ON THE LEAST FAVORABLE CONFIGURATION OF A
SELECTION PROCEDURE BASED ON RANKS

by

Shanti S. Gupta*
Purdue University

Takashi Matsui
Purdue University and Dokkyo University
Technical Report #87-2

Department of Statistics
Purdue University

January 1987

* The research of this author was supported in part by the Office of Naval Research Contract N00014-84-C-0167 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.
ON THE ASYMPTOTIC LEAST FAVORABLE CONFIGURATION OF
A SELECTION PROCEDURE BASED ON RANKS

Shanti S. Gupta*    Takashi Matsui**

ABSTRACT

The problem of selection of the population with the largest parameter is
considered using the subset selection as well as the indifference zone approach
for distributions that belong to a location or a scale parameter family. The
procedures are based on the sums of combined (Wilcoxon type) ranks and vector
(Friedman type) ranks. The least favorable configurations are obtained in an
asymptotic framework under certain order relations between the "gaps" of
parameters. The asymptotic theory is based on exact moments of the rank sum
statistics.

1. INTRODUCTION

Let \( \Pi_1, \Pi_2, \ldots, \Pi_k \) be \( k \) independent populations, where \( \Pi_i \) has the associ-
ated cumulative distribution function (c.d.f) \( G_{\theta_i}(y) \), \( i = 1, 2, \ldots, k \). It is
assumed that \( \{G_{\theta}\} \) is a location or a scale parameter family, i.e. \( G_{\theta}(y) = \)
\( G(y - \theta), \quad -\infty < \theta < \infty \), or \( G_{\theta}(y) = G(y/\theta), \quad \theta > 0 \). Let the ordered \( \theta_i \) be
denoted by \( \theta_1 \preceq \theta_2 \preceq \ldots \preceq \theta_k \). The population associated with \( \theta_{[k]} \) is
defined to be the best. Our procedures for selecting the best population are

* Department of Statistics, Purdue University, West Lafayette, Indiana 47907,
U.S.A. The research of this author was supported in part by the ONR Contract
N00014-84-C-0187 at Purdue University.
** Faculty of Economics, Dokkyo University, 600 Sakae-cho, Soka-shi, Saitama 340
based on ranks of observations from these populations in the location parameter case, and on ranks of the absolute values of the observations in the scale parameter case. For any observation \( Y \) from \( G_\theta (y) \) in the scale parameter case, the c.d.f. of \( X = |Y| \) is \( F_\theta (x) = G(x/\theta) - G(-x/\theta) \). For convenience of presentation, we use \( F_\theta (x) \) to denote the c.d.f. of \( |Y| \) in the scale parameter case as well as that of \( Y \) itself in the location parameter case [i.e. \( F_\theta (x) = G(x-\theta) \)]. Then we have
\[
F_{\theta_n}(x) \geq F_{\theta_{n-1}}(x) \geq \ldots \geq F_{\theta_1}(x) \tag{1.1}
\]
for all \( x \). We assume without loss of generality, that \( \Pi_k \) is the best population (i.e. \( \theta_k \geq \theta_i \), \( i = 1,2,\ldots,k-1 \)).

Let \( Y_{i1}, Y_{i2}, \ldots, Y_{in} \) be \( n \) independent observations from \( \Pi_i \), \( i = 1,2,\ldots,k \). Let \( X_{ij} = Y_{ij} \) in the location parameter case, and \( X_{ij} = |Y_{ij}| \) in the scale parameter case. We consider rank sum statistics based on combined (Wilcoxon type) ranks as well as vector (Friedman type) ranks. Let \( R_{ij}^{(1)} \) denote the rank of \( X_{ij} \) among all \( kn \) observations in the combined sample and \( R_{ij}^{(2)} \) denote the rank of \( X_{ij} \) among \( X_{1j}, X_{2j}, \ldots, X_{kj} \). In ranking observations in a group, the smallest is given rank 1. For \( \alpha = 1,2 \), define
\[
H_i^{(\alpha)} = c_{n\alpha} \sum_{j=1}^{R} R_{ij}^{(\alpha)}, \quad i = 1,2,\ldots,k \tag{1.2}
\]
and
\[
H^{(\alpha)} = \left( H_1^{(\alpha)}, H_2^{(\alpha)}, \ldots, H_k^{(\alpha)} \right), \tag{1.3}
\]
where \( c_{n1} = 1/n \) and \( c_{n2} = 1 \).

