Ann. Inst. Statist. Math.
37 (1985), Part A, 415-435

OPTIMAL CONSTRUCTION OF A SELECTION OF A SUBPOPULATION

KAZUO NODA

(Received Jan. 30, 1978; revised Apr. 10, 1985)

Summary

The problem of selecting a subpopulation from a given population Π is to be, on the basis of measurements of members of Π, achieved by choosing those members of Π who satisfy the standards determined by a given selection criterion and rejecting those who do not.

Since the optimum selection depends on the unknown parameter of the probability distribution of Π, it is here considered how to construct a decision function from the space of subsidiary sample having information on θ to the space of selections. Thus the existence of Bayes and minimax decision functions under the constraint defined by the selection criterion is proved. A necessary and sufficient condition for a decision function satisfying the constraint to be a Bayes decision function is also obtained.

1. Introduction

The problem of selecting a subpopulation from a given population have been studied by Birnbaum and Chapman [1], Cochran [3], Raj [10], and the others. A common feature in the selection problem treated in their works may be described in the following way. Consider a population Π whose members may be identified by realized values of a random variable (x, y) distributed according to a probability distribution P_θ with an unknown parameter θ (wherein the form of P_θ is known). Values of y characterize members of the population Π more suitably than those of x but they are to appear in the future (and cannot be measured at the time of selection). On the other hand, each value of the variable x is observable at the time of selection and distributed according to a known marginal distribution. Thus selecting a subpopulation Π' of individuals from Π must indirectly be achieved on the basis of observed values of x under a criterion given in connection with y and the parameter θ. This situation may be seen, for instance, in

Key words: Bayes and minimax decision functions, optimum selection, subpopulation.

personal selection, choice of a sire for breeding, and testing of materials. In personal selection, each value of x represents an applicant's score in an admission test for an establishment and each value of y is to appear as his achievement score after his entrance to the establishment. To get a feeling of this matter, we present here the selection criterion introduced by Cochran [3]. Let $f_\theta(x, y)$ be the joint frequency function of Π with its marginal frequency $g(x)$ assumed to be known, and α some constant satisfying a condition $0<\alpha<1$. Let ϕ be a measurable function such that $0\leqq\phi(x)\leqq 1$. We shall call such a function ϕ a selection (see [1]). The problem is to select a subpopulation Π' with a density $(\phi/\alpha)f$ which maximizes

$$\int y\phi(x)f_\theta(x, y)/\alpha dxdy \tag{1.1}$$

under the constraint

$$\int \phi(x)f_\theta(x, y)dxdy \qquad \left(=\int \phi(x)g(x)dx\right)=\alpha\ . \tag{1.2}$$

Thus, this problem requires us to find a selection such that the mean value of y becomes maximum in Π' selected under the constraint that the size of Π' is fixed to the given fraction α. We shall call such a selection optimum (see [1]). In his paper, "construction" of a preferable selection as a substitute for the optimum selection is also considered on the basis of a subsidiary sample, since the optimum selection depends on the unknown parameter θ.

In the present paper we consider a statistical decision theory of selections on the basis of a supplementary information on the joint probability distribution P_θ of x and y. We regard the term "construction" introduced by Cochran as a decision function from the space of these sample to the space of selections. In his paper it is insisted that any rule for constructing a best sample index can not be found (see Section 9 in [3]). This statement shows that it is difficult to find the optimality of each "construction" of a selection. On the contrary, our method provides a solution of the problem as to how we establish the optimalities of these decision functions. In this context our purpose is to obtain a necessary and sufficient condition for a decision function under the constraint to be a Bayes decision function.

We show, in Section 2, a fairly general formulation in order to discuss simultaneously the selection criteria introduced by the several authors as stated at the beginning. Each of the selection criteria is there obtained by specifying the form of a finite sequence $\{\psi_i(y, \theta);\ i=0, 1, \cdots, m\}$ given in connection with a constraint and an objective function (see (2.2) and (2.3)). In Section 3 we demonstrate the existence of Bayes and minimax decision functions satisfying the constraint (in Theorem 3.1 and Theorem 3.3). We further present a necessary and sufficient condition for a decision function to be such a Bayes decision function (in Theorem 3.2), using a Lagrange multiplier method in topological linear spaces (see [5]).

2. A decision theoretic approach to the construction problem of selections

Let $(\mathcal{X}\times\mathcal{Y}, \mathcal{A}\times\mathcal{B})$ and (Θ, Λ) be measurable spaces standing for a sample space and a parameter space, respectively. Let $\mathcal{P}=\{P_\theta: \theta\in\Theta\}$ be a family of probability distributions on $(\mathcal{X}\times\mathcal{Y}, \mathcal{A}\times\mathcal{B})$ each of which has a known common marginal distribution P^x on $(\mathcal{X}, \mathcal{A})$ (not depending on θ). Let P_θ^y denote the marginal distribution of P_θ on $(\mathcal{Y}, \mathcal{B})$. Let $\mathcal{R}$ denote the real line. Let Φ be the family of all $\mathcal{A}$-measurable functions ϕ on $\mathcal{X}$ such that $0\leqq\phi(x)\leqq 1$ for every $x\in\mathcal{X}$. A member of Φ is called a selection and Φ is called the space of selections. The probability distribution of the subpopulation Π' selected by a selection ϕ is defined as $[\phi(x)/\alpha]P_\theta$, where $\alpha=\int_{\mathcal{X}}\phi(x)P^x(dx)$. Let ψ_i $(i=0, 1,\cdots, m)$ be given real-valued functions on $\mathcal{Y}\times\Theta$ which are $\mathcal{B}\times\Lambda$-measurable and P_θ-integrable for each $\theta\in\Theta$. Let

$$\tau_i(\theta, \phi)=\int_{\mathcal{X}\times\mathcal{Y}}\psi_i(y, \theta)\phi(x)P_\theta(d(x, y))\,, \qquad (\theta, \phi)\in\Theta\times\Phi,\ (i=0, 1,\cdots, m)\,. \tag{2.1}$$

For given numbers $\alpha_1,\cdots,\alpha_m$ and for each fixed $\theta\in\Theta$, we shall denote by C_θ the constraint imposed on $\phi\in\Phi$ which is expressed as

$$\tau_i(\theta, \phi)\leqq\alpha_i \quad (\text{or } =\alpha_i) \qquad (i=1,\cdots, m)\,. \tag{2.2}$$

Further, we shall, for each θ, denote by $\Phi(C_\theta)$ the subset of all selections satisfying the constraint C_θ. We assume that $\Phi(C_\theta)$ is not empty. Then, under the constraint C_θ the minimization of the objective function defined by

$$\tau_0(\theta, \phi)=\int_{\mathcal{X}\times\mathcal{Y}}\psi_0(y, \theta)\phi(x)P_\theta(d(x, y)) \tag{2.3}$$

comes into question. Hence we consider:

DEFINITION 2.1. A selection ϕ_θ will be called optimum, if ϕ_θ satisfies, for each $\theta\in\Theta$, the constraint C_θ and the equality $\tau_0(\theta, \phi_\theta)=\inf\{\tau_0(\theta, \phi)\,|\,\phi\in\Phi(C_\theta)\}$. We shall arbitrarily take one of these optimum selections and denote it by ϕ_θ^* as a particular version.

We here consider the optimum selection ϕ_θ^* as a function of (x, θ) which becomes a selection whenever θ is fixed. Though the existence of an optimum selection can be proved under some regularity conditions (see [1], [3], [10] and [9] as specified cases), we do not discuss it for brevity. We shall only assume in the next section that there exists such an $\mathcal{A}\times\Lambda$-measurable optimum selection ϕ_θ^*.

We next consider the construction problem of selection when θ is unknown. We take the nth product space $(\mathcal{X}\times\mathcal{Y})^n$ associated with the product $(\mathcal{A}\times\mathcal{B})^n$ as the space of subsidiary samples s distributed independently of x and denote their combination by $(\mathcal{S}, \mathcal{U})$. Let P_θ^n denote the nth product of P_θ which stands for the probability distribution of s when θ is "true". Let E_θ denote the expectation with respect to P_θ^n. Let $\hat{\Phi}$ be the space of all $\mathcal{A}\times\mathcal{U}$-measurable functions $\hat{\phi}$ on $\mathcal{X}\times\mathcal{S}$ such that $0\leqq\hat{\phi}(x, s)\leqq 1$. Further, let $\hat{\Phi}(\hat{C})$ be the subset of all $\hat{\phi}$ satisfying the constraint $\hat{C}$ defined by either

$$\mathrm{E}_\theta\, \tau_i(\theta, \hat{\phi})\leqq\alpha_i+\beta_{in}(\theta) \qquad (i=1,\cdots, m) \tag{2.4}$$

or

$$\mathrm{E}_\theta\, \tau_i(\theta, \hat{\phi})=\alpha_i+\beta_{in}(\theta) \qquad (i=1,\cdots, m) \tag{2.4'}$$

for all $\theta\in\Theta$, where each β_{in} is a Λ-measurable real-valued function on Θ with the property that $\beta_{in}(\theta)\rightarrow 0$ $(n\rightarrow\infty)$ for every θ. Here $\mathrm{E}_\theta\,\tau_i(\theta, \hat{\phi})$ is, for every $\theta\in\Theta$ and for every $\hat{\phi}\in\hat{\Phi}$, expressed as

$$\mathrm{E}_\theta\, \tau_i(\theta, \hat{\phi})=\int_{\mathcal{S}}\int_{\mathcal{X}\times\mathcal{Y}}\psi_i(y, \theta)\hat{\phi}(x, s)P_\theta(d(x, y))P_\theta^n(ds) \qquad (i=0, 1,\cdots, m)\,. \tag{2.5}$$

We assume that $\hat{\phi}(\hat{C})$ is not empty.

