Title	An example of the optimal stopping rule problem : an approach to an adaptive decision process by dynamic programming
Sub Title	
Author	柳井, 浩(Yanai, Hiroshi)
Publisher	慶応義塾大学藤原記念工学部
Publication year	1965
Jtitle	Proceedings of the Fujihara Memorial Faculty of Engineering Keio University (慶応義塾大学藤原記念工学部研究報告). Vol.18, No.68 (1965.) ,p.9(9)- 24(24)
JaLC DOI	
Abstract	This article discusses some features of so-called optimal stopping rule problem by dynamic programming formulation. After introducing an example of the optimal stopping rule problem with finite horizon with known probabilities, the adaptive versions of the example will be discussed.
Notes	
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO50001004-00180068-0009

An Example of the Optimal Stopping Rule Problem

— An Approach to an Adaptive Decision Process by Dynamic Programming —

(Received June 2, 1965)

Hiroshi YANAI*

Abstract

This article discusses some features of so-called optimal stopping rule problem by dynamic programming formulation. After introducing an example of the optimal stopping rule problem with finite horizon with known probabilities, the adaptive versions of the example will be discussed.

I. Introduction

One of the most fundamental problems of the sequential decision process may be to make optimal decisions when to accept one of the successive proposals with the associated returns known in advance, so as to make the final gain as large as possible for the decision maker. In other words, it must be decided when to stop waiting for the favourable proposal. Thus, this kind of problem is sometimes called *the stopping rule problem.*

In some cases, (1°) it is known in advance what proposals will be made in future and in other cases, (2°) only the probabilities are known in advance. And there is also the case, (3°) even the probabilities are not known definitely but only *a priori* distributions of the probabilities are known and the probabilities become definite as the process proceeds.

The decision process associated with the case (1°) is called *the deterministic decision process*, (2°) —— *the stochastic decision process* and (3°) —— *the adaptive decision process*

There may not be remarkable difficulty in the *deterministic decision process* and will not be discussed in this article. An example of the decision process will be introduced and its stochastic version will be examined in II. The adaptive version of the example will be examined in III.

* 柳 井 浩, 助 手 Instructor at the Dep. of Administration Eng., Keio Uuiv.

II. The example and its stochastic version

Consider now a man waiting for a taxi-cab on a street to attend a weekly meeting, where he would have to pay penalty if he would be late. The penalty increases as he will be later as shown below.

Stage	Delay	Penalty
1	upto 0 min.	Q_1 (=0)
2	〃 5 〃	Q_2
⋮	⋮	⋮
N	〃 T 〃	Q_N

where

$$Q_1 < Q_2 < \cdots\cdots < Q_N$$

and the penalty for the delay in arrival over $5N$ minutes is so unreasonably that he would not be able to afford to pay.

His watch shows him that he would not be late, if he would take the first taxi-cab he would find. He knows that he would find empty taxi-cab at every 5 minutes thereafter. Let us denote by Q_i the penalty for him, if he would catch the i-th taxi-cab. There are two kinds of taxi-cab, A and B in the district. Although the times to reach the place of meeting are the same for the both kinds, the fee of the taxi-cab A is C_A, while the fee of the taxi-cab B is C_B ($>C_A$). Thus the total costs of taking the i-th taxi-cabs A and B are C_A+Q_i and C_B+Q_i respectively (cf. Fig. 1).

Let us assume, here, that the probabilities that the man would find the taxi-cabs A and B at every stage are p and $1-p$ and they are independent of what he has found before that stage.

The man wants to pay as little as possible. If he would find a taxi-cab A for the first, he would be so lucky that he would not have to pay any penalty while he pays the lowest fee. But in case the first taxi-cab is B, he must decide whether to pay no penalty paying a higher fee or to wait for the next taxi-cab in the hope that the latter would be a taxi-cab A, sacrifying the penalty. In general, there would be no difficulty if the i-th taxi-cab would be the first taxi-cab A —— he would catch it; while if it would be a taxi-cab B, he must decide whether to catch the taxi-cab B at that stage or to pass it over in the hope that he would be able to find *sooner* a taxi-cab A, sacrifying the increase of the penalty. Since, it is not always certain that a taxi-cab A would come sooner, he must decide so as to minimize the expected total cost he would pay.

