Γ-Minimax Selection Procedures in Simultaneous Testing Problems*

by

Klaus J. Miescke

University of Mainz and Purdue University

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #79-1

January 1979

Revised March 1979

*Research supported partly by NSF Grant MCS77-19640 at Purdue University

Γ-MINIMAX SELECTION PROCEDURES IN SIMULTANEOUS TESTING PROBLEMS*

By Klaus J. Miescke

University of Mainz

and

Purdue University

1. Introduction.

Suppose that for every $i \in \{0,1,\ldots,k\}$ we are given a family $\mathscr{F}_i = \{f_{i,\theta}\}_{\theta \in \Omega_i \subseteq \mathbb{R}}$ of densities with respect to either the Lebesgue measure on the real line $\mathbb{R}$ ("continuous case") or any counting measure on a finite or countably infinite subset of $\mathbb{R}$ ("discrete case"), which have monotone non-decreasing likelihood ratios (M.L.R.) in their arguments.

Let $\underline{X}_i = (X_{i1}, \ldots, X_{in_i})$, $i = 0, 1, \ldots, k$ be independent samples from populations $\mathscr{P}_0, \mathscr{P}_1, \ldots, \mathscr{P}_k$ about which we know that for every $i \in \{0,1,\ldots,k\}$ there exists a sufficient statistic Z_i with a distribution lying within the class of distributions given by $\mathscr{F}_i$, where the parameter θ_i only is unknown.

Our goal is to select a subset of $\{\mathscr{P}_1, \ldots, \mathscr{P}_k\}$ containing populations "good" compared to $\mathscr{P}_0$ and excluding "bad" ones, to be defined more precisely in the sequel. Finally, by definition, a selection procedure is a (random) probability measure over the set S of subsets of $\{1,\ldots,k\}$, depending on $\underline{X} = (\underline{X}_1, \ldots, \underline{X}_k)$ in the case where the control values are known and on $\underline{X} = (\underline{X}_0, \underline{X}_1, \ldots, \underline{X}_k)$ otherwise.

In the first part of this paper (section 3) we are dealing with the so-called "known controls problem": Ignoring $\mathscr{P}_0, \mathscr{F}_0$ and $\underline{X}_0$, for every $i \in \{1,\ldots,k\}$

*Research supported partly by NSF Grant MCS77-19640 at Purdue University.

we take values $\theta_{oi} \varepsilon \Omega_i$ and $\Delta_i > 0$ and call population $\mathscr{P}_i$ "good" if $\theta_i \geq \theta_{oi} + \Delta_i$ and "bad" if $\theta_i \leq \theta_{oi}$. Assume that the unknown parameter vector a priori varies randomly according to some prior distribution τ, say, and henceforth let us denote this random vector by $\underline{\Theta} = (\Theta_1, \ldots, \Theta_k)$ and its realizations by $\underline{\theta} = (\theta_1, \ldots, \theta_k)$.

Assume that we have at least the following partial knowledge about τ:

(1.1) $\tau\{\underline{\theta} \varepsilon \underline{\Omega} | \theta_i \geq \theta_{oi} + \Delta_i\} = \pi_i$

$\tau\{\underline{\theta} \varepsilon \underline{\Omega} | \theta_i \leq \theta_{oi}\} = \pi_i'$

where $\underline{\Omega} = \Omega_1 x \ldots x\Omega_k$, $\theta_{oi} \varepsilon \Omega_i$ and π_i, π_i' are non-negative numbers with $\pi_i + \pi_i' \leq 1$, $i = 1, \ldots, k$.

Let us denote the class of priors with property (1.1) by Γ. Now for any specific loss function chosen a Γ-minimax rule $\underline{\psi}^{\Gamma}$ is defined as having smallest supremal risk over Γ among all selection procedures. This definition is due to Blum and Rosenblatt (1967).

Let us adopt the following loss function:

(1.2) $$L(\underline{\theta}, s) = \sum_{i\varepsilon s} L_{i2}\, I_{(-\infty,\theta_{oi}]}(\theta_i) + \sum_{i\notin s} L_{1i}\, I_{[\theta_{oi}+\Delta_i,\infty)}(\theta_i)$$

$s \varepsilon S, \underline{\theta} \varepsilon \underline{\Omega}$, where L_{i1}, $L_{i2} \geq 0$ are fixed, $i = 1, \ldots, k$.

It is not difficult to see that for this type of loss function (and more general for every additive loss function,cf. (2.5)) and any prior distribution the conditional risk - given $\underline{X}$ - of a selection procedure depends only on its corresponding conditional probabilities of selecting population $\mathscr{P}_1, \ldots, \mathscr{P}_k$. This was stated already (but not proved) in Lehmann (1957) and Lehmann (1961). On the other hand, the idea of proof (implicitly in a different context) can be found in Nagel (1970) (section 1.2). Since our

Γ-minimax procedures can in fact be found with the help of these conditional probabilities alone, we shall henceforth restrict ourselves without loss of generality to the class $\mathscr{D}$ of selection procedures of the following form:

(1.3) $\quad \underline{\psi}(\underline{X}) = (\psi_1(\underline{X}), \ldots, \psi_k(\underline{X})), \ \underline{X} = (\underline{X}_1, \ldots, \underline{X}_k)$,

where $\psi_i(\underline{X})$ denotes the conditional probability - given $\underline{X}$ - of selecting $\mathscr{P}_i$, $i = 1, \ldots, k$.

Now for every procedure of type (1.3) each ψ_i can be viewed as being a test for the hypotheses "$\theta_i \leq \theta_{oi}$" versus "$\theta_i \geq \theta_{oi} + \Delta_i$", $i = 1, \ldots, k$. And since (as we shall see soon) the tests enter in the posterior risks only via their power functions, it is not surprising that uniformly most powerful (U.M.P.) tests will play a crucial role in this context: every component of a Γ-minimax procedure will be such a test whose level of significance is determined by the corresponding set of values $(L_{i1}, L_{i2}, \pi_i, \pi_i', \theta_{oi}, \Delta_i)$. By this reason we shall briefly remind the reader of some (mostly well known) properties of the power functions of tests in M.L.R.-situations at the beginning of section 2.

The main purpose of the first part of this paper is not so much to generalize results of Randles and Hollander (1971) and Huang (1974), but rather to demonstrate how close our problem is connected with the classical theory of testing hypotheses and how easily well known results of this theory can be transferred to solve problems of the type considered here. Within the scope of identification and selecting the best population alone, this was demonstrated already by Miescke (1979). Finally, it should be pointed out that one can find other interesting results in the papers by Gupta and Huang (1975, 1977) where Γ-minimax procedures are studied in a more general setup.

