

No. 74 (80 - 12)
A Sequential Selection Process
and
its Applications

by
Seizo Ikuta *)

April, 1980

This paper is released for limited circulation in order to facilitate discussion and invite criticism. It is wholly tentative in character and should not be referred to in publication without the permission of the author.

*) Institute of Socio-Economic Planning, University of Tsukuba, Sakura-mura, Niiharigun, Ibaraki-ken 305, Japan

ABSTRACT

Consider a decision model with a finite action space $A = \{A^r\}$. An m -vector θ , random variable with a known distribution function, is observed at each time point, and if an action A^r is chosen, then the reward of $x^r(\theta)$ is gained where $x^r(\theta)$ is a real-valued function of θ and z^r is a constant. The objective is to maximize the expected reward gained. The basic theorems are provided in two linear cases of $x^r(\theta) = a^r\theta$ for 1-dimensional vector $\theta = (\theta)$ and $x^r(\theta) = \theta_1 + \theta_2 + \dots + \theta_r$ where $\theta_1 \geq \theta_2 \geq \dots \geq \theta_r$ for all r ($\leq m$). A finite horizon discrete time decision process version of the above model will be defined which has a finite state space $I = \{i\}$ and a finite state-dependent action space $A_i = \{A_i^r\}$. In the process, if a vector θ is observed in state i , the immediate reward of $x_i^r(\theta)$ is gained and the next state becomes j with probability p_{ij}^r . Let v_j denote the total expected rewards starting from the next time with state j . Then an after-effect for choosing an action A_i^r when in state i becomes $z_i = \beta \sum p_{ij}^r v_j$. This version, called a sequential selection process, provides a very systematic approach to a certain class of dynamic stochastic decision problems, including as typical example an optimal stopping problem. It is shown that the dynamic model can be illustrated as a Markovian decision process (MDP) of special structure. In order to demonstrate the usefulness of the model, eleven real-world problems are taken as typical examples in its applications. By this it will be also revealed that the model gives more general approach, different from the conventional one, to these problems.

CONTENTS

	page
1. INTRODUCTION	1
2. LINEAR MODEL : CASE A	4
3. LINEAR MODEL : CASE B	9
4. SEQUENTIAL SELECTION PROCESS	11
4.1 Finite time horizon process	12
4.2 Infinite time horizon process	13
4.3 Policy improvement algorithm	15
5. APPLICATIONS	18
5.0 Preliminaries	18
5.1 Purchasing problem	22
5.2 Inventory problem	24
5.3 Warehouse problem	25
5.4 Buying-selling problem	30
5.5 Job search problem	34
5.6 Sequential stochastic assignment problem	36
5.7 Target attacking problem I	38
5.8 Target attacking problem II	39
5.9 Customer selection problem I	41
5.10 Customer selection problem II	44
5.11 Customer selection problem III	45
REFERENCES	48

DEFINITION 1	3
DEFINITION 2	3
COROLLARY 1	8
THEOREM 1	1
THEOREM 2	4
THEOREM 3	6
THEOREM 4	9
THEOREM 5	13
REMARK 1	2
REMARK 2	2
REMARK 3	6
REMARK 4	6
REMARK 5	9
REMARK 6	11
REMARK 7	13
REMARK 8	14
REMARK 9	15
REMARK 10	15
REMARK 11	15
REMARK 12	17
LEMMA 1	4
LEMMA 2	4
LEMMA 3	19
LEMMA 4	20

1. INTRODUCTION

Consider a decision model with a finite action space $A = \{A^r: r = 0, 1, \dots, k\}$ where k is a fixed non-negative integer. If an action A^r is chosen after observing an m -vector $\theta = (\theta_1, \theta_2, \dots, \theta_m)$, random variable with a known distribution function F , then the reward of $x^r(\theta) + z^r$ is gained where $x^r(\theta)$, called an immediate reward, is a real-valued measurable function and z^r , called an after-effect, is a given constant. Now by the term (k+1)-partition of m -dimensional space R^m we mean the set of $k+1$ mutually exclusive subsets B^r , $r = 0, 1, \dots, k$, of R^m which are Borel sets such that $\bigcup_{r=0}^k B^r = R^m$. Let us denote the space consisting of all $(k+1)$ -partitions of R^m by $G^m(k+1)$, called a partition space. By using a given $(k+1)$ -partition B in $G^m(k+1)$, we can prescribe a decision rule for the above decision model as follows. If an observed vector $\theta \in B^r$, choose an action A^r . Then the expected reward becomes $V(B) = \sum_{r=0}^k \int (x^r(\theta) + z^r) I(B^r) dF$ where $I(\cdot)$ is an indicator.

THEOREM 1

There exists the $(k+1)$ -partition attaining the maximum of the expected reward on $G^m(k+1)$ and that such $(k+1)$ -partition D , called an optimal partition or OP for brevity, is given by $D = \{D^r: r = 0, 1, \dots, k\}$ where

$$(1.1) \quad D^r = \left\{ \theta: \begin{aligned} &x^r(\theta) - x^s(\theta) \geq z^s - z^r \text{ for } 0 \leq s < r \text{ and} \\ &x^r(\theta) - x^s(\theta) > z^s - z^r \text{ for } r < s \leq k \end{aligned} \right\}$$

$$0 \leq r \leq k$$

PROOF Let $H^r = x^r(\theta) + z^r$. Then (1.1) is written as $D^r = \{H^r \geq H^s$ for $0 \leq s < r$ and $H^r > H^s$ for $r < s \leq k\}$. Two arbitrary sets D^r and D^q with $r > q$ are exclusive because $D^r \subset \{H^r \geq H^q\}$ and $D^q \subset \{H^q > H^r\}$. Hence D^0, D^1, \dots, D^k are also mutually exclusive. For any vector θ in R^m let $H^q = H^{q'} = \dots = H^{q''} > H^u, H^{u'}, \dots, H^{u''}$ where $(q, q',$

... , q'' , u , u' , ... , u'') is a permutation of integers $0, 1, \dots, k$ such as $q > q' > \dots > q''$ and $u > u' > \dots > u''$. Then $q = 0$ leads to $H^0 > H^s$ for $0 < s \leq k$, $0 < q < k$ to $H^q \geq H^s$ for $0 \leq s < q$ and $H^q > H^s$ for $q < s \leq k$, and $q = k$ to $H^k \geq H^s$ for $0 \leq s < k$. Thus the vector θ must belong to one of D^0, D^1, \dots, D^k . This implies $R^m = \bigcup_{r=0}^k D^r$. Hence it follows that D is a $(k+1)$ -partition of R^m . For any B in $G^{m(k+1)}$ we have $V(D) - V(B) = \sum_{r,s=0}^k \int (H^r - H^s) I(D^r \cap B^s) dF \geq 0$ owing to $H^r \geq H^s$ on $D^r \cap B^s$. Hence $V(D)$ becomes the maximum of the expected reward. Letting the maximum be denoted by v , we have

$$(1.2) \quad v = \sum_{r=0}^k \int (x^r(\theta) + z^r) I(D^r) dF$$

Q.E.D.

We shall call the above expectation (1.2) a fundamental expression or FE for brevity.

REMARK 1 It is clear that some of inequality signs \geq and $>$ in (1.1) may be replaced by $>$ and \geq , respectively, in such an arbitrary way as not to break the mutual exclusiveness of D^0, D^1, \dots, D^k . The simplest but most frequently used one of such replacements is to replace \geq and $>$ exhaustively.

REMARK 2 We will frequently encounter the models in which an immediate rewards have a constant term, that is, $x^r(\theta) + h^r$. It is easily shown by a glance at (1.1) and (1.2) that the model is just the same as one having an immediate reward $x^r(\theta)$ and after-effect $z^r + h^r$.

In the present paper only the two special cases below are discussed in which immediate rewards are of linear forms.

CASE A

$x^r(\theta) = a^r \theta$, $r = 0, 1, \dots, k$, where $\theta = (\theta)$ is an l -vector and θ is a random variable having a distribution function F with a finite expectation E . The model is said to be monotone if a^r is monotone in r and to be strictly monotone if a^r is strictly monotone in r .

CASE B

$x^0(\theta) = 0$ and $x^r(\theta) = \theta_1 + \theta_2 + \dots + \theta_r$ for $1 \leq r \leq k$ where $k \leq m$ and $\theta = (\theta_1, \theta_2, \dots, \theta_m)$, m -vector such that $\theta_1 \geq \theta_2 \geq \dots \geq \theta_m$. Let a distribution function of θ_r be denoted by F_r and its expectation by E_r , assumed to be finite.

Here we shall provide in advance two definitions and one lemma necessary for the discussions in the subsequent sections.

DEFINITION 1

Let θ be a random variable having a distribution function F with a finite expectation E . Then for any measurable real-valued functions h and g of θ , define

1. $T(g, h) = \int (\theta - g) I(h < \theta) dF$
2. $T(g) = T(g, g)$
3. $\hat{T}(g) = E(g) + T(g)$

where I and $E(\)$ represent an indicator and expectation operator of θ , respectively. In addition, for $g \equiv \infty$ or $g \equiv -\infty$ define

4. $T(-\infty) = \infty$, $T(\infty) = 0$, $\hat{T}(-\infty) = E$, $\hat{T}(\infty) = \infty$

Furthermore let T , \hat{T} , and E associated with a distribution function F_γ having a parameter γ be denoted by T_γ , \hat{T}_γ , and E_γ , respectively. In the present paper the above functionals T , \hat{T} , T_γ , \hat{T}_γ , are named T-functional generically. If g is a constant function, the functional T is called conventionally a shortage function.

DEFINITION 2

In general a given functional H is said to be increasing (strictly increasing) if $u < u'$ implies $H(u) \leq (<) H(u')$. Throughout the paper we shall write this as $H \nearrow (\nearrow)$ in u or simply as $H \nearrow (\nearrow)$. For a decreasing functional, symbols \searrow and \searrow are used.

The next lemma is clear.

LEMMA 1

Consider two sequences x_1, x_2, \dots, x_N and y_1, y_2, \dots, y_N where x_n are all either positive or negative. Then if $y_n/x_n \nearrow(\searrow)$ in n , we have

- (a) $y_1/x_1 \leq(\geq) (y_1 + y_2 + \dots + y_N)/(x_1 + x_2 + \dots + x_N) \leq(\geq) y_N/x_N$
- (b) $(y_1 + y_2 + \dots + y_n)/(x_1 + x_2 + \dots + x_n) \nearrow(\searrow)$ in n
- (c) $(y_n + y_{n+1} + \dots + y_N)/(x_n + x_{n+1} + \dots + x_N) \nearrow(\searrow)$ in n

LEMMA 2

$T(g) \geq T(g, h)$ for any h .

PROOF Since $\{h < \theta\} = \{g \leq h < \theta\} \cup \{h < g < \theta\} \cup \{h \leq \theta \leq g\}$ and $\{g < \theta\} = \{g \leq h < \theta\} \cup \{h < g < \theta\} \cup \{g \leq \theta \leq h\}$, we get $T(g) - T(g, h) = \int ((\theta - g)I(g \leq \theta \leq h) - (\theta - g)I(h \leq \theta \leq g))dF \geq 0$.

Q.E.D.

2. LINEAR MODEL : CASE A

THEOREM 2

Consider a strictly monotone model with $k \geq 1$. Then define

$$(2.1) \quad c^r = (z^{r-1} - z^r)/(a^r - a^{r-1}), \quad 1 \leq r \leq k$$

(a) If $a^r \nearrow(\searrow)$ in r and $c^r \nearrow(\searrow)$ in r , then the OP is reduced to a set of $k + 1$ intervals, $r = 0, 1, \dots, k$,

$$(2.2) \quad D^r = (c^r, c^{r+1})$$

$$(2.3) \quad (D^r = (c^{r+1}, c^r))$$

where $c^0 = -\infty(\infty)$ and $c^{k+1} = \infty(-\infty)$, and the FE to

$$(2.4) \quad v = a^0 E + z^0 + \sum_{r=1}^k (a^r - a^{r-1}) T(c^r)$$

$$(2.5) \quad (v = a^k E + z^k - \sum_{r=1}^k (a^r - a^{r-1}) T(c^r))$$

(b) If $a^r \nearrow(\nearrow)$ in r and $c^r \not\downarrow(\not\downarrow)$ in r , then the OP is provided by

$$(2.6) \quad D^0 = (-\infty, c], \quad D^k = (c, \infty)$$

$$(2.7) \quad (D^0 = (c, \infty), \quad D^k = (-\infty, c])$$

where D^r is empty for $0 < r < k$ and

$$(2.8) \quad c = (z^0 - z^k)/(a^k - a^0)$$

, and the FE by

$$(2.9) \quad v = a^0 E + z^0 + (a^k - a^0) T(c)$$

$$(2.10) \quad (v = a^k E + z^k - (a^k - a^0) T(c))$$

PROOF Define $c^{rs} = c^{sr} = (z^s - z^r)/(a^r - a^s)$ for $r \neq s$. Then we have $c^r = c^{r,r-1} = c^{r-1,r}$. Letting $\Delta z^r = z^{r-1} - z^r$ and $\Delta a^r = a^r - a^{r-1}$, we have

$$c^{rs} = (\Delta z^{s+1} + \Delta z^{s+2} + \dots + \Delta z^r) / (\Delta a^{s+1} + \Delta a^{s+2} + \dots + \Delta a^r) \quad s < r \quad (*)$$

$$c^{rs'} = (\Delta z^{r+1} + \Delta z^{r+2} + \dots + \Delta z^{s'}) / (\Delta a^{r+1} + \Delta a^{r+2} + \dots + \Delta a^{s'}) \quad r < s' \quad (**)$$

Let us prove only the case of $a^r \nearrow$. Then the OP (1.1) is transformed into

$$D^0 = \{-\infty < \theta < c^{0s'} \text{ for } 0 < s' \leq k\}$$

$$D^r = \{c^{rs} \leq \theta < c^{rs'} \text{ for } 0 \leq s < r < s' \leq k\} \quad 0 < r < k \quad (***)$$

$$D^k = \{c^{ks} \leq \theta < \infty \text{ for } 0 \leq s < k\}$$

(a) If $c^r = \Delta z^r / \Delta a^r \nearrow$, then from lemma 1(a) we get $c^{rs} \leq \Delta z^r / \Delta a^r = c^{r,r-1} = c^r$ for $s < r$ and $c^{rs'} > \Delta z^{r+1} / \Delta a^{r+1} = c^{r,r+1} = c^{r+1}$ for $r < s'$. Hence the sets (***) can be reduced to the intervals $D^0 = (-\infty, c^1)$, $D^r = (c^r, c^{r+1})$ for $0 < r < k$, $D^k = (c^k, \infty)$. The intervals as OP may be changed to (2.2), referring to remark 1. Arranging (1.2) by inserting these intervals in yields (2.4).

(b) If $c^r \not\downarrow$, then applying lemma 1(b,c) to (***) and (*) yields $c^{0s'} \geq (\Delta z^1 + \Delta z^2 + \dots + \Delta z^k) / (\Delta a^1 + \Delta a^2 + \dots + \Delta a^k) = c^{0k} = c$ and $c^{ks} \leq (\Delta z^1 + \Delta z^2 + \dots + \Delta z^k) / (\Delta a^1 + \Delta a^2 + \dots + \Delta a^k) = c^{k0} = c$ which

lead to $D^0 = (-\infty, c)$ and $D^k = (c, \infty)$. Consequently D^r must be empty for $1 < r < k$. The intervals D^0 and D^k may be changed to (2.6). The FE (2.9) is obtained by inserting the intervals in (1.1).

When $a^r \not\leq$ in r , the theorem is clear from reversing the sequence a^0, a^1, \dots, a^k as a^k, a^{k-1}, \dots, a^0 and applying (a) and (b) above.

Q.E.D

REMARK 3 When $k = 1$, the monotonicity requirement of c^r in theorem 2 becomes unnecessary. When $k = 0$, the OP is given by $D = \{D^0\}$ where $D^0 = R^m$, hence the FE becomes $v = a^0 E + z^0$.