Dynamic programming formulation

Let us now formulate the problem posed above by dynamic programming technique. Defining unknown functions

$f_i(p)$ = the minimum total cost expected at the stage i in case all the taxi-cabs arrived upto the stage i are of the kind B, $0 \leq p \leq 1$, $i=1, \cdots\cdots, N$, (2·1)

it would be clear that

$$f_N(p) = C_B + Q_N, \quad 0 \leq p \leq 1. \tag{2·2}$$

According to the principle of optimality [1]*), we have the following recurrence relation

$$f_i(p) = \min \begin{bmatrix} C_B + Q_i \\ p(C_A + Q_{i+1}) + (1-p) f_{i+1}(p) \end{bmatrix} \tag{2·3}$$

for $i=1, \cdots\cdots, N-1$.

Indeed, the optimal decision at the i-th stage would be to choose the way which will require him lower expectation of the cost he would have to pay. If he would catch the taxi-cab B before him, he would have to pay $C_B + Q_i$, while if he would pass it over and would be sure of taking the optimal decisions for $i+1$-st stage and thereafter, it would be expected that he would have to pay $p(C_A + Q_{i+1}) + (1-p) f_{i+1}(p)$. Thus if he would take the optimal decision at the i-th stage, too, the minimum expected total cost is represented by the relation (2·3).

It may be quite easy to solve the relations (2·3) beginning with $f_{N-1}(p)$ back to $f_1(p)$, given the values of p, C_A, C_B and Q_i. Moreover, in some cases, we may give a simple rule of optimal decision as shown in the following theorem.

Theorem II·1

If the difference

$$Q_{i+1} - Q_i \tag{2·4}$$

increases monotonously as to i and moreover if there exists such I that

$$Q_{I+1} - Q_I > p(C_B - C_A) > Q_I - Q_{I-1}, \tag{2·5}$$

then the optimal decision is to pass over the taxi-cabs B up to the $I-1$-st taxi-cab and to catch the taxi-cab whether it might be A or B thereafter. (Of course, there would be no difficulty if a taxi-cab A would come before the I-th stage.)

Proof

The theorem may be proved inductively from the N-th stage backward. For $i = N$, the only decision is to catch the taxi-cab B:

$$f_N(p) = C_B + Q_N, \quad \text{catch.} \tag{2·6}$$

Assume now for $i = N, N-1, \cdots\cdots, J > I$ that

$$f_i(p) = C_B + Q_i, \quad \text{catch.} \tag{2·7}$$

with this we have,

*) Numbers in brackets refer to the references cited at the end of the paper.

$$f_{J-1}(p)=C_B+Q_{J-1}, \quad \text{catch.} \tag{2·8}$$

In fact, since

$$\begin{aligned} p(C_A+Q_J)+(1-p)f_J(p) &= p(C_A+Q_J)+(1-p)(C_B+Q_J) \\ &= p(C_A-C_B)+C_B+Q_J>Q_{J-1}+C_B \end{aligned}$$

by assumptions (2·4) and (2·5), we obtain the relation (2·8) by the recurrence relation (2·3). Thus, the optimal decisions for $i=N, \cdots\cdots, I$ are to catch the taxi-cab before him even if it might be a taxi-cab B:

$$f_i(p)=C_B+Q_i, \quad \text{catch} \quad i=N, \cdots\cdots, I \tag{2·9}$$

Next, for $i=I-1$ we have

$$f_{I-1}(p)=p(C_A+Q_I)+(1-p)f_I(p), \quad \text{pass over.} \tag{2·10}$$

In fact, since

$$\begin{aligned} p(C_A+Q_I)+(1-p)f_I(p) &= p(C_A+Q_I)+(1-p)(C_B+Q_I) \\ &= p(C_A-C_B)+C_B+Q_I<Q_{I-1}+C_B \end{aligned}$$

by assumptions (2·4) and (2·5), we obtain the relation (2·10) by the recurrence relation (2·3).

Let us now assume that

$$f_i(p)=p(C_A+Q_{i+1})+(1-p)f_{i+1}(p), \quad \text{pass over} \tag{2·11}$$

for $i=I-1, \cdots\cdots, J+1$. With this, we have

$$f_J(p)=p(C_A+Q_{J+1})+(1-p)f_{J+1}(p), \quad \text{pass over} \tag{2·12}$$

In fact, since

$$\begin{aligned} p(C_A+Q_{J+1})+(1-p)f_{J+1}(p) &< p(C_A+Q_{J+1})+(1-p)(C_B+Q_{J+1}) \\ &= p(C_A-C_B)+Q_{J+1}+C_B<Q_J+C_B \end{aligned}$$

by assumptions (2·4) and (2·5), we see the relation (2·12) by the recurrence relation (2·3). Thus, the optimal decisions for $i=I-1, \cdots\cdots, 1$ are to pass over the taxi-cab if it is a taxi-cab B:

