The Annals of Mathematical Statistics
1971, Vol. 42, No. 1, 330–341

Υ-MINIMAX SELECTION PROCEDURES IN TREATMENTS VERSUS CONTROL PROBLEMS

BY RONALD H. RANDLES[1] AND MYLES HOLLANDER[2]

The Florida State University

1. Introduction and summary. Let X be a random variable with distribution function $F(x \mid \theta)$ over a sample space $\mathscr{X}$ for each θ in Θ, the parameter space. Consider the decision problem with loss function $L(\theta, a)$ for each a in $\mathscr{A}$, the action space. For a decision rule $\delta(x)$ mapping $\mathscr{X}$ into $\mathscr{A}$ the risk function is $R(\theta, \delta) = \int L(\theta, \delta(x))\, dF(x \mid \theta)$. If $\tau(\theta)$ is a distribution over Θ, the expected risk of using rule $\delta(x)$ is then $r(\tau, \delta) = \int_\Theta R(\theta, \delta)\, d\tau(\theta)$.

In many problems a priori information will be incomplete. Suppose that our prior information consists of a class Υ of distributions over Θ. One method of utilizing such partial prior information to obtain a decision rule is given by Blum and Rosenblatt [1].

DEFINITION 1.1. The rule $\delta_0(x)$ is a Υ-minimax decision rule if

$$\sup_{\tau \in \Upsilon} r(\tau, \delta_0) = \inf_\delta sup_{\tau \in \Upsilon} r(\tau, \delta).$$

The use of partial prior information by Menges [5] and Hodges and Lehmann [2] may also be considered as satisfying the Υ-minimax criterion for suitable choices of Υ.

In this paper the Υ-minimax principle is applied to the problem of selecting treatment populations which have larger translation parameters than that of a control population. Let $S_0, S_1, \cdots, S_k$ be $k+1$ independent random variables with respective probability density functions $f_0(s-\theta_0)$, $f_1(s-\theta_1), \cdots, f_k(s-\theta_k)$. The random variables $S_0, S_1, \cdots, S_k$ may represent sufficient statistics from the control and k treatment populations, respectively. We assume that each $f_i(s)$, $i = 0, 1, \cdots, k$, is a Pólya frequency function of order two (PF_2), that is, if $x_1 < x_2$ and $y_1 < y_2$ then

$$\begin{vmatrix} f_i(x_1 - y_1) & f_i(x_1 - y_2) \\ f_i(x_2 - y_1) & f_i(x_2 - y_2) \end{vmatrix} \geqq 0.$$

Hence $f_i(s-\theta_i)$ has a monotone likelihood ratio in its translation parameter.

In Section 2, necessary notation, the loss function, and the incomplete prior information are introduced. In Section 3, a Υ-minimax decision rule is found, for the case in which the control population parameter θ_0 is known, by finding a rule which is Bayes with respect to a least favorable prior in the class of prior distributions Υ. The case in which the control population parameter is unknown is treated

Received November 10, 1969.

[1] Supported by a National Institute of Health Traineeship (5-T01-GM00913). Now at the State University of Iowa, Iowa City.

[2] Supported by the Office of Naval Research under contract (NONR-988(08)NR042-004).

in Section 4. Rules are derived which are Υ-minimax among procedures for which the decision to select or reject the ith population depends only on S_i and S_0. When specialized to normal populations with common known variance σ^2, a Υ-minimax rule selects (rejects) the ith population as $\bar{X}_i - \bar{X}_0 \geqq (<)$ a constant, where the constant depends on Δ (a known constant used to define "positive" and "negative" populations), σ^2, the sample sizes, the ratio of the losses for the two kinds of incorrect decisions, and the ratio of the prior probabilities of negative and positive populations. (An analogous result is found in Section 3 for the known control case.) Section 5 gives comparisons of a Υ-minimax rule with a Bayes competitor based on independent normal priors for the case of normal populations with common known variance. Some comments on a theorem by Y. L. Tong [7] are given in Section 6.

Whenever a new criterion is being considered by the statistical community, it is important to see if the criterion is fruitful in a variety of situations. In the treatment versus control problems discussed here, the Υ-minimax criterion leads to simple explicit rules which compare favorably with Bayes rules that require stronger assumptions on the prior distribution. We hope that this work will encourage others to obtain and apply Υ-minimax rules in different settings.

2. Statement of the problem. For $i = 1, \cdots, k$, define the ith population to be positive if $\theta_i \geqq \theta_0 + \Delta$ and negative if $\theta_i \leqq \theta_0$ where Δ is a specified positive constant. The objective is to select all positive populations while rejecting all negative ones. This formulation is similar to that of Lehmann [4] and Tong [7].

Let L_1 denote the loss incurred if we fail to select a positive population and L_2, the loss for each negative population selected. If $\mathbf{S} = (S_0, S_1, \cdots, S_k)$, consider decision rules of the form

$$\boldsymbol{\psi}(\mathbf{s}) = (\psi_1(\mathbf{s}), \cdots, \psi_k(\mathbf{s})) \tag{2.1}$$

where $\psi_i(\mathbf{s})$ denotes the conditional probability of selecting the ith population given $\mathbf{S} = \mathbf{s}$. The loss function is then

$$L(\boldsymbol{\theta}, \boldsymbol{\psi}) = \sum_{i=1}^{k} L^{(i)}(\boldsymbol{\theta}, \psi_i) \tag{2.2}$$

where

$$\begin{aligned} L^{(i)}(\boldsymbol{\theta}, \psi_i) &= L_1(1-\psi_i) && \text{if } \theta_i \geqq \theta_0 + \Delta; \\ &= L_2 \psi_i && \text{if } \theta_i \leqq \theta_0; \\ &= 0 && \text{otherwise.} \end{aligned}$$

The risk function is $R(\boldsymbol{\theta}, \boldsymbol{\psi}) = \sum_{i=1}^{k} R^{(i)}(\boldsymbol{\theta}, \psi_i)$ where

$$R^{(i)}(\boldsymbol{\theta}, \psi_i) = \int_{\mathscr{S}_0} \int_{\mathscr{S}_1} \cdots \int_{\mathscr{S}_k} L^{(i)}(\boldsymbol{\theta}, \psi_i(\mathbf{s})) \prod_{i=0}^{k} [f_i(s_i - \theta_i)\, ds_i]$$

and $\mathscr{S}_i$ denotes the sample space of the random variable S_i. Thus $R(\boldsymbol{\theta}, \boldsymbol{\psi}) = L_1 N_1 + L_2 N_2$ where $N_1(N_2)$ is the expected number of positive (negative) populations rejected (selected). Note that there is no loss of generality in considering decision rules of the form given in (2.1). For any decision rule there exists a rule in the class (2.1) with the same risk function. If $\tau(\boldsymbol{\theta})$ is a distribution over Θ then the expected risk of a procedure ψ is $r(\tau,\psi) = \sum_{i=1}^{k} r^{(i)}(\tau,\psi_i)$ where $r^{(i)}(\tau,\psi_i) = \int_\Theta R^{(i)}(\boldsymbol{\theta},\psi_i)\,d\tau(\boldsymbol{\theta})$.