In the second part of this paper (section 4) we are concerned with the "unknown control case": instead of having the θ_{oi}'s as control values, $\mathcal{P}_o$ now plays the role of a control population with which the other populations are to be compared.

The only things which change with respect to our previous situation are obvious in nature and therefore can be sketched very briefly: Now we have $\underline{X} = (\underline{X}_o, \underline{X}_1, \ldots, \underline{X}_k)$, $\underline{\Theta} = (\Theta_o, \Theta_1, \ldots, \Theta_k)$ and $\underline{\Omega} = \Omega_o \times \Omega_1 \ldots \times \Omega_k$. Γ and L are given as before with the only change that now $\theta_{o1}, \ldots, \theta_{ok}$ coincide with θ_o, the unknown parameter of $\underline{X}_o$, which is viewed as realization of the random variable Θ_o. The really crucial assumption we have to add is as follows: for all populations we have the "continuous case", $\Omega_o, \Omega_1, \ldots, \Omega_k$ coincide with $\mathbb{R}$ and $\theta_o, \theta_1, \ldots, \theta_k$ are location parameters for $\mathcal{F}_o, \mathcal{F}_1, \ldots, \mathcal{F}_k$ in the usual sense.

In this setup Randles and Hollander (1971) tried to show that within $\mathcal{D}^\circ$, the class of procedures described by

(1.4) $$\underline{\psi}(\underline{X}) = (\psi_1(\underline{X}_o, \underline{X}_1), \ldots, \psi_k(\underline{X}_o, \underline{X}_k)),$$
$$\underline{X} = (\underline{X}_o, \underline{X}_1, \ldots, \underline{X}_k),$$

the Γ-minimax procedures have components which (as one may expect) turn out to be the U.M.P. one sample test based on $Z_i - Z_o$ for "$\theta_i - \theta_o \leq 0$" versus "$\theta_i - \theta_o \geq \Delta_i$" where the levels of significance of these tests are determined by the corresponding set of values $(L_{1i}, L_{2i}, \pi_i, \pi_i^*, \Delta_i)$.

Now in the terminology of Randles and Hollander (1971) or somewhat more generally in ours (cf. (2.5) - (2.7)) we state: in the proof of their Theorem 4.1 Randles and Hollander (1971) have assumed, apparently, that for any two rules $\underline{\psi} = (\psi_1, \ldots, \psi_k)$ and $\underline{\psi}^* = (\psi_1^*, \ldots, \psi_k^*)$ from $\mathcal{D}^\circ$

$$(1.5) \qquad \sup_{\tau\varepsilon\Gamma} r^{(i)}(\tau,\psi_i^*) \leq \sup_{\tau\varepsilon\Gamma} r^{(i)}(\tau,\psi_i),\ i = 1, \ldots ,k$$

implies

$$(1.6) \qquad \sup_{\tau\varepsilon\Gamma} r(\tau,\underline{\psi}^*) \leq \sup_{\tau\varepsilon\Gamma} r(\tau,\underline{\psi}).$$

But this clearly does not hold true, since in

$$(1.7) \qquad \sup_{\tau\varepsilon\Gamma} r(\tau,\underline{\psi}) = \sup_{\tau\varepsilon\Gamma} \sum_{i=1}^{k} r^{(i)} (\tau,\psi_i)$$

$$\leq \sum_{i=1}^{k} \sup_{\tau\varepsilon\Gamma} r^{(i)} (\tau,\psi_i),\ \underline{\psi}\ \varepsilon\mathscr{D}^o$$

the inequality may be strict.

What remains is that Randles and Hollander (1971) in fact have proved only Γ-minimaxity of the candidate with respect to the subclass of $\mathscr{D}^o$ consisting of all translation invariant rules.

The main purpose of the second part of this paper is to give an alternative proof of the theorem of Randles and Hollander (1971). The usual method of finding Γ-minimax procedures (i.e., finding Bayes-rules componentwise with respect to a common least favorable prior) does not work here, and in fact we have to generalize this standard method to admit also the use of improper priors. This will be done at the end of section 2. It should be pointed out that we do not use the Hunt-Stein argument of Randles and Hollander (1971) to derive our main result in section 4. Moreover, we hope that our general method is of interest in itself and may serve as a tool to derive further Γ-minimax results in the future.

Finally, in section 5 we give an example with normal populations. Especially two designs are compared which perhaps can be described best in short by the key-word "paired comparison versus random sampling" because of its analogy to a well-known problem arising in connection with the two-sample t-test (cf. Lehmann (1959), p. 206).

2. Preliminaries.

As announced already in section 1 let us briefly recall some relevant properties of the power functions of tests in M.L.R. - situations. For this purpose let $\underline{X} = (X_1, \dots, X_n)$ be any one of the samples $\underline{X}_1, \dots, \underline{X}_k$ given in section 1 with distribution properties as stated there, and for convenience let us omit subscript i in our considerations below. The testing problem of interest to us is as follows:

(2.1) $\quad H : \theta \leq \theta_o \quad$ versus $\quad K : \theta > \theta_o \quad$ where $\theta_o \in \Omega$ is fixed.

Then as is well known the U.M.P.-test φ^*_α at level $\alpha \in [0,1]$ is given by

(2.2) $\quad \varphi^*_\alpha(\underline{X}) = 1 \ (h(\alpha),\ 0) \quad$ if $\quad Z > (=, <) \quad c(\alpha)$,

where $h : [0,1] \to [0,1]$ and $E_{\theta_o} \varphi^*_\alpha(\underline{X}) = \alpha$.

The power function of this test has the following properties:

(P.1) (U.M.P. - property) For every test ψ and every $\alpha \in [0,1]$

$\sup_{\theta \leq \theta_o} E_\theta \psi(\underline{X}) \leq \alpha \quad$ implies $E_\theta \psi(\underline{X}) \leq E_\theta \varphi^*_\alpha(\underline{X})$

for all $\theta > \theta_o$.

(P.2) For $\alpha \in [0,1]$ fixed, $E_\theta \varphi^*_\alpha(\underline{X})$ is non-decreasing in $\theta \in \Omega$.

(P.3) For $\theta \in \Omega$ fixed, $E_\theta \varphi^*_\alpha(\underline{X})$ is non-decreasing in $\alpha \in [0,1]$.

