BULLETIN OF THE
INSTITUTE OF MATHEMATICS
ACADEMIA SINICA
Volume 3, Number 1, June 1975

BAYES APPROACH TO A PROBLEM OF PARTITIONING k NORMAL POPULATIONS

BY

WEN-TAO HUANG

Abstract. A problem of partitioning k univariate normal populations with respect to a standard based on the location parameter is set up in a Bayes formulation. Two kinds of loss function are considered. Bayes procedures and empirical Bayes procedures are obtained. An optimum property of the Bayes procedures is also shown.

0. **Introduction.** In many experimental situations the experimenter is confronted with the problem of partitioning k populations into two classes. Usually, one class is better than a standard and the other is worse. The terms better and worse are up to the experimenter and depend on his particular goal. For instance, suppose a pharmaceutical company is testing k new products of aspirin against the standard product which is on the market. It is desired to classify these k products into two groups. One group is more effective than the standard and the other group is less effective. Here, the so-called efficiency is up to the experimenter. It may depend on the length of time needed to cure the pain. Or it may be the frequency with which the drug cures the pain etc.

Recently Tong [4] considered the problem of partitioning k univariate normal populations according to their location parameters with respect to a control. Huang [2] also considered a problem of partitioning k multivariate normal populations in terms of the generalized variance with respect to a control.

However, in some practical situations the parameter of population changes from time to time. For instance, the fraction of defective products from a given new machine is different from what it is when the machine becomes old. And in general, a parameter of the

Received by the editors May 15, 1974 and, in revised form, October 16, 1974.

distribution of a population we are sampling is distributed according to some probability law. This motivates the consideration of our problem in terms of Bayes formulation.

In this paper we consider the problem of partitioning k univariate normal populations into disjoint exhaustive classes with respect to a given control in terms of location parameters.

1. **Definitions and notation.** Let $\pi_0, \pi_1, \pi_2, \cdots, \pi_k$ be $k+1$ normal populations with associated means $\theta_0, \theta_1, \cdots, \theta_k$ and associated variances $\sigma_0^2, \sigma_1^2, \cdots, \sigma_k^2$ respectively. For given positive values ρ_1 and ρ_2 $(\rho_2 < \rho_1)$ we define

$$
\begin{aligned}
\mathfrak{P} &= \{\pi_1, \pi_2, \cdots, \pi_k\}, \\
\mathfrak{P}_G &= \{\pi_i : \theta_i - \theta_0 \geq \rho_1\}, \\
\mathfrak{P}_B &= \{\pi_i : \theta_i - \theta_0 \leq -\rho_2\}, \\
\mathfrak{P}_I &= \{\pi_i : -\rho_2 < \theta_i - \theta_0 < \rho_1\}.
\end{aligned} \tag{1.1}
$$

For convenience, we also define

$$
\begin{aligned}
K &= \{1, 2, \cdots, k\}, \\
A &= \{i : \pi_i \in \mathfrak{P}_G\}, \\
B &= \{i : \pi_i \in \mathfrak{P}_B\}, \\
I &= \{i : \pi_i \in \mathfrak{P}_I\}, \\
E &= \{S : S \subset K\}.
\end{aligned} \tag{1.2}
$$

Let X_i denote the sample space of π_i for $i = 0, 1, 2, \cdots, k$, and let $Y = X_0 \times X_1 \times \cdots \times X_k$ be the cartesian product. A decision function d for our problem is a measurable function from Y into E, and for an observation $y \in Y$ we partition $\mathfrak{P}$ into $S_G = \{\pi_i : i \in S\}$ and $S_B = \mathfrak{P} \backslash S_G$ provided $d(y) = S$. We define

$$M(d, y) = \{\pi_i : i \in (d(y) \cap B) \cup (K - d(y)) \cap A\}. \tag{1.3}$$

If $\pi_i \in M(d, y)$, we say π_i is misclassified by d when y is observed. If $M(d, y) = \emptyset$, where $\emptyset$ denotes the empty set, we say $d(y)$ makes a correct decision (CD). A loss function $L(\cdot, \cdot)$ is a nonnegative function on $E \times \Omega$ (where Ω denotes the parameter space with element $(\theta_0, \theta_1, \cdots, \theta_k)$) such that

$$L(S, \omega) = 0 \quad \text{if } A \subset S \subset A \cup I, \qquad > 0 \quad \text{otherwise}$$

for all $\omega \in \Omega$.

