STATISTICS DEPARTMENT

THE GEORGE WASHINGTON UNIVERSITY

Topics in Information and Decision Processes

by

Minoru Sakaguchi

Visiting Professor of Statistics, 1964-1965

Foreword.

This publication is an expanded version of a series of lectures on Topics in Information and Decision Processes delivered by Professor Minoru Sakaguchi during the academic year 1964/65. Professor Sakaguchi's tenure as Visiting Professor of Statistics was jointly supported by The George Washington University and the Office of Naval Research under Contract Nonr (G)-00029-65. Publication of the lectures was made possible by funds made available by the Research Council of The George Washington University. We are grateful for all the support.

Acknowledgement is made of the editorial assistance of Mr. John Berterman and the fine typing of Mrs. Sue Golan.

S. Kullback
Professor of Statistics and
Director of Research under
contract Nonr (G) 00029-65.

Table of Contents

1. INDIVIDUAL DECISION-MAKING UNDER UNCERTAINTY

When an individual is required to make a decision in which a utility function for a problem is given, it occasionally happens that the true state of nature is not known to him and he has only an opinion regarding the probabilities of the relevant states of nature. Suppose that the latter are indexed by a variable x, to which probability density functions on some measure $\lambda(x)$ may be attributed, and that the opinion of the decision-maker is given by a probability density function $f(x)$. Let $\mathfrak{a}$ be the set of available decisions α. When the state of nature is x, let $u(\alpha, x)$ be the utility of decision α.

It is well-known that when the true state x is unknown to the decision-maker the expected-utility maxim would let him select that α maximizing

$$\int u(\alpha, x) f(x) \, d\lambda(x),$$

if he is a rational decision-maker.

Let us now introduce the concept of information structure. The <u>information structure</u> of the problem will be represented by a function $I(x)$ defined on the set of all possible states of nature. It is usually a non-stochastic many-to-one mapping defined over $\mathfrak{X}$ and sometimes it is a stochastic variable indexed by x. We shall exhibit several examples of conceivable information structures in the following.

(i) complete information (denoted by $\bar{I}(x)$)

$\bar{I}(x) = x$. The decision-maker is informed of the exact value of x.

(ii) null information (denoted by $I_0(x)$)

$I_0(x) = \mathcal{X}$. The value of $I_0(x)$ is identically equal to the whole space of the states of nature. No information is available to the decision-maker.

(iii) partition information (denoted by $I_{\mathcal{P}}(x)$)

Let

$$\mathcal{P} = \{R_1, \ldots R_N\}, \quad \mathcal{X} = \bigcup_{i=1}^{N} R_i, \quad R_i \cap R_j = \phi (i \neq j)$$

be a given finite or infinite decomposition of $\mathcal{X}$, i.e., $N = \infty$ is allowed.

Let

$$I_{\mathcal{P}}(x) = i, \quad \text{if} \quad x \in R_i \ (i=1, \ldots, N).$$

In this case the decision-maker is only informed which of the sets R_i the true state x of nature belongs to.

(iv) random information

Sometimes the decision-maker can observe the true state of nature only through the interference of some random distrubance. Suppose that the nature of the random disturbance is known and that it is represented by a probability density function $r(\cdot)$ on $\mathcal{X}$.

Let $I(x) = x + z$ where z is a random variable on x with the probability density function $r(z)$.

Intuitively speaking, the information structure $I(x)$ represents the information on which decision α is based. Let $A(I(x))$ denote any function of $I(x)$, which maps onto a, representing a decision rule for α. Let $f(x)\, d\lambda(x) = dF(x)$.

We shall next show that the value of the information structure $I(x)$ can be defined by the amount

$$v(I(x) \mid F(x)) \equiv \max_{A(\cdot)} \int u(A(I(x)), x)\, dF(x) - \max_{\alpha} \int u(\alpha, x)\, dF(x). \tag{1}$$

The above difference measures the advantage or profit obtained by a rational decision-maker under the information structure $I(x)$ compared to the case of complete ignorance.

<u>Theorem 1.</u> We have

$$v(\overline{I} \mid F) = \int \max_{\alpha} u(\alpha, x)\, dF(x) - \max_{\alpha} \int u(\alpha, x) dF(x) \tag{2}$$

$$v(I_{\rho} \mid F) = \max_{A(1), \ldots, A(N)} \sum_{i=1}^{N} \int_{R_i} u(A(i), x) dF(x) - \max_{\alpha} \int u(\alpha, x) dF(x) \tag{3}$$

(b) $0 = v(I_0 \mid F)$

<u>Proof:</u> (a) Since $\overline{I}(x) = x$ and

$$\max_{A(\cdot)} \int u(A(x), x) dF(x) = \int \max_{\alpha} u(\alpha, x) dF(x)$$

we have the stated equality (2). (3) is also easily verified from (1) and the definition of $I_{\rho}(x)$.

(b) Since $I_0(x)$ is independent of x, it follows by (1) that $v(I_0|F) = 0$. The restriction of the class of admissible decision functions to the smaller class $\{A(\cdot) \equiv \alpha \text{ identically } | \alpha \in u\}$ yields the proof of $v(I|F) \geq 0$. The last inequality $v(I|F) \leq v(\bar{I}|F)$ is verified from the easily proved relation

$$\max_{A(\cdot)} \int u(A(I(\ )), x)\, dF(x) \leq \int \max_{\alpha} u(\alpha, x)\, dF(x) .$$

Sometimes we denote the value of information structure by $v(I|F ; u)$ to emphasize the use of the utility function u. Then

<u>Corollary to Theorem 1.</u> <u>We have</u>

$$v(I_{\mathcal{P}'}|F ; u) \geq v(I_{\mathcal{P}}|F ; u) \text{ for all } F \text{ and } u,$$

<u>if and only if</u> $\mathcal{P}'$ <u>is a subpartition of</u> $\mathcal{P}$.

<u>Proof.</u> The if-part is evident from (3). The only-if-part follows from Example 3.

<u>Example 1.</u>

$$v(I|F(x_1, x_2)) = \max_{A(\cdot)} \int dF(x_1) \int u(A(x_1) ; x_1, x_2)\, dF(x_2|x_1)$$

$$= \max_{\alpha} \int u(a ; x_1, x_2)\, dF(x_1, x_2)$$

where we have set $dF(x_1, x_2) = dF(x_1)dF(x_2|x_1)$.

For a simple illustration assume that $a = (-\infty, \infty)$, $\mathfrak{X} = (-\infty, \infty)^2$ and $u(a ; x_1, x_2) = -(a - (x_1 + x_2)/2)^2$. Then we obtain the value

$$v(I|F(x_1, x_2)) = V\left(\frac{X_1+X_2}{2}\right) - \int V\left(\frac{x_1 + X_2}{2} \,\middle|\, x_1\right) dF(x_1)$$

and the optimal decision rule

TABLE 1

Examples of Values of Information Structures

Utility function	Information structure: Complete information $v(\bar{I}\mid F)$	Partitional information: $v(I_{\mathcal{P}}\mid F)$	Partitional information: Optimal decision rule $A^*(i)$
$\mathfrak{A} = \mathfrak{X} = (-\infty, \infty)$ $u = -(x - a)^2$	$\sigma^2(F)$ i.e., the variance of F	$\sigma^2(F) - \sum_{i=1}^{N} (\int_{R_i} dF)\, \sigma^2(F\mid R_i)$ where $\sigma^2(F\mid R_i)$ is the conditional variance of F given that $x \in R_i$	$\mu(F\mid R_i)$ i.e., the conditional mean of F given that $x \in R_i$
$\mathfrak{X} = [\alpha, \beta]$ $(\alpha < 0 < \beta)$ $\mathfrak{A} = [-1, 1]$ $u = ax$	$\int_{\alpha}^{\beta} \|x\| dF - \left\| \int_{\alpha}^{\beta} x\, dF \right\|$	$\sum_{i=1}^{N} \left(\int_{R_i} dF \right) \|\mu(F\mid R_i)\| - \mu(F)$ where $\mu(F)$ is the mean of F	$\operatorname{sgn} \int_{R_i} x\, dF$
$\mathfrak{A} = \mathfrak{X} = (-\infty, \infty)$ $u = -\|a - x\|$	$\int \|x - \tilde{\mu}(F)\| dF$ where $\tilde{\mu}(F)$ is the median of F	$\int \|x - \tilde{\mu}(F)\| dF - \sum_{i=1}^{N} \min_{a_i} \int \|a_i - x\| dF$	The number a_i^* satisfying $\int_{R_i} \|a_i - x\| dF \rightarrow \min_{a_i}$

$$A^*(x_1) = \int \frac{x_1 + x_2}{2} \, dF(x_2 | x_1)$$

where $V(\cdot)$ and $V(\cdot | x_1)$ are respectively the variance and the conditional variance of the random variable. For comparison it should be remarked here that in the complete-information case of $I(x_1, x_2) = (x_1, x_2)$ we have $v(\bar{I} | F) = V(X_1 + X_2)/2)$.

Example 2. Statistical decision problems.

We shall consider the statistical decision problem of discriminating among a homogeneous set of k distributions $F_1, \ldots, F_k$ from our information viewpoint.

Let $\mathfrak{A} = \mathfrak{X} = \{1, \ldots, k\}$ and

$$u(a, x) = \begin{cases} r_x, & \text{if } a = x \quad (x = 1, \ldots, k) \\ 0, & \text{if } a \neq x . \end{cases}$$

The information structure is given by a random information $I(x) = y$, where

$$K(a \mid y) \equiv \sum_{x=1}^{k} \xi_x u(a, x)\, dF_x(y) = \xi_a r_a dF_a(y) \rightarrow \max_{1 \leqq a \leqq k} (\xi_a r_a)$$

With

$$R_j \equiv \left\{ y \mid j \text{ is the minimum number of } j' \text{ such that } \xi_{j'} r_{j'} dF_{j'}(y) = \max_{1 \leqq a \leqq k} \xi_a r_a dF_a(y) \right\}$$

we have

$$A^*(y) = j, \text{ if } y \in R_j \qquad (j = 1, \ldots, k)$$

and

$$v(I \mid \xi ; u) = \sum_{j=1}^{k} r_j \xi_j \int_{R_j} dF_j(y) - \max_{1 \leqq a \leqq k} (\xi_a r_a)$$

One of the more instructive examples is the problem of point estimation. Let $\mathfrak{A} = \mathfrak{X} = (-\infty, \infty)$ and $u(a, x) = -(a-x)^2$. Let the information structure be given by a random information $I(x) = y$ where

$$\Pr \left\{ y < I(x) \leqq y + dy \mid X = x \right\} = p(y \mid x)dy .$$

Denote by $dF(x)$ the prior probability distribution over $\mathfrak{X}$. Then since

$$\int dF(x) \int u(A(y), x) p(y \mid x) dy = - \int dF(x) \int (A(y) - x)^2 p(y \mid x) dy =$$

$$= - \int q(y) dy \int (A(y) - x)^2 p(x \mid y) dx ,$$

where $q(y) = \int p(y|x)dF(x)$ and $p(x|y)dx = p(y|x)dF(x)/q(y)$, we have the optimal decision rule $A^*(y) = \int xp(x|y)dx = E[X\ y]$, and the value of information

$$v(I|F\ ;\ u) = E[u(A^*(Y), X)] - \max_a E[u(a, X)] = V(X) - E[V(X|Y)]\ .$$

<u>Example 3.</u> (Marschak and Radner, 1958, Chapter 3).

Suppose the price of a stock can change from this week to the next, by any amount in $\mathfrak{X} = \{-6, -5, \ldots, -1, 1, \ldots, 5, 6\}$. Suppose you can use the services of either of two informants A and B, each a faultless predictor of stock prices. Informant A sends only two kinds of messages: (1) stock will fall, (2) stock will rise. Informant B sends three kinds of messages: (1) stock will fall by 3 or more, (2) stock will move by 2 or less, (3) stock will rise by 3 or more. Your decision can be denoted by the three-valued variable a with

$$a = \begin{cases} +1 & \text{(buying)} \\ 0 & \text{(do nothing)} \\ -1 & \text{(selling)}\ , \end{cases}$$

so that the payoff function is given by

$$u(a, x) = ax \qquad (a = -1, 0, 1\ ;\ x = -6, \ldots, -1, 1, \ldots, 6).$$

Let us assume that the prior distribution $dF^0(x)$ over $\mathfrak{X}$ is uniform: $\Pr\{X = x\} = \frac{1}{12}$ for all $x \in \mathfrak{X}$. Informants A and B give you the partition information $I_{\mathcal{P}_A}$ and $I_{\mathcal{P}_B}$, respectively, where

$$\mathcal{P}_A = \{\{-6, \ldots, -1\}, \{1, \ldots, 6\}\}.$$

$$\mathcal{P}_B = \{\{-6, \ldots, -3\}, \{-2, -1, 1, 2\}, \{3, \ldots, 6\}\}.$$

The messages given by the information structures are

$$I_{\mathcal{P}_A}(x) = \begin{cases} y_+, & \text{if } x = 1, \ldots, 6 \\ y_-, & \text{if } x = -6, \ldots, -1 \end{cases}$$

$$I_{\mathcal{P}_B}(x) = \begin{cases} y_+, & \text{if } x = 3, \ldots, 6 \\ y_0, & \text{if } x = -2, \ldots, 2 \\ y_-, & \text{if } x = -6, \ldots, -3 \end{cases}$$

respectively.

We can find that

	Optimal decision rule	$v(I_{\mathcal{P}} \mid F^0 ; u)$
For $\mathcal{P}_A$	$A^*(y) = \begin{cases} 1, & \text{if } y = y_+ \\ -1, & \text{if } y = y_- \end{cases}$	7/2
For $\mathcal{P}_B$	$A^*(y) = \begin{cases} 1, & \text{if } y = y_+ \\ -1, & \text{if } y = y_- \\ \text{arbitrary}, & \text{if } y = y_0 \end{cases}$	3

Thus A can afford you the more valuable information than B, although the amounts of information (in Wiener-Shannon's sense) they provide are log 2 and log 3, respectively. This shows that information value and information amount do not necessarily go together.

selective information strategic information

Let us next consider a new payoff function

$$u'(a, x) = \begin{cases} ax - 2, & a = \pm 1 \\ 0, & a = 0 \end{cases}$$

(which means that you have to pay 2 on purchase or sale). For this payoff function we can find that

	Optimal decision rule	$v(I_\theta \mid F^0 ; u')$
For θ_A	$A^*(y) = \begin{cases} 1, & \text{if } y = y_+ \\ -1, & \text{if } y = y_- \end{cases}$	$\frac{3}{2}$ (= 1.5)
For θ_B	$A^*(y) = \begin{cases} 1, & \text{if } y = y_+ \\ -1, & \text{if } y = y_0 \\ -1, & \text{if } y = y_- \end{cases}$	$\frac{5}{3}$ (= 1.67)

This shows that the ranking of information structures by the information values depends in general on the payoff function used.

This example raises naturally the question whether there are pairs of information structures such that the ranking of their information values is not influenced by the payoff function. The corollary to Theorem 1 shows that such "objective" ranking is possible if one information structure is a subpartition of the other. Moreover we may have the question: Is there a class U of payoff functions such that there exists a numerical function $K(I|F)$ with the property that

$$v(I' \mid F ; u) \geqq v(I'' \mid F ; u) \text{ for all } u \in U \iff K(I' \mid F) \geqq K(I'' \mid F) ?$$

A positive partial answer to this question is provided by Example 3.

Example 4.

Let the decision-making problem under uncertainty be given by the following payoff matrix:

$$u(a, x) = \begin{array}{c|cc} \mathfrak{a} \diagdown \mathfrak{X} & x_1 & x_2 \\ \hline a_1 & r_1 & 0 \\ a_2 & 0 & r_2 \end{array} \qquad (r_1, r_2 > 0)$$

We easily see that for any prior distribution ξ, with the probability $\xi_1 = 1-\xi_2$ for x,

$$A^*(\bar{I}(x)) = A^*(x) = \begin{cases} a_1, & \text{if } x = x_1 \\ a_2, & \text{if } x = x_2 , \end{cases}$$

$$v(\bar{I} \mid \xi) = r_1\xi_1 + r_2\xi_2 - \max(r_1\xi_1, r_2\xi_2) = \min(r_1\xi_1, r_2\xi_2).$$

The maximum $\max\limits_{0 \leqq \xi_1 \leqq 1} v(\bar{I} \mid \xi)$ is attained by $\xi_1^* = \dfrac{r_2}{r_1+r_2}$, and equals $\dfrac{r_1 r_2}{r_1+r_2}$. Using the above fact let us consider the following two payoff matrices:

$$u'(a,x) = \begin{array}{c|ccc} \mathfrak{a} \diagdown \mathfrak{X} & x_1 & x_2 & x_3 \\ \hline a_1 & r & 0 & 0 \\ a_2 & 0 & r & r \end{array} \qquad (r > 0)$$

$$u''(a,x) = \begin{array}{|ccc|} r & r & 0 \\ 0 & 0 & r \end{array}$$

For any prior distribution ξ, with the probability $\xi_2 = \frac{1}{3}$ for x_2, we have

$$v(\bar{I}\,|\,\xi\,;\,u') = r\,\min(\xi_1,\ \tfrac{1}{3} + \xi_3).$$

$$v(\bar{I}\,|\,\xi\,;\,u'') = r\,\min(\xi_1 + \tfrac{1}{3},\ \xi_3).$$

The maximum values $\max\limits_{0 \leq \xi_1 \leq \frac{2}{3}} v(\bar{I}\,|\,\xi\,;\,u')$ and $\max\limits_{0 \leq \xi_1 \leq \frac{2}{3}} v(\bar{I}\,|\,\xi\,;\,u'')$ are attained by $\xi_1^* = \frac{1}{2}$ and $\frac{1}{6}$, respectively, not by $\frac{1}{3}$. The highest degree of uncertainty does not correspond to the highest information value to be paid to dispelling it.

Example 5.

Suppose that a dishonest expert tells that a stock will rise or fall. Suppose the probability that the expert tells the truth is p_1. The state of nature is denoted by x_2, with

$$x_2 = \begin{cases} 1, & \text{if stock rises,} \\ -1, & \text{if stock falls.} \end{cases}$$

The information structure is given by a random information $I(x_2) = y$, where

$$\Pr\{Y = 1\,|\,X_2 = 1\} = \Pr\{Y = -1\,|\,X_2 = -1\} = p_1,$$

$$\Pr\{Y = -1\,|\,X_2 = 1\} = \Pr\{Y = 1\,|\,X_2 = -1\} = q_1 = 1 - p_1.$$

The payoff function is given by

$u(a, x_2) =$

a \ x_2	$x_2 = 1$	$x_2 = -1$
$a=a_+$	r_1	s_2
$a=a_-$	s_1	r_2

$s_i < r_i$, $i = 1, 2$).

Let the prior probabilities over $\mathfrak{X}$ be $Pr(X_2=1) = p_2 = 1-q_2$. For a given message y, the function $A^*(y)$ is optimal if it yields the maximum conditional expected payoffs. That is,

$$A^*(1) = \begin{cases} a_+ \\ a_- \end{cases}, \text{ according as}$$

$$r_1 Pr.\{X_2=1 \mid Y=1\} + s_2 Pr.\{X_2=-1 \mid Y=1\} \begin{Bmatrix} > \\ < \end{Bmatrix} s_1 Pr.\{X_2=1 \mid Y\}$$

$$+ r_2 Pr.\{X_2 = -1 \mid Y=1\}, \text{ i.e. } p_1 \begin{Bmatrix} < \\ > \end{Bmatrix} \frac{q_2(r_2-s_2)}{p_2(r_1-s_1)+q_2(r_2-s_2)}$$

$$A^*(-1) = \begin{cases} a_+ \\ a_- \end{cases}, \text{ according as}$$

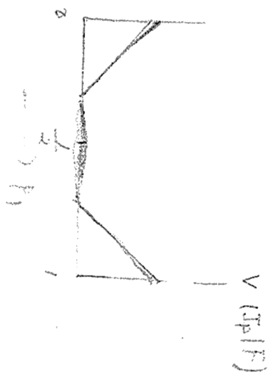
$$r_1 Pr.\{x_2=1 \mid Y=-1\} + s_2 Pr.\{X_2=-1 \mid Y=-1\} \begin{Bmatrix} > \\ < \end{Bmatrix} s_1 Pr.\{X_2=1 \mid Y=-1\}$$

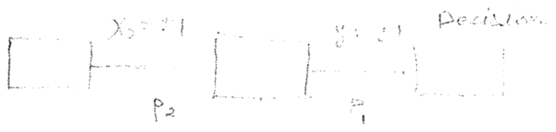
$$+ r_2 Pr.\{X_2=-1 \mid Y=-1\} \text{ i.e., } p_1 \begin{Bmatrix} < \\ > \end{Bmatrix} \frac{p_2(r_1-s_1)}{p_2(r_1-s_1)+q_2(r_2-s_2}$$

Combining these results we find the conditions on p_1 under which each of the four decision rules is optimal. We have the following table:

Cases	Optimal decision rule	Expected Payoff $E[u(A^*(Y), X_2)]$	Information values $v(I_p \mid F) = E[u(A^*(Y), X_2)] - \max_a E[u(a, X_2)]$
$A < B$ and $\left\|p_1 - \frac{1}{2}\right\| < \frac{\|A-B\|}{2(A+B)}$	$A^*(y) \equiv a_-$	$r_2q_2 + s_1p_2$	0
$A \geqq B$ and "	$A^*(y) \equiv a_+$	$s_2q_2 + r_1p_2$	0
$p_1 \leqq \frac{1}{2}$ and $\left\|p_1 - \frac{1}{2}\right\| \geqq \frac{\|A - B\|}{2(A+B)}$	$A_*(y) = \begin{cases} a_-, & \text{if } y = 1 \\ a_+, & \text{if } y = -1 \end{cases}$	$r_2q_1q_2 + s_1p_1p_2 + s_2p_1q_2 + r_1q_1p_2$	$(\frac{1}{2} - p_1)(A+B) - \frac{\|A-B\|}{2}$
$p_1 \geqq \frac{1}{2}$ and $\left\|p_1 - \frac{1}{2}\right\| \geqq \frac{\|A - B\|}{2(A+B)}$	$A^*(y) = \begin{cases} a_+, & \text{if } y = 1 \\ a_-, & \text{if } y = -1 \end{cases}$	$s_2q_{-1}q_2 + r_1p_1p_2 + r_2p_1q_{-2} + s_1q_1p_2$	$(p_1 - \frac{1}{2})(A+B) - \frac{\|A-B\|}{2}$

where $A \equiv p_2(r_1 - s_1)$ and $B \equiv q_2(r_2 - s_2)$.





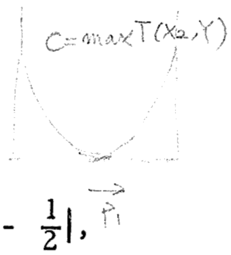
On the other hand, consider the informant as a channel whose input is x_2, and whose output is y. The transmitted information is equal to

$$T(X_2 : Y) = H(Y) - H(Y|X_2)$$

$$= \{-(p_1p_2+q_1q_2)\log(p_1p_2+q_1q_2) - (p_1q_2+q_1p_2)\log(p_1q_2+q_1p_2)\}$$

$$- \{-p_1 \log p_1 - q_1 \log q_1\} .$$

The capacity of the channel $C \equiv \max_{0 \leqq p_2 \leqq 1} T(X_2 ; Y)$ is attained by $p_2^* = \frac{1}{2}$, independently of p_1 and is equal to



$$C = \log 2 - (p_1 \log \frac{1}{p_1} + q_1 \log \frac{1}{q_1}).$$

Thus the information value $v(I_{p_1}|F)$ remains zero as long as $|p_1 - \frac{1}{2}|$, and hence C, does not exceed a certain critical level; $v(I_{p_1}|F)$ then rises as a non-linear function of C.

This example provides a partial positive answer to the question raised in Example 1. In fact, with

U = the class of payoff matrices $\begin{pmatrix} r_1 & s_2 \\ s_1 & r_2 \end{pmatrix}$ with $s_i < r_i$ (i=1, 2),

$$K(I_{p_1}|F) = \max_{0 \leqq p_2 \leqq 1} T(X_2 ; Y),$$

we have

$$v(I_{p_1'}|F ; u) \geqq v(I_{p_1''}|F ; u) \quad \text{for all } u \in U ,$$

if and only if

$$K(I_{p_1'}|F) \geqq K(I_{p_1''}|F) .$$

References

Marschak, J. (1954). Towards an economic theory of organization and information, in Decision Processes, pp. 187-220. Wiley, New York.

Marschak, J. (1959). Remarks on the economics of information, in Cont. Scientific Res. in Manag., pp. 79-100, Univ. of Calif. Printing Department.

Sakaguchi, M. (1963). Information pattern, learning structure, and optimal decision rule, Information and Control, 6, 218-229.

II. Games and Strategic Information - Duels.

The concept of information is one of the chief essentials in the game theory and it is quite distinct from that of Wiener and Shannon. If one speaks of information in Wiener-Shannon's sense as "selective", then von Neumann's sense could be called "strategic".

The theory of games may be viewed as a formal model embodying three principal elements: (1) the preferences of the players of a game; (2) the choices or decisions open to them; (3) their information regarding the choices made by the opponent player at previous moves.

Strategic information in a game situation is the means of expressing a player's state of knowledge, at any move of a game, regarding the choices which have been made at earlier moves. The problem of rational choice of a plan of action (- the optimal strategy) and the existence of equilibrium situations are both closely related to the nature of the information pattern of the game.

We will show, by using several examples of duels in this lecture and poker models in the subsequent lectures how the solution of a game is related to the information pattern of the game. Before doing this we shall explain some of the fundamentals of zero-sum infinite games in order to help understanding.

§0. Fundamentals of zero-sum infinite games.

In an infinite game, each player selects a strategy from an infinite set of strategies. We consider, in the present and the subsequent chapters, infinite games in which the strategies are represented by points on the closed interval [0, 1] . Player I chooses a strategy

$x \in [0, 1]$, and simultaneously Player II chooses $y \in [0, 1]$. These choices determine a play of the game, whose outcome is measured by a payoff $M(x, y)$ from I to II.

Suppose that I chooses his strategy x from $[0, 1]$ using the distribution function $F(x)$ and II chooses y by means of the distribution function $G(y)$. Then the expectation of I, if it exists, will be

$$M(F, G) = \int_0^1 \int_0^1 M(x, y)\, dF(x)\, dG(y) .$$

Suppose the following two expressions exist:

$$\max_F \min_G M(F, G) = v_1 ,$$

$$\min_G \max_F M(F, G) = v_2 .$$

In general $v_1 \leqq v_2$. However, if $v_1 = v_2$, then there exists a pair of distribution functions F^* , G^* , such that I can receive at least v_1 and II can lose at most v_2 at the same time: that is

$$M(F,G^*) \leqq v_1 = v \leqq v_2 \leqq M(F^*, G), \qquad \text{for all } F \text{ and } G.$$

Thus F^* , G^* are called <u>optimal</u> mixed strategies for I and II, respectively. The pair F^*, G^* is called a <u>solution</u> of the game, and $v = M(F^*, G^*)$ is called the <u>value</u> of the game. Every infinite game does not have a solution. There are examples of infinite games which do not have solutions. However it can be proven that if $M(x, y)$ is continuous in x and y, then both of v_1 and v_2 exist and are equal to each other, so that the optimal mixed strategies F^*, G^* exist and the value of the game is given by

$$v = \min_{0 \leq y \leq 1} \int_0^1 M(x, y)dF^*(x) = \max_{0 \leq x \leq 1} \int_0^1 M(x, y)dG^*(y) .$$

There are also many examples of infinite games with <u>discontinuous</u> payoff functions which do have a solution and a value. Duels and poker are examples.

Let

$$M(x, G^*) = \int_0^1 M(x, y)dG^*(y)$$

be I's expectation if he uses a pure strategy x and II uses an optimal strategy G^*. Similarly, let

$$M(F^*, y) = \int_0^1 M(x, y)\, dF^*(x)$$

be II's expectation if II uses a pure strategy y and I uses an optimal strategy F^*. We have

$$M(x, G^*) \leq v \leq M(F^*, y), \quad \text{for all } x, y \in [0, 1] .$$

The following properties can be readily proven:

(a) If $M(x_0, G^*) < v$ then $F^*(x_0 - 0) = F^*(x_0)$

If $M(F^*, y_0) > v$ then $G^*(y_0 - 0) = G^*(y_0)$.

A player's optimal mixed strategy contains no strategy which does not yield the value of the game when that strategy is used against the opponent's optimal strategy, and hence

(b) If $F^*(x_0 - 0) < F^*(x_0)$ then $M(x_0, G^*) = v.$

If $G^*(y_0 - 0) < G^*(y_0)$ then $M(F^*, y_0) = v.$

Each pure strategy in the optimal mixed strategy must yield the value of the game when used against the opponent's optimal mixed strategy.

§1. Duel as a game of timing.

A game of timing is a competitive environment, in which the actions the players may take are given in advance, but the timing of the actions is selected by the strategic decisions of the players. Such games are characterized by the following conflict of interests: each player wishes to delay his action as long as possible, but he may be penalized for waiting.

The duel is a good example of a game of timing. Each duelist wishes to hold his fire as long as possible, since his accuracy increases with time. However, if the duelist holds his fire too long, his opponent may win the duel.

Let us consider the mathematical formulation of the duels. The duelists starting at a distance 2 apart, approach each other at constant (unit) speed with no opportunity for retreat. The accuracies of firing are described by the <u>accuracy function</u>

$P_1(x)$ = the probability of I's hitting his oponent if he fires at time x .

Similarly, $P_2(y)$ is defined for the player II. Assume that $P_i(0) = 0$, $P_i(1) = 1$, and $P_i(x) \in \uparrow$, $(i = 1, 2)$. Let the payoff be +1 to the surviving duelist and 0 to each duelist if both survive or neither survives.

§2. Noisy and silent duels.

As in all games, we need to describe the information available to the players. If a duelist is informed about his opponent's firings as soon as they take place, we shall call the duel a noisy duel. If neither duelist ever learns when or whether his opponent has fired, we shall call the duel a silent duel.

Example 1. Noisy duel: one bullet each duelist.

