Journal of the Operations Research
Society of Japan
Vol. 39, No. 3, September 1996

A SEQUENTIAL ALLOCATION PROBLEM WITH SEARCH COST WHERE THE SHOOT-LOOK-SHOOT POLICY IS EMPLOYED

Masahiro Sato
University of Tsukuba

(Received November 24, 1994; Revised May 12, 1995)

Abstract Suppose a hunter starts hunting over t periods with i bullets. A distribution of the value of targets appearing and the hitting probability of a bullet are known. For shooting, he takes a strategy of a shoot-look-shoot scheme. The objective in this paper is to find an optimal policy which maximizes the total expected reward. In the case with no search cost, the optimal policy is monotone in the number of bullets remaining then, but not always monotone in the case with positive search cost. We show such examples of non-monotonicity and examine conditions for the monotonicity of the optimal policy.

1. Introduction

This paper presents a stochastic sequential allocation problem in which it is intended to maximize the total expected reward obtained from investment opportunities appearing at successive points in time by allocating available resources among them. A fixed cost, called a search cost, is incurred each time to find an investment opportunity. The value of each investment opportunity is a random sample from a known distribution. The resources are countable like person, machine, bullet, and so on. The allocation of resources follows a so-called *shoot-look-shoot* policy, implying that, if unsuccessful after investing one unit, then it is decided whether or not to invest one more unit. This problem can be applied to the following examples.

Hunting Problem Suppose a hunter sees a target of value w. Then, he must decide whether or not to shoot a bullet. If the value is rather small, it is wise to decide not to shoot any bullet expecting better chances to arrive later. Suppose the value is favorable and he decides to shoot one bullet. The bullet will hit the target or miss. In the former, he gets the value w; if bullets are still in hand, later chances are also available. In the latter, two cases are further possible: the target disappears immediately or stands still without any defense. If it does not escape, then he has to decide whether or not to shoot an additional bullet.

Advertising Problem Suppose a saleslady visits different stores with i samples to advertise new products. On coming across a store, she can estimate the long-run profit w obtained from the store on condition that the manager of the store fortunately decides to deal with the products. Customers who come to the store will try the samples. Seeing the customers' responses, the manager decides whether or not to sell them in the store. If he decides to sell, then it may be regarded as a success, and she can obtain w and begins searching for another store. If he decides not to sell, then she cannot get any profit from the store and must search for another store. It may be possible that he still hesitates to sell even after some customers have tried the samples. Then, she has to decide whether to continue advertising in the store or to quit and search for another store.

Here, if an opportunity is always successfully achieved with certainty by investing only one unit of resources, then the problem can be reduced to an optimal stopping problem with no recall where i opportunities can be obtained.

Now different types of sequential allocation problems have been discussed by many authors so far. Mastran and Thomas [4] treat it as a target attacking problem where the decision policy of a shoot-look-shoot scheme is discussed. They show the computational method for obtaining the optimal decision rule, but don't mathematically verify its structure. Kisi [3] considers a similar model to [4] and examines the relation between approximate solutions and exact ones. Sakaguchi [7] investigates the continuous-time version of [4]. He derives, for the shoot-look-shoot policy, the conclusion that the critical value, at which shooting and not shooting become indifferent in the optimal decision, is nonincreasing in the number of resources (torpedoes) remaining then. Namekata, Tabata and Nishida [5] also deal with a similar model where there exist two kinds of targets in a sense that the necessary number of resources (torpedoes) to get any of these targets are different. Derman, Lieberman and Ross [1], and Prastacos [6] consider them as investment problems. In [1], the case is discussed that the reward from an opportunity is nondecreasing and concave in the amount of resources invested, and in [6], the case that the reward depends on both the value (quality) of an opportunity appearing and the amount of resources invested. In both papers, however, the shoot-look-shoot policy was not discussed.

Now, it should be noted that, in all of the models above, a search cost was not introduced, among which there exist ones where a search cost must be assumed from a practical view point. The objective of the present paper is to pose a general model in which a search cost is an essential factor and examine properties of its optimal decision policy. One of the most distinctive results obtained in this paper is that *the critical value does not always become nonincreasing in the number of remaining units of resources.*

In Section 2 we define our model, and in Section 3 its fundamental equations are derived. In Section 4 that follows, the structure of the optimal decision policy is investigated. In Sections 5, 6 and 7, we consider cases with no search cost, with positive search cost, and with large search cost, respectively. In Section 8 the case is examined that an infinite amount of resources is available, and in Section 9 some numerical examples are given. In Section 10 the conclusions obtained are summarized. Finally, some limitations of the present model are stated in Section 11.

2. Model

Throughout this paper, we shall explain our model by using the following hunting problem. Suppose a hunter starts hunting over t periods with i bullets. His purpose is to maximize the total expected reward from the game that will be bagged over a given planning horizon, that is, t periods. In the woods, he can find a target with appearing probability $\theta \in (0, 1]$, assuming that more than one target cannot be found at the same time. The value of target is a random variable having a known probability distribution function $F_1(w)$, continuous or discrete, where $F_1(w) = 0$ for $w < 0$, $F_1(w) < 1$ for $w < 1$ and $F_1(w) = 1$ for $w \geq 1$. The distribution does not concentrate on only a point, that is, $\Pr(w) < 1$ for any w. The values of targets that are found at successive points are assumed to be stochastically independent. Now let $F_0(w)$ be a distribution function where $\Pr(w) = 0$ for any $w \neq 0$. Then, using θ, we can combined $F_1(w)$ and $F_0(w)$ into the following distribution function:

$$F(w) = (1 - \theta)F_0(w) + \theta F_1(w). \tag{2.1}$$

Below, we enumerate the parameters used in the paper:

q	:	hitting probability,	$q \in (0,1]$,
r	:	escaping probability,	$r \in [0,1]$,
c	:	search cost,	$c \geq 0$,
β	:	discount factor,	$\beta \in (0,1]$,
μ	:	mean of $F(w)$, $\mu = \int \xi dF(\xi)$,	$\mu \in (0,1)$,
p	:	$p = (1-q)(1-r)$,	$p \in [0,1)$.

The strict meanings of these parameters are as follows. Assume that the hunter can observe the value of a target as soon as he finds it. Then he has to decide whether or not to shoot a bullet. If the value w is rather small, it is wise for him not to shoot. Suppose the value is favorable and he decides to shoot a bullet. The bullet will either hit the target with hitting probability q or miss it. In the former, he gets the value w; if bullets are still in hand, later chances are also available. In the latter, 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; assuming that repeated firings waste no time. If he decides not to shoot any more,† then he comes home. On his way home, he must decide whether or not to go hunting on the next point. If it is profitable for him to go hunting, then he consumes a search cost c to prepare for the next hunting. Figure 1 illustrates the structure of this decision process where the rhombuses with " fire " or " go " are related to $u_t(i,w)$ or $z_t(i)$, respectively, which will be defined in the next section.

†Exactly speaking, it means that he decides not to shoot any more at the present target, need not shoot (get it), or cannot shoot (it escapes or $i = 0$). Below, we also employ this expression for simplicity.

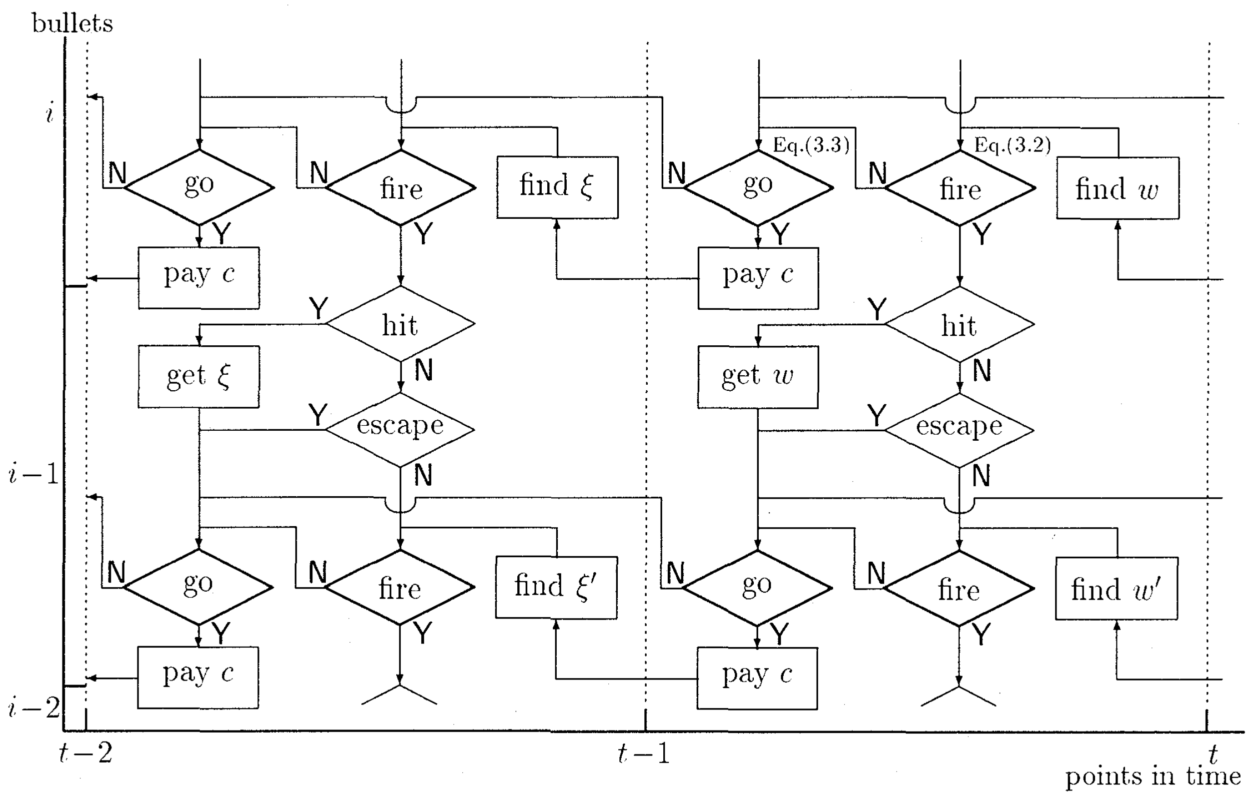


Figure 1. Graphic expression of the decision process
In the figure, (3.2) and (3.3) are the equation numbers of $u_t(i,w)$ and $z_t(i)$.

Finally we assume, except for Section 7, that

$$c \leq \beta q\mu, \tag{2.2}$$

which implies that it is favorable to go hunting even when one period remains and he has only one bullet in hand.