We consider the following two types of loss function:

$$(1.4) \qquad \begin{aligned} L_{1i}(S, \omega) &= \alpha && \text{if } i \in A \text{ and } i \notin S, \\ &= \beta && \text{if } i \in B \text{ and } i \in S, \\ &= 0 && \text{otherwise}, \end{aligned}$$

where α and β are some positive values.

$$(1.5) \qquad \begin{aligned} L_{2i}(S, \omega) &= \alpha(\theta_i - \theta_0) && \text{if } i \in A \text{ and } i \notin S, \\ &= \beta(\theta_0 - \theta_i) && \text{if } i \in B \text{ and } i \in S, \\ &= 0 && \text{otherwise}. \end{aligned}$$

We then have

$$(1.6) \qquad \begin{aligned} L_1(S, \omega) &= \sum_{i=1}^{k} L_{1i}(S, \omega), \\ L_2(S, \omega) &= \sum_{i=1}^{k} L_{2i}(S, \omega). \end{aligned}$$

2. **Bayes procedures.** For each i, $i = 0, 1, 2, \cdots, k$, we assume σ_i^2 is known and θ_i has a prior normal distribution with known mean λ_i and known variance φ_i^2. Let the configuration of $\pi_0, \pi_1, \cdots, \pi_k$ be denoted by $\Omega = \{(\theta_0, \theta_1, \cdots, \theta_k) : -\infty < \theta_i < \infty\}$, let G_i denote the normal distribution of θ_i and let G be the joint prior distribution on Ω such that G is the product of G_i. Let X_{ij} denote the jth observation from π_i and let $\bar{X}_i = \sum_{j=1}^{n} X_{ij}/n$ be the sample mean from π_i for $i = 0, 1, 2, \cdots, k$. Then we have the following Bayes procedures.

THEOREM 1. *Using loss functions L_1 and L_2 defined by* (1.6), *we have corresponding Bayes procedures R_{1G} and R_{2G} for our problem defined as follows.*

Define

$$(2.1) \qquad \begin{aligned} \alpha_i &= (n\varphi_i^2 \bar{x}_i + \sigma_i^2 \lambda_i)/(\sigma_i^2 + n\varphi_i^2), \\ \beta_i &= \sigma_i^2 \varphi_i^2/(\sigma_i^2 + n\varphi_i^2) \qquad (i = 0, 1, 2, \cdots, k), \\ C(\alpha, \beta, r, \delta) &= \exp[-(\alpha - r)^2/2(\beta + \delta)], \\ H(\alpha, \beta; r, \delta; a) &= \int_{-\infty}^{\infty} x\Phi(x + a; \alpha, \beta)\, d\Phi(x; r, \delta), \end{aligned}$$

where $\Phi(x;\ \alpha,\ \beta)$ *denotes the cdf of normal distribution with mean* α *and variance* β,

$$(2.2)\quad \begin{aligned} f_i &= f(\alpha_i,\ \beta_i,\ \alpha_0,\ \beta_0,\ \rho_1) \\ &\equiv [\beta_i C(\alpha_i - \rho_1,\ \beta_i,\ \alpha_0,\ \beta_0)/(2\pi(\beta_0+\beta_i))^{1/2}] \\ &\qquad + \alpha_i - \alpha_0 + H(\alpha_i,\ \beta_i;\ \alpha_0,\ \beta_0;\ \rho_1) \\ &\qquad - \alpha_i \Phi((\alpha_0 - \alpha_i + \rho_1)/(\beta_0+\beta_i)^{1/2}), \end{aligned}$$

