r-MINIMAX AND MINIMAX DECISION RULES
FOR COMPARISON OF TREATMENTS WITH A CONTROL*
by
Shanti S. Gupta and Woo-Chul Kim
Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #79-10

August 1979

*This research was supported by the Office of Naval Research contract NO0014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.
1. Introduction.

In many fields of research one is faced with the problem of comparing k experimental categories with reference to a 'standard' or a 'control'. Following the initial investigation by Paulson (1952), this problem has been studied in several different formulations by Dunnett (1955), Gupta and Sobel (1958) and Lehmann (1961) among others.

Let \( \pi_1, \ldots, \pi_k \) denote the k experimental categories or 'treatment' populations and let \( \pi_0 \) denote the 'control' population, where the quality of each population \( \pi_i \) is characterized by a real-valued parameter \( \theta_i \) (\( i = 0, 1, \ldots, k \)). Each treatment population \( \pi_i \) is said to be 'superior', 'equivalent' or 'inferior' to the control population \( \pi_0 \) if \( \theta_i - \theta_0 \geq \Delta \), \( -\Delta < \theta_i - \theta_0 < \Delta \), \( \theta_i - \theta_0 \leq -\Delta \), respectively, where \( \Delta \) is a given positive constant. We consider a problem in which the treatment populations are to be classified as one of the above three cases based on the observations from the populations. Bhattacharyya (1956, 1958) studied this problem for the normal populations with unknown means when the control population is assumed known. A similar problem has been considered by Seeger (1972). We apply the \( \gamma \)-minimax principle to this problem.

*This research was supported by the Office of Naval Research contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.
\[ \Gamma \text{-minimax principle is known as one of the techniques for the use of incomplete prior information. Such an idea was first used by Robbins (1951) and independently by Hodges and Lehmann (1952) and Menges (1966). The name } \Gamma \text{-minimax was first used by Blum and Rosenblatt (1967). Randles and Hollander (1971) applied such a principle to a problem of selecting the treatments 'better' than the control. It has been applied to various problems, and recently to selection problems by Gupta and Huang (1975, 1977), Berger (1977) and Miescke (1979). \]

In Section 2, necessary notations, definitions, a loss function and the incomplete prior are introduced. A lemma is given to help find \( \Gamma \)-minimax rules. Section 3 treats the case of known control population, and a \( \Gamma \)-minimax rule and a minimax rule are derived. In Section 4, the case in which the control population parameter \( \theta_0 \) is unknown is treated. Rules are derived which are \( \Gamma \)-minimax among rules for which the decision about the \( i \)-th population depends only on the observations from \( \pi_i \) and \( \pi_0 \). A minimax rule is also derived. A normal means problem and a normal variances problem are given as specific examples. Section 5 consists of comparisons of \( \Gamma \)-minimax rules with Bayes rules for independent normal priors for the normal means problem.

2. Formulation of the problem.

Let \( X_0, X_1, \ldots, X_k \) be \( k+1 \) independent random variables representing the control population \( \pi_0 \) and the \( k \) treatment populations \( \pi_1, \ldots, \pi_k \), respectively, with \( X_i \) having pdf \( f_i(x; \theta_i) \) with respect to the Lebesgue measure on the real line \( R \) where \( \theta_i \in \Theta = R \), \( i = 0, 1, \ldots, k \). The random variables \( X_0, \ldots, X_k \) may be sufficient statistics or other statistics based on which we wish to make
statistical decisions. We assume that each \( f_i(\cdot) (i = 0, 1, \ldots, k) \) is symmetric about the origin and strongly unimodal, i.e., \( f_i(\cdot) \) is log-concave on the real line. Hence \( f_i(x-\theta_i) \) has the monotone likelihood ratio (MLR) property.

Obviously, we do not need any observations from \( \pi_0 \) when \( \theta_0 \) is assumed known; therefore, it will be understood that, in such a case, the random variable \( X_0 \) is deleted from our consideration.

The action space \( G \) can be written as \( G = G_1 \times \ldots \times G_k \) where \( G_i = \{1, 2, 3\} \) for \( i = 1, \ldots, k \). The action \( a = (a_1, \ldots, a_k) \in G \) is to be interpreted in such a way that, for \( i = 1, \ldots, k \), the treatment population \( \pi_i \) is classified as 'inferior', 'equivalent' and 'superior' to \( \pi_0 \) for \( a_i = 1, 2, 3 \), respectively.

The loss \( L(\varphi, a) \) incurred by the action \( a \in G \) for \( \varphi = (\theta_0, \ldots, \theta_k) \) is assumed to be of the following form.

\[
L(\varphi, a) = \sum_{i=1}^{k} L_i(\varphi, a_i) \tag{2.1}
\]

where \( L_i(\varphi, a_i) \) is defined as in the following table;

<table>
<thead>
<tr>
<th>State of nature</th>
<th>( a_i )</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0_i - \varphi_0 \leq -\Lambda_2 )</td>
<td>0</td>
<td>( \varepsilon_1 )</td>
<td>( \varepsilon_1 + \varepsilon_3 )</td>
<td></td>
</tr>
<tr>
<td>( -\Lambda_2 &lt; 0_i - \varphi_0 \leq -\Lambda_1 )</td>
<td>0</td>
<td>0</td>
<td>( \varepsilon_4 ) ( (\varepsilon_i \geq 0, i=1, \ldots, 4) )</td>
<td></td>
</tr>
<tr>
<td>(</td>
<td>0_i - \varphi_0</td>
<td>&lt; \Lambda_1 )</td>
<td>( \varepsilon_2 )</td>
<td>0</td>
</tr>
<tr>
<td>( \Lambda_1 \leq 0_i - \varphi_0 &lt; \Lambda_2 )</td>
<td>( \varepsilon_4 )</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>( 0_i - \varphi_0 \geq \Lambda_2 )</td>
<td>( \varepsilon_1 + \varepsilon_3 )</td>
<td>( \varepsilon_1 )</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Here, \( \Lambda_1 = \Lambda - \varepsilon \), \( \Lambda_2 = \Lambda + \varepsilon \) for a given constant \( \varepsilon: 0 \leq \varepsilon < \Lambda \) and it will be
understood that the second row and the fourth row will disappear when $\epsilon = 0$.

Bhattacharyya (1956) derived a minimax rule assuming the above loss function
with $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 1$ and $\epsilon = 0$ when $\theta_0$ is assumed known and $\theta_1, \ldots, \theta_k$ are
the unknown means of normal distributions. However, the irregularity of such
a loss function has been pointed out in the sense that the minimax risk does not
tend to zero even if the sample sizes increase indefinitely, and the same
problem has been studied afresh by Bhattacharyya (1958) assuming the above loss
function with $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 1$ and $\epsilon > 0$. Note that the above loss
function with $\epsilon > 0$ assumes the indifference zones.

For given $\bar{x} = (x_0, x_1, \ldots, x_k)$ consider decision rules of the form

$$\delta(x) = (\delta_1(x), \ldots, \delta_k(x))$$

(2.2)

where $\delta_i(\bar{x}) = (\delta_i(1|\bar{x}), \delta_i(2|\bar{x}), \delta_i(3|\bar{x}))$ and, for $j = 1, 2, 3$, $\delta_i(j|\bar{x})$
denotes the conditional probability of taking action $j$ in the $i$-th component
decision problem. Note that there is no loss of generality in considering
decision rules of the form given in (2.3). The risk function of a rule
$\delta$ for fixed $\bar{q}$ is then $R(\bar{q}, \delta) = \sum_{i=1}^{k} R_i(\bar{q}, \delta_i)$ where $R_i(\bar{q}, \delta_i) = E_{\bar{q}}[L_i(\bar{q}, \delta_i(\bar{x}))].$

For a prior distribution $\tau(\bar{q})$ of $\bar{q}$, the overall risk of a rule $\delta$ wrt $\tau$ is
denoted by $r(\tau, \delta) = \sum_{i=1}^{k} r_i(\tau, \delta_i)$ where $r_i(\tau, \delta_i) = \int R_i(\bar{q}, \delta_i) d\tau(\bar{q}).$

It is assumed that partial prior information is available to a decision
maker such that, for each $i$, he can specify $\gamma_i = P[|\theta_i - \theta_0| \geq \Delta_2]$ and
$\gamma_i' = P[|\theta_i - \theta_0| < \Delta_1]$ where $\gamma_i + \gamma_i' \leq 1$ for $i = 1, \ldots, k$. Let $\Gamma$ denote the
class of all such prior distributions, i.e.,

$$\Gamma = \{\tau(\bar{q}) : \int_{|\theta_i - \theta_0| \geq \Delta_2} d\tau(\bar{q}) = \gamma_i, \int_{|\theta_i - \theta_0| < \Delta_1} d\tau(\bar{q}) = \gamma_i' \text{ for } i = 1, \ldots, k\}.$$  

(2.3)

Note that when $\epsilon = 0$, i.e., $\Delta_1 = \Delta_2$, $\gamma_i + \gamma_i' = 1$. 