For selecting the best population, we consider both the subset selection approach of Gupta (see Gupta and Panchapakesan [5]) and the indifference zone.
approach of Bechhofer [3]. For selecting a subset containing the best, we define the following procedures:

\[ R(\alpha, \beta, 1) : \text{Select } \Pi_i \text{ if and only if } H_i^{(\alpha)} \geq \max_{j \neq i} H_j^{(\alpha)} - d_\beta \]
\[ i = 1, 2, \ldots, k; \quad d_\beta \geq 0; \quad \alpha, \beta = 1, 2. \]  \hspace{1cm} (1.4)

Here \( \beta = 1 \) and 2 correspond to the location and scale cases, respectively.

The use of these rule is justified by Theorem 4.2. A correct selection (CS) is said to occur if and only if the best population (in our case \( \Pi_k \)) is included in the selected subset. Our aim is to select a subset satisfying

\[ \inf_{\Omega} P(\text{CS} | R(\alpha, \beta, 1)) \geq P^* \]  \hspace{1cm} (1.5)

where \( \alpha, \beta = 1, 2; \ 1/k < P^* < 1; \ \Omega = \{ \Theta = (\theta_1, \theta_2, \ldots, \theta_k); \ \theta_i \in \Theta, i = 1, 2, \ldots, k \}; \ \Theta \) is a real line. The constant \( d_\beta \) is the smallest nonnegative number satisfying (1.5), the so-called \( P^* \)-condition.

Using the indifference zone approach, we define the procedures:

\[ R(\alpha, \beta, 2) : \text{Select the population associate with } H_{[k]}^{(\alpha)} \text{ as the best.} \]  \hspace{1cm} (1.6)

In this case, the rule \( R(\alpha, \beta, 2), \alpha, \beta = 1, 2 \) are required to satisfy the following probability condition:

\[ P(\text{CS} | R(\alpha, \beta, 2)) \geq P^* \text{ whenever } \Phi_\beta (\theta_k, \theta_i) \geq c_\beta + \delta_\beta^* \]  \hspace{1cm} (1.7)

where \( \alpha, \beta = 1, 2, \ 1/k < P^* < 1, \ \delta_\beta > 0 \) is a given constant,

\[ \Phi_\beta (\theta_i, \theta_j) = \begin{cases} \theta_i - \theta_j & \text{when } \beta = 1 \\ \theta_i / \theta_j & \text{when } \beta = 2 \end{cases} \]  \hspace{1cm} (1.8)

and

\[ c_\beta = \begin{cases} 0 & \text{when } \beta = 1 \\ 1 & \text{when } \beta = 2. \end{cases} \]  \hspace{1cm} (1.9)
Selection procedures (using both subset selection and indifference zone approaches) based on this statistic $h^{(1)}$ have been studied by many authors including Lehmann [9], Bartlett and Govindarajulu [2], Gupta and McDonald [4], Puri and Puri [15,16], and Alam and Thompson [1]. Also procedures based on $h^{(2)}$ have been studied by McDonald [12,13], Watsui [10] and Lee and Dudewicz [8].

A review of procedures based on ranks is given by Gupta and McDonald [6], and Gupta and Panchapakesan [7].

A parametric configuration which gives the infimum of PCS, the probability of a correct selection, is called the least favorable configuration (LFC). It is fairly difficult to establish the LFC for both rules $R(\alpha, \beta, 1)$ and $R(\alpha, \beta, 2)$ using statistics $h^{(1)}, h^{(2)}$ and is still an open question in general ($\alpha, \beta = 1, 2$). The counter examples of Rizvi and Woodworth [17] and Lee and Dudewicz [8] show that the configuration $\theta_1 = \theta_2 = \ldots = \theta_k$ in the case of subset selection rules, and the configuration $\theta_1 = \theta_2 = \ldots = \theta_{k-1}$; $\varphi_\beta(\theta_k, \theta_{k-1}) = c_\beta + \delta_\beta^*$ in the case of indifference zone procedures, do not yield, in general the infimum of the PCS. A discussion on the LFC can be found in Gupta and McDonald [6].

Our purpose in this paper is to discuss the LFC in an asymptotic framework. An order relation is assumed to hold between the "gaps" of parameters. This assumption is similar to those considered by Puri and Puri [15,16], and Alam and Thompson [1]. The LFC's of the procedure are studied by using the exact moments of the combined and the vector rank statistics $H^{(\alpha)}$, $\alpha = 1, 2$, for location and scale parameter cases ($\beta = 1, 2$) for both subset selection and indifference zone approaches.
In Section 2, asymptotic distribution of $H^\alpha(\alpha), \alpha = 1, 2$, are considered under the assumption of order relation between gaps of parameters. The PCS and LFC are investigated in Section 3. Moments results are given in Section 4 as an appendix.