$$(2.3)\quad \begin{aligned} g_i &= g(\alpha_j,\ \beta_j,\ \alpha_0,\ \beta_0,\ \rho_2) \\ &\equiv H(\alpha_j,\ \beta_j;\ \alpha_0,\ \beta_0;\ -\rho_2) \\ &\qquad + \alpha_j \Phi((\alpha_0 - \alpha_j - \rho_2)/(\beta_0+\beta_j)^{1/2}) \\ &\qquad - [\beta_j C(\alpha_j + \rho_2,\ \beta_j;\ \alpha_0,\ \beta_0)/(2\pi(\beta_0+\beta_j))^{1/2}], \end{aligned}$$

where $\Phi(x)$ *denotes the cdf of the standard normal* $(i=0, 1, 2,\cdots, k)$.

(i) R_{1G}: *we classify* π_i *in* S_G *iff* $i\in S$ *such that, if* $S^c = K - S$,

$$(2.4)\quad \begin{aligned} &\alpha \sum_{i\in S^c} [1-\Phi(J(i,\ \rho_1))] + \beta \sum_{j\in S} \Phi(J(j,\ \rho_2)) \\ &\quad = \min_{T\in E}\Big\{\alpha \sum_{i\in T^c}[1-\Phi(J(i,\ \rho_1))] + \beta\sum_{j\in T}\Phi(J(j,\ \rho_2))\Big\}, \end{aligned}$$

where

$$(2.5)\quad J(i,\ \rho) = (\alpha_0 - \alpha_i + \rho)/(\beta_0 + \beta_i)^{1/2} \qquad (i = 1,\ 2,\cdots,\ k),$$

where α_j *and* β_j *are defined by* (2.1).

(ii) R_{2G}: $\pi_i \in S_G$ *iff* $i\in S$ *satisfying*

$$(2.6)\quad \alpha\sum_{i\in S^c} f_i + \beta\sum_{j\in S} g_j = \min_{T\in E}\Big(\alpha\sum_{i\in T^c} f_i + \beta\sum_{j\in T} g_j\Big).$$

Proof. Since the proofs of (i) and (ii) are analogous and computations for (i) are easier than (ii), we give the proof for (ii).

For a decision function d the Bayes risk is given by

$$(2.7)\quad R(d,\ G) = \int_\Omega dG(\omega)\int_Y L_2(d(y),\ \omega)\, f(y|\omega)\, dy,$$

where $f(y|\omega) = \prod_{j=0}^{k} f(x_j|\theta_j)$ and $f(x_j|\theta_j)$ is the common density of each random observation from π_j $(j = 0, 1, 2,\cdots, k)$. By Fubini's theorem, we have

$$R(d,\ G) = \int_Y \eta_G(d,\ y)\, dy,$$

where

$$(2.8)\quad \eta_G(d,\ y) = \int_\Omega L_2(d(y),\ \omega)\, f(y|\omega)\, dG(\omega).$$

To find a Bayes solution d_G, it suffices to require $\eta_G(d_G, y) \leq \eta_G(d, y)$ for almost all $y \in Y$ and any d. Hence we need

$$d_G(y) = S \text{ such that } \eta_G(S, y) = \min_{D \subset K} \eta_G(D, y) \tag{2.9}$$

for almost all $y \in Y$. Now

$$\begin{aligned}\eta_G(d, y) &= \sum_{i=1}^{k} \int_{\Omega} L_{2i}(d, \omega) f(y|\omega)\, dG(\omega) \\ &= \alpha \sum_{i \in S^c} \int_{\Omega_{1i}} (\theta_i - \theta_0) f(y|\omega)\, dG(\omega) \\ &\quad + \beta \sum_{j \in S} \int_{\Omega_{2j}} (\theta_0 - \theta_j) f(y|\omega)\, dG(\omega)\,,\end{aligned}$$

where $d(y) = S$ and

$$\begin{aligned}\Omega_{1i} &= \{(\theta_0, \theta_1, \cdots, \theta_k) | \theta_i - \theta_0 \geq \rho_1\}\,, \\ \Omega_{2i} &= \{(\theta_0, \theta_1, \cdots, \theta_k) | \theta_i - \theta_0 \leq -\rho_2\}\,.\end{aligned} \tag{2.10}$$

