HIROSHIMA MATH. J.
28 (1998), 55–93

Selection problems based on ranked data

Takashi MATSUI

(Received January 9, 1997)

ABSTRACT. The purpose of this paper is to investigate some topics on the selection problems based on the vector and the combined ranks, with special reference to the structures of population parameters. We pay more attention to the vector rank statistics. First, by obtaining the joint distribution of the rank sums, exact results on the problems of selecting the best population based on the vector rank are given. Then, we give asymptotic results, for both the vector rank and the combined rank cases, using their respective moments. We put the main emphasis on the fact that these results are given with reference to the distributions (parameters) of the underlying populations. One of the open questions of the selection problems based on the ranked data lies in the determination of the *LFC*, though this problem has been discussed in several places. Thus, finally, under the assumption of the parametric configuration, some asymptotic results on *LFC* are obtained.

1. Introduction

In the analysis of experimental data, there are so many occasions to test the significance of k treatments. Analysis of variance technique is one of the statistical methods to cope with such situations. The problem is that even though we have an analytical result such as certain hypothesis being significant (or not significant), this may not be necessarily satisfactory to the experimenters. R. E. Bechhofer [3] states, in his first pioneering paper relating to ranking and selection, as follows: "Thus in an agricultural problem the hypothesis that several essentially *different* varieties of grain have the *same* (population) mean yield is an unrealistic one since it is obvious that if the varieties actually are different, the (population) mean yields also be different, and a sufficiently large sample will establish this fact at any preassigned level of significance. Moreover, should a significant result be obtained, the experimenter's problem usually have just begun. For having established that the varieties are different he may now desire to select the one which is 'best'. Here the best variety might be defined as the one having the *largest* (population)

1991 *Mathematics Subject Classifications.* 62F07, 62G30.
Key words and phrases. Combined rank, indifference zone, selection problem, subset selection, vector rank

mean yield. Whenever the experimenter ultimately is faced by the prospect of having to choose a *best* variety, it seems reasonable that the experiment should have been designed with this outcome in mind."

This thought motivated the development of variety of multiple comparison (decision) procedures (see for example Gupta and Berger [9], Hsu [16]) and among which, ranking (ordering) and selection procedures, proposed first by Bechhofer, played very important roles. These procedures were proposed under the so called indifference zone formulation. Soon after the Bechhofer's proposal, Gupta proposed some selection procedures under the subset selection formulation (explicitly in [14], originally in his Thesis).

Bechhofer treated first the selection of t-best populations with respect to the population means [3] and population variances [4] of normal populations. Since then, varieties of methodologies and algorithms have been proposed for ranking and selection problems. Here we cite several keywords which illustrate the spread of ranking and selection problems—with respect to distributions (such as normal, binomial, exponential), with respect to parameters (such as location, scale), with respect to the selection of the best or worst, sequential method, two- or multi-stage procedures and so forth. Details of these spreads can be seen through the books by Bechhofer, Kiefer and Sobel [5], Gibbons, Olkin and Sobel [7] and Gupta and Panchapakesan [13]. Also we have a categorized guide of this field due to Dudewicz and Koo [6]. In [15], Gupta and Panchapakesan indicate and discuss some directions for future research in this field.

There are two streams of procedures based on the ranks, for the selection of populations with distributions having the location or scale parameters. One is based on the combined ranks (Wilcoxon type rank) and the other vector ranks (Friedman type rank). In the indifference zone approach, the first result using a statistic based on the combined ranks appears in Lehmann [18]. He obtained an asymptotic result under the assumption that the slippage configuration of population parameters is the least favorable configuration (*LFC*). While Rizvi and Woodworth [28] showed that the assumption is incorrect by giving a counter example. Further, under another assumptions, Puri and Puri [26], [27], Alam and Thompson [1] have given some relevant results.

The subset selection approach based on the combined and the vector ranks has been extensively studied by Gupta and McDonald [11], [12] and McDonald [24], [25].

The indifference zone approach based on the vector ranks was investigated from the exact theoretical points of view by Matsui [19]. He also treated the selection problem under the assumption of the slippage configuration of parameters to be *LFC* [20], but the assumption is not generally correct, as is shown by Lee and Dudewicz [17].

In most of the selection procedures, slippage configuration and equi-parameter configuration of population parameters play an important role as the parametric configurations giving the minimum of the probability of correctly selecting the target population(s). This is the so called least favorable configuration. The difficult part of the problem is how to obtain such a configuration.

In this paper we investigate the exact properties of selection procedures based on the vector rank, for both the indifference zone and the subset selection formulations, with paying attentions to the structure of population parameters. Our attention is also given to the structure of the least favorable configuration of population parameters. Further, using the same type of assumptions as Lehmann [18], behaviors of two procedures based on the combined and the vector ranks are investigated asymptotically for both the indifference zone and the subset selection formulations.

In Section 2, the distribution of the vector ranks in relation to the underlying distribution of the population is given. Applications to some main distributions are also given. In Section 3, exact results of selection procedures are given under both the indifference zone and the subset selection formulations. In general, the problem of determining the least favorable configuration on population parameters is an open question, for selection procedures based on the ranks. So, we attempt to give some relevent results. In Section 4, we give some preliminary results on the *PCS* using exact moments in the framework of slippage configuration. Finally in Section 5, asymptotic results of selection procedures based on the vector and the combined ranks are presented, under the given parametric assumption, for both the indifference zone and the subset selection formulations.

2. Distribution of the rank sum statistics

The purpose of this section is, first to explain the features of the vector ranks with reference to the underlying distributions of populations, then to give the distribution of the sums of the vector ranks.

2.1. Distribution

Let $\prod_1, \prod_2, \ldots, \prod_k$ denote k given independent populations. Suppose that $\prod_i$, $i = 1, 2, \ldots, k$ has a continuous cumulative distribution function (c.d.f.) $F_i(x) = F(x;\theta_i)$ specified by parameter θ_i. The probability density function (p.d.f.) of $F_i(x)$ is denoted by $f_i(x)$ or $f(x;\theta_i)$.

In this paper, we use two types of ranking systems to organize the statistics of concern. Since there are no popular wording to these ranking systems, we start to define one of the ranking systems "vector ranks". Another definition (combined ranks) is given in Section 4.1. Note that this wording is not necessarily used in general, but can be seen in Gupta and McDonald [12].

DEFINITION 2.1. Let $\boldsymbol{X}_j = (X_{1j}, X_{2j}, \dots, X_{kj})$, $j = 1, 2, \dots, n$ be n independent vectors of random variables corresponding to the j-th observation of the k populations $\prod_1, \prod_2, \dots, \prod_k$. For each vector $\boldsymbol{X}_j$, form a new vector of rank order statistics

$$R(\boldsymbol{X}_j) = (R(X_{1j}), R(X_{2j}), \dots, R(X_{kj})), \tag{2.1}$$

where for each j, $R(X_{ij}) = r$ if X_{ij} is the r-th smallest among the k components of $\boldsymbol{X}_j$, $i = 1, 2, \dots, k$, $j = 1, 2, \dots, n$. Then $R(\boldsymbol{X}_j)$ is called a vector rank.

The rank defined in this way is known as Friedman's ranks in a $n \times k$ table. The raws and columns indicate blocks and treatments, respectively. Such a rank occurs when n independent judges orders k populations according to some criteria of classification. Ordeing may be given to the objects, categories, etc. and sometimes useful in reducing the effects of extreme values or unknown (or lost) data with size information. Ranking processes can also be applied to some preference or sensory type of data and in such type of data, observations may be supposed to be distributed according to some continuous distribution.

Now, consider the rank sum T_i based on the n independent observations of populations $\prod_i$, $i = 1, 2, \dots, k$,

$$T_i = \sum_{j=1}^{n} R(X_{ij}). \tag{2.2}$$

We first consider the joint distribution of $T_1, T_2, \dots, T_k$ in connection with the distributions $F_i(x)$'s or parameter θ_i's, $i = 1, 2, \dots, k$. The approach by the joint distribution was proposed by Matsui [19]. Here we express the result in a general way, by introducing some vector and matrix notations.

For any given i, let $(m_1, \dots, m_i, \dots, m_k)$ be a permutation of numbers 1 through k, with m_i fixed to be k $(1 \le m_j \le k-1, j = 1, 2, \dots, k; j \ne i)$. Order $s \equiv (k-1)!$ such permutations lexicographically, namely $(m_1, m_2, \dots, m_k) < (m'_1, m'_2, \dots, m'_k)$ when $m_1 = m'_1$, $m_2 = m'_2, \dots, m_{r-1} = m'_{r-1}$ and $m_r < m'_r$ for some r. Denote the j-th vector in this order by

$$\boldsymbol{r}_{ij} = (m_{1j}, \dots, m_{i-1,j}, k, m_{i+1,j}, \dots, m_{kj})', \quad i = 1, 2, \dots, k, \quad j = 1, 2, \dots, s. \tag{2.3}$$

For example, when $k = 3$,

$$\boldsymbol{r}_{11} = (3,1,2)', \quad \boldsymbol{r}_{12} = (3,2,1)', \quad \boldsymbol{r}_{21} = (1,3,2)',$$
$$\boldsymbol{r}_{22} = (2,3,1)', \quad \boldsymbol{r}_{31} = (1,2,3)', \quad \boldsymbol{r}_{32} = (2,1,3)'.$$

By using these vectors, let us define the basic probabilities p_{kj}, $p_{k-1,j}, \dots, p_{1j}$, $(j = 1,2,\dots,s)$ as

(2.4) $$p_{ij} = \Pr(R(\boldsymbol{X}) = \boldsymbol{r}_{ij}), \quad i = 1,2,\dots,k, \quad j = 1,2,\dots,s,$$

where $R(\boldsymbol{X})$ is the vector of rank order statistics for any observation $\boldsymbol{X} = (X_1, X_2, \dots, X_k)'$.

By introducing the above basic probabilities p_{ij}'s, $(i = 1,2,\dots,k;$ $j = 1,2,\dots,s)$, we have the following theorem concerning the joint distribution of rank sums $(T_1, T_2, \dots, T_k)$. This theorem is essential to consider some exact theoretical results on selection procedures based on the vector ranks.

THEOREM 2.1. *In an n repetitions of experiments, let* $T_1, T_2, \dots, T_k$ *be the rank sums defined by* (2.2). *Then the joint distribution of* $\boldsymbol{T} = (T_1, T_2, \dots, T_k)$ *is given as follows.*

(2.5) $$\Pr(\boldsymbol{T} = \boldsymbol{t}) = \sum_{\boldsymbol{n} \in I} n! \prod_{i=1}^{k} \left(\prod_{j=1}^{s} \frac{1}{n_{ij}!} p_{ij}^{n_{ij}} \right).$$

Here $\boldsymbol{t} = (t_1, t_2, \dots, t_k)'$, t_i*'s are integers,* $n \le t_1, t_2, \dots, t_k \le nk$, $\sum_{i=1}^{k} t_i = nk(k+1)/2$, *and I is the index set such that*

(2.6) $$I = \{\boldsymbol{n}; \boldsymbol{K}\boldsymbol{n} = \boldsymbol{t}, \boldsymbol{J}'_{(k!)}\boldsymbol{n} = n\},$$

where $\boldsymbol{K}$ *is a* $k \times k!$ *matrix*

(2.7) $$\boldsymbol{K} = (\boldsymbol{r}_{k1}, \dots, \boldsymbol{r}_{ks}, \boldsymbol{r}_{k-1,1}, \dots, \boldsymbol{r}_{k-1,s}, \dots, \boldsymbol{r}_{11}, \dots, \boldsymbol{r}_{1s}),$$

(2.8) $$\boldsymbol{n} = (n_{k1}, \dots, n_{ks}, n_{k-1,1}, \dots, n_{k-1,s}, \dots, n_{11}, \dots, n_{1s})', \quad k! \times 1$$

and

(2.9) $$\boldsymbol{J}_{(k!)} = (1,1,\dots,1)', \quad k! \times 1.$$

PROOF. This distribution is a kind of multinomial probability but differs in summing up to form rank sums. By using the basic probabilities p_{ij} defined by (2.4), the ranks of any observation vector $\boldsymbol{X} = (X_1, X_2, \dots, X_k)'$ is specified by $\boldsymbol{r}_{ij} = (m_{1j,i}, m_{2j,i}, \dots, m_{kj,i})'$ where $m_{1j,i}, m_{2j,i}, \dots, m_{kj,i}$ takes values 1 through k exclusively and $m_{ij,i} = k$, $j = 1,2,\dots,s$; $l = 1,2,\dots,k$. Denoting the number of occurences of rank vector $R(\boldsymbol{X}) = \boldsymbol{r}_{ij}$ among n repetitions of observations by

n_{ij}, we have the rank sum of X_l to be $m_{lj}n_{ij}$ for given i and $j, i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, s$. Adding for every j and i, we have the rank sum of the l-th population as

$$\sum_{i=1}^{k} \sum_{j=1}^{s} m_{lj,i} n_{ij} = t_l.$$

Thus we have the theorem.

The cases $k = 2$ and 3 are of special interest from both the practical and the capable theoretical viewpoints. In the rest of the paper, distribution functions $F(x; \theta_i), i = 1, 2, \ldots, k$ are supposed to belong to the location or scale parameter family of distributions, i.e., $F(x; \theta_i)$ is expressed as $F(x, \theta_i) = F(x - \theta_i)$ or $F(x; \theta_i) = F(x/\theta_i)$ depending on whether they belong to the location or scale family of distributions. Here, we note that when we are dealing with the scale parameter, ranking should be given to the absolute values of the observations, if necessary. In this case, F_i's should be the distribution functions for the absolute values of variables.

