AN OPTIMAL ASSIGNMENT PROBLEM FOR MULTIPLE OBJECTS PER PERIOD-CASE OF A PARTIALLY OBSERVABLE MARKOV CHAIN

Nakai, Toru
Faculty of Economics, Kyushu University

https://doi.org/10.5109/13478

出版情報：Bulletin of informatics and cybernetics. 31 (1), pp.23-34, 1999-03. Research Association of Statistical Sciences
バージョン：
権利関係：

Bulletin of Informatics and Cybernetics, Vol. 31, No. 1, 1999

AN OPTIMAL ASSIGNMENT PROBLEM FOR MULTIPLE OBJECTS PER PERIOD – CASE OF A PARTIALLY OBSERVABLE MARKOV CHAIN

By

Tōru NAKAI*

Abstract

We deal with a sequential stochastic assignment problem on a partially observable Markov chain. This is a generalization of a problem treated in Derman, Lieberman and Ross (1972). Especially, not only one job will appear at one time period. Under several assumptions, we consider a partially observable Markov chain and several properties concerning a relation between observations and information. On basis of these properties, we investigate the sequential stochastic assignment problem on this partially observable Markov chain.

1. Introduction

We deal with a sequential stochastic assignment problem on a partially observable Markov chain. This is a generalization of problem treated in Derman, Lieberman and Ross (1972). There are m jobs which appear in a sequential order, and m persons to be assigned to these jobs. This problem is how to assign these persons to m jobs in order to maximize the total expected reward. Especially, not only one job will appear at one time period. Associated with each job, there is a random variable depending on the state of this chain, which indicates a value of a job. In order to investigate this problem, we initially treat an optimal selection problem. This is a problem to select a predetermined number of jobs in order to maximize the total expected reward of these selected. In both problems, the number of observations available at one time period is not known in advance, but only the probability is known beforehand.

When the state of this Markov chain indicates the economic condition, it affects the values of jobs available at each time period. Usually, this condition is not known directly, and there is only partial information about this condition. Theorems 1 and 2 give essential properties about posterior information after having obtained through the values of these jobs. As a learning procedure, we employ the Bayes' theorem. Albright (1974) and Nakai (1986b) also treated a sequential stochastic assignment problem on a partially observable Markov chain. Brown and Solomon (1973), Nakai (1985, 1990,

* Faculty of Economics, Kyushu University, Hakozaki 6-19-1, Fukuoka 812-8581, Japan

1995, 1996), Pinedo and Ross (1980) considered a partially observable Markov decision problems.

In Section 2, we deal with an optimal selection problem. In subsection 2.1, we make some preliminary notations and assumptions to treat these problems. In subsection 2.2, we summarize some properties about posterior information. For this problem, we use some results obtained in Nakai (1986a). In subsection 2.3, we summarize some essential properties about the optimal selection problem. Finally, we investigate a sequential stochastic assignment problem on this partially observable Markov chain.

2. Optimal Selection Problem

2.1. Some Preliminaries

An optimal selection problem is to select a predetermined number of jobs in order to maximize the total expected reward of these selected. Associated with each job, there is a random variable indicating its value. These random variables are independent and identically distributed. Let N be the planning horizon of this problem. On the other hand, there are m jobs having a random value $\{X_i\}_{i=1,\cdots,m}$. The objective of this problem is to maximize the total expected reward by selecting k jobs out of m during N time periods. If a job is not selected at one time period, then it is not available for future decisions.

In general, the number of jobs observable at one time period is assumed to be random with known distribution. When there are m jobs during N time periods, let $\{p_{N,m}(n)\}_{n=0,1,\cdots,m}$ be a probability of n jobs at the initial time period. When these jobs appear uniformly and independently, $p_{N,m}(n)$ is as follows.

$$p_{N,m}(n) = {}_mC_n \frac{(N-1)^{m-n}}{N^m}. \quad (0 \le n \le m, p_{1,m}(m) = 1) \tag{1}$$

We concentrate on this case, but it is possible to consider these problems for a more general case.