Let G_i and g_i denote respectively the cdf and pdf of θ_i. Then

$$\begin{aligned}\eta_G(d, x) &= \alpha \sum_{i \in S^c} \prod_{j \neq i} \int_{-\infty}^{\infty} f(x_j|\theta_j)\, dG_j(\theta) \\ &\quad \cdot \int_{-\infty}^{\infty} \int_{\theta_0+\rho_1}^{\infty} (\theta_i - \theta_0) f(x_i|\theta_i)\, dG_i(\theta) f(x_0|\theta_0)\, dG_0(\theta_0) \\ &\quad + \beta \sum_{j \in S} \prod_{i \neq j} \int_{-\infty}^{\infty} f(x_i|\theta_i)\, dG_i(\theta) \\ &\quad \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\theta_0-\rho_2} (\theta_0 - \theta_j) f(x_j|\theta_j)\, dG_j(\theta) f(x_0|\theta_0)\, dG_0(\theta_0)\,.\end{aligned} \tag{2.11}$$

We note that

$$\begin{aligned}\int_{-\infty}^{\infty} f(x_j|\theta_j)\, dG_j(\theta_j) &= \int_{-\infty}^{\infty} f(x_j|\theta)\, g_j(\theta)\, d\theta \\ &= \int_{-\infty}^{\infty} g_j(\theta|x_j)\, f_{G_j}(x_j)\, d\theta\,,\end{aligned} \tag{2.12}$$

where

$$\begin{aligned}f_{G_j}(x_j) &= \int_{-\infty}^{\infty} f(x_j|\theta)\, dG_j(\theta)\,, \\ g_j(\theta|x_j) &= f(x_j|\theta)\, g_j(\theta)/f_{G_j}(x_j)\,.\end{aligned} \tag{2.13}$$

It is seen that $g_j(\theta|x_j)$ is a normal density with mean α_j and variance β_j defined by (2.1). We also note that

$$\int_{\theta_0+\rho_1}^{\infty} \theta f(x_i|\theta)\, dG_i(\theta) = \int_{\theta_0+\rho_1}^{\infty} \theta g_i(\theta|x_i) f_{G_i}(x_i)\, d\theta$$

$$(2.14) \qquad = f_{G_i}(x_i) \left\{ \left(\frac{\beta_i}{2\pi}\right)^{1/2} \exp[-(\theta_0+\rho_1-\alpha_i)^2/2\beta_i] + \alpha_i [1 - \Phi((\theta_0+\rho_1-\alpha_i)/\sqrt{\beta_i})] \right\},$$

where α_i and β_i are defined by (2.1). We then have

$$\int_{-\infty}^{\infty} f(x_0|\theta_0)\, dG_0(\theta_0) \int_{\theta_0+\rho_1}^{\infty} \theta f(x_i|\theta)\, dG_i(\theta)$$

$$(2.15) \qquad = f_{G_i}(x_i) f_{G_0}(x_0) \{\beta_i C(\alpha_i - \rho_1, \beta_i; \alpha_0, \beta_0)/[2\pi(\beta_0+\beta_i)]^{1/2} + \alpha_i - \alpha_i \Phi((\alpha_0 - \alpha_i + \rho_1)/(\beta_0+\beta_i)^{1/2})\},$$

where $C(\alpha, \beta; \gamma, \delta)$ is defined by (2.1). Similarly, we have

$$(2.16) \qquad \int_{-\infty}^{\infty} \theta_0 f(x_0|\theta_0)\, dG_0(\theta_0) \int_{\theta_0+\rho_1}^{\infty} f(x_i|\theta)\, dG_i(\theta) = f_{G_i}(x_i) f_{G_0}(x_0)[\alpha_0 - H(\alpha_i, \beta_i; \alpha_0, \beta_0; \rho_1)],$$

where

$$H(\alpha, \beta; \gamma, \delta; a) = \int_{-\infty}^{\infty} x\Phi(x+a; \alpha, \beta)\, d\Phi(x; \gamma, \delta)$$

as defined by (2.1).

