On the Evaluation of the Joint Distribution of Order Statistics

Koon Shing KWONG
Singapore Management University, kskwong@smu.edu.sg

Yiu Man CHAN
National University of Singapore

Citation

KWONG, Koon Shing and CHAN, Yiu Man. On the Evaluation of the Joint Distribution of Order Statistics. (2008). *Computational Statistics and Data Analysis*. 52, (12), 5091-5099. Research Collection School Of Economics.
Available at: https://ink.library.smu.edu.sg/soe_research/199

On the evaluation of the joint distribution of order statistics

Koon-Shing Kwong [a,*], Yiu-Man Chan [b]

[a] *School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903, Singapore*

[b] *Department of Statistics and Applied Probability, National University of Singapore, 6 Science Drive 2, Singapore 117546, Singapore*

A B S T R A C T

Dunnett and Tamhane [Dunnett, C.W., Tamhane, A.C., 1992. A step-up multiple test procedure. J. Amer. Statist. Assoc. 87, 162–170.] proposed a step-up procedure for comparing k treatments with a control and showed that the step-up procedure is more powerful than its counterpart single step and step-down procedures. Since then, several modified step-up procedures have been suggested to deal with different testing environments. In order to establish those step-up procedures, it is necessary to derive approaches for evaluating the joint distribution of the order statistics. In some cases, experimenters may have difficulty in applying those step-up procedures in multiple hypothesis testing because of the computational limitation of existing algorithms in evaluating the critical values for a large number of multiple comparisons. As a result, most procedures are only workable when the design of the experiment is balanced with $k \leq 20$ or unbalanced with $k \leq 8$. In this paper, new algorithms are proposed in order to effectively compute the joint distribution of order statistics in various situations. An extensive numerical study shows that the proposed algorithms can easily handle the testing situations with a much larger k. Examples of applying the proposed algorithms to evaluate the critical values of two existing step-up procedures are also presented.

1. Introduction

The establishment of some statistical testing procedures depends on the evaluation of the joint cumulative distribution of order statistics *e.g.* Dalal and Mallows (1992) stopping rules for testing software faults and the step-up multiple test procedure proposed by Dunnett and Tamhane (1992), hereafter called DT. Recently, several modified step-up procedures have been proposed to deal with different testing environments. For instance, Dunnett and Tamhane (1997), Kwong (2001a) and Kwong et al. (2004) discussed the two-stage step-up multiple testing to establish superiority and equivalence of the efficacy between several treatments and a control. Besides, Tamhane et al. (1998) proposed the combination of step-up and step-down procedures to a generalized step-up-down procedure. Moreover, Kwong et al. (2007) proposed p-value consistent step-up procedure in the direction-mixed families which contain a combination of one- and two-sided inferences. Therefore, the evaluation of the joint distribution of order statistics plays a critical role for the applications and development of those stepwise testing procedures.

Several approaches for the evaluation of the joint distribution of order statistics in the independent and identically distributed (iid) case have been proposed. For example, based on the multinomial argument, Ma (1997) suggested a saddle-point approximation approach. Although the numerical study illustrates the great accuracy of the approximations in low dimensional cases, some existing approaches are also available to get the exact values in those cases. Therefore, this approximation approach is not desirable and will not be discussed any more in this paper. In order to evaluate the joint

---

* Corresponding author. Tel.: +65 68280834; fax: +65 68280833.
  E-mail address: kskwong@smu.edu.sg (K.-S. Kwong).

distribution exactly, DT proposed a recursive algorithm by successively conditioning the random variates. However, Kwong (2001b) pointed out that the approach is not efficient in computation processes under the iid case. This paper focuses on two exact approaches suggested by Finner and Roters (1994) and Kwong (2001b). Both approaches use the recursive equations to evaluate the joint distribution in two different ways. Based on multinomial argument, a new exact approach will be proposed. By applying these three approaches to evaluate the critical values of DT's step-up procedure under balanced design, their general performances can then be compared.

For the independent, but not identically distributed (ibnid) case, Dunnett and Tamhane (1995), Liu (1997), and Kwong and Liu (2000) have proposed different evaluation procedures. Among all the existing procedures, Kwong and Liu's (2000) approach, which is a modified DT's approach, is believed to be the most efficient and effective method to evaluate the joint distribution of order statistics in this case. In this paper, their algorithm is further modified in order to reduce the complexity in the computational process. In order to show the significant improvement of the modified algorithm compared with the existing algorithm, the modified algorithm is applied to evaluate the critical p-values of p-value consistent step-up procedure suggested by Kwong et al. (2007) in the direction-mixed families with the unbalanced designs.

2. Independent and identically distributed

2.1. Existing algorithms

Let $X_1, \ldots, X_n$ be iid random variables with a known cumulative distribution function G. Denote $(a_1, \ldots, a_k) < (b_1, \ldots, b_k)$ be $a_{(i)} < b_{(i)}$ for $i = 1, \ldots, k$, where $a_{(1)} \leq \cdots \leq a_{(k)}$ and $b_{(1)} \leq \cdots \leq b_{(k)}$ are the ordered values of a_i's and b_i's, respectively. For any given constants $\mathbf{b}_n = (b_1, \ldots, b_n)$ where $b_1 \leq b_2 \leq \cdots \leq b_n$, this paper discusses the evaluation of the joint distribution of order statistics

$$F_n(\mathbf{b}_n) = \Pr\left[(X_1, \ldots, X_n) \leq (b_1, \ldots, b_n)\right].$$

