The Annals of Mathematical Statistics
1969, Vol. 40, No. 4, 1300–1324

ON PARTITIONING A SET OF NORMAL POPULATIONS BY THEIR LOCATIONS WITH RESPECT TO A CONTROL[1]

BY YUNG LIANG TONG
University of Nebraska

0. The problem and the approaches. This paper is concerned with a problem of partitioning a set of normal populations into two subsets according to their locations with respect to a control population, based on indifference zone formulation. Let $\Pi_0, \Pi_1, \cdots, \Pi_k$ be $(k+1)$ normal populations with means $\mu_0, \mu_1, \cdots, \mu_k$ and a common variance σ^2; and let Π_0 denote the standard or control population. For arbitrary but fixed constants δ_1^* and δ_2^* such that $\delta_1^* < \delta_2^*$, we define three disjoint and exhaustive subsets Ω_B, Ω_I and Ω_G of the set

$$\Omega = (\Pi_1, \Pi_2, \cdots, \Pi_k) \tag{0.1}$$

by

$$\begin{aligned} \Omega_B &= (\Pi_i : \mu_i \leqq \mu_0 + \delta_1^*) \\ \Omega_I &= (\Pi_i : \mu_0 + \delta_1^* < \mu_i < \mu_0 + \delta_2^*) \\ \Omega_G &= (\Pi_i : \mu_i \geqq \mu_0 + \delta_2^*). \end{aligned} \tag{0.2}$$

After observations have been taken, the set Ω is partitioned into two disjoint subsets S_B and S_G.

DEFINITION 0.1. A decision is a correct decision (CD) if $\Omega_B \subset S_B$ and $\Omega_G \subset S_G$.

An equivalent definition to Definition 0.1 is that $S_B \subset (\Omega_B \cup \Omega_I)$ and $S_G \subset (\Omega_G \cup \Omega_I)$. It is noted that the open interval $(\mu_0 + \delta_1^*, \mu_0 + \delta_2^*)$ is considered as the indifference zone and a correct decision puts no restrictions on those populations in the set Ω_I. With this consideration, it will be consistent to give the following

DEFINITION 0.2. A population $\Pi_i \varepsilon \Omega$ is misclassified if $\Pi_i \varepsilon (\Omega_B \cap S_G) \cup (\Omega_G \cap S_B)$.

Let P^* be an arbitrary but preassigned constant such that $2^{-k} < P^* < 1$. The statistical problem is to find a procedure R which consists of a sampling procedure and a terminal decision rule such that the appropriate probability requirement below is satisfied.

(1) When σ^2 is known,

$$P[CD \mid \boldsymbol{\mu}, \sigma^2; R] \geqq P^* \quad \text{for every mean vector } \boldsymbol{\mu}. \tag{0.3}$$

(2) When σ^2 is unknown,

$$P[CD \mid \boldsymbol{\mu}, \sigma^2; R] \geqq P^* \quad \text{for every } \boldsymbol{\mu} \text{ and every } \sigma^2 > 0. \tag{0.4}$$

Received 18 December 1967; revised 3 January 1969.

[1] This paper forms a part of the author's doctoral thesis submitted to the University of Minnesota. The research was partially supported by the National Science Foundation under grant No. NSF-GP-3813.

The case of known σ^2 is considered in Section 1. A single-stage procedure is used there and the proposed decision rule is proved to be Bayes, minimax and admissible among a class of translation invariant decision rules. It is clear that when σ^2 is unknown, there is no single-stage procedure that can solve this problem. A two-stage procedure originally proposed by Stein [16] is used in Section 2; and a sequential procedure based on the idea of the random stopping rule developed by Chow and Robbins [3] is used in Section 3 to serve as an alternative to the two-stage procedure.

The (expected) sample size required, the expected misclassification size, the relative efficiency and their asymptotic behavior for the single-stage, two-stage and sequential procedures are investigated and are shown to be functions of the percentage points of a multivariate normal and a multivariate t distribution. Tables of these percentage points have been constructed and are attached with this paper.

The following assumptions are made throughout this paper:

(1) there is no a priori knowledge regarding the means of the populations;

(2) the observations are taken a vector at a time; and

(3) the observations are independent.

Unless mentioned otherwise, the following notations will be adopted:

$$\Phi(z) = \int_{-\infty}^{z} (2\pi)^{-\frac{1}{2}} e^{-x^2/2}\, dx, \qquad -\infty < z < \infty; \tag{0.5}$$

$$\begin{aligned} m &= k/2 && \text{if } k \text{ is even,} \\ &= (k+1)/2 && \text{if } k \text{ is odd;} \end{aligned} \tag{0.6}$$

$$d = (\delta_1^* + \delta_2^*)/2; \tag{0.7}$$

$$a = (\delta_2^* - \delta_1^*)/2; \quad \text{and} \tag{0.8}$$

$$\lambda = \sigma/a. \tag{0.9}$$

1. A single-stage procedure.

1.1 *The Procedure and its* PCD. Let

$$\{X_{0j}, X_{1j}, \cdots, X_{kj}\}_{j=1}^{\infty} \tag{1.1}$$

be a sequence of independent vector observations from the population with joint density

$$f(x_0, x_1, \cdots, x_k; \mu_0, \mu_1, \cdots, \mu_k, \sigma^2) = \prod_{i=0}^{k} (2\pi)^{-\frac{1}{2}} \sigma^{-1} e^{-\frac{1}{2}\sigma^{-2}(x_i-\mu_i)^2} \tag{1.2}$$

for $-\infty < x_i < \infty$ and with parameter spaces $-\infty < \mu_i < \infty$ $(i = 0, 1, \cdots, k)$. Throughout this section, we assume that σ^2 is a known constant.

The decision rule used is based on the differences of the sample means.

Procedure R_1. Observe the sequence defined in (1.1) for $j = 1, 2, \cdots, N_0$ where N_0 is to be determined below. Compute $\bar{X}_i = N_0^{-1} \sum_{j=1}^{N_0} X_{ij}$ for $i = 0, 1, \cdots, k$, and use the decision rule:

$$\begin{aligned} S_B &= \{\Pi_i : \bar{X}_i - \bar{X}_0 < d\}, \\ S_G &= \{\Pi_i : \bar{X}_i - \bar{X}_0 > d\}. \end{aligned} \tag{1.3}$$

To find the sample size N_0 such that (0.3) is satisfied, we first give the following

DEFINITION 1.1. A mean vector $\boldsymbol{\mu}^0 = (\mu_0^0, \mu_1^0, \cdots, \mu_k^0)$ is a least favorable (LF) configuration under a procedure R if

$$(1.4) \qquad P[CD \mid \boldsymbol{\mu}^0, \sigma^2; R] = \inf_{\boldsymbol{\mu}} P[CD \mid \boldsymbol{\mu}, \sigma^2; R].$$

It is clear that for a mean vector to be a LF configuration under R_1, the set Ω_I defined in (0.2) must be empty, all the populations in Ω_B must have a common mean $\mu_0 + \delta_1^*$, and all the populations in Ω_G must have a common mean $\mu_0 + \delta_2^*$. Without loss of generality, let $\boldsymbol{\mu}^0(r)$ be a configuration such that $\mu_i = \mu_0 + \delta_1^*$ and $\mu_j = \mu_0 + \delta_2^*$, $0 < i \leqq r$, $r < j \leqq k$ for some r such that $0 < r \leqq k$. Then it follows from Definition 0.1 that

$$\begin{aligned} P[CD \mid \boldsymbol{\mu}^0(r), \sigma^2; R_1] & \\ &= P[\bar{X}_i - \bar{X}_0 < d, \bar{X}_j - \bar{X}_0 > d(0 < i \leqq r, r < j \leqq k) \mid \boldsymbol{\mu}^0(r), \sigma^2] \\ &= P[Z_i - Z_0 < (\tfrac{1}{2}N_0)^{\frac{1}{2}}/\lambda, Z_j - Z_0 > -(\tfrac{1}{2}N_0)^{\frac{1}{2}}/\lambda(0 < i \leqq r, r < j \leqq k)] \\ &= P[Y_i \leqq (\tfrac{1}{2}N_0)^{\frac{1}{2}}/\lambda(i = 1, 2, \cdots, k)]; \end{aligned}$$

where $Z_i = (\bar{X}_i - \mu_i)/(\sigma\ (2/N_0)^{\frac{1}{2}})$ for $i = 0, 1, \cdots, k$, $Y_i = Z_i - Z_0$ for $0 < i \leqq r$ and $Y_i = Z_0 - Z_i$ for $r < i \leqq k$. Hence if we define the $(k \times k)$ covariance matrix $\boldsymbol{\Sigma}_r = (\sigma_{ij})$ by

$$(1.5) \qquad \begin{aligned} \sigma_{ij} &= 1 \quad \text{for} \quad i = j \\ &= 1/2 \quad \text{for} \quad i \neq j, \quad \text{and} \quad 0 < i, j \leqq r \quad \text{or} \quad r < i, j \leqq k \\ &= -1/2 \quad \text{for} \quad 0 < i \leqq r \quad \text{and} \quad r < j \leqq k; \end{aligned}$$

then

$$(1.6) \qquad P[CD \mid \boldsymbol{\mu}^0(r), \sigma^2; R_1] = \int_{-\infty}^{(\frac{1}{2}N_0)^{\frac{1}{2}}/\lambda} \int_{-\infty}^{(\frac{1}{2}N_0)^{\frac{1}{2}}/\lambda} \cdots \int_{-\infty}^{(\frac{1}{2}N_0)^{\frac{1}{2}}/\lambda} \cdot (2\pi)^{-k/2} \mid \boldsymbol{\Sigma}_r \mid^{-1/2} \exp\left(-\tfrac{1}{2}\mathbf{y}'\boldsymbol{\Sigma}_r^{-1}\mathbf{y}\right) \prod_{i=1}^{k} dy_i .$$

Equation (1.6) gives the infimum of the PCD under R_1 for the set of all configurations such that there are r populations in Ω_B and $(k - r)$ populations in Ω_G. To find the LF configuration under R_1 it suffices to find the integer where the rhs of (1.6) achieves a minimum over all $r(0 < r \leqq k)$.

LEMMA 1.1. *For every λ and every N_0, the* LF *configuration under R_1 is given by $\boldsymbol{\mu} = \boldsymbol{\mu}^0$ such that*

$$(1.7) \qquad \begin{aligned} \mu_1^0 = \mu_2^0 = \cdots = \mu_m^0 &= \mu_0 + \delta_1^*, \\ \mu_{m+1}^0 = \mu_{m+2}^0 = \cdots = \mu_k^0 &= \mu_0 + \delta_2^*, \end{aligned}$$

where m is defined in (0.6).

PROOF. The proof of this lemma follows from a general theorem given in the Appendix.

Now let $\Sigma = \Sigma_m$ denote the $(k \times k)$ covariance matrix defined in (1.5) for $\gamma = m$ i.e., Σ has the following structure:

$$\Sigma = \begin{array}{c} \begin{array}{cc} 1\ 2\cdots m & m+1\cdots k \end{array} \\ \left[\begin{array}{ccc|ccc} 1 & & \frac{1}{2} & & & \\ & \ddots & & & -\frac{1}{2} & \\ \frac{1}{2} & & 1 & & & \\ \hline & & & 1 & & \frac{1}{2} \\ & -\frac{1}{2} & & & \ddots & \\ & & & \frac{1}{2} & & 1 \end{array}\right] \end{array} \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ m \\ m+1 \\ \vdots \\ k \end{array} \tag{1.8}$$

Let $b = b(P, k)$ be the solution of the equation

$$P = \int_{-\infty}^{b} \int_{-\infty}^{b} \cdots \int_{-\infty}^{b} (2\pi)^{-k/2} |\Sigma|^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}\mathbf{y}'\Sigma^{-1}\mathbf{y}\right) \prod_{i=1}^{k} dy_i . \tag{1.9}$$

Then the sample size N_0 required under R_1 is given by

THEOREM 1.1. *Let λ be defined in* (0.9) *and b be the solution of* (1.9) *with $P = P^*$. If N_0 is the smallest integer satisfying*

$$N_0 \geqq 2\lambda^2 b^2 \tag{1.10}$$

then the probability requirement (0.3) *is satisfied.*

PROOF. For any mean vector $\mathbf{\mu}$, it follows from $b \leqq (\frac{1}{2}N_0)^{\frac{1}{2}}/\lambda$ that $P[CD \mid \mathbf{\mu}, \sigma^2; R_1] \geqq P[CD \mid \mathbf{\mu}^0, \sigma^2; R_1] \geqq P^*$.