$$f_i(p)=p(C_A+Q_{i+1})+(1-p)f_{i+1}(p), \quad \text{pass over.}$$

Q. E. D.

Remark II·1

In applying the theorem above, it may be useful to draw the graph

$$\frac{Q_{i+1}-Q_i}{C_B-C_A}.$$

Remark II·2

It follows from the proof of the theorem above, that if there exists such I that

$$p(C_B-C_A)=Q_I-Q_{I-1}$$

instead of the inequality relationship (2·5), we have

$$f_I(p)=Q_I+C_B=p(C_A+Q_I)+(1+p)(C_B+Q_{I+1}),\text{ either catch or pass over.}$$

Corollary

If the difference

$$Q_{i+1}-Q_i=c \qquad i=1, \cdots\cdots, N-1,$$

where c is a positive constant, then the optimal policy is to catch taxi-cab B, if

$$c\geqslant p(C_B-C_A),$$

and pass over taxi-cab B except the N-th one if

$$c\leqslant p(C_B-C_A).$$

Numerical example

Let us give a numerical example, with the parametres:

$$\begin{aligned} &N=5\\ &C_A=¥\,150,\ C_B=¥\,190\\ &Q_1=¥\,0\\ &Q_2=¥\,10\\ &Q_3=¥\,30\\ &Q_4=¥\,60\\ &Q_5=¥\,100 \end{aligned} \qquad (2\cdot 14)$$

The total costs of taking the i-th taxi-cabs A and B are shown in Fig. 1.

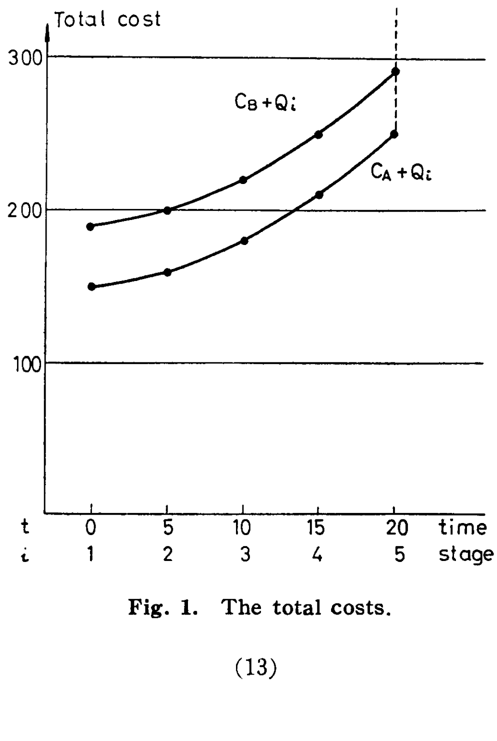

Fig. 1. The total costs.

It may be readily observed that we may apply the theorem above. Drawing the graph described in Remark II・1, we obtain the following chart of the optimal decisions.

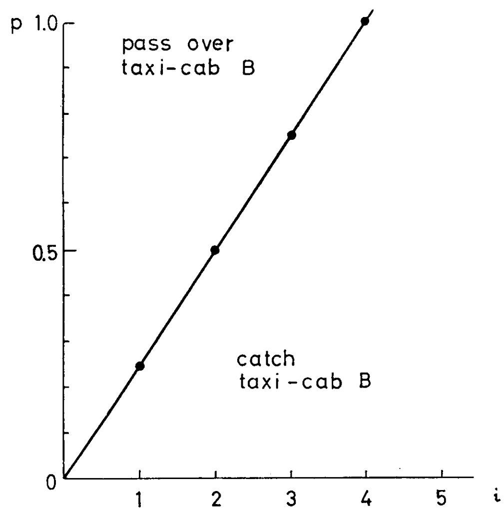

Fig. 2. Optimal decisions.

The minimal expected total costs $f_i(p)$ are shown in Fig. 3.

