z_{1:n_j} \le c_1, \ldots, Y^z_{n_j-k:n_j} \le c_{n_j-k}\big) \, dP^Z(z)$$

(A.1) $$= \int g_k\big(ct(z)\big) \, dP^Z(z)$$

$$= \int \lim_{j \to \infty} P\big(Y^z_{n_j-k:n_j} \le c_{n_j-k}\big) \, dP^Z(z)$$

$$= \lim_{j \to \infty} P\big(Y_{n_j-k:n_j} \le c_{n_j-k}\big).$$

Especially, this computation entails that $c < \infty$. If $c = 0$, then

$$\lim_{n \to \infty} n\overline{G}(c_n|z) = 0 \quad [P^Z].$$

Hence, Theorem 3.1 implies the desired result. So let $0 < c < \infty$ and suppose

$$\tilde{c} = \liminf_{n \to \infty} n\overline{G}(c_n|z_0) < c.$$

A little reflection reveals that we can select a subsequence $(m_j)_{j \in \mathbb{N}}$ of $\mathbb{N}$ such that $c_{m_j-k-1} < c_{m_j-k}$ for all $j \in \mathbb{N}$ and $\lim_{j \to \infty}(m_j-k)\overline{G}(c_{m_j-k}|z_0) = \tilde{c}$. Denote by F_n^k the joint distribution function of the first $n - k$ order statistics of n iid random variables having a uniform distribution on the unit interval $[0, 1]$. Define $\alpha_n = \alpha_n(z) = \overline{G}(c_n|z)$, $\beta_n = \beta_n(z) = \max\{\alpha_n(z), \min\{ct(z)/n, 1/2\}\}$, $n \in \mathbb{N}$, which implies $\lim_{n \to \infty} n\beta_n(z) = ct(z)$ and $\liminf_{n \to \infty} n\alpha_n(z) = \lim_{j \to \infty}(m_j - k)\alpha_{m_j-k}(z) = \tilde{c}t(z) \, [P^Z]$. Hence, we obtain

$$P\big(Y_{1:m_j}^z \leq c_1, \ldots, Y_{m_j-k:m_j}^z \leq c_{m_j-k}\big)$$

$$= F_{m_j}^k(1 - \alpha_1, \ldots, 1 - \alpha_{m_j-k})$$

$$\geq F_{m_j}^k(1 - \beta_1, \ldots, 1 - \beta_{m_j-k-1}, \, 1 - \alpha_{m_j-k})$$

$$= F_{m_j}^k(1 - \beta_1, \ldots, 1 - \beta_{m_j-k})$$

$$\quad + \binom{m_j}{k+1}\big(\beta_{m_j-k}^{k+1} - \alpha_{m_j-k}^{k+1}\big)F_{m_j-k-1}^0(1 - \beta_1, \ldots, 1 - \beta_{m_j-k-1}).$$

Together with the aforementioned assumptions and Theorem 2.1, taking the lim inf on both sides yields

$$\liminf_{n \to \infty} P\big(Y_{1:m_j}^z \leq c_1, \ldots, Y_{m_j-k:m_j}^z \leq c_{m_j-k}\big)$$

$$\geq g_k(ct(z)) + \frac{(t(z))^{k+1}}{(k+1)!}(c^{k+1} - \tilde{c}^{k+1})g_0(ct(z))$$

$$> g_k(ct(z)) \quad [P^Z].$$

On the other hand, we obtain with Fatou's lemma and (A.1) that

$$\beta = \lim_{n \to \infty} P(Y_{1:n} \leq c_1, \ldots, Y_{n-k:n} \leq c_{n-k})$$

$$= \liminf_{j \to \infty} P(Y_{1:m_j} \leq c_1, \ldots, Y_{m_j-k:m_j} \leq c_{m_j-k})$$

$$\geq \int \liminf_{j \to \infty} P(Y_{1:m_j}^z \leq c_1, \ldots, Y_{m_j-k:m_j}^z \leq c_{m_j-k})\, dP^Z(z)$$

$$> \int g_k(ct(z))\, dP^Z(z)$$

$$= \lim_{j \to \infty} P(Y_{n_j-k:n_j} \leq c_{n_j-k})$$

$$= \beta,$$

which is absurd. Hence, $\tilde{c} = c < \infty$, and similarly to part (a) we conclude that $\lim_{n \to \infty} n\overline{G}(c_n|z) = ct(z) \, [P^Z]$, so Theorem 3.1 implies the result for $\beta > 0$.