REMARK 4 When an immediate reward is deterministic, that is, $x^r = a^r$, it can be regarded as $x^r = a^r \theta$ in which θ is a random variable having a unit distribution function F with parameter $\lambda = 1$, that is, $F(\theta) = 0$ for $\theta < 1$ and $F(\theta) = 1$ for $1 \leq \theta$. Then for a real number c , we have $T(c) = 1 - c$ and $\hat{T}(c) = 1$ if $c < 1$, and $T(c) = 0$ and $\hat{T}(c) = c$ if $1 \leq c$. Then for $k \geq 1$,

(a) " $a^r \not\leq$ in r and $c^r \not\leq$ in r " leads to the choice of action A^r if $c^r < 1 \leq c^{r+1}$ ($c^{r+1} < 1 \leq c^r$), and

(b) " $a^r \not\leq$ in r and $c^r \leq$ in r " leads to the choice of A^0 if $1 \leq c$ ($c < 1$) and to the choice of A^k if $c < 1$ ($1 \leq c$).

THEOREM 3

Let the after-effects be a function of vector $a = (a^0, a^1, \dots, a^k)$, that is, $z^r(a)$, $0 \leq r \leq k$. Assume that

(1) For all r , c^r is a -independent for any strictly monotone model.

(2) $a^r = a^s$ leads always to $z^r(a) = z^s(a)$.

(3) If $a^r \not\leq$ in r , then $z^r(b_\epsilon)$ is right continuous at $\epsilon = 0$ for all r in which $b_\epsilon = a + \epsilon \mathbf{1}$ ($a - \epsilon \mathbf{1}$) for a sufficiently small positive number ϵ so that $b_\epsilon^r \not\leq$ in r where $\mathbf{1} = (1, 1, \dots, 1)$.

Then theorem 2 holds also for a monotone model.

PROOF Suppose a monotone model \mathcal{M} with $a^r \nearrow$ in r where let $a^{r(m)} = a^{r(m)+1} = \dots = a^{r(m+1)-1} < a^{r(m+1)}$, $m = 0, 1, \dots, M$ and $r(0) = 0$ and $a^{r(M+1)} = \infty$. Then for a strictly monotone model with $b_\varepsilon = (b_\varepsilon^0, b_\varepsilon^1, \dots, b_\varepsilon^k)$ where $b_\varepsilon^r = a^r + \varepsilon$, we have from (2.4),

$$v_\varepsilon = b_\varepsilon^0 E + z_\varepsilon^0 + \sum_{r=1}^k (b_\varepsilon^r - b_\varepsilon^{r-1}) T(c^r) \quad (*)$$

Then clearly

$$\lim_{\varepsilon \rightarrow 0^+} v_\varepsilon = a^0 E + z^0 + \sum_{r=1}^k (a^r - a^{r-1}) T(c^r) \quad (**)$$

$$= a^{r(0)} E + z^{r(0)} + \sum_{r=1}^M (a^{r(m)} - a^{r(m-1)}) T(c^{r(m)}) \quad (***)$$

Now since it is clear that the above expression (***) provides the FE for the strictly monotone model with the vector $\bar{a} = (a^{r(0)}, a^{r(1)}, \dots, a^{r(M)})$ and that this model is exactly the same as the model \mathcal{M} due to the assumption (2), we have $v = (**)$ where v represents the maximum of the expected reward for the model \mathcal{M} . Therefore it follows that exactly the same form of FE can be obtained also for the monotone model \mathcal{M} .

The OP for the above strictly monotone model with b_ε is given by $D^{r(m)} = (c^{r(m)}, c^{r(m+1)})$, $m = 0, 1, \dots, M$, from (2.2) where $c^{r(0)} = -\infty$ and $c^{r(M+1)} = \infty$. Since all actions A^r with $r(m) \leq r < r(m+1)$ are indifferent in the model due to the assumption (2), we may regard the subinterval $D^r = (c^r, c^{r+1})$, $r(m) \leq r < r(m+1)$, of $D^{r(m)}$ as the interval for choosing the action A^r .

Similarly for all other cases, that is, "both a^r and c^r in \searrow in r ", " $a^r \nearrow$ in r and $c^r \searrow$ in r ", " $a^r \searrow$ in r and $c^r \nearrow$ in r ".

Q.E.D.

Next we shall consider the case that θ is a random variable having a distribution function F_γ with a parameter γ and that after-effects depend on an observation θ . Such model is referred to as a γ -dependent model. Letting an after-effect be denoted by $z^{r\theta}$ for all r , we can easily obtain the corollary below, corresponding to theorem 2.

COROLLARY 1

Consider a strictly monotone model with $k \geq 1$, and define

$$(2.11) \quad c^{r\theta} = (z^{r-1,\theta} - z^{r\theta}) / (a^r - a^{r-1}) \quad 0 < r \leq k$$

(a) If $a^r \nearrow(\searrow)$ in r and $c^{r\theta} \nearrow(\searrow)$ in r , then the OP becomes for $0 \leq r \leq k$

$$(2.12) \quad D^r = \{\theta : c^{r\theta} < \theta \leq c^{r+1,\theta}\}$$

$$(2.13) \quad (D^r = \{\theta : c^{r+1,\theta} < \theta \leq c^{r\theta}\})$$

where $c^{0\theta} \equiv -\infty(\infty)$ and $c^{k+1,\theta} \equiv \infty(-\infty)$, and the FE is provided by

$$(2.14) \quad v^Y = a^0 E + E_Y(z^0) + \sum_{r=1}^k (a^r - a^{r-1}) T_Y(c^r)$$

$$(2.15) \quad (v^Y = a^k E_Y + E_Y(z^k) - \sum_{r=1}^k (a^r - a^{r-1}) T_Y(c^r))$$

Let

$$(2.16) \quad b^{r\theta} = \theta - c^{r\theta}$$

If $b^{r\theta} \nearrow$ in θ , define

$$(2.17) \quad c^r = \inf\{\theta : b^{r\theta} \geq 0\}$$

Then the OP (2.12)((2.13)) is reduced to the intervals

$$(2.18) \quad D^r = (c^r, c^{r+1})$$

$$(2.19) \quad (D^r = (c^{r+1}, c^r))$$

If $b^{r\theta} \searrow$ in θ , define

$$(2.20) \quad c^r = \sup\{\theta : b^{r\theta} \geq 0\}$$

Then the OP (2.12)((2.13)) is reduced to the intervals

$$(2.21) \quad D^r = (c^{r+1}, c^r)$$

$$(2.22) \quad (D^r = (c^r, c^{r+1}))$$

(b) If $a^r \nearrow(\searrow)$ in r and $c^{r\theta} \searrow(\nearrow)$ in r , define

$$(2.23) \quad c^\theta = (z^{0\theta} - z^{k\theta}) / (a^k - a^0)$$

Then the OP becomes

$$(2.24) \quad D^0 = \{ \theta : \theta \leq c^\theta \} \quad D^k = \{ \theta : c^\theta < \theta \}$$

$$(2.25) \quad (D^0 = \{ \theta : c^\theta < \theta \} \quad D^k = \{ \theta : \theta \leq c^\theta \})$$

where $D^r = \phi$ for $0 < r < k$, and the FE is provided by

$$(2.26) \quad v^Y = a^0 E_Y + E_Y(z^0) + (a^k - a^0) T_Y(c)$$

$$(2.27) \quad (v^Y = a^k E_Y + E_Y(z^k) - (a^k - a^0) T_Y(c))$$

Let

$$(2.28) \quad b^\theta = \theta - c^\theta$$

If $b^\theta \nearrow$ in θ , define

$$(2.29) \quad c = \inf\{ \theta : b^\theta \geq 0 \}$$

Then the OP (2.24)(2.25) is reduced to

$$(2.30) \quad D^0 = (-\infty, c) \quad D^k = (c, \infty)$$

$$(2.31) \quad (D^0 = (c, \infty) \quad D^k = (-\infty, c])$$

If $b^\theta \searrow$ in θ , define

$$(2.32) \quad c = \sup\{ \theta : b^\theta \geq 0 \}$$

Then the OP (2.24)((2.25)) is reduced to

$$(2.33) \quad D^0 = (c, \infty) \quad D^k = (-\infty, c)$$

$$(2.34) \quad (D^0 = (-\infty, c) , \quad D^k = (c, \infty))$$

REMARK 5 With slight modifications, remarks 3 and 4 and theorem 3 hold also for the above γ -dependent model.

3. LINEAR MODEL : CASE B

THEOREM 4

Suppose $k \geq 1$ and define

$$(3.1) \quad c^r = z^{r-1} - z^r, \quad 1 \leq r \leq k$$

Then if $c^r \nearrow$ in r , the OP becomes

$$(3.2) \quad \begin{cases} D^0 &= \{\theta : \theta_1 \leq c^1\} \\ D^r &= \{\theta : c^r < \theta_r \text{ and } \theta_{r+1} \leq c^{r+1}\} \\ D^k &= \{\theta : c^k < \theta_k\} \end{cases} \quad 0 < r < k$$

, and the FE is given by

$$(3.3) \quad v = z^0 + \sum_{r=1}^k T_r(c^r)$$

where $T_r(g) = \int (\theta - g) I(g \leq \theta) F_r$ (see definition 1)

PROOF In the case, the FE (1.2) is transformed into

$$v = z^0 + \sum_{r=1}^k \int (x^r - y^r) I(D^r) dF \quad (*)$$

where $y^r = z^0 - z^r = c^1 + c^2 + \dots + c^r$, and the OP (1.1) into $D^r = \{x^r - x^s > y^r - y^s \text{ for } 0 \leq s < r \text{ and } x^r - x^s \geq y^r - y^s \text{ for } r < s \leq k\}$, $0 \leq r \leq k$ (see remark 1). Now define $B^0 = \{\theta_1 \leq c^1\}$, $B^r = \{c^r < \theta_r, \theta_{r+1} \leq c^{r+1}\}$ for $0 < r < k$, and $B^k = \{c^k < \theta_k\}$. The next two hold from the monotonicity of c^r and θ_r . (a) If $0 \leq r < s \leq k$, then $\theta_{r+1} \leq c^{r+1}$ leads to $\theta_{r+1} + \theta_{r+2} + \dots + \theta_{s-1} + \theta_s \leq c^{r+1} + c^{r+2} + \dots + c^{s-1} + c^s$, hence to $x^s - x^r \leq y^s - y^r$, or $x^r - x^s \geq y^r - y^s$. (b) if $0 \leq s < r \leq k$, then $\theta_r > c^r$ leads to $\theta_{s+1} + \theta_{s+2} + \dots + \theta_{r-1} + \theta_r > c^{s+1} + c^{s+2} + \dots + c^{r-1} + c^r$, hence to $x^r - x^s > y^r - y^s$. Then, for " $r = 0$ in (a) ", " $0 < r < k$ in both (a) and (b) ", and " $r = k$ in (b) ", we can obtain, respectively, $D^0 \supset B^0$, $D^r \supset B^r$ for $0 < r < k$, and $D^k \supset B^k$. Now clearly we have $D^0 \subset \{x^0 - x^1 \geq y^0 - y^1\} = \{\theta_1 \leq c^1\} = B^1$, $D^r \subset \{x^r - x^{r-1} > y^r - y^{r-1} \text{ and } x^r - x^{r+1} \geq y^r - y^{r+1}\} = \{c^r < \theta_r \text{ and } \theta_{r+1} \leq c^{r+1}\} = B^r$ for $0 < r < k$, and $D^k \subset \{x^k - x^{k+1} > y^k - y^{k+1}\} = \{\theta_k > c^k\} = B^k$. Hence it follows that $D^r = B^r$ for all r .

Next define $S^r = D^r \cup D^{r+1} \cup \dots \cup D^k$ for $0 \leq r \leq k$. Obviously we have $S^k = D^k = \{c^k < \theta_k\}$. If $S^r = \{c^r < \theta_r\}$ for any r such as $1 < r \leq k$, then $S^{r-1} = D^{r-1} \cup S^r = \{c^{r-1} < \theta_{r-1} \text{ and } \theta_r \leq c^r\} \cup \{c^r < \theta_r\} = (\{c^{r-1} < \theta_{r-1}\} \cup \{c^r < \theta_r\}) \cap (\{\theta_r \leq c^r\} \cup \{c^r < \theta_r\}) = \{c^{r-1} < \theta_{r-1}\}$ because of $\{c^{r-1} < \theta_{r-1}\} \supset \{c^r < \theta_r\}$. Hence by induction we have $S^r = \{c^r < \theta_r\}$ for all r . Substituting $x^r = \theta_1 + \theta_2 + \dots + \theta_r$ and $y^1 = c^1 +$

$c^2 + \dots + c^r$ into (*) and arranging the resultant expressions by using S^r just defined above yield the second term of (3.3). Q.E.D

REMARK 6 Define $T_r(\infty) = 0$ for $r > m$ and $c_i^r = \infty$ for $r > k$. Then (3.3) can be written as

$$(3.4) \quad v = z^0 + \sum_{r=1}^{\infty} T_r(c^r)$$

, which is valid even for $k = 0$.

4. SEQUENTIAL SELECTION PROCESS

As a dynamic version of the decision model in the previous sections, we shall consider the following discrete-time stochastic decision process. It has a finite state space $I = \{i: i = 0, 1, 2, \dots, N\}$ and a finite state-dependent action space $A_i = \{A_i^r: r = 0, 1, \dots, k_i\}$, $i \in I$. At each time point an m -vector θ is observed. Here it is assumed that m -vectors $\theta, \theta', \theta'', \dots$ at successive time points are random variables, independently and identically distributed according to a given distribution function F . If an action A_i^r is taken when an m -vector θ is observed in state i , then an immediate reward $x_i^r(\theta)$ is gained and the state of the next time is chosen according to a known transition probability p_{ij}^r , $j \in I$. The objective is to find the optimal decision strategy, maximizing the total expected reward or expected reward per period over a given time horizon, possibly infinite. From now on we shall refer to the above decision process as a sequential selection process.

In section 5, we shall demonstrate how both linear models, CAESs A and B, for the sequential selection process can be successfully applied to many real-world problems of a certain class. A purchasing problem (18), asset disposing problem (11), and so on, are included among them as typical examples. Some of these problems have been

already investigated in detail by many authors, however, reviewing the original literatures cited in the section would reveal that the application of our sequential selection process provides a new approach, more systematic and general than the conventional one, to them. I must stress that honor for some of conclusions and results derived in these problems should be accorded to the authors who have initially introduced and developed them.

4.1 Finite time horizon process

We shall begin with a finite time horizon process. Here for convenience let a time be measured backward from the terminating time point $t = 0$ of the process. Let B_i represent a decision rule in state i . Then we shall call the set $\bar{B} = \{B_i : i \in I\}$ a decision policy and a time sequence of decision policies, $\bar{B}(t), \bar{B}(t-1), \dots, \bar{B}(0)$, a decision strategy over times t to 0 , denoted by $B(t)$. Here $\bar{B}(0)$ is given as the final(, or terminating) condition.

Now by $v_i(t)$ we shall denote the maximum total expected reward(if exist) starting from time t with state i , where $v_i(0)$ is assumed to be uniquely determined by a given final condition $\bar{B}(0)$. The decision strategy attaining $v_i(t)$ for all i is referred to as an optimal decision strategy, and the decision rules and decision policies associated with the optimal decision strategy as an optimal decision rule and optimal decision policy, respectively. Now suppose that $v_i(t-1), v_i(t-2), \dots, v_i(1)$ exist for all i . Then the after-effect for action A_i^r at time t becomes

$$(4.1) \quad z_i^r(t) = \beta \sum_{j \in I} p_{ij}^r v_j(t-1)$$

where β is a discount factor ($0 < \beta \leq 1$). It is easily proved by use of induction and theorem 1 that there exist $v_i(t)$ for all i and all t . From now on we shall very often suppress a time parameter t and vector θ in $x_i^r(\theta), v_i(t), D_i(t), D_i^r(t), z_i^r(t), \dots$ and simply write them as $x_i^r, v_i, D_i, D_i^r, z_i^r, \dots$, and furthermore in general let

functions $f(t)$, $f(t-1)$, $f(t-2)$, be written as f , f' , f'' ,

Then noticing z_i^r is expressed as

$$(4.2) \quad z_i^r = \beta \sum_{j=0}^N p_{ij}^r v_j^r \quad 0 \leq i \leq N, \quad 0 \leq r \leq k,$$

, the FE (1.2) can be transformed into

$$(4.3) \quad v_i = R_i(D_i) + \beta \sum_{j \in I} Q_{ij}(D_i) v_j^r$$

where

$$(4.4) \quad R_i(D_i) = \sum_{r=0}^{k_i} x_i^r I(D_i^r) dF$$

$$(4.5) \quad Q_{ij}(D_i) = \sum_{r=0}^{k_i} p_{ij}^r I(D_i^r) dF$$

The above expressions (4.3) to (4.5) indicate that our sequential selection process is the special case of a discrete time Markovian decision process(MDP) with a finite time horizon. Here an action space, immediate reward, transition probability, and optimal action usually defined in MDP correspond to $G_i^m(k_i+1)$, $R_i(B_i)$, $Q_{ij}(B_i)$, and $D_i(t)$, respectively.

