On the statistical decision function I.

Miyasawa, Koichi
Akita University

https://doi.org/10.5109/12947

出版情報：統計数理研究. 4 (1/2), pp.22-32, 1950-12. Research Association of Statistical Sciences
バージョン：
権利関係：

ON THE STATISTICAL DECISION FUNCTION[1] I.

Kôichi MIYASAWA

(*Akita Univ.*)

1. Introduction

Let $X=\{X_i\}$ $(i=1,2,3,\ldots)$ be an infinite sequence of chance variables. Any particular observation x on X is given by a sequence $x=\{x_i\}$ $(i=1,2,3,\ldots)$ of real values, where x_i denotes the observed value of X_i. Let the space of all the sample points x be M, the Borel field which contains all the sets $\{x;\ x_i<a_i,\ i=1,2,\ldots\}$ be K, where a_i are real numbers or $+\infty$, and the Lebesgue measure on K be m. Suppose that the distribution function $F(x)$ of X is not known, it is, however, known that $F(x)$ has the probability density function $p(F|x)$ and is an element of a given class Ω of distribution functions. There is, furthermore, a space D^* given whose elements d represent the possible decision that can be made in the problem under consideration. The problem is to construct a function $D=D(x)$ called the statistical decision function, which associates with each sample point x an element d of D^* so that the decision $d=D(x)$ is made when x is observed. Let $W(F,d)$ be the loss suffered by the statistician when F is the true distribution of X and the decision d is made. We assume that $W(F,d)$ is a non-negative bounded measurable function of F and d. Let $c(n)$ be the cost of making n observations, i. e., $c(n)$ is the cost of observing the values of $x_1,\ldots,x_n$. Thus, when the true distribution function of X is F, and if we decide the element of D^* by the decision function $d_n(x)$ which depends only on the first n coordinates $x_1,x_2,\ldots,x_n$ of the sample x, the loss is given by the following sum

$$r(F,\ d_n(x))=W(F,d_n(x))+c(n). \tag{1.01}$$

A sequential statistical decision function $\underline{D}$ is composed of the following two sequences $B=\{B_j\}$ and $D=\{d_j\}$.

(i) $\{B_j\}$ is the sequence of B_0 and disjoint subsets $B_1,B_2,\ldots B_j,\ldots$ of M, where B_j depends on the first j coordinates $x_1,\ldots,x_j$ of a sample x and B_j indicates that the sampling should stop at the j-th observation when $x\in B_j$ $(j=1,2,\ldots)$, B_0 is the event that we do not sample at all, but take some decision immediately and it will have probability either 0 or 1. It should be

$$\sum_{j=0}^{\infty}P_r(B_j|F)=1 \text{ for all } F\in\Omega. \tag{1.02}$$

This sequence $B=\{B_j\}$ $(j=0,1,2,\ldots)$ is called a sequential procedure.

(ii) $\{d_j\}$ is the sequence of d_0 and decision functions $d_1(x)$, $d_2(x),\ldots,$

1) Communicated at the Autumn-Meeting of Japanese Math. Soc., at Kôbe Univ., October 20, 1950.

$d_j(x),\ldots$, where d_0 is an element of D^* and $d_j(x)$ is the function of the first j coordinates $x_1,\ldots,x_j$ of $x \in B_j$ and its value is some decision, i.e., some element of D^*. This sequence $D=\{d_j\}$ is also called a decision function.

The sequential decision function $\underline{D}$ which is determined by two sequence B and D will be denoted by $\underline{D}=(B, D)$. Then the average loss caused by the sequential decision function $\underline{D}$ when F is the true distribution function of X is given by

$$(1.03) \qquad r(F,\underline{D})=\sum_{j=0}^{\infty}\int_{Bj} r(F,d_j(x))p(F \mid x)dx.$$

Here, we assume that the series on the right hand member of (1) is always convergent in our problem. Let ξ be an a priori distribution on Ω, i. e., ξ is a probability measure defined over a suitably chosen Borel field of subsets of Ω.

