SLIPPAGE PROBLEMS[1]

By Samuel Karlin and Donald Truax

Stanford University

1. Introduction. Slippage problems have been considered in the literature by Mosteller [6], Paulson [8], Truax [11], Doornbos and Prins [2], Kudo [5], and others. Roughly, the problem is as follows: We wish to compare n populations which have density functions $f(x, \theta_1), f(x, \theta_2), \cdots, f(x, \theta_n)$. On the basis of a sample from each population we want to decide if all the θ_i are equal, or, if not, which is the largest. Actually, a more restricted problem is considered in this paper, in which either all parameter values are equal, or all but one are equal and the exceptional one is larger. If the ith one is larger we will say it has slipped to the right. These slippage problems have certain similarities with the problem of ranking means considered by Bechhofer and others [1], but differ in that the latter deal mostly with procedures guaranteeing with prescribed probability the selection of the population with the largest parameter, where it is known in advance that one parameter exceeds all the others. These authors never allow the possibility that all parameters of the various populations are equal, which, in our situation, is called hypothesis zero. Other contrasts between the two problems will become apparent in our later discussions.

A slightly different problem can be formulated in which we have in addition a control population. The problem is then to compare the n populations with the control, and decide if all the parameters are equal to the parameter of the control population, or, if not, which of the n populations has the larger parameter. In order to obtain optimal solutions to the slippage problems, certain invariance restrictions will be imposed. Notice the obvious symmetry that states, if $X_1, X_2, \cdots, X_n$ is observed (X_i is an observation from the ith population) and if action j is appropriate (i.e., the jth parameter has slipped to the right) then if a permutation $X_{\pi 1}, X_{\pi 2}, \cdots, X_{\pi n}$ is observed, action πj is appropriate. This suggests restricting attention to symmetric procedures. That is, if $\varphi_i(X_1, X_2, \cdots, X_n)$ denotes the probability of taking action i when $X_1, X_2, \cdots, X_n$ is observed, then we will require $\varphi_{\pi i}(X_{\pi 1}, X_{\pi 2}, \cdots, X_{\pi n}) = \varphi_i(X_1, X_2, \cdots, X_n)$ for all permutations $(1, 2, \cdots, n) \to (\pi 1, \pi 2, \cdots, \pi n)$. We will further restrict attention exclusively to those problems in which it is possible to reduce the problem, by invariance, to a one parameter problem. In particular we will investigate several cases where the parameter is a translation or scale parameter.

The nature of the Bayes solutions will be examined for these problems. The Bayes solutions are usually fairly easy to characterize, and many problems lead us to complete classes of solutions. We will show that any symmetric Bayes solu-

---

Received June 16, 1959.

[1] This work was supported in part by an Office of Naval Research contract at Stanford University.

tion, which usually is Bayes against a symmetric distribution, can be explicitly evaluated.

One can conceive of more general problems than those we will discuss here. For example, we have considered only slippage, in the case of a one dimensional parameter, to the right. A more general problem would be that in which the direction of the slippage is not specified. Also, one might consider the problem in which a subset of the parameters has slipped, and we are to decide which subset it is, etc. Modifications of our arguments would apply to these more general problems.

In Section 2 we introduce the pertinent definitions and terminology. In Section 3 some preliminary lemmas are proved and the Bayes solutions are characterized in general form. In the following section the theory is applied to several examples, including the slippage problem of the means of normal populations with common variance and the slippage problem of the parameter of a Gamma family of distributions. Part of this discussion deals with known examples in a more direct manner, while other examples are new.

In Section 5 we study the slippage problem for populations having an unknown translation parameter. The problem is set up in a non-parametric form. The solutions of the slippage problems of normal variables and exponential variables are obtained by applying the theory of the translation parameter slippage problem in the case of the existence of a sufficient statistic. In a similar manner, the symmetric invariant Bayes solutions are explicitly determined in the case of slippage of a scale parameter possessing a sufficient statistic. Mixed translation and scale parameter problems are discussed in the following section.

In Section 8 it is shown under fairly general conditions that the symmetric Bayes procedures, characterized in the earlier sections, are uniformly most powerful amongst all symmetric procedures having the same error of rejecting hypothesis zero, when it is true.

In Section 9 we discuss a multivariate slippage problem. Two slippage problems for non-parametric situations are introduced in Section 10. Some decision procedures based on rank tests are proposed for their solution. In the last section a few remarks are offered about computing the critical numbers defining the symmetric Bayes solutions.

2. Preliminaries and definitions. The slippage problem can be formulated in the following decision theoretic way. We observe an n-dimensional random variable $X = (X_1, X_2, \cdots, X_n)$ distributed according to a density function $p(x_1, x_2, \cdots, x_n; \theta_1, \theta_2, \cdots, \theta_n)$ which is known except for the parameter point $\theta = (\theta_1, \theta_2, \cdots, \theta_n)$ where the θ_i are real numbers. We assume the following symmetry of the density:

$$p(x_{\pi 1}, x_{\pi 2}, \cdots, x_{\pi n}; \theta_{\pi 1}, \theta_{\pi 2}, \cdots, \theta_{\pi n}) = p(x_1, x_2, \cdots, x_n; \theta_1, \theta_2, \cdots, \theta_n)$$

for all permutations $(1, 2, \cdots, n) \to (\pi 1, \pi 2, \cdots, \pi n)$. There are $n + 1$ available actions which we will call $a_0, a_1, \cdots, a_n$, and the loss in taking action a_j

when θ is the true parameter point is $L_j(\theta)$. The loss functions are assumed to have the following properties.

(1) $L_{\pi j}(\theta_{\pi 1}, \theta_{\pi 2}, \cdots, \theta_{\pi n}) = L_j(\theta_1, \theta_2, \cdots, \theta_n)$ for all permutations $(1, 2, \cdots, n) \rightarrow (\pi 1, \pi 2, \cdots, \pi n)$ and for all j. (We may include the case $j = 0$ by defining $\pi 0 = 0$.)

(2) $L_j(\theta) < L_i(\theta)$ for all $i \neq j$ if $\theta_j = \omega + \Delta$ for some real ω and $\Delta > 0$, and $\theta_i = \omega$ for all $i \neq j$. $L_0(\theta) < \min_{1 \leqq i \leqq n} L_i(\theta)$ if $\theta_i = \omega$ for some real ω, $i = 1, 2, \cdots, n$. Otherwise, $L_0(\theta) = L_1(\theta) = \cdots = L_n(\theta)$.

For the problem in which there is a control population, a modification of the above is needed. In this case $(X_1, X_2, \cdots, X_n, Y)$ is observed according to a density

$$p(x_1, x_2, \cdots, x_n, y; \theta_1, \cdots, \theta_n, \theta),$$

which satisfies

$$p(x_{\pi 1}, x_{\pi 2}, \cdots, x_{\pi n}, y; \theta_{\pi 1}, \theta_{\pi 2}, \cdots, \theta_{\pi n}, \theta) = p(x_1, x_2, \cdots, x_n, y; \theta_1, \cdots, \theta_n, \theta)$$

for all permutations $\pi:(1, 2, \cdots, n) \rightarrow (\pi 1, \cdots, \pi n)$. The loss functions satisfy

(1′) $L_{\pi j}(\theta_{\pi 1}, \cdots, \theta_{\pi n}, \theta) = L_j(\theta_1, \cdots, \theta_n, \theta)$ for all permutations $\pi:(1, 2, \cdots, n) \rightarrow (\pi 1, \pi 2, \cdots, \pi n)$ and for all j. (We include $j = 0$ by defining $\pi 0 = 0$.)

(2′) $L_j(\theta_1, \cdots, \theta_n, \theta) < L_i(\theta_1, \cdots, \theta_n, \theta)$ for all $i \neq j$ if $\theta_j = \theta + \Delta$ for some $\Delta > 0$ and $\theta_i = \theta$ for all $i \neq j$. $L_0(\theta_1, \cdots, \theta_n, \theta) < L_i(\theta_i, \cdots, \theta_n, \theta)$ $(1 \leqq i \leqq n)$ if $\theta_i = \theta$ for all i. Otherwise, $L_0(\theta_1, \cdots, \theta_n, \theta) = \cdots = L_n(\theta_1, \cdots, \theta_n, \theta)$.

Definition 2.1: A symmetric decision function is a vector function $\varphi = (\varphi_0, \varphi_1, \cdots, \varphi_n)$ on R^n with

(i) $0 \leqq \varphi_i(x) \leqq 1$,
(ii) $\sum_{i=0}^n \varphi_i(x) = 1$,
(iii) $\varphi_{\pi j}(x_{\pi 1}, x_{\pi 2}, \cdots, x_{\pi n}) = \varphi_j(x_1, x_2, \cdots, x_n)$ for all permutations $(1, 2, \cdots, n) \rightarrow (\pi 1, \pi 2, \cdots, \pi n)$ and all j (including $j = 0$, with $\pi 0 = 0$).

3. Bayes solutions.

Definition 3.1: If Φ denotes the set of all possible decision functions, Θ the parameter space, the *risk function* ρ is a function defined on $\Phi \times \Theta$ by

$$\rho(\varphi, \theta) = \int \sum_{i=0}^{n} L_i(\theta)\varphi_i(x)p(x, \theta)\, dx$$

where $x = (x_1, \cdots, x_n)$, $\theta = (\theta_1, \cdots, \theta_n)$, and dx denotes ordinary Lebesgue measure on R^n.

Lemma 3.1: *The risk function of a symmetric decision function φ is a symmetric function of θ.*

PROOF: Let $(1, 2, \cdots, n) \to (\pi 1, \pi 2, \cdots, \pi n)$ be a permutation and let $\pi 0 = 0$. Then

$$\int \sum_{i=0}^{n} L_i(\theta)\varphi_i(x)p(x;\theta)\,dx = \int \sum_{i=0}^{n} L_{\pi i}(\theta_\pi)\varphi_{\pi i}(x_\pi)p(x_\pi;\theta_\pi)\,dx_\pi = \rho(\varphi, \theta_\pi),$$

where $\theta_\pi = (\theta_{\pi 1}, \theta_{\pi 2}, \cdots, \theta_{\pi n})$.

DEFINITION 3.2: A decision function φ^0 is said to be *Bayes against a distribution F* if

$$\int \rho(\varphi^0, \theta)\,dF(\theta) = \min_{\varphi \varepsilon \Phi} \int \rho(\varphi, \theta)\,dF(\theta).$$

THEOREM 3.1: *Any symmetric Bayes solution is Bayes against a symmetric distribution.*

PROOF: Let φ be symmetric. Then by Lemma 3.1, $\rho(\varphi, \theta) = \rho(\varphi, \theta_\pi)$ for all permutations $(1, 2, \cdots, n) \to (\pi 1, \pi 2, \cdots, \pi n)$. If φ is Bayes against a distribution F, φ minimizes

$$\int \rho(\varphi, \theta)\,dF(\theta) = \int \rho(\varphi, \theta_\pi)\,dF(\theta_\pi) = \int \rho(\varphi, \theta)\,dF(\theta_\pi).$$

Define the distribution function $F^*(\theta)$ by $F^*(\theta) = 1/n!\sum F(\theta_\pi)$ where the sum is taken over all permutations. F^* is symmetric and φ is Bayes against F^* since $\int\rho(\varphi, \theta)\,dF(\theta) = \int\rho(\varphi, \theta)\,dF^*(\theta)$.

