No.637

A STOCHASTIC SEQUENTIAL ALLOCATION PROBLEM WHERE THE RESOURCES CAN BE REPLENISHED

by

Masahiro Sato

August 1995

A STOCHASTIC SEQUENTIAL ALLOCATION PROBLEM WHERE THE RESOURCES CAN BE REPLENISHED

Masahiro Sato

University of Tsukuba

1995.8.2

Abstract Suppose a hunter starts hunting over certain given t periods with i bullets in hand. A distribution of the value of each appearing target and the hitting probability of a bullet are known. For shooting, he takes a strategy of *shoot-look-shoot* scheme, implying that if a bullet just fired does not hit the target, then the hunter must decide whether or not to shoot an additional one. At the end of each period, it is allowed to replenish a given number of bullets by paying a certain cost. The objective here is to examine the properties of the optimal policy which maximizes the total expected net reward. We get the following results: the optimal policy for shooting is monotone in the number of bullets in hand if it is always optimal either to replenish a certain number of bullets every period or not to replenish them at all; if only one bullet can be replenished per period, then both the optimal policies for shooting and replenishment are monotone in the number of bullets in hand; if more than one can be replenished per period, then there exist examples where the optimal policies for shooting are not monotone in the number of remaining bullets.

1. Introduction

Consider a problem of allocating countable resources to investment opportunities appearing one by one over a given planning horizon. At the beginning of each period, an opportunity comes with a certain value which is a random sample from a known probability distribution. Assume the resources are allocated to the opportunities pursuant to *shoot-look-shoot* policy, implying that, if investing one unit of resource results in unsuccessful, then it is decided whether or not to invest one more at once. At the end of each period, the resources can be replenished by paying a certain cost; it must be decided whether or not to replenish m units of resources then. The aim is to maximize the total expected reward obtained from the successful opportunities minus the total cost for replenishment.

In general, there exist two kinds of policies in sequential allocation problems: *shoot-look-shoot* policy [2,3,6,7] and *volley* [1,3,4,5,6]. In volley policy, it must be decided how much resources to invest in salvo. Mastran and Thomas [3] treat the problem as a target attacking one in which the computational method to obtain the optimal decision rules for the both policies are showed. Kisi [2] considers a model of shoot-look-shoot policy and examines the relation between the approximate solution and the exact. Sakaguchi [8] investigates the continuous-time version of [3]. Namekata, Tabata and Nishida [4] deal with a model of volley policy where there exist two kinds of targets in a sense that the necessary number of resources to get them are different. They also examine problems with volley policy in [5] and [6]. In [5], it is discussed how to allocate perishable resources, and in [6], a case with a random planning horizon is investigated. Derman, Lieberman and Ross [1], and Prastacos [7] deal with the problems as investment ones with volley policy. In [9], a problem with shoot-look-shoot policy, in which the search cost must be payed to find an investment opportunity, is discussed, and it is derived that the critical value, at which investing or not become indifferent in the optimal decision, is not always decreasing[†] in the number of remaining resources.

In models such as stated above, if all of resources are spent before the deadline, then later chances, which may be more attractive, will be unavailable. However, if the resources can be replenished by paying a certain cost, then he can continue investing activities in order to gain the total expected reward. In this paper, we discuss the problem where such replenishment is assumed.

In the following section, we exactly define our model and formulate fundamental equations. In Section 3, properties of the optimal policy are derived. In Section 4 and 5 that follow, a cases that it is optimal to replenish the resources every period and a case that not to replenish at all are investigated. The case that only one unit can be replenished per period is considered in Section 6, and a case for more than one bullet is examined and some numerical examples are shown in Section 7. The conclusion obtained are summarized in Section 8.

2. Model and Fundamental Equations

Now using the following hunting problem, we shall explain the model treated in this paper. Suppose a hunter starts hunting over a given planning horizon t with i bullets in hand. At the beginning of each period, he goes to hunt and can find only one target. The case that he cannot find any target is regarded as that he finds a target of value 0. The value of a target, w, is a random variable having a known probability distribution function $F(w)$ with a finite expectation μ, continuous or discrete where $F(w) = 0$ for $w < 0$, $F(w) < 1$ for $w < 1$, and $F(w) = 1$ for $1 \leq w$. The distribution does not concentrate on only a point, *i.e.*, $\Pr(w) < 1$ for any w. The values of successive targets are assumed to be stochastically independent.

He observes the value of a target as soon as finding it and has to immediately decide whether or not to shoot. If the value is rather small, then he may decide not to shoot and come home with no profit. Suppose the value is favorable and he decides to shoot a bullet. Then the bullet will hit the target with hitting probability q. If not hitting, then two cases are further possible: either the target disappears immediately with escaping probability r or still remains without any defense. If it stands still there, then he has to decide whether or not to fire an additional bullet. Assume that repeated firings waste no time. If he decides not to shoot any more, exactly speaking, if he decides not to shoot any more at the present target, need not shoot (get it), or cannot shoot (it escapes

[†]Throughout this paper, the following terms are used in order to avoid the expressions of double negatives; "increasing (decreasing)" means "nondecreasing (nonincreasing)".

or $i = 0$), then he comes home. On his way home, he must furthermore decide whether or not to replenish m bullets by paying a cost a; it is not permitted to supply more or less than m bullets. Thus, the period ends and the next comes. The objective is to maximize the total expected net reward over t periods. The flow of the decision process is illustrated in Figure 1.

