PÓLYA-TYPE DISTRIBUTIONS, III: ADMISSIBILITY FOR MULTI-ACTION PROBLEMS

By Samuel Karlin

Stanford University

In the previous publications, [1], [2], and [3], several types of decision problems associated with the general two-action problem (e.g., generalized testing problems) have been investigated under the condition that the underlying distributions have a density $p(x, \omega)$ with x the observed value, ω the state of nature, such that $p(x, \omega)$ is Pólya-type. In this paper we continue the study of the properties of best procedures in the case of multiple-action problems and problems of estimation. This part of the investigation is concerned principally with some detailed statistical queries for the special case when $p(x, \omega)$ is Pólya-type 2 or in more common statistical terminology when $p(x, \omega)$ possesses a monotone likelihood ratio. The main problem dealt with in the present manuscript is the question of admissibility of so-called monotone procedures.

To set up a common language, we summarize the statement of the general multi-action decision problem. The n action problem is usually formulated as follows: A real random variable X is observed (usually a sufficient statistic) whose distribution $P(x, \omega)$ has the form

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

where the density $p(\xi, \omega)$ has a monotone likelihood ratio and the parameter, ω, describes the state of nature. For a fixed value of one of the arguments $P(x, \omega)$ will be assumed to be a continuously differentiable function of the other argument. Throughout our discussion we may assume that ω ranges over an interval Ω of real values (for definiteness, let $\Omega = (-\infty, \infty)$) and that μ is a completely additive measure defined over the Borel field of subsets of the real line.

It is known from the theory of distributions with a monotone likelihood ratio that the set of possible observations $X_\omega = \{x \mid p(x, \omega) > 0\}$ form an interval. We shall further assume that X_ω is independent of ω. That $p(\xi, \omega)$ has a monotone likelihood ratio (strict) means that $p(x_1, \omega_1)p(x_2, \omega_2) - p(x_1, \omega_2)p(x_2, \omega_1) \geqq 0$ (>0) for $x_1 < x_2$ and $\omega_1 < \omega_2$ with x_i belonging to X and ω_i in Ω. Most of the standard densities occurring in statistical practice possess a monotone likelihood ratio. This class of densities includes, in particular, the exponential family, the non-central t, and the non-central F. The basic property of densities with a monotone likelihood ratio useful in our analysis is its variation diminishing nature. That is, if $h(\omega)$ changes sign once from, say, non-negative to non-positive values, then

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

Received March 25, 1957; revised June 17, 1957.

changes sign at most once. If, in fact, $g(x)$ does change signs, then as x increases, the values of $g(x)$ must also change in the direction of non-negative to non-positive values.

There exist n possible actions which a statistician may take. When taking action i, the loss is assumed to be measured by the function $L(i, \omega) = L_i(\omega)$, $i = 1, \cdots, n$, where ω represents the true state.

Throughout what follows the loss functions and the densities are assumed to satisfy enough smoothness properties to insure the existence of all integrals involving these quantities, as well as to justify all differentiation arguments. The order of the operations of differentiation and integration will be reversed on several occasions in the analysis. The assumption of validity for such interchanges is not overly stringent in view of the fact that except in Sects. 5 and 6 the loss functions are step-functions and the density is continuously differentiable.

In addition, it will be assumed henceforth that the loss functions $L_i(\omega)$ satisfy appropriate monotonicity assumptions. The precise statement of this is as follows: The functions $L_i(\omega) - L_{i+1}(\omega)$ $(i = 1, 2, \cdots, n - 1)$ as functions of ω have exactly one change of sign and the sets $S_i = \{\omega \mid L_i(\omega) = \min_j L_j(\omega)\}$ are non-degenerate intervals having the additional property that

$$S_1 < S_2 < \cdots < S_n$$

where $S_i < S_{i+1}$ means that S_i lies to the left of S_{i+1} with only the boundary points as common members for two successive S_i. In the case of such a loss structure, we say that the statistical problem has a monotone preference pattern. This is to suggest that if the parameter ω were known then the various actions $1, 2, \cdots$ up to n are preferred respectively for increasing values of the state of nature ω, a given action i being favored for known ω if and only if $L_i(\omega) < L_j(\omega)$ for every $j \neq i$. The fact that each of the sets S_i is a non-degenerate interval implies the existence of ω_i^0, $i = 0, 1, \cdots, n$, where $\omega_0^0 = -\infty$ and $\omega_n^0 = +\infty$, such that $\omega_i^0 < \omega_{i+1}^0$ and action i is definitely preferred for $\omega_{i-1}^0 < \omega < \omega_i^0$. The values ω_i^0 are necessarily the unique change points of $L_{i+1}(\omega) - L_i(\omega)$. For simplicity of exposition we have chosen to make the change points ω_i^0 distinct although the reader may supply appropriate modifications to the argument to extend our studies to the case when some of the ω_i^0 coincide. (See [1] and [2].)

A randomized decision procedure φ for the statistician is described by an n-tuple of functions

$$\varphi = (\varphi_1(x), \varphi_2(x), \cdots, \varphi_n(x)),$$

$\varphi_i(x) \geqq 0$ and $\sum \varphi_i(x) = 1$, where $\varphi_i(x)$ is interpreted as the probability of taking action i when x is observed. The expected risk becomes

$$\rho(\omega, \varphi) = \int p(x, \omega) \left(\sum_{i=1}^{n} L_i(\omega) \varphi_i(x) \right) d\mu(x).$$

A procedure φ is said to be admissible when there exists no other procedure φ^* such that $\rho(\omega, \varphi^*) \leqq \rho(\omega, \varphi)$ for every ω with inequality for at least one value of ω. In other words, a procedure is admissible if it cannot be improved upon in terms of expected risk—independent of the state of nature. Admissibility for a given decision procedure is an obvious prerequisite for its use. Hence, it is of some significance to be able to characterize all admissible procedures. A procedure φ^0 is said to be Bayes with respect to a distribution $F(\omega)$ (referred to as the a priori distribution of the state of nature) if

$$\rho(F, \varphi^0) = \int \rho(\omega, \varphi^0)\, dF(\omega) = \min_{\varphi} \int \rho(\omega, \varphi)\, dF(\omega).$$

It is readily seen that if φ^0 is unique Bayes with respect to a distribution $F(\omega)$ (i.e., φ^0 is the only procedure minimizing $\rho(F, \varphi)$), then φ^0 is admissible. Consequently, one method of establishing that a given procedure is admissible is to show that it is unique Bayes with respect to some distribution $F(\omega)$. In numerous cases, we shall actually verify this property.

In the n action problem a procedure $\varphi = (\varphi_1, \varphi_2, \cdots, \varphi_n)$ is said to be monotone if there exist critical numbers, $x_0 \leqq x_1 \leqq x_2 \leqq \cdots \leqq x_{n-1} \leqq x_n$ ($x_0 = -\infty$, $x_n = +\infty$) such that

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

and randomization may occur at $x = x_i, x_{i+1}$, i.e., $\varphi_i(x_i) = \lambda_i$ and $\varphi_{i+1}(x_i) = 1 - \lambda_i$ ($0 \leqq \lambda_i \leqq 1$). A monotone procedure is therefore fully specified by (x_1, $x_2, \cdots, x_{n-1}; \lambda_1, \cdots, \lambda_{n-1}$) provided $x_i \leqq x_{i+1}$ and $0 \leqq \lambda_i \leqq 1$ with appropriate modifications on the restrictions for λ_i when allowing for $x_i = x_{i+1}$. It was shown in [2] that the monotone procedures form a complete class. Moreover, the proof of completeness of the preceding reference contained an explicit construction which shows how to improve by a monotone procedure any specified non-monotone procedure. However, the question of determining when monotone procedures are admissible is of greater complexity. In the two-action problem, it was shown in [1] under almost negligible restrictions that all monotone procedures are admissible. In direct contrast, for the case of a general three-action problem the characteristic of admissibility ceases to be a property shared by all monotone decision procedures. For a counterexample see [2]. Apparently the explicit magnitudes of the loss functions and not only the preference regions have a direct influence on whether a procedure is admissible or not. Nevertheless, it is possible to characterize a wide class of multi-action monotone decision problems for which all monotone procedures are admissible. Consider a collection of loss functions $L_i(\omega)$ satisfying

$$\text{(I)} \qquad |L_i(\omega) - L_{i+1}(\omega)| = b_{ij} \text{ for } \omega \text{ in } S_j$$

($i = 1, 2, \cdots, n - 1$ and $j = 1, \cdots, n$) such that $b_{ij} \geqq 0$ for every i and j, $b_{ik} > 0$ for $k \geqq i$, and for $i = 1, 2, \cdots, n - 1$,

(II) $$\begin{vmatrix} b_{ij} & b_{ik} \\ b_{i+1,j} & b_{i+1,k} \end{vmatrix} \geqq 0$$

whenever $1 \leqq j \leqq i$ and $i + 1 \leqq k \leqq n$. (For instance, if $b_{ij} \equiv c > 0$ then (II) is certainly satisfied.) We will show later that for monotone loss functions satisfying these conditions all non-degenerate monotone procedures are admissible. In fact, all non-degenerate monotone procedures are found to be Bayes with respect to suitable finite distributions.