Let $X_1, \cdots, X_n$ be n i.i.d. random variables, and let further $X_{(1)}, \cdots, X_{(n)}$ be the order statistics corresponding to $X_1, \cdots, X_n$. $(X_{(1)} \ge \cdots \ge X_{(n)})$ We arrange these values from the largest to the smallest one, for convenience sake. We also use the notation $\{X_{(i)}\}_{i=1,\cdots,n}$ instead of $\{X_i\}_{i=1,\cdots,n}$. These n random variables $\{X_i\}_{i=1,\cdots,n}$ are absolutely continuous with density $f(x)$. It is well known that

$$g_{n,i}\left(x_{(i)}\right) = \frac{n!}{(i-1)!\,(n-i)!}\left(F\left(x_{(i)}\right)\right)^{n-i}\left(1-F\left(x_{(i)}\right)\right)^{i-1} f\left(x_{(i)}\right) \tag{2}$$

is the probability distribution of the i–th order statistic $X_{(i)}$ (See Wilks (1962)).

For a decreasing sequence $\{a_i\}$ $(a_0 = \infty)$ of positive numbers, define a function $U_n(a_i, a_{i-1}|k, y)$ by

$$U_n(a_i, a_{i-1}|k, y) = \int_0^{a_i \wedge y} a_i\, h_{n,k}\left(x_{(k)}\right) f\left(x_{(k)}\right) dx_{(k)}$$

$$+ \int_{a_i \wedge y}^{a_{i-1} \wedge y} x_{(k)} h_{n,k}\left(x_{(k)}\right) f\left(x_{(k)}\right) dx_{(k)}$$
$$+ \int_{a_{i-1} \wedge y}^{y} U_n\left(a_{i-1}, a_{i-2} \left| k+1, x_{(k)}\right.\right) f\left(x_{(k)}\right) dx_{(k)} \quad (3)$$

where $a \wedge b = \min\{a, b\}$, $U_n(a_i, a_{i-1}|n+1, y) = a_i \ (y \geq 0)$ and

$$h_{n,k}\left(x_{(k)}\right) = \frac{n!}{(n-k)!}\left(F\left(x_{(k)}\right)\right)^{n-k}. \quad (4)$$

For any decreasing sequence $\{a_i\}_{i \geq 0}$, we construct a sequence $\{\alpha_{i,n}\}_{i \geq 0}$ as

$$\alpha_{i,n} = U_n(a_i, a_{i-1}|1, \infty) \quad (i \geq 1), \quad (5)$$

where $\alpha_{0,n} = \infty$. We summarize some properties according to Nakai (1986a), which are necessary to investigate these problems.

Lemma 1

$$\alpha_{i,n} = \sum_{j=1}^{n \wedge i} \int_{a_{i-j+1}}^{a_{i-j}} x_{(j)} g_{n,j}\left(x_{(j)}\right) dx_{(j)} + \sum_{j=0}^{n \wedge (i-1)} a_{i-j}\, {}_nC_j \left(1 - F\left(a_{i-j}\right)\right)^j \left(F\left(a_{i-j}\right)\right)^{n-j}$$

Lemma 2 *If a sequence $\{a_i\}_{i=0,1,2,\cdots} (a_0 = \infty)$ is decreasing with respect to i, then the sequence $\{\alpha_{i,n}\}_{i,n=0,1,2,\cdots}$ defined by Equation* (5) *is also decreasing with respect to i.*

Lemma 3 *If $a_0 = \infty$ and $a_1 = a_2 = \cdots = 0$, then $\alpha_{i,n} = \mathrm{E}[X_{(i)}]\,(1 \leq i \leq n)$ and $\alpha_{i,n} = 0$. $(i > n)$*

2.2. A Partially Observable Markov Chain

Let's consider a countable state partially observable Markov chain. Let $\{0, 1, 2, \cdots\}$ be a set of states, and let further $\boldsymbol{P} = (p_{ss'})$ be a transition probability matrix of this chain. The random variables observed at each time period depend on a state of this chain. When the state of this chain is s and n jobs appear, we assume the random variables associated with these jobs to be independent and identically distributed, but dependent on s. Let S be a random variable indicating the state of this Markov chain. When a state of this chain is s, the conditional expectation μ_s is finite, and the distribution function $F_s(x)$ of these random variables are absolutely continuous with density $f_s(x)$. We consider two assumptions.

