Sankhyā : The Indian Journal of Statistics
1976, Volume **38**, Series B, Pt. 4, pp. 339-349.

A THREE-STAGE ELIMINATION TYPE PROCEDURE FOR SELECTING THE LARGEST NORMAL MEAN (COMMON UNKNOWN VARIANCE)*

By AJIT C. TAMHANE
Northwestern University, Evanston, Illinois

SUMMARY. The problem of selecting the population with the largest mean from several normal populations having a common unknown variance is considered in the context of the indifference-zone approach. A three-stage elimination type procedure is developed. For the choice of the design constants of the procedure a minimax criterion is proposed. A table of design constants is provided which can be implemented in practice. Monte Carlo sampling results are given to compare the performance of our procedure with the two-stage nonelimination type procedure of Bechhofer, Dunnett and Sobel (1954).

1. Introduction

Let $\Pi_i (1 \leqslant i \leqslant k)$ be $k \geqslant 2$ normal populations with means μ_i and a common unknown variance σ^2. Let $\boldsymbol{\mu} = (\mu_{[1]}, \ldots, \mu_{[k]})'$ where $\mu_{[1]} \leqslant \ldots \leqslant \mu_{[k]}$ denote the ordered means; we assume no prior knowledge concerning the correct pairing between Π_i and $\mu_{[j]} (1 \leqslant i, j \leqslant k)$. The experimenter's *goal* is to select the "best" population which is defined to be the population having the largest mean $\mu_{[k]}$; such a selection is referred to as a *correct selection* (CS). According to the usual *indifference-zone approach* we assume that the experimenter restricts consideration to only those procedures R which guarantee the *probability requirement*

$$P_{\boldsymbol{\mu}, \sigma}(\mathrm{CS} \mid R) \geqslant P^* \text{ whenever } \mu_{[k]} - \mu_{[k-1]} \geqslant \delta^*, \qquad \ldots \quad (1.1)$$

where $P^* (1/k < P^* < 1)$ and $\delta^* > 0$ are preassigned constants.

It is known (see Dudewicz, 1971) that a single-stage procedure, which guarantees a specified probability requirement, does not exist for this problem. Bechhofer, Dunnett and Sobel (1954) proposed a two-stage procedure R_2 in which the first stage outcome is only used to obtain an estimate of σ^2 and *not* used to eliminate the "noncontending" populations. We refer to R_2 as a *nonelimination type* procedure.

In many practical situations, viz., drug testing, the number of populations under study is often very large. Therefore a procedure incorporating a

* This work was partially supported by the Army Research Office—Durham and the Office of Naval Research at Cornell University.

preliminary screening stage to eliminate the noncontending populations from further sampling would be desirable. Paulson (1964) has given such a sequential *elimination type* procedure for the above selection problem. But sequential experiments become impractical when the time period that must elapse between the successive stages of the experiment is large, viz., agricultural field trials. Therefore what appears to be needed is a two or a three-stage procedure having the desirable elimination feature. The present paper is devoted to the development of such a *elimination type* procedure. We point out that a two-stage elimination type procedure for the above problem when σ^2 is known has been analyzed by Alam (1970) and Tamhane and Bechhofer (1977).

2. The proposed three-stage procedure R_3

In the procedure to be proposed below all populations are sampled in the first two stages; the first stage outcome is used to obtain a preliminary estimate of the unknown σ^2 and the cumulative outcome from the first two stages is used to eliminate the possible noncontenders. The selection of the "best" population is made at the end of the third stage. The rule used for retaining the "contending" populations at the end of the second stage is of the type proposed by Gupta (1965) for his subset selection problem. We remark here that a two-stage elimination type procedure was developed by us but was found to be less efficient in the Monte Carlo (MC) sampling studies; the interested reader may refer to Tamhane (1975) for further details.