By allowing the number of possible actions to become infinite, our multi-action problem approaches an estimation problem. That is, the estimation problem may be viewed both formally and practically as a limit of finite action problems. Therefore, aside from interest in itself, the n action decision problem also suggests and leads to consequences about estimation problems. In the case of estimation a non-randomized procedure is described by a mapping $a(x)$ of the observed value x into the space of actions. The loss function $L(a, \omega)$ is now a function of the action a taken and the state of nature ω. For example, $(a - \omega)^2$ would correspond to square error where a represents the estimate of ω used. Similarly, $|a - \omega|$ is the commonly used loss function measuring absolute error.

The case where

$$L_i(\omega) - L_{i+1}(\omega) = \begin{cases} -a & \text{if } \omega \text{ is in } S_j\,, j \leqq i \\ +a & \text{if } \omega \text{ is in } S_j\,, j \geqq i + 1, \end{cases}$$

or specifically $L_i(\omega) = a|\,j - i\,|$ for ω in S_j, may be considered as the discrete analog for n actions of the absolute error loss function. This last example satisfies the conditions of (II) so that all non-degenerate monotone procedures are admissible. In contrast, discrete analogs of the square error loss function do not satisfy (II), and it is unknown whether all monotone procedures are admissible. In Sec. 2, we shall investigate in detail the discrete absolute error loss functions. With the aid of suitable limiting arguments admissibility for some statistical procedures of the estimation problem with absolute error loss function will be presented in Sec. 4.

Also, in Sec. 3 we analyze the general monotone loss functions satisfying properties (I) and (II). In Sec. 5 we investigate a loss function for which the penalties are constant when we underestimate and a loss given by a monotone increasing function of the extent of overestimation when we overestimate. Minimal complete classes of statistical procedures may be fully determined. All monotone strategies are in this case admissible. The reader should bear in mind that these results are in sharp contrast to the general n action problem where all monotone procedures are not necessarily admissible.

Within a complete class the statistician obviously should not choose an inadmissible procedure. The principal result of this paper is the validation of the fact that at least for loss functions satisfying (I) and (II) the non-degenerate monotone strategies are admissible. The remaining deficiency of this theory is that in the n-action problem the class of monotone strategies represents an $n - 1$ parameter family of procedures. Thus the task of choosing a single procedure is still bewildering and some cogent principles are needed to reduce the size of the class.

In the following article [4], we advance some principles to guide the statistician in selecting a specific monotone strategy from the essentially complete class of all monotone procedures.

The study of admissibility concerning the square error loss function will be presented in a later publication.

Finally, I wish to express my indebtedness to Mr. Rupert Miller for his help in the preparation of this manuscript.

1. Some preliminary lemmas. Basic to our study of the question of admissibility for loss functions satisfying the conditions of (I) and (II) are the following propositions concerning solutions of systems of linear equations of a specific form which have special properties. These linear equation results are singled out here because of their independent mathematical interest. The reader interested only in their statistical relevance may on first reading pass over their proofs.

LEMMA 1. *The system of n homogeneous equations in $n + 1$ unknowns, $n \geqq 2$,*

$$(*) \qquad \sum_{j=1}^{i} a_{ij}x_j - \sum_{j=i+1}^{n+1} a_{ij}x_j = 0, \qquad i = 1, \cdots, n,$$

where A, the coefficient matrix of size $(n \times n + 1)$, satisfies the following properties:

(i) $a_{ij} \geqq 0$ *for all* i, j; $a_{ii} > 0$ *for* $i = 1, \cdots, n$

(ii) *For* $i = 1, 2, \cdots, n - 1$,

$$\begin{vmatrix} a_{ij} & a_{ik} \\ a_{i+1,j} & a_{i+1,k} \end{vmatrix} \geqq 0$$

for $1 \leqq j \leqq i$ *and* $i + 1 \leqq k \leqq n + 1$ *with strict inequality for some* j *for each* k,

has a unique (except for a multiplicative constant) solution $x^0 = (x_1^0, \cdots, x_{n+1}^0)$ *which has in addition the following properties:*

(a) $x_j^0 \neq 0, j = 1, \cdots, n + 1$

(b) $\operatorname{sgn} x_1^0 = \operatorname{sgn} x_2^0 = \cdots = \operatorname{sgn} x_{n+1}^0$.

An equivalent formulation of Lemma 1 in terms of non-homogeneous linear equations is as follows:

LEMMA 1a. *The system of n non-homogeneous equations in n unknowns*

$$\sum_{j=1}^{i} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}x_j = a_{i,n+1},$$

where the $n \times (n + 1)$ *matrix* $A = (a_{ij})$ *satisfies properties* (i) *and* (ii) *of Lemma* 1, *has a unique solution* $x^0 = (x_1^0, \cdots, x_n^0)$ *which has in addition the property that* $x_i^0 > 0, i = 1, \cdots, n$.

The proof of the equivalence of Lemma 1 and Lemma 1a is straightforward and will be omitted.

Proof of Lemma 1a (*by induction*). Suppose the result is true for $n - 1$ non-homogeneous equations in $n - 1$ unknowns. We prove it is also true for n non-homogeneous equations in n unknowns.

Consider the first $n - 1$ equations of the system of n equations. They can be written as

$$\sum_{j=1}^{i} a_{ij}\, x_j - \sum_{j=i+1}^{n-1} a_{ij}\, x_j - a_{in}\, \alpha - a_{i,n+1} = 0 \qquad i = 1, \cdots, n-1 \tag{1}$$

where $x_n = \alpha$, or as

$$\sum_{j=1}^{i} a'_{ij}\, x_j - \sum_{j=i+1}^{n-1} a'_{ij}\, x_j - a'_{in} = 0 \qquad i = 1, \cdots, n-1 \tag{2}$$

where $a'_{ij} = a_{ij}$ for $j = 1, 2, \cdots, n-1$, and $a'_{in} = a_{in}\alpha + a_{i,n+1}$. It is readily verified that the matrix $A' = (a'_{ij})$ satisfies property (ii) provided $\alpha \geqq 0$. Property (i) is also satisfied so by Lemma 1a for the case of $n - 1$ equations in $n - 1$ unknowns there exists a unique solution $x(\alpha) = (x_1(\alpha), \cdots, x_{n-1}(\alpha))$ to (1) for each $\alpha \geqq 0$ and $x_i(\alpha) > 0$, $i = 1, \cdots, n-1$.

Let $g(x(\alpha)) = a_{n1}x_1(\alpha) + \cdots + a_{n,n-1}x_{n-1}(\alpha) + a_{nn}\alpha - a_{n,n+1}$. For $\alpha > a_{n,n+1}/a_{nn}$, $g(x(\alpha)) > 0$ since $x_i(\alpha) > 0$, $i = 1, \cdots, n-1$. We assert that for $\alpha = 0$, $g(x(0))$ is <0. Suppose the contrary; i.e., suppose $g(x(0)) \geqq 0$. If the equation

$$a_{n-1,1}x_1(0) + \cdots + a_{n-1,n-1}x_{n-1}(0) - a_{n-1,n+1} = 0$$

is multiplied by $a_{n,n+1}$, the equation

$$a_{n1}x_1(0) + \cdots + a_{n,n-1}x_{n-1}(0) - a_{n,n+1} \geqq 0$$

is multiplied by $-a_{n-1,n+1}$, and the two equations are then added, the result is

$$\begin{vmatrix} a_{n-1,1} & a_{n-1,n+1} \\ a_{n1} & a_{n,n+1} \end{vmatrix} x_1(0) + \cdots + \begin{vmatrix} a_{n-1,n-1} & a_{n-1,n+1} \\ a_{n,n-1} & a_{n,n+1} \end{vmatrix} x_{n-1}(0) \leqq 0.$$

But each determinant is non-negative and at least one is strictly positive with $x_i(0) > 0$, $i = 1, \cdots, n-1$. This leads to an obvious contradiction. Therefore, $g(x(0)) < 0$. Since $g(x(\alpha))$ is a continuous function of α, there must exist an $\alpha_0 > 0$ such that

$$a_{n1}x_1(\alpha_0) + \cdots + a_{n,n-1}x_{n-1}(\alpha_0) + a_{nn}\alpha_0 - a_{n,n+1} = 0.$$

Consequently, one solution to the system of n non-homogeneous equations in the n unknowns which has the property that $x_i^0 > 0$, $i = 1, \cdots, n$, is $x^0 = (x_1(\alpha_0), \cdots, x_{n-1}(\alpha_0), \alpha_0)$.

To complete the induction proof of the existence of a positive solution we must verify that the lemma holds for the case $n = 2$. This task is reduced to routine enumeration of cases with direct use of the hypothesis.

It remains to establish the uniqueness of the solution. Suppose there exist two solutions $x^0 = (x_1^0, \cdots, x_n^0)$ and $y^0 = (y_1^0, \cdots, y_n^0)$ such that $x^0 \neq y^0$. Then

$$\sum_{j=1}^{i} a_{ij}\, z_j - \sum_{j=i+1}^{n} a_{ij}\, z_j = 0, \qquad i = 1, \cdots, n \tag{3}$$

for $z_i = x_i^0 - y_i^0$, $i = 1, \cdots, n$. Consider the first $n - 1$ equations. By the induction hypothesis the solution to this system of $n - 1$ equations in n unknowns is unique (except for a multiplicative constant) and possesses properties a and b. Since $z_i = c(x_i^0 - y_i^0)$ for $c \neq 0$ is the family of solutions, without loss of generality it can be assumed that $x_i^0 - y_i^0 > 0$, $i = 1, \cdots, n$. But for the nth equation this yields

$$a_{n1}(x_1^0 - y_1^0) + \cdots + a_{nr}(x_n^0 - y_n^0) > 0,$$

which contradicts (3). That uniqueness holds for the case $n = 2$ is easily checked. Thus, $x_i^0 \equiv y_i^0$.