In this duel, if a duelist fires and misses, the opponent can obtain a sure hit by waiting until they are together. Thus if x and y are strategies of I and II, respectively, the payoff to I is

$$M(x, y) = \begin{cases} 2P_1(x) - 1, & x < y \\ P_1(x) - P_1(y), & x = y \\ 1 - 2P_2(y), & x > y \end{cases}$$

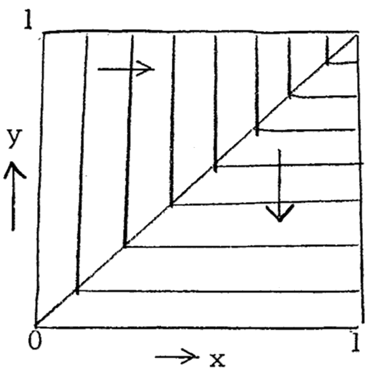


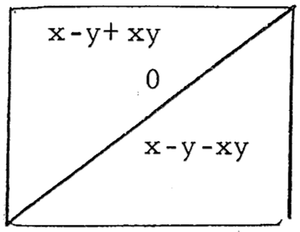
Fig. 1

We easily find that $M(x, y)$ has a saddle point at x_0, x_0 where x_0 satisfies $P_1(x_0) + P_2(x_0) = 1$. (Fig. 1). Thus the optimal strategy for the duelists is to fire their bullets simultaneously at time x_0. The value of the game is $P_1(x_0) - P_2(x_0)$.

Example 2. Silent duel: one bullet each duelist.

In this duel, each duelist is ignorant of firing by the other. Assume that $P_1(x) = x$ and $P_2(y) = y$. Then if x, y are pure strategies, the payoff to I is

$$M(x, y) = \begin{cases} x-(1-x)y, & x < y \\ 0 & x = y \\ -y+(1-y)x, & x > y \,. \end{cases}$$



We find that

$$2\sqrt{2} - 3 = \max_x \min_y M(x, y) < \min_y \max_x M(x, y) = 3-2\sqrt{2} \; (\doteqdot 0.172).$$

Therefore the game does not have a saddle point. Suppose that I and II use mixed strategies $F(x)$ and $G(y)$, respectively. Then we can show that

<u>The optimal strategy for</u> I <u>and</u> II <u>is the following mixed strategy:</u>

$$F^*(x) = \begin{cases} 0, & 0 \leqq x \leqq \frac{1}{3} \\ \int_{\frac{1}{3}}^{x} \frac{1}{4} y^{-3} dy = \frac{1}{8}(9-x^{-2}), & \frac{1}{3} \leqq x \leqq 1 \end{cases} \qquad (*)$$

<u>The value of the game is 0.</u>

<u>Proof.</u> $M(F, y) = \int_0^1 M(x, y)\, dF(x) = \int_0^{y-0} + \int_{y+0}^{1}$

$$= y(F(y) - F(y-0) - 1) + \left\{ (1+y) \int_0^{y-0} + (1-y) \int_{y+0}^{1} \right\} x dF(x) \,.$$

Supposing that $F(x)$ is a density function $f(x)$ over an interval $(\alpha, 1)$, we get

$$M(F, y) = \begin{cases} -y + (1-y)\displaystyle\int_\alpha^1 xf(x)\,dx\,, & 0 \leq y \leq \alpha \\ -y + \left\{(1+y)\displaystyle\int_\alpha^y + (1-y)\displaystyle\int_y^1\right\} xf(x)\,dx\,, & \alpha \leq y \leq 1\,. \end{cases}$$

Since the game is symmetric, the value of the game is 0 and both players have an optimal strategy in common, it follows that

$$M(F, y) \equiv 0\,, \quad \alpha \leqq y \leqq 1\,. \qquad (**)$$

Differentiating this equation two times we obtain $\frac{f'(y)}{f(y)} = -\frac{3}{y}$, from which we get $f(y) = Cy^{-3}$ $(\alpha \leqq y \leqq 1)$. This together with $\int_\alpha^1 f(y)\,dy = 1$ and (**), yields $\alpha = \frac{1}{3}$, $C = \frac{1}{4}$.

To show that an optimal strategy for I is actually given by (*), we must check

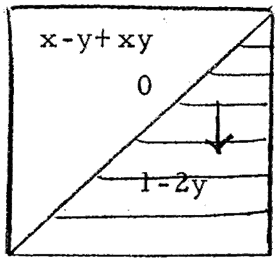
$$M(F^*, y) = \begin{cases} \frac{1}{2} - \frac{3}{2}y \;(\geqq 0), & 0 \leqq y \leqq \frac{1}{3} \\ 0\,, & \frac{1}{3} \leqq y \leqq 1\,. \end{cases}$$

Example 3. Silent-noisy duel: one bullet each duelist.

Let us consider the mixing case, i.e., the case in which I is the silent duelist and II is the noisy duelist. In this game, II who gives information to the opponent stands at a disadvantage.

Let us assume that $P_1(x) = x$, $P_2(y) = y$. Then the payoff is

$$M(x, y) = \begin{cases} x-(1-x)y, & x < y \\ 0 & x = y \\ 1-2y, & x > y \end{cases}$$



We can show that

Let $a = \sqrt{6} - 2$ ($\doteqdot 0.45$). The optimal strategy for I has the density function

$$f^*(x) = \begin{cases} 0 & 0 \leqq x < a \\ \sqrt{2}\, a\, (x^2 + 2x - 1)^{-3/2} & a \leqq x \leqq 1 . \end{cases}$$

The optimal strategy for II is described by

$$G^*(y) = \frac{2}{2+a} \int_0^y f^*(x)\, dx + \frac{a}{2+a} I_1(y) ,$$

where $I_1(y)$ is the unit-step function at $y = 1$. The value of the game is $v = 1 - 2a$ ($\doteqdot 0.101$).

Proof. Guided by heuristic considerations, we shall search for strategies $F(x)$ consisting of a density $f(x)$ on an interval $(a, 1)$ with a weight α at $x=1$, and $G(y)$ consisting of a density $g(y)$ on the same interval with a weight β at $y=1$. Then in analogy to the previous example.

$$\begin{cases} M(F, y) = \int_a^1 M(x, y) f(x)\, dx + \alpha M(1, y) \equiv v, & a \leqq y \leqq 1 \\ M(x, G) = \int_a^1 M(x, y) g(y)\, dy + \beta M(x, 1) \equiv v, & a \leqq x \leqq 1 \end{cases} \quad (*)$$

Two differentiations give

$$3(1+x)\,f(x) + (x^2+2x-1)\,f'(x) = 0 \qquad \therefore\; f(x) = C_1(x^2+2x-1)^{-3/2}$$
$$3(1+y)\,g(y) + (y^2+2y-1)\,g'(x) = 0 \qquad \therefore\; g(y) = C_2(y^2+2y-1)^{-3/2}\,. \qquad (**)$$

In order to determine the solution, we must evaluate the unknown constants a, α, β, C_1, C_2, and v. (**), together with (*) and the normalizing conditions $\int_a^1 f(x)\,dx = 1-\alpha$, $\int_a^1 g(y)\,dy = 1-\beta$ results in the strategies $F^*(x)$ and $G^*(y)$, stated in the theorem.

To show that the specified strategies are indeed optimal it is necessary to check

$$M(F^*, y) \geqq v\,, \quad 0 \leqq y < a$$
$$M(x, G^*) \leqq v\,, \quad 0 \leqq x < a$$

We note without proof that the optimal strategies F^*, G^* are actually unique. Moreover, note that the noisy duelist, who stands at a disadvantage has a positive probability of saving his bullet until the end (Fig. 2). And the silent duelist calls for firing at a later time than under the symmetric silent duel (Example 2). These results are not surprising in view of the unequal information pattern available to the two players.

Some interesting variants of the simple duel are:

(a) Each player is unaware when his opponent has fired until the time $\delta > 0$ has passed since his opponent had fired.

(b) Player I has m bullets and Player II has n bullets. They have different accuracy functions.

(c) In (b) one player does not know how many bullets his opponent has.

But the analyses of these games are really complicated.

Bibliographies

M. Dresher, Games of Strategy: Theory and Applications, Prentice Hall, New York, 1963.

S. Karlin, Mathematical Methods and Theory and Games, Programming and Economics, Vol. II. (in particular, Chapters 5 and 6), Addison-Wesley Publishing Co., London, 1959.

III. Games and Strategic Information - Poker

Following our previous chapters we continue to show how the solution of a game is related to the information pattern of the game. The three essentials in the theory of games are: (1) the preferences of the players of a game; (2) the choices or decisions open to them at each move; (3) their information regarding the choices made by the opponent player at previous moves.

If we set in a game-situation both the choices and preferences for the players symmetrically, then if, moreover, the information pattern is fair for all players in the game, that is, if each player, for example, is completely ignorant of the choices of his opponents, the value of the game is zero and the optimal strategy, when it existed, is common to all players. If we set the choices and preferences of the players symmetrically in a game-situation, and if we let the information pattern be unfair, then symmetry of the game disappears. Consider, for example, the case where the player I must take the first move in the game and his choice is told to the player II who can use this information and act optimally at the second move. It is, as our common sense tells us, clear that the player II stands to gain.

We shall, in this chapter, show somewhat numerically this type of information-unbalance by examples of continuous poker models.

Example 1. La relance (two-person stud poker with a single bet).

In our model the unit interval is taken as the representation of all possible hands that can be dealt to a player. Each hand is considered equally likely and therefore the operation of dealing a hand to a player may be considered as equivalent to selecting a random number from the unit interval according to the uniform distribution. The game proceeds as follows: An ante of 1 unit is required by each of the two players I and II. At the beginning of a play they receive fixed hands, x and y, chosen at random from the unit intervals $0 \leqq x \leqq 1$ and $0 \leqq y \leqq 1$. Then I either bets an amount a or drops out, losing his ante. If I bets, then II can either see the bet or fold, losing his ante. If II sees a bet the hands are compared with the higher card winning the total wager $1+a$.

The above procedure is summarized in the following diagram:

Player	Hand	1st Move	2nd Move	Payoff to I
I	x	drops out	. . . .	-1
		bets a →		
II	y		folds	1
			sees . . .	$(1+a)\operatorname{sgn}(x-y)$

A mixed strategy for I can be described as a function $\phi(x)$ which represents the probability with which I will bet the amount a when his hand is x. A strategy for II can be represented by a function $\psi(y)$ which expressed the probability with which the player II will see a bet when he folds the hand y.

Theorem 1. Let $b = a/(2+a)$. The optimal strategies in this model are as follows:

$$\phi^*(x) = \begin{cases} \text{arbitrary, but subject to the constraint that} & \\ \int_0^b \phi^*(x)\,dx = b(1-b), & \text{if } 0 \leqq x < b, \\ 1, & \text{if } b \leqq x \leqq 1, \end{cases}$$

$$\psi^*(y) = \begin{cases} 0, & \text{if } 0 \leqq y < 1, \\ 1, & \text{if } b \leqq y \leqq 1. \end{cases}$$

The value of the game is $-(a/(2+a))^2$.

Proof. By enumerating all the possibilities we find that the expected payoff to the player I is

$$(1)\qquad M(\phi, \psi) = -\int (1 - \phi(x))dx + \iint \phi(x)(1 - \psi(y))dx\,dy$$
$$+ (1 + a)\iint \phi(x)\psi(y)\operatorname{sgn}(x - y)\,dx\,dy\,.$$

The ranges of integrations are always from 0 to 1. Hence they are omitted here and hereafter. Common sense tells us to guess that the optimal $\psi^*(y)$ is of the form

$$(2)\qquad \psi^*(y) = \begin{cases} 0, & \text{if } y < b, \\ 1, & \text{if } y \geqq b \end{cases} \qquad \text{for some } b.$$

Under this assumption the part of M which involves ϕ can be reduced to

$$(3)\qquad \int \phi(x)L(x)dx\,,$$

where

$$L(x) \equiv \begin{cases} -a + b(2 + a), & \text{if } 0 \leqq x < b, \\ 2(a + 1)x - a(b + 1), & \text{if } b \leqq x \leqq 1. \end{cases}$$

If we set $b = a/(2 + a)$, we have

$$L(x) = \begin{cases} 0, & 0 \leqq x < b, \\ 2(1 + a)(x - b), & b \leqq x \leqq 1. \end{cases}$$

Thus it is clear that if I wants to maximize (3) he must take

$$\phi^*(x) = \begin{cases} \text{arbitrary}, & 0 \leq x < b, \\ 1, & b \leq x \leq 1. \end{cases}$$

The part of M which involves ψ can be reduced to

$$\int \psi(y)K(y)dy \,, \tag{4}$$

where

$$K(y) \equiv -\left(\int_0^y + \int_y^1\right)\phi^*(x)dx + (1+a)\left\{\int_y^1 \phi^*(x)\,dx - \int_0^y \phi^*(x)dx\right\}$$

$$= -(2+a)\int_0^y \phi^*(x)dx + a\int_y^1 \phi^*(x)dx \,.$$

It is clear that since $K(y)$ is monotonously decreasing and the function ψ which minimizes (4) is of the form $\psi = 1$ if $K(y) \leq 0$, and 0 if $K(y) > 0$, we must have the expression

$$-(2+a)\int_a^b \phi^*(x)dx + a\int_b^1 \phi^*(x)dx = 0,$$

in order to have the minimum of (4) with $\psi = \psi^*$ given by (2). We easily have from (1) and (3)

$$M(\phi^*, \psi^*) = -\left(\frac{a}{2+a}\right)^2 .$$

Thus we have shown that $\max_\phi M(\phi, \psi^*) = M(\phi^*, \psi^*) = \min_\phi M(\phi^*, \psi)$, completing the proof of the theorem.

Let us compare this result with that which will be obtained for the case of a one-person game and for the case of a simultaneous game. The following diagram represents the rule of play for the simplest version of Blackjack:

Player	Hand	1st Move	2nd Move	Payoff to I
I	x	fold	. . . .	-1
		bet a →		
II	y		→ see . . .	(1+a)sgn(x-y)

A mixed strategy for I can be described as a function $\phi(x)$ which represents the probability with which I will bet the amount a when his hand is x.. Then the expected payoff to I using the strategy $\phi(x)$ is

$$\tilde{M}(\phi) = -\int (1-\phi(x))dx + (1+a)\iint \phi(x)\,\mathrm{sgn}(x-y)\,dx\,dy$$

$$= -1 + \int (2(1+a)x - a)\,\phi(x)dx$$

which is maximized for $\phi(x) =$

$$\phi^+(x) \begin{cases} 0, & 0 \leqq x < a/2(1+a) \\ 1, & a/2(1+a) \leqq x \leqq 1, \end{cases}$$

and has the maximum value

$$\tilde{M}(\phi^+) = a^2/4(1+a)\ .$$

Next, consider the simultaneous version of La relance:

I ; x \ I ; y	fold	bet
fold	0	-1
bet	1	$(1+a)\,\mathrm{sgn}(x-y)$

where the rule is described by a 2×2 payoff matrix as above. It easily follows by symmetry that the value is 0 and the optimal strategy is common to both players and given by

$$\phi^0(x) = \begin{cases} 0, & 0 \leqq x < a/(1+a) \\ 1, & a/(1+a) \leqq x \leqq 1. \end{cases}$$

Note that $\tilde{M}(\phi^*) < 0 = M(\phi^0, \phi^0) < \tilde{M}(\phi^+)$ and that, for the betting threshold, $\frac{a}{2(1+a)}$ (in ϕ^+) $< \frac{a}{1+a}$ (in ϕ^0). These mean that the strategy ϕ^0 is more careful than ϕ^+ reflecting that Blackjack is essentially a one-person game without any competitor. ϕ^+ and ϕ^0 show that bluffing is useless for these games, as is seen by common sense, whereas I can bluff in ϕ^* such as, for example,

$$\phi^*(x) = \begin{cases} 1, & 0 \leqq x < b(1-b) \\ 0, & b(1-b) \leqq x < b \\ 1, & b \leqq x \leqq 1\,. \end{cases}$$

Even with this extreme bluffing, however, I stands at a disadvantage because of the rule of the game that he must move first and give information about his hand.

<u>Example 2.</u> Le her (two-person draw poker).

This game proceeds as follows: Before the play the two players I and II receive fixed hands, x and y, each being randomly (and independently) chosen from the unit interval.

Now if I is content with his hand he may keep it. But if I is not content with his hand he is allowed to change it for another taken out of the unit interval at random. The rule of the play is the same for player II, and I has to take the first move. The main object is for each to obtain a higher card than his opponent.

This procedure is summarized in the following diagram:

Player	Hand	1st Move	2nd Move	Payoff to I
I	x	keeps x	keeps y	sgn(x-y)
			changes to v	sgn(x-v)
		changes to u	keeps y	sgn(u-y)
II	y		changes to w ...	sgn(u-w)

Let $\alpha(x)$ be the probability that if I receives x he keeps it. Let $\beta(y)$ be the probability that if I keeps his hand and II receives y II keeps it. Let $\gamma(y)$ be the probability that II keeps his hand if he receives y and I has changed his hand. Clearly a mixed strategy for I can be represented by $\alpha(x)$ and that for II by $\beta(y)$ and $\gamma(y)$.

Theorem 2. The optimal strategies in this model are as follows:

$$\alpha^*(x) = \begin{cases} 0, & x < x_0 , \\ 1, & x \geqq x_0 , \end{cases}$$

$$\beta^*(y) = \begin{cases} 0, & y < b = (1 + x_0^2)/2 \doteqdot 0.65, \\ 1, & y \geqq b, \end{cases}$$

$$\gamma^*(y) = \begin{cases} 0, & y < 1/2, \\ 1, & y \geqq 1/2, \end{cases}$$

where $x_0 \doteq 0.56$ is the unique root lying in the unit interval of the equation $4x^3 + 4x - 3 = 0$.

The value of the game is

$$-\frac{1}{4}x_0^4 - \frac{1}{2}x_0^2 + \frac{3}{4}x_0 - \frac{1}{4} \doteq -0.015.$$

Proof. By enumerating all the possibilities we find the expected payoff to player I is

$$\begin{aligned} M(\alpha;\beta,\gamma) = & \iint \alpha(x)\beta(y)\operatorname{sgn}(x-y)dxdy + \iint \alpha(x)(1-\beta(y))dxdy \int \operatorname{sgn}(x-v)dv \\ & + \iint (1-\alpha(x))\gamma(y)dxdy \int \operatorname{sgn}(u-y)du \\ & + \iint (1-\alpha(x))(1-\beta(y))dxdy \iint \operatorname{sgn}(u-w)dudw . \end{aligned} \tag{5}$$

It is natural to guess that the optimal β^* and γ^* are of the forms

$$\beta^*(y) = \begin{cases} 0, & y < b, \\ 1, & y \geqq b, \end{cases} \qquad \gamma^*(y) = \begin{cases} 0, & y < 1/2 \\ 1, & y \geqq 1/2 , \end{cases} \tag{6}$$

for some b, since player II has no opportunity to bluff.

After some calculation the part of $M(\alpha;\beta^*,\gamma^*)$ which involves α becomes expressible as follows:

$$\int \alpha(x)L(x)dx , \tag{7}$$

where

$$L(x) \equiv \begin{cases} 2bx - 3/4 & x \leqq b, \\ 2(b+1)x - 2b - 3/4 & x \geqq b. \end{cases}$$

Thus it is clear that if I wants to maximize (7) he must take

$$\alpha^*(x) = \begin{cases} 0, & \text{if } x \leqq x_0, \\ 1, & \text{if } x \geqq x_0, \end{cases} \qquad \text{for some } 0 < x_0 < 1.$$

Now let us look at the part of $M(\alpha^*; \beta^*, \gamma^*)$ which involves β^*. This is found to be

$$\int \beta^*(y) K(y) dy, \tag{8}$$

where

$$K(y) \equiv \begin{cases} x_0^2 - 2x_0 + 1, & y \leqq x_0, \\ x_0^2 + 1 - 2y, & y \geqq x_0. \end{cases}$$

It is easily seen that we must have $b > x_0$, since II wants to minimize by the optimal choice of β^*. Hence x_0 and b must satisfy the equations $2bx_0 - 3/4 = 0$ and $x_0^2 + 1 - 2b = 0$ respectively.

From the derivations of α^* and β^* we know that

$$M(\alpha^*; \beta^*, \gamma^*) = \max_{\alpha} M(\alpha; \beta^*, \gamma^*) = \min_{\beta} (\alpha^*; \beta, \gamma^*),$$

but we must also check that

$$M(\alpha^*; \beta^*, \gamma^*) = \min_{\beta, \gamma} (\alpha^*; \beta, \gamma).$$

This is found out from (5) by reducing the part of $M(\alpha^*, \beta, \gamma)$ involving γ to

$$\int(1-\alpha^*(x))dx \int(1-2y)\gamma(y)dy .$$

This completes the proof of the theorem.

In this game the degree of disadvantage for I who moves first is very small, because he gives no information about his hand if he changes his hand at the first move.

The simultaneous version of Example 2 is described by the following payoff matrix

I;x \ II;y	keep y	change to v
keep x	sgn(x-y)	sgn(x-v)
change to u	sgn(u-y)	sgn (u-v)

An analysis using symmetry of this payoff function gives the value of the game 0 and the common optimal strategy

$$\phi^*(x) = \begin{cases} 0, & 0 \leqq x < (\sqrt{5} - 1)/2 \quad (\fallingdotseq 0.618) \\ 1, & (\sqrt{5} - 1)/2 \leqq x \leqq 1 . \end{cases}$$

Example 3. La relance with a single bet and a pass. (Kuhn)

Player	Hand	1st Move	2nd Move	3rd Move	Payoff to I
I	x	pass		fold	-1
		bet		see	(1+a)sgn(x-y)
II	y		pass		
			bet		
			fold		1
			see		(1+a)sgn(x-y)

This is the case in which I is allowed to pass when a bad hand is delivered to him.

I's strategy is a pair of functions $\alpha(x)$ and $\beta(x)$ representing the probability of betting at the 1st Move, and of seeing at the 3rd move, respectively, if his hand is x. Similarly, II's strategy can be represented by a pair of $\xi(y)$ and $\eta(y)$ which are the probabilities of betting and seeing, respectively, in the 2nd move, if his hand is y.

Expected payoff is given by

$$M(\alpha, \beta; \xi, \eta) = -\iint(1-\alpha(x))\xi(y)(1-\beta(x))dxdy + (1+a)\iint(1-\alpha(x))\xi(y)\beta(x)sgn(x-y)dxdy$$

$$+ \iint \alpha(x)(1-\eta(y))dxdy + (1+a)\iint \alpha(x)\eta(y)sgn(x-y)dxdy.$$

It would be interesting to investigate how much can I recover his dis-advantage due to moving first by the possibility of passing, and whether $\alpha^*(x)$ and $\xi^*(y)$, or $\beta^*(x)$ and $\eta^*(y)$ have the same forms or not.

Example 4. Indian poker.

In Indian poker each player must move knowing his opponent's hand, not his own. An interesting problem is to find whether it is true, as is commonly believed, or not, that the players in Indian pokers have to dare to bluff fearlessly.

Let us consider this problem for the simplest case, Example 1. Strategies of I and II are to choose their betting probabilities $\phi(y)$ and $\psi(x)$, respectively. The payoff is

$$M(\phi(y), \psi(x)) = -\int (1-\phi(y))dy + \iint \phi(y)(1-\psi(x))dxdy$$

$$+ (1+a)\iint \phi(y)\psi(x)\operatorname{sgn}(x-y)dxdy .$$

This is found equal to $M(\phi, \psi)$ in (1) of Example 1 by change of variables $\bar{x} = 1-x$, $\bar{y} = 1-y$. Hence the solution of the game is given by

$$\phi^*(y) = \begin{cases} 1, & 0 \leqq y < b \\ \text{arbitrary, but subject to the constraint that} & \\ \quad \int_b^1 \phi^*(y)dy = b(1-b), & b \leqq y \leqq 1 \end{cases}$$

$$\psi^*(x) = \begin{cases} 1, & 0 \leqq x < b \\ 0, & b \leqq x \leqq 1, \end{cases}$$

$$v = -(a/(2+a))^2 ,$$

where $b = a/(2+a))^2$. Therefore the strength of bluffing is the same as in Example 1. Situations, however, would be different for more complicated type of poker games.

Bibliography

S. Karlin, Mathematical Methods and Theory in Games, Programming and Economics, Vol. II (in particular Chapter 9). Addison-Wesley Publishing Co., London, 1959.

M. Sakaguchi, "Strategic information and non-cooperative games," KODAI Math. Seminar Report. Vol. 12(1960), pp. 91-101.

IV. Sequential Decisioning and Dynamic Programming

The ungrammatical name "decisioning" in the title of this lecture is quoted from an excellent article by M. M. Flood. In this age of rapid communication and transportation, people, particularly in the fields of management science and military operations, are very much involved in stressful loads of rational decision-making. Of course we cannot say that the mathematical scientists have solved the decision-making problem. However, it is true that there is a rapidly growing and impressively solid literature that deals mathematically with ingenious ideas and techniques for solving various decision-making situations. Dynamic programming techniques of Richard Bellman for multistage decision problems is really one of them. Our concern in this lecture will be with some basic concepts, rather than mathematical detail, and illustration of them by several examples to help understanding.

§1. Basic concepts of dynamic programming.

A multistage decision process may be described simply in the following form. We begin with

a set of positions p . . . . . position variable

a set of choices q . . . . . decision variable

successor state function $T(p, q)$

immediate return function $g(p, q)$

and length of the programming period N.

If $q_1, q_2, \ldots, q_N$ is a sequence of choices, a sequence of positions will be generated, starting with initial $p = p_1$,

$$p_1, \; p_2 = T(p_1, q_1), \;\; p_3 = T(p_2, q_2), \ldots, p_{N+1} = T(p_N, q_N).$$

We impose the condition that q_i are to be chosen to maximize the value of the criterion function

$$R(p_1, p_2, \ldots, p_{N+1}; q_1, \ldots, q_N) = \sum_{t=1}^{N} g(p_t, q_t) + h(p_{N+1}) \;,$$

where $h(p_{N+1})$ represents the utility of the final state p_{N+1}.

This assumption of additivity of utilities from particular steps of the process while apparently quite restrictive, is actually broad enough to permit us to handle many significant classes of processes, including such fields as the calculus of variations and the general theory of adaptive control processes.

Let a sequence of choices of the q's, $\{q_1, \ldots, q_N\}$, be called a policy and a policy which maximizes R an optimal policy. To determine optimal policies in an effective manner, we use the following intuitive principle:

Principle of Optimality. An optimal policy has the property that whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

The functional equations of dynamic programming are derived by a uniform and systematic application of the foregoing principle.

To derive an equation that will simultaneously permit the evaluation of the maximum value of R and the determination of the optimal policy, we begin with the observation that this maximum value depends upon p, the initial state, and N, the number of stages in the process. We therefore write, with $p = p_1$,

$$f_N(p) \equiv \max_{q_1, \ldots, q_N} \left[\sum_{t=1}^{N} g(p_t, q_t) + h(p_{N+1}) \right].$$

In so doing we are emphasizing the fact that p and N are not to be regarded as fixed constants, but rather as variable parameters. Now

$$f_N(p) = \max_{q_1} \left[g(p, q_1) + \max_{q_2, \ldots, q_N} \left\{ \sum_{t=2}^{N} g(p_t, q_t) + h(p_{N+1}) \right\} \right]$$

$$= \max_{q_1} \left[g(p, q_1) + f_{N-1}(T(p, q_1)) \right].$$

The original process with specified values of p and N has thus been imbedded within a family of similar processes described by the above functional equation. On the basis of this equation we can analyse the analytic structure of the maximum return and the optimal policy and compute numerical solutions.

The method described above allows us to transform the original N-dimensional maximization problem involving a choice of $\{q_1, \ldots, q_N\}$ into a succession of one-dimensional maximization problems involving choices of q_1, then q_2, and so on. There are some important advantages in this formulation, as is to be expected. For example,

(a) Absolute maximum, not the relative maximum , can be obtained without much difficulty.

(b) Even if there are constraints in choice of q, it will simplify the computations.

(c) We can formulate the discrete problems in the same fashion.

(d) $g(p, q)$ and $h(p)$ need not be smooth.

We shall illustrate the above concepts by several self-contained examples. The first one of these will be the simplest to show the idea most vividly.

Example 1. Consider the maximization problem:

$$x_1 x_2 \cdots x_N \longrightarrow \max,$$

subject to $$\sum_1^N x_i = a \qquad (a > 0),$$

$$x_i \geqq 0 \quad (i = 1, \ldots, N). \qquad \text{(B40)}$$

Of course we know the solution of this problem by the arithmetic-geometric mean inequality. But we try the functional equation approach.

Let the maximum value be denoted by $f_N(a)$. Then we get

$$f_N(a) = \max_{0 \leqq x \leqq a} \left[x f_{N-1}(a-x) \right], \qquad (N \geqq 2,\ f_1(a) = a).$$

Successively we obtain

$$f_N(a) = a^N / N^N, \qquad N \geqq 1.$$

We thus establish the well-known inequality

$$\frac{1}{N} \sum_1^N x_i \geqq (x_1 x_2 \cdots x_N)^{1/N} \quad \text{for } x_1, \ldots, x_N \geqq 0$$

with equality if and only if $x_1 = x_2 = \cdots = x_N$.

In the second example the dynamic programming formulation leads to the usual "working backward" method.

Example 2.

The manufacturing process for a perishable commodity is such that the cost of changing the level of production from one month to the next is twice the square of the difference in production levels. Any production not sold by the end of the month is wasted at a cost of $20 per unit. Given the sales forecast below, which must be met, determine a production schedule to minimize costs. Assume that December production was 200 units.

