AN ASYMPTOTIC DISTRIBUTION-FREE SELECTION PROCEDURE FOR A TWO-WAY LAYOUT PROBLEM*

by

Shanti S. Gupta
Purdue University

Lii-Yuh Leu
National Central University

Technical Report#83-41

Department of Statistics
Purdue University

Revised May 1987
September 1983

* The research was supported by the Office of Naval Research Contracts N00014-75-C-0455 and N00014-84-C-0167 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

o

AN ASYMPTOTIC DISTRIBUTION-FREE SELECTION PROCEDURE FOR A TWO-WAY LAYOUT PROBLEM

Shanti S. Gupta
Purdue University

Lii-Yuh Leu
National Central University

Key Words and Phrases: asymptotic distribution-free procedure; subset selection; two-way layout; treatment effect; nuisance parameter; asymptotic relative efficiency.

ABSTRACT

This paper deals with an asymptotic distribution-free subset selection procedure for a two-way layout problem. The treatment effect with the largest unknown value is of interest to us. The block effect is a nuisance parameter in this problem. The proposed procedure is based on the Hodges-Lehmann estimators of location parameters. The asymptotic relative efficiency of the proposed procedure with the normal means procedure is evaluated. It is shown that the proposed procedure has a high efficiency.

1. INTRODUCTION

Consider a two-factor complete block design with one observation per cell. Let the observable random variables be $X_{i\alpha}$, $i = 1, 2, \ldots, k$; $\alpha = 1, 2, \ldots, n$ and consider the linear model

$$X_{i\alpha} = \mu + \theta_i + \beta_\alpha + \varepsilon_{i\alpha}, \qquad \sum_{i=1}^{k} \theta_i = 0, \tag{1.1}$$

where $X_{i\alpha}$ is the observation under treatment i in the α^{th} block, μ is the mean-effect, θ_i is the effect of treatment i, β_α is the block effect for the α^{th} block (nuisance parameter), and the $\varepsilon_{i\alpha}$, $\alpha = 1, 2, \ldots, n$ are error components. It is assumed that the error components are independent and identically distributed with a continuous cumulative distribution function (cdf) $F(\underline{\varepsilon})$, $\underline{\varepsilon} \in \mathbb{R}^k$ (the real k-space), where $F(\underline{\varepsilon})$ is symmetric in its arguments. That is, for any $\underline{\varepsilon} \in \mathbb{R}^k$ and any permutation $(i_1, \ldots, i_k)$ of $(1, \ldots, k)$, we have

$$F(\varepsilon_1, \ldots, \varepsilon_k) = F(\varepsilon_{i_1}, \ldots, \varepsilon_{i_k}). \tag{1.2}$$

Let $\theta_{[1]} \le \cdots \le \theta_{[k]}$ be the ordered θ_i's. Suppose that we are interested in the treatment with the largest unknown parameter $\theta_{[k]}$ (if more treatments than one have θ_i equal to $\theta_{[k]}$, then exactly one of these treatments is "tagged" as the best treatment). Correct selection denotes the selection of any subset containing the population with $\theta_{[k]}$ (or the "tagged" population). For the nonparametric approach to the one-way layout problems, most previous authors have considered procedures based on a class of rank order statistics. It was pointed out by Rizvi and Woodworth (1970) that there are difficulties associated with these procedures, mainly because the least favorable configuration is usually not known. Randles (1970) and Ghosh (1973) have considered the procedures for the one-way layout based on the Hodges-Lehmann estimators. It was shown by them that the procedures based on the Hodges-Lehmann estimators have high efficiency. Hsu (1982) considered the two-way layout problem with independent errors. In this paper we assume the errors are equally correlated and use the results of Puri and Sen (1967) to derive a subset selection procedure for the largest unknown parameter $\theta_{[k]}$. Also, we use

Hodges-Lehmann estimators derived from signed-ranks rather than the ranks themselves as in Hsu (1982).

2. ROBUST COMPATIBLE ESTIMATION

In the model (1.1), for $1 \leq i, j \leq k$, $i \neq j$, let $X_{ij,\alpha} = X_{i\alpha} - X_{j\alpha}$, $e_{ij,\alpha} = \varepsilon_{i\alpha} - \varepsilon_{j\alpha}$, $\alpha = 1, 2, \ldots, n$ and $\Delta_{ij} = \theta_i - \theta_j$. Then for a fixed α, we can write

$$X_{ij,\alpha} = \Delta_{ij} + e_{ij,\alpha}. \tag{2.1}$$

From Assumption (1.2), $e_{ij,\alpha}$ have common distribution, say G, which is symmetric about zero. Hence $X_{ij,1}, \ldots, X_{ij,n}$ are i.i.d. with common cdf $G(x - \Delta_{ij})$. We assume that G is continuous, but otherwise unknown. Let $R_{ij,\alpha} =$ Rank of $|X_{ij,\alpha}|$ among $|X_{ij,1}|, \ldots, |X_{ij,n}|$ and let $\underline{X}_{ij} = (X_{ij,1}, \ldots, X_{ij,n})$. Consider the one-sample signed rank statistic