2. ASYMPTOTIC PROPERTY

2.1 Moments of Ranks

Let us define the mean vector and variance-covariance matrix of $H^{(1)}$ by $\mu_{\beta}^{(1)}$ and $\Lambda_{\beta}^{(1)}$ according as we are dealing with location ($\beta = 1$) or scale ($\beta = 2$) parameters. Under the population model we considered in Section 1, the elements of $\mu_{\beta}^{(1)}$ and $\Lambda_{\beta}^{(1)}$ are given as follows. These relations are obtained from more general results given in Theorem 4.1 of Appendix

\[
\mu_{\beta i}^{(1)} = kn \int G^{*}_i dF_i + 1/2, \quad i = 1, 2, \ldots, k, \quad (2.1)
\]

\[
\Lambda_{\beta ij}^{(1)} = \begin{cases}
 k(3n-1) \int G^{*}_i dF_i - 2k(2n-1) \int F_i G^{*}_i dF_i + k^2 n \int G^{*2}_i dF_i \\
 - k \int H^{*}_i dF_i - k^2 n(\int G^{*2}_i dF_i)^2 - (n-1) \sum_{m=1}^{k} (\int F_m dF_i)^2 - 1/12, \\
 i = j,
\end{cases}
\]

\[
\Lambda_{\beta ij}^{(1)} = \begin{cases}
 kn(2 - \int F_j dF_i) \int G^{*}_i dF_j + kn(2 - \int F_i dF_j) \int G^{*}_i dF_i \\
 - n \sum_{m=1}^{k} \int F_m dF_i \int F_m dF_j - 2kn \int F_i G^{*}_i dF_j - 2kn \int F_i G^{*}_i dF_j \\
 + \int F_i dF_j \int F_j dF_i + \int F_i dF_i + \int F_j dF_j + \int F_i dF_j - 1, \\
 i \neq j,
\end{cases}
\]

where

\[
G^{*}(x) = (1/k) \sum_{j=1}^{k} F_j(x), \quad (2.3)
\]

\[
H^{*}(x) = (1/k) \sum_{j=1}^{k} F_j^2(x). \quad (2.4)
\]
In case of vector rank $R_{ij}$, the moments results are given in Matsui [11] from which we obtain mean vector $\mu_\beta^{(2)}$, variance-covariance matrix $\Lambda_\beta^{(2)}$ of statistic $H_\beta^{(2)}$ as follows:

$$\mu_\beta^{(2)} = kn \int G^*dF_i + n/2, \quad i = 1, 2, \ldots, k, \quad (2.5)$$

$$\lambda_\beta^{(2)} = \begin{cases} 
    n[2k \int G^*dF_1 - 2k \int F_1 G^*dF_1 + k^2 \int G^{*2}dF_1 \\
    - k \int H^*dF_i - k^2 \int (G^*)^2dF_i - 1/12], \quad i = j, \\
    n[k(2 - \int F_jdF_i) \int F_j G^*dF_j + k(2 - \int F_idF_j) \int G^*dF_i \\
    - \sum_{m=1}^k \int F_mdF_i \int F_mdF_j - 2k \int F_j G^*dF_i - 2k \int F_idG^*dF_i \\
    + \int F_idF_j \int F_jdF_i + \int F_i^2dF_j + \int F_j^2dF_i - 1], \quad i \neq j, 
\end{cases} \quad (2.6)$$

### 2.2 Assumption

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

$$\varphi_\beta(\theta_i, \theta_j) = c_\beta + \kappa_\beta^{ij} n^{-1/2} + o(n^{-1/2}), \quad \beta = 1, 2, \quad (2.7)$$

where $c_\beta$ is given by (1.9); for each pair $(i, j)$, $\kappa_\beta^{ij}$ depends on $\theta_i$ and $\theta_j$ and is increasing in $\theta_i$ when $\theta_j$ is fixed, and decreasing in $\theta_j$ when $\theta_i$ is fixed; also, $\kappa_\beta^{ij} = c_\beta$ when $\theta_i = \theta_j$.

Then putting

$$I_\beta^{ij} = \sqrt{n} \{ \int F_j(x)dF_i(x) - \int F_i(x)dF_j(x) \}, \quad (2.8)$$

we obtain the following lemma.