Each element $\hat{\phi}(\cdot, s)$ of $\hat{\Phi}$ is regarded as a (non-randomized) decision function from $\mathcal{S}$ to Φ standing for the construction of a selection when a sample s is given. Here, since L is convex-linear in ϕ as is seen in (2.6), it is sufficient for this formulation only to consider non-randomized decision functions (see [4]). On the other hand, Φ can be regarded as a decision space. The loss function L of $\phi\in\Phi$ is defined by

$$L(\theta, \phi)=\int_{\mathcal{X}\times\mathcal{Y}}\psi_0(y, \theta)\phi(x)P_\theta(d(x, y)) -\int_{\mathcal{X}\times\mathcal{Y}}\psi_0(y, \theta)\phi_\theta^*(x)P_\theta(d(x, y))\,, \qquad \theta\in\Theta\,. \tag{2.6}$$

Hence the risk function r of $\hat{\phi}\in\hat{\Phi}$ is, under (2.5), defined as

$$r(\theta, \hat{\phi}) = \mathrm{E}_\theta L(\theta, \hat{\phi}) , \qquad \theta \in \boldsymbol{\Theta} .$$

Example 2.1. Let $\mathcal{X}=\mathcal{R}^q$, $\mathcal{Y}=\mathcal{R}$ and $\boldsymbol{\Theta}$ be an open set of $\mathcal{R}^p$. Let $\mathcal{A}$, $\mathcal{B}$ and Λ be their Borel σ-algebras, respectively. As $m=1$, let $\psi_1(y, \theta)=1$ and $\psi_0(y, \theta)=-y/\alpha_1$, where $0<\alpha_1<1$. Then, we have the criterion expressed by (1.1) and (1.2) in the previous section: Under the constraint

$$\tau_1(\theta, \phi) = \int_{\mathcal{X}} \phi(x) P^x(dx) = \alpha_1 , \tag{2.7}$$

consider how the objective function

$$\tau_0(\theta, \phi) = -\int_{\mathcal{X}\times\mathcal{Y}} y \frac{\phi(x)}{\alpha_1} P_\theta(d(x, y)) \tag{2.8}$$

is minimized.

If a sample $s \in \mathcal{S}$ is given, then for a decision function $\hat{\phi}$ satisfying $\mathrm{E}_\theta \tau_1(\theta, \hat{\phi})=\alpha_1$ the risk function $r(\theta, \hat{\phi})=\mathrm{E}_\theta[\tau_0(\theta, \hat{\phi})-\tau_0(\theta, \phi_\theta^*)]$ becomes a measurement examining how $\hat{\phi}$ is good. In other words, we set $\beta_{1n}(\theta) \equiv 0$ in this case.

Example 2.2. Let $\mathcal{X}=\mathcal{R}^q$, $\mathcal{Y}=\mathcal{R}$, and $\boldsymbol{\Theta}$ be an open set of $\mathcal{R}^p$. Let $\mathcal{A}$, $\mathcal{B}$ and Λ be their Borel σ-algebras, respectively. Here $\theta=(\theta_1, \theta_{(2)})$, where $\theta_{(2)}=(\theta_2, \cdots, \theta_{p-1})$ and θ_1 stands for the mean $\int_{\mathcal{Y}} y P_\theta^y(dy)$. As $m=2$, let $\psi_1(y, \theta)=1$, $\psi_2(y, \theta)=(y-\theta_1)/\alpha_1$ and $\psi_0(\theta, y)=(y-\theta_1-\alpha_2)^2/\alpha_1$, where $0<\alpha_1<1$.

Then we have a selection criterion: Under the constraint defined by

$$\begin{cases} \tau_1(\theta, \phi) = \int_{\mathcal{X}} \phi(x)\, P^x(dx) = \alpha_1 , \\ \tau_2(\theta, \phi) = \int_{\mathcal{X}\times\mathcal{Y}} (y-\theta_1) \frac{\phi(x)}{\alpha_1}\, P_\theta(d(x, y)) = \alpha_2 , \end{cases} \tag{2.9}$$

consider how the objective function

$$\tau_0(\theta, \phi) = \int_{\mathcal{X}\times\mathcal{Y}} [y-(\theta_1+\alpha_2)]^2 \frac{\phi(x)}{\alpha_1} P_\theta(d(x, y)) \tag{2.10}$$

is minimized.

This problem requires us to find a selection such that the variance of y becomes minimum in the subpopulation Π' selected under the constraint that the size of Π' is fixed to the fraction α_1 and the difference between the mean of y in Π' and the mean of y in the original population Π is fixed to the given number α_2.

3. On Bayes and minimax decision functions

We first consider Bayes decision functions under the following conditions:

(A1) Each value $P_\theta(E)$, $E \in \mathcal{A}\times\mathcal{B}$, is Λ-measurable in θ.

(A2) There exists a regular conditional probability distribution $P_\theta^{y|x}$ on $(\mathcal{Y}, \mathcal{B})$ given x such that for every θ and for every P_θ-integrable real-valued function f on $\mathcal{X}\times\mathcal{Y}$ the equality

$$\int_{\mathcal{X}\times\mathcal{Y}} f(x, y)P_\theta(d(x, y)) = \int_{\mathcal{X}}\int_{\mathcal{Y}} f(x, y)P_\theta^{y|x}(dy\,|\,x)P^x(dx)$$

holds and $P_\theta^{y|x}(B\,|\,x)$ is, for each fixed B, $\mathcal{A}\times\Lambda$-measurable.

Using (A2), we write

$$\bar{\varphi}_i(x, \theta) = \int_{\mathcal{Y}} \varphi_i(y, \theta)P_\theta^{y|x}(dy\,|\,x)\ , \qquad (x, \theta) \in \mathcal{X}\times\Theta \quad (i=0, 1, \cdots, m)\ . \tag{3.1}$$

Evidently, each $\bar{\varphi}_i$ is $\mathcal{A}\times\Lambda$-measurable. Hence, τ_i can be expressed as

$$\tau_i(\theta, \phi) = \int_{\mathcal{X}} \bar{\varphi}_i(x, \theta)\phi(x)P^x(dx) \qquad (i=0, 1, \cdots, m)\ . \tag{3.2}$$

Let ξ be a prior probability distribution on (Θ, Λ) with the property that for each i, $\bar{\varphi}_i$ is $P^x\times\xi$-integrable and β_{in} is ξ-integrable. Then, the Bayes risk function $\bar{r}(\xi, \hat{\phi})$ of $\hat{\phi}$ is expressed as

$$\begin{aligned}\bar{r}(\xi, \hat{\phi}) &= \int_\Theta r(\theta, \hat{\phi})\xi(d\theta)\\ &= \int_\Theta\int_{\mathcal{S}}\int_{\mathcal{X}} \bar{\varphi}_0(x, \theta)\hat{\phi}(x, s)P^x(dx)P_\theta^n(ds)\xi(d\theta)\\ &\quad -\int_\Theta\int_{\mathcal{X}} \bar{\varphi}_0(x, \theta)\phi_\theta^*(x)P^x(dx)\xi(d\theta)\ .\end{aligned} \tag{3.3}$$

Since the second term on the third expression in (3.3) is a constant, it is sufficient for Bayes decision functions to consider minimizing the first term independently of the second term. Hence, we write

$$\tilde{r}(\xi, \hat{\phi}) = \int_\Theta\int_{\mathcal{S}}\int_{\mathcal{X}} \bar{\varphi}_0(x, \theta)\hat{\phi}(x, s)P^x(dx)P_\theta^n(ds)\xi(d\theta)\ . \tag{3.4}$$

We here define, under (A1), the probability distribution $\bar{P}_\xi^n$ on $(\mathcal{S}\times\Theta, \mathcal{U}\times\Lambda)$ by

$$\bar{P}_\xi^n(E) = \int_\Theta\int_{\mathcal{S}} \chi_E P_\theta^n(ds)\xi(d\theta)\ , \qquad E \in \mathcal{U}\times\Lambda\ , \tag{3.5}$$

where χ_E denotes the indicator function of $E \in \mathcal{U}\times\Lambda$. Let $\bar{P}^n$ be the marginal probability distribution on $(\mathcal{S}, \mathcal{U})$ generated by $\bar{P}_\xi^n$. We also assume:

(A3) For $\bar{P}_\xi^n$ there exists a regular posterior probability distribution $\bar{\xi}(\cdot|s)$ on (Θ, Λ) given s such that for every $\bar{P}_\xi^n$-integrable real-valued function f on $\mathcal{S}\times\Theta$

$$\int_{\mathcal{S}\times\Theta} f(s, \theta)\bar{P}_\xi^n(d(s, \theta))=\int_{\mathcal{S}}\int_\Theta f(s, \theta)\bar{\xi}(d\theta\,|\,s)\bar{P}^n(ds)$$
$$=\int_\Theta\int_{\mathcal{S}} f(s, \theta)P_\theta^n(ds)\xi(d\theta)\ .$$

Since each $\hat{\phi}\in\hat{\Phi}$ is $\mathcal{A}\times\mathcal{U}$-measurable, (3.4) is, under (A1), (A2) and (A3), rewritten as

$$\tilde{r}(\xi, \hat{\phi})=\int_{\mathcal{X}\times\mathcal{S}} \bar{\bar{\varphi}}_0(x, s)\hat{\phi}(x, s)P^x(dx)\bar{P}^n(ds)\ , \tag{3.4'}$$

where

$$\bar{\bar{\varphi}}_0(x, s)=\int_\Theta \bar{\varphi}_0(x, \theta)\bar{\xi}(d\theta\,|\,s)\ , \qquad s\in\mathcal{S}\ . \tag{3.6}$$

To confirm the existence of a Bayes decision function, we further assume the following conditions.

