ON TWO RESULTS IN MULTIPLE TESTING

By Sanat K. Sarkar[1], Pranab K. Sen and Helmut Finner

Temple University, University of North Carolina at Chapel Hill and University of Duesseldorf

Two known results in multiple testing, one relating to the directional error control of augmented step-down procedure proved by Shaffer (1980) and the other on the monotonicity of the critical values of step-up procedure proved by Dalal and Mallows (1992), are revisited and given alternative proofs in this article.

1. Introduction. Testing of a null hypothesis against two-sided alternative is typically considered as a problem of making one of two kinds of decision, acceptance or rejection of the null hypothesis, and is designed in such a way that the Type I error rate associated with false rejection of the null hypothesis is controlled at a specified value. Once the null hypothesis is rejected, the direction of the alternative hypothesis is decided based on the value of the test statistic. However, a directional error or Type III error might occur in making such directional decisions. For instance, in testing $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$, with θ being the parameter of a random variable T and θ_0 being some known value, a rejection region of the form $T \leq a$ or $\geq b$ is used, where a and b are determined subject to a specified control of the Type I error rate, i.e., the probability of falsely rejecting H_0. Once H_0 is rejected, the decision regarding $\theta > \theta_0$ or $\theta < \theta_0$ is made by checking if $T \geq b$ or $T \leq a$. A Type III error occurs when, for example, $\theta < \theta_0$ (or $\theta > \theta_0$) is the true situation, but we falsely decide for $\theta > \theta_0$ (or $\theta < \theta_0$) after rejection of H_0. It is interesting to see, however, that in almost all testing situations where T stochastically increases with θ, controlling the Type I error rate will ensure the same control for the Type III error rate. This is because, when $\theta = \theta_0$, there is no Type III error. On the other hand, when $\theta < \theta_0$, the chance of Type III error, which is

[1]The research is supported by NSA Grant MDA-904-01-10082 and NSF Grant DMS-0306366

AMS 2000 subject classificatios. Primary 62F03, secondary 62G30.

Key Words and Phrases. Type III errors, augmented step-down procedure, TP_2 distributions, step-up-step-down procedure, monotonicity of critical values.

$$
\begin{aligned}
& P_\theta\{T \geq b\} \\
\leq \; & P_{\theta_0}\{T \geq b\} \\
\leq \; & P_{\theta_0}\{T \leq a \text{ or } T \geq b\},
\end{aligned}
$$

the chance of Type I error. Similarly, the chance of Type III error is less than that of Type I error, for any $\theta > \theta_0$. In other words, in testing a null hypothesis against two-sided alternative, directional decisions can be made following rejection of the null hypothesis without committing any additional error. Does this phenomenon hold when multiple null hypotheses are tested simultaneously against two-sided alternatives? This was first addressed by Shaffer (1980). She proved that with Holm's (1979) step-down procedure involving independent test statistics controlling the familywise error rate (FWER) at α, directional decisions can be made for the alternatives corresponding to the rejected null hypotheses without causing the probability of at least one of Type I or Type III errors to exceed α. Her proof, of course, relied on certain sufficient conditions related to the probability distributions of the statistics. We revisit this particular result in the present article and provide an alternative proof requiring the more common monotone likelihood ratio property of the underlying densities.

The other main result of this article concerns existence of increasing set of critical values in an FWER-controlling step-up-step-down procedure. Consider testing n null hypotheses $H_1, \ldots, H_n$ simultaneously against the corresponding one-sided alternatives $\bar{H}_1, \ldots, \bar{H}_n$ using right tailed tests based on the test statistics $X_1, \ldots, X_n$, respectively, which are identically distributed under the null hypotheses. Let $X_{1:n} \leq \cdots \leq X_{n:n}$ be the ordered versions of these statistics, and $H_{1:n}, \ldots, H_{n:n}$ be the corresponding ordering of the null hypotheses. Then, a step-up-step-down procedure of order r based on $(X_1, \ldots, X_n)$ and the critical values $c^r_{1:n} \leq \cdots \leq c^r_{n:n}$ accepts $H_{1:n}, \ldots, H_{j:n}$ and rejects the rest if $(X_1, \ldots, X_n) \in A^r_{j,n}$, where

$$
A^r_{j,n} = \begin{cases} \{X_{j:n} < c^r_{j:n}, X_{j+1:n} \geq c^r_{j+1:n}, \ldots, X_{r:n} \geq c^r_{r:n}\} & \text{for } j = 0, 1, \ldots, r-1, \\ \{X_{r:n} < c^r_{r:n}, \ldots, X_{j:n} < c^r_{j:n}, X_{j+1:n} \geq c^r_{j+1:n}\} & \text{for } j = r, \ldots, n, \end{cases}
$$

with $A^r_{0,n} = \{X_{1:n} \geq c^r_{1:n}, \ldots, X_{r:n} \geq c^r_{r:n}\}$ and $A^r_{n,n} = \{X_{r,n} < c^r_{r:n}, \ldots, X_{n:n} < c^r_{n:n}\}$. It reduces to a step-up procedure when $r = 1$, and to a step-down procedure when $r = n$

