STUDIES
ON
THE THEORY OF OPTIMAL STOPPING
AND ITS APPLICATIONS TO
BEST CHOICE PROBLEMS

1984

Masami YASUDA

CONTENTS

INTRODUCTION

The theory of optimal stopping was first formulated in connection with the sequential analysis and can be found in the book "Sequential Analysis" by A. Wald in 1947. A general theory of optimal stopping for stochastic processes was developed after the appearance of works by J.L. Snell in 1952. Snell's theory means the classical super martingale characterization of the value process. Afterwards, in Markov processes with continuous time parameter, the connection between optimal stopping and free boundary problems was discovered, and the methods to apply the theory of variational inequalities to optimal stopping problems have been studied. The formulation of a Markov decision process is fairly general, as it includes a broad class of models of sequential optimization. An optimal stopping problem can be formulated as a two-action Markov decision process, in which one may either stop and receive a reward, or pay a cost and go to the next state. If we ignore the finiteness of stopping times, then, the existence of an optimal stopping time and the methods for finding the stopping time can be discussed under the framework of Markov decision Processes.

In this thesis, the author studies the theory of optimal stopping in the discrete time parameter processes, which have a new structure described in terms of the observer's action and the system's decision. Under this situation of the problem the optimality equation and the optimal policy are discussed. The motivation of the model comes from the multi-variate stopping problem and from the uncertain employment problem on secretary choices. Concerning the best choice problem, which is a particular case of the optimal stopping problem, an integral equation is given as an asymptotic form of the solution for the problem with a random number of objects. Under conditions on the distribution of the number of objects the integral equation is solved and consequently the asymptotic forms of optimal value and optimal policy are explicitly obtained.

In chapter 1, the author considers a stopping problem in which the observer's action and the system's two decisions are introduced. The observer can select a strategy defined on an action space, and the decision of the system to stop or continue is determined by a prescribed conditional probability. For this model, it may happen that the strategy to stop is refused, or to continue is forcibly stopped.

One of the typical application of the above model is the multi-variate stopping problem. A monotone rule is introduced in chapter 2 to sum up individual declarations. This is a reasonable generalization of the simple majority, veto power and hierarchical rules. The rule is defined by a monotone logical function and turns out to be equivalent to the winning class of Kadane. The existence of an equilibrium stopping strategy and the associated gain are discussed for the finite and infinite horizon cases.

Chapter 3 treats the best choice problem with a random number of objects provided its distribution is known. The optimality equation of the problem is reduced to an integral equation by a scaling limit. The equation is explicitly solved under some conditions on the distribution, which closely relates to the conditions for an OLA policy to be optimal in Markov decision processes. Also this technique is applied to three different versions of the problem and an exact form for asymptotic optimal strategy is derived.

1. OPTIMAL STOPPING PROBLEM INVOLVING REFUSAL AND FORCED STOPPING

1.1. FORMULATION

Suppose that the discrete time parameter process $X_n (=X(n))$, $n=1,2,\ldots$ is observed, and one selects a strategy from an action space A at each period. This strategy determines stochastically the system's decision to stop or continue. If the decision is to stop, one gets a reward for interrupting the observation; if the desicion is to continue, one observes the next value, and pays the cost. To be explicit, let $X(n)$, $n=1,2,\ldots$ be a stochastic process with a state space $E \subset R$ and let an action space A be a finite topological space. To start the first observation, one must pay a cost $c \in R$. Then, observing $X(1)$, one selects a strategy $\sigma_1 \in A$. Observing $X(n)$ at the n-th period and selecting a strategy $\sigma_n \in A$, one gets a net gain $X(n) - nc$ if the process's decision is to stop. If not, one incurs the cost c and observes $X(n+1)$.

The strategy σ_n at the n-th period is an A-valued $\mathcal{B}(X_1,\ldots,X_n)$-measurable random variable with its distribution $\phi_n(a) = P(\sigma_n = a)$, $a \in A$, where $(\Omega, \mathcal{B}, P)$ is an underlying probability space and $\mathcal{B}(X_1,\ldots,X_n) \subset \mathcal{B}$. The strategy σ denotes the infinite sequence $(\sigma_1,\ldots,\sigma_n,\ldots)$, and Σ is the set of all strategies.

Let us denote the system's desicion by the variables (S_n);

$$(1.1) \qquad S_n = \begin{cases} 1 & \text{if the decision of the process is stop at the n-th period,} \\ 0 & \text{if continue.} \end{cases}$$

The decision $\{S_n\}$ is determined only by the strategy σ_n at the n-th period, with the a conditional probability

$$(1.2) \qquad \gamma_n(a) = P(S_n=1 \mid \sigma_n=a), \quad a \in A$$

where $\gamma_n(a)$ is a given amount.

<u>Assumption 1.1</u> We assume that $\gamma_n(a)$ is independent of n, so that

(1.3) $\gamma(a) = \gamma_n(a)$ for all n.

For the space A, there exist

(1.4) $\alpha = \min \gamma(a) = \gamma(a_0)$, $\beta = \max \gamma(a) = \gamma(a_1)$, $a_0, a_1 \in A$.

To avoid a trivial case, assume (a), a A is not constant so that

(1.5) $0 \leqq \alpha < \beta \leqq 1$.

According to the setup of our model in the finite N-horizon case, the stopping time is defined by

(1.6) $$t_N(\sigma)\begin{cases} = \text{first}\{n \leq N;\ S_n=1\} \\ = N \quad \text{if } \{\ \} \text{ is empty.}\end{cases}$$

where $\sigma \in \sum$ is a strategy.

Our aim in the finite-horizon stopping problem is to maximize the expected gain

$$E\left[X(t_N(\sigma)) - ct_N(\sigma)\right]$$

subject to the strategy $\sigma \in \sum$. The optimal value V_N is defined by

(1.7) $V_N = \sup E\left[X(t_N(\sigma)) - ct_N(\sigma)\right]$.

The optimal strategy $*\sigma$ is such that $E\left[X(t_N(*\sigma)-ct_N(*\sigma)\right] = V_N$.

The difference from the usual stopping problem is that a conditional probability $\gamma(a)$ has been introduced into the connection between the observer's strategy and system's decision. Roughly, the observer's strategy, which determines the system's decision is interrupted by this. Two extremal probabilities are significant: $1 - \beta = 1 - \max \gamma(a)$; that is, the probability of refusal to stop the process, and $\alpha = \min \gamma(a)$; that is, the probability of forced stopping. If $\alpha = 0$ and $\beta = 1$ (no interruption), then the problem reduces to the usual one. The model is motivated by the uncertain secretary choice problem of Smith(1975) with $\beta = p$ $(0 < p \leqq 1)$ and $\alpha = 0$, and also the multi-variate stopping problem of Kurano, Yasuda, Nakagami(1980) including both refusal and forced stopping. These secretary choice problem is discussed in section 1.4.

1.2. OPTIMAL STRATEGY

ASSUMPTION 2.1 (i) Let X, X(n), n=1,2,... denote independent identically distributed (i.i.d.) random variables with $E|X| < \infty$. Denote their distribution function by F and let $\mu = \int_{-\infty}^{\infty} x dF(x) = E(X)$. (ii) Assume that $\mu < \sup\{x; F(x) < 1\}$.

The first assumption is not essential to our argument and we shall treat the non-identically distributed case in the example of Section 1.4.

Using the notation:

(2.1) $\quad T_{\alpha,\beta}(x) = E(X-x)^{+}\beta - E(X-x)^{-}\alpha$

where $(a)^{+} = \max(a,0)$ and $(a)^{-} = (-a)^{+}$, define the sequence (μ_n) as follows:

(2.2) $\quad \mu_1 = E(X),$

$$\mu_n = \mu_{n-1} - c + T_{\alpha,\beta}(\mu_{n-1} - c),\ n=2,3...$$

In the special case, $\beta=1$ and $\alpha=0$, (2.1) implies $T_{0,1}(x) = E(X-x)^{+} = \int_{x}^{\infty}(y-x)dF(y)$. This appears frequently in the ordinary stopping problem. Clearly $T_{\alpha,\beta}(x) = (\beta-\alpha)T_{0,1}(x) + \alpha(\mu-x)$. Also $T_{\alpha,\beta}(x)$ is a continuous, convex function of x and has two asymptotes. If $\alpha=0$, then $T_{0,\beta}(x) \geqq 0$ but generally it varies over $(-\infty, \infty)$. Therefore we note that the sequence (μ_n) is not monotone increasing in the case of $\alpha \neq 0$.

THEOREM 2.1 The optimal strategy $*\sigma = (*\sigma_1, \ldots, *\sigma_N, \ldots) \in \Sigma$ is given by

(2.3) $$*\sigma_n(\omega) = \begin{cases} a_1 & \text{if } X_n(\omega) \geq \mu_{N-n} - c \\ a_0 & \quad < \end{cases}$$

for n=1,2,..., and the optimal value is

(2.4) $$V_N = \mu_N - c.$$

Proof. In the case of N=1, the optimal value is clearly

$$V_1 = E(X_1) - c = \mu_1 - c$$

because the reward is X_1 and the cost incurred for the observation is c. As the usual dynamic programming's procedure, we assume inductively that (2.4) holds and consider the parameter N as a time-period left in the sequential decision process. When N-n time-periods are left, one must select σ_n = a from $a \in A$. If $S_n = 1$ occurs, then one gets X_n, otherwise $V_{N-n} = \mu_{N-n} - c$ since it reduces to the (N-n)-th period problem. One selects a strategy σ_n at the n-th period so as to maximize

$$E\left[X_n P(S_n=1 \mid X_n, \sigma_n) + V_{N-n} P(S_n=0 \mid X_n, \sigma_n)\right].$$

Since $P(S_n=1 \mid X_n, \sigma_n) = P(S_n=1 \mid \sigma_n)$ and $\gamma(a) = P(S_n=1 \mid \sigma_n=a)$, one is to maximize

$$E\left[\sum_{a\in A} (X_n - V_{N-n})\, \gamma(a)\, \phi_n(a)\right] + V_{N-n}$$

for $0 \leqq \phi_n(a) \leqq 1$ over all the densities. Hence if $X_n - V_{N-n} \geqq 0$,

$$\phi_n(a) = 1 \text{ if } a = a_1, \text{ and } \phi_n(a) = 0 \quad \text{otherwise}$$

and if $X_n - V_{N-n} < 0$,

$$\phi_n(a) = 1 \text{ if } a = a_0, \text{ and } \phi_n(a) = 0 \quad \text{otherwise.}$$

That is, the pure strategy (2.3) is optimal. Its maximum equals

$$E[(X_n - V_{N-n})^+\beta - (X_n - V_{N-n})^-\alpha] + V_{N-n}$$
$$= T_{\alpha,\beta}(\mu_{N-n} - c) + \mu_{N-n} - c = \mu_{N-n+1}.$$

The total optimal value is, with a cost c per observation, is

$$V_{N-n+1} = \mu_{N-n+1} - c.$$

This proves the theorem by letting n=1.

1.3. INFINITE HORIZON PROBLEM

Define a stopping time $t(\sigma)$ by

(3.1)
$$t(\sigma) = \begin{cases} \inf\{n \geqq 1 : S_n = 1\}, \\ \infty \quad \text{if } \{\ \} \text{ is empty} \end{cases}$$

for the strategy $\sigma \in \Sigma$. Let $X(t(\sigma)) = X(n)$ on $t(\sigma)=n$, $X(t(\sigma)) = \limsup X(n)$ on $t(\sigma) = \infty$. The optimal value V^* is defined by

(3.2) $\quad V^* = \sup E\left[X(t(\sigma)) - ct(\sigma)\right].$

ASSUMPTION 3.1 We assume (i) $\alpha > 0$ and c is any real number or (ii) $\alpha = 0$ and $c > 0$.

LEMMA 3.1 Under Assumptions 1.1, 2.1 and 3.1, the limit of the sequence (μ_n) of (2.2) exists:

(3.3) $\quad \lim \mu_n = v^* + c$

where v^* is the unique solution of the equation:

(3.4) $\quad T_{\alpha,\beta}(v) = c.$

Proof. Let $v_n = \mu_n - c$. The iteration (2.2) implies $v_n = v_{n-1} + T_{\alpha,\beta}(v_{n-1}) - c$. It is clear that the function $v + T_{\alpha,\beta}(v)$ of v is continuous, convex and monotone increasing. Also $g(v)$, the asymptote of $T_{\alpha,\beta}(v)$ as $v \to \infty$, is $g(v) = \alpha\mu + (1-\alpha)v$. Therefore (3.4) has a unique finite solution for $\alpha > 0$ and for any c. Under the conditions $\alpha = 0$ and $c > 0$, it holds similarly.

The property (3.3) is called stable by Ross(1970); we can therefore say the forced stopping problem is stable.

A necessary and sufficient condition that the solution v^* of (3.4) satisfies $v^* \geqq \mu$ is that $E(X - \mu)^+ \geqq c/(\beta - \alpha)$. If $c = 0$, the result is trivial and it holds that

(3.5) $\quad \mu \leqq v^* \leqq \sup\{x;\ F(x) < 1\}.$

Examples of the solution v^* in (3.4) with $c = 0$ are as follows.

(i) Normal distribution $N(0,1)$; $\quad 0 \leqq v^* \leqq \infty,$

$$\Psi(v) = \alpha v/(\beta - \alpha)$$

where $\Psi(v) = \phi(v) - v\Phi(v)$, $\Phi(v) = \int_v^{\infty} \phi(x)dx$ and $\phi(x)$ is a density function.

(ii) Exponential distribution with a density function $\lambda\exp(-\lambda x)$, $\lambda > 0$;

$$1/\lambda \leqq v^* \leqq \infty, \quad (\exp(-\lambda v))/(1-\lambda v) = -\alpha/(\beta - \alpha).$$

(iii) Uniform distribution on a unit interval $(0,1)$; $0.5 \leqq v^* \leqq 1$,

$$v^* = 1/(1 + \sqrt{(\alpha/\beta)}).$$

<u>LEMMA 3.2</u> The functional equation of $V(x)$, $x \in R$:

$$V(x) = \max_a(\gamma(a)x + (1-\gamma(a))\{E(V(X)) - c\}) \tag{3.6}$$

where $\gamma(a)$, $a \in A$ is in (1.4), has a unique solution in a functional space $\{V(x),\ x \in R\ ;\ E(V(X)) < \infty\}$ under Assumption 3.1. It is given by

$$V(x) = (x - v^*)^+\beta - (x - v^*)^-\alpha + v^* \tag{3.7}$$

where v^* is determined by Lemma 3.1.

