Comparison of Several Populations with a Control Involving Folded Normal Distributions

Seung-Ho Lee* & Kang Sup Lee**

ABSTRACT

The problem of comparing k normal populations with a control (or a standard) in terms of the absolute values of their means is considered. Under the framework of indifference-zone formulation a single-stage and a two-stage procedures for selecting the best are proposed, according to their common variances known or unknown respectively. The procedures guarantee that the probability of correct selection is not less than some preassigned lower limit. Selected tables necessary to implement the procedures are provided.

1. Introduction

Suppose that we have $k+1$ different procedures for computing a value of some particular function. And let X_i be the error resulted by i-th method, which is assumed to follow a normal distribution with mean μ_i and common variance σ^2, $i=0, 1, \cdots, k$. We want to identify the one that consistently gives better accuracy in the results. In this case, we are interested in the bias that is observed by the absolute value of error. In other words, we are interested in comparing the quantity,

$$P(|X_i|<d)=P\left(\frac{-d-\mu_i}{\sigma}<Z<\frac{d-\mu_i}{\sigma}\right)$$
$$=\Phi\left(\frac{d}{\sigma}-\frac{\mu_i}{\sigma}\right)-\Phi\left(-\frac{d}{\sigma}-\frac{\mu_i}{\sigma}\right)$$

* Department of Computer Science and Applied Mathematics, Ajou University.

** Department of Mathematics, Dan-Kuk University.

Research in this paper was supported in part by the Ministry of Educations, Korean Government, through the Research Institute of Basic Science, Seoul National University.

The authors are grateful to Dr. Woo-Chul Kim and Dr. Moon-Sup Song for their suggestion of the problem and many valuable discussions throughout the preparation of this work.

$$=\Phi\left(\frac{d}{\sigma}-\frac{|\mu_i|}{\sigma}\right)-\Phi\left(-\frac{d}{\sigma}-\frac{|\mu_i|}{\sigma}\right)$$

where $\Phi(\cdot)$ is the *c.d.f.* of standard normal random variable, or equivalently in comparing $|\mu_i|$. Hence our goal is to select the population with the smallest value of $|\mu_i|$.

A procedure for selecting the t best normal populations in terms of $|\mu_i|$ was proposed by Rizvi (1971). Motivations for considering such a quantity can be found also in Rizvi (1971) and Gibbons, Olkin and Sobel (1977). In this paper we are concerned with the problem of selecting one from $k+1$ independent normal populations Π_0, Π_1, $\cdots$, Π_k as the best, where the population Π_0 plays a special role as a control (or a standard) under the framework of the indifference-zone formulation. For a detailed exposition of this formulation, see Bechhofer(1954).

The criterion of selecting a population is minimizing the bias or absolute mean. Procedures for selecting the best population are proposed, which guarantee that (i) with probability at least a given P_0^* none of the new populations is selected (i.e., to select the standard Π_0) when the smallest absolute population mean is sufficiently larger than that of the standard, and (ii) with probability at least a given P_1^* the population having the smallest absolute mean is to be selected when this mean is sufficiently smaller than its closest competitor's and the standard's in absolute value.

In Section 2, we propose a single-stage procedure for the case of known common variance, and in Section 3, a two-stage procedure for the case of unknown common variance. Selected tables necessary to implement the procedures are provided in each case.

Bechhofer and Turnbull (1978) considered the similar problem of selecting the best one from $k+1$ normal populations in terms of μ_i, where a standard is comletely specified in terms of μ_0.

2. The Procedure when σ^2 is known

Suppose we have k independent normal populations with known common variance σ^2 :

$$\Pi_i : N(\mu_i,\ \sigma^2),\ i=1,\ \cdots,\ k$$

and a specified standard population $\Pi_0 : N(\mu_0, \sigma^2)$ up to which Π_i's are to be compared. Let the ordered values of parameters, $\theta_i=|\mu_i|$, $i=1,\cdots, k$, be denoted by

$$0\leq\theta_{[1]}\leq\theta_{[2]}\leq\cdots\leq\theta_{[k]}$$

It is assumed that there is no a priori information available about how many or which (if any) populations have θ_i smaller than $\theta_0=|\mu_0|$

Our goal is to select the population associated with $\theta_{[1]}$ provided that $\theta_{[1]}$ is sufficiently smaller than $\theta_{[2]}$ and θ_0 or to select none if $\theta_{[1]}\geq\theta_0$. Such a selection is regarded as a correct selection (CS). And let $\underline{\theta}=(\theta_{[1]}, \theta_{[2]}, \cdots, \theta_{[k]})$ denote a point in the parameter space Ω, which is partitioned tnto a "preference-zone" Ω_0 and $\Omega_1(\delta^*)$ defined by

$$\Omega_0=\{\underline{\theta} : \theta_{[1]}\geq\theta_0\} \tag{2. 1. a}$$

and

$$\Omega_1(\delta^*)=\{\underline{\theta} : \theta_{[1]}\leq\theta_0-\delta^*\sigma,\ \theta_{[1]}\leq\theta_{[2]}-\delta^*\sigma\} \tag{2. 1. b}$$

respectively, and its complement called "indifference-zone". The quantity $\delta^*(>0)$ is a prespecified value by experimenter, and is called the "separation threshhold". In addition to specifying δ^*, the experimenter also specifies two probability constants P_0^* and P_1^*, and then he is seeking for a procedure R_1 which satisfies

$$P_{\underline{\theta}}(\text{select } \Pi_0)\geq P_0^* \qquad \text{if } \underline{\theta}\varepsilon\Omega_0 \tag{2. 2. a}$$

and

$$P_{\underline{\theta}}(\text{select } \Pi_{(1)})\geq P_1^* \qquad \text{if } \underline{\theta}\varepsilon\ \Omega_1(\delta^*) \tag{2. 2. b}$$

where $\Pi_{(1)}$ is the population with parameter $\theta_{[1]}$.