(Sarkar, 2002a, b, 2004; Tamhane, Liu and Dunnett, 1998). The $c^r_{j:n}$'s providing a control of the FWER at α are determined from the following set of conditions

$$\begin{aligned} \min_{I_j} P\{X_{j:I_j} \le c^r_{j:n}\} &\ \ge 1-\alpha, \text{ for } j=1,\ldots,r, \\ \min_{I_j} P\{X_{r:I_j} \le c^r_{r:n},\ldots,X_{j:I_j} \le c^r_{j:n}\} &\ \ge 1-\alpha, \text{ for } j=r+1,\ldots,n, \end{aligned} \tag{1.1}$$

where I_j is an ordered subset of $\{1,\ldots,n\}$ with cardinality j, $j=1,\ldots,n$, and $X_{1:I_j} \le \cdots \le X_{j:I_j}$ are the ordered components of the subset $\{X_i : i \in I_j\}$. The probabilities are determined assuming null distributions of the underlying test statistics. Note that, for $I_n = \{1,\ldots,n\}$, we are using n, instead of I_n, in the subscripts of the notations for the ordered components. The critical values satisfying (1.1) with the equalities, of course, will provide the least conservative procedure.

Although it is required that the critical values of a step-up-step-down procedure be increasing, the existence of such critical values in any distributional setting is not always immediate, especially when they are determined to yield the least conservative procedure (Finner and Roters, 1998; Sarkar, 2000). For instance, with $1 \le r \le n-1$, it is not obvious that there exist increasing critical values satisfying (1.1) with the equalities. On the other hand, it is not difficult to see that, when $r=n$, the critical values of the least conservative step-down procedure are indeed increasing, as they are the $100(1-\alpha)\%$ points of a stochastically increasing sequence of distributions. The problem of verifying the increasing property of the critical values satisfying (1.1) with the equalities for $1 \le r \le n-1$ is actually complicated by the intricate relationship that exists between probability distributions of the ordered components of two successively increasing subsets of the X_i's. The problem has been solved by Dalal and Mallows (1992) in the situation where $r=1$ and the X_i's are iid . The second main objective of this paper is to extend this result to a general r, of course still assuming that the X_i's are iid. Bai and Kwong (2002) considered a more general version of Dalal-Mallows' result. However, their proof of this version seems to be incorrect (Finner and Roters, 2003, private communication).

The two main results of this article are described in Section 2 and proved in Section 3. The first main result (Result 1) relates to the directional errors control in a step-down procedure and the other main result (Result 2) is on the monotonicity of the critical values of a step-up-step-down procedure. Our proofs in Section 3 require some supporting results which will be proved in the Appendix.

2. The main results. The two main results of this paper are stated in this section and will be proved in the next section.

2.1. *An improvement of Shaffer's result.* Suppose that random variables $X_1, \ldots, X_n$ are independently distributed, with X_i having a probability density $f_{\theta_i}(x)$, $i = 1, \ldots, n$. Assume that, for each i, $f_{\theta_i}(x)$ is TP_2 in (x, θ_i) (Karlin, 1968); i.e., has the monotone likelihood ratio property in x (Lehmann, 1986). As mentioned in the introduction, in testing a single null hypothesis, say $H_1 : \theta_1 = \theta_{10}$, against the corresponding two-sided alternative $\bar{H}_1 : \theta_1 \neq \theta_{10}$, a level α two-tailed test in terms of X_1 will control the Type III error rate at α if rejection of H_1 is concluded by deciding $\theta_1 > \theta_{10}$ or $< \theta_{10}$ according as X_1 is large or small.

Consider now n null hypotheses $H_i : \theta_i = \theta_{i0}$, $i = 1, \ldots, n$, which are to be tested simultaneously against the corresponding two-sided alternatives $\bar{H}_i : \theta_i \neq \theta_{i0}$, $i = 1, \ldots, n$. As described in Shaffer (1980), Holm's (1979a) step-down procedure controlling the FWER at α can be augmented to make directional decisions regarding the alternatives corresponding to rejected null hypotheses as follows. Determine constants a_{ij}, b_{ij}, $i, j = 1, \ldots, n$, such that under $H_J = \cap_{i \in J} H_i$

$$P_{H_J}\{a_{i|J|} < X_i \leq b_{i|J|}, i \in J\} \geq 1 - \alpha, \text{ for all } J \subseteq \{1, \ldots, n\}. \tag{2.2}$$