Now we shall formulate the fundamental equations of the model. Let points of time be numbered backward from the final point of the planning horizon as 0, 1, and so on; an interval between time t and time $t-1$ is called period t. We define $u_t(i, w)$ to be the maximum of the total expected net reward starting from time t when i bullets are in hand and a target of value w is found, and $v_t(i)$ to be its expectation in terms of w, that is;

$$v_t(i) = \int_0^1 u_t(i, \xi) dF(\xi), \quad t \geq 0. \tag{2.1}$$

Furthermore, $z_t(i)$ is defined as the maximum of the total expected net reward starting from time t

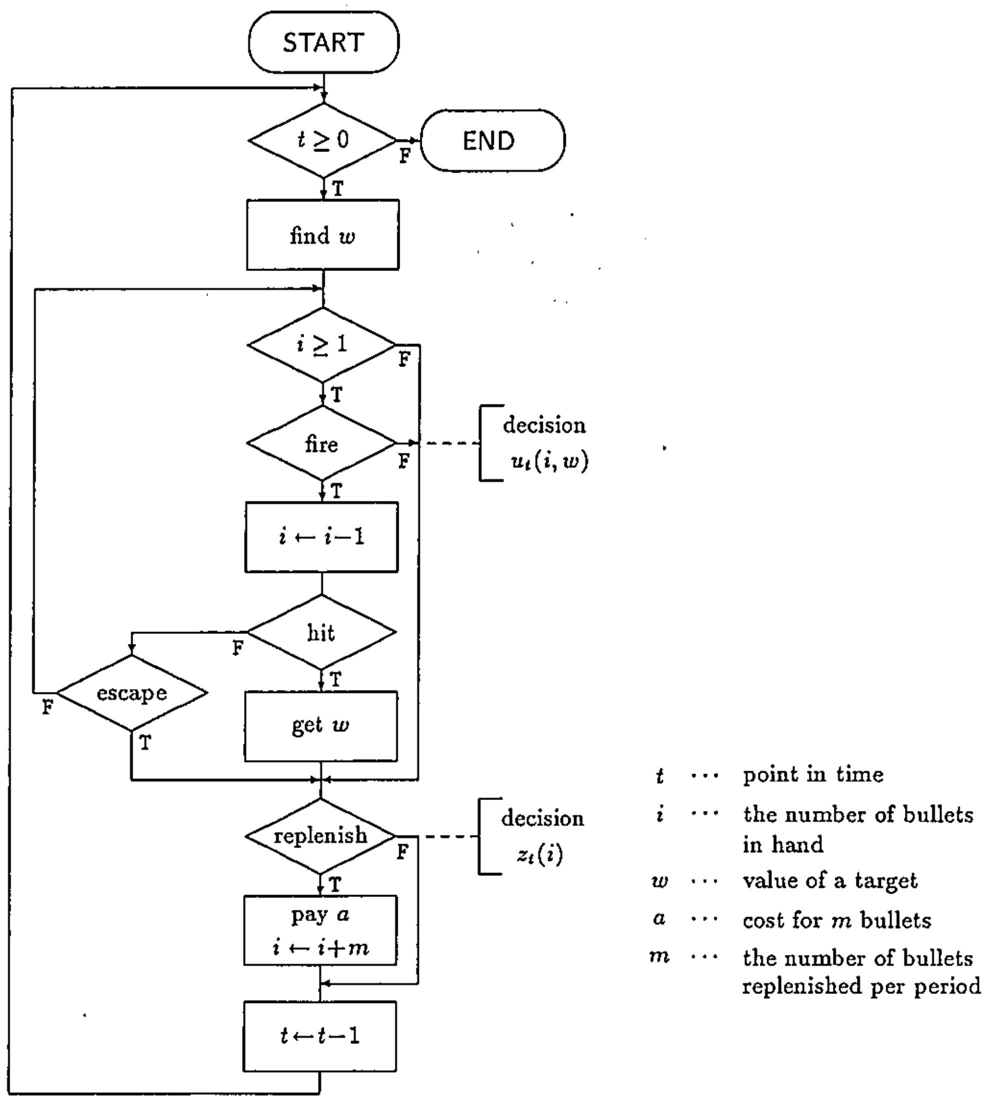

Figure 1 Flowchart of the Decision Process

when he decide not to shoot at the present target any more, provided that i bullets remain. Then we have the following relations:

$$u_t(i,w) = \max\{z_t(i),\ q(w+z_t(i-1)) + (1-q)(rz_t(i-1) + (1-r)u_t(i-1,w))\}$$
$$= \max\{z_t(i),\ pu_t(i-1,w) + qw + (1-p)z_t(i-1)\}, \quad t \geq 0,\ i \geq 1, \tag{2.2}$$
$$u_t(0,w) = v_t(0) = z_t(0), \quad t \geq 0, \tag{2.3}$$
$$z_t(i) = \max\{\beta v_{t-1}(i),\ \beta v_{t-1}(i+m) - a\}, \quad t \geq 1, i \geq 0 \tag{2.4}$$

where $p = (1-q)(1-r) \in [0,1)$ and $\beta \in (0,1]$, a discount factor. The first (second) term inside the braces in the right hand side of (2.2) represents the maximum of the total expected reward when it is decided not to shoot at the present target (to shoot at the present target), and the first (second) term inside the braces in the right hand side of (2.4) denotes the maximum of the total expected reward when it is optimal not to replenish m bullets (to replenish m bullets). Further, we immediately have the following final conditions:

$$u_0(i,w) = qw + pu_0(i-1,w) = \frac{1-p^i}{1-p}qw, \quad i \geq 1, \tag{2.5}$$
$$v_0(i) = \frac{1-p^i}{1-p}q\mu, \quad i \geq 1, \tag{2.6}$$
$$z_0(i) = 0, \quad i \geq 0. \tag{2.7}$$

Here (2.5), hence also (2.6), hold for $i \geq 0$.

Below, we examine the properties of fundamental equations.

Lemma 1.

