THE THEORY OF DECISION PROCEDURES FOR DISTRIBUTIONS WITH MONOTONE LIKELIHOOD RATIO

By Samuel Karlin and Herman Rubin

Stanford University

0. Introduction and Summary. In many statistical decision problems, the observations can be summarized in a single sufficient statistic such that the likelihood ratio for any two distributions in the family under consideration is a monotone function of that statistic. This paper assumes, accordingly, that the statistician's decision is to be based upon a single observation of a random variable X, whose distribution is given by (1) and satisfies the inequality (2) in Section 1. As examples of this family of distributions, we have the exponential family such as the normal, binomial, and Poisson. Other kinds of examples are given in Section 1.

In connection with the ordinary testing problem, Allen [1] showed that for the composite testing problem of the one-sided type for the special case of the exponential family of distributions, an admissible minimax procedure must be of the form: choose action 1 (accept the hypothesis) if $x < x_0$ and choose action 2 (accept the alternative) if $x > x_0$. If $x = x_0$, randomization may be required. Sobel [2] and Chernoff obtained partial results for the same class of distributions when the set of decisions is finite.

This paper unifies, extends, and strengthens these results and treats of a wide variety of stati stical decision problems for which the densities have a monotone likelihood ratio·

In Section 1 the fundamental definition and preliminaries are introduced. In particular, the conditions imposed on the loss functions and the densities are delimited and some simple properties of these quantities are developed. In Section 2 we establish some of the basic lemmas. Noteworthy are Lemmas 1 and 2 which express the variation of sign diminishing properties of the densities which possess a monotone likelihood ratio.

The essential completeness of the set of all monotone strategies (see Section 3 for the definition) in the class of all statistical procedures is demonstrated in Section 3 for the case of a finite number of actions. Section 4 deals with the problem of determining the form of all Bayes strategies for the statistician. The important problem of admissibility is studied in detail in Section 5. In the next section a study of the Bayes strategies for nature is made for the case of two actions. In Section 7 the complete class theory is carried through for the case of an infinite number of actions. This is accomplished by employing an argument involving a limiting procedure from the case of finite actions as treated in Section 3. The eighth section presents an analysis of the nature of the Bayes strategies for the case of an infinite number of actions. The final

Received February 16, 1955.

section entails a brief discussion of the connection of invariance theory and the conditions of monotonicity as are required throughout this paper.

Further extensions of these ideas in a different direction, which involves relaxing the conditions on the loss functions and strengthening the requirements on the densities, can be found in [3].

1. Definitions and preliminaries for the case of a finite number of actions. Let the observed random variable, usually a sufficient statistic, be denoted by x and the unknown state of nature by ω. It is supposed that both variables traverse subsets of the real line, X and Ω, respectively. The set Ω can be taken as an interval, without loss of generality, as will be shown in Lemma 1 below. Let the cumulative distribution, when the true state of nature is described by the parameter value ω, have the form

$$(1) \qquad P(x \mid \omega) = \int_{-\infty}^{x} p(t \mid \omega)\, d\mu(t) \qquad (\mu \text{ is a } \sigma\text{-finite measure on } X),$$

where if $x_1 > x_2$ and $\omega_1 > \omega_2$, then the density function satisfies

$$(2) \qquad p(x_1 \mid \omega_1)p(x_2 \mid \omega_2) - p(x_1 \mid \omega_2)p(x_2 \mid \omega_1) \geqq 0.$$

Without loss of generality, the spectrum of μ is assumed to be all of X.

Any distribution of the form (1) which satisfies (2) will be said to possess a monotone likelihood ratio (M.L.R.). Throughout this paper we shall be concerned only with such distributions. The most noteworthy such class of distributions consists of the exponential family of distributions, e.g.,

$$p(x \mid \omega) = \beta(\omega)e^{x\omega}.$$

Then,

$$p(x_1 \mid \omega_1)p(x_2 \mid \omega_2) - p(x_1 \mid \omega_2)p(x_2 \mid \omega_1) = \beta(\omega_1)\beta(\omega_2)[e^{(x_1-x_2)(\omega_1-\omega_2)} - 1]\cdot e^{x_1\omega_2+x_2\omega_1},$$

which is positive if $x_1 > x_2$ and $\omega_1 > \omega_2$. A more general class of distributions for which (1) and (2) hold is given in [5]. This class includes as special cases the noncentral t and noncentral F densities. Other examples of considerable interest, which occur in many practical situations and possess a M.L.R., are as follows: $d\mu(x) = dx$,

$$p(x \mid \omega) = \begin{cases} nx^{n-1}/\omega^n, & 0 < x < \omega, \quad \omega > 0, \quad n \text{ a fixed positive integer,} \\ 0, & \text{elsewhere,} \end{cases}$$

$$p(x \mid \omega) = \begin{cases} n(n-1)\dfrac{x^{n-2}}{\omega^n}(\omega - x) & 0 < x < \omega, \quad \omega > 0 \quad n \text{ an integer} \geqq 2 \\ 0, & \text{elsewhere,} \end{cases}$$

$$p(x, \omega) = \tfrac{1}{2}e^{-|x-\omega|}, \qquad -\infty < x < \infty, \quad -\infty < \omega < \infty,$$

$$p(x, \omega) = \begin{cases} e^{-(x-\omega)}, & x > \omega, \\ 0, & x \leqq \omega. \end{cases}$$

This last distribution is known as the exponential, or waiting time, distribution and occurs in some models of life-testing experiments.

The first few sections treat the statistical decision problem where the statistician has only a finite number n of actions. Let the loss function corresponding to the ith action be denoted by $L_i(\omega)$ when the true parameter value is ω. The following requirements are imposed upon L_i :

A. The $L_i(i = 1, \cdots, n)$ are defined throughout Ω.

B. The number of changes of sign of $L_i - L_{i+1}$ is at most one for $i = 1, \cdots, n - 1$. (A point ω_0 is called a change point for a function h if in some neighborhood of ω_0,

$$h(\omega)h(\omega^*) \leqq 0,$$

whenever $\omega \leqq \omega_0 \leqq \omega^*$, and for some $\omega_1 \leqq \omega_0 \leqq \omega_1^*$, $h(\omega_1) \neq 0$ and $h(\omega_1^*) \neq 0$ with $\omega_1 \neq \omega_1^*$.) The number $N(h)$ of changes of sign of the function h is the $\sup_m N(h(\omega_i))$, $i = 1, \cdots, m$, where $N(h(\omega_i))$ is the number of changes of sign of the sequence $h(\omega_1), h(\omega_2), \cdots, h(\omega_n)$ with $\omega_i < \omega_{i+1}$ and otherwise arbitrary.

C. Let S_i denote the set of ω in Ω where $L_i(\omega) = \min_j L_j(\omega)$. We assume for each S_i that $S_i < S_j$ for $i < j$, where $S < T$ means that the part of S not in T lies to the left of T.

D. Let each function $L_i - L_{i+1}$ have precisely one change point, which we denote by ω_i. .

Let us define the spectrum of p by

$$\sigma_\omega = \{x \mid p(x \mid \omega) > 0\}.$$

Then we may observe

LEMMA A. *If $x < y < z$ and $x \in \sigma_{\omega_1}$, $y \notin \sigma_{\omega_1}$, and $y \in \sigma_{\omega_2}$, then $\omega_1 < \omega_2$ and $z \notin \sigma_{\omega_1}$ and similarly with ">" for "<".*

PROOF. Note that $p(x \mid \omega_1)p(y \mid \omega_2) > 0 = p(x \mid \omega_2)p(y \mid \omega_1)$, and since $x < y$, we must have $\omega_1 < \omega_2$. Also $0 \leqq p(z \mid \omega_2)p(y \mid \omega_1) - p(y \mid \omega_2)p(z \mid \omega_1)$, but this can happen only if $p(z \mid \omega_1) = 0$. A similar argument holds for the opposite sign.

COROLLARY. *The set σ_ω for any ω is a relative interval in X, i.e., the intersection of an interval and X.*

A direct consequence of Lemma A is

LEMMA B. *If $\omega_1 < \omega_2$, then $\sigma_{\omega_1} < \sigma_{\omega_2}$.*

DEFINITION. Ω is called statistically connected if whenever $\omega < \omega'$ there is a sequence $\omega = \omega_0 < \omega_1 < \cdots < \omega_n = \omega'$ such that

$$P(\sigma_{\omega_{i+1}} \mid \omega_i) > 0 \qquad \text{for } i = 0, 1, \cdots, n - 1.$$

We see that if Ω is not statistically connected, it can be decomposed into statistically connected sets Ω_α. If $X_\alpha = \bigcup_{\omega \in \Omega_\alpha} \sigma_\omega$, then $P(X_\alpha \mid \omega)$ is zero if $\omega \notin \Omega_\alpha$ and one if $\omega \in \Omega_\alpha$. Thus, for statistical purposes, we may as well deal with each Ω_α separately, since from any observation x we can recognize in which X_α it lies and hence in which Ω_α our unknown parameter value lies. It is also clear from this that, without loss of generality, Ω_α can be considered an interval. Throughout the remainder of this paper we assume that Ω is statistically connected.

LEMMA C. *If* $\omega < \omega'$ *and* $P(\sigma_{\omega'} \mid \omega) > 0$, *then there exists a constant* K *such that for all* x *and for all* θ, $\omega \leqq \theta \leqq \omega'$,

$$p(x \mid \theta) \leqq K[p(x \mid \omega) + p(x \mid \omega')].$$

PROOF. Since (by the corollary to Lemma A) σ_ω and $\sigma_{\omega'}$ are relative intervals, the hypothesis yields the existence of an interval I such that $\sigma_\omega \cap \sigma_{\omega'} = I \cap X$. Moreover, as $P(\sigma_\omega \cap \sigma_{\omega'} \mid \omega) > 0$, there exists an x and a $y > x$ (fixed from now on) in $\sigma_\omega \cap \sigma_{\omega'}$ such that $p(x \mid \omega)$, $p(x \mid \omega')$, $p(y \mid \omega)$, and $p(y \mid \omega')$ are all positive. Let $A = (-\infty, x)$, $B = [x, y]$, $C = (y, \infty)$ and let $\omega < \theta < \omega'$. If z is in A, then

$$\text{(a)} \qquad p(z \mid \theta) \leqq \frac{p(z \mid \omega)p(x \mid \theta)}{p(x \mid \omega)} = c(\theta)p(z \mid \omega),$$

and if z in C,

$$\text{(b)} \qquad p(z \mid \theta) \leqq \frac{p(z \mid \omega')p(y \mid \theta)}{p(y \mid \omega')} = d(\theta)p(z \mid \omega').$$

Also if z is in B,

$$\text{(c)} \qquad p(z \mid \theta) \geqq \frac{p(z \mid \omega)p(x \mid \theta)}{p(x \mid \omega)} = c(\theta)p(z \mid \omega)$$

and

$$\text{(d)} \qquad p(z \mid \theta) \geqq \frac{p(z \mid \omega')p(y \mid \theta)}{p(y \mid \omega')} = d(\theta)p(z \mid \omega').$$

We obtain from (c) $1 \geqq P(B \mid \theta) \geqq c(\theta)p(B \mid \omega)$ or $c(\theta) \leqq 1 / P(B \mid \omega)$ for all $\omega \leqq \theta \leqq \omega'$. Similarly, from (d) it can be inferred that $d(\theta) \leqq 1 / P(B \mid \omega')$. Now if z is in B

$$\text{(e)} \qquad \begin{aligned} p(z \mid \theta) &\leqq \frac{p(z \mid w)p(y \mid \theta)}{p(y \mid \omega)} \leqq p(z \mid \omega)\, \frac{p(y \mid \theta)}{p(y \mid \omega')}\, \frac{p(y \mid \omega')}{p(y \mid \omega)} \\ &\leqq d(\theta)\alpha p(z \mid \omega) \leqq \frac{\alpha}{P(B \mid \omega')}\, p(z \mid \omega), \end{aligned}$$

where $\alpha = p(y \mid \omega') / p(y \mid \omega)$. Finally, (a) and (b) become for z in A

$$\text{(f)} \qquad p(z \mid \theta) \leqq \frac{1}{P(B \mid \omega)}\, p(z \mid \omega),$$

and for z in C

$$\text{(g)} \qquad p(z \mid \theta) \leqq \frac{1}{P(B \mid \omega')}\, p(z \mid \omega').$$

The three inequalities (e), (f), and (g) readily imply the result of our lemma.