For example, one can choose a_{ij} (or b_{ij}) to be the maximum (or minimum) of those values for which

$$P_{H_i}\{X_i < a_{ij}\} \leq (1 - \beta_i)\{1 - (1 - \alpha)^{\frac{1}{j}}\} \text{ (or } P_{H_i}\{X_i > b_{ij}\} \leq \beta_i\{1 - (1 - \alpha)^{\frac{1}{j}}\}),$$

for some $0 \leq \beta_i \leq 1/2$. Note that, for every fixed i, $a_{in} \leq \cdots \leq a_{i1} < b_{i1} \leq \cdots \leq b_{in}$. Define

$$B_J = \{a_{i|J|} < X_i \leq b_{i|J|}, i \in J\}. \tag{2.3}$$

Then, the augmented version of Holm's step-down procedure consists of the following steps.

STEP 1. Start with $J = J_n \equiv \{1, \ldots, n\}$. If $(X_1, \ldots, X_n) \in B_{J_n}$, stop by accepting all the hypotheses. Otherwise, reject the subset of null hypotheses $\{H_i : X_i \in B^c_{J_n}\}$ and go to the next step.

STEP j ($j \geq 2$). Let K_j^c be the subset of those indices i for which H_i is rejected in one of the previous $j-1$ stages. If $(X_1, \ldots, X_n) \in B_{K_j}$, stop by accepting the set of null hypotheses $\{H_i : X_i \in B_{K_j}\}$. Otherwise, reject the set of null hypotheses $\{H_i : X_i \in B_{K_j}^c\}$ and go to the next step.

Continue this way until each null hypothesis is either accepted or rejected. Decision regarding the direction of the alternative to a rejected null hypothesis is made based on the value of the corresponding test statistic; i.e., upon rejection of H_i, decide $\theta_i > \theta_{i0}$ or $< \theta_{i0}$ according as X_i is large or small. To be more specific, let us suppose that, for some $J \subseteq \{1, \ldots .n\}$, the above procedure results in rejection of the subset of null hypotheses $\{H_i : i \in J\}$ and acceptance of the rest. Then, regarding the directions of the alternatives corresponding to the rejected set of hypotheses, $\{H_i : i \in J^c\}$, one can decide $\theta_i > \theta_{i0}$ or $< \theta_{i0}$, for every $i \in J^c$, according as $X_i > b_{i|J|}$ or $< a_{i|J|}$.

RESULT 2.1. *For the above procedure,*

$$Pr\{\text{no Type I and Type III errors}\} \geq 1 - \alpha. \tag{2.4}$$

REMARK 2.1. It is interesting to note that the above result, originally proved by Shaffer (1980), does actually hold only under the TP_2 condition of the density of each X_i. This is a natural multiple testing analog of the corresponding result known for testing a single hypothesis. In Shaffer's (1980) proof, although a slightly less restrictive condition than the TP_2 condition has been assumed, i.e., the cdf, $F_{\theta_i}(x)$, of X_i is non-increasing in θ_i, some additional assumptions regarding $F_{\theta_i}(x)$ have also been made. These are: (i) $\lim_{\theta_i \to \underline{\theta}_i} F_{\theta_i}(x) = 1$, and $\lim_{\theta_i \to \overline{\theta}_i} F_{\theta_i}(x) = 0$, for every x in the support of $F_{\theta_{i0}}(x)$, where $[\underline{\theta}_i, \overline{\theta}_i]$ is the interval of possible values of θ_i, and (ii) $\partial[1 - F_{\theta_i}(x)]/\partial\theta_i$ is TP_2 in (x, θ_i). Location families of distributions with TP_2 densities, scale families of distributions of positive-valued random variable with TP_2 densities and exponential families of distributions satisfy the conditions assumed by Shaffer (1980); see also Finner (1999). However, the exponential families of distributions considered in Shaffer (1980) have TP_2 densities, and, as our proof in the next section suggests, once the TP_2 condition is known for all of these families of distributions, the other two conditions are redundant. Shaffer (1980) used an example involving Cauchy distribution to bring home the point that the above derivative condition is unavoidable; without this the result does not hold.

In fact, it is not surprising that this condition does not hold for the Cauchy distribution as it is not TP_2. The TP_2 condition actually appears to be unavoidable in this result.