(A4) Each of the σ-algebras $\mathcal{A}$ and $\mathcal{B}$ has a countable number of generators.

(A5) The parameter space Θ can be topologized by a topology $\mathcal{T}(\Theta)$ for which the following properties hold:

1° The first axiom of countability holds.

2° The family Λ contains the Borel σ-algebra.

3° The measure ξ is regular and $\xi(E)>0$ for every non-void open set E of Θ.

4° For each $\theta\in\Theta$

$$\sup_{E\in\mathcal{A}\times\mathcal{B}} |P_{\theta'}(E)-P_\theta(E)|\to 0 \qquad (\theta'\to\theta)\ . \tag{3.7}$$

5° For each i and for each fixed y, $\varphi_i(y, \theta)$ is continuous in θ. Further, for each $\theta\in\Theta$ there exists a neighborhood $V(\theta)$ of θ and a positive-valued function K_θ on $\mathcal{Y}$ which satisfies

$$|\varphi_i(y, \theta')|\leqq K_\theta(y) \qquad (i=0, 1\cdots, m) \tag{3.8}$$

for all $(y, \theta')\in\mathcal{Y}\times V(\theta)$ and is $P_{\theta'}^y$-integrable uniformly in $\theta'\in V(\theta)$ in the sense that

$$\lim_{c\to\infty}\int_{K_\theta(y)\geqq c} K_\theta(y)P_{\theta'}^y(dy)=0 \tag{3.9}$$

uniformly in $\theta'\in V(\theta)$.

6° For each i, β_{in} is continuous in θ.

LEMMA 3.1. *Under the conditions* (A2) *and* 1°, 4° *and* 5° of (A5), *each* $\mathrm{E}_\theta \tau_i(\theta, \hat{\phi})$ $(i=0, 1, \cdots, m)$ *is, for each fixed* $\hat{\phi} \in \hat{\Phi}$, *continuous in* θ *as a real-valued function of* θ.

PROOF. We first show that each $\tau_i(\theta, \hat{\phi})$ is, for each fixed $s \in \mathcal{S}$, continuous in θ. Because of (A5) 1°, it is sufficient for the proof to show that for each s, for each θ_0 and for each sequence $\{\theta_k\}$ converging to θ_0 the sequences $\{\tau_i(\theta_k, \hat{\phi})\}$ converge to $\tau_i(\theta_0, \hat{\phi})$ as $k \to \infty$. Now let s be fixed. There exists a probability measure Q on $(\mathcal{Y}, \mathcal{B})$ such that for each θ_k $(k=0, 1, 2, \cdots)$ $P^y_{\theta_k}$ is dominated by Q (see Appendix in [7]). We denote by dP_k/dQ $(k=0, 1, 2, \cdots)$ the Radon-Nikodym derivative of $P^y_{\theta_k}$ with respect to Q. From (3.7) it follows that

$$\sup_{B \in \mathcal{B}} |P^y_{\theta'}(B) - P^y_{\theta}(B)| \to 0 \qquad (\theta' \to \theta). \tag{3.10}$$

On the other hand, we have by (A5) 5° and the Lebesgue convergence theorem (see [11], for instance) that for each i

$$\int_{\mathcal{Y}} |\psi_i(y, \theta_k) - \psi_i(y, \theta_0)| P^y_{\theta_0}(dy) \to 0 \qquad (k \to \infty). \tag{3.11}$$

Further, we have by (3.10) and (A5) 5° that for each i and for each $\varepsilon > 0$ there exists a constant $c > 0$ such that

$$\begin{aligned} &\int_{\mathcal{Y}} |\psi_i(y, \theta_k)| \left| \frac{dP_k}{dQ} - \frac{dP_0}{dQ} \right| Q(dy) \\ &\leqq \int_{\mathcal{Y}} K_{\theta_0}(y) \left| \frac{dP_k}{dQ} - \frac{dP_0}{dQ} \right| Q(dy) \\ &\leqq \int_{K_{\theta_0}(y) \geqq c} K_{\theta_0}(y) P^y_{\theta_k}(dy) + \int_{K_{\theta_0}(y) \geqq c} K_{\theta_0}(y) P^y_{\theta_0}(dy) \\ &\quad + c \int_{\mathcal{Y}} \left| \frac{dP_k}{dQ} - \frac{dP_0}{dQ} \right| Q(dy) < \varepsilon \end{aligned} \tag{3.12}$$

holds for large k. Hence, it follows from (3.11) and (3.12) that

$$\begin{aligned} &|\tau_i(\theta_k, \hat{\phi}) - \tau_i(\theta_0, \hat{\phi})| \\ &= \left| \int_{\mathcal{X} \times \mathcal{Y}} \psi_i(y, \theta_k) \hat{\phi}(x, s) P_{\theta_k}(d(x, y)) \right. \\ &\qquad \left. - \int_{\mathcal{X} \times \mathcal{Y}} \psi_i(y, \theta_0) \hat{\phi}(x, s) P_{\theta_0}(d(x, y)) \right| \\ &\leqq \int_{\mathcal{Y}} |\psi_i(y, \theta_k)| \cdot \left| \frac{dP_k}{dQ} - \frac{dP_0}{dQ} \right| Q(dy) \\ &\quad + \int_{\mathcal{Y}} |\psi_i(y, \theta_k) - \psi_i(y, \theta_0)| P^y_{\theta_0}(dy) \to 0 \qquad (k \to \infty). \end{aligned} \tag{3.13}$$

Thus, $\tau_i(\theta, \hat{\phi})$ is, for each fixed $s \in \mathcal{S}$, continuous in θ.

We here notice that the convergence shown in (3.13) is uniform in $s \in \mathcal{S}$. Also we have by (3.7) that

$$\sup_{D \in \mathcal{U}} |P_{\theta'}^n(D) - P_\theta^n(D)| \to 0 \qquad (\theta' \to \theta) .$$

Therefore, by the similar method to deriving (3.13) we can also have that for each $\hat{\phi} \in \hat{\Phi}$

$$\mathrm{E}_{\theta_k} \tau_i(\theta_k, \hat{\phi}) \to \mathrm{E}_{\theta_0} \tau_i(\theta_0, \hat{\phi}) \qquad (k \to \infty) .$$

Thus this result completes the proof of Lemma 3.1.

DEFINITION 3.1. Let $(\mathcal{Z}, \mathcal{H})$ be a measurable space associated with a σ-finite measure μ. Let $B(\mathcal{Z})$ (for short, B) be the real linear space of all μ-essentially bounded μ-measurable real-valued functions on $\mathcal{Z}$, and let $L(\mathcal{Z})$ (for short, L) be that of all μ-integrable real-valued functions on $\mathcal{Z}$. We denote by $\mathcal{T}(B, L)$ and $\mathcal{T}(L, B)$ that weak topologies (see [2] or [6] for the general definition) on $B(\mathcal{Z})$ and $L(\mathcal{Z})$ which are both determined by the natural paring

$$\langle f, g \rangle = \int_{\mathcal{Z}} f(z) g(z) \mu(dz) , \qquad f \in B,\ g \in L . \tag{3.14}$$

In other words, $\mathcal{T}(B, L)$ (or $\mathcal{T}(L, B)$) is the weakest topology (see [2] for its existence) among topologies on $B(\mathcal{Z})$ (or $L(\mathcal{Z})$) each of which makes, for every $g \in L(\mathcal{Z})$ (or $f \in B(\mathcal{Z})$), the mapping $\langle \cdot, g \rangle$ from $B(\mathcal{Z})$ (or $\langle f, \cdot \rangle$ from $L(\mathcal{Z})$) to $\mathcal{R}$ continuous.

We now pair the real linear spaces $\mathcal{V}_1 = B(\mathcal{X} \times \mathcal{S})$ and $\mathcal{V}_2 = L(\mathcal{X} \times \mathcal{S})$ by the bilinear functional defined in (3.14) with respect to the probability measure $P^x \times \bar{P}^n$. We assign the topology $\mathcal{T}(\mathcal{V}_1, \mathcal{V}_2)$ to $\mathcal{V}_1$. Let $\tilde{\phi} = \{\tilde{\phi}\colon \tilde{\phi} \in \mathcal{V}_1,\ 0 \leqq \tilde{\phi}(x, s) \leqq 1$ for $P^x \times \bar{P}^n$—a.e. $(x, s)\}$.

LEMMA 3.2. *Assume the conditions* 2°, 3° *and* 4° *of* (A5). *If* $\tilde{\phi}_1$ *and* $\tilde{\phi}_2$ *are two versions of the same member of* $\tilde{\Phi}$, *that is,* $\tilde{\phi}_1(x, s) = \tilde{\phi}_2(x, s)$ *for* $P^x \times \bar{P}^n$—a.e. (x, s), *then for every* $\theta \in \Theta$

$$\tilde{\phi}_1(x, s) = \tilde{\phi}_2(x, s) \qquad \text{for } P^x \times P_\theta^n\text{—a.e. } (x, s) .$$

PROOF. Let $N = \{(x, s) \mid \tilde{\phi}_1(x, s) \neq \tilde{\phi}_2(x, s)\}$. Suppose that there exists some $\theta_0 \in \Theta$ such that $(P^x \times P_{\theta_0}^n)(N) > 0$. Then it follows from (3.7) that there exists a neighborhood $V(\theta_0)$ of θ_0 such that for every $\theta \in V(\theta_0)$, $(P^x \times P_\theta^n)(N) > 0$. Hence, as 2° and 4° of (A5) imply the Λ-measurability of $(P^x \times P_\theta^n)(N)$, it follows from (3.5) and (A5) 2° and 3° that

$$(P^x \times \bar{P}^n)(N) = \int_\Theta \int_{\mathcal{X} \times \mathcal{S}} \chi_{N \times \Theta} (P^x \times P_\theta^n)(d(x, s)) \xi(d\theta)$$

$$\geqq \int_{V(\theta_0)} [(P^x \times P_\theta^n)(N)]\xi(d\theta) > 0\,,$$

where $\chi_{N\times\Theta}$ denotes the indicator function of the set $N\times\Theta$, and the proof is complete.