The constants $1-P_0^*$ and $1-P_1^*$ might be considered as roughly analogous to the probabilities of type I and type II errors in the Neyman-Pearson framework of testing hypotheses. However, since our problem is of a multiple decision type, it is not equivalent to testing hypotheses proplem unless $k=1$. Note that P_0^* and P_1^* are specified as being strictly greater than $(k+1)^{-1}$, since this latter probability can be achieved by selecting any population without taking any observations.

We propose a procedure R_1 as below:

Take n independent observations X_{ij} $(j=1, \cdots, n)$ from each Π_i $(i=1, \cdots, k)$. Compute $|\overline{X}_i|=|\sum_{j=1}^{n} X_{ij}/n|$ and let $0\leq|\overline{X}|_{[1]}\leq|\overline{X}|_{[2]}\leq\cdots\leq|\overline{X}|_{[k]}$ denote the order statistics of the $|\overline{X}_i|$, $i=1, \cdots, k$.

Then our selection rule R_1 decides to

$$\text{select } \Pi_0, \text{ if } |\overline{X}|_{[1]}\geq|\overline{X}_0|-d\frac{\sigma}{\sqrt{n}} \tag{2. 3. a}$$

and

$$\text{select } \Pi_{[1]}, \text{ if } |\overline{X}|_{[1]}<|\overline{X}_0|-d\frac{\sigma}{\sqrt{n}}, \tag{2. 3. b}$$

where $\Pi_{[1]}$ is the population associated with $|\overline{X}|_{[1]}$, i.e., the population with the smallest sample absolute mean, and d is a nonnegative constant.

The selection procedure R_1 is completely defined once the values of the design constants (n, d) are assigned; n is the smallest sample size which will guarantee (2.2.a) and (2.2.b) with an appropriate d.

The procedure R_1 is based on the absolute value of the sample mean. Let $W_i = \frac{\sqrt{n}}{\sigma} \times |\overline{X}_i|$ denote a typical statistic equivalent to the absolute value of the sample mean, and let $\xi_i = \frac{\sqrt{n}}{\sigma}\theta_i = \frac{\sqrt{n}}{\sigma}|\mu_i|$. Then W_i has the "folded normal" distribution with folding at the origin, and its cumulative distribution function is

$$\begin{aligned} F(w, \xi_i) &= P\left(\frac{\sqrt{n}}{\sigma}|\overline{X}_i| \leq w\right) \qquad (2.4.a) \\ &= \Phi\left(w - \frac{\sqrt{n}}{\sigma}\mu_i\right) - \Phi\left(-w - \frac{\sqrt{n}}{\sigma}\mu_i\right) \\ &= \Phi\left(w - \frac{\sqrt{n}}{\sigma}|\mu_i|\right) - \Phi\left(-w - \frac{\sqrt{n}}{\sigma}|\mu_i|\right) \\ &= \Phi(w - \xi_i) - \Phi(-w - \xi_i), \quad w > 0 \end{aligned}$$

and its probability density function is

$$f(w, \xi_i) = \phi(w - \xi_i) + \phi(w + \xi_i), \quad w > 0 \qquad (2.4.b)$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the c.d.f. and p.d.f. of the standard normal random variable, respectively.

Lemma 1 :

$$\frac{-\overset{0}{F}(w, \xi)}{f(w, \xi)} \qquad (2.5)$$

is increasing in $w\xi$, where $\overset{0}{F}(w, \xi) = \frac{\partial}{\partial \xi} F(w, \xi)$.

Proof : $\quad -\overset{\bullet}{F}(w, \xi)/f(w, \xi) = \{\phi(w-\xi) - \phi(w+\xi)\} / \{\phi(w-\xi) + \phi(w+\xi)\}$

$= \tanh(w\xi)$, increasing in $w\xi$.