2.2. *An extension of Dalal-Mallows' result.* Consider simultaneous testing of null hypotheses $H_1, \ldots, H_n$ using right-tailed tests based on the corresponding test statistics $X_1, \ldots, X_n$ that are assumed to be iid with the common cdf F. The least conservative generalized step-up-step-down procedure of order r controlling the FWER at $\alpha \in (0,1)$ based on the X_i's requires existence of critical values $c_1 \leq \cdots \leq c_n$ satisfying the following conditions:

$$\begin{aligned} P\{X_{j:j} \leq c_j\} &= 1-\alpha, \text{ for } j = 1, \ldots, r, \\ P\{X_{r:j} \leq c_r, \ldots, X_{j:j} \leq c_j\} &= 1-\alpha, \text{ for } j = r+1, \ldots, n. \end{aligned} \tag{2.5}$$

While it is clear that the critical values obtained from the first r equations in (2.5) are increasing, as they are the solutions to the following equations

$$F(c_j) = (1-\alpha)^{\frac{1}{j}}, \ j = 1, \ldots, r, \tag{2.6}$$

it is relatively less obvious, however, that there exist solutions to the remaining $n-r$ equations that will continue to be increasing. Since, for any given $c_1 \leq \cdots \leq c_j$,

$$\begin{aligned} & P\{X_{1:j} \leq c_1, \ldots, X_{j:j} \leq c_j\} \\ = \ & F(c_1) \sum_{i=0}^{j-1} \bar{F}^i(c_1) - \sum_{i=1}^{j-1} \begin{pmatrix} j \\ i \end{pmatrix} P\{X_{1:i} \leq c_1, \ldots, X_{i:i} \leq c_i\} \bar{F}^{j-i}(c_{i+1}), \end{aligned} \tag{2.7}$$

where $\bar{F}(\cdot) = 1 - F(\cdot)$, the existence of $c_1 \leq \cdots \leq c_k$ satisfying (2.5) for $j = 1, \ldots, k$, for some $r \leq k \leq n-1$, would imply the existence of $c_{k+1} \geq c_k$ satisfying (2.5) for $j = k+1$ provided we can show that c_{k+1} obtained from the following:

$$\bar{F}(c_{k+1}) = \frac{1}{k+1} \Big[\sum_{i=1}^{k} \bar{F}^i(c_1) - \sum_{i=1}^{k-1} \begin{pmatrix} k+1 \\ i \end{pmatrix} \bar{F}^{k-i+1}(c_{i+1}) \Big], \text{ with } c_1 = \cdots = c_r, \tag{2.8}$$

is greater than or equal to c_k, which would ultimately prove the desired monotonicity property of all the critical values satisfying (2.5). But, this is the main hurdle in this problem. When $r = 1$, Dalal and Mallows (1992) proved the existence of an increasing sequence of c_j's satisfying (2.5). We extend this result by proving it for a general r, of course using a completely different line of arguments.

Result 2.2. *There exists an increasing sequence of critical values* $c_1, \dots, c_n$ *satisfying* (2.5).

Remark 2.2. Bai and Kwong (2002) considered the following conjecture. There exist $c_1 \le \cdots \le c_n$ satisfying the following conditions:

$$P\{X_{j+1:m+k} \le c_1, \dots, X_{j+k:m+k} \le c_k\} = 1 - \alpha, \text{ for } k = 1, \dots, n, \tag{2.9}$$

for any fixed $0 \le j \le m$. This is a more general version of Dalal-Mallows result than what we consider here. However, as mentioned before, the proof given by Bai and Kwong (2002) seems to be incorrect.

3. Proofs of the main results

3.1. *Proof of Result 2.1.* Let us assume without any loss of generality that $\theta_i = \theta_{i0}$ for $i = 1, \dots, k$, and $> \theta_{i0}$ for $i = k+1, \dots, n$. Then, neither a Type I nor a Type III error occurs if and only if, for some J such that $\{1, \dots, k\} \subseteq J \subseteq \{1, \dots, n\}$, H_i is accepted for all $i \in J$, and rejected because of X_i being large for all $i \in J^c$. Thus, with $J_1 = \{1, \dots, k\}$ and $\theta = (\theta_{10}, \dots, \theta_{k0}, \theta_{k+1}, \dots, \theta_n)$, we have

$$\begin{aligned} &P_\theta\{\text{no Type I and Type III errors}\} \\ &= \sum_{j=k}^{n} \sum_{J:|J|=j, J \supseteq J_1} P_\theta\{a_{ij} < X_i \le b_{ij}, i \in J;\ X_i > b_{il_i}, i \in J^c, \text{ for some permutation} \\ &\qquad\qquad (l_{j+1}, \dots, l_n) \text{ of } (j+1, \dots, n)\} \end{aligned} \tag{3.1}$$

Now, use the following lemma related to TP_2 property, whose proof will be provided in the Appendix:

Lemma 3.1. *Let* $Y \sim f_\theta(y)$, *which is* TP_2 *in* (y, θ). *Then, for any fixed* $a < b$, *and* $\theta \ge \theta_0$,

$$P_\theta\{a \le Y \le b\} \ge P_{\theta_0}\{a \le Y \le b\} P_\theta\{Y \le b\}. \tag{3.2}$$