A rule \( \delta' \) is called a \( \tau \)-minimax rule if 
\[
\sup_{\tau \in \Gamma} r(\tau, \delta') = \inf_{\delta} \sup_{\tau \in \Gamma} r(\tau, \delta),
\]
and \( \sup_{\tau \in \Gamma} r(\tau, \delta') \) is called the \( \tau \)-minimax value. The next result is useful to find the \( \tau \)-minimax rule.

**Lemma 2.1.** Suppose \( \{\tau_n, n = 1, 2, \ldots\} \) is a sequence of priors in \( \Gamma \). If 
\[
\lim_{n} \inf_{\delta} r(\tau_n, \delta) \geq c \quad \text{and if} \quad \sup_{\tau \in \Gamma} r(\tau, \delta') \leq c,
\]
then \( \delta' \) is a \( \tau \)-minimax rule and \( c \) is the \( \tau \)-minimax value.

**Proof.** The result follows from the following inequalities.

\[
\sup_{\tau \in \Gamma} \inf_{\delta} r(\tau, \delta) \geq \lim_{n} \inf_{\delta} r(\tau_n, \delta) \geq c \geq \sup_{\tau \in \Gamma} r(\tau, \delta') \geq \inf_{\delta} \sup_{\tau \in \Gamma} r(\tau, \delta) \geq \sup_{\tau \in \Gamma} \inf_{\delta} r(\tau, \delta).
\]

3. **Known control population**

In this section \( \theta_0 \) is assumed known and thus we may assume \( \theta_0 = 0 \) without loss of generality. Hence \( \bar{x} \) and \( \underline{x} \) in this section denote \( (x_1, \ldots, x_k) \) and \( (u_1, \ldots, u_k) \), respectively. Let us consider a rule \( \delta(x) \) of the form in (2.2) where \( \delta_i(j|x) \) \( (j = 1, 2, 3) \) is given by

\[
\begin{align*}
\delta_i(1|x) &= I(\infty, -d_i](x_i), \\
\delta_i(2|x) &= I(-d_i, d_i) (x_i), \\
\delta_i(3|x) &= I[d_i, \infty)(x_i), \\
\end{align*}
\]

for \( 0 < d_i \leq \infty \) and \( i = 1, \ldots, k \).
Lemma 3.1. Suppose that a decision rule $\delta(x)$ is given by (2.2) and (3.1).

Then, for $i = 1, \ldots, k$,

$$\sup_{\tau \in \Gamma} r_i(\tau, \delta_i) \leq v_i$$

where $v_i = \int_{-\infty}^{d_i} [\varepsilon_3 \gamma_1 f_i(x+\Delta_2) + \varepsilon_2 \gamma_1 f_i(x-\Delta_1) + f_i(x+\Delta_1) + \varepsilon_4 (1-\gamma_1 - \gamma_1^* f_i(x+\Lambda_1)] dx$

$$+ \int_{d_i}^{\infty} \varepsilon_1 \gamma_1 f_i(x-\Delta_2) dx.$$

Proof. It follows from the definition of $L_1(\theta, a_i)$ and the symmetry of $f_i(\cdot)$ that, for $|\theta_i| < \Delta_1$,

$$R_i^\prime(\theta, \delta_i) = \varepsilon_2 [f_i(d_i-\theta_i) - f_i(-d_i-\theta_i)]$$

$$= \varepsilon_2 f_i(\theta_i + d_i) \left[ \frac{f_i(\theta_i - d_i)}{f_i(\theta_i + d_i)} - 1 \right],$$

where $R_i^\prime$ denotes the derivative of $R_i$ wrt $\theta_i$.

It follows from the MLR property of $f_i(x-\theta_i)$ that $R_i^\prime(\theta, \delta_i)$ has at most one change of sign, from negative to positive if there is any sign change at all; therefore, $R_i(\theta, \delta_i)$ attains the supremum over $\theta_i \in (-\Delta_1, \Delta_1)$ at either $\theta_i = -\Delta_1$ or $\theta_i = \Delta_1$. Hence, for $|\theta_i| < \Delta_1$,

$$R_i(\theta, \delta_i) \leq \varepsilon_2 \int_{d_i}^{\infty} [f_i(x-\Delta_1) + f_i(x+\Delta_1)] dx.$$

It can be easily shown that

$$R_i(\theta, \delta_i) \begin{cases} \leq \varepsilon_1 \int_{-\infty}^{d_i} f_i(x-\Delta_2) dx + \varepsilon_3 \int_{d_i}^{\infty} f_i(x+\Delta_2) dx \quad \text{for } |\theta_i| > \Lambda_2, \\
\leq \varepsilon_4 \int_{d_i}^{\infty} f_i(x+\Delta_1) dx \quad \text{for } \Lambda_1 \leq |\theta_i| \leq \Lambda_2. \end{cases}$$
Therefore, it follows from (2.3) that \( \sup_{\tau \in \Gamma} r_i(\tau, \delta_i) \leq \nu_i \) which completes
the proof.

Now we derive a \( \Gamma \)-minimax rule for the case where \( \theta_0 \) is known.

**Theorem 3.1.** Assume that independent random variables \( X_1, \ldots, X_k \) have
\( f_1(x_1 - \theta_1), \ldots, f_k(x_k - \theta_k) \), respectively, with \( f_i(\cdot) \) being symmetric and
strongly unimodal, and that the loss function is given by (2.1). Then
the \( \Gamma \)-minimax rule \( \delta^\Gamma \) is given by (2.2) and (3.1) where each \( d_i = d^\Gamma_1 \) in
(3.1) is defined by \( d_1^\Gamma = \max(c_i, 0) \) with \( c_i \) being determined by

\[
\gamma_1 \gamma_i f_i(x + \Delta_2)^{\gamma_2} \gamma_1^i (f_i(x - \Delta_1) + f_i(x + \Delta_1)) + \gamma_1^i (1 - \gamma_i - \gamma_i^i) f_i(x + \Delta_1) \\
\leq \gamma_1 \gamma_i f_i(x - \Delta_2) \quad \text{as} \quad x \geq c_i.
\]

(3.2)

**Proof.** The existence of a \( c_i \) satisfying (3.2) follows from the MLR
property of \( f_i(x - \theta_i) \). Therefore, the decision rule \( \delta^\Gamma \) is well defined.