By successive conditioning the values of X_i, DT proposed an algorithm to evaluate the F_n. But the algorithm is not so efficient especially when n is large. Finner and Roters (1994) provided the following recursive formula for evaluating F_n:

$$F_n(\mathbf{b}_n) = 1 - \sum_{j=0}^{n-1} \binom{n}{j} F_j(\mathbf{b}_j)[1 - G(b_{j+1})]^{n-j} \tag{1}$$

where $F_0 = 1$. Without being aware of their recursive formula, Kwong (2001b) proposed another similar recursive formula as follows:

$$F_n(\mathbf{b}_n) = \sum_{j=1}^{n} (-1)^{j-1} \binom{n}{j} F_{n-j}(\mathbf{b}_{n-j})[G(b_{n-j+1})]^j. \tag{2}$$

Note that under the formulas (1) and (2), the calculation of F_n must be carried out recursively, i.e. one must obtain the values of $F_1, \ldots, F_{j-1}$ in order to determine F_j for $j = 2, \ldots, n$.

2.2. Proposed algorithm

Based on the multinomial argument, the other exact approach is now derived. Define N_j be the number of X_i's having the value between $(b_{j-1}, b_j]$ and e_j be $G(b_j) - G(b_{j-1})$ for $j = 1, \ldots, n$, where $b_0 = -\infty$. Then, $(N_1, \ldots, N_n)$ is a multinomial distribution with parameters $(n, e_1, \ldots, e_n)$. Since the joint event of $(X_1, \ldots, X_k) < (b_1, \ldots, b_k)$ is equivalent to the other joint event of $\sum_{j=1}^{k} N_j \geq k$ for $k = 1, \ldots, n-1$ and $\sum_{j=1}^{n} N_j = n$, the function $F_n(\mathbf{b}_n)$ can be written as

$$F_n(\mathbf{b}_n) = \Pr\left[\bigcap_{k=1}^{n-1}\left(\sum_{j=1}^{k} N_j \geq k\right), \sum_{j=1}^{n} N_j = n\right]. \tag{3}$$

By conditioning N_1, (3) is then modified to

$$F_n(\mathbf{b}_n) = \sum_{i_1=1}^{n} \Pr(N_1 = i_1) \Pr[A_1(s_1)]$$

where the event

$$A_r(s_r) = \left[\bigcap_{k=r+1}^{n-1}\left(\sum_{j=r+1}^{k} N_j \geq j - s_r\right), \sum_{j=r+1}^{n} N_j = n - s_r \,\middle|\, \sum_{j=1}^{r} N_j = s_r\right]$$

and $s_r = \sum_{l=1}^{r} i_l$ for $r = 1, \ldots, n-1$. As the marginal distribution of N_1 is binomial with parameters (n, e_1), (3) is further reduced to

$$F_n(\mathbf{b}_n) = \sum_{i_1=1}^{n} \binom{n}{i_1} (e_1)^{i_1} H_1(s_1),$$

where

$$H_r(s_r) = \begin{cases} 1 & \text{for } s_r = n \\ \left(1 - \sum_{j=1}^{r} e_j\right)^{n-s_r} \Pr[A_r(s_r)] & \text{for } r \leq s_r < n \\ 0 & \text{for } s_r < r. \end{cases}$$

It is well known that $(N_{r+1}, \ldots, N_n \mid \sum_{j=1}^{r} N_j = s_r)$ is also a multinomial distribution with parameters $(n - s_r, e_{r+1}^{(r)}, \ldots, e_n^{(r)})$, where $e_j^{(r)} = e_j/(1 - \sum_{i=1}^{r} e_i)$ for $j = r+1, \ldots, n$ and $r = 1, \ldots, n-1$. By conditioning N_2 in $\Pr[A_1(s_1)]$, the function $H_1(s_1)$ becomes

$$H_1(s_1) = (1 - e_1)^{n-s_1} \sum_{i_2=0}^{n-s_1} \Pr(N_2 = i_2) \Pr[A_2(s_2)]$$

$$= (1 - e_1)^{n-s_1} \sum_{i_2=0}^{n-s_1} \binom{n - s_1}{i_2} (e_2^{(2)})^{i_2} (1 - e_2^{(2)})^{n-s_2} \Pr[A_2(s_2)]$$

$$= \sum_{i_2=0}^{n-s_1} \binom{n - s_1}{i_2} (e_2)^{i_2} H_2(s_2)$$

for $1 \leq s_1 < n$. After continuing to condition N_j in this way, the general recursive equation can be obtained

$$H_r(s_r) = \begin{cases} 1 & \text{for } s_r = n \\ \sum_{i_{r+1}=0}^{n-s_r} \binom{n - s_r}{i_{r+1}} (e_{r+1})^{i_{r+1}} H_{r+1}(s_{r+1}) & \text{for } r \leq s_r < n \\ 0 & \text{for } s_r < r \end{cases}$$

and

$$H_n(s_n) = \begin{cases} 1 & \text{for } s_n = n \\ 0 & \text{for } s_n < n. \end{cases}$$

As a result, the new approach establishes the function F_n in terms of $e_1, \ldots, e_n$ through the functions $H_n, \ldots, H_1$ while the recursive formulas in (1) and (2) express function F_n through the functions $F_1, \ldots, F_{n-1}$.

An algorithm for computing F_n is then obtained as follows:

Step 0. Set $H_i(n) = 1$ and $H_i(i - 1) = 0$ for $i = 1, \ldots, n$.