This lemma can be expressed in terms of appropriate subdeterminants as follows:

COROLLARY 1. *The signs of the $n + 1$ $n \times n$ subdeterminants of the coefficient matrix A of the system of equations* (*) *obtained by deleting successive columns must alternate.*

LEMMA 2. *The system of $n + 1$ equations in n unknowns*

$$\sum_{i=1}^{j-1} a_{ij} y_i - \sum_{i=j}^{n} a_{ij} y_i = c_j, \qquad j = 1, \cdots, n+1, \tag{4}$$

where

(i) $a_{ij} \geqq 0$ *all* i, j; $a_{ii} > 0$ *for* $i = 1, \cdots, n$
(ii) *For* $j = 2, \cdots, n - 1$,

$$\begin{vmatrix} a_{kj} & a_{k,j+1} \\ a_{ij} & a_{i,j+1} \end{vmatrix} \geqq 0$$

for $j + 1 \leqq i \leqq n$, $1 \leqq k \leqq j$ *with strict inequality for some* i *for each* k.

(iii) $A = (a_{ij})$ *satisfies condition* (ii) *of Lemma* 1.
(iv) $c_j \geqq 0, j = 1, \cdots, n + 1$

has a solution only if $c_j = 0, j = 1, \cdots, n + 1$, *and in this case the only solution is the trivial one* $y_i = 0$, $i = 1, \cdots, n$.

PROOF. Suppose $c_j > 0$ for some j and there exists a solution $y^0 = (y_1^0, \cdots, y_n^0)$ to the system of equations (4), which can be written in matrix notation as $yA^* = c$. $c = (c_1, \cdots, c_{n+1})$ and $A^* = (a_{ij}^*)$ where $a_{ij}^* = a_{ij}$, $i < j$, and $a_{ij}^* = -a_{ij}$, $i \geqq j$. Since the conditions of Lemma 1 are satisfied for the system of equations $-A^*x = 0$, there exists a solution x^0, all of whose components are positive. Since $y^0A^* = c$ and $A^*x^0 = 0$,

$$0 < (c, x^0) = (y^0A^*, x^0) = (y^0, A^*x^0) = (y^0, 0) = 0,$$

where (α, β) denotes the inner product of the vectors α and β, which is a contradiction. Therefore, $c_j = 0$ for $j = 1, \cdots, n + 1$.

If we omit the first and last columns of the matrix A^* and consider the system of equations

$$\begin{array}{llllll} a_{nn}\, y_n & - a_{n-1,n}\, y_{n-1} & \cdots & -a_{2n}\, y_2 & - a_{1n}\, y_1 & = 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ a_{n2}\, y_n & + a_{n-1.2}\, y_{n-1} & \cdots & +a_{22}\, y_2 & - a_{12}\, y_1 & = 0 \end{array}$$

properties (i) and (ii) reduce to the conditions of Lemma 1 for this system of equations, so that any solution $y^0 = (y_1^0, \cdots, y_n^0)$ to (4) has its components of the same sign and unequal to zero. But this is impossible since

$$a_{n1}y_n + a_{n-1,1}y_{n-1} + \cdots + a_{11}y_1$$

will be unequal to zero if $y_i^0 > 0$ for all i or $y_i^0 < 0$ for all i. Thus, $y_i^0 = 0$ for all i.

In closing this section, we remark that an alternative, more complicated proof of Lemma 1 was given independently by P. Braumann.

2. Discrete absolute error loss functions. In the n-action problem under consideration in this section, the loss functions $L_i(\omega)$, $i = 1, \cdots, n$, will have the following form: There exist $n - 1$ values $\omega_1^0, \cdots, \omega_{n-1}^0$ such that $L_i(\omega) = c\,|j - i|$ for ω in the interval $(\omega_{j-1}^0, \omega_j^0]$, $j = 1, \cdots, n$. (By definition $\omega_0^0 = -\infty$ and $\omega_n^0 = +\infty$.) This system of loss functions may be viewed as the discrete analog of the absolute error loss function in estimation problems. The loss for any action is proportional to the distance from the best action. As n, the number of actions, tends to $+\infty$ and $|\omega_i^0 - \omega_{i-1}^0| \to 0$ suitably as $n \to +\infty$ for all i, then $L_{i_a}(\omega) \to c\,|a - \omega|$, where i_a is defined by $a \in (\omega_{i_a-1}^0, \omega_{i_a}^0]$. Thus the absolute error loss function is an appropriate limit of discrete absolute error loss functions.

The real random variable X will be assumed to be distributed according to

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

which depends on the real parameter ω. $p(\xi, \omega)$ will be assumed to possess a strict monotone likelihood ratio. The requirement of strictness may be relaxed in many cases, but to avoid inessential, tedious details we have preferred to impose the slightly stronger assumption of strictness.

For the type of loss function described above, it will be shown in this section that any non-degenerate monotone procedure characterized by $n - 1$ points $x_1^0, x_2^0, \cdots, x_{n-1}^0$ and $n - 1$ probabilities $\lambda_1^0, \cdots, \lambda_{n-1}^0$ is admissible. As was pointed out in the introduction, this is not true for general loss functions when $n \geqq 3$.

LEMMA 3. *Any non-degenerate monotone procedure is Bayes against a discrete a priori distribution F^* which concentrates all its probability at n points; each interval $(\omega_{i-1}^0, \omega_i^0]$, $i = 1, \cdots, n$, contains a mass-point of F^*, but the location of the mass-point in the interval is arbitrary. The non-degenerate monotone procedure is uniquely Bayes with respect to F^* except for the randomizations $\lambda_1^0, \cdots, \lambda_{n-1}^0$.*

PROOF. For any observed x, the a posteriori risk of taking action i, with respect to the a priori distribution F, is

$$r_x(i) = K \int_{-\infty}^{\infty} L_i(\omega) p(x, \omega)\, dF(\omega),$$

where $K^{-1} = \int_{-\infty}^{\infty} p(x, \omega)\, dF(\omega)$. Action i would be preferred to action $i + 1$ for those values of x for which

$$(5) \qquad \tau_x(i + 1) - \tau_x(i) = K \int_{-\infty}^{\infty} [L_{i+1}(\omega) - L_i(\omega)] p(x, \omega)\, dF(\omega)$$

is > 0, and $i + 1$ would be preferred to i when $\tau_x(i + 1) - \tau_x(i) < 0$. But $L_{i+1}(\omega) - L_i(\omega)$ changes sign exactly once and changes from positive to negative as ω increases. Since $p(x, \omega)$ is strictly Pólya-type 2, by Theorem 3 of [3] $\tau_x(i + 1) - \tau_x(i)$ has at most one change of sign and at most one zero counting multiplicities. Furthermore, in the event that $\tau_x(i + 1) - \tau_x(i)$ does change sign, then it must change from positive to negative as x increases [1].

Let ω_i be any arbitrary point in the interval $(\omega_{i-1}^0, \omega_i^0]$, $i = 1, \cdots, n$. It is a consequence of Lemma 1 that n constants $\xi_1, \cdots, \xi_n$ with $\xi_i > 0$, $\sum_1^n \xi_i = 1$, can be chosen such that the monotone procedure described by $(x_1^0, \cdots, x_{n-1}^0;\ \lambda_1^0, \cdots, \lambda_{n-1}^0)$ is Bayes against the distribution F^* which concentrates probability ξ_j at the point ω_j:

$$\int_{-\infty}^{\infty} [L_{i+1}(\omega) - L_i(\omega)] p(x, \omega)\, dF^* = c \left\{ \sum_{j=1}^{i} p(x, \omega_j)\xi_j - \sum_{j=i+1}^{n} p(x, \omega_j)\xi_j \right\}.$$

Indeed, consider the system of equations

$$\sum_{j=1}^{i} p(x_i^0, \omega_j) f_j - \sum_{j=i+1}^{n} p(x_i^0, \omega_j) f_j = 0, \quad i = 1, \cdots, n - 1.$$

Since $p(x, \omega)$ has a strict monotone likelihood ratio, the conditions of Lemma 1 are satisfied so there exists a solution to the above system of equations such that $\xi_i > 0$ for all i and $\sum_1^n \xi_i = 1$. Thus

$$(6) \qquad \int_{-\infty}^{\infty} [L_{i+1}(\omega) - L_i(\omega)] p(x, \omega)\, dF^*(\omega)$$

has a zero, its only zero, at the point x_i^0. Consequently, action i is preferred for $x < x_i^0$ and action $i + 1$ is preferred for $x > x_i^0$. Similarly, action $i + 1$ is preferred to action $i + 2$ for $x < x_{i+1}^0$. Since $x_i^0 < x_{i+1}^0$, action i is preferred to action $i + 2$ for $x < x_i^0$. Repetition of this argument shows that i is preferred to all $j > i$ for $x < x_i^0$. A similar argument shows that i is preferred to all actions $j < i$ for $x > x_{i-1}^0$. Thus action i is the best action for $x \in (x_{i-1}^0, x_i^0)$. At $x = x_i^0$ it is immaterial whether action i or $i + 1$ is taken since $\tau_i(x_i^0) = \tau_{i+1}(x_i^0)$. All randomizations between i and $i + 1$ at $x = x_i^0$ will produce the same overall risk. This implies that the monotone procedure $(x_1^0, \cdots, x_{n-1}^0;\ \lambda_1^0, \cdots, \lambda_{n-1}^0)$ is Bayes against F^*. Furthermore, it is the unique Bayes strategy against F^* (except for randomization allowed for action i and $i + 1$ at the points x_i^0) since x_i^0 is the unique zero of (6).