First, we will consider the case when \( \epsilon > 0 \), i.e., \( \Delta_2 > \Delta_1 \). For \( n > \Delta_1^{-1} \),
let \( \nu_i \) be a prior distribution in \( \Gamma \) under which \( \theta_1, \ldots, \theta_k \) are independent,
\( P(\nu_i = \Delta_2) = P(\nu_i = -\Delta_2) = \gamma_i / 2 \), \( P(\nu_i = \Delta_1) = P(\nu_i = -\Delta_1) = (1 - \gamma_i - \gamma_i^i) / 2 \) and
\( P(\nu_i = \Delta_1^{-1}) = P(\nu_i = -\Delta_1^{-1}) = \gamma_i^i / 2 \) for \( i = 1, \ldots, k \). Then it can be
easily verified that \( \inf_{\delta} r_i(\nu_i, \delta) = \sum_{i=1}^k \inf_{\delta_i} r_i(\nu_i, \delta_i) \) and, for \( i = 1, \ldots, k, \)

\[
\inf_{\delta_i} r_i(\nu_i, \delta_i) = \int_{-\infty}^{\infty} p_i(x) dx / 2,
\]

where \( p_i(x) = \min\{p_i(1, x), p_i(2, x), p_i(1, -x)\} \) with \( p_i(2, x) = \gamma_1 \gamma_i^i f_i(x + \Delta_2) + f_i(x - \Delta_2) \) and \( p_i(1, x) = \gamma_2 \gamma_i^i f_i(x - \Delta_1^{-1}) + f_i(x + \Delta_1^{-1}) \) + \( \gamma_i (1 - \gamma_i - \gamma_i^i) f_i(x - \Delta_1) + (\gamma_i + \gamma_i^i) \gamma_i f_i(x - \Delta_2) \). Since \( f_i(\cdot) \) is strongly unimodal on the real line, \( f_i(\cdot) \) is
continuous and thus \( p_n(x) \) converges, as \( n \to \infty \), to \( p(x) = \min(p(1,x), p(2,x), p(1,-x)) \) where \( p(1,x) = \lim_{n \to \infty} p_n(1,x) \). Note that \( p(1,x) \geq p(1,-x) \) if and only if \( x \geq 0 \). This follows from the fact that, for any \( t > 0 \) \( f_i(x-t) \geq f_i(x+t) \) if and only if \( x \geq 0 \). Since \( p_n(x) \) is bounded above by \( p(2,x) \) which is integrable, it follows from the Lebesgue convergence theorem that

\[
\liminf_{n \to \infty} r_i(\tau_n, \delta_i) = \frac{1}{2} \int_{-\infty}^{\infty} p(x) dx = \int_{0}^{\infty} \min(p(2,x), p(1,-x)) dx.
\tag{3.3}
\]

Note that \( \int_{0}^{\infty} \min(p(2,x), p(1,-x)) dx \) can be written as

\[
\int_{0}^{\infty} \min(p(2,x), p(1,-x)) dx = \int_{0}^{\infty} \min(\varepsilon_3 \gamma_i f_i(x+\Delta_{2}) + \varepsilon_4 (1-\gamma_i - \gamma_i') f_i(x+\Delta_{1}) + \varepsilon_2 \gamma_i' (f_i(x+\Delta_{1}) + f_i(x-\Delta_{1})) \] 
\tag{3.4a}
\]

\[
+ \int_{-\infty}^{0} 2 \gamma_i f_i(x-\Delta_2) dx + \int_{-\infty}^{0} 2 \gamma_i' f_i(x-\Delta_2) dx
\]

\[
= \int_{0}^{\infty} \left[ \varepsilon_3 \gamma_i f_i(x+\Delta_{2}) + \varepsilon_4 (1-\gamma_i - \gamma_i') f_i(x+\Delta_{1}) + \varepsilon_2 \gamma_i' (f_i(x+\Delta_{1}) + f_i(x-\Delta_{1})) \right] dx
\tag{3.4b}
\]

\[
+ \int_{-\infty}^{0} \gamma_i f_i(x-\Delta_2) dx,
\]

where \( d_i = \max(0, c_i) \) with \( c_i \) defined as in (3.2).

It follows from Lemma 3.1 that \( \liminf_{n \to \infty} r_i(\tau_n, \delta_i) \geq \sup_{\tau \in \Gamma} r_i(\tau, \delta_i') \). Therefore,

\[
\liminf_{n \to \infty} r(\tau_n, \delta) = \lim_{k \to \infty} \sum_{i=1}^{k} \inf_{\delta_i} r_i(\tau_n, \delta_i)
\geq \sum_{i=1}^{k} \sup_{\tau \in \Gamma} r_i(\tau, \delta_i')
\geq \sup_{\tau \in \Gamma} r_i(\tau, \delta_i').
\]
Hence Lemma 2.1 yields that $\delta^r$ is a $\Gamma$-minimax rule. This completes the proof of the case when $\epsilon > 0$. Note that $\Delta_1 = \Delta_2 = \Delta$ and $\gamma_i + \gamma_i^r = 1$ for $i = 1, \ldots, k$ if $\epsilon = 0$. When $\epsilon = 0$, let us consider a sequence of prior distributions, $\{\pi_n, n > n^{-1}\}$, under which $\theta_1, \ldots, \theta_k$ are independent, $P(\theta_i = \Delta) = P(\theta_i = -\Delta) = \gamma_i/2$ and $P(\theta_i = -\Delta - n^{-1}) = P(\theta_i = -\Delta + n^{-1}) = \gamma_i^r/2$ for $i = 1, \ldots, k$. Then we can prove in the exactly same manner as the above that $\delta^r$ is a $\Gamma$-minimax rule.

Now we discuss the derivation of the minimax rule for some special cases.

A minimax rule can be derived from the arguments in the proof of Theorem 2.1. For this purpose, assume that $\varepsilon_1 = \varepsilon_2$, $\varepsilon_4 \leq 2\varepsilon_1$ and $\varepsilon_3 \leq \varepsilon_1$. We may assume that $\varepsilon_1 = \varepsilon_2 = 1$ without loss of generality. Let us consider a rule $\delta^*$ of the type given by (2.2) and (3.1) where each $d_i = d_i^*$ in (3.1) is determined so that, for $F_i(x) = \int_{-\infty}^{\infty} f_i(t)dt$,

$$F_i(d_i - \Delta_2) + \gamma_3 F_i(-d_i + \Delta_2) = F_i(-d_i - \Delta_1) + F_i(-d_i + \Delta_1).$$

(3.3)

Note that the existence of such a non-negative $d_i^*$ follows from the strong unimodality and the symmetry of $f_i(\cdot)$. Let us define $\gamma_i$ and $\gamma_i^r = 1 - \gamma_i$ for $i = 1, \ldots, k$ by

$$\gamma_i = \frac{[f_i(d_i^* - \Delta_1) + f_i(d_i^* + \Delta_1)]/[f_i(d_i - \Delta_2) + \gamma_3 f_i(d_i + \Delta_2) + f_i(d_i^* - \Delta_1) + f_i(d_i^* + \Delta_1)]}{[f_i(d_i^* - \Delta_1) + f_i(d_i^* + \Delta_1)]/[f_i(d_i - \Delta_2) + \gamma_3 f_i(d_i + \Delta_2) + f_i(d_i^* - \Delta_1) + f_i(d_i^* + \Delta_1)]}.$$

Since $\gamma_i \in [0, 1]$, we can consider a family of prior distributions, $\Gamma$, given by (2.3). Then it follows from Theorem 2.1 that the corresponding $\Gamma$-minimax rule is of the same type as $\delta^*$ except that now $d_i^r = \max(c_i, 0)$ where $c_i$ is determined so that

$$H(c_i) = \gamma_i [\gamma_3 f_i(c_i + \Delta_2) - f_i(c_i - \Delta_2)] + \gamma_i^r [f_i(c_i - \Delta_1) + f_i(c_i + \Delta_1)] = 0.$$

Since $H(d_i^r) = 0$ and $d_i^r \geq 0$, $d_i^* = d_i^r$, i.e., the rule $\delta^*$ is the $\Gamma$-minimax.
rule; therefore it follows from the arguments in the proof of Theorem 2.1 that
\[
\liminf_{n \to \infty} r(\tau_n, \delta) = \frac{1}{k} \sum_{i=1}^{k} \gamma_i \left[ F_i (d_{i1}^x - \Delta_2) + \gamma_3 F_i (d_{i2}^x - \Delta_2) \right] + \gamma_i \left[ F_i (-d_{i1}^x - \Delta_1) + F_i (-d_{i2}^x + \Delta_1) \right]
\]
\[
= \frac{1}{k} \sum_{i=1}^{k} \left[ F_i (-d_{i1}^x - \Delta_1) + F_i (-d_{i2}^x + \Delta_1) \right]
\]
\[
= \frac{1}{k} \sup_{\theta, \delta^*} R_i(\theta, \delta^*)
\]
\[
\geq \sup_{\theta, \delta^*} R(\theta, \delta^*)
\]
Therefore, we have the next result which includes the results in Bhattacharyya (1956, 1958) as special cases.