REMARK 7

Let

$$(4.6) \quad y_i(t) = \beta(v_i(t) - v_{i-1}(t-1)) \quad 1 \leq i \leq N,$$

Then both $z_i^r(t) - z_i^s(t)$ and $y_i(t)$ can be expressed by use of linear combinations of $y_1(t)$, $y_2(t)$,, $y_N(t)$.

4.2 Infinite time horizon process

For an infinite time horizon process, the corollary 6.6 and theorems 6.17 and 6.18 in (20) are available, which can be rewritten as follows for our sequential selection process.

THEOREM 5

Suppose a discount factor $\beta < 1$. Then the limit of $v_i(t)$ as $t \rightarrow \infty$ exists for all i , denoted by v_i . Assume that the partition space

$G_i^m(k_i+1)$ is finite for all i . Then

(a) (see theorem 6.6 in [20]) The limit vector (v_0, v_1, \dots, v_N) is the unique solution to

$$(4.7) \quad v_i = R_i(D_i) + \beta \sum_{j=0}^N Q_{ij}(D_i) v_j, \quad 0 \leq i \leq N$$

in which D_i is given by (1.1) where

$$(4.8) \quad z_i^r = \beta \sum_{r=0}^{k_i} p_{ij}^r v_j$$

Then the optimal stationary strategy is provided by d^∞ where $d = \{D_i : 0 \leq i \leq N\}$.

(b) (see theorem 6.17 and 6.18 in [20]) For the limits v_i define $Y_i = \beta(v_i - v_{i-1})$, and suppose $\beta^{-1}Y_i$ is bounded uniformly in β for all i . Then there exist u_i and g satisfying

$$(4.9) \quad g + u_i = R_i(D_i) + \sum_{j=0}^N Q_{ij}(D_i) u_j, \quad 0 \leq i \leq N$$

, in which D_i is given by (1.1) where

$$(4.10) \quad z_i^r = \sum_{r=0}^{k_i} p_{ij}^r u_j$$

Furthermore we have $\beta_n(v_i^n - v_{i-1}^n) \rightarrow u_i - u_{i-1}$ for some sequence $\beta_n \rightarrow 1^-$ where v_i^n is a limit of $v_i(t)$ associated with the discount factor β_n . Then for any i we have

$$(4.11) \quad g = \lim_{n \rightarrow \infty} (1 - \beta_n) v_i^n$$

, which provides the maximum expected reward per period for an infinite time horizon process without discounting (that is, $\beta = 1$). The g is attained by d^∞ , called an optimal stationary strategy, where $d = \{D_i : i = 0, 1, \dots, N\}$.

We shall refer to D_i in the above theorem as a limit OP, which can be obtained in the finite number of steps by using the Howard's policy iteration algorithm (8).

REMARK 8 If a distribution function F takes positive value only on a finite domain, then the partition space $G_i^m(k_i+1)$ can be defined to be finite for all i .

REMARK 9 In order to avoid the confusions from employing the same term " limit " in both discounted process ($\beta < 1$) and undiscounted process ($\beta = 1$), two terms β -limit and l-limit will be very often used in the subsequent sections, for example, as β -limit y_i , β -limit OP, l-limit OP, Particularly in (b) in theorem 5 we shall define $y_i = u_i - u_{i-1}$ and call it l-limit y_i .

REMARK 10 Limit y_i inherits the structural form of $y_i(t)$. For example, if $y_i(t)$ is increasing in i (is convex) for all t , then it follows that both β -limit y_i and l-limit y_i (if exist) are also increasing in i (is convex). Moreover limit $z_i^r - z_i^s$ can be expressed by use of linear combinations of limits y_1, y_2, \dots, y_N (see remark 7).

REMARK 11 For any $i \neq i'$, the difference expressions $v_i - v_{i'}$, derived from (4.7) and $u_i - u_{i'}$, from (4.9) become of just the same form.

4.3 Policy iteration algorithm

Here the policy iteration algorithm is provided for seeking the limit OP in both CASES A and B. Whenever an infinite time horizon problem is discussed in the present paper, we will assume the next two things by implication.

(1) A limit c_i^r satisfies the monotonicity requirement in theorems 2 or 4.

(2) The partition space $G_i^m(k_i+1)$ is finite for all i .

Under these conditions we may restrict the partition space to its subspace $\bar{G}_i^m(k_i+1)$ consisting of such (k_i+1) -partitions B_i as defined below.

CASE A

Let the three requirements (1), (2), and (3) in theorem 3 be satisfied. Then

If both a_i^r and $c_i^r \nearrow$ in r , $B_i^r = (b_i^r, b_i^{r+1})$ $0 \leq r \leq k_i$
 (If both a_i^r and $c_i^r \searrow$ in r , $B_i^r = (b_i^{r+1}, b_i^r)$ $0 \leq r \leq k_i$)
 where $b_i^r \nearrow(\searrow)$ in r , $b_i^0 = -\infty(\infty)$ and $b_i^{k_i+1} = \infty(-\infty)$
 If $a_i^r \nearrow$ in r and $c_i^r \searrow$ in r , $B_i^0 = (-\infty, b_i)$ and $B_i^{k_i} = (b_i, \infty)$
 (If $a_i^r \searrow$ in r and $c_i^r \nearrow$ in r , $B_i^0 = (b_i, \infty)$ and $B_i^{k_i} = (-\infty, b_i]$)
 where $B_i^r = \phi$ for $0 < r < k_i$.

CASE B

$$(4.12) \quad \begin{cases} B_i^0 = \{ \theta : \theta_1 \leq b_i^1 \} \\ B_i^r = \{ \theta : b_i^r < \theta_r \text{ and } \theta_{r+1} \leq b_i^{r+1} \}, \quad 0 < r < k_i \\ B_i^{k_i} = \{ \theta : b_i^{k_i} < \theta_{k_i} \} \end{cases}$$

where $b_i^r \nearrow$ in r .

Then the policy iteration algorithm for both CASES becomes as follows.

In discounting case($0 < \beta < 1$)

Step 1 Set $v_i = 0$ for all i .

Step 2(policy improvement)

In CASE A, compute c_i^r using (2.1). When both a_i^r and $c_i^r \nearrow(\searrow)$ in r , find the b_i^r maximizing

$$(4.13) \quad \sum_{r=1}^{k_i} (a_i^r - a_i^{r-1}) T(c_i^r, b_i^r)$$

$$(4.14) \quad \left(\sum_{r=1}^{k_i} (a_i^{r-1} - a_i^r) T(c_i^r, b_i^r) \right)$$

subject to $b_i^r \nearrow(\searrow)$ in r . When $a_i^r \nearrow(\searrow)$ in r and $c_i^r \searrow(\nearrow)$ in r , compute c_i from (2.8) and set $b_i = c_i$ for all i .

In CASE B, compute c_i^r using (3.1) and find the b_i^r maximizing

$$(4.15) \quad \sum_{r=1}^{k_i} T_r(c_i^r, b_i^r)$$

subject to $b_i^r \nearrow$ in r .

Step 3 (value determination) Solve a system of equations (4.7) in which D_i is replaced by $B_i \subset \bar{G}_i^m(k_i+1)$. Then $R_i(B_i)$ and $Q_{ij}(B_i)$ are given as below where $E(b) = \int \theta I(\theta \leq b) dF$.

In CASE A, if both a_i^r and $c_i^r \nearrow(\searrow)$ in r ,

$$(4.16) \quad R_i(B_i) = \sum_{r=0}^{k_i} a_i^r (E(b_i^{r+1}) - E(b_i^r))$$

$$(4.17) \quad Q_{ij}(B_i) = \sum_{r=0}^{k_i} p_{ij}^r (F(b_i^{r+1}) - F(b_i^r))$$

$$(4.18) \quad \left(\begin{array}{l} R_i(B_i) = \sum_{r=0}^{k_i} a_i^r (E(b_i^r) - E(b_i^{r+1})) \\ Q_{ij}(B_i) = \sum_{r=0}^{k_i} p_{ij}^r (F(b_i^r) - F(b_i^{r+1})) \end{array} \right)$$

$$(4.19)$$

If $a_i^r \nearrow(\searrow)$ in r and $c_i^r \searrow(\nearrow)$ in r ,

$$(4.20) \quad R_i(B_i) = a_i^0 E(b_i) + a_i^{k_i} (E - E(b_i))$$

$$(4.21) \quad Q_{ij}(B_i) = p_{ij}^0 F(b_i) + p_{ij}^{k_i} (1 - F(b_i))$$

$$(4.22) \quad \left(\begin{array}{l} R_i(B_i) = a_i^0 (E - E(b_i)) + a_i^{k_i} E(b_i) \\ Q_{ij}(B_i) = p_{ij}^0 (1 - F(b_i)) + p_{ij}^{k_i} F(b_i) \end{array} \right)$$

$$(4.23)$$

In CASE B, we have, referring to the proof of theorem 4,

$$(4.24) \quad R_i(B_i) = \sum_{r=1}^{k_i} (E_r - E_r(b_i^r))$$

$$(4.25) \quad Q_{ij}(B_i) = \sum_{r=0}^{k_i} p_{ij}^r (F_{r+1}(b_i^{r+1}) - F_r(b_i^r))$$

where $F_0(b_i^0) = 0$ and $F_{k_i+1}(b_i^{k_i+1}) = 1$.

Go to step 2. If the solutions v_i from two successive steps 2 are identical, then the B_i associated with the solutions v_i provides the limit OP.

REMARK 12 The expressions (4.13), (4.14), and (4.15) are derived from arranging the right-hand of (4.7) and (4.9) by substituting (4.16) to (4.25) for $R_i(D_i)$ and $Q_{ij}(D_i)$. Now let c_i^r be said to be

regular if it has the monotonicity in r . It is clear from lemma 2 that the regularity of c_i^r yields the solution $b_i^r = c_i^r$, $0 \leq r \leq k_i$, to the maximization problems in step 2. On the other hand if c_i^r is not regular, then it can be directly solved by use of dynamic programming technique [1]. In step 2 for CASE A, when $a_i^r \nearrow(\searrow)$ in r and $c_i^r \searrow(\nearrow)$ in r , the right-hand of (4.7) are arranged as $(a_i^{k_i} - a_i^0)T(c_i, b_i) ((a_i^0 - a_i^{k_i})T(c_i, b_i))$. Clearly this is maximized at $b_i = c_i$ from lemma 2.

In undiscounting case($\beta = 1$)

In the case we shall assume that a chain with transition matrix $(Q_{ij}(B_i))$ is completely ergodic for all $B_i \in \bar{G}_i^m(k_i+1)$, $i = 0, 1, \dots, N$. Then the above algorithm may be only slightly changed as follows.

Step 1 Let $u_i = 0$ for all i .

Step 2(policy improvement) c_i^r is computed by use of (4.10).

Step 3(value determination) A system of equations to be solved is (4.9) instead of (4.7), in which any one of u_0, u_1, \dots, u_N is set zero.

5. APPLICATIONS

5.0 Preliminaries

(1) In a finite time horizon problem, the monotonicity of c_i^r for all t will be proved by use of structural forms, that is, non-negativity, non-positivity, or monotonicity, of y_i . And the structural form of y_i for all t will be in turn verified inductively by starting with the terminating conditions $y_i(0) = \beta(v_i(0) - v_{i-1}(0))$, inherent in each of the examples.

(2) In an infinite time horizon problem, the monotonicity of limit c_i^r will be determined using theorem 2 as a limiting monotonicity

of c_i^r for a discounted finite time horizon problem with the final condition $v_i(0) = 0$ for all i . Here notice that $y_i(0) = 0$ as the starting point of induction is \nearrow and \searrow in i and bounded uniformly in β for all i . Finiteness of partition space $G_i^m(k_i+1)$ is always assumed by implication.

(3) Distribution functions of all random variables defined in each example take the value 0 on $(-\infty, 0)$, hence the expectations of them always becomes ≥ 0 .

(4) Except for a γ -dependent model, random variables θ' , θ'' , θ''' , at successive times are independently and identically distributed.

(5) We frequently encounter a problem in which there may exist multiple action spaces $A_i^w = \{A_i^{rw} : r = 0, 1, \dots, k_i^w\}$, $0 \leq w \leq W_i$, and one of them appears with probability p_i^w when in state i where $\sum_w p_i^w = 1$. In the problem it is natural to consider that the immediate reward and transition probability may depend on an appearing action space. Consequently the after-effect and OP also may depend on the action space. Thus the FE is rewritten as $v_i = \sum_{w=0}^{W_i} p_i^w x$ (right-hand of the FE in the previous sections, dependent on w). In the case, however, any major changes in the relative formulas dose not take place and any new difficulties in its mathematical treatments do not occurs. Index w is always attached as superscript to the related notations.

(6) Next we shall clarify some properties of T-functional (see definition 1). They will play a central role in proving the monotonicity of c_i^r , essential to applications of theorems 2, 3, and 4, and corollary 1 to the examples in the sections 5.1 to 5.11. Let $()$ represent Gauss' symbol and $(x)^+ = \max\{0, x\}$.

LEMMA 3 (E.L. Lehmann [13])

For F_Y , $E_Y(g)$, and g in definition 1,

- (a) If $F_Y \searrow(\nearrow)$ in y and $g \nearrow$ in θ , then $E_Y(g) \nearrow(\searrow)$ in y , and
- (b) If $F_Y \searrow(\nearrow)$ in y and $g \searrow$ in θ , then $E_Y(g) \searrow(\nearrow)$ in y ,

LEMMA 4

Suppose $F(0) = 0$ and let L be a space of all real-valued non-negative measurable functions of θ . Then we have on L

- (a)
1. $T(g) \downarrow$ in g
 2. $T(0) = E$
 3. $E \geq T(g) \geq E - E(g)$
 4. $T(g) \geq 0$

- (b)
1. $\hat{T}(g) \uparrow$ in g
 2. $\hat{T}(0) = E$
 3. $E + E(g) \geq \hat{T}(g) \geq E$
 4. $\hat{T}(g) \geq E(g)$

- (c)
1. $|T(g) - T(g')| \leq E(|g - g'|)$
 2. $|\hat{T}(g) - \hat{T}(g')| \leq E(|g - g'|)$

(d) If $F(M) = 0$ for an $M > 0$, then $T(M) = 0$ and $\hat{T}(M) = M$.

(e) Letting $T_{ab}(g) = aT(g/a) - bT(g/b)$ for real numbers a and b such as $0 < a \leq b$, we have

1. $T_{ab}(g) \leq 0$
2. $T_{ab}(g) \uparrow$ in g
3. $T_{ab}(g) \geq (a - b)E$.

(f) For any non-negative real number g , define

$$\Delta T_r(g) = T_{r-1}(g) - T_r(g)$$

where T_r is a T -function(shortage function) associated with a distribution function F_r of θ_r in order statistics $\theta_1 \geq \theta_2 \geq \dots \geq \theta_m$ (m is fixed). Then we have

1. $\Delta T_r(g) \downarrow$ in g
2. $\Delta T_r(g) \geq 0$

(g) Suppose $F_r \downarrow(\uparrow)$ in y . Then we have

1. $T_y(g) \uparrow(\downarrow)$ in y if $\theta - g \uparrow$ in θ
2. $T_y(g) \downarrow(\uparrow)$ in y if $\theta - g \downarrow$ in θ
3. $\hat{T}_y(g) \uparrow(\downarrow)$ in y if $g + (\theta - g)^+ \uparrow$ in θ
4. $\hat{T}_y(g) \downarrow(\uparrow)$ in y if $g + (\theta - g)^+ \downarrow$ in θ
5. $T_y(g_y) \uparrow(\downarrow)$ in y if $g_y \downarrow(\uparrow)$ in y and $(\theta - g_y)^+ \uparrow$ in θ
6. $\hat{T}_y(g_y) \uparrow(\downarrow)$ in y if $g_y \uparrow(\downarrow)$ in y and $g_y + (\theta - g_y)^+ \uparrow$ in θ

PROOF (a1) For $g \geq g'$, we have $T(g) - T(g') \leq T(g) - T(g',g) = \int (g' - g)I(g \leq \theta)dF \leq 0$.