The significant result of this section which is deduced from Lemma 3 is as follows:

THEOREM 1. *All non-degenerate monotone procedures are admissible.*

PROOF. Suppose φ^0 is not admissible. Then there exists a decision procedure $\varphi^* = (\varphi_1^*, \cdots, \varphi_n^*)$ such that $\rho(\omega, \varphi^*) \leqq \rho(\omega, \varphi^0)$ for all ω with strict inequality for some ω'. Suppose ω' falls in the interval $(\omega_{h-1}^0, \omega_h^0]$. Select $n - 1$ specific parameter points of Ω so that $(\omega_1, \cdots, \omega_{h-1}, \omega', \omega_{h+1}, \cdots, \omega_n)$ satisfies $\omega_i \, \varepsilon \, (\omega_{i-1}^0, \omega_i^0]$, $i = 1, \cdots, h - 1, h + 1, \cdots, n$. Then, by Lemma 3 a discrete probability distribution F^* can be constructed which has positive probability at each of these points and only at these points, and against which φ^0 is Bayes. But since $\rho(\omega_i, \varphi^*) \leqq \rho(\omega_i, \varphi^0)$ for $i = 1, \cdots, h - 1, h + 1, \cdots, n$ and

$$\rho(\omega', \varphi^*) < \rho(\omega', \varphi^0),$$

it follows that

$$\int \rho(\omega, \varphi^*) \, dF^*(\omega) < \int \rho(\omega, \varphi^0) \, dF^*(\omega),$$

which contradicts the fact that φ^0 is Bayes against F^*. Therefore, φ^0 must be admissible.

The following lemma strengthens slightly the results of Theorem 1.

LEMMA 4. *If φ^* and φ^0 are two non-degenerate monotone procedures and $\rho(\omega, \varphi^*) \equiv \rho(\omega, \varphi^0)$, then $x_i^* = x_i^0$, $i = 1, \cdots, n - 1$, and $\lambda_i^* = \lambda_i^0$ for all i such that $\mu(x_i^0) > 0$.*

PROOF. Since monotone procedures are uniquely Bayes except for randomization at the endpoints of the intervals, $\rho(\omega, \varphi^*) \equiv \rho(\omega, \varphi^0)$ evidently implies that $x_i^* = x_i^0$, $i = 1, \cdots, n - 1$. It can be easily verified (cf. Theorem 1 of [2]) that

$$\begin{aligned}
\rho(\omega, \varphi^0) - \rho(\omega, \varphi^*) &= \int_{-\infty}^{\infty} p(x, \omega) \left\{ \sum_{i=1}^{n-1} [L_i(\omega) - L_{i+1}(\omega)] \left[\sum_{j=1}^{i} \varphi_j^0(x) - \sum_{j=1}^{i} \varphi_j^*(x) \right] \right\} d\mu(x) \\
&= \sum_{i=1}^{n-1} [L_i(\omega) - L_{i+1}(\omega)] p(x_i^0, \omega)(\lambda_i^0 - \lambda_i^*)\mu(x_i^0) \\
&= \sum_{i=1}^{n-1} [L_i(\omega) - L_{i+1}(\omega)] p(x_i^0, \omega)\eta_i
\end{aligned}$$

where $\eta_i = (\lambda_i^0 - \lambda_i^*)\mu(x_i^0)$.

Evaluation of $\rho(\omega, \varphi^0) - \rho(\omega, \varphi^*)$ at n points $\omega_1, \cdots, \omega_n$ which satisfy

$$\omega_i \, \varepsilon \, (\omega_{i-1}^0, \omega_i^0]$$

yields the system of equations

$$c \left\{ \sum_{i=1}^{j-1} p(x_i^0, \omega_j)\eta_i - \sum_{i=j}^{n-1} p(x_i^0, \omega_j)\eta_i \right\} = 0, \qquad j = 1, \cdots, n.$$

Since $p(x, \omega)$ has a strict monotone likelihood ratio, the conditions of Lemma 2 are satisfied, and therefore $\eta_i = 0$, $i = 1, \cdots, n - 1$. Hence $\lambda_i^* = \lambda_i^0$ whenever $\mu(x_i^0) > 0$.

A monotone procedure is said to be degenerate if $x_i^0 = x_{i+1}^0$ for some i. Several intervals can be missing as well, in which case $x_i^0 = x_{i+1}^0 = \cdots = x_{i+k}^0$ for the appropriate combinations of i and k. Theorem 1 does not extend to the case of degenerate procedures; i.e., there exist inadmissible degenerate monotone procedures. However, the degenerate monotone procedures do possess analogous Bayes properties.

LEMMA 5. *If the a priori distribution F concentrates no measure in the interval $(\omega_{i-1}^0, \omega_i^0]$, then $x_{i-1}^0 = x_i^0$ in the (monotone) Bayes strategy with respect to F. That is, action i is never taken except possibly at the point x_i^0 where randomization occurs.*

PROOF. Since $L_i(\omega) - L_{i+1}(\omega) = L_{i-1}(\omega) - L_i(\omega)$ for $\omega \not\in (\omega_{i-1}^0, \omega_i^0]$,

$$\int_{-\infty}^{\infty} [L_i(\omega) - L_{i+1}(\omega)]p(x, \omega)\, dF(\omega) = \int_{-\infty}^{\infty} [L_{i-1}(\omega) - L_i(\omega)]p(x, \omega)\, dF(\omega).$$

But since these two integrals are identically equal they must have the same zero (if one exists) at some point x'. Action i is preferred to action $i + 1$ for $x < x'$ and reciprocally for $x > x'$. Action $i - 1$ is preferred to action i for $x < x'$ and vice versa for $x > x'$. Combining these two facts, it is seen that action $i - 1$ is preferred to action $i + 1$ for $x < x'$ and vice versa for $x > x'$. Action i is preferred nowhere except possibly at the point x'. Thus, $x_{i-1}^0 = x_i^0 = x'$.

The analog of Lemma 3 for degenerate monotone procedures is as follows:

LEMMA 3a. *Any monotone procedure φ^0 with k degenerate intervals, $1 \leqq k \leqq n - 1$, is Bayes with respect to a discrete a priori distribution F^* which concentrates all its probability at $n - k$ points; each interval $(\omega_{i-1}^0, \omega_i^0]$ corresponding to a non-degenerate action interval in the X-space contains a mass-point of F^*, but the location of the mass-point in the interval is arbitrary. The monotone procedure is uniquely Bayes against F^* except for the determination of the randomizations $\lambda_1^0, \cdots, \lambda_{n-1}^0$.*

The proof of this is more elaborate. Nevertheless, since the techniques are similar to the preceding, we omit the details.

As mentioned previously, there exist inadmissible degenerate monotone procedures. An example, for which the author is indebted to the referee, can be constructed as follows: Let $n = 3$, and consider the strategy $(x_1^0, x_2^0; \lambda_1^0, \lambda_2^0)$ defined by $x_1^0 = x_2^0 = x^0$, $\lambda_1^0 = \frac{1}{2}$, $\lambda_2^0 = 0$. If $\mu(x^0) > 0$, then the strategy

$$(x_1^*, x_2^*; \lambda_1^*, \lambda_2^*)$$

with $x_1^* = x_2^* = x^0$, $\lambda_1^* = 0$, $\lambda_2^* = 1$ constitutes an improvement.

The following theorem describes some of the properties of degenerate monotone procedures.

THEOREM 1a. *Let φ^0 be a degenerate monotone procedure for which*

$$x_1^0 < \cdots < x_{i_0}^0 = x_{i_0+1}^0 = \cdots = x_{i_0+k}^0 < \cdots < x_{n-1}^0$$

(1) *If $\mu(x_{i_0}^0) = 0$, φ^0 is admissible.*

(2) *If $\mu(x_{i_0}^0) > 0$ and φ^* satisfies $\rho(\omega, \varphi^*) \leqq \rho(\omega, \varphi^0)$ for all ω, then φ^* is a degenerate monotone procedure with the following three properties:*

(a) $x_i^* = x_i^0, \quad i = 1, \cdots, n - 1;$

(b) $\lambda_i^0 = \lambda_i^*$ *if* $\mu(x_i^0) > 0, \quad i = 1, \cdots, i_0 - 1, i_0 + k + 1, \cdots, n - 1;$

(c) $(\lambda_{i_0+k}^0 - \lambda_{i_0+k}^*) + 2(\lambda_{i_0+k-1}^0 - \lambda_{i_0+k-1}^*) + \cdots + k(\lambda_{i_0}^0 - \lambda_{i_0}^*) = 0.$

PROOF. (1) and (2a) are trivial consequences of Lemma 3a.