3. Fundamental Equations

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. Now we define $u_t(i,w)$ as the maximum of the total expected present discounted reward starting from time t when the hunter is seeing a target of value w with i bullets in hand. Furthermore, let the expectation of $u_t(i,w)$ in terms of w be designated by

$$v_t(i) = \int_0^1 u_t(i,\xi)dF(\xi), \quad t \geq 0. \tag{3.1}$$

Then, we have the following recursive relation:

$$\begin{aligned} u_t(i,w) &= \max\{z_t(i),\ q(w + z_t(i-1)) + (1-q)(rz_t(i-1) + (1-r)u_t(i-1,w))\} \\ &= \max\{z_t(i),\ pu_t(i-1,w) + qw + (1-p)z_t(i-1)\}, \quad t \geq 1,\ i \geq 1. \end{aligned} \tag{3.2}$$

Here, the first (second) term inside the braces represents the maximum of the total expected reward when there remain t periods and i bullets and it is decided not to shoot (to shoot one bullet) at the present target. Now, if he decides not to shoot any more, he must decide whether or not to go hunting on the next point. Therefore, $z_t(i)$ can be expressed by

$$z_t(i) = \max\{\beta v_{t-1}(i) - c,\ \beta z_{t-1}(i)\}, \quad t \geq 1, i \geq 0. \tag{3.3}$$

In the right side of above, the first (second) term inside the braces represents the maximum of the total expected reward when it is decided to go (not to go) hunting on the next point. Furthermore, from the definition of the model, we have the following final conditions:

$$u_t(0,w) = v_t(0) = z_t(0) = z_0(i) = 0, \quad t \geq 0, i \geq 0, \tag{3.4}$$

$$u_0(i,w) = qw + pu_0(i-1,w), \quad i \geq 1. \tag{3.5}$$

Here, $u_0(i,w)$ can be expressed as

$$u_0(i,w) = \frac{1-p^i}{1-p}qw, \tag{3.6}$$

which holds for $i \geq 0$. Thus, we also get

$$v_0(i) = \frac{1-p^i}{1-p}q\mu, \quad i \geq 0. \tag{3.7}$$

Hereafter in this section, we clarify the properties of $u_t(i,w)$ and $v_t(i)$.

Lemma 1.

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

(b) *If $t \geq 1$ and $i \geq 1$, then $z_t(i) = \beta v_{t-1}(i) - c$.*

(c) *Both $u_t(i,w)$ and $v_t(i)$ are nondecreasing in i for $t \geq 0$. In particular for $p > 0$, they are strictly increasing in i for $t \geq 0$ and $w > 0$.*

(d) *If $w < 1$, then $u_t(i,w) - u_t(i-1,w) < q$ for $t \geq 0$ and $i \geq 1$. In addition, $v_t(i) - v_t(i-1) < q$ also holds for $t \geq 0$ and $i \geq 1$.*

(e) *$u_t(i,w)$ is nondecreasing in w for any t and any i.*

(f) *For $t \geq 0$ and $i \geq 1$, $u_t(i,1) = pu_t(i-1,1) + q + (1-p)z_t(i-1) > z_t(i)$.*

Proof: (a) From (3.4), we have for any i

$$z_1(i) = \max\{\beta v_0(i) - c,\ \beta z_0(i)\} \geq \beta z_0(i) = 0 = z_0(i), \tag{3.8}$$

hence we get for any i

$$\begin{aligned} u_1(i,w) &\geq pu_1(i-1,w)+qw+(1-p)z_1(i-1) \\ &\geq pu_1(i-1,w)+qw \\ &\geq p(pu_1(i-2,w)+qw)+qw \\ &\vdots \\ &\geq \frac{1-p^i}{1-p}qw = u_0(i,w). \end{aligned} \tag{3.9}$$

Immediately from above, it follows that $v_1(i) \geq v_0(i)$. Now, assuming $v_t(i) \geq v_{t-1}(i)$ and $z_t(i) \geq z_{t-1}(i)$ for all $i \geq 0$ as inductive assumption in terms of t, we have

$$\begin{aligned} z_{t+1}(i) &= \max\{\beta v_t(i)-c,\ \beta z_t(i)\} \\ &\geq \max\{\beta v_{t-1}(i)-c,\ \beta z_{t-1}(i)\} = z_t(i), \end{aligned} \tag{3.10}$$

accordingly, we get

$$\begin{aligned} u_{t+1}(1,w) &= \max\{z_{t+1}(1),\ qw\} \\ &\geq \max\{z_t(1),\ qw\} = u_t(1,w). \end{aligned} \tag{3.11}$$

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:

$$\begin{aligned} u_{t+1}(i,w) &= \max\{z_{t+1}(i),\ pu_{t+1}(i-1,w)+qw+(1-p)z_{t+1}(i-1)\} \\ &\geq \max\{z_t(i),\ pu_t(i-1,w)+qw+(1-p)z_t(i-1)\} = u_t(i,w). \end{aligned} \tag{3.12}$$

Thus, it follows by double induction that $u_t(i,w)$ is nondecreasing in t for any i and any w, so also are $v_t(i)$ and $z_t(i)$ for any i.

(b) It is clear from (2.2), (3.4) and (3.7) that $\beta v_0(i)-c = \beta(1-p^i)q\mu/(1-p)-c \geq 0 = \beta z_0(i)$ for $i \geq 1$. Assume $\beta v_{t-1}(i)-c \geq \beta z_{t-1}(i)$ for $i \geq 1$, hence $z_t(i) = \beta v_{t-1}(i)-c$ from (3.3). Then, we have for $i \geq 1$

$$\begin{aligned} \beta v_t(i)-c-\beta z_t(i) &= \beta v_t(i)-c-\beta(\beta v_{t-1}(i)-c) \\ &\geq \beta v_t(i)-c-(\beta v_{t-1}(i)-c) \\ &= \beta(v_t(i)-v_{t-1}(i)) \geq 0 \end{aligned} \tag{3.13}$$

from Lemma 1(a). Therefore, we get $z_{t+1}(i) = \beta v_t(i)-c$. Thus, it follows by induction that $z_t(i) = \beta v_{t-1}(i)-c$ for $t \geq 1$ and $i \geq 1$. Consequently, it follows for $t \geq 1$ that

$$u_t(i,w) = \begin{cases} \max\{\beta v_{t-1}(i)-c,\ pu_t(i-1,w)+qw+(1-p)(\beta v_{t-1}(i-1)-c)\}, & i \geq 2, \\ \max\{\beta v_{t-1}(1)-c,\ qw\}, & i=1. \end{cases} \tag{3.14}$$

(c) We shall only prove the case of $p > 0$. The proof for $p = 0$ is almost same as below. Now for $p > 0$, it is obvious from (3.6) and (3.7) that both $u_0(i,w)$ and $v_0(i)$ are strictly increasing in i. Let $v_{t-1}(i)$ be strictly increasing in i. Then, clearly $u_t(1,w) > u_t(0,w) = 0$ for $w > 0$. Furthermore, suppose that $u_t(i,w) > u_t(i-1,w)$ for $w > 0$. Then we have for $w > 0$

$$\begin{aligned} u_t(i+1,w) &= \max\{\beta v_{t-1}(i+1)-c,\ pu_t(i,w)+qw+(1-p)(\beta v_{t-1}(i)-c)\} \\ &> \max\{\beta v_{t-1}(i)-c,\ pu_t(i-1,w)+qw+(1-p)(\beta v_{t-1}(i-1)-c)\} \\ &= u_t(i,w). \end{aligned} \tag{3.15}$$

By integrating the equation in terms of w, we have $v_t(i+1) > v_t(i)$. Thus, it is proven by double induction that $u_t(i,w)$ and $v_t(i)$ are strictly increasing in i for $t \geq 0$ and $w > 0$.

(d) If $i = 1$, then $u_0(1,w)-u_0(0,w) = qw-0 < q$ for $w < 1$, so $v_0(1)-v_0(0) < q$. Suppose $u_{t-1}(1,w)-u_{t-1}(0,w) < q$ for $w < 1$, so $v_{t-1}(1)-v_{t-1}(0) < q$. Then, we have for $w < 1$

$$u_t(1,w)-u_t(0,w) = \max\{\beta v_{t-1}(1)-c,\ qw\}-0 < \max\{\beta q-c,\ q\} = q, \tag{3.16}$$

from which we obtain $v_t(1)-v_t(0) = v_t(1) < q$.

If $i = 2$, then $u_0(2, w) - u_0(1, w) = (1+p)qw - qw = pqw < q$, hence $v_0(2) - v_0(1) < q$. Let $u_{t-1}(2, w) - u_{t-1}(1, w) < q$, so $v_{t-1}(2) - v_{t-1}(1) < q$. Then we get

$$\begin{aligned} u_t(2, w) - u_t(1, w) &\leq \max\{\beta(v_{t-1}(2) - v_{t-1}(1)), pu_t(1, w) + (1-p)(\beta v_{t-1}(1) - c)\}^{\ddagger} \\ &< \max\{\beta q, pq + (1-p)(\beta q - c)\} \leq q, \end{aligned} \tag{3.17}$$

from which we obtain $v_t(2) - v_t(1) < q$.

If $i \geq 3$, then $u_0(i, w) - u_0(i-1, w) = ((1-p^i) - (1-p^{i-1}))qw/(1-p) = p^{i-1}qw < q$, accordingly $v_0(i) - v_0(i-1) < q$. Assume $u_{t-1}(i, w) - u_{t-1}(i-1, w) < q$, so $v_{t-1}(i) - v_{t-1}(i-1) < q$. Then we have

$$\begin{aligned} u_t(i, w) - u_t(i-1, w) \leq \max\{&\beta(v_{t-1}(i) - v_{t-1}(i-1)), \\ &p(u_t(i-1, w) - u_t(i-2, w)) + \beta(1-p)(v_{t-1}(i-1) - v_{t-1}(i-2))\}. \end{aligned} \tag{3.18}$$

Because $u_t(i-1, w) - u_t(i-2, w) < q$ for $i = 3$, we get

$$u_t(3, w) - u_t(2, w) < \max\{\beta q, pq + \beta(1-p)q\} \leq q \tag{3.19}$$

from (3.17). Furthermore, suppose $u_t(i-1, w) - u_t(i-2, w) < q$ as the second inductive assumption, then it yields $u_t(i, w) - u_t(i-1, w) < q$. Accordingly, it follows by double induction that $u_t(i, w) - u_t(i-1, w) < q$ for $i \geq 3$. Thus, we obtain $u_t(i, w) - u_t(i-1, w) < q$ for $t \geq 0$, $i \geq 1$ and $w < 1$, so that $v_t(i) - v_t(i-1) < q$ for $t \geq 0$ and $i \geq 1$.

(e) It is easily proven by induction.