It is well known ([4], p. 279) that the Bayes solutions have the form

$$\varphi_j(x) = 1 \text{ if } \int L_j(\theta)p(x;\theta)\,dF(\theta) < \int L_i(\theta)p(x;\theta)\,dF(\theta) \quad \text{for all } i \neq j.$$

In computing the Bayes procedures we examine expressions of the type

$$\int [L_i(\theta) - L_j(\theta)]p(x;\theta)\,dF(\theta)$$

for changes in sign. Since $L_i(\theta) - L_j(\theta) = 0$ for all those points θ not of the form $\theta_k = \omega$ for all k except possibly a single value, where $\theta_i = \omega + \Delta$, $\Delta \geqq 0$, we may restrict attention to those distributions F whose spectrum is contained in this set of points. If we denote this set by Ω we can identify the points of Ω with $\{(i, \omega, \Delta)\}$ by the correspondence $(\theta_1, \theta_2, \cdots, \theta_n) \leftrightarrow (i, \omega, \Delta)$ if $\theta_j = \omega$ for some real ω for indices j with the exception of i where $\theta_i = \omega + \Delta$, $\Delta \geqq 0$.

Let F denote a distribution function on Ω. Then define ξ_0 as the probability (under F) that $\Delta = 0$. Define ξ_i as the conditional probability when the exceptional index is i given $\Delta > 0$. Let $F_0(\omega)$ be the conditional distribution of ω given $\Delta = 0$, and $F_i(\omega, \Delta)$ the conditional distribution of (ω, Δ) given the exceptional index i where $\Delta > 0$. Note that F is symmetric if and only if $\xi_1 = \xi_2 = \cdots = \xi_n$, $F_1 = F_2 = \cdots = F_n$.

We now proceed to examine the Bayes procedures. In order to facilitate the study we state some lemmas concerning the loss functions on Ω.

LEMMA 3.2: *If* $i \neq j$, $i \neq k$, $j, k > 0$, *then* $L_j(i, \omega, \Delta) - L_k(i, \omega, \Delta) = 0$.[2] (*This includes the case* $\Delta = 0$.)

PROOF: Consider any permutation $(1, 2, \cdots, n) \to (\pi 1, \pi 2, \cdots, \pi n)$, such that $\pi i = i$, $\pi j = k$, $\pi k = j$. Then

$$L_j(i, \omega, \Delta) - L_k(i, \omega, \Delta) = L_{\pi j}(\pi i, \omega, \Delta) - L_{\pi k}(\pi i, \omega, \Delta)$$
$$= L_k(i, \omega, \Delta) - L_j(i, \omega, \Delta)$$

LEMMA 3.3: *If* $i \neq j$, $k \neq l$, *then* $L_i(j, \omega, \Delta) = L_k(l, \omega, \Delta)$.

PROOF: Consider any permutation $(1, 2, \cdots, n) \to (\pi 1, \pi 2, \cdots, \pi n)$ such that $\pi i = k$, $\pi j = l$. Then $L_i(j, \omega, \Delta) = L_{\pi i}(\pi j, \omega, \Delta) = L_k(l, \omega, \Delta)$.

If we let $L_i(\omega)$ denote the loss function when action i is taken and $\Delta = 0$, and let $p_j(x; \omega, \Delta)$ denote the density when the parameter is (j, ω, Δ), then the computation of the Bayes solution reduces to consideration of the expressions

$$\int \sum_{k=0}^{n} [L_i(k, \omega, \Delta) - L_j(k, \omega, \Delta)] p_k(x; \omega, \Delta)\, dF(k, \omega, \Delta)$$
$$= \xi_0 \int [L_i(\omega) - L_j(\omega)] p_0(x; \omega)\, dF_0(\omega)$$
$$+ \sum_{k=1}^{n} \xi_k \int [L_i(k, \omega, \Delta) - L_j(k, \omega, \Delta)] p_k(x; \omega, \Delta)\, dF_k(\omega, \Delta).$$

When we are seeking only symmetric Bayes procedures, we may, by Theorem 3.1, take $\xi_1 = \cdots = \xi_n = \xi$, $F_1 = F_2 = \cdots = F_n = F$.

A detailed study of the Bayes procedures will require an extension of the notion of densities possessing a monotone likelihood ratio. The usual idea of a monotone likelihood ratio is given below.

DEFINITION 3.3: A function $f(x, y)$ defined on R^2 is said to have a monotone likelihood ratio if for any choice $x_1 < x_2$, $y_1 < y_2$, we have $\det \| f(x_i, y_j) \| \geqq 0$.

The fundamental property possessed by these functions is summarized in the following lemma. A proof may be found in [4].

LEMMA 3.4: *If* $h(y)$ *is a function which changes sign at most once from positive to negative values, and* $f(x, y)$ *has a monotone likelihood ratio, then*

$$g(x) = \int h(y) f(x, y)\, d\mu(y)$$

changes sign at most once in the same direction as h. *Here,* μ *denotes any positive* σ*-finite measure.*

We now generalize the concept of a monotone likelihood ratio for joint densities of n variables that depend on n parameters. (See also Pratt [9].)

DEFINITION 3.4: Let A and B be arbitrary sets. For each $\alpha \varepsilon A$, $\beta \varepsilon B$, let $x_\alpha(t) = (x_\alpha^{(1)}(t), x_\alpha^{(2)}(t), \cdots, x_\alpha^{(n)}(t))$ and $\theta_\beta(s) = (\theta_\beta^{(1)}(s), \theta_\beta^{(2)}(s), \cdots,$

[2] The notation $L_j(i, \omega, \Delta)$ means that we are evaluating L_j at the parameter points in the set $\{i, \omega, \Delta\}$.

$\theta_\beta^{(n)}(s))$ be curves in R^n. A density $f(x_1, x_2, \cdots, x_n; \theta_1, \theta_2, \cdots, \theta_n)$ is said to have a monotone likelihood ratio with respect to the family of pairs of curves $\{(x_\alpha(t), \theta_\beta(s)); \alpha \varepsilon A, \beta \varepsilon B\}$ if for every $\alpha \varepsilon A$, $\beta \varepsilon B$, $f(x_\alpha(t); \theta_\beta(s))$ has a monotone likelihood ratio in the variables t and s.

Most of our discussion of Bayes procedures will be based on the following assumption on the densities.

(A) For $j > 0$, $k > 0$,

$$p_j(x_1, x_2, \cdots, x_n; \omega, \Delta) \geqq p_k(x_1, x_2, \cdots, x_n; \omega, \Delta)$$

if and only if $x_j \geqq x_k$, or strict inequality in both instances.

The following theorem provides one condition which implies the validity of assumption (A).

THEOREM 3.2: *Let $\gamma(s)$ and $\delta(t)$ be continuous strictly increasing functions defined for all real s and t, $\gamma(0) = \delta(0) = 0$, and such that $\gamma(s)$ and $\delta(t)$ range from $-\infty$ to $+\infty$. Define a family of pairs of curves with the following properties: the curves $x(t)$, $\theta(s)$ belong to the family if and only if for some j, k and real numbers a and b,*

$$x_j(t) = a + \delta(t), \qquad x_k(t) = a + \delta(-t),$$

$x_i(t)$ constant for $i \neq j$, $i \neq k$;

$$\theta_j(s) = b + \gamma(s), \qquad \theta_k(s) = b + \gamma(-s),$$

$\theta_i(s)$ constant for $i \neq j$, $i \neq k$. If the density $p(x_1, x_2, \cdots, x_n; \theta_1, \theta_2, \cdots, \theta_n)$ has a strict monotone likelihood ratio with respect to this family of pairs of curves, then assumption (A) is satisfied.

REMARK: Our main applications involve the curves $\delta(t) = t$ and $\gamma(s) = s$.

PROOF: Suppose $x_j \geqq x_k$. We can find a pair of curves $(x(t), \theta(s))$ in the family so that for some $t_0 \geqq 0$, $x = x(t_0)$ where $x_j = a + \delta(t_0)$, $x_k = a + \delta(-t_0)$ and such that for some $s_0 > 0$, $\theta(s_0)$ corresponds to the density $p_j(x; \omega, \Delta)$ and $\theta(-s_0)$ corresponds to the density $p_k(x; \omega, \Delta)$. Then $[p(x(t); \theta(s_0))]/[p(x(t); \theta(-s_0))]$ is a monotone strictly increasing function of t. But at $t = 0$ this ratio is one, since if π is the permutation which interchanges j and k and leaves all others fixed,

$$p(x(0); \theta(-s_0)) = p(x_\pi(0); \theta_\pi(-s_0)) = p(x(0); \theta(s_0)).$$

Hence, the above ratio is $\geqq 1$ if $t \geqq 0$, in particular for $t = t_0$.

Conversely, if $p_j(x; \omega, \Delta) \geqq p_k(x; \omega, \Delta)$ we can choose the curve $\theta(s)$ described above and an arbitrary curve $x(t)$ such that $x_j(t) = a + \delta(t)$, $x_k(t) = a + \delta(-t)$ for some real number a and arbitrary (but fixed) coordinate x_i when $i \neq j$, $i \neq k$. Then, as before,

$$\frac{p(x(t); \theta(s_0))}{p(x(t); \theta(-s_0))} \geqq 1$$

implies that $t \geqq 0$, which in turn implies $x_j(t) \geqq x_k(t)$. Since this holds for all curves $x(t)$ of this form, the theorem is proved.

THEOREM 3.3: *If $p(x; \theta)$ satisfies assumption (A), then any symmetric Bayes* *procedure for the slippage problem has the form* $\varphi_0(x) = 1$ *if* $x \varepsilon R_0$, *a symmetric set, and* $\varphi_i(x) = 1$ *if* $x \not\varepsilon R_0$ *and* $x_i > x_j$ *for all* $j \neq i$, $i = 1, 2, \cdots, n$. *If* $x \not\varepsilon R_0$ *and* $\max_{1 \leqq i \leqq n} x_j = x_{i_1} = x_{i_2} = \cdots = x_{i_r}$, *then* $\varphi_{i_1}(x) + \varphi_{i_2}(x) + \cdots + \varphi_{i_r}(x) = 1$.

PROOF: If φ is Bayes against an *a priori* symmetric distribution F, then $\varphi_0(x) = 1$ if

$$\xi_0 \int [L_0(\omega) - L_j(\omega)]p_0(x; \omega)\, dF_0(\omega) + \xi \sum_{k=1}^{n} \int [L_0(k, \omega, \Delta) - L_j(k, \omega, \Delta)]\cdot p_k(x; \omega, \Delta)\, dF(\omega, \Delta) < 0$$

for $j = 1, 2, \cdots, n$. The symmetry assumption on p and L clearly show that this set is symmetric. Call this set R_0.