Following are the steps in procedure R_3. (1) Take $n_1 \geqslant 2$ independent observations X_{ij} $(1 \leqslant j \leqslant n_1)$ from each Π_i and compute the sample means $\bar{X}_i^{(1)} (1 \leqslant i \leqslant k)$ and a pooled estimate of the variance $S_1^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_1} (X_{ij} - \bar{X}_i^{(1)})^2 / k(n_1 - 1)$. (2) Take additional $(N_2 - n_1)$ independent observations X_{ij} from each Π_i $(1 \leqslant i \leqslant k,\ n_1 + 1 \leqslant j \leqslant N_2)$ where

$$N_2 = \max \left\{ n_1,\ \left\lceil 2 \left(\frac{S_1 h_1}{\delta^*} \right)^2 \right\rceil \right\}. \qquad \ldots \quad (2.1)$$

In (2.1), $h_1 > 0$ is a constant defined in (2.3) below and $[x]$ is the smallest integer $\geqslant x$. Compute the cumulative sample means $X_i^{(1)}$, and a two-stage pooled estimate of the variance $S_2^2 = \sum_{i=1}^{k} \sum_{j=1}^{N_2} (X_{ij} - \bar{X}_i^{(2)})^2 / k(N_2 - 1)$. Select a subset I of populations where Π_i enters the subset iff $\bar{X}_i^{(2)} \geqslant \max_{1 \leqslant j \leqslant k} \bar{X}_j^{(2)} - \lambda S_1 (2/N_2)^{\frac{1}{2}}$ and where $\lambda > 0$ is a constant defined in (2.3) below. If I consists of a single population stop sampling and assert that, that population is best. (3) If I consists of more than one population

then take N_3-N_2 additional independent observations $X_{ij}(N_2+1 \leqslant j \leqslant N_3)$ for $i \in I$ where

$$N_3 = \max\left\{ N_2, \left[2\left(\frac{S_2 h_2}{\delta^*}\right)^2 \right] \right\}, \qquad \ldots \quad (2.2)$$

and $h_2 = h_2(N_2) > 0$ is defined in (2.4) below. Compute the cumulative sample means $\bar{X}_i^{(3)}$ and assert that $\max_{i \in I} \bar{X}_i^{(3)}$ is best.

We now show how to choose h_1, $h_2(N_2)$, and λ to guarantee (1.1). In the following, $F_{\nu,p}(x; \{\rho\})$ denotes the c.d.f. at the equicoordinate point x of a p-variate central t-distribution with ν degrees of freedom (d.f.) and having a common correlation $= \rho$.

Proposition 2.1 : If $h_1 > 0$ and $\lambda > 0$ are chosen to satisfy

$$F_{k(n_1-1),k-1}(h_1+\lambda; \{1/2\}) = \beta_1, \qquad \ldots \quad (2.3)$$

and having observed $N_2 = n_2$ in (2.1), $h_2 = h_2(n_2) > 0$ is chosen to satisfy

$$F_{k(n_2-1),k-1}(h_2; \{1/2\}) = \beta_2, \qquad \ldots \quad (2.4)$$

where β_1 and β_2 are preassigned constants such that $P^* < \beta_1, \beta_2 < 1$, and

$$\beta_1+\beta_2-1 = P^*, \qquad \ldots \quad (2.5)$$

then R_3 guarantees (1.1).

Proof : See the Appendix.

3. A CHOICE OF DESIGN CONSTANTS FOR R_3

In R_3 we can regard n_1, h_1, λ (and thus β_1) and β_2 as design constants. Note that we do not regard h_2 as a design constant since it is a random variable whose values depend on β_2 and N_2. Since an infinite number of combinations of design constants can guarantee (1.1), to make a choice among these we propose to use the criterion of minimizing the maximum (w.r.t.$\boldsymbol{\mu}$ for fixed σ^2) of the expected total sample size. We refer to this as the *minimax criterion.*

Let N denote the total sample size associated with R_3, $N = kN_2 + T(N_3-N_2)$, where T denotes the size of the subset retained by R_3 for sampling in the third stage. In an unabridged version of this paper, which is available from the author, we have derived an expression for $E_{\boldsymbol{\mu},\sigma}(N \mid R_3)$. We have also shown that for fixed σ^2, δ^*, k, and for fixed values of design constants, $E_{\boldsymbol{\mu},\sigma}(N \mid R_3)$ is maximized w.r.t.$\boldsymbol{\mu}$ at the equal means (EM-) configuration and if $\boldsymbol{\mu}$ is restricted to lie in the preference zone $= \{\boldsymbol{\mu} \mid \mu_{[k]}-\mu_{[k-1]} \geqslant \delta^*\}$ then $E_{\boldsymbol{\mu},\sigma}(N \mid R_3)$ is maximized at the slippage configuration $\mu_{[1]} = \mu_{[k-1]} = \mu_{[k]}-\delta^*$