(f) The statement is obvious for $t = 0$. Now we get $v_{t-1}(i) < q + v_{t-1}(i-1)$ for $t \geq 1$ and $i \geq 1$ from Lemma 1(d), and $u_t(i-1, 1) \geq \beta v_{t-1}(i-1) - c$ from (3.14). Hence we have for $t \geq 1$ and $i \geq 1$

$$\begin{aligned} u_t(i, 1) &= pu_t(i-1, 1) + q + (1-p)(\beta v_{t-1}(i-1) - c) \\ &\geq q + \beta v_{t-1}(i-1) - c \\ &\geq \beta(q + v_{t-1}(i-1)) - c \\ &> \beta v_{t-1}(i) - c = z_t(i). \end{aligned} \tag{3.20}$$

Thus, we get the statement. ∎

Lemmas 1(b) and (f) mean, respectively, that if $\beta mq \geq c$ and $i \geq 1$, then it is always optimal to go hunting and that if $w = 1$ and $i \geq 1$ then it is always optimal to shoot at the target. Using (3.1) and (3.14) recursively, we can calculate $v_t(i)$ starting with the final conditions (3.4), (3.6) and (3.7).

4. Structure of Optimal Policy

For $t \geq 1$, define $g_t(i, w)$ as follows:

$$g_t(i, w) = \begin{cases} pu_t(i-1, w) + qw + (1-p)(\beta v_{t-1}(i-1) - c) - (\beta v_{t-1}(i) - c), & i \geq 2, \\ qw - (\beta v_{t-1}(1) - c), & i = 1. \end{cases} \tag{4.1}$$

Then the lemma below holds true.

Lemma 2. *For $t \geq 1$ and $i \geq 1$, $g_t(i, w) = 0$ has a unique solution $w = h_t(i) \in [0, 1)$.*

Proof: First, using induction and Lemma 1(c), we $g_t(i, 0) \leq 0$ for $t \geq 1$ and $i \geq 1$. Next, it is also obtained that $g_t(i, 1) > 0$ for $t \geq 1$ and $i \geq 1$ from Lemma 1(f). Furthermore, $g_t(i, w)$ is a continuous function of w and strictly increasing in w from Lemma 1(e). Thus, from above, it is proven that $g_t(i, w) = 0$ has a unique solution $w = h_t(i) \in [0, 1)$. ∎

Remark: We call the $h_t(i)$ a critical value when the hunter has i bullets and t periods remain. From Lemma 2, the optimal decision policy becomes as follows. If $w \geq h_t(i)$, then

‡Here, we use the relation $\max\{a_i\} - \max\{b_i\} \leq \max\{a_i - b_i\}$, which is often used in the proofs of this paper.

fire, or else don't fire. When $w = h_t(i)$ for given t and i, shooting and not shooting become indifferent; that is,

$$u_t(i, h_t(i)) = \beta v_{t-1}(i) - c = pu_t(i-1, h_t(i)) + qh_t(i) + (1-p)(\beta v_{t-1}(i-1) - c) \quad (4.2)$$

for $t \geq 1$ and $i \geq 2$. Now, since $u_t(i-1, w) \geq \beta v_{t-1}(i-1) - c$, we have for $t \geq 1$ and $i \geq 2$

$$0 = g_t(i, h_t(i)) \geq qh_t(i) + \beta(v_{t-1}(i-1) - v_{t-1}(i)) \quad (4.3)$$

due to (4.1), from which we obtain

$$h_t(i) \leq \beta(v_{t-1}(i) - v_{t-1}(i-1))/q. \quad (4.4)$$

Furthermore, we get for $t \geq 1$ and $i = 1$

$$u_t(1, h_t(1)) = \beta v_{t-1}(1) - c = qh_t(1) \quad (4.5)$$

from (4.1), and since $c \geq 0$ and $v_t(0) = 0$, it follows that

$$h_t(1) = (\beta v_{t-1}(1) - c)/q \leq \beta v_{t-1}(1)/q = \beta(v_{t-1}(1) - v_{t-1}(0))/q. \quad (4.6)$$

Thus, (4.4) holds for $t \geq 1$ and $i \geq 1$.

Now we detail relations between $h_t(i)$ and $v_{t-1}(j)$ in Lemmas 3, 4, 5 and 6 and Corollary 1:

Lemma 3. *If $p > 0$, then the following hold for $i \geq 1$ and $t \geq 1$:*

(a) $h_t(i) \geq h_t(i+1) \quad \Leftrightarrow \quad h_t(i+1) = \beta(v_{t-1}(i+1) - v_{t-1}(i))/q,$

(b) $h_t(i) < h_t(i+1) \quad \Leftrightarrow \quad h_t(i+1) < \beta(v_{t-1}(i+1) - v_{t-1}(i))/q.$

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

Proof: F, we shall verify the case of $p > 0$. If $h_t(i) \geq h_t(i+1)$, then $u_t(i, h_t(i+1)) = \beta v_{t-1}(i) - c$ according to the optimal decision policy, hence we have for $i \geq 1$

$$\begin{aligned} 0 = g_t(i+1, h_t(i+1)) &= pu_t(i, h_t(i+1)) + qh_t(i+1) \\ &\qquad + (1-p)(\beta v_{t-1}(i) - c) - (\beta v_{t-1}(i+1) - c) \\ &= qh_t(i+1) + \beta(v_{t-1}(i) - v_{t-1}(i+1)). \end{aligned} \quad (4.7)$$

Thus, it follows that $h_t(i+1) = \beta(v_{t-1}(i+1) - v_{t-1}(i))/q$ for $i \geq 1$. Conversely, if $h_t(i+1) = \beta(v_{t-1}(i+1) - v_{t-1}(i))/q$, then we get

$$0 = g_t(i+1, h_t(i+1)) = pu_t(i, h_t(i+1)) - p(\beta v_{t-1}(i) - c), \quad (4.8)$$

from which we have $u_t(i, h_t(i+1)) = \beta v_{t-1}(i) - c$, so $h_t(i) \geq h_t(i+1)$. It is immediate from above and (4.4) that $h_t(i) < h_t(i+1) \;\Leftrightarrow\; h_t(i+1) < \beta(v_{t-1}(i+1) - v_{t-1}(i))/q$.

When $p = 0$, the assertion is easily verified because

$$0 = g_t(i+1, h_t(i+1)) = qh_t(i+1) + \beta(v_{t-1}(i) - v_{t-1}(i+1)) \quad (4.9)$$

for $i \geq 1$ and $t \geq 1$. ∎

Lemma 4. *For $t \geq 1$ and $p > 0$,*

$$\beta(2v_{t-1}(1) - v_{t-1}(2)) - c \begin{Bmatrix} \geq \\ = \\ < \end{Bmatrix} 0 \;\Leftrightarrow\; h_t(1) \begin{Bmatrix} \geq \\ = \\ < \end{Bmatrix} h_t(2).$$

Proof: Using (4.5), we have

$$\begin{aligned} g_t(2, h_t(1)) &= pu_t(1, h_t(1)) + qh_t(1) + (1-p)(\beta v_{t-1}(1) - c) - (\beta v_{t-1}(2) - c) \\ &= \beta(2v_{t-1}(1) - v_{t-1}(2)) - c, \end{aligned} \quad (4.10)$$

from which we get the statement in the lemma. ∎

Lemma 5. *For a given $i \geq 1$ and $t \geq 1$,*

$$h_t(i) > (=)\, h_t(i+1) \;\Rightarrow\; 2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) > (\geq)\, 0,$$

Proof: If $h_t(i) > (=)\, h_t(i+1)$, then we have for $i \geq 1$

$$\begin{aligned} 0 < (=)\, g_t(i+1, h_t(i)) &= pu_t(i, h_t(i)) + qh_t(i) + (1-p)(\beta v_{t-1}(i) - c) - (\beta v_{t-1}(i+1) - c) \\ &= qh_t(i) + \beta(v_{t-1}(i) - v_{t-1}(i+1)) \\ &\leq \beta(2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1)) \end{aligned} \quad (4.11)$$

from (4.2) and (4.4). ∎

Corollary 1. *For a given $i \geq 1$ and $t \geq 1$.*

$$2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) < 0 \quad \Rightarrow \quad h_t(i) < h_t(i+1).$$

Proof: The statement is the contraposition of Lemma 5. ∎

Lemma 6. *If $h_t(i) = \beta(v_{t-1}(i) - v_{t-1}(i-1))/q$ for a given $i \geq 2$, then*

$$2v_{t-1}(i) - v_{t-1}(i+1) - v_{t-1}(i-1) > (=) 0 \ \Rightarrow \ h_t(i) > (\geq) h_t(i+1).$$

Proof: From the assumption and (4.4), we immediately have

$$\begin{aligned} 0 < (=) &\ \beta(2v_{t-1}(i) - v_{t-1}(i+1) - v_{t-1}(i-1)) \\ = &\ qh_t(i) - \beta(v_{t-1}(i+1) - v_{t-1}(i)) \leq q(h_t(i) - h_t(i+1)). \end{aligned} \tag{4.12}$$

Hence, it follows that $h_t(i) > (\geq) h_t(i+1)$. ∎

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

Theorem 1. *The critical value $h_t(i)$ is strictly decreasing (nonincreasing) in i for a given t if and only if*

$$\beta(2v_{t-1}(1) - v_{t-1}(2)) - c > (\geq) 0 \tag{4.13}$$

and for all $i \geq 2$

$$2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) > (\geq) 0. \tag{4.14}$$

Proof: If $h_t(i)$ is strictly decreasing in i, then $\beta(2v_{t-1}(1) - v_{t-1}(2)) - c > 0$ from Lemma 4, and $2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) > 0$ for $i \geq 1$ from Lemma 5. The sufficient condition can be proven as follows. From Lemma 4, $h_t(1) > h_t(2)$ holds true, hence we have

$$h_t(2) = \beta(v_{t-1}(2) - v_{t-1}(1))/q \tag{4.15}$$

from Lemma 3(a). Accordingly, we get $h_t(2) > h_t(3)$ using Lemma 6, so

$$h_t(3) = \beta(v_{t-1}(3) - v_{t-1}(2))/q. \tag{4.16}$$

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 nonincreasing in i. ∎

So far we have considered the relations between $h_t(i)$ and $v_{t-1}(j)$. In the following lemma, let us show a relation between $h_t(i)$ and $v_t(j)$.

Lemma 7. *If $h_t(i)$ is strictly decreasing (nonincreasing) in i for a given $t \geq 1$, then*

$$2v_t(i) - v_t(i-1) - v_t(i+1) > (\geq) 0, \quad i \geq 1. \tag{4.17}$$

Proof: See Appendix. ∎

In the lemma below, we describe a condition for which $h_t(i)$ is decreasing in i for any t.

Lemma 8. *A necessary and sufficient condition for which $h_t(i)$ is nonincreasing in i for any $t \geq 1$ is*

$$\beta(2v_t(1) - v_t(2)) - c \geq 0 \tag{4.18}$$

for all $t \geq 0$. Particularly for $p > 0$, $h_t(i)$ is strictly decreasing in i for any $t \geq 1$ if and only if $\beta(2v_t(1) - v_t(2)) - c > 0$ for all $t \geq 0$.