COROLLARY. If Ω is statistically connected and θ_1, θ_2 are in Ω with $\theta_1 < \theta_2$, then there exist $\omega_1, \cdots, \omega_n$ in Ω and a constant K such that for all z and all $\theta_1 < \theta < \theta_2$

$$p(z \mid \theta) \leqq K \sum_{i=1}^{n} p(z \mid \omega_i).$$

This result will be used in section 6.

2. Fundamental lemmas. The following lemma is a fundamental tool to be used extensively throughout the sequel.

LEMMA 1. *If h changes sign at most once, then for F a measure,*

$$g(x) = \int p(x \mid \omega) h(\omega)\, dF(\omega)$$

changes sign at most once.

PROOF. Let ω_0 denote a change point for h. Suppose for definiteness that $h(\omega) \leqq 0$ for $\omega \leqq \omega_0$ and $h(\omega) \geqq 0$ for $\omega > \omega_0$. Define $h_1(\omega) = h(\omega)$ for $\omega > \omega_0$ and $h_1(\omega) = 0$ for $\omega \leqq \omega_0$ and $h_2(\omega) = h_1(\omega) - h(\omega)$. Let

$$g_i(x) = \int p(x \mid \omega) h_i(\omega)\, dF(\omega).$$

Clearly,

$$g_i(x) \geqq 0.$$

Consider for $x_1 > x_2$

$$\begin{aligned} &g_1(x_1)g_2(x_2) - g_2(x_1)g_1(x_2) \\ &\quad = \int_{-\infty}^{\omega_0} \int_{\omega_0}^{\infty} [p(x_1 \mid \omega)p(x_2 \mid \theta) - p(x_1 \mid \theta)p(x_2 \mid \omega)] h_1(\omega)h_2(\theta)\, dF(\omega)\, dF(\theta) \geqq 0 \end{aligned} \tag{3}$$

on account of (2). As a consequence of (3), we cannot have $g(x_2) > 0$ while $g(x_1) < 0$. Otherwise, $0 \leqq g_1(x_1) < g_2(x_1)$ and $0 \leqq g_2(x_2) < g_1(x_2)$. These last two inequalities lead to an obvious contradiction of (3). Let x_0 be the supremum of the set of all x^* such that $g(x) \leqq 0$ for $x \leqq x^*$ $(-\infty \leqq x_0 \leqq \infty)$. In view of the facts established above, we find that $g(x) \leqq 0$ for $x < x_0$ and $g(x) \geqq 0$ for $x > x_0$. This clearly implies that g changes sign at most once Q.E.D.

REMARK 1. It is useful to note that g changes sign in the same direction as h if it changes sign at all.

From now on, unless stated to the contrary, when a function changes sign, then it will be assumed that the function changes from nonpositive values to nonnegative values as the independent variable increases.

A careful study of the proof involved in Lemma 1 also shows

COROLLARY 1. *If $g(x_0) = 0$, but $g_1(x_0) = g_2(x_0) > 0$, then $g(x) \geqq 0$ for $x \geqq x_0$ and $g(x) \leqq 0$ for $x \leqq x_0$.*

In an analogous manner by defining ϕ_1 and ϕ_2 from ϕ, we obtain

LEMMA 2. *If ϕ changes sign at most once in X, then*

$$\psi(\omega) = \int p(x \mid \omega)\phi(x)\, d\mu(x)$$

changes sign at most once. Moreover, if $\psi(\omega_0) = 0$ while

$$\psi_i(\omega_0) = \int p(x \mid \omega)\phi_i(\omega) d\mu(x) > 0$$

where $\psi(\omega) = \psi_1(\omega) - \psi_2(\omega)$, then $\psi(\omega) \geqq 0$ for $\omega \geqq \omega_0$ and $\psi(\omega) \leqq 0$ for $\omega \leqq \omega_0$.

COROLLARY 2. *Assuming the integrals are well defined, then the functions*

$$g_i(x) = \int p(x \mid \omega)[L_i(\omega) - L_{i+1}(\omega)]\, dF(\omega)$$

with F a measure, change signs at most once. Moreover, if some ω in Ω, where $L_i - L_{i+1} \neq 0$, belongs to the spectrum of F and strict inequality holds in (2), then g_i is zero at most once.

Since $L_i(\omega)$ satisfy assumptions B, C and D, an application of Lemma 1 implies the statement of this corollary.

REMARK 2. If $\sigma_\omega = X$ for all ω in Ω and strict inequality takes place in (2) for x_i in X and ω_i in Ω, then the proof of Lemma 1 shows easily that g can have at most one zero, provided F does not concentrate its full mass in the set of zeros or change points of h. A similar comment can be made concerning Lemma 2.

LEMMA 3. *Let $0 \leqq \phi^* \leqq 1$ and $\int p(x \mid \bar{\omega})\phi^*(x)\, d\mu(x) = c(0 \leqq c \leqq 1)$. There exists an x_0 and $0 \leqq \lambda_0 \leqq 1$ such that if*

$$\phi^0(x) = \begin{cases} 1, & x < x_0, \\ \lambda_0, & x = x_0, \\ 0, & x > x_0, \end{cases} \tag{3a}$$

then

$$\int p(x \mid \omega)[\phi^* - \phi^0]\, d\mu(x) \begin{cases} \geqq 0, & \text{for} \quad \omega \geqq \bar{\omega}, \\ \leqq 0, & \qquad\ \ \omega \leqq \bar{\omega}. \end{cases} \tag{4}$$

If $\sigma_\omega = X$ for all ω, then the monotone strategy of the form (3a) satisfying (4) is unique except for at most one point.

PROOF. *Case* 1: $0 < c \leqq 1$. Let x_0 be a such that for

$$\int_{-\infty}^{x_0+} p(x \mid \bar{\omega})\, d\mu(x) \geqq c \geqq \int_{-\infty}^{x_0-} p(x \mid \bar{\omega})\, d\mu(x).$$

Such a value clearly exists. Define $\lambda_0(0 \leqq \lambda_0 \leqq 1)$ so that

$$\int_{-\infty}^{x_0-} p(x \mid \bar{\omega})\, d\mu(x) + \lambda_0 p(x_0 \mid \bar{\omega})\mu\{x_0\} = c. \tag{4a}$$

(λ_0 represents the amount of randomization necessary at x_0 for equality.) The randomization is only necessary when $\mu\{x_0\} > 0$. In the case when $\mu\{x_0\} = 0$, we always take $\lambda_0 = 0$. Define ϕ^0 in terms of x_0 and λ_0 as in (3a). The number of changes of sign of $\phi = \phi^* - \phi^0$ is at most one. Furthermore, $\phi = \phi_1 - \phi_2$, where $\phi_1 = \phi^*$ for $x \geqq x_0$ and $\phi_1 = 0$ for $x < x_0$, while $\phi_2 = \phi_1 - \phi$ is zero for $x > x_0$. If $\int p(x \mid \bar{\omega})\phi_1(x)\, d\mu(x) > 0$. then by virtue of Lemma 2 the result (4) is confirmed. Consider now the possibility of $\int p(x \mid \bar{\omega})\phi_1(x)\, dx = 0$. It follows from Lemma B that $\sigma_{\bar{\omega}} = X \cap [x_1, x_2]$, where $x_1 < x_0 < x_2$ and $\phi_1(x) = \phi(x) = 0$ for $x_0 \leqq x < x_2$. Since

$$\int p(x \mid \bar{\omega})[\phi^* - \phi^0]\, d\mu(x) = 0,$$

we must have $\phi^*(x) = 1$ for $x_1 \leqq x \leqq x_0$. As $\sigma_\omega < \sigma_{\bar{\omega}}$ for $\omega < \bar{\omega}$ and $\phi^0 \geqq \phi^*$ when $x \leqq x_0$, we conclude that

$$\int p(x \mid \omega)(\phi^* - \phi^0)\, d\mu(x) \leqq 0.$$

On the other hand, for $\omega > \bar{\omega}$, $\sigma_\omega > \sigma_{\bar{\omega}}$. As $\phi^*(x) = 1$ for $x_1 \leqq x \leqq x_0$, we infer that $\phi^*(x) \geqq \phi^0(x)$ for $x \geqq x_1$ and thus

$$\int p(x \mid \omega)(\phi^* - \phi^0)(x)\, d\mu(x) \geqq 0.$$

This completes the proof of Case 1.

Case 2: $c = 0$. Define $\phi^0 \equiv 1$ for $x < x_1$, $\equiv 0$ for $x > x_1$, where $\sigma_{\bar{\omega}} = X \cap [x_1, x_2]$, while $\phi^0(x_1) = 0$ or 1, according as x_1 is in $\sigma_{\bar{\omega}}$ or not. The condition $c = 0$ implies that the set $S = \{x \mid \phi^*(x) > 0\}$ is disjoint from $\sigma_{\bar{\omega}}$. Again, for $\omega < \bar{\omega}$, $\sigma_\omega < \sigma_{\bar{\omega}}$, so $\phi^0 \geqq \phi^*$ for $x \leqq x_2$. Consequently,

$$\int p(x \mid \omega)(\phi^* - \phi^0)\, d\mu(x) \leqq 0 \qquad (\omega \leqq \bar{\omega}).$$

If $\omega > \bar{\omega}$, then $\sigma_\omega > \sigma_{\bar{\omega}}$, and since $\phi^*(x) \geq \phi^0(x)$ whenever $x > x_1$, we conclude that

$$\int p(x \mid \omega)(\phi^* - \phi^0)\, d\mu(x) \geqq 0 \qquad \omega \geqq \bar{\omega}.$$

The proof of the last part of the lemma is obvious.

The greater detail in the above proof is necessitated only by the zeros possible in (2). If the conditions of Remark 2 are satisfied, the conclusion of (4) is immediate by virtue of the construction in (4a) and Lemma 2.

3. Essential completeness. A strategy for the statistician is a collection of functions $\phi = \{\phi_i, i = 1, \cdots, n\}$ depending on the observed variable where $\phi_i(x)$ equals the probability of taking action i when x was observed. Of course, $0 \leqq \phi_i(x) \leqq 1$ and $\sum_{i=1}^n \phi_i(x) = 1$. A strategy ϕ is called monotone if there exists a set of numbers $x_i(x_i \leqq x_{i+1}(i = 0, \cdots, n), x_0 = -\infty$, and $x_{n+1} = +\infty)$ with

$$\phi_i(x) = \begin{cases} 1, & x_i < x < x_{i+1}, \\ 0, & x < x_i \quad \text{or} \quad x > x_{i+1}, \end{cases} \qquad i = 1, \cdots, n.$$

Randomization for a monotone strategy thus can only occur at the boundary values x_i which describe the particular monotone strategy. Note that if $x_i = x_{i+1}$ and $\phi_i(x_i) = 0$, then action i could never be taken according to strategy ϕ. We shall frequently describe a monotone strategy merely by the set of boundary values (x_i).