For $\underline{\theta} \varepsilon \Omega_0 = \{\underline{\theta} : \theta_{[1]} \geq \theta_0\} = \{\underline{\xi} : \xi_{[1]} \geq \xi_0\}$, the probability of correct selection by the procedure R_1 is given by

$$\begin{aligned} P_{\underline{\theta}}(\text{CS}) &= P_{\underline{\theta}}\left(|\overline{X}|_{[1]} \geq |\overline{X}_0| - d\frac{\sigma}{\sqrt{n}}\right) \qquad (2.6.a) \\ &= P_{\underline{\theta}}(W_0 \leq d) + P_{\underline{\theta}}(W_{[1]} \geq W_0 - d,\ W_0 > d) \\ &= F(d, \xi_0) + \int_d^\infty \prod_{i=1}^k [1 - F(w-d, \xi_i)] f(w, \xi_0) dw, \end{aligned}$$

and for $\underline{\theta} \varepsilon \Omega_1 = \{\theta : \theta_{[1]} \leq \theta_0 - \delta^*\sigma,\ \theta_{[1]} \leq \theta_{[2]} - \delta^*\sigma\} = \{\xi : \xi_{[1]} \leq \xi_0 - \delta,\ \xi_{[1]} \leq \xi_{[2]} - \delta,\ \delta = \sqrt{n}\ \delta^*\}$,

$$P_{\underline{\theta}}(\mathrm{CS})=P_{\underline{\theta}}\left(|\bar{X}|_{(1)}<|\bar{X}_0|-d\frac{\sigma}{\sqrt{n}},\quad |\bar{X}|_{(1)}=|\bar{X}|_{[1]}\right) \qquad (2.6.\mathrm{b})$$
$$=\mathrm{p}_{\underline{\theta}}(W_{[1]}<W_0-d,\quad W_{(1)}=W_{[1]})$$
$$=\int_0^\infty [1-F(w+d,\ \xi_0)]\prod_{j=2}^{k}[1-F(w,\xi_j)]\ dF(w,\xi_1).$$

where $|\bar{X}|_{(1)}$, $W_{(1)}$ denote the sample mean from the population associated with $\theta_{[1]}$.

To satisfy the requirements (2.2.a) and (2.2.b), we need to find the infima of these probabilities and also $\underline{\theta}$ for which these infima attained; these $\underline{\theta}$ will be called the "least favorable configurations" (LFC) of the parameters.

Lemma 2 :
$$I(\xi)=F(d,\ \xi)+\int_d^\infty [1-F(w-d,\xi)]^k\ f(w,\xi)dw \qquad (2.7.\mathrm{a})$$

is a nonincreasing function of ξ, and

$$J(\xi)=\int_0^\infty [1-F(w+d,\xi+\delta)][1-F(w,\xi+\delta)]^{k-1}dF(w,\xi) \qquad (2.7.\mathrm{b})$$

is a nondecreasing function of ξ.

Proof : Using integration by parts and Lemma 1,

$$\frac{d}{d\xi}I(\xi)=\overset{0}{F}(d,\ \xi)+\int_d^\infty k[1-F(w-d,\xi)]^{k-1}[-\overset{0}{F}(w-d,\xi)]\ f(w,\xi)dw$$
$$+\int_d^\infty [1-F(w-d,\xi)]^k\ \overset{0}{f}(w,\xi)dw \qquad (2.8.\mathrm{a})$$
$$=k\int_d^\infty [1-F(w-d,\xi)]^{k-1}f(w-d,\xi)\ f(w,\xi)$$
$$\left\{\frac{-\overset{0}{F}(w-d,\xi)}{f(w-d,\xi)}-\frac{-\overset{0}{F}(w,\xi)}{f(w,\xi)}\right\}dw\leq 0$$

and

$$\frac{d}{d\xi}J(\xi)=\int_0^\infty (k-1)[1-F(w,\xi+\delta)]^{k-2}[1-F(w+d,\xi+\delta)]$$
$$[-\overset{0}{F}(w,\xi+\delta)]\ f(w,\xi)dw \qquad (2.8.\mathrm{b})$$
$$+\int_0^\infty [1-F(w,\xi+\delta)]^{k-1}[-\overset{0}{F}(w+d,\xi+\delta)]f(w,\xi)dw$$
$$+\int_0^\infty [1-F(w+d,\xi+\delta)][1-F(w,\xi+\delta)]^{k-1}\ \overset{0}{f}(w,\xi)dw$$
$$=\int_0^\infty [1-F(w,\xi+\delta)]^{k-1}\ f(w+d,\ \xi+\delta)\ f(w,\xi)$$
$$\left\{\frac{-\overset{0}{F}(w+d,\xi+\delta)}{f(w+d,\xi+\delta)}-\frac{-\overset{0}{F}(w,\xi)}{f(w,\xi)}\right\}dw$$
$$+\int_0^\infty (k-1)[1-F(w,\xi+\delta)]^{k-1}[1-F(w+d,\xi+\delta)]$$

$$f(w,\xi+\delta)\ f(w,\xi)\left\{\frac{-\overset{0}{F}(w,\xi+\delta)}{f(w,\xi+\delta)}+\frac{-\overset{0}{F}(w,\xi)}{f(w,\xi)}\right\}dw\geq 0.$$

Theorem 1 : Given δ^*, P_0^* and P_1^*, the design constants (n, d) of the procedure R_1 which guarantee(2. 2. a) and (2. 2. b) are determined by the following simultaneous equations:

$$\int_{-\infty}^{\infty}[1-\Phi(t-d)]^k\phi(t)dt=P_0^* \qquad (2.9.a)$$

$$2\int_0^{\infty}[1-\Phi(t+d-\delta)+\Phi(-t-d-\delta)][1-\Phi(t-\delta)+\Phi(-t-\delta)]^{k-1}\times$$

$$\phi(t)dt=P_1^* \qquad (2.9.b)$$

$$n=[(\delta/\delta^*)^2]+1 \qquad (2.9.c)$$

where $[x]$ denotes the largest integer less than or equal to x.