$$h_{ij,n}(\underline{X}_{ij}) = n^{-1} \sum_{\alpha=1}^{n} E_{n,\alpha} Z_{n,\alpha} \tag{2.2}$$

where $Z_{n,\alpha}$ is either one or zero as follows: if the α^{th} smallest observation among $|X_{ij,1}|, \ldots, |X_{ij,n}|$ corresponds to $X_{ij,t}$ (for some t), then $Z_{n,\alpha} = 1$ if $X_{ij,t} > 0$ or 0 if $X_{ij,t} < 0$. $E_{n,\alpha}$ is the expected value of the α^{th} order statistic of a sample of size n from a distribution $\psi^*(x)$ given by

$$\psi^*(x) = \begin{cases} \psi(x) - \psi(-x) & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{cases} \tag{2.3}$$

Throughout this paper, we shall assume that $\psi(x)$ and $G(x)$ satisfy the following assumptions (see Puri (1964) and Puri and Sen (1967)):

(1) $\psi(x)$ is a distribution function symmetric about $x = 0$, that is $\psi(x) + \psi(-x) = 1$.

(2) $\frac{1}{n} \sum_{\alpha=1}^{n} [E_{n,\alpha} - \psi^{-1}(\frac{\alpha}{n+1})] Z_{n,\alpha} = 0_p(n^{-\frac{1}{2}})$.

(3) $J(u) = \psi^{-1}(u)$, $0 \leq u \leq 1$ is absolutely continuous and $|J^{(i)}(u)| = |d^i J(u)/du^i| \leq M[u(1-u)]^{-i-\frac{1}{2}+\delta}$, $i = 0, 1, 2$, for some M and some $\delta > 0$.

(4) G is a continuous cdf, differentiable in each of the open intervals $(-\infty, a_1)$, (a_1, a_2), $\ldots$, (a_{s-1}, a_s), (a_s, ∞), for some $a_1, \ldots, a_s$ and the derivative of G is bounded in each of these intervals.

(5) The function $\frac{d}{dx}J(G(x))$ is bounded as $x \to \pm\infty$.

It is easy to see that $h_{ij,n}(x_{ij,1}+a,\ldots,x_{ij,n}+a)$ is a non-decreasing function of a for fixed $\underline{x}_{ij}$ and when $\Delta_{ij}=0$, the distribution of $h_{ij,n}$ is symmetric about a fixed point $\mu = \frac{1}{2}E_{\psi}|V|$, where V has cdf ψ.

Let

$$\begin{aligned} \Delta_{ij}^{*} &= \sup\{\Delta : h_{ij,n}(\underline{x}_{ij} - \Delta\underline{1}) > \mu\}, \\ \Delta_{ij}^{**} &= \inf\{\Delta : h_{ij,n}(\underline{x}_{ij} - \Delta\underline{1}) < \mu\} \end{aligned} \tag{2.4}$$

and let

$$\hat{\Delta}_{ij} = \tfrac{1}{2}(\Delta_{ij}^{*} + \Delta_{ij}^{**}).$$

Now from Hodges and Lehmann (1963), $\hat{\Delta}_{ij}$ is a translation invariant robust estimator of Δ_{ij} and has a distribution which is symmetric about Δ_{ij}. Note that the estimates $\hat{\Delta}_{ij}$ are incompatible (see Lehmann (1964)) in the sense that they do not satisfy the linear relations satisfied by the differences they estimate. This leads to certain ambiguities. To derive compatible estimators, let

$$\hat{\Delta}_{i} = \frac{1}{k}\sum_{\ell=1}^{k}\hat{\Delta}_{i\ell}, \qquad \hat{\Delta}_{ii} \equiv 0 \qquad \text{for } i = 1, 2, \ldots, k. \tag{2.5}$$

Then by minimizing $\sum\limits_{i\neq j}(\hat{\Delta}_{ij} - \Delta_{ij})^2$ with respect to θ's, we obtain the compatible or adjusted estimators of Δ_{ij} as

$$Z_{ij} = \hat{\Delta}_{i} - \hat{\Delta}_{j}, \qquad i \neq j. \tag{2.6}$$

Note that $E(\hat{\Delta}_i) = \theta_i$ since $\sum\limits_{i=1}^{k}\theta_i = 0$, hence $E(Z_{ij}) = \Delta_{ij}$. Puri and Sen (1967) have proved the following theorem:

<u>Theorem 2.1</u>. The joint distribution of $\{n^{\frac{1}{2}}(Z_{ik} - \Delta_{ik});\ i = 1, 2, \ldots, k-1\}$ is asymptotically normal with zero means and a covariance matrix $\Gamma = (\gamma_{ij})$, $i, j =$