Let ω_i denote the change point of $L_i - L_{i+1}$. By assumptions C and D we have that $\omega_i \leqq \omega_{i+1}$, $i = 1, 2, \cdots, n - 1$.

LEMMA 4. *For any strategy $\phi = \{\phi_i\}$ there exists a monotone strategy $\phi^0 = \{\phi_i^0\}$ such that*

$$\int p(x \mid \omega) \left(\sum_{j=1}^{i} \phi_j^0(x) - \sum_{j=1}^{i} \phi_j(x) \right) d\mu(x) \qquad \begin{cases} \geqq 0, & \text{for } \omega \leqq \omega_i, \\ \leqq 0, & \text{for } \omega \geqq \omega_i. \end{cases} \tag{5}$$

PROOF. For each i with $\omega = \omega_i$ and $\phi^* = \sum_{j=1}^{i} \phi_j(x)$ by virtue of Lemma 3 there exists a strategy

$$\psi_i(x) = \begin{cases} 1, & x < x_i, \\ \lambda_i, & x = x_i, \\ 0, & x > x_i, \end{cases}$$

such that

$$\int p(x \mid \omega) \left(\psi_i(x) - \sum_{j=1}^{i} \phi_j(x) \right) d\mu(x) \qquad \begin{cases} \geqq 0, \omega \leqq \omega_i, \\ \leqq 0, \omega \geqq \omega_i. \end{cases} \tag{6}$$

As $\omega_{i+1} \geqq \omega_i$, (5) yields

$$\begin{aligned} 0 &\geqq \int p(x \mid \omega_{i+1}) \left(\psi_i(x) - \sum_{j=1}^{i} \phi_j(x) \right) d\mu(x) \\ &\geqq \int p(x \mid \omega_{i+1}) \left(\psi_i(x) - \sum_{j=1}^{i+1} \phi_j(x) \right) d\mu(x). \end{aligned} \tag{7}$$

Since

$$\psi_{i+1}(x) = \begin{cases} 1, & x < x_{i+1}, \\ \lambda_{i+1}, & x = x_{i+1}, \\ 0, & x > x_{i+1}, \end{cases}$$

and on account of (6) and (7), we see that $\psi_{i+1}(x) \geqq \psi_i(x)$ and $x_{i+1} \geqq x_i$. Define $\phi_{i+1}^0(x) = \psi_{i+1}(x) = \psi_i(x)$, $i = 0, 1, \cdots, n - 1$, where we have set $\psi_0(x) \equiv 0$. It is an easy matter to verify that $\phi^0 = \{\phi_i^0\}$ is a monotone strategy characterized by the boundary points x_i and that the inequalities (5) are satisfied.

THEOREM 1. *The collection of all monotone procedures constitutes an essentially complete class of strategies.*[1]

[1] For the exponential case, this theorem was proved in [4].

PROOF. Let $\phi = \{\phi_i\}$ denote an arbitrary strategy, then the risk when nature's choice is ω becomes

$$\begin{aligned}\rho(\omega, \phi) &= \int p(x \mid \omega) \sum_{i=1}^{n} \phi_i(x) L_i(\omega)\, d\mu(x) \\ &= \int p(x \mid \omega) \left\{ \sum_{i=1}^{n-1} \phi_i(x)[L_i(\omega) - L_n(\omega)] + L_n(\omega) \right\} d\mu(x).\end{aligned}$$

Writing $L_i(\omega) - L_n(\omega) = \sum_{j=i}^{n-1} [L_j(\omega) - L_{j+1}(\omega)]$ and interchanging orders of summation, we get

$$(8) \quad \rho(\omega, \phi) = \int p(x \mid \omega) \left\{ \sum_{i=1}^{n-1} (L_i(\omega) - L_{i+1}(\omega)) \sum_{j=1}^{i} \phi_j(x) \right\} + L_n(\omega)\, d\mu(x).$$

We seek to find a monotone strategy $\{\phi^0\}$ so that $\rho(\omega, \phi^0) \leqq \rho(\omega, \phi)$. The difference becomes

$$(9) \quad \begin{aligned} &\rho(\omega, \phi) - \rho(\omega, \phi^0) \\ &\quad = \sum_{i=1}^{n-1} [L_i(\omega) - L_{i+1}(\omega)] \int p(x \mid \omega) \left(\sum_{j=1}^{i} \phi_j(x) - \sum_{j=1}^{i} \phi_j^0(x) \right) d\mu(x).\end{aligned}$$

By Lemma 4, the monotone strategy $\phi_j^0(x)$ is constructed so that

$$\int p(x \mid \omega) \left[\sum_{j=1}^{i} \phi_j(x) - \sum_{j=1}^{i} \phi_j^0(x) \right] d\mu(x) \begin{cases} \geqq 0, & \omega \geqq \omega_i, \\ \leqq 0, & \omega \leqq \omega_i. \end{cases}$$

But, also, by assumption

$$L_i(\omega) - L_{i+1}(\omega) \begin{cases} \geqq 0, & \omega > \omega_i, \\ \leqq 0, & \omega < \omega_i. \end{cases}$$

Consequently, every term in the sum of (9) is nonnegative and hence $\rho(\omega, \phi) \geqq \rho(\omega, \phi^0)$.

REMARK 3. It is important to observe for later use that the monotone strategy ϕ^0 constructed to dominate ϕ depends only on the specified points ω_i and in no other manner on the loss functions L_i.

More precise results can be obtained if the number of actions n equal 2.

THEOREM 2. *Let $\sigma_\omega = X$ for all ω and suppose $L_1 - L_2$ has precisely one change point ω_1 and no other possible zeros, then every nonmonotone procedure $\phi = (\phi_1, 1 - \phi_1)$ is dominated by a unique monotone strategy.*

PROOF. In order for any strategy ϕ^0 to dominate ϕ, then according to (9)

$$(10) \quad 0 \leqq \rho(\omega, \phi) - \rho(\omega, \phi^0) = [L_1(\omega) - L_2(\omega)] \int p(x \mid \omega)(\phi_1 - \phi_1^0)(x)\, d\mu(x).$$

Since both factors must change signs at the same point ω_1, we must have

$$\int p(x \mid \omega)(\phi_1 - \phi_1^0)(x) \begin{cases} \geqq 0, & \omega \geqq \omega_1, \\ \leqq 0, & \omega \leqq \omega_1. \end{cases}$$

Applying Lemma 3 leads immediately to the conclusion.

4. Bayes strategies. In this section we further assume that $L_i - L_j$ $(j > i)$ has precisely one sign change. Also, we suppose that L_i possess enough smoothness properties to ensure the existence of all integrals involving these quantities. Furthermore, assume that $\sigma_\omega = X$ for all ω. Let F denote a distribution function possessing more than one point in its spectrum. Define

$$\lambda_{ij}(x) = \int p(x \mid \omega)[L_i(\omega) - L_j(\omega)]\, dF(\omega) \qquad (j > i).$$

Let us suppose that the relation (2) holds with strict inequality. By virtue of Corollary 2, λ_{ij} changes sign at most once in the direction of negative to positive values and λ_{ij} has at most one zero. Therefore, if F represents a given distribution for nature, then action i is preferred to action j whenever $\lambda_{ij}(x) < 0$, and x was observed, while the reverse situation holds when $\lambda_{ij}(x) > 0$. If $\lambda_{ij}(x) = 0$, then for that observed x the statistician is indifferent in choosing between action i and j. Summarizing, action i is chosen over j provided $x < x_0$ and action j is desired over action i when $x > x_0$ for some appropriate x_0 (x_0 is the change point of λ_{ij}). This analysis is valid provided $j > i$. Therefore, for a given distribution of nature the optimal Bayes strategy requires that if x is observed and action i is favored over some other action $j (j > i)$, then for all larger x the same is true. This leads readily to the result that if action 1 is ever taken, then it must be taken when $x < x_1$ for some suitable x_1. Continuing the same reasoning yields the existence of values x_i such that

$$\phi_i(x) = \begin{cases} 1, & x_{i-1} < x < x_i, \\ 0, & x < x_{i-1} \quad \text{or} \quad x > x_i, \end{cases}$$

$(x_0 = -\infty)$ represents the unique Bayes strategy against F. This optimal monotone procedure is unique with the exception of $n - 1$ possible values $\{x_i, i = 1, \cdots, n - 1\}$ where randomization for the statistician may be allowed. Summing up, we have established

THEOREM 3. *If F is an a priori distribution for nature which does not concentrate its full mass at a value ω, where $L_i(\omega) = L_j(\omega)$, for some $i < j$, and $L_i - L_j$ changes sign precisely once, then the Bayes strategy is monotone and uniquely determined except for at most n values of the variable x.*

(It was assumed here that $\sigma_\omega = X$, and that strict inequality holds in (2)).

Examples can be given of Bayes strategies in which some actions are never taken. Let $L_1(\omega) = \omega - \theta_1$ for $\omega > \theta_1$, 0 elsewhere; $L_2(\omega) = \theta_1 - \omega$ for $\omega < \theta_1$, $\omega - \theta_2$ for $\omega > \theta_2$, 0 elsewhere; $L_3(\omega) = \theta_2 - \omega$ for $\omega < \theta_2$, 0 elsewhere; where $\theta_1 = 2, \theta_2 = 4, p(x \mid \omega) = e^{-1/2\omega^2}e^{x\omega}, d\mu(x) = e^{-1/2x^2} / \sqrt{2\pi}\, dx, X = \Omega = (-\infty, \infty)$. If F concentrates at 5 and -5 with probability $\frac{1}{2}$ each, then it can be easily shown that the Bayes strategy has the form that action 1 is taken if $e^{-10x} > \frac{1}{3}$ and action 3 is taken when $e^{-10x} < \frac{1}{3}$. Here action 2 is never taken.

5. Admissibility of monotone procedures. To determine when the monotone procedures form a minimal complete class is, in general, very complicated. In this section we obtain several results which provide the answers to the question of admissibility in some specific circumstances. These examples point out the dependence of the question of minimality on the loss functions involved.

