DECISION THEORY FOR PÓLYA TYPE DISTRIBUTIONS. CASE OF TWO ACTIONS, I

SAMUEL KARLIN
STANFORD UNIVERSITY

1. Introduction

In a recent paper by the author and H. Rubin [1] a thorough analysis was made of the form of essentially complete classes when the densities $p(x|\omega)$ have monotone likelihood ratios. The form of the Bayes strategies was determined and complete classes of strategies were characterized. In order to carry out this analysis, very stringent conditions were imposed on the loss functions. For example, in the case of two actions it was assumed that $L_1(\omega) - L_2(\omega)$ possesses at most one sign change. This condition of one sign change is intrinsically connected to the fact that the density $p(x|\omega)$ has a monotone likelihood ratio. In this paper we deal with the general case of two actions where no assumptions are made concerning the number of sign changes of $L_1(\omega) - L_2(\omega)$. However, the class of densities $p(x|\omega)$ is specialized to that of the exponential family in the first half of this paper and then the results are extended to a class of distributions which we call Pólya type distributions. Although the exponential distributions constitute a subfamily of the Pólya type distributions, we have chosen to treat all the decision theory associated with this class of distributions separately in order to best illustrate the scope of the ideas and methods developed below. In section 8 the results for the exponential family of distributions are then extended to the case when the underlying distributions are Pólya type.

A thorough study is made of form of complete classes, Bayes strategies and admissibility. Further results are obtained which relate Liapounoff's theorem on the range of a vector measure to the family of distributions under consideration.

2. Preliminaries and definitions

Let the observed random variable be denoted by x ranging over X and the unknown state of nature by ω in Ω. The sets X and Ω are one-dimensional and will be identified with subsets of the real line. To expedite the discussion we take Ω to be an interval.

The cumulative distribution of x, when the state of nature is ω, is assumed to belong to the exponential family, that is,

$$P(x, \omega) = \beta(\omega) \int_{-\infty}^{x} e^{t\omega} d\mu(t), \tag{1}$$

where $\beta(\omega) > 0$ for ω in Ω and μ is a σ-finite measure defined on the real line with the spectrum of μ equal to X. (The spectrum of μ is defined to be the set of all x such that for

This paper (Technical Report No. 28, Department of Statistics, Stanford University) was prepared under Contract N6onr-251, Task Order III (NR-042-993) for the Office of Naval Research.

any open set W containing x, $\mu\{W\} > 0$.) This family of distributions includes many well known examples, for instance, Normal with known variance, Poisson, Gamma, and Binomial.

We shall be concerned throughout this study with the general case of two actions. Let the loss functions be denoted by $L_1(\omega)$ and $L_2(\omega)$ when taking actions 1 and 2, respectively, and the true parameter value is ω.

For purposes of presenting the basic requirements on $L_i(\omega)$ we introduce the following concepts.

The number of sign changes $V(h)$ of a function $h(\omega)$ is taken to be $\sup_{\omega_1, \cdots, \omega_m} N[h(\omega_i)]$, where $N[h(\omega_i)]$ is the number of changes of sign of the sequence $h(\omega_1), h(\omega_2), \cdots, h(\omega_m)$, where $\omega_i < \omega_{i+1}$. A point ω_0 is called a change point for $h(\omega)$ if

$$h(\omega)\, h(\omega') \leqq 0\,, \qquad \omega, \omega' \text{ essentially near } \omega_0\,, \tag{2}$$

whenever $\omega \leqq \omega_0 \leqq \omega'$ with $\omega \neq \omega'$ and definite inequality occurs for some specific choice of ω and ω' or $h(\omega_0)h(\omega)h(\omega') \leqq 0$ for $\omega < \omega_0 < \omega'$.

We require that $L_1 - L_2 = h$ changes sign n times with n a fixed finite number. Furthermore, L_i, $i = 1, 2$, are defined everywhere in Ω and have sufficient smoothness properties to insure the existence of all integrals involving these quantities. Finally, we assume that $L_1 - L_2$ has at most a countable number of points of discontinuity.

The set of change points of $L_1(\omega) - L_2(\omega)$ is denoted by S^0 and arranged in order consists of $\omega_1^0, \omega_2^0, \cdots, \omega_n^0$. We allow the possibility that two successive ω_i^0 may be equal, but no more than two ω_i^0 can coincide. If $\omega_i^0 = \omega_{i+1}^0$, then we must have

$$[L_1(\omega) - L_2(\omega)]\,[L_1(\omega') - L_2(\omega')] > 0 \tag{3}$$

for $\omega < \omega_i^0 < \omega'$ with ω and ω' sufficiently close to ω_i^0. Of course, $[L_1(\omega) - L_2(\omega)] \cdot [L_1(\omega_i^0) - L_2(\omega_i^0)] < 0$ for the same choice of ω.

This corresponds to the case where one of the two actions is preferred to the other in an open neighborhood of ω_i^0 except for $\omega = \omega_i^0$ where the reverse preference exists. The loss functions L_i are numbered so that $L_1(\omega) - L_2(\omega) < 0$ for $\omega < \omega_1^0$.

Two important special cases are worth exhibiting.

(1) If $n = 1$, that is, there is only one sign change, then we are dealing with the usual one-sided decision problem. For $\omega < \omega_1^0$ action 1 is preferred (say), and for $\omega > \omega_1^0$ action 2 is preferred. The complete class theory, Bayes solutions and other aspects of the decision problem were treated in detail for this special case in [1].

(2) Another important example is the two-sided problem. This corresponds to the case of $n = 2$ change points. Heuristically speaking, for small and large parameter values action 1 is favored while if the parameter values are moderate, action 2 is favored.

3. Basic lemmas

The following lemma is fundamental in all that follows.

LEMMA 1. *If $h(\omega)$ changes sign n times, then*

$$g(x) = \int e^{x\omega} \beta(\omega)\, h(\omega)\, dF(\omega)\,, \tag{4}$$

where F is a measure, has at most n zeros counting multiplicities or is identically zero. The function g is identically zero only if the spectrum of F is contained in the union of the set of change points of h and the zeros of h.