If $i \neq 0$, $\varphi_i(x) = 1$ provided

$$\xi_0 \int [L_i(\omega) - L_j(\omega)]p_0(x; \omega)\, dF_0(\omega) + \xi \sum_{k=1}^{n} \int [L_i(k, \omega, \Delta) - L_j(k, \omega, \Delta)]p_k(x; \omega, \Delta)\, dF(\omega, \Delta) < 0$$

for $j = 0, 1, \cdots, n$, $j \neq i$. If $j = 0$, the above inequality says that $x \not\varepsilon R_0$. If $j > 0$, $j \neq i$, by Lemma 3.2, $L_i(\omega) - L_j(\omega) = 0$, and $L_i(k, \omega, \Delta) - L_j(k, \omega, \Delta) = 0$ if $i \neq k$, $j \neq k$. Thus, the above inequality reduces to

$$\xi \int [L_i(i, \omega, \Delta) - L_j(i, \omega, \Delta)]p_i(x; \omega, \Delta)\, dF(\omega, \Delta) + \xi \int [L_i(j, \omega, \Delta) - L_j(j, \omega, \Delta)]p_j(x; \omega, \Delta)\, dF(\omega, \Delta) < 0.$$

The symmetry of the loss function shows that $L_i(i, \omega, \Delta) - L_j(i, \omega, \Delta) = -[L_i(j, \omega, \Delta) - L_j(j, \omega, \Delta)]$ so that the inequality becomes

$$\xi \int [L_i(i, \omega, \Delta) - L_j(i, \omega, \Delta)][p_i(x; \omega, \Delta) - p_j(x; \omega, \Delta)]\, dF(\omega, \Delta) < 0.$$

Since $L_i(i, \omega, \Delta) - L_j(i, \omega, \Delta) < 0$, assumption (A) says that this inequality holds for $x_i > x_j$. Thus, $\varphi_i(x) = 1$ if $x \not\varepsilon R_0$ and $x_i = \max_{1 \leqq j \leqq n} x_j$ (except possibly $\varphi_i(x) < 1$ on the boundary).

Theorem 3.3 gives the general form of the Bayes procedures, but does not make explicit the nature of the set R_0 (the set where H_0 will be accepted) aside from the fact that R_0 is symmetric. The character of this set depends strongly on the form of the density function. In the next section we will consider several specific densities of importance in their own right, and the slippage problems connected with them. Some of these examples have been studied previously in the literature, while the others are new. All of them are easy examples of our general unified approach.

4. Specific examples. In this section we will examine the form of the Bayes solutions to specific slippage problems. The discussion of these cases are written in detail to exemplify the method of analysis. In most of the situations we will consider two cases which we will label the uncontrolled case and the controlled case. The controlled case refers to the situation where in addition to an observation from each of the n densities, we have an observation from a density whose parameter is known not to have slipped.

For reasons of symmetry, the controlled case will be easier to treat, and we will usually examine that case in detail and merely state the result in the uncontrolled case. Moreover, whenever the problem possesses a natural invariance structure with respect to a group of transformations, we will then automatically restrict ourselves exclusively to those procedures invariant under the induced group of transformations acting on the decision space.

4.1. *The normal density with known variance.*

(a) *The controlled case.* Suppose we have independent observations $X_1, X_2, \cdots, X_n, Y$ with X_i (each X_i and Y usually represent a sufficient statistic, the average sample values, based on several observations) having a normal distribution with mean θ_i and variance 1, and Y having a normal distribution with mean θ and variance 1. In addition to the assumptions made on the loss function in Section 2 we will assume that for every real number c,

$$L_i(\theta_1 + c, \cdots, \theta_n + c, \theta + c) = L_i(\theta_1, \cdots, \theta_n, \theta), \qquad i = 0, 1, \cdots, n.$$

Since the problem is invariant under translation of each observation by a fixed amount it is reasonable to look only at those procedures which depend upon

$$U_1 = X_1 - Y, U_2 = X_2 - Y, \cdots, U_n = X_n - Y.$$

(The statistic U represents a maximal invariant under the translation group.) For a discussion of invariance we refer to ([12], Chaps. 6–7).

The joint density of $U_1, U_2, \cdots, U_n$ is given by

$$p(u_1, \cdots, u_n; \omega_1, \cdots, \omega_n) = C \exp[-\tfrac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda^{ij}(u_i - \omega_i)(u_j - \omega_j)],$$

where C is a constant independent of the $\omega_i(\omega_i = \theta_i - \theta)$, and

$$\lambda^{ij} = \begin{cases} \dfrac{n}{n+1} & \text{if } i = j \\ -\dfrac{1}{n+1} & \text{if } i \neq j. \end{cases}$$

Assumption (A) can be verified by showing that the density has a monotone likelihood ratio along the curves $u(t)$, $\omega(s)$ defined by

$$\begin{aligned} u_1 &= a + t, & u_2 &= a - t, & &\text{all other } u_i \text{ fixed;} \\ \omega_1 &= b + s, & \omega_2 &= b - s, & &\text{all other } \omega_i \text{ fixed.} \end{aligned}$$

In fact,

$$p(u(t); \omega(s)) = f(t)g(s)e^{\kappa st} \qquad \kappa > 0$$

which clearly has a monotone likelihood ratio. Thus, the form of the Bayes solutions are as described in Theorem 3.3. Only the set R_0 must be characterized. In order to characterize R_0 we will make an additional assumption concerning the losses, namely,

$$\text{(B)} \qquad L_0(k, \omega, \Delta) = L_j(k, \omega, \Delta) \quad \text{for all } j \neq k \text{ and } \Delta \neq 0.^3$$

In the present problem the real parameter ω reduces to the single value 0, on account of the translation invariance; and we may write $L_j(k, \omega, \Delta)$ simply as $L_j(k, \Delta)$.

R_0 is an intersection of sets of the form

$$\int \sum_{k=0}^{n} [L_0(k, \Delta) - L_j(k, \Delta)]p(u; k, \Delta)\, dF(k, \Delta) < 0.$$

Using the notation of Section 3 this can be written

$$\xi_0(L_0 - L_j)p_0(u) + \xi \sum_{k=1}^{n} \int_{0+}^{\infty} [L_0(k, \Delta) - L_j(k, \Delta)]p_k(u; \Delta)\, dG(\Delta) < 0.$$

(the first term $L_0 - L_j$ corresponds to the parameter point $\Delta = 0$). Under assumption (B) this reduces to

$$\xi_0(L_0 - L_j)p_0(u) + \xi \int_{0+}^{\infty} [L_0(j, \Delta) - L_j(j, \Delta)]p_j(u; \Delta)\, dG(\Delta) < 0.$$

Using the fact that $p_j(u; \Delta) = p_0(u) \exp[-\frac{1}{2}\lambda^{jj}\Delta^2 + \Delta h_j(u)]$, where

$$h_j(u) = \lambda^{jj}u_j + \sum_{\substack{i \neq j \\ i=1}}^{n} \lambda^{ji}u_i = u_j - \frac{n}{n+1}\bar{u},$$

the inequality becomes

$$p_0(u)\, \xi_0(L_0 - L_j) + \xi \int_{0+}^{\infty}$$

$$\cdot [L_0(j, \Delta) - L_j(j, \Delta)] \exp[-\tfrac{1}{2}\lambda^{jj}\Delta^2 + \Delta h_j(u)]\, dG(\Delta) < 0.$$

We see (by virtue of condition 2′ of Section 2) that the quantity in brackets is a monotone function of $h_j(u)$, so the inequality is equivalent to $h_j(u) < c$, or $u_j - (n/(n + 1))\bar{u} < c$. Thus, every Bayes solution has the form

$$\varphi_0(u) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} \left(u_j - \frac{n}{n+1}\bar{u}\right) < c,$$

$$\varphi_i(u) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} \left(u_j - \frac{n}{n+1}\bar{u}\right) > c \text{ and } u_i > u_j \text{ for all } j \neq i.$$

[3] This assumption is reasonable in many situations and leads to a tractable explicit solution to the problem. Nonetheless our method can be employed in the general case without this assumption, but then the solution is only expressible implicitly.

In terms of the original observations, the procedure is of the form

$$\varphi_0(x_1, \cdots, x_n, y) = 1 \text{ if } \max_{1 \leq j \leq n} \left(x_j - \frac{n\bar{x} + y}{n + 1}\right) < c,$$

$$\varphi_i(x_1, \cdots, x_n, y) = 1 \text{ if } \max_{1 \leq j \leq n} \left(x_j - \frac{n\bar{x} + y}{n + 1}\right) > c \text{ and } x_i > x_j \text{ for all } j \neq i.$$

(b) *Uncontrolled case.* (Paulson [8].) This example was first treated by Paulson and now emerges as a special case of our theory. For the uncontrolled problem we assume that, for all real numbers c,

$$L_i(\theta_1 + c, \cdots, \theta_n + c) = L_i(\theta_1, \cdots, \theta_n),$$

and also condition (B) of part (a). Restricting attention to invariant procedures leads to considering decision functions on the variables $U_1 = X_1 - \bar{X}$, $U_2 = X_2 - \bar{X}, \cdots, U_n = X_n - \bar{X}$ where $\bar{X} = (1/n) \sum_{i=1}^n X_i$. The analysis proceeds in a manner similar to case (a) above, and actually factors in simpler terms. Every symmetric Bayes procedure for the slippage problem in terms of the u_i has the form

$$\varphi_0(u_1, \cdots, u_n) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} u_j < c,$$

$$\varphi_i(u_1, \cdots, u_n) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} u_i > c \text{ and } u_i > u_j \text{ for all } j \neq i.$$

In terms of the original observations

$$\varphi_0(x_1, \cdots, x_n) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} (x_j - \bar{x}) < c,$$

$$\varphi_i(x_1, \cdots, x_n) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} (x_j - \bar{x}) > c \text{ and } x_i > x_j \text{ for all } j \neq i.$$

4.2. *The normal density with unknown variance.*

(a) *Controlled case.* We have independent observations $X_{1j}, X_{2j}, \cdots, X_{nj}$, Y_j, $j = 1, 2, \cdots, k_i$, where the X_{ij} are normally distributed with unknown mean θ_i and unknown common variance σ^2, and the Y_j are normally distributed with unknown mean θ and unknown variance σ^2. For convenience of exposition we take $k_i = k$, and for reasons of invariance we assume that, for all real numbers α and real $\beta > 0$, the loss functions satisfy

$$L_i\left(\frac{\theta_1 + \alpha}{\beta}, \cdots, \frac{\theta_n + \alpha}{\beta}, \frac{\theta + \alpha}{\beta}, \frac{\sigma}{\beta}\right) = L_i(\theta_1, \cdots, \theta_n, \theta, \sigma).$$

For reasons of invariance it is reasonable to consider only those procedures which depend on

$$U_1 = (\bar{X}_1 - \bar{Y})/S, \cdots, U_n = (\bar{X}_n - \bar{Y})/S,$$

where

$$\bar{X}_i = \sum_{j=1}^k X_{ij}/k, \qquad \bar{Y} = \sum_{j=1}^k Y_j/k, \qquad S^2 = \sum_{i=1}^n \sum_{j=1}^k (X_{ij} - \bar{X}_i)^2 + \sum_{j=1}^k (Y_j - \bar{Y})^2.$$

The U_i constitute a maximal invariant with respect to the affine group of transformations of the real line into itself, under which the problem is invariant. The joint density of $U_1, U_2, \cdots, U_n$ can be written as

$$p(u; \eta) = c \int_0^\infty \exp\left[-\tfrac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \lambda^{ij}(u_i s - \eta_i)(u_j s - \eta_j) - s^2/2 \right] \cdot s^{(n+1)(k-1)-1} \, ds,$$

where

$$\lambda^{ij} = \begin{cases} \dfrac{nk}{(n+1)} & \text{if } i = j \\ -\dfrac{k}{(n+1)} & \text{if } i \neq j \end{cases} \qquad \text{and} \quad \eta_i = (\theta_i - \theta)/\sigma.$$

For this problem let us verify assumption (A) directly. If $p_i(u; \delta)$ denotes the density when H_i is true and $\delta = \Delta/\sigma$, we may write

$$p_i(u; \delta) = C \cdot f(\delta) g(u) \int_0^\infty \exp\,[t\delta Z_i - \tfrac{1}{2}t^2] t^{(n+1)(k-1)-1} \, dt,$$

where

$$f(\delta) = \exp\left[-\frac{nk}{n+1} \delta^2 \right], \qquad g(u) = 1 \Big/ \left[\left(\sum_{i=1}^n \sum_{j=1}^n \lambda^{ij} u_i u_j + 1 \right)^{\frac{(n+k)(k-1)}{2}} \right]$$

and

$$Z_i = \left(\lambda^{ij} u_i + \sum_{\substack{j \neq i \\ j=1}}^n \lambda^{ji} u_j \right) \Big/ \left(\sum_{i=1}^n \sum_{j=1}^n \lambda^{ij} u_i u_j + 1 \right)^{\frac{1}{2}}.$$

Thus,

$$p_i(u; \delta) - p_j(u; \delta) = C \cdot f(\delta) g(u) \int_0^\infty (e^{t\delta Z_i} - e^{t\delta Z_j}) e^{-\frac{1}{2}t^2} t^{(n+1)(k-1)-1} \, dt$$

so that $p_i(u; \delta) - p_j(u; \delta) \geqq 0$ if and only if $Z_i \geqq Z_j$ and it is easy to establish that the latter is equivalent to $u_i \geqq u_j$. Hence, assumption (A) holds and the form of the Bayes solutions is determined except for the set R_0.