**Lemma 2.1**

For $\varphi_\beta(\theta_i, \theta_j)$ ($\beta = 1, 2$) given by (2.7), we have the following:

$$I_\beta^{ij} = K_\beta^{ij} + o(1), \quad (2.9)$$

where
\[ k_{\beta ij} = \begin{cases} \kappa_{1ij} \int f^2(x)dx, & \text{when } \beta = 1, \\ \kappa_{2ij} \int xf^2(x)dx, & \text{when } \beta = 2, \end{cases} \tag{2.10} \]

\[ i,j = 1,2,\ldots,k; \ i \neq j. \]

**Example:**

When \( F(x) \) is \( N(0,1) \), \( \varphi_\beta(\theta, \theta) \) given by (2.8), we get

\[ I_{11ij} = (1/2\sqrt{\pi}) \kappa_{1ij} + o(1) \tag{2.11} \]

and

\[ I_{21ij} = (1/\pi) \kappa_{2ij} + o(1). \tag{2.12} \]

### 2.3 Asymptotic Distribution

Let us define

\[ \psi_1(\alpha) = (1/\sqrt{n})(H_k(\alpha) - H_1(\alpha)), \ \alpha = 1,2. \tag{2.13} \]

Then

\[ \psi(\alpha) = (1/\sqrt{n}) A \eta(\alpha), \ \alpha = 1,2, \tag{2.14} \]

where \( \psi(\alpha) = (\psi_1(\alpha), \psi_2(\alpha), \ldots, \psi_k(\alpha)) \), \( A = (-E(k-1), J(k)) \) \( (k-1) \times k \) and \( E(k-1) \) is a unit matrix of order \( k-1 \), \( J(k) = (1, 1, \ldots, 1)' \). \( \psi(\alpha) \) has mean vector \( \eta(\alpha) \), variance-covariance matrix \( \Sigma(\alpha) \) such that

\[ \eta(\alpha) = (1/\sqrt{n}) A \mu(\alpha), \tag{2.15} \]

and

\[ \Sigma(\alpha) = (1/n) A \Lambda(\alpha) A'. \tag{2.16} \]

Elements of \( \eta(\alpha) \) and \( \Sigma(\alpha) \) are given by

\[ \eta(\alpha)_{\beta i} = (1/\sqrt{n})(\mu(\alpha)_{\beta k} - \mu(\alpha)_{\beta i}), \ i = 1,2,\ldots,k-1, \tag{2.17} \]

\[ \sigma(\alpha)_{\beta ij} = (1/n)(\lambda(\alpha)_{\beta ij} - \lambda(\alpha)_{\beta ik} - \lambda(\alpha)_{\beta kj} + \lambda(\alpha)_{\beta kk}), \ i,j = 1,2,\ldots,k-1 \tag{2.18} \]
where \( \mu_\beta_i \) and \( \lambda_{\beta ij} \) are given by (2.1) through (2.6).

Now, under the assumption (2.7), using Lemma 2.1, we have for \( \beta = 1,2 \) and \( \alpha = 1,2 \),

\[
\eta_\beta i^{(\alpha)} = \frac{1}{\sqrt{n}} \left\{ n \int \sum_{j=1}^{k} F_j dF_k - n \int \sum_{j=1}^{k} F_i dF_i \right\} \\
= \frac{k-1}{\sum_{j=1}^{k} K_\beta kj} - \frac{k}{\sum_{j=1}^{k} K_\beta ij} \quad (= \gamma_\beta i^{(\alpha)}) \tag{2.19}
\]
as \( n \to \infty \), where \( K_\beta ij \) is given by (2.10). Also, under (2.7), we have

\[
\lambda_1 ij \to \begin{cases} 
-k/12 & \text{for } i \neq j \\
(k^2 - k)/12 & \text{for } i = j 
\end{cases}
\]

\[
\lambda_2 ij \to \begin{cases} 
-(k + 1)/12 & \text{for } i \neq j \\
(k^2 - 1)/12 & \text{for } i = j.
\end{cases}
\]

Consequently,

\[
\sigma_\beta ij \to \begin{cases} 
\nu_\alpha & \text{for } i \neq j \\
2\nu_\alpha & \text{for } i = j
\end{cases} \tag{2.20}
\]

where

\[
\nu_\alpha = \begin{cases} 
k^2/12 & \text{when } \alpha = 1 \\
k(k + 1)/12 & \text{when } \alpha = 2.
\end{cases} \tag{2.21}
\]

Thus by applying the central limit theorem, we have the following asymptotic distribution of \( \overline{y}^{(\alpha)} \):

\[
\overline{y}^{(\alpha)} \sim N(\gamma^{(\alpha)}_\beta, \Sigma^{(\alpha)}_\beta), \quad \beta = 1,2 \tag{2.22}
\]

where \( \gamma^{(\alpha)}_\beta = (\gamma^{(\alpha)}_\beta 1, \gamma^{(\alpha)}_\beta 2, \ldots, \gamma^{(\alpha)}_\beta (k-1))^\prime \) with elements given by (2.19) and

\[
\Sigma^{(\alpha)}_\beta = \nu_\alpha (F(k-1) + G(k-1)) \tag{2.23}
\]