Then the expected value of $r(F, \underline{D})$ is given by

$$(1.04) \qquad r(\xi,\underline{D})=\int_{\Omega} r(F,\underline{D})d\xi$$

$r(\xi,\underline{D})$ is called the risk when ξ is the a priori distribution on Ω and $\underline{D}$ is the decision function adopted. The sequential decision function $\underline{D}^*$ is called a Bayes solution relative to the a priori distribution ξ, if

$$(1.05) \qquad r(\xi,\underline{D}^*)=\inf_{\underline{D}} r(\xi,\underline{D}).$$

Our object is to determine the necessary and sufficient condition so that a sequential decision function will be a Bayes solution relative to the a priori distribution ξ.

2. Theorem and its Proof

Let D_j^* be the set of all decision functions $d_j(x)$ which depends only on the first j coordinates $x_1,\ldots,x_j$ of $x(j=1,2,\ldots)$. For $j=0$, d_0 is some element of D^*, i.e., $D_0^*=D^*$. Let $\{d_{jn}\}$ and d_j^* be a sequence of elements of D_j^* and an element of D_j^*, respectively. If it is valid that

$$W(F,d_{jn}(x))\rightarrow W(F,d_j^*(x)),$$

as $n\rightarrow\infty$, for all $F\in\Omega$ all $x\in M$, then we say that $\{d_{jn}\}$ converges to d_j^*. We assume that D_j^* is compact in the sense of the convergence of the above definition.

LEMMA.

For any a priori distribution ξ of F and for any j, there exists a decision function $d_j^\circ(x)$ such that

$$(2.01) \qquad r(\xi,d_j^\circ(x))=\inf_{d_j\in D_j^*} r(\xi,d_j(x)) \qquad \text{for all } x.$$

where $d_j(x)$ is an element of D_j^* and

(2.02) $$r(\xi, d_j(x)) = \int_\Omega r(F, d_j(x)) p(F \mid x) d\xi.$$

As this lemma will be proved in the same way as done in GIRSHICK'S[1], we do not refer to it here.

We write

(2.03) $$r_j(\xi, x) = r(\xi, d_j^\circ(x)) = \inf_{dj} r(\xi, d_j(x))$$

Let N be any fixed integer, and by the induction backwards we define functions $\alpha_{jN}(\xi, x)$ $(j=0, 1, \dots, N)$ which depends only on the first j coordinates of x.

That is

(2.04) $$\alpha_{NN}(\xi, x) = r_N(\xi, x)$$

and, for $j<N$

(2.05) $$\alpha_{jN}(\xi, x) = \min\ (r_j(\xi, x),\ E_j\{\alpha_{j+1,N}(\xi, x)\})$$

where E_j is the conditional expectation given $x_1, \dots, x_j$, i. e.,

(2.06) $$E_j\{\alpha_{j+1,N}(\xi, x)\} = \int_{-\infty}^{\infty} \alpha_{j+1,N}(\xi, x) dx_{j+1}.$$

It can be seen easily that when j is fixed, $\alpha_{jN}(\xi, x)$ is non-negative and non-increasing as N increases. Therefore there exists $\lim_{N\to\infty} \alpha_{jN}(\xi, x)$, and we represent this limit as $\alpha_j(\xi, x)$, i. e.,

(2.07) $$\alpha_j(\xi, x) = \lim_{N\to\infty} \alpha_{jN}(\xi, x).$$

Then the following relation will hold

(2.08) $$\alpha_j(\xi, x) = \min\{r_j(\xi, x),\ E_j\{\alpha_{j+1}(\xi, x)\}\},$$

and consequently

(2.09) $$\alpha_j(\xi, x) \leqq r_j(\xi, x) \quad (j=0, 1, 2, \dots).$$

By means of these functions $\alpha_j(\xi, x)$ we define subsets S_j of M as follows:

(2.10) $$S_j = \{x;\ r_i(\xi, x) > \alpha_i(\xi, x) \text{ for } i<j, \text{ and } r_j(\xi, x) = \alpha_j(\xi, x)\}.$$

It is clear that thus defined sequence of subsets $\{S_j\}$ $(j=0, 1, 2, \dots)$ forms a sequential procedure.