(referred to as the least favorable (LF-) configuration hereafter). Therefore to determine the minimax choice of the design constants we have to minimize $E_{\mathrm{EM},\sigma}\ (N \mid R_3)$ subject to conditions (2.3), (2.4) and (2.5). Not only is this optimization problem very difficult to solve but, more importantly, even if it can be solved on a computer, the solution of the optimization problem would depend on σ which is unknown to the experimenter. Therefore the solution would be of little use in practice.

To obviate some of these difficulties in deriving the minimax choice of the design constants we shall assume that n_1 is chosen so that d.f. $= k(n_1-1)$ is sufficiently large and hence $S_1^2,\ S_2^2 \to \sigma^2$ a.s. Then (2.3) and (2.4) become respectively,

$$\Phi_{k-1}(h_1+\lambda; \{1/2\}) = \beta_1, \qquad \ldots \quad (3.1)$$

and

$$\Phi_{k-1}(h_2; \{1/2\}) = \beta_2 = 1+P^*-\beta_1, \qquad \ldots \quad (3.2)$$

where $\Phi_p(x; \{\rho\})$ represents the c.d.f. value at the equicoordinate point x of a p-variate standard normal distribution with a common correlation $= \rho$. Further

$$E_{\mathrm{EM},\sigma}(N \mid R_3) \cong 2k \left(\frac{\sigma}{\delta^*}\right)^2 [h_1^2+(h_2^2-h_1^2) \int_{-\infty}^{\infty} \{\Phi^{k-1}(x+\lambda\sqrt{2})-\Phi^{k-1}$$

$$(x-\lambda\sqrt{2})\} d\Phi(x)]+\text{terms not involving } h_1, h_2 \text{ and } \lambda, \quad \ldots \quad (3.3)$$

where $\Phi(\cdot)$ denotes the standard normal c.d.f. Thus approximately minimax values of h_1, λ and β_2 (or equivalently h_2) can be obtained by minimizing (3.3) subject to (3.1) and (3.2). Notice that the solution to this optimization problem does not depend on σ and δ^* but depends only on k and P^*. We also point out that this optimization problem is the same as the one encountered in determining the minimax choice of the design constants for the conservative two-stage procedure for the known σ^2 case (see Tamhane and Bechhofer, 1977) for which the solutions have already been tabulated for selected values of k and P^*. However, to implement these solutions in the case of unknown σ^2 would require tables of the equicoordinate points of equicorrelated multivariate t-distribution with $\rho = 1/2$ for arbitrary values of upper tail areas $(1-\beta_1)$ and $(1-\beta_2)$—not just the usual 10%, 5%, 2·5% and 1% areas. Such tables are not currently available.

To obviate the above difficulty we have to put an additional restriction (which leads to a suboptimal choice of the design constants) that $\beta_1 = \beta_2 = (1+P^*)/2$. This restriction allows us to construct tables of implementable

design constants for standard values of $P^* = 0{\cdot}90$, $0{\cdot}95$ and $0{\cdot}99$. Now by substituting $h_2 = h_1 + \lambda$ in (3.3) and differentiating it w.r.t.λ we find that the "optimum" values of h_1 and λ, say $\hat{h}_1$ and $\hat{\lambda}$, satisfy

$$2\hat{h}_1\Big[1-\int_{-\infty}^{\infty}\{\Phi^{k-1}(x+\hat{\lambda}\sqrt{2})-\Phi^{k-1}(x-\hat{\lambda}\sqrt{2})\}d\Phi(x)\Big]$$

$$= \hat{\lambda}(2\hat{h}_1+\hat{\lambda})(k-1)\varphi(\hat{\lambda})\int_{-\infty}^{\infty}\Big\{\Phi^{k-1}\Big(\frac{x+\hat{\lambda}}{\sqrt{2}}\Big)-\Phi^{k-1}\Big(\frac{x-\hat{\lambda}}{\sqrt{2}}\Big)\Big\}d\Phi(x), \quad \ldots \quad (3.4)$$

and $$\Phi_{k-1}(\hat{h}_1+\hat{\lambda};\{1/2\}) = (1+P^*)/2 = \hat{\beta}_1 = \hat{\beta}_2, \quad \ldots \quad (3.5)$$

where $\varphi(\cdot)$ denotes the standard normal p.d.f. In Table 3.1 we give the values of constants $(\hat{\lambda}, \hat{h}_1, \hat{\beta}_2)$ for selected values of k and P^*.