2.2. Exact forms of distribution

By using Theorem 2.1 we attempt to give some exact forms of the vector rank distributions in more reduced expressions. As we mentiond before, some types of the ranked data which are categorized based on sensory preference or empirical knowledge may well be considered to be derived from continuous distributions. We consider the cases when the observations are from populations with normal and exponential distributions, both belong to the location and scale parameter families of distributions.

More precisely, we consider the four seperate cases when populations $\prod_i$ have the following main distributions ($i = 1, \ldots, k$):

A-1: Normal distribution with the location parameters; $N(\theta_i, \sigma^2)$

A-2: Normal distribution with the scale parameters; $N(0, \theta_i^2)$

B-1: Exponential distribution with the location parameters; $E(\theta_i, \sigma)$

B-2: Exponential distribution with the scale parameters; $E(0, \theta_i)$

Here $N(\mu, \sigma^2)$ expresses the normal distribution with location and scale parameters μ and σ. Also, $E(\theta, \sigma)$ expresses exponential distribution with location and scale parameters θ and σ respectively.

First, we will give some notations and preliminary results. Let the probability density function (p.d.f.) and cumulative distribution function (c.d.f.)

of the standard normal distribution $N(0,1)$ denoted by

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \tag{2.10}$$

$$\Phi(x) = \int_{-\infty}^{x} \phi(x)\,dx. \tag{2.11}$$

respectively. Also the upper probability of the bivariate normal distribution with means 0, variances 1 and correlation coefficient ρ is expressed as

$$L(u,v;\rho) = \int_u^\infty \int_v^\infty \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right\} dx\,dy, \tag{2.12}$$

which is called L-function. L-function has the following properties.

$$L(u,v;\rho) = L(v,u;\rho), \tag{2.13}$$

$$L(-u,v;\rho) = 1 - \Phi(v) - L(u,v;-\rho), \tag{2.14}$$

$$L(0,0;\rho) = \frac{1}{4} + \frac{1}{2\pi}\sin^{-1}\rho. \tag{2.15}$$

Now, by the transformation of variables, we have following relations. This lemma is useful to evaluate the joint distribution of $\boldsymbol{T}$ when the underlying population is normal.

LEMMA 2.1. *For any real a_i and w_i, $i = 1,2,\ldots,k-1$, let*

$$I_k = \underbrace{\iint\cdots\int}_{D} \prod_{i=1}^{k} \phi(x_i)\,dx_i, \tag{2.16}$$

where $D = \{(x_1,x_2,\ldots,x_k);\ x_i \le a_i x_k + w_i,\ i = 1,2,\ldots,k-1; -\infty < x_k < \infty\}$. Then the k-fold integral can be expressed as

$$I_k = \underbrace{\iint\cdots\int}_{D_0} \frac{|\boldsymbol{\Sigma}|^{-1/2}}{(2\pi)^{(k-1)/2}} \exp\left(-\frac{1}{2}\boldsymbol{u}'\boldsymbol{\Sigma}^{-1}\boldsymbol{u}\right) d\boldsymbol{u}, \tag{2.17}$$

where $\boldsymbol{u} = (u_1,u_2,\ldots,u_{k-1})'$, $D_0 = \{\boldsymbol{u}; u_i \le w_i, i = 1,2,\ldots,k-1\}$ and $\boldsymbol{\Sigma} = (\sigma_{ij})$, $\sigma_{ii} = 1 + a_i^2$, $i = 1,2,\ldots,k-1$; $\sigma_{ij} = a_i a_j$, $i,j = 1,2,\ldots,k-1$, $i \ne j$.

As a special case of Lemma 2.1, we have the following. This lemma is important to evaluate the PCS function numerically, and used frequently in later section.

LEMMA 2.2. *In the expression* (2.17), *if* $D_0 = \{\boldsymbol{u}; u_i \leq h, i = 1, 2, \dots, k-1\}$ *and* $\Sigma = (\sigma_{ij})$, $\sigma_{ii} = 1$, $i = 1, 2, \dots, k-1$; $\sigma_{ij} = \rho, i, j = 1, 2, \dots, k-1, i \neq j$, *then*

$$I_k = \int \Phi\left(\frac{\sqrt{\rho}x + h}{\sqrt{1-\rho}}\right)^{k-1} \phi(x)\, dx. \tag{2.18}$$

By letting $k = 2$ and 3 in the above lemma, we have the following relations.

$$I_2 = \Phi(c_1), \tag{2.19}$$

$$I_3 = L(-c_1, -c_2; \rho), \tag{2.20}$$

where $c_i = w_i/\sqrt{1 + a_i^2}$, $i = 1, 2$ and $\rho = a_1 a_2/\sqrt{1 + a_1^2}\sqrt{1 + a_2^2}$.

In the same way, we have the following relations for any real $a_1, a_2 \geq 0$,

$$\int_0^\infty \int_0^{a_1 x} \phi(x)\phi(y)\, dy\, dx = \frac{1}{2\pi} \sin^{-1} \rho_1, \tag{2.21}$$

$$\int_0^\infty \int_0^{a_2 x} \int_0^{a_1 x} \phi(x)\phi(y)\phi(z)\, dz\, dy\, dx = \frac{1}{8\pi} \sin^{-1} \rho_2, \tag{2.22}$$

where $\rho_1 = a_1/\sqrt{1 + a_1^2}$ and $\rho_2 = a_1 a_2/\sqrt{1 + a_1^2}\sqrt{1 + a_2^2}$.

The case $k = 2$

In this case, two rank sums T_1 and T_2 are dependent, and their joint distribution can be expressed as

$$\Pr(T_1 = t_1, T_2 = t_2) = \frac{n!}{(t_2 - n)!(2n - t_2)!} p_{21}^{t_2 - n}(1 - p_{21})^{2n - t_2}, \tag{2.23}$$

where $t_1 = 3n - t_2, t_2 = n, n+1, \dots, 2n$ and

$$p_{21} = \Pr(X_2 > X_1) = \int F_1(x)\, dF_2(x)$$

$$= \begin{cases} \displaystyle\int F(x + \theta_2 - \theta_1)\, dF(x), & \text{(location parameter case)} \\ \displaystyle\int_0^\infty F((\theta_2/\theta_1)x)\, dF(x). & \text{(scale parameter case)} \end{cases} \tag{2.24}$$

Again, turning to the four cases above, we have the following reduced forms of p_{21}.

$$\text{(2.25)} \qquad \begin{aligned} &\text{A-1:} \quad p_{21} = \Phi\{(\theta_2 - \theta_1)/\sqrt{2}\sigma\}. \\ &\text{A-2:} \quad p_{21} = (2/\pi)\sin^{-1}\rho, \quad \rho = (\theta_2/\theta_1)/\sqrt{1 + (\theta_2/\theta_1)^2}. \\ &\text{B-1:} \quad p_{21} = 1 - (1/2)\exp\{-(\theta_2 - \theta_1)/\sigma\}. \\ &\text{B-2:} \quad p_{21} = 1 - \theta_1/(\theta_1 + \theta_2). \end{aligned}$$

The case $\boldsymbol{k} = 3$

When $\boldsymbol{k} = 3$, the probability distribution of $\boldsymbol{T}$ is given by (2.5) for $\boldsymbol{t} = (t_1, t_2, t_3)'$ and $\boldsymbol{n} = (n_{31}, n_{32}, n_{21}, n_{22}, n_{11}, n_{12})$ such that $t_1 + t_2 + t_3 = 6n$, $n_{31} + n_{32} + n_{21} + n_{22} + n_{11} + n_{12} = n$, $\boldsymbol{t}$ and $\boldsymbol{n}$ have the relation

$$\text{(2.26)} \qquad \boldsymbol{t} = \begin{pmatrix} 1 & 2 & 1 & 2 & 3 & 3 \\ 2 & 1 & 3 & 3 & 1 & 2 \\ 3 & 3 & 2 & 1 & 2 & 1 \end{pmatrix} \boldsymbol{n}.$$

Further, there are six cases of basic probabilities p_{31} through p_{12}, whose original forms are given by (2.4). For example, p_{31} has the following form.

$$\text{(2.27)} \qquad \begin{aligned} p_{31} &= \Pr(X_3 \geq X_2 \geq X_1) \\ &= \iiint_{x_3 \geq x_2 \geq x_1} f_3(x_3) f_2(x_2) f_1(x_1)\, dx_1\, dx_2\, dx_3. \end{aligned}$$

By evaluating these integrals for the location and the scale parameter cases, we have the following basic probabilities p_{31} through p_{21} for cases A-1 through B-2.

A-1: Normal distribution with location parameters

$$\text{(2.28)} \qquad \begin{aligned} p_{31} &= L(-w_1, -w_2; \rho), & p_{32} &= L(w_1, -w_1 - w_2; \rho), \\ p_{21} &= L(w_2, -w_1 - w_2; \rho), & p_{22} &= L(-w_1, w_1 + w_2; \rho), \\ p_{11} &= L(-w_2, w_1 + w_2; \rho), & p_{12} &= L(w_1, w_2; \rho). \end{aligned}$$

where $w_1 = (\theta_2 - \theta_1)/\sqrt{2}\sigma$, $w_2 = (\theta_3 - \theta_2)/\sqrt{2}\sigma$ and $\rho = -1/2$

A-2: Normal distribution with scale parameters

$$\text{(2.29)} \qquad \begin{aligned} p_{31} &= \frac{2}{\pi}(\sin^{-1}\rho_1 - \sin^{-1}\rho_2), & p_{32} &= \frac{2}{\pi}(\sin^{-1}\rho_3 - \sin^{-1}\rho_4), \\ p_{21} &= \frac{2}{\pi}(\sin^{-1}\rho_5 - \sin^{-1}\rho_6), & p_{22} &= \frac{2}{\pi}(\sin^{-1}\rho_7 - \sin^{-1}\rho_8), \\ p_{11} &= \frac{2}{\pi}(\sin^{-1}\rho_9 - \sin^{-1}\rho_{10}), & p_{12} &= \frac{2}{\pi}(\sin^{-1}p_{11} - \sin^{-1}\rho_{12}), \end{aligned}$$

where

$$\rho_1 = \rho_{11} = -\delta_1/\sqrt{1+\delta_1^2}\sqrt{1+\delta_2^2}, \qquad \rho_2 = -\delta_1/\sqrt{1+\delta_1^2},$$
$$\rho_3 = \rho_7 = -1/\sqrt{1+\delta_1^2}\sqrt{1+\delta_1^2\delta_2^2}, \qquad \rho_4 = -\sqrt{1-\rho_2^2},$$
$$\rho_5 = \rho_9 = -\delta_1\delta_2^2/\sqrt{1+\delta_2^2}\sqrt{1+\delta_1^2\delta_2^2}, \quad \rho_6 = -\delta_1\delta_2/\sqrt{1+\delta_1^2\delta_2^2},$$
$$\rho_8 = -\sqrt{1-\rho_6^2}, \qquad \rho_{10} = -\delta_2/\sqrt{1+\delta_2^2},$$
$$\rho_{12} = -\sqrt{1-\rho_{10}^2},$$

and $\delta_1 = \theta_2/\theta_1, \delta_2 = \theta_3/\theta_2$.

B-1: Exponential distribution with location parameters

(2.30)
$$p_{31} = 1 - \frac{1}{2}e^{-w_1} - \frac{1}{2}e^{-w_2} + \frac{1}{6}e^{-(w_1+2w_2)}, \quad p_{32} = \frac{1}{2}e^{-w_1} - \frac{1}{2}e^{-(w_1+w_2)} + \frac{1}{6}e^{-(w_1+2w_2)},$$
$$p_{21} = \frac{1}{2}e^{-w_2} - \frac{1}{3}e^{-(w_1+2w_2)}, \qquad p_{22} = \frac{1}{6}e^{-(w_1+2w_2)},$$
$$p_{11} = \frac{1}{2}e^{-(w_1+w_2)} - \frac{1}{3}e^{-(w_1+2w_2)}, \qquad p_{12} = \frac{1}{6}e^{-(w_1+2w_2)},$$

where $w_1 = (\theta_2 - \theta_1)/\sigma$ and $w_2 = (\theta_3 - \theta_2)/\sigma$.

B-2: Exponential distribution with scale parameters

(2.31)
$$p_{31} = \frac{\delta_2}{1+\delta_2} - \frac{\delta_2}{1+\delta_2+\delta_1\delta_2}, \qquad p_{32} = \frac{\delta_1\delta_2}{1+\delta_1\delta_2} - \frac{\delta_1\delta_2}{1+\delta_2+\delta_1\delta_2},$$
$$p_{21} = \frac{1}{1+\delta_2} - \frac{1}{1+\delta_2+\delta_1\delta_2}, \qquad p_{22} = \frac{\delta_1}{1+\delta_1} - \frac{\delta_1\delta_2}{1+\delta_2+\delta_1\delta_2},$$
$$p_{11} = \frac{1}{1+\delta_1\delta_2} - \frac{1}{1+\delta_2+\delta_1\delta_2}, \qquad p_{12} = \frac{1}{1+\delta_1} - \frac{\delta_2}{1+\delta_2+\delta_1\delta_2},$$

where $\delta_1 = \theta_2/\theta_1$ and $\delta_2 = \theta_3/\theta_2$.

2.3. Moments of the vector ranks

In this section, we give moments of the vector ranks with reference to the underlying distributions of populations as given in Section 2.1. Since moments of the vector ranks $R(X_{ij})$, $i = 1, 2, \dots, k$; $j = 1, 2, \dots, n$, defined in Definition 2.1 do not depend on j, each repetition of observations, we denote $R(X_{ij})$, the rank of X_{ij}, by R_i. Then we have the following results, which can be seen in

Matsui [22] and play an important role in considering behaviors of selection procedures in later sections.