where \( G(k-1) = \frac{1}{2}(k-1)^{1/2} (k-1) \).
3. PCS AND LFC

Using the asymptotic distribution of \( W^\alpha \beta \) \( (\alpha, \beta = 1, 2) \) given by (2.22), the PCS for the rule \( R(\alpha, \beta, m) \) \( (\alpha, \beta, m = 1, 2) \) is given by

\[
P(\text{CS}|R(\alpha, \beta, m)) = \Pr(\mathcal{U}_{\beta}^{\alpha} \approx -\delta(\beta, m) \mathcal{J}_{(k-1)})
\]

\[
\approx \Pr(\mathcal{U}_{\beta}^{\alpha} \approx (\mathcal{Y}_{\beta}^{\alpha} + \delta(\beta, m) \mathcal{J}_{(k-1)})/\sqrt{\alpha}) \tag{3.1}
\]

where

\[
\delta(\beta, m) = \begin{cases} 
\frac{d_{\beta}}{\sqrt{n}} & \text{when } m = 1 \\
0 & \text{when } m = 2
\end{cases} \tag{3.2}
\]

\[
\mathcal{U}_{\beta}^{\alpha} = (\mathcal{Y}_{\beta}^{\alpha} - \mathcal{Y}_{\beta})/\sqrt{\alpha}. \tag{3.3}
\]

\[
\mathcal{U}_{\beta}^{\alpha} \sim N(0_{(k-1)}, E_{(k-1)} + G_{(k-1)}), \tag{3.4}
\]

and \( a_n \approx b_n \) means that \( \lim a_n/b_n = 1 \) as \( n \to \infty \).

For the subset selection approach \( (m = 1) \), since

\[
\kappa_{\beta kj} - \kappa_{\beta ij} \not\approx 0
\]

and

\[
\kappa_{\beta kj} \not\approx 0
\]

for large \( n \), we have

\[
\mathcal{Y}_{\beta}^{(1)} \approx 0_{(k-1)}. \tag{1}
\]

Also, for indifference zone approach \( (m = 2) \), with the specification

\[
\Phi_{\beta}(\theta_k, \theta_i) \approx c_{\beta} + \delta_{\beta}^*,
\]

we have

\[
\mathcal{Y}_{\beta}^{(2)} \approx \begin{cases} 
(\int f^2(x)dx) \sqrt{n} \delta_1^*/\mathcal{J}_{(k-1)} & \text{when } \beta = 1 \\
(\int x f^2(x) dx) \sqrt{n} (\delta_2^*/(1 + \delta_2^*)) \mathcal{J}_{(k-1)} & \text{when } \beta = 2.
\end{cases}
\]

Thus we have the following
Theorem 3.1

Under the assumption of order restriction (2.7) and for large $n$, the (asymptotic) LFC of the PCS for rules $R(\alpha, \beta, 1)$ ($\alpha, \beta = 1, 2$) are given by

$$K_{\beta ji} = 0, \quad i, j = 1, 2, \ldots, k; \quad \alpha, \beta = 1, 2 \quad (3.5)$$

and for rules $R(\alpha, \beta, 2)$ ($\alpha, \beta = 1, 2$) are given by

$$K_{\beta ji} = c_{\beta}, \quad i, j = 1, 2, \ldots, k-1, \ i \neq j;$$

$$K_{\beta ki} = c_{\beta} + \delta_{\beta}^*, \quad i = 1, 2, \ldots, k-1; \quad \alpha, \beta = 1, 2. \quad (3.6)$$

Under the asymptotic LFC

$$P(CS|R(\alpha, \beta, m) \geq \Pr\{\bigcup_{\beta}^{(\alpha)} \left( \left( \gamma(\beta, m) + \delta(\beta, m) \right) / \sqrt{\alpha} \right) \bigcup_{k-1} \} \quad (3.7)$$

where $\nu_{\alpha}$ is given by (2.21), $\delta(\beta, m)$ is given by (3.2) and $\gamma(\beta, m)$ is defined by

$$\gamma(\beta, 1) = 0 \quad \text{for } \beta = 1, 2$$

$$\gamma(\beta, 2) = \begin{cases} 
( k \int f^2(x)dx ) \sqrt{n} \delta_1^* & \text{for } \beta = 1 \\
( k \int x f^2(x)dx ) \sqrt{n} ( \delta_2^* / (1 + \delta_2^*) ) & \text{for } \beta = 2.
\end{cases} \quad (3.8)$$

The expression in (3.7) can be rewritten as

$$P(CS|R(\alpha, \beta, m) \geq \int \Phi^{-1}\{x + (\gamma(\beta, m) + \delta(\beta, m)) / \sqrt{\alpha}\}d\Phi(x) \quad (3.9)$$

for $\alpha, \beta, m = 1, 2$, where $\Phi(x)$ is the standard normal c.d.f.