I. *Admissibility for $n = 2$ actions.* Throughout this section we assume that $\sigma_\omega = X$ and that for $x_1 > x_2$, $\omega_1 > \omega_2$, x_i in X, and ω_i in Ω,

$$p(x_1 \mid \omega_1)p(x_2 \mid \omega_2) > p(x_1 \mid \omega_2)p(x_2 \mid \omega_1).$$

THEOREM 4. *If the loss functions $L_1 - L_2$ change signs precisely once with ω_0, the change point, such that ω_0 is interior to Ω, and $L_1 - L_2$ possesses no other possible zero in the neighborhood of ω_0 except possibly ω_0, then every monotone procedure is admissible.*

First we establish

LEMMA 5. *Under the assumptions of Theorem 4 every monotone strategy involving both actions is unique Bayes (except for at most one value of x) against a two-point distribution F for nature. Moreover, the two points ω_1 and ω_2 can be prescribed arbitrarily subject only to the condition $\omega_1 < \omega_0$ and $\omega_2 > \omega_0$.*

PROOF. If $\omega_1 < \omega_0$ and $\omega_2 > \omega_0$, then by hypothesis $L_1(\omega_1) - L_2(\omega_1) < 0$ and $L_1(\omega_2) - L_2(\omega_1) > 0$. Let $\phi^0 = (\phi_1^0, 1 - \phi_1^0)$ denote a monotone procedure determined by the point x_0 in X, viz.,

$$\phi_1^0 = \begin{cases} 1, & x < x_0, \\ \lambda, & x = x_0, \quad 0 \leqq \lambda \leqq 1, \quad x_0 \text{ interior to } X \\ 0, & x > x_0, \end{cases} \tag{11}$$

Determine μ by the relation

$$\mu p(x_0 \mid \omega_1)[(L_1 - L_2)(\omega_1)] = (1 - \mu)p(x_0 \mid \omega_2)[(L_2 - L_1)(\omega_2)]. \tag{12}$$

It follows from (12) that $0 < \mu < 1$. Let F be the distribution concentrating at ω_1 with weight μ and at ω_2 with weight $1 - \mu$. Equations (8) and (9) yield the result that

$$g(x) = \int p(x \mid \omega)[L_1(\omega) - L_2(\omega)]\, dF(\omega)$$

vanishes for $x = x_0$. By Corollary 2 and Theorem 3, we infer that $g(x) > 0$ for $x > x_0$ and $g(x) < 0$ for $x < x_0$. This means that the unique procedure for the statistician in minimizing the risk is to take action 2 for $x > x_0$ and action 1 for $x < x_0$. If $x = x_0$, then either action yields the same expected return.

PROOF OF THEOREM 4. It will now be shown that the monotone strategy described by (11) is admissible. To this end, if ϕ^* is a strategy dominating ϕ^0 for all ω in Ω, then on account of Lemma 5, $\phi^* = \phi^0$ except possibly for $x = x_0$. Let the value of $\phi^*(x_0) = \lambda^*$. Observe that for all ω

$$0 \geqq \rho(\phi^*, \omega) - \rho(\phi, \omega) = p(x_0 \mid \omega)(\lambda^* - \lambda)[L_1(\omega) - L_2(\omega)]\mu\{x_0\}.$$

Since $L_1 - L_2$ changes sign in the interior of Ω and our assumptions imply that $p(x_0, \omega) > 0$ except at the left endpoint of Ω, $(\lambda^* - \lambda)\mu\{x_0\} = 0$, so that $\rho(\phi^*, \omega) =$

$\rho(\phi, \omega)$ for all ω. The admissibility of the special monotone strategies $\phi \equiv \{1, 0\}$ and $\phi \equiv \{0, 1\}$ is a trivial fact to establish. This completes the proof.

II. *Admissibility for $n = 3$ actions.* The objective here is to study the case of 3 actions. Such problems occur in many practical situations and are of interest. For instance, the two-sided testing problem is of this form where the loss for the alternative hypothesis depends on which side of the hypothesis the true parameter value lies.

Throughout the remainder of this section we specialize the class of densities to the exponential family, i.e., $p(x \mid \omega) = \beta(\omega)e^{\omega x}$.

The following further restrictions are placed upon L_1, L_2, and L_3.

ASSUMPTION E. It is required that L_i, $i = 1, 2, 3$, do not grow exponentially and that they do not simultaneously in any region exponentially, approach zero.

LEMMA 6. *If Assumption E is satisfied and Ω contains infinitely many ω values tending either to + or − infinity (for definiteness let $\omega_i \to +\infty$ belong to Ω), then every monotone strategy involving all 3 actions is unique Bayes (except for at most two values of x) against a three-point distribution F for nature.*

PROOF. Let ω_1, ω_2, and ω_3 be chosen from the sets S_1, S_2, and S_3 (see condition C) such that $(L_i - L_{i+1})(\omega_j) \neq 0$ for $i = 1, 2$ and $j = 1, 2, 3$. Consider the following system of equations in the unknowns λ_1, λ_2, and λ_3 with $x_2 > x_1$ prescribed and ω_1, ω_2, and ω_3 selected as above.

$$(13)\quad \begin{aligned} &\lambda_1(L_1 - L_2)(\omega_1)e^{\omega_1 x_1} + \lambda_2(L_1 - L_2)(\omega_2)e^{\omega_2 x_1} + \lambda_3(L_1 - L_2)(\omega_3)e^{\omega_3 x_1} = 0, \\ &\lambda_1(L_2 - L_3)(\omega_1)e^{\omega_1 x_2} + \lambda_2(L_2 - L_3)(\omega_2)e^{\omega_2 x_2} + \lambda_3(L_2 - L_3)(\omega_3)e^{\omega_3 x_2} = 0. \end{aligned}$$

The determinant

$$\begin{vmatrix} (L_1 - L_2)(\omega_1)e^{\omega_1 x_1} & (L_1 - L_2)(\omega_2)e^{\omega_2 x_1} & (L_1 - L_2)(\omega_3)e^{\omega_3 x_1} \\ (L_2 - L_3)(\omega_1)e^{\omega_1 x_2} & (L_2 - L_3)(\omega_2)e^{\omega_2 x_2} & (L_2 - L_3)(\omega_3)e^{\omega_3 x_2} \\ a_1 & a_2 & a_3 \end{vmatrix}$$

has the property that when the vector $a = (a_1, a_2, a_3)$ is equal to the first or second row vector, then it vanishes. From this fact, we deduce that λ_1, λ_2, λ_3 can be chosen proportional to the co-factors of the last row, respectively.

Noting that $(L_1 - L_2)(\omega_1) < 0$, $(L_1 - L_2)(\omega_2) > 0$, $(L_1 - L_2)(\omega_3) > 0$, $(L_2 - L_3)(\omega_1) < 0$, $(L_2 - L_3)(\omega_2) < 0$, and $(L_2 - L_3)(\omega_3) > 0$, we readily find that the co-factors of a_1 and a_3 are positive. The co-factor of a_2 is

$$\begin{aligned} -[(L_1 - L_2)(\omega_1)(L_2 - L_3)(\omega_3)e^{\omega_1 x_1 + \omega_3 x_2} - (L_1 - L_2)(\omega_3)(L_2 - L_3)(\omega_1)e^{\omega_1 x_2 + \omega_3 x_1}] \\ = e^{+\omega_1 x_2 + \omega_3 x_1}[(L_2 - L_1)(\omega_1)(L_2 - L_3)(\omega_3)e^{(\omega_3 - \omega_1)(x_2 - x_1)} \\ - (L_1 - L_2)(\omega_3)(L_3 - L_2)(\omega_1)]. \end{aligned}$$

In view of Assumption E, if ω_3 is chosen sufficiently large then this last expression is positive and hence $\lambda_2 > 0$. Put $\mu_1 = k\lambda_1 / \beta(\omega_1)$, $\mu_2 = k\lambda_2 / \beta(\omega_2)$, and $\mu_3 = k\lambda_3 / \beta(\omega_3)$, and normalize μ_i by suitable choice of k so that $\sum \mu_i = 1$

with $\mu_i \geqq 0$. Let F be a distribution concentrating μ_i at ω_i; then equations (13) become

$$\int e^{\omega x_1}\beta(\omega)[L_1(\omega) - L_2(\omega)]\, dF(\omega) = 0,$$

$$\int e^{\omega x_2}\beta(\omega)[L_2(\omega) - L_3(\omega)]\, dF(\omega) = 0.$$

Corollary 2 and Theorem 3 imply that

$$\int e^{\omega x}\beta(\omega)[L_1(\omega) - L_2(\omega)]\, dF(\omega)\begin{cases} > 0, & x > x_1, \\ < 0, & x < x_1, \end{cases}$$

and

$$\int e^{\omega x}\beta(\omega)[L_2(\omega) - L_3(\omega)]\, dF(\omega)\begin{cases} > 0, & x > x_2, \\ < 0, & x < x_2. \end{cases}$$

Consequently, the optimal procedure is to take action 1 when $x < x_1$, action 2 for $x_1 < x < x_2$, and action 3 for $x > x_2$. Indifference exists between actions 1 and 2 for $x = x_1$ and between actions 2 and 3 for $x = x_2$. Thus, if a monotone strategy $\phi = \{\phi_1, \phi_2, \phi_3\}$ with

$$\phi_1 = \begin{cases} 1, & x < x_1, \\ \lambda_1, & x = x_1, \\ 0, & x > x_1, \end{cases} \qquad \phi_3 = \begin{cases} 0, & x < x_2, \\ 1, & x_2 < x, \\ \lambda_3, & x = x_2 \end{cases}$$

is prescribed, then we have constructed a three-point distribution F against which the given ϕ is the unique Bayes strategy with the exception of two possible values for the random observation x. The proof of the lemma is hereby complete.

THEOREM 5. *Under the assumptions of Lemma 6 every monotone strategy involving all 3 actions is admissible.*

PROOF. If ϕ^* is any strategy which dominates the given ϕ, then by Lemma 6 $\phi = \phi^*$ except possibly at $x = x_1$ and x_2. Suppose $\mu\{x_1\}$ or $\mu\{x_2\}$ is positive (otherwise, Theorem 5 is established). By (9)

$$\begin{aligned} 0 \geqq \rho(\phi^*, \omega) - \rho(\phi, \omega) = \beta(\omega)\{&e^{x_1\omega}(\lambda_1^* - \lambda_1)\mu\{x_1\}[L_1(\omega) - L_2(\omega)] \\ &+ e^{x_2\omega}(\lambda_2^* - \lambda_2)\mu\{x_2\}[L_2(\omega) - L_3(\omega)]\}. \end{aligned} \tag{14}$$

Letting $\omega \to +\infty$ and observing that the second term dominates, we conclude that $(\lambda_2^* - \lambda_2)\mu\{x_2\} \leqq 0$. Examining ω in S_1 on account of (14) compels $(\lambda_1^* - \lambda_1)\mu\{x_1\} \geqq 0$. Finally, evaluating relation (14) for ω in S_2, using the established facts that $(\lambda_2^* - \lambda_2)\mu\{x_2\} \leqq 0$ and $(\lambda_1^* - \lambda_1)\mu\{x_1\} \geqq 0$, yields that $(\lambda_1^* - \lambda_1)\mu\{x_1\} = 0 = (\lambda_2^* - \lambda_2)\mu\{x_2\}$, whence $\rho(\phi^*, \omega) - \rho(\phi, \omega) = 0$ for all ω, and the proof is complete.

The above theorem does not treat the special monotone strategies which involve only two possible actions. These strategies will now be shown to be admissible For. instance, suppose ϕ represents a monotone strategy which involves only actions 1 and 3. Precisely, let $\phi_1(x) = 1$ for $x < x_1$, 0 for $x > x_1$; $\phi_3(x) = 1$ for $x > x_1$, 0 for $x < x_1$; and $\phi_2 \equiv 0$. Suppose ϕ^0 is a monotone strategy which dominates ϕ determined by the critical values x_1^0, x_2^0. Three cases will now be considered.