THEOREM 2.2. *The mean, variance and covariance of the vector rank statistics* $\boldsymbol{R} = (R_1, R_2, \dots, R_k)$ *are given as follows.*

$$E(R_i) = k\int G(x)\,dF_i(x) + \frac{1}{2}, \quad i = 1, 2, \dots, k, \tag{2.32}$$

$$\begin{aligned} Var(R_i) = {} & 2k\int G(x)\,dF_i(x) - 2k\int G(x)F_i(x)\,dF_i(x) - k\int H(x)\,dF_i(x) \\ & + k^2\int G(x)^2\,dF_i(x) - k^2\left(\int G(x)\,dF_i(x)\right)^2 - \frac{1}{12}, \end{aligned} \tag{2.33}$$

$$\begin{aligned} Cov(R_i, R_j) = {} & k\left(2 - \int F_j(x)\,dF_i(x)\right)\int G(x)\,dF_j(x) \\ & + k\left(2 - \int F_i(x)\,dF_j(x)\right)\int G(x)\,dF_i(x) \\ & - \sum_{l=1}^{k}\int F_l(x)\,dF_i(x)\int F_l(x)\,dF_j(x) \\ & - 2k\left(\int G(x)F_i(x)\,dF_j(x) + \int G(x)F_j(x)\,dF_i(x)\right) \\ & + \int F_j(x)\,dF_i(x)\int F_i(x)\,dF_j(x) - 2\int F_i(x)F_j(x)\,dF_i(x) \\ & - 2\int F_i(x)F_j(x)\,dF_j(x) + 1, \end{aligned} \tag{2.34}$$

where

$$G(x) = \frac{1}{k}\sum_{j=1}^{k} F_j(x), \quad H(x) = \frac{1}{k}\sum_{j=1}^{k} F_j(x)^2. \tag{2.35}$$

One of the important implications of this theorem is the close relationship between ranks and parameters stated in the following theorem.

THEOREM 2.3. *If the distribution functions* $F(x; \theta_i)$, $i = 1, 2, \dots, k$ *belong to the stochastically ordered family of distributions, then we have*

$$E(R_i) \geq E(R_j) \text{ if and only if } \theta_i \geq \theta_j. \tag{2.36}$$

PROOF. We have

$$\int F(x;\theta_l)\,dF(x;\theta_i) = 1 - \int F(x;\theta_i)\,dF(x;\theta_l), \quad (i,l = 1,2,\dots,k).$$

Thus, from (2.32) and (2.35), we have the following.

$$E(R_i) - E(R_j) = \int \sum_{l=1}^{k} \{F(x;\theta_j) - F(x;\theta_i)\}\,dF(x;\theta_l). \tag{2.37}$$

Since c.d.f. $F(x;\theta)$ is stochastically ordered, we have $F(x;\theta_i) \le F(x;\theta_j)$ for $\theta_i \ge \theta_j$, thus the theorem follows.

From Theorem 2.3 we can expect that the selection procedures of the best population (with largest parameter, say) based on the ranks can be warranted in appropriate conditions.

In Tables 1 and 2, we will give some values of mean and standard deviation of the rank statistic R_i for the cases A-1 and A-2 with $k = 3$. Some behaviors of the rank statistic can be seen from these tables.

Table 1. Mean and Standard Deviation of Ranks; Normal Population with Location Parameters and $k = 3$

w_1	0.0			1.0			2.0		
w_2	0.0	1.0	2.0	0.0	1.0	2.0	0.0	1.0	2.0
$E(R_3)$	2.00	2.52	2.84	2.26	2.68	2.90	2.42	2.74	2.92
$E(R_2)$	2.00	1.74	1.58	2.26	2.00	1.84	2.42	2.16	2.00
$E(R_1)$	2.00	1.74	1.58	1.48	1.32	1.26	1.16	1.10	1.08
$SD(R_3)$	0.82	0.69	0.42	0.75	0.56	0.32	0.61	0.46	0.28
$SD(R_2)$	0.82	0.75	0.61	0.75	0.67	0.54	0.61	0.54	0.40
$SD(R_1)$	0.82	0.75	0.61	0.69	0.56	0.46	0.42	0.32	0.28

Table 2. Mean and Standard Deviation of Ranks; Normal Population with Scale Parameters and $k = 3$

δ_1	1.0			3.0			5.0		
δ_2	1.0	3.0	5.0	1.0	3.0	5.0	1.0	3.0	5.0
$E(R_3)$	2.00	2.59	2.75	2.30	2.72	2.83	2.37	2.75	2.85
$E(R_2)$	2.00	1.70	1.63	2.30	2.00	1.92	2.37	2.08	2.00
$E(R_1)$	2.00	1.70	1.63	1.41	1.28	1.25	1.25	1.17	1.15
$SD(R_3)$	0.82	0.70	0.58	0.74	0.56	0.46	0.68	0.51	0.42
$SD(R_2)$	0.82	0.70	0.64	0.74	0.62	0.56	0.68	0.56	0.50
$SD(R_1)$	0.82	0.70	0.64	0.61	0.49	0.46	0.49	0.40	0.37

3. Selection problems based on the vector ranks—exact results

In this section, we deal with selection procedures based on the vector ranks. First, we give two formulations of the selection procedures. Then the exact results of the procedures are given by using the distributional results of Section 2.

3.1. Two formulations

Consider k populations $\prod_1, \prod_2, \ldots, \prod_k$ $(k \geq 2)$ where $\prod_i$'s are characterized by the distributions with parameters $\theta_1, \theta_2, \ldots, \theta_k$. Here θ_i, $i = 1, 2, \ldots, k$, take the values of some interval Θ in the real line and denote the parameter space $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$ by Ω. Also, the ordered values of parameters are denoted as $\theta_{[1]} \leq \theta_{[2]} \leq \cdots \leq \theta_{[k]}$, and the population associated with $\theta_{[i]}$ is written as $\prod_{(i)}$, $i = 1, 2, \ldots, k$.

We assume that the ordered relation of populations is determined by the corresponding parameter $\theta_{[i]}$. We say $\prod_{(j)}$ is better than $\prod_{(i)}$ if $i < j$. Then, the population $\prod_{(k)}$ is the best population, and the populations $\prod_{(k-t+1)}, \ldots, \prod_{(k)}$ are called the t-best populations.

Indifference zone formulation

As we mentioned in Section 1, Bechhofer [3] considered the problem of selecting the t-best normal populations. Here we give the so-called indifference zone formulation due to Bechhofer for the case of selecting the t-best populations.

Our goal is to select the t-best populations associated with the parameters $\theta_{[k-t+1]}, \ldots, \theta_{[k]}$. To attain this goal, i.e., to select correctly the target population(s) is called "correct selection" and denoted by CS. Let $\mathscr{R}$ be a selection (decision) rule. The probability of a correct selection by $\mathscr{R}$ is written by $\Pr(CS|\mathscr{R})$, which is abbreviated as PCS.

Further, we specify the following two constants in relation to the probability requirement,

$$(3.1) \qquad \varphi^*_{k-t+1,k-t} \text{ and } P^*, \quad \text{where } 0 < \varphi^*_{k-t+1,k-t} < \infty, \quad 1\Big/\binom{k}{t} < P^* < 1.$$

For a rule $\mathscr{R}$, we impose the probability requirement:

$$(3.2) \qquad \Pr(CS|\mathscr{R}) \geq P^* \text{ whenever } \varphi_{k-t+1,k-t} \geq \varphi^*_{k-t+1,k-t},$$

where $\varphi_{i,j}$ is a distance function relating to two parameters such that $\varphi_{i,j} = \varphi(\theta_{[i]}, \theta_{[j]})$. We try to guarantee the probability $\Pr(CS|\mathscr{R})$ to be greater than P^* over the subspace of the parameters $\Omega_0 = \{\boldsymbol{\theta}; \varphi_{k-t+1,k-t} \geq \varphi^*_{k-t+1,k-t}\}$, and this region of parameters is called the *preference zone*. While if the distance $\varphi_{k-t+1,k-t} < \varphi^*_{k-t+1,k-t}$, then we do not have much concern to select (or to identify) the relevent population—i.e., we dare not to discriminate the population—and the region is called the *indifference zone*.

Our aim is to investigate the behavior of $\Pr(CS|\mathscr{R})$ under the probability requirements as above.

Subset selection formulation

The subset selection formulation proposed by Gupta (see e.g., [14]) is to select s_0 ($s_0 \geq t$) population(s) from k populations $\prod_1, \prod_2, \ldots, \prod_k$, in order that this subset S contain the t-best populations, i.e., the populations corresponding to $\theta_{[k-t+1]}, \ldots, \theta_{[k]}$. Here we note that the size of the subset S is a random variable and this formulation has the aim of considering a procedure which guarantees the probability of a correct selection (PCS) with as small a sample size as possible. In this case, the probability requirement is given by

$$\Pr(CS|\mathscr{R}) \geq P^* \quad \text{where } \boldsymbol{\theta} \in \Omega, \tag{3.3}$$

where $P^*\left(1/\binom{k}{t} < P^* < 1\right)$ is a preassigned constant. Note that the $\Pr(CS|\mathscr{R})$ is considered over the whole parameter space Ω.

3.2. Selection procedures based on the vector ranks

Let us consider the ranking and selection problem of selecting the population with the largest (or smallest) parameter value, based on the vector ranks defined in Definition 2.1.

For the descriptions on populations, we use the same ones as in Section 2. Further, in this section, we suppose that the distribution functions $F(x; \theta_i)$, $i = 1, 2, \ldots, k$ belong to the location or scale parameter family of distributions.

By the same procedure (2.2) as in Section 2, we construct the rank sum statistic

$$T_i = \sum_{j=1}^{n} R(X_{ij}), \ i = 1, 2, \ldots, k \tag{3.4}$$

obtained by the n repetitions of observation vectors. Again, as we stated in Section 2, note that the ranking procedure is carried out for the absolute values of observations, if it is necessary, when we are dealing with the scale parameters.

More specifically, the problem here is stated as follows. Let the ordered parameters be

$$\theta_{[1]} \leq \theta_{[2]} \leq \cdots \leq \theta_{[k]}. \tag{3.5}$$

We consider a ranking and selection procedure $\mathscr{R}$ based on the rank sum T_i, $i = 1, 2, \ldots, k$ or $\boldsymbol{T} = (T_1, T_2, \ldots, T_k)$ for selecting the population associated with $\theta_{[k]}$. Here, we assume without loss of generality that the population $\prod_k$ is associated with the parameter $\theta_{[k]}$.

Now, let us consider the following two procedures of selecting the best population based on $\boldsymbol{T}$. These types of rules are used for selecting the best population with either the largest location or scale parameters, using statistic $\boldsymbol{T}$.

3.3. *PCS* for the indifference zone formulation

First we consider the indifference zone approach due to Bechhofer [3] under the above framework. We denote this procedure as $\mathscr{R}_I(\alpha)$, which is given as follows.

(3.6) $\mathscr{R}_I(\alpha)$: Select the population associated with T_k as the best,

where α is the index for two cases, i.e., we set $\alpha = 1$ for the location parameter case and $\alpha = 2$ for the scale parameter case.

In this case, the rule $\mathscr{R}_I(\alpha)$, $(\alpha = 1, 2,)$ is requested to satisfy the following probability requirement.

$$\Pr(CS|\mathscr{R}_I(\alpha)) \geq P^* \text{ whenever } \varphi_\alpha(\theta_k, \theta_i) \geq \gamma_\alpha + \delta^*_\alpha \tag{3.7}$$

where $1/k < P^* < 1$, δ^*_α is a given constants, φ_α is a distance function,

$$\varphi_\alpha(\theta_i, \theta_j) = \begin{cases} \theta_i - \theta_j & \text{when } \alpha = 1, \\ \theta_i/\theta_j & \text{when } \alpha = 2, \end{cases} \tag{3.8}$$

and

$$\gamma_\alpha = \begin{cases} 0 & \text{when } \alpha = 1, \\ 1 & \text{when } \alpha = 2. \end{cases} \tag{3.9}$$

Note that this is a little different from the usual notation because we are treating the location and the scale parameter simultaneously. We sometimes call (3.8) as the location gap (when $\alpha = 1$) and the scale gap (when $\alpha = 2$).

By using Theorem 2.1 in Section 2, we have the following form of the probability of a correct selection (*PCS*) for the procedure $\mathscr{R}_I(\alpha)$.

THEOREM 3.1. *Let $\mathscr{R}_I(\alpha)$ be the selection procedure defined in* (3.6). *Then the probability of a correct selection is given as follows.*

$$\Pr(CS|\mathscr{R}_I(\alpha)) = \sum_{g=1}^{k} \frac{1}{g} \sum_{\boldsymbol{n} \in I_g} n! \prod_{i=1}^{k} \left(\prod_{j=1}^{s} \frac{1}{n_{ij}!} p_{ij}^{n_{ij}} \right), \tag{3.10}$$

where I_g is the index set

$$(3.11)\qquad I_g = \left\{ \boldsymbol{n};\ \boldsymbol{QKn} \underset{(g-1)}{\leq} \boldsymbol{O},\quad \boldsymbol{J}'_{(k!)}\boldsymbol{n} = n \right\},$$

where $\boldsymbol{Q}$ is a $(k-1) \times k$ matrix

$$(3.12)\qquad \boldsymbol{Q} = (\boldsymbol{E}_{(k-1)}, -\boldsymbol{J}_{(k-1)}).$$

Here $\boldsymbol{E}_{(k)}$ denotes the unit matrix of order k, g corresponds to the "tied" rank case and "$\boldsymbol{a} \underset{(g)}{\leq} \boldsymbol{O}$" means that some g elements of the vector $\boldsymbol{a}$ are equal to zero and the remaining elements are negative.