To represent the set R_0, we again postulate that (B) is valid for the problem in terms of the U_i. Then, R_0 is an intersection of sets of the form

$$\xi_0(L_0 - L_j) p_0(u) + \xi \int_{0+}^\infty [L_0(j, \delta) - L_j(j, \delta)] p_j(u; \delta) \, dF(\delta) < 0.$$

Inserting the explicit expressions of p_j the inequality becomes

$$g(u) \int_{0+}^\infty \int_\infty^\infty \{ \xi [L_0(j, \delta) - L_j(j, \delta)] f(\delta) e^{t\delta Z_i} - \xi_0 (L_j - L_0) \} \cdot e^{-\frac{1}{2}t^2} t^{(n+1)(k-1)-1} \, dt \, dF(\delta) < 0.$$

Since the integral (by virtue of condition 2′ of Section 2) is a strictly increasing function of Z_j, the inequality is equivalent to $Z_j < c$,

$$\frac{u_j - \dfrac{n}{n+1}\bar{u}}{\left\{1 + k\left[\sum_{k=1}^{n} u_i^2 - \dfrac{(\sum u_i)^2}{n+1}\right]\right\}^{\frac{1}{2}}} < c.$$

Thus, every Bayes procedure for the problem in the variables u_i has the form

$$\varphi_0(u) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} \frac{u_j - \dfrac{n}{n+1}\bar{u}}{\left[1 + k\sum_1^n u_i^2 - k\dfrac{(\sum u_i)^2}{n+1}\right]^{\frac{1}{2}}} < c,$$

$$\varphi_i(u) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} \frac{u_j - \dfrac{n}{n+1}\bar{u}}{\left[1 + k\sum_1^n u_i^2 - k\dfrac{(\sum u_i)^2}{n+1}\right]^{\frac{1}{2}}} > c$$

$$\text{and} \quad u_i > u_j \quad \text{for all} \quad j \neq i.$$

In terms of the original observations a slight calculation shows that $\varphi_0(X_{ij}, Y_j) = 1$ if

$$\max_{1 \leq j \leq n} \frac{\bar{X}_j - \dfrac{\sum_{i=1}^n \bar{X}_i + \bar{Y}}{n+1}}{\left[S^2 + k\sum\left(\bar{X}_i - \dfrac{n\bar{X} + \bar{Y}}{n+1}\right)^2 + k\left(\bar{Y} - \dfrac{n\bar{X} + \bar{Y}}{n+1}\right)^2\right]^{\frac{1}{2}}} < c$$

$\varphi_i(X_{ij}, Y_j) = 1$ if the above max is $>c$ and $\bar{X}_i > \bar{X}_j$ for all $j \neq i$.

(b) *Uncontrolled case* (compare with Paulson [8]). Here we have independent observations X_{ij}, $i = 1, 2, \cdots, n$; $j = 1, 2, \cdots, k$, where X_{ij} is normally distributed with unknown mean θ_i and unknown common variance σ^2. Again, we assume that the losses are invariant if a constant is added to each θ_i, and if each θ_i and σ are multiplied by the same positive constant. Then, invariance considerations tell us to look at procedures based on

$$U_1 = (\bar{X}_1 - \bar{X})/S, \cdots, U_n = (\bar{X}_n - \bar{X})/S$$

where $S^2 = \sum_{i=1}^n \sum_{j=1}^k (X_{ij} - \bar{X}_i)^2$. An analogous and simpler analysis as in (a) above shows that the symmetric Bayes procedures for the problem in terms of the variables U_i are all of the form

$$\varphi_0(u) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} u_j \Big/ \left(k\sum_{i=1}^n u_i^2 + 1\right)^{\frac{1}{2}} < c,$$

$$\varphi_i(u) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} u_j \Big/ \left(k\sum_{i=1}^n u_i^2 + 1\right)^{\frac{1}{2}} > c \quad \text{and} \quad u_i > u_j \quad \text{for} \quad j \neq i.$$

In terms of the original variables this becomes

$$\varphi_0(x) = 1 \quad \text{if} \quad \max_{1 \leqq j \leqq n} (\bar{X}_j - \bar{X})/[\sum \sum (X_{ij} - \bar{X})^2]^{\frac{1}{2}} < c,$$

$$\varphi_i(x) = 1 \quad \text{if} \quad \max_{1 \leqq j \leqq n} (\bar{X}_j - \bar{X})/[\sum \sum (X_{ij} - \bar{X})^2]^{\frac{1}{2}} < c,$$

$$\text{and} \quad \bar{X}_i > \bar{X}_j \quad \text{for all} \quad j \neq i.$$

4.3. *The gamma distributions.*

(a) *Controlled case.* Suppose we obtain observations $X_1, X_2, \cdots, X_n, Y$ such that the X_i and Y are independent and X_i has density

$$\frac{1}{\Gamma(p)} \frac{1}{\theta_i^p} e^{-x/\theta_i} x^{p-1}$$

and Y has density

$$\frac{1}{\Gamma(p)} \frac{1}{\theta^p} e^{-y/\theta} y^{p-1}.$$

(Here p is a fixed positive parameter.) We assume that the losses under scale transformations satisfy the invariance condition

$$L_i(\alpha\theta_1, \cdots, \alpha\theta_n, \alpha\theta) = L_i(\theta_1, \cdots, \theta_n; \theta)$$

for all $\alpha > 0$. In addition, we will assume that condition (B) holds. We see clearly that the problem remains invariant under the transformation which multiplies each observation by the same positive real number α, and as usual we will consider only invariant procedures. That is, procedures depending only on

$$U_1 = X_1/Y, \cdots, U_n = X_n/Y.$$

The joint density of $U_1, \cdots, U_n$ is

$$p(u; \omega) = \left(C \prod_{i=1}^{n} \frac{1}{\omega_i^p} u_i^{p-1}\right) \Big/ \left[\left(1 + \sum_{i=1}^{n} \frac{u_i}{\omega_i}\right)^{(n+1)p}\right],$$

where $\omega_i = \theta_i/\theta$. Assumption (A) may be checked by showing that $p(u; \omega)$ has a monotone likelihood ratio with respect to the curves

$$u_i = a + t, \quad u_j = a - t \quad (0 \leqq t \leqq a), \quad \text{while all other coordinates stay fixed;}$$

$$\omega_i = b + s, \quad \omega_j = b - s \quad (0 \leqq s \leqq b), \quad \text{while all other coordinates stay fixed.}$$

This is readily established by direct calculation. To characterize the set R^0 we take the intersection of the sets defined by the inequalities

$$\xi_0(L_0 - L_j)p_0(u) + \int_{0+}^{\infty} [L_0(j, \delta) - L_j(j, \delta)]p_j(u; \delta)\, dF(\delta) < 0$$

where $\delta = \Delta/\theta$. ($L_0 - L_j$ denotes the difference of the loss functions, when taking action 0 and j where the true hypothesis is zero.)

This inequality can be written as

$$p_0(u) \int_{0+}^{\infty} \left\{ [L_0(j, \delta) - L_j(j, \delta)] \left\{ \frac{1}{(1+\delta)^p} \Big/ \left[1 - \frac{\delta}{\delta + 1} \frac{u_j}{1 + \sum_{i=1}^n u_i} \right]^{(n+1)p} \right\} - \xi_0(L_j - L_0) \right\} dF(\delta) < 0.$$

The integral expression is clearly a strictly increasing function of $u_j/(1 + \sum_{i=1}^n u_i)$ and hence the inequality is equivalent to

$$u_j \Big/ \left(\sum_{i=1}^n u_i + 1 \right) < c \quad \text{or} \quad x_j \Big/ \left(\sum_{j=1}^n x_j + y \right) < c.$$

Thus, the symmetric Bayes solution has the form

$$\varphi_0(x, y) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} x_j \Big/ \left(\sum_{j=1}^n x_j + y \right) < c,$$

$$\varphi_i(x, y) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} x_j \Big/ \left(\sum_{j=1}^n x_j + y \right) > c \quad \text{and} \quad x_i > x_j \quad \text{for all} \quad j \neq i.$$

(b) *Uncontrolled case.* A special case of this example was treated in [11]. The corresponding symmetric Bayes solutions in the uncontrolled case have the form

$$\varphi_0(x) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} x_j / \textstyle\sum_{i=1}^n x_i < c,$$

$$\varphi_i(x) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} x_j / \textstyle\sum_{i=1}^n x_i > c \quad \text{and} \quad x_i > x_j \quad \text{for all} \quad j \neq i.$$

5. Translation parameter slippage problem.

5.1. *The general form of the invariant Bayes solutions.* Assume that $X_1, X_2, \cdots, X_n$ are independently distributed according to the densities $p(x - \theta_1), \cdots, p(x - \theta_n)$ respectively, and suppose that Y is independent of $(X_1, \cdots, X_n)$ and has density $p(y - \theta)$. Here, the variable x and the parameter θ traverse the real line. It is possible to develop a corresponding theory in the case where x and θ are integer valued. However, for the sake of exposition, we have limited our discussion to the case of continuous variables. The densities differ only in their location parameter.

The assumptions on the losses are the same as in the preceding problems which dealt with the controlled case. In addition the losses will be assumed to satisfy the invariance property

$$L_i(\theta_1 + c, \cdots, \theta_n + c, \theta + c) = L_i(\theta_1, \cdots, \theta_n, \theta)$$

for all real numbers c. A maximal invariant for the problem is then $U_1 = X_1 - Y, \cdots, U_n = X_n - Y$.