Proof : For $\underline{\theta}\varepsilon\Omega_0$, LFC is $\xi_0=\xi_{[1]}=\cdots=\xi_{[k]}$; hence

$$P_{\underline{\theta}}(\text{CS})=F(d,\xi_0)+\int_d^{\infty}\prod_{i=1}^{k}[1-F(w-d,\xi_i)]\,f(w,\xi_0)dw$$

$$\geq F(d,\xi_0)+\int_d^{\infty}[1-F(w-d,\xi_0)]^k\ f(w,\xi_0)dw. \qquad (2.10)$$

Therfore it follows from Lemma 2 that

$$\inf_{\theta\varepsilon\Omega_0} P_{\underline{\theta}}(\text{CS})=\inf_{\xi_0\geq 0}\{F(d,\xi_0)+\int_d^{\infty}[1-F(w-d,\xi_0)]^k f(w,\xi_0)dw\}$$

$$=\lim_{\xi_0\to\infty} I(\xi_0) \qquad \text{for } I(\xi) \text{ given in (2.7.a).}$$

Thus,

$$\inf_{\theta\varepsilon\Omega_0} P_{\underline{\theta}}(\text{CS})=\lim_{\xi_0\to\infty}\{\int_d^{\infty}[1-\Phi(w-d-\xi_0)+\Phi(-w+d-\xi_0)]^k\phi(w-\xi_0)dw$$

$$+\int_d^{\infty}[1-\Phi(w-d-\xi_0)+\Phi(-w+d-\xi_0)]^k\phi(w+\xi_0)dw\}$$

$$=\lim_{\xi_0\to\infty}\int_{d-\xi_0}^{\infty}[1-\Phi(t-d)+\Phi(-t+d-2\xi_0)]^k\phi(t)dt$$

$$+\lim_{\xi_0\to\infty}\int_{d+\xi_0}^{\infty}[1-\Phi(t-d-2\xi_0)+\Phi(-t+d)]^k\phi(t)dt$$

$$=\int_{-\infty}^{\infty}[1-\Phi(t-d)]^k\phi(t)dt. \qquad (2.11)$$

similarly, for $\underline{\theta}\varepsilon\Omega$, LFC is $\xi_0=\xi_{[1]}+\delta=\xi_{[2]}=\cdots=\xi_{[k]}$.

$$P_{\underline{\theta}}(\text{CS})=\int_0^{\infty}[1-F(w+d(\xi_0))]\prod_{j=2}^{k}[1-F(w,\xi_j)]dF(w,\xi_1)$$

$$>\int_0^{\infty}[1-F(w+d,\xi_1+\delta)][1-F(w,\xi_1+\delta)]^{k-1}dF(w,\xi_1)$$

(2. 12)

It follows from Lemma 2 that

$$\inf_{\theta\in\Omega_1} P_\theta(\text{CS})=\inf_{\xi_1\geq 0}\int_0^\infty[1-F(w+d,\xi_1+\delta)][1-F(w,\xi_1+\delta)]^{k-1}dF(w,\xi_1)$$
$$=\lim_{\xi_1\to 0} J(\xi_1)$$

Thus,

$$\inf_{\theta\in\Omega_1} P_\theta(\text{CS})=\int_0^\infty[1-F(w+d,\delta)][1-F(w,\delta)]^{k-1}dF(w,0) \qquad (2.13)$$
$$=\int_0^\infty[1-\Phi(w+d-\delta)+\Phi(-w-d-\delta)][1-\Phi(w-\delta)$$
$$+\Phi(-w-\delta)]^{k-1}[\phi(w)+\phi(-w)]dw$$
$$=2\int_0^\infty[1-\Phi(t+d-\delta)+\Phi(-t-d-\delta)][1-\Phi(t-\delta)$$
$$+\Phi(-t-\delta)]^{k-1}\phi(t)dt$$

From these two infima probabilities and the relationship $\delta=\sqrt{n}\delta^*$, the results follow.

Values of (δ, d) for $k=2(1)5$ and selected (P_0^*, P_1^*) are given in Table B, while Table A gives values of (P_0^*, P_1^*) for $k=2, 3, 5, 10$ and selected (δ, d). Computer programs to solve (2.9.a) and (2.9.b) have been prepared using 32 point Gauss-Laguerre quadrature formular, which is available upon request.

It should be noted that, the left side of (2.9.a) is increasing in d, and the left side of (2.9.b) is decreasing in d and increasing in δ. It should be also noted that, by setting $d=0$ and $n=0$ (and hence $\delta=0$) in the left side of (2.9.a) and (2.9.b), we obtain $(k+1)^{-1}$, the probability that can be achieved by randomly selecting a population without taking any obrervations.

As an example for using Table B, suppose that we are interested in selecting one from the three alternative measuring instruments if it is superior in accuracy to the others and the standard instrument. Suppose that the measurements are known to follow normal distributions with different means but with a common variance $\sigma^2=100$. Also, let $P_0^*=0.90$, $P_1^*=0.90$ and $\delta^*=1$ be specified. Then from Table B for $k=4$, $\delta=4.91230$ and $d=2.59970$ and hence $n=[(\delta/\delta^*)^2]+1=[24.1307]+1=25$.