**Corollary 3.1.** Under the assumptions in Theorem 3.1, if \( \lambda_1 = \lambda_2 = 1, \lambda_3 \leq 1 \)
and \( \lambda_4 \leq 2 \), then a rule \( \delta^M \) of the type given by (2.2) and (3.1) with \( d_i = d_i^M \)
in (3.1) being determined by (3.3) is minimax.

4. **Unknown control population.**

In this section we will consider the case when \( \theta_0 \) is unknown and will derive a \( \tau \)-minimax decision rule \( \delta^\tau \) in the class \( \phi_0 \) of decision rules for which \( \delta_i(x) \) in (2.2) depends only on \( x_0 \) and \( x_i \) for \( i = 1, \ldots, k \).

Let us consider rules \( \delta_0(x) \) in \( \phi_0 \) where \( \delta_i(j|x) \) \((j = 1,2,3) \) are given by
\[
\begin{align*}
\delta_i(1|x) &= I(-\infty, -d_i)(x_i - x_0), \\
\delta_i(2|x) &= I(-d_i, d_i)(x_i - x_0), \\
\delta_i(3|x) &= I(d_i, \infty)(x_i - x_0),
\end{align*}
\]
(4.1)
for \( 0 \leq d_i \leq \infty \) and \( i = 1, \ldots, k \).
Note that the pdf of \( Y_i = X_i - X_0 \) is given by
\[
g_i(y-(\theta_i-\theta_0)) = \int_{-\infty}^{\infty} f_i(t+y-\theta_i)f_0(t-\theta_0)dt,
\]
and that \( g_i(\cdot) \) is strongly unimodal by the result of Ibragimov (1956) and symmetric about the origin. Therefore, the next follows from this fact and Lemma 3.1.

**Lemma 4.1.** Suppose that a rule \( \delta(x) \) in \( \Delta_0 \) is given by (2.2) and (4.1). Then, for \( i = 1, \ldots, k \),
\[
\sup_{\tau \in \Gamma} r_i(\tau, \delta_1) \leq \omega_i,
\]
where, for \( i = 1, \ldots, k \),
\[
\omega_i = \int \left[ \kappa_1 \gamma_1 g_i(y+\Delta_1) + \kappa_2 \gamma_1^4 g_i(y-\Delta_1) + \kappa_4 (1-\gamma_1^4) g_i(y+\Delta_1) \right] dy
\]
\[
+ \int \kappa_1 \gamma_1 g_i(y-\Delta_2) dy.
\]

We now proceed as in Theorem 3.1 by considering the following sequence \( \{\tau_n, n \geq 1\} \) of prior distributions in \( \Gamma \) for the case when \( \epsilon > 0 \). Under \( \tau_n \),

(i) \( \theta_1 - \theta_0, \ldots, \theta_k - \theta_0 \) are independent,

(ii) \( P[\theta_i - \theta_0 = \Delta_2] = P[\theta_i - \theta_0 = -\Delta_2] = \gamma_i/2 \),
\[
P[\theta_i - \theta_0 = \Delta_1] = P[\theta_i - \theta_0 = -\Delta_1] = (1-\gamma_1^4)/2,
\]
\[
P[\theta_i - \theta_0 = \Delta_1 - n^{-1}] = P[\theta_i - \theta_0 = -\Delta_1 + n^{-1}] = \gamma_i/2 \text{ and}
\]

(iii) \( \theta_0 \) has uniform distribution over \([-n,n]\) and is independent of \( \theta_1 - \theta_0, \ldots, \theta_k - \theta_0 \).

It can be easily shown that the overall risk of the Bayes rule is given by
\[
\inf_{\delta \in \partial_0} r(\tau_n, \delta) = \frac{1}{4n} \sum_{i=1}^{k} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_n(i, x, y) dx dy
\]

(4.3)

where \( p_n(i, x, y) = \min(s_n(i, x, y), t_n(i, x, y), s_n(i, -x, -y)) \) with

\[
s_n(i, x, y) = \varepsilon_2 \gamma_i \int_{-n}^{n} [f_i(x-u-\Delta_1+n^{-1})f_i(x-u+\Delta_1-n^{-1})]f_0(y-u)du + \\
+ \varepsilon_4(1-\gamma_1-\gamma_i) \int_{-n}^{n} f_i(x-u-\Delta_1)f_0(y-u)du + \\
+ (\varepsilon_1+\varepsilon_3)\gamma_i \int_{-n}^{n} f_i(x-u-\Delta_2)f_0(y-u)du \text{ and}
\]

\[
t_n(i, x, y) = \varepsilon_1 \gamma_i \int_{-n}^{n} [f_i(x-u+\Delta_2)+f_i(x-u-\Delta_2)]f_0(y-u)du.
\]

From change of variables \( x = nv-w \) and \( y = nv+w \), it follows that

\[
\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_n(i, x, y) dx dy/4n = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p_n(i, nv-w, nv+w) dv dw/2
\]

\[
\geq \int_{-\infty}^{\infty} [\int_{-\infty}^{\infty} p_n(i, nv-w, nv+w) dv] dw/2.
\]

(4.4)

Note that

\[
s_n(i, nv-w, nv+w) = \varepsilon_2 \gamma_i \int_{n(v-1)}^{n(v+1)} [f_i(z-w-\Delta_1+n^{-1})f_i(z-w+\Delta_1-n^{-1})]f_0(z+w)dz + \\
+ \varepsilon_4(1-\gamma_1-\gamma_i) \int_{n(v-1)}^{n(v+1)} f_i(z-w-\Delta_1)f_0(z+w)dz + \\
+ (\varepsilon_1+\varepsilon_3)\gamma_i \int_{n(v-1)}^{n(v+1)} f_i(z-w-\Delta_2)f_0(z+w)dz \text{ and}
\]

\[
t_n(i, nv-w, nv+w) = \varepsilon_1 \gamma_i \int_{n(v-1)}^{n(v+1)} [f_i(z-w+\Delta_2)+f_i(z-w-\Delta_2)]f_0(z+w)dz.
\]

Therefore, for any \((v, w) \in (-1,1) \times R\), \( p_n(i, nv-w, nv+w) \) converges, as \( n \to \infty \),
to $p(i,w) = \min(s(i,w), t(i,w), s(i,w))$ where

$$s(i,w) = e_2 \gamma_i \int_{-\infty}^{\infty} \left[ f_i(z-w-\Delta_1) + f_i(z-w+\Delta_1) \right] f_0(z+w) dz +$$

$$+ e_4 (1-\gamma_1-\gamma'_1) \int_{-\infty}^{\infty} f_i(z-w-\Delta_1) f_0(z+w) dz +$$

$$+ (e_1+\varepsilon_3) \gamma_i \int_{-\infty}^{\infty} f_i(z-w-\Delta_2) f_0(z+w) dz \text{ and}$$

$$t(i,w) = \varepsilon_1 \gamma_i \int_{-\infty}^{\infty} \left[ f_i(z-w+\Delta_2) + f_i(z-w-\Delta_2) \right] f_0(z+w) dz.$$

It follows from (4.4) that, for $i = 1, \ldots, k$,

$$\lim_{n \to \infty} \int_{-\infty}^{\infty} p_n(i,x,y) dx dy / 4n \geq \int_{-\infty}^{\infty} p(i,w) dw$$

$$= \int_0^\infty h(i,y) dy,$$

where $h(i,y) = \min(e_2 \gamma_i [g_i(y-\Delta_1) + g_i(y+\Delta_1)] + e_4 (1-\gamma_1-\gamma'_1) g_i(y+\Delta_1) +$

$$+ (e_1+\varepsilon_3) \gamma_i g_i(y+\Delta_2), e_1 \gamma_i [g_i(y+\Delta_2) + g_i(y-\Delta_2)])$.$$

Then from (4.3), we have

$$\lim_{n \to \infty} \inf_{\delta_0} r(\tau_n, \delta) \geq \sum_{i=1}^{k} \int_0^\infty h(i,y) dy. \quad (4.5)$$