Assume that partial prior information is available in the selection problem. Define $\Theta_P(i) = \{\boldsymbol{\theta} \mid \theta_i \geqq \theta_0+\Delta\}$ and $\Theta_N(i) = \{\boldsymbol{\theta} \mid \theta_i \leqq \theta_0\}$ and assume that we are able to specify $\pi_i = P[\boldsymbol{\theta}\in\Theta_P(i)]$ and $\pi_i' = P[\boldsymbol{\theta}\in\Theta_N(i)]$ so that $\pi_i+\pi_i' \leqq 1$ for $i = 1,\cdots,k$. Define

$$\Upsilon = \{\tau(\boldsymbol{\theta}) \mid \textstyle\int_{\Theta_P(i)} d\tau(\boldsymbol{\theta}) = \pi_i \text{ and } \int_{\Theta_N(i)} d\tau(\boldsymbol{\theta}) = \pi_i' \text{ for } i = 1,\cdots,k\}. \tag{2.3}$$

3. Known control population. Consider the ith component problem, that is, the selection or rejection of the ith population when θ_0, the parameter of the control population is known. The loss function for the component problem is $L^{(i)}(\boldsymbol{\theta},\psi_i)$.

LEMMA 3.1. *If $\psi_i = 1,0$ as $S_i \geqq, < d_i$ and if $\tau\in\Upsilon$ as defined by* (2.3) *then $r^{(i)}(\tau,\psi_i) \leqq r^{(i)}(\tau_0,\psi_i)$ where $\tau_0\in\Upsilon_0$ and*

$$\Upsilon_0 = \{\tau(\boldsymbol{\theta}) \mid P[\theta_i = \theta_0+\Delta] = \pi_i, \quad P[\theta_i = \theta_0] = \pi_i' \text{ and } \textstyle\int_{D_i} d\tau(\boldsymbol{\theta}) = 1-\pi_i-\pi_i',\ i = 1,\cdots,k\} \tag{3.1}$$

where D_i is the complement of $(\Theta_P(i)\cup\Theta_N(i))$.

PROOF. We consider three cases. If $\boldsymbol{\theta}$ is such that $\theta_i \geqq \theta_0+\Delta$, then

$$R^{(i)}(\boldsymbol{\theta},\psi_i) = L_1 \int_{-\infty}^{d_i} f_i(s-\theta_i)\,ds \leqq L_1 \int_{-\infty}^{d_i} f_i(s-\theta_0-\Delta)\,ds = R^{(i)}(\boldsymbol{\theta}^*(\boldsymbol{\theta}),\psi_i)$$

where $\boldsymbol{\theta}^*(\boldsymbol{\theta}) = (\theta_0, q(\theta_1,\theta_0,\Delta),\cdots,q(\theta_k,\theta_0,\Delta))$ and

$$\begin{aligned} q(\theta_j,\theta_0,\Delta) &= \theta_0+\Delta && \text{if } \theta_j \geqq \theta_0+\Delta;\\ &= \theta_j && \text{if } \theta_0 < \theta_j < \theta_0+\Delta;\\ &= \theta_0 && \text{if } \theta_j \leqq \theta_0. \end{aligned}$$

If $\boldsymbol{\theta}$ is such that $\theta_i \leqq \theta_0$, then $R^{(i)}(\boldsymbol{\theta},\psi_i) = L_2\int_{d_i}^{\infty} f_i(s-\theta_i)\,ds \leqq L_2\int_{d_i}^{\infty} f_i(s-\theta_0)\,ds = R^{(i)}(\boldsymbol{\theta}^*(\boldsymbol{\theta}),\psi_i)$. Finally, if $\boldsymbol{\theta}$ is such that $\theta_0 < \theta_i < \theta_0+\Delta$, then $R^{(i)}(\boldsymbol{\theta},\psi_i) = 0 = R^{(i)}(\boldsymbol{\theta}^*(\boldsymbol{\theta}),\psi_i)$. It follows that $r^{(i)}(\tau,\psi_i) \leqq r^{(i)}(\tau_0,\psi_i)$.

LEMMA 3.2. *If for each $i = 1,\cdots,k$, ψ_i^0 is a Bayes rule for the* ith *component problem with respect to the same $\tau_0(\boldsymbol{\theta})\in\Upsilon$, and if $\sup_{\tau\in\Upsilon} r^{(i)}(\tau,\psi_i^0) = r^{(i)}(\tau_0,\psi_i^0)$ for $i = 1,\cdots,k$ then $\psi^0 = (\psi_1^0,\cdots,\psi_k^0)$ is a Υ-minimax decision rule.*

PROOF. Let $\psi = (\psi_1,\cdots,\psi_k)$ be an arbitrary decision rule, then

$$\begin{aligned} \sup_{\tau\in\Upsilon} r(\tau,\psi) &\geqq r(\tau_0,\psi) = \textstyle\sum_{i=1}^{k} r^{(i)}(\tau_0,\psi_i) \geqq \sum_{i=1}^{k} r^{(i)}(\tau_0,\psi_i^0)\\ &= \textstyle\sum_{i=1}^{k} \sup_{\tau\in\Upsilon} r^{(i)}(\tau,\psi_i^0) \geqq \sup_{\tau\in\Upsilon} r(\tau,\psi^0). \end{aligned}$$