By Lemma 3.2 each $\mathrm{E}_\theta\,\tau_i(\theta, \tilde{\phi})$ $(i=0, 1, \cdots, m)$ is well-defined for all $\tilde{\phi}\in\tilde{\Phi}$ when $\hat{\phi}$ is replaced by $\tilde{\phi}$ in (2.5). Hence we denote by $\tilde{\phi}(\hat{C})$ the set $\{\tilde{\phi}\,|\,\tilde{\phi}\in\tilde{\Phi},\ \mathrm{E}_\theta\,\tau_i(\theta, \tilde{\phi})\leqq\alpha_i+\beta_{in}(\theta)$ (or $=\alpha_i+\beta_{in}(\theta)$) for all $\theta\in\Theta$ $(i=1, 2, \cdots, m)\}$.

As we have seen the Λ-measurability of $P_\theta^n(E)$, $E\in\mathcal{U}$, in the proof of Lemma 3.2, we here notice that (A1) follows from (A5) 2° and 4°. Therefore, if (A5) is assumed, then we can dispense with (A1). We now have:

THEOREM 3.1. *Under the conditions stated in* (A2) *through* (A5), *there exists a Bayes decision function* $\hat{\phi}_\xi\in\hat{\Phi}(\hat{C})$ *minimizing* $\tilde{r}(\xi, \hat{\phi})$ *under the constraint* $\hat{C}$ *defined by* (2.4) *or* (2.4').

PROOF. From Lemma 3.2 it follows that for each $\tilde{\phi}\in\tilde{\Phi}$ there exists a $\hat{\phi}\in\hat{\Phi}$ such that $\mathrm{E}_\theta\,\tau_i(\theta, \tilde{\phi})=\mathrm{E}_\theta\,\tau_i(\theta, \hat{\phi})$ and $\tilde{r}(\theta, \tilde{\phi})=\tilde{r}(\theta, \hat{\phi})$ hold, where $\tilde{r}(\xi, \tilde{\phi})$ is defined by (3.4) if in the expression $\hat{\phi}$ is replaced by $\tilde{\phi}$. Therefore, it is sufficient for the proof to consider $\tilde{\Phi}$ as a substitute for $\hat{\Phi}$ in the formulation of the problem. Also, let $\tilde{\Phi}(\tilde{C})=\{\tilde{\phi}\,|\,\tilde{\phi}\in\tilde{\Phi},\ \mathrm{E}_\theta\,\tau_i(\theta, \tilde{\phi})\leqq\alpha_i+\beta_{in}(\theta)$ (or $=\alpha_i+\beta_{in}(\theta)$) for ξ—a.a. θ $(i=1, 2, \cdots, m)\}$. Clearly, $\tilde{\Phi}(\hat{C})\subset\tilde{\Phi}(\tilde{C})$. We show that its converse also holds. For each $\tilde{\phi}\in\tilde{\Phi}(\tilde{C})$ let $\Theta(\tilde{\phi})=\{\theta\,|\,\theta\in\Theta,\ \mathrm{E}_\theta\,\tau_i(\theta, \tilde{\phi})>\alpha_i+\beta_{in}(\theta)$ (or $\neq\alpha_i+\beta_{in}(\theta)$) $(i=1, 2, \cdots, m)\}$. Then, it follows from (A5) 6° and Lemma 3.1 that $\Theta(\tilde{\phi})$ is open with respect to $\mathcal{T}(\Theta)$. Therefore, (A5) 2° and 3° and the definition of $\tilde{\Phi}(\tilde{C})$ imply that $\Theta(\tilde{\phi})$ is empty, and hence $\tilde{\Phi}(\tilde{C})\subset\tilde{\Phi}(\hat{C})$ as well. Thus, we can consider the minimization of $\tilde{r}(\xi, \tilde{\phi})$ under the constraint $\tilde{C}$ instead of $\hat{C}$.

We now regard each $\mathrm{E}.\,\tau_i(\cdot, \tilde{\phi})$ $(i=1, 2, \cdots, m)$ as a mapping from $\tilde{\Phi}$ to $L(\Theta, \Lambda)$ with respect to the probability measure ξ. The conditions (A2), (A3) and (A5) imply that for every $h_i\in B(\Theta)$ with respect to ξ

$$\int_\Theta h_i(\theta)\,\mathrm{E}_\theta\,\tau_i(\theta, \tilde{\phi})\xi(d\theta)=\int_{\mathcal{X}\times\mathcal{S}}\left[\int_\Theta h_i(\theta)\bar{\varphi}_i(x, \theta)\bar{\xi}(d\theta\,|\,s)\right]\tilde{\phi}(x, s)P^x(dx)\bar{P}^n(ds)$$

holds and $\int_\Theta h_i(\theta)\bar{\varphi}_i(x, \theta)\bar{\xi}(d\theta\,|\,s)$ is contained in $\mathcal{V}_2$ as a function of (x, s). Hence, $\mathrm{E}.\,\tau_i(\cdot, \tilde{\phi})$ is a continuous mapping. Since (A4) implies that $\tilde{\Phi}$ is sequentially compact (see Appendix in [7]), it follows from the continuity of $\mathrm{E}.\,\tau(\cdot, \tilde{\phi})$ and the closedness of the set $\{h\,|\,h\in B(\Theta),\ h(\theta)\leqq\alpha_i+\beta_{in}(\theta)$ (or $=\alpha_i+\beta_{in}(\theta)$) for ξ—a.e. $\theta\}$ that $\tilde{\Phi}(\tilde{C})$ is also sequentially compact.

Similarly, we can see that $\tilde{r}(\xi, \cdot)$ is continuous when we regard it as a mapping from $\tilde{\Phi}$ to $\mathcal{R}$. Thus, there exists a $\tilde{\phi}_\xi \in \tilde{\Phi}(\tilde{C})$ minimizing $\tilde{r}(\xi, \tilde{\phi})$, which gives the assertion of this theorem.

To obtain a necessary and sufficient condition for a decision function to be a Bayes decision function, we apply Theorem 2 [5] to our problem. For convenience, we here present this theorem as the following lemma (the notions stated there will be defined immediately after the description of the lemma):

LEMMA 3.3 (Isii [5]). *Let $\mathcal{X}$ be a nonempty convex set of a real linear space, and $\mathcal{W}$ be a locally convex linear topological space preordered by a closed convex cone C. Let f be a concave mapping from $\mathcal{X}$ to $\mathcal{W}$ satisfying the conditions that*

1° *f becomes C-upper semicontinuous when $\mathcal{X}$ is equipped with some topology $\mathcal{T}(\mathcal{X})$ and*

2° *there exists a neighborhood V of the origin of $\mathcal{W}$ such that the closure of the set $f^{-1}(V+C)$ is compact with respect to $\mathcal{T}(\mathcal{X})$.*

Let g be a real-valued upper semicontinuous concave functional on $\mathcal{X}$ with respect to $\mathcal{T}(\mathcal{X})$.

If the set $\{x \mid f(x) \in C\}$ is not empty, then the following propositions hold: 1)

$$\sup\{g(x) \mid x \in \mathcal{X},\ f(x) \in C\} = \inf_{w^* \in C^+} \sup_{x \in \mathcal{X}} \{g(x) + w^*(f(x))\}, \tag{3.15}$$

where C^+ is the conjugate cone of C.

2) A point x_0 satisfying $f(x_0) \in C$ attains the supremum of the left-hand side of (3.15), if and only if there exists a sequence $\{w_l^*\}$ of members w_l^* $(l=1, 2, \cdots)$ of C^+ satisfying

$$\begin{cases} w_l^*(f(x_0)) \to 0\,, \\ \{[g(x_0)+w_l^*(f(x_0))] - \sup_{x \in \mathcal{X}} [g(x)+w_l^*(f(x))]\} \to 0\,, \end{cases}$$

as $l \to \infty$.

We here define the notions used in Lemma 3.3 in the following way:

1) C induces a preordering relation $\preceq$ on $\mathcal{W}$ such that $w_1 \preceq w_2$ if and only if $w_2 - w_1 \in C$. A preordering on a space $\mathcal{W}$ is a binary relation $\preceq$ on $\mathcal{W}$ satisfying (a) $u \preceq u$ for all $u \in \mathcal{W}$, (b) $u \preceq v$ and $v \preceq w$ imply $u \preceq w$ for $u, v, w \in \mathcal{W}$. A preordering induced by a cone C is a partial ordering if and only if $C \cap (-C) = \{0\}$.

2) A mapping f from $\mathcal{X}$ to $\mathcal{W}$ is said to be concave with respect to

the preorder induced by C, if $x_1, x_2 \in \mathcal{X}$ and $0 \leqq \lambda \leqq 1$ imply $\lambda f(x_1) + (1-\lambda)f(x_2) \preceq f(\lambda x_1 + (1-\lambda)x_2)$.