Let us denote the sequential procedure which consists of the sequence $\{S_j\}$ by S_ξ and the decision function which consists of the sequence $\{d_j^\circ\}$ by D°. Let D_ξ be the sequential decision function which is determined by S_ξ and D°, that is,

(2.11) $$D_\xi = (S_\xi, D^\circ)$$

We assume that if for any fixed element x of some subset $A \subset M$,

$$r(\xi, d_j(x)) = \int_\Omega r(F, d_j(x)) p(F \mid x) d\xi, \tag{2.12}$$

has the minimum value with respect to d_j, for $d_j = d_j^\circ$ and $d_j = d_j^*$, then it holds $d_j^\circ(x) = d_j^*(x)$ almost everywhere on A.
Then the following theorem holds.
THEOREM .

There exists a Bayes solution relative to any a priori distribution ξ. The necessary and sufficient condition for that the sequential decision function $\underline{D} = (T, D)$, where $T = \{B_j\}$, $D = \{d_j\}$, will be a Bayes solution relative to the a priori distribution ξ is that

(i) *$d_j(x) = d_j^\circ(x)$, almost everywhere on B_j, $(j = 0, 1, 2, \dots)$*

(ii) *the following relation holds except for a set of measure 0*

$$B_0 = S_0,\ B_1 \subset S_1,\ B_2 \subset S_1 + S_2, \dots,\ B_j \subset S_1 + S_2 + \dots + S_j, \dots \tag{2.13}$$

Consequently, if we write

$$B_i \cap S_i = D_i,\ S_i - D_i = S_i',\ B_l \cap S_k' = B_l^k\ (l = 2, 3, \dots;\ k = 1, 2, \dots l-1) \tag{2.14}$$

then it holds

$$B_1 = D_1,\ B_2 = D_2 + B_2^1,\ B_3 = D_3 + B_3^1 + B_3^2, \dots,\ B_j = D_j + B_j^1 + B_j^2 + \dots + B_j^{j-1}, \dots \tag{2.15}$$

and

$$S_i' = B_{i+1}^i + B_{i+2}^i + B_{i+3}^i + \dots,\ (i = 1, 2, 3, \dots). \tag{2.16}$$

(iii) *if there exists some k such that $m(S_k') > 0$, then except for the set of measure 0 it holds*

$$r_{k+1}(\xi, x) = E_{k+1}\{r_{k+2}(\xi, x)\}, \tag{2.17}$$

on the complement $C\, B_{k+1}^k$ of B_{k+1}^k with respect to S_k'

$$r_{k+2}(\xi, x) = E_{k+2}\{r_{k+3}(\xi, x)\}, \tag{2.18}$$

on the complement $C[B_{k+1}^k + B_{k+2}^k]$ of $[B_{k+1}^k + B_{k+2}^k]$ with respect to S_k', and so on.

PROOF: It can be shown that the above defined sequential decision function $D_\xi = (S_\xi, D^\circ)$ is a Bayes solution relative to the a priori distribution ξ in a analogous way as done in Girshick's[1], so we will omit it here. Let $\underline{D} = (T, D)$, $T = \{B_j\}$, $D = \{d_j\}$ be a Bayes solution relative to the a priori distribution ξ.

At first we will prove that the condition (i) is necessary. It follows from the lemma that

$$r(\xi, d_j(x)) \geqq r(\xi, d_j^\circ(x)). \tag{2.19}$$

Now let us put

(2.20) $$R_j=\{x;\ x\in B_j \text{ and } r(\xi,d_j(x))>r(\xi,d_j^\circ(x))\},$$

and suppose that $m(R_j)>0$.
Then, we can choose $\delta>0$ and $R_j'\subset R_j$ such that $m(R_j')>0$, and the following relation holds for $x\in R_j'$

(2.21) $$r(\xi,d_j(x))>r(\xi,d_j^\circ(x))+\delta.$$

Consequently it follows

(2.22) $$\begin{aligned} r(\xi;\ B_j,d_j)&=\int_{Bj}\int_\Omega r(F,d_j(x))\,p(F|x)d\,\xi\,dx=\int_{Bj}r(\xi,d_j(x))dx \\ &>\int_{R'j}r(\xi,d_j^\circ(x))dx+\delta\int_{R'j}dx+\int_{CR'j}r(\xi,d_j^\circ(x))dx \\ &=\int_{Bj}r(\xi,d_j^\circ(x))px+\delta m(R_j') \\ &>\int_{Bj}r(\xi,d_j^\circ(x))dx=r(\xi;\ B_j,d_j^\circ). \end{aligned}$$