Month	Jan.	Feb.	Mar.	Apr.
Sales Forecast	210	220	195	180

Let $f_n(p)$ be the minimum achievable cost when last month's production was p and there are n months to go. Then we get

$$f_n(p) = \min_{x \geq s} \left[2(x-p)^2 + 20(x-s) + f_{n-1}(x) \right]$$

$(n = 1, \ldots, 4 \; ; f_0(p) = 0)$, where

s ..... sales forecast for the current month

x ..... production for the current month

Using this formula for successive months starting with $f_1(p)$ for April and working backward to $f_4(p)$ for January, we get

$$f_1(p) = \min_{x \geq 180} \left[2(x-p)^2 + 20(x-180) \right]$$

$$= \begin{cases} 20p - 3650, & p \geq 185 \\ 2(180-p)^2, & p < 185, \end{cases}$$

$$f_2(p) = \min_{x \geqq 195} [2(x-p)^2 + 20(x - 195) + f_1(x)]$$

$$= \min_{x \geqq 195} [2(x - p)^2 + 40x - 7750]$$

$$= \begin{cases} 2(195 - p)^2 + 250, & p < 205 \\ 40p - 7750, & p \geq 205 \end{cases}$$

$$f_3(p) = \min_{x \geqq 220} [2(x-p)^2 + 20(x-220) + f_2(x)]$$

$$= \min_{x \geqq 220} [2(x-p)^2 + 60x - 12150]$$

$$= \begin{cases} 2(220 - p)^2 + 1050, & 210 \leqq p \leqq 235 \\ 60(p - 210), & p \geqq 235 \end{cases}$$

$$f_4(200) = \min_{x \geqq 210} [2(x - 200)^2 + 20(x - 210) + f_3(x)]$$

$$= \min \begin{cases} \min_{210 \leqq x \leqq 235} [2(x-200)^2 + 20(x-210) + 2(220-x)^2 + 1050] \\ \min_{x \geqq 235} [2(x - 200)^2 + 80(x - 210)] \end{cases}$$

and the answer is

Jan.	Feb.	March	April
210	220	210	205

The following three examples deal with mathematical formulations and solutions for some of the interesting decision-making problems we often encounter in real life.

Example 3. (a search problem). Suppose that we are given the information that a ball is in one of N boxes, and a priori probability, p_k, that it is in the k-th box. Let

t_k ..... time consumed in examining the k-th box.

q_k ..... probability that an examination of the k-th box gives no information concerning its contents.

Find the optimal search procedure which minimizes the expected time required to find the ball.

Let $f(p_1, \ldots, p_N)$ be the expected time required to obtain the ball using an optimal policy. Then

$$f(p_1, \ldots, p_N) = \min_{1 \leq k \leq N} \Big[t_k + q_k f(p_1, \ldots, p_N) + (1-q_k)(1-p_k) f\Big(\frac{p_1}{1-p_k}, \ldots, \overset{k}{0}, \ldots, \frac{p_N}{1-p_k}\Big)\Big]$$

$$= \min_{1 \leq k \leq N} \Big[\frac{t_k}{1-q_k} + (1-p_k) f\Big(\frac{p_1}{1-p_k}, \ldots, \overset{k}{0}, \ldots, \frac{p_N}{1-p_k}\Big)\Big]$$

With

f_{rs} = Expected time required if the examinations of the r-th box, first, and then, the s-th, are followed by optimal continuation,

we get

$$f_{rs} = \frac{t_r}{1-q_r} + (1-p_r)\frac{t_s}{1-q_s} + (1-p_r-p_s) f\Big(\frac{p_1}{1-p_r-p_s}, \ldots, \overset{r}{0}, \ldots, \overset{s}{0}, \ldots, \frac{p_N}{1-p_r-p_s}\Big)$$

Therefore

$$f_{rs} - f_{sr} \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} 0 \Longleftrightarrow \frac{t_r}{p_r(1-q_r)} \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} \frac{t_s}{p_s(1-q_s)}$$

Thus the optimal procedure consists of looking in the box which minimizes

$$\frac{t_k}{p_k(1-q_k)} .$$

If we arrange $\frac{t_k}{p_k(1-q_k)}$ in increasing order obtaining

$$\frac{t_{(1)}}{p_{(1)}(1 - q_{(1)})} \leqq \frac{t_{(2)}}{p_{(2)}(1 - q_{(2)})} \leqq \cdots \leqq \frac{t_N}{p_{(N)}(1-q_{(N)})}$$

then we have

$$f(p_1, \ldots, p_N) = \sum_{k=1}^{N} \frac{p_{(k)} + p_{(k+1)} + \cdots + p_{(N)}}{1 - q_{(k)}} t_{(k)} .$$

Example 4.

Assume that we are a contestant on a quiz program where we have an opportunity to win a substantial amount of money provided that we answer a series of questions correctly. Let, for $k = 1, \ldots, N$,

r_k ..... amount of money obtained if the k-th question is answered correctly.

p_k ..... probability that the k-th question can be answered correctly.

Assume that we have a choice at the end of each question of attempting to answer the next question, or of stopping with the amount already won. Find the optimal policies to pursue under the following conditions.

(a) Any wrong answer terminates the process with a total return of zero.

(b) One wrong answer is allowed.

Solution for (a): Let, for $k = 1, \ldots, N$,

f_k = expected return from the optimal policy given that he has already answered k-1 questions correctly and is now facing the k-th question.

Then

$$f_k = \max \begin{bmatrix} S : r_1 + \cdots + r_{k-1} \\ A : p_k f_{k+1} \end{bmatrix}, \quad (k = 2, \ldots, N \; ; \; f_{N+1} = \sum_1^N r_k),$$

where S and A in the right hand side represent, symbolically, the alternatives, available for him, "stop" and "answer", respectively.

Therefore

$$f_N = \max \begin{bmatrix} S : r_1 + \cdots + r_{N-1} \\ A : p_N \sum_1^N r_k \end{bmatrix}$$

$$\text{Stopping region : } r_1 + \cdots + r_{N-1} > \frac{p_N r_N}{1 - p_N}$$

$$f_{N-1} = \max \begin{bmatrix} S : r_1 + \cdots + r_{N-2} \\ A : p_{N-1} f_N \end{bmatrix}$$

$$= \max \begin{bmatrix} S : r_1 + \cdots + r_{N-2} \\ AS : p_{N-1}(r_1 + \cdots + r_{N-1}) \\ AA : p_{N-1} p_N \sum_1^N r_k \end{bmatrix}$$

$$\text{Stopping region : } r_1 + \cdots + r_{N-2} \geqq \frac{p_{N-1} r_{N-1}}{1 - p_{N-1}} \text{ and } \frac{p_{N-1} p_N (r_{N-1} + r_N)}{1 - p_{N-1} p_N}$$

Thus the stopping region at the second stage (i.e., for f_2) turns out to be

$$r_1 \geqq \frac{p_2 r_2}{1-p_2}, \frac{p_2 p_3(r_2+r_3)}{1-p_2 p_3}, \ldots, \text{ and } \frac{p_2 \cdots p_N(r_2+\cdots+r_N)}{1-p_2 \cdots p_N} .$$

Solution for (b) : Let f_k be defined as in the previous part. Let, for $1 \leqq s < k \leqq N$,

$g(s, k) =$ expected return from the optimal policy given that he has failed the s-th question and answered the other questions correctly and is now facing the k-th question.

Then we obtain

$$g(s, k) = \max \begin{bmatrix} S : \sum_{i=1}^{k-1} r_i - r_s \\ \\ A : p_k g(s, k+1) \end{bmatrix} .$$

$$(k = s+1, \ldots, N ; g(s, N+1) = \sum_1^N r_k - r_s) \text{ and}$$

$$f_s = p_s f_{s+1} + (1-p_s) g(s, s+1)$$

$$(s = 1, \ldots, N ; f(N+1) = \sum_1^N r_k , g(N, N+1) = \sum_1^N r_k).$$

From the upper recurrence relation we can find $g(s, s+1)$ for each s. This and the lower recurrence relation give values of f_s .

Example 5. (Marriage Problem)

A young man wishes to marry the finest young girl he can find. A known number n, of young girls are coming up one at a time before you in a random order. After inspecting any number $r (1 \leqq r \leqq n)$ of them he must decide whether to marry the r-th girl or to continue his search for a still better girl. What is his sequential scheme?

In order to give a mathematical formulation to this problem we must assign a utility function: After inspecting any number r $(1 \leqq r \leqq n)$ of the girls he is able to rank them from best (by rank 1) to worst (by rank r). Let U_i be the utility of marrying the girl with the i-th true rank. Thus $U_1 \geqq U_2 \geqq \cdots \geqq U_n$. Note that a girl's true rank is distinct from her apparent rank that she has when only some have appeared before him. Let

$U(s, r)$ = the expected utility of the optimal sequential scheme when starting from a situation where the r-th girl has an apparent rank s.

Then we obtain

$$U(s, r) = \max \left[\begin{array}{ll} S: & \sum_{i=s}^{n+s-r} \binom{i-1}{s-1}\binom{n-i}{r-s} U_i / \binom{n}{r} \\ \\ C: & \sum_{s'=1}^{r+1} U(s', r+1)/(r+1) \end{array} \right]$$

for all $1 \leqq s \leqq r \leqq n-1$, with $U(s, n) = U_s$, since

$$\text{Prob.} \{\text{Her true rank is } i \mid \text{The r-th girl has apparent rank } s\}$$

$$= \binom{i-1}{s-1}\binom{n-i}{r-s} / \binom{n}{r} .$$

Theorem. Assume that $U_1 = 1$, $U_2 = \cdots = U_n = 0$. Let q be the minimum integer such that $\sum_{r=q}^{n-1} r^{-1} < 1$. Then for a fixed n, the optimal policy is to inspect the first $q-1$ girls and then marry any subsequent girl of apparent rank 1. The expected utility is

$$\frac{1}{n} \left(1 + (q-1) \sum_{r=q}^{n-1} r^{-1}\right) .$$

(For large n, $q \simeq ne^{-1} \doteq 0.368n$, and the expected utility $\simeq e^{-1}$).

<u>Proof.</u> Since, by our assumptions,

$$\sum_{i=s}^{n+s-r} \binom{i-1}{s-1}\binom{n-i}{r-s} U_i/\binom{n}{r} = \begin{cases} \binom{n-1}{r-1}/\binom{n}{r} = \frac{r}{n}, & \text{if } s = 1 \\ 0, & \text{if } s \geqq 2 . \end{cases}$$

we have

$$U(1, r) = \max \begin{bmatrix} S : & r/n \\ C : & u_r \end{bmatrix}, \quad (r = 1, \ldots, n-1, U(1, n) = 1) \qquad (*)$$

and $U(s, r) = u_r$, if $2 \leqq s \leqq r$, where $u_r \equiv \sum_{s'=1}^{r+1} U(s', r+1)/(r+1)$.

The optimal policy requires to continue inspection if $2 \leqq s \leqq r$. Now

$$u_r = \frac{1}{r+1}(U(1, r+1) + ru_{r+1}). \qquad (**)$$

We can first show that if

$$U(1, r_0) = \frac{r_0}{n} \qquad (+)$$

for some r_0 then

$$U(1, r) = \frac{r}{n}, \qquad \text{for all } r \geqq r_0 .$$

For suppose that

$$\exists\, r > r_0 ; \quad U(1, r) = U_r > \frac{r}{n}.$$

Then by (**)

$$u_{r-1} = \frac{1}{r}(U(1, r) + (r-1)u_r) = u_r > \frac{r}{n}.$$

which implies $u_{r-1} > \frac{r-1}{n}$, and hence from (*)

$$U(1, r-1) = u_{r-1} > \frac{r-1}{n}.$$

This argument continues until we get a contradiction to (+).

So next we have to find out how large r must be. Equations (+) and (**) give

$$\frac{u_{r_0}}{r_0} = \frac{1}{nr_0} + \frac{u_{r_0+1}}{r_0+1} = \frac{1}{nr_0} + \frac{1}{n(r_0+1)} + \frac{u_{r_0+2}}{r_0+2}$$

$$= \cdots = \frac{1}{n}\left(\frac{1}{r_0} + \frac{1}{r_0+1} + \cdots + \frac{1}{n-2}\right) + \frac{u_{n-1}}{n-1}$$

$$= \frac{1}{n}\left(\frac{1}{r_0} + \frac{1}{r_0+1} + \cdots + \frac{1}{n-1}\right).$$

Therefore, since $u_{r_0} < \frac{r_0}{n} = U(1, r_0)$, we obtain

$$\frac{1}{r_0} + \frac{1}{r_0+1} + \cdots + \frac{1}{n-1} < 1.$$

Let q be the minimum integer such that $\sum_{r=q}^{n-1} r^{-1} < 1$. Then

$$\frac{1}{q} + \frac{1}{q+1} + \cdots + \frac{1}{n-1} < 1 \leqq \frac{1}{q-1} + \frac{1}{q} + \cdots + \frac{1}{n-1}.$$

From (**) we have

$$u_{q-1} = \frac{1}{q}\left(\frac{q}{n} + (q-1)u_q\right) = \frac{q-1}{n}\left(\frac{1}{q-1} + \frac{1}{q} + \cdots + \frac{1}{n-1}\right). \qquad (++)$$

Again from (**) with $r < q-1$ we have $u_r = u_{r+1}$, so that $U(1,1) = u_1 = \cdots = u_{q-1}$, the utility of the optimal policy starting at the beginning, is given by (++).

If, in real life, this process works between 18 and 40 (i. e., for 22 years) one should never propose until age $18 + 0.368 \times 22 \doteqdot 26$, A sound conclusion is that either many people do not pursue an optimal policy, or else they have a different utility function.

References

Bellman, R., Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

Flood, M. M., "Sequential decisioning", in Information and Decision Processes (R. E. Machol, edt.), McGraw-Hill Book Co., New York 1960.

Lindley, D. V., "Dynamic programming and decision theory", Applied Statistics, Vol. 10(1961), 39-51.

Sasieni, M., Yaspan, A., and Friedman, L., Operations Research - Methods and Problems, John Wiley and Sons, London, 1959.

V. Interpretations of Information by Decision-Making Models

§1. Optimal betting policy of a gambler.

Consider the following two decision-making problems:

(a). Consider a coin with the probability p of heads. A gambler knowing the probability is required to place a bet on the event of head. He is allowed to bet a quantity a, subject to the restriction $0 \leqq a \leqq x$, where x is his capital at the present stage. If he bets correctly he then wins, otherwise he loses. Continuing this betting process for N stages and assuming that the tossings of the coin are independent at these stages and that the gambler wishes to maximize the expected value of the logarithm of the final total at the end of the process, find an optimal betting policy.

(b) Consider the situation in which the gambler makes bets of amount Z_i on the random event E_i, subject to the restriction that $Z_i \geqq 0$ $(i = 1, \ldots, n)$, $\sum_1^n z_i = x$. Assume that he has the following information

p_i ..... the probability of occurring E_i $(i = 1, \ldots, n)$.

r_i (>1) ..... the return from a unit winning bet on E_i. We assume that $\sum r^{-1} < 1$.

Find the optimal betting policy for the N-stage problem.

We show the solutions as follows.

(a) Let the maximum expected value of the logarithm of the final total be denoted by $f_N(x)$ Then, with $q = 1-p$,

$$f_N(x) = \max_{0 \leqq a \leqq x} (pf_{N-1}(x+a) + qf_{N-1}(x-a))$$

$$(N = 1, 2, \ldots ; f_0(x) = \log x).$$

Therefore

$$f_1(x) = \max_{0 \leqq a \leqq x} (p \log (x+a) + q \log(x-a)),$$

in which the maximum in the right hand side is reached at $a =$

$$a^*(x) = \begin{cases} (2p-1)x, & \text{if } p > \frac{1}{2} \\ 0, & \leqq \end{cases} \quad . \qquad (*)$$

If $p > \frac{1}{2}$, then

$$f_1(x) = p \log(2px) + q \log(2qx)$$

$$= \log x + C(p)$$

where

$$C(p) \equiv p \log \frac{p}{1/2} + q \log \frac{q}{1/2} > 0.$$

It is easy to show that, for each stage N, the optimal betting strategy is given by (*) and

$$f_N(x) = \begin{cases} \log x + NC(p), & \text{if } p > \frac{1}{2} \\ \log x, & \leqq \end{cases} \quad .$$

(b) With the same definition of $f_N(x)$ as in (a) we have the recurrence relation

$$f_N(x) = \max_{\substack{\sum_1^n z_i = x \\ z_i \geqq 0}} \sum_{i=1}^{n} p_i f_{N-1}(r_i z_i) ,$$

$(N = 1, 2, \ldots ; f_0(x) = \log x)$.

Therefore

$$f_1(x) = \max_{\substack{\sum_1^n z_i = x \\ z_i \geqq 0}} \sum_1^n p_i \log(r_i z_i)$$

$$= \max_{\substack{\sum_1^n z_i/x = 1 \\ z_i/x \geqq 0}} \left(\sum_{i=1}^n p_i \log \frac{z_i}{x} \right) + \sum_1^n p_i \log r_i + \log x .$$

Since

$$\sum_{i=1}^n p_i \log \frac{z_i}{x} \leqq \sum_{i=1}^n p_i \log p_i ,$$

for all $z_i \geqq 0$, with $\sum_1^n z_i = x$, with equality if and only if

$$z_i = z_i^*(x) = p_i x \qquad (i = 1, \ldots, n). \tag{+}$$

we get

$$f_1(x) = \log x + K$$

where

$$K \equiv \sum_1^n p_i \log r_i - \sum_1^n p_i \log(1/p_i)$$

$$= \sum_{1}^{n} p_i \log \frac{p_i}{r_i^{-1}(\sum r_i^{-1})^{-1}} + \log (\sum r_i^{-1})^{-1}$$

which is positive because of the assumption $\sum r_i^{-1} < 1$. It is easy to show that, for each stage N, the optimal betting strategy is given by (+) and

$$f_N(x) = \log x + NK \qquad (N = 1, 2, \ldots)$$

The optimal policy is independent of the number of stages remaining, independent of the amount of money available, and independent of the interest factors r_i. It always divides the available money proportionally to p_i.

We have seen that in each of these two models, if a gambler bets on the input symbol to a communication channel optimally, his capital will grow exponentially, and the maximum value of this exponential rate of growth is equal to the rate of transmission of information. It is clear, however, that this is decisively due to the logarithmic nature of the payoff function chosen.

The foregoing models naturally suggests two rays of generalization, the one using another utility function $\phi(x)$ instead of $\log x$ and the other converting into the game situation, i.e., the analysis of max-min equations.

§2. A multi-stage game in which Nature is introduced.

Consider the same gambler treated in (b) of the previous section again. Situations are the same except in the following two points. The probabilities p_i of the events are assumed to be unknown by the gambler, and he wishes to maximize the expected value of a function $\phi(w)$ of the final total at the end of the process. $\phi(w)$ is a continuous, non-decreasing and concave function defined for $w > 0$.

Let us define the sequence of functions

$f_N(x)$ = the expected value of $\phi(w)$ obtained using an optimal N-stage policy, starting with a capital of x,

for $N = 1, 2, \ldots$ and $x > 0$, and $= 0$ for $x \leqq 0$.

The min-max principle forces the gambler to act as if Nature's choice of $\{p_i\}$ were the one least favorable to him. Then the principle of optimality yields the recurrence relations

$$(1) \qquad f_1(x) = \max_{\substack{\sum z_i = x \\ z_i \geqq 0}} \min_{\{p_i\}} \sum p_i \phi(r_i z_i),$$

$$(2) \qquad f_k(x) = \max_{\substack{\sum z_i = x \\ z_i \geqq 0}} \min_{\{p_i\}} \sum p_i f_{k-1}(r_i z_i), \quad 2 \leqq k \leqq N$$

for $x > 0$.

We have the

Theorem. (i) We have

$$(3) \qquad f_N(x) = \phi\left[x\left(\sum_1^n r_i^{-1}\right)^{-N}\right], \quad N = 1, 2, \cdots$$

The max-min policy of the gambler is independent of the number of stages remaining, and of the quantity of money available. It leads to: always divide the quantity available proportionally to r_i^{-1}; $z_i = x r_i^{-1}(\sum r_i^{-1})^{-1}$ $(i = 1, \ldots, N)$.

(ii) We have

(4) $$\max_{\substack{\sum z_i = 1 \\ z_i \geqq 0}} \min_{\{p_i\}} \sum p_i \phi(r_i z_i) = \min_{\{p_i\}} \max_{\substack{\sum z_i = 1 \\ z_i \geqq 0}} \sum p_i \phi(r_i z_i) = \phi\left[\left(\sum r_i^{-1}\right)^{-1}\right] .$$

Nature's min-max policy is determined by

(5) $$p_i = r_i^{-1}\left(\sum r_i^{-1}\right)^{-1} .$$

and this is the unique min-max policy if $\phi(w)$ is strictly concave.

Before proceding to the proof of the theorem we show a

Lemma. If r_i's $(i = 1, \ldots, n)$ are the given positive numbers, we have

(6) $$\max_{\substack{\sum z_i = 1 \\ z_i \geqq 0}} \min_{1 \leqq i \leqq n} (r_i z_i) = \min_{\substack{\sum z_i = 1 \\ z_i \geqq 0}} \max_{1 \leqq i \leqq n} (r_i z_i) = \left(\sum r_i^{-1}\right)^{-1} .$$

Proof. We shall show that max-min $= (\sum r_i^{-1})^{-1}$. Since the other half of the lemma can be readily proved by the similar way we shall omit it. Take any z with $z_i \geqq 0$ and $\sum z_i = 1$ and let

$$K(z) \equiv \min_{1 \leqq i \leqq n} (r_i z_i) .$$

Then we have for every i

$$z_i \geqq r_i^{-1} K(z),$$

and so

$$1 = \sum z_i \geqq K(z) \sum r_i^{-1} .$$

Hence we get

$$K(z) \leqq \left(\sum r_i^{-1} \right)^{-1}$$

for any z. The equality sign is attained by the special $z : z_i \equiv r_i^{-1}(\sum r_i^{-1})^{-1}$ $(i = 1, \ldots, n)$. Thus we have proved that

$$\max_{\substack{\sum z_i = 1 \\ z_i \geqq 0}} K(z) = \left(\sum r_i^{-1} \right)^{-1} .$$

Proof of the Theorem.

(i) We have from (1)

$$f_1(x) = \max_{\substack{\sum z_i = 1 \\ z_i \geqq 0}} \min_{p_i} \sum p_i \phi(r_i z_i x)$$

$$= \max_{z} \min_{1 \leqq i \leqq n} \phi(r_i z_i x)$$

$$= \phi[\max_{z} \min_{1 \leqq i \leqq n} (r_i z_i x)]$$

(by monotonicity and continuity of ϕ)

(7) $$= \phi[x(\sum r_i^{-1})^{-1}]$$ (by the lemma).

Next we have from (2) similarly

$$f_2(x) = \max_{\substack{\sum z_i = 0 \\ z_i \geqq 0}} \min_{\{p_i\}} \sum p_i \phi [r_i z_i \left(\sum r_i^{-1} \right)^{-1}]$$

$$= \phi [x \left(\sum r_i^{-1} \right)^{-2}]$$

Thus inductively we get (3). It is evident from the derivation that the max-min policy of the gambler is just determined by the max-min strategy in (6).

(ii) To show the validy of (4) we shall first show that max-min $\geqq$ min-max. The inverse inequality is evident.

We have

the middle term of (4) $= \min_p \max_z \sum p_i \phi (r_i z_i)$

$$\leqq \min_p \max_z \phi \left(\sum p_i r_i z_i \right) \quad \text{(by concavity of } \phi \text{)}$$

$$= \phi (\min_p \max_z \sum p_i r_i z_i)$$

(by monotonicity and continuity of ϕ)

$$= \phi [\min_p \max_i (p_i r_i)]$$

$$= \phi [\left(\sum r_i^{-1} \right)^{-1}] \quad \text{(by the lemma),}$$

and we have

the left-hand side of (4) $= \max_z \min_p \sum p_i \phi (r_i z_i) = \phi [\left(\sum r_i^{-1} \right)^{-1}]$

by (7). Thus we have established the inequality wanted.

The remaining parts of the theorem is again evident.

The max-min expected value is $\log\left\{x(\sum_i r_i^{-1})^{-N}\right\}$, if $\phi(w) = \log w$. In (b) of the previous model in which $\{p_i\}$ were known to the gambler the maximum expected value was $\log(xe^{KN})$. Since

$$K = \sum_1^n p_i \log \frac{p_i}{r_i^{-1}(\sum r_i^{-1})^{-1}} + \log\left(\sum r_i^{-1}\right)^{-1} \geqq \log\left(\sum r_i^{-1}\right)^{-1}$$

we can see that

$$\log\left\{x\left(\sum_i r_i^{-1}\right)^{-N}\right\} \leqq \log(xe^{KN}).$$

The difference of the both sides of this expression measures value of informations of probabilities $\{p_i\}$.

§3. Incentive fee to the forecaster.

In an article "Measures of the value of information", J. McCarthy (1956) attaches to this term a different meaning. He defines what we might call an efficient incentive function and discusses about "a payoff rule to keep the forecaster honest." The payoff in question is not the value of information in our sense, but his approach might give the promise of a fruitful analysis of the "economic" theory of information.

Denote by $p = (p_1, \ldots, p_n) \in \Sigma^n$ the probability-n vector of alternative events $i = 1, \ldots, n$. A client does not know p but has an a priori knowledge $\xi = (\xi_1, \ldots, \xi_n) \in \Sigma^n$. An expert will tell the client an estimate y of p, $y = (y_1, \ldots, y_n) \in \Sigma^n$, and receive a fee $h(y_i)$ if event i happens. We shall call the function h an <u>efficient</u>

<u>incentive function</u> if the following properties are satisfied:

(a) the expected fee is largest when the expert's estimates are correct:

$$\sum_{i=1}^{n} p_i h(y_i) \leqq \sum_{i=1}^{n} p_i h(p_i), \qquad \text{for all } y \in \Sigma^n \text{, with equality if and only if } y = p.$$

(b) the expected fee is zero, if the client's knowledge is correct and the expert does not know more than the client:

$$\sum_{i=1}^{n} \xi_i h(\xi_i) = 0.$$

[Theorem] <u>The function</u>

$$h(y_i) = A(\log y_i + \sum_{i=1}^{n} \xi_i \log \frac{1}{\xi_i}) \qquad A > 0 \; ; i = 1, \ldots, n) \qquad (*)$$

<u>is an efficient incentive function.</u>

<u>Proof.</u> Any function

$$h(y_i) = A \log y_i + B \qquad (A > 0 \; ; i = 1, \ldots, n)$$

satisfies (a). Condition (b) requires that $B = A \sum_{i=1}^{n} \xi_i \log \frac{1}{\xi_i}$.

This efficient incentive function yields the maximum expected fee

$$A\left(\sum_{i=1}^{n} \xi_i \log \frac{1}{\xi_i} - \sum_{i=1}^{n} p_i \log \frac{1}{p_i} \right),$$

a linearly decreasing function of the entropy that characterizes the true probability distribution of the events. In the particular case where the client has a belief a priori that all events are equally probable, the maximum expected fee to the expert becomes $A \sum_{i=1}^{n} p_i \log \frac{p_i}{1/n}$.

It must be noted that the logarithmic incentive function is not the only efficient one. M. Beckmann constructed the following example with $n = 2$:

$$h(y_i) = \int_{1/2}^{y_i} \frac{g(|t - \frac{1}{2}|)}{t} dt \qquad (i = 1, 2, \; ; y \in \Sigma^2),$$

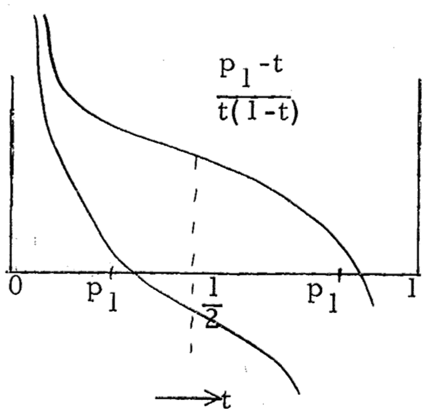
where g is an arbitrary positive-valued function. The expression

$$p_1 h(y_1) + (1-p_1)h(1-y_1) = \int_{1/2}^{y_1} \frac{p_1 - t}{t(1-t)} g(|t - \frac{1}{2}|)dt$$

is maximized when and only when $y_1 = p_1$. Thus

$$h(y_i) = A\left\{\int_{1/2}^{y_i} \frac{g(|t - \frac{1}{2}|)}{t} dt - \sum_{1}^{2} \xi_i \int_{1/2}^{\xi_i} \frac{g(|t - \frac{1}{2}|)}{t} dt\right\} \qquad (A > 0 \; ; i=1, 2)$$

is another efficient incentive function, which reduces to (*) if we take $g(z) \equiv 1$.



References

Bellman, R., and Kalaba, R., "Dynamic programming and statistical communication theory," Proc. Nat. Acad. Sci., U.S.A., Vol. 43(1957), 749-751.

Kelly, J., "A new interpretation of information rate", Bell Syst. Tech. J., Vol. 35(1956), 917-926.

Marschak. J., "Remarks on the economics of information," Contrib. to Sci. Res. in Manag., University of California Press, 1951, pp. 79-98.

McCarthy, J., "Measures of the value of information," Proc. Nat. Acad. Sci., U.S.A., Vol. 42(1956), 654-655.

Sakaguchi, M., "On a certain multi-stage game", Rep. Stat. Appl. Res. JUSE., Vol. 6(1959), 1-4.

VI. Statistical Decision Problems and Bayes Solutions

1. Notation.

Suppose a choice has to be made from a set D of possible decisions and that the value of an "unknown" parameter ω determines the "negative-utility" or "loss" of a decision $d \in D$ by a known loss function $L(\omega, d)$. To assist the choice of d, there is the observation of a r. v. x with the probability density $f(x, \omega)$ depending on the unknown "true" parameter value $\omega \in \Omega$. The observation of x can give some information about the unknown true value of ω. How is d to be chosen on the basis of x, when given the decision problem (Ω, D, L) and the sample space $[(\mathfrak{X}, \mathbb{B}) \{f(x, \omega) | \omega \in \Omega\ ]\}$? This is a usual formulation of statistical decision problems.

Let $\delta(x)$ be, for each $x \in \mathfrak{X}$, a probability distribution over D, so that $\delta(x)$ describes a <u>decision rule</u> by which a decision d is selected with probability $\Pr\{\delta(x) = d | X = x\}$. The <u>risk function</u> of the decision rule δ is defined by

$$r(\omega, \delta) \equiv E_\omega[L(\omega, \delta(X))] = \int_{\mathfrak{X}} \sum_{d \in D} L(\omega, d) \Pr\{\delta(x) = d | X = x\} f(x, \omega) dx,$$

i. e., the expected loss of using the decision rule δ, when the true parameter value is ω.