<u>Proof.</u> We can show by straightforward calculation that (3.7) satisfies $E(V(X)) < \infty$ and (3.6). The uniqueness can be proved from the fundamental property of 'max' mapping in (3.6), as in Bellman(1957).

<u>THEOREM 3.3</u> In the infinite horizon case under Assumptions 1.1, 2.1 and 3.1, the strategy $^*\sigma = (^*\sigma_1, \ldots, ^*\sigma_n, \ldots) \in \Sigma$ with

$$^*\sigma_n(\omega) = \begin{cases} a_1 & X_n(\omega) \geqq v^*, \\ a_0 & \qquad < v^*, \end{cases} \tag{3.8}$$

$n=1,2,\ldots$ is optimal and the optimal value V^* is given by

$$V^* = v^*. \tag{3.9}$$

<u>Proof.</u> Let $V(x)$ denotes the optimal value when the first $X_1 = x$ is observed. By the optimality principle, $V(x)$ satisfies the optimality equation (3.6). It follows that, with the incurred cost c, the optimal value equals $V^* = E[V(X)] - c$. Hence (3.9) is immediately obtained from (3.7) and $E[V(X)] = v^* + c$.

THEOREM 3.4 In the case of $c = 0$, a sufficient condition that $P(t(*_\sigma) < \infty) = 1$ is that $\alpha > 0$.

Proof. Since $X(k)$, $k=1,2,..$ are i.i.d.,

$$P(t(*_\sigma) = n) = P(X(k) < v^*,\ k=1,..,n-1,\ \ X(n) \geqq v^*)$$
$$= (1 - F(v^*-))(F(v^*-))^{n-1}.$$

Now $c=0$ implies that $E(X-v^*)^+/E(X-v^*)^- = \alpha/\beta$. If $\alpha > 0$, then $E(X-v^*)^+ > 0$ yields $v^* < \sup\{x;F(x) < 1\}$ and so $F(v^*) < 1$. From these, the conclusion is immediate.

1.4. APPLICATION TO A BEST CHOICE PROBLEM

Let the observation cost $c = 0$ and let $X(n)$, $n=1,..,N$ be independent random variables such that

(4.1) $$X(n) = \begin{cases} n/N & \text{with probability } 1/n, \\ 0 & \text{with probability } 1-1/n. \end{cases}$$

The stopping problem for this process (4.1) is called the secretary choice problem by Chow, Robbins and Siegmund(1971). We will not assume (1.3) and use α_n, β_n instead of α, β in (1.4). It is seen, indexing the time parameter, that results similar to Theorem 2.1 hold. Define

(4.2) $$T^{(n)}(x) = T^n_{\alpha_n,\beta_n}(x) = E(X_{N-n} - x)^+\beta_n - E(X_{N-n} - x)^-\alpha_n$$

for $n=0,..,N-1$ in place of (2.1). From (4.1),

$$T^{(n)}(x) = \begin{cases} \beta_n/N - \beta_n x & \text{if } x \leq 0, \\ \beta_n/N - (\alpha_n + (\beta_n-\alpha_n)/N_n)x & \text{if } 0 \leq x \leq N_n/N, \\ \alpha_n/N - \alpha_n x & \text{if } N_n/N \leq x, \end{cases}$$

where $N_n = N - n$. Since c=0, $V_n = \mu_n$ holds by (2.4), and so consider the following sequence similar to (2.2):

$$V_1 = E[X_N] = 1/N, \tag{4.3}$$
$$V_n = V_{n-1} + T^{(n-1)}(V_{n-1}),\ n=2,3,\ldots$$

This is different from the usual problem; if $\alpha_n \neq 0$, we note that the sequence V_n is not generally monotone increasing.

ASSUMPTION 4.1 Let α_n, β_n satisfy the conditions:

(i) $0 \leqq \alpha_n < \beta_n \leqq 1$,

(ii) $\beta_n \geqq \beta_{n+1}$,

(iii) $\alpha_n - \alpha_{n+1} + \alpha_n\alpha_{n+1} \leqq 0$

for each n.

LEMMA 4.1 Under Assumption 4.1, if $n/N \geqq V_{N-n}$ for some n, then it holds also for later n.

Proof. If V_{N-n} is concave in n, the lemma is immediately proved since the boundary V_N at n=0 is strictly positive, that is, the initial position is above the straight line n/N. To prove this, it is enough to show that

$$T^{(n)}(V_n) - T^{(n-1)}(V_{n-1}) \leqq 0. \tag{4.4}$$

First, we show that $T^{(n)}(x) \leqq T^{(n-1)}(x)$, $0 \leqq x < \infty$; this follows because $T^{(n)}(x)$ is a convex function of x and is composed of three line segments. Hence it is sufficient to consider the inequality at $x = N_{n+1}/N$ and $x = N_n/N$. The result is immediate at these points for the increasing α_n and decreasing β_n following from Assumptions 4.1 (i),(iii).

To prove (4.4), we restrict β_n to be a constant in n, without loss of generality. Because, for a general β_n, the gradient of $T^{(n)}(x)$ on $0 \leqq x \leqq N_n/N$ decreases, the above arguments hold independently of β_n on $x \geqq N_n/N$. Consider a function of x:

$$S^{(n)}(x) = T^{(n+1)}(x+y) - T^{(n)}(x)$$

where $y = T^{(n)}(x)$. On $0 \leqq x \leqq N_n/N$, if $y=T^{(n)}(x) \leqq 0$, $S^{(n)}(x) \leqq 0$ follows by considering

$$S^{(n)}(N_n/N) = \alpha_{n+1}/N - \alpha_{n+1}(N_n/N+y) - (-\alpha_n N_{n+1}/N)$$
$$= (\alpha_n - \alpha_{n+1} + \alpha_n \alpha_{n+1}) N_{n+1}/N \leqq 0.$$

If $y \geqq 0$, clearly $T^{(n+1)}(x+y) \leqq T^{(n+1)}(x) \leqq T^{(n)}(x)$ holds by the monotone decreasing property of $T^{(n)}(x)$ in n and x. For $x > N_n/N$, we easily see that $y=T^{(n)}(x) < 0$ and

$$S^{(n)}(x) = \alpha_{n+1}/N - \alpha_{n+1}(x+y) - y = (\alpha_n - \alpha_{n+1} + \alpha_n \alpha_{n+1})(x-1/N) \leqq 0$$

by Assumption 4.1(iii). We have thus obtained $S^{(n)}(x) \leqq 0$ on $0 \leqq x$ and so completed the proof of the lemma.

The optimal policy $*\sigma$ is, by (2.3) in Theorem 2.1, such that $*\sigma_n = a_1$ if $X_n \geqq V_{N-n}$ occurs or $n/N \geqq V_{N-n}$; that is, we declare "stop" if the relative best applicant has appeared. Define

(4.5) $\qquad n^* = \inf\{n \,;\, n/N \geqq V_{N-n}\}.$

By Assumption 4.1 and Lemma 4.1, the optimal strategy of the considering problem is the OLA policy(refer to Ross(1970)). The result is summarized as follows.

<u>Theorem 4.1</u> The optimal strategy of the secretary choice problem is $*\sigma_n = a_0$ for $n=1,..,n^*-1$ and $*\sigma_n = a_1$ for $n=n^*,..,N$. That is, observe applicants until n^*-1 and then declare "stop" if an appeared one is relatively the best among the previous ones.

In the rest of the section, we study the limiting procedure by allowing N tend to infinity. Two special cases of the coefficients α_n and β_n are considered.

(I) REFUSAL AND NO-FORCED STOPPING

Let

(4.6) $\beta_n = p$ and $\alpha_n = 0$

where p is a constant($0 < p \leqq 1$). Since $\alpha_n = 0$, there occurs no-forced stopping, and this is the uncertain employment case considered by Smith(1975). By (4.3) and (4.5), we have

(4.7) $$n^* = \inf\{ n \; ; \; p(\frac{1}{n} + (\frac{n+\bar{p}}{n})\frac{1}{n+1} + \dots + (\frac{n+\bar{p}}{n})(\frac{n+1+\bar{p}}{n+1})\dots(\frac{N-3+\bar{p}}{N-3})\frac{1}{N-2}) + (\frac{n+\bar{p}}{n})\dots(\frac{N-2+\bar{p}}{N-2})\frac{1}{N-1} \leq 1 \}$$

where $\bar{p} = 1 - p$. If $p=1$, (4.7) becomes

$$n^* = \inf\{n \; ; \; 1/n + 1/(n+1) + \dots + 1/(N-1) \leqq 1 \}$$

as is well-known.

If $p < 1$ and $v_1 = p/N$, (4.3) and (4.5) imply

(4.8) $n^* = \inf\{n \; ; \; p(1+\bar{p}/n)(1+\bar{p}/(n+1))\dots(1+\bar{p}/(N-1)) \leqq 1 \}$.

This result is obtained by Smith(1975). The limit is

(4.9) $\lim n^*/N = p^{1/(1-p)}$.

This value holds for both the cases (4.7) and (4.8). This is seen in the next generalized situation.

(II) REFUSAL AND FORCED STOPPING

Let

(4.10) $\beta_n = p$ and $\alpha_n = q/(N-n)$

where p and q are constants with $0 \leqq q < p \leqq 1$.

The situation in this secretary choice problem is that there are two observers, one is a young man who wants to choose a secretary and the other is his grandmother who also observes applicants. Each of the applicants ranks indedendently and also assume that there are no relation between two components of the rank. The problem is to find the best one with respect to the young man's rank. As a stopping rule, he could choose a candidate if he thinks she is best, in accordance with the possibility of refusal p. Aside from this case, there occurs forced stopping. That is, although he thinks that a candidate is not the best, he is forcibly stopped and must accept her when his grandmother thinks her the best one. The factor q denotes the strength of this effect.

Clearly this reduces to case(I) if $q = 0$ and (4.10) satisfies Assumption 4.1. Now we proceed to calculate $\lim n^*/N$ as before where n^* is given in (4.5). By (4.3), if $N_n/N < V_n$,

$$V_{n+1} = V_n + \left(\frac{1}{N} - \frac{V_n}{N_n}\right) p - \frac{N_{n+1}}{N_n}\alpha_n V_n$$

$$= p/N + \eta_n V_n$$

where $\eta_n = \bar{\alpha}_n + (\alpha_n - p)/N_n$ and $\bar{\alpha}_n = 1 - \alpha_n$. Hence we have, from the iteration (4.3) and the property of the optimal strategy, that

(4.11) $V_{n+1} = p(1+\eta_n+\eta_n\eta_{n-1}+\cdots+\eta_n\eta_{n-1}\cdots\eta_1)/N + (1-p)\eta_n\eta_{n-1}\cdots\eta_1/N$

(4.12) $\eta_n = 1 - (p+q)/N_n + q/N_n^2 = \delta_n N_{n+1}/N_n$

where

$$\delta_n = 1 + (\bar{p}-q)/N_{n+1} + q/(N_n N_{n+1}) = 1 + \bar{p}/N_{n+1} - q/N_n$$

and $\bar{p} = 1 - p$. Substituting (4.12) in (4.11), we obtain

$$V_{n+1} = \frac{N_{n+1}}{N}\left\{ p\left(\frac{1}{N_{n+1}} + \frac{\delta_n}{N_n} + \frac{\delta_n\delta_{n-1}}{N_{n-1}} + \cdots + \frac{\delta_n\delta_{n-1}\cdots\delta_1}{N_1}\right) + \bar{p}\frac{\delta_n\delta_{n-1}\cdots\delta_1}{N_1} \right\}.$$

By (4.5), we must find $\{$first n such that $n/N \geq V_{N-n}\}$ so

$$(4.13)\quad \inf\{\, n\ ;\ p\Big(\frac{1}{n} + \delta_{N_{n+1}}\frac{1}{n+1} + \dots + \delta_{N_{n+1}}\delta_{N_{n+2}}\cdots\delta_1\frac{1}{N-1}\Big) + \bar{p}\,\frac{\delta_{N_{n+1}}\cdots\delta_1}{N-1} \le 1 \,\}$$

In the limiting procedure, it is enough to consider the relation between n and N:

$$(4.14)\quad 1/p = 1/n + \Big(1 + \frac{\bar{p}-q}{n} + \frac{q}{n(n+1)}\Big)/(n+1) + \dots$$
$$+ \Big(1 + \frac{\bar{p}-q}{n} + \frac{q}{n(n+1)}\Big)\cdots\Big(1 + \frac{\bar{p}-q}{N-2} + \frac{q}{(N-2)(N-1)}\Big)/(N-1) + \frac{\bar{p}}{p}\,\frac{\delta_{N_{n+1}}\cdots\delta_1}{N-1}.$$

From the principal terms of δ_n, we can write

$$\delta_{N_{n+1}} = 1 + (\bar{p}-q)/n + o(1/n)$$

where $o(1/n)$ denotes terms of order smaller than $1/n$. Hence (4.14) implies

$$1/p = 1/n + \Big(1+\frac{\bar{p}-q}{n}\Big)/(n+1) + \dots + \Big(1+\frac{\bar{p}-q}{n}\Big)\cdots\Big(1+\frac{\bar{p}-q}{N-2}\Big)/(N-1) + \frac{\bar{p}}{p}\,\frac{\delta_{N_{n+1}}\cdots\delta_1}{N-1} + o(1)$$

where $o(1)$ is a term of negligible order as $n \to \infty$. Rearranging the sum in (4.14), we have

$$(1-q)/p = \Big(1 + \frac{\bar{p}-q}{n}\Big)\cdots\Big(1 + \frac{\bar{p}-q}{N-1}\Big) + \frac{\bar{p}(\bar{p}-q)}{p}\,\frac{\delta_{N_{n+1}}\cdots\delta_1}{N-1} + o(1)$$

provided $p + q \ne 1$. The last two terms of the above equality are negligible. Using the approximation $1+x \simeq \exp(x)$,

$$\log((1-q)/p) = (\bar{p}-q)\sum_{k=n}^{N-1} k^{-1} + o(1)$$

Therefore we have obtained the result that

$$(4.15)\quad \lim n^*/N = (p/(1-q))^{1/(1-p-q)} \quad \text{for } p + q \ne 1.$$

If $p + q = 1$, by (4.14), we have

$$1/p = 1/n + \dots + 1/(N-1) + o(1)$$

which implies

$$(4.16) \qquad \lim n^*/N = \exp(-1/p).$$

In (4.15), since

$$(p/(1-q))^{1/(1-p-q)} = \exp\left(- \frac{\log(p) - \log(1-q)}{p - (1-q)} \right)$$

when $1 - q \to p$, we have $\exp(-1/p)$. So there is no gap between (4.15) and (4.16). Letting $q = 0$ in (4.15), this reduces to $p^{1/(1-p)}$ as in Smith's (1975) refusal and no-forced stopping case, while letting $p = 1$ in (4.15), it reduces to $(1-q)^{1/q}$ as in the forced stopping case. From this, we see that p and $1-q$ in $\beta_n = p$ and $\alpha_n = q/N$ have a dual property.