The solution $b = b(P, k)$ of (1.9) is the equi-coordinate percentage point of a k-dimensional multivariate normal distribution with mean vector $\mathbf{0}$ and the covariance matrix Σ given in (1.8). The values of b as a function of P and k have been tabulated in Table 1 for $P = 0.50, 0.75, 0.90, 0.95, 0.975, 0.99$ and $k = 1(1)10(2)20$. It should be noted that for $k = 1$ the table reduces to the univariate standard normal table. The numerical solution was obtained by first changing (1.9) to a form of simple integral, this simple integral is then approximated by a Gaussian quadrature summation formula given in [17]. To be conservative, the entries in the table have all been rounded to the next higher value (in the 7th decimal) and should be in error by at most one unit in the last digit given.

1.2. *An upper bound on the sample size required.* We now give an upper bound on the sample size N_0 under R_1 based on the following

LEMMA 1.2. *For any given $P \in [0, 1]$ and any two events A and B,*

$$P(A) + P(B) = 1 + P \Rightarrow P(A \cap B) \geqq P, \tag{1.11}$$

and the equality holds iff $P(A \cup B) = 1$.

PROOF. It is an immediate consequence of the inequality

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 1 + P - P(A \cup B) \geq P,$$

and the equality holds iff $P(A \cup B) = 1$.

For any real number c and positive integer q, let

TABLE 1

Equi-coordinate percentage points b of a multivariate normal distribution with mean vector 0 and covariance matrix Σ

k	P 0.50	0.75	0.90	0.95	0.975	0.99
1	0.0000000	0.6744898	1.2815516	1.6448537	1.9599640	2.3263479
2	0.6423429	1.1462928	1.6445631	1.9599246	2.2413975	2.5758290
3	0.8370415	1.3192980	1.8003977	2.1057358	2.3786364	2.7033911
4	0.9938965	1.4528031	1.9162111	2.2121205	2.4775016	2.7942727
5	1.0890009	1.5389483	1.9950311	2.2865328	2.5480781	2.8604419
6	1.1742510	1.6140189	2.0620112	2.3489679	2.6067571	2.9149993
7	1.2356655	1.6702228	2.1138621	2.3981570	2.6536084	2.9591380
8	1.2928724	1.7214957	2.1602823	2.4417695	2.6948543	2.9977379
9	1.3376440	1.7627532	2.1985565	2.4781993	2.7296540	3.0306329
10	1.3801626	1.8012938	2.2337781	2.5114623	2.7612442	3.0603289
12	1.4486915	1.8643484	2.2921627	2.5669962	2.8142841	3.1104789
14	1.5048107	1.9162434	2.3404131	2.6130015	2.8583140	3.1522062
16	1.5521539	1.9601991	2.3814180	2.6521756	2.8958693	3.1878656
18	1.5929848	1.9982343	2.4169983	2.6862232	2.9285565	3.2189529
20	1.6288041	2.0316945	2.4483726	2.7162884	2.9574557	3.2464759

$$(1.12)\quad H_q(c) = \int_{-\infty}^{c}\int_{-\infty}^{c}\cdots\int_{-\infty}^{c}(2\pi)^{-q/2}\,|\Sigma'|^{-\frac{1}{2}}\exp\left[-\tfrac{1}{2}\mathbf{y}'(\Sigma')^{-1}\mathbf{y}\right]\prod_{i=1}^{q}dy_i$$

where the $(q \times q)$ covariance matrix $\Sigma' = (\sigma_{ij})$ is such that

$$(1.13)\quad \begin{aligned}\sigma_{ij} &= 1 \quad \text{if } i = j;\\ &= \tfrac{1}{2} \quad \text{if } i \neq j.\end{aligned}$$

Let b be the solution of equation (1.9) and b' be the solution of the equation

$$(1.14)\quad H_m(b') + H_{k-m}(b') = 1 + P.$$

THEOREM 1.2. *For every P and every k, we have*

$$(1.15)\quad b' > b.$$

PROOF. Let $(Y_1, Y_2, \cdots, Y_k)$ follow a multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix Σ, and let

$$A = [Y_i \leqq b' \ (i = 1, 2, \cdots, m)],$$
$$B = [Y_i \leqq b' \ (i = m+1, m+2, \cdots, k)];$$

then $H_m(b') + H_{k-m}(b') = 1 + P \Rightarrow P(A \cap B) = P[Y_i \leqq b'(i = 1, 2, \cdots, k)] > P$. It follows that $b' > b$ and this completes the proof.

COROLLARY. *If N_0' is the smallest integer satisfying*

$$(1.16)\quad N_0' \geqq 2\lambda^2 b'^2,$$

where b' is the solution of (1.14) with $P = P^*$, then $N_0' \geqq N_0$.

When k is even equation (1.14) reduces to

$$H_{k/2}(b') = \tfrac{1}{2}(1 + P); \tag{1.17}$$

the solution b' of (1.17) is the percentage point of an equi-correlated multivariate normal distribution. The numerical solutions have been tabulated by both Gupta [6] and Milton [12] at several probability levels. Let $\gamma = \gamma(P^*, k)$ denote the quantity $(b'/b)^2$ with $P = P^*$, then

$$N_0'/N_0 = \gamma. \tag{1.18}$$

These γ values have been computed for even k based on the b' values given by Milton and the b values given in Table 1 of this paper; an excerpt is given below.

Values of γ for Selected Values of P^* and k

P^*	k					
	2	4	6	8	12	16
0.50	1.102597	1.040598	1.026127	1.019653	1.013570	1.010628
0.90	1.000353	1.000126	1.000071	1.000047	1.000027	1.000018
0.95	1.000040	1.000013	1.000007	1.000005	1.000002	1.000001

The computation shows the bound given in (1.16) is quite good for most purposes since most of the γ values are very close to 1. Of course, the value of $b'/b = \gamma^{\frac{1}{2}}$ is even closer to one. It also appears that $\gamma(P^*, k)$ is monotonically decreasing both in P^* and in k; however, the author has not tried to prove this result.

1.3. *Some optimal properties of the decision rule.* We now prove some of the optimal properties of the proposed decision rule specified in (1.3). Let $\bar{X} = (\bar{X}_0, \bar{X}_1, \cdots, \bar{X}_k)$ where $\bar{X}_i$ is the sample mean from Π_i, $i = 0, 1, \cdots, k$. Since $\bar{X}$ is a sufficient statistic, there is no loss in considering only decision rules based on $\bar{X}$. Consider a group G of translations where $g \varepsilon G$ is defined by

$$g(x_0, x_1, \cdots, x_k) = (x_0 + c, x_1 + c, \cdots, x_k + c), \qquad -\infty < c < \infty. \tag{1.19}$$

This group of translations in turn induces the group $\bar{G}$ of translations on the parameter space of $(\mu_0, \mu_1, \cdots, \mu_k)$ with elements $\bar{g} \varepsilon \bar{G}$ given by

$$\bar{g}(\mu_0, \mu_1, \cdots, \mu_k) = (\mu_0 + c, \mu_1 + c, \cdots, \mu_k + c), \qquad -\infty < c < \infty. \tag{1.20}$$

Clearly, our problem remains invariant under G. It follows that (see [9: p. 216]) $\mathbf{Z} = (Z_1, Z_2, \cdots, Z_k)$, where $Z_i = \bar{X}_i - \bar{X}_0 (i = 1, 2, \cdots, k)$, is a maximal invariant wrt G. Also, the distribution of $\mathbf{Z}$ depends only on $(\mu_1 - \mu_0, \mu_2 - \mu_0, \cdots, \mu_k - \mu_0) = (\theta_1, \theta_2, \cdots, \theta_k) = \boldsymbol{\theta}$ (say), which is the maximal invariant of the induced group $\bar{G}$, as it should according to [9: Theorem 3, p. 220] (here $\mathbf{Z}$ has a multivariate normal distribution with mean vector $\boldsymbol{\theta}$ and known covariance matrix). Hence if we restrict our attention to the class of invariant decision rules under G, those rules must be a function of $\mathbf{Z}$ (see [9: Theorem 1, p. 216]), and the probabilities under those decision rules depend on $(\mu_0, \mu_1, \cdots, \mu_k)$ only through $\boldsymbol{\theta}$.

In the following we show that the decision rule specified in (1.3) is Bayes, minimax and admissible among the class of translation invariant decision rules. Denote

$$\psi = \{\boldsymbol{\theta}:\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_k)\}. \tag{1.21}$$

Following from the above discussion, it is sufficient to consider the only decision rules based on $\mathbf{Z}$ and the induced parameter space ψ. We first formulate our problem under the general framework of multiple decision problems developed by Lehmann [8].

DEFINITION 1.2. For $i = 1, 2, \cdots, k$, let (ψ_i, D_i, L_i) be k component statistical decision problems where ψ_i is the parameter space, D_i is the decision space and $L_i:\psi_i \times D_i \rightarrow (-\infty, \infty)$ is the loss function for the ith component decision problem. The decision problem (ψ, D, L) is said to be the corresponding product decision problem if

(1) $\psi = \times_{i=1}^{k} \psi_i = \{\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_k):\theta_i \varepsilon \psi_i, i = 1, 2, \cdots, k\}$ is the product parameter space,

(2) $D = \times_{i=1}^{k} D_i = \{\mathbf{a} = (a_1, a_2, \cdots, a_k):a_i \varepsilon D_i, i = 1, 2, \cdots, k\}$ is the product decision space, and

(3) $L = L(\boldsymbol{\theta}, \mathbf{a})$ is the loss function defined on $\psi \times D$.

REMARK. The problem of incompatibility of two component decision rules, which was discussed by Lehmann, does not arise in our formulation; hence it is ignored here.

Now for $i = 1, 2, \cdots, k$, let (ψ_i, D_i, L_i) be the component decision problem dealing with the population mean of Z_i (note that Z_i has a normal distribution with mean θ_i and known variance). With the term "misclassification" defined in Definition 0.2, we will consider the loss function for the product decision problem to be the total number of populations misclassified; i.e., let

$$\begin{aligned} L_i(\boldsymbol{\theta}, \mathbf{a}) = L_i(\theta_i, a_i) &= 1 \qquad \text{if } \Pi_i \text{ is misclassified,} \\ &= 0 \qquad \text{otherwise;} \end{aligned} \tag{1.22}$$

$$r_i(\theta_i, a_i) = EL_i(\theta_i, a_i) = P[\Pi_i \text{ is misclassified} \mid \theta_i, a_i] \tag{1.23}$$

for $i = 1, 2, \cdots, k$ (note that $L_i(\boldsymbol{\theta}, \mathbf{a})$ depends on $\boldsymbol{\theta}$ and $\mathbf{a}$ only through θ_i and a_i), then

$$L(\boldsymbol{\theta}, \mathbf{a}) = \textstyle\sum_{i=1}^{k} L_i(\theta_i, a_i) \tag{1.24}$$

and the corresponding risk function

$$r(\boldsymbol{\theta}, \mathbf{a}) = EL(\boldsymbol{\theta}, \mathbf{a}) = \textstyle\sum_{i=1}^{k} r_i(\theta_i, a_i) \tag{1.25}$$

is the expected misclassification size. For Φ defined in (0.5) and b defined in (1.9) with $P = P^*$, we note the obvious

LEMMA 1.3. *Under the procedure* R_1,

(1) $r(\boldsymbol{\theta}, R_1) \leqq k[1 - \Phi(b)]$ *for every* $\boldsymbol{\theta}$;

(2) *the equality holds when* θ_i *is either* δ_1^* *or* δ_2^* *for every* i.