2. The evaluation of decision rules.

Roughly speaking, we consider a decision rule δ "good" if $r(\omega, \delta)$ is "small" for all $\omega \in \Omega$. To be more precise, suppose we are considering two different decision rules $\delta^{(1)}$ and $\delta^{(2)}$, and suppose that

$$r(\omega, \delta^{(1)}) \leqq r(\omega, \delta^{(2)}), \qquad \text{for all } \omega \in \Omega \text{ and } < \text{ holds for at least one } \omega.$$

Then we say that $\delta^{(1)}$ is better than $\delta^{(2)}$, and we would not use $\delta^{(2)}$. A decision rule δ is called admissible if

$$\nexists\, \delta' \;;\; \delta' \text{ is better than } \delta.$$

Whatever decision is finally used should be an admissible decision rule and therefore a method for finding admissible decision rules is needed.

3. Bayes decision rules.

The Bayes principle of statistical decision problems makes ω random with a priori probability distribution $\xi(\omega)$ and chooses δ which minimizes the average risk $\int_\Omega r(\omega, \delta)\, d\xi(\omega)$. As is readily found, the Bayes principle leads to the Bayes decision rule which selects any decision, with probability 1, which minimizes

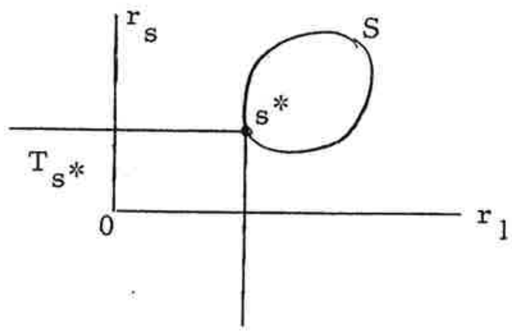
$$\int_\Omega L(\omega, d) f(x, \omega) d\xi(\omega) \propto \int_\Omega L(\omega, d) f(x, \omega) d\xi(\omega) \Big/ \int_\Omega f(x, \omega) d\xi(\omega).$$

i. e., the expectation of $L(\omega, d)$ in the posterior distribution of ω given x, evaluated by Bayes' theorem. (If for some x, this expression is minimized by more than one value of $d \in D$, then there is more than one decision rule which is Bayes r. t. ξ).

Suppose that there is a finite number k of possible values of ω, so that we assume that $\Omega = \{1, \ldots, k\}$. The prior distribution $\xi(\omega)$ over Ω is a probability k-vector $\xi = (\xi_1, \ldots, \xi_k)$. Suppose that the statistical decision problem is given such that the set

$$S \equiv \{(r(1,\delta), \ldots, r(k,\delta)) \mid \delta \in \mathcal{D}\}$$

of all risk vectors attained by all possible decision rules, is a bounded, convex and closed subset in Euclidean k-space. Then some of the basic theorems state



[Theorem 1] For any point $s^* \in S$, let

$$T_{s^*} = \{t \mid t_i > s_i^* \ (i = 1, \ldots, k)\}.$$

Then s^* is Bayes r.t. some $\xi \in \Xi^k$, if and only if

$$T_{s^*} \cap S = \emptyset .$$

[Theorem 2] Let

$$\Xi^+ = \{\xi \mid \xi_i > 0 \ (i = 1, \ldots, k), \ \sum_1^k \xi_i = 1\}$$

$\mathcal{A}$ = set of all admissible points in S

$\mathcal{B}$ = set of all Bayes points in S

$\mathcal{B}^+$ = set of all Bayes points in S r.t. some $\xi \in \Xi^+$.

Then (a) $\mathcal{B}^+ \subseteqq \mathcal{A} \subseteqq \mathcal{B}$.

(b) Both of $\mathcal{A}$ and $\mathcal{B}$ are complete:

Completeness of $\mathcal{A}$ is defined by $(\forall a' \notin \mathcal{A}) \ \exists a \in \mathcal{A}$; a is better than a'. Theorem 2 says thatto find the Bayes decision rules is a useful step in searching for the admissible decision rules. Besides it is so simple to find its Bayes decision rules.

4. Bayes principle leads to likelihood ratio.

In many decision problems in which Ω is finite, the Bayes decision rule is a rule based on the likelihood ratio.

Consider the following simple example. $\Omega = \{1, 2\}$, $D = \{d_1, d_2\}$ and the loss is given by

ω \ d	d_1	d_2
1	0	1
2	1	0

Here nature's states are represented by one of two densities $f_1(x)$ and $f_2(x)$. Statistician has to decide which is the true p.d.f. when observing the value x. The Bayes decision rule r.t. ζ, $1-\zeta$ is the rule which

$$\text{takes } \begin{Bmatrix} d_1 \\ d_2 \end{Bmatrix} \text{ according as } (1-\zeta)f_2(x) \begin{Bmatrix} \leq \\ > \end{Bmatrix} \zeta f_1(x).$$

This is a probability-ratio decision rule. As ζ varies $0 \sim 1$, $\frac{\zeta}{1-\zeta}$ varies $0 \sim \infty$. Thus by varying $\frac{\zeta}{1-\zeta}$ from 0 to ∞ we generate all Bayes rules. It is known that a statistical decision rule based on the likelihood-ratio often provides a seemingly peculiar example.

Let us give two examples.

[Proposition 1] Suppose that the data consists of a single observation X with Cauchy density

$$f(x, \omega) = \frac{1}{\pi} \frac{1}{1 + (x-\omega)^2}, \quad -\infty < x < \infty.$$

This density is symmetric and centered about ω and resembles the normal density. The Bayes test of

$$H_{-1} : \omega = -1$$

$$H_1 : \omega = 1$$

r.t. ζ, $1-\zeta$ has the acceptance region of H_1 :

$$\begin{cases} (-\infty, \infty), & \text{if } \zeta < \frac{2-\sqrt{2}}{4} \doteqdot 0.147 \\ (-\infty, \lambda_{-1}(\zeta)) \cup (\lambda_1(\zeta), \infty), & \text{if } \frac{2-\sqrt{2}}{4} < \zeta < \frac{1}{2} \\ (\lambda_{-1}(\zeta), \lambda_1(\zeta)), & \text{if } \frac{1}{2} < \zeta < \frac{2+\sqrt{2}}{4} \doteqdot 0.854 \\ \text{null set}, & \text{if } \zeta > \frac{2+\sqrt{2}}{4} . \end{cases}$$

where $\lambda_{\pm 1}(\zeta) = \frac{-1 \pm \sqrt{1-2(1-2\ )^2}}{1 - 2\zeta}$

The Bayes test r.t. $\frac{1}{3}$, $\frac{2}{3}$, for example, accepts H_1 when X is large enough $(\lambda_1(\frac{1}{3}) \doteqdot -0.35)$ or <u>very small</u> $(\lambda_1(\frac{1}{3}) \doteqdot -5.65)$.

<u>Proof.</u> Let ζ be the prior probability assigned to H_{-1}. The Bayes test r.t. ζ, $1-\zeta$ accepts H_1 if and only if

$$\frac{f(x, 1)}{f(x, -1)} = \frac{x^2 + 2x + 2}{x^2 + 2x + 2} \geqq \frac{\zeta}{1-\zeta}$$

<u>i.e.,</u> $(1 - 2\zeta)x^2 + 2x + 2(1-2\zeta) \geq 0$.

It should be remarked that, although an intuitively appealing strategy, such as "accept H_1 if and only if $X > -0.35$" is not admissible, it is almost so, and its risk point lies close to the boundary of the set S.

The above peculiarity is based, to some extent, on the property of Cauchy densities. But the following example applies for a considerably wider class of densities.

Suppose that we have to decide between two hypotheses H_1 and H_2 and that we can have the choice of observing a variable X alone or of observing two variables X and Y. While observation of the two variables is more informative than observation of the one variable, it is also more expensive.

	X	Y
(ζ) H_1	$f_1(x)$	$g_1(y)$
(1-ζ) H_2	$f_2(x)$	$g_2(y)$

The question is whether it is worthwhile for us to pay the additional cost $c(> 0)$ for the variable Y.

Let ζ and $1-\zeta$ be the prior probabilities of H_1 and H_2, respectively. And let us assume 0-or-1 loss function. It is easily conjectured that if ζ is small or large the Bayes decision rule will be based on observing X alone. The question is whether the value of ζ for which the Bayes rules are based on X and Y constitute an interval.

[Proposition 2] (a) Let the average risk of the Bayes decision rule based on X be denoted by $R_X(\zeta)$, and that based on both of X and Y, by $R_{X,Y}(\zeta)$. Then for any f_1, f_2, g_1 and g_2 we have

$$R_{X,Y}(\zeta) \leqq R_X(\zeta), \quad \text{for all} \quad 0 \leqq \zeta \leqq 1 .$$

(b) Suppose that the observation of Y needs the additional cost $c(> 0)$. Then the set $\{\zeta \mid R_{X,Y}(\zeta) + c < R_X(\zeta)\}$ does not necessarily constitute an interval.

Proof of (a).

$$R_{X,Y}(\zeta) = \zeta \iint\limits_{\frac{f_2(x)g_2(y)}{f_1(x)\,g_1(y)} > \frac{\zeta}{1-\zeta}} f_1(x)g_1(y)d\lambda(x)d\nu(y)$$

$$+ \quad (1-\zeta) \iint\limits_{\frac{f_2(x)g_2(y)}{f_1(x)\,g_1(y)} \leqq \frac{\zeta}{1-\zeta}} f_2(x)g_2(y)d\lambda(x)d\nu(y)$$

$$= 1 - \iint \max(\zeta f_1(x)g_1(y),\ (1-\zeta)f_2(x)g_2(y))d\lambda(x)d\nu(y).$$

But

$$\max(\zeta f_1(x)g_1(y),\ (1-\zeta)f_2(x)g_2(y))$$

$$\geqq \begin{cases} \zeta f_1(x)\,g_1(y) = \max(\zeta f_1(x),\ (1-\zeta)f_2(x))g_1(y), & \text{if } x \notin W \\ (1-\zeta)f_2(x)g_2(y) = \max(\zeta f_1(x),\ (1-\zeta)f_2(x))g_2(y) & \text{if } x \in W \end{cases}$$

where $W \equiv \left\{ x \mid \frac{f_2(x)}{f_1(x)} > \frac{\zeta}{1-\zeta} \right\}$. Thus we get

$$\iint \max(\zeta f_1(x)g_1(y),\ (1-\zeta)f_2(x)g_2(y))\, d\lambda(x)d\nu(y)$$

$$\geqq \int \max(\zeta f_1(x),\ (1-\zeta)f_2(x))d\lambda(x) = 1 - R_X(\zeta).$$

<u>Example of</u> (b). An example is given as follows (Anderson, 1964).

Hyp. \ Exp.		$\mathcal{E}_X$	$\mathcal{E}_Y$
ζ	H_1	$X = \begin{cases} a, & \text{w. p. } 1/3 \\ b, & " \\ c, & " \end{cases}$	$Y = \begin{cases} A, & \text{w. p. } 1/2 \\ B, & " \end{cases}$
$(1-\zeta)$	H_2	$X = \begin{cases} a, & \text{w. p. } 1/6 \\ b, & " \quad 1/3 \\ c, & " \quad 1/2 \end{cases}$	$Y = \begin{cases} A, & \text{w. p. } 1/3 \\ B, & " \quad 2/3 \end{cases}$

Both of

$$R_X(\zeta) = \zeta \Pr\left\{\frac{f_2(X)}{f1(X)} > \frac{\zeta}{1-\zeta} \,\Big|\, H_1\right\} + (1-\zeta)\Pr\left\{\frac{f_2(X)}{f_1(X)} \leqq \frac{\zeta}{1-\zeta} \,\Big|\, H_2\right\}$$

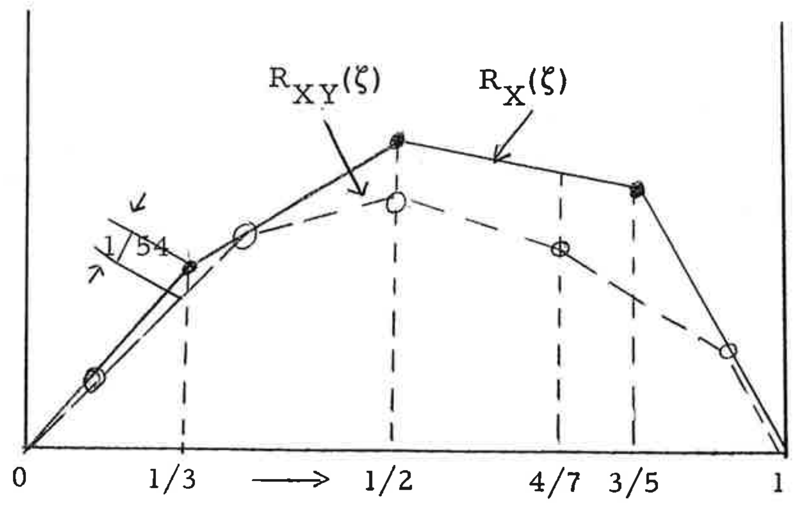
and

$$R_{XY}(\zeta) = \zeta \Pr\left\{\frac{f_2(X)g_2(Y)}{f_1(X)g_1(Y)} > \frac{\zeta}{1-\zeta} \,\Big|\, H_1\right\} + (1-\zeta)\Pr\left\{\frac{f_2(X)g_2(Y)}{f_1(X)g_1(Y)} \leqq \frac{\zeta}{1-\zeta} \,\Big|\, H_2\right\} + c,$$

and hence, $R_X(\zeta) - R_{XY}(\zeta)$ are all linear, in ζ, in intervals. The values are shown in the following table.

$\frac{\zeta}{1-\zeta}$	ζ	$R_X(\zeta)$	$R_{XY}(\zeta)-c$	$R_X(\zeta)-R_{XY}(\zeta)+c$
0	0	0	0	0
1/3	1/4	1/3	1/4	0
1/2	1/3			1/54
2/3	2/5	5/12	11/30	0
1	1/2		7/18	1/36
4/3	4/7		8/21	1/42
3/2	3/5	2/5		1/30
2	2/3		1/3	0
∞	1	0	0	0

Both of $R_X(\zeta)$ and $R_{X,Y}(\zeta)$ are concave in $0 \leqq \zeta \leqq 1$ and have the graphs shown below.



Hence if $0 < c < \frac{1}{54}$, $\left\{\zeta \mid R_X(\zeta) - R_{X,Y}(\zeta) - c > 0\right\}$ does not constitute an interval.

5. Robustness of Bayes solutions.

Let us turn our attention to the more practical view point. In most cases the Bayes decision rules are derived on the basis of two assumptions:

(i) $f(x, \omega)$ has a mathematically ideal, but in many cases practically impossible shape, for example, normal frequency function with mean ω and variance 1.

(ii) The prior distribution $\xi(\omega)$ is exactly known and has a shape mathematically convenient to treat.

We can, in many practical situations, agree with the opinion that $\xi(\omega)$ is unlikely to be known with much accuracy because the effort necessary for its accurate information is too great. Moreover, even in the case in which $\xi(\omega)$ is used to denote the subjective probability distribution for ω, "the vagueness associated with judgements of the magnitude of personal probability" (quoted from L. J. Savage) will usually preclude realization of $\xi(\omega)$. So, in practice, one has often to use a supposed distribution on $\tilde{\xi}(\omega)$ as an approximation to $\xi(\omega)$. If $\tilde{\xi}$ is used in place of ξ, the Bayes solution r. t. $\tilde{\xi}$, $\delta_{\tilde{\xi}}$, yields the average risk $\int_{\Omega} r(\omega, \delta_{\tilde{\xi}}) d\xi(\omega)$. Thus

$$\Delta(\xi : \tilde{\xi}) \equiv \int_{\Omega} (r(\omega, \delta_{\tilde{\xi}}) - r(\omega, \delta_{\xi})) d\xi(\omega) \quad (\geqq 0)$$

if the average increase of risk which the statistician will suffer through using $\tilde{\xi}$ instead of ξ.

Let us take as an example a typical problem of estimation. Suppose that each month a company sells a certain amount of its product, and that the amounts sold in the various months are independent, identically distributed r. v. with mean μ and variance σ^2 both of which are unknown. The problem is to estimate this unknown mean, using the observations available.

The following proposition is readily found:

[Proposition 3] Let $x_1, \ldots, x_n$ be independent observations of a r. v. having mean μ and variance σ^2, both unknown. We are required to estimate μ with squared-error loss functions. Suppose that $\xi(\omega)$ is an a priori distribution over $\Omega = \{(\mu, \sigma^2)\}$ and we consider the estimators linear in $y = \sum_1^n x_i/n$. Then the Bayes estimator of μ, r. t. $\xi(\omega)$ is given by

$$d_\xi(y) = (1-\alpha)y + \alpha E(\mu)$$

and the Bayes risk, by

$$\int_\Omega r(\omega, d_\xi) d\xi = (1-\alpha)^2 E\left(\frac{\sigma^2}{n}\right) + \sigma^2 V(\mu)$$

where $E(\cdot)$ and $V(\cdot)$ are expectations and variances in the distribution $\xi(\omega)$ and $\alpha \equiv E\left(\frac{\sigma^2}{n}\right) \Big/ \left(V(\mu) + E\left(\frac{\sigma^2}{n}\right)\right)$.

Proof. The average risk of the estimator $ay+b$ is

$$\int d\xi(\mu, \sigma^2) \int (ay+b-\mu)^2 g(y ; \mu, \sigma^2) dy = a^2 E\left(\frac{\sigma^2}{n}\right) + (a-1)^2 V(\mu) + ((a-1)E(\mu)+b)^2$$

which is minimized by choosing $a = 1-\alpha$, $b = \alpha E(\mu)$.

Apparently we require only partial information about $\xi(\omega)$. This is because of the rather unrestricted shape of $f(x ; \mu, \sigma^2)$. More precise knowledge of ξ would, of course, be required if the shape of $f(x ; \mu, \sigma^2)$ is specified.

[Proposition 4] In Proposition 3, let $d_{\tilde{\xi}}(y)$ be the Bayes estimator r.t. $\tilde{\xi}$. Then the increase of risk by using $\tilde{\xi}$ is given by

$$\Delta(\xi : \tilde{\xi}) \equiv \int_{\Omega} (r(\omega, d_{\tilde{\xi}}) - r(\omega, d_{\xi}))d\xi(\omega)$$

$$= ((1-\tilde{\alpha})^2 - (1-\alpha)^2)E\left(\frac{\sigma^2}{n}\right) + (\tilde{\alpha}^2 - \alpha^2)V(\mu) + \tilde{\alpha}^2(E(\mu) - \tilde{E}(\mu))^2$$

where

$$d_{\tilde{\xi}} = (1-\tilde{\alpha})y + \tilde{\alpha}\tilde{E}(\mu)$$

and $\tilde{\alpha} \equiv \tilde{E}\left(\frac{\sigma^2}{n}\right) / (\tilde{V}(\mu) + \tilde{E}\left(\frac{\sigma^2}{n}\right))$, in which $\tilde{E}(\cdot)$ and $\tilde{V}(\cdot)$ are taken in the distribution $\tilde{\xi}(\omega)$.

Several deductions may be made from this result:

(i) If $\tilde{\alpha} = 1$ (i.e., we use $\tilde{E}(\mu)$ as the Bayes estimator) then

$$[\Delta(\xi ; \tilde{\xi})]_{\tilde{\alpha}=1} = \frac{V^2(\mu)}{V(\mu) + E\left(\frac{\sigma^2}{n}\right)} + (E(\mu) - \tilde{E}(\mu))^2 .$$

(ii) If $\tilde{\alpha} = 0$ (i.e., we use the usual estimator y) then

$$[\Delta(\xi : \tilde{\xi})]_{\tilde{\alpha}=0} = \frac{E^2\left(\frac{\sigma^2}{n}\right)}{V(\mu) + E\left(\frac{\sigma^2}{n}\right)}$$

(iii) $[\Delta(\xi : \tilde{\xi})]_{\tilde{\alpha}=1} < [\Delta(\xi : \tilde{\xi})]_{\tilde{\alpha}=0}$ i.e., $\tilde{E}(\mu)$ is "better than" y, if and only if

$$E[(\mu - \tilde{E}(\mu))^2] < E\left(\frac{\sigma^2}{n}\right) .$$

In other words, if the prior estimator $\tilde{E}(\mu)$ is chosen good enough, the estimate using no observations is better than the usual estimator y.

(iv) $[\Delta(\xi : \tilde{\xi})]_{\tilde{\alpha}=0} < \Delta(\xi : \tilde{\xi})$, i.e., y is "better than" $d_{\tilde{\xi}}$, if and only if

$$E[(\mu - \tilde{E}(\mu))^2] < 2 \frac{E(\sigma^2)}{\tilde{E}(\sigma^2)} \tilde{V}(\mu) + E\left(\frac{\sigma^2}{n}\right).$$

Thus, if n increases the range of $\tilde{\xi}$ in which y is better then $d_{\tilde{\xi}}$ is reduced.

References

Anderson, T. W., "On Bayes procedures for a problem with choice of observations", Ann. Math. Stat., Vol. 35(1964), 1128-1135.

Chernoff, H. and Moses, L. E., Elementary Decision Theory, John Wiley and Sons, New York, 1959.

Stone, M., "Robustness of non-ideal decision procedures", Jour. Amer. Stat. Assoc., Vol. 58(1963), No. 302, 480-486.

Weiss, L., Statistical Decision Theory, McGraw-Hill Book Co., New York, 1961.

VII. Information, Experience and Learning.

Today's topic is information, experience and learning. The concept of information is a basic one in many fields of modern science. It is particularly penetrating and illuminating in the domain of mathematical statistics. Add to this classical theme the modern notion of "decision" and we have the basic ingredients of a new field of statistics, the theory of information and decision processes. To make good decisions is always hard work. In this age of rapid growth of industry and technology, engineers and scientists spend much of their time trying to obtain more information about the problem at hand, and finally making a decision based on the information which is available. It would be fatuous to assert that they use mathematical optimization techniques in their cases, or even that they might hope to do so. But certainly there is ample room in this direction, and we are moving steadily toward the quantification and objectivization of our decision-making.

I illustrate some central ideas of utilization by showing several examples in mathematical statistics. We start our consideration with a very simple example. Suppose that a 2 x 2 game against Nature is repeatedly played by a rational decision-maker. In each play the loss is given by

	Decision-Maker	
States \ Decision	d_1	d_2
w_1	0	1
w_2	1	0

Suppose that the decision-maker knows that Nature uses a fixed policy by which the state w_1 is selected by a fixed but unknown frequency $0 < p < 1$ at each stage. If the frequency p were known to the decision-maker the situation would be very easy: His best choice at each stage is to

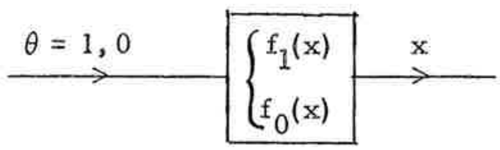
$$\text{choose} \begin{Bmatrix} d_1 \\ d_2 \end{Bmatrix}, \text{ according as } p \begin{Bmatrix} \geqq \\ < \end{Bmatrix} 1/2 ,$$

yielding the expected payoff $\min(p, 1-p)$ at each stage. This optimal behavior, which is a Bayes decision rule, is crucially based on his exact knowledge about whether the value of p is greater than or smaller than $1/2$.

Reasoning by use of statistical theory, however, enables us to act optimally in some sense, even if we have no information originally about the value of p, provided that we can play the game repeatedly under the invariant situation gathering the increasing information about the unknown value of p. The following Example 1 show this fact.

Example 1

Consider the guessing game



with the loss function

decision / state	d_1:guess "$\theta = 1$"	d_0:guess "$\theta = 0$"
$\theta = 1$	0	$a_1 (>0)$
$\theta = 0$	$a_0 (> 0)$	0

Let the decision function be denoted by $\delta(x)$: that is, choose d_1 with prob $\delta(x)$, if the observation is x. If the prior distribution of the unknown signal θ is binomial with parameters 1, p, then the average risk of the decision function δ is

$$r(p, \delta) \equiv pr(1, \delta) + (1-p)r(0, \delta)$$

$$= pa_1 \int (1-\delta(x))f_1(x)d\mu + (1-p)a_0 \int \delta(x)f_0(x)d\mu$$

$$= pa_1 - \int \phi_p(x)\delta(x)d\mu$$

where

$$\phi_p(x) \equiv pa_1 f_1(x) - (1-p)a_0 f_0(x) .$$

Hence the Bayes solution r. t. this prior distribution is given by

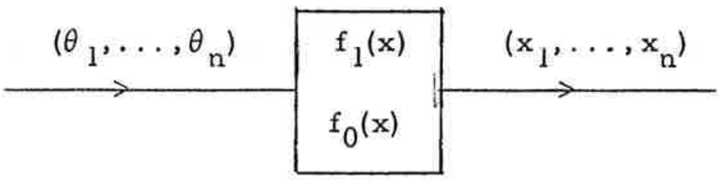
$$\delta_p^* = \begin{cases} 1, & \text{if } \phi_p(x) \geq 0 \\ 0, & \text{if } \quad\quad\;\; < \end{cases} .$$

The Bayes risk is

$$r(p, \delta_p^*) \equiv pr(1, \delta_p^*) + (1-p)r(0, \delta_p^*) = pa_1 - \int [\phi_p(x)]^+ d\mu .$$

Now suppose that the value of p is unknown so that we cannot adopt the Bayes decision rule δ^*_p, but that the problem of guessing the value of θ from observation of x is presented to us a total of n times:

$$\begin{cases} \theta_1 & \theta_2 & \cdots & \theta_{n-1} & \theta_n \\ x_1 & x_2 & \cdots & x_{n-1} & x_n \end{cases}$$



$\theta_1, \ldots, \theta_n$ are considered as independently and identically distributed, each with binomial distribution with parameters 1, p. The observations $x_1, \ldots, x_{n-1}$ are independent of θ_n, but they give some information about the unknown value of p.

Let the decision function for the n-th component problem be denoted by $\delta_n(x_1, \ldots, x_n)$ (with the same meaning as in the first component problem). It has the average risk

$$r(p, \delta_n) \equiv p r(1, \delta_n) + (1-p) r(0, \delta_n)$$

$$= p a_1 - \int \phi_p(x_n) \prod_{i=1}^{n-1} (p f_1(x_i) + (1-p) f_0(x_i)) \delta_n(x_1, \ldots, x_n) \, d\mu$$

and hence the corresponding Bayes solution is given by

$$\text{Bayes decision function} = \begin{cases} 1, & \text{if } \phi_p(x_n) \geqq 0 \\ 0, & \text{if } \qquad\quad < , \end{cases}$$

if the value of p is known to the decision-maker. The Bayes risk is, as before,

$$pa_1 - \int [\phi_p(x_n)]^+ d\mu(x_n) .$$

[Theorem] <u>Define the decision function for the</u> n-th <u>component problem by</u>

$$\delta_n^0(x_1, \ldots, x_n) \equiv \begin{cases} 1, & \text{if } \phi_{\hat{p}_n}(x_n) \geqq 0 \\ 0, & \text{if } \qquad\quad < , \end{cases}$$

<u>where</u> $\hat{p}_n = \hat{p}_n(x_1, \ldots, x_n)$ <u>is any consistent estimater of</u> p. <u>Then</u> $\left\{\delta_n^0(x_1, \ldots, x_n)\right\}_{n=1}^{\infty}$ <u>is asymptotically optimal in the sense that</u>

$$\lim_{n \to \infty} r(p, \delta_n^0) = r(p, \delta_p^*), \quad \underline{\text{for all}} \quad 0 \leqq p \leqq 1.$$

<u>Proof.</u>

$$r(p, \delta_n^0) = pa_1 - \int \phi_p(x) d\mu(x) \int \prod_{i=1}^{n-1} (pf_1(x_i) + (1-p)f_0(x_i))\delta_n^0(x_1, \ldots, x_{n-1}, x) d\mu^{n-1}.$$

For any fixed value of x,

$$\int \prod_{i=1}^{n-1} (pf_1(x_i) + (1-p)f_0(x_i))\, \delta_n^0(x_1, \ldots, x_{n-1}, x) d\mu^{n-1}$$

$$= \Pr.\left\{\phi_{\hat{p}_n}(x) \geqq 0\right\} \xrightarrow[(n \to \infty)]{} \begin{cases} 1, & \text{if } \phi_p(x) \geqq 0\,, \\ 0, & \text{if } \qquad < \end{cases} .$$

Since for the integral with respect to x,

$$|\text{integrand}| \leqq |\phi_p(x)|$$

and

$$\int |\phi_p(x)|\, d\mu(x) \leqq pa_1 + (1-p)a_0 \leqq \max(a_1, a_0),$$

we get, by the Lebesque theorem of dominated convergence,

$$\lim_{n\to\infty} r(p, \delta_n^0) = pa_1 - \int \lim_{n\to\infty} [\, \phi_p(x) \int \cdots \,]\, d\mu(x)$$

$$= pa_1 - \int [\phi_p(x)]^+ d\mu(x) = r(p, \delta_p^*) \,.$$

And we can actually find a consistent estimater of p. For, with an arbitrary unbiased estimater $h(x)$ of θ:

$$E_\theta\{h(X)\} \equiv \int h(x) f_\theta(x) d\mu = \theta \qquad (\theta = 0, 1)$$

let us define

$$\hat{p}_n(x_1, \ldots, x_n) \equiv \begin{cases} 0, & \text{if } \frac{1}{n} \sum_{i=1}^{n} h(x_i) < 0 \\ 1, & \text{if } \qquad > 1 \\ \frac{1}{n} \sum_{1}^{n} h(x_i), & \text{if otherwise.} \end{cases}$$

Since $\frac{1}{n}\sum_{1}^{n} h(X_i)$ is a consistent estimater of p ($\because$ $E[h(X_i)] = pE_1[h(X_i)] + (1-p)E_0[h(X_i)] = p$). $\hat{p}_n(x_1, \ldots, x_n)$ is also a consistent estimater of p. An example of h(x) is given, for any set A with $\int_A (f_1(x) - f_0(x))\, d\mu \neq 0$, by

$$h(x) = (\Delta_A(x) - \int_A f_0(x)d\mu) / (\int_A (f_1(x) - f_0(x))d\mu),$$

where $\Delta_A(x)$ is the indicator function of the set A. The important conclusion of this theorem is that the knowledge of the successive empirical distribution of past states makes it constructively possible to do almost as well at reducing the average risk as in the case where the value of p is known in advance.