2. MULTI-VARIATE STOPPING PROBLEMS WITH A MONOTONE RULE

2.1. STATEMENT OF THE PROBLEM

Let X_n, $n=1,2,..$ be p-dimentional random vectors on a probability space $(\Omega, \mathcal{B}, P)$. The process $\{X_n\}$ can be interpreted as the payoff to a group of p players. Each of p players observes sequentially values of X_n. Its distribution is assumed to be known to all of them. Players must make a declaration to either "stop" or "continue" on the basis of the observed value at each stage. A group decision whether to stop the process or not is summed up from the individual declarations by using a prescribed rule.

If the decision is to stop at stage n, then player i's net gain is

(1.1) $$Y_n^i = X_n^i - nc^i$$

where c^i is a constant observation cost. According to the individual declarations, let define random variables d_n^i, $n=1,2,..$, $i=1,..,p$ by

(1.2) $$d_n^i = \begin{cases} 1 & \text{if player i declares to stop,} \\ 0 & \text{continue.} \end{cases}$$

We assume, for each n and i,

(1.3) $$d_n^i \in \mathcal{B}(X_n)$$

where $\mathcal{B}(X_n)$ denotes the σ-algebra generated by X_n.

<u>DEFINITION 1.1.</u> An individual stopping strategy(abr. by ISS) is a sequence of random variables

(1.4) $$d^i = (d_1^i, d_2^i, ..., d_n^i, ..)$$

satisfying (1.3). $\mathcal{D}^i$ denotes the set of all ISS's for player i. A p-dimentional $\{0,1\}$-valued random vector

(1.5) $$d_n = (d_n^1, d_n^2, .., d_n^p)$$

denotes the declarations of p players at stage n. A stopping strategy(abr. by SS) is the sequence

(1.6) $\quad d = (d_1, d_2, .., d_n, ..)$

and $\mathfrak{D}$ denotes the whole set of the SS's.

Now we shall define a stopping rule by which a group decision is determined from the declarations of p players at each stage. A p-variate $\{0,1\}$-valued logical function

(1.7) $\quad \pi = \pi(x^1, .., x^p) \quad : \{0,1\}^p \rightarrow \{0,1\}$

is said to be monotone (cf. Fishburn(1971)) if

(1.8) $\quad \pi(x^1, .., x^p) \leq \pi(y^1, .., y^p)$

whenever $x^i \leq y^i$ for each i.

<u>DEFINITION 1.2.</u> A stopping rule(abr. by SR) is a non-constant logical function π and a monotone SR is an SR π with

(i) monotone and

(ii) $\pi(1,1,..,1) = 1$.

In this paper an SR means not "when to stop" the process but "how to sum up" the whole players' declarations. The property (ii) is called unanimity in Fishburn(1971). Its dual property $\pi(0,0,..,0) = 0$ is not needed to assume here. A constant function makes the problem trivial because the decision is always to stop from (ii).

The monotone SR has a wide variety in choice systems of our real life and shows a natural requirement in the analysis of our problem. Some examples for the monotone SR are given as follows.

<u>EXAMPLE 1.1.</u> (i) (Equal majority rule) In the group of p players, if no less than $r(\leq p)$ members declare to stop, then the group decision is to stop the process. That is,

(1.9) $\quad \pi(d_n^1, ..., d_n^p) = 1 \ (0) \quad \text{if} \sum_{i=1}^{p} d_n^i \geq (<) \quad r.$

For instance, a simple majority for three players, $(p,r)=(3,2)$, is

$$\pi(d_n^1, d_n^2, d_n^3) = d_n^1 \cdot d_n^2 + d_n^2 \cdot d_n^3 + d_n^3 \cdot d_n^1$$

where + is a logical sum and $\cdot$ is a logical product. The stopping problem of the majority rule was discussed in Kurano, Yasuda and Nakagami(1980).

(ii) (Unequal majority rule) A straightforward extension of (1.9) is

$$(1.10) \qquad \pi(d_n^1,\ldots,d_n^p) = \begin{cases} 1 & \text{if } \sum_{i=1}^{p} w^i d_n^i \geq r \\ 0 & \text{if } \qquad\quad < r \end{cases}$$

where $w^i \geq 0$, $i=1,..,p$ are given weighting constants. Including these cases, monotone rules have wide varieties. See Table 3.1 in Section2.3 for several rules with p=3.

(iii) (Hierarchical rule) A hierarchical system or Murakami's representative system(cf. Fishburn(1971)) is regarded as a composed rule. Since a composition of two monotone logical functions is monotone and (ii) of Def.1.2 holds, the hierarchical rule is also a monotone SR.

<u>DEFINITION 1.3.</u> For an SS $d=(d_1,d_2,..) \in \mathcal{D}$ with $d_n=(d_n^1,..,d_n^p)$, $n=1,2,..$ and a monotone SR π, a stopping time $t_\pi(d)$ is defined by

$$(1.11) \qquad t_\pi(d) = \begin{cases} \text{first } n \text{ such that } \pi(d_n^1,..,d_n^p) = 1, \\ \infty \quad \text{if no such } n \text{ exists.} \end{cases}$$

For any stopping time $t_\pi(d)$, let

$$(1.12) \qquad Y^i_{t_\pi(d)} = \begin{cases} Y_n^i & \text{on } t_\pi(d) = n, \\ \limsup_{n\to\infty} Y_n^i & \text{on } t_\pi(d) = \infty. \end{cases}$$

When the group decision is to stop at the time $t_\pi(d)$, player i gets $Y^i_{t_\pi(d)}$ as a net gain.

<u>DEFINITION 1.4.</u> Let π be a monotone SR. We call $*d = (*d^1,\ldots,*d^p)$ an equilibrium SS with respect to π if, for each i and any $d^i \in \mathcal{D}^i$,

$$(1.13) \qquad E\left[Y^i_{t_\pi(*d)}\right] \geq E\left[Y^i_{t_\pi(*d(i))}\right]$$

where $*d(i) = (*d^1,\ldots,*d^{i-1},d^i,*d^{i+1},\ldots,*d^p)$.

In this paper we treat the vector valued expected net gain

$$(1.14) \qquad E\left[Y_{t_\pi(d)}\right], \quad d \in \mathcal{D}$$

and our objective is to fined an equilibrium SS $^*d \in \mathcal{D}$ for a given monotone SR π. The notion of equilibrium owes to the non-cooperative game theory by Nash(1951).

In order to denote a stopping event of the system for a given SR, we need a set valued function on $\mathcal{B}^p(X_n)$. For an SS $d=(d_1,d_2,..)$, we call

$$(1.15) \qquad D_n^i = \{ \omega \in \Omega \mid d_n^i(\omega) = 1 \} \in \mathcal{B}(X_n)$$

an individual stopping event(abr. by ISE) for player i at stage n. If D_n^i occurs, i.e., $\omega \in D_n^i$, then player i declares to stop. So

$$(1.16) \qquad d_n^i = I_{D_n^i}$$

where I_D is an indicator of a set D on Ω. Hence there exists a set valued function Π on $\mathcal{B}^p(X_n)$ corresponding to a logical function π on $\{0,1\}^p$, such that

$$(1.17) \qquad \pi(d_n^1,..,d_n^p) = \pi(I_{D_n^1},..,I_{D_n^p}) = I_{\Pi(D_n^1,..,D_n^p)}.$$

Clearly two functions π and Π are related to each other. For example, $\pi(d_n^1,d_n^2,d_n^3) = d_n^1 + d_n^2 \cdot d_n^3$ corresponds to $\Pi(D_n^1,D_n^2,D_n^3) = D_n^1 \cup (D_n^2 \cap D_n^3)$.

The stopping event(abr. by SE) of the process at stage n is denoted by

$$(1.18) \qquad D_n = \{ \omega \in \Omega \mid \pi(d_n^1,..,d_n^p)=1 \} = \Pi(D_n^1,..,D_n^p).$$

We note that, if an SR π is monotone, $A^i \subset B^i$ for each i implies

$$(1.19) \qquad \Pi(A^1,..,A^p) \subset \Pi(B^1,..,B^p)$$

from (1.8).

DEFINITION 1.5. For a given (monotone) SR π, a corresponding set valued function Π is called a (monotone) stopping event rule(abr. by SER).

Next, a one-stage stopping model is considered to clarify an SS of our problem. Each player observes a random variable $X = (X^1,..,X^p)$ with $E|X^i| < \infty$ and player i receives a net gain $X^i - c^i$ if the group decision is to stop, or $v^i - c^i$ if not, where v^i is a given constant. If they use a monotone SR π, the SE of the system becomes $\Pi(D^1,..,D^p)$ for ISE D^i, i=1,..,p. Then the expected net gain for player i is expressed by

$$(1.20)\quad E\left[(X^i-c^i)I_{\Pi(D^1,\dots,D^p)}\right] + P(\overline{\Pi(D^1,\dots,D^p)})(v^i-c^i)$$
$$= E\left[(X^i-v^i)I_{\Pi(D^1,\dots,D^p)}\right] + v^i - c^i.$$

Since a logical function can be written generally as

$$\pi(x^1,\dots,x^p) = x^i\cdot\pi(x^1,\dots,\overset{i}{1},\dots,x^p) + \overline{x^i}\cdot\pi(x^1,\dots,\overset{i}{0},\dots,x^p),$$

it holds

$$(1.21)\quad \Pi(D^1,\dots,D^p) = \{D^i \cap \Pi(D^1,\dots,\overset{i}{\Omega},\dots,D^p)\}$$
$$\cup\{\overline{D^i}\cap\Pi(D^1,\dots,\overset{i}{\phi},\dots,D^p)\}$$

in terms of the SER. Substituting this into the last expression of (1.20), it becomes

$$(1.22)\quad \int_{D^i}(X^i-v^i)\{I_{\Pi(D^1,\dots,\Omega,\dots,D^p)} - I_{\Pi(D^1,\dots,\phi,\dots,D^p)}\}\,dP$$
$$+ \int_{\Pi(D^1,\dots,\ \ ,\dots D^p)}(X^i-v^i)\,dP + v^i - c^i.$$

By (1.19), it is clear that $I_{\Pi(D^1,\dots,\ \ ,\dots,D^p)} - I_{\Pi(D^1,\dots,\ \ ,\dots,D^p)} \gneq 0$.

Therefore we can derive the next proposition.

<u>PROPOSITION 1.1.</u> When $D^1,\dots,D^{i-1},D^{i+1},\dots,D^p$ are fixed, player i's maximum expected net gain subject to $D^i \in \mathcal{B}(X)$ is attained by

$$(1.23)\quad {}^*D^i = \{X^i \geq v^i\},$$

and it equals

$$(1.24)\quad \int_{\Omega}(X^i-v^i)^+ I_{\Pi(D^1,\dots,\Omega,\dots,D^p)}\,dP - \int_{\Omega}(X^i-v^i)^- I_{\Pi(D^1,\dots,\phi,\dots,D^p)}\,dP$$
$$+ v^i - c^i,$$

where $x^+=\max(x,0)$ and $x^-=\max(-x,0)$. Especially, when $\Pi(D^1,\dots,\Omega,\dots,D^p) = \Pi(D^1,\dots,\phi,\dots,D^p)$, player i's expected net gain (1.22) or (1.24) is constant not depending on D^i.

By Prop.1.1, we have solved a one-stage problem where the seeking equilibrium SS is given as (1.23) and we showed that player i's ISS depends on the i-th component X^i only among the p-dimensional vector X. In fact, it is seen intuitively as follows. Because the larger he observes his value, the larger he obtains his net gain, so he is eager to declare to stop. This situation holds under a monotonicity of the rule, but does not hold under another rule including negation. The negation is quite the opposite of one's intention. It is known that the monotone logical function does not include negation and vice versa. Other essential one is "non-cooperative" character in a reward, so other players' net gains do not affect his gain. Therefore, he observes his own value closely.

In the end of this section we refer to the winning class of Kadane (19778). He proved the conjecture of Sakaguchi(1978), that is, the reversibility in the juror problem by the choice of many persons. To prove the reversibility affirmative, he used a notion of the winning class as a choice rule.

DEFINITION 1.6. Let p denotes a number of players. A family $\mathcal{W}$ of subsets of integers $\{1,2,..,p\}$ is called a winning class if

(i) $\{1,2,..,p\} \in \mathcal{W}$

(ii) $W \in \mathcal{W}$, $W' \supset W$ implies $W' \in \mathcal{W}$.

Assume that r players, e.g., player $i_1,..,i_r$ declare to stop. Then the process must be stopped if a set $\{i_1,..,i_r\}$ is an element of $\mathcal{W}$, or continued if otherwise.

For a non-empty subset $W=\{i_1,..,i_r\}$ of $\{1,2,..,p\}$ there corresponds a vertex x of the p-dimensional unit cube whose i_1-, i_2-,.. and i_r-th component are equal to 1 and remaining components 0. For two correspondences between W_1, W_2 and x_1, x_2 respectively, a necessary and sufficient condition that $W_1 \subset W_2$ is that $x_1 \leqq x_2$ (component-wise). Let V be a set of vertices corresponding to a winning class $\mathcal{W}$. Define a logical function π by

$$\pi(x^1,..,x^p)=1 \quad \text{if } (x^1,..,x^p) \in V,$$
$$=0 \quad \text{otherwise.}$$

Then the following proposition holds immediately.

PROPOSITION 1.2. The stopping rule by a winning class of players, Def.1.6, is equivalent to the one by a monotone logical function, Def.1.2.