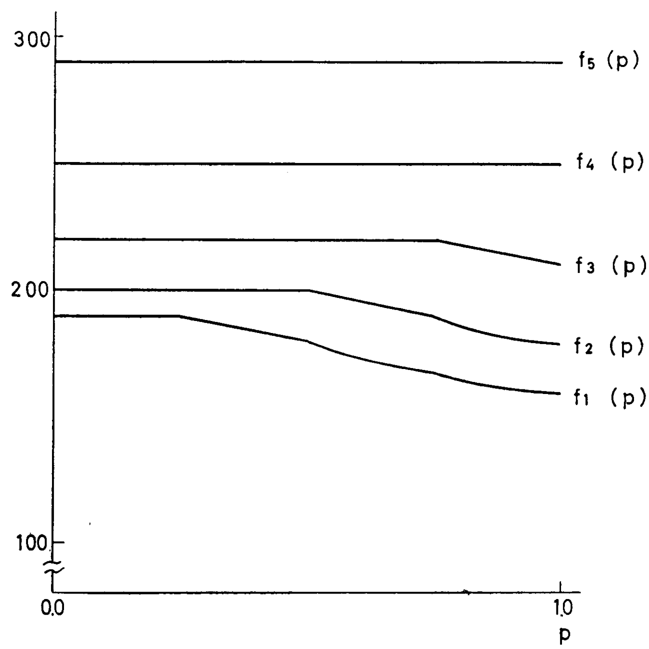

Fig. 3. Minimal expected total costs.

So far as we have considered, the differences between total costs associated with catching taxi-cabs A or B at the i-th stage were constant, $(C_B - C_A)$. We may circumvent this restriction as follows. Suppose that we are given the total costs associated with catching taxi-cabs A or B respectively as i-th stage by T_i^A and T_i^B, for $i=1, \cdots\cdots, N$ and the T_{N+1}^A and T_{N+1}^B is considered to be so large that the man would not be able to afford to pay. Moreover, assume that

$$T_i^A \leq T_i^B \quad \text{for} \quad i=1, \cdots\cdots, N. \tag{2·15}$$

In this case we have the recurrence relation

$$f_i(p) = \min \begin{bmatrix} T_i^B \\ p\,T_{i+1}^A + (1-p) f_{i+1}(p) \end{bmatrix} \quad \text{for} \quad i=1, \cdots\cdots, N \tag{2·16}$$

with the same definition for $f_i(p)$ as (2·3). And it would be clear that

$$f_N(p) = T_N^B. \tag{2·17}$$

In this case we have the following theorem. And it may be readily seen that Theorem II·1 follows from Theorem II·2.

Theorem II·2

If the difference

$$pT_i^A + (1-p)T_{i+1}^B - T_i^B \tag{2·18}$$

increases monotonously as to i and moreover if there exists such I that

$$pT_{I+1}^A + (1-p)T_{I+1}^B - T_I^B > 0 > pT_I^A + (1-p)T_I^B - T_{I-1}^B \tag{2·19}$$

then the optimal decision is to pass over the taxi-cabs B up to the I-1-st taxi-cab and to catch the taxi-cab before him whether it might be A or B thereafter. (Of course there would be no difficulty if a taxi-cab A would come before the I-th stage.)

Proof

The theorem may be proved inductively in an analogous way as the proof of Theorem II·1.

In fact, for $i=N$,

$$f_N(p) = T_N^B \text{ , catch.} \tag{2·20}$$

Let us assume that

$$f_i(p) = T_i^B \text{ , catch.} \tag{2·21}$$

for $i=N, N-1, \cdots\cdots, J>I$. Then we have for $i=J-1$,

$$f_{J-1}(p) = T_{J-1}^B \text{ , catch.} \tag{2·22}$$

In fact, since

$$pT_J^A + (1-p) f_J(p) = pT_J^A + (1-p) T_J^B > T_{J-1}^B$$

by the assumptions (2·18) and (2·19), we obtain the relation (2·22) by the recurrence relation (2·16). Thus, the optimal decision for $i=N, \cdots\cdots, I$ is to catch the taxi-cab before him even if it might be a taxi-cab B:

$$f_i(p)=T_i^B, \quad \text{catch} \quad \text{for } i=N, \cdots\cdots, I. \tag{2·23}$$

On the other hand, for $i=I-1$, we have.

$$f_{I-1}(p)=pT_{I-1}^A+(1-p)f_I(p)=pT_{I-1}^A+(1-p)T_I^B \quad \text{pass over.} \tag{2·24}$$

In fact, since

$$pT_{I-1}^A+(1-p)T_I^B<T_{I-1}^B,$$

(2·24) follows by the assumptions (2·18) and (2·19) and the recurrence relation (2·16).