Note that $\int_0^\infty h(i,y) dy$ can be written as

$$\int_0^\infty h(i,y) dy = \min(e_3 \gamma_i g_i(y+\Delta_2) + e_4 (1-\gamma_1-\gamma'_1) g_i(y+\Delta_1) + e_2 \gamma_i [g_i(y+\Delta_1) + g_i(y-\Delta_1)]) +$$

$$e_1 \gamma_i g_i(y-\Delta_2) dy + \int_{-\infty}^{0} e_1 \gamma_i g_i(y-\Delta_2) dy$$
\[
= \int_{d_i}^{\infty} \left[ \gamma_1 g_i(y + \Delta_2) + \frac{\gamma_1}{2} g_i(y - \Delta_1) + g_i(y + \Delta_1) + g_i(1 - \gamma_1 - \gamma_1') g_i(y + \Delta_1) \right] dy \\
\quad + \int_{-\infty}^{d_i} \gamma_1 g_i(y - \Delta_2) dy,
\]

where \( d_i = \max(c_i, 0)\) with \( c_i \) being determined so that

\[
\leq \frac{\ell_3 \gamma_1 g_i(y + \Delta_2)}{\gamma_1} + \frac{\ell_2 \gamma_1 g_i(y - \Delta_1) + g_i(y + \Delta_1)}{\gamma_1} + \frac{\ell_4 (1 - \gamma_1 - \gamma_1') g_i(y + \Delta_1)}{\gamma_1}
\leq \frac{\ell_1 \gamma_1 g_i(y - \Delta_2)}{\gamma_1} \quad \text{as} \quad y \geq c_i.
\]

Let \( \delta^R \) be the rule given by (2.2) and (4.1) where \( d_i = d_i^R \) in (4.1) is defined by \( d_i = \max(c_i, 0) \) with \( c_i \) determined by (4.6). Then it follows from (4.5) and Lemma 4.1 that

\[
\lim_{n \to \infty} \inf_{\delta \in \mathcal{O}_0} r(\tau, \delta) \geq \frac{1}{2} \sup_{i=1}^{k} \sup_{\tau \in \Gamma} r_i(\tau, \delta_R^i) \\
\quad \geq \sup_{\tau \in \Gamma} r(\tau, \delta_R^i).
\]

Therefore, Lemma 2.1 yields the next result.

**Theorem 4.1.** Assume that independent random variables \( X_0, \ldots, X_k \) have pdf's \( f_0(x_0 - \theta_0), \ldots, f_k(x_k - \theta_k) \), respectively, with \( f_i(\cdot) \) being strongly unimodal and symmetric, and that the loss function is given by (2.1). Then the \( r \)-minimax rule \( \delta^R \) in \( \mathcal{O}_0 \) is given by (2.2) and (4.1) where \( d_i = d_i^R \) in (4.1) is defined by \( d_i^R = \max(c_i, 0) \) with \( c_i \) being determined by (4.6), for all \( i = 1, 2, \ldots, k \).

**Remark 4.1.** It can be easily shown that the symmetry of \( f_i(\cdot) \) in Theorem 4.1 can be replaced by that of \( g_i(\cdot) \). It should be noted that the symmetry of \( g_i(\cdot) \) follows when \( f_0(\cdot), \ldots, f_k(\cdot) \) are identical.
The next result follows in exactly the same manner as Corollary 3.1 was proved.

Corollary 4.1. Under the assumptions in Theorem 4.1, if \( \xi_1 = \xi_2 = 1, \xi_3 \leq 1 \) and \( \xi_4 \leq 2 \), then a minimax rule \( \delta^M \) in \( \mathcal{W}_0 \) is given by (2.2) and (4.1) where \( d_i = d_i^M \) in (4.1) is determined so that, for \( G_i(x) = \int_{-\infty}^{x} g_i(t) dt \),

\[
G_i(d_i - \Delta_2^i) + \xi_3 G_i(-d_i - \Delta_2^i) = G_i(-d_i - \Delta_1^i) + G_i(-d_i + \Delta_1^i).
\]

Now we provide some examples to illustrate the application of the above results.

Example 4.1. Suppose \( \pi_i \) represents a normal population \( N(\theta_i, \sigma_i^2) \) for \( i = 0, \ldots, k \) with \( \sigma_i^2 \) \( (i = 0, \ldots, k) \) known. We assume that a random sample of size \( n_i \) is taken from each of the \( k+1 \) populations \( \pi_0, \ldots, \pi_k \). By sufficiency we can restrict our attention to the decision rules depending only on the sample means \( X_0, \ldots, X_k \) where \( X_i \) has normal distribution with mean \( \theta_i \) and variance \( n_i^2 / n_i \) for \( i = 0, 1, \ldots, k \).

(A) r-minimax rule: The r-minimax rule \( \delta^r \) in \( \mathcal{W}_0 \) in Theorem 4.1 is determined by \( d_i^r = (n_i^2 + \eta_0^2)^{1/2} \max(c_i, 0) \) where \( c_i \) is defined so that

\[
\begin{align*}
\xi_3 y_i^e + \xi_2 y_i^e & = -2 \lambda_i (x - \epsilon_i) + \xi_2 y_i^e \left[ e^{-2 \lambda_i (x - \epsilon_i)} + e^{-2 \lambda_i (x - \epsilon_i)} e^{-2 \lambda_i (x - \epsilon_i)} \right] + \\
+ \xi_4 (1 - y_i^e - y_i^e) & = 0 \text{ as } x \leq, \geq c_i,
\end{align*}
\]

where \( \lambda_i = \lambda(n_i^2 + \eta_0^2)^{-1/2} \) and \( \epsilon_i = \epsilon(n_i^2 + \eta_0^2)^{-1/2} \).

(B) Minimax rule: Assume \( \xi_1 = \xi_2 = 1, \xi_3 \leq 1 \) and \( \xi_4 \leq 2 \). Then the minimax rule \( \delta^M \) in \( \mathcal{W}_0 \) in Corollary 4.1 is determined by \( d_i^M = (n_i^2 + \eta_0^2)^{1/2} t_i \) where \( t_i \) is defined so that
\[ \phi(t_i - \lambda_i - \epsilon_i) + \epsilon_3 \phi(-t_i - \lambda_i - \epsilon_i) = \phi(-t_i - \lambda_i + \epsilon_i) + \phi(-t_i + \lambda_i - \epsilon_i) \] (4.8)

with \( \lambda_i \) and \( \epsilon_i \) defined as in (A) and \( \phi \) denoting the cdf of the standard normal distribution.

**Example 4.2.** Assume that \( \pi_i \) represents a normal population \( N(0, \sigma_i^2) \) for \( i = 0, 1, \ldots, k \) with \( \sigma_i^2 \) unknown, and that we have a random sample of size \( n \) taken from each population \( \pi_i \). Consider a problem of partitioning the treatment populations in terms of variances with a loss structure analogous to that given by (2.1), i.e., a loss function obtained from the latter by substituting \( \log \sigma_i^2 \), \( \log \Delta \) and \( \log \epsilon \) for \( \theta_i \), \( \Delta \) and \( \epsilon \), respectively. Thus \( \Delta \) and \( \epsilon \) are assumed such that \( 1 \leq \epsilon < \Delta \). By sufficiency we need to consider only the decision rules depending on \( s_0^2, \ldots, s_k^2 \) where \( s_i^2 \) denotes the sample variance corresponding to \( \pi_i \). Since \( n s_i^2 / \sigma_i^2 \) (\( i = 0, 1, \ldots, k \)) are independently distributed chi-square random variables with degrees of freedom \( n \), it can be easily seen that the associated location parameter problem satisfies the assumptions in Theorem 4.1 except the symmetry which is not necessary in this problem because of Remark 4.1. Therefore, with obvious modifications we have the following results. Let \( \mathcal{D}_0 \) denote the class of decision rules \( \delta = (\delta_1, \ldots, \delta_k) \) for which \( \delta_i \) depends only on \( s_0^2 \) and \( s_i^2 \) and let \( x_i \) denote \( s_i^2 / s_0^2 \) for \( i = 1, \ldots, k \).