2.2. A FINITE HORIZON CASE

Consider the finite horizon case restricted by a prescribed number $N < \infty$. Our object is to find an equilibrium SS for a given SR and determine the associated expected net gain under the situation formulated in the previous section.

ASSUMPTION 2.1.

(a) For any SS $d=(d_1,..,d_n,..) \in \mathcal{D}$, $d_N^i = 1$ for $i=1,..,p$ with prob.1.

(b) Random vectors $X_1,..,X_N$ are independent and $E|X_n^i| < \infty$ for each n,i.

(c) A logical function π is a monotone SR.

Let us consider a sequence of vectors $V_n=(v_n^1,..,v_n^p)$ defined by

$$(2.1) \quad v_{n+1}^i = v_n^i - c^i + E[(X_{N-n}^i - v_n^i)^+ \beta_n^{\Pi\{i\}}(V_n^{\{i\}}|X_{N-n}^i)]$$
$$- E[(X_{N-n}^i - v_n^i)^- \alpha_n^{\Pi\{i\}}(V_n^{\{i\}}|X_{N-n}^i)], \quad n \geq 1,$$

$$(2.2) \quad v_1^i = E[X_N^i] - c^i$$

where $V_n^{\{i\}}=(v_n^1,\dots,v_n^{i-1},v_n^{i+1},\dots,v_n^p) \in R^{p-1}$, $i=1,\dots,p$,

$$(2.3) \quad \beta_n^{\Pi\{i\}}(V_n^{\{i\}}|X_{N-n}^i) = P(\Pi(*D_{N-n}^1,\dots,*D_{N-n}^{i-1},\Omega,*D_{N-n}^{i+1},\dots,*D_{N-n}^p)|X_{N-n}^i)$$

$$(2.4) \quad \alpha_n^{\Pi\{i\}}(V_n^{\{i\}}|X_{N-n}^i) = P(\Pi(*D_{N-n}^1,\dots,*D_{N-n}^{i-1},\phi,*D_{N-n}^{i+1},\dots,*D_{N-n}^p)|X_{N-n}^i)$$

and Π is the SER corresponding to the SR π and

$$*D_{N-n}^i = \{X_{N-n}^i \geq v_n^i\} \in \mathcal{B}(X_{N-n}), \quad i=1,\dots,p.$$

From Assump.2.1 (a) and (c), $P(t_\pi(d) \leq N) = 1$ holds for all SS $d \in \mathcal{D}$ even if the observation cost is negative.

THEOREM 2.1. By a sequence $V_n=(v_n^1,..,v_n^p)$, $n \geq 1$ in (2.1) and (2.2), let us define an SS $*d \in \mathcal{D}$ as follows: For n=1,..,N-1,

$$(2.5) \quad *d_n^i(\omega) = \begin{cases} 1 & \text{if } \omega \in *D_n^i, \text{ i.e., } X_n^i(\omega) \geq v_{N-n}^i, \\ 0 & \text{otherwise} \end{cases}$$

and

(2.6) $\quad {}^*d_N^i(\omega) = 1$, a.e. $\omega \in \Omega$.

Then $*d$ is an equilibrium SS under the monotone SR π and

(2.7) $\quad E\left[Y_{t_\pi(*d)}\right] = V_N$

holds. That is, v_N^i is the equilibrium expected net gain for player i.

Proof. Define

$$t_n^* = t_n(*d) = \text{first}\{m \geq n \text{ such that } \pi(*d_m) = 1\}$$

for $n=1,..,N$. Clearly $n \leqq t_n^* \leqq N$ and $t_1^* = t(*d)$. Where $t(*d) = t_\pi(*d)$ and π is fixed. We will show that

(2.8) $\quad E\left[Y^i_{t_n^*}\right] = v^i_{N-n+1} - (n-1)c^i$, $i=1,..,p$

by backward induction on n.

From $t_N^* = N$ and (2.2), it is trivial for n=N. Assume that it is true for n+1. From the definition of SE $*D_n = \prod(*D_n^1,..,*D_n^p) \in \mathcal{B}(X_n)$,

$$\begin{aligned} t_n^* &= n && \text{on } *D_n, \\ &= t_{n+1}^* && \text{on } \overline{*D_n}. \end{aligned}$$

Hence

$$E[Y^i_{t_n^*}] = E[Y_n^i;*D_n] + E[Y^i_{t^*_{n+1}};\overline{*D_n}]$$

where $E[X;D] = \int X\cdot I_D dP$. Since X_{n+1}, X_{n+2},.. are independent of X_n, $E[Y^i_{t^*_{n+1}};\overline{*D_n}] = P(\overline{*D_n})E[Y^i_{t^*_{n+1}}]$. Therefore we have the iteration:

$$E[Y^i_{t_n^*}] = E[Y_n^i;*D_n] + P(\overline{*D_n})E[Y^i_{t^*_{n+1}}].$$

It equals, by induction,

$$E[X_n^i - nc^i;*D_n] + P(\overline{*D_n})(v^i_{N-n} - nc^i) = E[X_n^i - v^i_{N-n};*D_n] + (v^i_{N-n} - c^i) - (n-1)c^i.$$

The first term of the right hand side in the above equation is rewritten as

$$E[(X_n^i - v^i_{N-n})^+;\ \Pi(*D_n^1,\dots,\Omega,\dots,*D_n^p)] - E[(X_n^i - v^i_{N-n})^-;\ \Pi(*D_n^1,\dots,\emptyset,\dots,*D_n^p)]$$
$$= E[(X_n^i - v^i_{N-n})^+ \beta^{\pi\{i\}}_{N-n}(V^{\{i\}}_{N-n}|X_n^i)] - E[(X_n^i - v^i_{N-n})^- \alpha^{\pi\{i\}}_{N-n}(V^{\{i\}}_{N-n}|X_n^i)]$$

So, from (2.1),

$$v^i_{N-n+1} = E[X^i_n - v^i_{N-n};{}^*D_n] + v^i_{N-n} - c^i.$$

This implies (2.8) and we have proved the latter part of the theorem by letting n=1 in (2.8).

Next we must show that, for fixed i,

$$(2.9) \qquad E[Y^i_{t(*d(i))}] \leq E[Y^i_{t(*d)}]$$

where $*d(i)=(*d^1,\ldots,d^i,\ldots,*d^p)$ and $d^i=(d^i_1,\ldots,d^i_N)$ is any ISS for player i. Define ${}^nd^i$,n=0,1,..,N by

$$\begin{aligned} {}^nd^i &= (d^i_1,\ldots,d^i_n,*d^i_{n+1},\ldots,*d^i_N) \text{ if } n=1,\ldots,N \\ &= *d^i \qquad \text{if } n=0 \end{aligned}$$

using d^i and $*d^i$. This ISS for player i is consistent with $*d^i$ after n-th period. Also define a strategy ${}^nd(i)$ by

$${}^nd(i) = (*d^1,\ldots,{}^nd^i,\ldots,*d^p).$$

Clearly ${}^Nd(i)= *d(i)$ and ${}^0d(i)= *d$.

We show

$$(2.10) \qquad E[Y^i_{t({}^nd(i))}] \leq E[Y^i_{t({}^{n-1}d(i))}]$$

for n=1,..,N because (2.9) is proved immediately from (2.10). By the strategy ${}^nd(i)$, it is enough to consider a stopping time t_n instead of t. It is seen that

$$E[Y^i_{t_n}({}^nd(i))] = E[Y^i_n;D_n] + P(\overline{D}_n)E[Y^i_{t_{n+1}}({}^nd(i))]$$

where D_n is an SE with respect to $*d^1,\ldots,{}^nd^i,\ldots,*d^p$. Since $t_{n+1}({}^nd(i)) = t_{n+1}(*d)$ on $\overline{D}_n$ and $E\left[Y^i_{t_{n+1}}(*d)\right]= v^i_{N-n} -nc^i$, it becomes

$$E[X^i_n-c^i;D_n] + P(\overline{D}_n)(v^i_{N-n}-c^i) - (n-1)c^i \leq v^i_{N-n+1} -(n-1)c^i = E[Y^i_{t_n}({}^{n-1}d(i))].$$

Q.E.D.

This is an extension of Theorem 3.1 in our previous work, Kurano, Yasuda Nakagami(1980). In the result, the player i's region for declaring to stop has the form of $X_n^i \geqq \{$a certain value$\}$. It is intuitively natural and this rule is called a critical level strategy. In the proof of the theorem we can see the following corollary.

COROLLARY 2.1. A necessary condition for

$$\{{*d}_n^i = 1\} = \{X_n^i \geqq \text{a certain value}\},\ n \geqq 1$$

is that an SR π satisfies $\pi({*d}_n^1,\ldots,0,\ldots,{*d}_n^p) \leqq \pi({*d}_n^1,\ldots,1,\ldots,{*d}_n^p)$, $n \geqq 1$, for the equilibrium SS $*d$.

If we impose further assumptions, then next two corollaries are obtained immediately.

COROLLARY 2.2. For each n, if components of $(X_n^1,\ldots,X_n^p)$ are mutually independent and identically distributed with X_p^0, then (2.1) implies

$$\text{(2.11)} \qquad v_{n+1}^i = v_n^i - c^i + \beta_n^{\pi\{i\}} E(X_{N-n}^0 - v_n^i)^+ - \alpha_n^{\pi\{i\}} E(X_{N-n}^0 - v_n^i)^-$$

where

$$\beta_n^{\pi\{i\}} = \beta_n^{\pi\{i\}}(v_n^{\{i\}}) = P(\pi({*D}_{N-n}^1,\ldots,\Omega,\ldots,{*D}_{N-n}^p))$$

and

$$\alpha_n^{\pi\{i\}} = \alpha_n^{\pi\{i\}}(v_n^{\{i\}}) = P(\pi({*D}_{N-n}^1,\ldots,\phi,\ldots,{*D}_{N-n}^p)).$$

COROLLARY 2.3. In additon, if the stopping rule π is symmetric for i and j, that is,

$$\text{(2.12)} \qquad \pi(\ldots,d^i,\ldots,d^j,\ldots) = \pi(\ldots,d^j,\ldots,d^i,\ldots)$$

and if $c^i = c^j$, then $v_n^i = v_n^j$ for each n. If π is symmetric for any pairs, this leads to the majority case discussed in Kuramo, Yasuda and Nakagami (1980).

EXAMPLE 2.1. Similar to Example 4.2 in Kurano, Yasuda and Nakagami (1980), we consider a variant of the the secretary problem(cf. Chow, Robbins and Siegmund(1971), Gilbert and Mosteller(1966)) with a monotone rule. Three players want to choose one secretary and we impose the following unequal SR:

$$\text{(2.13)} \qquad \pi(x^1,x^2,x^3) = x^1 + x^2x^3, \quad x^i \in \{0,1\},\ i=1,2,3.$$

This means that a secretary is accepted only when either player 1 says "yes", or both of player 2 and player 3 say "yes".

From Thm.2.1, the equilibrium SS *d is determined by the sequence of $\{v_n^i ; n=1,2,..\}$ in (2.11) where $c^i=0$ and $v_1^i=1/N$. Since the SR π of (2.13) is symmetric for players 2 and 3, $v_n^2=v_n^3$ from Cor.2.3. Define

$$r^1 = \inf\{r \; ; \; v_{N-r}^1 \leqq r/N\}, \quad r^2 = \inf\{r \; ; \; v_{N-r}^2 \leqq r/N\}.$$

The strategy for player 1 is that he observes until the (r^1-1)th stage and then declares to accept if the relative best one appears. For players 2 and 3, the strategy is similar. Numerical results are as follows.

N	r^1	v_N^1	r^2	v_N^2
10	3	.3642	1	.1685
30	10	.3649	2	.0801
100	36	.3673	3	.0322
300	110	.3677	4	.0135
1000	367	.3678	5	.0050
10000	3678	.3679	6	.0007

We have applied our result to a secretary problem with an unequal SR and showed the equilibrium SS is a critical level strategy. But, as a remark, the asymptotic numerical results for $N=\infty$ is non-interesting. Under the SR (2.13), player 1 behaves as if it were a one-person-game and player 2, 3 are neglected. A modified setting of the secretary problem is discussed by Presman and Sonin(1975) and Sakaguchi(1980).

2.3. AN INFINITE HORIZON CASE

In this section we treat an infinite horizon case $N = \infty$. The SS's class is therefore $\{d \in \mathfrak{D}; P(t_\pi(d) \leq \infty) = 1\}$. The problem is worth studying when the observation cost is non-negative. Theorem 3.1 discusses the case of c^i 0 for all i, in which case the stopping time is finite. In case of $c^i = 0$, $i=1,..,p$, some trouble occures in the multi-variate problem. Though we have defined $Y^i_\infty = \limsup_{n\to\infty} Y^i_n$ in (1.12) on the analogy of one dimentional problem, apparently this definition is not natural for all players under some SR's. To avoid this, we assume that the equilibrium stopping time is finite. Then we can establish the continuity from the finite horizon case and compare the expected gains between rules and between players. From the formulation of our model, this assumption is often satisfied because the process is forced to stop by the conflict among players.

ASSUMPTION 3.1.

(a) Random vectors X_1, X_2,.., $X=(X^1,..,X^p)$ are independent and identically distributed with $E|X^i|<\infty$ for all i.

(b) Each element of a cost vector $C=(c^1,..,c^p)$ is strictly positive.

(c) The SR π is monotone and let Π be the corresponding SER.

(d) The following simultaneous equation of $V=(v^1,..,v^p)$:

$$E[(X^i-v^i)^+\beta^{\Pi\{i\}}(V^{\{i\}}|X^i)] - E[(X^i-v^i)^-\alpha^{\Pi\{i\}}(V^{\{i\}}|X^i)] = c^i \tag{3.1}$$

$i=1,..,p$ has a solution. Where $V^{\{i\}}=(v^1,..,v^{i-1},v^{i+1},..,v^p)\in R^{p-1}$,

$$\beta^{\Pi\{i\}}(V^{\{i\}}|X^i)=P(\Pi(D^1,...,\Omega,...,D^p)|X^i),$$
$$\alpha^{\Pi\{i\}}(V^{\{i\}}|X^i)=P(\Pi(D^1,...,\phi,...,D^p)|X^i)$$

and $D^i=\{X^i\geq v^i\}$, $i=1,2,..,p$.