Thus we need 25 observations from each population and select the one that gives the smallest absolute sample mean if the value is less than that of the standard minus $2.59970\times\frac{10}{\sqrt{25}}=5.1994$, and none otherwise.

Table A-1. Values of P_0^* given by (2.9.a)

d \ k	2	3	5	10
0.0	0.333333	0.250000	0.166667	0.090909
0.5	0.482593	0.393318	0.292527	0.185216
1.0	0.633702	0.552031	0.449365	0.323202
1.5	0.765812	0.701863	0.613555	0.490115
2.0	0.865767	0.822793	0.758452	0.657630
3.0	0.968795	0.956374	0.935305	0.895580
4.0	0.995496	0.993470	0.989742	0.981740
5.0	0.999599	0.999407	0.999037	0.998176
6.0	0.999978	0.999967	0.999946	0.999894
7.0	0.999999	0.999999	0.999998	0.999996

Table A-2. Values of P_1^* given by (2.9.b) for $k=2$

d \ δ	0.0	0.5	1.0	2.0	3.0	5.0	7.0
0.0	0.333333	0.367067	0.466021	0.755655	0.938738	0.999198	0.999999
0.5	0.182946	0.218094	0.323456	0.657017	0.900407	0.998156	0.999995
1.0	0.083697	0.111639	0.199583	0.525458	0.829833	0.994979	0.999977
2.0	0.009620	0.017484	0.047539	0.233124	0.569210	0.966147	0.999592
3.0	0.000469	0.001256	0.005310	0.054622	0.242597	0.848768	0.995327
5.0	0.000000	0.000000	0.000005	0.000279	0.006160	0.249994	0.848886
7.0	0.000000	0.000000	0.000000	0.000000	0.000005	0.006186	0.250000

for $k=3$

d \ δ	0.0	0.5	1.0	2.0	3.0	5.0	7.0
0.0	0.249999	0.278139	0.367404	0.680899	0.914466	0.998815	0.999998
0.5	0.140783	0.168949	0.259613	0.597644	0.879187	0.997791	0.999994
1.0	0.065954	0.088355	0.163367	0.433113	0.812691	0.994637	0.999976
2.0	0.007892	0.014385	0.040403	0.218780	0.561045	0.965905	0.999592
3.0	0.000397	0.001067	0.004652	0.052085	0.245299	0.848650	0.995327
5.0	0.000000	0.000000	0.000004	0.000271	0.006134	0.249988	0.848886
7.0	0.000000	0.000000	0.000000	0.000000	0.000005	0.006186	0.250000

for $k=5$

d \ δ	0.0	0.5	1.0	2.0	3.0	5.0	7.0
0.0	0.166666	0.187079	0.257339	0.573140	0.873460	0.998075	0.999996
0.5	0.096488	0.116422	0.185818	0.509515	0.842730	0.997077	0.999993
1.0	0.046420	0.062385	0.119636	0.418243	0.782665	0.993974	0.999975
2.0	0.005825	0.010639	0.031037	0.195429	0.546167	0.965430	0.999591
3.0	0.000304	0.000821	0.003731	0.047745	0.240973	0.848417	0.995326
5.0	0.000000	0.000000	0.000004	0.000256	0.006085	0.249975	0.848886
7.0	0.000000	0.000000	0.000000	0.000000	0.000005	0.006186	0.250000

for $k=10$

d \ δ	0.0	0.5	1.0	2.0	3.0	5.0	7.0
0.0	0.090908	0.102674	0.146127	0.417010	0.796810	0.996353	0.999993
0.5	0.054087	0.065462	0.107686	0.377065	0.772991	0,995410	0.999989
1.0	0.026746	0.035981	0.071304	0.316502	0.723571	0.992414	0.999972
2.0	0.003536	0.006460	0.019661	0.155435	0.514990	0.964294	0.999587
3.0	0.000194	0.000524	0.002497	0.039715	0.231385	0.847848	0.995324
5.0	0.000000	0.000000	0.000003	0.000227	0.005967	0.249945	0.848886
7.0	0.000000	0.000000	0.000000	0.000000	0.000005	0.006185	0.250000

Table B. Values for the design constants(δ, d) in the procedure R_1 for selected(P_0^*, P_1^*)