(a) *$u_t(i,w)$, $v_t(i)$ and $z_t(i)$ are increasing in t for any i and w.*

(b) *If $p > 0$, then $u_t(i,w)$, $v_t(i)$ and $z_t(i)$ are strictly increasing in i for any t and w except $z_0(i)$ and $u_0(i,0)$. If $p = 0$, then they are increasing in i for any t and w.*

(c) *$u_t(i+1,w) - u_t(i,w) \leq q$ for any t, i and w where the equal sign holds only for $i = 0$ and $w = 1$. In addition, $v_t(i+1) - v_t(i) < q$ and $z_t(i+1) - z_t(i) < q$ also hold for any t and i.*

(d) *$u_t(i,w)$ is increasing in w for any t and i.*

Proof: (a) Since $v_0(i) = (1-p^i)q\mu/(1-p) \geq 0$ for any i, it follows that $z_1(i) \geq \beta v_0(i) \geq 0 = z_0(i)$. Using (2.2), we get for any i

$$u_1(i,w) \geq pu_1(i-1,w) + qw + (1-p)z_1(i-1) \geq pu_1(i-1,w) + qw. \tag{2.8}$$

Therefore we have

$$u_1(i,w) \geq p(pu_1(i-2,w) + qw) + qw$$
$$\vdots$$

$$\geq p^i u_1(0,w) + \frac{1-p^i}{1-p} qw$$
$$\geq \frac{1-p^i}{1-p} qw = u_0(i,w), \tag{2.9}$$

leading to $v_1(i) \geq v_0(i)$ for any i. Assume $v_t(i) \geq v_{t-1}(i)$ for $i \geq 0$ as the first inductive assumption in terms of t. Then we have for $i \geq 0$

$$z_{t+1}(i) = \max\{\beta v_t(i),\ \beta v_t(i+m) - a\}$$
$$\geq \max\{\beta v_{t-1}(i),\ \beta v_{t-1}(i+m) - a\} = z_t(i). \tag{2.10}$$

Accordingly we get $u_{t+1}(0,w) = z_{t+1}(0) \geq z_t(0) = u_t(0,w)$. Furthermore, suppose $u_{t+1}(i-1,w) \geq u_t(i-1,w)$ for any w as the second inductive assumption in terms of i. Then the following can be obtained;

$$u_{t+1}(i,w) = \max\{z_{t+1}(i),\ pu_{t+1}(i-1,w) + qw + (1-p)z_{t+1}(i-1)\}$$
$$\geq \max\{z_t(i),\ pu_t(i-1,w) + qw + (1-p)z_t(i-1)\} = u_t(i,w), \tag{2.11}$$

which yields $v_{t+1}(i) \geq v_t(i)$. Thus, it is proven by double induction that $u_t(i,w)$ is increasing in t for any i and w, so also are $v_t(i)$ and $z_t(i)$ for any i.

(b) When $p > 0$, it is obvious from (2.5) that $u_0(i,w)$ is strictly increasing in i, so also is $v_0(i)$. Let $v_{t-1}(i)$ be strictly increasing in i. Then, we get for any i

$$z_t(i) = \max\{\beta v_{t-1}(i), \beta v_{t-1}(i+m) - a\}$$
$$> \max\{\beta v_{t-1}(i-1), \beta v_{t-1}(i+m-1) - a\} = z_t(i-1), \tag{2.12}$$

which yields

$$u_t(1,w) \geq z_t(1) > z_t(0) = u_t(0,w). \tag{2.13}$$

Furthermore, assuming $u_t(i,w) > u_t(i-1,w)$ for $w > 0$, we have $u_t(i+1,w) > u_t(i,w)$ from (2.2), so $v_t(i+1) > v_t(i)$. Thus, it is proven by double induction that $u_t(i,w)$, $v_t(i)$ and $z_t(i)$ are strictly increasing in i for any t and w except $u_0(i,0)$ and $z_0(i)$.

For $p = 0$, the proof is almost the same as above.

(c) It is clear that $z_0(i+1) - z_0(i) = 0 < q$. From (2.5), it follows that

$$u_0(i+1,w) - u_0(i,w) = p^i qw \leq q \tag{2.14}$$

where the equal sign holds only for $i = 0$ and $w = 1$. From this, we get $v_0(i+1) - v_0(i) < q$ for any i. Now suppose $z_t(i+1) - z_t(i) < q$ for any i. Then, we easily obtain from (2.13)

$$u_t(1,w) - u_t(0,w) = u_t(1,w) - z_t(0) = \max\{z_t(1) - z_t(0),\ qw\} \leq q \tag{2.15}$$

where the equal sign holds only for $w = 1$. Furthermore, assume $u_t(i,w) - u_t(i-1,w) \le q$. Then we have

$$u_t(i+1,w) - u_t(i,w) \le \max\{z_t(i+1) - z_t(i),$$
$$p(u_t(i,w) - u_t(i-1,w)) + (1-p)(z_t(i) - z_t(i-1))\} < q \tag{2.16}$$

using the general formula

$$\max_{0\le i\le k} a_i - \max_{0\le i\le k} b_i \le \max_{0\le i\le k}(a_i - b_i). \tag{2.17}$$

Thus, we get $u_t(i+1,w) - u_t(i,w) < q$ for any i and w except for $i = 1$ and $w = 1$, which yields $v_t(i+1) - v_t(i) < q$ for any i. From above, it follows that

$$z_{t+1}(i+1) - z_{t+1}(i) \le \beta \max\{v_t(i+1) - v_t(i),\ v_t(i+1+m) - v_t(i+m)\}$$
$$< \beta q \le q. \tag{2.18}$$

By double induction, we come to the statement.

(d) It is easily proven by induction. ∎

Using these properties, in the next section, we shall discuss the structure of the optimal decision policy.

3. Properties of Optimal Policy

Now define $g_t(i,w)$ and $\phi_t(i)$ as follows:

$$g_t(i,w) = pu_t(i-1,w) + qw + (1-p)z_t(i-1) - z_t(i), \quad i \ge 1,\ t \ge 0, \tag{3.1}$$

$$\phi_t(i) = \beta(v_{t-1}(i+m) - v_{t-1}(i)) - a, \quad i \ge 0,\ t \ge 1. \tag{3.2}$$

Then, the lemma below holds true.

Lemma 2.

(a) *For $t \ge 1$ and $i \ge 1$, $g_t(i,w)$ is strictly increasing in w, which is also true for $t \to \infty$.*

(b) *$g_t(i,w) = 0$ has a unique solution $w = h_t(i) \in (0,1)$ for $p > 0$ ($\in [0,1)$ for $p = 0$).*

Proof: (a) It is immediate from Lemma 1(d).