Define $\theta(j) = \theta_0+\Delta,\ \theta_0+\Delta/2,\ \theta_0$ as $j = 1,2,3$ and set $\Theta(j_1,\cdots,j_k) = \{\boldsymbol{\theta} \mid \theta_i = \theta(j_i)$ for $i = 1,\cdots,k\}$ where each $j_i = 1,2,3$. Define $\pi_i(j) = \pi_i,\ 1-\pi_i-\pi_i',\ \pi_i'$ as $j = 1, 2, 3$ and let

$$(3.2)\qquad \tau_0(\boldsymbol{\theta}) \in \{\tau(\boldsymbol{\theta}) \mid P[\boldsymbol{\theta} \in \Theta(j_1,\cdots,j_k)] = \textstyle\prod_{i=1}^k \pi_i(j_i) \text{ for each } (j_1,\cdots,j_k) \text{ with } 1 \leqq j_i \leqq 3\}.$$

Any such $\tau_0(\boldsymbol{\theta})$ is in Υ_0 as defined by (3.1) since $\sum_{(j_1,\cdots,j_k),\, j_i=r}[\prod_{t=1}^k \pi_t(j_t)] = \pi_i(r)$.

THEOREM 3.1. *Assume $S_1,\cdots,S_k$ are independent random variables with respective probability density functions $f_i(s-\theta_i)$ which are PF_2 densities. If the loss function is given by* (2.2) *the Υ-minimax decision rule, ψ^Υ, is of the form: $\psi_i^\Upsilon = 1,0$ as $S_i \geqq, < d_i$ for $i = 1,\cdots,k$ where each d_i is determined so that*

$$(3.3)\qquad L_2\,\pi_i' f_i(s-\theta_0) - L_1\,\pi_i f_i(s-\theta_0-\Delta) \leqq, > 0$$

as $s \geqq, < d_i$.

PROOF. For $\tau_0(\boldsymbol{\theta})$ given by (3.2) and $\psi_i(\mathbf{s})$,

$$\begin{aligned} r^{(i)}(\tau_0,\psi_i) &= \textstyle\int_{\mathscr{S}_1}\cdots\int_{\mathscr{S}_k}[L_1(1-\psi_i(\mathbf{s}))\{\sum_{(j_1,\cdots,j_k),\, j_i=1}[\prod_{t=1}^k \pi_t(j_t) f_t(s_t-\theta(j_t))]\} \\ &\quad + L_2\,\psi_i(\mathbf{s})\{\textstyle\sum_{(j_1,\cdots,j_k),\, j_i=3}[\prod_{t=1}^k \pi_t(j_t) f_t(s_t-\theta(j_t))]\}]\, ds_1\cdots ds_k \\ &= \textstyle\int_{\mathscr{S}_1}\cdots\int_{\mathscr{S}_k}[L_1(1-\psi_i(\mathbf{s}))\pi_i f_i(s_i-\theta_0-\Delta) \\ &\qquad \cdot\{\textstyle\sum_{(j_1,\cdots,j_k),\ \text{no } j_i \text{ term}}[\prod_{t=1,\, t\neq i}^k \pi_t(j_t) f_t(s_t-\theta(j_t))]\} \\ &\quad + L_2\,\psi_i(\mathbf{s})\pi_i' f_i(s_i-\theta_0) \\ &\qquad \cdot\{\textstyle\sum_{(j_1,\cdots,j_k),\ \text{no } j_i \text{ term}}[\prod_{t=1,\, t\neq i}^k \pi_t(j_t) f_t(s_t-\theta(j_t))]\}]\, ds_1\cdots ds_k. \end{aligned}$$

Thus the Bayes rule for the ith component problem with respect to $\tau_0(\boldsymbol{\theta})$ sets $\psi_i^\Upsilon = 1,0$ as

$$L_2\,\pi_i' f_i(s-\theta_0) - L_1\,\pi_i f_i(s-\theta_0-\Delta) \leqq, > 0.$$

Since $f_i(s-\theta_i)$ has a monotone likelihood ratio, there exists a d_i such that (3.3) holds as $s \geqq, < d_i$. It follows from Lemma 3.1 that

$$\sup_{\tau\in\Upsilon} r^{(i)}(\tau,\psi_i^\Upsilon) = r^{(i)}(\tau_0,\psi_i^\Upsilon).$$

Lemma 3.2 then yields that $\psi^\Upsilon = (\psi_1^\Upsilon,\cdots,\psi_k^\Upsilon)$ is Υ-minimax among all decision rules.

Note that the procedure depends on the prior probability $\pi_i(\pi_i')$ that the ith population is positive (negative) only through the ratio (π_i/π_i'). Υ-minimax procedures for this problem can also be derived by replacing the loss function given in (2.2) by an alternative selection criterion. We may desire a procedure ψ which will minimize the expected number of negative populations selected while insuring that the expected number of positive populations selected remains above some preassigned level. Minimax procedures under such a criterion have been considered by Lehmann [4].

The following example illustrates the application of Theorem 3.1. Consider $k+1$ normal populations $N(\theta_i, \sigma^2)$ $i = 0, \cdots, k$ with θ_0 and σ^2 known. A random sample of size n_i is taken from each of the k populations $N(\theta_i, \sigma^2)$ $i = 1, \cdots, k$. By the principle of sufficiency we can reduce consideration to $\bar{X}_1, \cdots, \bar{X}_k$ which are independent with normal distributions $N(\theta_i, \sigma^2/n_i)$, respectively. Since

$$f(s_i - \theta_0 - \Delta)/f(s_i - \theta_0) = \exp[-\tfrac{1}{2}\Delta^2 n_i \sigma^{-2} + (s_i - \theta_0) n_i \Delta \sigma^{-2}],$$

applying Theorem 3.1 yields $d_i = \theta_0 + \Delta/2 + \sigma^2(\Delta n_i)^{-1} \ln(L_2 \pi_i'/L_1 \pi_i)$. Thus the Υ-minimax decision rule selects the ith population if and only if

$$\bar{X}_i \geqq \theta_0 + \Delta/2 + \sigma^2(\Delta n_i)^{-1} \ln(L_2 \pi_i'/L_1 \pi_i).$$