Now let $\rho_i = \rho_i(\theta_i)$ be an a priori distribution of θ_i ($i = 1, 2, \cdots, k$), and let $\rho = \rho(\boldsymbol{\theta})$ be the product probability measure defined on the product parameter space ψ. We first observe the following lemma noted in [8]:

LEMMA 1.4. *For* $i = 1, 2, \cdots, k$, *let* a_i *be a Bayes decision rule in* D_i *for the* ith *component decision problem wrt* ρ_i *and loss function* $L_i(\theta_i, a_i)$. *If the loss function for the product decision problem has the form*

$$L(\boldsymbol{\theta}, \mathbf{a}) = \sum_{i=1}^k c_i L_i(\theta_i, a_i), \qquad c_i > 0 \tag{1.26}$$

then the product decision rule $\mathbf{a} = (a_1, a_2, \cdots, a_k)$ *is a Bayes decision rule in* D *wrt* ρ *for the product decision problem.*

PROOF. Let $\mathbf{a}' = (a_1', a_2', \cdots, a_k')$ be any other decision rule in D, then

$$\int r_i(\theta_i, a_i')\rho_i(\theta_i)\, d\theta_i \geqq \int r_i(\theta_i, a_i)\rho_i(\theta_i)\, d\theta_i \qquad \text{for} \quad i = 1, 2, \cdots, k;$$

hence the corresponding Bayes risks of $\mathbf{a}'$ and $\mathbf{a}$ satisfy

$$\begin{aligned} B(\rho, \mathbf{a}') &= \sum_{i=1}^k c_i \int r_i(\theta_i, a_i')\rho_i(\theta_i)\, d\theta_i \\ &\geqq \sum_{i=1}^k c_i \int r_i(\theta_i, a_i)\rho_i(\theta_i)\, d\theta_i = B(\boldsymbol{\theta}, \mathbf{a}). \end{aligned}$$

REMARK. It is clear that if a_i is the unique Bayes decision rule wrt ρ_i in D_i for every $i = 1, 2, \cdots, k$, then a is the unique Bayes decision rule wrt ρ in D.

Now for $i = 1, 2, \cdots, k$, let the a priori distribution of ρ_i be

$$\begin{aligned} \rho_i(\theta_i) &= \tfrac{1}{2} && \text{for} \quad \theta_i = \delta_1^* \quad \text{or} \quad \delta_2^* \\ &= 0 && \text{otherwise} \end{aligned} \tag{1.27}$$

and

$$\begin{aligned} \rho(\boldsymbol{\theta}) &= 2^{-k} && \text{if} \quad \theta_i \quad \text{is either } \delta_1^* \quad \text{or} \quad \delta_2^*, \quad i = 1, 2, \cdots, k \\ &= 0 && \text{otherwise.} \end{aligned} \tag{1.28}$$

We have

THEOREM 1.3. *The decision rule* $\mathbf{a}$ *specified in* (1.3) *is* (*a*) *minimax,* (*b*) *admissible and* (*c*) *unique Bayes wrt* ρ *among the class of* G *invariant decision rules.*

PROOF. The assertion that $\mathbf{a}$ is minimax among the class of G translation invariant rules can be proved by a lemma of Lehmann [11: p. 4–19], and the conditions of that lemma are justified by Lemma 1.3. Since the admissibility follows from the fact that a unique Bayes rule is admissible, it suffices to show that the rule $\mathbf{a}$ specified in (1.3) is unique Bayes wrt ρ among the class of G translation invariant rules. In view of Lemma 1.4 and the fact that any G translation invariant rule is a function of $\mathbf{Z}$, it in turn suffices to show that the ith component decision rule

$$\begin{aligned} a_i(Z_i) &= 0 && \text{if} \quad Z_i = (\bar{X}_i - \bar{X}_0) < d = \tfrac{1}{2}(\delta_1^* + \delta_2^*) \\ &= 1 && \text{if} \quad Z_i > d \end{aligned} \tag{1.29}$$

(here $\Pi_i \varepsilon S_B$ iff $a_i = 1$) is the unique Bayes decision rule wrt ρ_i among the class of all decision rules based on Z_i with the loss function $L_i(\theta_i, a_i)$ given in (1.22).

We first observe a well-known result in testing hypothesis problems: if Z_i is any random variable with density $f(z)$; under the hypotheses $H_1:f(z) = f_1(z)$ and $H_2:f(z) = f_2(z)$, H_2 is accepted iff $Z_i \varepsilon W$ for some region W. Then the sum of the two types of errors $(\alpha + \beta)$ is minimized iff W is taken to be (except on a set of probability measure zero)

$$W_0 = \{z:f_2(z)/f_1(z) > 1\}. \tag{1.30}$$

If Z_i has a normal distribution with mean θ_i and variance σ_0^2, and the corresponding hypotheses are $H_1:\theta_i = \delta_1^*$, $H_2:\theta_i = \delta_2^*$, then W_0 in (1.30) reduces to

$$W_0 = \{z:z > d). \tag{1.31}$$

which does not depend on σ_0^2.

Now let $a_i'(Z_i)$ be any decision rule about θ_i based on Z_i, i.e., we put Π_i into S_G iff $Z_i \varepsilon W$ for some region W specified by a_i'. For the a priori distribution $\rho_i(\theta_i)$ defined in (1.27), its corresponding Bayes risk is

$$B(\theta_i, a_i') = \tfrac{1}{2}\{P[Z_i \varepsilon W \mid \theta_i = \delta_1^*] + P[Z_i \not\varepsilon W \mid \theta_i = \delta_2^*]\}. \tag{1.32}$$

The infimum on the rhs of (1.32) is achieved iff $W = W_0$ given in (1.31). This implies that the decision rule a_i defined in (1.29) is the unique Bayes rule wrt ρ_i among the class of decision rules based on Z_i. This completes the proof of the theorem.

2. A two-stage procedure. In this section a two-stage procedure is given for the problem when σ^2 is unknown. It is specified in the following

PROCEDURE R_2. (1) Let $n_0 \geqq 2$ be a preassigned positive integer. We observe the sequence defined in (1.1) for $j = 1, 2, \cdots, n_0$. Compute

$$S^2 = \nu^{-1} \textstyle\sum_{i=0}^{k} \sum_{j=1}^{n_0} [X_{ij} - n_0^{-1} (\sum_{j=1}^{n_0} X_{ij})]^2 \tag{2.1}$$

with $\nu = (k + 1)(n_0 - 1)$.

(2) Observe the sequence defined in (1.1) for $j = (n_0 + 1), (n_0 + 2), \cdots, N$ where N is to be determined below.

(3) Compute the $(k + 1)$ overall sample means

$$\bar{X}_i = N^{-1} \textstyle\sum_{j=1}^{N} X_{ij} \qquad \text{for} \quad i = 0, 1, \cdots, k \tag{2.2}$$

and apply the decision rule defined in (1.3).

To determine the sample size N in the above procedure, we first observe that the LF configuration $\mathbf{\mu}^0$ given in (1.7') does not depend on σ^2 and N. Hence to satisfy the probability requirement (0.4) we can again restrict our attention to $\mathbf{\mu}^0$. We first note that

$$\begin{aligned} P[\mathrm{CD} \mid \mathbf{\mu}^0, \sigma^2; R_2] &= P[\bar{X}_i - \bar{X}_0 < d, \bar{X}_j - \bar{X}_0 > d(1 \leqq i \leqq m, m < j \leqq k) \mid \mathbf{\mu}^0, \sigma^2] \\ &= P[Y_i/U_\nu \leqq (N/2)^{\frac{1}{2}}a/S_\nu \ (i = 1, 2, \cdots, k)] \\ &= P[t_i \leqq (N/2)^{\frac{1}{2}}\, a/S_\nu \ (i = 1, 2, \cdots, k)] \end{aligned}$$

where $(Y_1, Y_2, \cdots, Y_k)$ follows a multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$ given in (1.8), νU_ν^2 follows a chi-square distribution with ν degrees of freedom and U_ν^2 is independent of $(Y_1, Y_2, \cdots, Y_k)$; $t_i = Y_i/U_\nu$ $(i = 1, 2, \cdots, k)$ are Student's t variables with ν degrees of freedom each, and they are correlated with correlation matrix $\boldsymbol{\Sigma}$. It can be seen from [4] that the joint distribution of $(t_1, t_2, \cdots, t_k)$ follows a multivariate t distribution with joint density function

$$(2.3)\quad f_{k,\nu,\boldsymbol{\Sigma}}(t_1, t_2, \cdots, t_k) = \Gamma(\tfrac{1}{2}(k+\nu))[(\nu\pi)^{k/2}\,|\boldsymbol{\Sigma}|^{\frac{1}{2}}\Gamma(\tfrac{1}{2}\nu)]^{-1}\,[1+\nu^{-1}\mathbf{t}'\boldsymbol{\Sigma}^{-1}\mathbf{t}]^{-\frac{1}{2}(k+\nu)}$$

for $t_i \,\varepsilon\, (-\infty, \infty)$, $i = 1, 2, \cdots, k$.

Let h_ν be the solution of the equation

$$(2.4)\qquad P = \int_{-\infty}^{h_\nu}\int_{-\infty}^{h_\nu}\cdots\int_{-\infty}^{h_\nu} f_{k,\nu,\boldsymbol{\Sigma}}(t_1, t_2, \cdots, t_k)\prod_{i=1}^{k} dt_i\,.$$

Then the sample size N can be determined in the following

THEOREM 2.1. *If N is the smallest integer satisfying*

$$(2.5)\qquad N \geqq \max\{n_0, 2h_\nu^2 S_\nu^2/a^2\},$$

where h_ν is the solution of (2.4) *with $P = P^*$ and a is defined in* (0.8), *then the probability requirement* (0.4) *is satisfied.*

PROOF.

$$P[\mathrm{CD} \mid \boldsymbol{\mu}, \sigma^2; R_2] \geqq P[\mathrm{CD} \mid \boldsymbol{\mu}^0, \sigma^2; R_2] = P[t_i \leqq (N/2)^{\frac{1}{2}}a/S_\nu\ (i = 1, 2, \cdots, k)] \geqq P^*$$

where the last inequality follows from the fact that for any observed S_ν in the first stage, we have $(N/2)^{\frac{1}{2}}a/S_\nu \geqq h_\nu$ from (2.5).

The values of h_ν have been computed and tabulated in Table 2 for $P = 0.50$, 0.75, 0.90, 0.95, 0.975 and 0.99; $k = 2(1)6(2)12(4)20$; and $\nu = 5(1)10(2)20(4)$-60(30)120. For every fixed P and k, h_ν converges to the corresponding b value given in Table 1 when ν is large, and those b values are repeated there under $\nu = \infty$. Table 2 is obtained by a double summation based on Gaussian quadrature formula given in [17]. To be conservative, the entries have all been rounded to the next higher value (in the 5th decimal) and should be in error by at most one unit in the last digit given.

2.1 *The expected sample size and relative efficiency.* Let N be the random sample size defined in (2.5). It is easily seen that

$$(2.6)\qquad \begin{aligned} P[N = n] &= 0 && \text{for } n < n_0,\\ &= P[2h_\nu^2 S_\nu^2/a^2 \leqq n_0] && \text{for } n = n_0,\\ &= P[n-1 < 2h_\nu^2 S_\nu^2/a^2 \leqq n] && \text{for } n \geqq n_0 + 1. \end{aligned}$$

Let

$$(2.7)\qquad \theta = (2\lambda^2 h_\nu^2)^{-1}.$$

TABLE 2

Equi-coordinate percentage points h of a multivariate t distribution with correlation matrix Σ *for*
$P = 0.50, 0.75, 0.90, 0.95, 0.975, 0.99$;
$k = 2(1)6(2)12(4)20$; *and*
$\nu = 5(1)10(2)20(4)60(30)120$.