Step 1. Calculate $H_{n-1}(i) = \sum_{j=0}^{n-i} \binom{n - i}{j} (e_n)^j H_n(i + j)$ for $i = n - 1$.

$\vdots$

Step r. Calculate $H_{n-r}(i) = \sum_{j=0}^{n-i} \binom{n - i}{j} (e_{n-r+1})^j H_{n-r+1}(i + j)$ for $i = n - r, \ldots, n - 1$.

$\vdots$

Step $n - 1$. Calculate $H_1(i) = \sum_{j=0}^{n-i} \binom{n - i}{j} (e_2)^j H_2(i + j)$ for $i = 1, \ldots, n - 1$.

Step n. Calculate $F_n(\mathbf{b}_n) = \sum_{j=1}^{n} \binom{n}{j} (e_1)^j H_1(j)$.

2.3. Practical application

Many studies involve a comparison of the efficacy of treatments with a specified treatment, called the control treatment, in terms of certain response variables. Assume an experiment comparing k $(k \geq 2)$ treatments with one control treatment, where the sample sizes of all k treatments and control treatment are equal to n_i for $i = 1, \ldots, k$ and n_0, respectively. Let the subscripts $1, \ldots, k$ denote the corresponding k treatments and subscript 0 denote the control treatment. Under the standard one-way lay-out of the fixed effect model, the response variable X_{ij} of j-th experimental unit under treatment i is always assumed to be normally distributed with an unknown mean μ_i for $j = 1, \ldots, n_i$ and $i = 0, 1, \ldots, k$ and a common unknown variance σ^2, denoted as $X_{ij} \sim N(\mu_i, \sigma^2)$. Thus the sample treatment means $\bar{X}_i$ and $\bar{X}_0$ are mutually independent and have distributions $N(\mu_i, \sigma^2/n_i)$ and $N(\mu_0, \sigma^2/n_0)$, respectively. Assume that the variance σ^2 is estimated by the pooled sample variance S^2 which is independent to all the sample means and is distributed as $\sigma^2 \chi_\nu^2/\nu$, where $\nu = \sum_{i=0}^{k} n_i - (k + 1)$.

Table 1
Critical values c_m for the DT's step-up procedure in one-sided tests with $\alpha = 0.05$, $\nu = 60$, $\rho = 0.3$

m	c_m	m	c_m	m	c_m	m	c_m	m	c_m	m	c_m
1	1.671	11	2.623	21	2.839	31	2.963	41	3.049	51	3.115
2	1.988	12	2.653	22	2.854	32	2.973	42	3.057	52	3.121
3	2.148	13	2.680	23	2.869	33	2.983	43	3.064	53	3.127
4	2.258	14	2.705	24	2.882	34	3.992	44	3.071	54	3.133
5	2.341	15	2.729	25	2.895	35	3.001	45	3.078	55	3.138
6	2.408	16	2.750	26	2.908	36	3.010	46	3.084	56	3.143
7	2.464	17	2.770	27	2.920	37	3.018	47	3.091	57	3.149
8	2.512	18	2.789	28	2.931	38	3.026	48	3.097	58	3.154
9	2.554	19	2.807	29	2.942	39	3.034	49	3.103	59	3.159
10	2.590	20	2.823	30	2.953	40	3.042	50	3.109	60	3.164

Denote $\theta_j = \mu_j - \mu_0$ for $j = 1, \ldots, k$ to be the difference between the j-th treatment and the control treatment. Assume that a larger value of θ_j means that the j-th treatment is more effective. Consider the testing of the one-sided null hypotheses $H_i : \theta_i = 0$ against the alternatives $H_i' : \theta_i > 0$, for $i = 1, 2, \ldots, k$. Then, under the null hypotheses the pivotal random variables

$$T_i = \frac{\bar{X}_i - \bar{X}_0}{S\sqrt{1/n_i + 1/n_0}}$$

for $i = 1, 2, \ldots, k$ have a joint k-variate t-distribution with ν degrees of freedom and the correlation matrix $\{\rho_{ij}^{(k)} = \rho_i\rho_j\}$ which is denoted as a matrix with entry $\rho_i\rho_j$ in the i-th row and j-th column for $i \neq j$ and entry 1 for $i = j$, where $\rho_i = \sqrt{n_i/(n_i + n_0)}$ and $1 \leq i, j \leq k$. Let t_i be the observed values of the corresponding random variables T_i for $i = 1, \ldots, k$. The details of the step-up multiple test are described as follows:

(i) Denote $t_{(1)} \leq \cdots \leq t_{(k)}$ as the ordered values of the t_i test statistics and $H_{(1)}, \ldots, H_{(k)}$ as the corresponding hypothesis.
(ii) A set of k critical values $c_1 \leq \cdots \leq c_k$ is determined such that the Type I familywise error (FWE) is strongly controlled at a prespecified level α (see Hochberg and Tamhane (1987)).
(iii) Compare $t_{(1)}$ with c_1. If $t_{(1)} \geq c_1$, reject all the hypotheses and terminate the test; otherwise accept $H_{(1)}$ and go to the next comparison of $t_{(2)}$ and c_2. In general, the test continues until the first occurrence of $t_{(i)} \geq c_i$, say $i = m$, then the $H_{(1)} \ldots, H_{(m)}$ hypotheses are accepted and $H_{(m+1)}, \ldots, H_{(k)}$ are rejected.