An easy computation verifies (cf. Theorem 1 of [2]) that

$$0 \leqq \rho(\omega, \varphi^0) - \rho(\omega, \varphi^*) = \sum_{i=1}^{n-1} [L_i(\omega) - L_{i+1}(\omega)] p(x_i^0, \omega)\eta_i',$$

where

$$\eta_i' = \sum_{K_i} (\lambda_j^0 - \lambda_j^*)\mu(x_i^0), \qquad K_i = \{j \mid j \leqq i, x_j^0 - x_i^0\}.$$

Evaluation of the last expression at $\omega_1, \cdots, \omega_{i_0}, \omega_{i_0+k+1}, \cdots, \omega_n$, $\omega_i \in (\omega_{i-1}^0, \omega_i^0]$, $i = 1, \quad i_0, i_0 + k + 1, \cdots, n$, produces the system of equations

$$c\left\{\sum_{i=1}^{j-1} p(x_i^0, \omega_j)\eta_i'' - \sum_{i=j}^{i_0} p(x_i^0, \omega_j)\eta_i'' - \sum_{i=i_0+k+1}^{n-1} p(x_i^0, \omega_j)\eta_i''\right\} \geqq 0,$$

$$j = 1, \cdots, i_0,$$

$$c\left\{\sum_{i=1}^{i_0} p(x_i^0, \omega_j)\eta_i'' + \sum_{i=i_0+k+1}^{j-1} p(x_i^0, \omega_j)\eta_i'' - \sum_{i=j}^{n-1} p(x_i^0, \omega_j)\eta_i''\right\} \geqq 0,$$

$$j = i_0 + k + 1, \cdots, n,$$

$$\eta_i'' = \eta_i' \qquad i = 1, \cdots, i_0 - 1, i_0 + k + 1, \cdots, n - 1,$$

and

$$\eta_{i_0}'' = \sum_{\nu=i_0}^{i_0+k} \eta_\nu' = (\lambda_{i_0+k}^0 - \lambda_{i_0+k}^*) + 2(\lambda_{i_0+k-1}^0 - \lambda_{i_0+k-1}^*) + \cdots + k(\lambda_{i_0}^0 - \lambda_{i_0}^*).$$

The conditions of Lemma 2 are satisfied so

$$\eta_i'' = 0, \qquad i = 1, \cdots, i_0, i_0 + k + 1, \cdots, n - 1.$$

This implies properties (2b) and (2c).

The statement of the analogous results when there are several groups of degenerate intervals is left to the reader.

It is worthwhile giving an explicit statement to the following general result.

COROLLARY 1. *If the measure μ is atomless, all monotone procedures are admissible.*

3. Some extensions. The results of the previous section concerning admissibility for multi-action problems extend immediately to a more general type of loss function, as indicated in the introduction. Specifically, we suppose that

(I) $$L_i(\omega) = c_{ij} \quad \text{for} \quad \omega \quad \text{in} \quad S_j;$$

that is, the loss is constant when taking action i instead of action j, which was preferred. Aside from the usual requirement that the $L_i(\omega)$ give rise to a monotone preference pattern, the further important assumption is that

$$|L_i(\omega) - L_{i+1}(\omega)| = b_{ij} \geqq 0 \text{ for } \omega \text{ in } S_j, \qquad b_{ik} > 0, \quad k \geqq i,$$

such that for $i = 1, 2, \cdots, n - 1$,

$$\text{(II)} \qquad \begin{vmatrix} b_{ij} & b_{ik} \\ b_{i+1,j} & b_{i+1,k} \end{vmatrix} \geqq 0$$

whenever $1 \leqq j \leqq i$ and $i + 1 \leqq k \leqq n$.

Again we asume that the density $p(x, \omega)$ possesses a strict monotone likelihood ratio.

The place of the monotone procedures for these loss functions is summarized in the following propositions.

LEMMA 6. *If the loss functions satisfy conditions* (I) *and* (II), *any non-degenerate monotone procedure is Bayes against a discrete a priori distribution F^* which concentrates all its probability at n points; each interval $(\omega_{i-1}, \omega_i^0]$, $i = 1, \cdots, n$, contains a mass-point of F^*, but the location of the mass-point in the interval is arbitrary. The non-degenerate monotone procedure is uniquely Bayes with respect to F^* except for the randomizations $\lambda_1^0, \cdots, \lambda_{n-1}^0$.*

The proof of this lemma is completely analogous to that of Lemma 3. We sketch the argument. Selecting n points ω_j, $j = 1, \cdots, n$, where $\omega_j \in (\omega_{j-1}^0, \omega_j^0]$, we seek to determine a discrete distribution $F^*(\omega)$ with weights located exclusively at ω_j such that

$$(7) \quad \int [L_{i+1}(\omega) - L_i(\omega)]p(x, \omega)\, dF^*(\omega) = \left\{\sum_{j=1}^{i} p(x, \omega_j)b_{ij}\xi_j - \sum_{j=i+1}^{n} p(x, \omega_j)b_{ij}\xi_j\right\}$$

vanishes only at x_i^0, $i = 1, \cdots, n - 1$, the critical values describing the specified non-degenerate monotone procedure. The system of equations (7) is exactly of the form of (*) of Sec. 1 with

$$a_{ij} = p(x_i, \omega_j)b_{ij}.$$

The hypotheses (I) and (II) and the fact that $p(x, \omega)$ possesses a strict monotone likelihood ratio immediately imply that the conditions of Lemma 1 are fulfilled. Consequently, we may conclude that an $F^*(\omega)$ with the desired properties exists. From here on the proof is a paraphrase of that of Lemma 3.

Paralleling the method of obtaining Theorem 1 from Lemma 3, we deduce from Lemma 6:

THEOREM 2. *If the loss functions satisfy conditions* (I) *and* (II), *all non-degenerate monotone procedures are admissible.*

The arguments of Lemma 3a dealing with degenerate monotone procedures do not extend directly to this more general loss function. In fact, for this case the conditions of II will be strengthened so that $b_{ij} > 0$, all i, j, and

$$\text{(II}'\text{)} \qquad B_{jk}^i = \begin{vmatrix} b_{ij} & b_{ik} \\ b_{i+1,j} & b_{i+1,k} \end{vmatrix} \geqq 0$$

is satisfied for every choice of $1 \leqq j < k \leqq n$. Under this more stringent condition it is now possible to show that every monotone degenerate procedure is Bayes with respect to a finite discrete distribution F. The method of proof is an extension of the ideas of Lemma 5. For example, let us consider the case where a monotone strategy φ is specified with critical numbers $(x_1, x_2, \cdots, x_{n-1})$ such that $x_{i_0} = x_{i_0+1}$ and all other x_i's are distinct. We must distinguish two cases: (a) If $B_{jk}^{i_0} = 0$ for all values of j and k satisfying $j < k$, then it is possible to find a distribution $F(\omega)$ whose full mass concentrates at $n - 1$ values ω_j where ω_j belongs to $(\omega_{j-1}^0, \omega_j^0]$, for all $j \neq i$, against which φ is Bayes. The values ω_j may be selected arbitrarily provided only that they belong to the appropriate intervals. This assertion can be verified along the lines of Lemmas 5 and 3a. (b) If for some $j_0 < k_0$, $B_{j_0,k_0}^{i_0} > 0$, then it follows from (II′) that for all $j < j_0$ and $k > k_0$ also $B_{jk}^{i_0} > 0$. By selecting arbitrarily n values of ω_j subject only to the condition that ω_j belongs $(\omega_{j-1}^0, \omega_j^0]$, an n-point distribution $F(\omega)$ with weights at ω_j may be constructed so that φ is Bayes with respect to $F(\omega)$. The proof of this statement, as in Lemma 3, reduces to an application of Lemma 1.

We summarize the conclusions of this analysis in the statement of our next theorem.

THEOREM 2a. *If the loss functions satisfy properties* (I) *and* (II′) *with* $b_{ij} > 0$ *for all* i, j, *then any degenerate monotone procedure* φ^0 *is Bayes.* φ^0 *is uniquely Bayes except for the randomizations* $\lambda_1^0, \cdots, \lambda_{n-1}^0$ *at* $x_1^0, \cdots, x_{n-1}^0$ *so that if* φ^0 *is inadmissible the decision procedure* φ^* *which improves on* φ^0 *differs from* φ^0 *only in the randomizations at* $x_1^0, \cdots, x_{n-1}^0$.

COROLLARY 2. *If* μ *is atomless, all monotone procedures are admissible.*

4. Estimation with absolute error loss function. In the previous section it was mentioned that the absolute error loss function $L(a, \omega) = c \mid a - \omega \mid$ is the limiting case of the discrete absolute error loss function in an n-action problem. This fact will be utilized in this section to prove that all bounded, continuous, monotone estimates in the estimation problem for absolute error loss functions are admissible.

Since the loss function $L(a, \omega) = c \mid a - \omega \mid$ is a convex function of a for each ω, it is only necessary to consider non-randomized estimates, i.e., single-valued functions $a(x)$ which map the space X into the space Ω. It was shown in [2] that the class of monotone estimates is essentially complete when the loss is absolute error. (An estimate is monotone if $x_1 < x_2$ implies $a(x_1) \leqq a(x_2)$.) This result will be strengthened by showing that all bounded, continuous, monotone estimates are in addition admissible. It would appear that this is about as general a result for admissibility as can be obtained since it is not true in general that arbitrary unbounded monotone estimates are admissible. Consider the problem of estimating the parameter ω when x is normally distributed with mean ω and variance 1. The estimate $a(x) = x + k$ for constant $k \neq 0$ is a monotone estimate, but it is strictly dominated by the estimate $a_0(x) = x$. The investigation of the admissibility of certain natural unbounded monotone estimates on absolute error loss functions is deferred to a subsequent publication.