(b) Assume $p > 0$. It can be easily proven by induction that $u_t(i,0) = z_t(i)$ for any t and i. Accordingly we get

$$g_t(i,0) = pu_t(i-1,0) + (1-p)z_t(i-1) - z_t(i) = z_t(i-1) - z_t(i) < 0 \tag{3.3}$$

from Lemma 1(b). In addition, it is obvious from Lemma 1(c) that

$$g_t(i,1) = pu_t(i-1,1) + q + (1-p)z_t(i-1) - z_t(i) \ge q + z_t(i-1) - z_t(i) > 0 \tag{3.4}$$

for $i \ge 1$ and $t \ge 0$. From (3.3), (3.4) and the continuity of $g_t(i,w)$ in w, it follows that $g_t(i,w) = 0$ has a unique solution $h_t(i) \in (0,1)$ for $p > 0$. For $p = 0$, the proof is almost the same as above. ∎

Remark: We call $h_t(i)$ a critical value when the hunter has i bullets and t periods remain. From Lemma 2, the optimal decision policy for shoot becomes as follows; if $g_t(i,w) \geq 0$ ($w \geq h_t(i)$), then fire, or else don't fire. The optimal policy for replenishment becomes as follows; if $\phi_t(i) \geq 0$, then replenish m bullets, or else don't replenish them.

Because $g_t(i, h_t(i)) = 0$, it follows that

$$0 = g_t(i, h_t(i)) = pu_t(i-1, h_t(i)) + qh_t(i) + (1-p)z_t(i-1) - z_t(i)$$
$$\geq qh_t(i) + z_t(i-1) - z_t(i), \tag{3.5}$$

from which we get

$$h_t(i) \leq (z_t(i) - z_t(i-1))/q. \tag{3.6}$$

In particular for $i = 1$, it is true from $g_t(1, h_t(1)) = 0$ that

$$h_t(1) = (z_t(1) - z_t(0))/q. \tag{3.7}$$

The following lemma describes the relation between $h_t(i)$ and $z_t(j)$ more detailed.

Lemma 3.

(a) *If $p > 0$, then for $i \geq 1$ and $t \geq 1$,*

$$h_t(i) \geq (<)\, h_t(i+1) \iff h_t(i+1) = (<)\, (z_t(i+1) - z_t(i))/q.$$

When $p = 0$, it always holds true for $i \geq 1$ and $t \geq 1$ that $h_t(i+1) = (z_t(i+1) - z_t(i))/q$.

(b) *For $t \geq 1$ and $p > 0$,*

$$h_t(1) \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} h_t(2) \iff 2z_t(1) - z_t(0) - z_t(2) \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} 0.$$

(c) *For given $i \geq 1$ and $t \geq 1$, if $h_t(i) > (=)\, h_t(i+1)$, then $2z_t(i) - z_t(i-1) - z_t(i+1) > (\geq) 0$.*

(d) *For given $i \geq 1$ and $t \geq 1$, if $2z_t(i) - z_t(i-1) - z_t(i+1) < 0$, then $h_t(i) < h_t(i+1)$.*

(e) *Assume $h_t(i) = (z_t(i) - z_t(i-1))/q$ for given $i \geq 1$ and $t \geq 1$. Then*

$$2z_t(i) - z_t(i+1) - z_t(i-1) > (=)\, 0 \implies h_t(i) > (\geq) h_t(i+1).$$

Proof: (a) It is immediate from (3.6) and Lemma 2.

(b) Using (3.7), we have

$$g_t(2, h_t(1)) = pu_t(1, h_t(1)) + qh_t(1) + (1-p)z_t(1) - z_t(2) = 2z_t(1) - z_t(0) - z_t(2), \tag{3.8}$$

which yields the statement.

(c) From the assumption and (3.6), we have

$$0 < (=)\, g_t(i+1, h_t(i)) = qh_t(i) + z_t(i) - z_t(i+1) \leq 2z_t(i) - z_t(i-1) - z_t(i+1). \tag{3.9}$$

(d) The statement is the contraposition of (c).

(e) Because $qh_t(i) = z_t(i) - z_t(i-1)$ for $i \geq 1$ from the assumption and $qh_t(i+1) \leq z_t(i+1) - z_t(i)$ from (3.6), we have

$$\begin{aligned} 0 < (=)\ & 2z_t(i) - z_t(i+1) - z_t(i-1) \\ = \ & qh_t(i) + z_t(i) - z_t(i+1) \leq q(h_t(i) - h_t(i+1)), \end{aligned} \tag{3.10}$$

from which we get the statement. ∎

In the next lemma, a condition for $h_t(i)$ being decreasing in i for a given t is revealed.

Lemma 4. *The critical value $h_t(i)$ is strictly decreasing (decreasing) in i for a given t if and only if for all $i \geq 1$* $2z_t(i) - z_t(i-1) - z_t(i+1) > (\geq)\, 0.$

Proof: If $h_t(i)$ is strictly decreasing in i, then $2z_t(i) - z_t(i-1) - z_t(i+1) > 0$ for all $i \geq 1$ from Lemma 3(c). The sufficient condition can be proven as follows. From Lemma 3(b), $h_t(1) > h_t(2)$ holds true, hence we have

$$h_t(2) = (z_t(2) - z_t(1))/q \tag{3.11}$$

due to Lemma 3(a). Accordingly, we get $h_t(2) > h_t(3)$ using Lemma 3(e), so

$$h_t(3) = (z_t(3) - z_t(2))/q. \tag{3.12}$$

Repeating the same procedure, we obtain $h_t(i) > h_t(i+1)$ for all $i \geq 1$. In a similar way, we can prove the case that $h_t(i)$ is decreasing in i. ∎

Next, we clarify the relation between $h_t(i)$ and $v_t(j)$.