From the lemma we note that the probability in (3.1) is greater than or equal to

$$
\begin{aligned}
& \sum_{j=k}^{n} \sum_{J:|J|=j, J\supseteq J_1} P_{(\theta_{i0}:i\in J)}\{a_{ij} < X_i \le b_{ij}, i \in J\} \times \\
& \qquad P_{\theta_{k+1},\ldots,\theta_n}\{X_i \le b_{ij}, i \in J - J_1; X_i > b_{il_i}, i \in J^c, \text{ for some permutation} \\
& \qquad\quad (l_{j+1},\ldots,l_n) \text{ of } (j+1,\ldots,n)\} \\
\ge\ & (1-\alpha)\sum_{j=k}^{n} \sum_{J:|J|=j, J\supseteq J_1} P_{\theta_{k+1},\ldots,\theta_n}\{X_i < b_{ij}, i \in J - J_1; X_i > b_{il_i}, i \in J^c, \\
& \qquad \text{for some permutation } (l_{j+1},\ldots,l_n) \text{ of } (j+1,\ldots,n)\} \\
=\ & 1-\alpha, \qquad (3.3)
\end{aligned}
$$

as

$$
\begin{aligned}
& \sum_{j=k}^{n} \sum_{J:|J|=j, J\supseteq J_1} P_{\theta_{k+1},\ldots,\theta_n}\{X_i \le b_{ij}, i \in J - J_1; X_i > b_{il_i}, i \in J^c, \\
& \qquad \text{for some permutation} (l_{j+1},\ldots,l_n) \text{ of } (j+1,\ldots,n)\} \\
=\ & 1. \qquad (3.4)
\end{aligned}
$$

A proof of (3.4) is given in the Appendix. This proves Result 2.1. □

3.2. *Proof of Result 2.2.* Replacing X_i by $U_i = F(X_i)$, which is a $U(0,1)$ random variable, the result can be restated as that of proving the existence of constants $\alpha_1 \le \cdots \le \alpha_n$ satisfying the following conditions:

$$
\begin{aligned}
P\{U_{j:j} \le \alpha_j\} &= \alpha_1, \text{ for } j = 1,\ldots,r, \\
P\{U_{r:j} \le \alpha_r, \ldots, U_{j:j} \le \alpha_j\} &= \alpha_1, \text{ for } j = r+1,\ldots,n.
\end{aligned} \qquad (3.5)
$$

As pointed out in Section 2.2, there exist critical values $\alpha_1,\ldots,\alpha_r$ satisfying the first r conditions in (3.5) that are increasing. The fact that the critical values satisfying the last $n-r$ equations continue to be increasing, i.e., $\alpha_r \le \alpha_{r+1} \le \cdots \le \alpha_n$, is proved in the following.

First, we prove the following lemma.

Lemma 3.2. *Let there exist $\alpha_r \le \cdots \le \alpha_j < 1$ satisfying (3.5) for all $j = r,\ldots,k$. where $r+1 \le k \le n-1$. Then, for an α_{k+1} satisfying (3.5) for $j = k+1$, we have $\alpha_{k+1} \ge \alpha_k$ if and only if*

$$\mathrm{Var}(U_{k-1:k-1}) \ge \mathrm{Var}(U_{k-1:k-1} | U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1}), \qquad (3.6)$$

PROOF. First note that

$$\begin{aligned}
& P\{U_{r:k} \le \alpha_r, \ldots, U_{k:k} \le \alpha_k\} \\
= \;& kE\{(\alpha_k - U_{k-1:k-1}) I(U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1}))\} \\
= \;& kE\{(\alpha_k - U_{k-1:k-1}) | I(U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1})\} \times \\
& \quad P\{U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1}\}. \qquad (3.7)
\end{aligned}$$

Since this is equal to $P\{U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1}\}$, we have

$$\alpha_k = \frac{1}{k} + E(U_{k-1:k-1} | U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1}). \qquad (3.8)$$