Let us assume now that

$$f_i(p)=pT_i^A+(1-p)f_{i+1}(p) \quad \text{pass over} \tag{2·25}$$

for $i=I-1, \cdots\cdots, J+1$. Then, we may evaluate $f_J(p)$ as

$$f_J(p)=pT_J^A+(1-p)f_{J-1}^B(p) \quad \text{pass over.} \tag{2·26}$$

In fact, since

$$pT_J^A+(1-p)f_{J+1}^B(p)\leq pT_J^A+(1-p)T_{J+1}^B<T_J^B$$

(2·26) follows by the assumptions (2·18) and (2·19) and the recurrence relation (2·16). Thus, the optimal decisions for $i=I-1, \cdots\cdots, 1$ are to pass over the taxi-cab before him if it is a taxi-cab B:

$$f_i(p)=pT_i^A+(1-p)T_{i+1}^B, \quad \text{pass over} \tag{2·27}$$

Q. E. D.

III. Adaptive version of the example

So far, we have considered the situation where the probabilities p and $1-p$ that the man would find the taxi-cab A and B at every stage are known definitely in advance. Using different words, this means, in a sense, that at least somebody has observed sufficiently many times which kinds of the taxi-cabs came along the street. But it may not be always the case. In most cases, he only has a vague notion about the probabilities at the beginning of the decision process. Even with this poor knowledge, the man must make his decision anyhow. As he proceeds on to make decisions observing which one comes, he enlarges his experience, and his vague notion becomes more definite, and he would be able to make more appropriate decisions.

This type of decision process is called the adaptive decision process. In this section we would like to examine the adaptive version of the example of decision process posed in the previous section.

As the vague notion of the probabilities, we assume that he knows the a priori distribution of the probability p :

$$dH_0(p) \tag{3·1}$$

before the man begins with the decision process. The 'vague notion' given in (3·1) would become clearer as he observes the kinds of taxi-cabs coming along the street. In other words, the vague notion given by the a priori distribution (3·1) will be changed into another a priori distribution :

$$dH_0(p) \to dH'(p)$$

after he observes a taxi-cab. And the a priori distribution $dH'(p)$ will be again changed into another a priori distribution after he observes the second taxi-cab, and so on.

Let us look for a moment how this will be done, before we consider the associated decision process. Assume that we have the a priori distribution

$$dH_i(p) \tag{3·2}$$

just before we observe the i-th taxi-cab. According to the Theorem of Bayes [2], the a priori distribution is changed into

$$dH_{i+1}(p) = \frac{(1-p)dH_i(p)}{\int (1-p)dH_i(p)} \tag{3·3}$$

after a taxi-cab B has been observed for the i-th taxi-cab. If he would observe a 'A' for the $i+1$-st taxi-cab, the new a priori distribution would be

$$dH'_{i+1}(p) = \frac{p\,dH_i(p)}{\int p\,dH_i(p)}\,. \tag{3·4}$$

But the latter will not be empoyed later, since, in that case, his decision would be just to catch the taxi-cab A.

As an example, if we assume that the a priori distribution belongs to β–family :

$$dH_i(p) = \frac{1}{B(m,n)} p^{m-i}(1-p)^{n-1}dp \tag{3·5}$$

$$0 \leq p \leq 1,\ m, n > 0$$

where

$$B(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1}\,dx \tag{3·6}$$

with the expectation

$$\bar{p}_i = \int_0^1 p\,dH_i(p) = \frac{m}{m+n}, \tag{3·7}$$

then the new a priori eistribution

$$dH_{i+1}(p)=\frac{(1-p)\,dH_i(p)}{\int(1-p)\,dH_i(p)}$$

$$=\frac{1}{B(m,n+1)}p^{m-1}(1-p)^n\,dp \quad 0\leq p\leq 1$$

again belongs to β-family with the parametres $m, n+1$ and the expectation

$$\bar{p}_{i+1}=\frac{m}{m+n+1}\,. \qquad (3\cdot 9)$$

Moreover, in the case of integral parametres m, n, we may interprete as if m taxi-cabs A and n taxi-cabs B were observed up to the i-th stage, and the expectation to be the direct average of them.

Fig. 4 shows an example how an a priori distribution (β ($m=4$, $n=1$)) is changed after repeated observations of taxi-cabs B. The expectations $\bar{p}_i$ are shown by upward arrows. It may be observed that the modes of the distributions move to the left and at the same time they become sharper as the repeated observations of the taxi-cabs B. (Remark that the variance of the distribution $\beta(m, n)$ is $\frac{mn}{(m+n+1)(m+n)}$.)