4. Unknown control population. Define $\mathscr{D}^0$ to be the set of all decision rules $\psi = (\psi_1, \cdots, \psi_k)$ for which ψ_i depends only on S_0 and S_i. Consider the component problem, the selection or rejection of the ith population, when the control population parameter θ_0 is unknown. If $\mathbf{S}^i = (S_0, S_i)$ and $\boldsymbol{\theta}^i = (\theta_0, \theta_i)$, let G be the group of transformations which map $\mathbf{S}^i$ into $\mathbf{S}^i + b\mathbf{1}$ where $\mathbf{1}$ denotes the vector $(1, 1)$ and b is a real number. The induced group of transformations on the parameter space is then $\bar{G} = \{\bar{g} \mid \bar{g}: \boldsymbol{\theta}^i \to \boldsymbol{\theta}^i + b\mathbf{1}\}$. A maximal invariant for this group of transformations is $S_i - S_0$. Hence, among rules in $\mathscr{D}^0$ we now consider the invariant ones for which ψ_i is only a function of $R_i = S_i - S_0$.

Let $\Theta^i = \{(\theta_0, \theta_i)\}$, $\Theta_P{}^i = \{(\theta_0, \theta_i) \mid \theta_i \geqq \theta_0 + \Delta\}$, and $\Theta_N{}^i = \{(\theta_0, \theta_i) \mid \theta_i \leqq \theta_0\}$. If $\tau(\boldsymbol{\theta})$ is a distribution over Θ, denote by $\tau^i(\boldsymbol{\theta}^i)$ its marginal distribution over Θ^i. The expected risk corresponding to $\tau(\boldsymbol{\theta})$ of using the rule $\psi_i(\mathbf{S}^i)$ in the ith component problem is then

$$r^{(i)}(\tau, \psi_i) = L_1 \textstyle\int_{\Theta_P{}^i} E_{\boldsymbol{\theta}^i}[1 - \psi_i]\, d\tau^i(\boldsymbol{\theta}^i) + L_2 \int_{\Theta_N{}^i} E_{\boldsymbol{\theta}^i}[\psi_i]\, d\tau^i(\boldsymbol{\theta}^i) \tag{4.1}$$

where $E_{\boldsymbol{\theta}^i}[\psi_i] = \int_{\mathscr{S}_0} \int_{\mathscr{S}_i} \psi_i(\mathbf{s}^i) f_i(s_i - \theta_i) f_0(s_0 - \theta_0)\, ds_i\, ds_0$.

LEMMA 4.1. *Given $\psi_i(\mathbf{s}^i)$, a decision rule for the selection or rejection of the ith population, and a loss function defined by* (2.2), *then*

$$\sup_{\tau \in \Upsilon} r^{(i)}(\tau, \psi_i) = L_1 \pi_i - L_1 \pi_i \inf_{\boldsymbol{\theta}^i \in \Theta_P{}^i} E_{\boldsymbol{\theta}^i}[\psi_i] + L_2 \pi_i' \sup_{\boldsymbol{\theta}^i \in \Theta_N{}^i} E_{\boldsymbol{\theta}^i}[\psi_i].$$

PROOF. The result follows directly.

The following lemma is well known.

LEMMA 4.2. *If S_0, S_i are independent random variables having respective probability densities $f_j(s - \theta_j)$ for $j = 0$, i which are PF_2 functions, then $R_i = S_i - S_0$ has a probability density $g_i(r - (\theta_i - \theta_0))$ which is a PF_2 function.*

PROOF. The density of R_i is given by

$$g_i(r - (\theta_i - \theta_0)) = \textstyle\int_{-\infty}^{\infty} f_i(r + u - (\theta_i - \theta_0)) f_0(u)\, du. \tag{4.2}$$

This is the convolution of two PF_2 densities $f_0(-s - \theta_0)$ and $f_i(s - \theta_i)$ which is known to yield a PF_2 density. See Schoenberg [6].

We now interpret the component problem as a test of the hypothesis $H_0: \boldsymbol{\theta}^i \in \Theta_N{}^i$ against the alternative $H_1: \boldsymbol{\theta}^i \in \Theta_P{}^i$ with ψ_i denoting the probability of rejecting the null hypothesis.

THEOREM 4.1. *If $S_0, S_1, \cdots, S_k$ are independent random variables having respective probability densities $f_i(s-\theta_i)$ which are PF_2 functions, and if the loss function is given by* (2.2), *then a Υ-minimax decision rule in $\mathscr{D}^0$, ψ^Υ, is of the form: $\psi_i{}^\Upsilon = 1, 0$ as $R_i = S_i - S_0 \geqq, < d_i'$ where the d_i' are determined so that*

$$L_2 \pi_i' g_i(r) - L_1 \pi_i g_i(r-\Delta) \leqq, > 0 \tag{4.3}$$

as $r \geqq, < d_i'$, with $g_i(r)$ defined by (4.2).

PROOF. For the ith component test given above, application of the Hunt–Stein Theorem (see Lehmann [3] page 225 and page 336) yields that the invariant tests form an essentially complete class in $\mathscr{D}^0$. Hence if $\psi_i(\mathbf{s}^i)$ is a decision rule for the ith component problem and if we let $\alpha = \sup_{\boldsymbol{\theta}^i \in \Theta_N{}^i} E_{\boldsymbol{\theta}^i}[\psi_i]$, there exists a rule $\psi_i{}^*$ based only on R_i such that $\sup_{\boldsymbol{\theta}^i \in \Theta_N{}^i} E_{\boldsymbol{\theta}^i}[\psi_i{}^*] \leqq \alpha$ and such that $\inf_{\boldsymbol{\theta}^i \in \Theta_P{}^i} E_{\boldsymbol{\theta}^i}[\psi_i{}^*] \geqq \inf_{\boldsymbol{\theta}^i \in \Theta_P{}^i} E_{\boldsymbol{\theta}^i}[\psi_i]$. From Lemma 4.1 it follows that $\sup_{\tau \in \Upsilon} r^{(i)}(\tau, \psi_i{}^*) \leqq \sup_{\tau \in \Upsilon} r^{(i)}(\tau, \psi_i)$. If $\tau_0(\boldsymbol{\theta})$ is given by (3.2) and ψ_i is only a function of R_i, it follows that

$$r^{(i)}(\tau_0, \psi_i) = \int_{\mathscr{R}_i} [L_1 \pi_i (1-\psi_i(r_i)) g_i(r_i - \Delta) + L_2 \pi_i' \psi_i(r_i) g_i(r_i)]\, dr_i$$

where $\mathscr{R}_i$ is the sample space of R_i. Since $g_i(r)$ is a PF_2 density the Bayes rule with respect to $\tau_0(\boldsymbol{\theta})$ is of the form $\psi_i{}^\Upsilon = 1, 0$ as $R_i \geqq, < d_i'$ where d_i' is determined by (4.3). From Lemma 3.1 and Lemma 3.2 it follows that ψ^Υ is Υ-minimax among rules for which ψ_i is a function of R_i. Since such rules are essentially complete for $\mathscr{D}^0$, ψ^Υ is Υ-minimax among rules in $\mathscr{D}^0$.

