On dependence assumption in p-value based multiple test procedures (DRAFT)

Jiangtao Gou

May 5, 2020

Abstract. The Hochberg procedure and Benjamini-Hochberg procedure are widely used in confirmatory clinical studies and exploratory research for multiplicity adjustment. A common misconception is that these procedures control the type I error rate properly if the test statistics are independent or positively correlated. In fact, a much stronger positive dependence assumption needs to be satisfied to guarantee the type I error rate control. We give a comprehensive review of the dependence conditions used in multiple testing procedures. We show that a weaker positive dependence assumption may result an inflation of type I error rate by a factor of 2, and discuss the type I error rate control under certain negative dependence conditions.

Keywords. Benjamini-Hochberg procedure, Dependence assumption, Hochberg procedure, Multiple testing procedure, Simes test

1 Introduction

Multiple test procedures are often applied in the analysis of clinical trials when more than one null hypothesis has to be tested simultaneously. The multiplicity can be adjusted in various ways depending on the type of controlled error rate. For example, the familywise error rate (FWER) is the probability of rejecting one or more true null hypothesis under any combination of true and false hypotheses (Hochberg and Tamhane, 1987). The Hochberg (1988) procedure is an FWER controlling method for adjusting multiplicity when testing multiple primary endpoints in confirmatory clinical trials. Another error rate is the false discovery rate (FDR) that is the expected proportion of the rejected null hypotheses among all rejections. The most popular FDR controlling method is the Benjamini and Hochberg (1995) procedure. Both the Hochberg (1988) FWER controlling procedure and Benjamini and Hochberg (1995) FDR controlling procedure are based on the Simes (1986) test. The Simes-test-based methods are usually more powerful than the Bonferroni-test-based procedures. The Simes test is a valid α-level test if the Simes inequality holds, which is

$$\Pr\left(\cup_{i=1}^{n}\left\{P_{(i)} \leq i\alpha/n\right\}\right) \leq \alpha,$$

where $P_{(1)} \leq \cdots \leq P_{(n)}$ are ordered p-values from testing n null hypotheses. However, the Simes inequality does not always hold. The Simes inequality is true under independence and certain dependence structures, for example, positive dependence through stochastic ordering (PDS) (Sarkar and Chang, 1997; Sarkar, 1998; Benjamini and Yekutieli, 2001; Sarkar, 2002). Comparing with the positive correlation assumption, the PDS assumption is much stricter and more difficult to be understood for practitioners. As a common misconception about the precise conditions under which the Simes-based tests are valid, some researchers think the Simes inequality holds when test statistics are positively correlated. For the Hochberg procedure, the draft guideline on multiple endpoints in clinical trials by the US Food and Drug Administration (FDA) makes a clear statement: "*Hochberg procedure usually will, but is not guaranteed to, control the overall type I error rate for positively-correlated endpoints*" (FDA, 2017). A conservative suggestion for the Hochberg procedure is to apply it on "*two positively-correlated dependent tests with standard test statistics, such as the normal Z, student's t, and 1 degree of freedom chi-square*" (Huque, 2016; FDA, 2017).

For dependent test statistics, the Hochberg procedure, Benjamini-Hochberg procedure and other Simes-test-based methods (e.g., Hommel (1988) procedure) control the type I error rate under a positive dependence condition that is considerably stricter than the positive correlation condition (Tamhane and Gou, 2018). The PDS condition is a sufficient condition that guarantees the truth of the Simes inequality. However, verifying the PDS condition based on the observed data is not an easy task, and currently there is no standard statistical test to evaluate the PDS condition. This constraint may limit the application of the Simes-test-based methods, especially in confirmatory clinical trials. Considering the Bonferroni-based procedures require no dependence assumption, we wonder if it is possible to find a new multiple testing procedure which is less powerful than the Simes method but more powerful than the Bonferroni method, and the type I error rate control of this new procedure is guaranteed under a general dependence structure that can be verified using existing statistical tests, e.g., the t-test of correlation coefficient.

This article investigates the possibility of weakening the dependence assumption for the Simes test and some tests whose statistical powers are between the Simes and the Bonferroni methods, and is organized as follows. Section 2 provides a summary of five difference families of dependence structures of random variables, including the PDS, and discusses their relationships between each other. In Section 3, we first investigate the type I error rate control of the Simes test under a type of positive dependence that is less strict than the PDS, and find that the nominal level can be inflated significantly under the positive quadrant dependence (PQD). Next, we consider the multiple testing procedures which control the type I error rate under PQD, and find these procedures can only be slightly more powerful than the Bonferroni methods. Section 4 includes some discussion and directions for future work. The proofs of all propositions are included in Appendix.

2 Dependence properties

This section presents a summary of five families of dependence structures which are closely related to the Simes (1986) test, including the correlation dependence, quadrant dependence, association dependence, stochastic monotonicity dependence, and monotonic likelihood ratio dependence. In order to be precise when describing the monotonicity, we call a function ϕ increasing if $\phi(u) \geq \phi(v)$ whenever $u > v$. A function is called strictly increasing if $\phi(u) > \phi(v)$ whenever $u > v$. The corresponding concepts decreasing function and strictly decreasing function are defined by inverting the order.

(1) Correlation dependence PC/NC. The random vector $\mathbf{X} = (X_1, \ldots, X_n)$ is *positively correlated* (PC) if $\operatorname{cov}(X_i, X_j) \geq 0$ for all $i \neq j$ (Tong, 1990). Analogously, $\mathbf{X}$ is *negatively correlated* (NC) if $\operatorname{cov}(X_i, X_j) \leq 0$ for all $i \neq j$. The PC/NC dependence is defined based on the Pearson product-moment correlation coefficient r, which is the most commonly

used dependence measure. Correlation dependence can also be defined through other correlation coefficients such as Kendall (1938,1945)'s τ and Spearman (1904)'s ρ. Generally speaking, any correlation coefficient which satisfies Rényi (1959)'s axioms or their relaxations (Schweizer and Wolff, 1981) can be applied as a measure of dependence, including the supremum correlation (Gebelein, 1941; Kimeldorf and Sampson, 1978).

Note that even if $\mathbf{X}$ is PC based on Pearson's r, $\mathbf{X}$ is not necessarily PC based on other correlation coefficients. To show the nonequivalence among correlation dependence based on different types of correlation coefficients, it suffices to give a counterexample. Let bivariate discrete random vector $(X_1, X_2) \in \{0, 1, 10\} \times \{0, 1\}$ have the probability mass function $\Pr(X_1 = 0, X_2 = 0) = \Pr(X_1 = 0, X_2 = 1) = \Pr(X_1 = 1, X_2 = 1) = \Pr(X_1 = 10, X_2 = 1) = 0.1$, $\Pr(X_1 = 1, X_2 = 0) = 0.4$, and $\Pr(X_1 = 10, X_2 = 0) = 0.2$. Calculation shows that Pearson's r is $1/\sqrt{1533} \approx 0.026$, Kendall's τ is -0.04 (τ_a, without adjustment for ties) or $-2/\sqrt{651} \approx -0.078$ (τ_b, with adjustment for ties), and Spearman's ρ is $-\sqrt{21}/49 \approx -0.094$. Thus (X_1, X_2) is PC based on Pearson's r, but NC based on Kendall's τ or Spearman's ρ.