Determination of the $d_{\beta}$ Values

By Theorem 3.1, the $d_{\beta}$ values can be asymptotically expressed as

$$d_{\beta}(n) = \sqrt{n/2} \ d \sqrt{\alpha} + o(n^{1/2}), \quad \alpha, \beta = 1, 2 \quad (3.10)$$

where $d$ is the solution of the following equation:

$$Q(d/\sqrt{2}, \ d/\sqrt{2}, \ldots, \ d/\sqrt{2}) = p^* \quad (3.11)$$

$Q$ is the joint cdf of a normally distributed vector $(V_1, V_2, \ldots, V_{k-1})$ with

$E(V_i) = 0$, $\text{Var}(V_i) = 1$ and $\text{Cov}(V_i, V_j) = 1/2$, $i \neq j$. 

10
Determination of the Minimum Common Sample Size

In the subset selection approach, let $S$ denote the size of the selected subset and $E_{\theta} [S|R]$ the corresponding expected value when subset selection rule $R$ is applied and $\theta$ is the true nature of status. Then

$$E_{\theta} [S|R] = \sum_{j=1}^{k} P\{ \Pi_j \text{ is selected } | R \}$$

(3.12)

Having determined $d_1(n)$ from (3.10), one may determine the common sample size $n$ by imposing the additional requirement that

$$E_{\theta} [S|R] \leq 1 + \varepsilon$$

(3.13)

for some $0 < \varepsilon < k-1$, where $\theta$ lies in a given proper subset of the parameter space under study, for example, the subset defined by $\theta[1] = \theta[2] = \ldots = \theta[k-1] = \theta[k] - \varepsilon(n)$ for location parameters case and $\theta[1] = \theta[2] = \ldots = \theta[k-1] = \theta[k]/(1 + \Delta(n))$ for scale parameters, where $\varepsilon(n) > 0$ and $\Delta(n) > 0$.

Asymptotically Relative Efficiency

Given two selection procedures $R_1$ and $R_2$, the (asymptotic) efficiency of $R_2$ relative to $R_1$ is defined by

$$\text{Eff}(R_1,R_2) = n_1/n_2$$

(3.14)

where the $n_i$ are the sample sizes required to satisfy $P(\text{CS}|R_i)_LFC = P^*$, $i = 1,2$.

Then, by Theorem 3.1 and (3.10), we have the following result:

$$\text{Eff}(R(1,\beta,1),R(2,\beta,1)) = 1, \quad \beta = 1,2.$$  

$$\text{Eff}(R(\alpha,1,1),R(\alpha,2,1)) = 1, \quad \alpha = 1,2.$$  

$$\text{Eff}(R(1,\beta,2),R(2,\beta,2)) = k/(k + 1), \quad \beta = 1,2.$$  

$$\text{Eff}(R(\alpha,1,2),R(\alpha,2,2))$$

$$= \left( \frac{S_1^* S_2^*}{1 + S_2^*} \right)^2 \left( \frac{\int xf^2(x)dx}{\int f^2(x)dx} \right)^2, \quad \alpha = 1,2.$$  

(3.15)
4. APPENDIX

Let \( k \) population \( \Pi_1, \Pi_2, \ldots, \Pi_k \) be given, where \( \Pi_i \) has the associated continuous distribution \( F_s(x) \) (\( s = 1, 2, \ldots, k \)). Take \( n_s \) observations \( X_{s1}, X_{s2}, \ldots, X_{sn_s} \) from population \( \pi_s \) (\( s = 1, 2, \ldots, k \)) and consider the combined (Wilcoxon type) rank \( R_{sj} \) of \( X_{sj} \) as stated in Section 1. Then the means, variances and covariances of the ranks \( R_{sj} \) are given in the following.

**Theorem 4.1**

\[
E(R_{sj}) = N \int GdF_s + 1/2 \tag{4.1}
\]

\[
V(R_{sj}) = 2N \int GdF_s - 2N \int F_s GdF_s + N^2 \int G^2 dF_s
- N \int H dF_s - N^2 (\int GdF_s)^2 - 1/12 \tag{4.2}
\]

\[
\text{Cov}(R_{si}, R_{sj}) = 3N \int GdF_s - 4N \int F_s GdF_s - \sum_{m=1}^{k} n_m (\int F_m dF_s)^2 - 1/12 \tag{4.3}
\]

\[
\text{Cov}(R_{si}, R_{tj}) = N(2 - \int F_t dF_s) \int GdF_t + N(2 - \int F_s dF_t) \int GdF_s
- \sum_{m=1}^{k} n_m \int F_m dF_s \int F_m dF_t - 2N \int F_t GdF_s - 2N \int F_s GdF_t
+ \int F_s dF_t \int F_t dF_s + \int F_s^2 dF_t + \int F_t^2 dF_s - 1 \tag{4.4}
\]