Lemma 5. *If $h_t(i)$ is strictly decreasing (decreasing) in i for a given t, then for all $i \geq 1$*

$$2v_t(i) - v_t(i-1) - v_t(i+1) > (\geq)\, 0.$$

Proof: We only prove the case that $h_t(i)$ is strictly decreasing in i. For the case that $h_t(i)$ is decreasing in i, it can be proven in a similar way.

From Lemma 3(a) and the assumption of this lemma, we get $h_t(i) = (z_t(i) - z_t(i-1))/q$ for all $i \geq 1$. Hence, we can express $2v_t(i) - v_t(i-1) - v_t(i+1)$ as follows;

$$2v_t(i) - v_t(i-1) - v_t(i+1) = \int_0^{h_t(i+1)} A_t(i,\xi)dF(\xi) + \int_{h_t(i+1)}^{h_t(i)} B_t(i,\xi)dF(\xi) + \int_{h_t(i)}^{1} C_t(i,\xi)dF(\xi) \tag{3.13}$$

where

$$A_t(i,\xi) = 2z_t(i) - z_t(i-1) - z_t(i+1), \tag{3.14}$$

$$B_t(i,\xi) = 2z_t(i) - z_t(i-1) - (pz_t(i) + q\xi + (1-p)z_t(i)) = z_t(i) - z_t(i-1) - q\xi, \tag{3.15}$$

$$\begin{aligned} C_t(i,\xi) &= 2u_t(i,\xi) - u_t(i-1,\xi) - (pu_t(i,\xi) + q\xi + (1-p)z_t(i)) \\ &= (2-p)(pu_t(i-1,\xi) + q\xi + (1-p)z_t(i-1)) - u_t(i-1,\xi) - q\xi - (1-p)z_t(i) \\ &= -(1-p)^2 u_t(i-1,\xi) + (1-p)q\xi - (1-p)z_t(i) + (2-p)(1-p)z_t(i-1). \end{aligned} \tag{3.16}$$

From Lemma 4, we get $A_t(i,\xi) > 0$ for $0 \le \xi \le h_t(i+1)$ and $B_t(i,\xi) > 0$ for $h_t(i+1) \le \xi < h_t(i)$.

Below, using induction, we shall verify $C_t(i,\xi) > 0$ for $h_t(i) < \xi \le 1$ and all $i \ge 1$. If $i = 1$, then we get for $h_t(1) < \xi \le 1$

$$\begin{aligned} C_t(1,\xi) &= -(1-p)^2 u_t(0,\xi) + (1-p)q\xi - (1-p)z_t(1) + (2-p)(1-p)z_t(0) \\ &> -(1-p)^2 z_t(0) + (1-p)qh_t(1) - (1-p)z_t(1) + (2-p)(1-p)z_t(0) \\ &= (1-p)(qh_t(1) + z_t(0) - z_t(1)) = 0. \end{aligned} \tag{3.17}$$

Assume $C_t(i-1,\xi) > 0$ for $h_t(i-1) \le \xi \le 1$. Then, we have for $h_t(i-1) < \xi \le 1$

$$\begin{aligned} C_t(i,\xi) &= -(1-p)^2(pu_t(i-2,\xi) + q\xi + (1-p)z_t(i-1)) \\ &\quad +(1-p)q\xi - (1-p)z_t(i) + (2-p)(1-p)z_t(i-1) \\ &= pC_t(i-1,\xi) + (1-p)(2z_t(i-1) - z_t(i-2) - z_t(i)) > 0. \end{aligned} \tag{3.18}$$

Further, we obtain for $h_t(i) < \xi \le h_t(i-1)$

$$\begin{aligned} C_t(i,\xi) &= -(1-p)^2 z_t(i-1) + (1-p)q\xi - (1-p)z_t(i) + (1-p)(2-p)z_t(i-1) \\ &= (1-p)(q\xi + z_t(i-1) - z_t(i)) > 0. \end{aligned} \tag{3.19}$$

Therefore, we get $C_t(i,\xi) > 0$ for $h_t(i) < \xi \le 1$ and $i \ge 1$. In addition, it follows by direct calculation that $B_t(i,h_t(i)) = C_t(i,h_t(i)) = 0$. Finally, from the fact that the distribution does not concentrate on only $w = h_t(i)$, we get $2v_t(i) - v_t(i-1) - v_t(i+1) > 0$ for $i \ge 1$. ∎

We have investigated the basic structure of the optimal decision policy for shooting. In the following sections, the properties of $h_t(i)$ for some special cases will be discussed.

4. Case that Replenishing Every Period is Optimal

In this section, suppose $\phi_t(i) \ge 0$ for all $t \ge 1$ and $i \ge 0$, implying that it is always optimal to replenish m bullets. Then we shall clarify the monotonicity of $h_t(i)$ in i and the condition for $\phi_t(i) \ge 0$ for all $t \ge 1$ and $i \ge 0$.

Theorem 1.

(a) *On the above condition, the critical value $h_t(i)$ is decreasing in i for any $t \ge 1$. Particularly for $p > 0$, it is strictly decreasing in i.*

(b) *It holds true if and only if $a = 0$ that $\phi_t(i) \ge 0$ for any $t \ge 1$ and $i \ge 0$.*

Proof: (a) We only verify the case for $p > 0$. The proof for $p = 0$ is almost the same as below. Since $\phi_t(i) \ge 0$ all $t \ge 1$ and $i \ge 0$, $z_t(i) = \beta v_{t-1}(i+m) - a$ always holds true. Therefore, using (2.6), we get for $i \ge 1$

$$2z_1(i) - z_1(i-1) - z_1(i+1) = \beta(2v_0(i+m) - v_0(i-1+m) - v_0(i+1+m)) > 0. \tag{4.1}$$

Hence it follows that $2v_1(i) - v_1(i-1) - v_1(i+1) > 0$ for all $i \geq 1$ due to Lemmas 4 and 5. Accordingly we obtain

$$2z_2(i) - z_2(i-1) - z_2(i+1) = \beta(2v_1(i+m) - v_1(i-1+m) - v_1(i+1+m)) > 0 \tag{4.2}$$

for all $i \geq 1$. Repeating the procedure above yields $2z_t(i) - z_t(i-1) - z_t(i+1) > 0$ for all $t \geq 1$ and $i \geq 1$. Hence $h_t(i)$ is strictly decreasing in i for any $t \geq 1$ from Lemma 4.