This example is observed since the sign of Pearson's r is not invariant under the positively monotonic transformation of random variables. In a special situation where the random vector $\mathbf{P}$ has the joint distribution function with margins Unif$(0, 1)$, Pearson's r and Spearman's ρ are the same, although the sign of Kendall's τ and the sign of Spearman's ρ still can be different (Daniels, 1950; Durbin and Stuart, 1951).

Correlation dependence is the least strict and the simplest requirement of positive/negative dependence, since this dependence is defined pairwise through a finite number of coefficients.

(2) Quadrant dependence PQD/POD/NQD/NOD. The random vector $\mathbf{X}$ is *positively orthant dependent* (POD) if

$$\Pr(\mathbf{X} > \mathbf{x}) \geq \prod_{i=1}^{n} \Pr(X_i > x_i) \tag{1}$$

and

$$\Pr(\mathbf{X} \leq \mathbf{x}) \geq \prod_{i=1}^{n} \Pr(X_i \leq x_i) \tag{2}$$

holds for all $\mathbf{x} = (x_1, \ldots, x_n)$, where $\{\mathbf{X} > \mathbf{x}\} = \cap_{i=1}^{n} \{X_i > x_i\}$ and $\{\mathbf{X} \leq \mathbf{x}\} = \cap_{i=1}^{n} \{X_i \leq x_i\}$ (Dykstra et al., 1973). If only (1) holds for all $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{X}$ is said to be *positively upper orthant dependent* (PUOD). Similarly, if only (2) holds for all $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{X}$ is said to be *positively lower orthant dependent* (PLOD) (Joe, 1997). Equivalently, $\mathbf{X}$ is said to be POD if

$$\mathbb{E}\left[\prod_{i=1}^{n} \phi_i(X_i)\right] \geq \prod_{i=1}^{n} \mathbb{E}[\phi_i(X_i)] \tag{3}$$

for any set of positive and comonotonic functions $\{\phi_1, \ldots, \phi_n\}$ (Belzunce and Semeraro, 2004). In two dimensions, PLOD and PUOD are equivalent, and we use quadrants instead of orthants: $\mathbf{X} = (X_1, X_2)$ is said to be *positively quadrant dependent* (PQD) if either (1) or (2) is satisfied (Lehmann, 1966). Analogously, *negative orthant dependence* (NOD) (Ebrahimi and Ghosh, 1981; Block et al., 1982) and *negative quadrant dependence* (NQD) are defined by reversing the direction of the inequalities in (1), (2) and (3).

Either PLOD or PUOD implies PC. In fact, if $\mathbf{X}$ is pairwise PQD, say, the inequality $\Pr(\{X_i \le x_i\} \cap \{X_j \le x_j\}) \ge \Pr(X_i \le x_i)\Pr(X_j \le x_j)$ holds for all possible pairs (X_i, X_j), then $\mathbf{X}$ is PC. This can be shown by using Hoeffding (1940)'s identity. Furthermore, pairwise PQD implies PC based on various correlation coefficients besides Pearson's r, including Kendall's τ, Spearman's ρ, Gini (1914)'s γ and medial correlation coefficient (Blomqvist, 1950). We can treat correlation coefficients as some measures of average quadrant dependence, so the quadrant dependence is a stricter dependence concept than the correlation dependence.

Quadrant dependence is a basic property of dependent random variables. Almost all commonly-used dependence structures fulfill the property of either the quadrant dependence or the pairwise quadrant dependence. Pairwise PQD is included in a list of conditions that any type of multivariate positive dependence should satisfy which were proposed by Kimeldorf and Sampson (1989) and Pellerey and Semeraro (2003).

(3) Association dependence PA/NA. The random vector $\mathbf{X}$ is *positively associated* (PA) if

$$\operatorname{cov}(\phi(\mathbf{X}), \psi(\mathbf{X})) \ge 0 \tag{4}$$

holds, whenever functions ϕ and ψ are increasing in each component (Esary et al., 1967). Negative association is defined on disjoint subsets of random variables. Let set λ and set η partition $\{1, \ldots, n\}$, then the random vector $\mathbf{X} \in \mathbb{R}^n$ is said to be *negatively associated* (NA) if

$$\operatorname{cov}(\phi(\mathbf{X}_\lambda), \psi(\mathbf{X}_\eta)) \le 0 \tag{5}$$

holds for all componentwise increasing functions ϕ and ψ, where $\mathbf{X}_\lambda = (X_i)_{i\in\lambda}$, $\mathbf{X}_\eta = (X_i)_{i\in\eta}$, $\lambda \cap \eta = \emptyset$ and $\lambda \cup \eta = \{1, \ldots, n\}$ (Joag-Dev and Proschan, 1983).

In two dimensions, PA/NA is equivalent to PQD/NQD. In general $n \ge 3$, association dependence strictly implies orthant dependence (Lehmann, 1966; Joag-Dev and Proschan, 1983). Association dependence has an equivalent form which is closely related to quadrant dependence. Esary et al. (1967) showed that association dependence can be established by using binary increasing function ϕ and ψ in (4) and (5). The proof is directly followed by Hoeffding's identity and its generalization (Cuadras, 2002). Quadrant dependence is implied by association dependence since the collection of set $\{\mathbf{X} > \mathbf{x}\}$ for all $\mathbf{x} \in \mathbb{R}^n$ belongs to the collection of $\{\phi(\mathbf{X}) = 1\}$ for all binary increasing function ϕ's.

(4) Stochastic monotonicity dependence PDS/NDS/CIS/CDS/PRDS. Let λ and η be disjoint subsets of $\{1, \ldots, n\}$, then the random vector $\mathbf{X}_\lambda$ is *stochastically increasing* (SI) in $\mathbf{X}_\eta$ if

$$\Pr(\mathbf{X}_\lambda > \mathbf{x}_\lambda \mid \mathbf{X}_\eta = \mathbf{x}_\eta) \tag{6}$$

is increasing in $\mathbf{x}_\eta$ for all $\mathbf{x}_\lambda$ (Shaked, 1977; Joe, 1997). Analogously, $\mathbf{X}_\lambda$ is *stochastically decreasing* (SD) in $\mathbf{X}_\eta$ if the conditional probability in (6) is decreasing in $\mathbf{x}_\eta$ for all $\mathbf{x}_\lambda$. If $n = 2$, Tukey (1958) said X_2 is *complete positive (negative) regression* on X_1 if $\Pr(X_2 > x_2 \mid X_1 = x_1)$ is increasing (decreasing) in x_1 for all x_2. For the same definition, Lehmann (1966) called X_2 is *positively (negatively) regression dependent* on X_1.