THEOREM 3. *All bounded, continuous, monotone estimates are admissible.*

PROOF. Let $a_0(x)$ be a monotone and continuous estimate for which $\lim_{x\to-\infty} a_0(x) = -b_1$ and $\lim_{x\to\infty} a_0(x) = b_2$, where $-b_1 < b_2$ and $b_1, b_2 < \infty$.

It will be shown that a_0 is the unique Bayes estimate with respect to some a priori distribution F^* and therefore is admissible. The problem is first reduced to the finite action case with loss functions corresponding to discrete absolute error functions.

To this end, let the Ω-space, the real line, be divided into $2N + 3$ half-open intervals $(\omega_{i-1}^N, \omega_i^N]$, $i = 1, 2, \cdots, 2N + 3$, where

$$\omega_0^N = -\infty,$$

$$\omega_1^N = -b_1 - \frac{b_1 + b_2}{2N},$$

$$\omega_i^N = -b_1 + (i - 2)\frac{b_1 + b_2}{2N}, \qquad i = 2, \cdots, 2N + 2,$$

$$\omega_{2N+3}^N = +\infty,$$

and consider the $(2N + 3)$-action problem which is defined by

$$L_j^N(\omega) = \frac{b_1 + b_2}{2N} | i - j | \qquad \text{for } \omega \, \varepsilon \, (\omega_{i-1}^N, \omega_i^N],$$

$i, j = 1, \cdots, 2N + 3$. Define the discrete decision procedure $a_0^N(x)$ as follows:

$$a_0^N(x) = \begin{cases} 2 & x < -N \\ A_i & -N + \dfrac{i}{N} \leqq x < -N + \dfrac{i+1}{N} \qquad i = 0, 1, \cdots, 2N^2 - 1 \\ & \text{where } \left| a_0\left(-N + \dfrac{i}{N}\right) + b_1 - (A_i - 2)\dfrac{b_1 + b_2}{2N} \right| \\ & \quad = \min\limits_{2 \leqq j \leqq 2N+2} \left| a_0\left(-N + \dfrac{i}{N}\right) + b_1 - (j - 2)\dfrac{b_1 + b_2}{2N} \right| \\ 2N + 2 & x \geqq N. \end{cases}$$

a_0^N should be interpreted as a monotone decision procedure which for each value of x specifies that the action $a_0^N(x)$ should be taken. Since a_0^N is monotone and never involves taking actions 1 or $2N + 3$, by Lemma 3a a_0^N is Bayes against a discrete probability distribution F_N whose spectrum is contained in the interval $(-b_1 - (b_1 + b_2)/2N, b_2]$. Since the spectrum of F_N is contained in a finite interval and as $N \to \infty$ each interval becomes a subset of the previous interval, the sequence $\{F_n\}$ has a subsequence $\{F_{n_i}\}$ such that $F_{n_i} \xrightarrow{L} F^*$, where F^* is a distribution function. Without loss of generality, assume $F_N \xrightarrow{L} F^*$ as $N \to \infty$. It is also clear that as $N \to \infty$ $L_{a_0^N(x)}^N(\omega) \to | a_0(x) - \omega |$ uniformly for $\omega \, \varepsilon \, [-b_1, b_2]$. Therefore, since $p(x, \omega)$ is continuous in ω for each x, as $N \to \infty$,

$$\int_{-\infty}^{\infty} L_{a_0^N(x)}^N(\omega) p(x, \omega) \, dF_N(\omega) \to \int_{-\infty}^{\infty} | a_0(x) - \omega | \, p(x, \omega) \, dF^*(\omega).$$

Consider any other estimate $a(x)$ whose range is contained in the interval $[-b_1, b_2]$. Without loss of generality $a(x)$ is monotone and for simplicity assume $a(x)$ is continuous. At a point of discontinuity $a(x)$ can be approximated pointwise by a continuous estimate. A limit argument gives the general case. Define $a^N(x)$ analogously to $a_0^N(x)$. Since a_0^N is Bayes against F_N,

$$\int_{-\infty}^{\infty} L^N_{a_0^N(x)}(\omega)p(x, \omega)\, dF_N(\omega) \leqq \int_{-\infty}^{\infty} L^N_{a^N(x)}(\omega)p(x, \omega)\, dF_N(\omega).$$

As $N \to \infty$, $L^N_{a^N(x)}(\omega) \to |a(x) - \omega|$ uniformly so

$$\int_{-\infty}^{\infty} |a_0(x) - \omega|\, p(x, \omega)\, dF^*(\omega) \leqq \int_{-\infty}^{\infty} |a(x) - \omega|\, p(x, \omega)\, dF^*(\omega).$$

Thus, a_0 is Bayes against F^* when the class of possible estimates is restricted to those whose ranges are contained in $[-b_1, b_2]$. But any estimate $b(x)$ which assumes values outside the interval $[-b_1, b_2]$ can obviously be improved upon by an estimate whose range is contained in $[-b_1, b_2]$ since the spectrum of F^* is contained in $[-b_1, b_2]$. Therefore, a_0 is Bayes against F^*. It is also clear that the spectrum of F^* comes arbitrarily close to the extreme values $-b_1$ and b_2. This fact is utilized below in the discussion of admissibility.

To prove admissibility it must be shown that a_0 is *uniquely* Bayes against F^*. For a fixed x in the positive sample space

$$\rho(F^*, a) = \int_{-\infty}^{\infty} |a - \omega|\, p(x, \omega)\, dF^*(\omega)$$

is a convex function of a. If it can be shown that $\rho(F^*, a)$ is strictly convex over $[-b_1, b_2]$, then the minimum will be unique. For $a_1, a_2 \,\varepsilon\, [-b_1, b_2]$, $a_1 < a_2$, and $0 < \lambda < 1$,

$$\rho(F^*, \lambda a_1 + (1 - \lambda)a_2) = \int_{-\infty}^{\infty} |\lambda(a_1 - \omega) + (1 - \lambda)(a_2 - \omega)|\, p(x, \omega)\, dF^*(\omega).$$

$|\lambda(a_1 - \omega) + (1 - \lambda)(a_2 - \omega)| \leqq \lambda\, |a_1 - \omega| + (1 - \lambda)\, |a_2 - \omega|$ with strict inequality for $a_1 < \omega < a_2$. Thus,

$$\rho(F^*, \lambda a_1 + (1 - \lambda)a_2) < \lambda\rho(F^*, a_1) + (1 - \lambda)\rho(F^*, a_2)$$

will follow if F^* assigns positive measure to every open set in the interval $[-b_1, b_2]$. Hence, it suffices to show that F^* has positive measure throughout the whole interval. Suppose the contrary; there exist constants, b_3 and b_4, such that $-b_1 < b_3 < b_4 < b_2$ and $0 < F^*(b_4) = F^*(b_3 + 0) < 1$. For $a \,\varepsilon\, [b_3, b_4]$,

$$\begin{aligned}\rho(F^*, a) &= \int_{-\infty}^{b_3} (a - \omega)p(x, \omega)\, dF^*(\omega) + \int_{b_4}^{\infty} (\omega - a)p(x, \omega)\, dF^*(\omega) \\ &= a\left[\int_{-\infty}^{b_3} p(x, \omega)\, dF^*(\omega) - \int_{b_4}^{\infty} p(x, \omega)\, dF^*(\omega)\right] + K(x, \omega),\end{aligned}$$

where $K(x, \omega)$ is a function of x and ω independent of a. If the expression in brackets is positive, $\min_{b_3 \leqq a \leqq b_4} \rho(F^*, a) = \rho(F^*, b_3)$, and if the expression is negative, $\min_{b_3 \leqq a \leqq b_4} \rho(F^*, a) = \rho(F^*, b_4)$. Thus, $a_0(x)$ can assume values between b_3 and b_4 only if the expression in brackets is zero. For each x in the non-degenerate interval $\{x : b_3 < a_0(x) < b_4\}$, we must have

$$\int_{-\infty}^{b_3} p(x, \omega)\, dF^*(\omega) - \int_{b_4}^{\infty} p(x, \omega)\, dF^*(\omega) = \int_{-\infty}^{\infty} h(\omega) p(x, \omega)\, dF^*(\omega) = 0,$$

where

$$h(\omega) = \begin{cases} 1 & \omega \leqq b_3 \\ -1 & \omega > b_3 \end{cases}.$$

But since $h(\omega)$ changes sign exactly once, $\int_{-\infty}^{\infty} h(\omega) p(x, \omega)\, dF^*(\omega)$ has at most one zero by Theorem 3 of [3] and cannot equal zero for an interval of x's. Thus, F^* must assign measure to every open interval in $[-b_1, b_2]$ and the theorem is proved for $-b_1 < b_2$.

It is a trivial verification that all estimates of the form $a(x) = c$, c constant, are admissible. (In fact, it is the unique Bayes strategy with respect to the distribution concentrating all its probability at the point $\omega = c$.)

This completes the proof of the theorem.