Proof: The necessary condition is obvious from Theorem 1. The sufficient condition is verified as follows. We have, for $i \geq 1$, $2v_0(i) - v_0(i-1) - v_0(i+1) = (1-p)p^{i-1}q\mu \geq 0$ from (3.7) and $\beta(2v_0(1) - v_0(2)) - c \geq 0$ from the assumption. Hence, it follows from Theorem 1 that $h_1(i)$ is nonincreasing in i for all $i \geq 1$, so $2v_1(i) - v_1(i-1) - v_1(i+1) \geq 0$ for $i \geq 1$ from Lemma 7. Repeatedly applying the above procedure, we obtain $h_t(i)$ is nonincreasing in i for all $t \geq 1$. For $p > 0$, it can be proven by replacing ≥ 0 by > 0 in the above. ∎

5. Case with No Search Cost ($c = 0$)

The theorem below and Corollary 2 that follows hold for $c = 0$.

Theorem 2. *For $t \geq 1$, $h_t(i)$ is nonincreasing in i. In particular, if $p > 0$, then $h_t(i)$ is strictly decreasing in i.*

Proof: Because $c = 0$, substituting $i = 1$ into (4.17) produces

$$\beta(2v_t(1) - v_t(2) - v_t(0)) = \beta(2v_t(1) - v_t(2)) - c > (\geq) 0, \quad 0 < p < 1\,(p < 1), \tag{5.1}$$

which is the same as (4.18) in Lemma 8. The remaining is similar to the proof of the sufficient condition for Lemma 8. ∎

Corollary 2. *For $t \geq 0$, $v_t(i)$ is concave in i. Particularly if $p > 0$, then it is strictly concave in i.*

Proof: It is from the fact that the difference of $v_t(i)$ as to i is nonincreasing in i (strictly decreasing in i for $p > 0$) from Lemma 7 and Theorem 2. ∎

Corollary 3 tells a property of $h_t(i)$ in terms of t for $\beta = 1$ and $c = 0$.

Corollary 3. *If $\beta = 1$ and $c = 0$, then $h_t(i)$ is nondecreasing in t for any $i \geq 1$.*

Proof: Due to the assumption of $\beta = 1$ and $c = 0$, we have $u_t(i, w) = \max\{v_{t-1}(i),\ pu_t(i-1, w) + qw + (1-p)v_{t-1}(i-1)\}$ for $i \geq 2$ and $u_t(1, w) = \max\{v_{t-1}(1),\ qw\}$. Now we get $u_t(1, w) - u_t(0, w) = u_t(1, w) \geq v_{t-1}(1) = v_{t-1}(1) - v_{t-1}(0)$. Suppose

$$u_t(i-1, w) - u_t(i-2, w) \geq v_{t-1}(i-1) - v_{t-1}(i-2). \tag{5.2}$$

Then, we obtain for $0 \leq w \leq h_t(i)$

$$u_t(i, w) - u_t(i-1, w) = v_{t-1}(i) - v_{t-1}(i-1), \tag{5.3}$$

for $h_t(i) < w \leq h_t(i-1)$

$$u_t(i, w) - u_t(i-1, w) = u_t(i, w) - v_{t-1}(i-1) > v_{t-1}(i) - v_{t-1}(i-1), \tag{5.4}$$

and for $h_t(i-1) < w \leq 1$

$$\begin{aligned} u_t(i, w) - u_t(i-1, w) &= p(u_t(i-1, w) - u_t(i-2, w)) \\ &\quad + (1-p)(v_{t-1}(i-1) - v_{t-1}(i-2)) \\ &\geq v_{t-1}(i-1) - v_{t-1}(i-2) \geq v_{t-1}(i) - v_{t-1}(i-1) \end{aligned} \tag{5.5}$$

using Corollary 2. From (5.3), (5.4) and (5.5), it follows for $t \geq 1$ and $i \geq 1$ that

$$v_t(i) - v_t(i-1) \geq v_{t-1}(i) - v_{t-1}(i-1). \tag{5.6}$$

Since $h_t(i)$ is nonincreasing in i for any $t \geq 1$ from Theorem 2, it is always equal to $\beta(v_{t-1}(i) - v_{t-1}(i-1))/q$ from Lemma 3. Therefore, $h_t(i)$ is nondecreasing in t for any $i \geq 1$. ∎

6. Case with Positive Search Cost ($0 < c \leq \beta q\mu$)

We can conjecture that the more bullets the hunter has, the smaller the value of target he may decide to shoot at; that is, $h_t(i)$ is decreasing in i. However, $h_t(i)$ is not always decreasing in i if $c > 0$. Below let us show such a simple example.

Using (3.7), (4.2) and (4.6), we have

$$\begin{aligned} g_1(2, h_1(1)) &= pu_1(1, h_1(1)) + qh_1(1) + (1-p)(\beta v_0(1) - c) - (\beta v_0(2) - c) \\ &= \beta(2v_0(1) - v_0(2)) - c = \beta(1-p)q\mu - c. \end{aligned} \tag{6.1}$$

If $\beta(1-p)q\mu < c \leq \beta q\mu$, then $g_1(2, h_1(1)) < 0$, implying that there may exist a certain interval of i in which $h_t(i)$ becomes increasing in i as being depicted in Figure 2[§] which illustrates the following structure in terms of the optimal decision policy.

[§]This is a conceptual figure illustrating that $h_t(i)$ may become partially increasing in i; ones for certain specified parameters are demonstrated in Section 9.

(a) Suppose a present target value is w_a. If he has more than eleven bullets, then he should continue to fire until at least one of the following three events occurs: he gets the target, the remaining number of bullets becomes less than twelve or it escapes. If he starts with less than twelve, then he should not fire and search for the next target.

(b) Suppose a present target value is w_b. If he has more than six bullets, then he should continue to fire until at least one of the following three events occurs: he gets it, the number of bullets in hand becomes less than seven or it runs away. If he starts with more than two and less than seven, then he should not fire and search for the next target. If he has one or two bullets, then he should fire until at least one of the following three events occurs: he gets the target, spends all the bullets or it escapes. We have defined $h_t(i)$ as a critical point in terms of the targets of the value w and called a critical value. In a similar way, we can also define a critical point in terms of i. The above explanation for w_b tells us the existence of double critical points i_* and i^* $(i_* < i^*)$ in terms of i in a sense that if $i \leq i_*$ or $i \geq i^*$, then fire, or else don't fire. It goes without saying that such a thing never occurs if $h_t(i)$ is nonincreasing in i.

(c) Suppose $w = w_c$. In this case, he should continue to fire until at least one of the following three events occurs: he gets the target, spends all the bullets or it runs away.

Now the following lemma says a property of $h_t(i)$ for $c > 0$.

Lemma 9. *If $h_t(i_a) < h_t(i_a + 1)$ for a certain i_a, then there exists a certain i_b such that $i_b > i_a$ and $h_t(i_b) > h_t(i_b + 1)$.*

Proof: It is obvious from the fact that $\lim_{i\to\infty} h_t(i) = 0$ because $\beta(v_{t-1}(i) - v_{t-1}(i-1))/q$ converges to zero as $i \to \infty$. ∎

Regarding $v_t(i)$ as a function of c, we shall use the symbol $v_t(i, c)$ instead of $v_t(i)$. In the same way, $h_t(i, c)$ will be also used. Let

$$D_t(c) = g_{t+1}(2, h_{t+1}(1)) = \beta(2v_t(1, c) - v_t(2, c)) - c, \quad t \geq 0. \tag{6.2}$$

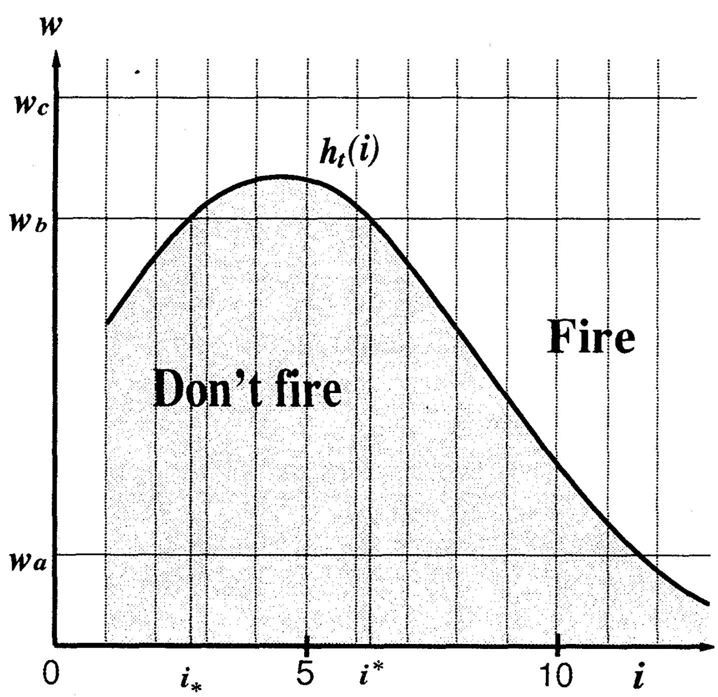


Figure 2. Existence of i_* and i^*

Because the number of the game that he will get over the whole planning horizon starting with *finite* i bullets is at most i, there exists the limit of $v_t(i,c)$ as $t \to \infty$, hence the limits of $h_t(1,c)$ and $D_t(c)$ as $t \to \infty$ also exist. We designate them by their symbols without the subscript t, i.e., $v(i,c)$, $h(1,c)$ and $D(c)$. Now, we shall verify the existence of c for which $h_t(i,c)$ is nonincreasing in i and for which $h_t(i,c)$ is not monotone in i.

Theorem 3. *If $\beta < 1$, then there exists a positive number c_* for which $h_t(i,c)$ is nonincreasing in i for any t and $c \in [0, c_*]$. In addition, if $\beta < 1$ and $p > 0$, then there exists $c^* \in [c_*, \beta q\mu)$ for which $h_t(1) < h_t(2)$ for any t and $c \in [c^*, \beta q\mu]$.*

Proof: See Appendix. ∎

Remark: When $\beta = 1$, since $v(1,0) = \int \max\{v(1,0), q\xi\} dF(\xi)$ and $F(\xi) < 1$ for $\xi < 1$, we have $v(1,0) \geq q\xi$ for all ξ, hence $v(1,0) = q$. Similarly we have $v(2,0) = 2q$. Thus we get $D(0) = 0$ for $\beta = 1$.