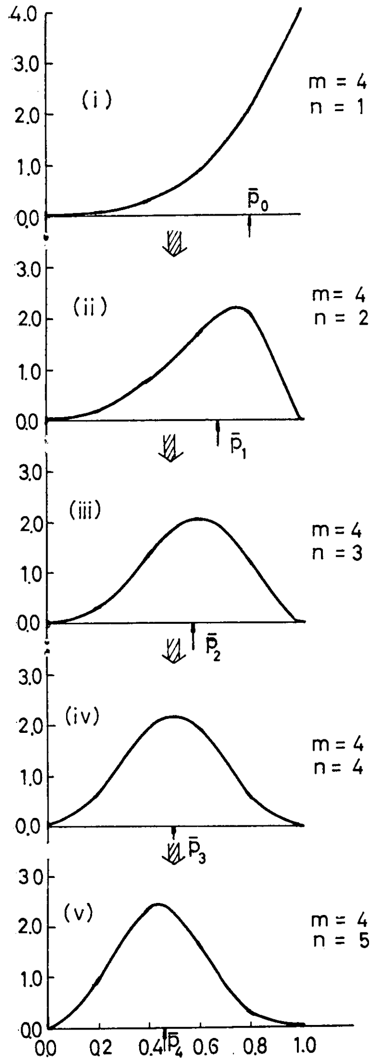

Fig. 4. Change of a priori distribution $\beta(4, 1)$.

Let us now turn to consider the optimal decisions under the circumstance described above. There may be various approaches to this problem. Two of them will be described below.

1° Parametric method

If every a priori distribution under consideration belongs to a class, we may reformulate the problem as an optimal decision process defined on the parametre space corresponding to the class. Let us examine the case of β-distributions.

We may apply the method of dynamic programming to the formulation on the

m-n space Let us define the unknown function as follows.

$f_i(m, n)$: the minimum total cost expected at the i-th stage after finding a taxi-cab B for the i-th taxi-cab making an optimal decision using the a priori distribution $\beta\ (p\,;\ m, n)$ obtained from the a priori distribution $\beta\ (p\,;\ m, n-1)$ we had before the i-th taxi-cab B was observed. (3·10)

It would be clear that

$$f_N(m, n) = C_B + Q_N. \tag{3·11}$$

According to the principle of optimality, we have the following recurrence relation for $i=1, \cdots\cdots, N-1$

$$f_i(m, n) = \min \begin{bmatrix} C_B + Q_i \\ \dfrac{m}{m+n}(C_A + Q_{i+1}) + \dfrac{n}{m+n} f_{i+1}(m, n+1) \end{bmatrix} \tag{3·12}$$

The representation in the second line in the right-hand-side is the total cost expected at the i-th stage after passing over the i-th taxi-cab B and using the optimal decisions thereafter, that is

$$\int_0^1 [p(C_A + Q_{i+1}) + (1-p) f_{i+1}(m, n+1)]\, dH_i(p)$$

$$= \frac{m}{m+n}(C_A + Q_{i+1}) + \frac{n}{m+n} f_{i+1}(m, n+1) \tag{3·13}$$

where $dH_i(p)$ is the a priori distribution

$$\beta(p\,;\ m, n)$$

we have after observing a taxi-cab B at the i-th stage.

Given the values of C_A, C_B and Q_i we may evaluate the functions $f_i(m, n)$, from $f_{N-1}(m, n)$ back to $f_1(m, n)$ and give the optimal decision for every m, n and i. In some cases, moreover, we may give a simple rule of optimal decisions as shown in the following theorem which corresponds to Theorem II·1.

Theorem III·1

If the difference

$$Q_{i+1} - Q_i \tag{3·14}$$

increases monotonously as to i, the optimal policy is to catch the i-th taxi-cab B if

$$\frac{m}{m+n} \leqslant \frac{Q_{i+1} - Q_i}{C_B - C_A} \tag{3·15}$$

and to pass over the i-th taxi-cab B if

$$\frac{m}{m+n} \geqslant \frac{Q_{i+1}-Q_i}{C_B-C_A}\,. \tag{3·16}$$

Proof

Since this theorem follows from Theorem III·2 the proof will not given here.

Numerical example

Let us now give a numerical example of the theorem. The parametres excluding the value of p are the same as those given in the numerical example in the previous section. The chart giving optimal decisions are drawn on the $m-n$ plane. (Fig. 5), in which an observation of a taxi-cab B corresponds to the unit upward movement of the point representing the distribution of p.