Also,

$$\begin{aligned}
& P\{U_{r:k+1} \le \alpha_r, \ldots, U_{k+1:k+1} \le \alpha_{k+1}\} \\
= \;& (k+1)E\{(\alpha_{k+1} - U_{k:k}) I(U_{r:k} \le \min(\alpha_r, \alpha_{k+1}) \ldots, U_{k:k} \le \min(\alpha_k, \alpha_{k+1}))\} \\
\le \;& (k+1)E\{(\alpha_{k+1} - U_{k:k}) I(U_{r:k} \le \alpha_r, \ldots, U_{k:k} \le \alpha_k)\} \\
= \;& (k+1)(\alpha_{k+1} - \alpha_k) P\{U_{r:k} \le \alpha_r, \ldots, U_{k:k} \le \alpha_k\} \\
& \quad + \frac{k(k+1)}{2} E\{(\alpha_k - U_{k-1:k-1})^2 I(U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1})\} \\
= \;& (k+1)(\alpha_{k+1} - \alpha_k) P\{U_{r:k} \le \alpha_r, \ldots, U_{k:k} \le \alpha_k\} \\
& \quad + \frac{k(k+1)}{2} [\mathrm{Var}\{(\alpha_k - U_{k-1:k-1}) | U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1}\} \\
& \quad + E^2\{(\alpha_k - U_{k-1:k-1}) | U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1}\}] \\
& \quad \times P\{U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1}\} \\
= \;& (k+1)(\alpha_{k+1} - \alpha_k) P\{U_{r:k} \le \alpha_r, \ldots, U_{k:k} \le \alpha_k\} \\
& \quad + \frac{k(k+1)}{2} [\mathrm{Var}\{(\alpha_k - U_{k-1:k-1} | U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1}\} \\
& \quad + \frac{1}{k^2}] P\{U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1}\}. \qquad (3.9)
\end{aligned}$$

The eqn. (3.8) has been used in the last equality in (3.9). Since $P\{U_{r:j} \le \alpha_r, \ldots, U_{j:j} \le c_j\}$ is the same for $j = k-1, k$ and $k+1$, we get

$$\begin{aligned}
\alpha_{k+1} - \alpha_k \;\ge\;& \frac{k}{2}\{(\frac{2}{k(k+1)} - \frac{1}{k^2}) \\
& \quad - \mathrm{Var}((\alpha_k - U_{k-1:k-1} | U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1})\} \\
=\;& \frac{k}{2}\{\mathrm{Var}(U_{k-1:k-1}) - \\
& \quad \mathrm{Var}(U_{k-1:k-1} | U_{r:k-1} \le \alpha_r, \ldots, U_{k-1:k-1} \le \alpha_{k-1})\}. \qquad (3.10)
\end{aligned}$$

Now, if (3.10) is greater than or equal to zero, we have $\alpha_{k+1} \geq \alpha_k$, which proves the 'if' part of the lemma. Conversely, if $\alpha_{k+1} \geq \alpha_k$, we will have equalities in (3.9) and (3.10), and hence (3.10) must be greater than or equal to zero. Thus, the lemma is proved. □

Next, we will prove the following lemma.

LEMMA 3.3. *Let* $0 < \alpha_r \leq \cdots \leq \alpha_j < 1$ *satisfy the condition (3.5) for all* $j = r, \ldots, k$, where $r \leq k \leq n$. *Then,*

$$\mathrm{Var}(U_{j:j}) \geq \mathrm{Var}(U_{j:j} | U_{r:j} \leq \alpha_r, \ldots, U_{j:j} \leq \alpha_j), \tag{3.11}$$

for all $j = r, \ldots, k$.

PROOF. A proof for $j = k$ will be enough. The conditional variance in this lemma for $j = k$ is the variance corresponding to the following distribution function

$$F_k(x) = \begin{cases} \alpha_1^{-1} P\{U_{r:k} \leq \alpha_r, \ldots, U_{k-1:k} \leq \alpha_{k-1}, U_{k:k} \leq x\} & \text{if } x \leq \alpha_k \\ 1 & \text{if } x > \alpha_k, \end{cases} \tag{3.12}$$

with the density given by

$$f_k(x) = k F_{k-1}(\min(x, \alpha_{k-1})) I(x \leq \alpha_k). \tag{3.13}$$

The lemma then follows from the result (Lemma A.1), proved in the Appendix, that the variance of this distribution is less than that of

$$G_k(x) = \begin{cases} 0 & \text{if } x < 0 \\ \min(x^k, 1) & \text{if } x \geq 0, \end{cases} \tag{3.14}$$

the distribution of $U_{k:k}$, with

$$g_k(x) = k x^{k-1} I(0 < x < 1). \tag{3.15}$$

being the corresponding density. □

From Lemmas 3.2 and 3.3, we see that, if there exist $\alpha_r \leq \cdots \leq \alpha_j < 1$ satisfying (3.5) for all $j = r, \ldots, k$, then there exist $\alpha_{k+1} \geq \alpha_k$ satisfying (3.5) for $j = k + 1$, where $r + 1 \leq k \leq n - 1$. This holds also for $k = r$, which is easy to check. Thus Result 2.2 holds by induction. □