(where CR_j' denotes the complement of R_j' with respect to B_j)
And generally it holds

(2.23) $$r(\xi;\ B_i,d_i)\geqq r(\xi;\ B_i,\ d_i^\circ).$$

Consequently, it folds

(2.24) $$r(\xi,\underline{D})=\sum_{j=0}^{\infty}r(\xi,B_j,d_j)>\sum_{j=0}^{\infty}r(\xi;\ B_j,d_j^\circ)=r(\xi;\ T,D^\circ)$$

Thus we have a sequential decision function $\underline{D}'=(T,D^\circ)$ such that $r(\xi,\underline{D})>r(\xi,\underline{D}')$. This contradicts with the assumption that the sequential decision function $\underline{D}=(T,D)$ is a Bayes solution relative to ξ.

Therefore, it holds on B_j

(2.26) $$r(\xi,d_j(x))=r(\xi,d_j^\circ(x))\ \ (j=0,1,2,\dots)$$

except for a set of measure 0. Consequently it follows from our assumption that $d_j(x)=d_j^\circ(x)$ almost everywhere on B_j ($j=0,1,2,\dots$).

Next we will prove that the condition (ii) is necessary.

We put

(2.27) $$B_i'=(B_i-D_i)\cap C[S_1+S_2+\dots+S_i],$$

where $C[S_1+\dots+S_i]$ is the complement of $[S_1+\dots+S_i]$ with respect to M,

$$B_i'' = (B_i - D_i) \cap [S_1 + S_2 + \dots + S_i] \tag{2.28}$$

$$S_i^l = B_l' \cap S_i \quad (i = l+1,\ l+2, \dots;\ l = 1, 2, \dots)$$

Let $S_i^l(1, 2, \dots, l)$ be the intersection of S_i^l and the subset $C(x_1, \dots, x_l)$ of M defined by $x_1 = \text{const.}, \dots, x_l = \text{const.}$ and $CS_i^l(1, 2, \dots, l)$ be the complement of $S_i^l(1, 2, \dots, l)$ with respect to $C(x_1, \dots, x_l)$.

To prove that $m(B_i') = 0$ $(i = 1, 2, \dots)$, we assume that $m(B_l') > 0$ for some l.

Now, by the definition of S_j, it follows that if $x \in B_l'$

$$r_l(\xi, x) > \alpha_l(\xi, x) = E_l\{\alpha_{l+1}(\xi, x)\} = \int_{-\infty}^{\infty} \alpha_{l+1}(\xi, x) dx_{l+1} \tag{2.29}$$

$$= \int_{S_{l+1}^l(1,2,\dots,l)} \alpha_{l+1}(\xi, x) dx_{l+1} + \int_{CS_{l+1}^l(1,2,\dots,l)} \alpha_{l+1}(\xi, x) dx_{l+1}.$$

And if $x \in S_{l+1}^l(1, 2, \dots, l)$, then we have

$$\alpha_{l+1}(\xi, x) = r_{l+1}(\xi, x). \tag{2.30}$$

Therefore it follows

$$r_l(\xi, x) > \int_{S_{l+1}^l(1,2,\dots,l)} r_{l+1}(\xi, x) dx_{l+1} + \int_{CS_{l+1}^l(1,2,\dots,l)} \alpha_{l+1}(\xi, x) dx_{l+1}. \tag{2.31}$$

If $x \in CS_{l+1}^l(1, 2, \dots, l)$, it holds by the definition of S_j

$$r_{l+1}(\xi, x) > \alpha_{l+1}(\xi, x) \tag{2.32}$$

hence

$$\alpha_{l+1}(\xi, x) = E_{l+1}\{\alpha_{l+2}(\xi, x)\}.$$

Consequently it follows

$$\int_{CS_{l+1}^l(1,2,\dots,l)} \alpha_{l+1}(\xi, x) dx_{l+1} = \int_{CS_{l+1}^l(1,2,\dots,l)} \int_{-\infty}^{\infty} \alpha_{l+2}(\xi, x) dx_{l+2}\, dx_{l+1} \tag{2.33}$$