PROOF. Let $\omega_1 \leqq \omega_2 \leqq \cdots \leqq \omega_n$ denote the change points of $h(\omega)$. Suppose for definiteness $(-1)^n h(\omega) \leqq 0$ for $\omega < \omega_1$. Consider the quantity

$$(5) \quad e^{\omega_1 x} \frac{d}{dx}\left\{ e^{-\omega_n x+\omega_{n-1}x} \frac{d}{dx} \cdots \frac{d}{dx}\left[e^{-\omega_2 x+\omega_1 x} \frac{d}{dx} (e^{-\omega_1 x} g(x))\right]\right\} = \int e^{x\omega} \prod_{i=1}^{n} (\omega - \omega_i) \beta(\omega) h(\omega) dF(\omega) .$$

Since $\prod_{i=1}^{n} (\omega - \omega_i)$ and $h(\omega)$ change signs at the same points in the same direction, the integral in (5) is everywhere positive provided F does not concentrate fully at the zeros of h and the set of change points ω_i. Thus, in this latter case

$$(6) \qquad \frac{d}{dx} e^{-\omega_n x+\omega_{n-1}x} \frac{d}{dx} \cdots \frac{d}{dx} e^{-\omega_2 x+\omega_1 x} \frac{d}{dx} e^{-\omega_1 x} g(x) > 0 .$$

Hence, by virtue of Rolle's theorem

$$(7) \qquad e^{-\omega_n x+\omega_{n-1}x} \frac{d}{dx} e^{-\omega_{n-1}x+\omega_{n-2}x} \cdots \frac{d}{dx} e^{-\omega_1 x} g(x)$$

has at most one zero. Therefore, the same is true of

$$(8) \qquad \frac{d}{dx} e^{-\omega_{n-1}x+\omega_{n-2}x} \cdots \frac{d}{dx} e^{-\omega_1 x} g(x) .$$

Another application of Rolle's theorem implies that

$$(9) \qquad e^{-\omega_{n-1}x+\omega_{n-2}x} \frac{d}{dx} e^{-\omega_{n-2}x+\omega_{n-3}x} \cdots \frac{d}{dx} e^{-\omega_1 x} g(x)$$

has at most 2 zeros counting multiplicities. Repeating this reasoning n times leads to the conclusion of the lemma.

Remark 1. A careful examination of the argument of the lemma also shows that if $g(x)$ has n changes of sign, then the first region of x values where g is different from zero has the property that $(-1)^n g(x) < 0$. In other words, if g and h have the same number of sign changes then they change signs in the same order.

In an analogous manner, it is shown

LEMMA 2. *If r changes sign at most n times then*

$$(10) \qquad s(\omega) = \int e^{x\omega} r(x)\, d\mu(x) , \qquad d\mu(x) > 0 \textit{ for } x \textit{ in } X ,$$

has at most n zeros counting multiplicities.

COROLLARY 1. *For any distribution F whose spectrum is not contained in the set consisting of $\omega_1^0, \omega_2^0, \cdots, \omega_n^0$ and the zeros of $L_1 - L_2$, then*

$$(11) \qquad \Lambda(x) = \int e^{x\omega} \beta(\omega) [L_1(\omega) - L_2(\omega)]\, dF(\omega)$$

has at most n zeros counting multiplicities.

4. Bayes strategies

A strategy for the statistician can be described by a pair of functions $\varphi = [\varphi_1(x), 1 - \varphi_1(x)]$ where $\varphi_1(x)$ represents the probability of taking action 1 when x is observed. Two strategies $\varphi = (\varphi_1, 1 - \varphi_1)$ and $\psi = (\psi_1, 1 - \psi_1)$ are identified if $\varphi_1 = \psi_1$ almost everywhere with respect to the measure μ. The risk corresponding to a given strategy $\varphi = (\varphi_1, 1 - \varphi_1)$ with nature's state described by ω is given by

$$\rho(\omega, \varphi) = \beta(\omega) \int e^{\omega x} \{ \varphi_1(x) L_1(\omega) + [1 - \varphi_1(x)] L_2(\omega) \} d\mu(x) . \tag{12}$$

Our first task is to characterize the form of the Bayes strategies. In order to avoid minor technical difficulties we make the assumption that the zeros of $L_1 - L_2$ are contained in the set of change points S^0.

THEOREM 1. *For any F either all strategies are Bayes or the Bayes strategy is uniquely determined except for at most n points and consists of at most $n + 1$ disjoint intervals where either action 1 or 2 is taken.*

PROOF. If the spectrum of F does not concentrate fully in S^0, then corollary 1 shows that for any distribution as in (11), Λ vanishes at most n times. The set of x where $\Lambda(x) < 0$ requires the optimum procedure to take action 1. The optimum strategy requires the statistician to take action 2 for any observed x for which $\Lambda(x) > 0$. Whenever $\Lambda(x) = 0$ both actions produce the same expected yield. Since Λ is continuous and has at most n zeros, the conclusion of the theorem is evident for this case. If the spectrum of F is contained in S^0 and Λ does not vanish identically, then Λ is a nonzero exponential polynomial with at most n terms. It follows that Λ vanishes at most $n - 1$ times and the argument proceeds as before. The proof of theorem 1 is hereby complete.

In view of theorem 1, all Bayes strategies can be described by at most n numbers x_i with $x_i \leqq x_{i+1}$ such that

$$\varphi_1(x) = \begin{cases} 1, & x_{2i} < x < x_{2i+1}, \quad i = 0, \cdots, \left(\frac{n-1}{2}\right) \\ \lambda_i, & x = x_i,\ 0 < \lambda_i < 1 \\ 0, & \text{elsewhere} \end{cases} \tag{13}$$

and $\varphi_2(x) = 1 - \varphi_1(x)$ where always $x_0 = -\infty$. The fact that $L_1(\omega) - L_2(\omega) < 0$ for $\omega < \omega_1^0$ and the implication of remark 1 enables us to conclude that $\varphi_1(x) = 1$ for $x < x_1$ and thus the representation of the Bayes strategies as in (13) is accurate. The collection of such strategies shall be denoted by $\mathcal{M}$. The collection of all possible strategies $\varphi = (\varphi_1, 1 - \varphi_1)$ will be denoted by $\mathcal{S}$. In the description of a strategy in $\mathcal{M}$ the possibility that some of the x_i are $\pm\infty$ is not excluded.