(P.4) For $\theta > \theta_o$ $(\theta < \theta_o)$ fixed, $E_\theta \varphi^*_\alpha(\underline{X})$ is a concave (convex) function in $\alpha \in [0,1]$.

The first three properties are assumed to be well known to the reader. To prove (P.4) let us first look at the "continuous case." There we have

(2.3) $$E_\theta \varphi^*_\alpha(\underline{X}) = 1 - F_\theta(F^{-1}_{\theta_o}(1 - \alpha)), \ \alpha \in [0,1], \ \theta \in \Omega,$$

where F_θ and F_{θ_o} denote the cumulative distribution functions of f_θ and f_{θ_o}.

Thus for $\theta > \theta_o$ ($\theta < \theta_o$) the M.L.R.-property implies that

(2.4) $$\frac{\partial}{\partial\alpha} E_\theta \varphi^*_\alpha(\underline{X}) = f_\theta(F^{-1}_{\theta_o}(1 - \alpha)) / f_{\theta_o}(F^{-1}_{\theta_o}(1 - \alpha))$$

is a non-increasing (non-decreasing) function in $\alpha \in [0,1]$, a fact which clearly implies (P.4).

The proof of (P.4) in the "discrete case" proceeds analogously if one takes into account that for every fixed $\theta \in \Omega$ $E_\theta\varphi^*_\alpha(\underline{X})$ now is a piecewise linear function in $\alpha \in [0,1]$ and that, instead of derivations, one has to take ratios of differences.

As already mentioned in section 1, we need a generalization of the usual method of finding Γ-minimax selection procedures which is the well known trick of looking componentwise for k Bayes-rules with respect to a common prior τ_o which, at the same time, is least favorable for these k Bayes-rules. This result is due to Randles and Hollander (1971) (cf. Lemma 3.2 there). Since in section 4 (in the "unknown control case") we have to take into considerations improper priors, too, we propose now a generalized version of this method. It should be pointed out that, to the author's knowledge, this idea of approach is primarily due to Lehmann (cf. Lehmann (1957) 6.(ii)).

Our next considerations up to the end of this section are valid in both the "known controls" - and the "unknown control" - setup. In this sense $\underline{X}$ is understood either to be equal to $(\underline{X}_1, \ldots, \underline{X}_k)$ or to be equal to

$(\underline{X}_0, \underline{X}_1, \ldots, \underline{X}_k)$ throughout the following, and $\underline{\theta}$, $\underline{\Omega}$ and Γ of course are to be interpreted analogously.

Let us somewhat more generally admit an additive loss function:

$$(2.5) \qquad L(\underline{\theta}, s) = \sum_{i \in s} L_2^{(i)}(\underline{\theta}) + \sum_{i \notin s} L_1^{(i)}(\underline{\theta}), s \in S,\ \underline{\theta} \in \underline{\Omega},$$

where the $L_j^{(i)}$'s all are non-negative functions.

Moreover let us denote the overall risk of a selection procedure $\underline{\psi} = (\psi_1, \ldots, \psi_k)$ with respect to a prior τ by

$$(2.6) \qquad r(\tau, \underline{\psi}) = \sum_{i=1}^{k} r^{(i)} (\tau, \psi_i),$$

where for every $i \in \{1, \ldots, k\}$

$$(2.7) \qquad r^{(i)}(\tau, \psi_i) =$$

$$= \int_{\underline{\Omega}} [L_2^{(i)} (\underline{\theta})\ E_{\underline{\theta}} \psi_i(\underline{X}) + L_1^{(i)} (\underline{\theta})\ (1 - E_{\underline{\theta}} \psi_i(\underline{X}))]\ d\tau(\underline{\theta})$$

is the overall risk of $\underline{\psi}$ in the corresponding i th component problem.

Let $\mathcal{G}$ denote any class of procedures. Then we can state:

<u>Lemma:</u> <u>A selection procedure</u> $\underline{\psi}^{\Gamma} = (\psi_1^{\Gamma}, \ldots, \psi_k^{\Gamma})$ <u>is</u> Γ-<u>minimax with respect to</u> $\mathcal{G}$ <u>if there exists a sequence of priors</u> $\tau_n \in \Gamma$, $n \in \mathbb{N} = \{1,2,\ldots\}$, <u>such that for every</u> $i \in \{1,\ldots,k\}$ <u>the following holds true:</u> <u>for the i th component problem there exists a sequence of Bayes-rules</u> ψ_{in}^B <u>with respect to</u> τ_n, $n \in \mathbb{N}$, <u>and</u> $\mathcal{G}$ <u>with</u>

$$(2.8) \qquad \liminf_{n \to \infty} \ r^{(i)} (\tau_n, \psi_{in}^B) \geq \sup_{\tau \in \Gamma} r^{(i)} (\tau, \psi_i^{\Gamma}).$$

Proof: Let $\underline{\psi} = (\psi_1, \ldots, \psi_k)$ be a selection procedure. Then

$$\sup_{\tau\in\Gamma} r(\tau,\underline{\psi}) = \sup_{\tau\in\Gamma} \sum_{i=1}^{k} r^{(i)}(\tau,\psi_i)$$

$$\geq \sup_{n} \sum_{i=1}^{k} r^{(i)}(\tau_n,\psi_i) \geq \sup_{n} \sum_{i=1}^{k} r^{(i)}(\tau_n,\psi_{in}^{B})$$

$$\geq \liminf_{n\to\infty} \sum_{i=1}^{k} r^{(i)}(\tau_n,\psi_{in}^{B})$$

$$\geq \sum_{i=1}^{k} \liminf_{n\to\infty} r^{(i)}(\tau_n,\psi_{in}^{B})$$

$$\geq \sum_{i=1}^{k} \sup_{\tau\in\Gamma} r^{(i)}(\tau,\psi_i^{\Gamma}) \geq \sup_{\tau\in\Gamma} \sum_{i=1}^{k} r^{(i)}(\tau,\psi_i^{\Gamma})$$

$$= \sup_{\tau\in\Gamma} r(\tau,\underline{\psi}^{\Gamma}).$$

3. Known Controls.

Within the framework given in the first part of section 1, we shall derive now Γ-minimax selection procedures with respect to $\mathscr{D}$ using techniques and results of the classical (Neyman and Pearson) theory of testing hypotheses. For every $i \in \{1,\ldots,k\}$ let $\varphi^*_{i,\alpha}$, $\alpha \in [0,1]$, denote the U.M.P. level α test - based on $\underline{X}_i$ - for the testing problem:

(3.1) $H_i : \theta_i \leq \theta_{oi}$ versus $K_i : \theta_i > \theta_{oi}$.