1, 2, ..., $k-1$ where

$$\gamma_{ij} = \begin{cases} 2\sigma_0^2 & \text{if } i = j \\ \sigma_0^2 & \text{if } i \neq j \end{cases} \tag{2.7}$$

and $\sigma_0^2 = [A^2 + (k-2)\lambda_J(G)]/kB^2$, where

$$A^2 = \int_0^1 J^2(u)\,du, \qquad B = \int_{-\infty}^{\infty} \frac{d}{dx} J\big(G(x)\big)\,dG(x),$$
$$\text{and } \lambda_J(G) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J\big(G(x)\big) J\big(G(y)\big)\,dG^*(x,y), \tag{2.8}$$

$G^*(x,y)$ is the joint cdf of $e_{ij,\alpha}$ and $e_{i\ell,\alpha}$ $(j \neq \ell)$ whose marginal cdf's are $G(x)$ and $G(y)$, respectively.

Moreover, using the translation invariant property of $\hat{\Delta}_{ij}$, we have the following lemma:

Lemma 2.2.

(1) $\hat{\Delta}_i(x_{11} - c_1, \ldots, x_{1n} - c_1, \ldots, x_{k1} - c_k, \ldots, x_{kn} - c_k)$
$= \hat{\Delta}_i(x_{11}, \ldots, x_{1n}, \ldots, x_{k1}, \ldots, x_{kn}) + \bar{c} - c_i$, where $\bar{c} = \frac{1}{k}\sum_{i=1}^{k} c_i$.

(2) The distribution of $\hat{\Delta}_i - \theta_i$ is independent of $\underline{\theta}$.

3. A NONPARAMETRIC PROCEDURE FOR SELECTING THE BEST TREATMENT

Based on the estimators defined in (2.5), we propose a selection procedure R_1 as follows:

$$R_1\text{: Select treatment } i \text{ iff } \hat{\Delta}_i \geq \max_{1 \leq j \leq k} \hat{\Delta}_j - d_1, \tag{3.1}$$

where $d_1 > 0$ is determined so as to satisfy the basic probability requirement. The value of d_1 can be determined asymptotically and will be discussed later.

Let $\Omega = \{\underline{\theta} = (\theta_1, \ldots, \theta_k) \mid \sum_{i=1}^{k} \theta_i = 0\}$ be the parameter space and let CS stand for a correct selection which means that the selected subset contains the best treatment. For a given constant $P^*(k^{-1} < P^* < 1)$, the basic probability requirement is

$$\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS \mid R_1) \geq P^*.$$

Let $\hat{\Delta}_{[1]} \leq \cdots \leq \hat{\Delta}_{[k]}$ denote the ordered $\hat{\Delta}_i$'s and $\hat{\Delta}_{(i)}$ denote the unknown estimator associated with the parameter $\theta_{[i]}$, $1 \leq i \leq k$. Let $P_j(\underline{\theta} \mid R_1)$ denote the probability that the treatment (j) is selected (treatment (j) is associated with parameter $\theta_{[j]}$) for the selection procedure R_1 when $\underline{\theta}$ is the true state of nature. We have the following lemma:

Lemma 3.1. For $\underline{\theta}$, $\underline{\theta}^* \in \Omega$ and fixed j, if $\theta^*_{[j]} - \theta^*_{[i]} \geq \theta_{[j]} - \theta_{[i]}$, $i \neq j$, then $P_j(\underline{\theta}^* \mid R_1) \geq P_j(\underline{\theta} \mid R_1)$.

Proof. $P_j(\underline{\theta} \mid R_1) = P_{\underline{\theta}}(\hat{\Delta}_{(j)} \geq \max_{1\leq i\leq k} \hat{\Delta}_{(i)} - d_1)$

$$= P_{\underline{\theta}}(\hat{\Delta}_{(i)} - \theta_{[i]} - \hat{\Delta}_{(j)} + \theta_{[j]} \leq d_1 + \theta_{[j]} - \theta_{[i]},\ i \neq j, i = 1, \ldots, k).$$

By Lemma 2.2, the distribution of $\hat{\Delta}_{(i)} - \theta_{[i]} - \hat{\Delta}_{(j)} + \theta_{[j]}$, $i \neq j$, $i = 1, 2, \ldots, k$ is independent of $\underline{\theta}$. Hence $P_j(\underline{\theta} \mid R_1) \leq P_j(\underline{\theta}^* \mid R_1)$ if $\theta^*_{[j]} - \theta^*_{[i]} \geq \theta_{[j]} - \theta_{[i]}$, $i \neq j$.

Corollary 3.2. $\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS \mid R_1) = P_{\underline{0}}(CS \mid R_1)$ where $\underline{0} = (0, \ldots, 0)$ (k-tuple).