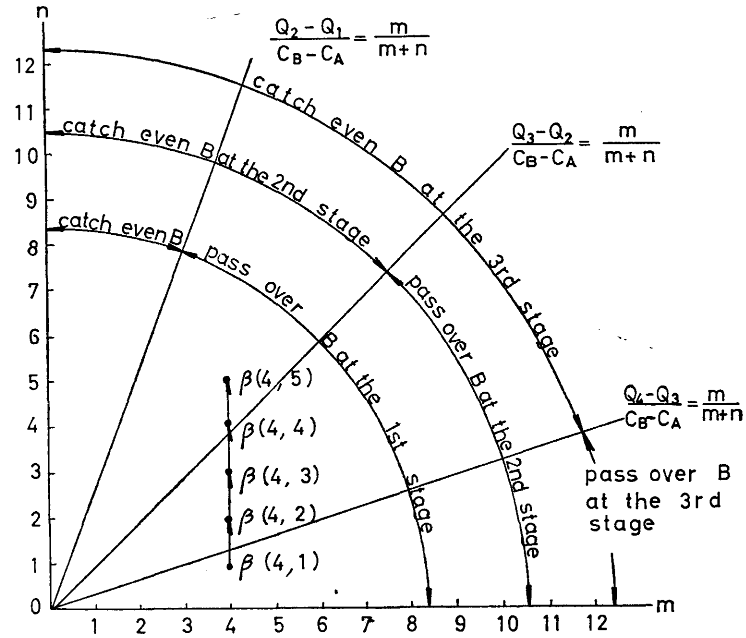

Fig. 5. Optimal decisions on $m-n$ plane.

Remark III·1

If we give the total costs associated with catching taxi-cabs A or B at the i-th stage by T_i^A and T_i^B, respectively for $i=1, \cdots\cdots, N$ instead of C_A+Q_i and C_B+Q_i, the corresponding recurrence relation is

$$f_i(m, n) = \min \begin{bmatrix} T_i^B \\ \frac{m}{m+n} T_{i+1}^A + \frac{n}{m+n} f_{i+1}(m, n+1) \end{bmatrix} \tag{3·17}$$

as it might be readily seen. Under the assumption as to the total cost at the $N+1$ st stage, we see

$$f_N(m, n) = T_N^B . \tag{3·18}$$

Now, we have the following theorem, which takes more general form than Theorem III·1.

Terorem III·2

If the difference

$$\frac{m}{m+n} T_{i+1}^A + \frac{n}{m+n} T_{i+1}^B - T_i^B \tag{3·19}$$

increases monotonously as to i, the optimal policy is to catch the i-th taxi-cabs B if

$$\frac{m}{m+n} T_{i+1}^A + \frac{n}{m+n} T_{i+1}^B \geqslant T_i^B \tag{3·20}$$

and to pass over the i-th taxi-cab B if

$$\frac{m}{m+n} T_{i+1}^A + \frac{n}{m+n} T_{i+1}^B \leqslant T_i^B . \tag{3·21}$$

Proof

This theorem may be proved inductively from the N-th stage backward in an analogous way as Theorem II·2.

In fact, without loss of generality, we assume that there exists a positive integer I for which

$$\frac{m}{m+n} T_{I+1}^A + \frac{n}{m+n} T_{I+1}^B - T_I^B > 0 > \frac{m}{m+n} T_I^A + \frac{n}{m+n} T_I^B - T_{I-1}^B .$$

For $i=N$, clearly

$$f_N(m, n) = T_N^B .$$

Let us assume that

$$f_i(m, n) = T_i^B \quad \text{catch}$$

for $i=N, N-1, \cdots\cdots, J>I$. Then we have for $i=J-1$

$$f_{J-1}(m, n) = T_{J-1}^B \quad \text{catch,}$$

since

$$\frac{m}{m+n} T_J^A + \frac{n}{m+n} f_J(m, n+1) = \frac{m}{m+n} T_J^A + \frac{n}{m+n} T_J^B > T_{J-1}^B .$$

Thus the optimal decision for $i=N, \cdots\cdots, I$ is to catch the taxi-cab before him even if it might be a taxi-cab B:

$$f_i(m, n) = T_i^B, \quad \text{catch} \quad \text{for } i=N, \cdots\cdots, I.$$

On the other hand, for $i=I-1$, we have

$$f_{I-1}(m, n)=\frac{m}{m+n}T_{I-1}^{A}+\frac{n}{m+n}f_{I}(m, n+1)$$

$$=\frac{m}{m+n}T_{I-1}^{A}+\frac{n}{m+n}T_{I}^{B}, \text{ pass over.}$$

In fact

$$\frac{m}{m+n}T_{I-1}^{A}+\frac{n}{m+n}T_{I}^{B}<T_{I-1}^{B}$$

by the assumptions above.