(A) \( r \)-minimax rule: A \( r \)-minimax rule \( \delta^r \) in \( \mathcal{D}_0 \) is given by

\[ \delta^r_i(1|x_i) = I_{(0, d_i^{-1})}(x_i), \delta^r_i(2|x_i) = I_{(d_i^{-1}, d_i)}(x_i) \text{ and } \delta^r_i(3|x_i) = I_{[d_i, \infty)}(x_i) \]

for \( i = 1, \ldots, k \) where \( d_i = \max(c_i, 1) \) with \( c_i \) being determined so that
\[ v_3 y_1 \left( \frac{\Lambda_2 + y}{1 + \Lambda_2 y} \right)^n + v_2 y_1 \left[ \left( \frac{\Lambda_2 + y}{1 + \Lambda_2 y} \right)^n + \left( \frac{\Lambda_2 + y}{1 + \Lambda_2 y} \right)^{\Delta_2} \right] \left( \frac{\Delta_1}{\Delta_2} \right)^{n/2} + \]

\[ + v_4 (1 - y_1 - y_1) \left( \frac{\Lambda_2 + y}{1 + \Lambda_2 y} \right)^n \left( \frac{\Delta_1}{\Delta_2} \right)^{n/2} \]

\[ \leq v_1 y_1 \text{ as } y \geq \leq c_i. \]

Here \( \Lambda_1 = \Lambda \xi^{-1}, \Lambda_2 = \Lambda \xi. \)

(B) \textbf{Minimax rule:} Assume \( v_1 = v_2 = 1, v_3 \leq 1 \) and \( v_4 \leq 2. \) The minimax rule \( \delta^M \) in \( \omega_0 \) is the same as \( \delta^R \) in (A) except that \( d_i = d \) is determined so that

\[ G_n \left( d/\Delta_2 \right) + v_3 \left[ 1 - G_n \left( d/\Delta_2 \right) \right] = G_n \left( \Delta_1/d \right) + 1 - G_n \left( d/\Delta_1 \right) \]

where \( G_n \) denotes the cdf of \( F \)-distribution with degrees of freedom \( n \) and \( n. \)

We note that if \( \omega_i \) represents \( N(\mu_i, \sigma_i^2) \) with both \( \mu_i \) and \( \sigma_i^2 \) unknown, then the above results still hold with \( n-1 \) replacing \( n. \)

5. \textbf{Comparison of }\Gamma\textbf{-minimax rules with Bayes rules.}

When we represent our a prior information about the parameters by prior distributions over the parameter space, one method for the use of such information is to find a rule which is \( \Gamma \)-minimax with respect to the class, \( \Gamma, \) of such prior distributions.

Another way is to select one such prior distribution and use the corresponding Bayes rule. Thus Bayes rules wrt prior distributions in \( \Gamma \) are natural competitors of a \( \Gamma \)-minimax rule.

In this section we consider \( k+1 \) normal populations \( N(\theta_i, \sigma_i^2) \) with \( \sigma_i^2 \) known, and derive Bayes rules wrt normal priors and then compare them with the corresponding \( \Gamma \)-minimax rules from both points of view.
For this purpose, assume that \((\theta_0, \ldots, \theta_k)\) have prior distribution \(\tau_0\) under which \(\theta_0, \ldots, \theta_k\) are independent and each \(\theta_i\) has a normal distribution with mean \(\mu_i\) and variance \(\nu_i^2\). Let \(x_0, \ldots, x_k\) denote the observed sample means based on samples of size \(n_i\) \((i = 0, 1, \ldots, k)\). To simplify forthcoming formulas, let us introduce the following notations:

\[
\sigma_i^2 = \sigma^2 / n_i, \quad b_i = [ (\sigma_0^{-2} + v_i^{-2})^{-1} + (\sigma_0^{-2} + v_0^{-2})^{-1} ]^{-1}, \\
m_i = (\sigma_i^{-2} x_i + v_i^{-2} \mu_i)(\sigma_i^{-2} + v_i^{-2})^{-1}, \quad y_i = (m_i - m_0)b_i.
\] (5.1)

The following theorem describes the Bayes rule.

**Theorem 5.1.** Assume the loss function is given by (2.1). Then the Bayes rule \(\delta^B\) wrt \(\tau_0\) is given by \(\delta^B_i(1|y) = I_{(-\infty, -d_i]}(y_i)\), \(\delta^B_i(2|y) = I_{(-d_i, d_i)}(y_i)\) and \(\delta^B_i(3|y) = I_{[d_i, +\infty)}(y_i)\) for \(i = 1, \ldots, k\) where \(d_i = \max(c_i, 0)\) with \(c_i\) being determined so that

\[
\begin{align*}
&\varepsilon_3 \phi(-\Delta_2 b_i^{-1} - y) + \varepsilon_4 [ \phi(-\Delta_1 b_i^{-1} - y) - \phi(-\Delta_2 b_i^{-1} - y) ] \\
&+ \varepsilon_2 [ \phi(\Delta_1 b_i^{-1} - y) - \phi(-\Delta_1 b_i^{-1} - y) ] - \varepsilon_1 \phi(-\Delta_2 b_i^{-1} + y) \\
&> 0 \quad \text{as} \quad y \leq c_i.
\end{align*}
\]

**Proof.** It suffices to find the Bayes rule for each of the \(k\) component decision problems. This reduces to the comparison of posterior risks of three possible actions. We will do this for the first component decision problem without loss of generality. Let \(p_1(y_i), p_2(y_i)\) and \(p_3(y_i)\) denote the posterior risks of the actions 1, 2 and 3, respectively, in the first component problem. Then it can be shown that
\[ p_1(y) = (\varepsilon_1 + \varepsilon_3) \phi(-\Delta_2 b_1^{-1} y) + \varepsilon_4 [\phi(\Delta_2 b_1^{-1} y) - \phi(\Delta_1 b_1^{-1} y)] + \varepsilon_2 [\phi(\Delta_1 b_1^{-1} y) - \phi(-\Delta_1 b_1^{-1} y)], \]

\[ p_2(y) = \varepsilon_1 [\phi(-\Delta_2 b_1^{-1} y) + \phi(-\Delta_2 b_1^{-1} y)] \text{ and } p_3(y) = p_1(-y). \]

Note that \( p_1(y) - p_3(y) \) can be written as \( E_H(Z) \) where \( Z \) has a normal distribution with mean \( y \) and variance \( 1 \) and \( H(\cdot) \) is given by

\[
H(z) = \begin{cases} 
\varepsilon_1 + \varepsilon_3 & \text{if } z \geq \Delta_2 b_1^{-1} \\
\varepsilon_4 & \text{if } \Delta_1 b_1^{-1} < z < \Delta_2 b_1^{-1} \\
0 & \text{if } -\Delta_1 b_1^{-1} \leq z \leq \Delta_1 b_1^{-1} \\
-\varepsilon_4 & \text{if } -\Delta_2 b_1^{-1} < z < -\Delta_1 b_1^{-1} \\
-(\varepsilon_1 + \varepsilon_3) & \text{if } z \leq -\Delta_2 b_1^{-1}.
\end{cases}
\]

Since the density of the normal distribution \( N(y,1) \) has the MLR property, it follows that \( p_1(y) - p_3(y) \) has at most one sign change.

Furthermore, it can be shown that \( p_1(y) - p_3(y) \) is strictly increasing on \((-\Delta_1 b_1^{-1}, \Delta_2 b_1^{-1})\) and \( p_1(0) - p_3(0) = 0 \). Thus \( p_1(y) - p_3(y) \geq 0 \) as \( y \geq 0 \) and \( p_1(y) - p_3(y) \leq 0 \) as \( y \leq 0 \).

Similarly, we can show that \( p_3(y) - p_2(y) \geq 0 \) as \( y \leq c_1 \) for some real number \( c_1 \) unless \( p_3(y) - p_2(y) \leq 0 \) for all \( y \). Therefore the result follows.