(b1) For $g \geq g'$, it follows that $\hat{T}(g) - \hat{T}(g') = E(g - g') + T(g) - T(g') \geq E(g - g') + T(g, g') - T(g') = E(g - g') + \int (g' - g)I(g' \leq \theta)dF = \int (g - g')I(\theta < g')dF \geq 0$.

(e1) We have $T_{ab}(g) = \int (a\theta - g)I(g < a\theta)dF - \int (b\theta - g)I(g < b\theta)dF \dots (*)$. Since $I(g < a\theta) \leq I(g < b\theta)$ due to $\{g < a\theta\} \subset \{g < b\theta\}$, we get

$T_{ab}(g) \leq \int (a\theta - g)I(g < a\theta)dF - \int (b\theta - g)I(g < a\theta)dF = (a - b) \int \theta I(g < a\theta)dF \leq 0$. (e2)* Since $\{g < b\theta\} = \{g < a\theta\} \cup \{a\theta \leq g < b\theta\}$, it

follows from (*) that $T_{ab}(g) = (a - b) \int \theta I(g \leq a\theta)dF - \int (b\theta - g)I(a\theta \leq g < b\theta)dF$. Now for any two functions g' and g such as $g' \geq g \geq 0$ we have

$\{g' < a\theta\} = \{g < a\theta\} \cup \{g' < a\theta \leq g\}$ and $\{g' < b\theta\} = \{g < a\theta\} \cup \{g' < a\theta \leq g\} \cup \{a\theta \leq g' < b\theta\}$. Hence it follows that $T_{ab}(g') = (a - b) \int \theta I(g < a\theta)dF + (a - b) \int \theta I(g' < a\theta \leq g)dF - \int (b\theta - g')I(a\theta \leq g' < b\theta)dF$. Thus

we have $T_{ab}(g) - T_{ab}(g') = -(a - b) \int \theta I(g' < a\theta \leq g)dF + \int (b\theta - g')I(a\theta \leq g' < b\theta)dF - \int (b\theta - g)I(a\theta \leq g < b\theta)dF \dots (**)$. Now letting $A = \{g' < a\theta \leq g < b\theta\}$, $B = \{g' < a\theta \leq b\theta \leq g\}$, $C = \{a\theta \leq g' \leq g < b\theta\}$, $D = \{a\theta \leq g' < b\theta \leq g\}$, we get $\{g' < a\theta \leq g\} = A \cup B$, $\{a\theta \leq g' < b\theta\} = C \cup D$,

and $\{a\theta \leq g < b\theta\} = A \cup C$. Applying these relationships to (**) yields $T_{ab}(g) - T_{ab}(g') = \int (g - a\theta)I(A)dF + (b - a) \int \theta I(B)dF + \int (g - g')I(C)dF + \int (b\theta - g')I(D)dF \geq 0$.

(f) In general, if g is a real number, then $T(g)$ can be transformed into $T(g) = \int_{g+}^{\infty} (1 - F)d\theta$ by using integration by parts. Hence we have $\Delta T_r(g) = \int_{g+}^{\infty} (F_r - F_{r-1})d\theta$. Thus it follows that $\Delta T_r(g) \geq 0$ and ∇ in g because of $F_r \geq F_{r-1}$.

(g) Suppose $F_Y \nabla$ in Y . Since $T_Y(g) = \int (\theta - g)^+ dF_Y$ and $\hat{T}_Y(g) = \int (g + (\theta - g)^+) dF_Y$, g1 to g4 are clear from lemma 3. Furthermore if $g_Y \nabla (\nearrow)$ in Y and $\theta - g_Y \nearrow$ in θ , then $Y' < Y$ leads to $T_{Y'}(g_{Y'}) \leq (\geq) T_Y(g_Y) \leq (\geq) T_Y(g_Y)$. If $g_Y \nearrow (\nabla)$ in Y and $g_Y + (\theta - g_Y)^+ \nearrow$ in θ , then $Y' < Y$ leads to $\hat{T}_{Y'}(g_{Y'}) \leq (\geq) \hat{T}_Y(g_Y) \leq (\geq) \hat{T}_Y(g_Y)$. Hence g5 and g6 hold. The proof is similar for case of $F_Y \nearrow$ in Y .

* For this proof using an indicator, the author is much indebted to Dr. Y. Nogami.

5.1 Purchasing problem (3) (12) (18)

Let N units of a certain kind of material have to be purchased within a given number of days so as to minimize the total expected purchase price. It is assumed that a per-unit price of the material at each day, θ , is a random variable with a known distribution function F_γ dependent on a price γ at the previous day. If r units are bought at the day when a price is θ , then the immediate total purchase price of $h_r \theta$ is gained where $h_0 = 0$ and $h_r \nearrow$ in r , hence we have $\Delta h_r = h_r - h_{r-1} > 0$. Let a discount factor $\beta = 1$. We shall here discuss the next two cases:

1. case of quantity-discount, i.e., $\Delta h_r \searrow$ in r
2. case of quantity-premium, i.e., $\Delta h_r \nearrow$ in r

In the problem the state space is given by $I = \{0, 1, \dots, N\}$ where a state i means the number of units having already purchased, and the action space by $A_i^r = \{A_i^r: r = 0, 1, \dots, N-i\}$ ($k_i = N - i$) in which A_i^r denotes the action of purchasing r units at the day when in state i . Then for an action A_i^r we have the immediate reward $x_i^r = -h_r \theta$ (hence $a_i^r = -h_r \nearrow$ in r) and after-effect $z_i^{r\theta} = v_{i+r}^\theta$, where v_i^θ represents the maximum total expected reward starting from day t in state i , provided that the price at the previous day, i.e., day $t + 1$, is θ . Noticing $v_N^\theta \equiv 0$ (hence $z_i^{k_i} \equiv 0$) for all t and $h_0 = 0$, (2.11) and (2.23) become, respectively, $c_i^{r\theta} = y_{i+r}^\theta / \Delta h_r$ and $c_i^\theta = -v_i^\theta / h_{N-i}$ where $y_i^\theta = v_i^\theta - v_{i-1}^\theta$ and $v_i^\theta(0) = -h_{N-i} E_\theta$, hence $y_i^\theta(0) = \Delta h_{N-i+1} E_\theta$.

Case of quantity-discount

In the case we have $y_i^\theta(0) \nearrow$ in i and ≥ 0 and $c_i^\theta(1) = E_\theta \geq 0$. Here notice that $c_i^\theta(1)$ is independent of i . For any t suppose that $y_i^\theta \nearrow$ in i and ≥ 0 and that $c_i^\theta \geq 0$ and independent of i , hence let $c^\theta = c_i^\theta$ for all i . Then since we have $c_i^{r\theta} \nearrow$ in r , from (2.27) the FE becomes

$$(5.1.1) \quad v_i^Y = -h_{N-i} E_Y + h_{N-i} T_Y(c_i) = -h_{N-i} (E_Y - T_Y(c(t))) \quad 0 \leq i \leq N$$

, and the OP is given by (2.25) where $k_i = N - i$. From (5.1.1) we get

$y_i^Y = \Delta h_{N-i+1}(E_Y - T_Y(c))$. Hence we have $y_i^Y \geq 0$ from lemma 4(a3) and $y_i^Y \nearrow$ in i . Moreover from (5.1.1) we get $c_i^Y(t+1) = -v_i^Y/h_{N-i} = E_Y - T_Y(c) \geq 0$, independent of i . Hence let $c^Y(t+1) = c_i^Y(t+1)$ for all i . Thus by induction it follows that for all $t \geq 1$ we have the OP (2.25) and the formula

$$(5.1.2) \quad c^Y(t+1) = E_Y - T_Y(c(t))$$

where $c^\theta(1) = E_\theta$. Now we have $c^Y(2) \leq E_Y = c^Y(1)$ for all γ . Supposing $c^\theta(t) \leq c^\theta(t-1)$ for any $t \geq 2$ yields $c^Y(t+1) - c^Y(t) = T_Y(c(t-1)) - T_Y(c(t)) \leq 0$ from lemma 4(a1). Hence it follows by induction that $c^Y(t) \searrow$ in t for all γ .

Now from (2.28) and (5.1.2) we obtain $b^Y(t+1) = \gamma - E_Y + T_Y(c(t))$. Suppose $F_Y \searrow$ in γ and $\gamma - E_Y \nearrow(\searrow)$ in γ . Then applying lemma 4(g1 and g2) yields $b^Y \nearrow(\searrow)$ in γ for all t by induction, noticing $b^Y(1) = \gamma - E_Y$. Hence the OP is given by (2.31)((2.34)) for all t .

Finally since c_i^θ is independent of a_i^r 's for all t , from remark 5 all the above things hold also in the case of $h_r \nearrow$ in r (that is, monotone model).

Case of quantity-premium

In the case we get $y_i^\theta(0) \searrow$ in i and ≥ 0 . Suppose $y_i^\theta \searrow$ in i and ≥ 0 for any $t \geq 1$. Then we have $c_i^{r\theta} \geq 0$ and \searrow both in r and in i . Hence the FE becomes from (2.15)

$$(5.1.3) \quad v_i^Y = -h_{N-i}E_Y + \sum_{r=1}^{N-i} \Delta h_r T_Y(c_i^r)$$

, and the OP is provided by (2.13). We can derive from (5.1.3)

$$(5.1.4) \quad y_i^Y = \Delta h_{N-i+1}(E_Y - T_Y(c_i^{N-i+1})) + \sum_{r=1}^{N-i} \Delta h_r (T_Y(c_i^r) - T_Y(c_{i-1}^r))$$

$$(5.1.5) \quad = \Delta h_{N-i+1}E_Y - \Delta h_1 T_Y(y_i^!/\Delta h_1) + \sum_{r=1}^{N-i} T_Y: \Delta h_r, \Delta h_{r+1} (y_{i+r}^!)$$

It is clear from (5.1.4) that $y_i \geq 0$. From (5.1.5) we have

$$\begin{aligned}
(5.1.6) \quad Y_i^Y - Y_{i-1}^Y &= \Delta h_1 (T_Y(Y_{i-1}^Y / \Delta h_1) - T_Y(Y_i^Y / \Delta h_1)) \\
&+ ((\Delta h_{N-i+1} - \Delta h_{N-i+2}) E_Y - T_{Y: \Delta h_{N-i+1}, \Delta h_{N-i+2}}(Y_N^Y)) \\
&+ \sum_{r=1}^{N-i} (T_{Y: \Delta h_r, \Delta h_{r+1}}(Y_{i+r}^Y) - T_{Y: \Delta h_r, \Delta h_{r+1}}(Y_{i+r-1}^Y))
\end{aligned}$$

All of the three terms in the right-hand of (5.1.6) become ≤ 0 , applying lemma 4(a1), (e3), and (e2), respectively. Hence we obtain $Y_i^Y \searrow$ in i . Thus it follows by induction that $Y_i^Y \geq 0$ and \searrow in i for all t . This yields $c_i^{r\theta} \geq 0$ and \searrow both in i and in r for all t .

5.2 Inventory problem (3) (12)

We shall here discuss the problem of purchasing a certain kind of item consumed at a constant rate of $w (\geq 1)$ units every day where w is a fixed positive integer. Assume that (1) a price per unit at each day, θ , is a random variable with a given distribution function F , (2) any shortage cannot be permitted, that is, a shortage cost is infinite, (3) the item is delivered immediately after being ordered, and (4) the maximum $N (\geq w)$ units can be stocked.

The objective is to minimize the total expected purchasing price over an infinite number of days. The state of the process can be characterized by the number of units in stock at the morning of each day. Hence the state space is given by $I = \{0, 1, \dots, N\}$. Let A_i^r , $(w - i)^+ \leq r \leq N - i$, represent an action of buying r units when in state i , where $(w - i)^+$ and $N - i$ correspond to $r = 0$ and $r = k$, respectively, in the action space of the decision model in section 1. In the problem the immediate reward becomes $x_i^r = -r\theta$, that is, $a_i^r = -r$, and the after-effect is given by $z_i^r = \beta v_{i+r-w}$. From (2.1) we get $c_i^r = Y_{i+r-w}^Y$ for $(w - i)^+ < r \leq N - i$.

Suppose $\beta E \geq Y_i^Y \geq 0$ and $Y_i^Y \searrow$ in i for any $t \geq 1$. Then since $a_i^r \searrow$ in r and $c_i^r \searrow$ in r , we get from (2.5)

$$(5.2.1) \quad v_i = -(N - i)E + \beta v_{N-w}^Y + \sum_{r=(w-i)^++1}^{N-i} T(Y_{i+r-w}^Y) \quad 0 \leq i \leq N$$

, from which the next expressions can be easily derived

$$(5.2.2) \quad \begin{cases} Y_i = \beta E & \text{for } 1 \leq i \leq w \\ Y_i = \beta(E - T(Y_{i-w}')) & \text{for } w < i \leq N \end{cases}$$

The above expressions yield $\beta E \geq Y_i \geq 0$ for all i . Now since $Y_w = \beta E \geq \beta(E - T(Y_1')) = Y_{w+1}$ and $Y_i \searrow$ in i on $w < i \leq N$ due to $Y_{i-w}' \searrow$ in i from the inductive hypothesis, it follows that $Y_i \searrow$ in i on $1 \leq i \leq N$. Therefore we have $\beta E \geq Y_i \geq 0$ and $Y_i \searrow$ in i for all t . Consequently if $\beta < 1$, for β -limit Y_i we have also $\beta E \geq Y_i \geq 0$ and $Y_i \searrow$ in i which, for β -limit c_i^r , leads to $\beta E \geq c_i^r \geq 0$ and $c_i^r \searrow$ both in i and in r . It should be noted here that, using (5.2.2) in which Y_{i-w}' is replaced by Y_{i-w} , the β -limit Y_i can be computed in order of Y_1, Y_2, \dots , independent of N , or capacity of storage. Then the β -limit OP for each i is given by the intervals (2.3).

Now since $E \geq \beta^{-1} Y_i \geq 0$ for β -limit Y_i , that is, $\beta^{-1} Y_i$ is bounded uniformly in β , the maximum expected reward per day for an undiscounting process becomes $g = \lim_{n \rightarrow \infty} \beta_n (v_{N-i}^n - v_i^n)$, noticing the FE for a discounting process $v_N = \beta v_{N-w}$, associated with β -limits v_i . Hence the minimum expected purchase price per day, g , is given by, using l -limits Y_i ,

$$(5.2.3) \quad g' = -g = Y_N + Y_{N-1} + \dots + Y_{N-w+1}$$

5.3 warehouse problem (26)

Consider the problem of buying and/or selling a certain kind of item with the aim of profit-maximization over a given number of days, taking a lease of a warehouse with the maximum capacity of N units. Now let p_w , $w = (b, s)$, denote the probability that at each day an offer of selling b units and offer of buying s units come where both b and s are non-negative integer. We shall call the vector w an offer vector where $\sum_{w \geq 0} p_w = 1$. Assume that a selling price θ_s for the buying offer and buying price θ_b for the selling offer are always

identical and that a price $\theta (= \theta_s = \theta_b)$ at each day is a random variable with a known distribution function F determined by the market mechanism. Moreover suppose no commission is charged on any transaction.

The objective is to maximize the total expected marginal profit from buying and/or selling actions over a finite number of days. In the problem it is clear that a substantial decision variable for a given offer vector $w = (b, s)$ are not the number of units to be sold, $s' (\leq s)$, and the number of units to be bought, $b' (\leq b)$, but the difference $r = s' - b'$. Notice that the difference implies selling r units if $r > 0$ and buying $-r$ units if $r < 0$.

The state of the process can be provided by the number of units being stocked at the morning of each day, i , hence the state space becomes $I = \{0, 1, \dots, N\}$. Now let h_i^W and k_i^W represent the maximum permissible buying number and maximum permissible selling number, respectively, provided that an offer vector is w when in state i . Then we have $h_i^W = \min\{N - i, b\}$ and $k_i^W = \min\{i, s\}$. Hence the action space becomes $A_i^W = \{A_i^{Wr} : -h_i^W \leq r \leq k_i^W\}$ in which A_i^{Wr} indicates an action of selling r units if $r > 0$, otherwise buying $-r$ units. Then the immediate reward and after-effect for action A_i^{Wr} become $x_i^{Wr} = r\theta$ and $z_i^{Wr} = \beta v_{i-r}^W$, respectively, where $v_i(0)$ is given as the final conditions for all i . Thus from (2.1) we have $c_i^{Wr} = y_{i-r+1}^W$ for $-h_i^W < r \leq k_i^W$.

