Bayes Procedures for Truncated Sequential Game Identifying an Alternative of Different Sampling Plans

By

Hisashi MINE*, Harunori SHISHIDO* and Munehiro OHMOTO**

(Received December 28, 1974)

Abstract

Recently, problems of decision making process have been discussed in many fields, and in particular, several mathematical approaches have been proposed for economic problems.

Decision making for management contains many complex situations and has theoretically unsolved problems. However, there are effective approaches for some problems.

One approach is a sequential decision analysis by a decision tree. This approach is often used because the analysis is applicable even if the decision stages are mutually dependent on each other or are subject to certain constraints.

On the other hand, it is known that a sequential analysis by the statistical decision theory can be used for problems of decision making even under uncertain information.

This paper investigates a truncated sequential game, an N-truncated sequential analysis.

1. Introduction

One kind of specialization of two-person games leads to a class of games known as statistical games. In these statistical games, the two players will be referred to as nature and the statistician. In statistical games, nature cannot be considered as a conscious opponent who can take advantage of mistakes made by the statistician. The statistician has at his disposal a class A of possible actions which he can take in the face of the unknown state of nature ω. If he decides to take an action without experimentation, we assume that he incurs a numerical loss $L(\omega, a)$, a known function of the state ω and the action a which he selects from A.

* Department of Applied Mathematics and Physics.
** Department of Applied Mathematics and Physics.
Present address: the Ministry of Labor, Tokyo, Japan.

The possibility of "spying" on the opponent by performing experiments is a distinguishing characteristic of all statistical games. The possibility of performing experiments and thus reducing the loss by gaining at least partial information about ω is open to the statistician. He must decide which experiments he is to perform, in what sequence he is to perform them, when he is to terminate experimentation, and what action he is to take once experimentation is terminated. What prevents the statistician from getting a full knowledge of ω by unlimited experimentation is the cost of the experiments.

At each stage he can choose one of two sampling methods as a subexperiment. A different sampling method has a different sampling cost. The total number of possible subexperiments does not exceed a certain preassigned integer N. The N-truncated sequential game described above can also be considered as an N-truncated sequential sampling plan.

2. Formulation of *N*-Truncated Sequential Sampling Plan

An N-truncated sequential sampling plan is a procedure by which a decision is made as to whether a sampling is to be continued or not; and a terminal decision is made when the sampling is truncated. Let us define the state space $\Omega=(\omega_1, \omega_2, \cdots, \omega_m)$ and the decision space $A=(a_1, a_2, \cdots, a_n)$, where ω_i is the index of state i and a_j is a terminal action.

Next, the following loss function or opportunity loss is introduced;

$$L(\omega_i, a_j)=\max_k w_{ik}-w_{ij}, \qquad i=1, 2, \cdots, m \quad j=1, 2, \cdots, n \tag{1}$$

where w_{ij} is the payoff if nature is in state i and the statistician's decision is j. An optimal solution is a solution which minimizes the total expectation of losses. Let Ξ be the space of a priori probability distributions on Ω. Then Ξ can be respresented as an $(m-1)$ dimensional simplex with point,

$$\xi=(\xi_1, \xi_2, \cdots, \xi_m), \qquad \xi_i \geq 0 \text{ for all } i \text{ and } \sum_{i=1}^{m}\xi_i=1. \tag{2}$$

It is assumed that two sampling methods can be used. One sampling method is to observe $\{Y\}$, which has a probability distribution $f_\omega(Y)$, with a cost C_1. The other is to observe $\{Z\}$, which has a probability distribution $f_\omega(Z)$, with a cost C_2. This paper treats only discrete probability distributions.

Assumptions

1. Sampling cost is C_1 or C_2 for each of the observations of $\{Y\}$ or $\{Z\}$, respectively. Only one of two observations is permitted at each stage.

2. Sampling is truncated within N stages.

3. When sampling is truncated before the N-th stage, a terminal action should be made.

4. Probability distribution at each stage of the same sampling is mutually independently and identically distributed

Let us define the observation vector $x(N)=(x_1, x_2, \cdots, x_N)$, where x_i represents the i-th observation, that is

$$x_i=\begin{cases} y_i, & y: \text{observation of } \{Y\} \\ z_i. & z: \text{observation of } \{Z\} \end{cases}$$

Now $x(N)$ is considered to be a point in N-product space $\Omega_x(N)=\Omega_x\times\cdots\times\Omega_x$ where $\Omega_x=\Omega_y$ or $\Omega_x=\Omega_z$, and Ω_y consists of all observation values of $\{Y\}$ and Ω_z consists of those of $\{Z\}$.