(b) Now suppose $\phi_t(i) \geq 0$ for all $t \geq 1$ and $i \geq 0$. Because the number of targets the hunter gets over the whole planning horizon is at most $t+1$, $v_t(i)$ is upper bound for any t, implying that $v_t(i)$ converges as $i \to \infty$. Hence we get

$$\lim_{i\to\infty} \phi_t(i) = -a \geq 0. \tag{4.3}$$

Therefore, it must be that $a = 0$. To go the other way, if $a = 0$, then $\phi_t(i) = \beta(v_{t-1}(i+m) - v_{t-1}(i)) \geq 0$ for all $t \geq 1$ and $i \geq 0$ since $v_t(i)$ is increasing in i ∎

5. Case that Not Replenishing is always Optimal

Next, suppose $\phi_t(i) \leq 0$ for all $t \geq 1$ and $i \geq 0$, implying that it is always optimal not to replenish m bullets. The case is the same as the model in [7] with $c = 0$, in which the conclusion that $h_t(i)$ is strictly decreasing (decreasing) in i for $p > 0$ ($p = 0$) is obtained. Using the fact, we examine the condition for which it is always optimal not to replenish at all.

Theorem 2. *If $\beta mq \leq a$, then $\phi_t(i) \leq 0$ for all $t \geq 1$ and $i \geq 0$. In particular for $\beta = 1$, $\phi_t(i) \leq 0$ for all $t \geq 1$ and $i \geq 0$ if and only if $mq \leq a$.*

Proof: Now, we define the limits of $v_t(i)$, $z_t(i)$, $\phi_t(i)$ and $h_t(i)$ as $t \to \infty$, if exist, by $v(i)$, $z(i)$, $\phi(i)$ and $h(i)$, respectively. Using Lemma 1(c), we obtain $v_t(i+m) - v_t(i) < mq$ for all t and i, from which we get for all $t \geq 1$ and $i \geq 0$

$$\phi_t(i) = \beta(v_{t-1}(i+m) - v_{t-1}(i)) - a < \beta mq - a \leq 0. \tag{5.1}$$

Thus, the former part of the theorem, which is also the sufficient condition for the latter part, is proven. Now assume $\beta = 1$ and $\phi_t(i) \leq 0$ for all $t \geq 1$ and $i \geq 0$. Then, noting $z(i) = v(i)$, we get for $i \geq 1$

$$\begin{aligned} v(i) &= \int_0^1 \max\{v(i),\ pu(i-1,\xi) + q\xi + (1-p)v(i-1)\} dF(\xi) \\ &= \int_0^{h(i)} v(i) dF(\xi) + \int_{h(i)}^1 (pu(i-1,\xi) + q\xi + (1-p)v(i-1)) dF(\xi), \end{aligned} \tag{5.2}$$

which is rewritten

$$\int_{h(i)}^1 v(i) dF(\xi) = \int_{h(i)}^1 (pu(i-1,\xi) + q\xi + (1-p)v(i-1)) dF(\xi). \tag{5.3}$$

Now suppose $h(i) < 1$. Then from Lemma 2(a), we obtain $v(i) < pu(i-1,\xi) + q\xi + (1-p)v(i-1)$ for $h(i) < \xi \le 1$, which contradicts (5.3) because of $F(w) < 1$ for $w < 1$. Therefore, $h(i)$ must be equal to 1. Thus, it follows from Lemmas 3(a) and 4 that

$$\begin{aligned} 1 = h(1) &= (z(i) - z(i-1))/q \\ &= (v(i+1) - v(i))/q, \quad i \ge 1, \end{aligned} \tag{5.4}$$

which yields $v(i+m) - v(i) = mq$ for any i. Thus, we have

$$\phi(i) = v(i+m) - v(i) - a = mq - a \le 0, \tag{5.5}$$

that is, $mq \le a$. ▮

6. Case of $m = 1$

Now suppose that only one bullet can be replenished each period. Then, the following property can be said.

Theorem 3. *When $m = 1$, both $\phi_t(i)$ and $h_t(i)$ are always decreasing in i for any $t \ge 1$.*

Proof: It is clear for $m = 1$ that $\phi_t(i)$ is decreasing in i for any $t \ge 1$ if and only if $2v_t(i) - v_t(i-1) - v_t(i+1) \ge 0$ for any $t \ge 0$ and $i \ge 1$. From (2.6), it is true that

$$2v_0(i) - v_0(i-1) - v_0(i+1) = (1-p)p^{i-1}q\mu \ge 0 \quad i \ge 1; \tag{6.1}$$

accordingly for $i \ge 1$,

$$\begin{aligned} &2z_1(i) - z_1(i-1) - z_1(i+1) \\ &= 2\max\{\beta v_0(i), \beta v_0(i+1) - a\} - \max\{\beta v_0(i-1), \beta v_0(i) - a\} - \max\{\beta v_0(i+1), \beta v_0(i+2) - a\} \\ &= \begin{cases} \beta(2v_0(i+1) - v_0(i) - v_0(i+2)), & 0 \le a < \beta(v_0(i+2) - v_0(i+1)), \\ 2(\beta v_0(i+1) - a) - (\beta v_0(i) - a) - \beta v_0(i+1), & \\ & \beta(v_0(i+2) - v_0(i+1)) \le a < \beta(v_0(i+1) - v_0(i)), \\ 2\beta v_0(i) - (\beta v_0(i) - a) - \beta v_0(i+1), & \beta(v_0(i+1) - v_0(i)) \le a < \beta(v_0(i) - v_0(i-1)), \\ \beta(2v_0(i) - v_0(i-1) - v_0(i+1)), & \beta(v_0(i) - v_0(i-1)) \le a. \end{cases} \end{aligned} \tag{6.2}$$