We will assume throughout this section that the density $p(x - \theta)$ has a monotone likelihood ratio (abbreviated M.L.R.). i.e.,

$$p(x_1 - \theta_1)p(x_2 - \theta_2) \geqq p(x_1 - \theta_2)p(x_2 - \theta_1)$$

whenever

$$x_1 < x_2 \quad \text{and} \quad \theta_1 < \theta_2 .$$

The class of distributions which possess a M.L.R. with respect to a translation parameter include all P.F.F.'s [4], any non-central χ^2, any non-central t, etc. Most distributions arising in statistical practice are of this kind. For convenience of exposition we shall assume henceforth that $p(x)$ is strictly positive. All our discussion will remain valid if we merely take $p(x)$ non-negative and positive on some interval. This involves a tedious consideration of cases with no essential new ideas.

The joint density of $U_1, \cdots, U_n$ is

$$q(u; \omega) = \int_{-\infty}^{\infty} \prod_{i=1}^{n} p(u_i - \omega_i + t)p(t)\, dt,$$

where $\omega_i = \theta_i - \theta$. To check assumption (A) we note that

$$q_j(u; \Delta) - q_k(u; \Delta) = \int_{-\infty}^{\infty} \prod_{i=1}^{n} p(u_i + t)$$
$$\cdot[(p(u_j - \Delta + t)/p(u_j + t)) - (p(u_k - \Delta + t)/p(u_k + t))]p(t)\, dt.$$

Since $p(u - \theta)$ has a monotone likelihood ratio, then for every t, the quantity in brackets is greater than or equal to zero, if and only if $u_j \geqq u_k$. Thus, assumption (A) holds, and we know the form of the Bayes solutions except for the set R_0. The set R_0 is an intersection of the sets of u values satisfying the inequalities

$$\xi_0(L_0 - L_j)q_0(u) + \xi \int_{0+}^{\infty} [L_0(j, \Delta) - L_j(j, \Delta)]q_j(u; \Delta)\, dF(\Delta) < 0,$$

or

$$\xi_0(L_0 - L_j) \int_{-\infty}^{\infty} \prod_{i=1}^{n} p(u_i + t)p(t)\, dt + \xi \int_{0+}^{\infty} [L_0(j, \Delta) - L_j(j, \Delta)]$$
$$\cdot \int_{-\infty}^{\infty} \prod_{i=1}^{n} p(u_i + t)[(p(u_j - \Delta + t)/p(u_j + t))]p(t)\, dt\, dF(\Delta) < 0.$$

If we interchange the order of integration and set

$$\Phi(u) = \int_{0+}^{\infty} \{\xi[L_0(j, \Delta) - L_j(j, \Delta)][p(u - \Delta)/p(u)] - \xi_0(L_j - L_0)\}\, dF(\Delta)$$

we may write the inequality as

$$\int_{-\infty}^{\infty} \Phi(u_j + t) \prod_{i=1}^{n} p(u_i + t)p(t)\, dt < 0,$$

where $\Phi(u)$ is monotone increasing. Now consider the curve $u_i = \lambda$, $i = 1, 2, \cdots, n$. For u on this curve the inequality becomes

$$(1) \quad \int_{-\infty}^{\infty} \Phi(\lambda + t) \left\{ \prod_{i=1}^{n} p(\lambda + t) \right\} p(t)\, dt = \int_{-\infty}^{\infty} \Phi(u) \left\{ \prod_{i=1}^{n} p(u) \right\} p(u - \lambda)\, du < 0.$$

Now, $\Phi(u)[p(u)]^n$ changes sign at most once, and $p(u - \lambda)$ has a monotone likelihood ratio. Hence, by Lemma 3.4, $\int_{-\infty}^{\infty} \Phi(u)p^n(u)p(u - \lambda)\, du$ changes sign at most once in λ (from negative to positive values). Let λ_0 be the value at which it changes sign. If there is an interval of λ values where the integral is zero, then define λ_0 to be the smallest value of this set. Let $\lambda_0 = \pm\infty$ respectively if the integral is always negative or always positive.

THEOREM 5.1:

$$R_0 \subset \{\mathbf{u} \mid \max_{1 \leq i \leq n} u_i < \lambda_0\},$$

where λ_0 is defined in the preceding paragraph.

PROOF:

CASE I: $\lambda_0 = +\infty$. The theorem is trivially true in this case.

CASE II: $\lambda_0 = -\infty$. This means that $\int_{-\infty}^{\infty} \Phi(u)p^n(u)p(u - \lambda)\, du$ is positive for all real λ. We must show that R_0 is the empty set. Suppose that $\mathbf{u} \in R_0$ and that u_1 is the maximum coordinate of $\mathbf{u}$. Now,

$$\int_{-\infty}^{\infty} \Phi(u_1 + t) \prod_{i=1}^{n} p(u_i + t)p(t)\, dt = \int_{-\infty}^{\infty} p(z)\Phi(z) \prod_{i=2}^{n} p(u_i - u_1 + z)p(z - u_1)\, dz$$

changes sign at most once in u_1 when the other variables are held fixed, since $\prod_{i=2}^{n} p(u_i + z - u_1)p(z - u_1)$ has a M.L.R. in the variables z and u_1. Hence, there exists a vector point $\mathbf{u}'$, having the same components as $\mathbf{u}$, except for the first component, where u_1' is determined so that $u_1' = \max_{2 \leq j \leq n} u_j$, and belongs to R_0. This is the case, since the above integral is negative for the point $\mathbf{u}$, and remains negative when the first component u_1 is decreased. Continuing in this fashion we may show that a point, having all coordinates equal, belongs to R_0 which contradicts the assumption that $\lambda_0 = -\infty$.

CASE III: λ_0 is finite.

Consider a point $\mathbf{u} \in \{\mathbf{u} \mid \max_{1 \leq j \leq n} u_i < \lambda_0\}$. For definiteness suppose $u_1 = \max_{1 \leq i \leq n} u_i$. Consider a curve $\Gamma(u_j = u_j(s), j = 1, \cdots, n)$, passing through the points $(\lambda_0, \lambda_0, \cdots, \lambda_0)$ and $\mathbf{u}$, with the property $u_i(s) - u_1(s)$ is decreasing for $i = 2, 3, \cdots, n$, and $u_1(s)$ is increasing. Along this curve it is easily verified that the density $\prod_{i=1}^{n} p[z - (u_1(s) - u_i(s))]p(z - u_1(s))$ has a monotone likelihood ratio in the variables z and s. Then

$$\int \Phi[u_1(s) + t] \prod_{i=1}^{n} p[u_i(s) + t]p(t)\, dt = \int \Phi(z) \prod_{i=1}^{n} p[z - (v_1(s) - u_i(s)]p[z - u_1(s)]\, dz$$

changes sign at most once. But since it changes at $(\lambda_0, \lambda_0, \cdots, \lambda_0)$ the point $\mathbf{u} \not\in R_0$.

In general, we cannot make more explicit the nature of the set R_0. However when there exists a one-dimensional sufficient statistic, a more precise characterization of R_0 is possible. This is done in the following paragraph.

5.2. *Form of the Bayes solutions when there is a sufficient statistic.* In this section, by way of variation, we shall discuss the uncontrolled problem and state without proof the corresponding conclusions in the case of the controlled problem.

We suppose that $p(x)$ is bounded, and we may assume for convenience that p has its maximum at $x = 0$. For, if it has a maximum at $x = x_0$ we can relabel the parameters so that $\theta' = \theta + x_0$. In addition we will suppose that there is a statistic $T = T(x_1, \cdots, x_n)$ which is sufficient for θ when $\theta_1 = \theta_2 = \cdots = \theta_n = \theta$. That is, the likelihood function can be written

$$\prod_{i=1}^{n} p(x_i - \theta) = r(x)q(T; \theta).$$

Since the maximum likelihood estimate (M.L.E.), $\hat{\theta}$, is *a fortiori* sufficient we can take $T = \hat{\theta}$. The following lemma is now immediate.

LEMMA 5.2.1: *The maximum likelihood estimate of* θ, $\hat{\theta}$, *is translation invariant,* i.e., $\hat{\theta}(x_1 + c, \cdots, x_n + c) = \hat{\theta}(x_1, \cdots, x_n) + c$, *and* $\prod_{i=1}^{n} p(x_i - \theta)$ *can be written as* $r(x)q(\hat{\theta} - \theta)$.

LEMMA 5.2.2: $q(\hat{\theta} - \theta)$ *has a monotone likelihood ratio.*

PROOF: We assume for simplicity of exposition that p is positive everywhere. The general situation can be handled by a tedious enumeration of cases. Let $\theta_1 > \theta_2$, $\hat{\theta}_2 = \hat{\theta}(x_1, \cdots, x_n)$, $\hat{\theta}_1 = \hat{\theta}(x_1 + \delta, \cdots, x_n + \delta) = \hat{\theta}_2 + \delta$, where $\delta > 0$. Then

$$\begin{aligned}[q(\hat{\theta}_1 - \theta_1)/q(\hat{\theta}_1 - \theta_2)] &- [q(\hat{\theta}_2 - \theta_1)/q(\hat{\theta}_2 - \theta_2)] \\ &= \left[\prod_{i=1}^{n} p(x_i + \delta - \theta_1) \Big/ \prod_{i=1}^{n} p(x_i + \delta - \theta_2)\right] \\ &\quad - \left[\prod_{i=1}^{n} p(x_i - \theta_i) \Big/ \prod_{i=1}^{n} p(x_i - \theta_2)\right] > 0,\end{aligned}$$

since for each i, $p(x_i + \delta - \theta_1)/p(x_i + \delta - \theta_2) > p(x_i - \theta_1)/p(x_i - \theta_2)$ because of the monotonicity of the likelihood ratio for $p(x - \theta)$.

THEOREM 5.2.1: *For the uncontrolled problem the set* R_0 *has the form*

$$R_0 = \{x \mid \max_{1 \leq i \leq n} (x_i - \hat{\theta}) < c\}$$

where $\hat{\theta}$ *is the* M.L.E. *of* θ *under* H_0; $\theta_1 = \cdots = \theta_n = \theta$.

PROOF: The maximal invariant has a density which can be written in the symmetric form $\int_{-\infty}^{\infty} \prod_{i=1}^{n} p(x_i + t)\, dt$, and, as before, the set R_0 can be expressed as an intersection of sets defined by the inequalities

$$\int_{-\infty}^{\infty} \Phi(x_j + t) \prod_{i=1}^{n} p(x_i + t)\, dt < 0 \qquad j = 1, 2, \cdots, n$$

where $\Phi(u)$ is the same as in Paragraph 5.1. By assumption, $\hat{\theta}$ is sufficient for θ so that

$$\int_{-\infty}^{\infty} \Phi(x_j + t) \prod_{i=1}^{n} p(x_i + t)\, dt = \int_{-\infty}^{\infty} r(x)\Phi(x_j + t)q(\hat{\theta} + t)\, dt$$

$$= r(x) \int_{-\infty}^{\infty} \Phi(u)q(u - (x_j - \hat{\theta}))\, du.$$

Since Φ is monotone increasing and $q(u - (x_j - \hat{\theta}))$ has a monotone likelihood ratio, the above is less than zero if $x_j - \hat{\theta} < c$ for some appropriate constant c. Since the above is to hold for $j = 1, \cdots, n$, the set R_0 is determined as

$$\{x \mid \max_{1 \leq j \leq n} (x_j - \hat{\theta}) < c\}.$$

An identical result holds for the controlled problem. The formal arguments are similar.