Further, $x(j)$ represents the observation vector whose first j elements assume observation values, and the remaining $(N-j)$ elements are arbitrary.

Terminal action $a\in A$ is written as terminal decision function d of j and x;

$$a=d(j, x), \tag{3}$$

where

$$d(j, x)=d(j, x') \quad \text{if and only if } x, x'\in\Omega_x \text{ and } x_i=x_i' \text{ for } i=1, 2, \cdots, j.$$

$d(j, x)$ is an action, using the information $x_1, x_2, \cdots, x_j$. Let D be a set of terminal decision functions d.

Definition Let Σ be a class of partitions for an N-product space Ω_x. If $S_N\in\Sigma$, then $S_N=(S_{0N}, S_{1N}, \cdots, S_{NN})$. When $x, x'\in\Omega_x$ and $x_i=x_i'$ for $i=1, 2, \cdots, j$, then $x\in S_{jN}$ if and only if $x'\in S_{jN}$. That is, each S_{jN} is a cylinder set.

An element (S_N, d) of $\Sigma\times D$ determines a sequential sampling plan. The sets S_{jN} of S_N are sometimes referred to as "stopping regions".

Let us introduce other elements of decision process:

$C_j(x)$ = sum of costs for sampling $x_1, x_2, \cdots, x_j$,

$P_\omega(x)$ = probability of x when nature is in state ω,

$\rho(\omega, S_N, d)$ = risk when nature is in state ω and we use (S_N, d),

$$\rho(\omega, S_N, d)=\sum_{j=0}^{N}\sum_{x\in S_{jN}}[C_j(x)+L(\omega, d(j, x))]P_\omega(x). \tag{4}$$

If a priori probability distribution on Ω is given by ξ, then

$$\rho(\xi, S_N, d)=\sum_{\omega}\rho(\omega, S_N, d)\xi(\omega)$$

$$=\sum_{\omega}\sum_{j=0}^{N}\sum_{x\in S_{jN}}[C_j(x)+L(\omega, d(j, x))]P_\omega(x)\xi(\omega). \tag{5}$$

$E_{j\xi}(h)$ = conditional expectation of bounded function h on $\Omega\times\Omega_x$, when x_1, $x_2, \cdots, x_j$ are given.

For any x, the value of $E_{j\xi}(h)$ at x is

$$E_{j\xi}(h)=\frac{\sum_{y\in F_j(x)}\sum_{\omega}\xi(\omega)P_\omega(y)h(\omega, y)}{\sum_{y\in F_j(x)}P_\xi(y)}, \qquad P_\xi(x)=\sum_{\omega}P_\omega(x)\xi(\omega), \tag{6}$$

if ω has distribution ξ and x has distribution P_ω for fixed ω. $F_j(x)$ is a set of points, the first j elements of which have the same elements as those of x. Therefore,

$$S_{jN}=\bigcup_{x\in S_{jN}}F_j(x), \tag{7}$$

and

$$P_\xi(x)=\sum_{y\in F_j(x)}P_\xi(y), \qquad \text{for } x\in S_{jN}. \tag{8}$$

By using these equalities,

$$\rho(\xi, S_N, d)=\sum_{j=0}^{N}\sum_{x\in S_{jN}}[C_j(x)+E_{j\xi}[L(\omega, d(j, x))]]P_\xi(\omega). \tag{9}$$

Some definitions are introduced in the following:

$$\tau_j(x, a)=E_{j\xi}[L(\omega, a)]=\sum_{\omega}\xi_j(\omega)L(\omega, a), \tag{10}$$

$$\tau_j^*(x)=\min_{a\in A}\tau_j(x, a)=\psi(\xi_j), \tag{11}$$

where

$$\xi_j=(\xi_j(\omega_1), \xi_j(\omega_2), \cdots, \xi_j(\omega_m)), \qquad \omega_i\in\Omega, \tag{12}$$

$$\rho(\xi, S_N)=\min_{d}\rho(\xi, S_N, d)=\sum_{j=0}^{N}\sum_{x\in S_{jN}}[C_j(x)+\tau_j^*(x)]P_\xi(x). \tag{13}$$

Now, an optimal S_N^* of N-TSSP for a given a priori probability distribution ξ is constructed.