5. Essential complete classes and admissibility

Using the result of theorem 1, we now establish

THEOREM 2. *The set $\mathcal{M}$ is an essentially complete class of strategies.*

PROOF. Consider a dense set of ω enumerated as follows: $\omega_1, \omega_2, \omega_3, \cdots$. We include in this collection any point of discontinuity of $L_1 - L_2$. Consider the modified statistical problem where the states of nature are restricted to the finite number $\omega_1, \omega_2, \cdots, \omega_m$. Since in this modified problem the spaces of strategies for both the statistician and nature are compact, the Bayes strategies constitute an essentially complete class [2]. By theorem 1 the relevant Bayes strategies are contained in the set $\mathcal{M}$. This holds for every finite choice of m. Thus for any strategy $\varphi = (\varphi_1, 1 - \varphi_1)$, we can find a strategy $\varphi^m = (\varphi_1^m, 1 - \varphi_1^m)$ in $\mathcal{M}$ such that

$$\rho(\varphi, \omega_i) \geqq \rho(\varphi^m, \omega_i) \qquad i = 1, \cdots, m, \tag{14}$$

or

$$[L_1(\omega_i) - L_2(\omega_i)] \int e^{\omega_i x}[\varphi_1(x) - \varphi_1^m(x)]\, d\psi(x) \geqq 0 . \tag{15}$$

From the strategies φ^m a limit strategy φ^0 in $\mathcal{M}$ can be selected. Indeed, as φ^m are determined by a set of x_i^m and λ_i^m, $i = 1, \cdots, n$, with $x_i^m \leqq x_{i+1}^m$ and $0 \leqq \lambda_i^m \leqq 1$ [see (13)], one selects limit values from these quantities to determine φ^0. It is an easy matter to show that for any fixed ω_i

$$\int e^{\omega_i x}\varphi_1^m(x)\, d\mu(x) \to \int e^{\omega_i x}\varphi_1^0(x)\, d\mu(x) . \tag{16}$$

Consequently, in (14) we obtain

$$\rho(\varphi, \omega_i) \geqq \rho(\varphi^0, \omega_i) , \qquad \text{ior every } i . \tag{17}$$

Since ω_i constitute a dense set and $\omega \notin \{\omega_i\}$ is a point of continuity of $L_1 - L_2$, we obtain by a limiting argument that (17) holds for every ω. This completes the proof.

To obtain a more precise result, we need

LEMMA 3. *If $\varphi^1 = (\varphi_1, 1 - \varphi_1)$ and $\varphi^2 = (\varphi_2, 1 - \varphi_2)$ are two strategies in $\mathcal{M}$ then*

$$\int e^{\omega x}[\varphi_1(x) - \varphi_2(x)]\, d\mu(x) \tag{18}$$

has less than n zeros counting multiplicities.

PROOF. Let the given strategies be described by

$$\varphi_1 = \begin{cases} 1 , & x_{2i} < x < x_{2i+1} \\ \lambda_i , & x = x_i \\ 0 , & \text{elsewhere} \end{cases} \tag{19}$$

$$\varphi_2 = \begin{cases} 1 , & y_{2i} < x < y_{2i+1} \\ \mu_i , & x = y_i \\ 0 , & \text{elsewhere} \end{cases} \tag{20}$$

with $x_i \leqq x_{i+1}$, $y_i \leqq y_{i+1}$, $i = 0, \cdots, \left[\frac{n-1}{2}\right]$.

Two cases shall be considered. Suppose $x_1 \leqq y_1$; then $\varphi_1 - \varphi_2 \leqq 0$ for $x < x_2$ and thus the first change point of $\varphi_1 - \varphi_2$ can occur only for $x \geqq x_2$. But, only the values x_i can qualify as change points since elsewhere φ_1 achieves its maximum and minimum values, 1 and 0, respectively, while φ_2 lies everywhere between 0 and 1. Thus, in this circumstance $\varphi_1 - \varphi_2$ can only possibly change signs for $x = x_2, x_3, \cdots, x_n$ and thus $\varphi_1 - \varphi_2$ has at most $n - 1$ changes of sign. If on the other hand $x_1 > y_1$, then $\varphi_1 - \varphi_2 \geqq 0$ for $x < y_2$. By a symmetrical argument in terms of φ_2, $\varphi_1 - \varphi_2$ can only change signs at the possible points $x = y_2, y_3, \cdots, y_{n-1}$ and again we obtain that $\varphi_1 - \varphi_2$ has at most $n - 1$ changes of sign. Applying lemma 2 shows that

$$\int e^{\omega x}[\varphi_1(x) - \varphi_2(x)]\, d\mu(x) \tag{21}$$

can vanish counting multiplicities at most $n - 1$ times. Q.E.D.

THEOREM 3. *For any strategy $\varphi = (\varphi_1, 1 - \varphi_1)$ not in $\mathcal{M}$ there exists a unique member φ^0 of $\mathcal{M}$ such that*

$$\rho(\varphi^0, \omega) \leqq \rho(\varphi, \omega) \tag{22}$$

with inequality everywhere except for ω in S^0. Moreover, the set $\mathcal{M}$ constitutes a minimal complete system.