It follows from (2.15) and (2.16) that

$$\int_{-\infty}^{\infty} \int_{\theta_0+\rho_1}^{\infty} (\theta - \theta_0) f(x_i|\theta)\, dG_i(\theta) f(x_0|\theta_0)\, dG_0(\theta_0)$$

$$(2.17) \qquad = f_{G_i}(x_i) f_{G_0}(x_0) \{\beta_i C(\alpha_i - \rho_1, \beta_i; \alpha_0, \beta_0)/[2\pi(\beta_0+\beta_i)]^{1/2} + \alpha_i - \alpha_i \Phi((\alpha_0 + \rho_1 - \alpha_i)/(\beta_0+\beta_i)^{1/2}) - \alpha_0 + H(\alpha_i, \beta_i; \alpha_0, \beta_0; \rho_1)\}.$$

Similarly, we also have

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\theta_0-\rho_2} (\theta_0 - \theta) f(x_i|\theta) f(x_0|\theta_0)\, dG_0(\theta_0)\, dG_i(\theta)$$

$$(2.18) \qquad = f_{G_i}(x_i) f_{G_0}(x_0) \{H(\alpha_i, \beta_i; \alpha_0, \beta_0; -\rho_2) - \beta_i C(\alpha_i + \rho_2, \beta_i; \alpha_0, \beta_0)/[2\pi(\beta_0+\beta_i)]^{1/2} + \alpha_i \Phi((\alpha_0 - \alpha_i - \rho_2)/(\beta_0+\beta_i)^{1/2})\}.$$

It follows from (2.9), (2.11), (2.12), (2.17) and (2.18) that (ii) holds. This completes the proof.

THEOREM 2. *The Bayes procedures R_{1G} and R_{2G} of our problem defined by* (i) *and* (ii) *of Theorem* 1 *are admissible.*

Proof. Since $r(\omega, d) = \int_Y L_2(d(y), \omega) f(y|\omega)\, dy$ is continuous in ω for each $d \in D^*$ (D^* denotes the set of randomized decision functions) and $R(R_{2G}, G)$ (defined by (2.7)) is finite and the support of G is the whole $(k+1)$-dimensional euclidean space, it follows from the following Lemma 1 that the result holds true. The proof for the admissibility of R_{1G} is analogous.

LEMMA 1. *Let $\Omega = R^{k+1}$ and let $r(\omega, d) = E_Y L(d(y), \omega)$ be continuous in ω for each $d \in D^*$, the set of randomized decision rules. If d_0 is Bayes with respect to the loss function $L(\cdot, \cdot)$ and a probability distribution G on Ω for which $E_G r(\omega, d_0)$ is finite, and if the support of G is the whole R^{k+1}, then d_0 is admissible.*

Proof. Suppose d_0 is not admissible. Then there exists $d' \in D^*$ such that $r(\omega, d') \leq r(\omega, d)$ for all $\omega \in \Omega$ and $r(\omega_0, d') < r(\omega_0, d_0)$ for some $\omega_0 \in \Omega$. For some $\eta > 0$ there exists an $\varepsilon > 0$ such that $r(\omega, d') \leq r(\omega, d_0) - \eta/2$ whenever $|\omega - \omega_0| < \varepsilon$. Then

$$E_G(\omega, d_0) - E_G(\omega, d') \geq \frac{\eta}{2} G(\omega_0 - \varepsilon, \omega_0 + \varepsilon) > 0,$$

where $G(\omega_0 - \varepsilon, \omega_0 + \varepsilon)$ is the probability over the cube $(\theta_0 - \varepsilon, \theta_0 + \varepsilon) \times (\theta_1 - \varepsilon, \theta_1 + \varepsilon) \times \cdots \times (\theta_k - \varepsilon, \theta_k + \varepsilon)$ with respect to G. This produces a contradiction, since d_0 is Bayes with respect to G. This completes the proof.