Important results concerning the same kind of effects of the knowledge of the successive empirical distribution of past states are also given in the "non-Bayesian" approach. In this approach we consider $\theta_1, \ldots, \theta_n, \ldots$ as a sequence of unknown constants, instead of a sequence of independent random variables each with a common unknown distribution. See Samuel's papers cited in References.

Example 2.

Suppose in the previous example, that we can not assume that Nature selects "$\theta = 1$" with a fixed probability p. Nature can choose $\theta = 1$ or 0 quite arbitrarily, so that we must consider $\theta_1, \ldots, \theta_n$ as an unknown sequence of numbers.

For simplicity let the two densities be specified as follows: We are given n independent, unrelated decision problems, each with the same simple-deciding problem (Ω, D, L), where $\Omega = \{\mathbb{N}(-1, 1), \mathbb{N}(1, 1)\}$ or equivalently $\Omega = \{-1, 1\}$; $D = \{d_-, d_+\}$, with

$d_-(d_+)$: decision that the time mean of the normal distribution is $-1(+1)$, and $L(w, d)$ is the usual 0-1 loss function. Let

$$\begin{cases} \theta_1 & \theta_2 & \cdots & \theta_n \\ x_1 & x_2 & \cdots & x_n \end{cases} .$$

The problem is to guess the whole sequence $\theta_1, \ldots, \theta_n$ after observing $x_1, \ldots, x_n$. Let $\delta(x_1, \ldots, x_n)$ be the decision function with values

$$(d_1(x_1, \ldots, x_n), \ \ldots, \ d_n(x_1, \ldots, x_n)),$$

where $d_i(x_1, \ldots, x_n) = \pm 1$ is the decision about θ_i. The risk function is given by the average number of wrong decisions, that is,

$$r(\theta, \ \delta) = E_\theta \left[\frac{1}{2n} \sum_{i=1}^{n} |\theta_i - d_i(X_1, \ldots, X_n)| \right]$$

where θ stands for $(\theta_1, \ldots, \theta_n)$.

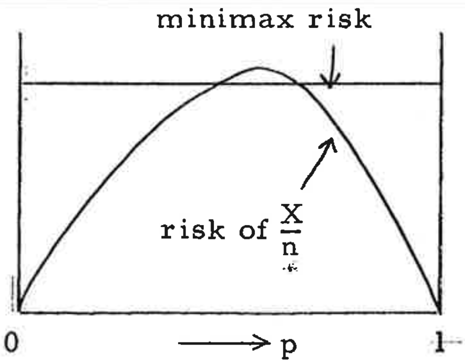
We can see in the following theorem, that the minimax solution for our problem is not so reasonable. Generally speaking, there are some classes of problems in which minimax solutions do not provide reasonable answers. For example, the minimax estimator of the parameter p of a binomial distribution with n and p is given by $\tilde{\delta} = \frac{X}{n} \frac{\sqrt{n}}{\sqrt{n}+1} + \frac{1}{2(\sqrt{n}+1)}$.

Comparing this estimator with the usual one $\frac{X}{n}$ we find that

$$(\text{risk of } \frac{X}{n}) = \frac{p(1-p)}{n} \leqq \frac{1}{4(1+\sqrt{n})^2} = (\text{risk of } \tilde{\delta}), \text{ if and only if}$$

$$|p - \frac{1}{2}| \geqq \frac{\sqrt{1+2\sqrt{n}}}{2(1+\sqrt{n})}$$

Hence when n is large $\frac{X}{n}$ has the smaller risk than $\tilde{\delta}$ except in a small interval about $p = \frac{1}{2}$. The solution like this estimator $\frac{X}{n}$ is called an asymptotically subminimax solution.



[Theorem] <u>The decision rule</u>

$$\bar{\delta} : \bar{d}_i(x_1, \ldots, x_n) = \operatorname{sgn} x_i , \quad i = 1, \ldots, n$$

<u>is minimax.</u> <u>And the decision rule</u>

$$\delta^* : d_i^*(x_1, \ldots, x_n) = \operatorname{sgn}(x_i - x^*), \quad i = 1, \ldots, n$$

<u>where</u>

$$x^* = \begin{cases} \infty , & \text{if } \bar{x} \equiv \frac{1}{n}\sum_1^n x_i \leqq -1 \\ \frac{1}{2}\log\frac{1-\bar{x}}{1+\bar{x}} , & \text{if } -1 < \bar{x} < 1 \\ -\infty , & \text{if } \bar{x} > 1 \end{cases}$$

<u>is asymptotically subminimax.</u>

<u>Sketch of proof.</u> Evidently $\bar{\delta}$ has a constant risk

$$\frac{1}{2}\int_{-\infty}^{\infty} |\,\theta_i - \operatorname{sgn} x_i\,|\ \frac{1}{\sqrt{2\pi}}\ e^{-\frac{(x_i-\theta_i)^2}{i}}\ dx_i = \int_{-\infty}^{-1} \frac{1}{\sqrt{2\pi}}\ e^{-t^2/2}\ dt \equiv F(-1).$$

Consider the prior distribution ξ over $\left\{\theta_1, \ldots, \theta_n \,|\, \theta_i = \pm 1\ ;\ i=1,\ldots,n\right\}$ such that

$$(\theta_1, \ldots, \theta_n) = b_k\ ,\ \text{if } k \text{ of } \theta_i\text{'s are } 1$$

for $k = 0, 1, \ldots, n$. It can be shown that the Bayes decision rule is given by

$$\delta\ :\ d_i(x_1, \ldots, x_n) = \operatorname{sgn}\left(x_i - \frac{1}{2}\log\frac{\sum_{k=0}^{n-1} b_k S_k^{(i)}}{\sum_{k=0}^{n-1} b_{k+1} S_k^{(i)}}\right),\quad i=1,\ldots,n$$

where

$$S_k^{(i)} \equiv \sum_{\{j_1,\ldots,j_k\}\subset\{1,\ldots,i-1,i+1,\ldots,n\}} e^{z(x_{j_1}+\cdots+x_{j_k})}\ .$$

$\bar{\delta}$ corresponds to the δ with $b_k \equiv 1$, $(k = 0, 1, \ldots, n)$. Computing the risk of δ^* for large n and comparing it with that of $\bar{\delta}$ we can find that δ^* is asymptotically minimax.

A striking fact seen in this example is that while values x_j $(j \neq i)$ give no information about θ_i, the good decision rule has $d_i(x_1, \ldots, x_n)$ a function of all variables $x_1, \ldots, x_n$, not of a single variable x_i.

References

H. Robbins, The empirical Bayes approach to testing statistical hypotheses, Rev. Intern. Stat. Inst., Vol. 31(1963), 195-208.

__________, Asymptotically subminimax solutions of compound statistical decision problems, Proc. of the 2nd Berkeley Symp. on Math. Stat. and Prob., 1951, 131-148.

E. Samuel, Asymptotic solutions of the sequential compound decision problem, Ann. Math. Stat., Vol. 34(1963), 1079-1094.

__________, Convergence of the losses of certain decision rules for the sequential compound decision problem, Ann. Math. Stat., Vol. 35(1964), 1606-1621.

VIII. Relationships Between Information Theory and Decision Theory

The problems I wish to discuss this evening are of importance in information theory, but are essentially problems in statistical decision theory. We shall consider some relationships between information theory and statistical decision theory. Inequalities of Wald type for testing hypothesis or Cramer-Rao type for point extimation will be derived for various statistical procedures.

§1. Statistical experiments and information provided by them.

An observation of a univariate random variable X is said to be a performance of an experiment with the random variable X. Hence the experiment will result in an observation x, belonging to a space $\mathfrak{X}$. The space $\mathfrak{X}$ has a σ-field $\mathbb{B}$ of subsets. We shall suppose to have a dominated parametric set of probability measures, each defined on the measurable space $(\mathfrak{X}, \mathbb{B})$. We shall describe it by $\{p(x|\theta) | \theta \in \Theta\}$, where $p(x|\theta)$ denotes the generalized p. d. f. with respect to a common dominating measure, and Θ is any parameter space. Then the couple

$$\mathcal{E} = [(\mathfrak{X}, \mathbb{B}), \{p(x|\theta) | \theta \in \Theta\}] \tag{1.1}$$

characterizes an experiment $\mathcal{E}$.

With this definition, the notion of the experiment corresponds to the following communication system with noise (Fig. 1.1). It consists of essentially two parts:

(i) The input space in the set Θ of symbols θ. These symbols are transmitted one by one by some discrete stochastic process, in which each choice of θ is made with probability $p(\theta)$, successive choices being independent.

(ii) The noisy channel is such that the output space is a set $\mathfrak{X}$. We assume that successive symbols are independently perturbed by the noise, and the channel, therefore, is described by the set of transition probabilities $p(x|\theta)$, $\theta \in \Theta$, the probability of transmitted symbol θ being received as $x \in \mathcal{X}$.

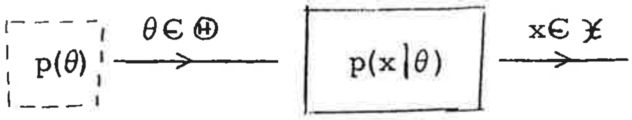


Fig. 1. 1

Statistical decision problem of deciding which θ is the "true" transmitted symbol will be represented by Fig. 1. 2.

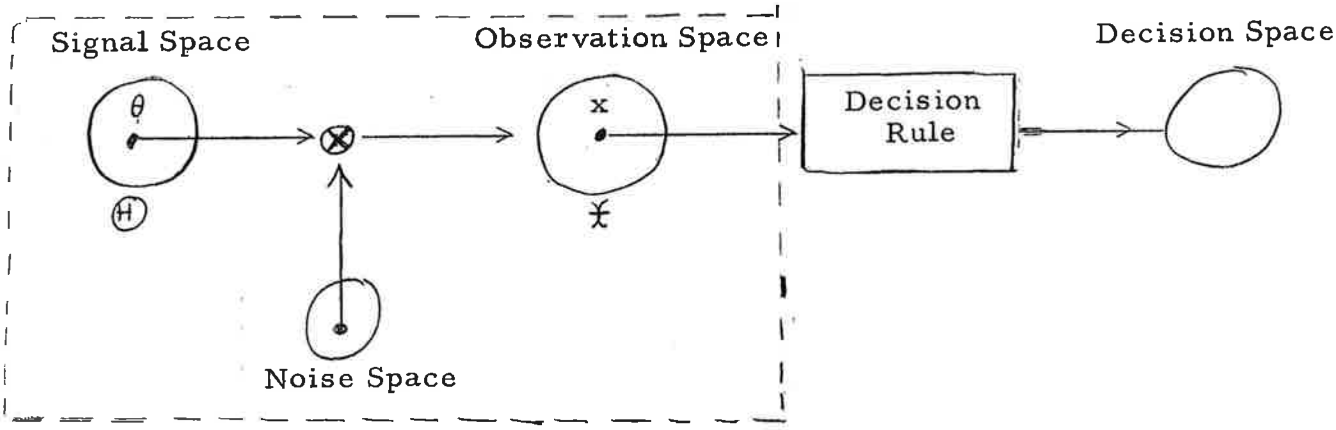


Fig. 1. 2

The part in the frame of the broken line in Fig. 1. 2 is the "communication channel with noise" and this part has no relation with statistical decision theory. Fig. 1. 1 is equivalent to this part. In order to clarify the input and output spaces we sometimes write (1. 1) as

$$\mathcal{E} = [\Theta, \{p(x|\theta) | \theta \in \Theta\}, \mathcal{X}] . \tag{1. 1'}$$

The transmission rate of the noisy channel is defined by Shannon as the amount of decrease of input entropy to the received conditional entropy. Hence, then if the a priori p. d. f. of input symbols θ is $p(\theta)$, it is given by

$$-\int p(\theta) \log p(\theta)\, d\theta + \iint p(\theta, x) \log p(\theta \mid x)\, d\theta\, dx$$

$$= \iint p(\theta)\, p(x \mid \theta) \log \frac{p(x \mid \theta)}{p(x)}\, d\theta\, dx$$

where $p(x) = \int p(\theta)\, p(x \mid \theta) d\theta$. In these notations for the sake of simplicity, we shall not distinguish between random variables and the values assumed by them, nor shall we attempt to be specific in describing the density functions. Thus $p(\theta)$ and $p(x)$ denote the density functions of the random variables θ and x, respectively, without any suggestion that they have the same density. Moreover we shall denote integration with respect to the dominating measures on $\mathfrak{X}$ and Θ by dx and $d\theta$ respectively, again for simplicity of notation. We can define the amount of information provided by the experiment $\mathcal{E}$ with the prior knowledge $p(\theta)$ by

(1. 2) $$\mathcal{I}[\mathcal{E}, p(\theta)] = \iint p(\theta)\, p(x \mid \theta) \log \frac{p(x \mid \theta)}{p(x)}\, d\theta\, dx$$

§2. Testing hypotheses.

Consider the problem of testing simple hypothesis

$$H_0 : f(x) = f_0(x)$$

against a simple alternative

$$H_1 : f(x) = f_1(x) .$$

The observation of x corresponds to the performance of the dichotomous experiment $\mathcal{E} = [(\mathfrak{X}, \mathbb{B}), \{f_1(x), f_0(x)\}]$. Suppose that a measurable subset R of $\mathfrak{X}$ is considered as the critical region, i.e., acceptance region of the alternative hypothesis H_1. Then the probabilities of two kinds of errors are

$$\alpha = \int_R f_0(x)\, d\lambda \quad \text{and} \quad \beta = \int_{R^c} f_1(x)\, d\lambda \quad .$$

Because of the above definitions, the two error probabilities α and β can not be made arbitrarily small simultaneously. One explanation of this fact is given by a pair of the well-known information inequalities

$$I(f_0 : f_1) \equiv \int f_0 \log \frac{f_0}{f_1}\, d\lambda \geqq \alpha \log \frac{\alpha}{1-\beta} + (1-\alpha) \log \frac{1-\alpha}{\beta} \;\; (\equiv F(\alpha, \beta), \text{ say})$$

$$I(f_1 : f_0) \equiv \int f_1 \log \frac{f_1}{f_0}\, d\lambda \geqq (1-\beta) \log \frac{1-\beta}{\alpha} + \beta \log \frac{\beta}{1-\alpha} .$$

For any fixed $0 < \alpha < 1$, $F(\alpha, \beta)$ is, as a function of β, strictly convex and tends to $+\infty$, as $\beta \longrightarrow 0$. Thus the upper one of the above inequalities implies that the possible value of β has a positive lower bound provided that $I(f_0 : f_1) < \infty$. Moreover this lower bound is monotone increasing as $\alpha \downarrow 0$.

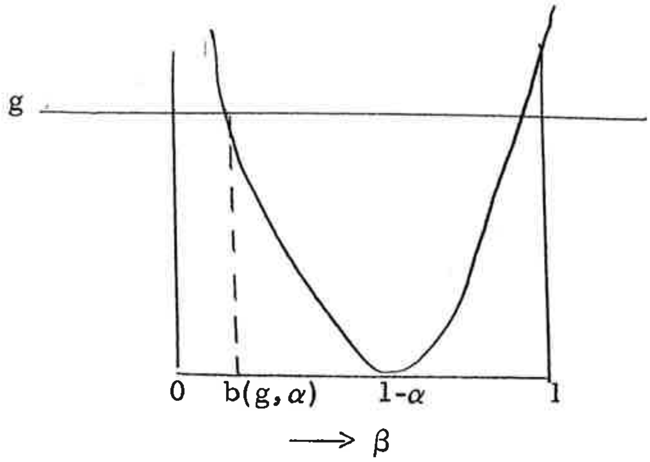
Proof. Let $b(g, \alpha)$ be the smaller one of the two roots of the equation

$$F(\alpha, \beta) = g \qquad (0 < g < \infty)$$

in β for each fixed $0 < \alpha < 1$. Then

$$\frac{\partial b(g,\alpha)}{\partial \alpha} = -\left[\frac{\frac{\partial F}{\partial \alpha}}{\frac{\partial F}{\partial \beta}}\right]_{\beta=b(g,\alpha)} = \left[\frac{\beta(1-\beta)}{1-\alpha-\beta}\log\frac{\alpha\beta}{(1-\alpha)(1-\beta)}\right]_{\beta=b(g,\alpha)} < 0,$$

Since $b(g,\alpha) < 1-\alpha$.



Graph of $F(\alpha,\beta)$ for a fixed α.

§3. Finite deciding problem.

Let

(3.1) $$\mathcal{E} = [(\mathcal{X}, \mathbb{B}), \{f_1(x), \ldots, f_k(x)\}]$$

be a finite experiment. Let $\xi = (\xi_1, \ldots, \xi_k)$ denote a probability-k vector, i.e., $\xi_i \geqq 0$ $(i = 1, \ldots, k)$ and $\sum_1^k \xi_i = 1$. The prior distribution ξ over $\textcircled{H} = \{1, \ldots, k\}$ has uncertainty, a priori, which is measured by Wiener-Shannon entropy

$$H(\xi) = -\sum_{i=1}^{k} \xi_i \log \xi_i .$$

The amount of information provided by the experiment $\mathcal{E}$ is defined by the expected reduction of Wiener-Shannon uncertainty after observing the random variable x

$$H(\xi) - E[H(\xi(X))|\xi],$$

where $\xi(x) = (\xi_1(x), \ldots, \xi_k(x))$ with

$$\xi_i(x) = \xi_i f_i(x) / \sum_{j=1}^{k} \xi_j f_j(x) \qquad (i=1,\ldots,k)$$

is the posterior distribution over Θ, and $E[h(X)|\xi] \equiv \sum_{i=1}^{k} \xi_i \int h(x) f_i(x)\, d\lambda$

denotes the average expectation taken under the prior distribution ξ. Let $\{R_1, \ldots, R_k\}$ be a measurable decomposition of $\mathfrak{X}$, i.e., a collection of measurable sets such that $R_i \cap R_j = \emptyset$, $(i \neq j)$ and $\bigcup_{i=1}^{n} R_i = \mathfrak{X}$.

The set R_j is considered as an acceptance region of hypothesis $H_j : f(x) = f_j(x)$. Let $\int_{R_j} f_i(x) d\lambda = \alpha_{ij}$. Then $\sum_{j=1}^{k} \alpha_{ij} \equiv 1$ $(i = 1, \ldots, k)$.

The probability of error due to the decision rule $\{R_1, \ldots, R_k\}$ is

$$P_e = 1 - \sum_{i=1}^{k} \xi_i \alpha_{ii} = \sum_{i=1}^{k} \int_{R_i^C} \xi_i f_i(x) d\lambda$$

Now the conditional entropy at the receiving end is

$$E[H(\xi(X))\ \xi] = -\sum_{i=1}^{k} \int \xi_i f_i(x) \log \frac{\xi_i f_i(x)}{\sum_{j=1}^{k} \xi_j f_j(x)}\, d\lambda$$

$$= -\left(\sum_{i=1}^{k} \int_{R_i^C} + \sum_{i=1}^{k} \int_{R_i} \right)$$

$$\leqq -\left(P_e \log \frac{P_e}{k-1} + (1-P_e)\log \frac{1-P_e}{1}\right)$$

$$= -P_e \log P_e - (1-P_e)\log(1-P_e) + P_e \log(k-1),$$

and hence we have

$$P_e \log \frac{P_e}{\frac{k-1}{k}} + (1-P_e)\log \frac{1-P_e}{1/k} \leqq \log k - E[H(\xi(X))|\xi] \tag{3.2}$$

with equality if and only if

$$\frac{\xi_i f_i(x)}{\sum \xi_j f_j(x)} = \frac{P_e}{k-1}, \quad \text{for all } i \text{ and } x \in R_i^C .$$

The above extreme case corresponds to a maximum likelihood rule since

$$\frac{\xi_i f_i(x)}{\sum \xi_j f_j(x)} = 1-P_e, \quad \frac{\xi_h f_h(x)}{\sum \xi_j f_j(x)} = \frac{P_e}{k-1} \qquad (h \neq i),$$

if $x \in R_i$.

We call the experiment

$$\mathcal{E}^{(n)} = [(\mathcal{X}^n, \mathbb{B}^n), \{f_i^{(n)}(x_1, \ldots, x_n) | i = 1, \ldots, k\}]$$

where $f_i^{(n)}(x_1, \ldots, x_n) = \prod_{j=1}^{n} f_i(x_j)$ $(i = 1, \ldots, k)$, an n replication of a common experiment $\mathcal{E}$ defined by (3.1). The amount of information provided by $\mathcal{E}^{(n)}$ is

$$\mathcal{I}[\mathcal{E}^{(n)}, \xi] \equiv H(\xi) - E[H(\xi(X_1, \ldots, X_n)) | \xi] .$$

[Theorem]. <u>For an</u> n <u>replication</u> $\mathcal{E}^{(n)}$ <u>of a common experiment</u> $\mathcal{E}$, <u>the information</u> $\mathcal{I}[\mathcal{E}^{(n)}, \xi]$ <u>is a concave, increasing function of</u> n <u>for any fixed</u> ξ.

Proof. It suffices to show that

$$0 \leqq e_n - e_{n+1} \leqq e_{n-1} - e_n$$

where $e_n \equiv e_n(\xi) \equiv E[H(\xi(X_1, \ldots, X_n)) | \xi]$. For any finite experiment (3. 1), it is an easy exercise to show that $E[H(\xi(X)) | \xi]$ is concave in ξ. Hence each function $e_n(\xi)$ $(n = 1, 2, \ldots \; ; e_0(\xi) = H(\xi))$ is concave in ξ. We have

$$\begin{aligned} e_n(\xi) &= E\{E[H(\xi(X_1)(X_2, \ldots, X_n)) | \xi(X_1)] | \xi\} \\ &= E[e_{n-1}(\xi(X_1)) | \xi] \\ &\leqq e_{n-1}(E[\xi(X_1) | \xi]) = e_{n-1}(\xi). \end{aligned}$$

Now the recurrence relation

$$e_n(\xi) - e_{n+1}(\xi) = E[e_{n-1}(\xi(X_1)) - e_n(\xi(X_1)) | \xi]$$

$n = 1, 2, \ldots \; ; e_0(\xi) = H(\xi))$, together with the concavity, in ξ, of

$$e_0(\xi) - e_1(\xi) = H(\xi) - E[H(\xi(X)) | \xi]$$

$$= \sum_{i=1}^{k} \xi_i \int f_i(x) \log \frac{f_i(x)}{\sum_j \xi_j f_j(x)} \, d\lambda ,$$

gives $e_n(\xi) - e_{n+1}(\xi) \leqq e_{n-1}(\xi) - e_n(\xi)$, for all n. For an n-replication of the finite experiment (3.1), we have, from (3.2),

$$P_e \log \frac{P_e}{\frac{k-1}{k}} + (1 - P_e) \log \frac{P_e}{1/k} \leqq \log k - E[H(\xi(X_1, \ldots, X_n)) | \xi]$$

where $P_e = \sum_{i=1}^{k} \int_{R_i^C} \xi_i f_i^{(n)} (x_1, \ldots, x_n) \, d\lambda^n$ is the probability of error.

Since $E[H(\xi(X_1, \ldots, X_n)) | \xi]$ is decreasing in n , the possible lower bound to P_e decreases in n.

§4. Point estimation.

For point estimation we introduce

$g(x)$ = point estimator of "true" θ

$L(\theta, g(x))$ = loss defined for $\theta \in \Theta$ and $x \in \mathcal{X}$.

If $d\xi(\theta)$ is the prior distribution of θ , the posterior probability element of θ given that x is observed is

$$p(\theta | x) d\theta = \frac{p(x|\theta) \, d\xi(\theta)}{\int p(x|\theta) d\xi(\theta)} .$$

The posterior risk from $g(\cdot)$ for a fixed received signal $x \in \mathcal{X}$ is

$$r(x) = \int L(\theta, g(x))\, p(\theta | x)\, d\theta .$$

Let

$$q(\theta | g(x)) \equiv e^{-\lambda L(\theta, g(x))} \Big/ \int_{\Theta} e^{-\lambda L(\theta, g(x))}\, d\theta , \qquad (\lambda > 0)$$

be a probability density of θ, and let

$$\Delta(x) \equiv \int_{\Theta} p(\theta | x) \log \frac{p(\theta | x)}{q(\theta | g(x))}\, d\theta .$$

Then we have

$$r(x) \equiv \int_{\Theta} L(\theta, g(x))\, p(\theta | x)\, d\theta = \frac{1}{\lambda} \left(H(\Theta | x) - \log \int_{\Theta} e^{-\lambda L(\theta, g(x))} d\theta + \Delta(x) \right).$$

where $H(\Theta | x)^{\Theta} \equiv - \int_{\Theta} p(\theta | x) \log p(\theta | x)\, d\theta$.

$$r(x) \geqq \sup_{\lambda > 0} \left(H(\Theta | x) - \log \int e^{-\lambda L(\theta, g(x))}\, d\theta \right) .$$

If Θ is s-dimensional and if $L(\theta, g(x))$ is a bilinear form

$$L(\theta, g(x)) = \sum_{i,j=1}^{s} a_{ij}(\theta_i - g_i(x))(\theta_j - g_j(x)) ,$$

where $(a_{ij} | i, j = 1, \ldots, s)$ is a positive-definite and symmetric matrix, then

$$(4.1) \qquad \max_{\lambda > 0} \frac{1}{\lambda} \left(H(\Theta | x) - \log \int_{\Theta} e^{-\lambda L(\theta, g(x))}\, d\theta \right) \doteqdot \frac{s}{2\pi e} |a_{ij}|^{1/s} e^{\frac{2}{s} H(\Theta | x)}$$

and hence

$$\text{(4. 2)} \qquad r(x) \geqq \frac{s}{2\pi e} |a_{ij}|^{1/s} e^{\frac{2}{s} H(\Theta | x)} .$$

The factor

$$\frac{1}{2\pi e} e^{\frac{2}{s} H(\Theta | x)} ,$$

in the right-hand side is the entropy power per dimension of the s-dimensional conditional density $p(\theta_1, \ldots, \theta_s | x)$.

Proof of (4. 1).

Note that if $Ce^{-\frac{1}{2} x \Lambda x'}$ (in vector notation: Λ is a positive-definite and symmetric s x s matrix) is an s-dimensional normal density. $C = \frac{1}{(2\pi)^{s/2} |\Lambda|^{-1/2}}$ and $(E(x_i x_j)) = \Lambda^{-1}$. The maximand of the left-hand side in (4. 1) is equal to

$$\frac{1}{\lambda} \left\{ H(\Theta | x) + \log(|a_{ij}|^{1/2} \pi^{-s/2}) + \frac{s}{2} \log \lambda \right\} .$$

The function $\frac{1}{\lambda}(a + b \log \lambda)$ $(b > 0, \lambda > 0)$ has a unique maximum at $\lambda = e^{1 - \frac{a}{b}}$ and the maximum value $be^{\frac{a}{b} - 1}$.

Jensen's inequality gives a lower bound to the average risk of the estimator $g(\cdot)$ as

$$\text{(4. 3)} \qquad \int_{\mathfrak{X}} r(x) p(x) dx = \frac{s}{2\pi e} |a_{ij}|^{1/s} e^{\frac{2}{s} H(\Theta | \mathfrak{X})} .$$

(4. 2) or (4. 3) is an inequality of Cramer-Rao type. This also shows the extremal property of normal distributions.

REFERENCES

[1] Kullback, S., Information Theory and Statistics, John Wiley, New York, 1959.

[2] Lindley, D. V., "On a measure of the information provided by an experiment", Ann. Math. Stat., Vol. 27(1956), 986-1005.

[3] Sakaguchi, M., Information Theory and Decision Making, Statistics Department, The George Washington University, Washington, D.C., June 1964.

[4] Seidler, J., "Relationships between information theory and decision function theory", Trans. Second Prague Conf. on Information Theory, Stat. Decision Function and Random Processes, Academic Press, New York, 1960, 579-592.

IX. AMOUNT OF INFORMATION PROVIDED BY A DICHOTOMOUS EXPERIMENT

This lecture considers the amount of information provided by dichotomous experiments under various uncertainty functions. Four types of concave, continuous uncertainty functions are investigated. Special attention is paid to the problem of maximizing information with respect to the prior probabilities. (Section 2). In Section 3, we derive conditions under which the information provided by a dichotomous experiment is convex with respect to mixtures of experiments and conditions under which information increases concavely to the original uncertainty, under independent replications of the same experiment. In the final section effects under grouping of observations are examined.

§1. Statistical experiments.

An observation of a univariate random variable X is said to be a performance of an experiment with the random variable X. Hence the experiment will result in an observation x, belonging to a space $\mathcal{X}$. The space $\mathcal{X}$ has a σ-algebra $\mathbb{B}$ of subsets. We assume a dominated parametric set of probability measures, each defined on the measurable space $(\mathcal{X}, \mathbb{B})$. We shall describe it by $\{p(x|\theta)\,|\,\theta \in \Theta\}$, where $p(x|\theta)$ denotes the generalized pdf with respect to a common dominating measure, and Θ is any parameter space. Then the couple

(1. 1) $$\mathcal{E} = [(\mathcal{X}, \mathbb{B}), \{p(x|\theta)\,|\,\theta \in \Theta\}]$$

characterizes an experiment.

With this definition, the notion of the experiment corresponds to the following communication system with noise. It consists of essentially two parts:

(1) The input space is the set Θ of input symbols θ. These symbols are transmitted one by one by some discrete atochastic process, in which each choice of θ is made with probability $p(\theta)$, successive choices being independent.

(2) The noisy channel is such that the output space is a set $\mathfrak{X}$. We assume that successive symbols are independently perturbed by the noise and the channel, therefore, is described by the set of conditional probabilities $p(x|\theta)$, $\theta \in \Theta$, the probability of a transmitted symbol θ being received as $x \in \mathfrak{X}$.