We now apply Theorem 4.1 to the following problem. A sample of size n_i is taken from each of $k+1$ $N(\theta_i, \sigma^2)$ populations with σ^2 known. We confine attention to rules ψ where ψ_i is based on $\bar{X}_0$ and $\bar{X}_i$. Application of Theorem 4.1 proves that a Υ-minimax decision rule in $\mathscr{D}^0$ selects the ith population if and only if

$$\bar{X}_i - \bar{X}_0 \geqq \frac{\Delta}{2} + \frac{\sigma^2}{\Delta}\left(\frac{1}{n_i} + \frac{1}{n_0}\right) \ln (L_2 \pi_i' / L_1 \pi_i). \tag{4.4}$$

5. Comparison of a Υ-minimax rule with a Bayes rule. If a priori considerations yield a class of prior distributions over Θ, one method of utilizing such information is to select a member of the class and use the corresponding Bayes decision rule. Another approach is to find a rule which is Υ-minimax with respect to the class of priors given. Thus Bayes rules corresponding to prior distributions in Υ are natural competitors for a Υ-minimax rule.

As in the previous section, consider $k+1$ normal distributions $N(\theta_i, \sigma^2)$ with σ^2 known. Assume that $\theta_0, \theta_1, \cdots, \theta_k$ have independent normal priors $\tau_i(\theta_i)$ with known mean α_i and variance $\gamma_i{}^2$, respectively. The prior distribution over Θ is then

$$\tau^*(\boldsymbol{\theta}) = \textstyle\prod_{i=0}^k \tau_i(\theta_i). \tag{5.1}$$

As in the Υ-minimax derivation above, by the principle of sufficiency we reduce consideration to $\bar{\mathbf{X}} = (\bar{X}_0, \bar{X}_1, \cdots, \bar{X}_k)$. To find the Bayes rule with respect to $\tau^*(\boldsymbol{\theta})$, ψ^B, we find the Bayes procedure for each of the k component problems. This is accomplished by minimizing

$$\int_{\Theta_{P^i}} (1-\psi_i(\bar{\mathbf{x}}^i))\, dv_i^*(\boldsymbol{\theta}^i \mid \bar{\mathbf{x}}^i) + \int_{\Theta_{N^i}} \psi_i(\bar{\mathbf{x}}^i)\, dv_i^*(\boldsymbol{\theta}^i \mid \bar{\mathbf{x}}^i)$$

for $\bar{\mathbf{x}}^i$ where $v_i^*(\boldsymbol{\theta}^i \mid \bar{\mathbf{x}}^i)$ is the cumulative distribution function of the posterior distribution of $\boldsymbol{\theta}^i$ given $\bar{\mathbf{X}}^i = (\bar{X}_0, \bar{X}_i) = \bar{\mathbf{x}}^i$. But $v_i^*(\boldsymbol{\theta}^i \mid \bar{\mathbf{x}}^i) = v_i(\theta_i \mid \bar{x}_i) \times v_0(\theta_0 \mid \bar{x}_0)$ where $v_j(\theta_j \mid \bar{x}_j)$, the posterior distribution of θ_j given $\bar{x}_j$, is normal with mean $a_j = \{(n_j\bar{x}_j/\sigma^2)+(\alpha_j/\gamma_j^2)\}/\{(n_j/\sigma^2)+(1/\gamma_j^2)\}$ and variance $b_j^2 = (\sigma^2\gamma_j^2/n_j)/\{(\sigma^2/n_j)+\gamma_j^2\}$. The Bayes procedure will thus be of the form $\psi_i^B(\bar{x}_i, \bar{x}_0) = 1, 0$ as

$$L_2 P[\theta_i \leqq \theta_0 \mid \bar{x}_i, \bar{x}_0] - L_1 P[\theta_i \geqq \theta_0 + \Delta \mid \bar{x}_i, \bar{x}_0] \leqq, > 0$$

or as

$$L_2\, \Phi((a_0 - a_i)/(b_0^2 + b_i^2)^{\frac{1}{2}}) - L_1\, \Phi((a_i - a_0 - \Delta)/(b_0^2 + b_i^2)^{\frac{1}{2}}) \leqq, > 0$$

where $\Phi(z)$ denotes the cdf of an $N(0, 1)$ variate.

When $L_1 = L_2 = 1$, the procedure becomes $\psi_i^B(\bar{x}_i, \bar{x}_0) = 1, 0$ as

$$a_i - a_0 \geqq, < \Delta/2. \tag{5.2}$$

Under these independent normal priors, the probability that the ith population is positive is given by $\pi_i = \Phi((\alpha_i - \alpha_0 - \Delta)/(\gamma_0^2 + \gamma_i^2)^{\frac{1}{2}})$ and the probability that it is negative is given by $\pi_i' = \Phi((\alpha_0 - \alpha_i)/(\gamma_0^2 + \gamma_i^2)^{\frac{1}{2}})$.

Comparison is thus made between the Bayes procedure given in (5.2) which is based on the specific prior information of independent normal distributions with the Υ-minimax procedure which is based on the less stringent prior specifications of π_i and π_i' for $i = 1, \cdots, k$. Assume that $L_1 = L_2 = 1$ so that the expected risk becomes the expected number of wrong decisions where a wrong decision is defined to be the rejection of a positive population or the selection of a negative one. Since $r(\tau, \psi) = \sum_{i=1}^k r^{(i)}(\tau, \psi_i)$, it suffices to compare these procedures with respect to the selection or rejection of one population. Without loss of generality, we confine attention to the population corresponding to θ_1.