THEOREM 5.2.2: *If the maximum likelihood estimate θ is sufficient, then the class of procedures of form*

$$\varphi_0(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} (x_i - \hat{\theta}) < c,$$

$$\varphi_j(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} (x_i - \hat{\theta}) > c \text{ and } x_j > x_i \text{ for all } i \neq j,$$

constitute a minimal essentially complete class of symmetric invariant procedures.

PROOF: Let $\Delta_0 = 0, \Delta_1, \Delta_2, \cdots$ be a dense set of points which includes all points of discontinuity of the function $\gamma(\Delta) - L_0(i, \Delta) - L_i(i, \Delta)$ which is clearly independent of i. Consider any symmetric invariant Bayes procedure $\varphi^{(m)}$ which improves on φ at $\Delta_0, \Delta_1, \cdots, \Delta_m$. That is,

$$\rho(\varphi, \Delta_i) - \rho(\varphi^{(m)}, \Delta_i) \geq 0 \qquad \text{for } i = 0, 1, \cdots, m,$$

where

$$\rho(\varphi, 0) - \rho(\varphi^{(m)}, 0) = -\gamma_0 \int (\varphi_0 - \varphi_0^{(m)})p_0(x)\, dx,$$

for $\gamma_0 = L_j(0) - L_0(0)$, and for $\Delta > 0$

$$\rho(\varphi, \Delta) - \rho(\varphi^{(m)}, \Delta) = -\gamma(\Delta) \int (\varphi_i - \varphi_i^{(m)})p_i(x; \Delta)\, dx$$

where $p_i(x; \Delta)$ refers to the density for which the ith parameter has slipped an amount Δ and the right hand side is clearly independent of i by symmetry considerations. Since we have a sequence $\{\varphi^{(m)}\}$ of procedures, which improves on φ in terms of risk, at $\Delta_0, \Delta_1, \cdots, \Delta_m$, and $\varphi^{(m)}$ is determined by a real number c_m, i.e.,

$$\varphi_0^{(m)}(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} x_i - \hat{\theta} < c_m,$$

$$\varphi_j^{(m)}(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} x_i - \hat{\theta} > c_m \text{ and } x_j > x_i \text{ for all } i \neq j$$

we can choose a limiting procedure which is of the same form, and which improves at all points of the enumerable dense set. Since any $\Delta \not\in \{\Delta_m\}$ is a point of continuity of $\gamma(\Delta)$, it will follow that this limiting procedure has risk no larger than the risk of φ at all points Δ. This shows that the class of procedures given above is essentially complete.

To establish minimal completeness we must show that no two procedures in this class can dominate the other. Let φ^1 and φ^2 be two procedures determined by the critical numbers c_1 and c_2. For definiteness, suppose $c_1 > c_2$. Then

$$\rho(\varphi^1, 0) - \rho(\varphi^2, 0) = -\gamma_0 \int (\varphi_0^1 - \varphi_0^2) p_0(x)\, dx$$
$$= -\gamma_0 P\{c_2 < \max_{1 \leq i \leq n} (X_i - \hat{\theta}) < c_1\} < 0,$$

and

$$\rho(\varphi^1, \Delta) - \rho(\varphi^2, \Delta) = -\gamma(\Delta) \int (\varphi_i^1 - \varphi_i^2) p_i(x; \Delta)\, dx$$
$$= \gamma(\Delta) P\{X_i = \max_{1 \leq j \leq n} X_j, \text{ and } c_2 < \max_{1 \leq j \leq n} (X_j - \Delta)$$
$$\text{given that the } i\text{th parameter has slipped by } \Delta\} > 0.$$

We state without proof that the corresponding result to Theorems 5.2.1 and 5.2.2 apply to the symmetric two sided slippage problem. Namely, every symmetric invariant Bayes procedure is characterized as follows:

$$\varphi_0(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} | x_i - \hat{\theta} | < c,$$

and

$$\varphi_j(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} | x_i - \hat{\theta} | > c$$
$$\text{and } | x_j - \hat{\theta} | > | x_i - \hat{\theta} | \text{ for all } i \neq j.$$

We now indicate two illustrations of Theorems 5.2.1 and 5.2.2. We discuss first an example treated earlier by direct methods. Example 2 below offers a new example of our theory. Except for small variations, these are the unique examples of the theory since the only distributions which admit a sufficient statistic under independent observations and for which the parameter occurs in translation form are the distributions of these illustrations.

EXAMPLE 1: *Normal distribution with known variance.*

For both the controlled and uncontrolled problem the maximum likelihood estimate is sufficient. Consequently, by virtue of Theorem 5.2.2 a minimal complete class of procedures in the uncontrolled case consists of all procedures having the form:

$$\varphi_0(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} x_i - \bar{x} < c,$$

$$\varphi_j(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} x_i - \bar{x} > c \text{ and } x_j > x_i \text{ for all } i \neq j.$$

In the controlled case the minimal complete class of procedures were characterized in terms of the statistic $\max_{1 \leq i \leq n} u_i - [n/(n+1)]\bar{u}$ where $u_i = x_i - y$. We observed here as mentioned earlier, that this can be expressed as $\max x_i - \hat{\theta}$, where $\hat{\theta} = (n\bar{x} + y)/(n+1)$ is the maximum likelihood estimate of θ when $\theta_1 = \cdots = \theta_n = \theta$.

EXAMPLE 2: *Exponential distribution.*

Let us take $p(x - \theta) = e^{-(x-\theta)}\psi(x - \theta)$ where

$$\psi(u) = \begin{cases} 0 & \text{if } u < 0 \\ 1 & \text{if } u \geqq 0. \end{cases}$$

For the non-controlled problem $\hat{\theta} = \min_{1 \leq i \leq n} x_i$. The minimal complete class of symmetric invariant procedures consists of all procedures having the form:

$$\varphi_0(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} x_i - \min_{1 \leq i \leq n} x_i < c,$$

and

$$\varphi_j(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} x_i - \min_{1 \leq i \leq n} x_i > c \text{ and } x_j > x_i \text{ for all } i \neq j.$$

For the controlled case

$$\hat{\theta} = \begin{cases} y & \text{if} \quad \min_{1 \leq i \leq n} x_i \geqq y \\ \min_{1 \leq i \leq n} x_i & \text{if} \quad \min_{1 \leq i \leq n} x_i \leqq y. \end{cases}$$

The minimal complete class of symmetric invariant procedures consists of all procedures of the form:

$$\varphi_0(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} x_i - \min\,(\min_{1 \leq i \leq n} x_i, y) < c,$$

$$\varphi_j(x) = 1 \quad \text{if} \quad \max_{1 \leq i \leq n} x_i - \min\,(\min_{1 \leq i \leq n} x_i, y) > c \text{ and } x_j > x_i \text{ for all } i \neq j.$$

6. Scale parameter problem. Here we assume that $X_1, X_2, \cdots, X_n$ are independently distributed according to the densities $(1/\theta_1)p(x/\theta_1), \cdots, (1/\theta_n)p(x/\theta_n)$ respectively. $p(x)$ is defined for $x \geqq 0$ and taken for convenience to be strictly positive. The densities differ only in their scale parameter. In addition to the usual assumptions on the loss functions we will assume

$$L_i(c\theta_1, \cdots, c\theta_n) = L_i(\theta_1, \cdots, \theta_n).$$

We also hypothesize that where $\theta_1 = \cdots = \theta_n = \theta$, the maximum likelihood estimate $\hat{\theta}$ exists and is sufficient for θ. Finally, we assume that $p(x/\theta)$ possesses a M.L.R. in the variables x and θ. As usual, we restrict attention to symmetric procedures invariant under scale transformations where x and θ traverse the positive real line. An argument entirely analogous to that given in Section 5 leads to the following result.

THEOREM 6.1: *Under the above assumptions the class of procedures of the form*

$$\varphi_0(x) = 1 \quad \text{if} \quad \max x_i/\hat{\theta} < c,$$

$$\varphi_j(x) = 1 \quad \text{if} \quad \max x_i/\hat{\theta} > c \text{ and } x_j = \max_{1 \leqq i \leqq n} x_i \text{ and } x_j > x_i \text{ for all } i \neq j$$

constitute a minimal essentially complete class of symmetric invariant procedures.

The Γ-family of distributions provide us an example. In the uncontrolled case,

$$\hat{\theta} = (1/np) \sum_{i=1}^{n} x_i$$

and in the controlled case

$$\hat{\theta} = (1/(n+1)p)(\sum_{i=1}^{n} x_i + y).$$

Thus, for the uncontrolled problem and the controlled problem the minimal complete class of symmetric invariant procedures as characterized in Theorem 6.1 agree with the class of procedures described in Section 4.3.

7. Combined translation and scale parameter slippage problem. Let X_1, $X_2, \cdots, X_n$ be independent, and let X_i have density $(1/\sigma)p((x - \theta_i)/\sigma)$ for $i = 1, 2, \cdots, n$. The density is known except for the location parameter θ_i and the scale parameter $\sigma > 0$. The slippage problem refers to the location parameters θ_i. Again, we make the usual assumptions about the losses with the added restriction that

$$L_i((\theta_1 + b)/a, \cdots, (\theta_n + b)/a, \sigma/a) = L_i(\theta_1, \cdots, \theta_n, \sigma)$$

for all real numbers b and all positive numbers a. A slight extension of the methods of the previous sections enables us to prove

THEOREM 7.1: *Let* $\prod_{i=1}^{n} p((x_i - \theta)/\sigma) = h(x)q(\hat{\sigma}/\sigma)r((\hat{\theta} - \theta)/\sigma)$ *where* $q(\hat{\sigma}/\sigma)$ *has a monotone likelihood ratio, and* $r((\hat{\theta} - \theta)/\sigma)$ *has, for each* σ, *a monotone likelihood ratio. Then the class of procedures of the form*:

$$\varphi_0(x) = 1 \quad \text{if} \quad \max_{1 \leqq i \leqq n} (x_i - \hat{\theta})/\hat{\sigma} < c,$$

$$\varphi_j(x) = 1 \quad \text{if} \quad \max_{1 \leqq i \leqq n} (x_i - \hat{\theta})/\hat{\sigma} > c \text{ and } x_j > x_i \text{ for all } j \neq i$$

is a minimal complete class of invariant symmetric procedures.

Theorem 7.1 may be illustrated by Paulson's result for Normal variates involving slippage of the mean parameter with unknown common variance. Another application of Theorem 7.1 is on the density

$$p(x - \theta)/\lambda) = \begin{cases} e^{-\frac{(x-\theta)}{\lambda}} & \text{for} \quad x \geqq \theta \\ 0 & \text{for} \quad x < \theta \end{cases} \qquad (\lambda > 0 \text{ and } \theta \text{ real})$$

There are discrete analogues of the results of Sections 5–7 valid for the Pascal family of distributions to which our methods apply.

8. Uniformly most powerful procedures. In this section we prove, subject to slight smoothness restrictions, that each symmetric invariant Bayes procedure involving a single critical parameter is uniformly most powerful amongst the class of all symmetric invariant procedures, having a prescribed error associated with hypothesis H_0. The loss functions are assumed to be zero or one, according, as a correct or incorrect decision was made. All relevant distributions in this section are assumed to derive from continuous densities.