TABLE 3.1. VALUES OF $(\hat{\lambda}, \hat{h}_1, \hat{\beta}_1, \hat{\beta}_2)$ FOR SELECTED VALUES OF k AND P^*

k	P^*	$\hat{\lambda}$	$\hat{h}_1$	$\hat{\beta}_1 = \hat{\beta}_2$
2	0.99	2.57806	0	0.995
	0.95	1.96231	0	0.975
	0.90	1.64692	0	0.950
3	0.99	1.05983	1.73444	0.995
	0.95	1.02113	1.19101	0.975
	0.90	0.99112	0.92521	0.950
4	0.99	0.95636	1.95864	0.995
	0.95	0.92951	1.41947	0.975
	0.90	0.90956	1.15252	0.950
5	0.99	0.93227	2.06547	0.995
	0.95	0.90881	1.53297	0.975
	0.90	0.89180	1.26853	0.950
10	0.99	0.96788	2.25107	0.995
	0.95	0.94726	1.73897	0.975
	0.90	0.93301	1.48401	0.950
25	0.99	1.09494	2.37042	0.995
	0.95	1.07451	1.87843	0.975
	0.90	1.06084	1.63316	0.950

For $k = 2$, $\hat{h}_1 = 0$ implies R_3 reduces to a two-stage procedure. The entries in the table are correct to 4 decimal places.

Since $F_{k(n_1-1),k-1}(\hat{h}_1+\hat{\lambda};\{1/2\}) < \Phi_{k-1}(\hat{h}_1+\hat{\lambda};\{1/2\}) = \hat{\beta}_1$ for $k(n_1-1) < \infty$, for implementation purposes we must adjust $\hat{h}_1$ and $\hat{\lambda}$ to some new values,

say $\hat{h}_1'$ and $\hat{\lambda}'$, such that $F_{k(n_1-1),k-1}(\hat{h}_1'+\hat{\lambda}'; \{1/2\}) = \hat{\beta}_1$ and hence $\hat{h}_1'+\hat{\lambda}' > \hat{h}_1+\hat{\lambda}$. Given $\hat{h}_1'+\hat{\lambda}'$, which can be found from Krishniah and Armitage (1966) or Gupta and Sobel (1957) for selected values of k, d.f. $= k(n_1-1)$, and $\hat{\beta}_1$, we suggest that $\hat{h}_1'$ and $\hat{\lambda}'$ be chosen so that $\hat{h}_1'/\hat{\lambda}' = \hat{h}_1/\hat{\lambda}$.

4. MONTE-CARLO SAMPLING RESULTS

If R_2 and R_3 are both designed to guarantee (1.1) then to compare their performances we define the relative efficiency (RE) of R_2 w.r.t. R_3 by taking the ratio of their expected total sample sizes as follows

$$\mathrm{RE}_{\mu,\sigma}(\delta^*, P^*, k \mid n_0, n_1, \lambda, h_1, \beta_2; R_2/R_3) = \frac{E_{\mu,\sigma}(N \mid R_3)}{E_{\mu,\sigma}(N \mid R_2)} \qquad \ldots \quad (4.1)$$

where n_0 = the first stage sample size for R_2. An expression for the expected total sample size associated with R_2, $E_{\mu,\sigma}(N \mid R_2)$, was given by Bechhofer *et al.* (1954). Note that $E_{\mu,\sigma}(N \mid R_2)$ does not depend on μ.

Due to the complicated nature of the expressions for $E_{\mu,\sigma}(N \mid R_2)$ and $E_{\mu,\sigma}(N \mid R_3)$, a direct analytical comparison between R_2 and R_3 appears to be difficult. Therefore our comparison is based on the results of MC simulations. In these simulation experiments we were particularly interested in studying (1) the performance of R_3 relative to R_2 in terms of the RE-values at the EM- and LF-configurations for selected values of δ^*, P^*, k, σ^2 and $n_0 = n_1$, and (2) the extent of overprotection in terms of excess $P(\mathrm{CS})$ over guaranteed P^* afforded by R_3 when the μ_i are in the LF-configuration.