Moreover for every $i \in \{1,\ldots,k\}$ let $A_i \subseteq [0,1]$ be the set of values $\beta \in [0,1]$ satisfying

(3.2) $$L_{2i}\, \pi_i' \beta - L_{1i}\, \pi_i\, E_{\theta_{oi}} + \Delta_i\, \varphi^*_{i,\beta}(\underline{X}_i)$$

$$= \min_{\alpha \in [0,1]} \{L_{2i}\, \pi_i' \alpha - L_{1i}\, \pi_i\, E_{\theta_{oi}} + \Delta_i\, \varphi^*_{i,\alpha}(\underline{X}_i)\}.$$

The non-emptiness of $A_1, \ldots, A_k$ is guaranteed by (P.3) and (P.4). Now we can state

Theorem 1. Every selection procedure $\underline{\psi}^\Gamma(\underline{X}) = (\psi_1^\Gamma(\underline{X}), \ldots, \psi_k^\Gamma(\underline{X}))$ with $\psi_i^\Gamma(\underline{X}) = \varphi^*_{i,\alpha_i}(\underline{X}_i)$, $\alpha_i \in A_i$, $i = 1, \ldots, k$, is Γ-minimax with respect to $\mathscr{D}$.

Proof: We apply our Lemma for $\mathscr{G} = \mathscr{D}$ in the simple version of Randles and Hollander (1971), where the sequences of priors and Bayes-rules reduce to a single prior τ_o and to single Bayes-rules $\psi_1^B, \ldots, \psi_k^B$. Thus we have to find a $\tau_o \in \Gamma$ such that for every $i \in \{1, \ldots, k\}$ and every $\alpha_i \in A_i$

(a) φ^*_{i,α_i} is Bayes-rule with respect to τ_o for the i th component problem, and simultaneously

(b) τ_o is least favorable for φ^*_{i,α_i}.

Here we choose a τ_o similar to that considered by Randles and Hollander (1971): Under τ_o let $\Theta_1, \ldots, \Theta_k$ be independent and moreover for every $i \in \{1, \ldots, k\}$ let Θ_i assume the value $\theta_{oi} + \Delta_i$ $(\theta_{oi}, \theta_{oi} + \Delta_i/2)$ with probability π_i $(\pi_i', 1 - \pi_i - \pi_i')$.

Let $i \in \{1, \ldots, k\}$ be fixed and let $\tau_o^{(i)}$ denote the marginal distribution of $(\Theta_1, \ldots, \Theta_{i-1}, \Theta_{i+1}, \ldots, \Theta_k)$ under τ_o. Then for every selection procedure $\underline{\psi}$ the Bayes-risk for the i th component problem in view of (2.7) is given by

(3.3) $$r^{(i)}(\tau_o, \psi_i)$$

$$= E_{\tau_o^{(i)}} \{L_{2i}\, \pi_i'\, E_{(\Theta_1,\ldots,\Theta_{i-1},\theta_{oi},\Theta_{i+1},\ldots,\Theta_k)}\, \psi_i(\underline{X})$$

$$+ L_{1i}\, \pi_i\, [1 - E_{(\Theta_1,\ldots,\Theta_{i-1},\theta_{oi}+\Delta_i,\Theta_{i+1},\ldots,\Theta_k)}\, \psi_i(\underline{x})]\}$$

$$= L_{2i}\, \pi_i'\, E_{\theta_{oi}}\, \tilde{\psi}_i(\underline{X}_i) + L_{1i}\, \pi_i\, [1 - E_{\theta_{oi}+\Delta_i}\, \tilde{\psi}_i(\underline{X}_i)]$$

where

$$(3.4) \qquad \tilde{\psi}_i(\underline{X}_i) = E_{\tau_o}(i)\, E_{(\theta_1,\ldots,\theta_{i-1},\theta_{i+1},\ldots,\theta_k)}\, \psi_i(\underline{X})$$

can now be viewed as being a test - based on $\underline{X}_i$ - for the testing problem (3.1). Taking $\beta = E_{\theta_{oi}}\, \tilde{\psi}_i(\underline{X}_i)$, by (P.1) we get

$$(3.5) \qquad r^{(i)}\,(\tau_o,\psi_i) \geq r^{(i)}\,(\tau_o,\varphi^*_{i,\beta})$$

$$= L_{2i}\,\pi_i'\beta + L_{1i}\,\pi_i\,[1 - E_{\theta_{oi}+\Delta_i}\,\varphi^*_{i,\beta}(\underline{X}_i)].$$

Thus, the class of Bayes-rules for the i th component problem consists just of those φ^*_{i,α_i} with $\alpha_i \in A_i$, $i = 1,\ldots,k$. Thus (a) is proved. To prove (b) let $\alpha_i \in A_i$, $i = 1,\ldots,k$ and $\tau \in \Gamma$ be fixed. Then

$$(3.6) \qquad r(\tau,(\varphi^*_{1,\alpha_1},\ldots,\varphi^*_{k,\alpha_k})) = \sum_{i=1}^{k} r^{(i)}\,(\tau,\varphi^*_{i,\alpha_i})$$

$$= \sum_{i=1}^{k} E_\tau\,\{L_{2i}\, I_{(-\infty,\theta_{oi}]}(\theta_i)\, E_{\theta_i}\,\varphi^*_{i,\alpha_i}(\underline{X}_i)$$

$$+ L_{1i}\, I_{[\theta_{oi}+\Delta_i,\infty)}(\theta_i)\,[1 - E_{\theta_i}\,\varphi^*_{i,\alpha_i}(\underline{X}_i)]\}$$

$$\leq \sum_{i=1}^{k} \{ L_{2i}\,\pi_i'\, E_{\theta_{oi}}\,\varphi^*_{i,\alpha_i}(\underline{X}_i) + L_{1i}\,\pi_i\,[1 - E_{\theta_{oi}+\Delta_i}\,\varphi^*_{i,\alpha_i}(\underline{X}_i)]\}$$

$$= r(\tau_o,(\varphi^*_{1,\alpha_1},\ \ldots\ ,\varphi^*_{k,\alpha_k})),$$

where the inequality follows from (P.2) and the condition $\tau \in \Gamma$.