3. **Empirical Bayes procedures.** When the mean λ_i of the random variable associated with the prior distribution of θ_i is unknown, the Bayes procedures R_{1G} and R_{2G} are not available. This is most of the cases in practical situations. However, the observed set of data itself reveals something about λ_i. Robbins [3] first proposed the idea of the asymptotic optimal rule, which achieves the Bayes risk as the sample size increases to infinity.

Here we treat the problem from an empirical Bayes approach, using loss functions L_1 and L_2 respectively. We assume λ_i and φ_i^2 are both finite but unknown.

Let X be a random variable distributed according to the cumulative distribution function $F(x;\ \theta)$ where θ belongs to a certain parameter space, say Θ, on which a distribution G is assigned, and G may or may not be known to us. Let $d(\cdot)$ be a decision function and $L(\cdot,\ \cdot)$ be a loss function defined, respectively, on the sample spaces Z and $\mathfrak{A}\times\Theta$ respectively ($\mathfrak{A}$ is the action space). Then the risk of d is defined by

$$R(d,\ G)=\int_{\Theta}\int_{Z}L(d(z),\ \theta)\ dF(z;\ \theta)\ dG(\theta)\,.$$

The rule d_G is called a Bayes decision rule if

$$R(d_G,\ G)=\inf_{d} R(d,\ G)=R(G)\,,$$

say. We recall a definition of Robbins [3].

Let $T=\{d_n\}$ be a sequence of decision functions such that $d_n(x)=d_n(x_1,\ x_2,\cdots,\ x_n;\ x)$, which is a function of x whose form depends on the preceding n observations $x_1,\ x_2,\cdots,\ x_n$. Usually x is the present or new observation. Let

$$R_n(T,\ G)=\int_{\Theta}\int_{Z}E(L(d_n(x_1,\ x_2,\cdots,\ x_n;\ z),\ \theta))\ dF(z;\ \theta)\ dG(\theta)\,,$$

where the expectation is taken with respect to $x_1,\ x_2,\cdots,\ x_n$. If $\lim_{n\to\infty}R_n(T,\ G)=R(G)$, the Bayes risk, we say T is asymptotically optimal (a.o.) relative to G. We call such an a.o. procedure an empirical Bayes procedure. Under the formulation and notation of our problem, we restate a result of Deely [1] as follows.

LEMMA 2. *Let $X_1,\ X_2,\cdots$ be independent identically distributed random vectors with X_i having density $f(X|\omega)$ and consisting of $(k+1)$-components which are independent random variables. Let G be a cumulative distribution function on the parameter space Ω. Suppose $\overline{G}_n$ is a function of $X_1,\ X_2,\cdots,X_n$ and let d_G denote a Bayes solution with respect to G. If*

(i) *$\lim_{n\to\infty}\overline{G}_n(\omega;\ x_1,\ x_2,\cdots,\ x_n)=G(\omega)$ wp 1 for every continuity point ω of G, where probability is taken with respect to $x_1,\ x_2,\cdots,\ x_n$;*

(ii) *$L(S,\ \omega)\,f(x|\omega)$ is continuous in ω and finite with respect to G for every $S\subset K$ and $x\in Y$, the sample space;*

(iii) $\int_{\Omega} L(S, \omega)\, dG(\omega) < \infty$ *for every* $S \subset K$;

then $T = \{d_{\bar{G}_n}\}$ *is a.o. relative to* G.