If we restrict $\lambda = \{i\}$ and $\eta = \{1, \ldots, i-1\}$ for all $i = 1, \ldots, n$, then when $\mathbf{X}_{\{i\}}$ is SI in $\mathbf{X}_{\{1,\ldots,i-1\}}$, $\mathbf{X}$ is said to be *conditionally increasing in sequence* (CIS) (Veinott, 1965; Barlow and Proschan, 1975); when $\mathbf{X}_{\{i\}}$ is SD in $\mathbf{X}_{\{1,\ldots,i-1\}}$, $\mathbf{X}$ is said to be *conditionally* *decreasing in sequence* (CDS) (Joag-Dev and Proschan, 1983). Correspondingly, when we restrict $\eta = \{i\}$ and $\lambda = \{1, \ldots, i-1\}$ for all $i = 1, \ldots, n$, then if $\mathbf{X}_{\{1,\ldots,i-1\}}$ is SI in $\mathbf{X}_{\{i\}}$, $\mathbf{X}$ is said to be *positively dependent in sequence*; if $\mathbf{X}_{\{1,\ldots,i-1\}}$ is SD in $\mathbf{X}_{\{i\}}$, $\mathbf{X}$ is said to be *negatively dependent in sequence* (Joag-Dev and Proschan, 1983).

The term *positive (negative) dependence in sequence* has not been acronymed PDS (NDS), since these initialisms have been in use for other dependence properties. When $\mathbf{X}_\lambda$ is SI in $\mathbf{X}_\eta$ for all singleton index sets $\eta = \{i\}$, $i = 1, \ldots, n$, $\mathbf{X}$ is said to be *positively dependent through stochastic ordering* (PDS) (Block et al., 1985). Equivalently, $\mathbf{X}$ is said to be PDS if for any $i = 1, \ldots, n$, the conditional expectation

$$\mathbb{E}\left[\phi\left(\mathbf{X}\right) \mid X_i = x\right] \tag{7}$$

is increasing in x for any increasing function ϕ. The *negatively dependence through stochastic ordering* (NDS) can be defined analogously by replacing the term SI in the definition of PDS with the term SD. Benjamini and Yekutieli (2001) defined *positive regression dependency on each one from a subset* I_0 (PRDS on I_0) by using binary increasing function ϕ in (7), and restricting $i \in I_0$, where I_0 is a prespecified index set. The PRDS requirement is a relaxed condition of the *positive regression dependency* (Sarkar, 1969). The PDS property is equivalent to PRDS on $\{1, \ldots, n\}$.

CIS implies PA (Tong, 1990), CDS implies NA. PDS implies POD (Joe, 1997), NDS implies NOD. PDS and PA do not imply one another (Benjamini and Yekutieli, 2001).

When proving the Simes inequality under positive dependence, we can replace the PDS condition by a slightly weaker condition (Finner et al., 2009): $\mathbf{X}$ satisfies that the conditional expectation $\mathbb{E}\left[\phi\left(\mathbf{X}\right) \mid X_i \leq x\right]$ is increasing in x for any increasing function ϕ and for any $i = 1, \ldots, n$. This version of PDS condition belongs to the *tail monotonicity* dependency property (Esary and Proschan, 1972; Brindley Jr. and Thompson Jr., 1972).

(5) Monotonic likelihood ratio dependence TP_2/RR_2/MTP_2/S-MRR_2. Let $\mathbf{x} \wedge \mathbf{y}$ be the elementwise minimum between $\mathbf{x}$ and $\mathbf{y}$, say, $\mathbf{x} \wedge \mathbf{y} = (\min\{x_1, y_1\}, \ldots, \min\{x_n, y_n\})$, and let $\mathbf{x} \vee \mathbf{y}$ be the elementwise maximum between $\mathbf{x}$ and $\mathbf{y}$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *multivariate totally positive of order two* (MTP_2) if the inequality

$$f\left(\mathbf{x} \wedge \mathbf{y}\right) f\left(\mathbf{x} \vee \mathbf{y}\right) \geq f\left(\mathbf{x}\right) f\left(\mathbf{y}\right) \tag{8}$$

holds for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (Karlin and Rinott, 1980a). The condition in (8) is known as the FKG condition (Fortuin et al., 1971). The random vector $\mathbf{X}$ is said to be *positively likelihood ratio dependent*, or is said to have an MTP_2 distribution, if its probability density function $f\left(\mathbf{x}\right)$ satisfies the FKG condition. Karlin and Rinott (1980b) defined the *multivariate reverse rule of order two* (MRR_2) by reversing the sense of the FKG condition in (8). If $n = 2$, $\mathbf{X} = (X_1, X_2)$ is said to be *totally positive of order two* (TP_2) if the FKG condition holds (Karlin, 1968). In two dimensions, if the inequality in (8) is reversed then we have a condition of negative dependence, called *reverse regular of order two* (RR_2) (Karlin, 1968). However TP_2 and RR_2 behave differently by considering the relation with other dependence conditions. For example, TP_2 in pairs implies POD, but RR_2 in pairs does not imply NOD. The MTP_2 property is preserved under marginalization, but MRR_2 is not (Burkschat, 2009). Thus, a stronger condition is needed for negative dependence.

Block et al. (1982) proposed *condition N* for negative dependence, which implies RR_2 in pairs and NOD. Condition N is also called the *structural condition* (Samuel et al., 2001). The random vector $\mathbf{X} \in \mathbb{R}^n$ fulfills *condition N* if the distribution of $\mathbf{X}$ is related to the distribution of a set of independent random variable conditioning on a fixed sum, say, we can find a fixed number s and a set of $n+1$ random variables $\{S_1, \ldots, S_{n+1}\}$ with PF_2 density functions, which satisfies

$$\mathbf{X} \overset{d}{=} (S_1, \ldots, S_n) \Big| \sum\nolimits_{i=1}^{n+1} S_i = s \ , \tag{9}$$

where a positive function ϕ is a *Pólya frequency function of order two* (PF_2 function) if $\phi(x-y)$ is TP_2 in (x, y) (Karlin and Proschan, 1960).

A concept which is less strict than condition N but stricter than MRR_2 is strong MRR_2. The random vector $\mathbf{X}$ satisfies the *strongly multivariate reverse rule of order two* (S-MRR_2), if its density function $f_{\mathbf{X}}(\mathbf{x})$ satisfies MRR_2, and the marginal density function

$$f_{\mathbf{X}_\lambda}(\mathbf{x}_\lambda) = \int_{\mathbb{R}^{|\eta|}} f_{\mathbf{X}}(\mathbf{x}) \prod\nolimits_{i \in \eta} \phi_i(x_i) \, d\mathbf{x}_\eta \tag{10}$$

is MRR_2 in $\mathbf{x}_\lambda$, for any PF_2 functions $\{\phi_i\}_{i=1}^n$, and for any λ and η which are complementary sets of indexes $\{1, \ldots, n\}$ (Karlin and Rinott, 1980b).