When $n_i = n$ for $i = 1, \ldots, k$, the random variables $T_1, \ldots, T_k$ have, under the null hypotheses, a joint k-variate t-distribution with correlation matrix $\{\rho_{ij}^{(k)} = \rho\}$ where $\rho = n/(n + n_0)$. In order to control the Type I FWE at α level, DT showed that the c_i for the step-up procedure is determined by solving $c_1, c_2, \ldots$ recursively from

$$P[(T_1, \ldots, T_m) < (c_1, \ldots, c_m); \{\rho_{ij}^{(m)} = \rho\}] = 1 - \alpha, \quad m = 1, \ldots, k. \tag{4}$$

Note that $c_1 = t_\nu(\alpha)$, the upper α-percentage point of a t-distribution with ν degrees of freedom. Several studies show that the critical values c_i always satisfy the monotonicity $c_1 < \cdots < c_k$, even though an analytical proof is not available.

Let Z_i for $i = 0, \ldots, k$ be independent standard normal random variables with density ϕ and distribution function Φ. Let U be a $\sqrt{\chi_\nu^2/\nu}$ random variable independent of all the Z_i. Then the probability in (4) for $2 \leq m \leq k$ is given by

$$P[(T_1, \ldots, T_m) < (c_1, \ldots, c_m); \{\rho_{ij}^{(m)} = \rho\}] = \int_0^\infty \int_{-\infty}^\infty J_m \phi(z_0) dz_0 f_\nu(u) du$$

where $f_\nu(u)$ is the density function of U, and $J_m = P[(Z_1, \ldots, Z_m) < (d_1, \ldots, d_m)]$ with $d_i = (c_i u + \sqrt{\rho}z_0)/\sqrt{1-\rho}$ for $i = 1, \ldots, m$.

Based on one of recursive formulas (1), (2), and the proposed algorithm to compute J_m for any $m \geq 2$, the critical values c_m can be determined. In order to compare the performances of the three approaches, three Fortran programs were constructed using these three different approaches for evaluating J_m to determine c_m in this case. All the programs are available from the first author upon request. After extensive numerical study, all three approaches are found to produce same critical values for the case $k \leq 20$. Due to the significant accumulation of the rounding errors for large value of k, Kwong and Liu's (2000) and Finner and Roters (1994) approaches begin to break down when k reaches 21 and 38, respectively. However, the proposed algorithm uniformly has a very stable performance even when k is as large as 60. Therefore, it is concluded that the proposed algorithm provides a more effective and stable approach for evaluating the joint distribution under the iid case when compared with the existing approaches. To illustrate the effectiveness and stability of the proposed algorithm, its program was used to calculate an example of the critical values c_i for one-sided step-up test with $\alpha = 0.05$, $\nu = 60$, $\rho = 0.3$ and $k = 1, \ldots, 60$ and the result is presented in Table 1.

3. Independent but not identically distributed

3.1. Existing algorithm

Let $X_1, \ldots, X_k$ be ibnid random variables with known cumulative distribution functions $G_1, \ldots, G_k$, respectively. For any given constants $d_1, \ldots, d_k$ where $d_1 \leq d_2 \leq \cdots \leq d_k$, the paper now focuses on determining

$$H_{X_1,\ldots,X_k}(d_1, \ldots, d_k) = P[(X_1, \ldots, X_k) < (d_1, \ldots, d_k)].$$

Similar to Lemma 3.1 in DT, an approach proposed by Kwong and Liu (2000) evaluates the function $H_{X_1,\ldots,X_k}(d_1, \ldots, d_k)$ by successively conditioning each of X_i. After first conditioning X_k to fall in the intervals $(-\infty, d_1), [d_1, d_2), \ldots, [d_{k-1}, d_k]$, $H_{X_1,\ldots,X_k}(d_1, \ldots, d_k)$ can be obtained as

$$H_{X_1,\ldots,X_k}(d_1, \ldots, d_k) = H_{X_1,\ldots,X_{k-1}}(d_2, \ldots, d_k)G_k(d_1) + H_{X_1,\ldots,X_{k-1}}(d_1, d_3, \ldots, d_k)[G_k(d_2) - G_k(d_1)]$$
$$+ \cdots + H_{X_1,\ldots,X_{k-1}}(d_1, \ldots, d_{k-1})[G_k(d_k) - G_k(d_{k-1})]. \tag{5}$$

Then, by successively conditioning one by one from X_{k-1} to X_2 in (5), $H_{X_1,\ldots,X_k}(d_1, \ldots, d_k)$ can be expressed in terms of G_i for $i = 1, \ldots, k$. Therefore, similar to the algorithm given in DT for the iid case, an algorithm for computing $H_{X_1,\ldots,X_k}(d_1, \ldots, d_k)$ is then obtained by Kwong and Liu (2000) as follows:

Step 1. Calculate $H_{X_1}(d_h) = G_1(d_h)$, for $h = 1, \ldots, k$, a total of k terms.

Step 2. Calculate $H_{X_1,X_2}(d_h, d_i) = H_{X_1}(d_i)G_2(d_h) + H_{X_1}(d_h)[G_2(d_i) - G_2(d_h)]$, for $1 \leq h < i \leq k$, a total of $\binom{k}{2}$ terms.

Step 3. Calculate $H_{X_1,X_2,X_3}(d_h, d_i, d_j) = H_{X_1,X_2}(d_i, d_j)G_3(d_h) + H_{X_1,X_2}(d_h, d_j)[G_3(d_i) - G_3(d_h)] + H_{X_1,X_2}(d_h, d_i)[G_3(d_j) - G_3(d_i)]$, for $1 \leq h < i < j \leq k$, a total of $\binom{k}{3}$ terms.