One meaningful comparison is found by examining the increase in expected risk which results from the use of the Υ-minimax procedure when $\boldsymbol{\theta}$ is distributed over Θ according to $\tau^*(\boldsymbol{\theta})$, as given by (5.1). Let $\omega_i^2 = \gamma_i^4/(\gamma_i^2 + (\sigma^2/n_i))$, $\alpha = \alpha_1 - \alpha_0$, $t_1 = \gamma_0^2 + \gamma_1^2$, $t_2 = \omega_0^2 + \omega_1^2$, and $t_3 = \sigma^2((1/n_0)+(1/n_1))$. The expected risk of the Bayes procedure is then

$$r^{(1)}(\tau^*, \psi_1^B) = F_0(-\alpha t_1^{-\frac{1}{2}}, (\alpha - \Delta/2)t_2^{-\frac{1}{2}}; -(t_2/t_1)^{\frac{1}{2}}) + F_0((\alpha - \Delta)t_1^{-\frac{1}{2}}, (-\alpha + \Delta/2)t_2^{-\frac{1}{2}}; -(t_2/t_1)^{\frac{1}{2}})$$

where $F_0(u_1, u_2; \rho)$ is the cumulative distribution function of a bivariate normal distribution with zero means, unit variances, and correlation coefficient ρ. The Υ-minimax procedure, ψ^{Υ}, given in (4.4) has expected risk

$$\begin{aligned} r^{(1)}(\tau^*, \psi_1^{\Upsilon}) = {} & F_0(-\alpha t_1^{-\frac{1}{2}}, (\alpha - \Delta/2 - t_3 \ln(\pi_1/\pi_1'))(t_1+t_3)^{-\frac{1}{2}}; -(t_1/(t_1+t_3))^{\frac{1}{2}}) \\ & + F_0((\alpha-\Delta)t_1^{-\frac{1}{2}}, -(\alpha - \Delta/2 - t_3 \ln(\pi_1/\pi_1'))(t_1+t_3)^{-\frac{1}{2}}; \\ & -(t_1/(t_1+t_3))^{\frac{1}{2}}). \end{aligned}$$

Table 1 exhibits $r^{(1)}(\tau^*, \psi_1^{B})$ and $r^{(1)}(\tau^*, \psi_1^{\Upsilon})$ when $\gamma_i^2 = \gamma^2$ and $n_i = n$, $i = 0, 1$. Here

$$\beta_1 = \frac{\alpha_1 - \alpha_0}{(2\gamma^2)^{\frac{1}{2}}} \quad \beta_2^2 = \frac{n\gamma^2}{\sigma^2} \quad \text{and} \quad \beta_3 = \frac{\Delta}{(2\gamma^2)^{\frac{1}{2}}}\,. \tag{5.3}$$

The Υ-minimax procedure is seen to have only slightly higher expected risk under independent normal priors in the cases given.

TABLE 1
Expected risks under normal priors

	$\beta_2^2 = .25$		$\beta_2^2 = 1.0$		$\beta_2^2 = 10.0$	
	Υ-*Minimax*	*Bayes*	Υ-*Minimax*	*Bayes*	Υ-*Minimax*	*Bayes*
			$\beta_3 = .2$			
$\beta_1 = -1.0$	.1151	.1142	.1151	.0968	.0922	.0333
$\beta_1 = -\ .3$	.3085	.2701	.3055	.1907	.0895	.0576
$\beta_1 = \ \ 0.0$	.3861	.3104	.2497	.2104	.0648	.0623
$\beta_1 = \ \ \ .3$	.3815	.3018	.3089	.2063	.0708	.0613
$\beta_1 = \ \ 1.0$	.1587	.1555	.1587	.1250	.1067	.0410
			$\beta_3 = 1.0$			
$\beta_1 = -1.0$	.0228	.0227	.0226	.0191	.0028	.0015
$\beta_1 = -\ .3$	.0967	.0910	.0782	.0598	.0044	.0036
$\beta_1 = \ \ 0.0$	.1532	.1361	.0948	.0793	.0048	.0044
$\beta_1 = \ \ \ .3$	.1825	.1717	.0962	.0925	.0050	.0049
$\beta_1 = \ \ 1.0$	.1532	.1367	.0948	.0793	.0048	.0044

Another important consideration is the comparison of $\sup_{\tau \in \Upsilon} r(\tau, \psi^B)$ and $\sup_{\tau \in \Upsilon} r(\tau, \psi^{\Upsilon})$. The Υ-minimax procedure is defined to be that procedure which minimizes the supremum of the expected risk when τ is in Υ. We know from the proof of Theorem 4.1 that

$$\begin{aligned} \sup_{\tau \in \Upsilon} r^{(1)}(\tau, \psi_1^{\Upsilon}) = r^{(1)}(\tau_0, \psi_1^{\Upsilon}) = {} & \pi_1 \Phi(\Delta^{-1} t_3^{\frac{1}{2}} \ln(\pi_1'/\pi_1) - \tfrac{1}{2}\Delta t_3^{-\frac{1}{2}}) \\ & + \pi_1' \Phi(-\Delta^{-1} t_3^{\frac{1}{2}} \ln(\pi_1'/\pi_1) - \tfrac{1}{2}\Delta t_3^{-\frac{1}{2}}) \end{aligned}$$

where $\tau_0 \in \Upsilon_0$ as defined in (3.1).