Proof. Since $P_{\underline{\theta}}(CS \mid R_1) = P_k(\underline{\theta} \mid R_1)$ and $\theta_{[k]} - \theta_{[i]} \geq 0$, $i \neq k$, by using Lemma 3.1, the result follows.

For large sample we can define d_1 as in Theorem 3.3 given below.

Theorem 3.3. For given $P^*(k^{-1} < P^* < 1)$, if $\sigma_0^2 < \infty$, we have

$$d_1(n) = n^{-\frac{1}{2}} d\sigma_0 + o(n^{-\frac{1}{2}}) \qquad \text{as } n \to \infty, \tag{3.2}$$

where d is the solution to the equation

$$Q(d/\sqrt{2}, \ldots, d/\sqrt{2}) = P^*, \tag{3.3}$$

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

$$E(V_i) = 0, \quad \text{Var}(V_i) = 1 \quad \text{and} \quad \text{Cov}(V_i, V_j) = 1/2, \qquad i \neq j. \tag{3.4}$$

Proof. By Theorem 2.1 and Corollary 3.2,

$$\begin{aligned}\lim_{n\to\infty} \inf_{\underline{\theta}\in\Omega} P_\theta(CS \mid R_1) &= \lim_{n\to\infty} P_{\underline{0}}(\hat{\Delta}_{(i)} - \hat{\Delta}_{(k)} \leq d_1(n),\ i = 1, 2, \ldots, k-1) \\ &= P\big(V_i \leq \lim_{n\to\infty} n^{\frac{1}{2}} d_1(n)/\sqrt{2}\,\sigma_0,\ i = 1, 2, \ldots, k-1\big).\end{aligned}$$

Therefore, if d is the solution of (3.3), then

$$d_1(n) = n^{-\frac{1}{2}} d\sigma_0 + o(n^{-\frac{1}{2}}) \qquad \text{as } n \to \infty.$$

Remark. The solution of (3.3) is also a solution of $\int_{-\infty}^{\infty} \Phi^{k-1}(x+d)d\Phi(x) = P^*$, where Φ is the cdf of standard normal. This has been shown by many authors (see for example, Gupta (1963)).

Determination of the Minimum Common Sample Size

Let $E_{\underline{\theta}}(S \mid R_1)$ denote the expected size of the selected subset using rule R_1 given $\underline{\theta}$. Then $E_{\underline{\theta}}(S \mid R_1) = \sum_{j=1}^{k} P_j(\underline{\theta} \mid R_1)$. Having determined $d_1(n)$ from (3.2), one may determine the common sample size n by imposing the additional requirement that $E_{\underline{\theta}}(S \mid R_1) \leq 1 + \varepsilon$, for some $\varepsilon > 0$, whenever $\underline{\theta}$ lies in a given proper subset of Ω, for example, the subset defined by

$$\theta_{[1]} = \cdots = \theta_{[k-1]} = \theta_{[k]} - \delta^*, \qquad \delta^* > 0. \tag{3.5}$$

It will be convenient in the sequel to replace (3.5), when the sample size is n, by

$$\theta_{[1]} = \cdots = \theta_{[k-1]} = \theta_{[k]} - \delta^{(n)}. \tag{3.6}$$

(see Bartlett and Govindarajulu (1968)).

Theorem 3.4. For given $\varepsilon > 0$, with $d_1(n)$ given by (3.2) and n determined by $E_{\underline{\theta}}(S \mid R_1) \le 1 + \varepsilon$ for $\underline{\theta}$ satisfying (3.6). Then as $n \to \infty$,

$$\delta^{(n)} = n^{-\frac{1}{2}} c(\varepsilon)\sigma_0 + o(n^{-\frac{1}{2}}), \tag{3.7}$$

where $c(\varepsilon)$ is the solution to the equation

$$Q((c+d)/\sqrt{2}, \ldots, (c+d)/\sqrt{2}) + (k-1)Q(d/\sqrt{2}, \ldots, d/\sqrt{2}, (d-c)/\sqrt{2}) = 1+\varepsilon, \tag{3.8}$$

where Q is defined as in Theorem 3.3.

Proof.

$$\begin{aligned} E_{\underline{\theta}}(S \mid R_1) = \sum_{j=1}^{k-1} P_{\underline{\theta}}\Big(& n^{\frac{1}{2}}(\hat{\Delta}_{(i)} - \hat{\Delta}_{(j)})/\sqrt{2}\sigma_0 \le n^{\frac{1}{2}} d_1(n)/\sqrt{2}\sigma_0, \; i = 1, \ldots, k-1, \\ & i \ne j, \; n^{\frac{1}{2}}(\hat{\Delta}_{(k)} - \hat{\Delta}_{(j)} - \delta^{(n)})/\sqrt{2}\sigma_0 \le n^{\frac{1}{2}}(d_1(n) - \delta^{(n)})/\sqrt{2}\sigma_0\Big) \\ + P_{\underline{\theta}}\Big(& n^{\frac{1}{2}}(\hat{\Delta}_{(i)} - \hat{\Delta}_{(k)} + \delta^{(n)})/\sqrt{2}\sigma_0 \le n^{\frac{1}{2}}(d_1(n) + \delta^{(n)})/\sqrt{2}\sigma_0, \\ & i = 1, \ldots, k-1\Big). \end{aligned}$$