PROOF. If the population associated with $\prod_k$ is the best, then a correct selection occurs if and only if $T_k > T_i$, $i = 1, 2, \ldots, k-1$. Thus, PCS is given by using Theorem 2.1. concerning the joint distribution of $(T_1, T_2, \ldots, T_k)$. While ties may occur among rank sums including T_k, and tied cases must be broken by certain chance mechanism with equal probabilities. Thus, considering the number g of tied to the maximum cases, we have the theorem.

Note that the PCS for $\mathscr{R}_I(\alpha)$ will be given to various kinds of goals—such as selection of the population with the smallest parameter—only by changing the index set given above.

Form Theorem 3.1, we have the following results for $k = 2$ and 3.

The case $k = 2$

By using the distributional results (2.23) of the rank sum and the relation known as the sum of binomial probabilities

$$(3.13)\qquad \sum_{\nu=0}^{n} \binom{N}{\nu} p^{\nu}(1-p)^{N-\nu} = N\binom{N-1}{n} \int_p^1 x^n (1-x)^{N-n-1}\, dx$$

we have

(3.14)

$$\Pr(CS|\mathscr{R}_I(\alpha)) = \begin{cases} B_p(\nu+1, \nu+1)/B(\nu+1, \nu+1) & \text{for } n = 2\nu+1, \\ B_p(\nu+1, \nu+1)/2B(\nu+1, \nu+2) & \text{for } n = 2\nu+2, \end{cases} (\nu = 1, 2, \ldots)$$

where $B(m, n)$ and $B_p(m, n)$ are the complete and the incomplete beta functions respectively, and p is given by the p_{21} in (2.24).

Now we have to give an idea related to the parameter configuration. The least favorable configuration (LFC) for the procedure $\mathscr{R}_I(\alpha)$ is defined as the configuration of the population parameters which attains the minimum of the PCS under the requirement (3.2) or (3.7). In this case, it is easily shown that

LFC is given by

$$\varphi_\alpha(\theta_k, \theta_i) = \gamma_\alpha + \delta_\alpha^* \tag{3.15}$$

for $\alpha = 1, 2$.

The case $k = 3$

In general, as k increases, the form of the index set I_g and the variable relationship becomes tremendously complicated. For $k = 3$, the index set I_g is given as follows.

$$\begin{aligned} I_1 &= \{\mathbf{n}; D_1 > 0, D_2 > 0\} \\ I_2 &= \{\mathbf{n}; D_1 = 0, D_2 > 0 \text{ or } D_1 > 0, D_2 = 0\} \\ I_3 &= \{\mathbf{n}; D_1 = D_2 = 0\} \end{aligned} \tag{3.16}$$

where

$$\begin{aligned} D_1 &= n_{31} + 2n_{32} - n_{21} - 2n_{22} + n_{11} - n_{12} \\ D_2 &= 2n_{31} + n_{32} + n_{21} - n_{22} - n_{11} - 2n_{12} \end{aligned} \tag{3.17}$$

and $n_{31} + n_{32} + n_{21} + n_{22} + n_{11} + n_{12} = n$.

PCS can be calculated by using Theorem 3.1, for various populations. For the location and scale parameter cases of normal and exponential distribution given in A-1, A-2, B-1 and B-2 of Section 2.2, we can obtain the *PCS* values by using the relations (2.28) through (2.31).

However, to obtain the forms of *LFC* is generally a hard nut to crack. We will discuss on this topic in Section 3.5.

3.4. *PCS* for the subset selection formulation

Another approach is based on the subset selection formulation due to Gupta [14], which selects the subset of populations using the following selection procedures.

$$\begin{aligned} &\mathscr{R}_S(\alpha) : \text{Select } \textstyle\prod_i \text{ if and only if } T_i \geq \max T_j - d_\alpha, \\ &\qquad i = 1, 2, \ldots, k; \quad d_\alpha \geq 0, \quad \alpha = 1, 2. \end{aligned} \tag{3.18}$$

Recall that a correct selection (CS) occurs if and only if the best population (in our case $\prod_k$) is included in the selected subset. Our aim is to find a selection procedure satisfying

$$\inf_{\Omega} \Pr(CS \mid \mathscr{R}_S(\alpha)) \geq P^*, \tag{3.19}$$

where $\alpha = 1, 2$, $1/k < P^* < 1$ and $\Omega = \{\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k); \theta_i \in \Theta, i = 1, 2, \ldots, k\}$.

We have the following theorem.

THEOREM 3.2. *Let $\mathscr{R}_S(\alpha)$ be the selection procedure defined by* (3.18). *Then the probability of a correct selection is given as follows.*

$$\Pr(CS|\mathscr{R}_S(\alpha)) = \sum_{\boldsymbol{n} \in I} n! \prod_{i=1}^{k} \left(\prod_{j=1}^{s} \frac{1}{n_{ij}!} p_{ij}^{n_{ij}} \right). \tag{3.20}$$

where $s = (k-1)!$, I is the index set such that

$$I = \{\boldsymbol{n}; \boldsymbol{QKn} \leq d_\alpha \boldsymbol{J}'_{(k-1)}, \boldsymbol{J}'_{(k!)}\boldsymbol{n} = n\} \tag{3.21}$$

and $\boldsymbol{Q}$, $\boldsymbol{J}_{(k-1)}$ are given by (3.12) *and* (2.9).

PROOF. If $\prod_k$ is the best population, then a correct selection occurs if and only if $T_k \geq T_i - d_\alpha$, $i = 1, 2, \ldots, k-1$. Thus, using Theorem 2.1, we have the above result.

By the same way as in the indifference zone formulation, we have the following results for $k = 2$ and 3.

The case $k = 2$

We can write the index set in the *PCS* formula (3.20).

$$I = \{(n_{21}, n_{11}); n_{21} - n_{11} \geq -d_\alpha, \ n_{21} + n_{11} = n\}. \tag{3.22}$$

Thus we have

$$n_{21} \geq (n - d_\alpha)/2. \tag{3.23}$$

Letting $n^* = \{(n - d_\alpha)/2\} - 1$, where $\{x\}$ is the smallest integer not less than x, from Theorem 3.1 we have

$$\Pr(CS|\mathscr{R}_S(\alpha)) = \frac{B_p(n - n^*, n^*)}{B(n - n^*, n^* + 1)}, \tag{3.24}$$

where p is p_{21} in (2.24). The least favorable configuration is given when $\theta_2 = \theta_1$, i.e., when $p = \frac{1}{2}$ in (3.24).

The case $k = 3$

The index set in the *PCS* formula (3.20) is given for $\boldsymbol{n} = (n_{31}, n_{32}, n_{21}, n_{22}, n_{12}, n_{11})$ which satisfy the relations

$$\begin{aligned} n_{31} + 2n_{32} - n_{21} - 2n_{22} + n_{11} - n_{12} &\geq -d_\alpha \\ 2n_{31} + n_{32} + n_{21} - n_{22} - n_{11} - 2n_{12} &\geq -d_\alpha \end{aligned} \tag{3.25}$$

and $n_{31}+n_{32}+n_{21}+n_{22}+n_{11}+n_{12}=n$. Further, $p_{ij}(i=1,2,3;j=1,2)$ is given in Section 2.2, for seperate cases of corresponding populations.

3.5. *LFC* and the slippage configuration

Since the probability of a correct selection (*PCS*) is a function of the unknown parameters, we try to evaluate its value under the scheme of probability requirements given by (3.2) for the indifference zone formulation and by (3.3) for the subset selection formulation.

As we mentioned in Section 3.3, it is sufficient to evaluate the *PCS* under the *LFC*, that is, under the configuration of parameters attaining the minimum of the *PCS* with respect to the respective probability requirements.

Various discussions have been made to get the least favorable configuration of some selection procedures. Among which, the slippage configuration played a significant role in giving the least favorable configuration of parameters. Actually, for example, for the selection of the best of location or scale parameter of normal populations, the slippage configuration is shown to be the *LFC* (Bechhofer [3], [4]).

Now we define the slippage configuration of distributions as follows:

DENITION 3.1. For given c.d.f.'s $F_1(x), F_2(x), \dots, F_k(x)$, a slippage configuration of distributions is defined by

$$(3.26) \qquad F_1(x)=\cdots=F_{k-t}(x)(\equiv F_0(x)) \geq F_{k-t+1}(x)=\cdots=F_k(x)(\equiv F(x)).$$

for all x, where t is any given integer such that $1 \leq t \leq k-1$.

Since we are dealing with the location and scale parmeter families of distributions, the above definition is the same as

$$(3.27) \qquad \theta_1=\cdots=\theta_{k-t}(\equiv\theta_0) \leq \theta_{k-t+1}=\cdots=\theta_k(\equiv\theta).$$

Then one of the important problems is to find the *LFC* for each of the selection procedures based on the ranks. Selection procedure based on the combined ranks (For the definition, see Section 4.1.) is first considered by Lehmann [18]. He obtained the *PCS* by assuming the slippage configuration to be a *LFC*. But Rizvi and Woodworth [28] showed the assumption to be incorrect by a counter example. Matsui [19] also considered the procedure using vector ranks and assuming the slippage configuration to be the *LFC*. But this assumption is also shown to be incorrect by Lee and Dudewicz [17], by giving a counter example.

These two counter examples are very tricky ones. So it may be noted that still slippage configuration is one of the nearest paths to give *LFC* for most of the selection procedures based on ranks. Actually, for selecting the location parameter of normal populations based on vector ranks, it is partially proved and partially numerically confirmed that the slippage configuration is the *LFC* (Alam and Thompson [1], Matsui [21]).

Finally, we give a result concerning the distribution of $\boldsymbol{T}$ given in (2.2), under the slippage configuration of $t = 1$, (see Definition 3.1). In this case, the probabilities $p_{ij}, i = 1, 2, \ldots, k; j = 1, 2, \ldots, (k-1)!$, defined in (2.4), have the same forms for every j. Thus for every j, we have

$$(3.28) \qquad p_{ij} = \Pr(\overbrace{X_{l_1} < \cdots < X_{l_{i-1}}}^{i-1} < X_k < \overbrace{X_{l_{i+1}} < \cdots < X_{l_{k-1}}}^{k-i}) \\ = \int \binom{k-1}{i-1} F_0(x)^{i-1}(1 - F_0(x))^{k-i}\, dF(x) \quad \text{for } i = 1, 2, \ldots, k.$$

This simplifies the distribution (2.5) and hence also the expressions for the *PCS* (3.10) and (3.20). We discuss this case again in Section 4.3.

3.6. Sufficient condition on the *LFC*

Now we give a sufficient condition for a configuration to be *LFC* for some selection procedures. This result can be used to examine the *LFC* of some selection procedures based on the ranks.

Let k-populations $\prod_1, \prod_2, \ldots, \prod_k$ be given with stochastically ordered distribution function as follows:

$$(3.29) \qquad F(x; \theta_i) > F(x, \theta_k), \quad i = 1, 2, \ldots, k-1, \quad \text{for all } x.$$

Let $\boldsymbol{X}_j = (X_{1j}, X_{2j}, \ldots, X_{kj}), j = 1, 2, \ldots, n$, be the j-th independent sample from each population, where $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_k$ are assumed to be independent. Let $R_{ij} = U(X_{ij})$ be functions of Xij, which are defined with respect to X_{ij} themselves, or with respect to relations among the components of $\boldsymbol{X}_j$. Now,

$$(3.30) \qquad T_i = \sum_{j=1}^{n} R_{ij}$$

is the statistic associated with the i-th population.

Our aim is to choose, among k populations, a population $\prod_k$ with the stochastically smallest c.d.f. $F_k(x) = F(x; \theta_k)$ by means of the statistics $(T_1, T_2, \ldots, T_k)$. A natural procedure is to choose $\prod_s$ if T_s is the largest among $T_1, T_2, \ldots, T_k$.

Now, the probability of a correct selection is given by

$$(3.31) \qquad \Pr_{\Omega}\left(\max_{1<i<k-1} T_i < T_k\right),$$

where

(3.32) $$\Omega = \{(F_1(x), F_2(x), \ldots, F_k(x));\ F_i(x) > F_k(x),\ i = 1, 2, \ldots, k-1\}.$$

The least favorable configuration of populations is defined to be a configuration Ω_0 such that

(3.33) $$\Pr_{\Omega}\left(\max_{1<i<k-1} T_i < T_k\right) \geq \Pr_{\Omega_0}\left(\max_{1<i<k-1} T_i < T_k\right)$$

for all Ω.

The following theorem gives us a sufficient condition for

(3.34) $$\Omega_0 = \{F_1(x), F_2(x), \ldots, F_k(x)); F_1(x) = F_2(x) = \cdots = F_{k-1}(x) > F_k(x)\}$$

to be the least favorable configuration to our selection problem (see Matsui and Choi [23]).

THEOREM 3.3. *If the c.d.f.* $G_\Omega(y) = G_\Omega(y_1, y_2, \ldots, y_{k-1})$ *of* $y = (R_{1j} - R_{kj}, R_{2j} - R_{kj}, \ldots, R_{k-1j} - R_{kj})$, $j = 1, 2, \ldots, n$, *satisfies the relation*

(3.35) $$G_\Omega(y) \geq G_{\Omega_0}(\mathrm{y}) \quad \textit{for all } y,$$

then the configuration Ω_0 *in* (3.34) *is a LFC.*

4. Selection problem based on the ranks—asymptotic case

As we mentioned in above sections, especially in Section 3.5, *LFC* is a key notion to the selection procedure based on the ranks. We have seen that several investigations have been made.