7. Case with Large Search Cost

Here, using (3.1), (3.3) and (3.2), we examine the case of $c > \beta q\mu$. First, suppose $c \geq \beta(1-p^i)q\mu/(1-p)$. Then we get $\beta v_0(j) - c = \beta(1-p^j)q\mu/(1-p) - c \leq 0$ for $j \leq i$, hence

$$z_1(j) = \max\{\beta v_0(j) - c,\ \beta z_0(j)\} = \max\{\beta v_0(j) - c,\ 0\} = 0,\ j \leq i. \tag{7.1}$$

Therefore, we obtain for $1 \leq j \leq i$

$$\begin{aligned} u_1(j,w) &= \max\{z_1(j), pu_1(j-1,w) + qw + (1-p)z_1(j-1)\} \\ &= pu_1(j-1,w) + qw \\ &= p\max\{z_1(j-1), pu_1(j-2,w) + qw + (1-p)z_1(j-2)\} + qw \\ &\ \vdots \\ &= \frac{1-p^j}{1-p}qw. \end{aligned} \tag{7.2}$$

The above also holds for $j = 0$. Accordingly, it follows for $j \leq i$ that

$$z_2(j) = \max\left\{\beta \int u_1(j,\xi)dF(\xi) - c,\ \beta z_1(j)\right\} = \max\left\{\beta\frac{1-p^j}{1-p}q\mu - c, 0\right\} = 0. \tag{7.3}$$

Repeatedly applying the same procedure yields $z_t(j) = \beta z_{t-1}(j) = 0$ for $t \geq 1$ and $j \leq i$, that is, $\beta v_t(j) - c \leq \beta z_t(j)$ for all t and $j \leq i$. As a result, in this case, the optimal policy is not to go shooting at all.

Next, suppose $\beta q\mu < c < \beta(1-p^i)q\mu/(1-p)$. Let $c = \beta(1-p^{\kappa(c)})q\mu/(1-p)$, from which we have

$$\kappa(c) = \log_p\left(1 - \frac{(1-p)c}{\beta q\mu}\right). \tag{7.4}$$

Then, for $i \geq \kappa(c)$, we get $z_1(i) = \max\{\beta v_0(i) - c,\ \beta z_0(i)\} = \beta v_0(i) - c$, hence

$$z_2(i) = \max\{\beta v_1(i) - c,\ \beta z_1(i)\} = \max\{\beta v_1(i) - c,\ \beta(\beta v_0(i) - c)\} = \beta v_1(i) - c \tag{7.5}$$

because $v_t(i)$ is nondecreasing in i. Repeating the same procedure, we get $z_t(i) = \beta v_{t-1}(i) - c$ for $t \geq 1$ and $i \geq \kappa(c)$, i.e., $\beta v_t(i) - c \geq \beta z_t(i)$ for all t and $i \geq \kappa(c)$. If the number of bullets remaining decreases less than $\kappa(c)$ during hunting, then it follows from the same reason as in the case of $c \geq \beta(1-p^i)q\mu/(1-p)$ that $z_t(i) = 0$ for all t.

Hence, in this case, the optimal decision policy becomes identical to the case of $c \leq \beta q\mu$ until the number of bullets remaining, i, decreases less than $\kappa(c)$ during shooting. If the number i becomes less than $\kappa(c)$, then it is optimal not to go shooting from the next time point on. However, if he has not got the present target yet and it still remains, then he should continue to fire until at least one of the three events stated in Section 6(c) occurs.

8. Case of Infinite Bullets

Since both $u_t(i,w)$ and $v_t(i)$ are increasing and upper-bounded in i for any t, they have *finite* limits as $i \to \infty$ for any t. Let $u_t(w) = \lim_{i\to\infty} u_t(i,w)$ and $v_t = \lim_{i\to\infty} v_t(i)$. Now we have for $t \geq 1$

$$\begin{aligned} &pu_t(w) + qw + (1-p)(\beta v_{t-1} - c) \\ &\quad = p\max\{\beta v_{t-1} - c,\ pu_t(w) + qw + (1-p)(\beta v_{t-1} - c)\} + qw + (1-p)(\beta v_{t-1} - c) \\ &\quad \geq qw + \beta v_{t-1} - c \geq \beta v_{t-1} - c, \end{aligned} \tag{8.1}$$

which means that it is always optimal to continue shooting until he gets the target or it escapes. Accordingly, it follows that

$$u_t(w) = pu_t(w) + qw + (1-p)(\beta v_{t-1} - c). \tag{8.2}$$

Hence we get for $t \geq 1$

$$u_t(w) = \frac{qw}{1-p} + \beta v_{t-1} - c. \tag{8.3}$$

Here, as final conditions, we immediately obtain $u_0(w) = qw/(1-p)$ and $v_0 = q\mu/(1-p)$ from (3.6) and (3.7), respectively. By induction starting with them, the following can be shown for $t \geq 1$

$$u_t(w) = \begin{cases} \dfrac{qw}{1-p} + t\Big(\dfrac{q\mu}{1-p} - c\Big), & \beta = 1, \\ \dfrac{qw}{1-p} + \dfrac{1-\beta^t}{1-\beta}\Big(\dfrac{\beta q\mu}{1-p} - c\Big), & \beta < 1, \end{cases} \tag{8.4}$$

$$v_t = \begin{cases} \dfrac{(t+1)q\mu}{1-p} - tc, & \beta = 1, \\ \dfrac{1-\beta^{t+1}}{1-\beta}\dfrac{q\mu}{1-p} - \dfrac{1-\beta^t}{1-\beta}c, & \beta < 1. \end{cases} \tag{8.5}$$

Now, in the proof of Lemma 9, we have already mentioned that $h_t = \lim_{i\to\infty} h_t(i) = 0$ for $t \geq 1$.

9. Results of Numerical Examples

In this section, using numerical examples, we will examine the relationship of critical value $h_t(i)$ with t, i, and parameters, β, q, r, and c, where a discrete uniform distribution function with 101 mass points, equally spaced on $[0,1]$, is used. Below, let us summarize the implications that are deduced from Figures 3,4, and A1 to A4 (see Appendix) illustrating the relation between $h_t(i)$ (vertical axis) and i (horizontal axis).

(a) If $c = 0$, then the critical value $h_t(i)$ is always nonincreasing in i for all $t \geq 1$ (Theorem 2, Figure 3(a,d)).

(b) If $c > 0$, then $h_t(i)$ is not always decreasing in i (Figure 3(b,c,e)). We see that the position of the maximal value,[¶] if it exists, shifts to the right as the planning horizon becomes longer. The maximal value does not always exist; in fact, as seen in Figure 3(c), it doesn't appear if the planning horizon is less than or equal to five.

(c) Figure 3(f) shows that $h_t(i)$ is nonincreasing in i even if c is positive. Now we proved in Lemma 8 that $h_t(i)$ is nonincreasing in i for all $t \geq 1$ if and only if $D_t(c) \geq 0$ for all t (see (6.2)). The results of the calculation of $D_t(c)$ for $t = 0, 1, \ldots, 5000$ are as shown in Table 1[‖], illustrating that $D_t(c)$ seems to converge to a positive number $0.104\cdots$ for $\beta = 0.9$ and $c = 10^{-2}$. However, it becomes negative for large t when $\beta = 1$ and $c = 10^{-2}$. Even if $\beta = 1$ and $c = 10^{-5}$, $D_{3500}(c)$ is negative.

[¶]If $h_t(i)$ takes a maximal value at a certain $i = i_a$, then $h_t(i_a - 1) < h_t(i_a)$ and $h_t(i_a + 1) < h_t(i_a)$.

[‖]The distribution used in the example is a continuous uniform one on $[0,1]$.

(d) If $q = 1$, then $h_t(i)$ becomes a nonincreasing function of i for each case in Figure 4; however, it is not proven that the property always holds.

(e) Regarding $h_t(i)$ as a function of β, q, r, and c, let $h_t(i, \beta, q, r, c) = h_t(i)$. The critical value $h_t(i, \beta, q, r, c)$ is nondecreasing in β and nonincreasing in r and c in Figures A1, A3, and A4. These properties don't conflict with our intuition; however, they are not also verified.

(f) Figure A2 shows that $h_t(i, \beta, q, r, c)$ is not always monotone in q. In fact, in Figure A2(b), if $t = 5$, $i = 3$, $\beta = 1$ and $c = 0.1$, then we have $h_t(i, \beta, 0.3, r, c) = 0.419 \cdots$,

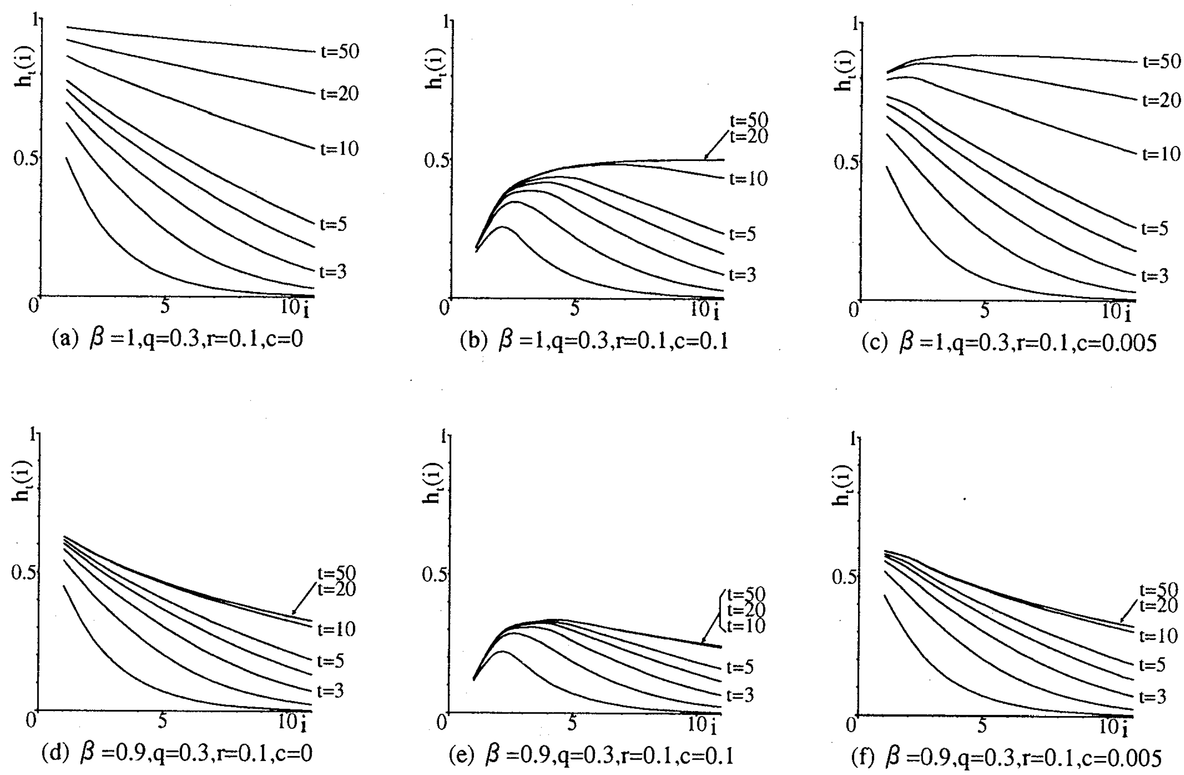


Figure 3. Relations between $h_t(i)$ and i

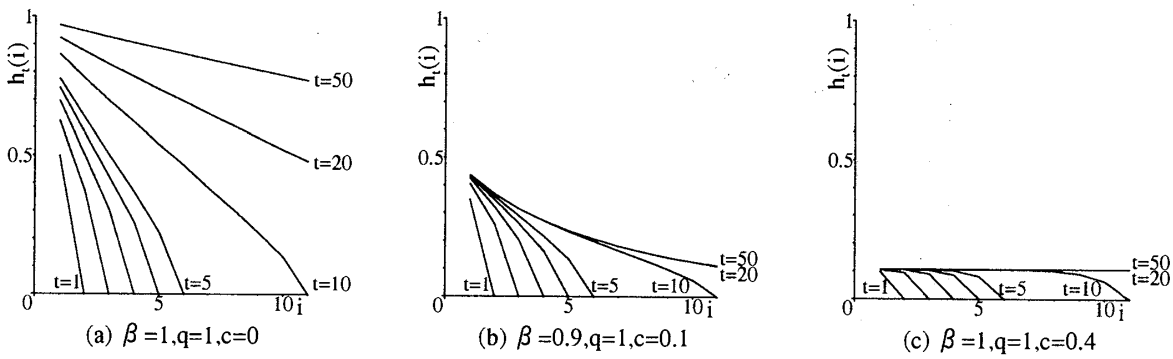


Figure 4. Relations between $h_t(i)$ and i for $q = 1$

Table 1. Values of $D_t(c)$ ($q = 0.9$, $r = 0.1$)