CASE 1. $x_1^0 < x_1 < x_2^0$. In view of (9)

$$0 \geqq \rho(\phi^0, \omega) - \rho(\phi, \omega) = -[L_1(\omega) - L_2(\omega)]\beta(\omega) \int_{x_1^0}^{x_1} e^{\omega x}\, d\mu(x)$$

$$+ [L_2(\omega) - L_3(\omega)]\beta(\omega) \int_{x_1}^{x_2^0} e^{\omega x}\, d\mu(x).$$

Since the second term dominates as $\omega \to +\infty$, we arrive at a contradiction.

CASE 2. $x_1 \leqq x_1^0 < x_2^0$. By (9)

$$0 \geqq \rho(\phi^0, \omega) - \rho(\phi, \omega) = [L_1(\omega) - L_2(\omega)]\beta(\omega) \int_{x_1}^{x_1^0} e^{\omega x}\, d\mu(x)$$

$$+ [L_2(\omega) - L_3(\omega)]\beta(\omega) \int_{x_1}^{x_2^0} e^{\omega x}\, d\mu(x).$$

However, this inequality is impossible for ω in S_3.

CASE 3. $x_1^0 < x_2^0 \leqq x_1$. This case can be handled similarly to that of case 2 above by examining ω in S_1.

The other types of monotone strategies involving at most two possible actions are treated similarly. This argument can also be extended to yield the conclusion of Theorem 5. The result of Lemma 6 is, however, stronger and possesses independent interest.

We now produce an example to show that when the state space of nature ω is restricted to a finite range, the conclusion of Theorem 5 is not valid. Let $\phi = (\phi_1, \phi_2, \phi_3)$ be a monotone strategy given by $\phi_1(x) = 1$ on $x \leqq x_1$, 0 elsewhere; $\phi_2(x) = 1$ on $x_1 < x \leqq x_2$, and zero elsewhere, with $\phi_3 = 1 - \phi_1 - \phi_2$. We desire to construct a monotone strategy $\phi^0 = \{\phi_1^0, \phi_2^0, \phi_3^0\}$ which dominates ϕ. Let ϕ^0 be determined by the critical values x_1^0 and x_2^0 with $x_1^0 < x_1$ and $x_2^0 > x_2$. Consider, according to (9),

$$\rho(\phi^0, \omega) - \rho(\phi, \omega)$$

$$= \beta(\omega)\Big\{ (L_1 - L_2)(\omega) \int e^{x\omega}[\phi_1^0(x) - \phi_1(x)]\, d\mu(x)$$

$$+ (L_2 - L_3)(\omega) \int e^{x\omega}[\phi_1^0 + \phi_2^0 - \phi_1 - \phi_2[\, d\mu(x)\Big\} \tag{15}$$

$$= \beta(\omega)\Big\{ -(L_1 - L_2)(\omega) \int_{x_1^0}^{x_1} e^{x\omega}\, d\mu(x) + [(L_2 - L_3)(\omega)] \int_{x_2}^{x_2^0} e^{x\omega}\, d\mu(x)\Big\}.$$

In the region S_2, (15) is automatically negative. In region S_3, we choose the loss functions so that

$$|(L_2 - L_3)(\omega)| > \frac{|(L_1 - L_2)(\omega)| \int_{x_1}^{x_1^0} e^{x\omega}\, d\mu(x)}{\int_{x_2}^{x_2^0} e^{x\omega}\, d\mu(x)}$$

and $L_3(\omega) > L_2(\omega)$. This can be done since ω ranges over a finite interval and hence the integrals are continuous and bounded away from 0 and ∞. Similarly, we determine in S_1 the loss functions L_1, L_2, and L_3 so that

$$|(L_1 - L_2)(\omega)| > \frac{\left|[(L_2 - L_3)(\omega)] \int_{x_2}^{x_2^0} e^{x\omega}\, d\mu(x)\right|}{\int_{x_1^0}^{x_1} e^{x\omega}\, d\mu(x)} \text{ with } L_1(\omega) < L_2(\omega).$$

With these determinations of the loss functions. L_1, L_2, and L_3, we see that (15) is always negative for the region of ω under consideration. Thus ϕ is not admissible. Intuitively, for any monotone strategy defined by critical values x_1 and x_2, loss functions can be chosen so that one wants to take action 2 more often than prescribed by the given strategy. An example, where the natural range of the distribution is a finite interval and where the above construction is valid, is obtained by setting $d\mu(x) = e^{-|x|}\, dx$. In connection with Assumption E, examples can be constructed to show that the growth restriction is essential in order to ensure the validity of Theorem 5.

III. *Admissibility for $n = 4$ actions.* Our next task is to analyze the case of four possible actions. Again assumption E is to hold.

LEMMA 7. *If the parameter space ω contains arbitrarily large values of ω tending to + and − infinity, then every monotone strategy involving all four actions is unique Bayes, except for three possible values, against a distribution $F(\omega)$ involving four points.*

PROOF. Let a monotone strategy ϕ be given, defined by the critical dividing numbers x_1, x_2, and $x_3 (x_1 < x_2 < x_3)$. Choose $\omega_1 < \omega_2 < \omega_3 < \omega_4$ from the four regions S_1, S_2, S_3, and S_4. Consider the system of equations in the unknowns λ_1, λ_2, λ_3, and λ_4 given by

$$\sum_{j=1}^{4} \lambda_j (L_i - L_{i+1})(\omega_j) e^{\omega_j x_i} = 0, \qquad i = 1, 2, 3. \tag{16}$$

As in the proof of Lemma 6, the solutions λ_i are proportional to the co-factors of the last row in the determinant

$$\begin{vmatrix} (L_1 - L_2)(\omega_1) e^{\omega_1 x_1} & (L_1 - L_2)(\omega_2) e^{\omega_2 x_1} & \cdots & (L_1 - L_2)(\omega_4) e^{\omega_4 x_1} \\ (L_2 - L_3)(\omega_1) e^{\omega_1 x_2} & (L_2 - L_3)(\omega_2) e^{\omega_2 x_2} & \cdots & \\ (L_3 - L_4)(\omega_1) e^{\omega_1 x_3} & & \cdots & (L_3 - L_4)(\omega_4) e^{\omega_4 x_3} \\ a_1 & a_2 & a_3 & a_4 \end{vmatrix}$$

Choosing ω_1 sufficiently large negatively and ω_4 sufficiently large positively, the signs of the subdeterminants are determined by the signs of the principle diagonals. It readily follows that by such a choice of ω_1 and ω_4 we get $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 > 0$, and $\lambda_4 > 0$. Define $\mu_i = k\lambda_i / \beta(\omega_i)$ with $\sum \mu_i = 1$ and F a distribution concentrating μ_i at ω_i. It is easy to see by (16) that the given strategy ϕ is unique Bayes against F with the exception of three possible values of x.

THEOREM 6. *Under the conditions of Lemma* 7, *every monotone strategy is admissible.*

The proof can be patterned, with some slight modification, after that of Theorem 5, and the remarks following Theorem 5. We omit the details.

The assumption in Lemma 7 concerning the parameter values ω extending to $\pm\infty$ is essential. If this condition is removed, an example can be produced which permits ω to range to $+\infty$, and is bounded on the left such that suitable monotone strategies are not admissible. The construction is similar to that shown in connection with three actions where we make use of the exponential for $\omega \to \infty$. This example will be left as an exercise for the interested reader.

In the case of 5 actions, even if the parameter range for ω is the full infinite interval, loss functions can be defined so that not all monotone procedures are admissible.

IV. *Admissibility for n actions.* We now present some examples of general interest for n actions where the monotone strategies are all admissible. Let the loss functions L_i satisfy the conditions of A, B, C, and D and the further property

$$L_i(\omega) = L_{i+1}(\omega) \qquad \text{for } \omega \in S_i \quad \text{or} \quad S_{i+1}. \tag{17}$$

Of course, by condition D, $L_i(\omega) < L_j(\omega)$ for $\omega \in S_i$. We shall now show that every monotone strategy $\phi = (\phi_i)$ is unique Bayes except for $n - 1$ possible values against an n-point distribution. Indeed, select $\omega_1 < \omega_2 < \omega_3 < \cdots \omega_n$, and ω_i in S_i. Note, $L_i(\omega_i) - L_{i+1}(\omega_i) < 0$ and $L_i(\omega_{i+1}) - L_{i+1}(\omega_{i+1}) > 0$. Let the strategy (ϕ_i) be described by the critical values $x_i (x_i \leqq x_{i+1})$. Consider the system of linear equations in the unknowns λ_j:

$$\sum_{j=1}^{n} \lambda_j (L_i - L_{i+1})(\omega_j) e^{\omega_j x_i} = 0, \qquad i = 1, \cdots, n-1. \tag{18}$$

The solutions λ_j are proportional to the co-factors of the last row of the determinant

$$\begin{vmatrix} (L_1 - L_2)(\omega_1)e^{\omega_1 x_1} & (L_1 - L_2)(\omega_2)e^{\omega_2 x_1} & 0 & \cdots & 0 \\ 0 & (L_2 - L_3)(\omega_2)e^{\omega_2 x_2} & (L_2 - L_3)(\omega_3)e^{\omega_3 x_3} & \cdots & 0 \\ & 0 & (L_3 - L_4)(\omega_3)e^{\omega_3 x_3} & \cdots & \\ 0 & 0 & 0 & & \\ \vdots & \vdots & \vdots & & (L_{n-2} - L_{n-1})(\omega_{n-1})e^{\omega_{n-1} x_{n-1}} \\ 0 & 0 & 0 & \cdots & (L_{n-1} - L_n)(\omega_n)e^{\omega_n x_{n-1}} \\ a_1 & a_2 & a_3 & \cdots & a_n \end{vmatrix}$$

The zeros appear in this determinantal expression as a consequence of (17). We deduce easily that λ_i are all of one sign. Put $\mu_i = k\lambda_i \,/\, \beta(\omega_i) > 0$ and $\sum \mu_i = 1$. Define F to concentrate weight μ_i at ω_i. Equations (18), Corollary 2, and Theorem 3 imply that ϕ is unique Bayes, except for $n - 1$ possible values of x, against $F(\omega)$. Thus we have established

THEOREM 7. *If the loss functions L_i satisfy the additional property* (17), *then every monotone strategy $\phi = \{\phi_i\}$ is unique Bayes, except for $n - 1$ values of x, against a distribution concentrating at n points ω_i. The parameter values ω_i can be chosen arbitrarily provided only that ω_i is in S_i.*

THEOREM 8. *Under the conditions of Theorem 7, every monotone strategy is admissible.*

PROOF. If $\phi = \{\phi_i\}$ is a monotone procedure determined by the critical values $x_i(x_i \leqq x_{i+1})$, then by Theorem 6, ϕ is Bayes against a distribution concentrating on n values. If ϕ^* is a monotone procedure which dominates ϕ, then by Theorem 7, $\phi = \phi^*$ except possibly for $x = x_i$. Let $\phi^*(x_i) = \alpha_i^*$ and $\phi(x_i) = \alpha_i$. Considering ω in S_1 we find, since

$$0 \leqq \rho(\phi, \omega) - \rho(\phi^*, \omega) = \beta(\omega)e^{\omega x_1}(\alpha_1 - \alpha_1^*)[L_1(\omega) - L_2(\omega)]\mu\{x_1\},$$

that $(\alpha_1 - \alpha_1^*)\mu\{x_1\} \leqq 0$.