5. n-action problem for a special loss function. Consider the n-action problem which is defined by the $n + 1$ values $\omega_0^0, \omega_1^0, \cdots, \omega_n^0$ ($\omega_0^0 = -\infty$ and $\omega_n^0 = +\infty$) and the following set of loss functions:

$$L_i(\omega) = \begin{cases} \psi_i(\omega) & \omega \leqq \omega_{i-1}^0 \\ 0 & \omega \in (\omega_{i-1}^0, \omega_i^0] \\ c > 0 & \omega > \omega_i^0 \end{cases}$$

where $\psi_i(\omega)$ is a monotone decreasing function of ω, $i = 1, \cdots, n$, and where $\psi_i(\omega) - \psi_j(\omega) > 0$ for $i > j$ and ω in their common domain of definition. (It will be assumed that $\psi_i(\omega)$ is sufficiently smooth to justify differentiation inside the integral sign of the a posteriori risk function.) The loss for an action i is constant if the corresponding ω-interval is underestimated and increases as the magnitude of the error increases if the ω-interval is overestimated. Such a family of loss functions is suggested by the following problem. It is desired to determine how much material will be required to construct a bridge across a certain river. If insufficient material is ordered, the bridge will not be completed and the loss is the same regardless of how small or large the discrepancy is. On the other hand, if there is an overabundance of material, the loss is proportional to the amount of excess, wasted material.

The assumptions concerning the distribution functions

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

remain the same as in Sec. 2.

The main result of this investigation is embodied in Theorem 4 given below. The proof is quite similar to the proof of Theorem 1.

LEMMA 7. *A monotone procedure (non-degenerate or degenerate) is Bayes with respect to a discrete a priori distribution F^* which concentrates its probability at n points; each interval $(\omega_{i-1}^0, \omega_i^0]$, $i = 1, \cdots, n$, contains a mass-point of F^*, but the location of the mass-point in the interval is arbitrary. The non-degenerate monotone procedure is uniquely Bayes with respect of F^* except for the randomizations $\lambda_1^0, \cdots, \lambda_{n-1}^0$.*

PROOF. Consider the difference of the a posteriori risks for actions $i + 1$ and i:

$$r_x(i + 1) - r_x(i) = \int_{-\infty}^{\infty} [L_{i+1}(\omega) - L_i(\omega)]p(x, \omega)\, dF(\omega),$$

where F is some a priori distribution. When $r_x(i + 1) - r_x(i) < 0$, action $i + 1$ is preferred to action i, and when $r_x(i + 1) - r_x(i) > 0$, action i is preferred to action $i + 1$. Since $L_{i+1}(\omega) - L_i(\omega)$ changes sign once, from positive to negative as ω increases, by Theorem 3 of [3] $r_x(i + 1) - r_x(i)$ has at most one zero counting multiplicities, and if it changes sign once as x increases, it changes from positive to negative. Thus, there exists an x' such that for $x < x'$ action i is preferred to $i + 1$ and for $x > x'$ action $i + 1$ is preferred to action i.

A monotone decision procedure characterized by $(x_1^0, \cdots, x_{n-1}^0; \lambda_1^0, \cdots, \lambda_{n-1}^0)$ will thus be Bayes against F if the system of equations

$$\int_{-\infty}^{\infty} [L_{i+1}(\omega) - L_i(\omega)]p(x_i^0, \omega)\, dF(\omega) = 0, \tag{8}$$

$i = 1, \cdots, n - 1$, is satisfied. Let ω_i be an arbitrary point in the interval

$$(\omega_{i-1}^0, \omega_i^0], \qquad i = 1, \cdots, n.$$

We assert that weights f_j, $f_j > 0$, $\sum_1^n f_j = 1$, can be determined such that if F^* is the distribution which assigns probability f_j to the point ω_j, then

$$(x_1^0, \cdots, x_{n-1}^0; \lambda_1^0, \cdots, \lambda_{n-1}^0)$$

is Bayes against F^*. This can be seen as follows. The system of equations (8) becomes

$$\sum_{j=1}^{i+1} [L_{i+1}(\omega_j) - L_i(\omega_j)]p(x_i^0, \omega_j)f_j = 0, \tag{9}$$

$i = 1, \cdots, n - 1$. The $(n - 1) \times n$ coefficient matrix $A = (a_{ij})$ of the system (9) has the form:

(1) $a_{ij} > 0$ for $j \leqq i$, $\quad i = 1, \cdots, n - 1$,
(2) $a_{i,i+1} < 0$ for $i = 1, \cdots, n - 1$,
(3) $a_{ij} = 0$ for $j > i + 1$, $\quad i = 1, \cdots, n - 1$.

But any such system of equations has a solution $(f_1, \cdots, f_n)$ such that $f_j > 0$, $\sum f_j = 1$. Consider the first equation,

$$a_{11}f_1 - |a_{12}|f_2 = 0.$$

Choose any positive value whatsoever for f_1, and solve for f_2. Clearly, $f_2 > 0$. Substitute these values of f_1 and f_2 into

$$a_{21}f_1 + a_{22}f_2 - |a_{23}|f_3 = 0$$

and solve for f_3. Clearly, $f_3 > 0$, and so on. The solution $(f_1, \cdots, f_n)$ can be normalized so that $\sum f_j = 1$. Thus, $(x_1^0, \cdots, x_{n-1}^0; \lambda_1^0, \cdots, \lambda_{n-1}^0)$ is Bayes against F^*.

The uniqueness except for $\lambda_1^0, \cdots, \lambda_{n-1}^0$ is established by the fact that the zeros of

$$\int_{-\infty}^{\infty} [L_{i+1}(\omega) - L_i(\omega)]p(x, \omega)\, dF^*(\omega),$$

$i = 1, \cdots, n - 1$, are unique.

THEOREM 4. *All monotone procedures are admissible.*

Since the proof of this theorem duplicates that of Theorem 1, it will be omitted. Lemma 8 below is the analog of Lemma 4. It strengthens slightly the results of Theorem 4.

LEMMA 8. *If φ^* and φ^0 are two monotone procedures and $\rho(\omega, \varphi^*) \equiv \rho(\omega, \varphi^0)$, then $x_i^* = x_i^0$, $i = 1, \cdots, n - 1$, and $\lambda_i^* = \lambda_i^0$ for all i for which $\mu(x_i^0) > 0$.*

PROOF. Since φ^* and φ^0 are uniquely Bayes except for randomizations, $\rho(\omega, \varphi^*) \equiv \rho(\omega, \varphi^0)$ trivially implies that $x_i^* = x_i^0$, $i = 1, \cdots, n - 1$.

As in Lemma 4,

$$\rho(\omega, \varphi^0) - \rho(\omega, \varphi^*)$$
$$= \int_{-\infty}^{\infty} p(x, \omega) \left\{ \sum_{i=1}^{n-1} [L_i(\omega) - L_{i+1}(\omega)] \left[\sum_{j=1}^{i} \varphi_j^0(x) - \sum_{j=1}^{i} \varphi_j^*(x) \right] \right\} d\mu(x)$$

and

$$0 = \sum_{i=1}^{n-1} [L_i(\omega) - L_{i+1}(\omega)]p(x_i^0, \omega)\eta_i$$

where

$$\eta_i = \sum_{K_i} (\lambda_j^0 - \lambda_j^*)\mu(x_i^0), \qquad K_i = \{j \mid j \leqq i, x_j^0 = x_i^0\}.$$

When this expression is evaluated at $\omega_1, \cdots, \omega_n$, the system of equations $\eta A = 0$ is produced where $\eta = (\eta_1, \cdots, \eta_{n-1})$ and $A = (a_{ij})$ is an $(n - 1) \times n$ matrix satisfying

(1) $a_{ij} < 0$ for $j \leqq i$, $\quad i = 1, \cdots, n - 1$,
(2) $a_{i,i+1} > 0$, $\quad i = 1, \cdots, n - 1$,
(3) $a_{ij} = 0$ for $j > i + 1$, $\quad i = 1, \cdots, n - 1$.

But the only solution to such a system of equations is $\eta = (0, 0, \cdots, 0)$. The last equation $a_{n-1,n}\eta_{n-1} = 0$ implies $\eta_{n-1} = 0$. If $\eta_{n-1} = 0$ is substituted into the next to last equation, the resulting equation $a_{n-1,n-1} \cdot 0 + a_{n-2,n-1}\eta_{n-2} = 0$ implies $\eta_{n-2} = 0$, and so on. But $\eta_i = 0$, $i = 1, \cdots, n - 1$ implies the stated result.

6. Estimation with the special loss function of Section 5. If $\omega \varepsilon \Omega$ is the true parameter point and ω is estimated to be a, then the loss is $L(\omega, a)$ where

$$L(\omega, a) = \begin{cases} c > 0 & a < \omega \\ \psi(a - \omega) & a \geqq \omega \end{cases}$$

where $\psi(\xi)$ is a monotone increasing function of ξ with $\psi(0) = 0$.

The method of proof of Theorem 5 below closely parallels that of the first part of Theorem 3. Unlike the absolute error loss function, this loss function does not admit an easy proof of the uniqueness of the Bayes estimate which is required for this type of admissibility proof. In fact, the indications are that uniqueness fails to hold, but no proof of the inadmissibility of a bounded, continuous, monotone estimate has been obtained as yet. However, a weaker positive result in this direction is the following:

THEOREM 5. *All bounded, continuous, monotone estimates are Bayes estimates.*

The proof is omitted.