THEOREM 3.1. Under Assump.3.1, an SS $*d=(*d^1,..,*d^p)$ determined by

$$*d^i_n(\omega) = 1\ (0) \quad \text{if} \quad X^i_n(\omega) \geq (<) \ *v^i, \tag{3.2}$$

for each n and i, is the equilibrium SS for the class $\{d\in\mathfrak{D}; P(t_\pi(d)\leq\infty)=1\}$

and

(3.3) $\quad P(t_\pi(*d) < \infty) = 1,$

(3.4) $\quad E[Y^i(t_\pi(*d))] = *v^i \quad , \; i=1,\ldots,p$

hold where $*V = (*v^1,\ldots,*v^p)$ is a solution of (3.1).

By (3.4), $*V$ is called an equilibrium expected net gain. The proof is similar to that of Thm.5.3, 5.4 of Kurano, Yasuda and Nakagami(1980). So we omit it here.

In the rest of the section we restrict our attension to the case of

(b') $\quad C = 0.$

Under the assumption (b'), it may happen that the equilibrium stopping time is not finite. But if we assume the next (e), the following corollaries hold.

(e) $\quad P(t_\pi(*d) < \infty) = 1$ where $*d$ is defined by (3.2).

It is seen in Example 3.2 that there are cases which satisfy (e).

COROLLARY 3.1. Assume (a), (b'), (c), (d) and (e). If X is bounded with prob.1, then $*d$ is an equilibrium SS for the restricted class $\{d \in \mathcal{D};\ P(t_\pi(d) < \infty) = 1\}$ and (3.4) holds.

The proof is immediate by Thm.5.3, 5.4 of Kurano, Yassuda and Nakagami (1980). Hereafter we assume that

(a') (a) and components of $(X^1,\ldots,X^p)$ are independent.

COROLLARY 3.2. Under assumptions (a'), (b'), (c), (d) and (e), if $P(X^i = y) = 0$ where $y = \sup\{y; P(X^i > y) > 0\}$, then $*d$ is an equilibrium SS for the class $\{d \in \mathcal{D};\ P(t_\pi(d) \leq \infty) = 1\}$ and (3.4) holds.

Proof. By Assumption(e), $P(\pi(*D^1,\ldots,*D^p)) > 0$ where $*D^i = \{X^i \geq *v^i\}$. If we assume that $P(\pi(*D^1,\ldots,\phi,\ldots,*D^p)) = 0$, then $P(\pi(*D^1,\ldots,\Omega,\ldots,*D^p)) > 0$ from the monotonicity of the rule. From (3.1), (a') and $P(\pi(*D^1,\ldots,\Omega,\ldots,*D^p)) > 0$ and $P(\pi(*D^1,\ldots,\phi,\ldots,*D^p)) = 0$ implies $(X^i - *v^i)^+ = 0$ a.e., that is, $*v^i \geq y$. This means $*D^i = \phi$, a.e.. by the assumption. We have

$$P(\prod(*D^1,\ldots,*D^i,\ldots,*D^p))=P(\prod(*D^1,\ldots,\phi,\ldots,*D^p)).$$

This is a contradiction because the left hand side is ">0" but the right one equals zero. Hence we obtain $P(\prod(*D^1,\ldots,\phi,\ldots,*D^p))>0$. For the SS $*d^{\{i\}}=(*d^1,\ldots,d^i,\ldots,*d^p)$ where d^i is any ISS, it is seen that $P(t_\pi(*d^{\{i\}})<\infty)$ $=1$. Hence the proof is immediately completed from Thm.3.1.

Q.E.D.

In the case of SR π with $P(t_\pi(*d)=\infty)=1$, there is a player i such that

(3.5) $\quad *v^i = \sup\{\, y; P(X^i>y)>0\,\}.$

Clearly (3.4) is satisfied for player i by (1.12). But for other player $j(\neq i)$, $*v^j$ does not necessarily satisfy (3.4). Therefore the solution of (3.1) does not always consist with the equilibrium expected gain in this case. In order to discuss the associated gain including this case, we simply call an expected gain (omitting "equilibrium") by the solution $*V$ which is the limiting value as $N\to\infty$ in the finite horizon case. Refer to Figure 4.1 in Kurano, Yasuda and Nakagami(1980) and Table 3.1.

Now we shall give the bound of the expected gain by varying SR. The expected gain $v^i=v^i(\pi)$ associated with an SR satisfies that

(3.6) $\quad EX^i \leq v^i \leq \sup\{\, y; P(X^i>y)>0\,\}$

for any monotone SR π. In fact, this is proved by using a ratio (3.8) as follows. By (a') and (b'), the equation (3.1) implies

(3.7) $\quad E[(X^i-v^i)^+]/E[(X^i-v^i)^-] = \rho_\Pi^{\{i\}}(v^{\{i\}})$

where

(3.8)
$$\rho_\Pi^{\{i\}}(v^{\{i\}}) = \alpha^{\Pi\{i\}}(v^{\{i\}})/\beta^{\Pi\{i\}}(v^{\{i\}})$$
$$= P(\Pi(D^1,\ldots,\phi,\ldots,D^p))/P(\Pi(D^1,\ldots,\Omega,\ldots,D^p))$$

provided the denominator is non-zero. Since the SR π is monotone,

(3.9) $\quad 0 \leq \rho_\Pi^{\{i\}}(v^{\{i\}}) \leq 1$

holds. Therefore (3.9) implies (3.6) immediately.

From the above argument, $\rho_{\Pi}^{\{i\}}(V^{\{i\}})=1$ implies $v^i=EX^i$, and $\rho_{\Pi}^{\{i\}}(V^{\{i\}})=0$ implies $v^i=\sup\{y; P(X^i>y)>0\}$. The second assertion corresponds to $P(t_{\pi}(*d)=\infty)=1$ as remarked at (3.5). Here these two extreme cases are interpreted as follows.

Firstly $\rho_{\Pi}^{\{i\}}(V^{\{i\}})=0$ is equivalent to $\Pi(D^1,..,\phi,..,D^p)=\phi$ a.e. and also to $\pi(d^1,..,0,..,d^p)=0$ with prob.1. This means that whenever player i declares to continue, the decision process surely continues. But it does not mean that declaring to stop causes to stop the process. Player i is endowed the veto power. This brings him the maximum expected gain. Secondly $\rho_{\Pi}^{\{i\}}(V^{\{i\}})=1$ is equivalent to $\Pi(D^1,..,\phi,..,D^p)=\Pi(D^1,..,\Omega,..,D^p)$ a.e. and also to $\pi(d^1,..,0,..,d^p)=\pi(d^1,..,1,..,d^p)$ with prob.1. For player i, declaring to stop or to continue does not affect to the resulting process. He is ranked as the outsider of the game and his expected gain EX^i is the least one.

Now we shall make a comparison of gains between players under a fixed rule in Cor.3.3 and also between two different rules in Cor.3.4. The next theorem is immediately proved from (3.7).

THEOREM 3.2. Let $V_{\Pi}=(v_{\Pi}^1,..,v_{\Pi}^p)$ and $V_{\tilde{\Pi}}=(v_{\tilde{\Pi}}^1,..,v_{\tilde{\Pi}}^p)$ be expected gains corresponding to SER's Π and $\tilde{\Pi}$ respectively. For player i, j, assume $\rho_{\Pi}^{\{i\}}$ and $\rho_{\tilde{\Pi}}^{\{j\}}$ are defined by (3.8). If X^i and X^j are identically distributed, we have

$$(3.10)\qquad v_{\Pi}^i \begin{Bmatrix} > \\ = \\ < \end{Bmatrix} v_{\tilde{\Pi}}^j$$

if and only if

$$(3.11)\qquad \rho_{\Pi}^{\{i\}}(V_{\Pi}^{\{i\}}) \begin{Bmatrix} < \\ = \\ > \end{Bmatrix} \rho_{\tilde{\Pi}}^{\{j\}}(V_{\tilde{\Pi}}^{\{j\}}).$$

COROLLARY 3.3. Under a fixed SER Π, if X^i and X^j are identically distributed and if

$$\rho_{\Pi}^{\{i\}}(V_{\Pi}^{\{i\}}) \leqq \rho_{\Pi}^{\{j\}}(V_{\Pi}^{\{j\}}) \tag{3.12}$$

then $v_{\Pi}^{i} \geq v_{\Pi}^{j}$.

<u>COROLLARY 3.4.</u> If, for player i,

$$\Pi(D^1,\ldots,\Omega,\ldots,D^p) \supset \tilde{\Pi}(D^1,\ldots,\Omega,\ldots,D^p) \tag{3.13}$$

and $\qquad \Pi(D^1,\ldots,\phi,\ldots,D^p) \subset \tilde{\Pi}(D^1,\ldots,\phi,\ldots,D^p)$

for every $D^k \in \mathcal{B}(X)$, $k\neq i$, or

$$\rho_{\Pi}^{\{i\}}(U^{\{i\}}) \leqq \rho_{\tilde{\Pi}}^{\{i\}}(U^{\{i\}}) \tag{3.14}$$

for every $U^{\{i\}}=(u^1,\ldots,u^{i-1},u^{i+1},\ldots,u^p)$ such that $EX^k \leqq u^k \leqq \sup\{y;P(X^k > y) > 0\}$, $k\neq i$, then $v_{\Pi}^{i} \geqq v_{\tilde{\Pi}}^{i}$ holds.

2.4. EXAMPLES

<u>EXAMPLE 3.1.</u> Consider a majority rule $\Pi[r]=(p,r)$ of p players, where r ($1\leqq r\leqq p$) is a majority level. Let X^i, i=1,..,p be independent, identically distributed with X. If $EX<\sup\{y;P(X>y)>0\}$, then the equilibrium expected gains for each rule are

$$v_{\Pi[p]} > v_{\Pi[p-1]} > \cdots > v_{\Pi[1]} . \tag{3.15}$$

In fact, since the SR is symmetric, we can set the players' gains being equal:

$$v_{\Pi[r]} = v_{\Pi[r]}^{i} , \quad i=1,2,\ldots,p.$$

Hence

$$\rho_{\Pi[r]}^{\{i\}}(V^{\{i\}}) = \rho_{\Pi[r]}(v) = 1 - \eta(r,\bar{v}),$$

where

$$\eta(r,\bar{v}) = \binom{p-1}{r-1}(1-\bar{v})^{r-1}\,\bar{v}^{p-r} \Big/ \sum_{k=r-1}^{p-1} \binom{p-1}{k}(1-\bar{v})^{k}\,\bar{v}^{p-k-1}$$

and $\bar{v}=P(X \leqq v)$. Since $\eta(r,\bar{v})$ is increasing in $\bar{v}$ and $\eta(r,\bar{v}) < \eta(r+1,\bar{v})$, we can see $\rho_{\pi[r]}(v)$ is decreasing in v and $\rho_{\pi[r]}(v) > \rho_{\pi[r+1]}(v)$ for each v. Similarly as Cor.3.4, it implies (3.15).

Figure 4.1, in Kurano, Yasuda and Nakagami, shows each expected gain of (3.15) for p=5 players. For r=1,..,p-1, $\pi[r]$ is an equilibrium SR and $v_{\pi[r]}$ is an equilibrium expected gain from Cor.3.2. But for r=p , each player has the veto power and so $v_{\pi[p]} = \sup\{y; P(X > y)>0\}$. Thougth its stopping time is $P(t_{\pi[p]}=\infty)=1$, the associated expected gain is equilibrium directly from (1.12), (1.13) and (3.4).

Example 3.2. Let components of random processes be independent, identically distributed with a common uniform distribution U(0,1). Table 3.1 shows a numerical example with p=3 for non-trivial monotone SR's. In the first four rules $P(t_{\pi}(*d) < \infty)=1$, but in other cases not so. From (3.5), there exist players who attain its maximum expected gain unity in the last four rules. Each expected gain is the limiting value of the finite horizon case. Except 5-th,6-th,7-th rule, the value is an equilibrium one by Cor.3.2.

Table 3.1 Monotone SR's with p=3.

Monotone SR $\pi(x^1,x^2,x^3)$		$x^1+x^2+x^3$	x^1+x^2	$x^1+x^2x^3$	$x^1x^2+x^2x^3+x^3x^1$	x^1	$x^1x^2+x^1x^3$	x^1x^2	$x^1x^2x^3$
Comments for the rule		majority rule for (p,r)=(3,1)	pl.3 is an outsider	asymmetric case	majority rule for (p,r)=(3,2)	pl.1 is a dictator	pl.1 has veto power	pl.3 is an outsider	unanimity (p,r)=(3,3)
(Equi-librium) expected gain v^i	v^1	0.5437	$(\sqrt{5}-1)/2$ ≒0.6180	$\sqrt{2}/2$ ≒0.7071	$\sqrt{2}/2$	1	1	1	1
	v^2	0.5437	$(\sqrt{5}-1)/2$	$2-\sqrt{2}$ ≒0.5858	$\sqrt{2}/2$	0.5	$(\sqrt{5}-1)/2$	1	1
	v^3	0.5437	0.5	$2-\sqrt{2}$	$\sqrt{2}/2$	0.5	$(\sqrt{5}-1)/2$	0.5	1

3. ASYMPTOTIC RESULTS FOR THE BEST CHOICE PROBLEM

3.1. STATEMENT OF THE PROBLEM

An optimal stopping problem is related to a Markov decision process with two actions: stop and continue. The equation for $v(i)$, the expected reward under an optimal policy when starting from state i, is given by

$$v(i) = \max\left\{r(i),\ -c(i) + \sum_j p(i,j)v(j)\right\}, \quad i \in \{1,2,..\} \tag{1.1}$$

where $r(i)$ is an immediate reward, $c(i)$ is a paying cost and $p(i,j)$ is a transition probability on the state space, $\{1,2,...\}$. The best choice problem, variously called the secretary problem, Googol, Dowry problem in Chow et al.(1964), in Gilbert and Mosteller(1966) and else, is an optimal stopping problem based on relative ranks for objects arriving in a random fashion; the objective is to find the stopping rule that maximizes the probability of attaining the best object of the sequence.

To consider the problem as a Markov decision process, suppose that the model is in state i iff the i^{th} object to be examined is better than all its predecessors(the relatively best object) and the two actions are to accept this object, or reject it and wait for the successors. The immediate reward $r(i)$ is a probability that the object accepted in state i is the absolutely best one. And the transition probability $p(i,j)$ is a conditional probability that the next relatively best object to appear will be the j^{th} object in the sequence, given that the i^{th} object in the sequence was relatively best.