(ii) Let $\lim_{n \to \infty} P(Y_{1:n} \leq c_1, \ldots, Y_{n-k:n} \leq c_{n-k}) = 0$ and suppose $c = \liminf_{n \to \infty} n\overline{G}(c_n|z_0) < \infty$. Since $(c_n)_{n \in \mathbb{N}}$ is nondecreasing we obtain $\lim_{n \to \infty} c_n = \infty$ and $\limsup_{n \to \infty} P(Y_{1:n}^{z_0} \leq c_1, \ldots, Y_{n-k:n}^{z_0} \leq c_{n-k}) = g_k(c) > 0$ from Theorem 2.1. Select a subsequence $(n_j)_{j \in \mathbb{N}}$ of $\mathbb{N}$ such that $\lim_{j \to \infty} P(Y_{1:n_j}^{z_0} \leq c_1, \ldots, Y_{n_j-k:n_j}^{z_0} \leq c_{n_j-k}) = g_k(c)$. Hence, due to assumption (3.2) we obtain together with $\lim_{n \to \infty} c_n = \infty$ and the definition $D = \{z \in \mathscr{D}: \overline{G}(x|z_0) \geq \overline{G}(x|z)$ for eventually all large $x \in \mathbb{R}\}$ by using reasoning similar to that in Finner and Roters [(1994), pages 345–346],

$$0 = \lim_{j \to \infty} P\big(Y_{1:n_j} \leq c_1, \ldots, Y_{n_j-k:n_j} \leq c_{n_j-k}\big)$$

$$\geq \int_D \liminf_{j \to \infty} P\big(Y_{1:n_j}^z \leq c_1, \ldots, Y_{n_j-k:n_j}^z \leq c_{n_j-k}\big) \, dP^Z(z)$$

$$\geq \int_D \lim_{j \to \infty} P\big(Y_{1:n_j}^{z_0} \leq c_1, \ldots, Y_{n_j-k:n_j}^{z_0} \leq c_{n_j-k}\big) \, dP^Z(z)$$

$$= P(Z \in D)g_k(c) > 0,$$

a contradiction. So $\lim_{n \to \infty} n\overline{G}(c_n|z_0) = \infty$ which entails by the same argumentation as at the beginning of the proof of part (b)(i) that

$$\lim_{n \to \infty} n\overline{G}(c_n|z) = \infty \quad [P^Z].$$

So Theorem 3.1 completes the proof. □

PROOF OF THEOREM 3.3. Let $B_{z_0} = \{z \in \mathscr{D} : t(z|z_0) = 0\}$, $z_0 \in \mathscr{D}$, and define $f(z) = P(Z \in B_z)$ for all $z \in \mathscr{D}$. Then the assumptions of the theorem imply that

(A.2) $f(Z)$ has a uniform distribution on the unit interval $[0, 1]$.

The rather technical proof of (A.2) will be given at the end of the main proof.

We now show the following facts: Either (i) the $(n-k)$-limit property (1.1) or (ii) the up-to-$(n-k)$-limit property (1.2) together with the assumption that $(c_n)_{n \in \mathbb{N}}$ is nondecreasing implies

(A.3) $\lim_{n \to \infty} n\overline{G}(c_n|z) = \begin{cases} \infty, & \text{if } f(z) > \beta, \\ 0, & \text{if } f(z) < \beta, \end{cases} \quad z \in \mathscr{D}.$

Hence $P(Z \in A \cap B) = 1$, and the assertion follows from Theorem 3.1.

To prove that (ii) implies (A.3), first choose $z_1 \in \mathscr{D}$ such that $f(z_1) > \beta$, $\beta \in [0, 1)$. Then there exists a $c \in [0, \infty]$ such that

$$\limsup_{n \to \infty} P\big(Y_{1:n}^{z_1} \leq c_1, \ldots, Y_{n-k:n}^{z_1} \leq c_{n-k}\big) = g_k(c).$$

Theorem 2.1 and the monotonicity of $(c_n)_{n \in \mathbb{N}}$ yield $\liminf_{n \to \infty} n\overline{G}(c_n|z_1) = c$. In the following we prove that $c = \infty$. To this end, suppose that $c < \infty$. Then by virtue of the properties of the t-function we have

$$\liminf_{n \to \infty} n\overline{G}(c_n|z) = 0 \quad \text{for all } z \in B_{z_1}.$$