$\vdots$

Step k. Calculate $H_{X_1,\ldots,X_k}(d_1, \ldots, d_k) = H_{X_1,\ldots,X_{k-1}}(d_2, \ldots, d_k)G_k(d_1) + H_{X_1,\ldots,X_{k-1}}(d_1, d_3, \ldots, d_k)[G_k(d_2) - G_k(d_1)] + \cdots + H_{X_1,\ldots,X_{k-1}}(d_1, \ldots, d_{k-1})[G_k(d_k) - G_k(d_{k-1})]$.

3.2. Modified algorithm

The major drawback of the existing algorithm discussed in Section 3.1 is the multi-dimensional array function H obtained in each step. For example, in step 6 of the algorithm with $k = 8$, the function H is a six-dimensional array with the total number of elements $8^6 = 262{,}144$. However, in step 7 of the algorithm, the function H is a seven-dimensional array and the total number of elements increases to $8^7 = 2{,}097{,}152$. Therefore, the storage spaces of all the elements in the function H increases exponentially with k in each step of the algorithm. As a result, the algorithm is only workable for $k \leq 8$.

In order to address this issue, the existing algorithm is modified by reducing the multi-dimensional array function H to an one-dimensional array function. Define the subsets $\mathcal{L}_m = \{l_1, l_2, \ldots, l_m\} \subset \{1, 2, \ldots, k\}$ with cardinality m for $m = 1, \ldots, k$, where $1 \leq l_1 < l_2 < \cdots < l_m \leq k$. Note that there are $n_m = \binom{k}{m}$ distinct subsets, say $S_1^{(m)}, S_2^{(m)}, \ldots, S_{n_m}^{(m)}$ in $\mathcal{L}_m$, where $S_j^{(m)} = \{l_1^{(j)}, l_2^{(j)}, \ldots, l_m^{(j)}\}$ for $j = 1, \ldots, n_m$. Let the function

$$I(S_j^{(m)}) = \sum_{i=1}^{m} 2^{l_i^{(j)}-1} \tag{6}$$

for $m = 1, 2, \ldots, k$ and $j = 1, \ldots, n_m$. It is straightforward to show that the function I is an one-to-one transformation of subsets $S_j^{(m)}$ into a positive integer space $\{1, 2, \ldots, 2^k - 1\}$.

As the result of (6), the multi-dimensional array function H can be uniquely transformed to the one-dimensional array function F after setting

$$H_{X_1,\ldots,X_m}(d_{l_1^{(j)}}, \ldots, d_{l_m^{(j)}}) = F[I(S_j^{(m)})] = F\left[\sum_{i=1}^{m} 2^{l_i^{(j)}-1}\right]$$

for $m = 1, 2, \ldots, k$ and $j = 1, \ldots, n_m$. Note that the required probability $H_{X_1,\ldots,X_k}(d_1, \ldots, d_k)$ is just equal to $F[I(S_1^{(k)})]$. In order to determine all the subsets $\mathcal{L}_m$, for $m = 1, \ldots, k$, the inverse function of I is derived. For any given integer $N \in \{1, 2, \ldots, 2^k - 1\}$ expressed in the form of $\sum_{i=1}^{m} 2^{l_i-1}$, the inverse function of I is defined as $I^{-1}(N) = I^{-1}(\sum_{i=1}^{m} 2^{l_i-1}) = \{l_1, \ldots, l_m\}$, e.g. $I^{-1}(45) = I^{-1}(2^0 + 2^2 + 2^3 + 2^5) = \{1, 3, 4, 6\}$.

After incorporating function I and its inverse function I^{-1} into Kwong and Liu's (2000) algorithm, a modified algorithm is now proposed as follows:

Step 0. Use the inverse function I^{-1} to identify all the subsets $\mathcal{L}_i$ for $i = 1, \ldots, k$.

Step 1. Calculate $F[I(S_j^{(1)})] = G_1(d_{l_1^{(j)}})$ for all $S_j^{(1)} \in \mathcal{L}_1$.

Table 2

Values of functions $H_{X_1,\ldots,X_k}(d_1,\ldots,d_k)$ where $X_1,\ldots,X_k$ have independent normal distributions with $\mu_i = 0.01i$ and $\sigma_i = 1 + 0.01i$ for $i = 1,\ldots k$, respectively and $d_i = 1.6 + 0.05i$ for $i = 1,\ldots,k$

k	$H_{X_1,\ldots,X_k}(d_1,\ldots,d_k)$	k	$H_{X_1,\ldots,X_k}(d_1,\ldots,d_k)$
9	0.7697190	19	0.7695573
10	0.7637749	20	0.7736962
11	0.7597622	21	0.7781366
12	0.7573876	22	0.7828088
13	0.7564033	23	0.7876531
14	0.7565950	24	0.7926183
15	0.7577794	25	0.7976596
16	0.7597968	26	0.8027388
17	0.7625098	27	0.8078243
18	0.7657977	28	0.8128865

Step 2. Calculate $F[I(S_j^{(2)})] = F[I(S_j^{(2)}) - 2^{l_1^{(j)}-1}]G_2(d_{l_1^{(j)}}) + F[I(S_j^{(2)}) - 2^{l_2^{(j)}-1}][G_2(d_{l_2^{(j)}}) - G_2(d_{l_1^{(j)}})]$ for all $S_j^{(2)} \in \mathcal{L}_2$.