The Markov chain formulation is considered, for example, by Dynkin and Yushevich(1969) and so its details are omitted. The practical situation for the well-known problem of one choice among n objects then becomes: The state space is a set of integers $\{1,2,..,n\}$, the reward $r(i) = i/n$ and the transition probability $p(i,j) = i/((j-1)j)$ for $1 \leq i < j \leq n$, $p(i,j) = 0$, otherwise. Hence (1.1) implies

$$(1.2) \quad v(i) = \max\{i/n,\ i\textstyle\sum_{j=i+1}^{n} v(j)/((j-1)j)\},\ i=1,2,..,n-1, \quad v(n) = 1.$$

By solving this equation, one obtains the optimal value, i.e., the maximal probability of attaining the best object, and the optimal strategy, i.e., how to accept or reject an object.

Although the solution can be obtained easily in this case, let us consider the following alternative method. We investigate the conditional optimal value when the decision-maker rejects all objects until and including the i^{th} relatively best, instead of the optimal value. Denote by $w(i;n)$ the second term on the right hand side of (1.2). Since this term corresponds to the rejection and $v(i)$ is the optimal value, $w(i;n)$ will be the conditional optimal value. That is, let $w(i;n) = w(i) = i\sum_{j=i+1}^{n} v(j)/((j-1)j)$, $i=1,2,..,n-1$ and $w(n) = 0$. Then clearly $w(i)-w(i+1) = (v(i+1)-w(i+1))/(i+1)$ and so

$$(1.3) \quad w(i) - w(i+1) = ((i+1)/n - w(i+1))^{+}/(i+1), \quad i=1,2,..,n-1$$

where $a^{+} = \max(a,0)$.

Following Mucci(1973) and Lorenzen(1981), we consider a scaling limit of (1.3), $f(x)= \lim_{n\to\infty} w(i;n)$ as i and n tend to infinity subject to $i/n=x$. This leads to the differential equation:

$$(1.4) \quad df(x)/dx = -x^{-1}(x-f(x))^{+}, \quad 0 < x < 1$$

with boundary condition $f(1)=0$. Immediately we obtain $f(x) = -x\log(x)$ on $e^{-1} \leq x \leq 1$, $f(x) = e^{-1}$ on $0 < x \leq e^{-1}$. From this solution, we can determine the optimal value and the stopping island named after Pressman and Sonin(1972). A relatively best object is accepted iff the time of occurrence of this objects belong to the stopping set. If $k,k+1,...,m$ belong to this set, then the interval $[k,m]$ is a stopping island. The optimal value equals $v^* = \lim v(1;n) = \lim \max\{1/n,\ w(1;n)\} = f(0+) = e^{-1}$

and the stopping island is the interval $[\alpha^*, 1]$ where $\alpha^* = \inf\{x;\ x \geq f(x)\}$ $= e^{-1}$.

The aim of this chapter is to apply this method to the best choice problem with a random number of objects, and obtain some explicit solutions in the asymptotic form. Instead of the differential equation, an integral equation is considered so as to treat the case with a general distribution of the number of objects. But here we assume that the total number of objects is a bounded random variable with known distribution. Presman and Sonin(1972) considered this problem by an approximation method of the parameter associated with its distribution, rather than by using the scaling limit. For another problem of minimizing the expected rank of the individual selected, Gianini(1979) has used a differential equation method.

In section 3.2 an integral equation with a general distribution of the number of objects is derived by adapting the above method. However, if the distribution is absolutely continuous, it reduces to a differential equation, the simplest one being (1.4). To find an optimal strategy, we determine the stopping island. A certain condition implies that the stopping set is a single island of which the lower bound can be found, and of which the upper bound is 1. This condition is fundamental to our discussion and contributes to obtaining a solution of the integral equation exactly. As an extension of the uniformly distributed case, we obtain an intermediate result between the non-random case and the Rasmussen and Robbins(1975) problem. Another intermediate case of a distribution, which is not absolutely continuous, is also considered. The next three sections are devoted to discussing three different variants of the best choice problem.

In section 3.3 the result of Smith(1979) involving a refusal probability is extended to that of a uniformly distributed number of objects with non-

non- constant refusal. For the variation of the multiple choice permitting r offers, Gilbert and Mosteller(1966) had formulated and Tamaki(1979a) had obtained the result for r=2 in the uniform case. In section 3.4, we give a further result of the optimal value of r in an iterative form for the same situation. For the multiple choice problem, the aim is to select the best and the second best objects, a problem solved by Nikolaev(1977) and Sakaguchi (1979). We consider this problem with a random number of objects and calculate results for the uniformly distributed case in section 3.5.

In the rest of this section we set out notations and preliminaries. For integration with respect to the probability measure $d\Phi$ on the unit interval $[0,1]$: $V(A) = \int_A v(x)d\Phi(x)$ for all intervals A in $[0,1]$, we shall use the abbreviation:

(1.5) $$dV(x) = v(x)d\Phi(x).$$

For any bounded function u(x) the relation (1.5) obviously implies $u(x)dV(x) = u(x)v(x)d\Phi(x)$ (p.137, Feller(1966)). Using this short hand notation, an integral equation of the form:

$$f(y) - f(x) = \int_x^y a(t,f(t))d\Phi(t) + \int_x^y b(t,f(t))dt$$

for all $0 < x < y < 1$ is equivalent to

(1.6) $$df(x) = a(x,f(x))d\Phi(x) + b(x,f(x))dx, \quad 0 < x < 1.$$

Let f(x) and g(x) be two functions of bounded variation over $[0,1]$, right continuous and with left-hand limits, then, by Fubini's Theorem,

(1.7) $$d(fg)(x) = f(x)dg(x) + g(x)df(x) - \{f(x)-f(x-)\}dg(x)$$

holds(p.336, Brèmaud(1981)). If f(x) is continuos in $0 < x < 1$, then

$$d(fg)(x) = f(x)dg(x) + g(x)df(x)$$

follows immediately.

3.2. A SCALING LIMIT OF THE OPTIMALITY EQUATION

The probability model for the best choice problem with a random number of objects has been considered by Presman and Sonin(1972). We therefore omit details of its construction here. To take a scaling limit, we restrict ourselves to the case where the number of objects is bounded.

ASSUMPTION(I). A random number of objects N is bounded with a probability one, that is, there is a positive integer n such that

(2.1) $n = \inf\{k \geq 1 \ ; \ P(N > k) = 0\}$.

The state space is a set of integers $\{1, 2, .., n\}$. State i in the model means that the i^{th} object appearing is the relatively best one (better that all its predecessors). The meanings of the transition probability and reward are similar to those for the deterministic case introduced in the previous section, with some learning procedures included. Let us denote $p_i = P(N = i)$ and $\pi_i = \sum_{k=i}^{n} p_k$. The transition probability matrix $P = (p(i,j);\ 1 \leq i,\ j \leq n)$ is defined by

$$(2.2) \qquad \begin{aligned} p(i,j) &= i\pi_j/(j(j-1)\pi_i), \quad 1 \leq i < j \leq n, \\ p(i,n) &= \sum_{k=i+1}^{n} ip_k/(k\pi_i), \quad 1 \leq i < n \quad \text{and} \quad p(n,n) = 1. \end{aligned}$$

The expected reward r(i) is

(2.3) $r(i) = r(i;n) = \sum_{k=i}^{n} ip_k/(k\pi_i)$,

and the cost is c(i)=0 for each i. From the general equation(1.1), the optimal value v(i) = v(i;n) satisfies an optimality equation:

(2.4) $v(i) = \max\{r(i),\ Pv(i)\}$, $i=1,2,..,n-1$, $v(n) = 1$

where P, r are defined as in (2.2), (2.3) respectively.

ASSUMPTION(II). There is a probability measure $d\Phi$ on $[0,1]$ such that for any sequence $s(k;n)$, $k=1,2,..,n$ with $\lim s(k;n) = s(x)$ for $k/n = x$

(2.5) $\lim\limits_{i,j,n} \sum_{k=j+1}^{n} s(k;n)p_k = \int_x^y s(t)d\Phi(t) = \int_{(x,y]} s(t)d\Phi(t)$

where $i/n = x$, $j/n = y$ for $x,\ y \in [0,\ 1]$. Further we assume that $d\Phi$

satisfies the conditions

(2.5i) $\quad (1-\Phi(x))^{-1}\int_x^1 y^{-1}d\Phi(y) \rightarrow 1 \quad \text{as } x \rightarrow 1,$

(2.5ii) $\quad x\int_x^1 y^{-1}d\Phi(y) \rightarrow 0 \text{ as } x \rightarrow 0.$

Hereafter Assumptions (I) and (II) will always hold. But, in section 3.5, (2.5i) and (2.5ii) are slightly strengthened to discuss multiple choice problems.

Let us define

$$
\text{(2.6)} \qquad \begin{aligned} w(k;n) &= w(k) = Pv(k) = \sum_{j=k+1}^{n} k\pi_j v(j)/(j(j-1)\pi_k), \quad k=1,..,n-1, \\ w(n;n) &= w(n) = 0. \end{aligned}
$$

As in the previous section, this corresponds to the conditional optimal value when the decision-maker rejects all objects until and including the i^{th} relatively best. Since

$$w(k) = \{v(k+1)/(k+1) + w(k+1)k/(k+1)\}\pi_{k+1}/\pi_k$$

holds, (2.4) implies that

$$\text{(2.7)} \qquad w(k+1) - w(k) = w(k+1)\,\pi_k^{-1}p_k - (k+1)^{-1}\{r(k+1)-w(k+1)\}^{+}\pi_{k+1}/\pi_k.$$

<u>PROPOSITION 2.1.</u> A scaling limit of the sequence, $f(x) = \lim_{k,n\to\infty} w(k;n)$ for $k/n=x$ exists. Using the abbreviation (1.6), $f(x)$ satisfies the equation

$$
\text{(2.8)} \qquad \begin{aligned} df(x) &= f(x)(1-\Phi(x))^{-1}d\Phi(x) - x^{-1}(R(x)-f(x))^{+}dx, \quad 0 < x < 1, \\ f(1) &= 0, \end{aligned}
$$

where $R(x) = x(1-\Phi(x))^{-1}\int_x^1 y^{-1}d\Phi(y)$, $0 < x < 1$, is well defined by (2.5i) and (2.5ii).

<u>Proof.</u> The standard Picard iteration method implies the existence of the equation and the scaling limit. As k and n tend to infinity, provided k/n =x, we see that $\pi_k^{-1}p_k \rightarrow (1-\Phi(x))^{-1}d\Phi(x)$, $r(k+1) \rightarrow R(x)$, $\pi_{k+1}/\pi_k \rightarrow 1$ and $(k+1)^{-1}=n(k+1)^{-1}(1/n) \rightarrow x^{-1}dx$. Thus (2.8) is immediately obtained by taking the sum of (2.7).

THEOREM 2.2. The optimal value v^* of the problem in the asymptotic form is given by $v^* = f(0+)$.

Proof. Since $w(1;n) \geqq 0$, $v^* = \lim_{n\to\infty} v(1;n) = \lim_{n\to\infty} \max(r(1;n), w(1;n)) = \lim_{n\to\infty} w(1;n) = f(0+)$.

Now let $h(x) = \int_x^1 y^{-1} d\Phi(y)$ and

(2.9) $\quad H(x) = h(x) - \int_x^1 y^{-1}h(y)dy = (1+\log(x))h(x) + \int_x^1 \log(y)dh(y)$

for $0 \leqq x \leqq 1$.

CONDITION(Φ). $H(x) = H(x;\Phi)$ changes its sign once from $-$ to $+$ as x varies from 0 to 1.

Define

(2.10) $$\alpha^* = \begin{cases} \inf\{x;\ H(x) \geqq 0\} \\ 1 \text{ if empty.} \end{cases}$$

Then Condition(Φ) implies that $H(x) \geqq 0$ on $[\alpha^*, 1]$. This is important for our argument to obtain the solution exactly, and is closely related to the condition for an OLA policy in Markov decision processes. In the discrete parameter problem, a similar condition was imposed in Presman and Sonin(1972), Derman et al.(Unpublished) and Rasmussen and Robbins(1975).

PROPOSITION 2.3. If $\Phi(x)$ satisfies Condition(Φ), and if $\Phi(x)$ is continuous for $0 \leqq x < 1$, then the optimal value is given by

(2.11) $\quad v^* = (1-\Phi(\alpha^*))f(\alpha^*) = \alpha^* h(\alpha^*)$.

The stopping island $[\alpha^*, 1]$ is determined by the unique solution of the equation:

(2.12) $\quad H(x) = 0, \quad 0 < x < 1$.

Proof. By (1.6), (2.8) is equivalent to

$$(1-\Phi(x))df(x) = f(x)d\Phi(x) - x^{-1}(1-\Phi(x))(R(x)-f(x))^+dx,\ 0 < x < 1.$$

Since $\Phi(x)-\Phi(x-) = 0$ for every $0 < x < 1$, (1.7) implies that $(1-\Phi(x))f(x)$ is differentiable in $0 < x < 1$ and $g(x) = x^{-1}(1-\Phi(x))f(x)$ satisfies the equation

$$dg(x) = -x^{-1}\max\{h(x),\ g(x)\}\,dx,\ 0 < x < 1,$$

$$g(1) = 0.$$

Condition(Φ) implies that (2.12) has a unique solution and this differential equation is explicitly solved as

$$g(x) = \begin{cases} \int_x^1 y^{-1}h(y)dy & \text{on } \{H(x) \geqq 0\} = [\alpha^*,\ 1], \\ (\text{const})/x & \text{on } \{H(x) < 0\} = (0,\ \alpha^*). \end{cases}$$

Therefore, using Theorem 2.2, (2.11) is obtained immediately.

This proposition provides a solution of the problem with the random structure under Condition(Φ). From equation(2.12), the lower bound of the stopping island, or the threshold of the acceptance region for the relatively best object is determined; the optimal value is also calculated from this threshold in (2.11).