Due to the general assumptions of the theorem and (A.2) we may choose a $z_2 \in \mathscr{Q}$ such that $\beta < f(z_2) < f(z_1)$. Then the comparability of the functions $t(\cdot|z)$, $z \in \mathscr{Q}$, yields

$$t(\cdot|z_1) \leq t(\cdot|z_2) \; [P^Z] \quad \text{and} \quad P^Z\big(t(\cdot|z_1) < t(\cdot|z_2)\big) > 0.$$

So there exists a $z_3 \in \mathscr{Q}$ such that $t(z_3|z_1) = 0$, $t(z_3|z_2) = \infty$, which implies $t(z_2|z_1) = 0$, that is, $z_2 \in B_{z_1}$. Since, by Theorem 2.1, $\limsup_{n \to \infty} P(Y^{z_2}_{1:n} \leq c_1, \ldots, Y^{z_2}_{n-k:n} \leq c_{n-k}) = 1$ we may choose a subsequence $(n_j)_{j \in \mathbb{N}}$ of $\mathbb{N}$ such that

$$\lim_{j \to \infty} P(Y^{z_2}_{1:n_j} \leq c_1, \ldots, Y^{z_2}_{n_j-k:n_j} \leq c_{n_j-k}) = 1.$$

This, together with the monotonicity of the sequence $(c_n)_{n \in \mathbb{N}}$, implies that $\lim_{n \to \infty} c_n = \infty$. Thus, it follows for all $z \in B_{z_2}$ that $G(c_n|z_2) \leq G(c_n|z)$ for eventually all $n \in \mathbb{N}$, so that a reasoning similar to that in Finner and Roters [(1994), pages 345–346] (cf. the proof of Theorem 3.2), yields

$$\lim_{j \to \infty} P(Y^z_{1:n_j} \leq c_1, \ldots, Y^z_{n_j-k:n_j} \leq c_{n_j-k}) = 1 \quad \text{for all } z \in B_{z_2}.$$

Finally,

$$\begin{aligned}
\beta &= \lim_{j \to \infty} P(Y_{1:n_j} \leq c_1, \ldots, Y_{n_j-k:n_j} \leq c_{n_j-k}) \\
&\geq \int_{B_{z_2}} \lim_{j \to \infty} P\big(Y^z_{1:n_j} \leq c_1, \ldots, Y^z_{n_j-k:n_j} \leq c_{n_j-k}\big) \, dP^Z(z) \\
&= f(z_2) \\
&> \beta
\end{aligned}$$

yields a contradiction; hence, $c = \infty$.

For $z_1 \in \mathscr{Q}$ with $f(z_1) < \beta$, $\beta \in (0, 1]$, all the $\liminf$'s and $\limsup$'s have to be exchanged, and c has to be shown to be equal to 0. Since the argumentation is similar to that in the case just treated, a detailed proof of the remaining part is omitted. This concludes the proof of (A.3) under condition (ii).

If condition (i) is in force obvious modifications of the aforementioned arguments yield the validity of (A.3). $\square$

PROOF OF (A.2). Remember that $f(z) = P(Z \in B_z)$, $z \in \mathscr{Q}$. Then f is measurable since $\overline{G}(\cdot|\cdot)$ is measurable in $(x, z) \in \mathbb{R} \times \mathscr{Q}$. For the sake of notational simplicity we introduce the following abbreviations. For $z, z_0 \in \mathscr{Q}$ put $z \sim z_0$ if $t(\cdot|z) = t(\cdot|z_0) \; [P^Z]$, $z \preceq z_0$ if $t(\cdot|z) \leq t(\cdot|z_0) \; [P^Z]$, and $z \prec z_0$ if

$z \preceq z_0$ and $z \not\succ z_0$. Recalling that $B_{z_0} = \{z \in \mathscr{Z}: t(z|z_0) = 0\} \in \mathscr{C}$, $z_0 \in \mathscr{Z}$, we obtain the following facts. Let $z_0, z_1 \in \mathscr{Z}$. Then

(A.4) $\quad \{z \in \mathscr{Z}: \ z_0 \prec z\} \subseteq B_{z_0} \subseteq \{z \in \mathscr{Z}: \ z_0 \preceq z\},$

(A.5) $\qquad\qquad z_0 \preceq z_1 \quad \text{iff } f(z_1) \leq f(z_0),$

(A.6) $\qquad\quad f(z_1) \leq f(z_0) \quad \text{iff } B_{z_1} \subseteq B_{z_0} \quad [P^Z],$
$\qquad\qquad\qquad \text{i.e., } P\big(Z \in B_{z_1} \cap B_{z_0}^c\big) = 0,$