$$= \int_{S_{l+2}^l(1,2,\dots,l)} \int_{-\infty}^{\infty} \alpha_{l+2}(\xi, x) dx_{l+2}\, dx_{l+1} + \int_{C[S_{l+1}^l(1,2,\dots,l) + S_{l+2}^l(1,2,\dots,l)]} \int_{-\infty}^{\infty} \alpha_{l+2}(\xi, x) dx_{l+2} dx_{l+1}$$

$$= \int_{S_{l+2}^l(1,2,\dots,l)} \int_{S_{l+2}^l(1,2,\dots,l,l+1)} \alpha_{l+2}(\xi, x) dx_{l+2}\, dx_{l+1}$$

$$+ \int_{S_{l+2}^l(1,\dots,l)} \int_{CS_{l+2}^l(1,2,\dots,l,l+1)} \alpha_{l+2}(\xi, x) dx_{l+2}\, dx_{l+1}$$

$$+ \int_{C[S_{l+1}^l(1,\dots,l) + S_{l+2}^l(\dots,l)]} \int_{-\infty}^{\infty} \alpha_{l+2}(\xi, x) dx_{l+2}\, dx_{l+1}$$

$$= \int_{S_{l+2}^l(1,\dots,l)} r_{l+2}(\xi, x) dx_{l+2}\, dx_{l+1} + \int_{C[S_{l+1}^l(1,\dots,l) + S_{l+2}^l(1,\dots,l)]} \alpha_{l+2}(\xi, x) dx_{l+2}\, dx_{l+1}.$$

Since we have

$$(2.33)\qquad \alpha_{l+2}(\xi, x)=r_{l+2}(\xi, x) \text{ on } S^l_{l+2}(1,\dots, l, l+1).$$

From (2.31) and (2.33) it follows

$$(2.34)\quad r_l(\xi, x)>\int_{S^l_{l+1}(1,\dots,l)} r_{l+1}(\xi, x)dx_{l+1}+\int_{S^l_{l+2}(1,\dots,l)} r_{l+2}(\xi, x)dx_{l+2}\, dx_{l+1}$$
$$+\int_{C[S^l_{l+1}(1,\dots,l)+S^l_{l+2}(1,\dots,l)]} \alpha_{l+2}(\xi, x)dx_{l+2}\, dx_{l+1}.$$

If $x\in B_l'$, using the analogous method as above, we can conclude the following relation,

$$(2.35)\quad r_l(\xi, x)>\int_{S^l_{l+1}(1,\dots,l)} r_{l+1}(\xi, x)\, dx_{l+1}+\int_{S^l_{l+2}(1,\dots,l)} r_{l+2}(\xi, x)dx_{l+2}\, dx_{l+1}$$
$$+\int_{S^l_{l+3}(1,\dots,l)} r_{l+3}(\xi, x)dx_{l+3}\, dx_{l+2}\, dx_{l+1}+\dots .$$

Therefore, if $m(B_l')>0$, we have

$$(2.36)\qquad \int_{B_l'} r_l(\xi, x)dx>\int_{S^l_{l+1}} r_{l+1}(\xi, x)dx+\int_{S^l_{l+2}} r_{l+2}(\xi, x)dx+\int_{S^l_{l+3}} r_{l+3}(\xi, x)dx+\dots.$$

Now, we construct the sequential procedure $T^*=\{C_j\}$ as follows: $C_0=B_0,\ C_1=B_1,\dots,\ C_{l-1}=B_{l-1},\ C_l=D_l+B_l''\ C_{l+1}=B_{l+1}+S^l_{l+1},\ C_{l+2}=B_{l+2}S^l_{l+2},\dots$.

Then, it is clear that any two sets C_i and C_j $(i\neq j)$ are disjoint, and

$$(2{\cdot}37)\qquad \begin{aligned} &C_1+C_2+C_3+\dots\\ &=B_1+\dots+B_{l-1}+(D_l+B_l''+S^l_{l+1}+S^l_{l+2}+\dots)+B_{l+1}+B_{l+2}+\dots\\ &=M.\end{aligned}$$