4. Concluding remarks. The results in this article provide alternative proofs of two previously known results (Shaffer, 1980; Dalal and Mallows, 1992). We have given a much simpler proof of Shaffer's result based only on the TP_2 property of the underlying densities, and an alternative proof of Dalal and Mallows' result in a much more general context. Nevertheless, these proofs are still limited to the framework of independent test statistics. While it is believed that these results might hold for certain types of dependent test statistics, they still remain to be two of the most challenging problems in multiple testing. Some partial attempts, however, have been made to address these open problems, theoretically as well as empirically. For instance, Finner (1999) and Holm (1979b, 1981) extended Shaffer's result and Sarkar (2000) extended Dalal and Mallows' result, to some very special types of dependent test statistics. Also, extension of Dalal and Mallows' result to some other dependence situations have been empirically checked (Dunnett and Tamhane, 1992, 1995; Kwong and Liu, 2000; Liu, 1997; Tamhane, Liu and Dunnett, 1998). The method of Shaffer (1980) was adopted by Finner (1994) and Liu (1996) to prove directional error control for a step-up test with independent test statistics under the same distributional assumptions as those made by Shaffer. We conjecture that these assumptions can be relaxed and only the TP_2 condition will suffice. Finner (1999) generalized the method of proof under Shaffer's (1980) assumptions for a large class of procedures satisfying a unimodality property of acceptance regions, and gave a new but very simple and elegant proof under the assumption of TP_3 densities.

APPENDIX

Proofs

Proof of Lemma 3.1. Let $\phi(x, y) = 1$ if $x \geq y$, and $= 0$ if $x < y$. The function $\phi(x, y)$ is known to be TP_2 in (x, y) (see, for example, Karlin, 1968). The basic composition theorem of Karlin (1968) then implies that

$$P_\theta\{a \leq Y \leq b\} = \int [1 - \phi(y, b)]\phi(y, a) f_\theta(y) dy$$

is TP_2 in (a, θ) for fixed b. Therefore,

$$P_{\theta_0}\{-\infty \leq Y \leq b\} P_\theta\{a \leq Y \leq b\} \geq P_{\theta_0}\{a \leq Y \leq b\} P_\theta\{-\infty \leq Y \leq b\},$$

which yields the lemma. □

PROOF OF (3.4). Note that

$$\sum_{J:|J|=j, J\supseteq J_1} P\{X_i \le b_{ij}, i \in J - J_1; X_i > b_{il_i}, i \in J^c, \text{ for some permutation } (l_{j+1}, \ldots, l_n) \text{ of } (j+1, \ldots, n)\}$$

is the probability $P\{N = n - j\}$, where N represents the number of null hypotheses that are rejected when the $n - k$ null hypotheses in the set $\{H_i : i \in J_1^c\}$ are tested simultaneously against the corresponding right-sided alternatives using Holm's step-down procedure using the critical values b_{ij}, $i \in J_1^c$, $j = k+1, \ldots, n$. In terms of this N, the left-hand side of (3.4) is $\sum_{j=k}^{n} P\{N = n - j\}$, which is equal to 1. □

LEMMA A.1 *The variance of F_k in (3.12) is less than that of G_k in (3.14).*

PROOF. Given two distribution functions G and H, H is more dispersive than G, implying that H has larger variance than G, iff $H^{-1}(v) - H^{-1}(u) > G^{-1}(v) - G^{-1}(u)$, for any $0 \le u < v \le 1$. Let, for a function $\phi(x)$ defined on $A \subset \mathcal{R}$, $S^-(\phi)$ be the number of sign changes of ϕ as defined in Karlin (1968); that is,

$$S^-(\phi) = S^-[\phi(x)] = \sup S^-[\phi(x_1), \cdots, \phi(x_m)],$$

where $S^-(y_1, \cdots, y_m)$ is the number of sign changes of the indicated sequence, zero terms being discarded, and the supremum is taken over all sets $y_1, \cdots, y_m$, with $y_i \in A, m < \infty$. Shaked (1982) proved that when G and H are both continuous and strictly increasing on their supports $[0, \infty)$, a necessary and sufficient condition for H to be more dispersive than G is that, for every fixed $a > 0$, $S^-[G(x-a) - H(x)] \le 1$, with the sign sequence being −,+ in case of the equality, and, for every $x > 0$, $G(x) - H(x) \ge 0$. Furthermore, it follows from Karlin (1968), and also pointed out in Shaked (1982), that if g and h are the densities of G and H respectively, then the fact that $S^-[g(x-a) - h(x)] \le 2$, with the sign sequence being $-, +, -$ in case of the equality, implies that $S^-[G(x-a) - H(x)] \le 1$, with the sign sequence being $-, +$ in case of the equality. Using these results, we will show that the variance of F_k is less than that of G_k.