For each fixed n, we recall our notation, namely that $\bar{x}_{in} = \bar{x}_i = (1/n) \sum_{j=1}^{n} x_{ij}$ and $S_{in}^2 = S_i^2 = (1/(n-1)) \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2$ and $\Phi(x; \theta, \sigma^2)$ denotes the cdf of a normal distribution with mean θ and variance σ^2. Under the formulation of our problem we have

LEMMA 3. *Let* $G = \prod_{i=0}^{k} \Phi(x_i; \lambda_i, \sigma_i^2 + \varphi_i^2)$ *and* $\bar{G}_n = \prod_{i=0}^{k} \Phi(x_i; \bar{X}_i, S_i^2)$. *Then the condition* (i) *of Lemma* 1 *is satisfied where* $\omega = (\omega_1, \omega_2)$ *with*

$$\omega_1 = \{(\lambda_0, \lambda_1, \cdots, \lambda_k);\ \lambda_i \in R\},$$

$$\omega_2 = \{(\sigma_0^2 + \varphi_0^2, \sigma_1^2 + \varphi_1^2, \cdots, \sigma_k^2 + \varphi_k^2);\ \sigma_i > 0,\ \varphi_i > 0\}.$$

Proof. First we show the case $k = 0$. By simple calculation we see that the x_{0i} are iid with unconditional normal density with mean λ_0 and variance $\sigma_0^2 + \varphi_0^2$. It is well known that X_{0n} and S_{0n} are independent for every n, that $\bar{X}_{0n} \to \lambda_0$ wp 1 and that $S_{0n} \to \sigma_0^2 + \varphi_0^2$ wp 1 by the strong law of large numbers. Since $\Phi(x; \alpha, \beta)$ is a continuous function of α and $\beta > 0$, we then have

$$\Phi(x; \bar{X}_{0n}, S_{0n}^2) \longrightarrow \Phi(x; \lambda_0, \sigma_0^2 + \varphi_0^2) \quad \text{wp } 1 \quad \forall x \in R. \tag{3.1}$$

Similarly, we also have

$$\Phi(x; \bar{X}_{in}, S_{in}^2) \longrightarrow \Phi(x; \lambda_i, \sigma_i^2 + \varphi_i^2) \quad \text{wp } 1 \quad \forall x \in R, \quad \text{for } i = 1, 2, \cdots, k. \tag{3.2}$$

Define

$$H(x_0, x_1, \cdots, x_k) = \prod_{j=0}^{k} x_j \qquad \text{for } x_i \geq 0, \quad i = 1, 2, \cdots, k.$$

It is obvious that H is a continuous function of each variable. By the usual (ε, δ)-argument and by (3.1) and (3.2) we can show that

$$\bar{G}_n = H(\bar{G}_{n0}, \bar{G}_{n1}, \cdots, \bar{G}_{nk}) \longrightarrow H(G_0, G_1, \cdots, G_k) = G \quad \text{wp } 1,$$

where

$$G_{ni} = \Phi(x; \bar{X}_{in}, S_{in}^2) \quad \text{and} \quad G_i = \Phi(x; \lambda_i, \sigma_i^2 + \varphi_i^2), \quad i = 0, 1, 2, \cdots, k.$$

This completes the proof.

THEOREM 2. *Under the formulation of our problem, the following rules* $\tilde{R}_{1n}$ *and* $\tilde{R}_{2n}$, *using loss functions* L_1 *and* L_2, *respectively, are a.o. with respect to G, where G is any normal distribution with a finite mean and a finite variance.*

Define

$$
\begin{aligned}
&\tilde{\alpha}_i = (nS_{in}^2 - (n-1)\,\sigma_i^2 \bar{X}_{in})/(nS_{in}^2 - (n-1)\,\sigma_i^2)\,, \\
&\tilde{\beta}_i = \sigma_i^2 (S_{in}^2 - \sigma_i^2)/(nS_{in}^2 - (n-1)\,\sigma_i^2) \qquad (i = 0, 1, 2, \cdots, k)\,, \\
(3.3)\quad &\tilde{f}_i = f_i(\tilde{\alpha}_i, \tilde{\beta}_i, \tilde{\alpha}_0, \tilde{\beta}_0, \rho_1)\,, \\
&\tilde{g}_i = g_i(\tilde{\alpha}_i, \tilde{\beta}_i, \tilde{\alpha}_0, \tilde{\beta}_0, \rho_2) \qquad (i = 1, 2, \cdots, k)\,,
\end{aligned}
$$

where f_i *and* g_i *are defined, respectively, by* (2.2) *and* (2.3).