We suppose that the problem has been reduced by invariance so that the density $p_j(x; \Delta)$ under the condition that the jth population has slipped depends on a single positive parameter Δ. Moreover, as usual, the permutation π which interchanges j and k and leaves the other indices fixed satisfies

$$p_{\pi j}(x_\pi ; \Delta)\, dx_\pi = p_j(x; \Delta)\, dx.$$

Finally, we suppose, it has been demonstrated that all symmetric invariant Bayes procedures are of the form

$$\varphi_0(x) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} x_j - v(x) < c,$$

(1)

$$\varphi_i(x) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} x_j - v(x) > c \text{ and } x_i > x_j \text{ for all } j \neq i$$

$$i = 1, 2, \cdots, n$$

where $v(x)$ is a symmetric function of x, and c is a constant.

Consider a class of *a priori* symmetric distributions $F_{\Delta_0,\xi}$ depending on a parameter ξ constructed as follows. Let Δ_0 be an arbitrary positive but fixed real number. We define $F_{\Delta_0,\xi}$ as a discrete distribution which assigns mass ξ to $p_0(x)$ and mass $(1 - \xi)/n$ to $p_j(x; \Delta_0), j = 1, \cdots, n$. By our previous theory we know that the Bayes procedure against $F_{\Delta_0,\xi}$ is of the form (1) with an appropriate c depending on ξ. A simple continuity argument shows that when ξ varies between 1 and 0, the constant c varies continuously from its largest possible value to its smallest possible value. In particular, when $\xi = 1, \varphi_0(x) \equiv 1$, and when $\xi = 0, \varphi_0(x) \equiv 0$. For any prescribed c, by continuity, we obtain the existence of ξ_0 such that the given procedure of (1) defined by the constant c is Bayes against F_{Δ_0,ξ_0}. This can be done for any $\Delta_0 > 0$.

Let $\tilde{\varphi}$ denote a decision procedure characterized as in (1) for which the probability of accepting hypothesis zero when it is true has fixed size equal to

$$\int \tilde{\varphi}_0(x) p_0(x)\, dx = 1 - \alpha \qquad (0 < \alpha < 1)$$

Consider any other symmetric invariant procedure φ having the same prescribed error associated with action a_0. Since $\tilde{\varphi}$ is Bayes against F_{Δ_0,ξ_0} for a suitable ξ_0 we have

$$\int \rho(\tilde{\varphi}, \theta)\, dF_{\Delta_0 \xi_0}(\theta) \leq \int \rho(\varphi, \theta)\, dF_{\Delta_0 \xi_0}(\theta)$$

or

$$\xi_0 \int \tilde{\varphi}_0(x)p_0(x)\,dx + \sum_{j=1}^{n} [(1-\xi_0)/n] \int [1-\tilde{\varphi}_j(x)]p_j(x,\Delta_0)\,dx$$
$$\leqq \xi_0 \int \varphi_0(x)p_0(x)\,dx + [(1-\xi_0)/n] \sum_{j=1}^{n} \int [1-\varphi_j(x)]p_j(x,\Delta_0)\,dx.$$

(Remember that the loss due to a wrong decision is 1, independent of the nature of the error.)

Since φ and $\tilde{\varphi}$ are symmetric invariant we infer that $\int [1-\tilde{\varphi}_j(x)]p_j(x, \Delta_0)\,dx$ is independent of j. Since the error in rejecting H_0 is fixed, we obtain

$$\int [1-\tilde{\varphi}_j(x)]p_j(x,\Delta_0)\,dx \leqq \int [1-\varphi_j(x)]p_j(x,\Delta_0)\,dx \tag{2}$$

and this is true for any $\Delta_0 > 0$ by choosing in each case ξ suitably.

Thus, we have proved that, amongst all symmetric invariant procedures possessing a prescribed probability of rejecting hypothesis H_0 when it is true, there is a single member (up to equivalence in terms of risk) in the class (1) which is uniformly most powerful.

Another way to express the conclusion of (2) is as follows: Amongst all symmetric procedures with a prescribed probability of accepting hypothesis zero when it is true, there is a unique decision procedure of type (1) which maximizes the probability of making the correct decision, whatever the true state of nature. That there is only one is clear by virtue of the fact that the related distributions are all continuous densities.

9. A multivariate slippage problem. Let S be a $p \times p$ Wishart matrix with covariance matrix Σ, and let $X_1, X_2, \cdots, X_n$ be independent normal random p-vectors with mean vector θ_i respectively, and covariance matrix Σ. In addition, let S be independent of $(X_1, \cdots, X_n)$.

The loss functions satisfy conditions (1) and (2) of Section 2, where the θ_i are now vectors and Δ is a non-zero vector. Also assumption (B) of 4.1 is assumed to hold. The loss functions will also be required to satisfy the invariance property

$$L_i(C\theta_1 + \alpha, \cdots, C\theta_n + \alpha, C\Sigma C') = L_i(\theta_1, \cdots, \theta_n, \Sigma)$$

for all non-singular $p \times p$ matrices C and all p component vectors α. We restrict attention to symmetric procedures based on X, S which are invariant under the non-singular transformations $X \to CX$ and $S \to CSC'$, and under translations of X_i by the same vector.

It is known that the density of the maximal invariant when the jth mean has slipped can be calculated by integrating, with respect to the Haar measure of the full linear group, the joint density $p(CX, CSC')$ transformed by a general element of the group, viz;

$$f(X,S) \iint \exp -\tfrac{1}{2}\{\operatorname{tr}(CSC'\Sigma^{-1} + CXX'C'\Sigma^{-1} + \Sigma^{-1}\Theta_{(j)}\Theta'_{(j)} - 2\Sigma^{-1}CX\Theta_{(j)})\}(d\Theta\, dC/|C|^{q})$$

where $X = (X_1, \cdots, X_n)$ is the matrix with column vectors as indicated and $\Theta_{(j)} = (\theta, \cdots, \theta, \theta + \Delta, \theta, \cdots, \theta)$ is a matrix whose jth column is the vector $\theta + \Delta$, while the remaining columns are each composed of the vector θ. Here q denotes an appropriate real number. We will let Θ represent the $p \times n$ matrix every column of which is composed of the vector θ. Now completing the square and integrating with respect to Θ, we have $\exp[-\frac{1}{2}\operatorname{tr}\Delta'\Sigma^{-1}\Delta]f(X, S)\cdot$

$$\iint \exp[-\tfrac{1}{2}\operatorname{tr} CSC'\Sigma^{-1} - \tfrac{1}{2}\operatorname{tr} CXX'C'\Sigma^{-1} + \operatorname{tr}\Sigma^{-1}CX_j\Delta'$$

$$- \operatorname{tr}(n/2)\Sigma^{-1}(\theta\theta' - 2(C\bar{X} - (\Delta/n))\theta')](d\theta\, dC/|C|^q) = g(\Delta, \Sigma)f(X, S)$$

$$\cdot \int \exp[-\tfrac{1}{2}\operatorname{tr}\Sigma^{-1}C(S + XX' - n\bar{X}\bar{X}')C' + \operatorname{tr}\Sigma^{-1}C(X_j - \bar{X})\Delta'](dC/|C|^q).$$

Put $W = S + XX' - n\bar{X}\bar{X}'$ and the density is

$$g(\Delta, \Sigma)f(X, S) \int \exp[-\tfrac{1}{2}\operatorname{tr} C'\Sigma^{-1}CW + \operatorname{tr}\Delta'\Sigma^{-1}C(X_j - \bar{X})]\,(dC/|C|^q).$$

Introducing the new variable $D = \Sigma^{-\frac{1}{2}}C$, and defining $\eta' = \Delta'\Sigma^{-\frac{1}{2}}$ we have

$$g(\Delta, \Sigma)f(X, S) \int \exp[-\tfrac{1}{2}\operatorname{tr} DWD' + \operatorname{tr}\eta' D(X_j - \bar{X})]\,(dD/|D|^q).$$

(It should be understood that after each change of variables, the functions g and f may change by a multiplicative factor. However, since the explicit expressions of these functions are of no relevance, we will continue to use the same symbol and no ambiguities will arise.)

Since W is positive definite with probability one, we can reduce the integral further by the change of variable $DW^{\frac{1}{2}} = E$. The density becomes

$$g(\Delta, \Sigma)f(X, S) \int \exp[-\tfrac{1}{2}\operatorname{tr} EE' + \operatorname{tr} EW^{-\frac{1}{2}}(X_j - \bar{X})\eta']\,(dE/|E|^q).$$

Making use of polar decomposition of matrices we write

$$W^{-\frac{1}{2}}(X_j - \bar{X})\eta' = U[\eta(X_j - \bar{X})'W^{-1}(X_j - \bar{X})\eta']^{\frac{1}{2}}$$

where U is orthogonal. Now the change of variable $A = EU$ gives

$$g(\Delta, \Sigma)f(X, S) \int \exp[-\tfrac{1}{2}\operatorname{tr} AA'$$

$$+ \operatorname{tr} A[\eta(X_j - \bar{X})'W^{-1}(X_j - \bar{X})\eta']^{\frac{1}{2}}]\,(dA/|A|^q).$$

Let V be orthogonal and such that $V[\eta(X_j - \bar{X})'W^{-1}(X_j - \bar{X})\eta']^{\frac{1}{2}}V' = \Lambda$ is diagonal. The resulting matrix is clearly of rank one since

$$[\eta(X_j - \bar{X})'W^{-1}(X_j - \bar{X})\eta']^{\frac{1}{2}}$$

is of rank one. A further change of variable of A into $V'AV$ gives

$$g(\Delta, \Sigma)f(X, S) \int \exp[-\tfrac{1}{2}\operatorname{tr} AA' + \operatorname{tr} A\Lambda]\,(dA/|A|^q)$$

or

$$g(\Delta, \Sigma)f(X, S) \int \exp\left[-\tfrac{1}{2}\operatorname{tr} AA' + a_{11}(\eta'\eta)^{\frac{1}{2}}[(X_j - \bar{X})' \cdot W^{-1}(X_j - \bar{X})]^{\frac{1}{2}}\right](dA/|A|^q). \tag{1}$$

Let $Z_j = (X_j - \bar{X})'W^{-1}(X_j - \bar{X})$, and denote the integral by $p(Z_j, \delta)$ where $\delta = \eta'\eta$. We now show that $p(Z_j, \delta)$ has a monotone likelihood ratio, and moreover, is a monotone function of Z_j for each δ, where p denotes the integral expression (1) excluding the multiplying factors.

If we expand the integral in an infinite series and integrate term by term we obtain $p(Z_j, \delta) = \sum_{n=0}^{\infty} C_n\, \delta^{(n/2)}\, Z_j^{(n-2)}$. The even coefficients C_{2n} are nonnegative. We will prove that the odd coefficients are zero.

$$C_{2n+1} = \int \exp\left[-\tfrac{1}{2}\Sigma\Sigma\, a_{ij}^2\right] a_{11}^{2n+1}\,(dA/|A|^q).$$

In fact, the change of variable

$$D = \begin{bmatrix} -1 & & & & \\ & 1 & & 0 & \\ & & 1 & & \\ & & & \ddots & \\ & 0 & & & 1 \end{bmatrix} A,$$

leads to

$$C_{2n+1} = \int \exp\left[-\tfrac{1}{2}\Sigma\Sigma\, d_{ij}^2\right](-\,d_{11}^{2n+1})(dD/|D|^q) = -C_{2n+1},$$

so $C_{2n+1} = 0$.