If $\underline{\theta}$ satisfies (3.6), then

$$\begin{aligned} \lim_{n\to\infty} E_{\underline{\theta}}(S \mid R_1) &= \sum_{j=1}^{k-1} P\Big(V_i \le \lim_{n\to\infty} n^{\frac{1}{2}} d_1(n)/\sqrt{2}\sigma_0, \; i = 1, \ldots, k-1, i \ne j, \\ &\qquad V_{k-1} \le \lim_{n\to\infty} n^{\frac{1}{2}}(d_1(n) - \delta^{(n)})/\sqrt{2}\sigma_0\Big) \\ &\quad + P\Big(V_i \le \lim_{n\to\infty} n^{\frac{1}{2}}(d_1(n) + \delta^{(n)})/\sqrt{2}\sigma_0, \; i = 1, \ldots, k-1\Big) \\ &= (k-1)Q(d/\sqrt{2}, \ldots, d/\sqrt{2}, (d-c)/\sqrt{2}) + Q((d+c)/\sqrt{2}, \ldots, (d+c)/\sqrt{2}) \\ &= 1 + \varepsilon. \end{aligned}$$

Hence $c(\varepsilon)$ is the solution of (3.8) iff

$$\delta^{(n)} = n^{-\frac{1}{2}} c(\varepsilon)\sigma_0 + o(n^{-\frac{1}{2}}) \qquad \text{as } n \to \infty.$$

Remark. The common sample size n required to satisfy $\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS \mid R_1) = P^*$ and $E_{\underline{\theta}}(S \mid R_1) \leq 1+\varepsilon$ for $\underline{\theta}$ satisfying (3.5) is $(c(\varepsilon)\sigma_0/\delta^*)^2$. Note that n is a function of k, P^*, δ^*, and ε.

4. A SELECTION PROCEDURE FOR THE NORMAL CASE

In the following we assume that $(\varepsilon_{1\alpha},\ldots,\varepsilon_{k\alpha})$ are jointly normally distributed with zero means and the covariance matrix $\sigma^2\begin{pmatrix} 1 & & \rho \\ & \ddots & \\ \rho & & 1\end{pmatrix}$, $\alpha = 1, 2, \ldots, n$ where $-1/(k-1) < \rho < 1$ is known and $\sigma^2 < \infty$, may be known or unknown. Let $\overline{X}_i = \frac{1}{n}\sum_{\alpha=1}^{n} X_{i\alpha}$. Then the vector $(\overline{X}_1 - \overline{X}_k,\ldots,\overline{X}_{k-1} - \overline{X}_k)$ has a joint normal distribution with mean vector $(\theta_1-\theta_k,\ldots,\theta_{k-1}-\theta_k)$ and the covariance matrix

$$\frac{2\sigma^2(1-\rho)}{n}\begin{pmatrix} 1 & & \frac{1}{2} \\ & \ddots & \\ \frac{1}{2} & & 1\end{pmatrix}. \tag{4.1}$$

We assume that σ^2 is known and propose a selection procedure R_2 by

$$R_2\text{: Select treatment } i \text{ iff } \overline{X}_i \geq \max_{1\leq j\leq k} \overline{X}_j - d_2. \tag{4.2}$$

It is easy to see that $\inf_{\underline{\theta}\in\Omega} P_{\underline{\theta}}(CS \mid R_2) = P_{\underline{0}}(CS \mid R_2)$.
Thus, similar to Theorem 3.3 and Theorem 3.4 we have the following theorems:

Theorem 4.1. For given $P^*(k^{-1} < P^* < 1)$ and any sample size n, let $d_2(n)$ be chosen to satisfy $P_{\underline{0}}(CS \mid R_2) = P^*$. Then

$$d_2(n) = n^{-\frac{1}{2}} d\sqrt{1-\rho}\,\sigma, \tag{4.3}$$

where d is the solution of (3.3).

Theorem 4.2. For given $\varepsilon > 0$, let the sample size n be determined so that $d_2(n) = n^{-\frac{1}{2}} d\sqrt{1-\rho}\,\sigma$ and $E_{\underline{\theta}}(S \mid R_2) \leq 1+\varepsilon$ for $\underline{\theta}$ satisfying (3.6). Then as $n \to \infty$,

$$\delta^{(n)} = n^{-\frac{1}{2}} c(\varepsilon)\sqrt{1-\rho}\,\sigma, \tag{4.4}$$

where $c(\varepsilon)$ is the solution of (3.8).