ν	k = 2	3	4	5	6	8	10	12	16	20
					$P = 0.50$					
5	.68057	.89079	1.06270	1.16737	1.26195	1.39388	1.49125	1.56779	1.68341	1.76905
6	.67397	.88145	1.05075	1.15380	1.24685	1.37666	1.47249	1.54784	1.66173	1.74613
7	.66931	.87488	1.04234	1.14425	1.23620	1.36449	1.45921	1.53371	1.64633	1.72982
8	.66584	.87000	1.03610	1.13716	1.22829	1.35544	1.44931	1.52315	1.63480	1.71760
9	.66317	.86623	1.03129	1.13168	1.22218	1.34843	1.44163	1.51495	1.62583	1.70807
10	.66104	.86324	1.02746	1.12733	1.21731	1.34284	1.43551	1.50840	1.61865	1.70043
12	.65786	.85878	1.02176	1.12084	1.21005	1.33449	1.42633	1.49857	1.60784	1.68892
14	.65561	.85562	1.01771	1.11623	1.20489	1.32854	1.41978	1.49154	1.60009	1.68064
16	.65393	.85326	1.01470	1.11279	1.20103	1.32408	1.41486	1.48626	1.59425	1.67438
18	.65263	.85143	1.01236	1.11012	1.19804	1.32061	1.41104	1.48214	1.58969	1.66949
20	.65159	.84998	1.01049	1.10799	1.19565	1.31784	1.40797	1.47884	1.58602	1.66555
24	.65003	.84780	1.00770	1.10480	1.19207	1.31369	1.40337	1.47388	1.58050	1.65961
28	.64892	.84625	1.00571	1.10253	1.18952	1.31072	1.40008	1.47032	1.57653	1.65532
32	.64810	.84509	1.00423	1.10083	1.18760	1.30850	1.39761	1.46765	1.57354	1.65209
36	.64745	.84419	1.00307	1.09951	1.18612	1.30677	1.39568	1.46557	1.57120	1.64956
40	.64694	.84347	1.00215	1.09846	1.18493	1.30538	1.39414	1.46389	1.56933	1.64753
44	.64652	.84288	1.00140	1.09760	1.18396	1.30424	1.39288	1.46253	1.56779	1.64586
48	.64617	.84239	1.00077	1.09688	1.18315	1.30330	1.39182	1.46138	1.56651	1.64447
52	.64588	.84198	1.00024	1.09627	1.18246	1.30250	1.39093	1.46041	1.56542	1.64328
56	.64562	.84163	0.99979	1.09575	1.18188	1.30181	1.39017	1.45958	1.56448	1.64227
60	.64540	.84132	0.99939	1.09530	1.18137	1.30122	1.38950	1.45886	1.56367	1.64138
90	.64438	.83989	0.99756	1.09320	1.17900	1.29844	1.38640	1.45549	1.55986	1.63723
120	.64387	.83918	0.99664	1.09215	1.17781	1.29705	1.38485	1.45380	1.55795	1.63515
∞	.64235	.83705	0.99390	1.08901	1.17426	1.29288	1.38017	1.44870	1.55216	1.62881
					$P = 0.75$					
5	1.28547	1.49920	1.67149	1.78273	1.88216	2.02501	2.13209	2.21722	2.34734	2.44484
6	1.26035	1.46657	1.63178	1.73848	1.83356	1.97019	2.07255	2.15393	2.27832	2.37156
7	1.24291	1.44395	1.60427	1.70782	1.79988	1.93216	2.03122	2.10994	2.23027	2.32047
8	1.23009	1.42735	1.58409	1.68534	1.77517	1.90423	2.00083	2.07757	2.19487	2.28279
9	1.22027	1.41465	1.56866	1.66814	1.75626	1.88285	1.97754	2.05276	2.16769	2.25384
10	1.21251	1.40463	1.55647	1.65455	1.74132	1.86594	1.95913	2.03312	2.14617	2.23088
12	1.20102	1.38980	1.53845	1.63447	1.71923	1.84093	1.93185	2.00402	2.11422	2.19678
14	1.19293	1.37936	1.52577	1.62033	1.70367	1.82330	1.91262	1.98348	2.09165	2.17266
16	1.18693	1.37161	1.51636	1.60984	1.69212	1.81020	1.89832	1.96821	2.07484	2.15469
18	1.18229	1.36564	1.50910	1.60174	1.68321	1.80010	1.88728	1.95640	2.06185	2.14078
20	1.17860	1.36089	1.50333	1.59531	1.67612	1.79205	1.87850	1.94701	2.05150	2.12969
24	1.17311	1.35381	1.49473	1.58572	1.66556	1.78007	1.86540	1.93300	2.03605	2.11313
28	1.16921	1.34879	1.48864	1.57892	1.65807	1.77157	1.85610	1.92305	2.02507	2.10136
32	1.16631	1.34505	1.48409	1.57385	1.65248	1.76522	1.84915	1.91561	2.01687	2.09256
36	1.16405	1.34215	1.48056	1.56992	1.64815	1.76030	1.84377	1.90985	2.01050	2.08573

TABLE 2 (*continued*)

ν	k = 2	3	4	5	6	8	10	12	16	20
					P = 0.75 (*continued*)					
40	1.16226	1.33983	1.47775	1.56678	1.64470	1.75637	1.83948	1.90525	2.00543	2.08028
44	1.16079	1.33795	1.47546	1.56422	1.64188	1.75317	1.83597	1.90150	2.00128	2.07583
48	1.15957	1.33638	1.47355	1.56210	1.63954	1.75051	1.83305	1.89838	1.99783	2.07212
52	1.15854	1.33505	1.47194	1.56030	1.63755	1.74825	1.83059	1.89574	1.99491	2.06899
56	1.15766	1.33392	1.47056	1.55876	1.63586	1.74633	1.82848	1.89348	1.99241	2.06631
60	1.15689	1.33293	1.46937	1.55743	1.63439	1.74466	1.82666	1.89152	1.99025	2.06399
90	1.15334	1.32837	1.46382	1.55124	1.62757	1.73690	1.81816	1.88242	1.98019	2.05318
120	1.15158	1.32609	1.46106	1.54815	1.62417	1.73304	1.81393	1.87789	1.97518	2.04779
∞	1.14630	1.31930	1.45281	1.53895	1.61402	1.72150	1.80130	1.86435	1.96020	2.03170
					P = 0.90					
5	2.00771	2.24189	2.42873	2.55442	2.66588	2.82937	2.95328	3.05255	3.20554	3.32111
6	1.93795	2.15646	2.32910	2.44550	2.54818	2.69895	2.81313	2.90461	3.04560	3.15216
7	1.89059	2.09861	2.26174	2.37189	2.46867	2.61082	2.71840	2.80455	2.93732	3.03766
8	1.85635	2.05686	2.21320	2.31886	2.41140	2.54735	2.65015	2.73244	2.85922	2.95502
9	1.83046	2.02534	2.17657	2.27887	2.36821	2.49948	2.59867	2.67803	2.80026	2.89261
10	1.81019	2.00070	2.14796	2.24764	2.33449	2.46210	2.55846	2.63554	2.75419	2.84381
12	1.78053	1.96468	2.10617	2.20203	2.28526	2.40754	2.49977	2.57348	2.68688	2.77248
14	1.75987	1.93963	2.07713	2.17035	2.25106	2.36964	2.45899	2.53036	2.64009	2.72288
16	1.74465	1.92120	2.05578	2.14707	2.22594	2.34180	2.42904	2.49868	2.60570	2.68640
18	1.73298	1.90707	2.03943	2.12924	2.20670	2.32048	2.40610	2.47442	2.57936	2.65846
20	1.72375	1.89590	2.02650	2.11514	2.19149	2.30364	2.38798	2.45525	2.55854	2.63637
24	1.71006	1.87936	2.00737	2.09429	2.16900	2.27872	2.36117	2.42689	2.52775	2.60369
28	1.70041	1.86770	1.99389	2.07961	2.15316	2.26118	2.34229	2.40693	2.50606	2.58068
32	1.69324	1.85904	1.98388	2.06870	2.14141	2.24816	2.32828	2.39211	2.48997	2.56360
36	1.68770	1.85236	1.97616	2.06029	2.13233	2.23811	2.31748	2.38068	2.47756	2.55042
40	1.68329	1.84704	1.97002	2.05360	2.12512	2.23013	2.30889	2.37160	2.46769	2.53995
44	1.67970	1.84271	1.96502	2.04815	2.11925	2.22363	2.30190	2.36420	2.45966	2.53142
48	1.67672	1.83911	1.96087	2.04364	2.11438	2.21823	2.29610	2.35807	2.45299	2.52435
52	1.67420	1.83608	1.95737	2.03982	2.11027	2.21369	2.29121	2.35289	2.44737	2.51838
56	1.67205	1.83349	1.95438	2.03657	2.10676	2.20980	2.28703	2.34847	2.44257	2.51329
60	1.67019	1.83125	1.95179	2.03375	2.10373	2.20644	2.28342	2.34465	2.43842	2.50888
90	1.66157	1.82086	1.93980	2.02070	2.08967	2.19088	2.26668	2.32695	2.41920	2.48847
120	1.65729	1.81571	1.93386	2.01423	2.08270	2.18316	2.25838	2.31818	2.40967	2.47836
∞	1.64457	1.80040	1.91622	1.99504	2.06202	2.16029	2.23378	2.29217	2.38142	2.44838
					P = 0.95					
5	2.56513	2.82213	3.02640	3.16640	3.29014	3.47342	3.61302	3.72530	3.89898	4.03064
6	2.44344	2.67692	2.86038	2.98658	3.09743	3.26186	3.38705	3.48772	3.64352	3.76174
7	2.36222	2.58029	2.75012	2.86724	2.96961	3.12161	3.23724	3.33021	3.47406	3.58324
8	2.30424	2.51145	2.67169	2.78241	2.87879	3.02198	3.13082	3.21829	3.35361	3.45629
9	2.26079	2.45996	2.61310	2.71907	2.81101	2.94764	3.05142	3.13479	3.26370	3.36150
10	2.22705	2.42002	2.56770	2.67001	2.75851	2.89008	2.98995	3.07013	3.19408	3.28809
12	2.17806	2.36212	2.50195	2.59900	2.68257	2.80684	2.90105	2.97663	3.09338	3.18186

TABLE 2 (*continued*)