Step 3. Calculate $F[I(S_j^{(3)})] = F[I(S_j^{(3)}) - 2^{l_1^{(j)}-1}]G_3(d_{l_1^{(j)}}) + F[I(S_j^{(3)}) - 2^{l_2^{(j)}-1}][G_3(d_{l_2^{(j)}}) - G_3(d_{l_1^{(j)}})] + F[I(S_j^{(3)}) - 2^{l_3^{(j)}-1}][G_3(d_{l_3^{(j)}}) - G_3(d_{l_2^{(j)}})]$ for all $S_j^{(3)} \in \mathcal{L}_3$.

$\vdots$

Step k. Calculate $F[I(S_1^{(k)})] = F[I(S_1^{(k)}) - 1]G_k(d_1) + F[I(S_1^{(k)}) - 2^1][G_k(d_2) - G_k(d_1)] + \cdots + F[I(S_1^{(k)}) - 2^{k-1}][G_k(d_k) - G_k(d_{k-1})]$.

In contrast to the existing program, a Fortran program constructed under the modified algorithm can easily evaluate the joint distributions of the independent but not identically random variables with $k \leq 28$. For the purpose of illustration, the program is now used to evaluate functions H for $k = 9, \ldots, 28$. Let $X_1, \ldots, X_k$ be independent normal distributions with means $\mu_i = 0.01i$ and standard deviations $\sigma_i = 1 + 0.01i$ for $i = 1, \ldots k$, respectively. Assume that $d_i = 1.6 + 0.05i$ for $i = 1, \ldots, k$. The values of functions $H_{X_1,\ldots,X_k}(d_1,\ldots,d_k)$ are determined by the program and the results are presented in Table 2.

3.3. Practical application

Similar to the multiple hypothesis testing set-up stated in Section 2.3, the simultaneous tests of the k null hypotheses in the direction-mixed family are:

$$H_i : \theta_i = 0$$

for $i = 1, \ldots, k$ versus r one-sided alternative hypotheses

$$H_i' : \theta_i > 0$$

for $i = 1, \ldots, r$ and $(k - r)$ two-sided alternative hypotheses

$$H_j' : \theta_j \neq 0$$

for $j = r+1, \ldots, k$. With respect to the family of null hypotheses $\{H_1, \ldots, H_k\}$, a subset $\{H_1, \ldots, H_r\}$ $(r \leq k)$ is tested against one-sided alternatives, while the remaining null hypotheses $\{H_{r+1}, \ldots, H_k\}$ are tested against two-sided alternatives.

To test the k null hypotheses simultaneously in the direction-mixed family, the test statistics are $T_1, \ldots, T_r, |T_{r+1}|, \ldots, |T_k|$, respectively. As mentioned in Section 2.3, the variates $T_1, T_2, \ldots, T_k$ have a multivariate t-distribution with correlation matrix $\{\rho_{ij}^{(k)} = \rho_i \rho_j\}$ in the unbalanced design. Let t_i and $|t_j|$ be the observed test statistics of the corresponding random variables T_i for $i = 1, \ldots, r$ and $|T_j|$ for $j = r + 1, \ldots, k$, respectively. Then, the observed p-values for t_i and $|t_j|$ are $p_i = P(T_i > t_i)$ for $i = 1, \ldots, r$ and $p_j = P(|T_j| > |t_j|)$ for $j = r + 1, \ldots, k$, respectively. Assume that the set of ordered observed p-values from the test statistics t_i and $|t_j|$ are $p_{(1)} \leq \cdots \leq p_{(k)}$.

Similarly to DT's step-up procedure, the p-value consistent step-up procedure proposed by Kwong et al. (2007) determines k critical p-values, $p_k^u \geq p_{k-1}^u \geq \cdots \geq p_1^u$, for comparisons with k ordered observed p-values. The algorithm to conduct the procedure is as follows:

1. Obtain the set of ordered observed p-values, $p_{(1)} \leq \cdots \leq p_{(k)}$ with corresponding hypotheses $H_{(1)}, \ldots, H_{(k)}$, respectively.
2. Compare $p_{(k)}$ with p_k^u. If $p_{(k)} < p_k^u$, then reject all hypotheses and terminate the procedure; otherwise accept $H_{(k)}$ and proceed to the next step of comparing $p_{(k-1)}$ with p_{k-1}^u.
3. The testing procedure continues until the first occurrence of $p_{(i)} < p_i^u$, say $i = s \leq k$. Then, $H_{(s+1)}, \ldots, H_{(k)}$ are accepted while the remaining hypotheses $H_{(1)}, \ldots, H_{(s)}$ are rejected. If $p_{(i)} \geq p_i^u$ for $i = 1, \ldots, k$, then all the hypotheses are accepted.

Let $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$ and $\boldsymbol{\theta}_s \subset \boldsymbol{\theta}$ be a set of vectors that have elements $\theta_l = 0$ for $l \in \mathcal{S} = \{l_1, \ldots, l_s\}$, where $\{l_1, \ldots, l_s\} \subseteq \{1, \ldots, k\}$ and $\theta_l > 0$ for $l \notin \mathcal{S}$. To control the FWE at α, the critical values of the step-up procedure must be determined such that

$$P_{\boldsymbol{\theta}_s} \left[\text{Accept } H_{l_1}, \ldots, H_{l_s} \right] \geq 1 - \alpha$$

for $s = 1, \ldots, k$. According to the arguments given by Liu (1996) for establishing a step-up procedure, the least favorable parameter vector of $\boldsymbol{\theta}_s$ is $\boldsymbol{\theta}_s^*$ which has elements $\theta_l = 0$ for $l \in \mathcal{S}$ and $\theta_l \to \infty$ for $l \notin \mathcal{S}$. Therefore, in order to control the FWE at α, the critical p-values of the step-up procedure are determined by solving $p_k^u, p_{k-1}^u, \ldots, p_1^u$ recursively in this sequence from the following equation:

$$\min_{1 \leq l_1 < \cdots < l_s \leq k} P_{\boldsymbol{\theta}_s^*} \left[(P_{l_1}, \ldots, P_{l_s}) > (p_{k-s+1}^u, \ldots, p_k^u) \right] = 1 - \alpha \tag{7}$$

for $s = 1, \ldots, k$, where $P_{l_1}, \ldots, P_{l_s}$ are the p-values corresponding to true hypotheses $H_{l_1}, \ldots, H_{l_s}$, respectively, and the minimum is over all subsets $\{l_1, \ldots, l_s\} \subset \{1, \ldots, k\}$ with cardinality s. It is noted that $p_k^u = \alpha$.

Since directly inverting the p-values to the corresponding test statistics is not straightforward in (7), Kwong et al. (2007) modify the event $(P_{l_1}, \ldots, P_{l_s}) > (p_{k-s+1}^u, \ldots, p_k^u)$ in (7) before applying the procedure of inversion. For ease of exposition, only the computation of p_s^u is illustrated after assuming that $p_{s+1}^u, \ldots, p_k^u$ have already been obtained. Without loss of generality, assume that $P_{l_1}, \ldots, P_{l_w}$ and $P_{l_{w+1}}, \ldots, P_{l_s}$ are the p-values corresponding to true one-sided hypotheses $H_{l_1}, \ldots, H_{l_w}$ and true two-sided hypotheses $H_{l_{w+1}}, \ldots, H_{l_s}$, respectively.

By applying the law of total probability to P_{l_1} which has to fall into one of the intervals $(p_{k-s+1}^u, p_{k-s+2}^u)$, $(p_{k-s+2}^u, p_{k-s+3}^u)$, $\ldots, (p_k^u, \infty)$, the probability in (7) becomes

$$
\begin{aligned}
P_{\boldsymbol{\theta}_s^*} \left[(P_{l_1}, \ldots, P_{l_s}) > (p_{k-s+1}^u, \ldots, p_k^u) \right] &= P_{\boldsymbol{\theta}_s^*} \left[(P_{l_2}, \ldots, P_{l_s}) > (p_{k-s+2}^u, \ldots, p_k^u) \bigcap (p_{k-s+1}^u < P_{l_1} < p_{k-s+2}^u) \right] \\
&+ P_{\boldsymbol{\theta}_s^*} \left[(P_{l_2}, \ldots, P_{l_s}) > (p_{k-s+1}^u, p_{k-s+3}^u, \ldots, p_k^u) \bigcap (p_{k-s+2}^u < P_{l_1} < p_{k-s+3}^u) \right] + \cdots \\
&+ P_{\boldsymbol{\theta}_s^*} \left[(P_{l_2}, \ldots, P_{l_s}) > (p_{k-s+1}^u, \ldots, p_{k-1}^u) \bigcap (p_k^u < P_{l_1}) \right].
\end{aligned}
\tag{8}
$$

After successively applying the law of total probability from P_{l_2} to P_{l_s} in (8), all the p-values can then be inverted to the corresponding test statistics, i.e. the event $\{a < P_{l_i} < b\}$ for $i = 1, \ldots, s$ becomes $\{F^{-1}(1 - b) < T_{l_i} < F^{-1}(1 - a)\}$ for $i = 1, \ldots, w$ and $\{F^{-1}(1 - b/2) < T_{l_i} < F^{-1}(1 - a/2)\}$ for $i = w + 1, \ldots, s$, respectively, where F^{-1} is the inverse cumulative function of t-distribution with ν degrees of freedom. As the result, the probability in (7) can be evaluated with application of modified algorithm given Section 3.2. The modified FORTRAN program to evaluate the critical p-values of the p-value consistent step-up procedure is available from the first author upon request. For illustration, some examples of critical p-values with $k = 10, 11, 12, 13, 14$ for the p-value consistent step-up procedure are calculated by the program and the results are presented in Table 3. Contrast to the old program which only works for $k \leq 8$, the modified program under new algorithm can deal with the cases with $k \leq 14$. Therefore, our modified algorithm can significantly extend the p-value consistent step-up procedure to higher-dimensional multiple testing comparisons.

4. Final remarks

Existing algorithms, which have the dimensional restrictions of computing the joint distributions of k order statistics, are only workable for relatively small values of k. In order to evaluate the cases with the larger value of k, this paper proposes two new algorithms: one for the iid case and the other for ibnid case.

In the iid case, the new algorithm based on the multinomial approach provides a more stable procedure for evaluating the joint distribution by reducing the impact of rounding errors generated due to recursion in those existing algorithms. For the purpose of illustration, the discussed algorithms are applied to evaluate the critical values of the DT's step-up procedure under balanced designs. As stated in Section 2.3, the proposed algorithm is able to determine the critical values up to the cases $k = 60$ while the existing algorithms break down for $k > 38$.

In the ibnid case, the existing algorithm has the shortcoming of requiring a multi-dimensional array function to store information during the recursion procedure and is therefore constructed for $k \leq 8$. The modified algorithm is exactly like the existing algorithm except that the required information storage is significantly reduced during the evaluation process by transforming the multi-dimensional array function to a one-dimensional array function. As the result, it can be constructed to evaluate the cases with $k \leq 28$ as discussed in Section 3.2. For illustration, in Section 3.3, the modified algorithm is used to calculate the critical p-values of the p-value consistent step-up procedure in the direction-mixed family under unbalanced designs for $k = 10, \ldots, 14$. In conclusion, the proposed algorithms are comparable with the existing algorithms for small k in terms of computational time, and can effectively extend the applications of most step-up procedures to handle far higher dimensions.