If $b_1^2(\sigma^2/n_1)^{-1} \neq b_0^2(\sigma^2/n_0)^{-1}$ then it is easily shown using Lemma 4.1 that $\sup_{\tau\in\Upsilon} r^{(1)}(\tau, \psi_1^B) = \pi_1 + \pi_1'$. Unless $\pi_i = \pi$ and $\pi_i' = \pi'$ for $i = 1, \cdots, k$, however, it will not always be true that $\sup_{\tau\in\Upsilon} r(\tau, \psi^B) = \sum_{i=1}^k (\pi_i + \pi_i')$. When $\gamma_i^2 = \gamma^2$ and $n_i = n$, $i = 0, 1$, the Bayes procedure is of the form $\psi_1^B = 1, 0$ as $\bar{X}_1 - \bar{X}_0 \geqq, < c$ for some real number c. In this case it is invariant under the group of translations and $\sup_{\tau\in\Upsilon} r^{(1)}(\tau, \psi_1^B)$ is achieved by any $\tau_0 \in \Upsilon_0$ as defined in (3.1). Thus

$$r^{(1)}(\tau_0, \psi_1^B) = \pi_1' + (\pi_1 - \pi_1')\Phi(\tfrac{1}{2}\beta_3(\beta_2^{-1} - \beta_2) - (\beta_1/\beta_2)).$$

A comparison of $r^{(1)}(\tau_0, \psi_1^\Upsilon)$ and $r^{(1)}(\tau_0, \psi_1^B)$ when $\gamma_i^2 = \gamma^2$ and $n_i = n$, $i = 0, 1$ is given in Table 2. Little difference in the supremum of the expected risk is exhibited in these cases. Here β_1, β_2^2 and β_3 are defined as in (5.3).

TABLE 2
Expected risks under a least favorable distribution

	$\beta_2^2 = .25$		$\beta_2^2 = 1.0$		$\beta_2^2 = 10.0$	
	Υ-Minimax	*Bayes*	*Υ-Minimax*	*Bayes*	*Υ-Minimax*	*Bayes*
			$\beta_3 = .2$			
$\beta_1 = -1.0$	.1151	.1235	.1151	.1936	.1150	.2721
$\beta_1 = -\ .3$	.3085	.3608	.3085	.3813	.2901	.3343
$\beta_1 = \ 0.0$	.4200	.4361	.4109	.4207	.3420	.3452
$\beta_1 = \ .3$	.3821	.4195	.3805	.4125	.3306	.3430
$\beta_1 = \ 1.0$	.1587	.1776	.1587	.2501	.1583	.2933
			$\beta_3 = 1.0$			
$\beta_1 = -1.0$	.0228	.0232	.0227	.0383	.0102	.0198
$\beta_1 = -\ .3$	.0968	.1081	.0940	.1196	.0248	.0295
$\beta_1 = \ 0.0$	.1581	.1755	.1425	.1587	.0307	.0328
$\beta_1 = \ .3$	.2274	.2339	.1815	.1849	.0344	.0347
$\beta_1 = \ 1.0$	.1581	.1755	.1425	.1587	.0307	.0328

Consider a third distribution over Θ. Assume the prior distribution over the treatment population parameters is conditioned on θ_0, the value of the control population parameter. This situation may arise, for example, when a treatment which may be slightly effectual (such as a saline solution) is administered to the control population and the same preparation only with additional ingredients is given to the treatment populations. Assume a uniform prior for θ_0 over the interval $[-1, 1]$ and denote its cumulative distribution function by $\tau^0(\theta_0)$. Assume also that the distribution function of the prior over Θ given θ_0 is $\tau^1(\boldsymbol{\theta} \mid \theta_0) = \prod_{i=1}^k \tau_i^1(\theta_i \mid \theta_0)$ where $\tau_i^1(\theta_i \mid \theta_0)$ is the distribution function of a uniform distribution on the interval

$$\left[\theta_0 - \frac{\Delta\pi_i'}{1 - \pi_i - \pi_i'},\ \theta_0 + \frac{\Delta(1 - \pi_i')}{1 - \pi_i - \pi_i'}\right].$$

Hence $P[\theta_i \geqq \theta_0+\Delta]=\pi_i$ and $P[\theta_i \leqq \theta_0]=\pi_i'$ for $i=1,\cdots,k$. Let $\tau_1(\boldsymbol{\theta})=\int_{-\infty}^{\theta_0}[\prod_{i=1}^k \tau_i^1(\theta_i\,|\,\mu)]\,d\tau^0(\mu)$. For a procedure, $\psi_1{}^c$, which selects θ_1 if and only if

$$c_1\bar{X}_1-c_0\bar{X}_0 \geqq d_1 \tag{5.4}$$

where c_1, c_0 and d_1 are real numbers, the expected risk is

$$r^{(1)}(\tau_1,\psi_1{}^c)=\int_{-1}^{1}\int_{\theta_0-\Delta q_1'}^{\theta_0}\pi_1'(2\Delta q_1')^{-1}\Phi(h(\theta_1,\theta_0))\,d\theta_1\,d\theta_0$$
$$+\int_{-1}^{1}\int_{\theta_0+\Delta}^{\theta_0+\Delta q_1}\pi_1'(2\Delta q_1')^{-1}\Phi(-h(\theta_1,\theta_0))\,d\theta_1\,d\theta_0$$

where $q_1=(1-\pi_1')/(1-\pi_1-\pi_1')$, $q_1'=\pi_1'/(1-\pi_1-\pi_1')$ and

$$h(\theta_1,\theta_0)=(c_1\theta_1-c_0\theta_0-d_1)/\{(\sigma^2c_0{}^2/n_0)+(\sigma^2c_1{}^2/n_1)\}^{\frac{1}{2}}.$$

For the Υ-minimax procedure given in (4.4) $c_0=c_1=1$ and $d_1=(\Delta/2)+t_3\ln(\pi_1'/\pi_1)$. The parameters of the normal priors for the procedure in (5.2) must satisfy $t_1T^2=\Delta^2$ and $\alpha T=\Delta\Phi^{-1}(\pi_1')$ where $T=\Phi^{-1}(\pi_1)+\Phi^{-1}(\pi_1')$. Subject to these restrictions the Bayes procedure is then in the form of (5.4) with $c_i=b_i{}^2(\sigma^2/n_i)^{-1}$, $i=0,1$, and

$$d_1=\Delta/2-b_1{}^2\Phi^{-1}(\pi_1')\Delta(T\gamma_1{}^2)^{-1}+\alpha_0[b_1{}^2(\sigma^2/n_1)^{-1}-b_0{}^2(\sigma^2/n_0)^{-1}].$$

Assume $n_i=n$, $i=0,1$. Table 3 displays $r^{(1)}(\tau_1,\psi_1{}^B)$ and $r^{(1)}(\tau_1,\psi_1{}^{\Upsilon})$ for several values of π_1 and π_1' and levels of Δ, σ^2/n and the normal parameters. In this comparison the Bayes procedure shows a slight improvement over the Υ-minimax procedure in most cases. The improvement diminishes as π_1 approaches π_1'.