Now suppose that for any $t \geq 1$ we have $y_t^W \geq 0$ and θ in i . Then since $a_i^{Wr} = r \uparrow$ in r and $c_i^{Wr} \uparrow$ in r , the OP is given by (2.2) with $-h_i^W \leq r \leq k_i^W$ and the FE becomes

$$(5.3.1) \quad v_i = \sum_w p_w (-h_i^W E + \beta v_{i+h_i^W}^W + \sum_{r=-h_i^W+1}^{k_i^W} T(y_{i-r+1}^W))$$

where the third term = 0 if $k_i^W = -h_i^W$. Now the vector

$(h_i^W, h_{i-1}^W, k_i^W, k_{i-1}^W)$ equals

- (a) (b, b, s, s) if $b \leq N - i$ and $s < i$
- (b) $(N - i, N - i + 1, s, s)$ if $b > N - i$ and $s < i$
- (c) $(b, b, i, i - 1)$ if $b \leq N - i$ and $s \geq i$
- (d) $(N - i, N - i + 1, i, i - 1)$ if $b > N - i$ and $s \geq i$

Noticing the above vectors, we can get from (5.3.1)

$$(5.3.2) \quad y_i = \beta \sum_W p_W \times \begin{cases} \hat{T}(y_{i+b}^t) - T(y_{i-s}^t) & \text{for (a)} \\ E - T(y_{i-s}^t) & \text{for (b)} \\ \hat{T}(y_{i+b}^t) & \text{for (c)} \\ E & \text{for (d)} \end{cases}$$

It is easily seen from (5.3.2) that $y_i \geq 0$ for all i . Now, for convenience, for all t define $y_i = \infty$ for $i \leq 0$ and $y_i = -\infty$ for $N < i$. Then since we have $y_i^t \searrow$ in i on $-\infty < i < \infty$, (5.3.2) can be compiled into the one expression below, using definition 1(4),

$$(5.3.3) \quad y_i = \sum_W p_W (\hat{T}(y_{i+b}^t) - T(y_{i-s}^t)) \quad 1 \leq i \leq N$$

Hence we obtain $y_i \searrow$ in i because both $\hat{T}(y_{i+b}^t)$ and $-T(y_{i-s}^t) \searrow$ in i on $-\infty < i < \infty$.

Consequently if $y_i(0) \geq 0$ and \searrow in i , then it follows by induction that, for all t , $y_i \geq 0$ and \searrow in i . This yields $c_i^{wr} \geq 0$, \searrow in i , and \nearrow in r for all t .

Next we shall discuss the case of an infinite time horizon process without discounting. First suppose $\beta < 1$. If $0 \leq y_i^t \leq \beta M (< M)$ for a sufficiently large number M such that $\hat{T}(M) = M$ (see section 5.0 (2)), from (5.3.2) we have $0 \leq y_i^t \leq \beta \sum_W p_W \max\{E, \hat{T}(M)\} \leq \beta M < M$. Hence since it follows by induction that $0 \leq \beta^{-1} y_i < M$ for a β -limit y_i , the maximim expected marginal profit per day for an undiscounting process is provided by, noticing the limit $FE v_N = \beta v_N + \sum_W p_W \sum_{r=1}^{k_i^W} T(y_{i-r+1}^t)$ for a discounting process,

$$(5.3.4) \quad g = \sum_W p_W \sum_{r=1}^{k_N^W} T(y_{N-r+1})$$

in which y_i is an l-limit.

Next we shall apply the above results to three special cases below. Here let p_w , h_i^w , and k_i^w be denoted by $p_{b,s}$, $h_i^{b,s}$, and $k_i^{b,s}$, respectively.

Example 1 Let $p_{0,s} \geq 0$ and $p_{b,s} = 0$ for $b > 0$

Example 2 Let $p_{b,0} \geq 0$ and $p_{b,s} = 0$ for $s > 0$

Now for simplicity let $p_n = p_{0,n}$ for example 1 and $p_n = p_{n,0}$ for example 2. We shall suppose that both examples are of finite time horizon.

In example 1, no selling offer is made, hence no units are bought at all. This means that the example is equivalent to the problem of selling out inventory in the warehouse so as to maximize the total expected selling price over a given number of days.

On the other hand, in example 2, no buying offers come, in other words, no units in stock are sold. Now let the objective in the example be to buying the item to reach N units of storage capacity in such a way as to minimize the total expected purchasing price over a finite number of days under the following final conditions. Suppose that i units have been purchased by the end of day 1 and the offer vector is $w = (n, 0)$ at day 0. If $n \geq N - i$, then $N - i$ units must be bought at the current price at day 0. Conversely if $n < N - i$, first n units are purchased at the current price θ and then the additional $N - i - n$ units must be done at a higher price than the current price, $h\theta$, where h is a fixed number > 1 .

Now we have for $n \geq 0$

$$(5.3.5) \quad h_i^{0,n} = 0, \quad k_i^{0,n} = \min\{i, n\} \quad \text{for Ex. 1}$$

$$(5.3.6) \quad h_i^{n,0} = \min\{N - i, n\}, \quad k_i^{n,0} = 0 \quad \text{for Ex. 2}$$

The final conditions become

$$(5.3.7) \quad v_i(0) = \sum_{n=0}^{i-1} n p_n^E + i \sum_{n=i}^{\infty} p_n^E \quad \text{for Ex. 1}$$

$$(5.3.8) \quad v_i(0) = -\sum_{n=0}^{N-i-1} (n + h(N - i - n)) p_n E \\ + (N - i) \sum_{n=N-i}^{\infty} p_n E \quad \text{for Ex. 2}$$

Hence it follows that

$$(5.3.9) \quad Y_i(0) = \sum_{n=i}^{\infty} p_n E \geq 0 \quad \text{for Ex. 1}$$

$$(5.3.10) \quad Y_i(0) - Y_{i-1}(0) = -p_{i-1} E \leq 0$$

$$(5.3.11) \quad Y_i(0) = \sum_{n=N-i+1}^{\infty} p_n E + h \sum_{n=0}^{N-i} p_n E \geq 0 \quad \text{for Ex. 2}$$

$$(5.3.12) \quad Y_i(0) - Y_{i-1}(0) = (1 - h) p_{N-i+1} E \leq 0$$

From the above four expressions, for both examples we have $y_i(0) \geq 0$ and $y_i(0) - y_{i-1}(0) \leq 0$, hence $y_i(0) \searrow$ in i . Then (5.3.2) becomes

$$(5.3.13) \quad Y_i = \hat{T}(Y_i^t) - \sum_{n=0}^{i-1} p_n T(Y_{i-n}^t) \quad \left. \vphantom{Y_i} \right\} \text{for Ex. 1}$$

$$(5.3.14) \quad = Y_i^t + \sum_{n=i}^{\infty} p_n T(Y_i^t) + \sum_{n=0}^{i-1} p_n (T(Y_i^t) - T(Y_{i-1}^t))$$

$$(5.3.15) \quad Y_i = \sum_{n=0}^{N-i} p_n \hat{T}(Y_{i+n}^t) + \sum_{n=N-i+1}^{\infty} p_n E - T(Y_i^t) \quad \left. \vphantom{Y_i} \right\} \text{for Ex. 2}$$

$$(5.3.16) \quad = Y_i^t + \sum_{n=0}^{N-i} p_n (\hat{T}(Y_{i+n}^t) - \hat{T}(Y_i^t)) \\ + \sum_{n=N-i+1}^{\infty} p_n (E - \hat{T}(Y_i^t))$$

From (5.3.14) and (5.3.16) we obtain $y_i \geq y_i^t$ and $y_i \leq y_i^t$, respectively. Hence it follows that $y_i \searrow$ in t for example 1 and $y_i \nearrow$ in t for example 2. Thus eventually by induction we have $c_i^{wr} \geq 0$, \nearrow in r , and \searrow in i for all t , and $c_i^{wr} \nearrow$ in t for example 1 and \searrow in t for example 2.

Example 3 Let $p_{1,1} = 1$ and time horizon $= \infty$. For simplicity, the vector $w = (1, 1)$ as superscript is removed from all the notations below. We have $h_i = \min\{N - i, 1\}$ and $k_i = \min\{i, 1\}$ for the example. Hence the OP (2.2) becomes as follows.

$$(5.3.17) \quad \left\{ \begin{array}{l} D_0^{-1} = (-\infty, Y_1) \quad , \quad D_0^0 = (Y_1, \infty) \\ D_i^{-1} = (-\infty, Y_{i+1}), \quad D_i^0 = (Y_{i+1}, Y_i), \quad D_i^1 = (Y_i, \infty) \quad 0 < i < N \\ D_N^0 = (-\infty, Y_N), \quad D_N^1 = (Y_N, \infty) \end{array} \right.$$

where y_i satisfies (see (5.3.2))

$$(5.3.18) \begin{cases} \beta^{-1}y_1 = \hat{T}(y_2) \\ \beta^{-1}y_i = \hat{T}(y_{i+1}) - T(y_{i-1}) & 0 < i < N \\ \beta^{-1}y_N = E - T(y_{N-1}) \end{cases}$$

In an discounting process, (5.3.4) becomes $g = T(y_N)$, using 1-limit y_N .

5.4 Buying-selling problem (6) (23)

Let us call the action of buying one unit of a certain kind of item and selling it at another day a buying-selling action cycle or simply action cycle. A selling price θ_s is always assumed to be equal to buying price θ_b . Then let us set $\theta = \theta_s = \theta_b$, assumed to be a random variable with a known distribution function F . Furthermore postulate that no commission is charged on any transaction and that a discount factor $\beta = 1$. The goal is here to optimally repeat the action cycle the maximum $M(\geq 1)$ times in a sence of maximizing the total expected marginal profit over a given number of days.

For convenience, let action cycles be numbered backward from the final action cycle $n = 1$. Now by s (b) we shall denote a selling(buying) situation, in which there exists one unit(no unit) on hand. Then a state of the process can be characterized by the vector (n,h) in which n and h denote a cycle number and situation index, respectively. Hence a state space becomes $I = \{(n,h) : 1 \leq n \leq M \text{ and } h = s, b\}$. Then the state transition law can be described as follows. A state (n,b) changes into (n,s) if one unit of the item is bought and a state (n,s) into $(n-1,b)$ if one unit on hand is sold, otherwise both states (n,b) and (n,s) remain unchanged. Let us set the final conditions as follows. If the process is in state (n,b) at day 0, it terminates immediately, and if in state (n,s) at day 0, it stops after disposing of the one unit on hand.

When in state (n,b) , by A_{nb}^{-1} and A_{nb}^0 we shall denote a buying action and not buying action, respectively. On the other hand, when in state (n,s) , by A_{ns}^0 and A_{ns}^1 a not selling action and selling action, respectively. Then the immediate rewards for these actions become $x_{nb}^{-1} = -\theta$, $x_{nb}^0 = 0$, $x_{ns}^0 = 0$, and $x_{ns}^1 = \theta$, provided that the current price is θ . That is, we have $a_{nb}^{-1} = -1$, $a_{nb}^0 = 0$, $a_{ns}^0 = 0$, and $a_{ns}^1 = 1$. From the state transition law defined above, the after-effects for these actions become $z_{nb}^{-1} = v'_{ns}$, $z_{nb}^0 = v'_{nb}$, $z_{ns}^0 = v'_{ns}$, and $z_{ns}^1 = v'_{n-1,b}$ where v_{nh} represents the maximum of the total expected marginal profit starting from day t in state (n,h) . The final conditions defined above yield $v_{ns}(0) = E$ and $v_{nb}(0) = 0$ for all $n \geq 1$. Here, for convenience, define $v_{0h}(t) = 0$ for all $t \geq 1$ and $h = b, s$.

Now since $a_{nb}^{-1} < a_{nb}^0$ and $a_{ns}^0 < a_{ns}^1$, from (2.2) and remark 3 we have the OP

$$(5.4.1) \quad D_{nb}^{-1}(t) = (-\infty, c_{nb}(t)) \quad , \quad D_{nb}^0(t) = (c_{nb}(t), \infty) \quad \text{for } (n,b)$$

$$(5.4.2) \quad D_{ns}^0(t) = (-\infty, c_{ns}(t)) \quad , \quad D_{ns}^1(t) = (c_{ns}(t), \infty) \quad \text{for } (n,s)$$

where

$$(5.4.3) \quad c_{nb}(t) = v_{ns}(t-1) - v_{nb}(t-1) \quad n \geq 0$$

$$(5.4.4) \quad c_{ns}(t) = v_{ns}(t-1) - v_{n-1,b}(t-1) \quad n \geq 1$$

From (2.4) we have the FE for $t \geq 1$,

$$(5.4.5) \quad v_{nb}(t) = -E + v_{ns}(t-1) + T(c_{nb}(t))$$

$$(5.4.6) \quad \quad \quad = v_{nb}(t-1) + \hat{T}(c_{nb}(t)) - E \quad n \geq 1$$

$$(5.4.7) \quad v_{ns}(t) = v_{ns}(t-1) + T(c_{ns}(t)) \quad n \geq 1$$

$$(5.4.8) \quad \quad \quad = v_{n-1,b}(t-1) + \hat{T}(c_{ns}(t))$$

where notice $v_{1s}(t) = \hat{T}(c_{1s}(t))$ because of $c_{1s}(t) = v_{1s}(t-1)$ due to the definition $v_{0s}(t-1) = 0$. By induction it is clear from (5.4.6) and (5.4.7) that both $v_{ns}(t)$ and $v_{nb}(t) \geq 0$ for all $t \geq 1$. We obtain from (5.4.3) to (5.4.8),

$$(5.4.9) \quad c_{0b}(t+1) = 0$$

$$(5.4.10) \quad c_{nb}(t+1) = T(c_{ns}(t)) - T(c_{nb}(t)) + E \quad n \geq 1$$

$$(5.4.11) \quad c_{1s}(t+1) = \hat{T}(c_{1s}(t))$$

$$(5.4.12) \quad c_{ns}(t+1) = v_{ns}(t-1) - v_{n-1,s}(t-1) \\ + T(c_{ns}(t)) - T(c_{n-1,b}(t)) + E \quad n \geq 2$$

$$(5.4.13) \quad = \hat{T}(c_{ns}(t)) - \hat{T}(c_{n-1,b}(t)) + E \quad n \geq 2$$

Now define

$$(5.4.14) \quad \Delta v_{nh}(t) = v_{nh}(t) - v_{n-1,h}(t) \quad n \geq 1, h = s, b$$

$$(5.4.15) \quad \Delta c_{nsb}(t) = c_{ns}(t) - c_{nb}(t) (= \Delta v_{nb}(t-1)) \quad n \geq 1$$

$$(5.4.16) \quad \bar{\Delta} c_{nb}(t) = c_{nb}(t) - c_{nb}(t-1) \quad n \geq 0$$

$$(5.4.17) \quad \bar{\Delta} c_{ns}(t) = c_{ns}(t) - c_{ns}(t-1) \quad n \geq 1$$

$$(5.4.18) \quad \Delta c_{nb}(t) = c_{nb}(t) - c_{n-1,b}(t) \quad n \geq 1$$

$$(5.4.19) \quad \Delta c_{ns}(t) = c_{ns}(t) - c_{n-1,s}(t) \quad n \geq 2$$

For simplicity, in the expressions below we shall employ the definition of $\underline{H \bullet (g - g')} = H(g) - H(g')$ for a given function H . Its mathematical treatments, however, must be sufficiently carefull. Then for $t \geq 1$ we have

$$(5.4.20) \quad \Delta v_{1b}(t) = v_{1b}(t-1) + \hat{T}(c_{1b}(t)) - E$$

$$(5.4.21) \quad \Delta v_{nb}(t) = \Delta v_{nb}(t-1) + \hat{T} \bullet \Delta c_{nb}(t) \quad n \geq 2$$

$$(5.4.22) \quad \Delta v_{1s}(t) = \hat{T}(c_{1s}(t))$$

$$(5.4.23) \quad \Delta v_{ns}(t) = \Delta v_{ns}(t-1) + T \bullet \Delta c_{ns}(t) \quad n \geq 2$$

$$(5.4.24) \quad \Delta c_{nsb}(t+1) = \Delta v_{nb}(t) = (5.4.17), (5.4.18) \quad n \geq 1$$

$$(5.4.25) \quad \bar{\Delta} c_{nb}(t+2) = T \bullet \bar{\Delta} c_{ns}(t+1) - T \bullet \bar{\Delta} c_{nb}(t+1) \quad n \geq 1$$

$$(5.4.26) \quad \bar{\Delta} c_{1s}(t+2) = \hat{T} \bullet \bar{\Delta} c_{1s}(t+1)$$

$$(5.4.27) \quad \bar{\Delta} c_{ns}(t+2) = \hat{T} \bullet \bar{\Delta} c_{ns}(t+1) - \hat{T} \bullet \bar{\Delta} c_{n-1,b}(t+1) \quad n \geq 2$$

$$(5.4.28) \quad \Delta c_{1b}(t+1) = T(c_{1s}(t)) - T(c_{1b}(t)) + E$$

$$(5.4.29) \quad \Delta c_{nb}(t+1) = T \cdot \Delta c_{ns}(t) - T \cdot \Delta c_{nb}(t) \quad n \geq 2$$

$$(5.4.30) \quad \Delta c_{2s}(t+1) = \hat{T} \cdot \Delta c_{2s}(t) - \hat{T}(c_{1b}(t)) + E$$

$$(5.4.31) \quad \Delta c_{ns}(t+1) = \hat{T} \cdot \Delta c_{ns}(t) - \hat{T} \cdot \Delta c_{n-1,b}(t) \quad n \geq 3$$