From (6.1) and (6.2), we get $2z_1(i) - z_1(i-1) - z_1(i+1) \ge 0$ for any a and $i \ge 1$. Hence it follows from Lemma 4 that $h_1(i)$ is decreasing in i for any a. Now suppose $2z_t(i) - z_t(i-1) - z_t(i+1) \ge 0$, so $h_t(i)$ is decreasing in i and $h_t(i) = (z_t(i) - z_t(i-1))/q$. Accordingly, it follows from Lemma 5 that

$$2v_t(i) - v_t(i-1) - v_t(i+1) \ge 0, \quad i \ge 1. \tag{6.3}$$

Hence we have for $i \geq 1$

$$
\begin{aligned}
&2z_{t+1}(i) - z_{t+1}(i-1) - z_{t+1}(i+1) \\
&= 2\max\{\beta v_t(i), \beta v_t(i+1)-a\} - \max\{\beta v_t(i-1), \beta v_t(i)-a\} - \max\{\beta v_t(i+1), \beta v_t(i+2)-a\} \\
&= \begin{cases}
\beta(2v_t(i+1)-v_t(i)-v_t(i+2)), & 0 \leq a < \beta(v_t(i+2)-v_t(i+1)), \\
\beta(v_t(i+1)-v_t(i))-a, & \beta(v_t(i+2)-v_t(i+1)) \leq a < \beta(v_t(i+1)-v_t(i)), \\
-\beta(v_t(i+1)-v_t(i))+a, & \beta(v_t(i+1)-v_t(i)) \leq a < \beta(v_t(i)-v_t(i-1)), \\
\beta(2v_t(i)-v_t(i-1)-v_t(i+1)), & \beta(v_t(i)-v_t(i-1)) \leq a.
\end{cases}
\end{aligned}
\tag{6.4}
$$

From (6.3) and (6.4), we have $2z_{t+1}(i) - z_{t+1}(i-1) - z_{t+1}(i+1) \geq 0$ for any a. Thus by induction in terms of t, we obtain $2z_t(i) - z_t(i-1) - z_t(i+1) \geq 0$ for any $t \geq 1$, $i \geq 1$ and $a \geq 0$, so $h_t(i)$ is decreasing in i. On the other hand, since it is also verified in the above proof that $2v_t(i) - v_t(i-1) - v_t(i+1) \geq 0$ for any $t \geq 0$, $i \geq 1$ and $a \geq 0$, it also follows that $\phi_t(i)$ is decreasing in i. ∎

The monotonicity of $h_t(i)$ in i for any t and a is characteristic to the case for $m = 1$, however, this does not always hold true for $m \geq 2$. By the way, the property that $\phi_t(i)$ is decreasing in i leads us to the conclusion that, for a given t, the critical point for replenishment in terms of i where $\phi_t(i-1) \geq 0 > \phi_t(i)$ is at most one. Concretely speaking, if it is optimal to replenish m bullets with i bullets in hand, then it is also optimal to replenish m bullets with $j(< i)$ bullets in hand.

7. Case of $m \geq 2$ and Numerical Examples

Here we shall demonstrate an example that $h_t(i)$ is not always decreasing in i for $m \geq 2$. Let $p > 0$, $m = 2$ and $a = \beta(1+p)pq\mu$. Then, we get

$$z_1(0) = \max\{\beta v_0(0), \beta v_0(2) - a\} = \beta v_0(2) - a = \beta(1-p^2)q\mu, \tag{7.1}$$

$$z_1(1) = \max\{\beta v_0(1), \beta v_0(3) - a\} = \beta v_0(1) = \beta q\mu, \tag{7.2}$$

$$z_1(2) = \max\{\beta v_0(2), \beta v_0(4) - a\} = \beta v_0(2) = \beta(1+p)q\mu. \tag{7.3}$$

Accordingly we have

$$2z_1(1) - z_1(0) - z_1(2) = -\beta(1-p)pq\mu < 0, \tag{7.4}$$

which means $h_1(1) < h_1(2)$ due to Lemma 3(d).

Below, we depict the results of several numerical examples where a discrete uniform distribution function with 101 mass points equally spaced on $[0, 1]$ is used.

(a) When $m = 1$, $h_t(i)$ is decreasing in i even for $a > 0$ (Figure 2(a)).

(b) The non-monotonicity of $h_t(i)$ in i is shown in Figure 2(b,c,d), which also lead us to the conclusion that $h_t(i)$ is not always increasing in t. In [9], the monotonicity of $h_t(i)$ in t has been proven only for the case that it is always optimal not to replenish at all with $\beta = 1$.

(c) So far we have not investigated the relation of $h_t(i)$ to parameters a, q and r; it is quite intractable to reveal them theoretically. All of numerical examples we calculate show that $h_t(i)$ is increasing in a and r and decreasing in q. Figure 2(e) is an example of the relation of $h_t(i)$ to a.

(d) Figure 2(f) tells us the fact that $h_t(i)$ is not always monotone in β. We also get examples where $h_t(i)$ is not monotone in m.

By the way, such a non-monotonicity of $h_t(i)$ in i for $m \geq 2$ may fit our intuition in the following case where, for a certain j, $\phi_t(i) \geq 0$ if $i \leq j$ or else $\phi_t(i) < 0$. First, suppose the hunter has j bullets in hand. If it is decided not to shoot, then he needs not replenish m bullets at the period, or else he must replenish them according to the optimal policy for replenishment. Therefore, his behavior for shooting may become a little passive, that is, $h_t(j)$ becomes a little high. Next, suppose he has $j-1$ bullets. Then, his behavior may be more or less active since it is already decided to replenish the bullets at the period whether he decides or not to shoot, so $h_t(j-1)$ becomes a little low.

On the other hand, let us consider the optimal policy for replacement. It has been clarified that if $m = 1$, then the critical point for replenishment in terms of i is at most one. However, it has not been verified that whether the above property holds for $m \geq 2$ or not.