TP_2 in pairs implies PA (Block and Ting, 1981), RR_2 in pairs implies NOD (Block et al., 1982). MTP_2 implies PDS, but it is not known whether S-MRR_2 implies NDS or not (Block et al., 2008). Moreover, for negative dependence, if a set of random variables $\mathbf{X} = (X_1, \ldots, X_n)$ are NA, then any set of increasing functions defined on disjoint subsets of $\{X_1, \ldots, X_n\}$ are also NA (Joag-Dev and Proschan, 1983). NOD also satisfies this closure property. Note that NUOD/NLOD does not satisfy this property although NOD does. NDS and RR_2 in pairs do not possess this closure property (Joag-Dev and Proschan, 1983).

The relations among dependence properties can be summarized as in the following diagram, where pTP_2 denotes TP_2 in pairs, pRR_2 denotes RR_2 in pairs, pPQD denotes pairwise PQD, and pNQD denotes pairwise NQD.

$$\begin{tikzcd}
& \mathrm{pTP}_2 \arrow[r] & \mathrm{PA} \arrow[dr] & & & \\
\mathrm{MTP}_2 \arrow[ur] \arrow[dr] & & & \mathrm{POD} \arrow[r] & \mathrm{pPQD} \arrow[r] & \mathrm{PC} \\
& \mathrm{PDS} \arrow[urr] & & & &
\end{tikzcd}$$

$$\begin{tikzcd}
& \text{S-MRR}_2 \arrow[r] & \mathrm{pRR}_2 \arrow[dr] & & & \\
\text{condition N} \arrow[ur] \arrow[r] \arrow[dr] & \mathrm{NDS} \arrow[rr] & & \mathrm{NOD} \arrow[r] & \mathrm{pNQD} \arrow[r] & \mathrm{NC} \\
& & \mathrm{NA} \arrow[ur] & & &
\end{tikzcd}$$

In this section, we only introduce the concepts of dependence related to the Simes (1986) inequality. For other dependence structures, Tong (1990), Szekli (1995), Joe (1997), Samuel et al. (2001), and Colangelo et al. (2005) provide good summaries.

3 Dependence assumptions of the Simes test

The Bonferroni correction is based on the Bonferroni inequality $\Pr\left(\cup_{i=1}^n \left\{P_i \le \frac{\alpha}{n}\right\}\right) \le \sum_{i=1}^n \Pr\left(P_i \le \frac{\alpha}{n}\right) = \alpha$ (Dunn, 1961). The Bonferroni inequality is assumption-free and becomes an equality only if $\left\{P_i \leq \frac{\alpha}{n}\right\}$ and $\left\{P_j \leq \frac{\alpha}{n}\right\}$ are disjoint for all i and j. The Simes (1986) test is an improved Bonferroni procedure for multiple tests of significance, which is based on Simes' identity $\Pr\left(\cup_{i=1}^n \left\{P_{(i)} \leq \frac{i\alpha}{n}\right\}\right) = \alpha$, where the P_i's are mutually independent uniform random variables on $[0,1]$, and $P_{(1)} \leq \cdots \leq P_{(n)}$ are ordered marginal p-values. Sarkar and Chang (1997) and Sarkar (1998) showed that the Simes test is conservative or equivalently $\Pr\left(\cup_{i=1}^n \left\{P_{(i)} \leq \frac{i\alpha}{n}\right\}\right) \leq \alpha$ if the test statistics are MTP_2. Benjamini and Yekutieli (2001) and Sarkar (2002) later weaken the condition from MTP_2 to PDS. In addition, Block et al. (2013) showed that the standard multivariate-t with nonnegative correlations satisfies Simes inequality under certain sign restrictions, although the multivariate-t is not PDS. The example in Block et al. (2013) indicates that PDS only serves as a sufficient condition for the Simes inequality but not a necessary one. A natural question arises: can we further weaken the dependence assumption for the Simes inequality, for example, using PQD/POD or even PC? For PC, Samuel-Cahn (1996) has provided examples to show the Simes test can be anti-conservative for positively correlated random variables. For PQD/POD, unfortunately, the answer is also no: PQD/POD fails to guarantee the type I error control of the Simes test at level α.

We first consider the bivariate case. Assume that (P_1, P_2) is PQD, and $P_{(1)} \leq P_{(2)}$. The upper bound of $\Pr\left(\left\{P_{(1)} \leq \frac{\alpha}{2}\right\} \cup \left\{P_{(2)} \leq \alpha\right\}\right)$ is $\frac{3\alpha}{2} - \frac{\alpha^2}{2}$, which is greater than α when $\alpha > 0$. This inequality holds since

$$
\begin{aligned}
&\Pr\left(\left\{P_{(1)} \leq \alpha/2\right\} \cup \left\{P_{(2)} \leq \alpha\right\}\right) \\
&= \sum_{i=1}^{2} \Pr\left(P_i \leq \alpha/2\right) - \Pr\left(\cap_{i=1}^2 \{P_i \leq \alpha/2\}\right) + \Pr\left(\cap_{i=1}^2 \{\alpha/2 < P_i \leq \alpha\}\right) \\
&\leq \sum_{i=1}^{2} \Pr\left(P_i \leq \alpha/2\right) - \Pr\left(\cap_{i=1}^2 \{P_i \leq \alpha/2\}\right) + \Pr\left(\alpha/2 < P_1 \leq \alpha\right) \\
&\quad - \Pr\left(\{\alpha/2 < P_1 \leq \alpha\} \cap \{P_2 \leq \alpha/2\}\right) \\
&\leq \alpha/2 + \alpha/2 - (\alpha/2)^2 + \alpha/2 - (\alpha/2)^2 = 3\alpha/2 - \alpha^2/2.
\end{aligned}
$$

This upper bound is achievable. Consider a bivariate PQD distribution of (P_1, P_2) with joint distribution $C(p_1, p_2)$, where

$$
C(p_1, p_2) = \begin{cases}
p_1 p_2 & \text{(Region I)} \\
p_1 p_2 + 2\left(1 - \varepsilon\right)\left(\frac{1}{\alpha} - 1\right)\left(p_1 - \frac{\alpha}{2}\right)\left(p_2 - \frac{\alpha}{2}\right) & \text{(Region II)} \\
\left((p_1 \wedge p_2) - \frac{\alpha}{2}\right)\left(1 - \left(1 - (p_1 \vee p_2)\right)\varepsilon\right) + \frac{\alpha}{2}\left(p_1 \vee p_2\right) & \text{(Region III)} \\
p_1 p_2 + \frac{\varepsilon\alpha(1-\varepsilon\alpha)(1-\alpha)}{2} + \frac{(p_1 - 1 + \varepsilon\alpha(1-\alpha))(p_2 - 1 + \varepsilon\alpha(1-\alpha))}{2(1-\alpha)} & \text{(Region IV)} \\
\text{or } p_1 p_2 + \frac{(1-p_1)(1-p_2)}{2(1-\alpha)} + \frac{\varepsilon\alpha}{2}\left(p_1 + p_2 - 1 - \alpha\right) & \text{(Region IV)}
\end{cases}
$$

and ε is a small positive real number. This example is a generalization of an example from Gou and Tamhane (2018). The region for defining $C(p_1, p_2)$ is shown in Figure 1.