Suppose that the joint distribution F of $(\varepsilon_{1\alpha}, \ldots, \varepsilon_{k\alpha})$ is unknown, but the variance of $\varepsilon_{1\alpha}$ is finite. By the central limit theorem, the joint distribution of $\{n^{\frac{1}{2}}(\overline{X}_i - \overline{X}_k - \theta_i + \theta_k);\ i = 1, 2, \ldots, k-1\}$ is asymptotically normal with zero means and the covariance matrix (4.1). We can still use the procedure R_2 given by (4.2). For large samples we have $d_2(n) = n^{-\frac{1}{2}} d\sqrt{1-\rho}\,\sigma + o(n^{-\frac{1}{2}})$ and $\delta^{(n)} = n^{-\frac{1}{2}} c(\varepsilon)\sqrt{1-\rho}\,\sigma + o(n^{-\frac{1}{2}})$, where d is the solution of (3.3) and $c(\varepsilon)$ is the solution of (3.8).

5. ASYMPTOTIC RELATIVE EFFICIENCY OF R_1 TO R_2

For any two procedures R_1 and R_2 satisfying the basic P^*-condition, let us define the asymptotic relative efficiency, say $\text{ARE}(R_1, R_2) = \lim_{\varepsilon\downarrow 0} n_{R_2}(\varepsilon)/n_{R_1}(\varepsilon)$ for the given parametric configuration (3.5), where $n_{R_i}(\varepsilon)$, $i = 1, 2$ are the sample sizes required to achieve the same expected size, $1+\varepsilon$. Then we have the following theorem:

Theorem 5.1. $\text{ARE}(R_1, R_2) = \{2\sigma^2(1-\rho)B^2/A^2\}\{kA^2/2[A^2 + (k-2)\lambda_J(G)]\}$, where A^2, B and $\lambda_J(G)$ are defined in (2.8).

Proof. For procedure R_1, putting $\delta^{(n)} = \delta^*$, from (3.7) we have

$$n_{R_1}(\varepsilon) = \left(c(\varepsilon)\sigma_0/\delta^*\right)^2,$$

where $c(\varepsilon)$ is the solution of (3.8). (Note that $\varepsilon \downarrow 0$, then $n \to \infty$). Similarly, for procedure R_2, we have

$$n_{R_2}(\varepsilon) = \left(c(\varepsilon)\sqrt{1-\rho}\,\sigma/\delta^*\right)^2.$$

Hence

$$\begin{aligned}\text{ARE}(R_1, R_2) &= (1-\rho)\sigma^2/\sigma_0^2 \\ &= \{2\sigma^2(1-\rho)B^2/A^2\}\{kA^2/2[A^2 + (k-2)\lambda_J(G)]\}.\end{aligned}$$

Remarks.

(1) Barlow and Gupta (1969) define $\text{ARE}(R_1, R_2) = \lim_{\varepsilon \downarrow 0} n_{R_2}(\varepsilon)/n_{R_1}(\varepsilon)$ for the given parametric configuration (3.5), where $n_{R_i}(\varepsilon)$, $i = 1, 2$ are the sample sizes required to achieve the same expected size (say ε) of non-best populations selected. If we consider the case where expected size refers only to the number of non-best populations in the selected subset, we have $n_{R_1}(\varepsilon) = (c'(\varepsilon)\sigma_0/\delta^*)^2$, where $c'(\varepsilon)$ is the solution to the equation

$$(k-1)Q(d/\sqrt{2}, \ldots, d/\sqrt{2}, (d-c')/\sqrt{2}) = \varepsilon.$$

Similarly, we have $n_{R_2}(\varepsilon) = (c'(\varepsilon)\sqrt{1-\rho}\,\sigma/\delta^*)^2$, and hence

$$\text{ARE}(R_1, R_2) = (1-\rho)\sigma^2/\sigma_0^2$$

which is the same as in Theorem 5.1.

(2) Puri and Sen (1967) proved that $\lambda_J(G) \le \frac{1}{2}A^2$, hence $kA^2/2[A^2+(k-2)\lambda_J(G)] \ge 1$ and $\text{ARE}(R_1, R_2) \ge 2(1-\rho)\sigma^2 B^2/A^2$. The variance of G is $2(1-\rho)\sigma^2$, hence $2(1-\rho)\sigma^2 B^2/A^2$ is the ARE of the one-sample rank order tests (for location) with respect to the Student's t-test when the parent distribution is $G(x)$. If we use the normal scores estimator, we have $\text{ARE}(R_1, R_2) \ge 1$. If we use the Wilcoxon scores estimator, then for any F, we have $\text{ARE}(R_1, R_2) \ge 0.864$ and $\text{ARE}(R_1, R_2) = 3/\pi$ when F is normal. Hence the procedure given by (3.1) has "high" efficiency. In the above discussions, we consider the parameter points satisfying (3.6). When the condition (3.6) is not satisfied, but the ratio of sample sizes, m for R_1 and n for R_2, satisfies $\lim_{n\to\infty} \frac{n}{m} = (1-\rho)\sigma^2/\sigma_0^2$, then, for large n, the procedures R_1 and R_2 have approximately the same probability of a correct selection and expected size.