The MC experiments were performed for $k = 5$, 10, and 25 and $P^* = 0{\cdot}90$ and $0{\cdot}95$; the value of δ^* was kept fixed throughout at $0{\cdot}5$. For each (k, P^*)-combination we considered $\sigma^2 = 1$ and 5. The first stage sample sizes for R_2 and R_3 were chosen to obtain a reasonably large (at least 20 to 30) number of d.f. for the estimation of σ^2. Same values of $\bar{X}_i^{(l)}$ $(1 \leqslant i \leqslant k)$ and S_l^2 were used for simulating the l-th stage $(l = 1, 2)$ of both R_2 and R_3; the purpose being to eliminate the difference in the performances of R_2 and R_3 on this account. The values of $\bar{X}_i^{(l)}$ and S_l^2 were obtained by generating standard normal and chi-square random variables (r.v.'s) by usual techniques. For each run 1000 experiments were conducted. The results of the simulation studies are given in Table 4.1. A summary comparison of R_2 and R_3 in terms of the MC estimates of the RE-values at the LF- and EM-configurations is given in Table 4.2.

TABLE 4.1. MONTE CARLO ESTIMATES FOR R_2 AND R_3

($P^* = 0.90$)

k	n_1	σ^2	$\hat{h}_1'$	$\hat{\lambda}'$	prob. of correct selection		total sample size			subset size	
					$P_{LF,\sigma}(CS\|R_2)$	$P_{LF,\sigma}(CS\|R_3)$	$E_{\mu,\sigma}(N\|R_2)$	$E_{LF,\sigma}(N\|R_3)$	$E_{EM,\sigma}(N\|R_3)$	$E_{LF,\sigma}(T\|R_3)$	$E_{EM,\sigma}(T\|R_3)$
5	7	1.0	1.32118	0.92882	0.912	0.924	146.93	112.79	139.47	1.715	2.812
					(.0090)	(.0084)	(1.12)	(1.35)	(1.49)	(.049)	(.047)
5	7	5.0	1.32118	0.92882	0.900	0.919	717.81	561.61	678.11	1.820	2.847
					(.0095)	(.0095)	(6.04)	(5.99)	(6.02)	(.048)	(.047)
10	4	1.0	1.55952	0.98048	0.908	0.921	385.60	267.61	321.61	2.504	4.542
					(.0091)	(.0085)	(2.99)	(2.30)	(2.42)	(.069)	(.069)
10	4	5.0	1.55952	0.98048	0.894	0.909	1928.6	1334.8	1591.3	2.593	4.614
					(.0097)	(.0091)	(16.0)	(11.3)	(11.9)	(.071)	(.073)
25	5	1.0	1.65893	1.07757	0.889	0.913	1181.9	750.06	880.43	5.235	8.850
					(.0099)	(.0089)	(5.19)	(5.37)	(5.48)	(.133)	(.134)
25	5	5.0	1.65893	1.07757	0.902	0.909	5898.6	3668.7	4336.5	5.016	8.794
					(.0094)	(.0091)	(26.9)	(25.0)	(27.3)	(.127)	(.135)

The numbers in round brackets are standard errors.

TABLE 4.1. (*Contd.*) MONTE CARLO ESTIMATES FOR R_2 AND R_3

($P^* = 0.95$)