At first, a posteriori probability distribution on Ω, under the condition that the observation of $x(j)$ is made, is derived by Bayes' Theorem,

$$\xi_j(\omega)=\xi(\omega|x(j))=\frac{\xi(\omega)P_\omega(x(j))}{P_\xi(x(j))}. \tag{14}$$

It is noted that

$$q_\omega(x_i)=\begin{cases} f_\omega(y_i), & \text{if } \Omega_x{}^i=\Omega_y{}^i, \\ g_\omega(z_i), & \text{if } \Omega_x{}^i=\Omega_z{}^i \quad \text{for each } i. \end{cases} \tag{15}$$

Then, it is shown that

$$P_\omega(x(j))=\sum_{x_{j+1},\cdots,x_N}\prod_{i=1}^{N}q_\omega(x_i)=\prod_{i=1}^{j}q_\omega(x_i) \tag{16}$$

$$P_\xi(x(j))=\sum_\omega \xi(\omega)P_\omega(x(j))=\sum_\omega \xi(\omega)\prod_{i=1}^{j}q_\omega(x_i). \tag{17}$$

Among the j elements of $x(j)$, m elements are given by the observation of $\{Y\}$ and $(j-m)$ are by $\{Z\}$, Therefore,

if $x_j=y_j$,

$$\xi_j(\omega)=\xi_{j-m}{}^m(\omega)=\frac{\xi(\omega)\prod_{i=1}^{m}f_\omega(y_i)\prod_{k=1}^{j-m}g_\omega(z_k)}{\sum_\omega \xi(\omega)\prod_{i=1}^{m}f_\omega(y_i)\prod_{k=1}^{j-m}g_\omega(z_k)}=\frac{f_\omega(y_j)\xi_{j-m}{}^{m-1}(\omega)}{\sum_\omega f_\omega(y_j)\xi_{j-m}{}^{m-1}(\omega)}$$

$$=\frac{q_\omega(x_j)}{\sum_\omega q_\omega(x_j)\xi_{j-m}{}^{m-1}(\omega)}\xi_{j-m}{}^{m-1}(\omega)=T_{x_j=y_j}\xi_{j-1}(\omega), \tag{18}$$

and if $x_j=z_j$,

$$\xi_j(\omega)=\xi_{j-m}{}^m(\omega)=\frac{g_\omega(z_j)\xi_{j-m}{}^m(\omega)}{\sum_\omega g_\omega(z_j)\xi_{j-m}{}^m(\omega)}=\frac{q_\omega(x_j)}{\sum_\omega q_\omega(x_j)\xi_{j-m-1}{}^m(\omega)}\xi_{j-m}{}^{m-1}(\omega)$$

$$=T_{x_j=z_j}\xi_{j-1}(\omega). \tag{19}$$

In any case, ξ_j can be expressed by $T_{x_j}\xi_{j-1}$, i.e., by ξ_{j-1} and x_j. The a posteriori probability distribution ξ_j becomes the a priori distribution for the next observation.

Remarks Let h be a real-valued and bounded function on Ξ; and for any j and fixed $\xi_j\in\Xi$ let

$$E_j[h(T_{x_{j+1}}\xi_j)]=E_{j\xi}[h(T_{x_{j+1}}\xi_j)], \tag{20}$$

and

$$E_x[h(T_x\xi)]\equiv\sum_x\sum_\omega \xi(\omega)h(T_x\xi)q_\omega(x). \tag{21}$$

Now, the conditional expected value of ξ_{j+1}, given $x_1, x_2, \cdots, x_j$ is written as

$$E_j[h(T_{x_{j+1}}\xi_j)]=\sum_{x_{j+1}}\sum_\omega \xi_j(\omega)h(T_{x_{j+1}}\xi_j)q_\omega(x_{j+1}). \tag{22}$$

This expression can be derived from the definition of symbol E_j and the fact that x_i's are independent of each other. Since the x_i's are also identically distributed, the above equation shows that, if $\xi_j=\xi$ and $x_{j+1}=x$, then

$$E_j[h(T_{x_{j+1}}\xi_j)]=\underset{x}{E}_j[h(T_x\xi)]. \tag{23}$$

The following notation is used below:

$$E_j[h(T_x\xi_j)]=\underset{x}{E}_j[h(T_x\xi_j)]. \tag{24}$$

3. Theorem and its Proof

Theorem There exists an optimal solution $S_N^*=(S_{0N}^*, S_{1N}^*, \cdots, S_{NN}^*)$ for N-TSSP, given a priori probability distribution ξ.

And the Bayes risk is given as follows:

$$\rho_N^*(\xi)=\min[\psi(\xi), C_1+\underset{y}{E}[\rho_{N-1}^*(T_y\xi)], C_2+\underset{z}{E}[\rho_{N-1}^*(T_z\xi)]], \tag{25}$$

where

$$S_{jN}^*=\{x|\xi_r\notin\Xi_{N-j} \text{ for } r<j,\ \xi_j\in\Xi_{N-j}\}. \tag{26}$$

S_N^* is a partition of N-product space Ω_x.