To implement Γ-minimax procedures in concrete situations, note that we can always take values $\alpha_i \in A_i$ which result in non-randomized tests φ^*_{i,α_i}, $i = 1,\ldots,k$. More precisely standard analysis leads to

Corollary 1. Every selection procedure $\psi^\Gamma = (\psi_1^\Gamma, \ldots, \psi_k^\Gamma)$ with $\psi_i^\Gamma(\underline{X}) = 1\ (0)$ as $Z_i \geq (<)\ c_i$, where c_i satisfies

(3.8) $$[L_{2i}\, \pi_i'\, f_{i,\theta_{oi}}(c) - L_{1i}\, \pi_i\, f_{i,\theta_{oi}+\Delta_i}(c)]\,(c_i - c) \geq 0,$$

$c \neq c_i$, $i = 1,\ldots,k$, is Γ-minimax with respect to $\mathcal{D}$.

4. Unknown Control.

In this section we shall derive Γ-minimax procedures within the framework described in the second part of section 1, and again we shall use techniques and results from the classical (Neyman and Pearson) theory of testing hypotheses.

Since we now are dealing with location parameter M.L.R.-families, we can utilize the following well-known fact (for a proof see Randles and Hollander (1971), Lemma 4.2):

For every $i \in \{1,\ldots,k\}$, if θ_i and θ_o are the location parameters of Z_i and Z_o, then $Y_i = Z_i - Z_o$ has the density

(4.1) $$g_i(y - (\theta_i - \theta_o)) = \int_{-\infty}^{\infty} f_i(y+u-\theta_i) f(u-\theta_o)\, du \qquad y \in \mathbb{R},$$

which likewise has the M.L.R. property if $\delta_i = \theta_i - \theta_o$ is considered as location parameter for Y_i.

For every $i \in \{1,\ldots,k\}$ let $\overline{\varphi}^*_{i,\alpha}$, $\alpha \in [0,1]$, denote the U.M.P. level α test - based on Y_i - for the testing problem

(4.2) $$\overline{H}_i : \delta_i \leq 0 \text{ versus } \overline{K}_i : \delta_i > 0.$$

Moreover let for every $i \in \{1,\ldots,k\}$ $\overline{A}_i \subseteq [0,1]$ be the set of values $\beta \in [0,1]$ satisfying

$$(4.3)\qquad L_{2i}\,\pi_i'\,\beta - L_{1i}\,\pi_i\,E_{\Delta_i}\,\overline{\varphi}^*_{i,\beta}\,(Y_i)$$

$$= \min_{\alpha \in [0,1]} \{ L_{2i}\,\pi_i'\alpha - L_{1i}\,\pi_i\,E_{\Delta_i}\,\overline{\varphi}^*_{i,\alpha}\,(Y_i)\}.$$

The non-emptiness of $\overline{A}_1,\ldots,\overline{A}_k$ again is guaranteed by (P.3) and (P.4).

Now we can prove the following theorem stated in Randles and Hollander (1971):

Theorem 2. Every selection procedure $\underline{\psi}^\Gamma(\underline{X}) = (\psi_1^\Gamma(\underline{X}), \ldots, \psi_k^\Gamma(\underline{X}))$ with $\psi_i^\Gamma(\underline{X}) = \overline{\varphi}^*_{i,\alpha_i}(Y_i)$, $\alpha_i \in \overline{A}_i$, $i = 1,\ldots,k$, is Γ-minimax with respect to $\mathscr{D}^o$.

Proof: The basic tool for the proof will be our Lemma for $\mathscr{G} = \mathscr{D}^o$, now in its general version.

Let us represent every $\tau \in \Gamma$ by $\tau = (T,t)$, say, where t denotes the conditional distribution of $(\Theta_1, \ldots, \Theta_k)$ - given Θ_o - , and T denotes the marginal distribution of Θ_o. Especially let t be denoted by t_{θ_o} if $\Theta_o = \theta_o$ is given.

Our sequence $\tau_n \in \Gamma$, $n \in \mathbb{N}$, now is chosen as follows: for every $n \in \mathbb{N}$ let $\tau_n = (T_n, w)$, where

(i) T_n is the uniform distribution over $[-n,n]$, $n \in \mathbb{N}$, and

(ii) Given $\Theta_o = \theta_o$ - under w_{θ_o}, $\Theta_1, \ldots, \Theta_k$ are independent, where each Θ_i assumes the value $\theta_o + \Delta$ $(\theta_o, \theta_o + \Delta/2)$ with probability π_i $(\pi_i', 1 - \pi_i - \pi_i')$, $i = 1,\ldots,k$.

Finally let $w_{\theta_o}^{(i)}$ denote the marginal distribution of $(\Theta_1, \ldots, \Theta_{i-1}, \Theta_{i+1}, \ldots, \Theta_k)$ under w_{θ_o}, $\theta_o \in \mathbb{R}, i = 1,\ldots,k$.

First we derive within $\mathscr{D}^o$ appropriate Bayes-rules ψ_{in}^B with respect to τ_n, $n \in \mathbb{N}$ $i = 1,\ldots,k$:

Let $i \in \{1,\dots,k\}$ be fixed up to the end of the proof. Then for every selection procedure $\underline{\psi} = (\psi_1, \dots, \psi_k) \in \mathscr{D}^o$ and $n \in \mathbb{N}$ the overall risk with respect to τ_n for the i th component problem is given by

(4.4) $$r^{(i)}(\tau_n,\psi_i)$$

$$= \int_{\mathbb{R}} E_{w_{\theta_o}^{(i)}} \Big\{ L_{2i}\, \pi_i'\, E_{(\theta_o,\Theta_1,\dots,\Theta_{i-1},\theta_o,\Theta_{i+1},\dots,\Theta_k)}\, \psi_i\, (\underline{X}_o,\underline{X}_i)$$

$$+ L_{1i}\, \pi_i \left[1 - E_{(\theta_o,\Theta_1,\dots,\Theta_{i-1},\theta_o+\Delta_i,\Theta_{i+1},\dots,\hat{\Theta}_k)}\, \psi_i\, (\underline{X}_o,\underline{X}_i)\right] \Big\}\, dT_n(\theta_o)$$

$$= \int_{\mathbb{R}} \Big\{ L_{2i}\, \pi_i'\, E_{(\theta_o,\theta_o)}\, \psi_i\, (\underline{X}_o,\underline{X}_i)$$

$$+ L_{1i}\, \pi_i \left[1 - E_{(\theta_o,\theta_o+\Delta_i)}\, \psi_i\, (\underline{X}_o,\underline{X}_i)\right] \Big\}\, dT_n(\theta_o).$$

Thus the Bayes-rules (which by definition have to minimize this risk) turn out to be of the form

(4.5) $$\psi_{in}^B\, (\underline{X}_o,\underline{X}_i) = 1\ (h,\ 0) \text{ iff}$$

$$L_{2i}\, \pi_i' \int_{\mathbb{R}} f_o\,(Z_o - \theta_o)\, f_i\,(Z_i - \theta_o)\, d\, T_n\,(\theta_o)$$

$$< (=,>)\ L_{1i}\, \pi_i \int_{\mathbb{R}} f_o\,(Z_o - \theta_o)\, f_i\,(Z_i - \theta_o - \Delta_i)\, d\, T_n\,(\theta_o),$$

where $h = h(Z_o,Z_i) \in [0,1]$ may be chosen arbitrarily, since it has no influence upon the risk.