k	n_1	σ^2	$\hat{h}_1'$	$\hat{\lambda}'$	prob. of correct selection		total sample size			subset size	
					$P_{LF,\sigma}(CS\|R_2)$	$P_{LF,\sigma}(CS\|R_3)$	$E_{\mu,\sigma}(N\|R_2)$	$E_{LF,\sigma}(N\|R_3)$	$E_{EM,\sigma}(N\|R_3)$	$E_{LF,\sigma}(T\|R_3)$	$E_{EM,\sigma}(T\|R_3)$
5	7	1.0	1.61974	0.96026	0.944	0.961	205.19	146.17	185.25	1.401	2.862
					(.0073)	(.0061)	(1.62)	(1.53)	(1.65)	(.047)	(.045)
5	7	5.0	1.61974	0.96026	0.951	0.959	1003.9	716.84	904.16	1.386	2.905
					(.0068)	(.0063)	(7.9)	(6.88)	(7.04)	(.046)	(.047)
10	4	1.0	1.84499	1.00501	0.955	0.962	522.96	343.98	420.61	2.110	4.782
					(.0066)	(.0060)	(4.19)	(2.75)	(2.90)	(.068)	(.073)
10	4	5.0	1.84499	1.00501	0.945	0.953	2592.2	1682.1	2064.3	2.023	4.697
					(.0072)	(.0067)	(20.8)	(13.1)	(13.6)	(.066)	(.072)
25	5	1.0	1.91618	1.09610	0.951	0.949	1512.2	931.4	1115.9	4.546	9.237
					(.0068)	(.0070)	(6.74)	(5.62)	(6.02)	(.126)	(.138)
25	5	5.0	1.91618	1.09610	0.951	0.961	7515.0	4582.4	5497.1	4.430	9.064
					(.0068)	(.0061)	(34.2)	(28.2)	(31.0)	(.122)	(.135)

The numbers in round brackets are standard errors.

An examination of Table 4.2 reveals that for all the cases considered, the performance of R_3 is consistently better than that of R_2 even in the "worst" possible parameter configuration of equal means. For fixed P^* and σ^2, relative performance of R_3 improves with k. This is due to the fact that the screening aspect of R_3 becomes more effective for larger k which is also evident from the values of the expected retained subset sizes given in Table 4.1. In practice, the μ_i's would be spread somewhat further apart and therefore even greater savings would result by using R_3 than those indicated by the RE-values in the LF- and EM-configurations. We also note that the choice of design constants $(\hat{\lambda}', \hat{h}'_1, \hat{\beta}_2)$ which is determined without the knowledge of σ^2 appears to work well for a range of σ^2-values. Finally we note that R_3 provides a greater overprotection in terms $P(\mathrm{CS})$ whereas R_2 does not provide any appreciable overprotection. This conservative nature of R_3 is due to the Bonferroni inequality used in determining the conditions to be satisfied by (λ, h_1, β_2) for guaranteeing (1.1) (see the proof of Proposition 2.1).

TABLE 4.2. ESTIMATED VALUES OF THE RELATIVE EFFICIENCY OF R_2 w.r.t. R_3

k	P^*	σ^2	$\mathrm{RE}_{\mathrm{LF},\sigma}$	$\mathrm{RE}_{\mathrm{EM},\sigma}$
5	0.90	1.0	0.768	0.949
		5.0	0.782	0.945
5	0.95	1.0	0.713	0.903
		5.0	0.714	0.900
10	0.90	1.0	0.694	0.834
		5.0	0.692	0.825
10	0.95	1.0	0.658	0.804
		5.0	0.649	0.796
25	0.90	1.0	0.635	0.745
		5.0	0.622	0.735
25	0.95	1.0	0.616	0.738
		5.0	0.610	0.731

ACKNOWLEDGEMENTS

The author is very grateful to Professor R. E. Bechhofer, Cornell University from whose guidance and advice he has immensely benefitted. The author also wishes to thank Mr. Joseph Voelkel, Northwestern University who helped in computing the tables given in the paper.

Appendix

Proof of proposition 2.1 : Denote by $\bar{X}_{(i)}^{(l)}$ = the cumulative sample mean up to the l-th stage from the population having the mean $\mu_{[i]}(1 \leqslant i \leqslant k;$ $l = 2, 3)$ and $\delta_{ij} = \mu_{[i]} - \mu_{[j]}(1 \leqslant i, j \leqslant k)$. We have