Proof Proof consists of two parts.

[I] Construction of S_N^*.

Functions $h_0, h_1, \cdots, h_N$ on Ξ are defined as follows:

$$h_0(\xi)\equiv\psi(\xi)=\min_{a\in A}\sum_{\omega}\xi(\omega)L(\omega, a), \tag{27}$$

and by induction on j, $j=1, 2, \cdots, N$,

$$h_j(\xi)\equiv\min[\psi(\xi), C_1+\underset{y}{E}[h_{j-1}(T_y\xi)], C_2+\underset{z}{E}[h_{j-1}(T_z\xi)]]. \tag{28}$$

Let

$$U_j(\xi_j)\equiv C_1m_j+C_2(j-m_j)+\psi(\xi_j), \tag{29}$$

where

m_j= the number of observations of $\{Y\}$,

$j-m_j$= the number of observations of $\{Z\}$.

And by induction backward

$$\alpha_{NN}\equiv U_N, \qquad \alpha_{jN}\equiv\min[U_j, \underset{y}{E}_j(\alpha_{j+1}{}^y{}_{,N}), \underset{z}{E}_j(\alpha_{j+1}{}^z{}_{,N})]. \tag{30}$$

Finally, the following notations are introduced

$$\alpha_{jN}\equiv\alpha_{jN}(\xi_j)=\begin{cases}\alpha_{jN}(T_y\xi_{j-1})=\alpha_{jN}{}^y \\ \alpha_{jN}(T_z\xi_{j-1})=\alpha_{jN}{}^z \qquad \text{for } j\leq N.\end{cases} \tag{31}$$

It is noted that

$$\alpha_{NN}=\alpha_{NN}(\xi_N)=U_N(\xi_N)=C_1m_N+C_2(N-m_N)+\psi(\xi_N)$$

$$=\begin{cases} C_1(m_{N-1}+1)+C_2(N-m_{N-1}-1)+h_0(T_y\xi_{N-1})=\alpha_{NN}^y(\xi_{N-1}) \\ C_1m_{N-1}+C_2(N-m_{N-1})+h_0(T_z\xi_{N-1})=\alpha_{NN}^z(\xi_{N-1}), \end{cases} \tag{32}$$

$$\alpha_{N-1,N}=\alpha_{N-1,N}(\xi_{N-1})=\min[U_{N-1}, \underset{y}{E}_{N-1}(\xi_{NN}^y), \underset{z}{E}_{N-1}(\xi_{NN}^z)]$$
$$=C_1m_{N-1}+C_2(N-1-m_{N-1})+h_1(\xi_{N-1})$$
$$=\begin{cases} C_1(m_{N-2}+1)+C_2(N-1-m_{N-2}-1)+h_1(T_y\xi_{N-2})=\alpha_{N-1}^y{}_{,N} \\ C_1m_{N-2}+C_2(N-1-m_{N-2})+h_1(T_z\xi_{N-2})=\alpha_{N-1}^z{}_{,N}, \end{cases} \tag{33}$$

$$\alpha_{N-j,N}=\alpha_{N-j,N}(\xi_{N-j})=\min[U_{N-j}, \underset{y}{E}_{N-j}(\alpha_{N-j+1}^y{}_{,N}), \underset{z}{E}_{N-j}(\alpha_{N-j+1}^z{}_{,N})]$$
$$=C_1m_{N-j}+C_2(N-j-m_{N-j})+h_j(\xi_{N-j})$$
$$=\begin{cases} C_1(m_{N-j-1}+1)+C_2(N-j-m_{N-j-1}-1)+h_j(T_y\xi_{N-j-1})=\alpha_{N-j}^y{}_{,N} \\ C_1m_{N-j-1}+C_2(N-j-m_{N-j-1})+h_j(T_z\xi_{N-j-1})=\alpha_{N-j}^z{}_{,N} \end{cases} \tag{34}$$

$$\alpha_{0N}=C_1m_0+C_2(-m_0)+h_N(\xi_0)=h_N(\xi_0). \tag{35}$$

i) Let us define

$$S_{jN}^*=\{x|\alpha_{rN}<U_r \quad \text{for} \quad r<j,\ \alpha_{jN}=U_j\}. \tag{36}$$