t	$\beta = 0.9$, $c = 10^{-2}$	$\beta = 1$, $c = 10^{-2}$	$\beta = 1$, $c = 10^{-4}$	$\beta = 1$, $c = 10^{-5}$
0	0.36755000	0.40850000	0.40949000	0.40949900
1	0.19188098	0.21291981	0.21395330	0.21396271
2	0.15190272	0.16463288	0.16586718	0.16587843
3	0.13178598	0.13661378	0.13805264	0.13806577
4	0.12041073	0.11747495	0.11911669	0.11913169
5	0.11373292	0.10330834	0.10515039	0.10516724
10	0.10474813	0.06461562	0.06742139	0.06744734
50	0.10410564	0.00998829	0.01818530	0.01828057
100	0.10410564	0.00015102	0.00946883	0.00964782
500	0.10410564	-0.00194995	0.00126448	0.00194734
1000	0.10410564	-0.00194995	0.00007398	0.00085488
1500	0.10410564	-0.00194995	-0.00015226	0.00045219
2000	0.10410564	-0.00194995	-0.00018908	0.00024057
2500	0.10410564	-0.00194995	-0.00019427	0.00011673
3000	0.10410564	-0.00194995	-0.00019494	0.00004239
3500	0.10410564	-0.00194995	-0.00019502	-0.00000193
4000	0.10410564	-0.00194995	-0.00019503	-0.00002805
4500	0.10410564	-0.00194995	-0.00019503	-0.00004304
5000	0.10410564	-0.00194995	-0.00019503	-0.00005149

$h_t(i, \beta, 0.6, r, c) = 0.470\cdots$ and $h_t(i, \beta, 0.9, r, c) = 0.437\cdots$. Such property leads us to the existence of two critical values in terms of the hitting probability in the following sense. Suppose there exist several hunters whose abilities for hunting, evaluated by each hitting probability q, are different, and a target of value w appears. Let q_* and q^* be the two hitting probabilities corresponding to the target value w as been indicated in Figure 5. Then, the optimal decision policy is to shoot for hunters with $q \leq q_*$, and not to shoot for hunters with $q_* < q < q^*$, and again to shoot for hunters with $q^* \leq q$.

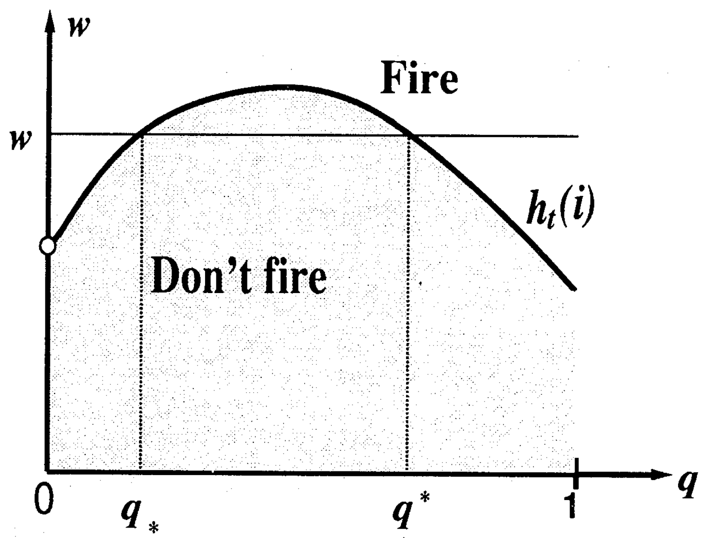


Figure 5. Existence of q_* and q^*

(g) The critical value $h_t(i)$ can be guessed to be always nondecreasing in t for all $i \geq 1$ as seen in Figures 3 and 4. This can be proven in the case of $\beta = 1$ and $c = 0$, but unfortunately, cannot be verified in other cases.

10. Conclusions

• *Optimal Decision Policy*

First, we outline the case of $c \leq \beta q\mu$. If the hunter has at least one bullet, then it is optimal to go hunting. Suppose he sees a target of value w when he has i bullets and t periods remain. Then, the following can be said:

(a) If $w < h_t(i)$, then it is optimal not to shoot and to search for the next target by paying a search cost c.
(b) If $w \geq h_t(i)$, $w \geq h_t(i-1)$, ..., $w \geq h_t(i-k)$ and $w < h_t(i-k-1)$ for a certain k such as $0 \leq k \leq i-1$, then it is optimal to continue to fire until at least one of the following three events occurs: he gets the target, the number of remaining bullets becomes less than $i-k$ (i.e., up to at most k bullets) or it runs away.
(c) If $w \geq h_t(k)$ for all $k = 1, 2, \ldots, i$, then it is optimal to continue to fire until at least one of the three events stated in Section 6(c) occurs.

Next, we summarize the policy in the case for $c > \beta q\mu$. When there remain t periods and i bullets, the optimal policy is as follows:

(a) If $i \leq \kappa(c)$, then it is optimal not to go hunting over the whole planning horizon.
(b) If $i > \kappa(c)$, then it is optimal to go hunting until the number of bullets in hand becomes less than $\kappa(c)$, and the optimal decision policy for firing is the same as the case of $c \leq \beta q\mu$. Suppose the number becomes less than $\kappa(c)$ during hunting. Then, he should not go shooting from the next time point on. However, if he has not got the present target yet and it still remains, then it is optimal to continue to fire until at least one of the three events stated in Section 6(c) occurs.

• *Properties of the Critical Value $h_t(i)$*

We have obtained the conclusion that, if $c = 0$, then $h_t(i)$ is nonincreasing in i for all $t \geq 1$, and if $c > 0$, then we get the following results:

(a) If $h_t(i_a) < h_t(i_a+1)$ for a certain i_a, then there exists $i_b > i_a$ such that $h_t(i_b) > h_t(i_b+1)$.
(b) If $\beta < 1$, then there exists $c_* > 0$ for which $h_t(i, c)$ is nonincreasing in i for all t and $c \in [0, c_*]$.
(c) If $\beta < 1$ and $p > 0$, then there exist $c^* \in [c_*, \beta q\mu)$ and $i_a \geq 2$ for which $h_t(i, c)$ strictly increases on $1 \leq i \leq i_a$ for all $t \geq 1$ and $c^* \leq c \leq \beta q\mu$.

11. Some Limitations

We have discussed a stochastic sequential allocation problem with search cost where a shoot-look-shoot policy is employed. However, it is sure that some of the assumptions restrict the problem. We will suggest the following provisions to relax these restrictions: (a) Hitting probability q and escaping probability r depend on the total number of bullets he has shot so far at a present target. (b) In this paper, the action of shooting results in only one of two outcomes; "get the target" or "don't get the target". As a variation of this problem, we can consider the case where part of the target value w can be got in firing a bullet. Then, there still remains a decision whether to continue to shoot in order to gain the rest of the target value or not. (c) It may happen that more than two targets appear at the same time. (d) He can replenish some bullets by paying a cost.

Acknowledgments

The author would like to thank Associate Professor S. Ikuta of University of Tsukuba and the referees for their constructive suggestions and comments.

Appendix

Proof of Lemma 7: We only prove the case that $h_t(i)$ is strictly decreasing in i. For the case that $h_t(i)$ is nonincreasing in i, it can be proven by replacing $>$ with $\geq$ in the proof below.

From the hypothesis of the lemma, Lemma 3, and Theorem 1, the following relations hold for a given t

$$h_t(i) = \beta(v_{t-1}(i) - v_t(i-1))/q, \ i \geq 2, \tag{A.1}$$

$$\beta(2v_{t-1}(1) - v_{t-1}(2)) - c > 0, \tag{A.2}$$

$$2v_{t-1}(i) - v_{t-1}(i-1) - v_{t-1}(i+1) > 0, \ i \geq 1. \tag{A.3}$$

First, we examine the case of $i = 1$. Since $v_t(i)$ is the expectation of $u_t(i, w)$ and $h_t(1) > h_t(2)$ from the assumption, we get

$$v_t(1) = \int_0^{h_t(1)} (\beta v_{t-1}(1) - c)dF(\xi) + \int_{h_t(1)}^1 q\xi dF(\xi), \tag{A.4}$$

$$\begin{aligned} v_t(2) = & \int_0^{h_t(1)} \max\{\beta v_{t-1}(2) - c, \, pu_t(1,\xi) + q\xi + (1-p)(\beta v_{t-1}(1) - c)\} dF(\xi) \\ & + \int_{h_t(1)}^1 (pu_t(1,\xi) + q\xi + (1-p)(\beta v_{t-1}(1) - c)) dF(\xi) \\ = & \int_0^{h_t(1)} \max\{\beta v_{t-1}(2) - c, \, q\xi + \beta v_{t-1}(1) - c\} dF(\xi) \\ & + \int_{h_t(1)}^1 ((1+p)q\xi + (1-p)(\beta v_{t-1}(1) - c)) dF(\xi). \end{aligned} \tag{A.5}$$

Then they yield

$$2v_t(1) - v_t(0) - v_t(2) = 2v_t(1) - v_t(2) = \int_0^{h_t(1)} A_t(1,\xi)dF(\xi) + \int_{h_t(1)}^1 B_t(1,\xi)dF(\xi), \tag{A.6}$$

where

$$\begin{aligned} A_t(1,\xi) &= 2(\beta v_{t-1}(1) - c) - \max\{\beta v_{t-1}(2) - c, \, q\xi + \beta v_{t-1}(1) - c\} \\ &= \min\{\beta(2v_{t-1}(1) - v_{t-1}(2)) - c, \, \beta v_{t-1}(1) - c - q\xi\}, \end{aligned} \tag{A.7}$$

$$B_t(1,\xi) = 2q\xi - (1+p)q\xi - (1-p)(\beta v_{t-1}(1) - c) = (1-p)(q\xi - ((\beta v_{t-1}(1) - c)). \tag{A.8}$$

From (A.2) and $h_t(1) = (\beta v_{t-1}(1) - c)/q$, we have $A_t(1,\xi) > 0$ for $0 \leq \xi < h_t(1)$ and $B_t(1,\xi) > 0$ for $h_t(1) < \xi \leq 1$. Furthermore, we easily obtain $A_t(1, h_t(1)) = B_t(1, h_t(1)) = 0$. Because $F(w)$ does not concentrate on only $w = h_t(1)$, we get $2v_t(1) - v_t(0) - v_t(2) > 0$.