$$
\begin{aligned}
&1 - P_{\mu,\sigma}(\mathrm{CS} \mid R_3) \\
&\leqslant P_{\mu,\sigma}\{\bar{X}_{(k)}^{(2)} < X_{(i)}^{(2)} - \lambda S_1(2/N_2)^{1/2} \text{ for some } i \neq k\} \\
&\qquad + P_{\mu,\sigma}\{\bar{X}_{(k)}^{(3)} < \bar{X}_{(i)}^{(3)} \text{ for some } i \neq k\} \\
&= 2 - P_{\mu,\sigma}\{T_i^{(2)} \leqslant (\delta_{ki}/S_1)(N_2/2)^{1/2} + \lambda \ \forall\ i \neq k\} \\
&\qquad - P_{\mu,\sigma}\{T_i^{(3)} \leqslant (\delta_{ki}/S_2)(N_3/2)^{1/2} \ \forall\ i \neq k\} \\
&\leqslant 2 - P_{\mu,\sigma}\{T_i^{(2)} \leqslant (\delta^*/S_1)(N_2/2)^{1/2} + \lambda \ \forall\ i \neq k\} \\
&\qquad - P_{\mu,\sigma}\{T_i^{(3)} \leqslant (\delta^*/S_1)(N_3/2)^{1/2} \ \forall\ i \neq k\} \\
&= 2 - P_1 - P_2 \text{ (say)}, \qquad \ldots \text{ (A.1)}
\end{aligned}
$$

whenever $\delta_{k,k-1} \geqslant \delta^*$. In (A.1) $T_i^{(l)} = \{\bar{X}_{(i)}^{(l)} - \bar{X}_{(k)}^{(l)} + \delta_{ki})\sqrt{N_l}\}/S_{l-1}\sqrt{2}$ $(1 \leqslant i \leqslant k-1;\ l = 2, 3)$. It is straightforward to check that $(T_1^{(2)}, \ldots, T_{k-1}^{(2)})$ and $(T_1^{(3)}, \ldots, T_{k-1}^{(3)})$ each have a $(k-1)$-variate central t-distribution with equal correlation $= 1/2$; the former has $k(n_1-1)$ d.f. and the latter has $k(N_2-1)$ d.f. (random) associated with it. Now from (2.1) we have $h_1 \leqslant (\delta^*/S_1)(N_2/2)^{1/2}$ and therefore $P_1 \geqslant F_{k(n_1-1),k-1}(h_1+\lambda; \{1/2\}) = \beta_1$. Next conditioning on $N_2 = n_2$ and from (2.2) noting that $h_2(n_2) \leqslant (\delta^*/S_2)(N_3/2)^{1/2}$ we have $P_2 \geqslant E\{F_{k(N_2-1),k-1}(h_2(N_2); \{1/2\}) \mid N_2 = n_2\} = \beta_2$. Substituting in (A.1) and using (2.5) we obtain $P_{\mu,\sigma}(\mathrm{CS} \mid R_3) \geqslant P^*$ whenever $\delta_{k,k-1} \geqslant \delta^*$.

References

Alam, K. (1970) : A two-sample procedure for selecting the population with the largest mean from k normal populations. *Ann. Inst. Statist. Math.*, **22**, 127–136.

Bechhofer, R. E., Dunnett, C. W., and Sobel, M. (1954) : A two-sample multiple decision rule for ranking means of normal populations with a common unknown variance. *Biometrika*, **41**, 170–176.

Dudewicz, E. J. (1971) : Nonexistence of a single-sample selection procedure whose $P(\mathrm{CS})$ is independent of the variances. *South African Statist. J.*, **5**, 37–39.

Gupta, S. S. (1965) : On some multiple decision (selection and ranking) rules. *Technometrics*, **7**, 225–245.

GUPTA, S. S., and SOBEL, M. (1957) : On a statistic which arises in selection and ranking problems. *Ann. Math. Statist.*, **28**, 957–967.

KRISHNIAH, P. R., and ARMITAGE, J. V. (1966) : Tables for multivariate t-distribution. *Sankhyā*, Ser. B, **28**, 31–56.

PAULSON, E. (1964) : A sequential procedure for selecting the population with the largest mean from k normal populations. *Ann. Math. Statist.*, **35**, 174–180.

TAMHANE, A. C. (1975) : On minimax multistage elimination type rules for selecting the largest normal mean. Tech. Rep. No. 259, Dept. of Operations Research, Cornell University, Ithaca, N.Y.

TAMHANE, A. C., and BECHHOFER, R. E. (1977) : A two-stage minimax procedure with screening for selecting the largest normal means. To appear in *Commun. Statist.*, Ser. A. (Special Issue on Ranking and Selection).

Paper received : June, 1976.
Revised : December, 1976.