	P_1^* \ P_0^*		0.5	0.75	0.80	0.90	0.95	0.99
	0.90	δ	3.02037	3.80571	4.00579	4.54273	4.99697	5.87283
		d	2.23020	2.23020	2.23020	2.23020	2.23020	2.23020
$k=2$	0.95	δ	3.48724	4.27926	4.48043	5.01945	5.47480	6.36186
		d	2.71010	2.71010	2.71010	2.71010	2.71010	2.71010
	0.99	δ	4.38850	5.18411	5.38577	5.92565	6.38141	7.25885
		d	3.61730	3.61730	3.61730	3.61730	3.61730	3.61730
	0.90	δ	3.24529	4.02809	4.22788	4.75433	5.21839	6.09414
		d	2.45157	2.45157	2.45157	2.45157	2.45157	2.45157
$k=3$	0.95	δ	3.69444	4.48564	4.68673	5.22562	6.68093	6.55797
		d	2.91623	2.91623	2.91623	2.91623	2.91623	2.91623
	0.99	δ	4.56824	5.36379	5.56544	6.15532	6.56107	7.43851
		d	3.79696	3.79696	3.79696	3.79696	3.79696	3.79696
	0.90	δ	3.39433	4.17622	4.37593	4.91230	5.36636	6.24217
		d	2.59970	2.59970	2.59970	2.59970	2.59970	2.59970
$k=4$	0.95	δ	3.83359	4.62456	4.82563	5.36452	5.81984	6.69689
		d	3.05517	3.05517	3.05517	3.05517	3.05517	3.05517
	0.99	δ	4.69087	5.48641	5.68806	6.22794	6.68370	7.56113
		d	3.91958	3.91958	3.91958	3.91958	3.91958	3.91958
	0.90	δ	3.50474	4.28625	4.48595	5.02235	5.47644	6.35233
		d	2.70995	2.70995	2.70995	2.70995	2.70995	2.70995
$k=5$	0.95	δ	3.93750	4.72840	4.92948	5.46838	5.92372	6.80080
		d	3.15909	3.15909	3.15909	3.15909	3.15909	3.15909
	0.99	δ	4.78336	5.57891	5.78056	6.32044	6.77620	7.65363
		d	4.01208	4.01208	4.01208	4.01208	4.01208	4.01208

3. The procedure σ^2 is unknown

When σ is not known, the preference-zone $\Omega_1(\delta^*)$ given in (2.1.b) is not completely determined, and hence it must be modified. Thus we define a "preference-zone" Ω_0^* and $\Omega_1^*(\delta^*)$ as follows:

$$\Omega_0^* = \{(\underline{\theta}, \sigma) : \theta_{[1]} > \theta_0\} \tag{3.1.a}$$

$$\Omega_1^*(\delta_*) = \{(\underline{\theta}, \sigma) ; \theta_{[1]} \leq \theta_0 - \delta^*, \theta_{[1]} \leq \theta_{[2]} - \delta^*\} \tag{3.1.b}$$

where $\underline{\theta}, \theta_{[i]}$ and δ^* are the same as those in Section 2.

Our goal is to find a selection procedure which guarantees;

$$P_{(\underline{\theta},\sigma)}\{\text{select } \Pi_0\} \geq P_0^*, \qquad \text{if } (\underline{\theta}, \sigma) \varepsilon \Omega_0^* \tag{3.2.a}$$

$$P_{(\underline{\theta},\sigma)}\{\text{select } \Pi_{(1)}\} \geq P_1^*, \qquad \text{if } (\underline{\theta}, \sigma) \varepsilon \Omega_1^* \tag{3.2.b}$$

for the specified probability bounds P_0^* and P_1^*. We propose a two-stage procedure R_2 as below;

(a) In the first stage, take $n_0(\geq 2)$ independent observations X_{ij} $(j=1, \cdots, n_0.)$ from each Π_i $(i=0, 1, \cdots, k)$ and compute

$$s^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_0} (X_{ij} - \overline{X}_i)^2/\nu, \quad \nu = k(n_0 - 1) \tag{3.3}$$

which is an unbiased estimate of σ^2 based on $\nu = k(n_0-1)$ degrees of freedom.

(b) In the second stage, take $n - n_0$ indepencent observations from each Π_i, $i=0, 1, 2, \cdots, k$

$$\text{where } n = \max\ \{n_0, [(cs/\delta^*)^2] + 1\} \tag{3.4}$$

and compute the overall (first stage plus second stage) absolute values of the sample means

$$|\overline{X}_i| = |\sum_{j=1}^{n} X_{ij}/n|, \qquad i=0, 1, \cdots, k$$

and let $0 \leq |\overline{X}|_{[1]} \leq |\overline{X}|_{[2]} \leq \cdots \leq |\overline{X}|_{[k]}$ denote the order statistics of $\overline{X}_i, i=1, \cdots, k$.

(c) Stlect Π_0, if $|\overline{X}|_{[1]} \geq |\overline{X}_0| - d\dfrac{s}{\sqrt{n}}$ (3.5.a)

and

Select $\Pi_{[1]}$, if $|\overline{X}|_{[1]} < |\overline{X}_0| - d\dfrac{s}{\sqrt{n}}$ (3.5.b)

For the selection procedure R_2 to be defined, the design constants n_0, n and d should be assigned. The choice of the initial sample size n_0 is optional, though some guidelines to assist in this choice are discussed at the end of this section. For a chosen n_0, n is completely determined in terms of c, and hence the design constants(n, d), or equivalently (c, d) depend on k, n_0 and the specified quantities (δ^*, P_0^*, P_1^*). Theorem 2 tells how to determine (c, d) so as to guarantee(3. 2. a) and (3. 2. b).