ν	k = 2	3	4	5	6	8	10	12	16	20
					P = 0.95 (*continued*)					
14	2.14423	2.32220	2.45668	2.55011	2.63031	2.74958	2.83991	2.91232	3.02410	3.10878
16	2.11947	2.29301	2.42361	2.51443	2.59218	2.70781	2.79531	2.86541	2.97357	3.05546
18	2.10056	2.27075	2.39840	2.48723	2.56313	2.67600	2.76134	2.82970	2.93509	3.01485
20	2.08566	2.25322	2.37856	2.46583	2.54027	2.65097	2.73463	2.80160	2.90483	2.98291
24	2.06367	2.22736	2.34932	2.43430	2.50660	2.61412	2.69529	2.76024	2.86027	2.93589
28	2.04822	2.20921	2.32881	2.41220	2.48300	2.58829	2.66774	2.73126	2.82906	2.90295
32	2.03678	2.19577	2.31363	2.39584	2.46554	2.56919	2.64736	2.70984	2.80598	2.87860
36	2.02796	2.18542	2.30195	2.38325	2.45211	2.55450	2.63168	2.69335	2.78823	2.85986
40	2.02095	2.17721	2.29267	2.37326	2.44145	2.54284	2.61924	2.68028	2.77415	2.84500
44	2.01526	2.17053	2.28514	2.36514	2.43278	2.53337	2.60914	2.66966	2.76271	2.83293
48	2.01053	2.16499	2.27889	2.35841	2.42560	2.52552	2.60077	2.66085	2.75323	2.82292
52	2.00655	2.16032	2.27362	2.35274	2.41956	2.51891	2.59371	2.65344	2.74525	2.81450
56	2.00315	2.15634	2.26913	2.34790	2.41439	2.51326	2.58769	2.64711	2.73843	2.80731
60	2.00021	2.15289	2.26525	2.34372	2.40993	2.50839	2.58250	2.64165	2.73255	2.80110
90	1.98661	2.13696	2.24729	2.32438	2.38931	2.48585	2.55846	2.61638	2.70534	2.77239
120	1.97987	2.12908	2.23840	2.31482	2.37912	2.47470	2.54658	2.60390	2.69190	2.75820
∞	1.95993	2.10574	2.21213	2.28654	2.34897	2.44177	2.51147	2.56700	2.65218	2.71629
					P = 0.975					
5	3.15876	3.44437	3.67090	3.82816	3.96687	4.17376	4.33198	4.45962	4.65777	4.80857
6	2.96601	3.21856	3.41633	3.55432	3.67516	3.85575	3.99376	4.10505	4.27776	4.40917
7	2.83955	3.07086	3.25020	3.37579	3.48516	3.64884	3.77385	3.87465	4.03111	4.15022
8	2.75037	2.96695	3.13355	3.25053	3.35193	3.50382	3.61976	3.71320	3.85821	3.96860
9	2.68419	2.88996	3.04725	3.15792	3.25348	3.39672	3.50597	3.59398	3.73053	3.83445
10	3.63315	2.83069	2.98088	3.08673	3.17784	3.31446	3.41859	3.50245	3.63250	3.73145
12	2.55965	2.74547	2.88560	2.98458	3.06935	3.19655	3.29337	3.37128	3.49201	3.58382
14	2.50931	2.68721	2.82053	2.91488	2.99537	3.11618	3.20804	3.28191	3.39631	3.48325
16	2.47269	2.64487	2.77332	2.86432	2.94172	3.05793	3.14622	3.21717	3.32699	3.41040
18	2.44487	2.61273	2.73750	2.82598	2.90106	3.01379	3.09938	3.16813	3.27448	3.35522
20	2.42301	2.58751	2.70941	2.79591	2.86918	2.97921	3.06268	3.12971	3.23335	3.31200
24	2.39088	2.55046	2.66818	2.75181	2.82244	2.92850	3.00889	3.07340	3.17307	3.24866
28	2.36840	2.52457	2.63939	2.72102	2.78981	2.89312	2.97137	3.03413	3.13104	3.20450
32	2.35179	2.50545	2.61814	2.69831	2.76575	2.86704	2.94372	3.00519	3.10007	3.17196
36	2.33903	2.49076	2.60183	2.68087	2.74728	2.84702	2.92249	2.98297	3.07630	3.14699
40	2.32891	2.47912	2.58890	2.66705	2.73265	2.83117	2.90569	2.96539	3.05749	3.12722
44	2.32069	2.46967	2.57841	2.65584	2.72078	2.81831	2.89206	2.95112	3.04223	3.11119
48	2.31388	2.46184	2.56972	2.64656	2.71096	2.80766	2.88077	2.93932	3.02960	3.09792
52	2.30815	2.45525	2.56241	2.63875	2.70269	2.79871	2.87128	2.92939	3.01897	3.08676
56	2.30326	2.44963	2.55618	2.63209	2.69564	2.79107	2.86319	2.92092	3.00991	3.07724
60	2.29903	2.44478	2.55079	2.62634	2.68955	2.78448	2.85620	2.91361	3.00209	3.06903
90	2.27951	2.42236	2.52594	2.59979	2.66146	2.75406	2.82396	2.87988	2.96602	3.03114
120	2.26987	2.41129	2.51367	2.58669	2.64760	2.73905	2.80807	2.86326	2.94824	3.01246
∞	2.24140	2.37864	2.47751	2.54808	2.60676	2.69486	2.76125	2.81429	2.89587	2.95746

TABLE 2 (*concluded*)

ν	k									
	2	3	4	5	6	8	10	12	16	20
					$P = 0.99$					
5	4.02790	4.36051	4.62385	4.80875	4.97154	5.21579	5.40324	5.55485	5.79095	5.97127
6	3.70523	3.98906	4.21075	4.36746	4.50437	4.71040	4.86846	4.99628	5.19528	5.34717
7	3.49824	3.75150	3.94718	4.08617	4.20682	4.38875	4.52823	4.64100	4.81657	4.95059
8	3.35462	3.58711	3.76519	3.89213	4.00177	4.16733	4.29419	4.39675	4.55641	4.67835
9	3.24933	3.46683	3.63226	3.75052	3.85222	4.00594	4.12366	4.21879	4.36687	4.47994
10	3.16892	3.37513	3.53105	3.64276	3.73848	3.88327	3.99408	4.08360	4.22289	4.32923
12	3.05435	3.24470	3.38732	3.48983	3.57717	3.70941	3.81051	3.89211	4.01901	4.11584
14	2.97673	3.15650	3.29029	3.38666	3.46843	3.59230	3.68690	3.76322	3.88182	3.97227
16	2.92071	3.09293	3.22044	3.31244	3.39023	3.50814	3.59811	3.67065	3.78331	3.86919
18	2.87839	3.04496	3.16779	3.25651	3.33134	3.44478	3.53128	3.60099	3.70920	3.79165
20	2.84531	3.00748	3.12669	3.21287	3.28540	3.39538	3.47919	3.54669	3.65144	3.73122
24	2.79692	2.95274	3.06669	3.14918	3.21839	3.32335	3.40325	3.46756	3.56729	3.64320
28	2.76325	2.91467	3.02501	3.10497	3.17188	3.27338	3.35058	3.41270	3.50896	3.58219
32	2.73848	2.88669	2.99439	3.07248	3.13773	3.23669	3.31193	3.37243	3.46616	3.53742
36	2.71948	2.86524	2.97093	3.04761	3.11158	3.20862	3.28235	3.34162	3.43341	3.50318
40	2.70446	2.84829	2.95240	3.02796	3.09093	3.18644	3.25899	3.31730	3.40756	3.47614
44	2.69228	2.83455	2.93738	3.01204	3.07420	3.16848	3.24008	3.29760	3.38663	3.45425
48	2.68221	2.82318	2.92497	2.99888	3.06037	3.15364	3.22445	3.28133	3.36933	3.43617
52	2.67374	2.81363	2.91453	2.98783	3.04875	3.14118	3.21132	3.26766	3.35481	3.42099
56	2.66651	2.80549	2.90564	2.97841	3.03886	3.13056	3.20014	3.25601	3.34244	3.40805
60	2.66029	2.79847	2.89798	2.97028	3.03032	3.12140	3.19049	3.24597	3.33177	3.39690
90	2.63157	2.76612	2.86267	2.93287	2.99103	3.07925	3.14612	3.19977	3.28269	3.34559
120	2.61743	2.75019	2.84529	2.91447	2.97171	3.05853	3.12430	3.17706	3.25857	3.32038
∞	2.57583	2.70340	2.79428	2.86045	2.91500	2.99774	3.06033	3.11048	3.18787	3.24648

Since $\nu S_\nu^2/\sigma^2$ is a chi-square variable with ν degrees of freedom, (2.6) can be rewritten as

$$
\begin{aligned}
P[N = n] &= 0 && \text{for } n < n_0, \\
&= P[\chi_\nu^2 \leqq \theta\nu n_0] && \text{for } n = n_0, \\
&= P[\theta\nu(n-1) < \chi_\nu^2 \leqq \theta\nu n] && \text{for } n \geqq n_0 + 1.
\end{aligned} \tag{2.8}
$$

Hence

$$
\begin{aligned}
EN &= n_0 P[N = n_0] + \textstyle\sum_{n=n_0+1}^{\infty} nP[N = n] \\
&= n_0 P[\chi_\nu^2 \leqq \theta\nu n_0] + \textstyle\sum_{n=n_0+1}^{\infty} n[2^{\nu/2}\Gamma(\nu/2)]^{-1} \int_{\theta\nu(n-1)}^{\theta\nu n} y^{\nu/2-1} e^{-y/2}\, dy.
\end{aligned} \tag{2.9}
$$

Consider any fixed summand in the second term on the rhs of (2.9). Since for $\theta\nu(n-1) \leqq y \leqq \theta\nu n$ n satisfies

$$y/\theta\nu \leqq n \leqq y/\theta\nu + 1, \tag{2.10}$$

using the first inequality in (2.10), the second term Q on the rhs of (2.9) can be bounded by

$$(2.11)\qquad \begin{aligned} Q &\geqq \textstyle\sum_{n=n_0+1}^{\infty} [2^{\nu/2}\Gamma(\nu/2)]^{-1} \int_{\theta\nu(n-1)}^{\theta\nu n} (\theta\nu)^{-1} y^{\nu/2} e^{-y/2}\, dy \\ &= \theta^{-1} P[\chi^2_{\nu+2} > \theta\nu n_0]; \end{aligned}$$

similarly, by the second inequality in (2.10) Q is upper bounded by

$$(2.12)\qquad Q \leqq \theta^{-1} P[\chi^2_{\nu+2} > \theta\nu n_0] + P[\chi_\nu^2 > \theta\nu n_0].$$

Combining (2.9), (2.11) and (2.12), we have

$$(2.13)\quad EN = n_0 P[\chi_\nu^2 \leqq \theta\nu n_0] + \theta^{-1} P[\chi^2_{\nu+2} > \theta\nu n_0] + rP[\chi_\nu^2 > \theta\nu n_0]$$

for some $r \varepsilon (0, 1)$.

The expected sample size EN given in (2.13) is a function of P^*, k, n_0 and λ, and it depends on P^* only through h_ν.

LEMMA 2.1. *For every P^*, k and first-stage sample size n_0,*

$$(2.14)\qquad EN \geqq 2\lambda^2 h_\nu^2 \qquad \textit{for every } \lambda,$$

$$(2.15)\qquad \lim_{\lambda\to\infty} EN/(2\lambda^2 h_\nu^2) = 1.$$

PROOF. By (2.5),

$$EN = E \max \{n_0, 2h_\nu^2 S_\nu^2/a^2\} \geqq E\, 2h_\nu^2 S_\nu^2/a^2 = 2\lambda^2 h_\nu^2,$$

this proves (2.14). The proof of (2.15) follows from (2.13).

Let N_0 be the sample size required under the single-stage procedure for the LF configuration, the following theorem investigates the asymptotic relative efficiency of the two-stage procedure wrt the single-stage procedure.

THEOREM 2.2. *For every P^*, k and first-stage sample size n_0.*

$$(2.16)\qquad \frac{EN}{N_0} \geqq \left(\frac{h_\nu}{b}\right)^2 \qquad \textit{for every } \lambda,$$

$$(2.17)\qquad \lim_{\lambda\to\infty} \frac{EN}{N_0} = \left(\frac{h_\nu}{b}\right)^2.$$

PROOF. The proof of this theorem follows from (1.10) and Lemma 2.1, if we disregard the fact that N_0 has to be an integer.

2.2 *The expected misclassification size.* Let EM_0 denote the expected misclassification size under the procedure R_2 when $\boldsymbol{\mu} = \boldsymbol{\mu}^0$, then it is easily seen that EM_0 is the supremum of the expected misclassification size over all mean vectors $\boldsymbol{\mu}$. Due to the difficulty that the sample size N is a chance variable, the exact mathematical expression for EM_0 can not be obtained. However, both lower and upper bounds on EM_0 can be derived and the asymptotic behavior of EM_0 (as $\lambda \to \infty$) can be examined based on those bounds.

LEMMA 2.2. *Let EN be the expected sample size under R_2, then*

$$(2.18)\qquad EM_0 \geqq k[1 - \Phi((\tfrac{1}{2}EN)^{\frac{1}{2}}/\lambda)] \qquad \textit{for every } \lambda.$$

PROOF. It is clear that

$$EM_0 = Ek[1 - \Phi((\tfrac{1}{2}N)^{\frac{1}{2}}/\lambda)]$$

where the expectation is taken over N space. Consider $g(N) = 1 - \Phi((\frac{1}{2}N)^{\frac{1}{2}}/\lambda)$ as a function of N. Since for $\lambda > 0$

$$(d^2/dN^2)g(N) = (8\lambda)^{-1}(\pi N)^{-\frac{1}{2}}[N^{-1} + (2\lambda^2)^{-1}] \exp(-N/(4\lambda^2)) > 0$$

$g(N)$ is a convex function of N. It follows from Jensen inequality that $Eg(N) \geqq g(EN)$, which completes the proof.