Table 3
Critical p-values for the p-value consistent step-up procedure at $\alpha = 0.05$

k	r	$(n_0, n_1, \ldots, n_r, n_{r+1}, \ldots n_k)$	p_i^u
10	4	(12, 7, 9, 10, 12, 12, 14, 15, 16, 20, 22)	$p_1^u = 0.006730$ $p_2^u = 0.007238$ $p_3^u = 0.007892$ $p_4^u = 0.008791$ $p_5^u = 0.009984$ $p_6^u = 0.011625$ $p_7^u = 0.014079$ $p_8^u = 0.018028$ $p_9^u = 0.025666$ $p_{10}^u = 0.050000$
11	5	(10, 6, 10, 11, 13, 10, 7, 10, 11, 19, 22, 24)	$p_1^u = 0.006387$ $p_2^u = 0.006778$ $p_3^u = 0.007279$ $p_4^u = 0.007958$ $p_5^u = 0.008793$ $p_6^u = 0.009946$ $p_7^u = 0.011585$ $p_8^u = 0.014007$ $p_9^u = 0.017929$ $p_{10}^u = 0.025557$ $p_{11}^u = 0.050000$
12	4	(16, 6, 8, 11, 11, 12, 15, 16, 17, 20, 22, 16, 14)	$p_1^u = 0.005357$ $p_2^u = 0.005712$ $p_3^u = 0.006148$ $p_4^u = 0.006703$ $p_5^u = 0.007403$ $p_6^u = 0.008303$ $p_7^u = 0.009509$ $p_8^u = 0.011165$ $p_9^u = 0.013598$ $p_{10}^u = 0.017602$ $p_{11}^u = 0.025418$ $p_{12}^u = 0.050000$
13	9	(8, 8, 9, 10, 11, 12, 19, 15, 18, 20, 22, 9, 15, 23)	$p_1^u = 0.006496$ $p_2^u = 0.006785$ $p_3^u = 0.007132$ $p_4^u = 0.007555$ $p_5^u = 0.008091$ $p_6^u = 0.008750$ $p_7^u = 0.009627$ $p_8^u = 0.010769$ $p_9^u = 0.012371$ $p_{10}^u = 0.014786$ $p_{11}^u = 0.018620$ $p_{12}^u = 0.025997$ $p_{13}^u = 0.050000$
14	4	(10, 8, 10, 12, 14, 9, 11, 13, 15, 17, 19, 21, 22, 23, 24)	$p_1^u = 0.005666$ $p_2^u = 0.005910$ $p_3^u = 0.006197$ $p_4^u = 0.006540$ $p_5^u = 0.006956$ $p_6^u = 0.007483$ $p_7^u = 0.008169$ $p_8^u = 0.009049$ $p_9^u = 0.010217$ $p_{10}^u = 0.011848$ $p_{11}^u = 0.014244$ $p_{12}^u = 0.018215$ $p_{13}^u = 0.025767$ $p_{14}^u = 0.050000$

Acknowledgements

References

Dalal, S.R., Mallows, C.L., 1992. Buying with exact confidence. Ann. Appl. Probab. 2, 752–765.

Dunnett, C.W., Tamhane, A.C., 1992. A step-up multiple test procedure. J. Amer. Stat. Assoc. 87, 162–170.

Dunnett, C.W., Tamhane, A.C., 1995. Step-up multiple testing of parameters with unequally correlated estimates. Biometrics 51, 217–227.

Dunnett, C.W., Tamhane, A.C., 1997. Multiple testing to establish superiority/equivalence of a new treatment compared with k standard treatments. Statist. Med. 16, 2489–2506.

Finner, H., Roters, M., 1994. On the limit behaviour of the joint distribution function of order statistics. Ann. Inst. Statist. Math. 46, 343–349.

Hochberg, Y., Tamhane, A.C., 1987. Multiple Comparison Procedures. Wiley, New York.

Kwong, K.S., 2001a. A modified Dunnett and Tamhane step-up approach for establishing superiority/equivalence of a new treatment compared with k standard treatments. J. Statist. Plann. Inference 97, 359–366.

Kwong, K.S., 2001b. An algorithm for construction of multiple hypothesis testing. Comput. Stat. 16, 165–171.

Kwong, K.S., Cheung, S.H., Chan, W.S., 2004. Multiple testing to establish superiority/equivalence of a new treatment compared with k standard treatments for unbalanced designs. Biometrics 60, 491–498.

Kwong, K.S., Cheung, S.H., Holland, B., Wang, Y., 2007. Three p-value consistent procedures for multiple comparisons with a control in direction-mixed families. Statist. Med. 26, 4253–4272.

Kwong, K.S., Liu, W., 2000. Calculation of critical values for Dunnett and Tamhane's step-up multiple test procedure. Statist. Probab. Lett. 49, 411–416.

Liu, W., 1996. Multiple tests of a non-hierarchical finite family of hypotheses. J. Roy. Statist. Soc. Ser. B 58, 455–461.

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

Ma, C., 1997. Saddlepoint approximations for multinomial distributions with applications to the joint distribution of order statistics. J. Stat. Comput. Simul. 59, 33–45.

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