Then by (28) and $h_j(\xi)\leq\psi(\xi)$,

$$\begin{aligned} S_{jN}^* &= \{x|\alpha_{N-(N-r),N}<U_r \quad \text{for} \quad r<j,\ \alpha_{jN}=U_j\} \\ &= \{x|h_{N-r}(\xi_r)<\psi(\xi_r) \quad \text{for} \quad r<j,\ h_{N-j}(\xi_j)=\psi(\xi_j)\} \\ &= \{x|\xi_r\notin\Xi_{N-r} \quad \text{for} \quad r<j,\ \xi_j\in\Xi_{N-j}\}, \end{aligned} \tag{37}$$

where

$$\Xi_j=\{\xi|h_j(\xi)=\psi(\xi)\}, \tag{38}$$

$$\Xi_{jy}=\{\xi|h_j(\xi)=C_1+\underset{y}{E}[h_{j-1}(T_y\xi)]\}, \tag{39}$$

$$\Xi_{jz}=\{\xi|h_j(\xi)=C_2+\underset{z}{E}[h_{j-1}(T_z\xi)]\}. \tag{40}$$

If $\alpha_{rN}<U_r$, then $\alpha_{rN}=\min[\underset{y}{E}_r[\alpha_{r+1}^y{}_{,N}], \underset{z}{E}_r[\alpha_{r+1}^z{}_{,N}]]$.

Therefore,

$$\Omega_x^{r+1}=\Omega_y, \quad \text{if} \quad \alpha_{rN}=\underset{y}{E}_r[\alpha_{r+1}^y{}_{,N}],$$
$$\Omega_x^{r+1}=\Omega_z, \quad \text{if} \quad \alpha_{rN}=\underset{z}{E}_r[\alpha_{r+1}^z{}_{,N}],$$

where

$$\Omega_x^{r+1}=\Omega_y \quad \text{if and only if} \quad \xi_r\in\Xi_{N-r,y},$$
$$\Omega_x^{r+1}=\Omega_y \quad \text{if and only if} \quad \xi_r\in\Xi_{N-r,z}.$$

ii) $S_{iN}^*\cap S_{jN}^*=\phi$ for any i, j $(0\leq i,\ j\leq N,\ i\neq j)$

Proof For any p, q $(p<q)$, $\xi_p\notin\Xi_{N-p}$ for all $x\in S_{qN}^*$. On the other hand, ξ_p

$\in \Xi_{N-p}$ for all $x \in S_{pN}^*$.

Therefore,

$$S_{qN}^* \cap S_{pN}^* = \phi.$$

iii) Any point in Ω_x belongs to only one of S_{jN}^*.

Proof If x does not belong to any $S_{jN}^*(j<N)$, then x belongs to S_{NN}^*, because $\alpha_{NN}=U_N$.

In consequence of i), ii) and iii), S_N^* is a partition of Ω_x.

The meaning of each notation is as follows:

$L_j(\xi_j)$ = the risk of optimal terminal action after j observations,

α_{jN} = the risk of j-th stage assuming that optimal actions are made after the j-th stage.

Therefore, at each stage, if $\alpha_{jN}=U_j$, then the optimal terminal action should be made, and if $\alpha_{jN}<U_j$, then a sampling of $\{Y\}$ or $\{Z\}$ should be used, according to $\alpha_{jN}=\alpha_{jN}^y$ or α_{jN}^z, respectively.

Ξ_{N-j} represents the stopping region. That is, if at the $(j+1)$th stage a priori probability distribution $\xi_j \in \Xi_{N-j}$, then the optimal terminal action should be made.

$\Xi_{N-j,\,y}$ represents the region where $\{Y\}$ should be observed.

$\Xi_{N-j,\,z}$ represents the region where $\{Z\}$ should be observed.

[II] It is shown in the following that S_N^* is Bayes' optimal.

It is sufficient to show that

$$\rho_N^*(\xi) = \min_{S_N} \rho(\xi, S_N) = \rho(\xi, S_N^*), \tag{41}$$

where S_N is a partition of Ω_x.

$\rho_N^*(\xi)$ is given as follows:

$$\rho_N^*(\xi) = \min[\phi(\xi), C_1 + \underset{y}{E}[\rho_{N-1}^*(T_y\xi)], C_2 + \underset{z}{E}[\rho_{N-1}^*(T_z\xi)]]. \tag{42}$$

If ξ_0 is fixed, then ξ_j depends on only $x_1, x_2, \cdots, x_j$.

Therefore, the following expression can be introduced;

$$\alpha_{jN} = \alpha_{jN}(\xi_j) = \alpha_{jN}(x) \qquad \text{for any } x \in S_{jN},$$

where $\alpha_{jN}(x)$ as well as $U_j(\xi_j) = U_j(x)$ mean that they are functions of only $x_1, x_2, \cdots, x_j$.