Examining the risks for ω in S_2, we get

$$0 \leqq \beta(\omega)\{e^{\omega x_1}(\alpha_1 - \alpha_1^*)[L_1(\omega) - L_2(\omega)]\mu\{x_1\} + e^{\omega x_2}(\alpha_2 - \alpha_2^*)[L_2(\omega) - L_3(\omega)]\mu\{x_2\}\}$$

which implies as $(\alpha_1 - \alpha_1^*)\mu\{x_1\} \leqq 0$ that $(\alpha_2 - \alpha_2^*)\mu\{x_2\} \leqq 0$. Continued application of this analysis by successively looking at ω in S_i $(i = 1, \cdots, n - 1)$ we conclude that $(\alpha_i - \alpha_i^*)\mu\{x_i\} \leqq 0$. Finally, for ω in S_n, we conclude that $(\alpha_{n-1} - \alpha_{n-1}^*)\mu\{x_n\} = 0$ and working back we find successively that $(\alpha_i - \alpha_i^*)\mu\{x_i\} = 0$. Consequently, $\rho(\phi, \omega) - \rho(\phi^*, \omega) = 0$ for all ω, and the proof of the theorem is complete.

An important application of Theorem 7 arises when we consider the situation where $L_i(\omega) = a$ for $\omega \not\in S_i$ and $L_i(\omega) = 0$ for $\omega \in S_i$. In other words, the statistician is penalized a fixed amount if the wrong decision is made independently of the action taken, with zero loss if the correct decision is made. Then the conclusion of Theorem 8 can be stated as follows: The collection of all monotone procedures form a minimal essentially complete class.

This section is closed with the further enumeration of some examples of minimal essentially complete classes. In an n action case if $L_i(\omega) = |\, i - j \,|\, a$ for ω in S_j, then the collection of monotone procedures constitutes a minimal essentially complete class. This is a situation where the penalty of a wrong decision is proportional to how far the decision is from the correct action. The proof of this last fact is omitted.

6. Minimax strategies for nature. In this section we characterize the form of the minimax strategies for nature in the case of two actions. The underlying distribution, as before, has the form

$$P(x \mid \omega) = \int_{-\infty}^{x} p(x \mid \omega) \, d\mu(x),$$

where p satisfies the condition given in Section 1. We now require further that $p(x \mid \omega)$ is continuous in ω for each fixed x. The loss functions L_1 and L_2 have the properties A through D (see Section 1) and are in addition continuous. Without loss of generality, they may take the form $L_1(\omega) = 0$ for $\omega \leqq \theta$, $L_1(\omega) > 0$, $\omega > \theta$, while $L_2(\omega) > 0$ for $\omega < \theta$, $L_2(\omega) = 0$, $\omega \geqq \theta$, where θ is interior to Ω. Let $\rho(\phi, F)$ denote the expected risk when nature chooses a distribution F, and the statistician follows the procedure ϕ. Lebesgue's convergence criterion, in view of the continuity assumptions on L_i, implies that $\rho(\phi, \omega)$ is continuous in ω. Let the smallest interval containing Ω be denoted by $[a, b]$. The points a, b may or may not belong to Ω and the values $\pm\infty$ are not excluded. If a does not belong to Ω, then we assume that for any monotone strategy ϕ for which $\phi \not\equiv 0$,

$$\lim_{\omega \to a} L_2(\omega) \int [1 - \phi(t)] p(t \mid \omega) \, d\mu(t) = 0. \tag{19}$$

Similarly, we assume that if b does not belong to Ω, then for any monotone strategy ϕ with $\phi \not\equiv 1$,

$$\lim_{\omega \to b} L_1(\omega) \int \phi(t) p(t \mid \omega) \, d\mu(t) = 0. \tag{20}$$

If a belongs to Ω, then we do not impose condition (19). A similar statement applies to the endpoint b.

If L_1 and L_2 do not grow exponentially and the family of distributions belongs to the exponential class, then it is easy to verify the validity of conditions (19) and (20).

Consider the game G_n defined as follows: Let ω range over the interval $[\omega_n, \omega_n^*]$ where $\omega_n < \theta < \omega_n^*$. If a belongs to Ω, we take $\omega_n = a$ and if b belongs to Ω take $\omega_n^* = b$. Otherwise let $\omega_n \to a$ and $\omega_n^* \to b$. Let the strategy space for the statistician consist of all monotone strategies and let the strategy space for nature $\mathcal{S}_n$ consist of all distributions on the closed interval $[\omega_n, \omega_n^*]$. The payoff is the risk $\rho(\phi, F)$ which is continuous in ϕ and in F in the usual weak topology imposed on these sets of strategies. The reason we can restrict ourselves to the monotone procedures in the consideration of the game G_n is a consequence of Theorem 1. The following facts will be used in the course of the subsequent analysis: Suppose $\phi_r(x) \to 1$ for every x, where $\phi_r(x) = 1, x < x_r, = 0, x > x_r$. For any set of ω where $\omega \in \Omega \cap [\tilde{\omega}, \tilde{\omega}^*] = W$ with both $\tilde{\omega}$ and $\tilde{\omega}^*$ interior to Ω, it follows that

$$\int [1 - \phi_r(x)] p(x \mid \omega) \, d\mu(x) \tag{21a}$$

converges uniformly to zero in W as $r \to \infty$. Similarly, if $\phi_r(x) \to 0$ where $\phi_r(x) = 0$, $x < x_r$, $= 1$, $x > x_r$, then for any set ω of the form W, it follows that

$$\int \phi_r(x) p(x \mid \omega) \, d\mu(x) \tag{21b}$$

converges uniformly to zero.

That we obtain pointwise convergence in (21a) and (21b) follows from Lebesgue's convergence criterion. The uniformity of the convergence can be secured with the aid of the Collorary to Lemma C in Section 1.

Since the spaces of strategies are both compact, it is a well-known result that the game is determined. Let ϕ_0^n and F_0^n denote minimax strategies for the statistician and nature respectively. If v_n denotes the value of the game, then

$$\begin{aligned} \rho(\phi_0^n, F) &\leqq v_n \qquad \text{all } F \text{ in } \mathcal{S}_n, \\ \rho(\phi, F_0^n) &\geqq v_n \qquad \text{all monotone } \phi. \end{aligned} \tag{22}$$

Let T_n denote the set of all ω satisfying $\rho(\phi_0^n, \omega) = v_n$. The set T_n cannot be fully contained in either of the intervals $I_1^n = [\omega_n, \theta]$ or $I_2^n = [\theta, \omega_n^*]$. Indeed, if $T_n \subseteq [\theta, \omega_n^*]$, then since the spectrum of F_0^n must lie in T_n by examining $\phi = (\phi_1, 1 - \phi_1)$, where $\phi_1 \equiv 0$, we secure, using the fact that $L_2(\omega) = 0$ for $\omega \geqq \theta$, that $\rho(\theta, F_0^n) \equiv 0$, an impossibility. Hence, T_n contains points of both intervals I_1^n and I_2^n. This last analysis also implies that the monotone procedure ϕ_0^n is not identically 1 or 0. Choose ω_1 in $I_1^n \cap T_n$ and ω_2 in $I_2^n \cap T_n$. In view of Lemma 5, there exists a distribution $\tilde{F}^n(\omega)$ with spectrum consisting of ω_1 and ω_2 such that ϕ_0^n is Bayes against $\tilde{F}^n$. Since ω_1 and ω_2 belong to T, we get that $v = \rho(\phi_0^n, \tilde{F}^n) = \min_\phi \rho(\phi, \tilde{F}^n)$. Hence $\tilde{F}^n$ is a minimax strategy for nature involving only two points. Allow n to go to infinity and select a limit monotone strategy $\phi_0 = \lim_{i \to \infty} \phi_0^{n_i}$. It will now be shown that $\phi^0 = (\phi_1^0, 1 - \phi_1^0)$ cannot have $\phi_1^0 \equiv 0$ or $\phi_1^0 \equiv 1$. First note that by choosing any two-point strategy $F(\omega)$ for nature, it follows that $v_n \geqq \alpha > 0$. Consider the case where $\phi_1^0 \equiv 1$. If x_{n_i} represents the critical dividing point for the strategy $\phi_0^{n_i}$, then x_{n_i} must converge to the right-hand endpoint of X. But,

$$\rho(\phi_0^{n_i}, \omega) = L_2(\omega) \int (1 - \phi_0^{n_i}) p(x \mid \omega) \, d\mu(x)$$

tends uniformly to zero for $\omega < \theta$. This is a consequence of assumption (19) and equation (21a). As $v_{n_i} \geqq \alpha > 0$, we deduce for n_i sufficiently large that T_{n_i} must lie wholly in the interval $I_2^{n_i}$, contradicting the fact established above that T_{n_i} intersects both $I_1^{n_i}$ and $I_2^{n_i}$. A similar argument using (20) and (21b) eliminates the possibility that $\phi_1^0 \equiv 0$. This completes the proof of the assertion made.

In view of conditions (19) and (20), we find easily that $\rho(\phi^0, \omega) \to 0$ as $\omega \to a$ and b, if a and b do not belong to Ω. It is clear that there exists infinitely many n_i which we enumerate through m such that $\phi_1^m \leqq \phi_1^0$ or $\phi_1^m \geqq \phi_1^0$. Without loss of generality, let us consider the case where $\phi_1^m \leqq \phi_1^0$. It follows now that for all $m > m_0$ there exists a subinterval $U = [\omega', \omega'']$ depending on $\epsilon < \alpha$ with ω', ω'' interior to Ω such that for $\omega \not\in U$ and $m > m_0$, $\rho(\phi^m, \omega) \leqq \varepsilon < \alpha$. Consequently, all the two-point distributions $\tilde{F}^m$ constructed above have the property that their spectrum is simultaneously contained in U. A limit two-point distribution $\tilde{F}$ can now be selected with spectrum in U as U is compact. By considering an appropriate subsequence $n_i = k$, it follows that $\lim_{k\to\infty} v_k = v$, $\rho(\phi^0, F) \leqq v$ for every F, and $\rho(\phi; \tilde{F}) \geqq v$ for any monotone strategy ϕ. These last inequalities can be expressed in the following theorem.

THEOREM 9. *If the loss functions and distributions satisfy* (19) *and* (20), *then the game with risk function* $\rho(\phi, F)$ *as* ϕ *ranges over all strategies and* F *ranges over all distributions on* Ω *is determined, i.e.,*

$$\min_{\phi} \max_{F} \rho(\phi, F) = \max_{F} \min_{\phi} \rho(\phi, F). \tag{23}$$

Moreover, there exists a minimax strategy F for nature involving only two points of increase.

A careful study of the proof actually shows

COROLLARY. *If* (23) *holds and there exists a minimax strategy for nature, then there exists a minimax strategy for nature with a two-point spectrum.*

The conditions (19) and (20) were imposed to ensure the determinateness of the game $\rho(\phi, F)$ as given in (23).

Another type of result is to study Bayes distribution for nature against a given monotone procedure. We limit ourselves for $P(x \mid \omega)$ to the exponential class and we assume that $L_i(\omega)$ are such that (19) and (20) hold. For simplicity put $\theta = 0$.