The joint distribution $C(p_1, p_2)$ is a copula since the marginal distributions are uniform

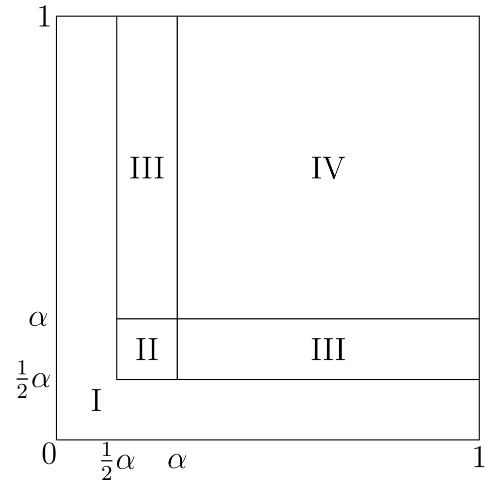

Figure 1: The region for defining $C(p_1, p_2)$

distribution on $[0, 1]$. The corresponding joint density function is

$$\frac{\partial^2 C(p_1,p_2)}{\partial p_1 \partial p_2} = \begin{cases} 1, & 0 \le p_1, p_2 \le \alpha/2 \text{ and } \alpha/2 \le p_1 \le 1, 0 \le p_2 \le \alpha/2 \\ & \quad \text{and } 0 \le p_1 \le \alpha/2, \alpha/2 \le p_2 \le 1 \text{ (Region I)}, \\ \frac{2(1-\varepsilon)}{\alpha} - (1-2\varepsilon), & \alpha/2 \le p_1, p_2 \le \alpha \text{ (Region II)} \\ \varepsilon, & \alpha \le p_1 \le 1, \alpha/2 \le p_2 \le \alpha \text{ and} \\ & \quad \alpha/2 \le p_1 \le \alpha, \alpha \le p_2 \le 1 \text{ (Region III)}, \\ \frac{2-\alpha}{2(1-\alpha)}, & \alpha \le p_1, p_2 \le 1 \text{ (Region IV)}, \end{cases}$$

and we compute $\Pr\left(\left\{P_{(1)} \le \alpha/2\right\} \cup \left\{P_{(2)} \le \alpha\right\}\right) = (3-\varepsilon)\alpha/2 - (1-\varepsilon)\alpha^2/2$, which tends to the upper bound $3\alpha/2 - \alpha^2/2$ as $\varepsilon \to 0$.

For the multivariate case, assuming $(P_1, \cdots, P_n)$ is POD, the asymptotic achievable upper bound of $\Pr\left(\cup_{i=1}^n \left\{P_{(i)} \le \frac{i\alpha}{n}\right\}\right)$ is at least $\left(2 - \frac{1}{n}\right)\alpha + o(\alpha)$. In order to find this upper bound, we consider an n-variate POD distribution of $(P_1, \ldots, P_n)$ with joint density function which is

$$\frac{\partial^n C(p_1,\ldots,p_n)}{\partial p_1 \ldots \partial p_2} = \begin{cases} 1, & (p_1,\ldots,p_n) \in \cup_{i=1}^n \{p_i \le \frac{\alpha}{n}\} \\ \frac{(n-\alpha)^{n-1}}{(n-1)^{n-1}\alpha^{n-1}}, & (p_1,\ldots,p_n) \in \cap_{i=1}^n \{\frac{\alpha}{n} < p_i \le \alpha\} \\ \frac{(n-\alpha)^{n-1}}{n^{n-1}(1-\alpha)^{n-1}}, & (p_1,\ldots,p_n) \in \cap_{i=1}^n \{p_i > \alpha\}, \\ 0, & \text{others.} \end{cases} \tag{11}$$

The calculation of the type I error rate of the Simes times under the POD distribution above is included in Appendix. This result indicates that the Simes test can be anti-conservative and its type I error can be inflated by a factor of 2 when n is large and α is small, if we only assume the dependence is POD.

Next we investigate the type I error control under PQD/POD from another direction. Consider a test on global intersection hypotheses based on ordered p-values with type I error rate $\Pr\left(\cup_{i=1}^n \{P_{(n-i+1)} \le c_i\alpha\}\right)$. The Bonferroni test is conservative under PQD/POD with $c_i = 1/n$, and the Simes test doesn't guarantee the type I error control under PQD/POD with

$c_i = (n-i+1)/n$. Therefore, there exist tests with critical values $1/n \le c_i \le (n-i+1)/n$ which control the type I error under PQD/POD. Holland and Copenhaver (1987, 1988) showed that Sidak (1967)'s test with $c_i = (1-(1-\alpha)^{1/n})/\alpha = \frac{1}{n} + \frac{n+1}{2n^2}\alpha + o(\alpha)$ is valid under POD. Proposition 1 considers a test with critical boundaries $c_2 = \cdots = c_n = 1/n$.

Proposition 1. *If* $\Pr\left(\cup_{i=1}^{n-1}\left\{P_{(i)} \le \frac{\alpha}{n}\right\} \cup \left\{P_{(n)} \le c\alpha\right\}\right) \le \alpha$ *holds for any random vector* $\boldsymbol{P} = (P_1, \cdots, P_n)$ *which is PQD/POD, then*

$$c \le \frac{1}{n} + \frac{(1-\alpha/n)^n - (1-\alpha)}{\alpha(1-\alpha/n)^{n-1}} = \frac{1}{n} + \frac{n-1}{2n}\alpha + o(\alpha).$$

Proposition 1 shows that we may only be able to use a procedure slightly more powerful than the Bonferroni test for positively dependent hypotheses which satisfy the PQD/POD assumption. This test has critical constants $c_1 = \frac{1}{n} + \frac{n-1}{2n}\alpha + o(\alpha), c_2 = \cdots = c_n = \frac{1}{n}$. For example, consider the case with two hypotheses. The critical constants are $c_1 = \frac{1}{2} + \frac{\alpha/4}{1-\alpha/2} \approx 0.513$ and $c_2 = 0.5$ when $\alpha = 5\%$, comparing with the Bonferroni test where $c_1 = c_2 = 0.5$ and the Simes test where $c_1 = 1$ and $c_2 = 0.5$.

Let us next consider the Simes test under negative dependence. Block et al. (2008) show that the Simes test is anti-conservative when the test statistic is NDS. The draft guideline on multiple endpoints by the US Food and Drug Administration includes a statement: "*the Hochberg procedure fails to control the overall type I error rate for some negatively-correlated endpoints*" (FDA, 2017). The conclusion under stochastic monotonicity dependence looks simple: the Simes test is conservative under PDS but anti-conservative under NDS. However, when considering the quadrant dependence, the conclusion turns out contrary to what we may expect: the control of error rate under PQD/POD could be worse than the control under NQD/NOD.