The transmission rate of the noise channel was defined by Shannon as the amount of decrease of input entropy to the received conditional entropy. Hence, if the a priori pdf of input symbols θ is $p(\theta)$, the rate is given by

$$-\int p(\theta) \log p(\theta)\, d\theta - \left(-\int p(x)\, dx \int p(\theta \,|\, x) \log p(\theta \,|\, x)\, d\theta\right)$$

$$= \iint p(\theta)\, p(x \,|\, \theta) \log \frac{p(x|\theta)}{p(x)}\, d\theta\, dx \tag{1. 2}$$

where $p(x) = \int p(\theta)\; p(x|\theta) d\theta$ and $p(\theta\,|x) = \dfrac{p(\theta)p(x|\theta)}{p(x)}$ (here, for simplicity in notation, we do not attempt to be specific in describing the density functions. Thus $p(\theta)$ and $p(x)$, denote the density functions of the random variables θ and x, respectively, without any suggestion that they have the same density. Moreover we denote integration with respect to the dominating measures on θ and x by $d\theta$ and dx

respectively.)

Lindley [5] proposed to define the amount of information provided by the experiment (1.1) with the prior knowledge $p(\theta)$ by the Shannon information (1.2). He studied various interesting properties of information (1.2) by exploiting certain additive properties of the functional

$$H(p(\cdot)) = -\int p(\theta)\log p(\theta)\, d\theta$$

for composite experiments. Recently DeGroot [2] showed that similar arguments can be made quite generally for information functions derived from any other concave function (in probability distributions) which plays the same role as of $H(p(\cdot))$. This paper discusses in some detail the amount of information provided by a dichotomous experiment, i.e., an experiment with an input space which consists of only two elements. Some of the main results (Theorems 3.1, 4.1 and 4.2) obtained can be extended without difficulty to the case of finite experiments, i.e., experiments with finite input space.

§2. Amount of Information Provided by a Dichotomous Experiment.

Let Θ be a two-element set $\{1, 0\}$, and let $\mathcal{E}$ be a dichotomous experiment

$$(2.1) \qquad \mathcal{E} = [(\mathcal{X}, \mathbb{B}), \{f_1(x), f_0(x)\}]$$

where $(\mathcal{X}, \mathbb{B})$ is a measurable space and μ_i $(i = 1, 0)$ are probability measures with $d\mu_i = f_i(x)\, d\lambda$ $(i = 1, 0)$. The ordered pair of the two densities $\{f_1, f_0\}$ means that $f_\theta(x)$ $(\theta = 1, 0)$ corresponds to the input symbol θ.

The two probability densities f_1 and f_0 generate a parametric family of densities

(2. 2) $$\mathcal{F} = \{ f_\zeta(x) = \zeta f_1(x) + (1-\zeta) f_0(x) \mid 0 \leqq \zeta \leqq 1 \} ,$$

where ζ may be interpreted as the proportion in the mixture of two densities or the prior probability for the hypothesis H_1 which states that $f_1(x)$ is the true probability density.

An uncertainty function $U(\zeta)$ is a non-negative measurable function defined on $0 \leqq \zeta \leqq 1$. Intuitively, the value $U(\zeta)$ is meant to represent the uncertainty of an experimenter (before performing an experiment) about which hypothesis is true when his prior knowledge is the probability ζ for the hypothesis H_1.

The information in this dichotomy with the uncertainty function U, is defined by

(2. 3) $$\mathcal{I}[\mathcal{E}, \zeta ; U] = U(\zeta) - E[U(\zeta(X)) \mid \zeta] ,$$

where $\zeta(x) = \zeta f_1(x)/f_\zeta(x)$ is the posterior probability for H_1 after observing x when using ζ as the prior probability for H and where the expectation $E[\cdot \mid \zeta]$ means $\zeta E_1[\cdot] + (1 - \zeta) E_0[\cdot]$.

It is shown by DeGroot [2] that

$$\mathcal{I}[\mathcal{E}, \zeta ; U] \geqq 0 , \quad \text{for all } 0 \leqq \zeta \leqq 1 ,$$

if and only if the uncertainty function U is concave.

We discuss some examples of concave uncertainty functions in the following. All of these uncertain functions are also continuous unimodal (in fact, symmetric about $\zeta = \frac{1}{2}$, except (d)) functions with $U(0) = U(1) = 0$.

(a) Shannon entropy function

$$U(\zeta) = -\zeta \log \zeta - (1-\zeta)\log(1-\zeta) \quad (\equiv H(\zeta), \text{ say}) \tag{2.4}$$

Then we have

$$\begin{aligned} E[H(\zeta|X))|\zeta] &= -\zeta \int f_1(x) \log \frac{\zeta f_1(x)}{f_\zeta(x)} d\lambda - (1-\zeta) \int f_0(x) \log \frac{(1-\zeta) f_0(x)}{f_\zeta(x)} d\lambda \\ &= H(\zeta) - \left(\zeta \int f_1(x) \log \frac{f_1(x)}{f_\zeta(x)} d\lambda + (1-\zeta) \int f_0(x) \log \frac{f_0(x)}{f_\zeta(x)} d\lambda\right). \end{aligned}$$

Hence

$$\mathcal{J}[\mathcal{E}, \zeta ; H] = \zeta I(f_1 : f_\zeta) + (1-\zeta) I(f_0 : f_\zeta) , \tag{2.5}$$

where

$$I(f_i : f_\zeta) = \int f_i(x) \log \frac{f_i(x)}{f_\zeta(x)} d\lambda \qquad (i = 1, 0) \tag{2.6}$$

is the Kullback-Leibler information number (Kullback [4]) for discriminating between two densities $f_i(x)$ and $f_\zeta(x)$.

Under some regularity conditions which guarantee the validity of interchanging the order of integration with respect to $d\lambda$ and differentiation with respect to ζ , we can find that the function $\mathcal{J}[\mathcal{E}, \zeta ; H]$ is twice-differentiable and concave with

$$\frac{d}{d\zeta}\mathcal{J}[\mathcal{E},\zeta;H]\Big|_{\zeta=0} = I(f_1 : f_0)\ ,\ \frac{d}{d\zeta}\mathcal{J}[\mathcal{E},\zeta;H]\Big|_{\zeta=1} = -I(f_0 : f_1)$$

and

$$\frac{d^2}{d\zeta^2}\mathcal{J}[\mathcal{E},\zeta;H] = -\int (f_1 - f_0)^2 / f_\zeta \, d\lambda\ ,$$

where the last expression is equal to the negative of the Fisher information in $\mathcal{F}$ for estimating the proportion ζ :

$$i(\zeta) \equiv \int \left\{ -\frac{\partial^2}{\partial\zeta^2} \log f_\zeta(x) \right\} f_\zeta(x)\, d\lambda$$

$$= \int (f_1 - f_0)^2 / f_\zeta \, d\lambda\ .$$

(b) The other important concave uncertainty function is

(2.7) $$U(\zeta) = \min(\zeta,\ 1-\zeta) \quad (\equiv \psi(\zeta),\ \text{say.})$$

Using this uncertainty measure,

$$\mathcal{J}[\mathcal{E},\zeta;\psi] = \psi(\zeta) - E[\psi(\zeta(X)) \mid \zeta]$$

$$= \psi(\zeta) - \left(\zeta \int_{f_0/f_1 > \zeta/(1-\xi)} f_1(x)\,d\lambda + (1-\zeta)\int_{f_0/f_1 \leqq \zeta/(1-\zeta)} f_0(x)\,d\lambda \right)$$

is the reduction of the risk of Bayes decision rule deciding between two densities f_1 and f_0 with usual zero-one loss. We can show that the above information attains its maximum value at $\zeta = 1/2$. The proof is as follows: we can rewrite as

$$\mathcal{I}[\mathcal{E}, \zeta ; \psi] = \min(\zeta - E[\psi(\zeta(X)) \mid \zeta] , 1-\zeta - E[\psi(\zeta(X)) \mid \zeta])$$

$$= \min(h_1(\zeta), h_2(\zeta)), \qquad \text{say.}$$

Since $E[\psi(\zeta(X)) \mid \zeta]$ is a continuous and concave function in $0 \leqq \zeta \leqq 1$, $h_1(\zeta)$ and $h_2(\zeta)$ are both continuous and convex, with $h_1(0) = h_2(1) = 0$. The desired result follows from the fact that the minimum of two convex functions has a maximum at one of the end points or at the point at which these functions have equal values.

(c) Let us take

$$U(\zeta) = \zeta(1 - \zeta) \quad (\equiv U^0(\zeta), \quad \text{say}). \tag{2.8}$$

Then

$$\mathcal{I}[\mathcal{E}, \zeta ; U^0] = \zeta(1-\zeta)(1 - \int f_1 f_0/f_\zeta \, d\lambda). \tag{2.9}$$

Since Fisher's information in $\mathcal{F}$ for estimating the proportion ζ is

$$i(\zeta) = \int (f_1 - f_0)^2/f_\zeta \, d\lambda = -\frac{1}{\zeta(1-\zeta)} \int (f_\zeta - f_1)(f_\zeta - f_0)/f_\zeta \, d\lambda$$

$$= \frac{1}{\zeta(1-\zeta)} (1 - \int f_1 f_0/f_\zeta \, d\lambda) ,$$

we obtain

$$\mathcal{I}[\mathcal{E}, \zeta ; U^0] = \zeta^2(1-\zeta)^2 i(\zeta) .$$

A power series expansion was obtained by Hill [3] for

$$S(\zeta) \equiv \int f_1 f_0 / f_\zeta \, d\lambda \tag{2.10}$$

which was then investigated in detail for the case of two exponential densities and for the cases of two normal densities with equal variances.

From (2.9) and (2.10) we see that the information as to which population an observation comes from in the mixture f_ζ , compared with the initial uncertainty, is reflected in the factor $\mathcal{I}[\mathcal{E}, \zeta ; U^0]/U^0(\zeta) = 1 - S(\zeta)$.

Under the usual regularity conditions $S(\zeta)$ is a twice-differentiable, convex function with $0 \leqq S(\zeta) \leqq S(0) = S(1) = 1$, and hence has a unique minimum value. Thus

$$\max_{0 \leqq \zeta \leqq 1} \frac{\mathcal{I}[\mathcal{E}, \zeta ; U^0]}{U^0(\zeta)} = 1 - \min_{0 \leqq \zeta \leqq 1} S(\zeta) \tag{2.11}$$

(d) For $0 \leq t \leq 1$, the function

$$U(\zeta) = \zeta^t(1-\zeta)^{1-t} \quad (\equiv U_t(\zeta), \quad \text{say}) \tag{2.12}$$

is also continuous and concave, and we have

$$\mathcal{I}[\mathcal{E}, \zeta ; U_t] = \zeta^t(1-\zeta)^{1-t}(1 - \int f_1^t f_0^{1-t} \, d\lambda).$$

In the above expression, the integral

(2.13) $$M_0(t) \equiv \int [f_1(x)]^t [f_0(x)]^{1-t} \, d\lambda$$

is the moment generating function of the random variable $z = \log f_1(X)/f_0(X)$ under H_0 and is continuous and convex in $0 \leqq t \leqq 1$, with

$$0 \leqq M_0(t) \leqq M_0(0) = M_0(1) = 1, \; M_0'(0) = -I(f_0:f_1), \; M_0'(1) = I(f_1:f_0).$$

The number $-\log \rho \equiv -\log(\min_{0 \leqq t \leqq 1} M_0(t))$ is called Chernoff information number [1] for deciding between $f_1(x)$ and $f_0(x)$. We at once obtain, for all $0 < t < 1$,

(2.14) $$\frac{\mathcal{R}[\mathcal{E}, \zeta ; U_t]}{U_t(\zeta)} = 1 - M_0(t) \leqq 1 - e^{-(-\log \rho)}$$

independently of $0 \leqq \zeta \leqq 1$.

In this connection it is natural to consider the family of pdf

$$\mathcal{H} \equiv \left\{ h_\zeta(x) = [f_1(x)]^\zeta \, [f_0(x)]^{1-\zeta} / M_0(\zeta) \;\middle|\; 0 \leqq \zeta \leqq 1 \right\}$$

regarding ζ as an unknown parameter. We can show

Theorem 2.1 (a) The Fisher information contained in $\mathcal{H}$ is

$$i(\zeta) \equiv \int \left\{ -\frac{\partial^2}{\partial \zeta^2} \log h_\zeta(x) \right\} h_\zeta(x) \, d\lambda = \frac{\partial^2}{\partial \zeta^2} \log M_0(\zeta).$$

(b) The maximum likelihood estimate of ζ is given by

$$\zeta = \begin{cases} 0, & \text{if } \bar{Z}_N \leqq -I(f_0 : f_1) \\ 1, & \text{if } \bar{Z}_N \geqq I(f_1 : f_0) \\ \text{The unique root of the equation} & \\ \quad \bar{Z}_N = M_0'(\zeta)/M_0(\zeta), & \text{if otherwise} \end{cases}$$

where

$$\bar{Z}_N = \frac{1}{N}\sum_{i=1}^{N} \log(f_1(X_i)/f_0(X_i)) .$$

Proof. Since

$$\log h_\zeta(x) = \zeta \log f_1(x) + (1-\zeta)\log f_0(x) - \log M_0(\zeta)$$

(a) is evident. The logarithm of the likelihood of the sample $x_1, \ldots, x_N$ is

$$L_\zeta(x_1, \ldots, x_N) \equiv \log \prod_{i=1}^{N} h_\zeta(x_i) = \zeta \sum_{i=1}^{N} \log f_1(x_i) + (1-\zeta)\sum_{i=1}^{N} \log f_0(x_i) - N \log M_0(\zeta).$$

Now $\log M_0(\zeta)$ is twice-differentiable and strictly convex. Therefore $L_\zeta(x_1, \ldots, x_N)$ is twice-differentiable and strictly concave with

$$\frac{\partial}{\partial\zeta} L_\zeta(x_1, \ldots, x_N) = \sum_{i=1}^{N} \log \frac{f_1(x_i)}{f_0(x_i)} - N\frac{M_0'(\zeta)}{M_0(\zeta)}.$$

The (b) part thus follows from the fact that $M_0'(\zeta)/M_0(\zeta)$ is strictly increasing in $0 \leqq \zeta \leqq 1$ from $-I(f_0 : f_1)$ to $I(f_1 : f_0)$.

As an illustration let us take

$$f_i(x) = (\sqrt{2\pi}\,\sigma)^{-1} \exp[-\frac{1}{2\sigma^2}(x-\mu_i)^2], \quad (i = 1, 0 ; \mu_0 < \mu_1).$$

Letting $\bar{\mu} = (\mu_0 + \mu_1)/2, \;\; \delta = (\mu_1 - \mu_0)/\sigma$ we readily obtain that

$$M_0(\zeta) = \exp.\,[-\,\zeta(1-\zeta)\delta^2/2]$$

$$i(\zeta) = \frac{\partial^2}{\partial\zeta^2} \log M_0(\zeta) = \delta^2$$

$$I(f_0 : f_1) = I(f_1 : f_0) = \delta^2/2$$

$$\frac{1}{N}\sum_{i=1}^{N} \log \frac{f_1(X_i)}{f_1(x_i)} = (\bar{x}_N - \bar{\mu})\,\delta/ \qquad\qquad (\bar{x}_N \equiv N^{-1}\sum_1^N x_i)$$

and hence

$$\hat{\zeta} = \begin{cases} 0, & \text{if } \bar{x}_N \leqq \mu_0 \\ (\bar{x}_N - \mu_0)/(\mu_1 - \mu_0), & \text{if } \mu_0 < \bar{x}_N < \mu_1 \\ 1\,, & \text{if } \bar{x}_N \geqq \mu_1 \,. \end{cases}$$

§3. Independent Replications of an Experiment.

Let $\mathcal{E}^{(n)}$ be an independent replication of a dichotomous experiment $\mathcal{E}$ i.e.,

$$\mathcal{E}^{(n)} = [\,(\mathcal{X}^n, \mathbb{B}^n), \{f_1^{(n)}(x_1, \ldots, x_n), f_0^{(n)}(x_1, \ldots, x_n)\}\,]$$

where $f_i^{(n)}(x_1, \ldots, x_n) = \prod_{j=1}^{n} f_i(x_j) (i = 1, 0)$. Then

(3.1) $$\mathcal{I}[\mathcal{E}^{(n)}, \zeta ; U] = U(\zeta) - E[U(\zeta(X_1,\ldots,X_n))|\zeta] .$$

First we show

Theorem 3.1. If $U(\zeta)$ is a concave uncertainty function, then $\mathcal{I}[\mathcal{E}^{(n)}, \zeta ; U]$ is increasing in n. Moreover, if $\mathcal{I}[\mathcal{E}, \zeta ; U]$ is concave in ζ then $\mathcal{I}[\mathcal{E}^{(n)}, \zeta ; U]$ is concave in n.

Proof. For any dichotomous experiment, as is easily proved, $E[U(\zeta(X))|\zeta]$ is concave in ζ if $U(\zeta)$ is concave (DeGroot [2]). Thus each function $e_n(\zeta) \equiv E[U(\zeta(X_1,\ldots,X_n))|\zeta]$, for $n = 1,2,\ldots$, is concave in ζ. By Bayes theorem and concavity of $e_{n-1}(\zeta)$ we have

$$\begin{aligned} e_n(\zeta) &= E[E\{U(\zeta(X_1)(X_2,\ldots,X_n))|\zeta(X_1)\}|\zeta] \\ &= E[e_{n-1}(\zeta(X_1))|\zeta] \\ &\leqq e_{n-1}(E[\zeta(X_1)|\zeta]) = e_{n-1}(\zeta), \quad (n=1,2,\ldots ; e_0(\zeta)=U(\zeta)). \end{aligned}$$

Thus the increasing property is evident since $\mathcal{I}[\mathcal{E}^{(n)}, \zeta ; U] = U(\zeta) - e_n(\zeta)$.

To prove the concavity of $\mathcal{I}[\mathcal{E}^{(n)}, \zeta ; U]$ in n, it suffices to show that

(3.2) $$e_n(\zeta) - e_{n+1}(\zeta) \leqq e_{n-1}(\zeta) - e_n(\zeta).$$

Since

(3.3) $$e_n(\zeta) - e_{n+1}(\zeta) = E[e_{n-1}(\zeta(X_1)) - e_n(\zeta(X_1))|\zeta] ,$$

we use mathematical induction. If $e_0(\zeta) - e_1(\zeta) = U(\zeta) - E[U(\zeta(X_1))|\zeta]$ is concave in ζ , then

$$e_1(\zeta) - e_2(\zeta) = E[\, e_0(\zeta(X_1)) - e_1(\zeta(X_1)) \,|\, \zeta\,]$$

is also concave, and

$$\leqq e_0(E[\, \zeta(X_1) \,|\, \zeta]) - e_1(E[\zeta(X_1) \,|\, \zeta\,]) = e_0(\zeta) - e_1(\zeta) .$$

The recurrence relation (3. 3) yields (3. 2).

Theorem 3. 2 If $U(\zeta)$ is continuous in $0 \leqq \zeta \leqq 1$ with $U(1) = 0$ and if $\lim_{\zeta \to 0} U(\zeta)/\zeta$ exists, then

$$\lim_{n \to \infty} \mathcal{J}[\mathcal{E}^{(n)}, \zeta\, ; U] = U(\zeta) .$$

Proof. Let $\{Z_n\}$ be a sequence of independent random variables identically distributed as a random variable Z with finite mean. Let $\{g_n(\cdot)\}$ be a sequence of functions which satisfies (a) $\{g_n(\cdot)\}$ is uniformly bounded, (b) $\{g_n(\cdot)\}$ is uniformly continuous in some interval about $E(Z)$, and (c) $\lim_{n \to \infty} g_n(E(Z))$ exists, then

$$\lim_{n \to \infty} E[\, g_n(n^{-1}(Z_1 + \ldots + Z_n))] = \lim_{n \to \infty} g_n(E(Z)) .$$

The proof of this fact is not difficult, so it will be omitted (see, for example Parzen [6]).

Now $e_n(\zeta)$ is rewritten as

$$e_n(\zeta) = E_1[\, U(\zeta/\zeta + (1-\zeta)e^{-\sum_1^n Z_i}))(\zeta + (1-\zeta)e^{-\sum_1^n Z_i})\,]$$

with $Z_i = \log \frac{f_1(X_i)}{f_0(X_i)}$. The sequence of functions

$$g_n(t) \equiv (\zeta + (1-\zeta)e^{-nt})U(\zeta/(\zeta+(1-\zeta)e^{-nt}))$$

satisfies (a), (b) and (c), by the assumption of the theorem. Thus

$$\lim_{n \to \infty} g_n(E_1(Z)) = \zeta U(1) = 0 .$$

This completes the proof of the theorem.

The discussions in the previous section on the four examples of uncertainty function $U(\zeta)$, give, by Theorems 3. 1 and 3. 2, the following

Corollary 3. 2. 1 $\mathcal{I}[\mathcal{E}^{(n)}, \zeta ; U]$ is (a) an increasing function of n, for $U(\zeta) = \psi(\zeta)$ or $U^0(\zeta)$, (b) an increasing and strictly concave function of n, for $U(\zeta) = H(\zeta)$ or $U_t(\zeta)$ and (c)

$$\lim_{n \to \infty} \mathcal{I}[\mathcal{E}^{(n)}, \zeta ; U] = U(\zeta)$$

for $U(\zeta) = \psi(\zeta)$ or $U^0(\zeta)$.

§4. Grouping of the Observations.

Let R be any measurable subset of $\mathfrak{X}$, and let us consider the decision rule which accepts H_1 is $x \in R$ and H_0 if $x \notin R$. Or equivalently, we define the random variable

$$y = Tx = \begin{cases} 1, & \text{if } x \in R \\ 0, & \text{if } x \in R^C \end{cases}$$

The probability densities of y under the two hypotheses are binomial:

(4. 1)
$$g_1(y) = (1-\beta)^y \beta^{1-y}, \quad \text{(under } H_1) \\ = \alpha^y (1-\alpha)^{1-y}, \quad \text{(under } H_0)$$

where $y = 1, 0$ and

(4. 2)
$$\alpha = \int_R f_0(x)\, d\lambda \qquad \text{and} \qquad \beta = \int_{R^C} f_1(x)\, d\lambda$$

Thus α and β are probabilities of the two kinds of errors due to the above decision rule. Let

(4. 3)
$$\mathcal{E}_Y = [(\mathcal{Y}, \mathbb{C}), \quad g_1(y), \ g_0(y) \ ]$$

be a dichotomous (binomial) experiment, where $\mathcal{Y} = \{1, 0\}$ and $\mathbb{C}$ is an induced σ-field of subsets. Then we obtain

Theorem 4.1. For any R and for any concave uncertainty function U we have

$$\mathcal{I}[\mathcal{E}_X, \zeta ; U] \geqq \mathcal{I}[\mathcal{E}_Y, \zeta ; U], \quad \text{for all } 0 \leqq \zeta \leqq 1$$

with uniform equality in ζ if and only if

$$\frac{f_1(x)}{f_0(x)} \equiv \begin{cases} (1-\beta)/\alpha, & \text{in } R \\ \beta/(1-\alpha), & \text{in } R^C \end{cases}.$$

Thus there is always some loss of information due to grouping of the observed values of x. The corresponding theorem for the case of $U(\zeta) = H(\zeta)$ is given in Kullback [4] .

Proof. Since

$$g_\zeta(y) = \zeta g_1(y) + (1-\zeta)g_0(y) = \begin{cases} \zeta(1-\beta)+(1-\zeta)\alpha\ , & y=1 \\ \zeta\beta+(1-\zeta)(1-\alpha)\ , & y=0 \end{cases}$$

$$\zeta(y)g_\zeta(y) = \zeta g_1(y) = \begin{cases} \zeta(1-\beta), & y=1 \\ \zeta\beta\ , & y=0\ . \end{cases}$$

we get

$$E[U(\zeta(Y))\,|\,\zeta] = (\zeta(1-\beta)+(1-\zeta)\alpha)\,U\left(\frac{\zeta(1-\beta)}{\zeta(1-\beta)+(1-\zeta)\alpha}\right)$$

(4. 4)
$$+ (\zeta\beta+(1-\zeta)(1-\alpha))U\left(\frac{\zeta\beta}{(\zeta\beta+(1-\zeta)(1-\alpha))}\right)$$

But

$$E[U(\zeta(X))\,|\,\zeta] = \left(\int_R + \int_{R^C}\right) U\left(\frac{\zeta f_1(x)}{f_\zeta(x)}\right) f_\zeta(x)\,d\lambda$$

$$\leqq \left(\int_R f_\zeta(x)\,d\lambda\right) U\left(\zeta\int_R f_1(x)\,d\ \Big/ \int_R f_\zeta(x)\,d\lambda\right)$$

$$+\left(\int_{R^C} f_\zeta(x)\,d\lambda\right) U\left(\zeta\int_{R^C} f_1(x)\,d\lambda \Big/ \int_{R^C} f_\zeta(x)\,d\lambda\right)$$

by concavity of U, and this last expression is, from (3. 1) and (4. 4), equal to $E[U(\zeta(Y))\,|\,\zeta]$. This completes the proof of the theorem.

Because of the definition (4. 2), the two error probabilities α and β cannot be made arbitrarily small simultaneously. One explanation of this fact is given by a pair of the well-known information inequalities (Kullback [4]).

$$I(f_0 : f_1) = \left(\int_R + \int_{R^C}\right) f_0 \log \frac{f_0}{f_1} \, d\lambda \geqq \alpha \log \frac{\alpha}{1-\beta} + (1-\alpha) \log \frac{1-\alpha}{\beta} \quad (\equiv F(\alpha, \beta), \text{ say}),$$

$$I(f_1 : f_0) = \left(\int_R + \int_{R^C}\right) f_1 \log \frac{f_1}{f_0} \, d\lambda \geqq (1-\beta) \log \frac{1-\beta}{\alpha} + \beta \log \frac{\beta}{1-\alpha} .$$

For any fixed $0 < \alpha \leq 1$, $F(\alpha, \beta)$ is, as a function of β, strictly convex and tends to $+\infty$, as $\beta \to 0$. Thus the upper one of the above inequalities implies that the possible value of β has a positive lower bound. Theorems 4. 3 and 4. 4 which we are going to state and prove later in this section give another explanation to that same fact.

Let

$$\mathcal{E}' = [(\mathcal{X}, \mathbb{B}), \{f_1(x), f_0(x)\}] ,$$

$$\mathcal{E}'' = [(\mathcal{X}, \mathbb{B}), \{h_1(x), h_0(x)\}]$$

be two dichotomous experiments with the common input space $\Theta = \{1, 0\}$ and the same output space $\mathcal{X}$. For any $0 \leqq t \leqq 1$, we call an experiment

$$[(\mathcal{X}, \mathbb{B}), \{tf_1(x) + (1-t)h_1(x), \ tf_0(x) + (1-t)h_0(x)\}]$$

denoted by $t\mathcal{E}' * (1-t)\mathcal{E}''$ symbolically, a <u>mixture</u> of $\mathcal{E}'$ and $\mathcal{E}''$ (with weight t on $\mathcal{E}'$). We shall show the following

<u>Theorem 4. 2</u> For any fixed $0 \leqq \zeta \leqq 1$ and concave uncertainty function $U(\zeta)$, the information $\mathcal{J}[\mathcal{E}, \zeta ; U]$ is convex in $\mathcal{E}$, <u>i. e.,</u> for all $0 \leqq t \leqq 1$,

$$\mathcal{I}[t\mathcal{E}'*(1-t)\,\mathcal{E}'',\ \zeta\,;\,U] \leqq t\mathcal{I}[\mathcal{E}',\ \zeta\,;\,U] + (1-t)\,\mathcal{I}[\mathcal{E}'',\ \zeta\,;\,U]$$

Proof. Writing the information $\mathcal{I}[\mathcal{E},\zeta\,;\,U]$ simply by $\mathcal{I}[\mathcal{E}]$,

$$\mathcal{I}[t\mathcal{E}'*(1-t)\,\mathcal{E}''] = U(\zeta) - \int U\left(\frac{\zeta(tf_1(x)+(1-t)h_1(x))}{tf_\zeta(x)+(1-t)h_\zeta(x)}\right)$$

(4. 6)

$$\cdot\ (tf_\zeta(x)+(1-t)h_\zeta(x))\,d\lambda\,.$$

Since

$$U\left(\frac{\zeta(tf_1+(1-t)h_1)}{tf_\zeta+(1-t)h_\zeta}\right) = U\left(\left(tf_\zeta\frac{\zeta f_1}{t_\zeta}+(1-t)h_\zeta\frac{\zeta h_1}{h_\zeta}\right)/(tf_\zeta+(1-t)h_\zeta)\right)$$

$$\geqq \left\{tf_\zeta U(\zeta f_1/f_\zeta)+(1-t)h_\zeta U(\zeta h_1/h_\zeta)\right\}/(tf_\zeta+(1-t)h_\zeta)$$

by concavity of U, it follows from (4. 6) that

$$\mathcal{I}[t\mathcal{E}'*(1-t)\mathcal{E}''] \leqq U(\zeta) - \int\left\{tf_\zeta(x)U\left(\frac{\zeta f_1(x)}{f_\zeta(x)}\right)+(1-t)h_\zeta(x)U\left(\frac{\zeta h_1(x)}{h_\zeta(x)}\right)\right\}\,d\lambda$$

$$= t\left\{U(\zeta) - E[U(\zeta(X))\mid\zeta\,;\,\mathcal{E}']\right\} + (1-t)\left\{U(\zeta) - E[U(\zeta(X))\mid\zeta\,;\,\mathcal{E}'']\right\}$$

$$= t\mathcal{I}[\mathcal{E}'] + (1-t)\,\mathcal{I}[\mathcal{E}'']\,.$$

The corresponding theorem for the case of $U(\zeta) = H(\zeta)$ is given by Lindley [5] .

The mixture $t\,\mathcal{E}'*(1-t)\,\mathcal{E}''$, defined by (4. 5), can be thought of as follows: A value x is obtained through the experiment $\mathcal{E}'$ or $\mathcal{E}''$ with probabilities t and $1-t$ respectively. The experimenter is informed only of x, and not of which event of probability t or $1-t$ took place. Using the above Theorem 4. 2 we shall prove the next Theorem 4. 3.

Consider the dichotomous binomial experiment $\mathcal{E}_Y$, defined by (4. 3), where two densities have probability of $Y = 1$, $1-\beta$ and α, respectively. Suppose that the experimenter performs $\mathcal{E}_Y$ and suffers cost d_{iy} $(i, y=1, 0)$ if he receives the output y when the input was i. If the prior probability of sending the input $i=1$ as ζ and the cost is given by

$$\begin{pmatrix} d_{11} & d_{10} \\ d_{01} & d_{00} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

then the expected cost due to performing the experiment $\mathcal{E}_Y$ is $\zeta\beta + (1-\zeta)\alpha$.