Let $S_N = (S_{0N}, S_{1N}, \cdots, S_{NN})$ be an arbitrary partition of Ω_x, and

$$T_r = S_{rN} \cup S_{r+1,\,N} \cup \cdots \cup S_{NN} \qquad \text{for} \quad r = 0, 1, \cdots, N,$$

then

$$g(r)=\sum_{j=0}^{r-1}\sum_{x\in S_{jN}}\alpha_{jN}(x)P_\xi(x)+\sum_{x\in T_r}\alpha_{rN}(x)P_\xi(x)$$
$$=\sum_{j=0}^{r}\sum_{x\in S_{jN}}\alpha_{jN}(x)P_\xi(x)+\sum_{x\in T_{r+1}}\alpha_{rN}(x)P_\xi(x), \tag{43}$$

$$g(N)=\sum_{j=0}^{N}\sum_{x\in S_{jN}}\alpha_{jN}(x)P_\xi(x), \tag{44}$$

$$g(0)=\sum_{x\in T_0}\alpha_{0N}(x)P_\xi(x)=\alpha_{0N}, \tag{45}$$

$$\alpha_{0N}(x)=\alpha_{0N}(\xi_0)=\alpha_{0N}=\text{constant}. \tag{46}$$

It can be shown that

$$\sum_{x\in T_{r+1}}\alpha_{r+1,N}(x)P_\xi(x)=\sum_{x\in T_{r+1}}E_r[\alpha_{r+1,N}(x)]P_\xi(x), \tag{47}$$

where

$$E_r[\alpha_{r+1,N}]=\min[\underset{y}{E_r}[\alpha_{r+1}{}^y{}_{,N}],\ \underset{z}{E_r}[\alpha_{r+1}{}^z{}_{,N}]]. \tag{48}$$

By (43), (47) and definition of α_{jN}, the inequality

$$g(r+1)\geq\sum_{j=0}^{r}\sum_{x\in S_{jN}}\alpha_{jN}(x)P_\xi(x)+\sum_{x\in T_{r+1}}\alpha_{rN}(x)P_\xi(x)=g(r) \tag{49}$$

holds with equality if $\alpha_{rN}=E_r[\alpha_{r+1,N}]$, and does

$$\rho(\xi,S_N)=\sum_{j=0}^{N}\sum_{x\in S_{jN}}U_j(x)P_\xi(x)\geq\sum_{j=0}^{N}\sum_{x\in S_{jN}}\alpha_{jN}(x)P_\xi(x)=g(N) \tag{50}$$

for any S_N.

If $S_N=S_N^*$, then $x\notin S_{jN}$ $(j=0,1,\cdots,r)$ for $x\in T_{r+1}$, and

$$\alpha_{rN}(x)<U_r(x). \tag{51}$$

Therefore

$$\alpha_{rN}(x)=E_r[\alpha_{r+1,N}(x)], \tag{52}$$

and

$$g(r+1)=g(r). \tag{53}$$

The inequality holds

$$\rho(\xi,S_N)\geq\rho(\xi,S_N^*), \tag{54}$$

because

$$\rho_N^*(\xi)=\rho(\xi,S_N^*)=\sum_{j=0}^{N}\sum_{x\in S_{jN}^*}U_j(x)P_\xi(x)=g(N)=g(0)=\alpha_{0N}. \tag{55}$$

Therefore, S_N^* is a Bayes solution with respect to ξ, and the following holds:

$$\rho_N^*(\xi)=\alpha_{0N}=h_N(\xi), \qquad \text{for} \quad N=0,1,2,\cdots. \tag{56}$$

Next, S_k^* for k-TSSP is considered. Then,

$$\rho_k^*(\xi)=\rho(\xi, S_k^*)=h_k(\xi). \tag{57}$$

It is shown that S_k^* is a part of S_N^*, in the following.

Let us consider $(N-j_0)$-TSSP, where $(N-j_0)$ observations are permitted after the j_0 observations.

A priori probability distribution is ξ_{j_0}. The total cost of j_0 observations is

$$C_{j_0}=C_1 m_{j_0}+C_2(j_0-m_{j_0}). \tag{58}$$

Let

$$\begin{aligned} U_{j-j_0}'\equiv U_j-C_{j_0} &= C_1 m_j+C_2(j-m_j)+\psi-\{C_1 m_{j_0}+C_2(j_0-m_{j_0})\} \\ &= C_1(m_j-m_{j_0})+C_2(j-j_0-m_j+m_{j_0})+\psi \\ &= C_1 m_{j-j_0}'+C_2(j-j_0-m_{j-j_0}')+\psi, \end{aligned} \tag{59}$$

where

$$m_{j-j_0}'=m_j-m_{j_0}, \tag{60}$$

and

$$\begin{aligned} &\alpha_{N-j_0,\, N-j_0}\equiv U_{N-j_0}' \\ &\alpha_{j,\, N-j_0}\equiv \min[U_j',\ \underset{y}{E_j}[\alpha_{j+1}{}^y{}_{,\, N-j_0}],\ \underset{z}{E_j}[\alpha_{j+1}{}^z{}_{,\, N-j_0}]]. \end{aligned} \tag{61}$$