Proposition 2. *If* (P_1, P_2) *is NQD, then* $\Pr\left(\left\{P_{(1)} \le \frac{\alpha}{2}\right\} \cup \left\{P_{(2)} \le \alpha\right\}\right) \le \alpha + \alpha^2$. *If* $(P_1, \cdots, P_n)$ *is NOD, then* $\Pr\left(\cup_{i=1}^{n}\left\{P_{(i)} \le \frac{i\alpha}{n}\right\}\right) \le \alpha + \frac{2(n-1)}{n}\alpha^2 + o(\alpha^2)$*, where* $P_{(1)} \le \cdots \le P_{(n)}$ *are ordered p-values.*

Therefore, when n is large and α is small, the upper bound of type I error rate of the Simes test under POD is about 2α and that under NOD is around $\alpha + 2\alpha^2$. The type I error inflation under NQD/NOD is actually less severe than that under PQD/POD.

We have an explanation of the behavior of the Simes test under positive dependence and negative dependence from the copula viewpoint. The distribution of $(P_1, \ldots, P_n)$ is an n-copula, which satisfies the Fréchet-Hoeffding bounds: if C is a n-copula, then for any $\boldsymbol{p} = (p_1, \ldots, p_n)$, the inequality $W^n(\boldsymbol{p}) \le C(\boldsymbol{p}) \le M^n(\boldsymbol{p})$ holds, where $M^n(\boldsymbol{p}) = \min\{p_1, \cdots, p_n\}$ is the Fréchet-Hoeffding upper bound describing the limit of positively dependent p-values, and $W^n(\boldsymbol{p}) = \max\{1 - \sum_{i=1}^n (1-p_i), 0\}$ is the Fréchet-Hoeffding lower bound describing the limit of negatively dependent p-values. In addition, the product n-copula $\Pi^n(\boldsymbol{p}) = \prod_{i=1}^n p_i$ describes the distribution of independent p-values. Nelsen (2006) calculates the ratio of the volume of the space between the upper bound and the independent case to that between the lower bound and the independence case, which is

$$\frac{\int_{[0,1]^n} (M^n(\boldsymbol{p}) - \Pi^n(\boldsymbol{p}))\, d\boldsymbol{p}}{\int_{[0,1]^n} (\Pi^n(\boldsymbol{p}) - W^n(\boldsymbol{p}))\, d\boldsymbol{p}} = \frac{1/(n+1) - 1/2^n}{1/2^n - 1/(n+1)!} \to \infty$$

as n goes to ∞. This result indicates that the surfaces of $z = W^n(\boldsymbol{p})$ and $\Pi^n(\boldsymbol{p})$ are much closer to one another than are the surfaces of $z = M^n(\boldsymbol{p})$ and $\Pi^n(\boldsymbol{p})$. In other words, the distribution under negative dependence is less flexible than that under positive dependence. This partially explains why the upper bound of the probability in Simes' inequality under general positive dependence is greater than that under general negative dependence.

4 Discussion

Without knowing the distribution of the collected data, the type I error control of the Simes-test-based methods under dependence requires a relatively strict condition which is PDS. Relaxing the positive dependence assumption in the Simes test may result a significant inflation of type I error rate. However, if the distribution of test statistics is known, the dependence assumption for the Simes test can be simplified and verified from the observed data. For example, if the test statistics follow a multivariate normal distribution with covariance matrix $\boldsymbol{\Sigma}$, then the Simes inequality holds when all the non-diagonal elements in the precision matrix $\boldsymbol{\Sigma}^{-1}$ are non-positive (Karlin and Rinott, 1980a; Sarkar, 1998). Bodnar and Dickhaus (2017) provided conditions for the validity of the Simes inequality where the test statistics follow the elliptically contoured distributions (Gupta et al., 2013). With these distribution assumptions, checking the non-negativity of correlation coefficients between test statistics becomes sufficient to validate the dependence requirement for the Simes-test-based methods.

The inflation of type I error rate can be significant if we assume PQD/POD which is a weaker condition than PDS. Modifying the Simes test under PQD/POD can only create a test that is similar to the Bonferroni test. However, the error rate inflation of the Simes test under a general negative dependence assumption is not as severe as that under a positive dependence structure. It is possible to develop powerful procedure to control the type I error rate at level α under negative dependence. For example, Gou and Tamhane (2014) presented some generalized Simes test which is conservative under certain negative dependence assumptions. It is also possible to develop specific procedures for some specific multiple testing questions (Tamhane et al., 2020). For example, Lynch et al. (2017) created an FDR-controlling fixed sequence multiple test procedure under certain negative dependence conditions.

It is feasible to check the dependence structure using the distribution information or creating some new multiple test procedures to replace the Simes-test-based methods. Alternative, we can develop a statistical test to test the PDS condition, or other positive dependence conditions which are less strict than PDS but still sufficient to guarantee the validity of Simes' inequality, and let data themselves decide whether the Simes-test-based procedures are suitable or not. This topic will be further investigated and discussed in a separate paper.

Acknowledgment

We thank Ajit C. Tamhane, Helmut Finner, Xin Wang, Lingyun Liu, Lie Leslie Liu, Tianmeng Lyu and Shuyi Wu for helpful discussion. Some preliminary results of this research

article were present at the 10th International Conference on Multiple Comparison Procedures in Riverside, California, Session WAM 2-3 Multiple testing, *Tests for the positive dependence assumption of Simes' inequality* on June 21, 2017.

Appendix

Derivation of the type I error rate of the Simes test under the distribution shown in (11) and proofs of Proposition 1 and 2 are included in this appendix.

Derivation of the type I error rate under the distribution in (11). We compute the probability of rejecting the overall null hypothesis using the Simes test, where the p-values follow the multivariate distribution in (11).

$$\begin{aligned}
\Pr\left(\cup_{i=1}^{n}\left\{P_{(i)} \leq \frac{i\alpha}{n}\right\}\right) &= 1\cdot\left(1-\left(1-\frac{\alpha}{n}\right)^{n}\right)+\frac{(n-\alpha)^{n-1}}{(n-1)^{n-1}\alpha^{n-1}}\cdot\left(\frac{n-1}{n}\cdot\alpha\right)^{n} \\
&= 1-(1-\alpha)\left(1-\frac{\alpha}{n}\right)^{n-1} \\
&= \alpha\cdot\left[\sum_{i=0}^{n-1}\binom{n-1}{i}\cdot\left(-\frac{\alpha}{n}\right)^{i}+\sum_{i=0}^{n-2}\frac{1}{n}\cdot\binom{n-1}{i+1}\cdot\left(-\frac{\alpha}{n}\right)^{i}\right] \\
&= \alpha\cdot\left[\sum_{i=0}^{n-1}\binom{n}{i+1}\frac{n(i+2)-(i+1)}{n^2}\cdot\left(-\frac{\alpha}{n}\right)^{i}\right] \\
&= \sum_{i=1}^{n}(-1)^{i-1}\cdot\binom{n}{i}\cdot\frac{1+i\cdot\left(1-\frac{1}{n}\right)}{n^{i}}\cdot\alpha^{i},
\end{aligned}$$

and the expansion is $\frac{2n-1}{n}\alpha-\frac{(n-1)(3n-2)}{2n}\alpha^2+o(\alpha^2)$. The inflation factor is almost two. For example, when $\alpha=5\%$ and $n=10$, the exact type I error rate $1-(1-\alpha)\left(1-\frac{\alpha}{n}\right)^{n-1}$ is 9.2% and the first order approximation $\frac{2n-1}{n}\alpha$ is 9.5%. The inflation factors are 1.84 and 1.90 correspondingly. □