Therefore $T^*=\{C_j\}$ is certainly a sequential procedure. Then, for the sequential decision function $\underline{D}^*$, which is determined by T^* and D°, we have

$$(2.38)\quad r(\xi, D^*)=r_0P_r(C_0)+\int_{C_1} r_1(\xi, x)dx+\dots+\int_{C_{l+1}} r_{l-1}(\xi,x)dx$$
$$+\int_{C_l} r_l(\xi, x)dx+\int_{C_{l+1}} r_{l+1}(\xi, x)dx+\dots$$
$$=r_0P_r(B_0)+\int_{B_1} r_1(\xi, x)dx+\dots+\int_{B_{l-1}} r_{l-1}(\xi, x)dx+\int_{D_l+B_l''} r_l(\xi, x)dx$$
$$+\int_{B_{l+1}+S^l_{l+1}} r_{l+1}(\xi, x)dx+\int_{B_{l+2}+S^l_{l+2}} r_{l+2}(\xi, x)dx+\dots$$

$$=r_0P_r(B_0)+\int_{B_1}r_1(\xi,x)dx+\dots+\int_{B_{l-1}}r_{l-1}(\xi,x)dx$$

$$+\left\{\int_{D_l+B_l''}r_l(\xi,x)dx+\int_{S_{l+1}^l}r_{l+1}(\xi,x)dx+\int_{S_{l+2}^l}r_{l+2}(\xi,x)dx+\dots\right\}$$

$$+\int_{B_{l+1}}r_{l+1}(\xi,x)dx+\int_{B_{l+2}}r_{l+2}(\xi,x)dx+\dots\quad .$$

So that, from (2.35) and (2.38), it follows

$$(2.39)\quad r(\xi,\underline{D}^*)<r_0P_r(B_0)+\int_{B_1}r_1(\xi,x)dx+\dots+\int_{B_{l-1}}r_{l-1}(\xi,x)dx$$

$$+\left\{\int_{D_l+B_l''}r_l(\xi,x)dx+\int_{B_l'}r_l(\xi,x)dx\right\}$$

$$+\int_{B_{l+1}}r_{l+1}(\xi,x)dx+\int_{B_{l+2}}r_{l+2}(\xi,x)dx+\dots$$

$$=r(\xi;\ T,D^\circ).$$

That is $r(\xi,\underline{D}^*)<r(\xi;\ T,D^\circ)$. This contradicts with the assumption that the sequential decision function (T,D°) is a Bayes solution relative to the a priori distribution ξ. Consequently we may conclude $m(B_l')=0$ $(l=1,2,\dots)$. This shows that the condition (ii) is necessary.

Lastly we will prove that the condition (iii) is necessary.

Let us assume that for some k, say $k=1$, $m(S_1')>0$. If $n\geqq N$, it can be shown easily that $r_N(\xi,x)\geqq E_N\{r_n(\xi,x)\}$, so that

$$(2.40)\qquad r_1(\xi,x)\geqq E_1\{r_2(\xi,x)\}.$$

On the other hand, if $x\in S_1'$, we have $r_1(\xi,x)=\alpha_1(\xi,x)$, so that

$$(2.41)\qquad r_1(\xi,x)\leqq E_1\{\alpha_2(\xi,x)\}\leqq E_1\{r_2(\xi,x)\}.$$

From (2.40) and (2.41), it follows

$$(2.42)\qquad r_1(\xi,x)=E_1\{r_2(\xi,x)\},\ \text{if}\ x\in S_1'.$$

Accordingly we have

$$(2.43)\qquad \int_{S_1'}r_1(\xi,x)dx_1=\int_{S_1'}\int_{-\infty}^{\infty}r_2(\xi,x)dx_2\,dx_1.$$

Let $B_l'(1,\dots,k)$ be the intersection of B_l' and $C(x_1,\dots,x_k)$, and let $CB_l'(1,2,\dots,k)$ be the complement of $B_l'(1,\dots,k)$ with respect to $C(x_1,\dots,x_k)$, where $C(x_1,\dots,x_k)$ is the subset of M defined by $x_1=$const.,..., $x_k=$const. Then

$$(2.44)\qquad \int_{-\infty}^{\infty} r_2(\xi, x)dx_2 = \int_{B_2'(1)} r_2(\xi, x)dx_2 + \int_{CB_2'(1)} r_2(\xi, x)dx_2.$$