Now we consider the information $\mathcal{I}[\mathcal{E}_Y, \zeta\ ; U]$ as a function of (α, β). Take any $0 \leqq t, \alpha', \beta', \alpha'', \beta' \leqq 1$ and let

$$\text{(4. 7)} \qquad \left.\begin{aligned} \mathcal{E}'_Y &= [(\mathcal{Y}, \mathbb{C}), \{g'_1(y), g'_0(y)\}] \\ \mathcal{E}''_Y &= [(\mathcal{Y}, \mathbb{C}), \{g''_1(y), g''_0(y)\}] \end{aligned}\right\}$$

where $(\mathcal{Y}, \mathbb{C})$ is the same measurable space defined in (4. 3) and $g'_1(y)$ and $g'_0(y)$ are binomial densities with the probability of $Y=1$, $1-\beta'$ and α' respectively. Similarly $g''_1(y)$ and $g''_0(y)$ are binomial densities with $1-\beta''$ and α'', respectively. Then from the definition (4. 5) of the mixture,

$$\text{(4. 8)} \qquad t\,\mathcal{E}'_Y * (1-t)\,\mathcal{E}''_Y = [(\mathcal{Y}, \mathbb{C}), \{h_1(y), h_0(y)\}],$$

where $h_1(y)$ and $h_0(y)$ are binomial densities with the probability of $Y=1$, $t(1-\beta') + (1-t)(1-\beta'') = 1 - (+\beta' + (1-t)\beta'')$ and $t\alpha' + (1-t)\alpha''$, respectively. Therefore it follows, from Theorem 4. 2, that

Corollary 4. 2. 1. $\mathcal{I}[\mathcal{E}_Y, \zeta; U]$ is a continuous convex function of $0 \leqq \alpha, \beta \leqq 1$, if the uncertainty function $U(\zeta)$ is continuous and concave in $0 \leqq \zeta \leqq 1$.

These considerations lead to the following minimization problem: For any fixed $0 < \zeta < 1$.

$$\mathcal{I}[\mathcal{E}_Y, \zeta ; U] = U(\zeta) - E[U(\zeta(Y)) | \zeta] \longrightarrow \min_{\alpha, \beta} .$$

where $E[U(\zeta(Y)) | \zeta]$ is given by (4. 2) and (4. 3), with the constraints that

$$\zeta\beta + (1-\zeta)\alpha \leqq d$$
$$0 \leqq \alpha, \beta, \alpha+\beta \leqq 1 .$$

Since from (4. 3), $\mathcal{I}[\mathcal{E}_Y, \zeta ; U]$ is symmetric about $\alpha = \beta = \frac{1}{2}$ we may consider the points with $\alpha + \beta \leqq 1$ only.

Let the minimizing value be denoted by $R(d ; \zeta, U)$. For the case in which $U(\zeta) = H(\zeta)$, this is the rate-distortion function, introduced by Shannon [8] , for the binomial dichotomous experiment. $\mathcal{E}_Y$.

Theorem 4. 3. For any fixed $0 < \zeta < 1$ and concave uncertainty function U. $R(d; \zeta, U)$ is a convex decreasing function in $0 \leqq d \leqq \max(\zeta, 1-\zeta)$, from $U(\zeta) - \zeta U(1) - (1-\zeta)U(0)$ at $d = 0$, to 0, at $d = \max(\zeta, 1-\zeta)$.

Proof. The decreasing property is clear. Hence we shall prove the convexity only. Take two arbitrary points on the curve. Let the minimizing choice of values for the problem corresponding to d be denoted by (α', β') , and for d'', by (α'', β''). Consider $t(\alpha', \beta') + (1-t)(\alpha'', \beta'')$, for $0 \leqq t \leqq 1$. Using the definition (4. 7) and the result (4. 8), we get, by Theorem 4. 2 and the definition of $R(d; \zeta, U)$,

$$\mathcal{I}[t\mathcal{E}'_Y * (1-t)\mathcal{E}''_Y, \zeta; U] \leqq t\mathcal{I}[\mathcal{E}'_Y, \zeta; U] + (1-t)\mathcal{I}[\mathcal{E}''_Y, \zeta; U]$$

(4. 9)

$$= tR(d'; \zeta, U) + (1-t)R(d''; \zeta, U)$$

for any concave function U.

Now since

$$\zeta(t\beta' + (1-t)\beta'') + (1-\zeta)(t\alpha' + (1-t)\alpha'')$$

$$= (t(\zeta\beta' + (1-\zeta)\alpha'') + (1-t)(\zeta\beta'' + (1-\zeta)\alpha'') \leqq td' + (1-t)d'' ,$$

and

$$0 \leqq t\alpha' + (1-t)\alpha'', \quad t\beta' + (1-t)\beta'', \quad t(\alpha' + \beta') + (1-t)(\alpha'' + \beta'') \leqq 1 ,$$

we obtain from (4. 9) that

$$R(td' + (1-t)d''; \zeta, U) = \min \mathcal{I}[t\mathcal{E}'_Y * (1-t)\mathcal{E}''_Y, \zeta; U]$$

$$\leqq tR(d'; \zeta, U) + (1-t)R(d''; \zeta, U).$$

Thus we have proved the convexity of $R(d; \zeta, U)$ in d.

We show, in the following, various amounts of information after applying the decision rule.

(a) From (2. 4), (2. 5) and (2. 6), we obtain

$$\mathcal{I}[\mathcal{E}_Y, \zeta; H] = \zeta I(g_1 : g_\zeta) + (1-\zeta)I(g_0 : g_\zeta)$$

$$= \zeta\left((1-\beta)\log\frac{1-\beta}{\zeta(1-\beta)+(1-\zeta)\alpha} + \beta\log\frac{\beta}{\zeta\beta+(1-\zeta)(1-\alpha)}\right)$$

$$+ (1-\zeta)\left(\alpha\log\frac{\alpha}{\zeta(1-\beta)+(1-\zeta)\alpha} + (1-\alpha)\log\frac{1-\alpha}{\zeta\beta+(1-\zeta)(1-\alpha)}\right)$$

(4.10) $$= H(\zeta(1-\beta) + (1-\zeta)\alpha) - \zeta H(1-\beta) - (1-\zeta) H(\alpha).$$

(b) From (2.7) and (3.2) we get

$$E[\psi(\zeta(Y)) | \zeta] = (\zeta(1-\beta) + (1-\zeta)\alpha)\psi\left(\frac{\zeta(1-\beta)}{\zeta(1-\beta)+(1-\zeta)\alpha}\right)$$

$$+ (\zeta\beta+(1-\zeta)(1-\alpha))\psi\left(\frac{\zeta\beta}{\zeta\beta+(1-\zeta)(1-\alpha)}\right)$$

$$= \min(\zeta(1-\beta), (1-\zeta)\alpha) + \min(\zeta\beta, (1-\zeta)(1-\alpha))$$

$$= \begin{cases} \zeta, & \text{if } 0 \leqq \zeta < \alpha/(1+\alpha-\beta) \\ \alpha + \zeta(\beta-\alpha), & \text{if } \alpha/(1+\alpha-\beta) \leqq \zeta < (1-\alpha)/(1-\alpha+\beta) \\ 1-\zeta, & \text{if } (1-\alpha)/(1-\alpha+\beta) \leqq \zeta \leqq 1, \end{cases}$$

if $\alpha+\beta < 1$. We thus have

$$\mathcal{I}[\mathcal{E}_Y, \zeta; \psi] = \psi(\zeta) - E[\psi(\zeta(Y)) | \zeta]$$

$$= \begin{cases} 0, & \text{if } 0 \leqq \zeta < \alpha/(1+\alpha-\beta) \\ \zeta(1+\alpha-\beta)-\alpha, & \text{if } \alpha/(1+\alpha-\beta) \leqq \zeta < 1/2 \\ 1-\alpha-\zeta(1-\alpha+\beta), & \text{if } \frac{1}{2} \leqq \zeta < (1-\alpha)/(1-\alpha+\beta) \\ 0, & \text{if } (1-\alpha)/(1-\alpha+\beta) \leqq \zeta \leqq 1. \end{cases}$$

(c) From (2.8), (2.9) and (2.10) we obtain

$$S_Y(\zeta) \equiv \int g_1 g_0/g_\zeta \, d\lambda T^{-1}(y)$$

$$= \frac{(1-\beta)\alpha}{\zeta(1-\beta)+(1-\zeta)\alpha} + \frac{\beta(1-\alpha)}{\zeta\beta+(1-\zeta)(1-\alpha)}$$

and

$$\mathcal{I}[\mathcal{E}_Y, \zeta ; U^0] = \zeta(1-\zeta)\left(1 - \frac{(1-\beta)\alpha}{\zeta(1-\beta)+(1-\zeta)\alpha} - \frac{\beta(1-\alpha)}{\zeta\beta+(1-\zeta)(1-\alpha)}\right) .$$

Since the convex function $C(Ax+B)^{-1} + D(1-Ax-B)^{-1}$, $(C, D > 0 \neq A)$, has the absolute minimum value $(\sqrt{C} + \sqrt{D})^2$ at x satisfying $\sqrt{C}(Ax+B)^{-1} = \sqrt{D}(1-Ax-B)^{-1}$, we get by (2. 11)

$$\max_{0 \leqq \zeta \leqq 1} \frac{\mathcal{I}[\mathcal{E}_Y, \zeta ; U^0]}{U^0(\zeta)} = 1 - \min_{0 \leqq \zeta \leqq 1} S_Y(\zeta)$$

$$= 1 - (\sqrt{(1-\beta)\alpha} + \sqrt{\beta(1-\alpha)})^2$$

(4. 11) $$= (\sqrt{(1-\alpha)(1-\beta)} - \sqrt{\alpha\beta})^2$$

(d) From (2. 12), (2. 13) and (2. 14) we have, for all $0 < t < 1$

$$\mathcal{I}[\mathcal{E}_Y, \ ; U_t] = 1 - ((1-\beta)^t \alpha^{1-t} + \beta^t (1-\alpha)^{1-t})$$

$$< 1 - \exp.\left\{- (p^* \log \frac{p^*}{\alpha} + q^* \log \frac{q^*}{1-\alpha})\right\}$$

independently of $0 \leqq \zeta \leqq 1$, where

$$p^* = 1-q^* = \left(\log \frac{\beta}{1-\alpha}\right) / \left(\log \frac{\beta}{1-\alpha} + \log \frac{\alpha}{1-\beta}\right) .$$

In what follows we want to give some examples of rate-distortion functions. Let us take the uncertainty function $H(\zeta)$. The information $\mathcal{I}[\mathcal{E}_Y, \zeta ; H]$ is given by (4. 10). Since $H(\zeta)$ is continuous and strictly concave, we find that, by Corollary 4. 2. 1, $\mathcal{I}[\mathcal{E}_Y, \zeta ; H]$ is a continuous and strictly convex function of (α, β).

Theorem 4.4. For any given ζ and d with $0 \leqq d \leqq \max(\zeta, 1-\zeta)$, the minimum value

$$\min_{\substack{\zeta\beta+(1-\zeta)\alpha = d \\ 0 \leqq \alpha, \beta \leqq 1}} \mathcal{J}[\mathcal{E}_Y, \zeta; H]$$

is attained at $\alpha^* = \beta^* = d$, independently of ζ.

Proof. Using the Lagrange multiplier ℓ, we compute equations

$$\frac{\partial}{\partial\alpha} \text{ and } \frac{\partial}{\partial\beta} \text{ of } [H(\zeta(1-\beta)+(1-\zeta)\alpha) - \zeta H(1-\beta)-(1-\zeta)H(\alpha) - \ell(\zeta\beta\ (1-\zeta)\alpha)] = 0$$

and obtain, by eliminating ℓ,

$$\zeta\beta + (1-\zeta)(1-\alpha) = \frac{\sqrt{\beta(1-\alpha)}}{\sqrt{\beta(1-\alpha)} + \sqrt{\alpha(1-\beta)}} .$$

This equation, together with the condition

(4.12) $$\zeta\beta + (1-\zeta)\alpha = d$$

gives after eliminating ζ

(4.13) $$d = \frac{\sqrt{\alpha\beta}}{\sqrt{(1-\alpha)(1-\beta)} + \sqrt{\alpha\beta}} ,$$

or equivalently,

(4.14) $$(\alpha^{-1} - 1)(\beta^{-1} - 1) = (d^{-1} - 1)^2 .$$

Equation (4. 14) represents a hyperbola

$$\left(\alpha - \frac{d^2}{2d-1}\right)\left(\beta - \frac{d^2}{2d-1}\right) = \frac{d^2(1-d)^2}{(2d-1)^2} \quad , \qquad \text{if } d \neq \frac{1}{2} .$$

or the straight line $\alpha + \beta = 1$, if $d = 1/2$. For any d one branch of the hyperbola passes through the three points (0, 1), (1. 0) and (d, d). We can show that this branch and the straight line (4. 12) have one point of intersection in the region $0 \leqq \alpha, \beta \leqq 1$. This proves the theorem.

We note the following remark.

Remark. Some minimizing problems in the same line as in Theorem 4. 4 have the same minimizing points. Let

$$F(\alpha, \beta) = \alpha \log \frac{\alpha}{1-\beta} + (1-\alpha) \log \frac{1-\alpha}{\beta} \quad .$$

This is the Kullback-Leibler information $I(g_0 : g_1)$ for binomial densities g_0 and g_1 defined by (4. 1). It is well known that $F(\alpha, \beta)$, together with $F(\beta, \alpha)$, represents the efficiency of the best sequential test deciding between two simple hypotheses with the probabilities of two kinds of errors, α and β. (For example, Wald [8]).

$F(\alpha, \beta)$ is strictly convex in (α, β), for the matrix of the second order derivatives

$$\begin{pmatrix} F_{\alpha\alpha} & F_{\alpha\beta} \\ F_{\beta\alpha} & F_{\beta\beta} \end{pmatrix} = \begin{pmatrix} (\alpha(1-\alpha))^{-1} & (\beta(1-\beta))^{-1} \\ (\beta(1-\beta))^{-1} & \dfrac{\alpha}{(1-\beta)^2} + \dfrac{1-\alpha}{\beta^2} \end{pmatrix}$$

is positive definite in $0 \leqq \alpha, \beta \leqq 1$. Now consider the minimizing problem

$$\min_{\substack{\zeta\beta+(1-\zeta)\alpha = d \\ 0 \leqq \alpha,\ \beta \leqq 1}} (\zeta F(\beta, \alpha) + (1-\zeta)F(\alpha, \beta)) .$$

Legrangian equations give, by straightforward differentiation and by eliminating the Lagrange multiplier,

$$\frac{\alpha(1-\alpha)}{\beta(1-\beta)} = \left(\frac{\zeta}{1-\zeta}\right)^2 . \tag{4.15}$$

Hence, from (4. 12), after eliminating ζ, we again obtain (4. 13), or equivalently, (4. 14).

We shall give one more example. As was already shown by (4. 11), the expression

$$G(\alpha, \beta) \equiv (\sqrt{(1-\alpha)(1-\beta)} - \sqrt{\alpha\beta})^2$$

represents the maximum relative information provided by the experiment (4. 3) when using the uncertainty function $U^0(\zeta) = \zeta(1-\zeta)$. $G(\alpha, \beta)$ is also strictly convex in (α, β), since the matrix of the second order derivatives $\begin{pmatrix} G_{\alpha\alpha} & G_{\alpha\beta} \\ G_{\beta\alpha} & G_{\beta\beta} \end{pmatrix}$, where

$$G_{\alpha\alpha} = \frac{(\beta(1-\beta))^{1/2}}{2(\alpha(1-\alpha))^{3/2}} , \qquad G_{\alpha\beta} = 2\left\{1 - \frac{(\alpha-1/2)(\beta-1/2)}{\{\alpha\beta(1-\alpha)(1-\beta)\}^{1/2}}\right\}$$

can be shown to be positive definite in $0 \leqq \alpha,\ \beta \leqq 1$. Now, Lagrangian equations for the minimizing problem

$$\min_{\substack{\zeta\beta+(1-\zeta)\alpha = d \\ 0=\alpha,\ \beta = 1}} G(\alpha, \beta)$$

give, by straightforward differentiation, the same equation (4. 15) as in the foregoing example. Thus the minimizing point is given by $\alpha^* = \beta^* = d$, independently of ζ.

REFERENCES

[1] Chernoff, H., "A measure of asymptotic efficiency for the tests of a hypothesis based on the sum of observations," Annals of Mathematical Statistics, 23(1952), 492-505.

[2] DeGroot, M. H., "Uncertainty, information and sequential experiments," Annals of Mathematical Statistics, 33(1962), 404-419.

[3] Hill, B. M., "Information for estimating the proportions in mixtures of exponential and normal distributions," Journal of the American Statistical Association, 58(1963), 918-932.

[4] Kullback, S., Information Theory and Statistics, John Wiley and Sons, New York, 1959.

[5] Lindley, D. V., "On a measure of the information provided by an experiment, Annals of Mathematical Statistics, 27(1956), 986-1005.

[6] Parzen, E., Modern Probability Theory and Its Applications, John Wiley and Sons, New York, 1960.

[7] Shannon, C. E. and Weaver, W., The Mathematical Theory of Communication. The University of Illinois Press, Urbana, Ill., 1949.

[8] Shannon, C. E., "Coding theorems for a discrete source with a fidelity criterion," IRE National Convention Record, 7(1959), Part 4, 142-163.

[9] Wald, A., Sequential Analysis, John Wiley and Sons, New York, 1947.

X. Some Examples of Bayesian Adaptive Programming

Abstract: The Bayes learning in adaptive control processes is defined by the learning structure in which the unknown probability distribution is reestimated a posteriori by use of the Bayes' theorem after the random variable is observed at each stage of the process. Three kinds of measures which evaluate the effect of the Bayes learning are presented. Giving three examples in which the adaptive control problems are completely solved it is shown that the Bayes learning is sometimes unreasonable, in a certain sense, if the programming horizon is not large.

This work was supported in part by the Office of Naval Research (Grant NRO42-247).

I. Purpose of This Note

In the last few years, the mathematical theory of control processes has attracted a great deal of attention. In many control processes or multistage decision processes we face the problem of dealing with random variables whose distributions are initially imperfectly known, but which become known with increasing accuracy as the process goes. The processes that occur are 'learning' or 'adaptive' processes in which it is required to act and learn simultaneously. "Adaptive control" consists in finding the best sequence of rules of action when properties of the relevant probability distribution, to be used in choosing actions, have to be inferred from the effects of previous actions and observations. In problems of this type, which we shall call adaptive model of control processes, the best sequence of rules of action can be determined in two steps of data processing: 1) reestimating the unknown probability distribution a posteriori, and 2) computing the action on the basis of new estimate.

The Bayes learning is defined by the learning structure in which the unknown probability distribution is reestimated a posteriori by use of the Bayes' theorem after the random variable is observed at each stage of the process. In this paper we specify the category of ignorance about the unknown true distribution as follows: the unknown true distribution is determined by specifying a single parameter θ. The purpose of this note is to give some examples of Bayesian adaptive processes with worked computations of their efficacies (Sakaguchi, 1963). Though, in general, the optimal action in the Bayesian adaptive programming, is itself a functional of the posterior probability distribution given past observations, it will be shown that the optimal action is characterized by a small number of parameters if the prior distribution of θ is appropriately chosen for the control problem. It will also be shown that, in some sense, the Bayes learning is not always reasonable but it is possible that it would be worse than no learning.

II. General Formulation and Solution of the Problem

The foundation for a general theory of adaptive control processes in a mathematical framework was laid by Bellman-Kalaba (1959) and Bellman (1961). In this paper we formulate the problem as follows: we are required to choose a sequence of decisions $\{a_t\}_{t=1}^{N}$ in a given space $\mathcal{A}$ of all available decisions. The state of the process is represented by a variable x. We are given a system of two random variables R, S and a real-valued function Φ,

$$[R(a, x ; r), S(s, x ; r), \Phi(x)]$$

where r represents a random variable and we interpret $R(a, x ; r)$ as the 'immediate return' to the action when decision a is employed in state x, $S(a, x ; r)$ as the 'successor state' following x at the beginning of the succeeding time period if decision a is employed, $\Phi(x)$ as the 'value' of the final state x of the process. We assume that $\{r_t\}_{t=1}^{N}$ is a <u>sequence of independent random variables</u> with a common distribution function $G(r)$. Of course, if $G(r)$ is known and if we consider an N-stage decision process, then we are led to the problem

$$\int \cdots \int \left[\sum_{t=1}^{N} R(A_t(x_t), x_t \; ; r_t) + \Phi(x_{N+1}) \right] \prod_{t=1}^{N} dG(r_t) \longrightarrow \max_{\left\{A_t(\cdot)\right\}_{t=1}^{N}} \qquad (2.1)$$

where $\left\{A_t(\cdot)\right\}_{t=1}^{N}$ is the 'policy', i.e., a sequence of decision functions $A_t(x_t)$ and $x_{t+1} = S(A_t(x_t), x_t \; ; r_t)$ $(t = 1, \ldots, N)$.

When we are forced to take action without a complete knowledge of $G(r)$ then the problem of adaptive control arises. Let us define an <u>information pattern</u> as a sequence of information structures (Sakaguchi, 1963), i.e.

$$P = \left\{I_t(r_1, \ldots, r_{t-1})\right\}_{t=2}^{N}$$

and define a <u>learning structure</u> as a sequence

$$L = \left\{\hat{G}_1(r), \hat{G}_2(r \mid I_2), \ldots, \hat{G}_N(r \mid I_N)\right\},$$

of each estimate of $G(r)$ based on the information I_t, available on the time period t. $\hat{G}_1(r)$ is an a priori estimate of $G(r)$ without use of any information about $G(r)$. Let us now specify the category of ignorance about $G(r)$ as follows: The unknown true distribution is determined by specifying a single parameter θ, so that

$$dG = p(r, \theta)dr ,$$

where $p(r, \theta)$ is the density function of r for fixed θ. We assume complete information [3] , [5] , that is, at each time t we can know and use all observations $r_1, \dots, r_{t-1}$ already observed. Moreover, we assume the Bayes learning structure, i. e., the unknown parameter is reestimated a posteriori by use of the Bayes theorem at each stage of the process. Thus, if $\xi(\theta)$ is the prior distribution of θ, at time t is given by

$$d\hat{G}_t(r_t | r_1, \dots, r_{t-1}) = dr_t \int p(r_t, \theta) d\xi_{t-1}, \tag{2.2}$$

where

$$d\xi_{t-1} \equiv d\xi(r_1, \dots, r_{t-1}) = \frac{p(r_1, \theta) \cdots p(r_{t-1}, \theta) d\xi(\theta)}{\int p(r_1, \theta) \cdots p(r_{t-1}, \theta) d\xi(\theta)},$$

is the posterior distribution of θ given that $r_1, \dots, r_{t-1}$ are observed. (For $t = 1$, let $d\hat{G}_1(r_1) = dr_1 \int p(r_1, \theta) d\xi$. This is an a priori estimate of $G(r)$ without use of any observations).

Let I_t stand for $(r_1, \dots, r_{t-1})$. Since

$$\prod_{t=k}^{N} d\hat{G}_t(r_t | I_t) = dr_k \cdots dr_N \int \left\{ \prod_{t=k}^{N} p(r_t, \theta) \right\} d\xi_{k-1} \tag{2.3}$$

$(k = 1, \dots, N ; \xi_0 = \xi)$, a rational decision-maker under the complete information would choose the Bayes learning structure

$$L^* = \left\{\hat{G}_1(r_1),\ \hat{G}_2(r_2 | I_2),\ \ldots,\ \hat{G}_N(r_N | I_N)\right\}$$

with $\hat{G}_t(r_t | I_t)$ given by (2. 2), and then solve the maximizing problem

$$\int \cdots \int \left[\sum_{t=1}^{N} R(A_t(x_t,\ I_t),\ x_t\ ;\ r_t) + \Phi(x_{N+1})\right] \prod_{t=1}^{N} d\hat{G}_t(r_t | I_t) \longrightarrow \max_{\left\{A_t(\cdot)\right\}_{t=1}^{N}} \qquad (2.4)$$

where $\left\{A_t(\cdot)\right\}_{t=1}^{N}$ is the 'policy', i. e., a sequence of decision functions $A_t(x_t,\ I_t)$ each based on x_t and on the information I_t, and $x_{t+1} = S(A_t(x_t,\ I_t),\ x_t : r_t)$, $(t = 1,\ \ldots,\ N\ ;\ I_1 \equiv$ null information).

In principle, problems (2. 1) and (2. 4) can be solved by the "working backward" technique of dynamic programming (Bellman, 1961). For the problem (2. 4), for instance, defining

$$h_{k,N}(x_k,\ I_k) \equiv \max_{\left\{A_t(\cdot)\right\}_{t=k}^{N}} \int \cdots \int \left[\sum_{t=k}^{N} R(A_t(x_t,\ I_t),\ x_t\ ;\ r_t) + \Phi(x_{N+1})\right] \prod_{t=k}^{N} d\hat{G}_t(r_t | I_t)$$

for $k = 1,\ \ldots,\ N$, the principle of optimality (Bellman, 1961) yields

$$h_{k,N}(x, I_k) = \max_a \int [R(a, x ; r_k) + h_{k+1,N}(S(a, x ; r_k), I_{k+1})]$$

$$d\hat{G}_k(r_k | I_k) \qquad (k = 1, \ldots, N ; h_{N+1,N}(x, I_{N+1}) \equiv \Phi(x)), \qquad (2.5)$$

The problem can be solved by computing $h_{N,N}(x, I_N)$ first, and then $h_{k,N}(x, I_k)$'s downwards recursively. The sequence of the maximizing a's at each stage determines the optimal policy $\{A_k^*(x_k, I_k)\}_{k=1}^N$.

III. Bayes Learning and its Efficacies

Let the optimal policy of the problem (2.4) be denoted by $\{A_t^*(\cdot)\}_{t=1}^N$ and let that of the problem (2.1) with G replaced by $\hat{G}_1$ be written as $\{A_t^0(\cdot)\}_{t=1}^N$. Then we are led to the following three types of definitions of the efficacy of the Bayes learning structure L^* under the complete information and the given 'true' distribution function $G(r)$ of the random parameter (Sakaguchi, 1963), as

$$e_1(L^* | P) \equiv \int \cdots \int [\sum_{t=1}^N R(A_t^*(x_t, I_t), x_t^* ; r_t) + \Phi(x_{N+1}^*)]$$

$$\prod_{t=1}^N d\hat{G}_t(r_t | I_t) - \int \cdots \int [\sum_{t=1}^N R(A_t^0(x_t^0), x_t^0 ; r_t)$$

$$+ \Phi(x^0_{N+1})] \prod_{t=1}^{N} d\hat{G}_1(r_t), \tag{3. 1}$$

$$e_2(L^* | P, G) \equiv \int \cdots \int [\left\{ \sum_{t=1}^{N} R(A^*_t(x_t, I_t), x^*_t, r_t) + \Phi(x^*_{N+1}) \right\}$$

$$- \left\{ \sum_{t=1}^{N} R(A^0_t(x^0_t), x^0_t : r_t) + \Phi(x^0_{N+1}) \right\}] \prod_{t=1}^{N} dG(r_t), \tag{3. 2}$$

and

$$e_3(L^* | P) \equiv e_2(L^* | P, G) \text{ with } G(r_t) \text{ replaced by } \hat{G}_t(r_t | I_t) \tag{3. 3}$$

where x^*_t and x^0_t are defined by

$$\left. \begin{aligned} x^*_{t+1} &= S(A^*_t(x^*_t, I_t), x^*_t ; r_t) \\ x^0_{t+1} &= S(A^0_t(x^0_t), x^0_t ; r_t) \end{aligned} \right\} \tag{3. 4}$$

for $t = 1, \ldots, N$; $x^*_1 = x_1$; I_1 = null information.

In the expression of definition (3. 1), the second term represents the maximum expected-overall-return when using the "null learning" structure, i. e.,

$$L^0 = \left\{ \hat{G}_1(r_1), \hat{G}_1(r_2), \ldots, \hat{G}_1(r_N) \right\} .$$

It immediately follows, by the definition and (2. 3) that $e_3(L^* | P) =$

$= \int e_2(L^* | P, G) d\xi \geqq 0.$ This represents the average increase of expected-over-all-return which the decision maker will get through using the Bayes learning L^* instead of null learning L^0. It will also be clear that the definition of $e_2(L^* | P, G)$ by (3. 2) has the similar meaning. It should be remarked that the Bayes learning is reasonable, if we could say so, only in the sense that $e_3(L^* | P) \geqq 0$. We may have, for Bayes learning L^*, $e_1(L^* | P) < 0$ and it is quite possible, as is shown in the later section, that $e_2(L^* | P, G) < 0$ for a class of G. Let us now compute the efficiencies e_1 and e_3 for some special models of adaptive processes. For somewhat trivial cases in which the state x remains constant we have the following

THEOREM 1. For the Bayes learning L^* and complete information pattern P, if $S(a, x ; r) \equiv x$ and $\Phi(x) \equiv 0$, then the optimal policy of the adaptive problem (2. 4) is given by

$$A_t^*(x_t, I_t) = \underset{(r_t | I_t)}{\text{the maximizing}} \; a_t \; \text{of} \int R(a_t, x_t ; r_t) d\hat{G}_t$$

Proof: It suffices to prove for N = 2. By (2. 5) we can get

$$h_{1,2}(x) = \max_{a_1} \int R(a_1, x ; r_1) d\hat{G}_1(r_1)$$

$$+ \int d\hat{G}_1(r_1) \cdot \max_{a_2} \int R(a_2, x ; r_2) d\hat{G}_2(r_2 | r_1)$$

Now the second term in the right hand side is not smaller than

$$\max_{a_2} \iint R(a_2, x ; r_2) d\hat{G}_1(r_1) d\hat{G}_2(r_2 | r_1) = \max_{a_2} \int R(a_2, x ; r_2) d\hat{G}_1(r_2) .$$

Hence we have

$$e_1(L^* | P) = h_{1,2}(x) - 2 \max_{a_1} \int R(a_1, x ; r_1) d\hat{G}_1(r_1) \geqq 0 .$$

Moreover, denoting the maximizing a_1 of the integral $\int R(a_1, x; r_1) d\hat{G}_1(r_1)$ by $a^0(x)$, we obtain

$$e_3(L^* | P) = h_{1,2}(x) - \iint \sum_{t=1}^{2} R(a^0(x), x ; r_t) d\hat{G}_1(r_1) d\hat{G}_2(r_2 | r_1)$$

$$= h_{1,2}(x) - 2 \int R(a^0(x), x ; r_1) d\hat{G}_1(r_1) = e_1(L^* | P) .$$

This completes the proof.