COROLLARY 2.4. If the measure $d\Phi(x)$ is absolutely continuous with respect to Lebesgue measure dx and $\phi(x)$ is its density function,

(2.13) $\quad d\Phi(x) = \phi(x)dx,$

then (2.8) is reduced to a differential equation:

$$(2.14)\quad \begin{aligned} & df(x)/dx = \phi(x)(1-\Phi(x))^{-1}f(x) - x^{-1}(R(x)-f(x))^{+},\ 0 < x < 1, \\ & f(1) = 0. \end{aligned}$$

Hence α^* is a solution of the equation:

(2.15) $\quad H(x) = \int_x^1 y^{-1}(1-\log(y)+\log(x))\phi(y)dy = 0.$

It is noted that $x \leqq R(x) \leqq 1$ for $0 \leqq x \leqq 1$. The case $R(x) = x$ for $0 < x < 1$, gives a model for the non random number of objects, that is, $p_k = 1$ for $k=n$, $p_k = 0$ otherwise. Since $\Phi(x) = 0$, $0 \leqq x < 1$, (2.14) becomes the differential equation (1.4), which is known to be the simplest case. The other equality, $R(x) = 1$ for $0 < x < 1$ and $R(0) = 0$, implies $f(x) = 0$ because no stopping occures. Generally, if $R_1(x) \leqq R_2(x)$, $0 < x < 1$ then the corresponding optimal value is $v_1^* \leq v_2^*$. Hence for the non-random case, $R(x) = x$, this gives the maximum value for the number of objects, when this has a distribution. The next two examples are intended to illustrate an intermediate result between the non-random and the uniformly distributed cases.

EXAMPLE 2.1. Let the number of objects be uniformly distributed on a partial interval $\{n-m, n-m+1, .., n\}$ of $\{1, .., n\}$ for some $m (0 \leq m \leq n)$. That is, $p_i = 1/(m+1)$ for $i=n-m, .., n$, and $p_i = 0$, otherwise. Let i, m, $n \to \infty$ with $\theta = m/n$ fixed. Taking the scaling limit (2.5) of Assumption(II), we have $\phi(x) = 1/\theta$ for $1-\theta \leq x \leq 1$, and $\phi(x) = 0$, otherwise, and it is seen that (2.5i) and (2.5ii) are satisfied. Instead of solving the differential equation (2.14), we obtain v^* and α^* directly from (2.11) and (2.15), because each distribution $\Phi(x) = \Phi(x;\theta)$; $0 < \theta < 1$ satisfies Condition(Φ). We conclude that

Table 1

case	stopping island	optimal value
$1-\theta \geq e^{-2}$	$[\sqrt{1-\theta}\, e^{-1}, 1]$	$-(\sqrt{1-\theta}/\theta)\log(1-\theta)e^{-1}$
$1-\theta \leq e^{-2}$	$[e^{-2}, 1]$	$2e^{-2}/\theta$

If $\theta \to 0$, the optimal value tends to e^{-1} (non-random case). If $\theta \to 1$, it tends to $2e^{-2}$ as discussed in Presman and Sonin(1972), and Rasmussen and Robbins(1975). Stewart(1981) treated the same distribution but his model was adapted in a Bayesian sense.

EXAMPLE 2.2. Now consider the limit distribution,

$$\Phi(\{1\}) = 1 - \theta \text{ and } d\Phi(x) = \theta dx \text{ for } 0 < x < 1 \text{ with some } 0 \leq \theta \leq 1.$$

There is a point mass of probability at the point 1. This is another intermediate example between the non-random case and the uniformly distributed case, which is not absolutely continuous. Since it satisfies Condition(Φ) and is continuous in $0 < x < 1$, we can apply Proposition 2.3. We see that

$$\alpha^* = \exp((1-2\theta-\sqrt{(1-2\theta+2\theta^2)})/\theta)$$

by solving equation (2.12). Hence the optimal value is

$$v^* = (\theta+\sqrt{(1-2\theta+2\theta^2)})\exp((1-2\theta-\sqrt{(1-2\theta+2\theta^2)})/\theta)$$

by (2.11). We observe that the optimal value is monotone decreasing as θ increases.

3.3. THE PROBLEM WITH A REFUSAL PROBABILITY

One of the variations in the best choice problem is a model which induces a refusal probability into the decision "acceptance". Smith (1975) calls the secretary problem with this change "uncertain employment". Sakaguchi(1979) generalized this model to the multiple choice problem, on which a random structure will also be imposed in section 3.5. The optimality equation for a finite(deterministic) number of objects n with a refusal probability p is

$$(3.1)\qquad v(i) = \max\left\{pi/n + (1-p)i\sum_{j=i+1}^{n} v(j)/(j(j-1)),\ i\sum_{j=i+1}^{n} v(j)/(j(j-1))\right\}$$

where p is a constant such that $0<p\leq 1$. Following the same procedure with the scaling limit, this leads to the differential equation:

$$(3.2)\qquad df(x)/dx = -px^{-1}(x-f(x))^{+},\quad 0<x<1,\quad f(1)=0.$$

Solving it, we obtain the optimal value $v^*_p = f(0+) = p^{1/(1-p)}$ and the stopping island $[p^{1/(1-p)}, 1]$, namely Smith(1975)'s result.

Now we consider a model with a random number of objects and inducing the non constant refusal probability $p(i) = p(i;n)$. We can describe the model by the optimality equation using the same notation as in section 3.2:

$$(3.3)\qquad \begin{aligned} v(i) &= \max\{p(i)r(i) + (1-p(i))Pv(i),\ Pv(i)\},\ i=1,..,n-1,\\ v(n) &= p(n). \end{aligned}$$

As in the previous section, we have the following theorem under the same assumptions.

Let $h(x) = \int_x^1 y^{-1}d\Phi(y)$ and

$$(3.4)\qquad H_p(x) = h(x) - q(x)\int_x^1 h(y)p(y)/(yq(y))dy$$

where $q(x) = \exp(\int_x^1 y^{-1}(1-p(y))dy)$ and $p(x)$ is a scaling limit of $p(i)= p(i;n)$ with $i/n=x$. From a realistic point view, the refusal probability should not depend on the order in which the objects are examined. In this case, (3.4) becomes as $H_p(x) = h(x) - px^{p-1}\int_x^1 y^{-p}h(y)dy$ where, as in example 3.1, the refusal probability is assumed to be constant.

CONDITION(Φ_p). $H_p(x)$ changes its sign once from - to + as x increases.

Define, similarly,

$$\alpha_p^* = \begin{cases} \inf\{x;\ H_p(x) \geq 0\} \\ 1 \text{ if empty.} \end{cases} \tag{3.5}$$

THEOREM 3.1. The integral equation of the problem is

$$\begin{aligned} &df(x) = f(x)(1-\Phi(x))^{-1}d\Phi(x) - x^{-1}p(x)(R(x)-f(x))^{+}dx,\ 0 < x < 1, \\ &f(1)=0. \end{aligned} \tag{3.6}$$

If $\Phi(x)$ is continuous for $0<x<1$ with Condition(Φ_p), then the optimal value v_p^* with a refusal probability $p(x)$, $0<x<1$ is given by

$$v_p^* = f(0+) = (1-\Phi(\alpha_p^*))f(\alpha_p^*) = \alpha_p^* q(\alpha_p^*)\int_{\alpha_p^*}^{1} h(y)p(y)/(yq(y))dy. \tag{3.7}$$

The stopping island $[\alpha_p^*, 1]$ is determined by the solution α_p^* of $H_p(x) = 0$.

EXAMPLE 3.1. We consider the case of $\Phi(x)=x$, $0<x<1$, where the number of objects is uniformly distributed on $\{1,2,..,n\}$ and $p(x) = p$ for $0<x<1$. Since $d\Phi(x) = dx$, (3.6) leads to a differential equation:

$$df(x)/dx = (1-x)^{-1}f(x) - px^{-1}(R(x)-f(x))^{+},\ 0 < x < 1,\ f(1) = 0$$

where $R(x)= -x(1-x)^{-1}\log(x)$. Since $h(x) = -\log(x)$ and $q(x) = x^{p-1}$, the equation $H_p(x) = 0$ becomes

$$p(x^{p-1}-1) + (1-p)\log(x)= 0.$$

Hence α_p^* is the unique solution of this transcendental equation in $0<x<1$. We see immediately that $\{x\ ;\ H_p(x) \geq 0\} = [\alpha_p^*, 1]$ holds, and hence $v_p^* = -\alpha_p^*\log(\alpha_p^*)$ by (3.7). Some numerical results are given in the Table 2. We note that p = 1.0 corresponds to the non-refusal case with a uniformly distributed number of objects discussed in section 3.2 (See Rasmussen and Robbins(1975)).

Table 2

refusal probability p	stopping island $[\alpha_p^*, 1]$	optimal value $-\alpha_p^* \log(\alpha_p^*)$
.5	$[.0810, 1]$	.2036
.7	$[.1052, 1]$	.2369
.9	$[.1260, 1]$	.2610
.99	$[.1344, 1]$	.2698
1.0	$[e^{-2}=.1353, 1]$	$2e^{-2}=.2707$

3.4. A MULTIPLE CHOICE PROBLEM (I)

Another variation in the best choice problem is the case where the decision is allowed to make r-object choices(i.e., r stops) and one wants to choose the best among these(See Gilbert and Mosteller(1966)). Sakaguchi (1978) has solved this by using the OLA policy and Tamaki(1979a) has discussed the case where the number of objects is a uniformly distributed random variable, and has obtained an explicit value in the asymptotic form for the case of $r=2$.

As in the previous sections, we derive an integral equation in the case of r-object choices with a random number of objects for the optimality equation. Following Presman and Sonin(1972) and Tamaki(1979a), the optimality equation becomes

$$(4.1)\qquad \begin{aligned} v_r(i) &= \max\{r(i) + Pv_{r-1}(i),\ Pv_r(i)\},\quad r=1,2,..,\\ v_0(i) &= 0. \end{aligned}$$

As in (2.4), let $w_r(k) = Pv_r(k)$ $k=1,2,..,n-1$ and $w_r(n)=0$ for each r. This denotes the conditional optimal value, as before. The same Assumptions (I) and (II) hold as in section 3.2.

<u>THEOREM 4.1.</u> A scaling limit $f_r(x)$ of $w_r(k;n)$ provided $k/n=x$ in the multiple choice problem satisfies the equation

$$(4.2)\qquad \begin{aligned} df_r(x) &= (1-\Phi(x))^{-1}f_r(x)d\Phi(x) - x^{-1}(R(x)+f_{r-1}(x)-f_r(x))^{+}dx,\\ f_r(1) &= 0,\quad r=1,2,...,\\ f_0(x) &= 0 \quad \text{for} \quad 0 < x < 1. \end{aligned}$$

The optimal value v_r^* equals $f_r(0+)$.

<u>PROPOSITION 4.2.</u> Let $g_r(x) = x^{-1}(1-\Phi(x))f_r(x)$ for $r=1,2,..$. If $\Phi(x)$ is continuous for $0 < x < 1$, then they are differentiable and satisfy

$$(4.3)\qquad dg_r(x) = -x^{-1}\max\{h(x) + g_{r-1}(x),\ g_r(x)\}\,dx,\qquad g_r(1) = 0$$

where $h(x)$ is defined in (2.9).

Let $h_r(x) = h(x) + g_{r-1}(x)$ and

(4.4) $\quad H_r(x) = h_r(x) - \int_x^1 y^{-1} h_r(y)dy \quad$ for $r=1,2,\ldots$.

CONDITION(Φ_r). $H_r(x)$ changes its sign one from − to + as x increases.

Let $\alpha_r^* = \inf\{x;\ H_r(x) \geq 0\}$.

THEOREM 4.3. The optimal value v_r^* of permitting r-object choices is $(1-\Phi(\alpha_r^*))f_r(\alpha_r^*)$, and the stopping islands are determined by the sequence $(\alpha_k^*;\ k=1,2,\ldots,r)$.

In the rest of this section it is restricted to the uniform distribution: $p_k = 1/n$, $k=1,2,\ldots,n$. Then (4.2) implies

(4.5)
$$\begin{aligned} &df_r(x)/dx = (1-x)^{-1}f_r(x) - x^{-1}(R(x)+f_{r-1}(x)-f_r(x))^+, \quad 0<x<1,\\ &f_r(1) = 0 \end{aligned}$$

where $R(x) = -x(1-x)^{-1}\log(x)$. We now use Proposition 4.2. From (4.3), we have that

(4.6) $\quad g_r(x) = \int_x^1 y^{-1}h_r(y)dy = -\int_x^1 y^{-1}g_{r-1}(y)dy + \int_x^1 y^{-1}(\int_y^1 z^{-1}d\Phi(z))dy$

on $\{x;\ -\log(x) + g_{r-1}(x) \geq g_r(x)\}$ and, in the neighborhood of x=0,

(4.7) $\quad g_r(x) = (\mathrm{const})/x.$

From (4.6) and (4.7), $f_r(x)$ is solved. To denote this solution explicitly, we set inductively

(4.8)
$$\begin{aligned} K_{i+1} &= L_i^3/(3!) + (c_{i-1}-c_i)\exp(L_i) + K_iL_i,\\ c_i &= c_{i-1} + L_i\exp(-L_i), \quad i=1,2,\ldots,r \end{aligned}$$

where $L_i = 1+\sqrt{(1-2K_i)}$ and $K_1=0$ and $c_0=0$. It is seen that, from the continuity of the solution, that

(4.9)
$$f_r(x) \begin{cases} = c_r/(1-x), & 0< x \leq x_r\\ = x/(1-x)*\{\log^2(x)/(2!) + c_{r-1}/x + K_r\}, & x_r \leq x \leq x_{r-1},\\ = x/(1-x)*\{-\log^3(x)/(3!) + \log^2(x)/(2!) & \\ \qquad -K_{r-1}\log(x)+c_{r-2}/x\}, & x_{r-1} \leq x \leq x_{r-2},\\ = \ldots & \end{cases}$$

where $x_i = \exp(-L_i)$, $i=1,2,..$ and $0 < x_r < x_{r-1} < \cdot\cdot < x_1 = e^{-2} < 1$. The optimal value v_r^* of r-object choices is $v_r^* = f_r(0+) = c_r$. Therefore we can determine the optimal value for every r by the iteration (4.8). For example, $c_1 = 2e^{-2} = .2707$ and $c_2 = c_1 + (1+\sqrt{21}/3)\exp(-(1+\sqrt{21}/3)) = .4725$. The first two terms are consistent with Presman and Sonin(1972), and Tamaki(1979a) respectively. Numerical calculation for different values of r gives the following results:

Table 3

Times of choice r :	1	2	3	4	5	6	7
Optimal value v_r^* :	.2707	.4725	.6208	.7149	.7552	.7609	.7610

It seems here as if the optimal value converges, but in the original model of the situation it must tend to unity as r increases. The cause of this may be that we have taken the limit n to infinity for a prefixed number r.