3) In 1°, f is called C-upper semicontinuous if for each open set W of $\mathcal{W}$, $f^{-1}(W-C)$ is an open set of $\mathcal{X}$. $W-C$ denotes the set $\{w-c \mid w \in W \text{ and } c \in C\}$.

We now consider applying Lemma 3.3 to obtaining a necessary and sufficient condition for $\hat{\phi}_\xi \in \hat{\Phi}(\hat{C})$ to be a Bayes decision function. Under the same formulation as that in the proof of Theorem 3.1, we can also replace $\hat{\Phi}$ and $\hat{\Phi}(\hat{C})$ by $\tilde{\Phi}$ and $\tilde{\Phi}(\tilde{C})$ and regard $\alpha_i + \beta_{in}(\cdot) - \mathrm{E}.\tau_i(\cdot, \tilde{\phi})$ and $\tilde{r}(\xi, \tilde{\phi})$ as a continuous convex-linear mapping from $\tilde{\Phi}$ to $L(\Theta)$ and a continuous convex-linear mapping from $\tilde{\Phi}$ to $\mathcal{R}$, respectively. Further, since $\tilde{\Phi}$ is sequentially compact, the Eberlein-Smulian theorem (see [6], pp. 310–315 [6]) implies that $\tilde{\Phi}$ is relatively compact with respect to the topology $\mathcal{T}(\mathcal{V}_1, \mathcal{V}_2)$.

We first consider the case where $\tilde{C}$ is defined by $\mathrm{E}_\theta \tau_i(\theta, \tilde{\phi}) \leqq \alpha_i + \beta_{in}(\theta)$ for ξ—a.a. θ. In the mth product space $L^m(\Theta)$, the set $\mathcal{K}^m$ defined by $\{(k_1, \cdots, k_m) \mid k_i \in L(\Theta),\ k_i(\theta) \geqq 0 \text{ for } \xi\text{—a.e. } \theta\ (i=1, \cdots, m)\}$ is a closed convex cone. Therefore, replacing $\mathcal{X}$, $\mathcal{W}$, C, f and g by $\tilde{\Phi}$, L^m, $\mathcal{K}^m$, $(\alpha_1 + \beta_{1n}(\cdot) - \mathrm{E}.\tau_1(\cdot, \tilde{\phi}), \cdots, \alpha_m + \beta_{mn}(\cdot) - \mathrm{E}.\tau_m(\cdot, \tilde{\phi}))$ and $-\tilde{r}(\xi, \tilde{\phi})$, we can see that the conditions assumed in Lemma 3.3 are all satisfied.

Let L^{m*} and $\mathcal{K}^{m*}$ denote the conjugate space of the product space of $L^m(\Theta)$ and the conjugate cone of $\mathcal{K}^m$ with respect to the product topology of L^m, respectively. For each $k^* \in \mathcal{K}^{m*}$ there exists a member $(h_1, h_2, \cdots, h_m) \in B^m$ such that $h_i(\theta) \geqq 0$ for ξ—a.e. θ $(i=1, \cdots, m)$ and

$$k^*(k_1, \cdots, k_m) = \sum_{i=1}^{m} \int_\Theta h_i(\theta) k_i(\theta) \xi(d\theta) \tag{3.16}$$

for every $(k_1, \cdots, k_m) \in L^m$. Hence, denoting for each i

$$\begin{cases} \bar{h}_i^{(1)}(x, s) = \displaystyle\int_\Theta h_i(\theta) \bar{\phi}_i(x, \theta) \bar{\xi}(d\theta \mid s)\ , \\ \bar{h}_i^{(2)} = \displaystyle\int_\Theta h_i(\theta) [\alpha_i + \beta_{in}(\theta)] \xi(d\theta)\ , \end{cases} \tag{3.17}$$

we have by the conditions (A2), (A3) and (A5) and (3.17) that

$$\begin{aligned} &k^*(\alpha_1 + \beta_{1n}(\theta) - \mathrm{E}_\theta \tau_1(\theta, \tilde{\phi}), \cdots, \alpha_m + \beta_{mn}(\theta) - \mathrm{E}_\theta \tau_m(\theta, \tilde{\phi})) \\ &\quad = -\sum_{i=1}^{m} \int_\Theta h_i(\theta) \int_{\mathcal{X} \times \mathcal{S}} \bar{\phi}_i(x, \theta) \tilde{\phi}(x, s) P^x(dx) P_\theta^n(ds) \xi(d\theta) \\ &\quad\quad + \sum_{i=1}^{m} \int_\Theta \{\alpha_i + \beta_{in}(\theta)\} h_i(\theta) \xi(d\theta) \\ &\quad = -\sum_{i=1}^{m} \int_{\mathcal{X} \times \mathcal{S}} \left[\int_\Theta h_i(\theta) \bar{\phi}_i(x, \theta) \bar{\xi}(d\theta \mid s) \right] \end{aligned} \tag{3.18}$$

$$\times\tilde{\phi}(x, s)P^x(dx)\bar{P}^n(ds)+\sum_{i=1}^{m}\bar{h}_i^{(2)}$$

$$=-\sum_{i=1}^{m}\int_{\mathcal{X}\times\mathcal{S}}\bar{h}_i^{(1)}(x, s)\tilde{\phi}(x, s)P^x(dx)\bar{P}^n(ds)+\sum_{i=1}^{m}\bar{h}_i^{(2)}.$$

Further, (3.15) is, in this case, expressed as

$$\begin{aligned}(3.19)\qquad &\inf[\tilde{r}(\xi, \tilde{\phi});\ \tilde{\phi}\in\tilde{\Phi}(\tilde{C})]\\ &=\sup\left[\int_{\mathcal{X}\times\mathcal{S}}\left\{\bar{\bar{\phi}}_0(x, s)+\sum_{i=1}^{m}\bar{h}_i^{(1)}(x, s)\right\}^{-}P^x(dx)\bar{P}^n(ds)\right.\\ &\quad -\sum_{i=1}^{m}\bar{h}_i^{(2)}\,|\,(h_1,\cdots,h_m)\in B^m,\ h_i(\theta)\geqq 0\\ &\qquad\qquad\left.\text{for } \xi-\text{a.e. }\theta\ (i=1, 2,\cdots, m)\right],\end{aligned}$$

which stands for the expression of the Bayes risk. Here, the notation $\{t(x, s)\}^{-}$ is defined to be $t(x, s)$ or 0 according to $t(x, s)\leqq 0$ or $t(x, s)>0$ for $P^x\times\bar{P}^n$—a.e. (x, s).

In the case where $\tilde{C}$ is defined by $E_\theta\,\tau_i(\theta, \tilde{\phi})=\alpha_i+\beta_{in}(\theta)$ for ξ—a.a. θ, we define the cone $\mathcal{K}^m$ as the set $\{0\}$ consisting of the sole point, the origin of L^m. Then, the constraint "$h_i(\theta)\geqq 0$ for ξ—a.e. θ" stated in (3.16) and (3.19) is eliminated.

Finally we notice that for each function $h_i\in B(\Theta)$ there exists a bounded Λ-measurable function h_i' on Θ such that $h_i'(\theta)=h_i(\theta)$ for ξ—a.e. θ. Therefore, we can use h_i' instead of h_i in case of defining $\bar{h}_i^{(1)}$ and $\bar{h}_i^{(2)}$ in (3.17) for the expression of the necessary and sufficient condition. Under this remark, we prepare the following definition for the purpose. For sequences $\{\bar{h}_{il}^{(1)}\}$ $(i=1, 2,\cdots, m)$ of $\mathcal{A}\times\mathcal{U}$-measurable and $P^x\times\bar{P}^n$-integrable real-valued function $\bar{h}_{il}^{(1)}$ $(l=1, 2,\cdots)$ on $\mathcal{X}\times\mathcal{S}$, let $\hat{\phi}_l$ be a member of $\hat{\Phi}$ satisfying

$$(3.20)\qquad \hat{\phi}_l(x, s)=\begin{cases}1 & \text{if } \bar{\bar{\phi}}_0(x, s)+\sum_{i=1}^{m}\bar{h}_{il}^{(1)}(x, s)<0,\\ 0 & \text{if } \bar{\bar{\phi}}_0(x, s)+\sum_{i=1}^{m}\bar{h}_{il}^{(1)}(x, s)>0\end{cases}$$

for $P^x\times\bar{P}^n$—a.e. (x, s). Then the following main theorem is straightforwardly obtained from the above-mentioned discussion.