To establish an upper bound on EM_0, we need the following inequality which is given in [5: p. 166].

LEMMA 2.3. (Feller-Laplace). *For every* $z > 0$,

$$(2.19)\quad (2\pi)^{-\frac{1}{2}}(z^{-1} - z^{-3}) \exp(-z^2/2) < \int_z^\infty (2\pi)^{-\frac{1}{2}} \exp(-x^2/2)\, dx < (2\pi)^{-\frac{1}{2}} z^{-1} \exp(-z^2/2).$$

When z *is not too small, we can write the approximation*

$$(2.20)\quad \int_z^\infty (2\pi)^{-\frac{1}{2}} \exp(-x^2/2)\, dx \sim (2\pi)^{-\frac{1}{2}} z^{-1} \exp(-z^2/2).$$

LEMMA 2.4. *For every* $\lambda > 0$ *and the first-stage sample size* n_0.

$$(2.21)\quad EM_0 < k\{[1 - \Phi((\tfrac{1}{2}n_0)^{\frac{1}{2}}/\lambda)]P[\chi_\nu^2 \leqq \theta\nu n_0] + (\nu + h_\nu^2)[(\nu - 1)h_\nu]^{-1} f_\nu(h_\nu)P[\chi_{\nu-1}^2 > \theta\nu n_0(1 + h_\nu^2/\nu)]\}$$

where $f_\nu(\cdot)$ *is the density function of Student's t-distribution with* ν *degrees of freedom and* θ *is defined in* (2.7).

PROOF:

$$\begin{aligned} EM_0 &= k\{[1 - \Phi((\tfrac{1}{2}n_0)^{\frac{1}{2}}/\lambda)]P[N = n_0] \\ &\quad + \textstyle\sum_{n=n_0+1}^\infty [1 - \Phi((\tfrac{1}{2}n)^{\frac{1}{2}}/\lambda)]P[N = n]\} \\ &= k\{[1 - \Phi((\tfrac{1}{2}n_0)^{\frac{1}{2}}/\lambda)]P[\chi_\nu^2 \leqq \theta\nu n_0] \\ &\quad + \textstyle\sum_{n=n_0+1}^\infty [1 - \Phi((\tfrac{1}{2}n)^{\frac{1}{2}}/\lambda)]P[\theta\nu(n-1) < \chi_\nu^2 \leqq \theta\nu n]\} \\ &=_{\text{def}} k\{I_1 + I_2\}. \end{aligned}$$

To find an upper bound on I_2, using Lemma (2.3) and the first inequality in (2.10) we have

$$\begin{aligned} I_2 &\leqq \textstyle\sum_{n=n_0+1}^\infty 2^{\frac{1}{2}}\lambda(2\pi n)^{-\frac{1}{2}} \exp(-n/(4\lambda^2)) \int_{\theta\nu(n-1)}^{\theta\nu n} (2^{\nu/2}\Gamma(\nu/2))^{-1} y^{\nu/2-1} e^{-y/2}\, dy \\ &< \textstyle\sum_{n=n_0+1}^\infty (\nu/\pi)^{\frac{1}{2}}[2^{\nu/2+1}\Gamma(\nu/2)h_\nu]^{-1} \int_{\theta\nu(n-1)}^{\theta\nu n} y^{\frac{1}{2}(\nu-3)} \exp(-\tfrac{1}{2}(1 + h_\nu^2/\nu)y)\, dy \\ &= (\nu + h_\nu^2)((\nu - 1)h_\nu)^{-1} f_\nu(h_\nu)P[\chi_{\nu-1}^2 > \theta\nu n_0(1 + h_\nu^2/\nu)], \end{aligned}$$

which yields the desired result.

THEOREM 2.3. *For every P^*, k and first-stage sample size n_0,*

$$k[1 - \Phi(h_\nu)] \leqq \lim_{\lambda\to\infty} EM_0 \leqq k(\nu + h_\nu^2)(\nu - 1)h_\nu)^{-1}f_\nu(h_\nu) \tag{2.22}$$

where $f_\nu(\cdot)$ is the density function of Student's t-distribution with ν degrees of freedom.

PROOF. The lower bound follows from Lemma 2.2 and (2.15), the upper bound follows from Lemma 2.4.

When ν is fairly large, the ratio $(\nu + h_\nu^2)/(\nu - 1)$ is approximately 1, $f_\nu(h_\nu)$ is approximately $\varphi(h_\nu)$ where $\varphi(\cdot)$ is the standard normal density function. Applying (2.20), the upper bound in (2.22) is approximately $k[1 - \Phi(h_\nu)]$. Hence $\lim_{\lambda\to\infty} EM_0$ is approximately $k[1 - \Phi(h_\nu)]$.

3. A sequential procedure. The relative efficiency of the two-stage procedure based on the idea of Stein [16] was investigated in Section 2. It can be seen that the relative efficiency, N_0/EN, is uniformly less than 1 (for all values of σ^2 and the first-stage sample size n_0), and it can be explained as (at least partly) due to the fact that the information of the observations in the second stage is not utilized in estimating the unknown parameter σ^2. This gives the idea of performing the experiment so that σ^2 can be estimated sequentially. A sequential procedure based on the idea of the random stopping rule developed by Chow and Robbins [3] is then considered in this section to serve as an alternative to the two-stage procedure. It should be pointed out that this sequential procedure provides only an "asymptotic" solution to our problem, and the PCD under this procedure may be slightly less than P^* for some values of the unknown parameter σ^2.

3.1 *The procedure and its asymptotic relative efficiency.*

PROCEDURE R_3.

(1) We observe the sequence $\mathbf{X}_j = \{X_{0j}, X_{1j}, \cdots, X_{kj}\}$ defined in (1.1), one vector at a time, stop with $\mathbf{X}_N$ where

$$N \text{ is the first integer } n \geqq 2 \text{ such that } S_\nu^2 \leqq na^2/(2h_\nu^2), \tag{3.1}$$

a is defined in (0.8), $\nu = (k + 1)(n - 1)$, h_ν satisfies (2.4) with $P = P^*$ and

$$S_\nu^2 = \nu^{-1} \textstyle\sum_{i=0}^{k} \sum_{j=1}^{n} [X_{ij} - n^{-1} (\sum_{j=1}^{n} X_{ij})]^2. \tag{3.2}$$

(2) Let the observed N value in (1) be n. Compute

$$\bar{X}_i = n^{-1} \textstyle\sum_{j=1}^{n} X_{ij} \quad \text{for} \quad i = 0, 1, \cdots, k \tag{3.3}$$

and apply the decision rule defined in (1.3).

LEMMA 3.1. *For every $\boldsymbol{\mu}$ and every σ^2,*

$$P[N < \infty \mid \boldsymbol{\mu}, \sigma^2; R_3] = 1. \tag{3.4}$$

PROOF. By the strong law of large numbers, $\lim_{\nu\to\infty} S_\nu^2 = \sigma^2$ a.s. Hence

$$P[N = \infty \mid \boldsymbol{\mu}, \sigma^2; R_3] = P\{\textstyle\bigcap_{n=2}^{\infty} [S_\nu^2/n > a^2/(2h_\nu^2)]\} = 0.$$

The following theorem states a relationship between the sample sizes required for the two-stage procedure and the sequential procedure. Let N_t and N_s denote the sample size required under R_2 and R_3, respectively, then

THEOREM 3.1. *For every first-stage sample size n_0 in R_2, we have*

$$[N_t = n_0] \subset [N_s \leqq N_t]. \tag{3.5}$$

PROOF. Let $\mathfrak{X} = \{\omega : \omega = (\mathbf{x}_1, \mathbf{x}_2, \cdots)\}$ be the sample space. Since for every $\omega \,\varepsilon\, \mathfrak{X}$ we have $N_t(\omega) \geqq n_0$, it suffices to show that $[N_t = n_0] \subset [N_s \leqq n_0]$.

Let $\{B_n\}$ and $\{C_n\}$ denote the terminal sets for R_2 and R_3, respectively; i.e.,

$$B_n = \{\omega : \omega \,\varepsilon\, \mathfrak{X}, N_t = n\} \quad \text{for} \quad n = n_0, n_0 + 1, \cdots,$$

$$C_n = \{\omega : \omega \,\varepsilon\, \mathfrak{X}, N_s = n\} \quad \text{for} \quad n = 2, 3 \cdots.$$

Then for $\nu = (k + 1)(n_0 - 1)$, it follows from (2.5) that

$$\omega \,\varepsilon\, B_{n_0} \Leftrightarrow S_\nu^2(\omega) \leqq n_0\, a^2/(2h_\nu^2).$$

this implies that either there exists an $n < n_0$ such that $\omega \,\varepsilon\, C_n$ or $\omega \,\varepsilon\, C_{n_0}$. Hence $\omega \,\varepsilon\, \bigcup_{n=2}^{n_0} C_n$ or equivalently, $\omega \,\varepsilon\, [N_s \leqq n_0]$.

COROLLARY. *For $\nu = (k + 1)(n_0 - 1)$,*

$$P[N_s \leqq N_t] \geqq P[N_t = n_0] = P[\chi_\nu^2 \leqq \nu n_0/(2\lambda^2 h_\nu^2)]. \tag{3.6}$$

In particular, $\lim_{\lambda\to 0} P[N_s \leqq N_t] = 1$ *for every $n_0 \geqq 2$.*

In the following we give the bounds on the cdf of the random sample size N under the procedure R_3. We first observe that for every $n \geqq 2$ and S_ν^2 given in (3.2), $\nu S_\nu^2/\sigma^2$ is distributed as $V_1 + V_2 + \cdots + V_{n-1}$ where the V's are independently identically distributed chi-square chance variables with $(k + 1)$ degrees of freedom each (in fact, the V's can be obtained by using Helmert transformation). Let the sequence of real numbers $\{q_j\}_1^\infty$ be such that

$$q_j = [(k + 1)/(2\lambda^2)][j(j + 1)/h^2_{(k+1)j} - (j - 1)j/h^2_{(k+1)(j-1)}] \quad \text{for} \quad j = 1, 2, \cdots \tag{3.7}$$

where $h_0 = 0$ and $h_{(k+1)j}$ satisfies (2.4) with $P = P^*$ for $j \geqq 1$, then

THEOREM 3.2. *For every fixed $n \geqq 2$,*

$$\chi_\nu^2(\nu n/(2\lambda^2 h_\nu^2)) \leqq P[N \leqq n] \leqq 1 - \textstyle\prod_{j=1}^{n-1} [1 - \chi^2_{(k+1)}(q_j)] \tag{3.8}$$

where $\nu = (k + 1)(n - 1)$.

PROOF. It follows from (3.1) that

$$\begin{aligned}[N > n] &= \textstyle\bigcap_{j=2}^{n} [(k + 1)(j - 1)\sigma^{-2} S^2_{(k+1)(j-1)} > (k + 1)(j - 1)j/(2\lambda^2 h_\nu^2)] \\ &= [V_1 > q_1, \textstyle\sum_{j=1}^{2} V_j > \sum_{j=1}^{2} q_j, \cdots, \sum_{j=1}^{n-1} V_j > \sum_{j=1}^{n-1} q_j].\end{aligned}$$

Since

$$\begin{aligned}&[\textstyle\sum_{j=1}^{n-1} V_j > \sum_{j=1}^{n-1} q_j] \\ &\quad\supset [V_1 > q_1, \textstyle\sum_{j=1}^{2} V_j > \sum_{j=1}^{2} q_j, \cdots, \sum_{j=1}^{n-1} V_j > \sum_{j=1}^{n-1} q_j] \\ &\quad\supset \textstyle\bigcap_{j=1}^{n-1} [V_j > q_j],\end{aligned}$$

it follows that

$$(3.9)\qquad 1 - \chi_\nu^2(\nu n/(2\lambda^2 h_\nu^2)) \geqq P[N > n] \geqq \textstyle\prod_{j=1}^{n-1} [1 - \chi^2_{(k+1)}(q_j)],$$

and the theorem is proved by taking complements.