Assumption 1 *If $t < s$, $(s, t = 0, 1, 2, \cdots)$ then, for $x \leq y$,*

$$f_t(y) f_s(x) \geq f_s(y) f_t(x), \text{ i.e., } \begin{vmatrix} f_s(x) & f_t(x) \\ f_s(y) & f_t(y) \end{vmatrix} \geq 0. \quad (6)$$

Assumption 2 *For any pair of t and t', if $t \geq t'$, $(t, t' = 0, 1, 2, \cdots)$ then*

$$p_{st'} p_{s't} \geq p_{s't'} p_{st}, \text{ i.e., } \begin{vmatrix} p_{s't} & p_{s't'} \\ p_{st} & p_{st'} \end{vmatrix} \geq 0, \quad (7)$$

where $s \leq s'$. $(s, s' = 0, 1, 2, \cdots)$

If the state of this chain corresponds to the condition of the economy, then Assumption 1 means that this condition becomes better as s becomes smaller. The state 0 is the best one among them. It is easy to show that $\{\mu_s\}_{s=0,1,2,\cdots}$ is decreasing with respect to s. From Assumption 1, if $s > t$ $(s,t = 0,1,2,\cdots)$, then $X(t)$ is greater than $X(s)$ by means of the likelihood ratio. This also implies that $X(s)$ satisfies the TP_2 condition (totally positivity of order two).

For this countable state Markov chain, assume that the state of this chain is not known directly. All information about this state is summarized by a probability distribution Φ on the state space, i.e., $\Phi \in \mathcal{S} = \{ \Phi | \Phi = (\phi_0, \phi_1, \phi_2, \cdots), \phi_s \geq 0, \sum_{s=0}^{\infty} \phi_s = 1\}$. At each time period, information is obtained through random variables depending on the state of this chain. If there is no observation at one time period, there is no information about it. Let prior information about the state of this chain be $\Phi (\in \mathcal{S})$. When n jobs appear with $\boldsymbol{x} = (x_1, \cdots, x_n)$, then, by learning, we improve information as $T(\overline{\Phi}|\boldsymbol{x})$. In this paper, we use Bayes' theorem as the learning procedure. Since this chain initially makes a transition according to $\boldsymbol{P}$, information $\overline{\Phi}$ changes to

$$\begin{cases} \overline{\phi}_{s'} &= \sum_{s=0}^{\infty} \phi_s p_{s\,s'}, \\ \overline{\Phi} &= (\overline{\phi}_0, \overline{\phi}_1, \overline{\phi}_2, \cdots). \end{cases} \tag{8}$$

Let $f_t(\boldsymbol{x})$ be a joint probability distribution function of n random variables $\boldsymbol{X} = (X_1, \cdots, X_n)$ when the state of the chain is t. Since the random variables are independent with each other and identically distributed, we note that

$$f_t(\boldsymbol{x}) = \prod_{i=1}^{n} f_t(x_i). \tag{9}$$

Information about the state of this chain is updated by using Bayes' theorem as

$$\begin{cases} T_{s'}\left(\overline{\Phi}\middle|\,\boldsymbol{x}\right) = \dfrac{\overline{\phi}_{s'} f_{s'}(\boldsymbol{x})}{\sum_{s=0}^{\infty} \overline{\phi}_s f_s(\boldsymbol{x})} \\ T(\overline{\Phi}|\boldsymbol{x}) \quad = \left(T_0\left(\overline{\Phi}\middle|\,\boldsymbol{x}\right), T_1\left(\overline{\Phi}\middle|\,\boldsymbol{x}\right), T_2\left(\overline{\Phi}\middle|\,\boldsymbol{x}\right), \cdots\right). \end{cases} \tag{10}$$

Initially, we introduce an order on $\mathcal{S}$ by using a likelihood ratio. This order is also treated in Nakai (1985, 1986b, 1990, 1993) and Ross (1983), etc.