Our next step is to verify (2.8). To begin with, note that for every $\alpha \in [0,1]$ we have

$$(4.6) \quad \sup_{\tau \in \Gamma} r^{(i)}(\tau, \overline{\varphi}^*_{i,\alpha})$$

$$= \sup_{T} \sup_{t,(T,t)\in\Gamma} r^{(i)}((T,t), \overline{\varphi}^*_{i,\alpha})$$

$$= \sup_{T} \sup_{t,(T,t)\in\Gamma} \int_{\mathbb{R}} E_{t_{\theta_o}} \Big\{ L_{2i}\, I_{(-\infty,0]}(\Theta_i - \theta_o) E_{\Theta_i - \theta_o} \overline{\varphi}^*_{i,\alpha}(Y_i)$$

$$+ L_{1i}\, I_{[\Delta_i,\infty)}(\Theta_i - \theta_o) \left[1 - E_{\Theta_i - \theta_o} \overline{\varphi}^*_{i,\alpha}(Y_i)\right] \Big\}\, d\,T(\theta_o)$$

$$= \sup_{T} \int_{\mathbb{R}} \Big\{ L_{2i}\, \pi_i'\, E_o\, \overline{\varphi}^*_{i,\alpha}(Y_i)$$

$$+ L_{1i}\, \pi_i \left[1 - E_{\Delta_i} \overline{\varphi}^*_{i,\alpha}(Y_i)\right] \Big\}\, dT(\theta_o)$$

$$= L_{2i}\, \pi_i' \alpha + L_{1i}\, \pi_i \left[1 - E_{\Delta_i} \overline{\varphi}^*_{i,\alpha}(Y_i)\right] .$$

Unfortunately, to verify (2.8) we have to leave now the scope of power functions and enter that of density functions (see also our Remark 2 stated below). For this purpose we note at first that for every $\alpha_i \in \overline{A}_i$ we have

$$(4.7) \quad L_{2i}\, \pi_i'\, \alpha_i + L_{1i}\, \pi_i \left[1 - E_{\Delta_i} \overline{\varphi}^*_{i,\alpha_i}(Y_i)\right]$$

$$= \int_{\mathbb{R}} \min \Big\{ L_{2i}\, \pi_i'\, g_i(u),\ L_{1i}\, \pi_i\, g_i(u - \Delta_i) \Big\}\, du.$$

This follows from the simple fact that every $\overline{\varphi}^*_{i,\alpha_i}$ with $\alpha_i \in \overline{A}_i$ can be viewed as being a Bayes-test for the following auxiliary Bayesian testing problem, where Y_i is known to have either the density $g_i(u)$ or the density $g_i(u - \Delta_i)$, the losses of errors of the first and second kind both are equal to one and the prior probabilities are $L_{2i}\, \pi_i' (L_{1i}\, \pi_i + L_{2i}\, \pi_i')^{-1}$ and $L_{1i}\, \pi_i (L_{1i}\, \pi_i + L_{2i}\, \pi_i')^{-1}$, respectively.

Thus to verify (2.8) it suffices to prove that we have

$$(4.8)\qquad \liminf_{n\to\infty} \; r^{(i)}\,((T_n,w),\,\psi^B_{in})$$

$$\geq \int_{\mathbb{R}} \min\,\{L_{2i}\,\pi_i'\,g_i(u),\; L_{1i}\,\pi_i\,g_i(u-\Delta_i)\}\,du.$$

Now by (4.4) and (4.5) and the fundamental principle of Bayesian analysis (i.e., interchanging the order of integration with respect to observables and parameters) we get

$$(4.9)\qquad r^{(i)}\,((T_n,w),\,\psi^B_{in})$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} \min\,\{L_{2i}\,\pi_i'\,\frac{1}{2n}\int_{-n}^{n} f_o(z_o-\eta)\,f_i(z_i-\eta)\,d\eta,$$

$$L_{1i}\,\pi_i\,\frac{1}{2n}\int_{-n}^{n} f_o(z_o-\eta)\;f_i(z_i-\Delta_i-\eta)\;d\eta\}\;dz_o\;dz_i.$$

Substituting first $z_o = nv + u$ and $z_i = nv - u$ in the outer integrals and then $\eta = -\xi + nv$ in the interior ones, we arrive at

$$(4.10)\qquad r^{(i)}\,((T_n,w),\,\psi^B_{in})$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} \min\left\{L_{2i}\,\pi_i'\int_{n(v-1)}^{n(v+1)} f_o(\xi+u)\;f_i(\xi-u)\;d\xi,\right.$$

$$\left.L_{1i}\,\pi_i\int_{n(v-1)}^{n(v+1)} f_o\,(\xi+u)\;f_i(\xi-\Delta_i-u)\;d\xi\right\}\,dv\,du$$

$$\geq \int_{\mathbb{R}}\int_{-1}^{1} \min\left\{L_{2i}\,\pi_i'\int_{n(v-1)}^{n(v+1)} f_o(\xi+u)\;f_i(\xi-u)\;d\xi,\right.$$

$$\left.L_{1i}\,\pi_i\int_{n(v-1)}^{n(v+1)} f_o(\xi+u)\;f_i(\xi-\Delta_i-u)\;d\xi\right\}\,dv\,du$$

Now, after all, we are in position to apply Lebesgue's dominated convergence theorem to the last double-integral, since the "min"-terms converge to

$\min \{L_{2i}\, \pi_i'\, g_i\, (2u),\ L_{1i}\, \pi_i\, g_i\, (2u - \Delta_i)$ (pointwise in $(u,v) \in \mathbb{R} \times (-1,1)$) which, at the same time, may serve as an integrable upper bound for them. Thus, we have

$$(4.11) \qquad \liminf_{n \to \infty} \; r^{(i)} \left((T_n, w),\ \psi_{in}^{B} \right)$$

$$\geq \int_{\mathbb{R}} \int_{-1}^{1} \min \{L_{2i}\, \pi_i'\, g_i\, (2u),\ L_{1i}\, \pi_i\, g_i\, (2u - \Delta_i)\}\, dv\, du$$

$$= \int_{\mathbb{R}} \min \{L_{2i}\, \pi_i'\, g_i\, (u),\ L_{1i}\, \pi_i\, g_i\, (u - \Delta_i)\}\, du$$

and therefore (4.8) is verified and our proof is completed.