COROLLARY. *For every fixed* $n \geqq 2$,

$$(3.10)\qquad \lim_{\lambda\to 0} P[N \leqq n] = 1,$$

$$(3.11)\qquad \lim_{\lambda\to\infty} P[N \leqq n] = 0$$

In particular, the cdf *of* N *converges to a degenerate distribution as* $\lambda \to 0$; *i.e.*,

$$(3.12)\qquad \lim_{\lambda\to 0} P[N = 2] = 1.$$

REMARK. (3.11) implies that as $\lambda \to \infty$, $N \to \infty$ in probability, which is also implied by [3], see (3.14) in Lemma 3.2 below.

For large values of λ, we first state the following lemmas which are due to Chow and Robbins [3]:

LEMMA 3.2. *Let* $y_n (n = 1, 2, \cdots)$ *be any sequence of random variables such that* $y_n > 0$ a.s., $\lim_{n\to\infty} y_n = 1$ a.s., *let* $f(n)$ *be any sequence of constants such that* $f(n) > 0$, $\lim_{n\to\infty} f(n) = \infty$, $\lim_{n\to\infty} f(n)/f(n-1) = 1$, *and for each* $t > 0$, *define*

$$(3.13)\qquad N = N(t) = \text{smallest } n \geqq 1 \text{ such that } y_n \leqq f(n)/t.$$

then N *is well defined and nondecreasing as a function of* t,

$$(3.14)\qquad \lim_{t\to\infty} N = \infty \text{ a.s.}, \lim_{t\to\infty} EN = \infty$$

and

$$(3.15)\qquad \lim_{t\to\infty} f(N)/t = 1 \text{ a.s.}$$

LEMMA 3.3. *If the conditions of Lemma* 3.2. *hold and if also* $E(\sup_n y_n) < \infty$, *then*

$$(3.16)\qquad \lim_{t\to\infty} Ef(N)/t = 1.$$

Let N_0 be the sample size required under the single-stage procedure for the LF configuration, the following theorem investigates the asymptotic relative efficiency of the sequential procedure wrt the single-stage procedure.

THEOREM 3.3. *Let* N_0 *be defined in* (1.10) *and* N *be the random sample size defined in* (3.1). *Then*

$$(3.17)\qquad \lim_{\lambda\to\infty} N/N_0 = 1 \text{ a.s.};$$

$$(3.18)\qquad \lim_{\lambda\to\infty} EN/N_0 = 1.$$

Using the terminology in [3], it follows from (3.18) that the procedure R_3 is asymptotically relatively efficient.

PROOF. For $\nu = (k + 1)(n - 1)$, set

$$y_n = S_\nu^2/\sigma^2, \quad f(n) = n(b/h_\nu)^2 \quad \text{and} \quad t = 2\lambda^2 b^2.$$

Since y_n is distributed as $[(k+1)(n-1)]^{-1}(V_1 + V_2 + \cdots + V_{n-1})$ where the V's are i.i.d. chi-square chance variables each with $(k+1)$ degrees of freedom, it follows from the strong law of large numbers that $\lim_{n\to\infty} y_n = 1$ a.s. Since $h_\nu \to b$, the rest of the conditions in Lemma 3.2 are easily seen to be satisfied. Hence (3.17) is proved.

To prove (3.18), by Lemma 3.3 it suffices to show that $E(\sup_n y_n) < \infty$. Let $c > 1$ be any real number. Then

$$\begin{aligned} P[\sup_n y_n > c] &= P\{\textstyle\bigcup_{n=1}^{\infty} [[(k+1)n]^{-1}\sum_{j=1}^{n} V_j > c]\} \\ &\leqq \textstyle\sum_{n=1}^{\infty} P[\sum_{j=1}^{n} V_j > (k+1)nc] \\ &\leqq \textstyle\sum_{n=1}^{\infty} P[|\sum_{j=1}^{n} V_j - (k+1)n| > (k+1)n(c-1)] \\ &\leqq \textstyle\sum_{n=1}^{\infty} E[\sum_{j=1}^{n} V_j - (k+1)n]^4/\{[(k+1)n]^4(c-1)^4\} \end{aligned}$$

where the last inequality follows from Markov Inequality. By elementary calculations it is easily seen that the fourth central moment of a chi-square chance variable with r degrees of freedom is $12r(r+4)$. Hence

$$E[\textstyle\sum_{j=1}^{n} V_j - (k+1)n]^4 = 12(k+1)n[(k+1)n+4] \leqq 60(k+1)^2n^2$$

which implies that for every $c > 1$,

$$(3.19)\quad P[\sup_n y_n > c] \leqq 60(k+1)^{-2}(c-1)^{-4}\textstyle\sum_{n=1}^{\infty} n^{-2} = M(c-1)^{-4}$$

for some finite number M that does not depend on c. Thus

$$\begin{aligned} E(\sup_n y_n) &\leqq 2 + \textstyle\sum_{j=1}^{\infty} (2+j)P[2+j-1 < \sup_n y_n \leqq 2+j] \\ &\leqq 2 + \textstyle\sum_{j=1}^{\infty} (2+j)P[\sup_n y_n > 2+j-1] \\ &\leqq 2 + \textstyle\sum_{j=1}^{\infty} M(2+j)j^{-4} \leqq 2 + 3M\sum_{j=1}^{\infty} j^{-3} < \infty \end{aligned}$$

which completes the proof of (3.18).

3.2 *The* PCD *function and its asymptotic behavior.*

The sequential procedure provides only an "asymptotic" solution to our problem in the sense that the PCD under this procedure may be slightly less than P^* for some values of the unknown parameter σ^2, or equivalently, λ. In the following we examine the *PCD* function and its asymptotic behavior as a function of λ.

For the covariance matrix $\boldsymbol{\Sigma}$ specified in (1.8), we first define

$$(3.20)\quad \beta((\tfrac{1}{2}n)^{\frac{1}{2}}/\lambda) = \int_{-\infty}^{(\frac{1}{2}n)^{\frac{1}{2}}/\lambda}\int_{-\infty}^{(\frac{1}{2}n)^{\frac{1}{2}}/\lambda}\cdots\int_{-\infty}^{(\frac{1}{2}n)^{\frac{1}{2}}/\lambda} (2\pi)^{-k/2}|\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp(-\tfrac{1}{2}\mathbf{y}'\boldsymbol{\Sigma}^{-1}\mathbf{y})\textstyle\prod_{i=1}^{k} dy_i$$

(note that β depends on n and λ only through the ratio $(\frac{1}{2}n)^{\frac{1}{2}}/\lambda)$. Then it is easily seen that for any mean vector $\boldsymbol{\mu}$, the conditional PCD given $N = n$ is lower bounded by

$$P[\mathrm{CD} \mid \boldsymbol{\mu}, \lambda, N = n] \geqq P[\mathrm{CD} \mid \boldsymbol{\mu}^0, \lambda, N = n] = \beta((\tfrac{1}{2}n)^{\frac{1}{2}}/\lambda)$$

for every n, where $\mathbf{u}^0$ is the LF configuration given in (1.7). Hence it follows that

$$P[\text{CD} \mid \mathbf{u}, \lambda, R_3] \geqq P[\text{CD} \mid \mathbf{u}^0, \lambda; R_3] = E\beta((\tfrac{1}{2}N)^{\frac{1}{2}}/\lambda) \tag{3.21}$$

where the expectation is taken over N space (it should be observed from (3.1) that the distribution of N here depends on the parameter λ).

THEOREM 3.4. *For every mean vector* $\mathbf{u}$,

$$\lim_{\lambda\to 0} P[\text{CD} \mid \mathbf{u}, \lambda; R_3] = 1, \tag{3.22}$$

$$\lim_{\lambda\to\infty} P[\text{CD} \mid \mathbf{u}, \lambda; R_3] \geqq P^*. \tag{3.23}$$

PROOF. By (3.21), we can restrict our attention to $\mathbf{u} = \mathbf{u}^0$ and work on $E\beta((\frac{1}{2}N)^{\frac{1}{2}}/\lambda)$.

Since β is continuous and monotonically increasing and $N \geqq 2$ a.s., it follows that $\beta((\frac{1}{2}N)^{\frac{1}{2}}/\lambda) \geqq \beta(1/\lambda)$ a.s. and

$$\lim_{\lambda\to 0} E\beta((\tfrac{1}{2}N)^{\frac{1}{2}}/\lambda) \geqq \lim_{\lambda\to 0} E\beta(1/\lambda) = \lim_{\lambda\to 0} \beta(1/\lambda) = 1. \tag{3.24}$$

This proves (3.22).

To prove (3.23), let $\{\lambda_j\}_1^\infty$ be an arbitrary but fixed monotonically increasing sequence such that $\lim_{j\to\infty} \lambda_j = \infty$. By (3.17), $\lim_{j\to\infty} N/(2\lambda_j^2 b^2) = 1$ a.s. where b is such that $\beta(b) = P^*$. Since a.s. convergence is preserved by continuous mapping, it follows

$$\lim_{j\to\infty} \beta((\tfrac{1}{2}N)^{\frac{1}{2}}/\lambda_j) = \beta(b) \text{ a.s.} \tag{3.25}$$

Let $F_j(\cdot)$ and $F(\cdot)$ be the cdf of $\beta((\frac{1}{2}N)^{\frac{1}{2}}/\lambda_j)(j = 1, 2, \cdots)$ and $\beta(b)$ respectively. Then it follows from (3.25) that

$$F_j(\cdot) \to_d F(\cdot). \tag{3.26}$$

But the sequence of chance variables $\{\beta((\frac{1}{2}N)^{\frac{1}{2}}/\lambda_j)\}$ is uniformly bounded by 0 and 1 and $\beta(b)$ is a degenerate variable, so $F_j(0) = F(0) = 0$ and $F_j(1) = F(1) = 1 (j = 1, 2, \cdots)$. Applying Helly-Bray Lemma,

$$\lim_{j\to\infty} E\beta((\tfrac{1}{2}N)^{\frac{1}{2}}/\lambda_j) = \lim_{j\to\infty} \int_0^1 y\,dF_j(y) = \int_0^1 y\,dF(y) = P^*.$$

Since the sequence $\{\lambda_j\}$ is arbitrarily chosen, the proof of (3.23) is completed.

3.3 *The expected misclassification size.* Let EM_0 be the expected misclassification size under R_3 when $\mathbf{u} = \mathbf{u}^0$, we first give a lower bound on EM_0 for all λ.

LEMMA 3.4. *Let EN be the expected sample size under R_3, then*

$$EM_0 \geqq k[1 - \Phi((\tfrac{1}{2}EN)^{\frac{1}{2}}/\lambda)] \quad \textit{for every } \lambda. \tag{3.27}$$

PROOF. The proof of this lemma is similar to that of Lemma 2.2.

The following theorem shows that for extreme values of λ, EM_0 under the sequential procedure is the same as that under the single-stage procedure.

THEOREM 3.5. *Let b be defined in (1.9) with $P = P^*$,*

$$\lim_{\lambda\to 0} EM_0 = 0, \tag{3.28}$$

$$\lim_{\lambda\to\infty} EM_0 = k[1 - \Phi(b)]. \tag{3.29}$$

PROOF. The proof of this theorem is similar to that of Theorem 3.4 with $\beta((\frac{1}{2}N)^{\frac{1}{2}}/\lambda) = k[1 - \Phi((\frac{1}{2}N)^{\frac{1}{2}}/\lambda)]$.

4. Acknowledgments. I wish to express my deep gratitude to Professor Milton Sobel for his guidance and continuing encouragement during the course of this research. I also wish to thank the referee whose helpful comments lead to an improved formulation of Theorem 1.3.