Let us now define the following seven statements for $n \geq 1$ and $t \geq 1$

- St. 1 $\Delta v_{nb}(t-1) \geq 0$
- St. 2 $\Delta v_{ns}(t-1) \geq 0$
- St. 3 $\Delta c_{nsb}(t) \geq 0$
- St. 4 $\bar{\Delta} c_{nb}(t+1) \leq 0$
- St. 5 $\bar{\Delta} c_{ns}(t+1) \geq 0$
- St. 6 $\Delta c_{nb}(t) \geq 0$
- St. 7 $\Delta c_{ns}(t) \leq 0$ where $n \geq 2$

Then we have

- (a) $c_{nb}(1) = c_{ns}(1) = E \geq 0$ for $n \geq 1$
- (b) Sts. 1 to 6 hold for $t = 1$ and $n \geq 1$
- (c) Sts. 1 to 7 hold for $t = 1$ and $n \geq 2$
- (d) $c_{nb}(t) \geq 0$ for $n \geq 1$ and $t \geq 1$
- (e) Sts. 1 to 6 hold for $t \geq 1$ and $n = 1$
- (f) Sts. 1 to 7 hold for $t \geq 1$ and $n \geq 2$
- (g) $c_{ns}(t) \geq 0$ for all $n \geq 1$ and all $t \geq 1$

PROOF

(a) is clear from (5.4.3) and (5.4.4).

(b) $\Delta v_{nb}(0) = 0$ for $n \geq 1$ (St. 1), $\Delta v_{1s}(0) = E \geq 0$ and $\Delta v_{ns}(0) = 0$ for $n \geq 2$ (St. 2), and $\Delta c_{nsb}(1) = 0$ for $n \geq 1$ (St. 3). Since $v_{ns}(1) = E + \hat{T}(E)$ and $v_{nb}(1) = T(E)$ for $n \geq 1$, we get $c_{nb}(2) = E$ for $n \geq 1$, $c_{1s}(2) = \hat{T}(E)$, and $c_{ns}(2) = E$ for $n \geq 2$. Hence it follows that $\bar{\Delta} c_{nb}(2) = 0$ for $n \geq 1$ (St. 4), $\bar{\Delta} c_{1s}(2) = T(E)$ and $\bar{\Delta} c_{ns}(2) = 0$ for $n \geq 2$ (St. 5), and $\Delta c_{1b}(1) = c_{1b}(1) = E \geq 0$ and $\Delta c_{nb}(1) = 0$ for $n \geq 1$ (St. 6).

(c) we get $\Delta c_{ns}(1) = 0$ for $n \geq 2$ (St. 7).

(d) $c_{nb}(1) = E \geq 0$ for $n \geq 1$ from (a). For any $t \geq 1$ suppose $c_{nb}(t) \geq 0$ for $n \geq 1$. Then from (5.4.10), we get $c_{nb}(t+1) \geq 0$ for $n \geq$

1, using lemma 4(a3).

(e) The assertion is true for $t = 1$ from (b). If it is true for any $t \geq 1$, then it becomes also true for $t+1$ by applying lemma 4(a,b) and the above (d) to (5.4.20) to (5.4.31).

(f) For $n = 2$ the proof is exactly similar to (e). Now suppose it holds for $n' = 2, 3, \dots, n-1$. Then the assertion for n becomes to be true for $t = 1$ from (c). If it is true for any $t \geq 1$, then so also is it for $t+1$ from (5.4.20) to (5.4.31). Hence by induction it holds for all $t \geq 1$. Thus eventually it holds also for all $n \geq 2$ and all $t \geq 1$, using induction again.

(g) $c_{ns}(1) = E \geq 0$ for all $n \geq 1$ from (a). We obtain $c_{1s}(t+1) \geq 0$ for all $t \geq 1$ from (5.4.11) and $c_{ns}(t+1) \geq 0$ for $n \geq 2$ from (5.4.12), using $\Delta v_{ns}(t-1) \geq 0$ and $c_{n-1,b}(t) \geq 0$. Q.E.D.

From the above (a) to (g) we obtain

- (1)' $c_{ns}(t) \geq c_{nb}(t)$
- (2) $c_{nb}(t) \geq 0$, \checkmark in t , and \nearrow in n
- (3) $c_{ns}(t) \geq 0$, \nearrow in t , and \checkmark in n

5.5 Job search problem [14]

This is the problem of determining one to be accepted among sequentially arriving job offers within a finite number of periods. Let $S = \{s: s = 0, 1, \dots, K\}$ be a finite space of the possible economy states. An economy state changes according to a discrete time Markov chain with one-step transition matrix (p_{su}) , assumed $\sum_{u=0}^q p_{su} \checkmark$ in s for all $q \in S$. A per-period wage of a job offered at each period, θ , is a random variable with a known s -dependent distribution function F_s with an expectation E_s . Let $F_s \checkmark$ in s , hence $E_s \nearrow$ in s from lemma 3(a). A job can be always offered at the next period if $h (> 0)$ dollar are spent for searching a job at the current period.

The objective is to maximize the total expected reward, or the difference between the total expected wage over the period and the

total searching cost. It should be noted here that the state space I of this search process as a sequential selection process must be distinguished from the economy state space S . In the process the state space is given by $I = \{0,1\}$ in which $i = 0$ represents the state of not yet accepting any job offer, and $i = 1$ the one of already having accepted a job offer. Then the action spaces are given by $A_0 = \{A_0^0, A_0^1\}$ for state $i = 0$ and $A_1 = \{A_1^0\}$ for state $i = 1$ where A_0^0 and A_1^0 denote an action of rejecting an arriving job offer and A_0^1 one of accepting it.

Let $v_{i,s}$ represent the maximum total expected reward starting from period t with an economy state s and process state i and let $\alpha_t = \sum_{n=0}^t \beta^n$. It is clear that $v_{0s}(0) = E_s$ and $v_{1s}(t) \equiv 0$ for all t . Then for a job offer of per-period wage θ made at period t with state $i = 0$, the immediate reward is provided by $x_0^0 = 0$ and $x_0^1 = \alpha_t \theta$ and the after-effect by $z_{0,s}^0 = \beta \sum_{u=0}^K q_{su} v'_{0,u} - h$ and $z_{0,s}^1 = 0$. Now since clearly $x_1^0 = 0$ and $z_{1,s}^0 = 0$, suffice it to discuss only the case of state $i = 0$. Since $a_0^0 = 0 < \alpha_t = a_0^1$, from (2.4) we have the FE $v_{0,s} = z_{0,s}^0 + \alpha_t T_s(c_{0,s}^1) = \alpha_t \hat{T}_s(c_{0,s}^1)$ where $c_{0,s}^1 = z_{0,s}^0 / \alpha_t$ from (2.1). Then letting $v_s = v_{0,s}$, $z_s = z_{0,s}^0$, and $c_s = c_{0,s}^1$, the above FE can be rewritten as

$$(5.5.1) \quad v_s = \alpha_t \hat{T}_s(c_s)$$

where $c_s = z_s / \alpha_t$ and $z_s = \beta \sum_{u=0}^K q_{su} v'_u - h$. Then from (2.2) the OP is given by $D_s^0 = (-\infty, c_s]$, $D_s^1 = (c_s, \infty)$.

Now notice $v_s(0) (= v_{0,s}(0)) = E_s \nearrow$ in s . Suppose $v'_s \nearrow$ in s for any $t \geq 1$, then we obtain $z_s \nearrow$ in s from lemma 3(a), which yields $c_s \nearrow$ in s . Moreover since $c_s + \{\theta - c_s\}^+ \nearrow$ in θ , we get $\hat{T}_s(c_s) \nearrow$ in s from lemma 4(g6). Thus we have $v_s \nearrow$ in s . Consequently it follows by induction that for all $t \geq 1$ we get $v_s \nearrow$ in s , hence $c_s \nearrow$ in s .

Now since $v_s(1) = \alpha_1 \hat{T}_s(c_s(1)) \geq \alpha_1 E_s$ from lemma 4(b3), we have $c_s(2) \geq (\beta \sum_{u=0}^K q_{su} \alpha_1 E_u - h) / \alpha_2 = \beta (\alpha_1 / \alpha_2) \sum_{u=0}^K q_{su} E_u - h / \alpha_2 \geq \beta (1 / \alpha_1) \sum_{u=0}^K q_{su} E_u - h / \alpha_1 = c_s(1)$. Suppose $c_s \geq c'_s$ for any $t \geq 1$. Then

we have $v_s/\alpha_t - v'_s/\alpha_{t-1} \geq \hat{T}_s(c_s) - \hat{T}_s(c'_s) \geq 0$, that is, $v_s/\alpha_t \geq v'_s/\alpha_{t-1}$. Applying this yields $c_s(t+1) = z_s(t+1)/\alpha_{t+1} \geq (\alpha_t/\alpha_{t+1}\alpha_{t-1})^\beta \sum_{u=1}^K q_{su} v'_u - h/\alpha_{t+1} \geq c_s (= c_s(t))$, noticing $\alpha_t/\alpha_{t+1}\alpha_{t-1} \geq 1/\alpha_t$. Hence we have $c_s \nearrow$ in t by induction.

5.6 Sequential stochastic assignment problem (2)

This section is concerned with another approach to the following sequential selection problem posed in (2). The t jobs arrive at a group of t men(workers) in a sequential order, that is, one by one. They are called, respectively, b_1 -man, b_2 -man, ..., b_t -man, assumed $b_r \nearrow$ in r . If b_r -man is assigned to an arriving job of value θ , then the reward of $b_r\theta$ is gained. A man assigned to a job is unavailable for future assignment. A value of each arriving job is a random variable according to a given distribution function F .

The objective is to maximize the total expected reward obtained. We shall refer to the problem as $b(t)$ -problem where $b(t) = (b_1, b_2, \dots, b_t)$, t -vector. Let $v(t)$, $v_i(t)$, and $v_{ir}(t-1)$ be the maximum total expected reward for $b(t)$ -, $b_i(t)$ -, and $b_{ir}(t-1)$ -problems, respectively, where $b_i(t)$ is a t -vector resulting from removing the i -th element of a vector $b(t+1)$ and where $b_{ir}(t-1)$ is a $(t-1)$ -vector from eliminating the r -th element of the vector $b_i(t)$. Let the r -th element of vector $b_i(t)$ be represented by b_{ir} .

Now the action space for $b(t)$ -problem is given by $A = \{A^r: r = 1, 2, \dots, t\}$ in which A^r represents the action of assigning b_r -man to the first arriving job. Then the immediate reward and after-effect for an action A^r are given by $x^r = b_r\theta$ and $z^r(t) = v_r(t-1)$. Hence from (2.1) we have $c^r(t) = (v_{r-1}(t-1) - v_r(t-1))/(b_r - b_{r-1})$ for $2 \leq r \leq t$. Since $a^r = b_r \nearrow$ in r (strictly monotone model), if $c^r(t) \nearrow$ in r , then the OP is given by (2.2) with $k = t$ where $c^1(t) = -\infty$ and $c^{t+1} = \infty$, and the FE become from (2.4)

$$(5.6.1) \quad v(t) = b_1 E + v_1(t-1) + \sum_{r=2}^t (b_r - b_{r-1}) T(c^r(t))$$

For $b(1)$ -problem, clearly $v(1) = bE$ where $b(1) = (b)$, 1-vector.

For $b(2)$ -problem, we have $v(2) = b_1 E + v_1(1) + (b_2 - b_1) T(c^2(2))$ in which $c^2(2) = (v_1(1) - v_2(1))/(b_2 - b_1) = E$, using $v_1(1) = b_2 E$ and $v_2(1) = b_1 E$. Here notice that $c^2(2)$ can be expressed as $c^2(2) = \hat{T}(c^1(1)) - T(c^2(1))$ by using definition 1(4) and that $c^r(2)$ is $b(2)$ -independent for $r = 1, 2, 3$ and $c^1(2) \leq c^2(2) \leq c^3(2)$, that is, $c^r(2) \nearrow$ in r .

For $b(t)$ -problem (for any $t \geq 2$) suppose $c^r(t)$ is independent of $b(t)$ and \nearrow in r . Then (5.6.1) becomes for the $b_i(t)$ -problem

$$(5.6.2) \quad v_i(t) = b_{i1} E + v_{i1}(t-1) + \sum_{r=2}^t (b_{ir} - b_{i,r-1}) T(c^r(t))$$

$1 \leq i \leq t+1$

Now $i > r$ leads to $b_{ir} = b_r$ and $v_{ir}(t-1) = v_{r,i-1}(t-1)$, and $i \leq r$ to $b_{ir} = b_{r+1}$ and $v_{ir}(t-1) = v_{r+1,i}(t-1)$. For simplicity let $v_i = v_i(t)$, $v'_{ir} = v'_{ir}(t-1)$, and $c^r = c^r(t)$. Then we have, noticing $b_i - b_{i-1} = b_{1,i-1} - b_{1,i-2}$ for $i \geq 3$,

$$(5.6.3) \quad \left\{ \begin{array}{l} (v_1 - v_2)/(b_2 - b_1) = \hat{T}(c^1) - T(c^2) \\ (v_{i-1} - v_i)/(b_i - b_{i-1}) = (v'_{1,i-2} - v'_{1,i-1})/(b_{1,i-1} - b_{1,i-2}) \\ \quad + T(c^{i-1}) - T(c^i) \quad 3 \leq i \leq t \\ (v_t - v_{t+1})/(b_{t+1} - b_t) = (v'_{1,t-1} - v'_{1,t})/(b_{1t} - b_{1,r-1}) \\ \quad + T(c^t) \end{array} \right.$$

Arranging (5.6.3) by applying the definitions of $c^r(t+1)$ and $c^{r-1}(t)$ produces

$$(5.6.4) \quad c^r(t+1) = \hat{T}(c^{r-1}(t)) - T(c^r(t)) \quad 2 \leq r \leq t+1$$

In the above expressions, since $c^{r-1}(t)$ and $c^r(t)$ are independent of $b_i(t)$ from the inductive hypothesis, $c^r(t+1)$ also becomes independent of $b(t+1)$. It is easy to prove that $c^r(t) \nearrow$ in r leads to $c^r(t+1) \nearrow r$. Therefore by induction it follows that $c^r(t)$ is $b(t)$ -independent and \nearrow in r for all t .

Now since for $2 \leq r \leq t$ we obtain $c^r(t+1) = c^r(t) - (\hat{T}(c^r(t)) - \hat{T}(c^{r-1}(t))) \leq c^r(t)$ owing to $c^r(t) \geq c^{r-1}(t)$, it follows that $c^r(t) \searrow$ in t for all r .

Out of three requirements in theorem 3, (1) is verified just above and (2) is clear from the formulation of the problem. It is easily proved by induction that the last requirement (3) is satisfied. Therefore the FE and OP for a model with $b_r \nearrow$ in r , that is, for a monotone model are given by just the same forms as in the model with $b_r \searrow$ in r as discussed just above, that is, a strictly monotone model.