Proof of Proposition 1. We only need to find a distribution that satisfies the PQD/POD condition. Under this distribution, the solution c to equation $\Pr\left(\cup_{i=1}^{n-1}\left\{P_{(i)}\leq\frac{\alpha}{n}\right\}\cup\left\{P_{(n)}=c\alpha\right\}\right)=\alpha$ provides an upper bound for all PQD/POD distributions. Consider a n-copula with density function $\partial^n C(p_1,\ldots,p_n)/\partial p_1\cdots\partial p_n=1$ at region $\cup_{i=1}^{n}\{p_i\leq\alpha/n\}$, $\partial^n C(p_1,\ldots,p_n)/\partial p_1\cdots\partial p_n=0$ at $\{\cup_{i=1}^{n}\{p_i\leq c\alpha\}\}\cap\{\cap_{i=1}^{n}\{p_i>\alpha/n\}\}\cup\{\cap_{i=1}^{n}\{\alpha/n<p_i\leq c\alpha\}$, $\partial^n C(p_1,\ldots,p_n)/\partial p_1\cdots\partial p_n=(1-\alpha/n)^{n-1}/\alpha^{n-1}/(c-1/n)^{n-1}$ at $\cap_{i=1}^{n}\{\alpha/n<p_i\leq c\alpha\}$, and $\partial^n C(p_1,\ldots,p_n)/\partial p_1\cdots\partial p_n=(1-\alpha/n)^{n-1}/(1-c\alpha)^{n-1}$ at $\cap_{i=1}^{n}\{p_i>c\alpha\}$. For a level-$\alpha$ test, the critical value c can be solved from $(c\alpha-\alpha/n)\cdot(1-\alpha/n)^{n-1}=(1-\alpha/n)^n-(1-\alpha)$, and we achieve $c=\frac{1}{n}+\frac{(1-\alpha/n)^n-(1-\alpha)}{\alpha(1-\alpha/n)^{n-1}}$ and $c=\frac{1}{n}+\frac{n-1}{2n}\alpha+o(\alpha)$. □

Proof of Proposition 2. For bivariate (P_1,P_2) which satisfies the NQD condition, the probability $\Pr\left(\left\{P_{(1)}\leq\frac{\alpha}{2}\right\}\cup\left\{P_{(2)}\leq\alpha\right\}\right)\leq\Pr\left(\cup_{i=1}^{2}\{P_i\leq\alpha/2\}\right)+\Pr\left(\cap_{i=1}^{2}\{P_i\leq\alpha\}\right)\leq\Pr\left(P_1\leq\alpha/2\right)+\Pr\left(P_2\leq\alpha/2\right)+\Pr\left(P_1\leq\alpha\right)\cdot\Pr\left(P_2\leq\alpha\right)=\alpha+\alpha^2$, where $\Pr\left(\cap_{i=1}^{2}\{P_i\leq\alpha\}\right)\leq\Pr\left(P_1\leq\alpha\right)\cdot\Pr\left(P_2\leq\alpha\right)$ since (P_1,P_2) is NQD. For multivariate $(P_1,\ldots,P_n)$ which satisfies the NOD condition, similarly, we can show that $\Pr\left(\cup_{i=1}^{n}\left\{P_{(i)} \leq \frac{i\alpha}{n}\right\}\right) \leq \sum_{i=1}^{n}\Pr\left(P_i \leq \alpha/n\right) + \sum_{i<j}\Pr\left(\{P_i \leq 2\alpha/n\} \cap \{P_j \leq 2\alpha/n\}\right) + o(\alpha^2) \leq n\cdot\alpha/n + \binom{n}{2}\cdot(2\alpha/n)^2 = \alpha + \frac{2(n-1)}{n}\alpha^2 + o(\alpha^2)$, where $\Pr\left(\cap_{j=1}^{k}\{P_{i_j} \leq k\alpha/n\}\right)$ are $o(\alpha^2)$ terms for $k \geq 3$ because $(P_1, \ldots, P_n)$ is NOD. $\square$

References

Barlow, R. and Proschan, F. (1975). *Statistical theory of reliability and life testing: probability models.* International series in decision processes. Holt, Rinehart and Winston, New York.

Belzunce, F. and Semeraro, P. (2004). Preservation of positive and negative orthant dependence concepts under mixtures and applications. *Journal of Applied Probability* **41,** 961–974.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57,** 289–300.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* **29,** 1165–1188.

Block, H. W., Savits, T. H., and Shaked, M. (1982). Some concepts of negative dependence. *The Annals of Probability* **10,** 765–772.

Block, H. W., Savits, T. H., and Shaked, M. (1985). A concept of negative dependence using stochastic ordering. *Statistics and Probability Letters* **3,** 81–86.

Block, H. W., Savits, T. H., and Wang, J. (2008). Negative dependence and the Simes inequality. *Journal of Statistical Planning and Inference* **138,** 4107–4110.

Block, H. W., Savits, T. H., Wang, J., and Sarkar, S. K. (2013). The multivariate-t distribution and the Simes inequality. *Statistics and Probability Letters* **83,** 227–232.

Block, H. W. and Ting, M.-L. (1981). Some concepts of mult1variate dependence. *Communications in Statistics - Theory and Methods* **10,** 749–762.

Blomqvist, N. (1950). On a measure of dependence between two random variables. *The Annals of Mathematical Statistics* **21,** 593–600.

Bodnar, T. and Dickhaus, T. (2017). On the Simes inequality in elliptical models. *Annals of the Institute of Statistical Mathematics* **69,** 215–230.

Brindley Jr., E. C. and Thompson Jr., W. A. (1972). Dependence and aging aspects of multivariate survival. *Journal of the American Statistical Association* **67,** 822–830.

Burkschat, M. (2009). Multivariate dependence of spacings of generalized order statistics. *Journal of Multivariate Analysis* **100,** 1093–1106.

Colangelo, A., Scarsini, M., and Shaked, M. (2005). Some notions of multivariate positive dependence. *Insurance: Mathematics and Economics* **37,** 13–26.

Cuadras, C. (2002). On the covariance between functions. *Journal of Multivariate Analysis* **81,** 19–27.

Daniels, H. E. (1950). Rank correlation and population models. *Journal of the Royal Statistical Society. Series B (Methodological)* **12,** 171–191.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association* **56,** 52–64.

Durbin, J. and Stuart, A. (1951). Inversions and rank correlation coefficients. *Journal of the Royal Statistical Society. Series B (Methodological)* **13,** 303–309.