PROOF. For any φ not in $\mathcal{M}$ there exists by theorem 2 a φ^0 in $\mathcal{M}$ such that (22) holds, that is,

$$[L_1(\omega) - L_2(\omega)] \int e^{\omega x} (\varphi_1 - \varphi_1^0) \, d\psi(x) \geqq 0 \tag{23}$$

for all ω. Since $L_1 - L_2$ changes sign at ω_i^0 and the integral expression is continuous, we find that

$$\int e^{\omega_i^0 x} (\varphi_1 - \varphi_1^0)(x) \, d\mu(x) = 0 \tag{24}$$

with $i = 1, \cdots, n$. If two of the ω_i^0 coincide, then

$$\int e^{\omega x} [\varphi_1(x) - \varphi_1^0(x)] \, d\mu(x) \tag{25}$$

must have a double root at ω_i^0. If there existed two members φ^0 and φ^* of $\mathcal{M}$ dominating φ, then (24) would hold for both φ_1^0 and φ_1^* which contradicts the conclusion of lemma 3. Moreover, an analysis as in lemma 3 shows that if φ is not in $\mathcal{M}$ and φ^0 is in $\mathcal{M}$, then $\varphi_1^0 - \varphi_1$ can have at most n changes of sign. In fact, $\varphi_1^0 - \varphi_1$ can change signs only at x_i where φ_1^0 changes from its minimum zero values to its maximum one values. There are n such points and hence at most n sign changes. Thus the relation (24) cannot hold for any ω excluding ω_i^0 since the relation (24) has at most n zeros counting multiplicities. From this we deduce that if $L_1(\omega) - L_2(\omega) \neq 0$, for $\omega \neq \omega_i^0$ then $\rho(\varphi^0, \omega) < \rho(\varphi, \omega)$. That $\mathcal{M}$ forms a minimal complete class is due to the fact that no strategy in $\mathcal{M}$ can be dominated. Indeed, if φ in $\mathcal{M}$ is dominated by some strategy, then on account of theorem 2 there exists a strategy φ^* in $\mathcal{M}$ which dominates φ. But then

$$\int e^{\omega_i^0 x} [(\varphi_1 - \varphi_1^*)(x)] \, d\mu(x) = 0, \qquad i = 1, \cdots, n, \tag{26}$$

with a double root occurring at ω_i^0 if $\omega_i^0 = \omega_{i+1}^0$. This last assertion is impossible by virtue of lemma 3. The proof of theorem 3 is hereby complete.

6. Liapounoff theorem and its connections

In this section we deduce some interesting consequences from theorem 3. It is well known and evident that if φ varies over all functions such that $0 \leqq \varphi \leqq 1$, then the set of all n-tuples

$$\int e^{\omega_i x} \varphi(x) \, d\mu(x), \qquad i = 1, \cdots, n, \tag{27}$$

spans a convex closed bounded set in n space. Let us denote this set by $\mathcal{C}$. [In considering these n-tuples given by (27) we would like to allow the possibility that two consecutive $\omega_i = \omega_{i+1}$ are equal; then in that case, in place of the two components

$$\int e^{\omega_i x} \varphi(x) \, d\mu(x), \quad \int e^{\omega_{i+1} x} \varphi(x) \, d\mu(x) \tag{28}$$

in (27), the component terms

$$\int e^{\omega_i x} \varphi(x) \, d\mu(x), \quad \int e^{\omega_i x} x \varphi(x) \, d\mu(x) \tag{29}$$

are used.] Furthermore, in what follows, a strategy $\varphi = (\varphi_1, 1 - \varphi_1)$ shall be referred to in terms of its first component. That is, for our purposes a strategy shall be identifiable with a function φ where $0 \leqq \varphi \leqq 1$.

In the process of demonstrating theorem 3 we have shown that to each φ not in $\mathcal{M}$, $0 \leqq \varphi \leqq 1$, there exists a unique φ^0 in $\mathcal{M}$ such that

$$\int e^{\omega_i x}\varphi(x)\,d\mu(x) = \int e^{\omega_i x}\varphi^0(x)\,d\mu(x), \qquad i=1,\cdots,n. \tag{30}$$

Thus, we obtain the interesting conclusion that the set of all n-tuples

$$\int e^{\omega_i x}\varphi^0(x)\,d\mu(x) \tag{31}$$

spans a convex closed bounded set identical with $\mathcal{C}$ as φ^0 varies over $\mathcal{M}$. Moreover, each point in $\mathcal{C}$ corresponds to a unique φ^0 of $\mathcal{M}$. This last fact is a consequence of lemma 3. This conclusion represents a generalization of the Liapounoff theorem on the range of a vector measure for a special collection of measures. We now proceed to characterize the boundary of $\mathcal{C}$.

Lemma 4. *If* $(a_i) = (\int \exp(\omega_i x)\varphi^0(x)dx)$ *is a boundary point of* $\mathcal{C}$, *then* φ^0 *belongs to* $\mathcal{M}$ *and* φ^0 *is determined by* (x_i), $i = 0,\cdots, n$, *such that for some* i *either* $x_i = x_{i+1}$ *or* $x_{i+1} = \pm\infty$. *Conversely, any such member of* $\mathcal{M}$ *is a boundary point of* $\mathcal{C}$.

Proof. A point a in the convex set $\mathcal{C}$ is a boundary point if and only if a belongs to a supporting hyperplane to $\mathcal{C}$. Supposing this is the case, then there exist constants α_i, $i = 1,\cdots, n$, and α_0 such that

$$\int \sum_{i=1}^{n} \alpha_i e^{\omega_i x}\varphi^0(x)\,d\mu(x) = \alpha_0 \tag{32}$$

while for any other φ

$$\int \sum_{i=1}^{n} \alpha_i e^{\omega_i x}\varphi(x)\,d\mu(x) \geqq \alpha_0. \tag{33}$$

Forming the difference of (32) and (33), we obtain for any φ

$$\int \left[\sum_{i=1}^{n} \alpha_i e^{\omega_i x}\right](\varphi-\varphi^0)(x)\,d\mu(x) \geqq 0. \tag{34}$$

Suppose the zeros of $\alpha(x) = \sum_{i=1}^{n} \alpha_i \exp(\omega_i x)$ occur at $x_1,\cdots, x_{n-1}$. (It is well known that $\sum_{i=1}^{n} \alpha_i \exp(\omega_i x)$ has at most $n-1$ zeros counting multiplicities.) By choosing φ identical to φ^0 except for $x_i < x < x_{i+1}$, we deduce that φ^0 is either identically zero or 1 for $x_i < x < x_{i+1}$ depending on whether $\alpha(x) > 0$ or <0, respectively, in that interval. By multiplying through by a suitable constant we may, without loss of generality, assume that $\alpha(x) < 0$ preceding its first zero. Thus the boundary point φ^0 is characterized and evidently is determined as stated in the lemma.