Theorem 5.2. Let n and $m = g(n)$ satisfy $\lim_{n\to\infty} \frac{n}{m} = (1-\rho)\sigma^2/\sigma_0^2$, then the procedures R_2 and R_1 have the same asymptotic probability of a correct selection and the same expected size for any parametric configuration.

Proof. For procedure R_2, consider any sequence of parameter points satisfying

$$\theta_{[k]}^{(n)} - \theta_{[i]}^{(n)} = \delta_{in} = n^{-\frac{1}{2}}\sqrt{1-\rho}\,\sigma\delta_i + o(n^{-\frac{1}{2}}),$$

$i = 1, \ldots, k-1$ and for some i, j, $\delta_i \neq \delta^*$, $\delta_j \neq 0$. Then

$$\begin{aligned}\lim_{n\to\infty} P_{\underline{\theta}^{(n)}}(CS \mid R_2) &= P\big(V_i \leq \lim_{n\to\infty} n^{\frac{1}{2}}\big(d_2(n) + \delta_{in}\big)/\sqrt{2(1-\rho)}\,\sigma, i = 1,\ldots,k-1\big) \\ &= P\big(V_i \leq (d+\delta_i)/\sqrt{2},\ i = 1, \ldots, k-1\big) \\ &= Q\big((d+\delta_1)/\sqrt{2}, \ldots, (d+\delta_{k-1})/\sqrt{2}\,\big)\end{aligned}$$

and for $1 \leq j \leq k-1$,

$$\begin{aligned}\lim_{n\to\infty} P_j(\underline{\theta}^{(n)} \mid R_2) &= P\big(V_i \leq (d+\delta_i-\delta_j)/\sqrt{2},\ i = 1, 2, \ldots, k, i \neq j\big) \\ &= Q\big((d+\delta_1-\delta_j)/\sqrt{2}, \ldots, (d+\delta_k-\delta_j)/\sqrt{2}\,\big).\end{aligned}$$

For procedure R_1, $m^{-\frac{1}{2}}\sigma_0 \sim n^{-\frac{1}{2}}\sqrt{1-\rho}\,\sigma$, so

$$\theta_{[k]}^{(m)} - \theta_{[i]}^{(m)} = m^{-\frac{1}{2}}\sigma_0\delta_i + o(m^{-\frac{1}{2}}) = \delta_{im}, \qquad i = 1, 2 \ldots, k-1.$$

Hence

$$\begin{aligned}\lim_{m\to\infty} P_{\underline{\theta}^{(m)}}(CS \mid R_1) &= P\big(V_i \leq (d+\delta_i)/\sqrt{2},\ i = 1, 2, \ldots, k-1\big) \\ &= Q\big((d+\delta_1)/\sqrt{2}, \ldots, (d+\delta_{k-1})/\sqrt{2}\,\big).\end{aligned}$$

and for $1 \leq j \leq k-1$,

$$\lim_{m\to\infty} P_j\big(\underline{\theta}^{(m)} \mid R_1\big) = Q\big((d+\delta_1-\delta_j)/\sqrt{2}, \ldots, (d+\delta_k-\delta_j)/\sqrt{2}\,\big).$$

In the above parameter points, we assume that $\theta_{[k]}^{(n)} - \theta_{[i]}^{(n)}$ tend to zero at the $n^{-\frac{1}{2}}$ rate. If any difference tends to zero more rapidly, we replace δ_i by 0, and if it tends to zero more slowly, or tends to a finite limit, then we replace δ_i by ∞, and still obtain the same asymptotic behavior. This completes the proof of the above theorem.

6. ESTIMATION OF B AND $\lambda_I(G)$

In practical application, for large n, the procedure R_1 can be rewritten as

$$R_1\text{: Select treatment } i \text{ iff } \hat{\Delta}_i \geq \max_{1\leq j\leq k} \hat{\Delta}_j - \frac{d\sigma_0}{\sqrt{n}}, \tag{6.1}$$

where d is the solution of (3.3) or $\int_{-\infty}^{\infty} \Phi^{k-1}(x+d)d\Phi(x) = P^*$. However, σ_0 is still unknown. We need to find a consistent estimator of σ_0^2. Since $\sigma_0^2 = \{A^2 + (k-2)\lambda_J(G)\}/kB^2$, where $A^2 = \int_0^1 J^2(u)\,du$ is known, but $B = \int_{-\infty}^{\infty} \frac{d}{dx}J(G(x))\,dG(x)$ and $\lambda_J(G) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J(G(x))J(G(y))\,dG^*(x,y)$ are unknown. Our problem is to find consistent estimator of B and $\lambda_J(G)$.