Remark 1. Since by (4.6) we know that for every $i \in \{1,\ldots,k\}$ and $\alpha_i \in \overline{A}_i$

$$(4.12) \qquad \sup_{\tau \in \Gamma} \; r^{(i)} (\tau,\ \overline{\varphi}^{*}_{i,\alpha_i}) = r^{(i)} ((T,w),\ \overline{\varphi}^{*}_{i,\alpha_i})$$

holds for all T and therefore in particular for $T_1, T_2, \ldots$, and since we moreover know that

$$(4.13) \qquad r^{(i)} \left((T_n,w),\ \psi_{in}^{B} \right) \leq r^{(i)} \left((T_n,w),\ \overline{\varphi}^{*}_{i,\alpha_i} \right)$$

holds for all $n \in \mathbb{N}$, we have in fact proved that for every ψ_i^{Γ} (as defined in Theorem 2)

$$(4.14) \qquad \lim_{n \to \infty} \; r^{(i)} (\tau_n,\ \psi_{in}^{B}) = \sup_{\tau \in \Gamma} \; r^{(i)} (\tau,\ \psi_i^{\Gamma}).$$

Remark 2. At no point prior to inequality (4.10) it was possible to complete the proof with the help of Lebesgue's dominated convergence theorem alone. A thorough analysis shows that one is always faced finally with the impossibility of finding an appropriate integrable upper bound. Clearly this is because the sequence T_n, $n \in \mathbb{N}$ approximates an improper distribution (the "uniform distribution over $\mathbb{R}$").

It should be pointed out that there is an interesting paper by C. Stein (1965) where admissibility of formal Bayes-rules with respect to improper priors is under concern. Clearly, the Γ-minimax procedures of this section are of this type. But, unfortunately, Stein's ideas do not apply in our framework, since he used Taylor expansions of $L(\underline{\theta},s)$ with respect to the second argument s, which of course is only possible in cases where the action space is an interval of $\mathbb{R}$, thus in estimation problems for example. On the other hand, it is not difficult to see that the Γ-minimax procedures of section 3 are in fact admissible.

In analogy to section 3 we conclude this section with

Corollary 2. Every selection procedure $\underline{\psi} = (\psi_1^\Gamma, \ldots, \psi_k^\Gamma)$ with $\psi_i^\Gamma(\underline{X}) = 1\ (0)$ as $Y_i \geq (<)\ c_i$, where c_i satisfies

$$(4.15) \qquad [L_{2i}\,\pi_i'\,g_i(c) - L_{1i}\,\pi_i\,g_i(c-\Delta_i)]\,(c_i - c) \geq 0,$$

$c \neq c_i$, $i = 1,\ldots,k$, is Γ-minimax with respect to $\mathscr{D}^o$.

5. An Example.

Let $\underline{X}_i = (X_{i1}, \ldots, X_{in_i})$ be independent samples from normal populations $N(\theta_i, \sigma^2)$, $i = 0,1,\ldots,k$. Then for given Δ_i, L_{1i}, L_{2i}, π_i, π_i', $i = 1,\ldots,k$, and $\sigma^2 > 0$, Theorem 2 provides us with the Γ-minimax procedure (see also Randles and Hollander (1971)):

$$(5.1) \qquad \overline{\psi}_i^\Gamma(\underline{X}) = 1 \quad \text{iff } \overline{X}_i - \overline{X}_o \geq \frac{\Delta_i}{2} + \frac{\sigma^2}{\Delta_i}\left(\frac{1}{n_i} + \frac{1}{n_o}\right) \ln \frac{L_{2i}\,\pi_i'}{L_{1i}\,\pi_i},$$

$i = 1,\ldots,k$, where $\overline{X}_o, \overline{X}_1, \ldots, \overline{X}_k$ are the corresponding sample means.

Now the natural question arises whether in case of $n_o \geq k$ it is better to split the sample $\underline{X}_o$ into k disjoint sub-samples $\underline{X}_{oi}$ of sizes m_i with $m_1 + \ldots + m_k = n_o$, then to switch over to the differences of sample means $\overline{X}_i - \overline{X}_{oi}$, i = 1,...,k, and finally to take the Γ-minimax procedure provided by Theorem 1:

$$\text{(5.2)} \qquad \psi_i^{\Gamma}(\underline{X}) = 1 \text{ iff } \overline{X}_i - \overline{X}_{oi} \geq \frac{\Delta_i}{2} + \frac{\sigma^2}{\Delta_i}\left(\frac{1}{n_i} + \frac{1}{m_i}\right) \ln \frac{L_{2i}\,\pi_i'}{L_{1i}\,\pi_i}, \quad i = 1,\ldots,k.$$

If comparison is made in terms of supremal risks over the corresponding Γ's, the answer, as one may expect, turns out to be in favor of $\overline{\underline{\psi}}^{\Gamma}$: Since the risk components to be compared with are given by (3.2) and (4.3), respectively, and since moreover the tests $\varphi^*_{i,\beta}$ and $\overline{\varphi}^*_{i,\beta}$, associated with ψ_i and $\overline{\psi}_i^{\Gamma}$, i = 1,...,k, in view of (2.3) satisfy

$$\text{(5.3)} \qquad E_{\Delta_i}\, \varphi^*_{i,\beta}\,(\overline{X}_i - \overline{X}_{oi})$$

$$= \Phi(\Phi^{-1}(\beta) + \Delta_i\,(\,(\frac{1}{n_i} + \frac{1}{m_i})\,\sigma^2)^{-1/2})$$

$$\leq \Phi(\Phi^{-1}(\beta) + \Delta_i\,(\,(\frac{1}{n_i} + \frac{1}{n_o})\,\sigma^2)^{-1/2})$$

$$= E_{\Delta_i}\, \overline{\varphi}^*_{i,\beta}\,(\overline{X}_i - \overline{X}_o),\ \beta \in [0,1],\ i = 1,\ldots,k,$$

where Φ denotes the c.d.f. of N(0,1), we conclude that

$$\text{(5.4)} \qquad \sup_{\tau \in \Gamma} r(\tau, \underline{\psi}^{\Gamma}) \geq \sup_{(T,t) \in \overline{\Gamma}} r(\,(T,t),\, \overline{\underline{\psi}}^{\Gamma})$$

holds, where Γ and $\overline{\Gamma}$ denote the corresponding classes of priors.