The expected risks of the Υ-minimax procedure given in (4.4) and the Bayes procedure given in (5.2) have been compared for three distributions in Υ. The Υ-minimax procedure compares favorably with the given Bayes procedure in terms of expected risk. Moreover, the Υ-minimax procedure has the advantage of requiring less prior information than the Bayes procedure.

6. Comments on a theorem by Y. L. Tong. One problem considered by Tong in reference [7] is the partitioning of a set of normal populations with known common variance into two subsets according to the relationship of their respective translation parameters to that of a control population. His formulation of the problem is similar to our own except he assumes no prior information about the translation parameters. In Theorem 1.3 of his paper, Tong states that (i) his decision rule (1.3) is minimax among invariant rules because (ii) in each component problem it is the Bayes rule among invariant rules with respect to the least favorable prior distribution (1.28). He has tacitly assumed, apparently, that the decision in the ith component problem depends only on $\bar{X}_i$ and $\bar{X}_0$. The following example shows that without such a restriction (ii) is not correct.

Using Tong's notation, let $k=2$, $\delta_2{}^*-\Delta=\delta_1{}^*=0$, $Z_i=\bar{X}_i-\bar{X}_0$ for $i=1,2$ and $\sigma^2=n/2$. Then $\mathbf{Z}=(Z_1,Z_2)$ has a bivariate normal distribution with respective means $\mu_1-\mu_0$ and $\mu_2-\mu_0$, unit variances, and a covariance of $\frac{1}{2}$. Denote the density of $\mathbf{Z}$ by $h_{\mu_1-\mu_0,\,\mu_2-\mu_0}(z_1,z_2)$. Any G invariant decision rule may be considered a function of $\mathbf{Z}$, a maximal invariant. It follows that among G invariant

TABLE 3
Expected risks under a uniform distribution

		Υ-*Minimax*	*Bayes*								
			$\alpha_0=-.75$			$\alpha_0=0.0$			$\alpha_0=.75$		
			$\gamma_1^2=\frac{\Delta^2}{8T^2}$	$\gamma_1^2=\frac{\Delta^2}{2T^2}$	$\gamma_1^2=\frac{7\Delta^2}{8T^2}$	$\gamma_1^2=\frac{\Delta^2}{8T^2}$	$\gamma_1^2=\frac{\Delta^2}{2T^2}$	$\gamma_1^2=\frac{7\Delta^2}{8T^2}$	$\gamma_1^2=\frac{\Delta^2}{8T^2}$	$\gamma_1^2=\frac{\Delta^2}{2T^2}$	$\gamma_1^2=\frac{7\Delta^2}{8T^2}$
			$\pi_1=.15$ $\pi_1'=.80$								
$\Delta=.15$	$\sigma^2/n=.10$	.150	.143	.089	.110	.132	.089	.087	.145	.089	.107
$\Delta=.50$	$\sigma^2/n=.10$	.046	.028	.017	.015	.024	.017	.015	.021	.017	.016
$\Delta=.50$	$\sigma^2/n=1.0$	.150	.135	.085	.066	.128	.085	.068	.121	.085	.076
			$\pi_1=.40$ $\pi_1'=.50$								
$\Delta=.15$	$\sigma^2/n=.10$	.208	.345	.179	.265	.332	.179	.199	.382	.179	.249
$\Delta=.50$	$\sigma^2/n=.10$	.033	.044	.032	.031	.039	.032	.029	.040	.032	.031
$\Delta=.50$	$\sigma^2/n=1.0$	.196	.285	.170	.142	.275	.170	.133	.285	.170	.147
			$\pi_1=.75$ $\pi_1'=.20$								
$\Delta=.15$	$\sigma^2/n=.10$	.200	.200	.094	.117	.151	.094	.091	.200	.094	.108
$\Delta=.50$	$\sigma^2/n=.10$	.036	.019	.016	.016	.020	.016	.015	.022	.016	.015
$\Delta=.50$	$\sigma^2/n=1.0$	.199	.131	.089	.079	.142	.089	.072	.154	.089	.071

rules, the Bayes decision rule for the first component problem with respect to the prior distribution (1.28), selects (rejects) the first population as

$$(6.1) \qquad h_{0,0}(z_1, z_2)+h_{0,\Delta}(z_1, z_2)-h_{\Delta,0}(z_1, z_2)-h_{\Delta,\Delta}(z_1, z_2) <(>)0.$$

Substituting $\Delta = (200/99)z_1 = -z_2 = (\frac{3}{2})^{\frac{1}{2}}$, the left-hand side of equation (6.1) is seen to be negative. This contradicts the fact that the first component of Tong's decision rule (1.3) is Bayes among invariant rules for the first component problem.

It is correct, however, to conclude that rule (1.3) is minimax among decision rules which are G invariant and for which D_i, the decision concerning the ith population, depends only on $\bar{X}_i$ and $\bar{X}_0$. Moreover, application of the Hunt–Stein theorem as in our Theorem 4.1 would yield that Tong's decision rule (1.3) is minimax among all rules for which D_i depends only on $\bar{X}_i$ and $\bar{X}_0$.

Acknowledgment. The authors would like to thank Professor I. R. Savage for contributing helpful comments and Professor W. J. Hall for useful suggestions. This paper is based on part of a dissertation written by R. H. Randles under the direction of M. Hollander at Florida State University.

REFERENCES

[1] BLUM, J. R. and ROSENBLATT, J. (1967). On partial a priori information in statistical inference. *Ann. Math. Statist.* **38** 1671–1678.

[2] HODGES, J. L., JR. and LEHMANN, E. L. (1952). The use of previous experience in reaching statistical decisions. *Ann. Math. Statist.* **23** 396–407.

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

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

[5] MENGES, G. (1966). On the "Bayesification" of the minimax principle. *Unternehmensforschung* **10** 81–91.

[6] SCHOENBERG, T. S. (1951). On Pólya frequency functions. *J. Analyse Math.* **1** 331–374.

[7] TONG, Y. L. (1969). On partitioning a set of normal populations by their locations with respect to a control. *Ann. Math. Statist.* **40** 1300–1324.