The yield, if the statistician employs the procedure $\phi_1(x) = 1$, $x \leqq x_0$, $= 0$, $x > x_0$, becomes

$$H(\omega) = \begin{cases} \beta(\omega)L_1(\omega) \int_{-\infty}^{x_0} e^{x\omega}\, d\mu(x) & \text{for } \omega \geqq 0, \\ \beta(\omega)L_2(\omega) \int_{x_0}^{\infty} e^{x\omega}\, d\mu(x) & \omega \leqq 0. \end{cases} \tag{23}$$

For definiteness, let us assume that the endpoints are $\pm\infty$ when open at that respective end. By (19) and (20), as $\omega \to \infty$ or $\omega \to -\infty$ $H(\omega) \to 0$. Also $H(0) = 0$. If L_i were analytic, then H could achieve its maximum at most a finite number of times and since it is a Bayes distribution with probability one at the maximum values, every Bayes distribution against a monotone procedure involves only a finite number of points. Examples can be produced which show that the number of points involved in a distribution may be more than 2. However, for certain distributions and suitable loss functions we can show that every Bayes distribution involves at most two points.

EXAMPLE 1. Suppose that $p(x \mid \omega) = 1 / (\sqrt{2\pi})e^{-1/2(x-\omega)^2}$ and for $\omega > 0$,

$L_1(\omega)$ satisfies the condition that $L_1(\omega) / L_1'(\omega)$ is nondecreasing and that $L_2(\omega) / L_2'(\omega)$ is nondecreasing, in $\omega < 0$. We now show that

$$(24) \qquad L_1(\omega) \frac{\int_{-\infty}^{x_0} e^{x\omega}\, d\mu(x)}{\int_{-\infty}^{\infty} e^{x\omega}\, d\mu(x)} = L_1(\omega) \frac{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x_0} e^{-1/2(x-\omega)^2}}{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-1/2(x-\omega)^2}\, dx} = L_1(\omega)\Phi(x_0 - \omega)$$

has a unique maximum for $\omega \geqq 0$ where Φ is the cumulative standard normal distribution. Differentiating (24) yields

$$(25) \qquad -L_1(\omega)\Phi'(x_0 - \omega) + L_1'(\omega)\Phi(x_0 - \omega).$$

Dividing by $L_1'(\omega)$ and $\Phi'(x_0 - \omega)$, we have

$$-\frac{L_1(\omega)}{L_1'(\omega)} + \frac{\Phi(x_0 - \omega)}{\Phi'(x_0 - \omega)}.$$

It is an easy matter to show for $\omega > 0$ that the second term is decreasing and thus (25) has at most one zero.

Hence for $\omega > 0$ (24) has a unique maximum. A similar analysis shows that

$$L_2(\omega)\beta(\omega) \int_{x_0}^{\infty} e^{x\omega}\, d\mu(x)$$

has a unique maximum for $\omega < 0$. This implies that for any monotone strategy for the statistician the Bayes strategy for nature concentrates at most at two points.

EXAMPLE 2. Imposing the same assumptions on the loss conditions, we now give a general sufficient condition on the distribution to ensure a unique maximum in each of the regions $\omega \geqq 0$ and $\omega \leqq 0$. Let

$$\beta_{x_0}(\omega) = \int_{-\infty}^{x_0} e^{x\omega}\, d\psi(x).$$

By (23)

$$H(\omega) = \beta(\omega)\beta_{x_0}(\omega)L_1(\omega) \qquad \text{for} \quad \omega > 0.$$

Let $m(\omega) = \beta(\omega)\beta_{x_0}(\omega)$. We deduce by analogous reasoning to that of Example 1, that if $-m(\omega) / m'(\omega)$ has the property that it is strictly monotone decreasing for $\omega > \theta$, then H has a unique maximum as $H(0) = 0$ and we make assumptions on L_1 so that $H(\omega) \to 0$ as $\omega \to \infty$. However, $m(\omega) / m'(\omega)$ is increasing if and only if $\log m$ is concave. But

$$\log m(\omega) = \log \beta(\omega) + \log \beta_{x_0}(\omega).$$

Let y denote the random variable with density $\beta(\omega)e^{x\omega}$ with respect to μ and y_{x_0} denote the random variable with the same density except that we alter μ

so that $\mu(x_1, \infty) = 0$. In other words, y_{x_0} is the truncated random variable where we do not allow observations $x > x_0$. The variances of these random variables will be denoted by

$$\sigma_\omega^2(y) \qquad \text{and} \qquad \sigma_\omega^2(y_{x_0})$$

when the true parameter value is ω. Since $d^2 / d\omega \log \beta(\omega) = \sigma_\omega^2(y)$, we obtain that $\log m$ has a negative second derivative if and only if

$$\sigma_\omega^2(y) > \sigma_\omega^2(y_{x_0}). \tag{26}$$

Thus if (26) holds for every x_0, then $H(\omega)$ has a unique maximum. In order for H in (23) to possess a unique maximum for $\omega < 0$, a corresponding condition to (26) must be satisfied when the random variable is truncated on the other side; namely, when we take $\mu(-\infty, x_0) = 0$. Thus a sufficient condition for at most two maxima for H is that (26) hold for every truncated variable on the left and right. Some instances where (26) holds for the exponential class are the normal, gamma, Poisson, and binomial. However, examples of exponential distributions can be constructed in which the seemingly natural condition (26 is not satisfied. We leave to the reader the task of constructing such examples.

7. Essentially complete classes for infinite number of actions. An analysis of essentially complete classes of decision procedures for distributions with monotone likelihood ratio for an infinite number of actions will now be carried out. The parameter space ω for nature will range over an interval Ω as before. The action space A will also consist of a closed subset of the real line. The loss function $L(\omega, a)(\omega$ in Ω, a in $A)$ is assumed to satisfy the following properties:

(i) For each ω, $L(\omega, a)$ attains its minimum as a function of a at a point $a = q(\omega)$ which is a monotone increasing function of ω.

(ii) For each ω, $L(\omega, a)$ as a function of a increases away from that minimum.

Without loss of generality we may take $L(\omega, q(\omega)) = 0$ for every ω.

Particularly important examples of decision problems whose loss functions satisfy (i) and (ii) are furnished by the estimation problem. Here a commonly used loss function is given by $L(\omega, a) = |\omega - a|^k (k > 0)$ where both ω and a traverse the infinite real line. The function $q(\omega)$ is evidently equal to ω.

A decision procedure or strategy for the statistician in this case is a probability measure $\nu(x)$ on A specified for every observation x. A monotone strategy is defined in this general situation as follows: If $x_1 > x_2$ and C_1 and C_2 are open sets with C_1 lying to the left of C_2, then either

$$\nu(C_1 \mid x_1) = 0 \qquad \text{or} \qquad \nu(C_2 \mid x_2) = 0.$$

This definition agrees with the meaning of monotone strategy for a finite number of actions given previously. In the case of convex loss functions (e.g., the estimation problem introduced above) where nonrandomized strategies are frequently employed, we obtain that a monotone decision procedure can be identified with a function ϕ with values in A such that ϕ is monotone nondecreasing.

The risk for ω and a given strategy ν has the form

$$\rho(\omega, \nu) = \iint L(\omega, a)\, d\nu(a \mid x) p(x \mid \omega)\, d\mu(x).$$

Our main objective in this section is to establish the essential completeness of the monotone strategies for the case where A is infinite and closed. The proof will be carried out in three stages: First, the theorem will be demonstrated for the case that $L(\omega, a)$ is continuous and bounded in a for each fixed ω; second, when $L(\omega, a)$ is 0 or 1, and then the general case.

The method of proof for the case where $L(\omega, a)$ is continuous in a entails a limiting argument by using the essential completeness theorem for a finite number of actions.

Let $\mathfrak{F}$ be the collection of all non-null finite subsets of A partially ordered by set inclusion, i.e., $B \geqq C$ if and only if $B \supseteq C$, where B and C are members of $\mathfrak{F}$. Let $\mathfrak{H}$ be any subfamily with the properties:

(a) If $B \in \mathfrak{H}$ and $C \in \mathfrak{H}$, there is a D in $\mathfrak{H}$ such that $B \cup C \leqq D$.

(b) $\bigcup_{B \in \mathfrak{H}} B$ is dense in A.

The family of finite sets $\mathfrak{H}$ in view of (a) and (b) form a directed system and therefore we can speak of convergence with respect to $\mathfrak{H}$.

We shall construct for every B in $\mathfrak{H}$ a new loss function $L_B(\omega, b)$, preserving the monotone properties of assumptions A through C for the case of a finite number of actions. For a given decision procedure ν, we shall then construct a new decision procedure ν_B concentrated on B. This will have the property that ν_B converges to ν as B gets large and that if L is bounded and continuous in a for each ω, $\rho(\omega, \nu_B)$ converges to $\rho(\omega, \nu)$. With the aid of Theorem 1 we can then produce a monotone strategy ν_B^* concentrating on B better than ν_B and if we take ν^* as a cluster point of ν_B^*, $\rho(\omega, \nu_B^*)$ will converge to $\rho(\omega, \nu^*)$ with $\rho(\omega, \nu^*) \leqq \rho(\omega, \nu)$ and ν^* a monotone procedure. The conclusions will then be extended to the case of loss functions $L(\omega, a)$ not necessarily continuous in a. The above discussion indicates the direction of the proof; we now proceed to develop the details.

For each a in A, define $I_B(a)$ to be the smallest closed interval whose endpoints are in B and which contains a (the endpoints may coincide), if there is such an interval. If there is no such interval, let $I_B(a)$ be the set consisting of the nearest element of B.

For a in A and b in B, put

$$f_B(a, b) = \begin{cases} 0 & \text{if } b \notin I_B(a), \\ \lambda & \text{if } I_B(a) = [b, c] \text{ or } [c, b], \\ & \qquad a = \lambda b + (1 - \lambda)c. \end{cases} \tag{27}$$

The function $f_B(a, b)$ will be used to distribute the probability of a over B. In fact, set

$$\nu_B(S \mid x) = \sum_{b \in S} \int f_B(a, b)\, d\nu(a \mid x). \tag{28}$$

The probability of $\nu(a \mid x)$ concentrated at a is distributed to the points b and c of B, where $a \varepsilon [b, c]$, proportional to the distance from a to b and c respectively. In forming $\nu_B(S \mid x)$ this is done for every a in A.

The loss function is now altered as follows: Set

$$L_B(\omega, b) = \begin{cases} L(\omega, b) & \text{if } b \not\varepsilon\, I_B(q(\omega)), \\ 0 & \text{if } b \,\varepsilon\, I_B(q(\omega)). \end{cases} \tag{29}$$

The function $L_B(\omega, b)$ is only different from $L(\omega, b)$ in the neighborhood of $q(\omega)$. Moreover, $L(\omega, b)$ is changed at most for two adjacent values of b and it is readily seen that for each fixed ω, $L_B(\omega, b)$ converges uniformly to $L(\omega, b)$ on B.

It is readily verified that the loss function $L_B(\omega, b)$ satisfies the conditions of (A) through (D). By Theorem 1 there exists a monotone procedure ν_B^* concentrating on B such that for the loss function $L_B(\omega, b)$

$$\rho_B(\omega, \nu_B^*) \leqq \rho_B(\omega, \nu^*).$$

It is important to note that the constructed monotone strategy ν_B^* depends only on ν_B and on q and not on the nature of the loss functions elsewhere. This was pointed out in the remark following Theorem 1, for the change points used there can be made to depend only on q.