So that (2.43) can be written as follows,

$$(2.44)\qquad \int_{S_1'} r_1(\xi, x)dx_1 = \int_{B_2'} r_2(\xi, x)dx_2\, dx_1 + \int_{CB_2'} r_2(\xi, x)dx_2\, dx_1\,,$$

where $C\ B_2'$ is the complement of B_2' with respect to S_1'.
Generally it holds that

$$(2.45)\qquad r_2(\xi, x) \geqq E_2\{r_3(\xi, x)\},$$

but now we assume that the condition (iii) does not hold and on some subset of positive measure of $C\ B_2'$ it holds

$$(2.46)\qquad r_2(\xi, x) > E_2\{r_3(\xi, x)\}.$$

Then, it follows

$$(2.47)\qquad \int_{CB_2'} r_2(\xi, x)dx_2\, dx_1 > \int_{CB_2'} \int_{-\infty}^{\infty} r_3(\xi, x)dx_3\, dx_2\, dx_1.$$

Now

$$(2.48)\qquad \int_{CB_2'(1)} \int_{-\infty}^{\infty} r_3(\xi, x)dx_3\, dx_2 = \int_{B_3'(1)} \int_{-\infty}^{\infty} r_3(\xi, x)dx_3\, dx_2 + \int_{C[B_2'(1)+B_3'(1)]} \int_{-\infty}^{\infty} r_3(\xi, x)dx_3\, dx_2$$

$$= \int_{B_3'(1)} \int_{B_3'(1,2)} r_3(\xi, x)dx_3\, dx_2 + \int_{B_3'(1)} \int_{CB_3'(1,2)} r_3(\xi, x)dx_3\, dx_2 + \int_{C[B_2'(1)+B_3'(1)]} \int_{-\infty}^{\infty} r_3(\xi, x)dx_3\, dx_2$$

$$= \int_{B_3'(1)} r_3(\xi, x)dx_3\, dx_2 + \int_{C[B_2'(1)+B_3'(1)]} r_3(\xi, x)dx_3\, dx_2\,.$$

Therefore it follows from (2.44), (2.47) and (2.48)

$$(2.49)\qquad \int_{S_1'} r_1(\xi, x)dx_1 > \int_{B_2'} r_2(\xi, x)dx + \int_{B_3'} r_3(\xi, x)dx + \int_{C[B_2'+B_3']} r_3(\xi, x)dx\,,$$

where $C[B_2'+B_3']$ is the complement of $B_2'+B_3'$ with respect to S_1'.
Generally it holds that

$$(2.50)\qquad r_3(\xi, x) \geqq E_3\{r_4(\xi, x)\},$$

but now we assume that on some subset of positive measure of $C[B_2'+B_3']$ it holds

$$(2.51) \qquad r_3(\xi, x) > E_3\{r_4(\xi, x)\}.$$

Then it follows

$$(2.52) \qquad \int_{C[B_2'+B_3']} r_3(\xi, x)dx > \int_{C[B_2'+B_3']} \int_{-\infty}^{\infty} r_4(\xi, x)dx_4\, dx_3\, dx_2\, dx_1.$$

Let $C[B_2'+B_3']$ $(1, 2)$ be the intersection of $C[B_2'+B_3']$ and $C(x_1, x_2)$, etc. Then we have

$$(2.53) \qquad \int_{C[B_2'+B_3'](1,2)} \int_{-\infty}^{\infty} r_4(\xi, x)dx_4\, dx_3$$
$$= \int_{B_4'(1,2)} \int_{-\infty}^{\infty} r_4(\xi, x)dx_4\, dx_3 + \int_{C[B_2'+B_3'+B_4'](1,2)} \int_{-\infty}^{\infty} r_4(\xi, x)dx_4\, dx_3$$
$$= \int_{B_4'(1,2)} \int_{B_4'(1,2,3)} r_4(\xi, x)dx_4\, dx_3 + \int_{B_4'(1,2)} \int_{CB_4'(1,2,3)} r_4(\xi, x)dx_4\, dx_3$$
$$+ \int_{C[B_2'+B_3'+B_4'](1,2)} \int_{-\infty}^{\infty} r_4(\xi, x)dx_4\, dx_3$$
$$= \int_{B_4'(1,2)} r_4(\xi, x)dx_4\, dx_3 + \int_{C[B_2'+B_3'+B_4'](1,2)} r_4(\xi, x)dx_4\, dx_3.$$