Definition 1 *For any pair of Φ and Ψ in $\mathcal{S}$, denote $\Phi >_l \Psi$ if and only if*

$$\phi_s \psi_t \leq \phi_t \psi_s, i.e., \begin{vmatrix} \psi_s & \psi_t \\ \phi_s & \phi_t \end{vmatrix} \geq 0 \tag{11}$$

for any s and t, $(t \leq s, s, t = 0, 1, 2, \cdots)$ and at least one pair of s and t, $\phi_s \psi_t < \phi_t \psi_s$. If $\phi_s = \psi_s$ for any $s = 0, 1, 2, \cdots$, then $\Phi =_l \Psi$. $\Phi \geq_l \Psi$, if and only if $\Phi =_l \Psi$ and $\Phi >_l \Psi$.

It is easy to show that this order is a partial order. We say a function $u(\Phi)$ on $\mathcal{S}$ to be increasing by means of this order, if and only if $u(\Phi) \geq u(\Psi)$ for any pair Φ and Ψ where $\Phi \geq_l \Psi$.

Definition 2 *For any $\boldsymbol{x}$ and $\boldsymbol{y}$ of k observations ($\boldsymbol{x}, \boldsymbol{y} \in \mathcal{R}^k$), $\boldsymbol{x} \prec \boldsymbol{y}$ if and only if $x_{(i)} \leq y_{(i)}$ $(1 \leq i \leq k)$.*

Assumptions 1 and 2 imply Lemma 4 (Nakai (1985, 1986b, 1990, 1993)).

Lemma 4 *For any pair of t and s, ($t < s, s, t = 0, 1, 2, \cdots$) if $\boldsymbol{x} \prec \boldsymbol{y}$, then*

$$f_t(\boldsymbol{y})f_s(\boldsymbol{x}) \geq f_s(\boldsymbol{y})f_t(\boldsymbol{x}), i.e., \begin{vmatrix} f_s(\boldsymbol{x}) & f_t(\boldsymbol{x}) \\ f_s(\boldsymbol{y}) & f_t(\boldsymbol{y}) \end{vmatrix} \geq 0. \tag{12}$$

As for posterior information, we obtain the following properties by a method similar to one used in Nakai (1985, 1986b, 1993). We have already obtained for the case where $n = 1$ in (1985, 1986b, 1993). We use the notation $\mathrm{E}_{\Phi}[\,\cdot\,]$ for the expectation when a probability distribution on the state space is Φ.

Theorem 1 *If $\boldsymbol{x} \prec \boldsymbol{y}$, then $T\left(\overline{\Phi}\middle|\,\boldsymbol{x}\right) \leq_l T\left(\overline{\Phi}\middle|\,\boldsymbol{y}\right)$ $(\Phi \in \mathcal{S})$.*

Theorem 2 *If $\Phi \geq_l \Psi$, then $\overline{\Phi} \geq_l \overline{\Psi}$ and $T\left(\overline{\Phi}\middle|\,\boldsymbol{x}\right) \geq_l T\left(\overline{\Psi}\middle|\,\boldsymbol{x}\right)$ $(\boldsymbol{x} \in \mathcal{R}^n)$.*

Lemma 5 *If $\boldsymbol{\Phi} \geq_l \boldsymbol{\Psi}$ $(\boldsymbol{\Phi}, \boldsymbol{\Psi} \in \mathcal{S})$, then*

$$\mathrm{E}_{\boldsymbol{\Psi}}[\varphi(\boldsymbol{X})] = \sum_{i=0}^{\infty} \phi_i \int \varphi(\boldsymbol{x}) f_i(\boldsymbol{x}) d\boldsymbol{x} \leq \sum_{i=0}^{\infty} \psi_i \int \varphi(\boldsymbol{x}) f_i(\boldsymbol{x}) d\boldsymbol{x} = \mathrm{E}_{\boldsymbol{\Phi}}[\psi(\boldsymbol{X})]$$

for any non-decreasing function $\varphi(\cdot)$ of $\boldsymbol{x}$.