In this section, we consider the selection procedures of selecting the largest (smallest) location parameter or the smallest (largest) scale parameter from an asymptotic viewpoint. Since we are considering two formulations and two types of rank sum statistics—vector rank type and combined rank type—let us restate our framework on the populations, parameters, statistics and procedures. Note that notations of statistics and procedures are different from the previous sections, since we are treating another (combined) type of statistics also.

4.1. Restatement of the problem

Let $\prod_1, \prod_2, \ldots, \prod_k$ be k independent populations where $\prod_i$ has the associated cumulative distribution function (*c.d.f.*) $F(x; \theta_i)$, $i = 1, 2, \ldots, k$. It is assumed that $\{F(x; \theta_i)\}$ is a location or scale parameter family, i.e., $F(x; \theta_i) = F(x - \theta_i)$, $-\infty < \theta_i < \infty$, or $F(x; \theta_i) = F(x/\theta_i)$, $\theta_i > 0$. Let the ordered θ_i be denoted by $\theta_{[1]} \leq \theta_{[2]} \leq \cdots \leq \theta_{[k]}$. The population associated with $\theta_{[k]}$ is defined to be the best. Note that selection of the worst can be treated in the same as the best case, as is explained below. Our procedure for selecting the best population is based on the ranks of observations from these populations.

Then we have

$$F(x; \theta_{[1]}) \geq F(x; \theta_{[2]}) \geq \cdots \geq F(x; \theta_{[k]}) \tag{4.1}$$

for all x. We assume, without loss of generality, that $\prod_k$ is the best population (i.e., $\theta_k \geq \theta_i$, $i = 1, 2, \ldots, k-1$).

We now consider the rank sum statistics based on the combined ranks.

DEFINITION 4.1. Let $X_{i1}, X_{i2}, \ldots, X_{in_i}$ be n_i independent observations from the population $\prod_i$, $i = 1, 2, \ldots, k$, and let $N = \sum_{i=1}^{k} n_i$. Then, the combined ranks of the observation X_{ij} is defined as

$$R(X_{ij}) = r \tag{4.2}$$

if X_{ij} is the r-th smallest among all N observations.

The combined rank is known as Wilcoxon type rank, and its rank sum is used in Mann-Whitney-Wilcoxon Test, or in Kruskal-Wallis Test for k-sample problem.

We are going to give a unified treatment for the selection procedures based on the vector ranks (Friedman type rank) defined in (2.1), and the combined ranks (Wilcoxon type rank) defined here. Let $R_{ij}^{(1)}$ denote the rank of X_{ij} as a vector rank, $R_{ij}^{(2)}$ denote the rank of X_{ij} as a combined rank and consider the rank sum in the following way.

For $s = 1, 2$, define

$$T_i^{(s)} = b_s \sum_{j=1}^{n} R_{ij}^{(s)}, \quad i = 1, 2, \ldots, k, \tag{4.3}$$

and

$$\boldsymbol{T}^{(s)} = (T_1^{(s)}, T_2^{(s)}, \ldots, T_k^{(s)})', \tag{4.4}$$

where $b_1 = 1$ and $b_2 = 1/n_i$.

We now consider the two selection procedures—the indifference zone formulation and the subset selection formulation—based on these two statistics. First, under the indifference zone formulation, we define the selection procedure $\mathscr{R}_I(s, \alpha)$ as follows:

$$\mathscr{R}_I(s, \alpha) : \text{Select the population associated with } T_{[k]}^{(s)} \text{ as the best.} \tag{4.5}$$

This is the selection procedure selecting the population with the best location parameter ($\alpha = 1$) or scale parameter ($\alpha = 2$), based on the statistic $\boldsymbol{T}^{(s)}$.

In this case, the rule $\mathscr{R}_I(s, \alpha)$, $(s, \alpha = 1, 2)$ is required to satisfy the following probability condition:

$$(4.6) \qquad \Pr(CS|\mathscr{R}_I(s,\alpha) \geq P^* \text{ whenever } \varphi_\alpha(\theta_k, \theta_i) \geq \gamma_\alpha + \delta_\alpha^*,$$

where $1/k < P^* < 1$, $\delta_\alpha^* > 0$ are specified constants, $\varphi_\alpha(\theta_i, \theta_j)$ and γ_α are given by (3.8) and (3.9) respectively.

Under the subset selection formulation for selecting a subset containing the best population, we define the following procedures:

$$(4.7) \qquad \mathscr{R}_S(s,\alpha) : \text{Select } \prod_i \text{ if and only if } T_i^{(s)} \geq \max T_j^{(s)} - d_\alpha,$$
$$i = 1, 2, \ldots, k; \quad d_\alpha \geq 0; \quad s, \alpha = 1, 2.$$

Again, $\alpha = 1$ and 2 correspond to the location and scale parameter cases, respectively. We say that a correct selection (CS) occurs if and only if the best population (in our case $\prod_k$) is included in the selected subset. Our aim is to obtain a rule $\mathscr{R}_S(s,\alpha)$ satisfying

$$(4.8) \qquad \inf_{\Omega} \Pr(CS|\mathscr{R}_S(s,\alpha)) \geq P^*,$$

where $s, \alpha = 1, 2$; $1/k < P^* < 1$; $\Omega = \{\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_k); \theta_i \in \Theta, i = 1, 2, \ldots, k\}$, Θ is a real line. The constant d_α is the smallest non-negative number satisfying (4.8), so called P^* condition.

As we mentioned before, it is fairly difficult to construct the *LFC* for both the rules $\mathscr{R}_I(1,\alpha)$ and $\mathscr{R}_S(2,\alpha)$ based on the ranks. This problem is still open in general (for $\alpha = 1, 2$). For example, as in the counter examples, it is pointed out that the configuration $\theta_1 = \theta_2 = \cdots = \theta_{k-1}$; $\varphi_\alpha(\theta_k, \theta_{k-1}) = \gamma_\alpha + \delta_\alpha^*$ in the indifference zone procedures, or the configuration $\theta_1 = \theta_2 = \cdots = \theta_k$ in the subset selection rules, do not yield, in general, the minimum of the *PCS*. A discussion on the *LFC* can be found in Gupta and McDonald [12].

Our purpose is to discuss the *LFC* under an asymptotic framework with reference to the underlying distribution of populations. We assume that an order relation holds between the "gaps" of parameters. This assumption is similar to those considered by Lehmann [18], and Alam and Thompson [1]. The *LFC*'s of the procedure are studied by using the exact moments of the combined and vector rank statistics $\boldsymbol{T}^{(s)}$, $s = 1, 2$ for the location and the scale parameter cases ($\alpha = 1, 2$), for both the subset selection and the indifference zone formulations.

Now we consider the asymptotic distribution of $\boldsymbol{T}^{(s)}$, $s = 1, 2$ with reference to the exact moments of population distributions. Then we will investigate the *PCS* under the slippage configuration of parameters.

4.2. Moments of the ranks

Let us denote the mean vector and variance-covariance matrix of $\boldsymbol{T}^{(s)}$ by $\boldsymbol{\mu}_\alpha^{(s)}$ and $\Lambda_\alpha^{(s)}$, ($s = 1, 2$), according as we are dealing with the location ($\alpha = 1$) or the scale ($\alpha = 2$) parameters. Under the distributional set-up we considered in Section 1, the elements of $\boldsymbol{\mu}_\alpha^{(s)} = (\mu_{\alpha 1}^{(s)}, \mu_{\alpha 2}^{(s)}, \dots, \mu_{\alpha k}^{(s)})$ and $\Lambda_\alpha^{(s)} = (\lambda_{\alpha ij}^{(s)})$, $i, j = 1, 2, \dots, k$ are given as follows.

For the vector rank case, the moments were given in Matsui [22] from which we obtain the elements of mean vector $\boldsymbol{\mu}_\alpha^{(1)}$ and variance-covariance matrix $\Lambda_\alpha^{(1)}$ of statistics $\boldsymbol{T}^{(1)}$ as follows:

$$\mu_{\alpha i}^{(1)} = nE(R_i), \tag{4.9}$$

$$\lambda_{\alpha ij}^{(1)} = \begin{cases} n\,Var(R_i), & \text{for } i = j, \\ n\,Cov(R_i, R_j), & \text{for } i \neq j, \end{cases} \tag{4.10}$$

where $E(R_i)$, $Var(R_i)$ and $Cov(R_i, R_j)$ are given in Theorem 2.2.

For the combined rank case, the mean vector $\boldsymbol{\mu}_\alpha^{(2)}$ and the variance-covariance matrix $\Lambda_\alpha^{(2)}$ are obtained from more general results given in the Appendix. In this case, since we are going to compare the two procedures based on the vector rank and the combined rank statistics, we consider the case

$$n_1 = n_2 = \cdots = n_k \equiv n \tag{4.11}$$

in Definition 4.1. Then we have

$$\mu_{\alpha i}^{(2)} = kn \int G(x)\, dF_i(x) + 1/2, \quad i = 1, 2, \dots, k, \tag{4.12}$$

$$\lambda_{\alpha ij}^{(2)} = \begin{cases} k(3n-1)\int G(x)\,dF_i(x) - 2k(2n-1)\int F_i(x)G(x)\,dF_i(x) \\ \quad + k^2 n \int G(x)^2\,dF_i(x) - k\int H(x)\,dF_i(x) - k^2 n\left(\int G(x)\,dF_i(x)\right)^2 \\ \quad - (n-1)\sum_{m=1}^{k}\left(\int F_m(x)\,dF_i(x)\right)^2 - 1/12, & \text{if } i = j, \\ kn\left(2 - \int F_j(x)\,dF_i(x)\right)\int G(x)\,dF_j(x) \\ \quad + kn\left(2 - \int F_i(x)\,dF_j(x)\right)\int G(x)\,dF_i(x) \\ \quad - n\sum_{m=1}^{k}\int F_m(x)\,dF_i(x)\int F_m(x)\,dF_j(x) - 2kn\int F_j(x)G(x)\,dF_i(x) \\ \quad - 2kn\int F_j(x)G(x)\,dF_j(x) + \int F_i(x)\,dF_j(x)\int F_j(x)\,dF_i(x) \\ \quad + \int F_i(x)^2\,dF_j(x) + \int F_j(x)^2\,dF_i(x) - 1, & \text{if } i \neq j, \end{cases} \tag{4.13}$$

where $G(x)$ and $H(x)$ are the same ones as in the equations (2.35).

4.3. *PCS* under the slippage configuration

In this section, we derive asymptotic results on the *PCS* for both the indifference zone and the subset selection formulation, under the slippage configuration of distributions.

The slippage configuration of distributions we consider here is the cases $t = 1$ in Definition 3.1 and hence we have the following distributional structure,

$$F_1(x) = F_2(x) = \cdots = F_{k-1}(\equiv F_0(x)) \geq F_k(x)(\equiv F(x)). \tag{3.14}$$

For the vector rank sum statistics (2.2), we have, from the expressions (2.32), (2.33) and (2.34), the following means, variances and covariances $(i = 1, 2, \ldots, k-1)$.

$$\begin{aligned}
(4.15)\qquad \mu_i^{(1)} &= n\{-a + (k+2)/2\},\\
\mu_k^{(1)} &= n\{(k-1)a + 1\},\\
\sigma_{ii}^{(1)} &= n\{(k-1)a - a^2 - (k-2)b + k(k-2)/12\},\\
\sigma_{kk}^{(1)} &= n\{(k-1)a - (k-1)^2a^2 + (k-1)(k-2)b\},\\
\sigma_{ki}^{(1)} &= \sigma_{ik}^{(1)} = n\{-a + (k-1)a^2 - (k-2)b\},\\
\sigma_{ij}^{(1)} &= n(-a - a^2 + 2b - k/12), \quad i \neq j, \quad i, j \neq k.
\end{aligned}$$

Also for the combined rank sum statistic (4.3), from (4.12) and (4.13) we have the following expressions.

$$\begin{aligned}
(4.16)\qquad \mu_i^{(2)} &= -na + (kn + n + 1)/2,\\
\mu_k^{(2)} &= (k-1)na + (n+1)/2,\\
\sigma_{ii}^{(2)} &= (kn-1)a - (2n-1)a^2 - (kn - 3n + 1)b - 2(n-1)c\\
&\quad + (k^2n - 3kn + 2n + k - 2)/12,\\
\sigma_{kk}^{(2)} &= (k-1)(2n-1)a - (k-1)(nk-1)a^2\\
&\quad + (k-1)(nk - n - 1)b - 2(k-1)(n-1)c,\\
\sigma_{ki}^{(2)} &= \sigma_{ik}^{(2)} = (-2n+1)a + (nk-1)a^2 - (nk - n - 1)b + 2(n-1)c,\\
\sigma_{ij}^{(2)} &= -na - na^2 + 2nb - (nk - n + 1)/12, \quad i \neq j, \quad i, j \neq k,
\end{aligned}$$

where

$$a = \int F_0(x)\, dF(x), \tag{4.17}$$

$$b = \int F_0(x)^2\, dF(x), \tag{4.18}$$

$$c = \int F_0(x) F(x)\, dF(x), \tag{4.19}$$

Thus for large n, the rank sum vector $\boldsymbol{T}^{(s)}$ is asymptotically distributed as a k-dimensional normal distribution with mean vector $\boldsymbol{\mu}_\alpha^{(s)}$ and variance covariance matrix $\Lambda_\alpha^{(s)}$ where elements of $\boldsymbol{\mu}_\alpha^{(s)}$ and $\Lambda_\alpha^{(s)}$ are given by (4.15) and (4.16), respectively for $s = 1, 2$.