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

\[
N = \sum_{m=1}^{k} n_m \tag{4.5}
\]

\[
G(x) = \frac{1}{N} \sum_{m=1}^{k} n_m F_m(x) \tag{4.6}
\]

\[
H(x) = \frac{1}{N} \sum_{m=1}^{k} n_m F_m^2(x) \tag{4.7}
\]

**Proof**

We sketch the proofs for (4.1) and (4.3) above. The remaining results are obtained similarly.
\[ \Pr(R_{11} = s) = \sum_{A} \Pr(a_1 \text{ of } X_1's, a_2 \text{ of } X_2's, \ldots, a_k \text{ of } X_k's \neq x_1 \neq (n_1-a_1-1) \text{ of } X_1's, (n_2-a_2) \text{ of } X_2's, \ldots, (n_k-a_k) \text{ of } X_k's) \]  
(4.8)

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

\[ 0 \leq a_i \leq n_i - 1, \quad 0 \leq a_i \leq n_i \quad (i = 2, 3, \ldots, k) \]  
(4.9)

\[ \sum_{j=1}^{k} a_j = s - 1 \]  
(4.10)

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, summation \( \sum_{A} \) is taken over all \( k \)-tuples \( (a_1, a_2, \ldots, a_k) \) of integers which satisfy the relations (4.9) and (4.10). From (4.8), we have

\[ E(R_{11}) = \int \sum_{s=1}^{N} \sum_{A} s \binom{n_1-1}{a_1} \binom{n_2}{a_2} \ldots \binom{n_k}{a_k} f_1^{a_1} f_2^{a_2} \ldots f_k^{a_k} \times (1-f_1)^{n_1-a_1-1} (1-f_2)^{n_2-a_2} \ldots (1-f_k)^{n_k-a_k} dF_1 \]  
(4.11)

By changing the order of summation, we have

\[ E(R_{11}) = \int \sum_{s=1}^{N} \sum_{A_i} s \binom{n_1-1}{a_1} \binom{n_2}{a_2} \ldots \binom{n_k-1}{a_k-1} f_1^{a_1} f_2^{a_2} \ldots f_k^{a_k-1} \times (1-f_1)^{n_1-a_1-1} (1-f_2)^{n_2-a_2} \ldots (1-f_k)^{n_k-a_k-1} (n_k F_k + \sum_{i=1}^{k} a_i + 1) dF_1 \]

where the summation \( \sum_{A_i} \) is taken over all \( (k-1) \)-tuples \( (a_1, a_2, \ldots, a_{k-1}) \) of integers which satisfy the relation (4.9). 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

\[ \Pr(R_{11} = s, R_{21} = t) = \sum_{B} \Pr(a_1 \text{ of } X_1's, a_2 \text{ of } X_2's, \ldots, a_k \text{ of } X_k's \neq X_1 \neq \neq b_1 \text{ of } X_1's, b_2 \text{ of } X_2's, \ldots, b_k \text{ of } X_k's \neq c_1 \text{ of } X_1's, c_2 \text{ of } X_2's, \ldots, c_k \text{ of } X_k's) \]  
(4.12)

where \( a_i, b_i, c_i \) (i = 1, 2, \ldots, k) are integers such that
\[ a_i + b_i + c_i = \nu_i, \quad i = 1, 2, \ldots, k \] (4.13)

\[ \frac{k}{j=1} a_j = s - 1, \quad \frac{k}{j=1} b_j = t - s - 1, \quad \frac{k}{j=1} c_j = n - t \] (4.14)

and \( \nu_i = n_i - 1 \) for \( i = 1, 2 \), \( \nu_i = n_i \) for \( i = 3, 4, \ldots, k \).

Summation \( \sum_{B_i} \) is taken over all tuples \((a_1, \ldots, a_k, b_1, \ldots, b_k, c_1, \ldots, c_k)\)

which satisfy the relations (4.13) and (4.14). Then

\[
I_1 = \sum_{s < t} \Pr(R_{11} = s, R_{12} = t)
= \int \int \sum_{x < y} \sum_{s < t} \prod_{i=1}^{k} P_i(x, y) dF_1(x) dF_2(y)
\] (4.15)

where

\[
P_i(x, y) = \left( \begin{array}{c} \nu_i \\ a_i, b_i, c_i \end{array} \right) F_i(x) (F_i(y) - F_i(x))^b_i (1 - F_i(y))^c_i, \quad i = 1, 2, \ldots, k.
\]