Let $W_i=\frac{\sqrt{n}}{\sigma}|\bar{X}_i|$ and $\xi_i=\frac{\sqrt{n}}{\sigma}\theta_i$ as in Section 2, then the selection rule R_2 can be represented in terms of W_i as follows;

$$\text{Select } \Pi_0,\ \text{if } W_{[1]}\geq W_0-du \tag{3. 6. a}$$

and

$$\text{Select } \Pi_{[1]},\text{if } W_{[1]}<W_0-du,\ \text{where } u=s/\sigma \tag{3. 6. b}$$

and the preference-zone $\Omega_1^*(\delta^*)$, in terms of ξ_i ;

$$\Omega_1^*(\delta)=\{\underline{\xi} : \xi_{[1]}\leq\xi_0-\delta,\ \ \xi_{[1]}\leq\xi_{[2]}-\delta\} \tag{3. 7}$$

where $\delta=\sqrt{n}\delta^*/\sigma$.

Note that u and W_i are independent.

Theorem 2 : Given δ, P_0^* and P_1^*, the design constants (c, d) of the procedure R_2 which guarantee (3. 2. a) (3. 2. b) are the pair satisfying the simultaneous equations;

$$\int_0^\infty\int_{-\infty}^\infty[1-\Phi(t-ud)]^k\phi(t)dt\ q_\nu(\mathrm{u})=ud\ P_0^* \tag{3. 8. a}$$

$$2\int_0^\infty\int_0^\infty[1-\Phi(t+ud-uc)+\Phi(-t-ud-uc)][1-\Phi(t-uc)$$
$$+\Phi(-t-uc)]^{k-1}\phi(t)dt\ q_\nu(u)ud=P_1^* \tag{3. 8. b}$$

where $q_\nu(u)$ is the density of $\sqrt{\chi_\nu^2/\nu}$, $\nu=k(n_0-1)$.

Proof: When$(\underline{\theta}, \sigma)\ \varepsilon\,\Omega_0^*$. LFC is $\xi_0=\xi_1=\cdots=\xi_k$ and hence by Lemma 2,

$$P_{(\underline{\theta},\sigma)}(\mathrm{CS})=P_{(\underline{\theta},\delta)}(W_{[1]}\geq W_0-ud) \tag{3. 9}$$
$$=\int_0^\infty\{F(ud,\ \xi_0)+\int_{du}^\infty\prod_{j=1}^k[1-F(w-ud, \xi_j)]dF(w, \xi_0)\}\,q_\nu(u)du$$
$$\geq\lim_{\xi_0\to\infty}\int_0^\infty\{F(ud, \xi_0)+\int_{du}^\infty[1-F(w-ud, \xi_0)]^k f(w, \xi_0)dw\}$$
$$q_\nu(u)du$$
$$=\int_0^\infty\int_{-\infty}^\infty[1-\Phi(t-ud)]^k\phi(t)dtq_\nu(u)du.$$

Similarly, when$(\underline{\theta}, \sigma)\varepsilon\,\Omega_1^*$, LFC is $\xi_0=\xi_{[1]}+\delta=\xi_{[1]}=\cdots=\xi_{[k]}$ and by Lemma 2,

$$P_{(\underline{\theta},\delta)}=P_{(\underline{\theta},\delta)}(W_{[1]}<W_0-ud,\ W_{[1]}=W_{(1)}) \tag{3. 10}$$
$$=\int_0^\infty\int_0^\infty[1-F(w+ud,\ \xi_0)]\prod_{j=2}^k[1-F(w, \xi_j)]dF(w, \xi_1)\cdot q_\nu(u)du$$

$$\geq \lim_{\xi_1\to 0}\int_0^\infty\int_0^\infty [1-F(w+ud,\xi_1+\delta)]^{k-1}dF(w,\xi_1)q_\nu(u)du$$
$$=2\int_0^\infty\int_0^\infty [1-\Phi(t+ud-\delta)+\Phi(-t-ud-\delta)][1-\Phi(t-\delta)$$
$$+\Phi(-t-\delta)]^{k-1}\phi(t)dt\cdot q_\nu(u)du.$$

Since $n>(cs/\delta^*)^2, \delta=\sqrt{n}\delta^*/\sigma\geq uc$
and the above integral (3. 10) is increasing in δ. Thus the infimum of $P_{(\underline{\theta},\sigma)}(\mathrm{CS})$ occurs when $\delta=cu$, i.e.,

$$\inf_{\Omega_1^*} P_{(\underline{\theta},\sigma)}(\mathrm{CS})=2\int_0^\infty\int_0^\infty [1-\Phi(t+ud-uc)+\Phi(-t-ud-uc)] \quad (3.11)$$
$$[1-\Phi(t-uc)+\Phi(-t-uc)]^{k-1}\phi(t)dt \;\; q_\nu(u)du.$$

From the two infima probabilities (3. 9) and (3. 11), the results follow.

It should be noted that, if $\nu\to\infty$, then $u\to 1$ by the "Strong law of large numbers" and hence the double integrals (3. 8. a) and (3. 8. b) are reduced to (2. 9. a) and (2. 9. b) in theorem 1, respectively. This means that when the degrees of freedom for estimating σ^2 is sufficiently large, the value of the design constants (c,d) in the unknown variance case are nearly the same as the (δ,d) of the known variance case.