Using $\boldsymbol{Q}$ defined by (3.12), let $\boldsymbol{Z}^{(s)} = \boldsymbol{Q}\boldsymbol{T}^{(s)}$. Then, $\boldsymbol{Z}^{(s)}$ is asymptotically distributed as a normal distribution

$$N(\boldsymbol{Q}\boldsymbol{\mu}_\alpha^{(s)}, \boldsymbol{Q}\Lambda_\alpha^{(s)}\boldsymbol{Q}'). \tag{4.20}$$

Here $\boldsymbol{Q}\boldsymbol{\mu}_\alpha$ is the vector with $k-1$ elements

$$\boldsymbol{Q}\boldsymbol{\mu}_\alpha^{(s)} = (\nu^{(s)}, \dots, \nu^{(s)}), \tag{4.21}$$

where

$$\nu^{(s)} = \mu_i^{(s)} - \mu_k^{(s)}, \quad i = 1, 2, \dots, k-1, \tag{4.22}$$

and $\boldsymbol{Q}\Lambda_\alpha^{(s)}\boldsymbol{Q}'$ is a $(k-1) \times (k-1)$ matrix

$$\boldsymbol{Q}\Lambda_\alpha^{(s)}\boldsymbol{Q}' = \begin{pmatrix} \tau_0^{(s)} & & \tau^{(s)} \\ & \ddots & \\ \tau^{(s)} & & \tau_0^{(s)} \end{pmatrix}, \tag{4.23}$$

with

$$\begin{aligned} \tau_0^{(s)} &= \sigma_{ii}^{(s)} - \sigma_{ik}^{(s)} - \sigma_{ki}^{(s)} + \sigma_{kk}^{(s)}, \quad i = 1, 2, \dots, k-1, \\ \tau^{(s)} &= \sigma_{ij}^{(s)} - \sigma_{ik}^{(s)} - \sigma_{ki}^{(s)} + \sigma_{kk}^{(s)}, \quad i, j = 1, 2, \dots, k-1, \ i \neq j. \end{aligned} \tag{4.24}$$

That is,

$$\begin{aligned} \nu^{(1)} &= kn(-a + 1/2), \\ \tau_0^{(1)} &= n\{2ka - k^2a^2 + k(k-2)b + k(k-2)/12\}, \\ \tau^{(1)} &= n\{ka - k^2a^2 + k(k-1)b - k/12\}, \end{aligned} \tag{4.25}$$

and

$$(4.26)\quad \nu^{(2)} = kn(-a + 1/2),$$

$$\tau_0^{(2)} = (3kn + 2n - k - 2)a - (k^2n + kn + 2n - k - 2)a^2 + (k^2n - kn + 2n - k - 2)b - 2(k+2)(n-1)c + (k^2n - 3kn + 2n + k - 2)/12,$$

$$\tau^{(2)} = (2kn + n - k - 1)a + (-k^2n - kn - n + k + 1)a^2 + (k^2n + n - k - 1)b - 2(kn + n - k - 1)c - (kn - n + 1)/12,$$

If the underlying population is normal, then from the statements given in Section 2.2, we have the following forms.

Location parameter case: Let the location gap of $F_0(x)$ and $F(x)$ be $\Delta(\geq 0)$, then

$$(4.27)\quad a = \Phi(\Delta/\sqrt{2}),\quad b = L(-\Delta/\sqrt{2}, -\Delta/\sqrt{2}; 0.5),\quad c = L(-\Delta/\sqrt{2}, 0; 0.5).$$

Scale parameter case: Let the scale gap of $F_0(x)$ and $F(x)$ be $\Delta(\geq 1)$, then

$$(4.28)\quad a = (2/\pi)\sin^{-1}\{\Delta/\sqrt{1+\Delta^2}\},\quad b = (2/\pi)\sin^{-1}\{\Delta^2/(1+\Delta^2)\},$$

$$c = (2/\pi)\sin^{-1}\{\Delta/\sqrt{2(1+\Delta^2)}\}.$$

According to the statements on *PCS*'s of two formulations, we have the following propositions. Where the notation $\approx$ means the (asymptotic) equivalence of both sides of the equation when n is large.

PROPOSITION 4.1. *The PCS for the selection procedure $\mathscr{R}_I(s,\alpha)$ under the slippage configuration is given for large n as follows* $(s, \alpha = 1, 2)$.

$$(4.29)\quad \Pr(CS \mid \mathscr{R}_I(s,\alpha)) \approx \int \Phi\left(\frac{-\nu^{(s)} + \sqrt{\tau^{(s)}}y}{\sqrt{\tau_0^{(s)} - \tau^{(s)}}}\right)^{k-1} \phi(y)\, dy.$$

where ϕ and Φ are given in (2.10) *and* (2.11) *respectively.*

PROOF. If the population $\prod_k$ is the best population, then a correct selection occurs if and only if $\boldsymbol{QT}^{(s)} \leq \boldsymbol{O}$. Thus using (4.20) through (4.23), we have for large n

$$(4.30)\quad \Pr(CS \mid \mathscr{R}_I(s,\alpha)) \approx \Pr(\boldsymbol{Z}^{(s)} = \boldsymbol{QT}^{(s)} \leq \boldsymbol{O}) = \int_{\boldsymbol{W}^{(s)} \leq -\boldsymbol{Q}\boldsymbol{\mu}_\alpha^{(s)}/\sqrt{\tau_0^{(s)}}} N(\boldsymbol{O}, \boldsymbol{B}_\alpha^{(s)})\, d\boldsymbol{W}^{(s)},$$

where $\boldsymbol{B}_\alpha^{(s)} = (1/\tau_0^{(s)})\boldsymbol{Q}\boldsymbol{\Lambda}_\alpha^{(s)}\boldsymbol{Q}'$ and $\boldsymbol{W}^{(s)} = (\boldsymbol{Z}^{(s)} - \boldsymbol{Q}\boldsymbol{\mu}^{(s)})/\sqrt{\tau_0^{(s)}}$. Here, using Lemma 2.2, we have (4.29).

The integral (4.29) is a special case of the integral investigated by Gupta [8]. This type of integral can be evaluated numerically by using the numerical integration method of Gauss-Hermite type.

PROPOSITION 4.2. *The PCS for the procedure $\mathscr{R}_S(s,\alpha)$ under the slippage configuration is given for large n as follows* $(s,\alpha = 1,2)$.

$$\Pr(CS|\mathscr{R}_S(s,\alpha)) \approx \int \Phi\left(\frac{d_\alpha - \nu^{(s)} + \sqrt{\tau^{(s)}}y}{\sqrt{\tau_0^{(s)} - \tau^{(s)}}}\right)^{k-1} \phi(y)\,dy. \tag{4.31}$$

PROOF. Since a correct selection occurs if and only if $\boldsymbol{Q}\boldsymbol{T}^{(s)} \leq d_\alpha \boldsymbol{J}_{(k-1)}$, thus we have for large n

$$\begin{aligned} \Pr(CS|\mathscr{R}_S(s,\alpha)) &\approx \Pr(\boldsymbol{Z}^{(s)} = \boldsymbol{Q}\boldsymbol{T}^{(s)} \leq d_\alpha \boldsymbol{J}_{(k-1)}) \\ &= \int_{\boldsymbol{W}^{(s)} \leq (d_\alpha \boldsymbol{J}_{(k-1)} - \boldsymbol{Q}\boldsymbol{\mu}_\alpha^{(s)})/\sqrt{\tau_0^{(s)}}} N(\boldsymbol{O}, \boldsymbol{B}_\alpha^{(s)})\,d\boldsymbol{W}^{(s)}, \end{aligned} \tag{4.32}$$

where $\boldsymbol{W}^{(s)} = (1/\tau_0^{(s)})\boldsymbol{Q}\boldsymbol{\Lambda}_\alpha^{(s)}\boldsymbol{Q}'$ and $\boldsymbol{W}^{(s)} = (\boldsymbol{Z}^{(s)} - \boldsymbol{Q}\boldsymbol{\mu}^{(s)})/\sqrt{\tau_0^{(s)}}$. Here, using Lemma 2.2, we have (4.31).

Note that in most of subset selection procedures, *LFC* is given by the equi-parameter case (but not proved in general, for the present case). In the equi-parameter case, $\nu^{(s)} = 0$ in (4.31).

5. Asymptotic properties

This section is devoted to obtain the *LFC* under the assumption of order relation between gaps of parameters. Several relevant results are also given. First, we will set an assumption in relation to the pairs of parameters.

5.1. Assumption

We assume the following relation to hold between the gaps of parameters:

$$\varphi_\alpha(\theta_i, \theta_j) = \gamma_\alpha + \kappa_{\alpha ij} n^{-1/2} + o(n^{-1/2}), \quad \alpha = 1,2, \tag{5.1}$$

where γ_α is given by (3.9); for each pair (i,j), $\kappa_{\alpha ij}$ depends on θ_i and θ_j and is increasing in θ_i when θ_j is fixed, and decreasing in θ_j when θ_i is fixed; also, $\kappa_{\alpha ij} = \gamma_\alpha$ when $\theta_i = \theta_j$.

Then letting

$$I_{\alpha ij} \equiv \sqrt{n}\left\{\int F_j(x)\,dF_i(x) - \int F_i(x)\,dF_i(x)\right\}, \tag{5.2}$$

we obtain the following lemma.

LEMMA5.1. *For* $\varphi_\alpha(\theta_i,\theta_j)$, $(\alpha = 1,2)$ *given by* (3.8), *we have the following*:

$$I_{\alpha ij} = K_{\alpha ij} + o(1), \tag{5.3}$$

where

$$K_{\alpha ij} = \begin{cases} \kappa_{1ij}\int f(x)^2\,dx, & \text{when } \alpha = 1, \\ \kappa_{2ij}\int xf(x)^2\,dx, & \text{when } \alpha = 2, \\ & i,j = 1,2,\ldots,k; \quad i \neq j. \end{cases} \tag{5.4}$$

PROOF. By using the Taylor expansion to the integral $I_{\alpha ij}$,

$$I_{\alpha ij} = \sqrt{n}\left[\int F(x + \varphi_\alpha(\theta_i,\theta_j))\,dF(x) - 1/2\right],$$

we have the lemma.

In the special cases when the underlying distribution is $N(0,1)$, we have the following:

$$I_{1ij} = \frac{1}{2\sqrt{\pi}}\kappa_{1ij} + o(1), \tag{5.5}$$

and

$$I_{2ij} = \sqrt{\frac{8}{\pi}}\kappa_{2ij} + o(1). \tag{5.6}$$

5.2. Asymptotic distribution

Let us define

$$W_i^{(s)} = \frac{1}{\sqrt{n}}\,(T_i^{(s)} - T_k^{(s)}), \quad s = 1,2. \tag{5.7}$$

Then

$$\boldsymbol{W}_\alpha^{(s)} = \frac{1}{\sqrt{n}}\,\boldsymbol{Q}\boldsymbol{T}^{(s)}, \quad s = 1,2, \tag{5.8}$$

where $\boldsymbol{W}_\alpha^{(s)} = (W_1^{(s)}, W_2^{(s)}, \ldots, W_k^{(s)})'$ and $\boldsymbol{Q}$ is given by (3.12), has the mean vector $\boldsymbol{\eta}_\alpha^{(s)}$ with elements

$$(5.9) \qquad \eta_{\alpha i}^{(s)} = \frac{1}{\sqrt{n}} (\mu_{\alpha i}^{(s)} - \mu_{\alpha k}^{(s)}), \quad i = 1, 2, \ldots, k-1,$$

and the variance-covariance matrix $\dot{\boldsymbol{\Sigma}}_\alpha^{(s)}$ with elements

$$(5.10) \qquad \sigma_{\alpha ij}^{(s)} = \frac{1}{n} (\lambda_{\alpha ij}^{(s)} - \lambda_{\alpha ik}^{(s)} - \lambda_{\alpha kj}^{(s)} + \lambda_{\alpha kk}^{(s)}), \quad i, j = 1, 2, \ldots, k-1,$$

where $\mu_{\alpha i}^{(s)}$ and $\lambda_{\alpha ij}^{(s)}$ are given by (4.9) throught (4.13).