Now we compare the \( \Gamma \)-minimax rule \( \delta^\Gamma \) given in Example 4.1 and the Bayes rule \( \delta^B \) given in Theorem 5.1 under the assumption that \( \varepsilon_1 = \varepsilon_2 = \varepsilon_4 = 1, \varepsilon_3 = \nu, n_i = n \) and \( \nu_i^2 = \nu^2 \) for \( i = 0, \ldots, k \). Note that we compare these rules under the relations

\[
\gamma^i_1 = \phi((-\Delta_2 + \mu_i - \mu_0)(2\nu_i^2)^{-\frac{1}{2}}) + \phi((-\Delta_2 - \mu_i + \mu_0)(2\nu_i^2)^{-\frac{1}{2}})
\]

\[
\text{and } \gamma^i_1 = \phi((\Delta_1 - \mu_i + \mu_0)(2\nu_i^2)^{-\frac{1}{2}}) - \phi((-\Delta_1 - \mu_i + \mu_0)(2\nu_i^2)^{-\frac{1}{2}}) \text{ for } i = 1, \ldots, k. \]
meaningful comparison of these rules. One way is to examine the increase in the overall risk wrt $\tau_0$ resulting from the use of $\delta^B$. Another way is to compare them in terms of $\sup_{\tau \in \Gamma} r(\tau, \delta)$. When $n_{i} = n$ and $v_{i}^2 = v^2$ for $i = 0, 1, \ldots, k$, the Bayes rule depends on $x$ only through $x_1 - x_0, \ldots, x_k - x_0$ and it can be shown that $\sup_{\tau \in \Gamma} r(\tau, \delta^B) = \sum_{i=1}^{k} \sup_{\tau \in \Gamma} r_i(\tau, \delta_i^B)$. Thus it suffices to compare these rules wrt classification of one population. We choose $\nu_1$ for this purpose without loss of generality.

Now we introduce the parameters used in the comparison as follows.

\[ \beta_1 = \frac{nv^2}{\sigma^2}, \beta_2 = \frac{\Delta}{\sqrt{2}v^2}, \beta_3 = \frac{\varepsilon}{\sqrt{2}v^2}, \text{ and } \beta_4 = \frac{\mu_1 - \mu_0}{\sqrt{2}v^2}. \]

It can be verified that the overall risk wrt $\tau_0$ of these rules can be written as

\[ \text{e} \phi(-A-B-C) + \phi(A-B-C) + \phi(D-E) + \phi(-D-E) \]

\[-\phi_0(-A-B-C, -D-E; \rho) + (1-\varepsilon)\phi_0(-A-B-C, D-E; \rho) \]

\[-\phi_0(-A+D-C, D-E; \rho) + \phi_0(A-B-C, -D-E; \rho) \]

\[-\phi_0(-A+D-C, -D-E; \rho) - \phi_0(A-B-C, D-E; \rho) \]

\[+ \phi_0(-A+D-C, D-E; \rho) - \phi_0(A-B-C, -D-E; \rho) \]

\[-\phi_0(A+D-C, D-E; \rho) - (1-\varepsilon)\phi_0(A+D-C, -D-E; \rho) \]

where $\phi_0(\cdot, \cdot; \rho)$ is the cdf of a bivariate normal distribution with zero means, unit variances and correlation coefficient $\rho$, and where $A = \beta_2$, $B = \beta_3$, $C = \beta_4$, $\rho = \beta_3(1+\beta_1)^{-\frac{1}{2}}$, $D = d_1\beta_1^{-\frac{1}{2}}$ for $\delta^B$, $D = \max(c_1, 0)(1+\beta_1)^{-\frac{1}{2}}$ for $\delta^L$, $E = \rho^{-1}\beta_4$ for $\delta^B$ and $E = \rho\beta_4$ for $\delta^L$ with $d_1$ and $c_1$ being those in Theorem 5.1 and Example 4.1, respectively. Also $\sup_{\tau \in \Gamma} r_i(\tau, \delta_i)$ for both rules $\tau \in \Gamma$. 

can be written as
\[ \gamma_1[\phi(R+|S|-T-U) - \phi(-R+|S|-T+U) + \phi(-R+|S|-T-U)] \\
+ \gamma_1'[\phi(-R+S+T-U) + \phi(-R+S-T+U)]V[\phi(-R+S-T-U) + \phi(-R-S+T+U)] \\
+ (1-\gamma_1')\phi(-R+|S|-T+U) \]

where \( x^y = \max(x,y) \), \( T = \beta_2\beta_1^2 \), \( U = \beta_3\beta_1^2 \), \( S = \beta_4\beta_1^2 \) for \( \delta^B \), \( S = 0 \) for \( \delta^A \), \( R = \beta_1\beta_1^2(1+\beta_1)^2d_1 \) for \( \delta^B \), \( R = \max(c_1, 0) \) for \( \delta^R \) with \( d_1 \) and \( c_1 \) being those in Theorem 5.1 and Example 4.1, respectively. For selected values of \( \beta_i \) (i = 1, ..., 4), Table I and Table II give \( r_1(\tau, \delta_1) \) and \( \sup_{\tau \in \Gamma} r_1(\tau, \delta_1) \) for \( \delta_1 = \delta_1^A \), \( \delta_1^B \) for \( \lambda = 0 \) and \( \lambda = 1 \), respectively. It can be observed from these tables that, in many cases, the increase in the overall risk wrt \( \tau_0 \) from the use of \( \delta_1^A \) is only slight compared to that in \( \sup_{\tau \in \Gamma} r_1(\tau, \delta_1) \) from the use of \( \delta_1^B \). In this sense, \( \delta^R \) is more robust against other formulation than \( \delta^B \). Such properties of \( \delta^R \) become more prominent as the difference between the prior means (\( \beta_4 \)) increases and the prior variance (\( \beta_1 \)) gets smaller. When we have the same prior means and the prior variance is large, both rules compare favorably with each other. In most cases, we can observe that \( \delta^R \) compares favorably with \( \delta^B \) in terms of the overall risk.
Table I
Overall risks and the values of $\sup_{\tau \in \mathbb{R}} r(\tau, \delta)$ of $\delta^B$ and $\delta^r$ when $\lambda = 0$.