Let us assume now that

$$f_{i}(m, n)=\frac{m}{m+n}T_{i}^{A}+\frac{n}{m+n}f_{i+1}(m, n+1), \text{ pass over}$$

for $i=I-1, \cdots\cdots, J+1$. Then, we have for $i=J$

$$f_{J}(m, n)=\frac{m}{m+n}T_{J}^{A}+\frac{n}{m+n}f_{J+1}(m, n+1), \text{ pass over.}$$

In fact,

$$\frac{m}{m+n}T_{J}^{A}+\frac{n}{m+n}f_{J+1}(m, n+1)\leqslant\frac{m}{m+n}T_{J}^{A}+\frac{n}{m+n}T_{J+1}^{B}<T_{J}^{B}$$

Thus the optimal policies are to pass taxi-cabs B upto $I-1$-st stage and to catch thereafter.

Q. E. D.

2° Non-parametric method

If we cannot assume a class of distributions defined by a finite number of parametres we have to define the unknown functions (or functionals) on some infinite-dimensional spaces or on some functional spaces. It is not possible, at least in practical sense, to treat recurrence relations as to such functions or functionals.

Thus, in this case, we must invent some appoximate approach to the problem. One method may be to use approximation of the a priori distributions by those which belong to a class decribed above like the β-family.

Another method may be to use the recurrence relation

$$f_{i}(p)=\min\left[\begin{array}{l}C_{B}+Q_{i}\\ \int_{0}^{1}[p(C_{A}+Q_{i+1})+(1-p)f_{i+1}(p)]\,dH_{i+1}(p).\end{array}\right. \tag{3・22}$$

where $dH_{i+1}(p)$ is the a priori distribution obtained from the a priori distribution $dH_{i}(p)$ after finding a taxi-cab B at the i-th stage by

$$dH_{i+1}(p)=\frac{(1-p)\,dH_{i}(p)}{\int_{0}^{1}(1-p)\,dH_{i}(p)} \tag{3・23}$$

And a further approximation

$$\int_0^1 [(C_A-Q_{i+1})+(1-p)f_{i+1}(p)]\,dH_{i+1}(p)$$

$$\simeq \bar{p}_{i+1}(C_A+Q_{i+1})+(1-\bar{p}_{i+1})f_{i+1}(\bar{p}_{i+1})$$

where

$$\bar{p}_{i+1}=\int_0^1 p\,dH_{i+1}(p)$$

leads to the simplest idea to approach the problem: to enter the chart of optimal decisions such like given in Fig. 2 with the value $p=\bar{p}_i$ which is the expectation of the probability of finding a taxi-cab A at the i-th stage. Some examples are given in Fig. 6 with the same paramtres as the previous example.

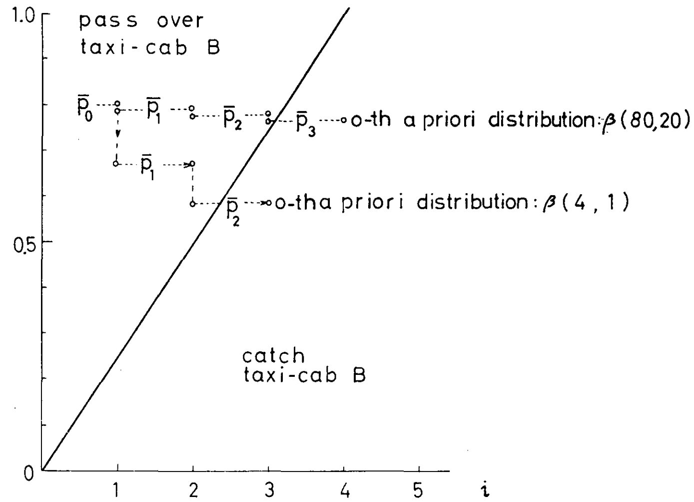

Fig. 6. Optimal decisions by simple approximate method.

If we take, for the sake of comparison, a β-distribution as the a priori distribution, the chart of optimal decisions ın Fig. 6 may be transformed into a chart on $m-n$ space which is exactly the same as Fig. 5. Although we have arrived at the same result via two different approaches for this particular example, it is not yet known the relationship between these two approaches. But, this might be an illustration for the second approach to be of some practical use.

Acknowledgement

The author wishes to show his acknowledgment to Associate. Profs. T. Washio and T. Sekine of Keio Univ. who gave the author many valuable suggestions.

References

[1] Bellman, R. E. and S. E. Dreyfus, Applied Dynamic Programming, Princeton, 1962.

[2] Гнеденко, Б. В., Курс Теории Вероятностей, Ф М. Москва 1961.