Consequently

$$(2.54) \qquad \int_{C[B_2'+B_3']} \int_{-\infty}^{\infty} r_4(\xi, x)dx_4\, dx_3\, dx_2\, dx_1 = \int_{B_4'} r_4(\xi, x)dx + \int_{C[B_2'+B_3'+B_4']} r_4(\xi, x)dx.$$

From (2.49), (2.52) and (2.54), it follows

$$(2.55) \qquad \int_{S_1'} r_1(\xi, x)dx > \int_{B_2'} r_2(\xi, x)dx + \int_{B_3'} r_3(\xi, x)dx + \int_{B_4'} r_4(\xi, x)dx$$
$$+ \int_{C[B_2'+B_3'+B_4']} r_4(\xi, x)dx.$$

Thus proceeding as above, if the condition (iii) is not satisfied for $k=1$, then we have

$$(2.56) \qquad \int_{S_1'} r_1(\xi, x)dx > \int_{B_2'} r_2(\xi, x)dx + \int_{B_3'} r_3(\xi, x)dx + \ldots + \int_{B_l'} r_l(\xi, x)dx + \ldots .$$

If the condition (iii) is satisfied here, the left and the right hand members of (2.56) are equal.

As when $k=1$, we have generally

$$(2.57) \qquad \int_{S'_k} r_k(\xi, x)dx \geq \int_{B^k_{k+1}} r_{k+1}(\xi, x)dx + \int_{B^k_{k+2}} r_{k+2}(\xi, x)dx + \dots .$$

On the otherhand, we have

$$(2.58) \qquad D_i + S_i' = S_i \quad (i=1,2,\dots)$$

$$D_1 = B_1,\ D_i + B_i^1 + B_i^2 + \dots + B_i^{i-1} = B_i \ (i=2,3,\dots).$$

Therefore if the condition (iii) is not satisfied for some k, say $k=1$, it follows from (2.56), (2.57) and (2.58)

$$(2.59) \qquad r(\xi;\ S_\xi, D^\circ) = r_0 P_r(S_0) + \left[\int_{D_1} r_1(\xi, x)dx + \int_{S'_1} r_1(\xi, x)dx\right]$$

$$+ \left[\int_{D_2} r_2(\xi, x)dx + \int_{S'_2} r_2(\xi, x)dx\right] + \dots$$

$$> r_0 P_r(S_0) + \int_{D_1} r_1(\xi, x) + \left[\int_{D_2} r_2(\xi, x)dx + \int_{B^1_2} r_2(\xi, x)dx\right]$$

$$+ \left[\int_{D_3} r_3(\xi; x)dx + \int_{B^1_3} r_3(\xi, x)dx + \int_{B^2_3} r_3(\xi, x)dx\right] + \dots$$

$$= r_0 P_r(S_0) + \int_{B_1} r_1(\xi, x)dx + \int_{B_2} r_2(\xi, x)dx + \int_{B_3} r_3(\xi, x)dx + \dots$$

$$= r(\xi;\ T, D^\circ).$$

That is

$$(2.60) \qquad r(\xi;\ S_\xi, D^\circ) > r(\xi;\ T, D^\circ).$$

This contradicts with the fact that the sequdntial decision function $D_\xi = (S_\xi, D^\circ)$ is a Bayes solution relative to the a priori distribution ξ. Therefore the condition (iii) is necessary.

On the otherhand if the conditions (i), (ii) and (iii) are satisfied for the sequential decision function $\underline{D} = (T, D)$, then we known from the above proof that $r(\xi, \underline{D}) = r(\xi, D_\xi)$. Therefore in this case the sequential decision function $\underline{D} = (T, D)$ is a Bayes solution relative to ξ.

Consequently, the conditions (i), (ii) and (iii) are the necessary and sufficient condition, so that a sequential decision function may be a Bayes solution relative to the a pripri distribution ξ.

References

1) A. J. Arrow, D. Blackwell, M. A. Girshick; Bayes and Minimax solutions on Sequential decision problems, *Eccnometrica*, **17**(1949), 213—244.

2) A. Wald and J. Wolfowitz; Bayes solutions of sequential decision problems. *Ann. Math. Stat.*, **21**(1950), 82—99.