2.3. Optimal Selection Problem on a Partially Observable Markov Chain

Let's consider an optimal selection problem on the partially observable Markov chain treated in the last subsection. There are m jobs, and each one of them appears uniformly and independently during N time periods. Associated with each job is a random variable indicating its value, which depends only on a state of this chain. If a job with a value x is selected, a reward x is earned. The density function of these random variables is known previously. If a state of this chain corresponds to an economic condition, the economic condition is not known precisely, but there exists only partial information about it. Concerning the number of jobs at the initial time period, the probability $\{p_{N,m}(n)\}_{n=0,1,\cdots,m}$ is given by Equation (1), which is independent of the state of this chain. The value of a job appeared at one time period, however, depends only on the current state s of this chain.

The optimal selection problem is to select k jobs out of m during N time periods when prior information about the state is Φ. We consider (N, m, k, Φ) as a state of this problem. The objective of this problem is to find an optimal policy in order to maximize the total expected reward. When the problem is in state (N, m, k, Φ), consider that n jobs appear with $\boldsymbol{x} = (x_{(1)}, \cdots, x_{(n)})$. If i jobs out of n are selected, then, at the next instance, the state of the problem becomes $\left(N-1, m-n, k-i, T(\overline{\Phi}|\boldsymbol{x})\right)$.

Let $v_{N,m,k,\Phi}$ be the total expected reward (under the optimal policy) for a problem in state (N, m, k, Φ). As in Ross (1970), the principle of optimality implies

$$v_{N,m,k,\Phi} = \sum_{n=0}^{m} v_{N,m,k,\Phi}(n) p_{N,m}(n), \tag{13}$$

$$v_{N,m,k,\Phi}(n) = \mathrm{E}\left[v_{N,m,k,\Phi}\left(n; X_{(1)}, \cdots, X_{(n)}\right)\right], \tag{14}$$

$$v_{N,m,k,\Phi}(n; \boldsymbol{x}) = \max\left\{\sum_{j=1}^{k} x_{(j)} + v_{N-1,m-n,k-i,T(\overline{\Phi},\boldsymbol{x})}\right\}. \tag{15}$$

Define sequences $\left\{d^i_{N,\Phi,m}\right\}$, $\left\{d^i_{N,\Phi,m}(n)\right\}$ and $\left\{e^i_{N,\Phi,m}\right\}$ of non-negative numbers recursively in the following manner. $(i, n = 0, 1, 2, \cdots, m,\ N = 0, 1, 2, \cdots,\ \Phi \in \mathcal{S})$

$$d^i_{N,\Phi,m} = \sum_{n=0}^{m} d^i_{N,j,m}(n) p_{N,m}(n), \tag{16}$$

$$d^i_{N,\Phi,m}(n) = \mathrm{E}\left[d^i_{N,\Phi,m}(n; \boldsymbol{X})\right], \tag{17}$$

$$d^i_{N,\Phi,m}(n; \boldsymbol{x}) = U_n\left(e^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}, e^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} \middle| 1, \infty\right), \tag{18}$$

$$d^i_{N,\Phi,m}(0) = e^i_{N-1,\overline{\Phi},m}, \quad e^i_{N-1,\Phi,m} = d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m}, \qquad (N \ge 2) \tag{19}$$

where $\overline{\Phi} = (\overline{\phi}_0, \overline{\phi}_1, \overline{\phi}_2, \cdots)$, $\overline{\phi}_s = \sum_{t=0}^{\infty} \phi_t p_{ts}$, $d^0_{N,\Phi,m} = d^0_{N,\Phi,m}(n) = \infty$, and $d^i_{0,\Phi,0} = 0$. When prior information about the state of the chain is Φ, Lemma 3 implies $d^i_{1,\Phi,m} = d^i_{N,j,m}(m)$ and $d^i_{1,\Phi,m}(m) = \mathrm{E}_\Phi\left[X_{(i)}\right]$. When $\Phi \le_l \Psi$ implies $a_\Phi \ge_l a_\Psi$, we call a function a_Φ to be decreasing with respect to Φ. The following properties are also obtained in Nakai(1996).