The converse is readily established as follows. We select α_i so that $\sum_{i=1}^{n} \alpha_i \exp(\omega_i x)$ vanishes at the prescribed x_i with the proper multiplicity. This can be accomplished since there are only at most $n-1$ distinct x_i counting multiplicity. Define

$$\alpha_0 = \int \left(\sum \alpha_i e^{\omega_i x}\right)\varphi_0(x)\,d\mu(x). \tag{35}$$

The relation (34) or its negative is now clearly satisfied. From this (32) and (33) can be deduced in a straightforward manner and these equations together show that the given $[\int \exp(\omega_i x)\varphi_0\, d\psi(x)]$ is a boundary point.

Combining all the previous discussion in the form of a theorem, we obtain

THEOREM 4. *If* $\omega_1, \omega_2, \cdots, \omega_n$ *are any n distinct numbers, then the set of all n-tuples*

$$\int e^{\omega_i x}\varphi_1(x)\, d\mu(x) \tag{36}$$

obtained by allowing $\varphi = (\varphi_1, 1 - \varphi_1)$ *to vary over* $\mathcal{M}$ *constitutes a closed bounded convex set. Moreover, distinct members of* $\mathcal{M}$ *yield distinct n-tuples. The boundary of this set is characterized in lemma* 4.

Another method of proof can be given involving the use of the fixed point theorem of Brouwer and lemma 3. We now proceed to this second approach.

Suppose for simplicity $d\mu(x) = r(x)dx$ (that is, μ is absolutely continuous) and, hence, in dealing with strategies in $\mathcal{M}$ no randomization is required. A strategy is now fully defined in terms of the critical x_i. We consider the mapping of the simplex Δ_δ in n space defined by

$$x_1 < x_2 < x_3 < \cdots < x_n \tag{37}$$

with

$$x_{i+1} - x_i \geqq \delta\,, \qquad x_1 \geqq -M,\ x_n \leqq M\,, \tag{38}$$

into the n-tuple $\int \exp(\omega_i x)\varphi_0(x)\, r(x)dx$, $i = 1, \cdots, n$, where

$$\varphi_0(x) = \begin{cases} 1, & x_{2i} \leqq x < x_{2i+1} \\ 0, & \text{elsewhere}\,. \end{cases} \tag{39}$$

The boundaries of Δ_δ are described by $x_{i+1} - x_i = \delta$, $x_1 = -M$ or $x_n = M$. This continuous transformation maps Δ_δ into $\mathcal{C}_\delta$, a subset of $\mathcal{C}$. Let the boundary of Δ_δ be denoted by B_δ, and the image of the boundary B_δ in $\mathcal{C}_\delta$ by D_δ. Since the mapping of B_δ to D_δ is 1.1 by virtue of lemma 3, a modified form of Brouwer's fixed point theorem implies that the full interior of $\mathcal{C}_\delta$ is covered by Δ_δ. In other words, every point in the set bounded by D_δ is the image of some point in Δ_δ. The next step is to show that the boundary of $\mathcal{C}$ is uniformly close to D_δ if δ is sufficiently small and M is sufficiently large. Indeed, this will follow from the fact that the boundary as characterized in lemma 4 corresponds to points where at least one i has $x_i = x_{i+1}$, or some j has $x_j = +\infty$, or $x_j = -\infty$. Now we choose δ sufficiently small so that

$$\int_y^{y+\delta} e^{\omega_i x} r(x)\, dx \leqq \epsilon \tag{40}$$

uniformly for all y and $i = 1, \cdots, n$ and finally M so large that

$$\int_{-\infty}^{-M} + \int_M^{\infty} e^{\omega_i x} r(x)\, dx \leqq \epsilon\,, \qquad i = 1, \cdots, n\,. \tag{41}$$

It follows readily from (40) and (41) that these choices of δ and M bring D_δ uniformly close to the boundary D of $\mathcal{C}$. Thus as $\delta \to 0$ and $M \to \infty$, D_δ approaches D, and, hence, every point interior to $\mathcal{C}$ is covered by some point of Δ_δ for suitable δ. This provides a new proof of theorem 4. In this special case where $\mu(x)$ is absolutely continuous and the strategies φ in $\mathcal{M}$ are completely determined by a set of n numbers x_i, with $x_i \leqq x_{i+1}$, then the set $\mathcal{C}$ can be parameterized by the n variables x_i. To every interior point of $\mathcal{C}$ there exists a unique set of $\{x_i\}$, $x_i < x_{i+1}$, such that if φ is determined by $\{x_i\}$ then $(\int \exp(\omega_i x)\varphi(x)\, d\mu(x))$ is equal to the prescribed point.

7. Construction of a distribution for nature with a given Bayes solution in the class of strategies of $\mathcal{M}$

Let the interval between ω_i^0 and ω_{i+1}^0 be denoted by I_i. A special convention is made where if $\omega_i^0 = \omega_{i+1}^0$, then I_i is taken to be the point itself. Let $\varphi = (\varphi_1, 1 - \varphi_1)$ be a member of $\mathcal{M}$ given by

$$\varphi_1(x) = \begin{cases} 1\,, & x_{2i} < x < x_{2i+1} \\ \lambda_i\,, & x = x_i \\ 0\,, & \text{elsewhere}\,. \end{cases} \tag{42}$$

Suppose to begin with that x_i are distinct. Choose any points ω_i from I_i, $i = 0,\cdots, n$, and consider the system of equations

$$\sum_{i=0}^{n} \lambda_i \left[L_1(\omega_i) - L_2(\omega_i)\right] e^{\omega_i x_j} = 0\,, \qquad j = 1,\cdots, n \tag{43}$$