If, on the other hand, comparison is made in terms of risks, this time pointwise over pairs of "comparable priors," then it is not possible to give an unique answer in favor of one of the two competing procedures. But there is an exception:

Suppose we are given $L_{1i}\ \pi_i = L_{2i}\ \pi_i'$, $1,\dots,k$. Then if τ is any prior of $(\Theta_1 - \Theta_o, \dots, \Theta_k - \Theta_o)$ with respect to $\underline{\psi}^\Gamma$ and if (T,t) is a "comparable prior" of $(\Theta_o, \Theta_1, \dots, \Theta_k)$ with respect to $\overline{\underline{\psi}}^\Gamma$ in such a way that under (T,t) the conditional distribution of $\Theta_i - \Theta_o$ - given $\Theta_o = \theta_o$ - does not depend on $\theta_o \in \mathbb{R}$ (i.e., θ_o acts here as location parameter) and coincides with the marginal distribution of $\Theta_i - \Theta_o$ under τ, denoted by λ_i, say, $i = 1,\dots,k$, then standard analysis leads us to

(5.5) $$r^{(i)}\ (\ (T,t),\ \overline{\underline{\psi}}^\Gamma) =$$

$$= L_{2i} \int_{-\infty}^{0} \Phi(\ (\eta - \frac{\Delta_i}{2})\ (\ (\frac{1}{n_i} + \frac{1}{n_o})\ \sigma^2)^{-1/2})\ d\lambda_i\ (\eta)$$

$$+ L_{1i} \int_{\Delta_i}^{\infty} \Phi(\ (\frac{\Delta_i}{2} - \eta)\ (\ (\frac{1}{n_i} + \frac{1}{n_o})\ \sigma^2)^{-1/2})\ d\lambda_i\ (\eta)$$

$$\leq L_{2i} \int_{-\infty}^{0} \Phi(\ (\eta - \frac{\Delta_i}{2})\ (\ (\frac{1}{n_i} + \frac{1}{m_i})\ \sigma^2)^{-1/2})\ d\lambda_i\ (\eta)$$

$$+ L_{1i} \int_{\Delta_i}^{\infty} \Phi(\ (\frac{\Delta_i}{2} - \eta)\ (\ (\frac{1}{n_i} + \frac{1}{m_i})\ \sigma^2)^{-1/2})\ d\lambda_i\ (\eta)$$

$$= r^{(i)}\ (\tau,\ \underline{\psi}^\Gamma)\ ,\ i = 1,\dots,k.$$

Acknowledgement: I wish to express my sincere thanks to Dr. Woo-Chul Kim for pointing me out a difficulty in the proof of Theorem 2 which makes it impossible to extend the results to a larger class of procedures. Moreover, he brought to my attention that Randle's and Hollander's proof is incomplete.

Klaus J. Miescke

Fachbereich Mathematik

Universität Mainz

Saarstrasse 21

6500 Mainz

Fed. Rep. Germany

SUMMARY

Γ-Minimax Selection Procedures in Simultaneous Testing Problems

Suppose we have to decide on the basis of appropriately drawn samples which of k treatment populations are good compared to a control population, where the populations can be identified by certain parameters $\theta_1,\ldots,\theta_k$ and θ_o. In the first part of this paper we assume that θ_o is known and that $\theta_1,\ldots,\theta_k$ vary randomly according to a certain prior distribution, about which we only have the partial knowledge that it is contained in a given class Γ of priors. In the second part θ_o additionally is assumed to vary at random. Though we derive in both cases (under the assumption of monotone likelihood ratios) Γ-minimax procedures which by definition attain minimal supremal risk over Γ, the emphases are different: while we try to demonstrate in the "known controls case" how well known results from the theory of testing hypotheses can be utilized to solve the problem, our main purpose in the "unknown control case" is to give a new proof for a theorem which was stated but only partially proved by Randles and Hollander (1971). Finally, an example in the "paired comparison versus random sampling" setup is given.

A M S subject classifications. - Primary 62 F 07

Secondary 62 F 15

Key words and phrases. Gamma minimax procedures, simultaneous testing, Bayesian procedures, improper prior distributions.

REFERENCES

[1] Blum, J. R. and Rosenblatt, J. (1967). On partial a priori information in statistical inference. Ann. Math. Statist. 38, 1671-1678.

[2] Gupta, S. S. and Huang, D. Y. (1975). A note on Γ-minimax classification procedures. In: Proceedings of the 40th Session of the International Statistical Institute. Bulletin of the I.S.I. Vol. 46, Book 3, 330-335.

[3] Gupta, S. S. and Huang, D. Y. (1977). On some Γ-minimax selection and multiple comparison procedures. In: Statistical Decision Theory and Related Topics II. (ed. Gupta, S. S. and Moore, D. S.), Academic Press, New York, 139-155.

[4] Huang, D. Y. (1974). Some contributions to fixed sample and sequential multiple decision (selecting and ranking) theory. Mimeo Series No. 363, Department of Statistics, Purdue University.

[5] Lehmann, E. L. (1957). A theory of some multiple decision problems I. Ann. Math. Statist. 28, 1-25.

[6] Lehmann, E. L. (1959). Testing Statistical Hypotheses. J. Wiley, New York.

[7] Lehmann, E. L. (1961). Some model I problems of selection. Ann. Math. Statist. 32, 990-1012.

[8] Miescke, K. J. (1979). Identification and selection procedures based on tests. Ann. Statist. 7, 1207-1219.

[9] Nagel, K. (1970). On subset selection rules with certain optimality properties. Mimeo Series No. 222, Department of Statistics, Purdue University.

[10] Randles, H. R. and Hollander, M. (1971). Γ-minimax selection procedures in treatment versus control problems. Ann. Math. Statist. 42, 330-341.

[11] Stein, C. (1965). Approximation of improper prior measures by prior probability measures. In: "Bernoulli, Bayes, Laplace" Anniversary Volume. Proceedings of an International Research Seminar, Statistical Laboratory, University of California, Berkeley, 1963. (ed. Neyman, J. and LeCam, L. M.), Springer Verlag, Berlin.