For other simple cases in which, for example,

$$R(a, x ; r) = 0 , \quad \Phi(x) = x$$

we cannot expect $e_1(L^* | P) \geqq 0$ unless we assume a special form of the function $S(a, x ; r)$. Let us consider an interesting model of adaptive control processes which was treated by Marschak (1963). Let

$$R(a, x ; r) \equiv 0, \quad S(a, x ; r) \equiv x - g(a-x-r), \quad \Phi(x) \equiv x , \tag{3.5}$$

so that the control problem be given by

$$\int \cdots \int x_{N+1} \prod_{t=1}^{N} d\hat{G}_t(r_t | I_t) \longrightarrow \max_{\{A_t(\cdot)\}_{t=1}^{N}} \qquad (3.6)$$

where

$$x_{t+1} = x_t - g(A_t(x_t, I_t) - x_t - r_t), \quad t = 1, \ldots, N \qquad (3.7)$$

We set the assumption: the function $g(y)$ is differentiable and satisfies some appropriate second order conditions. (3.A)

THEOREM 2. For the adaptive control problem given by (3.5) and (3.6), and for the Bayes learning L^* and complete information pattern P, the optimal policy is given by

$$A_t^*(x_t, I_t) - x_t = \text{the minimizing } a_t \text{ of } \int g(a_t - r_t) d\hat{G}_t(r_t | I_t)$$

$$(t = 1, \ldots, N \; ; I_1 = \text{null information})$$

and we have

$$e_1(L^* | P) = e_3(L^* | P) \geqq 0 \, .$$

Proof. It suffices to prove for $N=2$. Let us denote the minimizing a's of the integrals

$$\int g(a - r_1) d\hat{G}_1(r_1) \qquad \text{and} \qquad \int g(a - r_2) d\hat{G}_2(r_2 | r_1) \, ,$$

by a^0 and $a^*(r_1)$, respectively. Then by (2. 5) it is easy to show that

$$h_{1,2}(x_1) = x_1 - \int g(a^0 - r_1)d\hat{G}_1(r_1)$$

$$- \iint g(a^*(r_1) - r_2)d\hat{G}_1(r_1)d\hat{G}_2(r_2 | r_1) \qquad (3.8)$$

and that the optimal actions are given by

$$\left.\begin{aligned} A_1^*(x_1) &= x_1 + a^0 \\ A_2^*(x_2, I_2) &= x_2 + a^*(r_1) \end{aligned}\right\} \qquad (3.9)$$

Proceeding the analogous way we obtain, for the case of null learning L^0,

$$h_{1,2}^0(x_1) \equiv \iint x_3^0 \prod_{t=1}^{2} d\hat{G}_1(r_t)$$

$$= x_1 - 2\int g(a^0 - r_1)d\hat{G}_1(r_1) \, , \qquad (3.10)$$

and the corresponding optimal actions

$$A_t^0(x_t) = x_t + a^0 \, , \; t = 1, \, 2 \, . \qquad (3.11)$$

by a^0 and $a^*(r_1)$, respectively. Then by (2. 5) it is easy to show that

$$h_{1,2}(x_1) = x_1 - \int g(a^0 - r_1)d\hat{G}_1(r_1)$$

$$- \iint g(a^*(r_1) - r_2)d\hat{G}_1(r_1)d\hat{G}_2(r_2 | r_1) \qquad (3.8)$$

and that the optimal actions are given by

$$\left.\begin{aligned} A_1^*(x_1) &= x_1 + a^0 \\ A_2^*(x_2, I_2) &= x_2 + a^*(r_1) \end{aligned}\right\} \qquad (3.9)$$

Proceeding the analogous way we obtain, for the case of null learning L^0,

$$h_{1,2}^0(x_1) \equiv \iint x_3^0 \prod_{t=1}^{2} d\hat{G}_1(r_t)$$

$$= x_1 - 2\int g(a^0 - r_1)d\hat{G}_1(r_1) , \qquad (3.10)$$

and the corresponding optimal actions

$$A_t^0(x_t) = x_t + a^0 , \; t = 1, 2 . \qquad (3.11)$$

Thus we get

$$
\begin{aligned}
e_1(L^*|P) &= h_{1,2}(x_1) - h^0_{1,2}(x_1) = \int g(a^0 - r_1)d\hat{G}_1(r_1) \\
&- \iint g(a^*(r_1) - r_2)d\hat{G}_1(r_1)d\hat{G}_2(r_2|r_1) \\
&= \min_{a_1} \int g(a_1 - x_1 - r_1)d\hat{G}_1(r_1) \\
&- \int d\hat{G}_1(r_1)\left\{\min_{a_2} \int g(a_2 - x_2 - r_2)d\hat{G}_2(r_2|r_1)\right\} \geqq 0 .
\end{aligned}
\tag{3.12}
$$

Moreover since we get from (3. 7), (3. 9) and (3. 11)

$$
\left.
\begin{aligned}
x^*_3 &= x_1 - g(a^0 - r_1) - g(a^*(r_1) - r_2) \\
x^0_3 &= x_1 - g(a^0 - r_1) - g(a^0 - r_2)
\end{aligned}
\right\}
$$

it follows that

$$
\begin{aligned}
0 \leqq e_3(L^*|P) &= \iint (x^*_3 - x^0_3)d\hat{G}_1(r_1)d\hat{G}_2(r_2|r_1) \\
&= \iint \left\{g(a^0-r_2)-g(a^*(r_1)-r_2)\right\} d\hat{G}_1(r_1)d\hat{G}_2(r_2|r_1) \\
&= e_1(L^*|P) .
\end{aligned}
$$

This completes the proof of the theorem.

IV. Examples of the Efficacies of the Bayes Learning

To help understand the results in the previous section we shall give three worked examples. Examples 1 and 2 belong to the adaptive model treated in Theorem 2.

Example 1.

Assume that the control process is given by (3. 5) and let N=2. To give better insight, the adaptive model will be preceded by a corresponding stochastic model (dG(r) known).

(a) Stochastic model.

Let G(r) be the normal distribution function with known mean μ and known variance σ^2:

$$dG(r) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(r-\mu)^2}{2\sigma^2}} dr \qquad (4.\ 1)$$

Thus our control problem is described by (2. 1) together with (3. 5) and (4. 1).

Let us denote the root of the equation

$$0 = E_r\left\{g'(y-r)\right\} = \int_{-\infty}^{\infty} g'(y-r)\,\frac{1}{\sqrt{2\pi}}\, e^{-\frac{(r-\mu)^2}{2\sigma^2}} dr \qquad (4.\ 2)$$

by $gS'(\mu, \sigma^2)$, which is assumed to exist uniquely by our assumption (3. A).

By tracing exactly the proof of Theorem 2 we can easily find that the solution of the problem is given by

$$A_1^*(x_1) = x_1 + g^S(\mu, \sigma^{-2})$$

$$A_2^*(x_2) = x_2 + g^S(\mu, \sigma^{-2}) = x_1 - g(g^S(\mu, \sigma^{-2}) - r_1) + g^S(\mu, \sigma^{-2}) \qquad (4.3)$$

$$\max E(x_3) = x_1 - 2E\left\{g(g^S(\mu, \sigma^{-2}) - r_1)\right\}$$

Marschak (1963) called $g^S(\mu, \sigma^{-2})$ the "transform of g appropriate for the stochastic model". For example, if $g(y) \equiv \frac{y^2}{2}$, then $g^S(\mu, \sigma^{-2}) \equiv \mu$, $A_t^*(x_t) = x_t + \mu$ $(t = 1, 2)$ and $\max E(x_3) = x_1 - \sigma^{-2}$.

(b) <u>Adaptive model</u>

Assume that, in the previous model (a), we know the value of σ, but not the value of μ, and that only an <u>a priori</u> distribution ξ of μ is known before the first action a_1 is chosen. Let the priori distribution be ξ:

normal with mean m and variance v^2, where m and v are given constants.

Then, by Bayes' theorem, the a posteriori distribution $\xi(r_1)$ after observing the random variable r_1 is

$$\xi_1 = \xi(r_1) : \text{ normal with mean } \frac{mv^{-2} + r_1\sigma^{-2}}{v^{-2} + \sigma^{-2}} \text{ and variance } \frac{1}{v^{-2} + \sigma^{-2}}.$$

Hence by (2. 3)

$$\hat{G}_1(r_1) : \text{normal with mean } m \text{ and variance } \sigma^2 + v^2$$

$$G_2(r_2 | r_1) : \text{normal with mean } \frac{mv^{-2} + r_1\sigma^{-2}}{v^{-2} + \sigma^{-2}} \text{ and variance } \frac{\sigma^2/v^2 + 2}{v^{-2} + \sigma^{-2}} \tag{4.4}$$

Again, by (3. 8) and (3. 9) appearing in the proof of Theorem 2 we get the solution of the adaptive control problem described by

$$A_1^*(x_1) = x_1 + g^S(m, \sigma^2 + v^2),$$

$$A_2^*(x_2, I_2) = x_2 + g^S\left(\frac{mv^{-2} + r_1\sigma^{-2}}{v^{-2} + \sigma^{-2}}, \frac{\sigma^2/v^2 + 2}{v^{-2} + \sigma^{-2}}\right),$$

$$\max E_\mu E(x_3) = x_1 - E_\mu E[g\{g^S(m, \sigma^2 + v^2) - r_1\}]$$

$$- E_\mu E[g\{g^S\left(\frac{mv^{-2} + r_1\sigma^{-2}}{v^{-2} + \sigma^{-2}}, \frac{\sigma^2/v^2 + 2}{v^{-2} + \sigma^{-2}}\right) - r_2\}]. \tag{4.5}$$

where $E_\mu E[\cdots]$ stands for $\int [\cdots] d\hat{G}_1(r_1)$ or

$\iint [\cdots] d\hat{G}_1(r_1) d\hat{G}_2(r_2 | r_1)$.

We also obtain by Theorem 2 and (3. 12)

$$0 \leqq e_1(L^*|P) = e_3(L^*|P)$$

$$= E_\mu E\,[g\{g^S(m, \sigma^{-2}+v^2) - r_1\}]$$

$$- E_\mu E[g\{g^S\left(\frac{mv^{-2}+r_1\sigma^{-2}}{v^{-2}+\sigma^{-2}}, \frac{\sigma^{-2}/v^2+2}{v^{-2}+\sigma^{-2}}\right) - r_2\}],$$

and we can easily find that $e_2(L^*|P, G)$ is equal to the above expression with $E_\mu E$ replaced by E, i.e., $\iint [\cdots]\, G(r_1)dG(r_2)$.

If $g(y) = y^2/2$ then after some computations $e_1(L^*|P) = e_3(L^*|P$ is equal to

$$\frac{1}{2}(\sigma^{2} + v^2) - \frac{1}{2}\,\frac{\sigma^2/v^2+2}{v^{-2}+\sigma^{-2}} = \frac{v^2}{2(1+\sigma^2/v^2)} > 0, \qquad (4.6)$$

and

$$e_2(L^*|P,G) = \frac{1}{2}\iint\left\{(m-r_2)^2 - \left(\frac{mv^{-2}+r_1\sigma^{-2}}{v^{-2}+\sigma^{-2}} - r_2\right)^2\right\}\prod_{t=1}^{2} dG(r_t)$$

$$= \frac{1}{2(1+\sigma^2/v^2)^2}\left\{(\mu-m)^2(1+\frac{2\sigma^2}{v^2}) - \sigma^2\right\}, \qquad (4.7)$$

which is $\left\{\begin{matrix}>\\=\\<\end{matrix}\right\}$ 0, according as $(\mu-m)^2(\sigma^{-2}+2v^{-2})$ $\left\{\begin{matrix}>\\=\\<\end{matrix}\right\}$ 1. (By (4.6) and (4.7) we can check $e_3(L^*|P) = \int e_2(L^*|P, G)d\xi$). Thus, in this case, even with the "sufficiently good" learning structure, <u>i.e.,</u> a prior estimate very close to the true probability distribution (for example, $v = \mathcal{E}$ and $m = \mu + \mathcal{E}^2$ for sufficiently small $\mathcal{E} > 0$), and Bayesian inference, it is quite possible that $e_2(L^* | P, G) < 0$.

Example 2

Let us consider the same control problem as before for the case of a binomial distribution. Assume that r_1 and r_2 are identically and independently distributed, each with a binomial distribution with parameters 1 and unknown α. (Marschak, 1963).

Let the prior distribution of α be given by

$$\xi : \Pr\{\alpha = \alpha_i\} = p_i$$

$$(i = 1, \ldots, k ; p_1, \ldots, p_k > 0; \sum_1^k p_i = 1),$$

and so the posterior distribution after observing r_1, is

$$\xi_1 = \xi(r_1) ; \begin{cases} \Pr\{\alpha = \alpha_i \mid r_1 = 1\} = p_i\alpha_i / \sum_i p_i\alpha_i \\ \Pr\{\alpha = \alpha_i \mid r_1 = 0\} = p_i(1-\alpha_i) / \sum p_i(1-\alpha_i), \end{cases}$$

and thus by (2. 3)

$$\hat{G}_1(r_1) : \text{binomial with parameters } 1, \hat{\alpha}(\xi)$$

$$\hat{G}_1(r_2 | r_1) : \text{binomial with parameters } 1, \hat{\alpha}(\xi_1)$$

where $\hat{\alpha}(\xi_t)$ $(t = 0, 1 ; \xi_0 = \xi)$ is the mean of the distribution ξ_t, so that $\hat{\alpha}(\xi) = \sum_{i=1}^k p_i\alpha_i$ and

$$\hat{\alpha}(\xi_1) = \hat{\alpha}(\xi(r_1)) = \sum_{i=1}^{k} \alpha_i \Pr\{\alpha = \alpha_i \mid r_1\}$$

$$= \begin{cases} \sum p_i\alpha_i^2 / \sum p_i\alpha_i \,, & \text{if } r_1 = 1, \\ \sum p_i\alpha_i(1-\alpha_i) / \sum p_i(1-\alpha_i), & \text{if } r_1 = 0. \end{cases}$$

Let us denote the root of the equation

$$0 = E_r\{g'(y-r)\} = \alpha g'(y-1) + (1-\alpha)g'(y) = 0$$

by again $g^S(\alpha)$, which is assumed to exist uniquely by the assumption (3. A).

Then, from (3. 8), (3. 9) and (4. 8), the solution of this adaptive control problem is given by

$$\left.\begin{aligned} A_1^*(x_1) &= x_1 + g^S(\hat{\alpha}(\xi)) \\ A_2^*(x_2, I_2) &= x_2 + g^S(\hat{\alpha}(\xi_1)) \\ \max E_\alpha E(X_3) &= X_1 - E_\alpha E[g\{g^S(\hat{\alpha}(\xi)) - r_1\}] \\ &\quad - E_\alpha E[g\{g^S(\hat{\alpha}(\xi_1)) - r_2\}] . \end{aligned}\right\} \qquad (4.9)$$

We also obtain by Theorem 2 and (3. 12) that

$$0 \leqq e_1(L^*|P) = e_3(L^*|P)$$

$$= E_\alpha E[g\{g^S(\hat{\alpha}(\xi) - r_1\}] - E_\alpha E[g\{g^S(\hat{\alpha}(\xi_1) - r_2\}],$$

and that $e_2(L^*|P, G)$ is equal to the above expression with $E_\alpha E$ replaced simply by E, i.e., $\iint [\cdots]\, dG(r_1)dG(r_2)$. Let us put $\hat{\alpha} \equiv \sum_{i=1}^{k} p_i\alpha_i$ and $\tilde{\alpha} \equiv \sum_{i=1}^{k} p_i\alpha_i^2$. Then if the problem is given by the special successor stated function with $g(y) = y^2/2$, then $g^S(\alpha) = \alpha$ and we can find, after some computations,

$$e_1(L^*|P) = e_3(L^*|P) = \frac{\hat{\alpha}(1-\hat{\alpha})}{1} - \frac{(\hat{\alpha}-\tilde{\alpha})(\hat{\alpha}+\tilde{\alpha}-2\hat{\alpha}^2)}{2\hat{\alpha}(1-\hat{\alpha})}$$

$$= \frac{(\tilde{\alpha}-\hat{\alpha}^2)^2}{2\hat{\alpha}(1-\hat{\alpha})} = \frac{\left\{\sum_1^k (\alpha_i-\hat{\alpha})^2 p_i\right\}^2}{2\hat{\alpha}(1-\hat{\alpha})} > 0,$$

and

$$e_2(L^*|P,G) = \frac{1}{2}\iint\left\{(\hat{\alpha}-r_1)^2 - (\hat{\alpha}(\xi(r_1)) - r_2)^2\right\}\prod_1^2 dG(r_t)$$

$$= \left(\frac{\tilde{\alpha}}{\hat{\alpha}} - \frac{\hat{\alpha}-\tilde{\alpha}}{1-\hat{\alpha}}\right)\alpha^2 - \left[\hat{\alpha} - \frac{\hat{\alpha}-\tilde{\alpha}}{1-\hat{\alpha}} + \frac{1}{2}\left\{\left(\frac{\tilde{\alpha}}{\hat{\alpha}}\right)^2 - \left(\frac{\hat{\alpha}-\tilde{\alpha}}{1-\hat{\alpha}}\right)^2\right\}\right]\alpha$$

$$+ \frac{1}{2}\cdot\left\{\hat{\alpha}^2 - \left(\frac{\hat{\alpha}-\tilde{\alpha}}{1-\hat{\alpha}}\right)^2\right. .$$

Let the quadratic function above be denoted by $f(\alpha)$. Then since we have

$$0 < \hat{\alpha}(\xi(r_1 = 0)) = \frac{\hat{\alpha} - \tilde{\alpha}}{1 - \tilde{\alpha}} < \hat{\alpha} < \frac{\alpha}{\hat{\alpha}} = \hat{\alpha}(\xi(r_1 = 1)) < 1 \qquad (4.10)$$

by definitions of $\hat{\alpha}$ and $\tilde{\alpha}$, $f(\alpha)$ is strictly convex with $f(0)$ and $f(1) > 0$. After some infinitesimal calculations we can find that $f(\alpha)$ can have negative values in an interval of α values, if for example, $k = 2$, $\alpha_1 = \varepsilon$, and $\alpha_2 = 1 - \varepsilon$, for sufficiently small $\varepsilon > 0$.

Example 3.

Let us consider the control problem in which the random variable has a binomial distribution with parameters $1, \alpha$ $(0 < \alpha < 1)$ and

$$R(a, x; r) \equiv 0 \; ; \; S(a, x; r) = \begin{cases} x+a, & \text{if } r = 1 \; ; \\ x-a, & \text{if } r = 0 \end{cases} \; ; \; \Phi(x) = \log x$$

The action space when given x_k, I_k is the interval $0 \leqq a \leqq x_k$. This model corresponds to the following betting problem (Bellman 1961, Sakaguchi 1963). Consider a coin with probability α of heads. A gambler, without knowing this probability is required to place a bet on the event of head. He is allowed to bet a quantity a, subject to the restriction $0 \leqq a \leqq x$, where x is his capital at the present stage. If he bets correctly he then wins, otherwise he loses. Continuing this betting process for N stages, and assuming that tossings of the coin are independent at these stages and that a gambler wishes to maximize the expected value of the logarithm of the final total at the end of the process, the problem is then to derive an optimal betting policy.

Let the prior distribution of α be the same as in Example 2. With the functions $h_{k,N}(x, I_k)$ defined in Section 2. We have, for $N = 2$,

$$h_{2,2}(x_2, I_2) = \max_{0 \leqq a_2 \leqq x_2} \int \log S(a_2, x_2 ; r_2) d\hat{G}_2(r_2 | r_1)$$

$$= \max_{0 \leqq a_2 \leqq x_2} \hat{\alpha}(\xi(r_1)) \log(x_2 + a) + (1 - \hat{\alpha}(\xi(r_1)) \log(x_2 - a_2)$$

$$= \log x_2 + C(\max(\hat{\alpha}(\xi(r_1)), \tfrac{1}{2}))$$

in which the optimal choice is given by

$$A_2^*(x_2, I_2) = \max\{(2\hat{\alpha}(\xi(r_1)) - 1)x_2, 0\} \tag{4.11}$$

and $C(\alpha)$ is defined by

$$C(\alpha) \equiv \alpha \log \frac{\alpha}{\frac{1}{2}} + (1-\alpha) \log \frac{1-\alpha}{\frac{1}{2}}$$

which is $\geqq 0$ for all $0 \leqq \alpha \leqq 1$ and equals 0 if and only if $\alpha = \frac{1}{2}$.

Now

$$h_{1,2}(x_1) = \max_{0 \leqq a_1 \leqq x_1} \int h_{2,2}(S(a_1, x_1, ; r_1), I_2) d\hat{G}_1(r_1$$

$$= \max_{0 \leqq a_1 \leqq x_1} \int \log S(a_1, x_1 ; r_1) d\hat{G}_1(r_1)$$

$$+ \int C(\max(\hat{\alpha}(\xi(r_1), \tfrac{1}{2})) d\hat{G}_1(r_1)$$

where the first term in the right hand side is

$$\max_{0 \leqq a_1 \leqq x_1} [\hat{\alpha} \log(x_1 + a_1) + (1-\hat{\alpha}) \log(x_1 - a_1)]$$

$$= \log x_1 + C(\max(\hat{\alpha}, \tfrac{1}{2})$$

and the optimal action is given by

$$A_1^*(x_1) = \max \left\{(2\hat{\alpha} - 1)x_1, 0\right\}.$$

Considering the inequalities (4. 10) we finally get

$$h_{1,2}(x_1) = \log x_1 + C(\max(\hat{\alpha}, \tfrac{1}{2}) + \hat{\alpha} C(\max(\frac{\tilde{\alpha}}{\hat{\alpha}}, \tfrac{1}{2}))$$

$$+ (1-\hat{\alpha}) C(\max(\frac{\hat{\alpha} - \tilde{\alpha}}{1-\hat{\alpha}}, \tfrac{1}{2})).$$

In an analogous way we can find for the case of null learning L^0 that

$$\begin{cases} h^0_{1,2}(x_1) = \log x_1 + 2C(\max(\hat{\alpha}, \frac{1}{2}) \\ \\ A^0_t(x_t) = \max\left\{(2\hat{\alpha}-1)x_t, 0\right\}, \qquad t = 1, 2. \end{cases}$$

And thus

$$e_1(L^*|P) = \hat{\alpha}C(\max(\frac{\tilde{\alpha}}{\hat{\alpha}}, \frac{1}{2})) + (1-\hat{\alpha})C(\max(\frac{\hat{\alpha}-\tilde{\alpha}}{1-\hat{\alpha}}, \frac{1}{2})) - C(\max \hat{\alpha}, \frac{1}{2})) \tag{4.12}$$

which is $\geqq 0$, since the function $C(\max(\alpha, \frac{1}{2})$ is 0 for $0 \leqq \alpha \leqq \frac{1}{2}$ and strictly convex for $\frac{1}{2} < \alpha < 1$.

For another efficacy $e_2(L^*|P, G)$ of Bayes learning L^* we can find that with $x_1 = x$

$$\iint \log S(a_2^*, x_2^*; r_2) \prod_{t=1}^{2} dG(r_t)$$

$$= \alpha\left\{\log(x+a_1^*) + \alpha \log(1+\frac{a_2^*(1)}{x+a_1^*}) + (1-\alpha)\log(1 - \frac{a_2^*(1)}{x+a_1^*})\right\}$$

$$+ (1-\alpha)\left\{\log(x - a_1^*) + \alpha \log(1 + \frac{a_2^*(0)}{x-a_1^*})\right.$$

$$\left. + (1-\alpha)\log(1 - \frac{a_2^*(0)}{x-a_1^*})\right\}.$$

$$\iint \log S(a_2^0\,,\ x_2^0\ ;\ r_2) \prod_{t=1}^{2} dG(r_t)$$

$$= \log(x+a_1^0) + \alpha\left\{\log\left(1+\frac{a_2^0(1)}{x+a_1^0}\right) + (1-\alpha)\log\left(1-\frac{a_2^0(1)}{x+a_1^0}\right)\right\}$$

$$+ (1-\alpha)\left\{\log(x-a_1^0) + \alpha\log\left(1+\frac{a_2^0(0)}{x+a_1^0}\right)\right.$$

$$\left. + (1-\alpha)\log\left(1-\frac{a_2^0(0)}{x-a_1^0}\right)\right\},$$

and hence

$$e_2(L^*|P,G) = \iint \log S(a_2^*,\ x_2^*\ ;\ r_2)$$

$$- \log S(a_2^0\,,\ x_2^0\ ;\ r_2) \prod_{t=1}^{2} dG(r_t)$$

$$= \alpha\left\{\log\left(1+\frac{a_2^*(1)}{x+a_1^*}\right) + (1-\alpha)\log\left(1-\frac{a_2^*(1)}{x+a_1^*}\right) - \log\left(1+\frac{a_1^0}{x}\right)\right\}$$

$$+ (1-\alpha)\left\{\alpha\log\left(1+\frac{a_2^*(0)}{x-a_1^*}\right) + (1-\alpha)\log\left(1-\frac{a_2^*(0)}{x-a_1^*}\right)\right.$$

$$\left. - \log\left(1-\frac{a_1^0}{x}\right)\right\}.$$

where

$$\frac{a_1^*}{x} = \frac{a_1^0}{x} = \frac{a_2^0(1)}{x+a_1^0} = \frac{a_2^0(0)}{x-a_1^0} = \max\left\{2\hat{\alpha}-1,\ 0\right\},$$

$$\frac{a_2^*(1)}{x+a_1^*} = \max\left\{2\hat{\alpha}(\xi(r_1=1)) - 1,\ 0\right\},$$

$$\frac{a_2^*(0)}{x-a_1^*} = \max\left\{2\hat{\alpha}(\xi(r_1=0)) - 1,\ 0\right\}.$$

Consider the case in which the prior distribution of α satisfies $\frac{\hat{\alpha}-\tilde{\alpha}}{1-\hat{\alpha}} > \frac{1}{2}$ (e.g., $k = 2$, $\alpha_1 = \frac{3}{4} - \varepsilon$ and $\alpha_2 = \frac{3}{4} + \varepsilon$ for sufficiently small $\varepsilon > 0$). Then we have

$$e_2(L^*|P,G) = \alpha\left\{\alpha \log \frac{\tilde{\alpha}/\hat{\alpha}}{\hat{\alpha}} + (1-\alpha)\log\frac{1-\tilde{\alpha}/\hat{\alpha}}{1-\hat{\alpha}}\right.$$

$$\left. + (1-\alpha)\left\{\alpha \log\frac{(\hat{\alpha}-\tilde{\alpha})/(1-\hat{\alpha})}{\hat{\alpha}} + (1-\alpha)\log\frac{1-(\hat{\alpha}-\tilde{\alpha})/(1-\hat{\alpha})}{1-\hat{\alpha}}\right\}. \quad (4.13)$$

Let the quadratic function in the right hand side of the above expression be denoted by $f(\alpha)$. It is easily seen that $f(\alpha)$ is convex with $f(0)$ and $f(1) > 0$ and assumes negative values in some interval about $\alpha = \hat{\alpha}$.

It is interesting to find, after taking expectation of (4. 13) and comparing the result with (4. 12), that

$$e_1(L^*/P) = e_3(L^*|P) = \hat{\alpha}\,I\left(\frac{\tilde{\alpha}}{\hat{\alpha}} : \hat{\alpha}\right) + (1-\hat{\alpha})I\left(\frac{\hat{\alpha}-\tilde{\alpha}}{1-\hat{\alpha}} : \hat{\alpha}\right) \geqq 0$$

where

$$I(\omega : \omega') \equiv \omega \log \frac{\omega}{\omega'} + (1-\omega) \log \frac{1-\omega}{1-\omega'} , \quad (0 \leqq \omega, \omega' \leqq 1)$$

is the Kullback-Leibler information (Kullback, 1959) descriminating between two binomial distributions with parameters $(1, \omega)$ and $(1, \omega')$.

V. Concluding Remark

Summarizing the foregoing discussions it would not be useless to note the following remarks.

(a) The optimal policy of the adaptive control problem (2.4), generally, does not coincide with the one which would be obtained by, at first, finding the best estimators $\hat{\theta}_t$ of the unknown parameter θ at each stage t and then solving the stochastic control problem (2.1) with $\prod_{t=1}^{N} dG(r_t)$ replaced $\prod_{t=1}^{N} P(r_t, \hat{\theta}_t)dr_t$.

(b) From (2.2) and (2.5) we see that the optimal action in the Bayesian adaptive programming, is itself a functional of the posterior probability distribution given past observations. But as we have seen in the three examples ((4.5), (4.9) and (4.11)), the optimal action is characterized by one or two parameters if the prior distribution of θ is chosen from the family which is closed under sampling (i.e., ξ and $\xi(r_1)$ belong to the same family). This simplifies the computations very much.

(c) The Bayes learning L^* is reasonable only in the sense that $e_3(L^*|P) \geqq 0$ for all prior distributions. And at the same time it is unreasonable in the sense that $e_2(L^*|P, G)$ can be negative for the prior distributions very close to the true distribution if N is not large.

References

(1) Bellman, R. (1961) "Adaptive Control Processes - A Guided Tour". Princeton University Press, New Jersey.

(2) Bellman, R. and Kalaba, R. (1959), A mathematical theory of adaptive control processes. Proc. Nat. Acad. Sci., U.S.A., 45, 1288-1290.

(3) Kullback, S. (1959), Information Theory and Statistics, John Wiley and Sons, New York.

(4) Marschak, J., (1963), On adaptive programming, Manag. Sci., 9, 517-526.

(5) Sakaguchi, M. (1963), Information pattern, learning structure and optimal decision rules, Information and Control, 6, 218-220.