Consider the determinantal equation

$$0 = \begin{vmatrix} [L_1(\omega_0)-L_2(\omega_0)]\, e^{\omega_0 x_1}, & [L_1(\omega_1)-L_2(\omega_1)]\, e^{\omega_1 x_1}, & \cdots, & [L_1(\omega_n)-L_2(\omega_n)]\, e^{\omega_n x_1} \\ [L_1(\omega_0)-L_2(\omega_0)]\, e^{\omega_0 x_2}, & [L_1(\omega_1)-L_2(\omega_1)]\, e^{\omega_1 x_2}, & \cdots, & [L_1(\omega_n)-L_2(\omega_n)]\, e^{\omega_n x_2} \\ \cdots, & \cdots, & \cdots, & \cdots \\ [L_1(\omega_0)-L_2(\omega_0)]\, e^{\omega_0 x_n}, & \cdots, & \cdots, & \cdots \\ a_0, & a_1, & \cdots, & a_n \end{vmatrix}. \tag{44}$$

If one expands by the last row and denotes the cofactor of a_i by A_i, then one obtains

$$\sum_{i=0}^{n} a_i A_i = 0\,. \tag{45}$$

This equation is obviously satisfied for the choices of the vector $(a_i) = ([L_1(\omega_i) - L_2(\omega_i)]\exp(\omega_i x_j))$ for each j. A unique solution, except for a constant multiple, of (43) is obtained by choosing $\lambda_i = A_i$. It is evident that

$$A_i = (-1)^{n+1+i} \prod_{\substack{j\neq i\\ j=0}}^{n} |L_1(\omega_j) - L_2(\omega_j)|\, C_i\,, \tag{46}$$

where

$$C_i = \begin{vmatrix} e^{\omega_0 x_1}, & -e^{\omega_1 x_1}, \cdots, & (-1)^{i-1} e^{\omega_{i-1} x_1}, & (-1)^{i+1} e^{\omega_{i+1} x_1}, \cdots, & (-1)^n e^{\omega_n x_1} \\ \cdots, & \cdots, \cdots, & \cdots, & \cdots, \cdots, & \cdots \\ e^{\omega_0 x_n}, & -e^{\omega_1 x_n}, \cdots, & (-1)^{i-1} e^{\omega_{i-1} x_n}, & (-1)^{i+1} e^{\omega_{i+1} x_n}, \cdots, & (-1)^n e^{\omega_n x_n} \end{vmatrix}. \tag{47}$$

It is also clear as the ordinary Vandermonde determinant is positive that C_i alternate in sign. Thus A_i all have the same sign. Define $\mu_i = kA_i/\beta(\omega_i)$ so that $\mu_i > 0$ and $\sum \mu_i = 1$. Let $F(\omega)$ be a distribution concentrating at ω_i with mass μ_i.

Equation (43) implies that

$$\lambda(x) = \int e^{\omega x} [L_1(\omega) - L_2(\omega)] \beta(\omega) \, dF(\omega) \tag{48}$$

vanishes at $x = x_i$. By lemma 1 it cannot vanish anywhere else and thus the optimum strategy for the statistician is to take action 1 when $\lambda(x) < 0$ and action 2 when $\lambda(x) > 0$, so that the unique Bayes solution coincides with that of (42). If two of the x_j coincide, say $x_j = x_{j+1}$, then we replace the equation of (43) involving x_{j+1} by

$$\sum_{i=0}^{n} \lambda_i [L_1(\omega_i) - L_2(\omega_i)] \, \omega_i e^{\omega_i x_j} = 0 , \tag{49}$$

which represents the derivative with respect to x evaluated at x_j of the previous equation. From here on the proof proceeds as before. If several x_j are equal, then one must use similar linear equations of the form

$$\sum_{i=0}^{n} \lambda_i [L_1(\omega_i) - L_2(\omega_i)] \, \omega_i^r e^{\omega_i x_j} = 0 \tag{50}$$

for some suitable r. Finally, to deal with the case where several x_j are $\pm \infty$, one uses fewer choices of ω_i. The details are omitted.

Summing up, we have shown

THEOREM 5. *If φ denotes any strategy of $\mathcal{M}$, then there exists a distribution for nature concentrating at most at $n + 1$ points against which the Bayes strategy is unique and equal to φ except possibly at the n critical values x_i. For x_i the Bayes strategy is undetermined.*

8. Pólya type distributions

A class of distributions $P(x, \omega)$ of a real random variable x in X depending on a real parameter ω in Ω is said to belong to class $\mathcal{P}_n$ if

$$P(x, \omega) = \beta(\omega) \int_{-\infty}^{x} p(x, \omega) \, d\mu(x) , \tag{51}$$

and whenever $x_1 < x_2 < \cdots < x_m$, $\omega_1 < \omega_2 < \cdots < \omega_m$ with x_i in X and ω_i in Ω, then

$$P\begin{pmatrix} x_1, \cdots, x_m \\ \omega_1, \cdots, \omega_m \end{pmatrix} = \det p(x_i, \omega_j) \geqq 0 , \tag{52}$$

for every $m \leqq n$. The measure μ is taken to be σ-finite and $p(x, \omega)$ is assumed continuous in each variable. If strict inequality holds in (52) then we say that P belongs properly to class $\mathcal{P}_n$. If P belongs to $\mathcal{P}_n$ for every n, then we say that $P \in \mathcal{P}_\infty$. We shall also frequently use the terminology that the density $p(x, \omega)$ belongs to $\mathcal{P}_\infty$ in place of P in $\mathcal{P}_\infty$.

Pólya had introduced special positive functions $g(x)$ such that if we set $p(x, \omega) = g(x - \omega)$, then p satisfies the set of inequalities in (52) for every m [3]. The context in which Pólya studied such functions is related to the possibility of approximating continuous functions by polynomials with real zeros. Later Schoenberg investigated the relationship of such Pólya frequency functions $g(x)$ with the theory of variation diminishing transformation [4]. It is these results that are relevant to the statistical analysis presented below.