By changing the order of summation, first for \( s \) then for \( t \), we have

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

where

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

and

\[
\alpha_1 = n_k (n_k - 1) F_k(x) F_k(y) + 3n_k F_k(x) + n_k F_k(y) + 2
\]

\[
\beta_1 = n_k F_k(x) + n_k F_k(y) + 3
\]

\[
\gamma_1 = n_k F_k(x) + 1.
\]

Summation \( \sum_{B_i} \) is taken over all tuples \((a_1, \ldots, a_{k-1}, b_1, \ldots, b_{k-1}, c_1, \ldots, c_{k-1})\) which satisfies the condition (4.13). By carrying out the addition in turn for a set \((a_i, b_i, c_i)\), \( i = k-1, k-2, \ldots, 1 \), we have a reduced form of \( I_1 \). By proceeding on similar steps for \( \sum_{s > t} \Pr(R_{11} = s, R_{21} = t) \), we obtain \( \text{Cov}(R_{11}, R_{21}) \).

For rank sums

\[
T_s = \sum_{j=1}^{n_s} R_{sj}, \quad s = 1, 2, \ldots, k
\] (4.16)
we have

\[ E(T_s) = n_s E(R_{sj}), \quad s = 1, 2, ..., k \]  
(4.17)

\[ \text{Cov}(T_s, T_t) = n_s n_t \text{Cov}(R_{sj}, R_{tj}), \quad s, t = 1, 2, ..., k, \quad s \neq k, \]  
(4.18)

and

\[ V(T_s) = \sum_{j=1}^{n_s} V(R_{sj}) + \sum_{i \neq j}^{n_s} \text{Cov}(R_{si}, R_{sj}) \]

\[ = n_s (3n_s - 1) f GdF_s - 2n_s (2n_s - 1) f F_s GdF_s + n_s f G^2 dF_s \]
\[ - n_s f HdF_s - n_s^2 (f GdF_s)^2 - n_s (n_s - 1) \sum_{m=1}^{k} n_m^2 (f F_m dF_s)^2 - \frac{n_s^2}{12}. \]  
(4.19)

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

\[ E(T_s) = n_s (N+1)/2. \]  
(4.20)

\[ V(T_s) = n_s (N-n_s) (N+1)/12. \]  
(4.21)

\[ \text{Cov}(T_s, T_t) = -n_s n_t (N+1)/12. \]  
(4.22)

Also for \( k = 2 \), we have the following:

\[ E(T_1) = n_1 (n_1 + 1)/2 + n_1 n_j f F_j dF_1, \quad i, j = 1, 2; \quad j \neq i \]  
(4.23)

\[ V(T_1) = n_1 n_j (2n_1 - 1) f F_j dF_1 + n_1 n_j (n_j - 1) f F_j^2 dF_1 \]
\[ + n_1 n_j (n_j - 1) f F_j^2 dF_j - n_1 n_j (n_1 + n_j - 1) (f F_j dF_1)^2 - n_1 n_j (n_1 - 1) \]
\[ i, j = 1, 2; \quad i \neq j. \]  
(4.24)

\[ \text{Cov}(T_1, T_2) = n_1 n_2 [n_1 f F_1 dF_2 + n_2 f F_2 dF_1 - (n_1 + n_2 - 1) f F_1 dF_2 / F_2 dF_1 \]
\[ - (n_1 - 1) f F_2^2 dF_2 - (n_2 - 1) f F_2^2 dF_1 - 1]. \]  
(4.25)

Finally, we state an order relation for the expected rank sum. Let \( \{ F_\theta (x) \} \) be a family of distributions stochastically increasing in \( \theta \). Then we have the following.

**Theorem 4.2**

\[ E(R_s) \preceq E(R_t) \text{ if and only if } F_s \preceq F_t \text{ where } s, t = 1, 2, ..., k. \]
REFERENCES


Technical Report #87-2

ON THE LEAST FAVORABLE CONFIGURATION OF A SELECTION PROCEDURE BASED ON RANKS

Shanti S. Gupta and Takashi Matsui

Purdue University
Department of Statistics
West Lafayette, IN 47907

Office of Naval Research
Washington, DC

Approved for public release, distribution unlimited.

Selection Procedures, Combined-Ranks, Least Favorable Configuration, Moments of Ranks, Wilcoxon Type, Friedman Type

We consider two types of statistics based on the sums of combined (Wilcoxon type) ranks and vector (Friedman type) ranks. Underlying populations are supposed to belong to the location or scale parameter family of distributions.

Two approaches - subset selection and indifference zone - of ranking and selection procedures based on these statistics are considered in an asymptotic framework for selecting the population with the largest parameter value. The least favorable configurations of parameters are discussed by computing the
exact moments of these statistics and introducing an assumption of order relation between the gaps of parameters.