The required result is proved once we prove the following: (i) For any fixed $a > 0$, $S^-[F_k(x-a) - G_k(x)] \le 1$, with the sign sequence being $-$, $+$ in case of the equality, and (ii) $F_k(x) \geq G_k(x)$, for all $x > 0$. For any fixed $a > 0$,

$$f_k(x-a) - g_k(x) = \begin{cases} k[F_{k-1}(x-a) - G_{k-1}(x)] & \text{if } a < x \leq \alpha_{k-1} + a \\ k[1 - G_{k-1}(x)] & \text{if } \alpha_{k-1} + a < x \leq \alpha_k + a \\ -kG_{k-1}(x) & \text{if } \alpha_k + a < x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, $S^-[f_k(x-a) - g_k(x)] \leq 2$, with the sign sequence being $-$, $+$, $-$ in case of the equality, if $S^-[F_{k-1}(x-a) - G_{k-1}(x)] \leq 1$, with the sign sequence being $-$, $+$ in case of the equality. The result (i) then follows from induction because for $F_r(x)$, which is

$$F_r(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{\alpha_1} \min(x^r, \alpha_1) & \text{if } x \geq 0, \end{cases}$$

we see that $S^-[F_r(x-a) - G_r(x)] \leq 1$, with the sign sequence being $-$, $+$ in case of the equality. Result (ii) also follows from induction. To verify this, first note that $S^-[f_k(x) - g_k(x)] \leq 1$ with the sign sequence being $+$, $-$ when the equality holds, provided $F_{k-1}(x) - G_{k-1}(x) \geq 0$, for all $x > 0$. That is, F_k is stochastically smaller than G_k, implying that $F_k(x) - G_k(x) \geq 0$, for all $x > 0$, provided $F_{k-1}(x) - G_{k-1}(x) \geq 0$, for all $x > 0$. Thus the lemma is proved. □

REFERENCES

Bai, Z. D. and Kwong, K. S. (2002). Solution to Dalal and Mallows conjecture on monotone property of the joint distribution of order statistics, *Statist. & Prob. Lett.*, **59** 29-35.

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.

Dunnett, C. W. and Tamhane, A. C. (1993). Power comparisons of some step-up-multiple test procedures. *Statist. Probab. Lett.* **16** 55-58.

Finner, H. (1994). Testing multiple hypotheses: general theory, specific problems, and relationship to other multiple decision procedures. Habilitationsschrift, Fachbereich IV Mathematik, Univ. Trier.

FINNER, H. (1999). Stepwise multiple test procedures and control of directional errors. *Ann. Statist.* **27** 274-289.

FINNER, H. and ROTERS, M. (1998). Asymptotic comparisons of step-up and step-down multiple test procedures based on exchangeable test statistics. *Ann. Statist.* **26** 505-520.

HOLM, S. (1979a). A simple sequentially rejective multiple test procedure. *Scand. J. Statist.* **6** 65-70.

HOLM, S. (1979b). A stagewise directional test based on t statistics. Statistical Research Report 1979-3, Inst. Mathematics, Chalmers Univ. Technology, Gothenburg.

HOLM, S. (1981). A stagewise directional test for the normal regerssion situation. In *Proceedings of the Sixth Conference on Probability Theory* (B. Bereanu, S. Grigorescu, M. Josifescu and T. Postelnicu, eds.) 103-106. National Institute of Metrology, Bucharest.

KARLIN, S. (1968). *Total Positivity.* Stanford University Press, Stanford, California.

KWONG, K.S. and LIU, W. (2000). Calculation of critical values for Dunnett and Tamhane's step-up multiple test procedures. *Statist. Probab. Lett.* **49** 411-416.

LEHMANN, E. (1986). *Testing Statistical Hypotheses.* Wiley, New York.

LIU, W. (1996). Control of directional errors with step-up multiple tests. *Statist. Probab. Lett.* **31** 239-242.

LIU, W. (1997). Some results on step-up tests for comparing treatments with a control in unbalanced one-way layouts. *Biometrics* **53** 1508-1512.

SARKAR, S. K. (2000). A note on the monotonicity of the critical values of a step-up test. *J. Statist. Plann. Inf.* **87** 241 - 249.

SARKAR, S.K. (2002a). Some results on false discovery rate in multiple testing procedures. *Ann. Statist.* **30** 239-257.

SARKAR, S. K. (2002b). Recent advances in multiple testing. *J. Statist. Studies*, A special volume in honor of Prof. M.M. Ali's 65th birthday, 293-306.

SARKAR, S. K. (2004). FDR-controlling stepwise procedures and their false negatives rates, *J. Statist. Plann. Inf.*, To appear.

SHAKED, M. (1982). Dispersive ordering of distributions. *J. App. Prob.* **19** 310-320.

SHAFFER, J, P. (1980). Control of directional errors with stagewsie multiple test procedures. *Ann. Statist.* **8** 1342-1347.

Tamhane, A. C., Liu, W. and Dunnett, C. W. (1998). A generalized step-up-down multiple test procedure. *Canad. J. Statist.* **26** 353-363.