(A.7) $\qquad\qquad D_{z_0} = \{z \in \mathscr{Z}: z \sim z_0\} = \{z \in \mathscr{Z}: f(z) = f(z_0)\} \in \mathscr{C},$

(A.8) $\qquad\qquad f(z_0) = P\big(f(Z) \leq f(z_0)\big).$

It should be mentioned that these facts are valid due to the assumption $P(t(Z|z_0) \in \{0, \infty\}) = 1$, $z_0 \in \mathscr{Z}$, which implies that each $z_0 \in \mathscr{Z}$ is contained in a P^Z-null set because $t(z_0|z_0) = 1$. Since the proofs of these facts only use easy arguments concerning the shape of and the interrelation between the functions $t(\cdot|z_0)$, $z_0 \in \mathscr{Z}$, they are omitted and left to the reader.

All that really needs some deeper investigation is the uniform distributedness of $f(Z)$. To this end we observe that (A.8) means that on the image $f(\mathscr{Z}) \subseteq [0, 1]$ the cdf H (say) of $f(Z)$ is equal to that of the uniform distribution on the unit interval $[0, 1]$. Hence, it remains to be shown that $H(x) = x$ on $(0, 1) \setminus f(\mathscr{Z})$.

Since, by assumption, $P(f(Z) = f(z)) = P(Z \in D_z) = 0$ for all $z \in \mathscr{Z}$, H is continuous. Now let $x \in (0, 1) \setminus f(\mathscr{Z})$ and consider the sets $E_x = \{y \in f(\mathscr{Z}): y < x\}$ and $F_x = \{y \in f(\mathscr{Z}): y > x\}$. For all $y \in E_x$ we have $H(y) = P(f(Z) \leq y) = y < x$. Hence, $H(\underline{x}) = \underline{x} \leq x$, where $\underline{x} = \sup_{y \in E_x} y$ $[\sup \varnothing = 0$ (say)], and similarly $H(\overline{x}) = \overline{x} \geq x$, where $\overline{x} = \inf_{y \in F_x} y$ $[\inf \varnothing = 1$ (say)]. Now, by virtue of the definitions of E_x and F_x, we have $(\underline{x}, \overline{x}) \cap f(\mathscr{Z}) = \varnothing$. Hence, $\underline{x} = H(\underline{x}) = H(\overline{x}-) = H(\overline{x}) = \overline{x}$, which implies that $\underline{x} = x = \overline{x}$ and consequently $H(x) = x$ for all $x \in [0, 1]$.

REFERENCES

DALAL, S. R. and MALLOWS, C. L. (1992). Buying with exact confidence. *Ann. Appl. Probab.* **2** 752–765.

DUNNETT, C. W. and TAMHANE, A. C. (1992). A step-up multiple test procedure. *J. Amer. Statist. Assoc.* **87** 162–170.

FINNER, H. and GIANI, G. (1996). Duality between multiple testing and selecting. *J. Statist. Plann. Inference* **54** 201–227.

FINNER, H., HAYTER, A. J. and ROTERS, M. (1993). On the joint distribution function of order statistics with reference to step-up multiple test procedures. Technical Report 93-19, Univ. Trier.

FINNER, H. and ROTERS, M. (1994). On the limit behaviour of the joint distribution function of order statistics. *Ann. Inst. Statist. Math.* **46** 343–349.

FINNER, H. and ROTERS, M. (1997). Asymptotic comparison of the critical values of step-down and step-up multiple comparison procedures. To appear in *J. Statist. Plann. Inference.*

HOCHBERG, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* **75** 800–802.

LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

MARCUS, R., PERITZ, E. and GABRIEL, K. R. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika* **63** 655–660.

ROM, D. M. (1990). A sequentially rejective test procedure based on a modified Bonferroni inequality. *Biometrika* **77** 663–665.

SARKAR, S. K. and CHANG, C.-K. (1997). The Simes method for multiple hypothesis testing with positively dependent test statistics. *J. Amer. Statist. Assoc.* **92** 1601–1608.

TONG, Y. L. (1990). *The Multivariate Normal Distribution*. Springer, New York.

WELSCH, R. E. (1977). Stepwise multiple comparison procedures. *J. Amer. Statist. Assoc.* **72** 566–575.

UNIVERSITÄT TRIER
FB IV MATHEMATIK/STATISTIK
54286 TRIER
GERMANY
E-MAIL: finner@uni-trier.de
       roters@uni-trier.de