Table C gives the solution (c,d) of (3. 8. a) and (3. 8. b) for $k=2(1)5$ and selected ν, P_0^* and P_1^*. The tabulated values of c (upper entry) and d (lower entry) are calculated to an accuracy of $\pm 10^{-5}$ in the associated values of $(P_0^*,\ P_1^*)$ using 24 point Gaussian-Lagurre and Gaussian-Hermite quadratures, both in the IBM scientific subroutine package.

As an example for using the Table C, suppose that we have $k=4$ populations and specify $\delta^*=1$, $P_0^*=0.90$, $P_1^*=0.75$. And suppose that we take $n_0=11$ observations from each populations and obtain $s^2=16$ with 40 degrees of freedom. Then from Table C we have $(c,d)=(4.62953, 2.66130)$ and $(cs/\delta^*)^2=342.9208$. Hence we need 332 observations more from each populations and select the one that gives the smallest absolute sample mean if the value is less than $|\overline{X}_0|-2.66130\dfrac{4}{\sqrt{343}}$.

Although the initial sample size n_0 of the procedure R_2 is optional, some guidelines to assist in this choice can be suggested.

Follwing Stein (1945),

$$N=\max\{n_0, [(cs/\delta^*)^2]+1\}$$

is a random variable with expectation

$$E[N]=n_0\,P(\chi^2_\nu\leq n_0\nu\lambda)+\frac{1}{\lambda}P(\chi^2_{\nu+2}>n_0\nu\lambda)+\alpha P(\chi^2_\nu>n_0\nu\lambda)$$

where $\lambda=(\delta^*/c)^2/\sigma^2$ and $0<\alpha<1$.

If the experimenter has some idea as to the possible values of σ, Bechhoffer and Turnbull (1978) proposed that n_0 would be chosen to minimize the maximum expected loss in number of extra observations needed due to ignorance of σ. As an alternative to the minimax regret criterion, n_0 could be chosen so as to minimize the expected loss in number with respect to an appropriate prior distribution of σ. The results of Moshman (1958) and Wormleighton (1960) can also be extended to our problem.

Table C. Values for the design constants (c, d) in the procedure R_2 for selected (P_0^*, P_1^*)

	ν		$P_0^*=0.90$	0.90	0.95	0.95	0.99	0.99
			$P_1^*=0.75$	0.90	0.75	0.90	0.75	0.90
$k=2$	10	c	4.50068	6.66578	5.10692	7.26895	6.46304	8.62159
		d	2.42301	2.42301	3.04143	3.04143	4.40493	4.40493
	20	c	4.32475	6.15470	4.86077	6.69329	5.96760	7.80130
		d	2.32281	2.32281	2.86706	2.86786	3.97794	3.97794
	40	c	4.24179	5.92949	4.74633	6.43765	5.74715	7.44012
		d	2.27559	2.27559	2.78652	2.78652	3.79004	3.79004
$k=3$	15	c	4.63346	6.56740	5.18503	7.12015	6.36735	8.30256
		d	2.60281	2.60281	3.16632	3.16632	4.35514	4.35514
	30	c	4.50509	6.23747	5.00839	6.74485	6.02789	7.76623
		d	2.52510	2.52510	3.03673	3.03673	4.05996	4.05996
$k=4$	20	c	4.73175	6.56089	5.25429	7.08657	6.34869	8.18239
		d	2.72557	2.72557	3.25895	3.25895	4.35881	4.35881
	40	c	4.62953	6.31587	5.11470	6.80558	6.08671	7.77963
		d	2.66130	2.66130	3.15423	3.15423	4.12954	4.12954
$k=5$	25	c	4.80966	6.57942	5.31374	7.08753	6.35376	8.12946
		d	2.81848	2.81848	3.33239	3.33239	4.37710	4.37710
	50	c	4.72418	6.38381	5.19750	6.86177	6.13886	7.80513
		d	2.76328	2.76328	3.24379	3.24379	4.18809	4.18809

[Received August 1982; Revised October 1982]

REFERENCES

(1) Bechhofer, R.E. (1954). "A Single-Sample Multiple Decision Procedure for Ranking Means of Normal Populations with Known Variances." Ann. Math. Statist. Vol. 25, 16—39.

(2) Bechhofer, R.E. and Turnbull, B.W. (1978). "Two $(k+1)$—Decision Selection Procedures for

Comparing k Normal Means with a Specified Standard." J. Amer. Statist. Assoc., Vol. 73, 385—392.

(3) Gibbons, J.D., Olkin, I. and Sobel,M. (1977). "Selecting and Ordering Populations:". A New Statistical Methodology. John Wiley & Sons, New York.

(4) Moshman, J.(1958). "A Method for Selecting the Size of the Initial Sample in Stein's Two-Sample Procedure". Ann. Math. Statist. Vol. 29, 1271—1275.

(5) Rizvi, M.H. (1971). "Some Selection Problems Involving Folded Normal Distribution". Technometrics, Vol. 13, 355—369.

(6) Stein, C. (1945). "A Two-Sample Test for a Linear Hypothesis Whose Power Is Independent of the Variance." Ann. Math. Statist. Vol. 16, 243—258.

(7) Wormleighton,R. (1960). "A Useful Generalization of the Stein Two-Sample Procedure". Ann. Math. Statist. Vol. 31, 217—221.