<table>
<thead>
<tr>
<th>$\delta^r$</th>
<th>$\delta^B$</th>
<th>$\delta^r$</th>
<th>$\delta^B$</th>
<th>$\delta^r$</th>
<th>$\delta^B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_4$</td>
<td>.25</td>
<td>1.0</td>
<td>4.0</td>
<td>$\beta_2$ = 0.5</td>
<td>$\beta_3$ = 0.125</td>
</tr>
<tr>
<td>1.0</td>
<td>.4961</td>
<td>.9696</td>
<td>.4097</td>
<td>.7784</td>
<td>.2816</td>
</tr>
<tr>
<td>1.0</td>
<td>.3930</td>
<td>.2587</td>
<td>.2864</td>
<td>.2097</td>
<td>.2001</td>
</tr>
<tr>
<td>0.5</td>
<td>.5344</td>
<td>.8274</td>
<td>.4577</td>
<td>.5201</td>
<td>.3312</td>
</tr>
<tr>
<td>0.5</td>
<td>.4520</td>
<td>.3957</td>
<td>.3321</td>
<td>.2963</td>
<td>.2121</td>
</tr>
<tr>
<td>0.0</td>
<td>.5164</td>
<td>.5527</td>
<td>.4568</td>
<td>.4626</td>
<td>.3413</td>
</tr>
<tr>
<td>0.0</td>
<td>.4599</td>
<td>.4544</td>
<td>.3341</td>
<td>.3315</td>
<td>.2103</td>
</tr>
<tr>
<td>$\beta_2$ = 0.5</td>
<td>$\beta_3$ = 0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.5250</td>
<td>.9731</td>
<td>.4471</td>
<td>.8117</td>
<td>.3261</td>
</tr>
<tr>
<td>1.0</td>
<td>.4140</td>
<td>.2853</td>
<td>.3103</td>
<td>.2395</td>
<td>.2413</td>
</tr>
<tr>
<td>0.5</td>
<td>.5768</td>
<td>.8460</td>
<td>.5088</td>
<td>.6703</td>
<td>.3999</td>
</tr>
<tr>
<td>0.5</td>
<td>.4898</td>
<td>.4330</td>
<td>.3840</td>
<td>.3368</td>
<td>.2899</td>
</tr>
<tr>
<td>0.0</td>
<td>.5680</td>
<td>.5937</td>
<td>.5170</td>
<td>.5190</td>
<td>.4210</td>
</tr>
<tr>
<td>0.0</td>
<td>.5027</td>
<td>.4978</td>
<td>.3828</td>
<td>.3764</td>
<td>.2856</td>
</tr>
<tr>
<td>$\beta_2$ = 0.8</td>
<td>$\beta_3$ = 0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.4770</td>
<td>.9598</td>
<td>.3793</td>
<td>.7074</td>
<td>.2574</td>
</tr>
<tr>
<td>1.0</td>
<td>.3909</td>
<td>.2824</td>
<td>.2539</td>
<td>.1931</td>
<td>.1369</td>
</tr>
<tr>
<td>0.5</td>
<td>.3739</td>
<td>.7920</td>
<td>.3450</td>
<td>.5344</td>
<td>.2692</td>
</tr>
<tr>
<td>0.5</td>
<td>.3594</td>
<td>.3163</td>
<td>.2424</td>
<td>.2237</td>
<td>.1152</td>
</tr>
<tr>
<td>0.0</td>
<td>.3172</td>
<td>.5226</td>
<td>.3066</td>
<td>.3794</td>
<td>.2546</td>
</tr>
<tr>
<td>0.0</td>
<td>.3151</td>
<td>.3050</td>
<td>.2404</td>
<td>.2298</td>
<td>.1133</td>
</tr>
<tr>
<td>$\beta_2$ = 0.8</td>
<td>$\beta_3$ = 0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>.5365</td>
<td>.9699</td>
<td>.4541</td>
<td>.7840</td>
<td>.3573</td>
</tr>
<tr>
<td>0.0</td>
<td>.4380</td>
<td>.3309</td>
<td>.3190</td>
<td>.2411</td>
<td>.2407</td>
</tr>
<tr>
<td>0.5</td>
<td>.4357</td>
<td>.8308</td>
<td>.4278</td>
<td>.6286</td>
<td>.3960</td>
</tr>
<tr>
<td>0.0</td>
<td>.3789</td>
<td>.5827</td>
<td>.3785</td>
<td>.4658</td>
<td>.3679</td>
</tr>
<tr>
<td>0.0</td>
<td>.3789</td>
<td>.3658</td>
<td>.3509</td>
<td>.2887</td>
<td>.1939</td>
</tr>
</tbody>
</table>

The numbers on the first (second) row in each box are the values of $\sup_{\tau \in \mathbb{R}} r(\tau, \delta)$ ($r(\tau_0, \delta)$).
Table II

Overall risks and the value of \( \sup_{\tau \in \Gamma} r(\tau, \delta) \) of \( \delta^B \) and \( \delta^P \) when \( \epsilon = 1 \).

<table>
<thead>
<tr>
<th>( \beta_2 = 0.5 )</th>
<th>( \beta_3 = 0.125 )</th>
<th>( \beta_2 = 0.5 )</th>
<th>( \beta_3 = 0.05 )</th>
<th>( \beta_2 = 0.8 )</th>
<th>( \beta_3 = 0.2 )</th>
<th>( \beta_2 = 0.8 )</th>
<th>( \beta_3 = 0.08 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>( \beta_1 )</td>
<td>.25</td>
<td>1.0</td>
<td>4.0</td>
<td>.6655</td>
<td>1.6167</td>
<td>.5435</td>
</tr>
<tr>
<td>1.0</td>
<td>.5218</td>
<td>.2929</td>
<td>.2947</td>
<td>.2195</td>
<td>.1674</td>
<td>.1194</td>
<td>.6655</td>
</tr>
<tr>
<td>0.5</td>
<td>.5746</td>
<td>1.2157</td>
<td>.5098</td>
<td>.7758</td>
<td>.3496</td>
<td>.4498</td>
<td>.6655</td>
</tr>
<tr>
<td>0.0</td>
<td>.5299</td>
<td>.6991</td>
<td>.4865</td>
<td>.5298</td>
<td>.3529</td>
<td>.3560</td>
<td>.6655</td>
</tr>
<tr>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.2 )</td>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.08 )</td>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.2 )</td>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.08 )</td>
</tr>
<tr>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>( \beta_1 )</td>
<td>.1717</td>
<td>1.6582</td>
<td>.6084</td>
<td>1.1721</td>
<td>.4027</td>
<td>.6698</td>
</tr>
<tr>
<td>1.0</td>
<td>.5688</td>
<td>.3264</td>
<td>.3280</td>
<td>.2523</td>
<td>.2059</td>
<td>.1500</td>
<td>.6655</td>
</tr>
<tr>
<td>0.5</td>
<td>.6242</td>
<td>1.2762</td>
<td>.5775</td>
<td>.8566</td>
<td>.4372</td>
<td>.5520</td>
<td>.6655</td>
</tr>
<tr>
<td>0.0</td>
<td>.5817</td>
<td>.7627</td>
<td>.5545</td>
<td>.6032</td>
<td>.4452</td>
<td>.4502</td>
<td>.6655</td>
</tr>
<tr>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.08 )</td>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.08 )</td>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.08 )</td>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.08 )</td>
</tr>
<tr>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>( \beta_1 )</td>
<td>.5021</td>
<td>1.4114</td>
<td>.3982</td>
<td>.8181</td>
<td>.2580</td>
<td>.4413</td>
</tr>
<tr>
<td>1.0</td>
<td>.4248</td>
<td>.2883</td>
<td>.2465</td>
<td>.1938</td>
<td>.1355</td>
<td>.0923</td>
<td>.5021</td>
</tr>
<tr>
<td>0.5</td>
<td>.3742</td>
<td>.9929</td>
<td>.3470</td>
<td>.5702</td>
<td>.2693</td>
<td>.3451</td>
<td>.5021</td>
</tr>
<tr>
<td>0.0</td>
<td>.3157</td>
<td>.5774</td>
<td>.3070</td>
<td>.3900</td>
<td>.2546</td>
<td>.2667</td>
<td>.5021</td>
</tr>
<tr>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.08 )</td>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.08 )</td>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.08 )</td>
<td>( \beta_2 = 0.8 )</td>
<td>( \beta_3 = 0.08 )</td>
</tr>
<tr>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
<td>( \delta^F )</td>
<td>( \delta^B )</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>( \beta_1 )</td>
<td>.5665</td>
<td>1.4765</td>
<td>.4861</td>
<td>.9276</td>
<td>.3617</td>
<td>.5872</td>
</tr>
<tr>
<td>1.0</td>
<td>.4891</td>
<td>.3393</td>
<td>.2966</td>
<td>.2423</td>
<td>.2239</td>
<td>.1353</td>
<td>.5665</td>
</tr>
<tr>
<td>0.5</td>
<td>.4357</td>
<td>1.0748</td>
<td>.4289</td>
<td>.6799</td>
<td>.3960</td>
<td>.4894</td>
<td>.5665</td>
</tr>
<tr>
<td>0.0</td>
<td>.3789</td>
<td>.6540</td>
<td>.3785</td>
<td>.4823</td>
<td>.3679</td>
<td>.3960</td>
<td>.5665</td>
</tr>
</tbody>
</table>

The numbers on the first (second) row in each box are the values of \( \sup_{\tau \in \Gamma} r(\tau, \delta) \) (\( r(\tau_0, \delta) \)).
References


I-Minimax and Minimax Decision Rules for Comparison of Treatments with a Control

Shanti S. Gupta and Woo-Chul Kim

Purdue University
Department of Statistics
West Lafayette, IN 47907

Office of Naval Research
Washington, DC

August 1979
25

Approved for public release, distribution unlimited.
optimal selection rules are derived for classifying the treatments into superior, equivalent, or inferior groups. These $I$-minimax rules are compared with Bayes rules for the normal means problem. It is shown that the $I$-minimax rules compare quite favorably with the Bayes rules. Minimax rules are also derived for the same problem.