Since the space of measures is compact in the weak $*$ topology we can select a measure $\nu^*(a \mid x)$ which is a simultaneous cluster point of $\nu_B^*(a \mid x)$ for every x. In view of the continuity of $L(\omega, a)$ as a function of a, we get for every ω

$$\begin{aligned} \rho(\omega, \nu_B^*) &= \iint L(\omega, a)\, d\nu_B^*(a \mid x) p(x \mid \omega)\, d\mu(x) \\ &\to \iint L(\omega, a)\, d\nu^*(a \mid x) \cdot p(x \mid \omega)\, d\mu(x) = \rho(\omega, \nu^*). \end{aligned} \tag{30}$$

As $L_B(\omega, b)$ converges uniformly to $L(\omega, b)$, we also find that

$$\begin{aligned} \rho_B(\omega, \nu_B) &= \iint L_B(\omega, a)\, d\nu_B(a \mid x) p(x \mid \omega)\, d\mu(x) - \rho(\omega, \nu_B) \\ &= \iint L(\omega, a)\, d\nu_B(a \mid x) p(x \mid \omega)\, d\mu(x) \end{aligned} \tag{31}$$

can be made as small as we desire for each fixed ω by choosing B sufficiently large. Our next task is to show that $\rho(\omega, \nu_B) \to \rho(\omega, \nu)$. To this end, define for fixed ω

$$\bar{\nu}(s) = \int \nu(s \mid x) p(x \mid \omega)\, d\mu(x). \tag{32}$$

The set function $\bar{\nu}$ is a probability distribution on A induced by the action ν if ω is the state of nature. Then clearly

$$\rho(\omega, \nu) = \int L(\omega, a)\, d\bar{\nu}(a). \tag{33}$$

We also see from the construction of ν_B that for each b in B

$$\begin{aligned} \nu_B((-\infty, b) \mid x) &\leqq \nu((-\infty, b) \mid x), \\ \nu_B((-\infty, b] \mid x) &\geqq \nu((-\infty, b] \mid x). \end{aligned} \tag{34}$$

The same inequalities persist for $\bar{\nu}_B$ and $\bar{\nu}$. Hence for any c in $\bigcup_{b \in \mathfrak{K}} B$, we have

$$\begin{aligned} \limsup_B \bar{\nu}_B((-\infty, c)) &\leqq \bar{\nu}(-\infty, c), \\ \liminf_B \bar{\nu}_B((-\infty, c]) &\geqq \bar{\nu}(-\infty, c], \end{aligned} \tag{35}$$

and hence $\bar{\nu}_B$ converges to $\bar{\nu}$ in the sense of measures. As $L(\omega, a)$ is continuous in a, we now conclude from (35) that $\rho(\omega, \nu_B) \rightarrow \rho(\omega, \nu)$. Combining (30), (31), (35), and the fact that $\rho_B(\omega, \nu_B^*) \leqq \rho_B(\omega, \nu_B^*)$ yields finally that

$$\rho(\omega, \nu^*) \leqq \rho(\omega, \nu). \tag{36}$$

Again, we emphasize the fact that ν^* depends only on ν and q and in no other way on the nature of the loss functions. We next show that ν^* is monotone and completely additive.

(a) Proof that ν^* is monotone: Let $x_1 < x_2$, C_1 and C_2 be two open sets where C_1 lies to the right of C_2. For any given $\epsilon > 0$ there exists a B such that

$$\nu_B^*(C_i \mid x_i) > \nu^*(C_i \mid x_i) - \epsilon, \qquad i = 1, 2.$$

Since for every B either $\nu_B^*(C_1 \mid x_1)$ or $\nu_B^*(C_2 \mid x_2)$ is 0, it follows that $\nu^*(C_1 \mid x_1)$ or $\nu^*(C_2 \mid x_2)$ is 0.

(b) Proof that ν^* is completely additive for almost all x for each ω. Consider

$$K(\nu^*, x) = 1 - \lim_{N \to \infty} \nu^*([-N, N] \mid x).$$

Let $L_n(\omega, a)$ be a sequence of new loss functions with the properties:

(1) $L_n(\omega, a)$ increases to $L_\infty(\omega, a)$,

(2) $\lim_{a \to \pm\infty} L_n(\omega, a) = n$,

(3) $\int L_\infty(\omega, a)\, d\bar{\nu}(a) < \infty$,

(4) $L_n(\theta, a) = L(\theta, a)$ for $\theta \neq \omega$.

For any completely additive measure there always exists a function increasing to infinity at the endpoints such that (3) holds. The $L_n(\omega, a)$ are so constructed that each remain bounded but converge to $L_\infty(\omega, a)$. The loss functions for $L(\theta, a)$ are not altered for $\theta \neq \omega$. On the other hand, only $L(\omega, a)$ is replaced by a sequence of $L_n(\omega, a)$ tending to infinity at the endpoints preserving q. Since the ν^* depended only on ν and q, we obtain

$$\begin{aligned} \rho_n(\omega, \nu) \geqq \rho_n(\omega, \nu^*) &= \iint L_n(\omega, a)\, d\nu^*(a \mid x) p(x \mid \omega)\, d\mu(x) \\ &\geqq (n - \varepsilon) \int K(\nu^*, x) p(x \mid \omega)\, d\mu(x). \end{aligned}$$

In view of (3) and (1), we find that $\rho_n(\omega, \nu) < \infty$ and therefore

$$\int K(\nu^*, x)p(x \mid \omega) \, d\mu(x) = 0.$$

Consequently, $K(\nu^*, x) = 0$ for almost all x for each ω. But this is equivalent to the countable additivity.

Combining all the previous conclusions leads to

THEOREM 9. *Given any decision procedure* ν, *there is a monotone decision procedure* ν^* *depending only on* ν *and* q *such that if* $L(\omega, a)$ *is a bounded continuous function of* a *for each fixed* ω *satisfying* (i) *and* (ii) *with the prescribed* q, *then*

$$\rho(\omega, \nu^*) \leqq \rho(\omega, \nu)$$

for all ω.

From the easily seen fact that every function which is 0 on I_1 and 1 on $I_2 \cup I_3$, where I_1, I_2, and I_3 are disjoint intervals covering $(-\infty, \infty)$, can be approximated by a sequence of continuous functions which are all 0 at a specified point of I_1, monotone away from that point, and bounded by 1, and from the Lebesgue convergence theorem, it follows that the ν^*, whose existence was established in Theorem 9, also works for all loss functions which only assume the values 0 and 1 on intervals of the form I_1, I_2, and I_3 specified above.

Now if $\lambda > 0$, we define

$$L^\lambda(\omega, a) = \begin{cases} 0 & \text{if } L(\omega, a) < \lambda, \\ 1 & \text{if } L(\omega, a) \geqq \lambda. \end{cases}$$

On account of property (ii) for $L(\omega, a)$, we see that $L^\lambda(\omega, a)$ is 0 on an interval I_1 and 1 on two intervals I_2 and I_3, all disjoint, which together cover $(-\infty, \infty)$. Thus

$$\rho_\lambda(\omega, \nu^*) \leqq \rho_\lambda(\omega, \nu).$$

But,

$$L(\omega, a) = \int_0^\infty L^\lambda(\omega, a) \, d\lambda,$$

and since the order of integration can be reversed, we have

$$\rho(\omega, \nu^*) \leqq \rho(\omega, \nu).$$

Thus we have established

THEOREM 10. *Given any decision procedure* ν, *there is a monotone decision procedure* ν^* *such that for all monotone* $L(\omega, a)$ *satisfying* (i) *and* (ii) *with a given* q, ν^* *dominates* ν.[2]

[2] This theorem was proved by a more complicated method in [5].

8. Bayes strategies for the case of an infinite number of actions. In addition to the conditions imposed on $L(\omega, a)$ in the previous section, we assume that for any two actions $a_1 < a_2$, $L(\omega, a_1) - L(\omega, a_2)$ changes sign at most once in the direction of negative to positive values and has at most one zero. For example this condition is satisfied when $L(\omega, a) = |\omega - a|^k$ $(k > 0)$. The conditions (i) and (ii) almost imply this requirement. Furthermore, we require here that $\sigma_\omega = X$ and the inequality in (2) be strict.

By Lemma 1, for any two actions,

$$\rho_a(x) - \rho_{a'}(x) = \int [L(\omega, a) - L(\omega, a')]p(x \mid \omega)\, dF(\omega) \qquad (a < a') \tag{37}$$

changes sign at most once in the direction of negative to positive values. Hence, if, for a given x_0, $\min_a \rho_a(x_0)$ is achieved for a set $A(x_0)$ with g.l.b. equal to a_0, then for $a < a_0$ and $x > x_0$, by virtue of (37) we have

$$\rho_{a_0}(x) < \rho_a(x).$$

Thus the minimum of $\rho_a(x)$ for $x > x_0$ is attained for a set of values of a with $a \geqq a_0$. Since any Bayes strategy must for x concentrate its full measure in the set $A(x)$, we deduce from this last fact that the Bayes strategy must be a monotone procedure. Thus we have shown

THEOREM 11. *If $L(\omega, a)$ satisfies* (i) *and* (ii) *of Section* 7 *and if $L(\omega, a_1) - L(\omega, a_2)$ changes sign at most once for $a_1 < a_2$ and has at most one zero, then any Bayes strategy against a distribution F is a monotone procedure.*

9. Invariance and monotone strategies. Suppose a statistical procedure satisfying the monotonicity requirements is also invariant under a group of transformations, i.e., there is a group G such that each element g in G generates three mappings g_X, g_Ω, and g_A of X, Ω, and A, respectively, into themselves satisfying the following properties:

(a) The mapping of g in G into (g_X, g_Ω, g_A) is a homomorphism.

(b) $P(g_X(S) \mid g_\Omega(\omega)) = P(S \mid \omega)$ for any measureable set S in X and ω in Ω. Of course, g_X transforms measureable sets of X into measureable sets, and conversely.

(c) $L(g_\Omega(\omega), g_A(a)) = L(\omega, a)$ for every ω in Ω and a in A.

A decision procedure $\nu(a \mid x)$ is called invariant if for a T-set in A and x in X and any g in G

$$\nu(g_A(T) \mid g_X(x)) = \nu(T \mid x).$$

To relate the monotonicity hypothesis in this paper to the invariance, we also require that all three functions $g_X(x)$, $g_\Omega(\omega)$, and $g_A(a)$ be monotone in the same direction.

The proof of the next theorem and other results connecting monotonicity and invariance will be deferred to a future paper where extensive details will be given. This last theorem is stated only for the purpose of providing here a complete discussion of the theory of statistical decision problems with monotone likelihood ratio.

THEOREM. *If a statistical decision problem with monotone likelihood ratio is invariant under a group and the loss function* $L(\omega, a)$ *satisfies the monotonicity requirements of section 7 with* $g_A[q(\omega)] = q[g_\Omega(\omega)]$ *for each* g *in* G *and* ω *in* Ω, *then the class of monotone invariant procedures is essentially complete in the class of all invariant procedures.*

REFERENCES

[1] S. G. ALLEN, "A class of minimax tests for one-sided composite hypotheses," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 295–298.

[2] MILTON SOBEL, "An essentially complete class of decision functions for certain standard sequential problems," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 319–337.

[3] SAMUEL KARLIN, "Decision theory for Polyá type distributions. Case of two actions I," to appear in *Proceedings of the Third Berkeley Symposium on Probability and Statistics*.

[4] D. BLACKWELL AND M. A. GIRSHICK, *Theory of Games and Statistical Decisions*, John Wiley & Sons, New York.

[5] HERMAN RUBIN, "A complete class of decision procedures for distributions with monotone likelihood ratio," Technical Report No. 9, Applied Mathematics and Statistics Laboratory, Stanford University, 1951, pp. 1–12.