Now, under the assumption (5.1), using Lemma 5.1, we have for $\alpha = 1, 2$ and $s = 1, 2$

$$(5.11) \qquad \eta_{\alpha i}^{(s)} = \frac{1}{\sqrt{n}} \left(n \int \sum_{j=1}^{k} F_j \, dF_i - n \int \sum_{j=1}^{k} F_j \, dF_k \right) \rightarrow \sum_{\substack{j=1 \\ j \neq i}}^{k} K_{\alpha ij} - \sum_{j=1}^{k-1} K_{\alpha kj}, \quad (\equiv \tilde{\eta}_{\alpha i}^{(s)})$$

as $n \rightarrow \infty$, where $K_{\alpha ij}$ is given by (5.4). Also, under (5.1), we have

$$(5.12) \qquad \lambda_{1ij} \rightarrow \begin{cases} -(k+1)/12, & \text{for } i \neq j, \\ (k^2 - 1)/12, & \text{for } i = j, \end{cases}$$

$$(5.13) \qquad \lambda_{2ij} \rightarrow \begin{cases} -k/12, & \text{for } i \neq j, \\ (k^2 - k)/12, & \text{for } i = j. \end{cases}$$

Consequently,

$$(5.14) \qquad \sigma_{\alpha ij}^{(s)} \rightarrow \begin{cases} v_s, & \text{for } i \neq j, \\ 2v_s, & \text{for } i = j, \end{cases}$$

where

$$(5.15) \qquad v_s = \begin{cases} k(k+1)/12, & \text{when } s = 1, \\ k^2/12, & \text{when } s = 2. \end{cases}$$

Thus using the central limit theorem, we have the following asymptotic distribution of $\boldsymbol{W}_\alpha^{(s)}$:

$$(5.16) \qquad \boldsymbol{W}_\alpha^{(s)} \sim N(\tilde{\boldsymbol{\eta}}_\alpha^{(s)}, \tilde{\boldsymbol{\Sigma}}_\alpha^{(s)}), \quad \alpha = 1, 2,$$

where $\tilde{\boldsymbol{\eta}}_\alpha^{(s)} = (\tilde{\eta}_{\alpha 1}^{(s)}, \tilde{\eta}_{\alpha 2}^{(s)}, \ldots, \tilde{\eta}_{\alpha(k-1)}^{(s)})'$ with lements given by (5.11), and

$$\tilde{\boldsymbol{\Sigma}}_\alpha^{(s)} = v_s(\boldsymbol{E}_{(k-1)} + \boldsymbol{G}_{(k-1)}), \tag{5.17}$$

with $\boldsymbol{G}_{(k-1)} = \boldsymbol{J}_{(k-1)}\boldsymbol{J}'_{(k-1)}$.

5.3. *PCS* and *LFC*

In this section, we give the probability of a correct selection (*PCS*) for the indifference zone and the subset selection approaches and consider the least favorable configuration (*LFC*) of parameters for these two approaches.

Indifference zone formulation

Using the asymptotic distribution of $\boldsymbol{W}_\alpha^{(s)}$, $(s, \alpha = 1, 2)$ given by (5.16), the *PCS* for the rule $\mathscr{R}_I(s, \alpha)$ $(s, \alpha = 1, 2)$ is given for large n by

$$\begin{aligned} \Pr(CS|\mathscr{R}_I(s,\alpha)) &\approx \Pr(\boldsymbol{W}_\alpha^{(s)} \le 0) \\ &= \Pr(\boldsymbol{U}_\alpha^{(s)} \le -\tilde{\boldsymbol{\eta}}_\alpha^{(s)}/\sqrt{2v_s}), \end{aligned} \tag{5.18}$$

where

$$\boldsymbol{U}_\alpha^{(s)} = \frac{1}{\sqrt{2v_s}}(\boldsymbol{W}_\alpha^{(s)} - \tilde{\boldsymbol{\eta}}_\alpha^{(s)}), \tag{5.19}$$

$$\boldsymbol{U}_\alpha^{(s)} \sim N(\boldsymbol{O}_{(k-1)}, (\boldsymbol{E}_{(k-1)} + \boldsymbol{G}_{(k-1)})/2). \tag{5.20}$$

Since *LFC* is given as the minimum of the (5.18), we study behaviors of $\tilde{\boldsymbol{\eta}}_\alpha^{(s)}$ under the requirement

$$\varphi_\alpha(\theta_k, \theta_i) \ge \gamma_\alpha + \delta_\alpha^*, \tag{5.21}$$

we have the following theorem.

THEOREM 5.1. *Under the assumptian of order restriction* (5.1) *the* (*asymptotic*) *LFC of the PCS for the rules* $\mathscr{R}_I(s, \alpha)$, $s, \alpha = 1, 2$ *is given for large n by*

$$\varphi(\theta_k, \theta_i) = \gamma_\alpha + \delta_\alpha^*, \quad i = 1, 2, \ldots, k-1. \tag{5.22}$$

PROOF. By using (5.11) and (5.4), we have

$$\tilde{\boldsymbol{\eta}}_\alpha^{(s)} = \sum_{\substack{j=1 \\ j \ne i}}^{k-1} (\kappa_{\alpha ij} - \kappa_{\alpha kj})A_\alpha + (\kappa_{\alpha ik} - \kappa_{\alpha ki})A_\alpha, \tag{5.23}$$

where

$$A_\alpha = \begin{cases} \int f(x)^2\,dx, & \text{when } \alpha = 1, \\ \int x f(x)^2\,dx, & \text{when } \alpha = 2. \end{cases} \tag{5.24}$$

From assumption (5.1), we have

$$\kappa_{\alpha ij} - \kappa_{\alpha kj} = \sqrt{n}(\varphi_\alpha(\theta_i, \theta_j) - \varphi_\alpha(\theta_k, \theta_j)) + o(1). \tag{5.25}$$

Here we examine the location and the scale parameter cases seperately.
The case $\alpha = 1$ (location parameter): From the probability requirement (5.21), we have $\varphi_1(\theta_k, \theta_i) \geq \delta_1^*$, thus

$$\kappa_{1ij} - \kappa_{1kj} = -\sqrt{n}\varphi_1(\theta_k, \theta_i) + o(1) \leq -\sqrt{n}\delta_1^* + o(1). \tag{5.26}$$

Thus we have, for large n

$$\tilde{\eta}_{1i}^{(s)} \leq -\sqrt{n}(k-2)\delta_1^* A_1, \quad i = 1, 2, \ldots, k-1. \tag{5.27}$$

The case $\alpha = 2$ (scale parameter): Since

$$\kappa_{2ij} - \kappa_{2kj} = \sqrt{n}\varphi_2(\theta_k, \theta_j)\{1/\varphi_2(\theta_k, \theta_i) - 1\} + o(1) \tag{5.28}$$

and $\varphi_2(\theta_k, \theta_i) \geq 1 + \delta_2^*$ for scale parameter case, we have

$$\kappa_{2ij} - \kappa_{2kj} = \leq -\sqrt{n}\delta_2^* + o(1). \tag{5.29}$$

Thus we have, for large n

$$\tilde{\eta}_{2i}^{(s)} \leq -\sqrt{n}(k-2)\delta_2^* A_2, \quad i = 1, 2, \ldots, k-1. \tag{5.30}$$

In summary, for both the location and the scale parameter cases, we have

$$\tilde{\boldsymbol{\eta}}_\alpha^{(s)} \leq -\zeta_\alpha \boldsymbol{J}_{(k-1)}, \tag{5.31}$$

where

$$\zeta_\alpha = \begin{cases} \sqrt{n}(k-2)\int f(x)^2\,dx, & \text{for } \alpha = 1, \\ \sqrt{n}(k-2)\int x f(x)^2\,dx, & \text{for } \alpha = 2. \end{cases} \tag{5.32}$$

Since the region of integration for PCS is $\boldsymbol{U}_\alpha^{(s)} \leq -\tilde{\boldsymbol{\eta}}_\alpha^{(s)}/\sqrt{2v_s}$, the minimum of the PCS is given by the minimum of $\tilde{\eta}_{\alpha i}^{(s)}$, $i = 1, 2, \ldots, k-1$, that is when $\varphi_\alpha(\theta_k, \theta_i) = \gamma_\alpha + \delta_\alpha^*$.

Thus, neglecting the terms of $o(1)$, we have the following expressions for PCS.

COROLLARY 5.1. *Under the asymptotic LFC,*

$$\Pr(CS|\mathscr{R}_I(s,\alpha)) \approx \Pr\left(\boldsymbol{U}_\alpha^{(s)} \le \frac{1}{\sqrt{2v_s}}\zeta_\alpha \boldsymbol{J}_{(k-1)}\right). \tag{5.33}$$

Using lemma 2.2, *the expression* (5.33) *can be rewritten as*

$$\Pr(CS|\mathscr{R}_I(s,\alpha)) \approx \int \Phi\left(x + \frac{\zeta_\alpha}{2\sqrt{v_s}}\right)^{k-1} \phi(x)\,dx \tag{5.34}$$

for $s, \alpha = 1, 2$.

As we mentioned before, this type of integral can be evaluated by using the Gauss-Hermite method of numerical integration.

Subset selection formulation

Following the arguments similar to the ones for the rule $\mathscr{R}_I(s,\alpha)$, we can derive asymptotic results on the *PCS* for the rule $\mathscr{R}_S(s,\alpha)$, $(s,\alpha = 1,2)$. We have, for large n,

$$\begin{aligned}\Pr(CS|\mathscr{R}_I(s,\alpha)) &\approx \Pr\left(\boldsymbol{W}_\alpha^{(s)} \le \frac{d_\alpha}{\sqrt{n}}\boldsymbol{J}_{(k-1)}\right) \\ &= \Pr\left(\boldsymbol{U}_\alpha^{(s)} \le \left(-\tilde{\boldsymbol{\eta}}_\alpha^{(s)} + \frac{d_\alpha}{\sqrt{n}}\boldsymbol{J}_{(k-1)}\right)/\sqrt{2v_s}\right),\end{aligned} \tag{5.35}$$

where $\boldsymbol{U}_\alpha^{(s)}$ is same as (5.19).

By the same way as in the indifference zone formulation, the infimum of *PCS* is given when $\tilde{\boldsymbol{\eta}}_\alpha^{(s)} = \boldsymbol{O}_{(k-1)}$, that is $\varphi(\theta_k,\theta_i) = 0$, which implies the following.

THEOREM 5.2. *Under the assumption of order restriction* (5.1) *the* (*asymptotic*) *LFC of the PCS for the rules* $\mathscr{R}_S(s,\alpha), s,\alpha = 1,2$ *is given for large* n *by*

$$\varphi(\theta_k,\theta_i) = 0, \quad i = 1,2,\ldots,k-1. \tag{5.36}$$

Also we have the following corollary to evaluate the *PCS* function asymptotically.

COROLLARY 5.2. *Under the asymptotic LFC,*

$$\Pr(CS|\mathscr{R}_S(s,\alpha)) \approx \Pr\left(\boldsymbol{U}_\alpha^{(s)} \le \frac{d_\alpha}{\sqrt{2nv_s}}\boldsymbol{J}_{(k-1)}\right). \tag{5.37}$$

Using lemma 2.2, *the expression* (5.37) *can be rewritten as*

$$\Pr(CS|\mathscr{R}_S(s,\alpha)) \approx \int \Phi\left(x + \frac{d_\alpha}{2\sqrt{nv_s}}\right)^{k-1} \phi(x)\,dx \tag{5.38}$$

for $s, \alpha = 1, 2.$

Appendix: Moments of combined rank

Let k populations $\prod_1, \prod_2, \dots, \prod_k$ be given, where $\prod_i$ has the corresponding continuous distribution $F_s(x)$, $(s = 1, 2, \dots, k)$. Take n_s observations $X_{s1}, X_{s2}, \dots, X_{sn_s}$ from populations $\prod_s$, $(s = 1, 2, \dots, k)$ and consider the combined (Wilcoxon type) rank R_{sj} of X_{sj} among all kn observations. Then the means, variances and covariances of ranks R_{sj} are given in the following way.

THEOREM A.1.

$$E(R_{sj}) = N\int G(x)\,dF_s(x) + \frac{1}{2}, \tag{A.1}$$

$$V(R_{sj}) = 2N\int G(x)\,dF_s(x) - 2N\int F_s(x)G(x)\,dF_s(x) + N^2\int G(x)^2\,dF_s(x) - N\int H(x)\,dF_s(x) - N^2\left(\int G(x)\,dF_s(x)\right)^2 - \frac{1}{12}, \tag{A.2}$$

$$Cov(R_{si}, R_{sj}) = 3N\int G(x)\,dF_s(x) - 4N\int F_s(x)G(x)\,dF_s(x) - \sum_{l=1}^{k} n_l\left(\int F_l(x)\,dF_s(x)\right)^2 - \frac{1}{12}, \tag{A.3}$$

$$\begin{aligned} Cov(R_{si}, R_{tj'}) = {} & N\left(2 - \int F_t(x)\,dF_s(x)\right)\int G(x)\,dF_t(x) \\ & + N\left(2 - \int F_s(x)\,dF_t(x)\right)\int G(x)\,dF_s(x) \\ & - \sum_{l=1}^{k} n_l\int F_l(x)\,dF_s\int F_l\,dF_t \\ & - 2N\int F_t(x)G(x)\,dF_s(x) - 2N\int F_s(x)G(x)\,dF_t(x) \\ & + \int F_s(x)\,dF_t(x)\int F_t(x)\,dF_s(x) + \int F_s(x)^2\,dF_t(x) \\ & + \int F_t(x)^2\,dF_s(x) - 1, \end{aligned} \tag{A.4}$$

where $s, t = 1, 2, \ldots, k, s \neq t$; $i, j = 1, 2, \ldots, n_s, i \neq j$; $j' = 1, 2, \ldots, n_t$ and

$$N = \sum_{l=1}^{k} n_l, \quad G(s) = \frac{1}{N}\sum_{l=1}^{k} n_l F_l(x), \quad H(x) = \frac{1}{N}\sum_{l=1}^{k} n_l F_l(x)^2. \tag{A.5}$$

Proof. We sketch the proofs for (A.1) and (A.3) above. The remaining results are obtained similarly.