The fact that $p(Z_j, \delta)$ is monotone in Z_j is now clear. The monotonicity of the likelihood ratio follows from the general result in [4] which states that if $q_1(Z, y)$ has a monotone likelihood ratio and $q_2(y, \delta)$ has a monotone likelihood ratio, then $p(Z_j, \delta) = \int q_1(Z_i, y)\, q_2(y, \delta)\, d\psi(y)$ has a monotone likelihood ratio, where ψ represents any σ-finite measure. In our case

$$q_1(Z_j, y) = e^{y \log Z_j},$$
$$q_2(y, \delta) = e^{y \log \delta}$$

both of which clearly possess monotone likelihood ratios.

Finally, careful examination of the derivation of the distribution of the maximal invariant will show that the normalizing function $g(\Sigma, \Delta)$ is only a function of δ.

The Bayes procedures are now easily characterized as follows: Let $F(\delta)$ denote an *a priori* distribution. $\varphi_0(X, S) = 1$ if for every j

$$\int_{0+}^{\infty} [\xi_0(L_0 - L_j)h_0(X,S) + \xi(L_0(j,\delta) - L_j(j,\delta))h_j(X,S;\delta)]\, dF(\delta) < 0$$

where h_0 denotes the density of the maximal invariant when there is no slippage, and h_j is the density when the jth parameter has slipped. This can be written as

$$f(X,S) \int_{0+}^{\infty} g(\delta)[C_0\, \xi_0(L_0 - L_j) + \xi(L_0(\delta) - L_j(\delta))p_j(Z_j, \delta)]\, dF(\delta) < 0.$$

Since the integral is monotone in Z_j, the inequality is equivalent to $Z_j < c$.
Moreover, since $p_j(Z_j, \delta)$ is strictly increasing in Z_j

$$\int_{0+}^{\infty} g(\delta)\, [p_j(Z_j, \delta) - p_i(Z_i, \delta)]\, dF(\delta) \gtrless 0$$

according as $Z_j \gtrless Z_i$. Thus, applying Theorem 3.3 we conclude that the form of the Bayes solutions is

$$\varphi_0(X, S) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} (X_j - \bar{X})'W^{-1}(X_j - \bar{X}) < c,$$

$$\varphi_i(X, S) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} (X_j - \bar{X})'W^{-1}(X_j - \bar{X}) > c,$$

and

$$(X_i - \bar{X})'W^{-1}(X_i - \bar{X}) > (X_j - \bar{X})'W^{-1}(X_j - \bar{X}) \qquad \text{for all } j \neq i,$$

where $W = S + XX' - n\bar{X}\bar{X}'$.

If there is a control population each symmetric invariant Bayes solution can be determined in an analogous fashion. The explicit solution is: Let $Z_i = X_i - Y$, and let S be a Wishart matrix independent of X_i, Y.

$$\varphi_0(X, Y, S) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} \{Z_j - [n/(n+1)]\bar{Z}\}'W^{-1}\{Z_j - [n/(n+1)]\bar{Z}\} < c,$$

where $W = \{S + ZZ' - [n^2/(n+1)]\, \bar{Z}\bar{Z}'\}$

$$\varphi_i(X, Y, S) = 1 \quad \text{if} \quad \max_{1 \leq j \leq n} \{Z_j - [n/(n+1)]\bar{Z}\}'\, W^{-1}\{Z_j - [n/(n+1)]\bar{Z}\} > c$$

and

$$\{Z_i - [n/(n+1)]\bar{Z}\}'W^{-1}\{Z_i - [n/(n+1)]\bar{Z}\}$$
$$> \{Z_j - [n/(n+1)]\bar{Z}\}'W^{-1}\{Z_j - [n/(n+1)]\bar{Z}\} \quad \text{for all } j \neq i.$$

In most applications, the matrix S represents the sample covariance matrix based on several observations of a normal distribution with covariance matrix Σ.

The case where Σ is known can be handled by similar methods. The solution is to look at $\max_{1 \leq i \leq n} (X_i - \bar{X})'\Sigma^{-1}(X_i - \bar{X})$ in the non-controlled case, and $\max_{1 \leq j \leq n} \{Z_i - [n/(n+1)]\bar{Z}\}\Sigma^{-1}\{Z_i - [n/(n+1)]\bar{Z}\}$ in the case where there is a control and $Z_i = X_i - Y$.

10. Some results for non-parametric problems. Let X_{ij} be independent random variables with X_{ij} having a continuous c.d.f. F_j for $i = 1, \cdots, k$;

$j = 1, 2, \cdots, n$. We will define two notions of slippage and, for each, give a symmetric invariant procedure which has local optimum properties. These notions were discussed by Lehmann [3], and I. R. Savage [10] and the solutions we will give are direct applications of their results.

In general, we will say that the jth distribution has slipped if $F_i = F$ for all $i \neq j$, and $F_j = g(F)$ where $g(x) \gneqq x$ is a continuous distribution function on $[0, 1]$. The problem is invariant under monotone transformations, and hence, any invariant procedure will depend only on the ranks of the observations. First, let us take $g(F) = (1 - \lambda)F + \lambda F^2$. Lehmann has shown that if we let r_{ij} denote the rank of X_{ij} in the combined sample, and R the matrix of ranks, then if the jth distribution has slipped, the probability of R is

$$P_j^{(\lambda)}(R) = \left[1 \bigg/ \binom{nk}{nk - k}\right] E\left\{ \prod_{i=1}^{k} [(1 - \lambda) + 2\lambda U^{(r_{ij})}] \right\},$$

where $U^{(r_{ij})}$ denotes the r_{ij}th order statistic in a sample of nk for a distribution uniform on $[0, 1]$. We want to find regions $C_0, C_1, \cdots, C_n$ in the set of possible ranks so that $P_0(C_j) = 1 - \alpha$, the procedure is symmetric, and $P_j^{(\lambda)}(C_j)$ is maximized for small λ. Since

$$(d/d\lambda)P_j^{(\lambda)}(C_j)\,|_{\lambda=0} = \left\{ \sum_{R \varepsilon C_{lj}} \left[2 \sum_{i=1}^{k} r_{ij} / (nk + 1) - k \right] \right\} \bigg/ \binom{nk}{nk - k}$$

it is clear that the region C_i is of the form $\max_{1 \leq j \leq n} \sum_{l=1}^{k} r_{lj} = \sum_{l=1}^{k} r_{li} > \gamma$ where γ is chosen so that

$$P_0\left[\max_{1 \leq j \leq n} \sum_{l=1}^{k} r_{lj} \leq \gamma \right] = 1 - \alpha.$$

Now, consider slippage as follows. We let $g(F) = F^{1+\lambda}$ where $\lambda > 0$. We order the combined sample and define for each j, $Z_i^{(j)} = 0$ or 1 according as the ith member of the ordered sample is not, or is, from the jth distribution. Let Z denote the set of $Z_i^{(j)}$. The distribution of Z when the jth distribution has slipped is, following Savage [10]

$$P_j^{(\lambda)}(Z) = [(nk - k)!k!(1 + \lambda)^k] \bigg/ \left\{ \prod_{l=1}^{nk} \left(\sum_{i=1}^{l} [Z_i^{(j)} + (1 - Z_i^{(j)})(1 + \lambda)] \right) \right\}.$$

We want to choose regions $C_0, C_1, \cdots, C_n$ in the set of all Z so that $P_0(C_0) = 1 - \alpha$, the procedure is symmetric, and $P_j^{(\lambda)}(C_j)$ is maximized for small λ. Since,

$$(d/d\lambda)P_j^{(\lambda)}(C_j)\,|_{\lambda=0} = \sum_{Z \varepsilon C_j} [(nk - k)!k!/(nk)!] \left\{ k - \sum_{l=1}^{nk} \sum_{i=1}^{l} [(1 - Z_i^{(j)})/l] \right\}$$

the region C_j must be of the form

$$\max_{\beta} \sum_{l=1}^{nk} \sum_{i=1}^{l} (Z_i^{(\beta)}/l) = \sum_{l=1}^{nk} \sum_{i=1}^{l} (Z_i^{(\beta)}/l) > c$$

and C_0 of the form

$$\max_{\beta} \left(\sum_{l=1}^{nk} \sum_{i=1}^{l} (Z_i^{(\beta)}/l) \right) < c$$

where c is determined by the condition that $P_0(C_0) = 1 - \alpha$.

11. Selection of a procedure from a complete class. The complete classes of procedures obtained are always of the form:

$$\varphi_0(x) = 1 \quad \text{if} \quad \max_{1 \leqq i \leqq n} U_i(x) < c,$$

$$\varphi_j(x) = 1 \quad \text{if} \quad \max_{1 \leqq i \leqq n} U_i(x) > c \text{ and } U_j(x) > U_i(x) \text{ for all } i \neq j.$$

The usual way of selecting a procedure from the complete class is to control the probability of one of the errors. One may thus choose a number α, $0 < \alpha < 1$, and ask that $E(\varphi_0 \mid \theta_1 = \cdots = \theta_n) = 1 - \alpha$. Thus, the distribution of $\max_{1 \leqq i \leqq n} U_i(x)$ is needed to determine c. In most of the cases discussed, this distribution is not tabulated (or even worked out). It has been suggested by Paulson that, since the U_i are negatively correlated and have the same distribution, a reasonable approximation to the solution of

$$P[\max_{1 \leqq i \leqq n} U_i < c \mid \theta_1 = \cdots = \theta_n] = 1 - \alpha$$

is provided by the solution of

$$P[U_1(x) > c] = \alpha/n.$$

Another alternative for finding c is to use a multivariate Chebychev inequality proposed by Olkin and Pratt [7]. In this way one can put a lower bound on the probability of deciding there was no slippage when this indeed is the case. The bounds given by Olkin and Pratt can be evaluated explicitly in case the correlation matrix has equal non-diagonal elements.

Finally, the constant c can be approximated by direct sampling.

REFERENCES

[1] BECHHOFER, ROBERT E., "A single sample multiple decision procedure for ranking means of normal populations with known variances," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 16–39.

[2] DOORNBOS, R., AND PRINS, H. J., "Slippage tests for a set of gamma variates," *Indagationes Mathematicae*, Vol. 18 (1956), pp. 329–337.

[3] LEHMANN, ERICH L., "The power of rank tests," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 23–43.

[4] KARLIN, S. AND RUBIN, H., "The theory of decision procedures for distributions with monotone likelihood ratio," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 272–300.

[5] KUDO, A., "On the invariant multiple decision procedures," *Bull. Math. Stat.*, Vol. 6 (1956), pp. 57–69.

[6] MOSTELLER, F., "A k-sample slippage test for an extreme population," *Ann. Math. Stat.*, Vol. 19 (1948), pp. 58–65.

[7] OLKIN, I., AND PRATT, J. W., "A multivariate Tchebycheff inequality," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 226–234.
[8] PAULSON, E., "An optimum solution to the k-sample slippage problem for the normal distribution," *Ann. Math. Stat.*, Vol. 23 (1952), pp. 610–616.
[9] PRATT, J. W., "Some results in the decision theory of one parameter multivariate Polya type distributions," *Stanford Technical Report No.* 37, Oct. 1955.
[10] SAVAGE, I. R., "Contributions to the theory of rank order statistics—the two sample case," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 590–615.
[11] TRUAX, D., "An optimum slippage test for the variances of k-normal distributions," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 669–674.
[12] LEHMANN, E., *Testing Statistical Hypothesis*, John Wiley and Sons, New York, 1959.