3.5. A MULTIPLE CHOICE PROBLEM (II)

A multiple choice problem which is to select the best and the second best objects, permitting a 2-object choice, is considered by Nikolaev(1977) and Sakaguchi(1979). Sakaguchi treats the uncertain employment problem i.e. with a refusal probability, in our terminology, which we have discussed in section 3.3. While this model is not considered here, we shall discuss the case of a random number of objects, and calculate the uniformly distributed special case as previously.

The optimality equation obtained by Sakaguchi(1979) and Tamaki(1979b) is as follows:

$$u_1(j) = j(j-1)/(n(n-1)),$$
$$u_2(j) = \max\{u_1(j), \textstyle\sum_{k=j+1}^{n} j(j-1)/(k(k-1)(k-2))*\sum_{s=1}^{2} u_s(k)\},\ j=2,..,n-1,$$
$$u_1(n) = u_2(n) = 1,$$
$$v(1) = \max\{(u_1(2)+u_2(2))/2, \textstyle\sum_{k=1}^{n} 1/(k(k-1))*v(k)\},$$
$$v(i) = \max\{\textstyle\sum_{k=i+1}^{n} i(i-1)/(k(k-1)(k-2))*\sum_{s=1}^{2} u_s(k), \sum_{k=i+1}^{n} i/(k(k-1))*v(k)\},$$
$$i=2,..,n-2,\ v(n) = 0.$$

The solution of $v(1) = v(1;n)$ is the optimal value, that is, the maximum probability of stopping twice which includes the best and the second best objects.

Similarly as in the previous sections, we are concerned with the bounded random number of objects and the same notations are used. Because two stops are required, it is enough to assume that $N \geq 3$, that is, $p_1 = p_2 = 0$. Hence $\pi_1 = \pi_2$ holds. We then have the following optimality equation

$$u_1(j) = \textstyle\sum_{k=j}^{n} j(j-1)p_k/(k(k-1)\pi_j),\ j=2,..,n-1,$$
$$u_2(j) = \max\{u_1(j), \textstyle\sum_{k=j+1}^{n} j(j-1)\pi_k/(k(k-1)(k-2)\pi_j)*\sum_{s=1}^{2} u_s(k)\},$$
$$u_1(n)=u_2(n)=1, \tag{5.1}$$

$$v(1) = \max\left\{(u_1(2)+u_2(2))/2,\ \sum_{k=2}^{n}\pi_k/(k(k-1)\pi_1)*v(k)\right\},$$

$$v(i) = \max\left\{\sum_{k=i+1}^{n} i(i-1)\pi_k/(k(k-1)(k-2)\pi_i)*\sum_{s=1}^{2}u_s(k),\ \sum_{k=i+1}^{n} i\pi_k/(k(k-1)\pi_i)*v(k)\right\},\ i=2,..,n-1,$$

$$v(n) = 0.$$

Define the conditional optimal value $w(k) = w(k;n)$, $k=1,2,..,n$ by

$$w(1) = \sum_{s=1}^{2}u_s(2)/2,$$

$$w(k) = \sum_{s=k+1}^{n}k(k-1)\pi_s/(s(s-1)(s-2)\pi_k)*\sum_{t=1}^{2}u_t(s),\ k=2,..,n-1,$$

$$w(n) = 0.$$

Then

$$(5.2)\qquad w(k+1) - w(k) = p_k w(k+1)/\pi_k - \pi_{k+1}/((k+1)\pi_k)* * \sum_{s=k+1}^{n}k(k+1)p_s/(s(s-1)\pi_{k+1}) - w(k+1) + \left(\sum_{s=k+1}^{n}k(k+1)p_s/(s(s-1)\pi_{k+1}) - w(k+1)\right)^+ .$$

Also define

$$\tilde{w}(k) = \sum_{s=k+1}^{n}k\pi_s/(s(s-1)\pi_k)*v(s),\ k=1,..,n-1,$$

$$\tilde{w}(n) = 0.$$

Then this satisfies

$$(5.3)\qquad \tilde{w}(k+1) - \tilde{w}(k) = p_k\tilde{w}(k+1)/\pi_k - \pi_{k+1}/((k+1)\pi_k)*(w(k+1)-w(k+1))^+.$$

Hence, if Assumptions (I), (II) of section 3.2, and if

$$(1-\Phi(x))^{-1}\int_x^1 y^{-2}d\Phi(y) \to 1 \text{ as } x \to 1$$

and

$$x^2\int_x^1 y^{-2}d\Phi(y) \to 0 \text{ as } x \to 0$$

hold, we have the next two integral equations by taking the scaling limit.

<u>PROPOSITION 5.1.</u> Let $f(x) = \lim_{k,n} w(k;n)$ and $\tilde{f}(x) = \lim_{k,n}\tilde{w}(k;n)$ provided $k/n=x$. Then these satisfy

$$(5.4)\qquad \begin{aligned} df(x) &= f(x)(1-\Phi(x))^{-1}d\Phi(x) - x^{-1}\left\{R_2(x)-f(x) + (R_2(x)-f(x))^+\right\}dx, \\ f(1) &= 0 \end{aligned}$$

where $R_2(x) = x^2(1-\Phi(x))^{-1}\int_x^1 y^{-2}d\Phi(y)$, $0\leq x\leq 1$, and

$$(5.5)\quad \begin{aligned} d\tilde{f}(x) &= \tilde{f}(x)(1-\Phi(x))^{-1}d\Phi(x) - x^{-1}(f(x)-\tilde{f}(x))^{+}dx, \\ \tilde{f}(1) &= 0. \end{aligned}$$

THEOREM 5.2. The optimal value $\tilde{v}^* = \lim v(1;n)$ in the asymptotic form is given by the solution $\tilde{f}(0+)$ of (5.5).

Proof. From $v(1;n) = \max\{w(1;n), \tilde{w}(1;n)\}$, we have $\tilde{v}^* = \max\{f(0+), \tilde{f}(0+)\}$. By (5.4), $f(0+) = 0$ implies the result, $\tilde{v}^* = \tilde{f}(0+)$.

EXAMPLE 5.1. We calculate the optimal value and the stopping island for the case of $p_i = 1/n$ for $i=1,..,n$, that is, the uniform distribution $d\Phi(x) = dx$. By the same method as in previous sections,

$$(5.6)\quad f(x)\begin{cases} = 2x/(1-x)*(1+x\log(x)-x), & \alpha_1^* \leqq x < 1, \\ = x/(1-x)*\{1-\log(x)+x-2\alpha_1^*+\log(\alpha_1^*)\}, & 0 < x \leqq \alpha_1^* \end{cases}$$

where α_1^* (=.28467) is a unique solution x of $1-x = 2(1+x\log(x)-x)$ in $0 < x < 1$. Also,

$$(5.7)\quad \tilde{f}(x)\begin{cases} = -2x/(1-x)*((1+x)\log(x) - 2x + 2), & \alpha_1^* \leqq x < 1, \\ = x/(1-x)*\{\log^2(x)/2 - (\log(\alpha_1^*)-2\alpha_1^*+1)\log(x) - x \\ \qquad + \log^2(\alpha_1^*)/2 - 4\alpha_1^*\log(\alpha_1^*) - \log(\alpha_1^*) + 5\alpha_1^* - 4\}, & \alpha_2^* \leqq x \leqq \alpha_1^*, \\ = \alpha_2^*/(1-x)*\{1 - \log(\alpha_2^*) + \alpha_2^* - 2\alpha_1^* + \log(\alpha_1^*)\}, & 0 < x \leqq \alpha_2^* \end{cases}$$

where α_2^* (=.09610) is a unique solution x of

$$\begin{aligned} &x - \log(x) + \log(\alpha_1^*) - 2\alpha_1^* + 1 \\ &= \log^2(x)/2 - (\log(\alpha_1^*)-2\alpha_1^*+1)\log(x) - x + \log^2(\alpha_1^*)/2 - 4\alpha_1^*\log(\alpha_1^*) \\ &\quad - \log(\alpha_1^*) + 5\alpha_1^* - 4. \end{aligned}$$

Hence the optimal value $\tilde{v}^* = \tilde{f}(0+)$ equals

$$(5.8)\quad \tilde{f}(0+) = \alpha_2^*\{1-\log(\alpha_2^*)+\alpha_2^*-2\alpha_1^*+\log(\alpha_1^*)\} \ (= .15498).$$

The optimal strategy in the asymptotic form is that

(i) on $[0, \alpha_2^*]$, we pass

(ii) on $[\alpha_2^*, \alpha_1^*]$, we make the 1st stop if the relative best object appears

(iii) on $[\alpha_1^*, 1]$, we make the 2^{nd} stop if the relative best or 2^{nd} best object appears.

ACKNOWLEDGEMENTS

The author is deeply grateful to Professor N. Furukawa, Kyushu University, for his guidance and many valuable suggestions. He also wishes to express his gratitude to Professor M. Murakami, Chiba University, for his continuous supports and encouragements. Furthermore the auther wishes to thank Professor M. Kurano and J. Nakagami, Chiba University, for their helpful suggestions and discussions.

REFERENCES

1) Abdel-Hamid, A.R., Bather, J.A. and Thurstrum, G.B. : The Secretary Problem with Unkown Number of Candidates, J. Appl. Prob. 19(1982), 619-630.

2) Bartoszynski, R. and Govindarajulu, Z. : The Secretary Problem with Interview Cost, Sankhya, 40(1978), Series B, 11-28.

3) Bellman, R. : Dynamic Programming, 1957, Princeton University Press, Princeton.

4) Brèmaud, P. : Point Processes and Queues - Martingale Dynamics, 1981, Springer Verlag, New York.

5) Chow, Y.S., Robbins, H. and Siegmund, D. : Great Expectations: The Theory of Optimal Stopping, 1971, Houghton Mifflin Co., Massachusetts.

6) DeGroot, M.H. : Optimal Statistical Decisions, 1970, McGrow-Hill Book Co., New York.

7) Derman, C., Lieberman, G.J. and Ross, S.M.(Unpublished) : On the candidate problem with a random number of candidates.

8) Dynkin, E.B. and Yuskevich, A.A. : Theorems and Problems in Markov Processes, 1969, Plenum Press, New York.

9) Feller, W. : An Introduction to Probability Theory and Its Applications, vol. II, 1966, John Wiley, New York.

10) Fishburn, P.C. : The theoty of representative majority decision, Econometrica, 39(1971), 273-284.

11) Gianini, J. : Optimal selection based on relative ranks with a random number of individuals, Adv. Appl. Prob. 11(1979), 720-736.

12) Gilbert, J.P. and Mosteller, F. : Recognizing the maximum of a sequence, J. Amer. Stat. Assoc. 61(1966), 35-73.

13) Gusein-Zade, S.M. : The problem of choice and the optimal stopping rule for a sequence of independent trials, T. Prob. Appl. 11(1966), 472-476.

14) Irle, A. : On the Best Choice Problem with Random Population Size, Z. fur Oper. Res. 24(1980), 177-190.

15) Kadane, J.B. : Reversibility of a multilateral sequential game: proof of a conjecture of Sakaguchi, J. Oper. Res. Soc. Japan, 21(1978), 509-516.

16) Kurano, M., Yasuda, M. and Nakagami, J. : Multi-variate stopping problem with a majority rule, J. Oper. Res. Soc. Japan, 23(1980), 205-223.

17) Lorenzen, T.J. : Optimal stopping with sampling cost: The secretary problem, Ann. Prob. 9(1981), 167-172.

18) Mucci, A.G. : Differential equations and optimal choice problems, Ann. Stat. 1(1973), 104-113.

19) Nash, J. : Non-cooperative games, Annals of Math., 54(1951), 286-295.

20) Nikolaev, M.L. : On a generalization of the best choice problem, T. Prob. Appl. 22(1977), 187-190.

21) Presman, E.L. and Sonin, I.M. : The best choice problem for a random number of objects, T. Prob. Appl. 17(1972), 657-668.

22) Presman, E.L. and Sonin, I.M. : Equilibrium points in a game related to the best choice problem, T. Prob. Appl. 20(1975), 770-781.

23) Rasmussen, W.T. and Robbins, H. : The candidate problem with unknown population size, J. Appl. Prob. 12(1975), 692-701.

24) Ross, S.M. : Applied Probability Models with Optimization Applications, 1970, Holden Day, California.

25) Sakaguchi, M. : A bilateral sequential game for sums of bivariate random variables, J. Oper. Res. Soc. Japan, 21(1978), 486-508.

26) Sakaguchi, M. : A generalized secretary problem with uncertain employ-

ment, Math. Japonica 23(1979), 647-653.

27) Sakaguchi, M. : Non-zero-sum games related to the secretary problem, J. Oper. Res. Soc. Japan, 23(1980), 287-293.

28) Siryaev, A.N. : Statistical Sequential Analysis, Trans. Math. Mono. 38, 1973.

29) Snell, I.L. : Applications of martingale system theory, Trans. Amer. Math. Soc. 73(1952), 293-312.

30) Smith, M.H. : A secretary problem with uncertain employment, J. Appl. Prob. 12(1975), 620-624.

31) Stewart, T.J. : The secretary problem with an unknown number of options, Oper. Res. 29(1981), 130-145.

32) Tamaki, M. : OLA policy and the best choice problem with random number of objects, Math. Japonica 24(1979a), 451-457.

33) Tamaki, M. : Recognizing both the maximum and the second maximun of a sequence, J. Appl. Prob. 16(1979b), 803-812.

34) Yasuda, M., Nakagami N. and Kurano, M. : Multi-variate stopping problem with a monotone rule, J. Oper. Res. Soc. Japan, 25(1982), 334-350.

35) Yasuda, M. : On a stopping problem involving refusal and forced stopping, J. Appl. Prob. 20(1983), 71-81.

36) Yasuda, M. : Asymptotic results for the best choice problem with a random number of objects, To appear in J. Applied Probability, September, 1984.

37) Wald, A. : Sequential Analysis, 1947, Wiley, New York.