Finally, we remark that the same kind of results can be obtained for the case where the loss functions are such that the error is constant for overestimation and arbitrarily monotonically increasing for underestimation.

7. Minimax results for the discrete absolute error loss function. For the n-action problem with discrete absolute error loss function it is not true, in general, that $\min_\varphi \max_F \rho(F, \varphi) = \max_F \min_\varphi \rho(F, \varphi)$, although $\inf_\varphi \sup_F \rho(F, \varphi) = \sup_F \inf_\varphi \rho(F, \varphi)$ provided $L_i(\omega) \geqq 0$ and the value of the game is allowed to be infinite [5]. Most often the game fails to have the property that the F player has a minimax strategy. The difficulties stem from two sources. The space $\{F\}$ of all probability distributions on the real line is not compact, and the loss functions are discontinuous at ω_i^0, $i = 1, \cdots, n - 1$. However, the following result is true. Suppose Ω consists of a finite number of points $\{\omega_1, \cdots, \omega_N\}$ where $n \leqq N$ and $\omega_i < \omega_{i+1}$, $i = 1, \cdots, N - 1$, and $L(\omega_i, j) = c \mid i' - j \mid$ for $\omega_i \varepsilon S_{i'}$. For this game structure it is very easy to establish (cf. [3]) that

$$\max_F \min_\varphi \rho(F, \varphi) = \min_\varphi \max_F \rho(F, \varphi),$$

where $\{F\}$ is the space of all discrete probability distributions defined on Ω.

It will be assumed that all the previous assumptions of Sec. 2 apply equally well here.

Since the class of monotone procedures is essentially complete, there exists at least one monotone minimax strategy. The character of the monotone strategy has been well-defined, but nothing has been said as yet about the structure of nature's minimax strategies. The following lemmas have this as their aim.

LEMMA 9. *Let v be the value of the game, and let φ^0 be a monotone minimax strategy. If $T_{\varphi_0} = \{\omega \mid \rho(\omega, \varphi^0) = v\}$, then T_{φ_0} contains points in each region*

$S_i \subset \Omega$, $i = 1, \cdots, n$, for which φ^0 has a corresponding non-degenerate interval in X in which action i is taken.

PROOF. Suppose that for some i, $T_{\varphi_0} \cap S_i = \theta$ (the empty set). Then a minimax strategy F^0 for nature will not concentrate any probability in the region S_i. Since for all φ, $\rho(F^0, \varphi) \geqq \rho(F^0, \varphi^0) = v$, φ^0 is Bayes with respect to F^0. By Lemma 5, φ^0 cannot have a non-degenerate interval for action i.

LEMMA 10. *If the monotone minimax strategy φ^0 involves k non-degenerate intervals, there exists a minimax strategy F^0 for nature whose spectrum consists of k points, each point belonging to a region S_i for which the action i interval in X is non-degenerate.*

PROOF. By Lemma 9, in each region S_i for which action i has a non-degenerate interval there exists at least one $\omega \varepsilon \Omega$ such that $\rho(\omega, \varphi^0) = v$. Choose one such point from each of the eligible S_i. By Lemma 3a, it is possible to construct a distribution F^0 concentrating its probability at these points with respect to which φ^0 is Bayes. Since for all φ, $\rho(F^0, \varphi) \geqq \rho(F^0, \varphi^0) = v$, F^0 is minimax.

When $n = 3$ and $p(x, \omega)$ is Pólya-type 3, with continuous second derivatives, there is a constructive method of obtaining the monotone minimax strategy. Define

$$\begin{aligned} A_1(\omega) = c \int_{x_1}^{x_2} p(x, \omega)\, d\mu(x) + 2c \int_{x_2}^{\infty} p(x, \omega)\, d\mu(x) + c(1 - \lambda_1)p(x_1, \omega)\mu(x_1) \\ + [c\lambda_2 + 2c(1 - \lambda_2)]p(x_2, \omega)\mu(x_2), \end{aligned}$$

$$\begin{aligned} A_2(\omega) = c \int_{-\infty}^{x_1} p(x, \omega)\, d\mu(x) + c \int_{x_2}^{\infty} p(x, \omega)\, d\mu(x) + c\lambda_1 p(x_1, \omega)\mu(x_1) \\ + c(1 - \lambda_2)p(x_2, \omega)\mu(x_2), \end{aligned}$$

$$\begin{aligned} A_3(\omega) = 2c \int_{-\infty}^{x_1} p(x, \omega)\, d\mu(x) + c \int_{x_1}^{x_2} p(x, \omega)\, d\mu(x) \\ + [2c\lambda_1 + c(1 - \lambda_1)]p(x_1, \omega)\mu(x_1) + c\lambda_2 p(x_2, \omega)\mu(x_2). \end{aligned}$$

Let φ be the monotone procedure characterized by $(x_1, x_2; \lambda_1, \lambda_2)$. For $\omega \varepsilon S_1$, $\rho(\omega, \varphi) = A_1(\omega)$; for $\omega \varepsilon S_2$, $\rho(\omega, \varphi) = A_2(\omega)$; and for $\omega \varepsilon S_3$, $\rho(\omega, \varphi) = A_3(\omega)$. Let

$$\omega_{i_1} = \max \{\omega_i \mid \omega_i \varepsilon S_1\}, \qquad \omega_{i_2} = \min \{\omega_i \mid \omega_i \varepsilon S_2\},$$
$$\omega_{i_3} = \max \{\omega_i \mid \omega_i \varepsilon S_2\}, \qquad \omega_{i_4} = \min \{\omega_i \mid \omega_i \varepsilon S_3\}.$$

Since $p(x, \omega)$ is Pólya-type 3, $A_1(\omega)$ is a monotone increasing function of ω, $A_3(\omega)$ is a monotone decreasing function of ω, and $A_2(\omega) - \alpha$ as a function of ω has at most two changes of sign (for any α). Thus

$$\max_{\omega_i \varepsilon S_1} \rho(\omega, \varphi) = \rho(\omega_{i_1}, \varphi) = A_1(\omega_{i_1})$$

$$\begin{aligned} \max_{\omega_i \varepsilon S_2} \rho(\omega, \varphi) &= \max \{\rho(\omega_{i_2}, \varphi), \rho(\omega_{i_3}, \varphi)\} \\ &= \max \{A_2(\omega_{i_2}), A_2(\omega_{i_3})\} \end{aligned}$$

and

$$\max_{\omega_i \varepsilon S_3} \rho(\omega, \varphi) = \rho(\omega_{i_4}, \varphi) = A_3(\omega_{i_4}).$$

To obtain the monotone minimax strategy choose $(x_1, x_2; \lambda_1, \lambda_2)$ such that

$$\rho(\omega_{i_1}, \varphi) = \rho(\omega_{i_2}, \varphi) = \rho(\omega_{i_4}, \varphi) \geqq \rho(\omega_{i_3}, \varphi), \tag{10}$$

or, if this is impossible, choose $(x_1, x_2; \lambda_1, \lambda_2)$ such that

$$\rho(\omega_{i_1}, \varphi) = \rho(\omega_{i_3}, \varphi) = \rho(\omega_{i_4}, \varphi) \geqq \rho(\omega_{i_2}, \varphi). \tag{11}$$

It is clear that the monotone strategy $(x_1, x_2; \lambda_1, \lambda_2)$ where $x_1, x_2, \lambda_1, \lambda_2$ are defined by (10) or (11) is a minimax strategy.

Either the system (10) or the system (11) has a solution since the statistician has a minimax strategy and this strategy must involve three non-degenerate intervals. The latter statement is proved as follows:

LEMMA 11. $v < c$.

PROOF. Consider the strategy $\varphi = (\varphi_1, \varphi_2, \varphi_3)$ where $\varphi_i(x) = \frac{1}{3}$, $i = 1, 2, 3$. $\rho(\omega, \varphi) \leqq c$ for all ω. Therefore, $v \leqq c$. But there exists a monotone strategy φ^0 which improves uniformly on φ by Theorem 1 of [2]. Therefore, $v < c$.

Now consider the various cases for which there are at most two non-degenerate intervals.

Case 1. $x_2 = \infty$: $\rho(\omega, \varphi) \geqq c$ for $\omega \varepsilon S_3$
Case 2. $x_1 = x_2$: $\rho(\omega, \varphi) = c$ for $\omega \varepsilon S_2$
Case 3. $x_1 = -\infty$: $\rho(\omega, \varphi) \geqq c$ for $\omega \varepsilon S_1$

Therefore, a minimax strategy for nature must involve three non-degenerate intervals.

REFERENCES

[1] S. KARLIN, "Decision theory for Pólya type distributions, case of two actions, I," *Proceedings of the Third Berkeley Symposium on Probability and Statistics*, Vol. I, University of California Press, 1956, pp. 115–129.
[2] S. KARLIN AND H. RUBIN, "The theory of decision procedures for distributions with monotone likelihood ratio," *Ann. Math. Stat.*, Vol. 27 (1956), pp. 272–299.
[3] S. KARLIN, "Pólya type distributions, II," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 281–308.
[4] S. KARLIN, "Pólya type distributions, IV," to be published.
[5] L. LE CAM, "An extension of Wald's theory of statistical decision functions," *Ann. Math. Stat.*, Vol. 26 (1955), p. 69.