(i) $\tilde{R}_{1n}$: *we classify* π_i *in* S_G *iff* $i \in S$ *for some* $S \in E$ *satisfying*

$$
\begin{aligned}
(3.4)\quad &\alpha \sum_{i \in S^c} [1 - \Phi(\tilde{J}(i, \rho_1))] + \beta \sum_{j \in S} \Phi(\tilde{J}(j, \rho_2)) \\
&= \min_{T \in E} \left\{ \alpha \sum_{i \in T^c} [1 - \Phi(\tilde{J}(i, \rho_1))] + \beta \sum_{j \in T} \Phi(\tilde{J}(j, \rho_2)) \right\},
\end{aligned}
$$

where

$$
(3.5)\quad \tilde{J}(i, \rho) = (\tilde{\alpha}_0 - \tilde{\alpha}_i + \rho)/(\tilde{\beta}_0 + \tilde{\beta}_i)^{1/2} \qquad (i = 1, 2, \cdots, k)\,,
$$

where $\tilde{\alpha}_j$ *and* $\tilde{\beta}_j$ *are defined by* (3.3).

(ii) $\tilde{R}_{2n}$: $\pi_i \in S_G$ *iff* $i \in S$ *for some* $S \in E$ *satisfying*

$$
\alpha \sum_{i \in S^c} \tilde{f}_i + \beta \sum_{j \in S} \tilde{g}_j = \min_{T \in E} \left(\alpha \sum_{i \in T^c} \tilde{f}_i + \beta \sum_{j \in T} \tilde{g}_j \right).
$$

Proof. By Theorem 1, and Lemmas 2 and 3, it suffices to check the conditions (ii) and (iii) of Lemma 2. It is obvious that

$$
L_1(S, \omega)\, f(x|\omega) \le (\alpha + \beta)\,(1/2\pi)^{(k+1)/2} \left(\prod_{i=0}^{k} \sigma_i^{-n} \right) < \infty
$$

$$
\forall x \in R^{n(k+1)} \text{ and } S \subset K.
$$

Hence

$$
\int_{\Omega_1} L_1(S, \omega)\, dG(\omega) < \infty\,.
$$

We also note that

$$
\int_{\Omega} L_2(S, \omega)\, dG(\omega) \le 2(\alpha + \beta)\, E_G(|\theta|) < \infty,
$$

since $G_i(\theta)$ has finite absolute moment. Finally, we note that

$$L_2(S, \omega) f(x|\omega) \leq (\alpha + \beta) \max_{1 \leq i \leq k} |\theta_i - \theta_0| (1/2\pi)^{(k+1)/2} \left(\prod_{i=0}^{k} \sigma_i^{-n}\right) < \infty$$

with respect to G. This completes the proof.

Acknowledgment. The author would like to thank the referees for their very helpful suggestions and useful comments.

References

1. J. J. Deely, *Multiple decision procedures from an empirical Bayes approach*, Ph. D. Thesis, Department of Statistics, Purdue University, West Lafayette, Indiana, 1965.
2. W. T. Huang, *On partitioning k multivariate normal populations with respect to a control*, Bull. Inst. Math. Acad. Sinica **1** (1973), 191–266.
3. H. Robbins, *An empirical Bayes approach to statistics*, Proc. Third Berkeley Sympos. on Statist. and Prob. 1954/55, vol. 1, pp. 157–164.
4. Y. I. Tong, *On partitioning a set of normal populations by their locations with respect to a control*, Ann. Math. Statist. **40** (1969), 1300–1324.

INSTITUTE OF MATHEMATICS, ACADEMIA SINICA, TAIPEI