Next, we consider the case of $i \geq 2$. In this case, dividing the interval of integration, $[0, 1]$, into three subintervals, we can express $v_t(i-1)$, $v_t(i)$ and $v_t(i+1)$ as, respectively,

$$\begin{aligned} v_t(i-1) = & \int_0^{h_t(i)} (\beta v_{t-1}(i-1) - c)dF(\xi) \\ & + \int_{h_t(i)}^{h_t(i-1)} (\beta v_{t-1}(i-1) - c)dF(\xi) + \int_{h_t(i-1)}^1 u_t(i-1,\xi)dF(\xi), \end{aligned} \tag{A.9}$$

$$\begin{aligned} v_t(i) = & \int_0^{h_t(i)} (\beta v_{t-1}(i) - c)dF(\xi) \\ & + \int_{h_t(i)}^1 (pu_t(i-1,\xi) + q\xi + (1-p)(\beta v_{t-1}(i-1) - c))dF(\xi) \\ = & \int_0^{h_t(i)} (\beta v_{t-1}(i) - c)dF(\xi) + \int_{h_t(i)}^{h_t(i-1)} (q\xi + \beta v_{t-1}(i-1) - c)dF(\xi) \\ & + \int_{h_t(i-1)}^1 (pu_t(i-1,\xi) + q\xi + (1-p)(\beta v_{t-1}(i-1) - c))dF(\xi), \end{aligned} \tag{A.10}$$

$$
\begin{aligned}
v_t(i+1) &= \int_0^{h_t(i)} \max\{\beta v_{t-1}(i+1)-c,\ pu_t(i,\xi)+q\xi+(1-p)(\beta v_{t-1}(i)-c)\}dF(\xi) \\
&\qquad + \int_{h_t(i)}^1 (pu_t(i,\xi)+q\xi+(1-p)(\beta v_{t-1}(i)-c))dF(\xi) \\
&= \int_0^{h_t(i)} \max\{\beta v_{t-1}(i+1)-c,\ q\xi+\beta v_{t-1}(i)-c\}dF(\xi) \\
&\quad + \int_{h_t(i)}^1 ((1+p)q\xi+p^2u_t(i-1,\xi) \\
&\qquad +(1-p)p(\beta v_{t-1}(i-1)-c)+(1-p)(\beta v_{t-1}(i)-c))dF(\xi) \\
&= \int_0^{h_t(i)} \max\{\beta v_{t-1}(i+1)-c,\ q\xi+\beta v_{t-1}(i)-c\}dF(\xi) \\
&\quad + \int_{h_t(i)}^{h_t(i-1)} ((1+p)q\xi+p\beta v_{t-1}(i-1)+(1-p)\beta v_{t-1}(i)-c)dF(\xi) \\
&\quad + \int_{h_t(i-1)}^1 ((1+p)q\xi+p^2u_t(i-1,\xi) \\
&\qquad +(1-p)p\beta v_{t-1}(i-1)+(1-p)\beta v_{t-1}(i)-(1-p^2)c)dF(\xi). \qquad (A.11)
\end{aligned}
$$

Then we have for $i \geq 2$

$$
2v_t(i)-v_t(i-1)-v_t(i+1) = \int_0^{h_t(i)} A_t(i,\xi)dF(\xi)+\int_{h_t(i)}^{h_t(i-1)} B_t(i,\xi)dF(\xi)+\int_{h_t(i-1)}^1 C_t(i,\xi)dF(\xi), \qquad (A.12)
$$

where

$$
\begin{aligned}
A_t(i,\xi) &= 2(\beta v_{t-1}(i)-c)-(\beta v_{t-1}(i-1)-c)-\max\{\beta v_{t-1}(i+1)-c,\ q\xi+(\beta v_{t-1}(i)-c)\} \\
&= \min\{\beta(2v_{t-1}(i)-v_{t-1}(i-1)-v_{t-1}(i+1)),\ \beta(v_{t-1}(i)-v_{t-1}(i-1))-q\xi\}, \qquad (A.13) \\
B_t(i,\xi) &= 2(q\xi+(\beta v_{t-1}(i-1)-c))-(\beta v_{t-1}(i-1)-c) \\
&\quad -((1+p)q\xi+(1-p)\beta v_{t-1}(i)+p\beta v_{t-1}(i-1)-c) \\
&= (1-p)(q\xi+\beta v_{t-1}(i-1)-\beta v_{t-1}(i)), \qquad (A.14) \\
C_t(i,\xi) &= 2(q\xi+pu_t(i-1,\xi)+(1-p)(\beta v_{t-1}(i-1)-c))-u_t(i-1,\xi) \\
&\quad -((1+p)q\xi+p^2u_t(i-1,\xi)+(1-p)p\beta v_{t-1}(i-1)+(1-p)\beta v_{t-1}(i)-(1-p^2)c) \\
&= (1-p)q\xi-(1-p)^2u_t(i-1,\xi) \\
&\quad +(2-p)(1-p)\beta v_{t-1}(i-1)-(1-p)\beta v_{t-1}(i)-(1-p)^2c. \qquad (A.15)
\end{aligned}
$$

Immediately we get $A_t(i,\xi)>0$ for $0 \leq \xi < h_t(i)$ and $B_t(i,\xi)>0$ for $h_t(i)<\xi \leq h_t(i-1)$ from (A.1) and (A.3).

Below, using induction, we shall deduce $C_t(i,\xi)>0$ for $i \geq 2$ and $h_t(i-1)<\xi \leq 1$. If $i=2$, then we get for $h_t(1)<\xi \leq 1$

$$
\begin{aligned}
C_t(2,\xi) &= (1-p)q\xi-(1-p)^2u_t(1,\xi)+(2-p)(1-p)\beta v_{t-1}(1)-(1-p)\beta v_{t-1}(2)-(1-p)^2c \\
&= (1-p)q\xi-(1-p)^2q\xi+(2-p)(1-p)\beta v_{t-1}(1)-(1-p)\beta v_{t-1}(2)-(1-p)^2c \\
&= (1-p)p(q\xi-\beta v_{t-1}(1)+c)+(1-p)(\beta(2v_{t-1}(1)-v_{t-1}(2))-c)>0 \qquad (A.16)
\end{aligned}
$$

owing to (A.2), $h_t(1)=(\beta v_{t-1}(1)-c)/q$, and $u_t(1,\xi)=q\xi$ for $\xi \geq h_t(i)$. Assume $C_t(i-1,\xi)>0$ for $h_t(i-2) \leq \xi \leq 1$. Noting that $u_t(i-1,\xi)=pu_t(i-2,\xi)+q\xi+(1-p)(\beta v_{t-1}(i-2)-c)$ for $\xi \geq h_t(i-1)$, we get for $i \geq 3$

$$
\begin{aligned}
C_t(i,\xi) &= (1-p)q\xi-(1-p)^2(pu_t(i-2,\xi)+q\xi+(1-p)(\beta v_{t-1}(i-2)-c)) \\
&\qquad +(2-p)(1-p)\beta v_{t-1}(i-1)-(1-p)\beta v_{t-1}(i)-(1-p)^2c \\
&= (1-p)pq\xi-(1-p)^2pu_t(i-2,\xi)-(1-p)^3\beta v_{t-1}(i-2) \\
&\qquad +(2-p)(1-p)\beta v_{t-1}(i-1)-(1-p)\beta v_{t-1}(i)-(1-p)^2pc \qquad (A.17) \\
&= pC_t(i-1,\xi)+(1-p)\beta(2v_{t-1}(i-1)-v_{t-1}(i-2)-v_{t-1}(i)). \qquad (A.18)
\end{aligned}
$$

From the inductive assumption and (A.3), it follows that $C_t(i,\xi)>0$ for $h_t(i-2) \leq \xi \leq 1$. If $h_t(i-1)<\xi<h_t(i-2)$, then rearranging (A.17) by substituting $u_t(i-2,\xi)=\beta v_{t-1}(i-2)-c$

yields

$$\begin{aligned}C_t(i,\xi) &= (1-p)pq\xi - (1-p)^2p(\beta v_{t-1}(i-2)-c)\\&\quad -(1-p)^3\beta v_{t-1}(i-2) + (2-p)(1-p)\beta v_{t-1}(i-1) - (1-p)\beta v_{t-1}(i) - (1-p)^2pc\\&\geq (1-p)p\beta(v_{t-1}(i-1) - v_{t-1}(i-2)) - (1-p)^2p(\beta v_{t-1}(i-2)-c)\\&\quad -(1-p)^3\beta v_{t-1}(i-2) + (2-p)(1-p)\beta v_{t-1}(i-1) - (1-p)\beta v_{t-1}(i) - (1-p)^2pc\\&= (1-p)\beta(2v_{t-1}(i-1) - v_{t-1}(i-2) - v_{t-1}(i)) > 0 \qquad\qquad (A.19)\end{aligned}$$

from (A.1) and (A.3). Accordingly, we obtain by induction that $C_t(i,\xi) > 0$ for all $t \geq 0$, $i \geq 2$ and $h_t(i-1) \leq \xi \leq 1$. In addition, it follows that $A_t(i, h_t(i)) = B_t(i, h_t(i)) = 0$. From above, we have $2v_t(i) - v_t(i-1) - v_t(i+1) > 0$ for $i \geq 2$ since $F(w)$ does not concentrate on only $w = h_t(i)$. Eventually, it follows that $2v_t(i) - v_t(i-1) - v_t(i+1) > 0$ for $i \geq 1$ if $h_t(i)$ is strictly decreasing in i. ∎

Proof of Theorem 3: It is clear that $v_t(1,c)$ and $v_t(2,c)$ are continuous functions of $c \in [0, \beta q\mu]$ for any t. Now, since $v(1,c) \leq q \leq 1$ from Lemma 1(d), we have

$$\begin{aligned}v(1,c) - v_t(1,c) &= \int \max\{\beta v(1,c) - c, q\xi\}dF(\xi) - \int \max\{\beta v_{t-1}(1,c) - c, q\xi\}dF(\xi)\\&\leq \int \max\{\beta(v(1,c) - v_{t-1}(1,c)), 0\}dF(\xi)\\&= \beta(v(1,c) - v_{t-1}(1,c))\\&\ \ \vdots\\&\leq \beta^t(v(1,c) - v_0(1,c)) < \beta^t v(1,c) \leq \beta^t. \qquad\qquad (A.20)\end{aligned}$$

Because $\beta < 1$, $v_t(1,c)$ uniformly converges to $v(1,c)$ as $t \to \infty$, hence $v(1,c)$ is also a continuous function of c. Similarly, also $v(2,c)$ can be shown to be a continuous function of c. Eventually, it follows that $D_t(c)$ for any t and $D(c)$ are both continuous functions of c.