The type of kernel $p(x, \omega)$ obeying (52) dealt with here represents a generalization of Pólya frequency functions which retains all the necessary variation diminishing properties necessary to enable us to extend the statistical theory developed in connection with the exponential family to Pólya type distributions. Some examples of Pólya frequency functions are as follows:

$$(53)\quad \begin{aligned} &\text{(a)}\quad e^{-x^2}, \qquad \text{(b)}\quad \operatorname{sech} x\,, \qquad \text{(c)}\quad e^{x-e^x},\\ &\text{(d)}\quad \begin{cases} x^n e^{-x}, & \text{for } x \geqq 0\,,\ n \text{ a nonnegative integer}\\ 0\,, & x<0 \end{cases}\\ &\text{(e)}\quad e^{-|x|}\,, \qquad \text{(f)}\quad \begin{cases} \sum_{v=-\infty}^{\infty} (-1)^v e^{-xv^2}, & x>0\\ 0, & x \leqq 0\,. \end{cases} \end{aligned}$$

The corresponding densities depending on a real parameter ω become

$$(54)\quad \begin{aligned} &(\text{a}')\quad p(x,\omega) = e^{-(x-\omega)^2},\\ &(\text{b}')\quad p(x,\omega) = \operatorname{sech}(x-\omega)\\ &(\text{c}')\quad p(x,\omega) = e^{x-\omega-e^{x-\omega}},\\ &(\text{d}')\quad p(x,\omega) = \begin{cases} (x-\omega)^n e^{-(x-\omega)}, & x \geqq \omega\,,\ n \text{ a nonnegative integer}\\ 0\,, & x<\omega\,, \end{cases}\\ &(\text{e}')\quad p(x,\omega) = e^{-|x-\omega|},\\ &(\text{f}')\quad p(x,\omega) = \begin{cases} -\sum_{v=-\infty}^{\infty} (-1)^v v^2 e^{-(x-\omega)v^2}, & x>\omega\\ 0\,, & x<\omega\,. \end{cases} \end{aligned}$$

Thus all the densities of a′ through f′ furnish examples of distributions of class ρ_∞.

Another very important class of examples of distributions in ρ_∞ which are not derived from Pólya frequency functions is obtained by taking $p(x, \omega) = e^{x\omega}$. The corresponding distributions constitute the exponential family and belong properly to ρ_∞.

In order to verify that other common densities occurring in statistical theory, such as the noncentral t and noncentral F, also belong to ρ_∞ we shall need the following lemma.

LEMMA 5. *Let*

$$(55)\qquad r(x,\omega) = \int p(x,t)\, q(t,\omega)\, d\sigma(t)$$

where $p(x, t)$ belongs to ρ_m and $q(t, \omega)$ is in ρ_n with σ a positive measure. Then $r(x, \omega)$ belongs to $\rho_{\min(m,n)}$.

This lemma is an immediate consequence of the identity

$$(56)\; R\begin{pmatrix} x_1,\cdots,x_k\\ \omega_1,\cdots,\omega_k\end{pmatrix} = \frac{1}{k!}\int\cdots\int_T\int P\begin{pmatrix} x_1,\cdots,x_k\\ t_1,\cdots,t_k\end{pmatrix} Q\begin{pmatrix} t_1,\cdots,t_k\\ \omega_1,\cdots,\omega_k\end{pmatrix} d\sigma(t_1)\cdots d\sigma(t_k)\,,$$

where T is the subset $(t_1, \cdots, t_k)$ in Euclidean k space where $t_1 \leqq t_2 \leqq \cdots \leqq t_k$. For a proof of the identity see Pólya, Szegö in [5], p. 48, problem 68.

As an application of this lemma it will now be shown that the noncentral F distribution belongs to ρ_∞. To this end, let $G = u/v$ where u and v are independently distributed according to the noncentral χ^2 distribution with r degrees of freedom and the central χ^2 distribution with s degrees of freedom, respectively. The probability distribution of G with λ the noncentrality parameter has the form

$$(57)\qquad P(G\mid\lambda) = \sum_{m=0}^{\infty} e^{-\lambda}\left(\frac{G}{1+G}\right)^{(r+2m-2)/2}\left(\frac{1}{1+G}\right)^{(s+2)/2}\frac{\lambda^m}{m!}\,k(r, s, m),$$

where $k > 0$. Letting $u = G/(1 + G)$ which is an increasing function of G for $G > 0$, we have

$$(58)\quad p(G\mid\lambda) = p^*(u\mid\lambda) = e^{-\lambda}\varphi(u)\sum_{m=0}^{\infty} u^m\frac{\lambda^m}{m!}\,k(m)$$

$$= e^{-\lambda}\varphi(u)\int_{-\infty}^{\infty} e^{x\log u}e^{x\log\lambda}\,d\psi(x) = e^{-\lambda}\varphi(u)\,f(u, \lambda),$$

where $\psi(x)$ has a jump $k(m)/m!$ for $x = m$. From lemma 5 as $e^{x\log\lambda}$ is in ρ_∞, we infer that $f(u, \lambda)$ belongs to ρ_∞. That the same is true of $p^*(u, \lambda)$ and hence of $p(G, \lambda)$ readily follows.

A similar argument can be carried through to show that the noncentral generalized Hotelling t-distribution also belongs to ρ_∞.

The classical noncentral t-distribution is analyzed as follows: If z is distributed $N(\delta, 1)$, and y distributed chi-square with k degrees of freedom, then the distribution of $x = z/\sqrt{y/k}$ has a density of the form

$$(59)\qquad p(x\mid\delta) = c\int_0^{\infty} e^{-[(\sqrt{y}/\sqrt{k})x-\delta]^2/2}y^{(k-1)/2}e^{-y/2}dy.$$

For $x > 0$, let $u = (\sqrt{y}/\sqrt{k})x$; then we get

$$(60)\qquad p(x\mid\delta) = c'e^{-\delta^2/2}x^{-(k+1)}\int_0^{\infty} e^{\delta u}e^{(-k/2x^2)u^2}d\sigma(u).$$

Setting $\xi = -k/2x^2$, which is a monotone increasing function of x for $x > 0$, we get

$$(61)\qquad p(x\mid\delta) = q(\xi\mid\delta) = c'e^{-\delta^2/2}\varphi(\xi)\int_0^{\infty} e^{\delta u}e^{\xi u^2}d\sigma(u),$$

where $\varphi(\xi) > 0$, which on account of lemma 5 satisfies the determinant criterion (52) for $x_i \geqq 0$ and δ_j arbitrary. A similar result can be obtained for the case of $x_i \leqq 0$. A continuity argument enables us to verify condition (52) for any choice of x_i.

The above discussion has shown that the collection of distributions in ρ_∞ is quite large and includes most of the common distributions occurring in statistical applications.