The consistent estimators of B and $\lambda_J(G)$ can be found from (4.7) and Theorem 4.2 of Puri and Sen (1967); let these be $\hat{B}_n$ and $\hat{L}_n$, respectively. Then $\hat{\sigma}_0^2 = [A^2 + (k-2)\hat{L}_n]/k\hat{B}_n^2$ is a consistent estimator of σ_0^2. Hence, for large n, the procedure R_1 is defined by

$$R_1\text{: Select treatment } i \text{ iff } \hat{\Delta}_i \geq \max_{1\leq j\leq k} \hat{\Delta}_j - \frac{d\hat{\sigma}_0}{\sqrt{n}}. \tag{6.2}$$

Remark. If $\psi(x)$ is the cdf of $U(-1,1)$, then $J(u) = 2u-1,\ 0 \leq u \leq 1$, hence $A^2 = 1/3$. If $G'(x) = g(x)$ exists, then $B = 2\int g^2(x)\,dx$ and $\lambda_J(G) = 4\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x)G(y)\,dG^*(x,y) - 1$. It has been shown by Doksum (1967) that $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x)G(y)\,dG^*(x,y)$ can be estimated by $\frac{1}{n(n-1)(n-2)k(k-1)(k-2)}$ { number of sixtuples $(i,j,\ell,\alpha,\beta,\gamma)$ with i, j, ℓ distinct; α, β, γ distinct, and $X_{i\alpha} - X_{j\alpha} < X_{i\beta} - X_{j\beta}$, $X_{i\alpha} - X_{\ell\alpha} < X_{i\gamma} - X_{\ell\gamma}$}.

ACKNOWLEDGEMENTS

This research was supported by the Office of Naval Research Contracts N00014-75-C-0455 and N00014-84-C-0167 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

BIBLIOGRAPHY

Barlow, R. E. and Gupta, S. S. (1969). "Selection procedures for restricted families of probability distributions," *Ann. Math. Statist.*, 40, 905–917.

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

Doksum, K. (1967). "Robust procedures for some linear models with one observation per cell," *Ann. Math. Statist.*, 38, 878–883.

Ghosh, M. (1973). "Nonparametric selection procedures for symmetric location parameter populations," *Ann. Statist.*, 1, 773–779.

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

Hodges, J. L. and Lehmann, E. L. (1963). "Estimates of location based on rank tests," *Ann. Math. Statist.*, 34, 598–611.

Hsu, J. C. (1982). "Simultaneous inference with respect to the best treatment in block designs," *J. Amer. Statist. Assoc.*, 77, 461–467.

Lehmann, E. L. (1964). "Asymptotically nonparametric inference in some linear models with one observation per cell," *Ann. Math. Statist.*, 35, 726–734.

Puri, M. L. (1964). "Asymptotic efficiency of a class of c-sample tests," *Ann. Math. Statist.*, 35, 102–121.

Puri, M. L. and Sen, P. K. (1967). "On some optimum nonparametric procedures in two-way layouts," *J. Amer. Statist. Assoc.*, 62, 1214–1229.

Randles, R. H. (1970). "Some robust selection procedures," *Ann. Math. Statist.*, 41, 1640–1645.

Rizvi, M. H. and Woodworth, G. G. (1970). "On selection procedures based on ranks: Counterexamples concerning least favorable configurations," *Ann. Math. Statist.*, 41, 1942–1951.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	Report #83-41
2. GOVT ACCESSION NO.	
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle)	AN ASYMPTOTIC DISTRIBUTION-FREE SELECTION PROCEDURE FOR A TWO-WAY LAYOUT PROBLEM
5. TYPE OF REPORT & PERIOD COVERED	Technical
6. PERFORMING ORG. REPORT NUMBER	Technical Report #83-41
7. AUTHOR(s)	Shanti S. Gupta and Lii-Yuh Leu
8. CONTRACT OR GRANT NUMBER(s)	N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS	Purdue University Department of Statistics West Lafayette, IN 47907
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS	Office of Naval Research Washington, DC
12. REPORT DATE	September 1983
13. NUMBER OF PAGES	13
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	
15. SECURITY CLASS. (of this report)	UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report)	Approved for public release, distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)	
18. SUPPLEMENTARY NOTES	
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)	Nonparametric procedure, subset selection, two-way layout, treatment effect, nuisance parameter, asymptotic relative efficiency.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper deals with an asymptotic distribution-free subset selection procedure for a two-way layout problem. The treatment effect with the largest unknown value is of interest to us. The block effect is a nuisance parameter in this problem. The proposed procedure is based on the Hodges-Lehmann estimators of location parameters. The asymptotic relative efficiency of the proposed procedure with the normal means procedure is evaluated. It is shown that the proposed procedure has a high efficiency.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)