Now we obtain $0 \leq h_t(2,0) \leq h_t(1,0) < 1$ from Lemma 2 and Theorem 2. Hence, we can express $D_t(0)/\beta$ as the sum of three integrals similar to (A.12),

$$\begin{aligned}D_t(0)/\beta &= 2v_t(1,0) - v_t(2,0) - 0/\beta\\&= \int_0^{h_t(2,0)} \beta(2v_{t-1}(1,0) - v_{t-1}(2,0))dF(\xi)\\&\quad + \int_{h_t(2,0)}^{h_t(1,0)} (\beta v_{t-1}(1,0) - q\xi)dF(\xi) + \int_{h_t(1,0)}^1 (1-p)(q\xi - \beta v_{t-1}(1,0))dF(\xi)\\&= q\int_0^{h_t(2,0)} (h_t(1,0) - h_t(2,0))dF(\xi)\\&\quad + q\int_{h_t(2,0)}^{h_t(1,0)} (h_t(1,0) - \xi)dF(\xi) + q\int_{h_t(1,0)}^1 (1-p)(\xi - h_t(1,0))dF(\xi). \qquad (A.21)\end{aligned}$$

Since the first and second terms of the above expression are nonnegative and the third term is positive, we get $D_t(0) > 0$ for $t \geq 0$. Furthermore, (A.21) also holds for $t \to \infty$ because of $h(1,0) = \beta v(1,0)/q \leq \beta < 1$, so is also obtained $D(0) > 0$.

On the other hand, it can be easily shown for any t and any p

$$v_t(1, \beta q\mu) = q\mu \qquad\qquad (A.22)$$

by induction, hence $v(1, \beta q\mu) = q\mu$. If $p > 0$, then we get

$$\begin{aligned}v(2,\beta q\mu) &\geq v_t(2,\beta q\mu)\\&= \int \max\{\beta v_{t-1}(2,\beta q\mu) - \beta q\mu,\ (1+p)q\xi\}dF(w) \geq (1+p)q\mu. \qquad (A.23)\end{aligned}$$

From (A.22) and (A.23), it follows that

$$D(\beta q\mu) \leq D_t(\beta q\mu) \leq 2\beta q\mu - \beta(1+p)q\mu - \beta q\mu < 0. \qquad\qquad (A.24)$$

If $p = 0$, then we can inductively verify $v_t(2, \beta q\mu) = q\mu$ for all $t \geq 1$ using $v_0(2, \beta q\mu) = q\mu$ as a final condition, so $v(2, \beta q\mu) = q\mu$. Consequently, we obtain for any t

$$D(\beta q\mu) = D_t(\beta q\mu) = 2\beta q\mu - \beta q\mu - \beta q\mu = 0. \qquad\qquad (A.25)$$

Accordingly, we come to:

(a) Let $c_{*t} = \min\{c \mid D_t(c) = 0, c \in [0, \beta q\mu]\}$. Then, $c_{*t} > 0$ for all t.
(b) Let $c_t^* = \max\{c \mid D_t(c) = 0, c \in [0, \beta q\mu]\}$. Then, $c_t^* \in [c_{*t}, \beta q\mu)$ for all t and $p > 0$, and $c_t^* = \beta q\mu$ for all t and $p = 0$.

Furthermore, define $c_* = \inf_{t \geq 0} c_{*t}$ and $c^* = \sup_{t \geq 0} c_t^*$. Then, it follows for $p > (=) 0$ that $0 < c_* \leq c^* < (=) \beta q\mu$ due to $D(0) > 0$ and $D(\beta q\mu) < (=) 0$. Therefore, the length of the interval $[0, c_*]$ is not zero for any p and $\beta < 1$, and the length of the interval $[c^*, \beta q\mu]$ is not zero for $p > 0$ and $\beta < 1$. With this, the proof is complete. ∎

References

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

[2] S. Ikuta : The Optimal Stopping Problem in which the Sum of the Accepted Offer's Value and the Remaining Search Budget is an Objective Function, *Journal of the Operations Research Society of Japan*, Vol.35, (1992), 172-193.

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

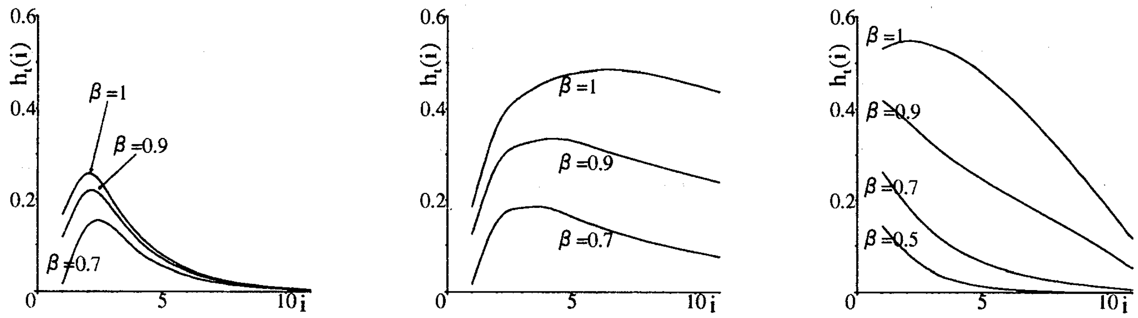
[4] D. V. Mastran and C. J. Thomas : Decision Rule for Attacking Targets of Opportunity, *Naval Research Logistics Quarterly*, Vol.20, (1973), 661-672.

[5] 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.

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

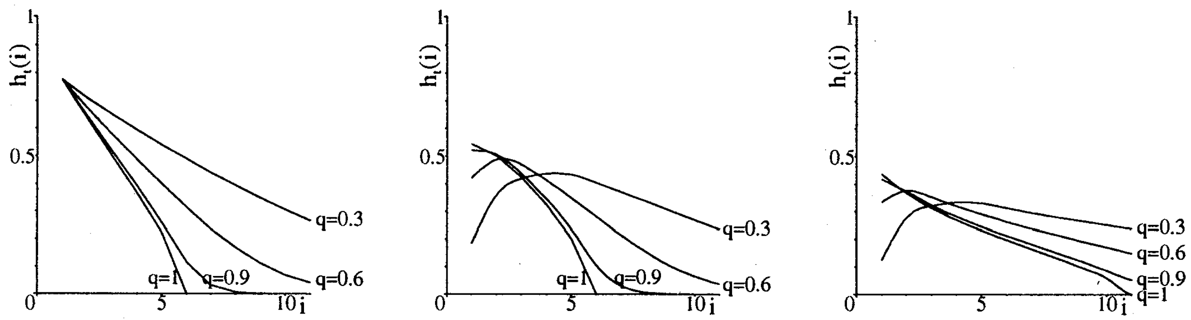
[7] M. Sakaguchi : A Sequential Allocation Problem for Randomly Appearing Targets, *Mathematica Japonica*, Vol.21, (1976), 89-103.

Masahiro SATO
Doctoral Program
in Policy and Planning Sciences
University of Tsukuba
Tsukuba, Ibaraki 305, Japan
e-mail: msato@shako.sk.tsukuba.ac.jp



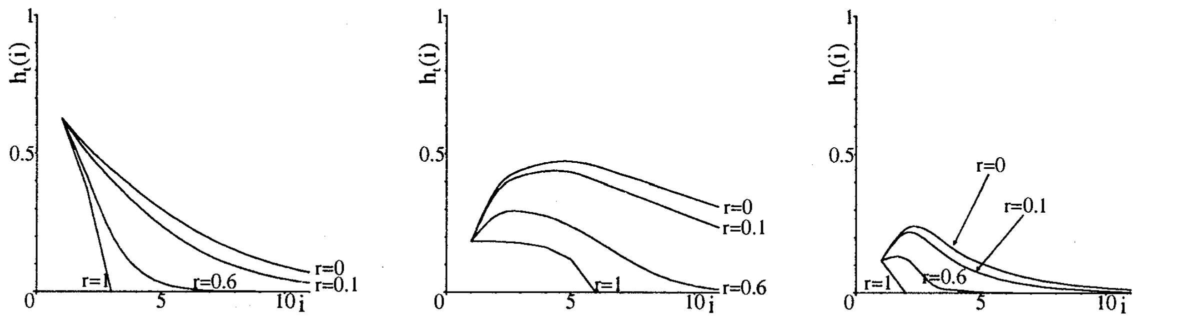
(a) q=0.3,r=0.1,c=0.1,t=1 (b) q=0.3,r=0.1,c=0.1,t=10 (c) q=0.9,r=0.1,c=0.1,t=10

Figure A1. Sensitivity of β



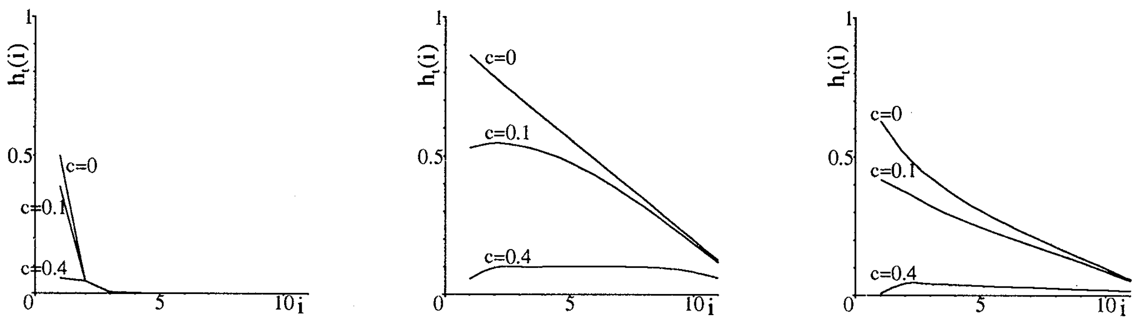
(a) β =1,r=0.1,c=0,t=5 (b) β =1,r=0.1,c=0.1,t=5 (c) β =0.9,r=0.1,c=0.1,t=10

Figure A2. Sensitivity of q



(a) β =1,q=0.3,c=0,t=2 (b) β =1,q=0.3,c=0.1,t=5 (c) β =0.9,q=0.3,c=0.1,t=1

Figure A3. Sensitivity of r



(a) β =1,q=0.9,r=0.1,t=1 (b) β =1,q=0.9,r=0.1,t=10 (c) β =0.9,q=0.9,r=0.1,t=10

Figure A4. Sensitivity of c