The statement that $p(x, \omega)$ belongs to ρ_1 is synonymous with the fact that $p(x, \omega)$ can qualify as a density function. The condition of (52) for $m = 2$ is equivalent to the fact that $p(x, \omega)$ has a monotone likelihood ratio. The fundamental property of Pólya type distributions is the following:

LEMMA 6. *If $p(x, \omega)$ belongs properly to ρ_∞ and h is continuous and changes sign at most n times, then*

$$(62)\qquad g(x) = \int p(x, \omega)\,h(\omega)\,dF(\omega)$$

has at most n distinct zeros or is identically zero. The function g is identically zero if and only if the spectrum of F is contained in the set of zeros of h.

PROOF. Suppose $h(\omega)$ changes sign at $\omega = \omega_1^0, \omega_2^0, \cdots, \omega_n^0$. Define

$$(63) \qquad G_i(x) = \left| \int_{\omega_i^0}^{\omega_{i+1}^0} p(x, \omega)\, h(\omega)\, dF(\omega) \right|, \qquad i = 0, \cdots, n,$$

where $\omega_0^0 = -\infty$ and $\omega_{n+1}^0 = +\infty$. If x_i is chosen arbitrarily but satisfying $x_1 < x_2 < \cdots < x_{n+1}$, then the det $[G_i(x_j)]$ equals

$$(64) \quad \int\int_{\omega_i^0}^{\omega_{i+1}^0} \cdots \int_{\omega_0^0}^{\omega_1^0} P\begin{pmatrix} x_1, \cdots, x_{n+1} \\ \omega_1, \cdots, \omega_{n+1} \end{pmatrix} | h(\omega_1) | \cdots | h(\omega_{n+1}) | \, dF(\omega_1) \cdots dF(\omega_{n+1})$$

and thus det $[G_i(x_j)] > 0$, provided only the spectrum of F is not contained in the zeros of $h(\omega)$. Moreover, as g is equal to either $+$ or $-$ the function

$$(65) \qquad G(x) = \sum_{i=0}^{n} (-1)^i G_i(x)$$

and G has at most n distinct zeros (since det $[G_i(x_j)] > 0$ for $n + 1$ x's), we conclude that g changes sign at most n times. Q.E.D.

Remark 1. If h is piece-wise continuous, then the conclusion of lemma 6 holds. This can be seen by approximating to h by continuous functions with the same number of sign changes.

Remark 2. The result of lemma 6 also is valid for the case when $p(x, \omega)$ does not necessarily belong properly to $\mathcal{P}_\infty$, but we require instead that for any n and any prescribed $\omega_1 < \cdots < \omega_n$ there exist a set of $x_1 < \cdots < x_n$ which may depend on $(\omega_1, \cdots, \omega_n)$ such that

$$(66) \qquad P\begin{pmatrix} x_1, \cdots, x_n \\ \omega_1, \cdots, \omega_n \end{pmatrix} > 0\,.$$

To demonstrate this fact we form for $\sigma > 0$

$$(67) \qquad q_\sigma(x, \omega) = \int \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-(x-u)^2/2\sigma^2} p(u, \omega)\, du\,.$$

By lemma 5 and our assumption, $q_\sigma(x, \omega)$ belongs properly to $\mathcal{P}_\infty$. Thus the number of variation of signs of

$$(68) \qquad g_\sigma(x) = \int q_\sigma(x, \omega)\, h(\omega)\, dF(\omega)$$

is at most $n =$ var (h). But, $q_\sigma(x, \omega) \to p(x, \omega)$ as $\sigma \to 0$ and $g_\sigma(x) \to g(x)$ uniformly in any finite x interval. Consequently, $g(x)$ also has at most n changes of sign.

Remark 3. If the number of sign changes of g is $n = V(h)$, then g and h change signs in the same order. The proof of this fact is involved and will be omitted.

In view of lemma 6, remarks 1 to 3, and its relationship to lemma 1, it now follows that all the preceding results established for the exponential class of distribution are valid for Pólya type distributions which properly belong to $\mathcal{P}_\infty$. The proofs are essentially unaltered. Detailed proofs can be found in [6].

We sum up the results in the form of a theorem. The assumptions on the loss functions are the same as given in section 3.

THEOREM 6. *If $P(x, \omega)$ is a class of distributions of a real variable x and a real parameter ω belonging properly to $\mathcal{P}_\infty$, then the strategies in $\mathcal{M}$ constitute a minimal complete class.*

Moreover, every member of $\mathcal{M}$ *is unique Bayes against a distribution F for nature which has* $n+1$ *points in its spectrum.*

A more general class of theorems can be developed for distributions of class $\mathcal{P}_m$, m finite. In this case it is further necessary to assume that the loss functions $L_1 - L_2$ change signs at most $m-1$ times. The form of the complete classes can be characterized; admissibility and the other investigations can be extended requiring only slight changes of our methods. For example, if $P(x, \omega)$ belongs to $\mathcal{P}_2$, then we must require for our theory that $L_1(\omega) - L_2(\omega)$ change signs at most once. This is precisely the case of distributions with a monotone likelihood ratio which is discussed in [1]. This generalized theory and a further study of the decision problems concerning more than two actions for Pólya type distributions will be presented elsewhere.

REFERENCES

[1] S. Karlin and H. Rubin, "The theory of decision procedures for distributions with monotone likelihood ratio," *Annals of Math. Stat.*, Vol. 27 (1956).

[2] D. Blackwell and M. A. Girshick, *Theory of Games of Statistical Decisions*, New York, John Wiley and Sons, 1954.

[3] G. Pólya, "Über Annäherung durch Polynome mit lauter reellen Wurzeln," *Rendiconti del Circ. Mat. di Palermo*, Vol. 36 (1913), pp. 279–295.

[4] I. S. Schoenberg, "On Pólya frequency functions," *Jour. d'Analyse Mathématique*, Vol. 1 (1951), pp. 331–374.

[5] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, Vol. I, Berlin, Julius Springer, 1925.

[6] S. Karlin, "Pólya type distributions, II," submitted to the *Annals of Math. Stat.*