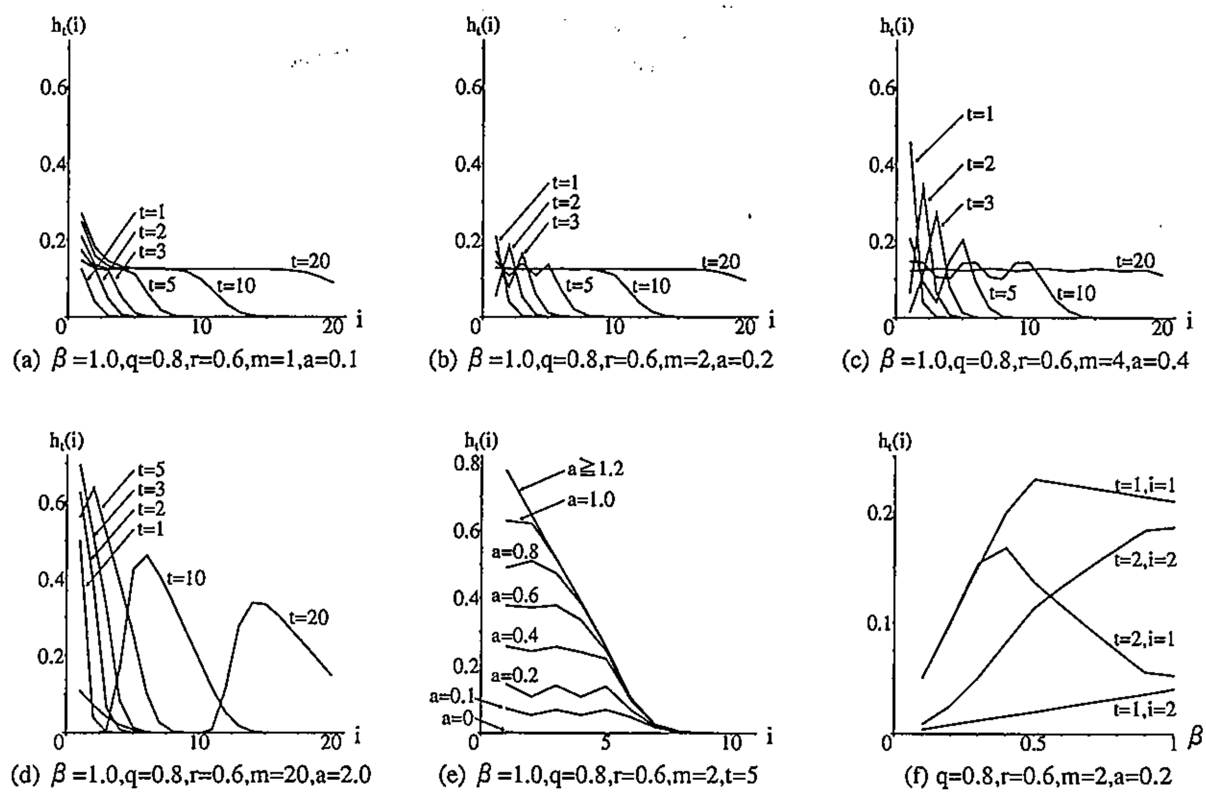

Figure 2 Numerical Examples

8. Conclusion

We have considered a discrete-time sequential allocation problem with countable resources which can be replenished each point in time, and the following conclusions are obtained:

(a) The necessary and sufficient condition for $\phi_t(i) \geq 0$ for all $t \geq 1$ and $i \geq 0$ is $a = 0$, for which $h_t(i)$ is always decreasing in i.

(b) If $\beta mq \leq a$, then $\phi_t(i) \leq 0$ for all $t \geq 1$ and $i \geq 0$, that is, it is optimal not to replenish at all. In particular for $\beta = 1$, $\phi_t(i) \leq 0$ for all $t \geq 1$ and $i \geq 0$ if and only if $mq \leq a$.

(c) If $m = 1$, then $h_t(i)$ is always decreasing in i and the critical point for replenishment is at most one, *i.e.*, if i^* such as $\phi_t(i^* - 1) \geq 0 > \phi_t(i^*)$ exists, then it is unique.

(d) If $m \geq 2$, then $h_t(i)$ is not always decreasing in i as well as not always monotone in t.

REFERENCES

[1] C. Derman, G. J. Lieberman and S. M. Ross : A Stochastic Sequential Allocation Model, *Operations Research*, Vol.23, (1975), 1120-1130.

[2] T. Kisi : Suboptimal Decision Rule for Attacking Targets of Opportunity, *Naval Research Logistics Quarterly*, Vol.23, (1976), 525-533.

[3] D. V. Mastran and C. J. Thomas : Decision Rule for Attacking Targets of Opportunity, *Naval Research Logistics Quarterly*, Vol.20, (1973), 661-672.

[4] T. Namekata, Y. Tabata and T. Nishida : A Sequential Allocation Problem with two kinds of Targets, *Journal of the Operations Research Society of Japan*, Vol.22, (1979), 16-28.

[5] T. Namekata, Y. Tabata and T. Nishida : A Sequential Allocation Problem with Perishable Goods, *Journal of the Operations Research Society of Japan*, Vol.22, (1979), 16-28.

[6] T. Namekata, Y. Tabata and T. Nishida : Some Remarks on Sequential Allocation Problem over Unknown Number of Periods, *Journal of the Operations Research Society of Japan*, Vol.22, (1979), 16-28.

[7] G. P. Prastacos : Optimal Sequential Investment Decisions under Conditions of Uncertainty, *Management Science*, Vol.29, (1983), 118-134.

[8] M. Sakaguchi : A Sequential Allocation Problem for Randomly Appearing Targets, *Mathematica Japonica*, Vol.21, (1976), 89-103.

[9] M. Sato and S. Ikuta : A Sequential Allocation Problem with Search Cost where the Shoot-Look-Shoot Policy is Employed, Discussion Paper, No.605 (Socio Economic Planning, University of Tsukuba), (1994).