Then

$$\alpha_{N-j_0,\, N-j_0}=U_{N-j_0}'=U_N-C_{j_0}=\alpha_{NN}-C_{j_0}, \tag{62}$$

$$\alpha_{j-j_0,\, N-j_0}=\alpha_{jN}-C_{j_0}, \qquad j=j_0,\ j_0+1,\cdots,N. \tag{63}$$

In this case, an optimal solution for $(N-j_0)$-TSSP can be constructed by

$$S_{N-j_0}^*=(S_0^*{}_{,\, N-j_0},\ S_1^*{}_{,\, N-j_0},\cdots,S_{N-j_0}^*{}_{,\, N-j_0}), \tag{64}$$

where

$$S_j^*{}_{,\, N-j_0}=\{x|\alpha_{r,\, N-j_0}<U_r' \text{ for } r<j,\ \alpha_{j,\, N-j_0}=U_j'\}. \tag{65}$$

$\Omega_x{}^{r+1}$ is decided for $r=0,1,\cdots,N-j_0-1$ similarly in N-TSSP according to

$$\alpha_{r,\, N-j}=\min[\underset{y}{E_r}(\alpha_{r+1}{}^y{}_{,\, N-j_0}),\ \underset{z}{E_r}(\alpha_{r+1}{}^z{}_{,\, N-j_0})]. \tag{66}$$

Then, S_{N-j}^* can be considered as a partition of $(N-j_0)$-product space

$$\Omega_x=\Omega_x{}^1\times\Omega_x{}^2\times\cdots\times\Omega_x{}^{N-j_0}.$$

By the following equalities,

$$\alpha_{r,\,N-j_0}=\alpha_{r,\,N-j_0}(\xi_r')=\alpha_{r+j_0,\,N}(\xi_{r+j_0})-C_{j_0},\quad \xi_r'=\xi_{r+j_0}, \tag{67}$$

$$\alpha_{r+1}{}^{y}{}_{,\,N-j_0}(\xi_r')=\alpha_{r+1,\,N-j_0}(T_y\xi_r')=\alpha_{r+1+j_0}{}^{y}{}_{,\,N}(\xi_{r+j_0})-C_{j_0}, \tag{68}$$

$$\alpha_{r+1}{}^{z}{}_{,\,N-j_0}(\xi_r')=\alpha_{r+1,\,N-j_0}(T_z\xi_r')=\alpha_{r+1+j_0}{}^{z}{}_{,\,N}(\xi_{r+j_0})-C_{j_0} \tag{69}$$

can be rewritten as

$$\alpha_{r+j_0,\,N}(\xi_{r+j_0})=\min[\underset{y}{E}_{r+j_0}[\alpha_{r+1+j_0}{}^{y}{}_{,\,N}(\xi_{r+j_0})],\ \underset{z}{E}_{r+j_0}[\alpha_{r+1+j_0}{}^{z}{}_{,\,N}(\xi_{r+j_0})]],$$
$$r=0,1,\cdots,N-j_0-1. \tag{70}$$

Furthermore,

$$\alpha_{kN}(\xi_k)=\min[\underset{y}{E}_k[\alpha_{k+1}{}^{y}{}_{,\,N}(\xi_k)],\ \underset{z}{E}_k[\alpha_{k+1}{}^{z}{}_{,\,N}(\xi_k)]],$$
$$k=j_0,j_0+1,\cdots,N-1. \tag{71}$$

If

$$\alpha_{r,\,N-j_0}=\underset{y}{E}_r[\alpha_{r+1}{}^{y}{}_{,\,N-j_0}], \tag{72}$$

or

$$\alpha_{r,\,N-j_0}=\underset{z}{E}_r[\alpha_{r+1}{}^{z}{}_{,\,N-j_0}],\qquad \text{for}\quad r=0,1,\cdots,N-j_0-1, \tag{73}$$

then equalities

$$\alpha_{kN}=\underset{y}{E}_k[\alpha_{k+1}{}^{y}{}_{,\,N}], \tag{74}$$

or

$$\alpha_{kN}=\underset{z}{E}_k[\alpha_{k+1}{}^{z}{}_{,\,N}], \tag{75}$$

hold respectively.