THEOREM 3.2. *Assume the same conditions as stated in Theorem 3.1. Then,* $\hat{\phi}_\xi\in\hat{\Phi}(\hat{C})$ *is a Bayes decision function minimizing* $\tilde{r}(\xi, \hat{\phi})$ *in the case where* $\hat{C}$ *is defined by (2.4′), if and only if for each* i $(i=1, 2,\cdots, m)$ *there exist a sequence* $\{h_{il}\}$ *of bounded* Λ*-measurable functions* h_{il} *on* Θ *satisfying*

$$\int_{\mathcal{X}\times\mathcal{S}}\left[\sum_{i=1}^{m}\bar{h}_{il}^{(1)}(x,s)\right]\hat{\phi}_{\xi}(x,s)P^{x}(dx)\bar{P}^{n}(ds)-\sum_{i=1}^{m}\bar{h}_{il}^{(2)}\to 0\ ,$$

$$\tilde{r}(\xi,\hat{\phi}_{\xi})+\int_{\mathcal{X}\times\mathcal{S}}\left[\sum_{i=1}^{m}\bar{h}_{il}^{(1)}(x,s)\right]\hat{\phi}_{\xi}(x,s)P^{x}(dx)\bar{P}^{n}(ds)$$
$$-\int_{\mathcal{X}\times\mathcal{S}}\left[\bar{\bar{\phi}}_{0}(x,s)+\sum_{i=1}^{m}\bar{h}_{il}^{(1)}(x,s)\right]\hat{\phi}_{l}(x,s)P^{x}(dx)\bar{P}^{n}(ds)\to 0\ ,$$

as $l\to\infty$, *where each* $\bar{h}_{il}^{(1)}$ *and each* $\bar{h}_{il}^{(2)}$ *are defined by*

$$(3.21)\qquad \begin{cases} \bar{h}_{il}^{(1)}(x,s)=\int_{\Theta}h_{il}(\theta)\bar{\varphi}_{i}(x,\theta)\bar{\xi}(d\theta\,|\,s)\ , \\ \bar{h}_{il}^{(2)}=\int_{\Theta}\{\alpha_{i}+\beta_{in}(\theta)\}h_{il}(\theta)\xi(d\theta) \end{cases}$$
$$(i=1,2,\cdots,m),\ (l=1,2,\cdots)\ ,$$

and each $\hat{\phi}_{l}$ *is the member of* $\hat{\Phi}$ *defined by* (3.20).

Theorem 3.2 also holds true if the constraint (2.4′) is replaced by (2.4) and if the following condition is added to the "if and only if" part: For each i and for each l, $h_{il}(\theta)\geqq 0$ for ξ—a.e. θ.

From Theorem 3.2 it directly follows that:

COROLLARY 3.2. *Assume the same conditions as stated in Theorem 3.1. Then, an element* $\hat{\phi}_{\xi}\in\hat{\Phi}(\hat{C})$ *is a Bayes decision function in the case where* $\hat{C}$ *is defined by* (2.4′), *if there exist bounded* Λ*-measurable functions* h_{i} *on* Θ *such that* $\hat{\phi}_{\xi}$ *satisfies*

$$(3.22)\qquad \hat{\phi}_{\xi}(x,s)=\begin{cases} 1 & \text{if } \bar{\bar{\phi}}_{0}(x,s)+\sum_{i=1}^{m}\bar{h}_{i}^{(1)}(x,s)<0\ , \\ 0 & \text{if } \bar{\bar{\phi}}_{0}(x,s)+\sum_{i=1}^{m}\bar{h}_{i}^{(1)}(x,s)>0 \end{cases}$$

for $P^{x}\times\bar{P}^{n}$—a.e. (x,s) *and*

$$(3.23)\qquad \sum_{i=1}^{m}\int_{\mathcal{X}\times\mathcal{S}}\bar{h}_{i}^{(1)}(x,s)\hat{\phi}_{\xi}(x,s)P^{x}(dx)\bar{P}^{n}(ds)-\sum_{i=1}^{m}\bar{h}_{i}^{(2)}=0\ ,$$

where $\bar{h}_{i}^{(1)}$ *and* $\bar{h}_{i}^{(2)}$ *are defined by*

$$(3.24)\qquad \begin{cases} \bar{h}_{i}^{(1)}(x,s)=\int_{\Theta}h_{i}(\theta)\bar{\varphi}_{i}(x,\theta)\bar{\xi}(d\theta\,|\,s)\ , \\ \bar{h}_{i}^{(2)}=\int_{\Theta}[\alpha_{i}+\beta_{in}(\theta)]h_{i}(\theta)\xi(d\theta) \qquad (i=1,2,\cdots,m)\ . \end{cases}$$

In the case where the constraint $\hat{C}$ is defined by (2.4), the same assertion as is aformentioned holds, if the following condition is added to the "if" part, that is, the condition of the functions $h_i: h_i(\theta)\geqq 0$ for ξ—a.e. θ ($i=1, 2, \cdots, m$).

We next consider minimax decision functions. The condition (A5) is, in this case, replaced by:

(A5′) Some distance can be assigned to $\boldsymbol{\Theta}$ such that $\boldsymbol{\Theta}$ becomes a separable metric space satisfying the conditions 2°, 4° and 5° mentioned in (A5).

We obtain the following result by the same method as stated in Theorem 5.1 [8], since $\mathrm{E}_\theta\tau_i(\theta, \hat{\phi})$ are, for each $\hat{\phi}$, continuous in θ as is seen in Lemma 3.1.

THEOREM 3.3. *Assume the conditions* (A2), (A4) *and* (A5′). *Let* Ξ *be the family of all prior probability distributions whose supports are finite subsets of* $\boldsymbol{\Theta}$.

Then, the following properties hold:

1) $\inf_{\hat{\phi}\in\hat{\Phi}(\hat{C})}\sup_{\theta\in\Theta} r(\theta, \hat{\phi}) = \inf_{\hat{\phi}\in\hat{\Phi}(\hat{C})}\sup_{\xi\in\Xi}\bar{r}(\xi, \hat{\phi}) = \sup_{\xi\in\Xi}\inf_{\hat{\phi}\in\hat{\Phi}(\hat{C})}\bar{r}(\xi, \hat{\phi})$.

2) If $\inf_{\hat{\phi}\in\hat{\Phi}(\hat{C})}\sup_{\theta\in\Theta} r(\theta, \hat{\phi})<\infty$, *then there exists a minimax decision function* $\hat{\phi}_0\in\hat{\Phi}(\hat{C})$ *such that* $\hat{\phi}_0$ *satisfies*

$$r(\theta, \hat{\phi}_0) = \inf_{\hat{\phi}\in\hat{\Phi}(\hat{C})}\sup_{\theta\in\Theta} r(\theta, \hat{\phi}) .$$

3) *Selecting a subfamily* Ξ' *of* Ξ *which consists of a countable number of elements,* $\hat{\phi}_0$ *can be expressed as a Bayes solution in the wide sense relative to* Ξ'.

Example 3.1. Let us consider Bayes decision functions under the formulation in Example 2.1. Under the condition (A3), let $\bar{\xi}(\cdot\,|\,s)$ be a regular posterior probability distribution on $(\boldsymbol{\Theta}, \Lambda)$ given s with respect to $\bar{P}^n$.

Considering the decision function $\hat{\phi}(x, s)\equiv\alpha_1$, we can see that $\hat{\Phi}(\hat{C})$ is not empty. Further, since $\bar{\phi}_1(x, \theta)\equiv 1$ and $\bar{\phi}_1$ does not depend on $x\in\mathcal{X}$, (3.18) is, in this case, expressed as

$$\begin{aligned}(3.25)\quad & k^*[\alpha_1-\mathrm{E}_\theta\tau_1(\theta, \tilde{\phi})]\\ &=\alpha_1\int_\Theta h_l(\theta)\xi(d\theta)-\int_\Theta h_l(\theta)\int_{\mathcal{X}\times S}1\cdot\tilde{\phi}(x, s)P^x(dx)P_\theta^n(ds)\xi(d\theta)\\ &=\alpha_1\int_\Theta h_l(\theta)\xi(d\theta)-\int_{\mathcal{X}\times S}\left[\int_\Theta h_l(\theta)\bar{\xi}(d\theta\,|\,s)\right]\\ &\quad\times\tilde{\phi}(x, s)P^x(dx)\bar{P}^n(ds)\ ,\end{aligned}$$

where $h_l\in B(\boldsymbol{\Theta})$ ($l=1, 2, \cdots$). Therefore, writing

$$(3.26)\qquad \bar{h}_l^{(1)}(s)=\alpha_1\int_\Theta h_l(\theta)\bar{\xi}(d\theta\,|\,s)\ ,\qquad \bar{h}_l^{(2)}=\alpha_1\int_\Theta h_l(\theta)\xi(d\theta)\ ,$$

we can reexpress (3.25) as

$$(3.27)\quad k^*[\alpha_1 - \mathrm{E}_\theta \tau_1(\theta, \tilde{\phi})] = \bar{h}_l^{(2)} - \frac{1}{\alpha_1}\int_{\mathcal{X}\times\mathcal{S}} \bar{h}_l(s)\tilde{\phi}(x, s)P^x(dx)\bar{P}^n(ds) .$$

On the other hand, let η be the regression function of y on x defined by

$$(3.28)\qquad \eta(x, \theta) = \int_{\mathcal{Y}} yP_\theta^{y|x}(dy\,|\,x) , \qquad \theta \in \Theta .$$

Then, (3.6) is, in this case, rewritten as

$$(3.29)\qquad \bar{\bar{\phi}}_0(x, s) = \frac{-1}{\alpha_1}\int_\Theta \eta(x, \theta)\bar{\xi}(d\theta\,|\,s) = -\frac{1}{\alpha_1}\bar{\eta}(x, s) ,$$

where

$$(3.30)\qquad \bar{\eta}(x, s) = \int_\Theta \eta(x, \theta)\bar{\xi}(d\theta\,|\,s) .$$

Thus, by (3.27) and (3.29) we can reexpress (3.19) as

$$(3.31)\qquad \begin{aligned} &\sup\,[-\tilde{r}(\xi, \tilde{\phi});\ \tilde{\phi} \in \tilde{\Phi}(\tilde{C})] \\ &= \inf\left[\int_{\mathcal{X}\times\mathcal{S}} \frac{1}{\alpha_1}[\bar{\eta}(x, s) - \bar{h}_l^{(1)}(s)]^+ P^x(dx)\bar{P}^n(ds) \right. \\ &\qquad \left. + \bar{h}_l^{(2)}:\ h_l \in B(\Theta)\right] . \end{aligned}$$