Dykstra, R. L., Hewett, J. E., and Thompson, W. A. (1973). Events which are almost independent. *The Annals of Statistics* **1,** 674–681.

Ebrahimi, N. and Ghosh, M. (1981). Multivariate negative dependence. *Communications in Statistics - Theory and Methods* **10,** 307–337.

Esary, J. D. and Proschan, F. (1972). Relationships among some concepts of bivariate dependence. *The Annals of Mathematical Statistics* **43,** 651–655.

Esary, J. D., Proschan, F., and Walkup, D. W. (1967). Association of random variables, with applications. *The Annals of Mathematical Statistics* **38,** 1466–1474.

FDA (2017). *Multiple endpoints in clinical trials guidance for industry (draft guidance).* U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). FDA-2016-D-4460.

Finner, H., Dickhaus, T., and Roters, M. (2009). On the false discovery rate and an asymptotically optimal rejection curve. *The Annals of Statistics* **37,** 596–618.

Fortuin, C. M., Kasteleyn, P. W., and Ginibre, J. (1971). Correlation inequalities on some partially ordered sets. *Communications in Mathematical Physics* **22,** 89–103.

Gebelein, H. (1941). Das statistische problem der korrelation als variations- und eigenwertproblem und sein zusammenhang mit der ausgleichsrechnung. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik* **21,** 364–379.

Gini, C. (1914). *L'ammontare e la composizione della ricchezza delle nazioni.* F. Bocca, Torino.

Gou, J. and Tamhane, A. C. (2014). On generalized Simes critical constants. *Biometrical Journal* **56,** 1035–1054.

Gou, J. and Tamhane, A. C. (2018). Hochberg procedure under negative dependence. *Statistica Sinica* **28,** 339–362.

Gupta, A. K., Varga, T., and Bodnar, T. (2013). *Elliptically Contoured Models in Statistics and Portfolio Theory.* Springer, New York, New York.

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

Hochberg, Y. and Tamhane, A. C. (1987). *Multiple Comparison Procedures.* John Wiley and Sons, New York, New York.

Hoeffding, W. (1940). Masstabinvariante Korrelationstheorie. *Schriften des Matematischen Institus und des Instituts fur Angewandte Mathematik der Universitat Berlin* **5,** 179–233.

Holland, B. S. and Copenhaver, M. D. (1987). An improved sequentially rejective Bonferroni test procedure. *Biometrics* **43,** 417–423.

Holland, B. S. and Copenhaver, M. D. (1988). Improved Bonferroni-type multiple testing procedures. *Psychological Bulletin* **104,** 145–149.

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. *Biometrika* **75,** 383–386.

Huque, M. F. (2016). Validity of the Hochberg procedure revisited for clinical trial applications. *Statistics in Medicine* **35,** 5–20.

Joag-Dev, K. and Proschan, F. (1983). Negative association of random variables with applications. *The Annals of Statistics* **11,** 286–295.

Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts.* Chapman and Hall, London.

Karlin, S. (1968). *Total Positivity*, volume 1. Stanford University Press, Stanford, California.

Karlin, S. and Proschan, F. (1960). Polya type distributions of convolutions. *The Annals of Mathematical Statistics* **31,** 721–736.

Karlin, S. and Rinott, Y. (1980a). Classes of orderings of measures and related correlation inequalities. I. multivariate totally positive distributions. *Journal of Multivariate Analysis* **10,** 467–498.

Karlin, S. and Rinott, Y. (1980b). Classes of orderings of measures and related correlation inequalities. II. multivariate reverse rule distributions. *Journal of Multivariate Analysis* **10,** 499–516.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika* **30,** 81–93.

Kendall, M. G. (1945). The treatment of ties in ranking problems. *Biometrika* **33,** 239–251.

Kimeldorf, G. and Sampson, A. R. (1978). Monotone dependence. *The Annals of Statistics* **6,** 895–903.

Kimeldorf, G. and Sampson, A. R. (1989). A framework for positive dependence. *Annals of the Institute of Statistical Mathematics* **41,** 31–45.

Lehmann, E. L. (1966). Some concepts of dependence. *The Annals of Mathematical Statistics* **37,** 1137–1153.

Lynch, G., Guo, W., Sarkar, S. K., and Finner, H. (2017). The control of the false discovery rate in fixed sequence multiple testing. *Electronic Journal of Statistics* **11,** 4649–4673.

Nelsen, R. B. (2006). *An Introduction to Copulas.* Springer Series in Statistics. Springer Science+Business Media, Inc., New York, New York, 2 edition.

Pellerey, F. and Semeraro, P. (2003). A positive dependence notion based on the supermodular order. Technical report, Dipartimento di Matematica, Politecnico di Torino, Torino, Italy.

Rényi, A. (1959). On measures of dependence. *Acta Mathematica Academiae Scientiarum Hungarica* **10,** 441–451.

Samuel, M., Mari, D., and Kotz, S. (2001). *Correlation and Dependence.* EBL-Schweitzer. Imperial College Press.

Samuel-Cahn, E. (1996). Is the Simes improved Bonferroni procedure conservative? *Biometrika* **83,** 928–933.

Sarkar, S. K. (1998). Some probability inequalities for ordered MTP_2 random variables: a proof of the Simes conjecture. *The Annals of Statistics* **26,** 494–504.

Sarkar, S. K. (2002). Some results on false discovery rate in stepwise multiple testing procedures. *The Annals of Statistics* **30,** 239–257.

Sarkar, S. K. and Chang, C. K. (1997). The Simes method for multiple hypothesis testing with positively dependent test statistics. *Journal of the American Statistical Asociation* **92,** 1601–1608.

Sarkar, T. K. (1969). Some lower bounds of reliability. Technical Report 124, Department of Operations Research and Department of Statistics, Stanford University, Stanford, California.

Schweizer, B. and Wolff, E. F. (1981). On nonparametric measures of dependence for random variables. *The Annals of Statistics* **9,** 879–885.

Shaked, M. (1977). A family of concepts of dependence for bivariate distributions. *Journal of the American Statistical Association* **72,** 642–650.

Sidak, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association* **62,** 626–633.

Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika* **73,** 751–754.

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology* **15,** 72–101.

Szekli, R. (1995). *Stochastic Ordering and Dependence in Applied Probability.* Lecture Notes in Statistics. Springer, New York.

Tamhane, A. C. and Gou, J. (2018). Advances in p-value based multiple test procedures. *Journal of Biopharmaceutical Statistics* **28,** 10–27.

Tamhane, A. C., Gou, J., and Dmitrienko, A. (2020). Some drawbacks of the Simes test in the group sequential setting. *Statistics in Biopharmaceutical Research* **12,** 390–393.

Tong, Y. L. (1990). *The Multivariate Normal Distribution.* Springer-Verlag, New York.

Tukey, J. W. (1958). A problem of berkson, and minimum variance orderly estimators. *The Annals of Mathematical Statistics* **29,** 588–592.

Veinott, A. F. (1965). Optimal policy in a dynamic, single product, nonstationary inventory model with several demand classes. *Operations Research* **13,** 761–778.