Mean:

$$\begin{aligned} Pr(R_{11} = s) = \sum_{A} \Pr(&a_1 \text{ of } X_1's, a_2 \text{ of } X_2's, \ldots, a_k \text{ of } X_k's \leq X_1 \\ &\leq (n_1 - a_1 - 1) \text{ of } X_1's, (n_2 - a_2) \text{ of } X_2's, \ldots, (n_k - a_k) \text{ of } X_k's) \end{aligned} \tag{A.6}$$

where a_i $(i = 1, 2, \ldots, k)$ is an integer such that

$$0 \leq a_1 \leq n_1 - 1, \quad 0 \leq a_i \leq n_i \ (i = 2, 3, \ldots, k), \quad \sum_{j=1}^{k} a_j = s - 1 \tag{A.7}$$

and "a_i of X_i's", "$(n_i - a_i)$ of X_i's" should be read as "a_i variables out of $(X_{i1}, X_{i2}, \ldots, X_{in_i})$ and remaining $(n_i - a_i)$ variables", and so forth. Further, the summation $\sum_A$ is taken over all k-tuples $(a_1, a_2, \ldots, a_k)$ of integers which satisfy the relation (A.7). From (A.6), we have

(A.8)

$$\begin{aligned} E(R_{11}) = \int \sum_{s=1}^{N} \sum_{A} s \binom{n_1 - 1}{a_1} \binom{n_2}{a_2} \cdots \binom{n_k}{a_k} F_1(x)^{a_1} F_2(x)^{a_2} \cdots F_k(x)^{a_k} \\ \times (1 - F_1(x))^{(n_1 - a_1 - 1)} (1 - F_2(x))^{(n_2 - a_2)} \cdots (1 - F_k(x))^{(n_k - a_k)} \, dF_1(x). \end{aligned}$$

By changing the order of summation, we have

$$\begin{aligned} E(R_{11}) = \int \sum_{s=1}^{N} \sum_{A} s \binom{n_1 - 1}{a_1} \binom{n_2}{a_2} \cdots \binom{n_{k-1}}{a_{k-1}} F_1(x)^{a_1} F_2(x)^{a_2} \cdots F_{k-1}(x)^{a_{k-1}} \\ \times (1 - F_1(x))^{(n_1 - a_1 - 1)} (1 - F_2(x))^{(n_2 - a_2)} \cdots (1 - F_{k-1}(x))^{(n_{k-1} - a_{k-1})} \\ \times \left(n_k F_k(x) + \sum_{i=1}^{k-1} a_i + 1 \right) dF_1(x), \end{aligned}$$

where the summation $\sum_{A_1}$ is taken over all $(k-1)$-tuples $(a_1, a_2, \ldots, a_{k-1})$ of integers which satisfy the relation (A.7). Adding in turn over $a_{k-1}, a_{k-2}, \ldots, a_1$, we obtain the result for $E(R_{11})$.

Covariance:
For $s < t$, we have

$$\text{(A.9)} \qquad \Pr(R_{11} = s, R_{21} = t) = \sum\nolimits_B \Pr(a_1 \text{ of } X_1's, a_2 \text{ of } X_2's, \ldots, a_k \text{ of } X_k's$$
$$\leq X_{11} \leq b_1 \text{ of } X_1's, b_2 \text{ of } X_2's, \ldots, b_k \text{ of } X_k's$$
$$\leq X_{21} \leq c_1 \text{ of } X_1's, c_2 \text{ of } X_2's, \ldots, c_k \text{ of } X_k's),$$

where a_i, b_i, c_i $(i = 1, 2, \ldots, k)$ are integers such that

$$\text{(A.10)} \qquad a_i + b_i + c_i = v_i, \quad i = 1, 2, \ldots, k,$$

$$\text{(A.11)} \qquad \sum_{j=1}^{k} a_j = s - 1, \quad \sum_{j=1}^{k} b_j = t - s - 1, \quad \sum_{j=1}^{k} c_j = n - t,$$

and $v_i = n_i - 1$ for $i = 1, 2$, $v_i = n_i$ for $i = 3, 4, \ldots, k$.

The summation $\sum_B$ is taken over all tuples $(a_i, \ldots, a_k, b_1, \ldots, b_k, c_1, \ldots, c_k)$ which satisfy the relations (A.9) and (A.10). Then

(A.12)

$$I_1 \equiv \sum_{s<t} st \Pr(R_{11} = s, R_{12} = t) = \iint_{x<y} \sum_{s<t} \sum_{B} \prod_{i=1}^{k} P_i(x, y)\, dF_1(x)\, dF_2(y),$$

where

$$P_i(x, y) = \binom{v_i}{a_i, b_i, c_i} F_i^{a_i}(x)(F_i(y) - F_i(x))^{b_i}(1 - F_i(y))^{c_i}, \quad i = 1, 2, \ldots, k.$$

By changing the order of summation, first for s and for t, we have

$$I_1 = \iint_{x<y} \sum_{s<t} \sum_{B_1} C_1 \prod_{i=1}^{k-1} P_i(x, y)\, dF_1(x)\, dF_2(y)$$

where

$$C_1 = \alpha_1 + \beta_1 \sum_{j=1}^{k-1} a_j + \gamma_1 \sum_{j=1}^{k-1} b_j + \left(\sum_{j=1}^{k-1} a_j\right)^2 + \left(\sum_{j=1}^{k-1} a_j\right)\left(\sum_{j=1}^{k-1} b_j\right)$$

and

$$\alpha_1 = n_k(n_k - 1)F_k(x)F_k(y) + 3n_kF_k(x) + n_kF_k(y) + 2,$$
$$\beta_1 = n_kF_k(x) + n_kF_k(y) + 3,$$
$$\gamma_1 = n_kF_k(x) + 1.$$

The summation $\sum_{B_1}$ is taken over all tuples $(a_1, \ldots, a_{k-1}, b_1, \ldots, b_{k-1}, c_1, \ldots, c_{k-1})$ which satisfy the condition (A.10). By carrying out the addition in turn for sets (a_i, b_i, c_i), $(i = k-1, k-2, \ldots, 2, 1)$, we have a reduced form of I_1. By proceeding on similar steps for $\sum_{s>t} st \Pr(R_{11} = s, R_{21} = t)$, we obtain $Cov(R_{11}, R_{21})$.

For rank sums

$$T_s = \sum_{j=1}^{n_s} R_{sj}, \quad s = 1, 2, \ldots, k, \tag{A.13}$$

we have

$$E(T_s) = n_s E(R_{sj}), \quad s = 1, 2, \ldots, k, \tag{A.14}$$

$$Cov(T_s, T_t) = n_s n_t Cov(R_{sj}, R_{tj'}), \quad s, t = 1, 2, \ldots, k, \quad s \neq t, \tag{A.15}$$

and

(A.16)

$$\begin{aligned} V(T_s) &= \sum_{j=1}^{n_s} V(R_{sj}) + \sum_{i \neq j}^{n_s} Cov(R_{si}, R_{sj}) \\ &= Nn_s(3n_s - 1) \int G(x)\, dF_s(x) - 2Nn_s(2n_s - 1) \int F_s(x) G(x)\, dF_s(x) \\ &\quad + N^2 n_s \int G(x)^2\, dF_s(x) - Nn_s \int H(x)\, dF_s(x) - N^2 n_s \left(\int G(x)\, dF_s(x) \right)^2 \\ &\quad - n_s(n_s - 1) \sum_{l=1^k}^{k} n_l \left(\int F_l(x)\, dF_s(x) \right)^2 - \frac{n_s^2}{12}. \end{aligned}$$

Especially, if $F_i(x) = F(x)$ for all i, then we have

$$E(T_s) = \frac{n_s(N+1)}{2}, \tag{A.17}$$

$$V(T_s) = \frac{n_s(N - n_s)(N+1)}{12}. \tag{A.18}$$

$$Cov(T_s, T_t) = -\frac{n_s n_t (N+1)}{12}. \tag{A.19}$$

Also for $k = 2$, we have the following:

$$E(T_i) = \frac{n_i(n_i + 1)}{2} + n_i n_j \int F_j(x)\, dF_i(x), \quad i, j = 1, 2; \quad j \neq i, \tag{A.20}$$

(A.21)

$$V(T_i) = n_i n_j (2n_i - 1) \int F_j(x)\, dF_i(x) + n_i n_j (n_j - 1) \int F_j(x)^2\, dF_i(x)$$
$$+ n_i n_j (n_i - 1) \int F_i(x)^2\, dF_j(x) - n_i n_j (n_i + n_j - 1) \left(\int F_j(x)\, dF_i(x) \right)^2$$
$$- n_i n_j (n_i - 1), \quad i, j = 1, 2; i \neq j,$$

(A.22)
$$Cov(T_1, T_2) = n_1 n_2 \Big\{ n_1 \int F_1(x)\, dF_2 + n_2 \int F_2(x)\, dF_1(x)$$
$$- (n_1 + n_2 - 1) \int F_1(x)\, dF_2(x) \int F_2(x)\, dF_1(x)$$
$$- (n_1 - 1) \int F_1(x)^2 dF_2(x) - (n_2 - 1) \int F_2(x)^2\, dF_1(x) - 1 \Big\}.$$

Acknowledgement

The author expresses his sincere thanks to Professor Yasunori Fujikoshi, Hiroshima University, for his encouragements and various comments and suggestions through this study. He also thanks to Professors Junjiro Ogawa, Sadao Ikeda, Sumiyasu Yamamoto and Shanti S. Gupta for their continuous encouragements.

References

[1] K. Alam, and J. R. Thompson, A selection procedure based on ranks, Ann. Inst. Statist. Math., **23** (1971), 253–262.

[2] N. S. Bartlett, and Z. Govindarajulu, Some distribution-free statistics and their applications to the selection problem, Ann. Inst. Statist. Math., **20** (1968), 79–97.

[3] R. E. Bechhofer, A single sample multiple decision procedure for ranking means of normal populations with known variances, Ann. Math. Statist., **25** (1954), 16–39.

[4] R. E. Bechhofer and M. Sobel, A single sample multiple decison procedure for ranking variances of normal populations, Ann. Math. Statist., **25** (1954), 273–289.

[5] R. E. Bechhofer, J. Kiefer and M. Sobel, Sequential Identification and Ranking Procedures. The University of Chicago Press, Chicago, 1968.

[6] E. J. Dudewicz and J. O. Koo, The Complete Categorized Guide to Statistical Selection and Ranking Procedures. American Science Press, Ohio, 1982.

[7] J. D. Gibbons, I. Olkin, and M. Sobel, Selecting and Ordering Populations. Wiley, New York, 1977.

[8] S. S. Gupta, Probability integrals of multivariate normal and multivariate *t*, Ann. Math. Statist., **34** (1963), 792–828.

[9] S. S. Gupta J. O. Berger (Eds.), Statistical Decision Theory and Related Topics, V. Springer-Verlag, New York, 1994.

[10] S. S. Gupta and T. Matsui, On the least favorable configuration of a selection procedure based on ranks, Tech. Rep. 87-2, Dept. of Stat., Purdue University, 1987.
[11] S. S. Gupta and G. C. McDonald, On some classes of selection procedures based on ranks, Nonparametric Techniques in Statistical Inference (Ed. M. L. Puri), Cambridge Univ. Press, London, 1970, 491–514.
[12] S. S. Gupta and G. C. McDonald, Nonparametric procedures in multiple decisions (ranking and selection procedures), Colloquia Mathematica Societatis János Bolyai, **32**, Nonparametric Statistical Inference (B. V. Gnedenko, M. L. Puri and I. Vince eds.), Vol. I, North-Holland Publishing Co., 1982, 361–389.
[13] S. S. Gupta and S. Panchapakesan, Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations, John Wiley & Sons, 1979.
[14] S. S. Gupta and S. Panchapakesan, Subset selection procedures: Review and assessment, Amer. J. of Math. Management Sciences, **5** (1985), 235–311.
[15] S. S. Gupta and S. Panchapakesan, Selection and screening procedures in multivariate analysis, Multivariate Analysis: Future Directions, 1993, 233–263.
[16] J. C. Hsu, Multiple Comparisons, Chapman & Hall, London, 1996.
[17] Y. J. Lee and E. J. Dudewicz, On the least favorable configuration of a selection procedure based on rank sums: Counter example to a postulate of Matsui, J. Japan Statist. Soc., **9** (1979), 65–69.
[18] E. L. Lehmann, A class of selection procedures based on ranks, Math. Ann., **150** (1963), 268–275.
[19] T. Matsui, On selecting the best one of k normal populations based on ranks, J. Japan Statist. Soc., **2** (1972), 71–81.
[20] T. Matsui, Asymptotic behavior of a selection procedures based on rank sums, J. Japan Statist. Soc., **4** (1974), 17–25.
[21] T. Matsui, Note on a selection procedure based on ranks, $k = 3$. Dokkyo Studies of Economics, **21** (1977), 79–94.
[22] T. Matsui, Moments of rank vector with applications to selection and ranking, J. Japan Statist. Soc., **15** (1985), 17–25.
[23] T. Matsui and K. Choi, A sufficient condition for *LFC* of certain selection procedures, Dokkyo Studies of Economics, **54** (1989), 109–116.
[24] G. C. McDonald, Some multiple comparison selection procedures based on ranks, Sankhya, Ser. A, **34** (1972), 53–64.
[25] G. C. McDonald, The distribution of some rank statistics with applications in block design selection problem, Sankhya, Ser. A. **35** (1973), 187–204.
[26] M. L. Puri and P. S. Puri, Multiple decison procedures based on ranks for certain problems in analysis of variance, Ann. Math. Statist., **40** (1969), 619–632.
[27] P. S. Puri and M. L. Puri, Selection procedures based on ranks: Scale parameter case, Sankhya, Ser. A., **30** (1968), 291–302.
[28] M. H. Rizvi and G. G. Woodworth, On selection procedure based on ranks: Counter examples concerning the least favorable configurations, Ann. Math. Statist., **41** (1970), 1942–1951.

Dokkyo University
Faculty of Economics
Soka, Saitama, 340-0042 *Japan*