5.7 Target attacking problem I [16]

Suppose that a submarine with N torpedos is starting to patrol a sea area over a given mission periods. Assume that a target can be detected every periods and that the value of a detected target, θ , is a random variable with a known distribution function F . By q ($0 < q < 1$) we shall denote a single shot hit probability, assumed to be independent of a target value θ .

The objective is to maximize the total expected value from enemy ships hit by the submarine over the given mission periods. The state space of the problem is given by $I = \{i : 0 \leq i \leq N\}$ in which i represents the number of torpedos remaining in the submarine on patrol, and the action space by $A_i = \{A_i^r : 0 \leq r \leq i\}$ where A_i^r denotes an action of firing r torpedos at a target detected when in state i . Then the immediate reward and after-effect for an action A_i^r become $x_i^r = (1 - (1 - q)^r)\theta$ and $z_i^r = v_{i-r}^r$, respectively, given that the detected enemy ships has a value of θ . It is clear that $v_i(0) = (1 - (1 - q)^i)E$. Since $a_i^r = 1 - (1 - q)^r$, it follows from (2.1) that $c_i^r = y_{i-r+1}^r/h(r)$ for $1 \leq r \leq i$ where $h(r) = (1 - q)^{r-1}q$. Obviously we have $y_i(0) = h(i)E \geq 0$ and $y_i(0) - y_{i-1}(0) = -(1 - q)^{i-2}q^2E \leq 0$.

Now suppose that $y_i^t \geq 0$ and $y_i^t - y_{i-1}^t \leq 0$ for any $t \geq 1$. Then

noticing $h(r) \searrow$ in r and $y_i^t \searrow$ in i , we have $c_i^r \nearrow$ in r , \searrow in i , and \geq 0. Since $a_i^r \nearrow$ in r , the OP is given by (2.2) with $k_i = i$, and the FE becomes from (2.4)

$$(5.7.1) \quad v_i = v_i^t + \sum_{r=1}^i h(r) T(c_i^r)$$

From (5.7.1) we obtain

$$(5.7.2) \quad y_i = y_i^t + h(i) T(y_i^t/h(i))$$

$$+ \sum_{r=1}^{i-1} h(r) (T(y_{i-r+1}^t/h(r)) - T(y_{i-r}^t/h(r)))$$

$$(5.7.3) \quad = h(1) \hat{T}(y_i^t/h(1)) + \sum_{r=1}^{i-1} T_{h(r+1), h(r)}(y_{i-r}^t) \quad 1 \leq i \leq N$$

, and from (5.7.3)

$$(5.7.4) \quad y_i - y_{i-1} =$$

$$h(1) (\hat{T}(y_i^t/h(1)) - \hat{T}(y_{i-1}^t/h(1))) + T_{h(i), h(i-1)}(y_i^t)$$

$$+ \sum_{r=1}^{i-2} (T_{h(r+1), h(r)}(y_{i-r}^t) - T_{h(r+1), h(r)}(y_{i-r-1}^t))$$

$$2 \leq i \leq r$$

We have $y_i \geq 0$ from (5.7.2) and $y_i \leq y_{i-1}$ from (5.7.4), hence $y_i \searrow$ in i . Thus it follows by induction that $y_i \geq 0$ and \searrow in i for all $t \geq 1$. Moreover since $y_i \geq y_i^t$ from (5.7.2), it follows that $y_i \nearrow$ in t . Hence we have $c_i^r \geq 0$, \nearrow in r , and \searrow in i for all $t \geq 1$, and $c_i^r \nearrow$ in t for all i and all r .

5.8 Optimal target attacking problem II

In this section a variation of the optimal target attacking problem I is discussed. It is here assumed that (1) a single shot hit probability $q = 1$, that is, a hundred percent in bulls-eye, (2) a fleet of m enemy ships can be detected at each period (m is a fixed positive integer), and (3) a value of each of m enemy ships detected is mutually independent random variable with a known distribution function F , where let these values be arranged as $\theta_1 \geq \theta_2 \geq \dots \geq \theta_m$.

The state space of this problem is just the same as one in the

preceding section, but the action space becomes $A_i = \{A_i^r: 0 \leq r \leq \min\{i, m\}\}$ where A_i^r represents an action of firing at r ones from the upper rank of values for m enemy ships detected in state i . Let $k_i = \min\{i, m\}$. For an action A_i^r the immediate reward is given by $x_i^0 = 0$ and $x_i^r = \theta_1 + \theta_2 + \dots + \theta_r$, $1 \leq r \leq k_i$, and the after-effect by $z_i^r = v_{i-r}'$ where clearly $v_i(0) = \sum_{r=1}^{k_i} E_r$ for $1 \leq i \leq N$. From (3.1) we have $c_i^r = y_{i-r+1}'$ for $0 < r \leq k_i$. Define $c_i^r = \infty$ for $r > k_i$. On $K = \{c_i^r : 0 < r \leq k_i\}$ we have $c_i^r = c_{i-1}^{r-1}$.

Now we have $y_i(0) = E_i \geq 0$ for $i \leq m$ and $y_i(0) = 0$ for $m < i \leq N$. Hence it follows that $y_i(0) - y_{i-1}(0) = E_i - E_{i-1} \leq 0$ for $i < m$, $y_{m+1}(0) - y_m(0) = -E_m \leq 0$, and $y_i(0) - y_{i-1}(0) = 0$ for $m < i \leq N$. Thus for all i we have $y_i(0) \geq 0$ and $y_i(0) \searrow$ in i . For any $t \geq 1$ suppose $y_i^t \geq 0$ and $y_i^t \searrow$ in i . Then we have $c_i^r \geq 0$, \nearrow in r , and \searrow in i , hence the OP is provided by (3.2), and the FE becomes from (3.3)

$$(5.8.1) \quad v_i = v_i' + \sum_{r=1}^{k_i} T_r(c_i^r) \quad 0 \leq i \leq N$$

from which we get (see remark 6),

$$(5.8.2) \quad Y_i = Y_i' + \sum_{r=1}^{\infty} (T_r(\underbrace{c_i^r}_B) - T_r(\underbrace{c_{i-1}^r}_D)) \quad 1 \leq i \leq N$$

Noticing $c_i^1 \in K$ and $c_{i-1}^1 \in K$ for $i \geq 2$, from (5.8.2) we obtain for $2 \leq i \leq N$

$$(5.8.3) \quad Y_i - Y_{i-1} = (T_1(Y_i') - T_1(Y_{i-1}'))^* - \sum_{r=2}^{\infty} ((T_{r-1}(\underbrace{c_{i-1}^{r-1}}_A) - T_r(\underbrace{c_i^r}_B)) - (T_{r-1}(\underbrace{c_{i-2}^{r-1}}_C) - T_r(\underbrace{c_{i-1}^r}_D)))^{**}$$

where A, B, C, and D in (5.8.2) and (5.8.3) are not always finite, in other words, not always in K.

Since $k_i \nearrow$ in i , $r \leq k_{i-1}$ leads to $r \leq k_i$. This means that if $D \in K$, then also $B \in K$. Hence we have always $B \leq D$, including \leq , which provides the second term ≥ 0 in (5.8.2), that is, $Y_i \geq 0$. Now if $r-1 > k_{i-1}$, then $r > k_i$, $r-1 > k_{i-2}$ and $r > k_{i-1}$, if $r > k_i$, then $r > k_{i-1}$, and if $r-1 > k_{i-2}$, then $r > k_{i-1}$. These imply, respectively, (1) $A = \infty \rightarrow B = \infty$, $C = \infty$, $D = \infty$, (2) $B = \infty \rightarrow$

$D = \infty$, (3) $C = \infty \rightarrow D = \infty$. Next let us show the impossibility of the joint event of " $r \leq k_i, r-1 \leq k_{i-2}, k_{i-1} < r$ ". If it is possible, then we have $k_{i-1} < k_i$, which leads to $m \geq i$. This yields $k_{i-2} = i - 2$, from which we have $r - 1 \leq i - 2$, or $r \leq i - 1$. Then we obtain $k_{i-1} < i - 1$ because of $k_{i-1} < r$. This leads to the contradiction $i < i$. Thus the above joint event, that is, (4) $B < \infty$ and $C < \infty$ and $D = \infty$, must be impossible. It is easily seen from (1) to (4) above that the possible combinations as to finiteness(F) or infiniteness(∞) of each of A, B, C, and D are the following five:

Table I

	A	B	C	D
(a)	∞	∞	∞	∞
(b)	F	∞	∞	∞
(c)	F	∞	F	∞
(d)	F	F	∞	∞
(e)	F	F	F	F

It is clear that the term $()^* \leq 0$ in (5.8.3). For each of (a) to (e) in the above table, the term $()^{**}$ becomes, respectively, 0 for (a), $T_{r-1}(c_{i-1}^{r-1}) \geq 0$ for (b), $T_{r-1}(c_{i-1}^{r-1}) - T_{r-1}(c_{i-2}^{r-1}) \geq 0$ for (c), $\Delta T_r(c_i^r) \geq 0$ for (d), and $\Delta T_r(c_i^r) - \Delta T_r(c_{i-1}^r) \geq 0$ for (e). Hence eventually we have $y_i - y_{i-1} \leq 0$. Thus it follows by induction that for all $t \geq 1$ we get $y_i \geq 0$ and $y_i - y_{i-1} \leq 0$, which produces $c_i^r \geq 0$, \nearrow in r , and \searrow in i for all t . Furthermore from (5.8.2) we obtain $y_i \geq y_i^t$, hence $y_i \nearrow$ in t , which yields $c_i^r \nearrow$ in t .

5.9 Customer selection problem I [17]

Here we shall consider the following discrete-type version of a continuous time customer selection problem in [17]. Suppose there are M customer classes and at most one customer arrives at the start of each period. Let p_m be now the probability that an arriving customer is of class m . Then the probability of no customer arriving is given by $p_0 = 1 - \sum_{m=1}^M p_m$. Assume a customer of class m has a fixed reward

$r_m > 0$. Without loss in generality we may rearrange the rewards as $r_1 \geq r_2 \geq \dots \geq r_m > 0$. Every customer is served in a single station according to probability q_s , by which we mean the probability of completing services for s customers in one period. Here assume $q_0 > 0$. Then notice that if there are i customers in the system, the probability that services for all of them are completed in one period is given by $\sum_{s=i}^{\infty} q_s$. Furthermore postulate that the maximum permissible number of customers in the system is $N \geq 1$ and that we can select the customers to be accepted.

Our objective is to maximize the total expected reward over an infinite time horizon. Then a state of the system can be characterized by the number of customers in the system at the start of each period, i , hence the state space is given by $I = \{0, 1, \dots, N\}$. Now for a customer of class m arriving when in state i , let A_i^{m0} and A_i^{m1} represent the action of rejecting the customer and the action of accepting it, respectively. For state N only one action A_N^{m0} , that is, the action of not accepting any customer, is available. Thus the action space becomes $A_i^m = \{A_i^{m0}, A_i^{m1}\}$ for $i < N$ and $A_N^m = \{A_N^{m0}\}$. Then for these actions the immediate rewards are given by $x_i^{m0} = 0$ and $x_i^{m1} = r_m \theta$, that is, $a_i^{m0} = 0$ and $a_i^{m1} = r_m$, where θ is a random variable distributed according to a unit distribution function F_m with parameter $\lambda = 1$ (see remark 4). The after-effects are provided by

$$(5.9.1) \quad z_i^{m0} = \beta \sum_{s=0}^{i-1} q_s v'_{i-s} + \beta v'_0 \sum_{s=i}^{\infty} q_s$$

$$(5.9.2) \quad z_i^{m1} = \beta \sum_{s=0}^i q_s v'_{i+1-s} + \beta v'_0 \sum_{s=i+1}^{\infty} q_s$$

Then from (2.1) we have

$$(5.9.3) \quad c_i^{m1} = -\sum_{s=0}^i q_s y'_{i+1-s} / r_m$$

Now since both z_i^{m0} and z_i^{m1} are independent of m , let $z_i^0 = z_i^{m0}$ and $z_i^1 = z_i^{m1}$. Furthermore let $c_i^m = c_i^{m1}$. Then we have

$$(5.9.4) \quad c_i^m - c_{i-1}^m = -(\sum_{s=0}^{i-1} q_s (y'_{i+1-s} - y'_{i-s}) + q_i y'_1) / r_m$$

Since $a_i^{m0} < a_i^{m1}$, the FE becomes for $0 \leq i \leq N$, letting $c_N^m = \omega$,

$$v_i = p_0 z_i^0 + \sum_{m=1}^M p_m (z_i^0 + r_m T(c_i^m))$$

$$(5.9.5) \quad = \beta \sum_{s=0}^{i-1} q_s v_{i-s}' + \beta v_0' \sum_{s=i}^{\infty} q_s + \sum_{m=1}^M p_m r_m T(c_i^m)$$

from which we obtain for $1 \leq i \leq N$

$$(5.9.6) \quad \beta^{-1} y_i = \sum_{s=0}^{i-1} q_s y_{i-s}' + \sum_{m=1}^M p_m r_m (T(c_i^m) - T(c_{i-1}^m))$$

$$(5.9.7) \quad \geq \sum_{s=0}^{i-1} q_s y_{i-s}' - \sum_{m=1}^M p_m r_m T(c_{i-1}^m)$$

Noticing $T(c_{i-1}^m) - T(c_{i-2}^m) = -(c_{i-1}^m - c_{i-2}^m) + \hat{T}(c_{i-1}^m) - \hat{T}(c_{i-2}^m)$ for $2 \leq i \leq N$, we get from (5.9.6).

$$(5.9.8) \quad \beta^{-1} (y_i - y_{i-1}) = p_0 \left(\sum_{s=0}^{i-2} q_s (y_{i-s}' - y_{i-s-1}') + q_{i-1} y_1' \right) \\ + \sum_{m=1}^M p_m r_m ((T(c_i^m) - T(c_{i-1}^m)) - (\hat{T}(c_{i-1}^m) - \hat{T}(c_{i-2}^m)))$$

$2 \leq i \leq N$

For any $t \geq 1$ suppose $y_i' \leq 0$ and $y_i' - y_{i-1}' \leq 0$ for all i . Then we have $c_i^m \geq 0$ and \nearrow in i . Thus we get $y_i \leq 0$ from (5.9.6) and $y_i - y_{i-1} \leq 0$ from (5.9.8). Hence it follows by induction that $y_i \leq 0$ and $y_i - y_{i-1} \leq 0$ for all t . This produces $c_i^m \geq 0$ and \nearrow in i for all t . Thus we have β -limit $y_i \leq 0$ and \nearrow in i and β -limit $c_i^m \geq 0$ and \nearrow in i .

Now noticing $T(c_{i-1}^m) \leq T(0) = 1$, for β -limit y_i we have from (5.9.7) $0 \geq y_i \geq \beta \left(\sum_{s=0}^{i-1} q_s y_{i-s}' - h \right)$ where $h = \sum_{m=1}^M p_m r_m \geq 0$. For $i = 1$, we obtain $0 \geq y_1 \geq \beta(q_0 y_1 - h)$. From this we get $0 \geq \beta^{-1} y_1 \geq M_1 = -h/(1 - q_0)$ and $0 \geq y_1 \geq M_1$, noticing $-\beta h/(1 - \beta q_0) \searrow$ in β on $(0, 1)$. For any $i \geq 2$, we have $0 \geq y_i \geq \beta(q_0 y_i + d_i - h)$ where $d_i = \sum_{s=1}^{i-1} q_s y_{i-1}' \leq 0$. Hence similarly it follows that $0 \geq \beta^{-1} y_i \geq M_i = (d_i - h)/(1 - q_0)$ and $0 \geq y_i \geq M_i$. Suppose y_1, y_2, \dots, y_{i-1} are all bounded uniformly in β . Then since d_i becomes bounded uniformly in β , it follows that both $\beta^{-1} y_i$ and y_i are also bounded uniformly in β . Thus by induction these hold also for all i . Therefore the maximum expected reward per period for an infinite time horizon problem without discounting becomes, noticing the FE $v_0 = \beta v_0 + \sum_{m=1}^M p_m r_m T(c_0^m)$

for β -limit v_i ,

$$(5.9.9) \quad g = \sum_{m=1}^M p_m r_m T(c_0^m)$$

where c_0^m is 1-limit.