Here, the notation $[t(x, s)]^+$ is defined to be $t(x, s)$ or 0 according to $t(x, s) \geqq 0$ or $t(x, s) < 0$ for $P^x \times \bar{P}^n$–a.e. (x, s). Hence, considering (3.31) into account, we obtain by Theorem 3.2 the following result. A neccessary and sufficient condition for $\hat{\phi}_\xi \in \hat{\Phi}(\hat{C})$ to be a Bayes decision function is that there exists a sequence $\{h_l\}$ of bounded Λ-measurable functions h_l on Θ satisfying

$$(3.32)\qquad \begin{cases} \dfrac{1}{\alpha_1}\displaystyle\int_{\mathcal{X}\times\mathcal{S}} \bar{h}_l^{(1)}(s)\hat{\phi}_\xi(x, s)P^x(dx)\bar{P}^n(ds) - \bar{h}_l^{(2)} \to 0 , \\ \dfrac{1}{\alpha_1}\displaystyle\int_{\mathcal{X}\times\mathcal{S}} [\bar{\eta}(x, s) - \bar{h}_l^{(1)}(s)]\hat{\phi}_\xi(x, s)P^x(dx)\bar{P}^n(ds) \\ \qquad - \dfrac{1}{\alpha_1}\displaystyle\int_{\mathcal{X}\times\mathcal{S}} [\bar{\eta}(x, s) - \bar{h}_l^{(1)}(s)]^+ P^x(dx)\bar{P}^n(ds) \to 0 , \end{cases}$$

as $l \to \infty$.

Example 3.2. In Example 3.1, let $\mathcal{X}\times\mathcal{Y} = \mathcal{R}^q \times \mathcal{R}$, $\theta_{(1)} = (\theta_0, \theta_1, \cdots, \theta_q)$ be a vector consisting of real numbers θ_0 and θ_i $(i = 1, \cdots, q)$ some of which is not zero, $P_\theta^{y|x}(dy|x)=(1/\sqrt{2\pi}\theta_{(2)})\exp\left[(-1/2\theta_{(2)}^2)\left(y-\theta_0-\sum_{i=1}^q\theta_i x_i\right)^2\right]$ dy with a positive number $\theta_{(2)}$, and P^x be a nonsingular normal distribution $N(0, \Sigma)$, where $\Sigma=(\sigma_{ij})$ is known.

The selection ϕ_0^* defined for every x by

$$\phi_\theta^*(x)=\begin{cases}1 & \text{if } \theta_0+\sum_{i=1}^q\theta_i x_i\geqq h(\theta)\,,\\ 0 & \text{if } \theta_0+\sum_{i=1}^q\theta_i x_i< h(\theta)\end{cases}\tag{3.33}$$

is optimum, where for every θ

$$h(\theta)=\sigma(\theta)\psi^{-1}(1-\alpha_1)+\theta_0 \quad \text{with} \quad \sigma(\theta)=\sqrt{\sum_{i,j=1}^q\sigma_{ij}\theta_i\theta_j}\tag{3.34}$$

and Ψ^{-1} denotes the inverse function of the normal distribution function Ψ with the mean 0 and the variance 1. We here restrict the parameter space Θ to the one defined by $\Theta=\Theta_1\times\Theta_2$, $\Theta_1=\{\theta_{(1)}|\theta_{(1)}=(\theta_0, \theta_1,\cdots,\theta_q),\ 0<a_i<\theta_i<b_i\ (i=0,1,\cdots,q)\}$ and $\Theta_2=\{\theta_{(2)}|0<c<\log\theta_{(2)}<d\}$, where a_i, b_i, c and d are given constants Also, let the prior distributions of $\theta_{(1)}$ and $\log\theta_{(2)}$ be those as are locally independent and uniform on Θ_1 and Θ_2, that is, $\xi_1(d\theta_{(1)})\propto d\theta_{(1)}$ on Θ_1 and $\xi_2(d\theta_{(2)})\propto(1/\theta_{(2)})d\theta_{(2)}$ on Θ_2, respectively (see [12]). Using matrix notations, we obtain by [12] that the joint posterior distribution of $\theta_{(1)}$ and $\theta_{(2)}$ is expressed by

$$\bar\xi(d(\theta_{(1)},\theta_{(2)})|s)\propto\theta_{(2)}^{-(n+1)}\exp\left[-\frac{1}{2\theta_{(2)}^2}\{(n-q-1)\hat\theta_{(2)}^2 +Q(\theta_{(1)},\hat\theta_{(1)},X'X)\}\right]d(\theta_{(1)},\theta_{(2)})\,,\qquad \theta=(\theta_{(1)},\theta_{(2)})\in\Theta\,,\tag{3.35}$$

and the marginal posterior distribution of $\theta_{(1)}$ is

$$\bar\xi_1(d\theta_{(1)}|s)\propto[1+Q(\theta_{(1)},\hat\theta_{(1)},X'X)/\{(n-q-1)\hat\theta_{(2)}^2\}]^{-n/2}d\theta_{(1)}\,,\qquad \theta_{(1)}\in\Theta_1\,,\tag{3.36}$$

where

$$s=(X,y)=\begin{bmatrix}1 & x_{11} & \cdots & x_{q1} & y_1\\ \vdots & \vdots & & \vdots & \vdots\\ 1 & x_{1n} & \cdots & x_{qn} & y_n\end{bmatrix},$$

$$\hat\theta_{(1)}=(X'X)^{-1}X'y\,,\qquad \hat\theta_{(2)}^2=\{(y-X\hat\theta_{(1)})'(y-X\hat\theta_{(1)})\}/(n-q-1)\,,$$

$$Q(\theta_{(1)},\hat\theta_{(1)},X'X)=(\theta_{(1)}-\hat\theta_{(1)})'\cdot X'X\cdot(\theta_{(1)}-\hat\theta_{(1)})\,.$$

According to Corollary 3.2, we here find a Bayes decision function

$\hat{\phi}_\xi \in \hat{\Phi}(\hat{C})$ for $\xi = \xi_1 \times \xi_2$. Since the marginal posterior distribution of each component $\theta_i - \hat{\theta}_i$ with respect to $\bar{\xi}_1$ may be expressed in terms of a univariate t-distribution with $n-q-1$ degrees of freedom (see [12]), it follows from Example 3.1 that in this case

$$(3.37) \qquad \bar{\eta}(x, s) = \hat{\theta}_0 + \sum_{i=1}^{q} \hat{\theta}_i x_i$$

for every (x, s). Since $\bar{\eta}$ is bilinear in $(1, x_1, \cdots, x_q)$ and $(\hat{\theta}_0, \hat{\theta}_1, \cdots, \hat{\theta}_q)$, except a difference by $\bar{P}^n$-null function there uniquely exists a function $\bar{h}(s)$ on $\mathcal{S}$ satisfying

$$(3.38) \qquad \int_{\{x;\, \bar{\eta}(x,s) > \bar{h}(s)\}} P^x(dx) = \alpha_1$$

for every s. In fact, it follows from (3.37) and (3.38) that $\bar{h}(s)$ can be expressed as

$$(3.39) \qquad \bar{h}(s) = \hat{\sigma}\Psi^{-1}(1-\alpha_1) + \hat{\theta}_0 \quad \text{with} \quad \hat{\sigma}(s) = \sqrt{\sum_{i,j=1}^{q} \sigma_{ij}\hat{\theta}_i\hat{\theta}_j}$$

for every s. Thus, in this case, the decision function $\hat{\phi}_\xi$ defined by

$$(3.40) \qquad \hat{\phi}_\xi(x, s) = \begin{cases} 1 & \text{if } \sum_{i=1}^{q} \hat{\theta}_i(s) x_i \geqq \hat{\sigma}(s)\Psi^{-1}(1-\alpha_1)\,, \\ 0 & \text{if } \sum_{i=1}^{q} \hat{\theta}_i(s) x_i < \hat{\sigma}(s)\Psi^{-1}(1-\alpha_1)\,, \qquad (x, s) \in \mathcal{X} \times \mathcal{S}\,, \end{cases}$$

is a Bayes decision function contained in $\hat{\Phi}(\hat{C})$.

Example 3.3. In Example 2.2, let $\mathcal{X} \times \mathcal{Y} = \mathcal{R} \times \mathcal{R}$, $\theta_{(2)} = \theta_2$, $P_\theta^{y|x}(dy \mid x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(y - \theta_1 - \theta_2 x)^2\right] dy$, and $P^x(dx) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right] dx$. Further let $\Theta = \Theta^+ \cup \Theta^-$, where Θ^+ and Θ^- are the sets of all θ ($\in \mathcal{R}^2$) satisfying the following inequalities, respectively:

$$(3.41) \qquad \theta_2 > 0\,, \quad \theta_2 > -(\alpha_1\alpha_2)/\psi(\Psi^{-1}(\alpha_1))\,, \quad \theta_2 > (\alpha_1\alpha_2)/\psi(\Psi^{-1}(1-\alpha_1))\,,$$

$$(3.42) \qquad \theta_2 < 0\,, \quad \theta_2 < -(\alpha_1\alpha_2)/\psi(\Psi^{-1}(\alpha_1))\,, \quad \theta_2 < (\alpha_1\alpha_2)/\psi(\Psi^{-1}(1-\alpha_1))\,.$$

Here ψ denotes the density function of $N(0, 1^2)$. Let $\bar{\Theta}$ be the closure of Θ. In this case, every indicator function of a bounded interval is contained in $\Phi(C_\theta)$ if and only if $\theta \in \Theta$ (see [9]). However, if θ is a boundary point of $\bar{\Theta}$, then the selection problem falls under the trivial case. In other words, if $\theta_2 = 0$, then every ϕ satisfying $\tau_1(\theta, \phi) = \alpha_1$ is optimum. If $\theta_2 \neq 0$ and θ is a boundary point of $\bar{\Theta}$, then $\Phi(C_\theta) = \{\phi_a\}$ or $\{\phi_b\}$. Here, ϕ_a and ϕ_b denote the indicator functions of $(-\infty, a]$ and

$[b, \infty)$ such that $P^x((-\infty, a]) = P^x([b, \infty)) = \alpha_1$, respectively. Therefore we define the parameter space by $\boldsymbol{\Theta}$.