Therefore,

$$\Omega_x{}^{r+1}=\Omega_x{}^{k+1}\qquad \text{for}\ \begin{cases} r=0,1,\cdots,N-j_0-1\\ k=j_0,j_0+1,\cdots,N-1,\end{cases} \tag{76}$$

and

$$\Omega_x{}^1\times\Omega_x{}^2\times\cdots\times\Omega_x{}^{N-j_0}=\Omega_x{}^{j_0+1}\times\Omega_x{}^{j_0+2}\times\cdots\times\Omega_x{}^{N}. \tag{77}$$

$(N-j_0)$-dim vector x in (65) can be extended to N-dim by adding fixed $(x_1,\cdots,x_{j_0})$ for the first j_0 elements. Denoting this,

$$\begin{aligned}S_j{}^*{}_{,\,N-j_0}&=\{x|\alpha_{r,\,N-j_0}<U_r' \text{ for } r<j,\ \alpha_{j,\,N-j_0}=U_j'\}\\ &=\{x|\alpha_{r+j_0,\,N}-C_{j_0}<U_{r+j_0}-C_{j_0} \text{ for } r<j_0,\ \alpha_{j+j_0,\,N}-C_{j_0}=U_{j+j_0}-C_{j_0}\}\\ &=\{x|\alpha_{r+j_0,\,N}<U_{r+j_0} \text{ for } r<j,\ \alpha_{j+j_0,\,N}=U_{j+j_0}\}\end{aligned}$$

$$= \{x|\alpha_{r,N} < U_r \text{ for } r < j+j_0,\ \alpha_{j+j_0,N} = U_{j+j_0}\}$$
$$= S_{j+j_0}{}^*{}_{,N}, \qquad j=0,1,\cdots,N-j_0. \tag{78}$$

Therefore,

$$S_{N-j_0}{}^* = (S_0{}^*{}_{,N-j_0}, S_1{}^*{}_{,N-j_0}, \cdots, S_{N-j_0}{}^*{}_{,N-j_0})$$
$$= (S_{j_0}{}^*{}_{,N}, S_{j_0+1}{}^*{}_{,N}, \cdots, S_{NN}{}^*). \tag{79}$$

This shows that $S_k{}^*$ is a part of $S_N{}^*$.

Hence

$$h_j(\xi) = \rho_j{}^*(\xi) \qquad \text{for } j=0,1,\cdots,N, \tag{80}$$

and

$$\rho_N{}^*(\xi) = \min[\psi(\xi), C_1 + \underset{y}{E}[\rho_{N-1}{}^*(T_y\xi)], C_2 + \underset{z}{E}[\rho_{N-1}{}^*(T_z\xi)]]. \tag{81}$$

This completes the proof of [II].

Corollary 1. $h_j(\xi) \geq h_{j+1}(\xi)$ and $\Xi_j \supset \Xi_{j+1}$, for fixed ξ.

Corollary 2. $\rho_N{}^*(\xi)$ is a concave function with respect to ξ.

Conclusion

A special kind of statistical game, an N-truncated sequential game, has been investigated where the statistician can choose one of two sampling methods as a subexperiment at each stage.

The existence of an optimal strategy for the statistician, that is, an optimal solution $S_N{}^*$ for N-TSSP, is proved. Then, given an a priori probability distribution ξ on Ω, Bayes risk $\rho_N{}^*(\xi)$ is derived in the Theorem.

Further, some properties of Bayes risk are shown in the Corollaries.

References

1) A. Wald: Statistical Decision Functions, John Wiley, 1959.
2) D. Blackwell and M. A. Dirshick: Theory of Games and Statistical Decisions, John Wiley, 1954.
3) H. Raiffa: Decision Analysis, Addison-Wesley, 1968.
4) T. R. Dickman, S. Smidt and A. K. McAdams: Management Decision Making under Uncertainty, Macmillan, 1969.
5) L. Weiss: Statistical Decision Theory, McGraw-Hill, 1961.
6) H. Chernoff and L. E. Moses: Elementary Decision Theory, John Wiley, 1959.
7) M. H. DeGroot: "Uncertainty, information and sequential experiments", The Annals of Mathematical Statistics, Vol. 33-I, 1962.
8) P. C. Fishburn: "Analysis of decision with incomplete knowledge of probability", Operations Research, Vol. 13, 1965.

9) E. Samuel: "Note on a sequential classification problem", The Annals of Mathematics, Vol. 34-II, 1963.
10) P. C. Fishburn: Utility Theory for Decision Making, John Wiley, 1970.