5.10 Customer selection problem II

Here the variation of the customer selection problem I is discussed where there exists only one class but multiple customers of the class arrive. Now let m customers arrive every period where m is a fixed positive integer, and assume that a reward of each of them is a random variable distributed according to a given distribution function where we shall rearrange the rewards of these customers as $\theta_1 \geq \theta_2 \geq \dots \geq \theta_m$, order statistics. Then the maximum permissible number of customers to be accepted when in state i is given by $k_i = \min\{m, N-i\} \leq m$. Denote an action of accepting r ($0 \leq r \leq k_i$) customers from the upper rank of rewards by A_i^r . Then for an action A_i^r the immediate reward is given by $x_i^r = \theta_1 + \theta_2 + \dots + \theta_r$ for $1 \leq r \leq k_i$ where $x_i^0 = 0$, and the after-effect z_i^r by

$$(5.10.1) \quad z_i^r = \beta \sum_{s=0}^{i+r-1} q_s v'_{i+r-s} + \beta v'_0 \sum_{s=i+r}^{\infty} q_s \quad 0 \leq r \leq k_i$$

From (3.1) we have $c_i^r = -\sum_{s=0}^{i+r-1} q_s y'_{i+r-s}$ for $0 < r \leq k_i$. Define $c_i^r = \infty$ for $r > k_i$ (see remark 6). On $K = \{c_i^r : 1 \leq r \leq k_i\}$ we have $c_i^r = c_{i-1}^{r+1} = c_{i+1}^{r-1}$ and $c_i^r - c_{i-1}^r = c_i^r - c_i^{r-1} = -\sum_{s=0}^{i+r-2} q_s (y'_{i+r-s} - y'_{i+r-s-1}) - q_{i+r-1} y'_1$. Suppose that $y'_i \leq 0$ and $y'_i - y'_{i-1} \leq 0$ for any $t \geq 1$. Then we get $c_i^r \geq 0$ and \nearrow both in i and in r . Hence the OP is given by (3.2) and the FE becomes from (3.3)

$$(5.10.2) \quad v_i = \beta \sum_{s=0}^{i-1} q_s v'_{i-s} + \beta v'_0 \sum_{s=i}^{\infty} q_s + \sum_{r=1}^{k_i} T_r(c_i^r) \quad 0 \leq i \leq N$$

Noticing remark 6 and lemma 4(a4), we have for $1 \leq i \leq N$

$$(5.10.3) \quad \beta^{-1} y_i = \sum_{s=0}^{i-1} q_s y'_{i-s} + \sum_{r=1}^{\infty} (T_r(c_i^r) - T_r(c_{i-1}^r))$$

$$(5.10.4) \quad \geq \sum_{s=0}^{i-1} q_s y'_{i-s} - \sum_{r=1}^{k_{i-1}} T_r(c_{i-1}^r)$$

Since $k_i \downarrow$ in i , $r \leq k_i$ leads to $r \leq k_{i-1}$. Hence if $c_i^r \in K$, then also $c_{i-1}^r \in K$. Thus we have always $c_i^r \geq c_{i-1}^r$. Consequently since the second term of (5.10.3) ≤ 0 , we have $y_i \leq 0$.

Furthermore for $2 \leq i \leq N$, both c_{i-1}^1 and c_{i-2}^1 are in K because of $1 \leq k_{i-1}$. Noticing this, we have

$$(5.10.5) \quad \beta^{-1}(y_i - y_{i-1}) = (\underbrace{T_1(c_{i-2}^1)}_A - \underbrace{T_1(c_{i-1}^1)}_B) * \\ - \sum_{r=2}^{\infty} ((\underbrace{T_{r-1}(c_{i-1}^{r-1})}_A - \underbrace{T_r(c_{i-2}^r)}_B)) - (\underbrace{T_{r-1}(c_i^{r-1})}_C - \underbrace{T_r(c_{i-1}^r)}_D) ** \\ 2 \leq i \leq N$$

It is clear that the term $(*) \leq 0$ in (5.10.5). Now it can be proved similarly for the problem that the table I in section 5.8 holds also for A to D in (5.10.5), hence the term $(**) \geq 0$. Thus we get $y_i - y_{i-1} \leq 0$. Consequently by induction we have $y_i \leq 0$ and $y_i - y_{i-1} \leq 0$ for all i , which yields β -limit $y_i \geq 0$ and \downarrow in i . This leads to β -limit $c_i^r \geq 0$ and \uparrow both in i and in r .

Since for β -limit y_i we have $y_i \geq \beta \left(\sum_{s=0}^{i-1} q_s y_{i-s} - h_i \right)$ from (5.10.4) where $h_i = \sum_{r=1}^{k_i-1} E_r \geq 0$, the uniform boundedness of y_i and $\beta^{-1} y_i$ in β can be proved in the exactly similar way as in section 5.9. Hence the maximum total expected reward per period for an infinite time horizon problem without discounting becomes, noticing the FE $v_0 = \beta v_0 + \sum_{r=1}^{k_0} T_r(c_0^r)$ associated with β -limits v_i ,

$$(5.10.6) \quad g = \sum_{r=1}^{k_0} T_r(c_0^r)$$

where c_0^r is an l -limit.

5.11 Customer selection problem III [10]

In the section we shall deal with the following variation of the customer selection problem II. Let time points be taken at an equally spaced interval and assume that every arriving customer has the same fixed service time w (periods), integer ≥ 1 . We shall refer to the total service time required for all customers now in the system as a

backlog and assume that the backlog must be equal to or less than N (integer ≥ 1) at any instant. Now since a state of the system can be given by the backlog at the start of each period, the state space becomes $I = \{0, 1, \dots, N\}$. The above definition of state implies that if r customers are accepted, then the current state i becomes $(i + r\omega - 1)^+$ $r = 0, 1, \dots$ at the next time. Since it must be $\leq N$ from the assumption above, the maximum permissible number of customers to be accepted in state i is given by $k_i = \min\{m, \lceil (N-i+1)/\omega \rceil\}$. Now the action, action space, and immediate reward for the problem are of quite the same form as in section 5.10, but the after-effect becomes $z_i^r = v_{(i+r\omega-1)^+}$. Now define

$$(5.11.1) \quad Y_i = \begin{cases} \beta(v_i - v_0) & 1 \leq i \leq \omega \\ \beta(v_i - v_{i-\omega}) & \omega < i \leq N \end{cases}$$

$$(5.11.2) \quad = \beta(v_i - v_{(i-\omega)^+}) \quad 1 \leq i \leq N$$

$$(5.11.3) \quad \Delta \tilde{Y}_i = \tilde{Y}_i - \tilde{Y}_{i-1} \quad 2 \leq i \leq N.$$

Then noticing $\omega \geq 1$, from (3.1) we have $c_i^r = -\tilde{Y}_{i+r\omega-1}^!$ for $0 < r \leq k_i$. Define $c_i^r = \infty$ for $r > k_i$. On $K = \{c_i^r : 0 < r \leq k_i\}$ we have $c_i^r = c_{i-\omega}^{r+1} = c_{i+\omega}^{r-1}$, $c_i^r - c_{i-1}^r = -\Delta \tilde{Y}_{i+r\omega-1}^!$, and $c_i^r - c_{i-1}^{r-1} = -\Delta \tilde{Y}_{i+r\omega-1}^! - \Delta \tilde{Y}_{i+r\omega-2}^! - \dots - \Delta \tilde{Y}_{i+r\omega-\omega}^!$.

Now suppose $\tilde{Y}_1^! \leq 0$ and $\Delta \tilde{Y}_t^! \leq 0$ for any $t \geq 1$. Then it follows that $c_i^r \geq 0$ and $c_i^r \nearrow$ both in i and in r . Hence the OP is given by (3.2), and the FE (3.3) becomes

$$(5.11.4) \quad v_i = \beta v_{(i-1)^+} + \sum_{r=1}^{k_i} T_r(c_i^r)$$

From this we get, noticing remark 6,

$$(5.11.5) \quad \beta^{-1} \tilde{Y}_1^! = \sum_{r=1}^{\infty} (T_r(c_1^r) - T_r(c_0^r))$$

$$(5.11.6) \quad \beta^{-1} \tilde{Y}_i^! = \tilde{Y}_{i-1}^! + \sum_{r=1}^{\infty} (T_r(c_i^r) - T_r(c_0^r)) \quad 2 \leq i \leq \omega$$

$$(5.11.7) \quad \beta^{-1} \tilde{Y}_i^! = \tilde{Y}_{i-1}^! + \sum_{r=1}^{\infty} (T_r(c_i^r) - T_r(c_{i-\omega}^r)) \quad \omega < i \leq N$$

Furthermore from (5.11.5) to (5.11.7) we have

$$(5.11.8) \quad \beta^{-1} \Delta \tilde{Y}_2 = \tilde{Y}_1^1 + \sum_{r=1}^{\infty} (T_r(c_2^r) - T_r(c_1^r))$$

$$(5.11.9) \quad \beta^{-1} \Delta \tilde{Y}_i = \Delta \tilde{Y}_{i-1}^1 + \sum_{r=1}^{\infty} (T_r(c_i^r) - T_r(c_{i-1}^r)) \quad 3 < i \leq \omega$$

Noticing $c_{i-\omega}^1 \in K$ and $c_{i-\omega-1}^1 \in K$ for $i \geq 2$, we have for $\omega < i \leq N$

$$(5.11.10) \quad \beta^{-1} \Delta \tilde{Y}_i = (\tilde{Y}_1^1(c_{i-\omega-1}^1) - \tilde{Y}_1^1(c_{i-\omega}^1))^* \\ - \sum_{r=2}^{\infty} ((\underbrace{T_{r-1}(c_{i-1}^{r-1})}_A - \underbrace{T_r(c_{i-\omega-1}^r)}_B)) - (\underbrace{T_{r-1}(c_i^{r-1})}_C - \underbrace{T_r(c_{i-\omega}^r)}_D))^{**}$$

Since $k_i \not\leq$ in i , $r \leq k_i$ leads to $r \leq k_j$ for $i > j$, in other words, $c_i^r \in K$ to $c_j^r \in K$ for $i > j$. Hence we have always $c_i^r \geq c_0^r$ and $c_i^r \geq c_{i-\omega}^r$. This yields $\tilde{Y}_i \leq 0$ for $1 \leq i \leq N$ from (5.11.5) to (5.11.7) and $\Delta \tilde{Y}_i \leq 0$ for $2 \leq i \leq \omega$ from (5.11.8) and (5.11.9). It is clear that the term $(*) \leq 0$ in (5.11.10). When noticing that (1) $k_i \not\leq$ in i , (2) $k_{i-\omega}$ is greater by at most 1 than k_i , and (3) $(x) + 1 > (x + 1)$ is impossible for any real number x , it is easily shown that table I in section 5.8 holds also for A to B in (5.11.10). Consequently we have the term $(**) \geq 0$, hence it follows that $\Delta \tilde{Y}_i \leq 0$ for $\omega < i \leq N$. Thus by induction we get $\tilde{Y}_i \leq 0$ and $\Delta \tilde{Y}_i \leq 0$ for all i and all $t \geq 1$. Therefore we have β -limit $\tilde{Y}_i \leq 0$ and $\not\leq$ in i which leads to β -limit $c_i^r \geq 0$ and \nearrow both in i and in r . Then letting $h_i = \sum_{r=1}^{k_i-1} E_r (\geq 0)$, for the β -limit \tilde{Y}_i we get from (5.11.5) to (5.11.7) $\tilde{Y}_1 \geq -\beta h_1 \geq -h_1$ and $\tilde{Y}_i \geq \beta(\tilde{Y}_{i-1} - h_i) \geq \tilde{Y}_{i-1} - h_i$ for $2 \leq i \leq N$. From these inequalities it follows that $0 \geq \tilde{Y}_i \geq \beta(-h_1 - h_2 - \dots - h_i) \geq \beta(-h_1 - h_2 - \dots - h_N)$ for $1 \leq i \leq N$. Hence $\beta^{-1} \tilde{Y}_i$ becomes bounded uniformly in β for all i . Thus the expected marginal profit per period for an infinite time horizon problem without discounting becomes, noticing the FE $v_0 = \beta v_0 + \sum_{r=1}^{k_0} T_r(c_0^r)$ for β -limit v_i ,

$$(5.11.11) \quad g = \sum_{r=1}^{k_0} T_r(c_0^r)$$

where c_0^r is an l-limit.

ACKNOWLEDGEMENT

The author is indebted to Professor S. Senju (Keio University) for stimulating his interest in this problem area, particularly in customer selection problem and purchasing problem. We are grateful to Dr. N. Matsubara, Dr. S. Nishimura and Dr. Y. Nogami for helpful suggestions and observations.

REFERENCES

- [1] Bellman, R., Dynamic programming, Princeton university press, (1957)
- [2] Derman, C., Lieberman, G.J. and Ross, S.M., A sequential stochastic assignment problem, Management Science Vol.18, NO.7, 349-355 (1972)
- [3] Fabian, T., Fisher, J.L., Sasieni, M.W. and Yardeni, A. Purchasing raw material on a fluctuating market, Operations Research, Vol.7, No.1, 107-122 (1959)
- [4] Gilbert, J.P. and Mosteller, F., Recognizing the maximum of a sequence, J. Amer. Statist. Assoc. Vol.16, 35-73 (1966)
- [5] Hayes, R.H., Optimal strategies for divestiture, Operations Research, Vol.17, No.2, 292-310 (1969)
- [6] Haggstrom, G.W., Optimal sequential procedures when more than one stop is required, Ann. Math. Statist. Vol.38, 1618-1626, (1969)
- [7] Hockman, E., An optimal stopping problem of a growing inventory, Management Science, Vol.19, No.11, 1289-1291 (1973)
- [8] Howard, R.A., Dynamic programming and Markov processes, MIT press, (1960)
- [9] Ikuta, S., A basic theory of optimal order-selection problem (Japanese), Journal of Japan industrial management association, Vol.46, 17-26 (1971)
- [10] Ikuta, S., A method of optimal order-selection (Japanese), Dissertation, Keio University, (1975)
- [11] Karlin, S., Stochastic models and optimal policy for selling an asset, Arrow, Karlin and Scarf (eds.) "Studies in applied probability and management science" Stanford university press, chapter 9, 148-158 (1962)
- [12] Kingsman, B.G. Commodity purchasing, Operations Research Quarterly, Vol.20, No.1, 59-79 (1969)
- [13] Lehmann, E.L., Ordered families of distributions, Ann. math. statist. Vol.26, 399-419, (1955).
- [14] Lippman, S.A. and McCall, J.J. Job search in a dynamic economy, Journal of Economic theory, Vol.12, 365-390 (1976)
- [15] MacQueen, J. and Miller, R.G. jr., Optimal persistence policies, Operations Research Vol.8; No.3, 362-380 (1960)

- [16] Mastran, D.V. and Thomas, C.J., Decision rules for attacking targets of opportunity, Naval Research Logistic Quarterly, Vol.20, 661-672, (1973)
- [17] Miller, B.L., A queueing reward system with several customer classes, Management Science, Vol.16, No.3, 234-245 (1969)
- [18] Morris, W.T., Some analysis of purchasing policy, Management Science, Vol.5, NO.4, 443-452 (1959)
- [19] Nishimura, S., Monotone optimal control of arrivals distinguished by reward and service time, Discussion paper series, NO.66 (80-4), University of Tsukuba, Institute of Socio-Economic planning (1980)
- [20] Ross, S.M, Applied probability models with optimization applications, Holden-Day, San Francisco (1970)
- [21] Sakaguchi, M., A sequential allocation problem for randomly appearing targets, Math. Japonicae, Vol.21, 89-103 (1976)
- [22] Sakaguchi, M., Optimal stopping problems for randomly arriving offers, Math. Japonicae, Vol.21, 201-217 (1976)
- [23] Sakaguchi, M., An investment problem: An optimal stopping problem in which two stops are required, Journal of Operations research Society of Japan, Vol.15, No.1, 45-52 (1972)
- [24] Sakaguchi, M., Dynamic programming of some sequential sampling design, Journal of Mathematical Analysis and Applications, Vol.2, 446-466, (1961)
- [25] Taylor, H.M., Evaluating a call option and optimal timing strategy in the stock market, Management Science, Vol.14, No.1, 111-120 (1967)
- [26] Wagner, H.M., Principles of operations research with applications to managerial decisions, Prentice-Hall, (1975)