A selection $\phi_\theta \in \Phi(C_\theta)$ is optimum (see also [9]), if and only if ϕ_θ satisfies

$$\phi_\theta(x) = \begin{cases} 1 & \text{if } x_1(\theta) < x < x_2(\theta)\,, \\ 0 & \text{if } x < x_1(\theta) \text{ or } x > x_2(\theta) \end{cases} \tag{3.43}$$

for P^x–a.e. x and for every $\theta \in \boldsymbol{\Theta}$, where $x_1(\theta)$ and $x_2(\theta)$ are the real solutions of

$$\zeta(x, \theta) \overleftarrow{=} 1 + (\theta_2 x - \alpha_2)^2 - h_1(\theta) - h_2(\theta)x = 0\,. \tag{3.44}$$

Here, $h_1(\theta)$ and $h_2(\theta)$ are uniquely determined as the real functions satisfying

$$\begin{aligned} &\Psi(x_2(\theta)) - \Psi(x_1(\theta)) = \alpha_1\,, \\ &\theta_2[\psi(x_1(\theta)) - \psi(x_2(\theta))]/\alpha_1 = \alpha_2\,, \qquad \theta \in \boldsymbol{\Theta}\,. \end{aligned} \tag{3.45}$$

We take one of these optimum selections and denote it by ϕ_θ^*.

We now consider decision functions $\hat{\phi}$ in $\hat{\Phi}$. We take as a prior probability distribution ξ a truncated normal distribution on $\boldsymbol{\Theta}$. We can obtain the form of the posterior probability distribution $\bar{\xi}(d\theta \mid s)$ with its mean $\hat{\theta}$ and covariance matrix $\hat{V} = (\hat{v}_{ij})$ by the same method as the one stated in [9]. Considering that the decision function $\phi_{\hat{\theta}}^*$ is standard, we set that $\beta_{1n}(\theta) \equiv 0$ and

$$\beta_{2n}(\theta) \overleftarrow{=} \mathrm{E}_\theta\, \tau_2(\theta, \phi_{\hat{\theta}}^*) - \alpha_2 = -\alpha_2 + \theta_2\, \mathrm{E}_\theta\, [(\theta_1 + \alpha_2 - \hat{\theta}_1)/\hat{\theta}_2]$$

for every $\theta \in \boldsymbol{\Theta}$.

Hence, under the constraint that $\mathrm{E}_\theta\, \tau_1(\theta, \hat{\phi}) = \alpha_1$ and $\mathrm{E}_\theta\, \tau_2(\theta, \hat{\phi}) = \alpha_2 + \beta_{2n}(\theta)$, we consider the Bayes decision functions minimizing the Bayes risk $\tilde{r}(\theta, \hat{\phi})$. The function $\bar{\bar{\varphi}}_0(x, s)$ defined in (3.6) is, in this case, expressed as

$$\bar{\bar{\varphi}}_0(x, s) = 1 + (\hat{\theta}_2(s)x - \alpha_2)^2 + \hat{v}_{11}(s)x^2\,. \tag{3.46}$$

To apply Theorem 3.2 to this case, we define the functions

$$\hat{\zeta}_l(x, s) = 1 + (\hat{\theta}_2(s)x - \alpha_2)^2 + \hat{v}_{11}(s)x^2 - \bar{h}_{1l}^{(1)}(s) - \bar{h}_{2l}^{(1)}(s)x \qquad (l = 1, 2, \cdots) \tag{3.47}$$

on $\mathcal{X} \times \mathcal{S}$, where each $\bar{h}_{il}^{(1)}$ $(i = 1, 2)$ has the form

$$(3.48)\qquad \begin{cases} \bar{h}_{1l}^{(1)}(s)=\int_{\Theta} h_{1l}(\theta)\bar{\xi}(d\theta\,|\,s)\ , \\ \bar{h}_{2l}^{(1)}(s)=\int_{\Theta} \theta_2 h_{2l}(\theta)\bar{\xi}(d\theta\,|\,s) \qquad (l=1,2,\cdots) \end{cases}$$

for some bounded Λ-measurable function h_{il} on Θ. We also consider the decision functions $\hat{\phi}_l\in\hat{\Phi}(\hat{C})$ satisfying for $P^x\times\bar{P}^n$—a.e. (x,s)

$$(3.49)\qquad \hat{\phi}_l(x,s)=\begin{cases} 1 & \text{if } \hat{x}_{1l}(s)<x<\hat{x}_{2l}(s)\ , \\ 0 & \text{if } x<\hat{x}_{1l}(s) \text{ or } x>\hat{x}_{2l}(s)\ , \qquad (l=1,2,\cdots)\ , \end{cases}$$

where $\hat{x}_{1l}(s)$ and $\hat{x}_{2l}(s)$ are the real solutions of the equations

$$(3.50)\qquad \hat{\zeta}_l(x,s)=0 \qquad (l=1,2,\cdots)\ .$$

We finally define for $h_{1l}(\theta)$ and $h_{2l}(\theta)$ given in (3.48) the constants

$$(3.51)\qquad \begin{aligned} \bar{h}_{1l}^{(2)}&=\alpha_1\int_{\Theta} h_{1l}(\theta)\bar{\xi}(d\theta)\ , \\ \bar{h}_{2l}^{(2)}&=\alpha_1\int_{\Theta} [\alpha_2+\beta_{2n}(\theta)]h_{2l}(\theta)\bar{\xi}(d\theta) \qquad (l=1,2,\cdots)\ . \end{aligned}$$

Thus we obtain by Theorem 3.2 the following result in this case: A decision function $\hat{\phi}_\xi\in\hat{\Phi}(\hat{C})$ is a Bayes decision function, if and only if there exist sequences $\{h_{1l}\}$ and $\{h_{2l}\}$ of bounded Λ-measurable functions h_{1l} and h_{2l} on Θ, respectively, satisfying

$$(3.52)\qquad \int_{\mathcal{X}\times\mathcal{S}} [\bar{h}_{1l}^{(1)}(s)+\bar{h}_{2l}^{(1)}(s)x]\hat{\phi}_\xi(x,s)P^x(dx)\bar{P}^n(ds)-[\bar{h}_{1l}^{(2)}+\bar{h}_{2l}^{(2)}]\to 0\ ,$$

$$(3.53)\qquad \int_{\mathcal{X}\times\mathcal{S}} \hat{\zeta}_l(x,s)\hat{\phi}_\xi(x,s)P^x(dx)\bar{P}^n(ds) - \int_{\mathcal{X}\times\mathcal{S}} \hat{\zeta}_l(x,s)\hat{\phi}_l(x,s)P^x(dx)\bar{P}^n(ds)\to 0\ ,$$

as $l\to\infty$.

Acknowledgement

The author would like to thank the referee for his valuable and helpful comments. In particular, the author is much obliged to the referee for presenting Lemma 3.1, Lemma 3.2, Theorem 3.1 and Theorem 3.2 and for proving them. The author is also grateful to Mrs. S. Terao for preparing the manuscript.

THE INSTITUTE OF STATISTICAL MATHEMATICS

REFERENCES

[1] Birnbaum, Z. W. and Chapman, D. G. (1950). On optimum selections from multinormal populations, *Ann. Math. Statist.*, **21**, 443-447.

[2] Bourbaki, N. (1965). Éléments de mathematique, *Topologie générale*, Chapitre let 2, Hermann, Paris.

[3] Chochran, W. G. (1951). Improvement by means of selection, *Proc. Second Berkeley Sympo. on Math. Statist. and Probability*, 449-470.

[4] Ferguson, T. S. (1967). *Mathematical Statistics*, Academic Press, New York.

[5] Isii, K. (1967). The Lagrange multiplier method and admissibility, *Surikagaku-Kokyuroku*, **27**, 73-87 (in Japanese).

[6] Köthe, G. (1966). *Topological Vector Space I*, Translated by D. J. H. Garling, Springer-Verlag, Berlin.

[7] Lehmann, E. L. (1959). *Testing Statistical Hypotheses*, John Wiley and Sons, New York.

[8] Noda, K. and Taga, Y. (1971). Minimax estimation method for the optimum decomposition of a sample space based on prior information, *Ann. Inst. Statist. Math.*, **23**, 1-29.

[9] Noda, K. (1980). Bayes decision functions for selections of subpopulations, Recent Development in Statistical Inference and Data Analysis, *Proc. Inter. Conference in Statist.*, in Tokyo, (ed. Matusita, K.), North-Holland Publishing Company, 227-236.

[10] Raj, D. (1954). On optimum selections from multivariate populations, *Sankhyā*, **14**, 363-366.

[11] Schwartz, L. (1967). *Cours d'analyse I*, Hermann, Paris.

[12] Tiao, G. C. and Zellner, A. (1964). Bayes's theorem and the use of prior knowledge in regression analysis, *Biometrika*, **51**, 219-230.