APPENDIX

ON A PROPERTY OF CERTAIN MULTIVARIATE NORMAL DISTRIBUTIONS

For arbitrary but fixed positive integers k and $r(0 \leqq r \leqq k)$, let the $(k \times k)$ positive definite covariance matrix $\Sigma_r = (\sigma_{ij})$ be such that

$$
\begin{aligned}
\sigma_{ij} &= \sigma^2 && \text{if } i = j, \\
&= \rho\sigma^2 && \text{if } i \neq j \text{ and } i, j \in \{1, 2, \cdots, r\} \text{ or} \\
&&& \quad i, j \in \{r+1, r+2, \cdots, k\}, \\
&= -\rho\sigma^2 && \text{if } i \in \{1, 2, \cdots, r\} \text{ and} \\
&&& \quad j \in \{r+1, r+2, \cdots, k\};
\end{aligned} \tag{A.1}
$$

where $\rho \in (0, 1)$ and $\sigma^2 > 0$. Let the multivariate normal probability integral $P(r)$ be defined by

$$
\begin{aligned}
P(r) &= P_{\rho,c}(r) \\
&= \int_{-\infty}^{c} \int_{-\infty}^{c} \cdots \int_{-\infty}^{c} (2\pi)^{-k/2} \, |\Sigma_r|^{-\frac{1}{2}} \exp\left(-\tfrac{1}{2}\mathbf{y}'\Sigma_r^{-1}\mathbf{y}\right) \prod_{i=1}^{k} dy_i
\end{aligned} \tag{A.2}
$$

for $c \in (-\infty, \infty)$. It should be observed that for either $r = 0$ or $r = k$, $P(r)$ is the probability integral of an equi-correlated multivariate normal distribution. Let $[x]$ denote the largest integer $\leqq x$. The purpose of this appendix is to prove the following

THEOREM. *For every positive integer k, $\rho \in (0, 1)$, $c \in (-\infty, \infty)$ and $\sigma^2 > 0$, we have*

$$P(r) = P(k - r) \quad \textit{for} \quad r = 0, 1, \cdots, k; \tag{A.3}$$

$$P(r + 1) < P(r) \quad \textit{for} \quad r = 0, 1 \cdots, [\tfrac{1}{2}(k - 2)]. \tag{A.4}$$

REMARK. It follows immediately from this theorem that

(1) $P(r + 1) > P(r)$ for $r = k - 1, k - 2, \cdots, [\frac{1}{2}(k + 1)]$; and

(2) $P(r)$ achieves an unique minimum at $r = k/2$ when k is even and a common minimum at $r = \frac{1}{2}(k - 1)$ and $r = \frac{1}{2}(k + 1)$ when k is odd.

COROLLARIES. (1) *Let $(U_1, U_2, \cdots, U_k)$ have the joint distribution $N(0, \Sigma_r)$ and let $(V_1, V_2, \cdots, V_k)$ have the joint distribution $N(0, \Sigma_s)$. Let $U = \max_{1 \leqq i \leqq k} U_i$, $V = \max_{1 \leqq i \leqq k} V_i$. If $|r - k/2| < |s - k/2|$, then U is stochastically larger than V.*

(2) *Let $(U_1, U_2, \cdots, U_k)$ and $(V_1, V_2, \cdots, V_k)$ follow multivariate t dis-*

tributions with common degrees of freedom ν and correlation matrices Σ_r and Σ_s, respectively. Let $U = \max_{1 \leq i \leq k} U_i$, $V = \max_{1 \leq i \leq k} V_i$. If $|r - k/2| < |s - k/2|$, then U is stochastically larger than V.

Before we prove this theorem we first prove a lemma dealing with symmetric functions. Let $f(z)$ and $G(z)$ be two real functions defined for $z \in (-\infty, \infty)$ such that $f(z) \geqq 0$, $\int_{-\infty}^{\infty} f(z)\, dz < \infty$ and $0 \leqq G(z) \leqq M$ for some $M > 0$. For arbitrary but fixed real numbers $\eta \in (0, \infty)$, $s \in (-\infty, \infty)$ and any positive integer k, we define

$$\beta(r) = \int_{-\infty}^{\infty} G^r(\eta z + s) G^{k-r}(-\eta z + s) f(z)\, dz, \quad \text{for} \quad r = 0, 1, \cdots, k \tag{A.5}$$

and its first difference

$$\Delta\beta(r) = \beta(r + 1) - \beta(r), \quad \text{for} \quad r = 0, 1, \cdots, k - 1. \tag{A.6}$$

Lemma. *If $f(z) = f(-z)$ and $G(z)$ is monotonically increasing, then for every $\eta \in (0, \infty)$ and $s \in (-\infty, \infty)$,*

$$\beta(r) = \beta(k - r) \quad \textit{for} \quad r = 0, 1, \cdots, k; \tag{A.7}$$

$$\Delta\beta(r) \leqq 0 \quad \textit{for} \quad r = 0, 1, \cdots, [\tfrac{1}{2}(k - 2)]. \tag{A.8}$$

Proof. Property (A.7) follows immediately by setting $u = -z$ in the integral on the rhs of (A.5).

To prove (A.8), first let

$$H(z) = [G(\eta z + s) - G(-\eta z + s)] f(z) \quad \text{for} \quad z \in (-\infty, \infty); \tag{A.9}$$

then it is easily seen that

$$H(z) = -H(-z) \quad \text{for} \quad z \in (-\infty, \infty), \tag{A.10}$$

and since $\eta > 0$,

$$H(z) \geqq 0 \quad \text{for} \quad z \in (0, \infty). \tag{A.11}$$

For every fixed $r \leqq \frac{1}{2}(k - 1)$,

$$\begin{aligned} \Delta\beta(r) = &\int_0^{\infty} G^r(\eta z + s) G^{k-r-1}(-\eta z + s) H(z)\, dz \\ &+ \int_{-\infty}^{0} G^r(\eta z + s) G^{k-r-1}(-\eta z + s) H(z)\, dz. \end{aligned} \tag{A.12}$$

Consider the second integral I_2 on the rhs of the above expression. Setting $u = -z$ and applying (A.10), we have

$$I_2 = -\int_0^{\infty} G^{k-r-1}(\eta u + s) G^r(-\eta u + s) H(u)\, du.$$

Substituting this in (A.12) gives

$$\begin{aligned} \Delta\beta(r) = &\int_0^{\infty} G^r(\eta z + s) G^r(-\eta z + s) H(z) \\ &\cdot [G^{k-2r-1}(-\eta z + s) - G^{k-2r-1}(\eta z + s)]\, dz. \end{aligned} \tag{A.13}$$

Since by (A.11) $G^r(\eta z + s) G^r(-\eta z + s) H(z) \geq 0$ and $G(-\eta z + s) \leqq G(\eta z + s)$

for $z \epsilon (0, \infty)$, it follows that $\Delta\beta(r) \leqq 0$ for $k - 2r - 1 > 0$ or equivalently, for $r \leqq [\frac{1}{2}(k-2)]$; and $\Delta\beta(r) = 0$ for $r = \frac{1}{2}(k-1)$. This proves (A.8).

REMARK. If the function $G(z)$ is strictly increasing in $(-\infty, \infty)$, then every inequality in the proof of the above lemma will be a strict inequality hence (A.8) will be a strict inequality.

PROOF OF THE THEOREM. Without loss of generality, we assume $\sigma^2 = 1$.

Let $Z_0, Z_1, \cdots, Z_k$ be independent standard normal chance variables, $\varphi(\cdot)$ and $\Phi(\cdot)$ be the density function and cdf, respectively, of the standard normal distribution. For arbitrary but fixed $\rho \, \varepsilon \, (0, 1)$ and $c \, \varepsilon \, (-\infty, \infty)$, let $\eta > 0$ satisfy $\rho = \eta^2/(\eta^2 + 1)$ and let $s = c(\eta^2 + 1)^{\frac{1}{2}}$.
For fixed $r (0 \leqq r \leqq k)$, we define

$$\text{(A.14)} \quad \begin{aligned} Y_i &= (Z_i - \eta Z_0)/(\eta^2 + 1)^{\frac{1}{2}} \qquad \text{for} \quad i = 1, 2, \cdots, r; \\ &= (\eta Z_0 - Z_i)/(\eta^2 + 1)^{\frac{1}{2}} \qquad \text{for} \quad i = r+1, r+2, \cdots, k. \end{aligned}$$

Then $(Y_1, Y_2, \cdots, Y_k)$ follows a multivariate normal distribution with mean vector 0 and covariance matrix Σ_r defined in (A.1) for $\sigma^2 = 1$. Hence

$$\begin{aligned} P(r) &= P[Z_i \leqq \eta Z_0 + s, Z_j > \eta Z_0 - s; 1 \leqq i \leqq r, r < j \leqq k] \\ &= \textstyle\int_{-\infty}^{\infty} \Phi^r(\eta z + s)\Phi^{k-r}(-\eta z + s)\varphi(z)\, dz. \end{aligned}$$

The rest of the argument follows from the lemma. This completes the proof.

EXAMPLE. We consider the special case $k = 2$ and $c = 0$. It is well-known that if (U_1, U_2) follows a bivariate normal distribution with means 0, a common but arbitrary variance σ^2 and correlation coefficient ρ, then

$$g(\rho) = P[U_1 \leqq 0, U_2 \leqq 0] = \tfrac{1}{4} + (2\pi)^{-1} \arcsin \rho.$$

If $\rho > 0$, then $g(\rho) > g(-\rho)$. Our result agrees with this statement because $g(\rho)$ corresponds to $P(r)$ for either $r = 0$ or $r = 2$ and $g(-\rho)$ corresponds to $P(r)$ for $r = 1$.

REFERENCES

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

[2] BECHHOFER, R. E., DUNNETT, C. W. and SOBEL, M. (1954). A two-sample multiple decision procedure for ranking means of normal populations with a common unknown variance. *Biometrika* **41** 170–176.

[3] CHOW, Y. S. and ROBBINS, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

[4] DUNNETT, C. W. and SOBEL, M. (1954). A bivariate generalization on Student's t-distribution, with tables for certain special cases. *Biometrika* **41** 153–169.

[5] FELLER, W. (1961). *An Introduction to Probability Theory and its Applications*, **1** (2nd ed.) Wiley, New York.

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

[7] GUPTA, S. S. and SOBEL, M. (1958). On selecting a subset which contains all populations better than a standard. *Ann. Math. Statist.* **29** 235–244.

[8] LEHMANN, E. L. (1957). A theory of some multiple decision problems, I, II. *Ann. Math. Statist.* **28** 1–25 and 547–572.

[9] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York.

[10] LEHMANN, E. L. (1961). Some model I problems of selection. *Ann. Math. Statist.* **32** 990–1012.

[11] LEHMANN, E. L. (1962). *Notes on the Theory of Estimation*. Univ. California, Berkeley.

[12] MILTON, R. C. (1963). Tables of the equally correlated multivariate normal probability integral. Technical Report No. 27, Department of Statistics, Univ. Minnesota.

[13] PAULSON, E. (1952). On the comparison of several experimental categories with a control. *Ann. Math. Statist.* **23** 239–246.

[14] ROBBINS, H., SOBEL, M. and STARR, N. (1968). A sequential procedure for selecting the largest of k means. *Ann. Math. Statist.* **39** 88–92.

[15] STARR, N. (1966). The performance of a sequential procedure for the fixed-width interval estimation of the mean. *Ann. Math. Statist.* **37** 36–50.

[16] STEIN, C. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Statist.* **16** 243–258.

[17] STROUD, A. H. and SECREST, D. (1966). *Gaussian Quadrature Formulas*. Prentice-Hall, New Jersey.

[18] TONG, Y. L. (1967). On partitioning a set of normal distributions by their locations with respect to a control using a single-stage, a two-stage and a sequential procedure. Technical Report No. 90, Department of Statistics, Univ. Minnesota.