Lemma 6 *$d^i_{N,\Phi,m}$, $d^i_{N,\Phi,m}(n)$ and $e^i_{N,\Phi,m}$ are decreasing with respect to Φ by means of the likelihood ratio. $\{d^i_{N,\Phi,m}\}_{0\le i\le m}$, $\{d^i_{N,\Phi,m}(n)\}_{0\le i,n\le m}$ and $\{e^i_{N,\Phi,m}\}_{0\le i\le m}$ are decreasing sequences of i.*

Theorem 3 *For a problem in state (N, m, k, Φ), assume that n random variables appear with $x_{(1)}, \cdots, x_{(n)}$. Let j be the largest number which satisfies $x_{(j)} \ge e^{k-j+1}_{N-1,\Phi,m-n}$ and $1 \le j \le k \wedge n$, then it is optimal to select the largest j observations out of n, i.e., $x_{(1)}, \cdots, x_{(j)}$.*

Theorem 4 *For an optimal selection problem in state (N, m, k, Φ),*

$$v_{N,m,k,\Phi} = \sum_{i=1}^{k} d^i_{N,\Phi,m}, \quad v_{N,m,k,\Phi}(n) = \sum_{i=1}^{k} d^i_{N,\Phi,m}(n). \tag{20}$$

$v_{N,m,k,\Phi}$ and $v_{N,m,k,\Phi}(n)$ are decreasing with respect to Φ.

Remark 1 *Theorem 4 implies* $v_{N,m,k,\Phi} - v_{N,m,k-1,\Phi} = d^k_{N,\Phi,m}$, *and it is possible to consider that a value of* $d^k_{N,\Phi,m}$ *indicates a value of a contribution from another option to select for a problem in state* $(N, m, k-1, \Phi)$. *Lemma 6 yields that* $d^k_{N,m,\Phi}$ *is a decreasing sequence of* k. *This implies that the value of a contribution is, therefore, decreasing with respect to* k. *Since Equation (20) implies* $v_{N,m,k,\Phi} = \sum_{i=1}^{k} d^i_{N,\Phi,m}$ *and* $v_{N,m,k,\Phi}(n) = \sum_{i=1}^{k} d^i_{N,\Phi,m}(n)$, *these values do not necessary decreases with respect to* N. *If the number of observations is known and even if the number of periods increases, one cannot obtain monotone results as before. On the other hand, Theorem 4 shows that* $\{v_{N,m,k,\Phi}\}$ *and* $\{v_{N,m,k,\Phi}(n)\}$ *are decreasing with respect to* Φ.

3. A Sequential Stochastic Assignment Problem

Let's consider a partially observable Markov chain treated in Section 2.2. There are m jobs, and each one of them appears uniformly and independently during N time periods. Associated with each job is a random variable which depends only on a state of this chain. Regarding the number of jobs at the initial time period, the probability $\{p_{N,m}(n)\}_{n=0,1,\cdots,m}$ is given by Equation (1). On the other hand, there are m persons to be assigned to these m jobs. The abilities of these m persons are indicated as $p_1, p_2, \cdots, p_m$. We assume $1 \geq p_1 \geq p_2 \geq \cdots \geq p_m \geq 0$ without loss of generality. Associated with the n appeared jobs are independent and identically distributed random variables. When a state of this chain is s, these random variables indicate the values of jobs and depend on s. If a person with an ability p is assigned to a job with x, a reward px is earned. We also assume that a person who is assigned to some job is not available for future decisions. This problem is to assign m persons to m appearing jobs, in order to maximize the total expected reward.

As usual, there exists threshold values independent to $p_1, p_2 \cdots, p_m$, which are closely related to the optimal policy and essential properties of this problem. Let m and m' be numbers of jobs and persons, respectively. We assume here that m' and m are equal. If $m < m'$, then it is sufficient to assign the m largest p, i.e., $\{p_1, \cdots, p_m\}$ to m jobs. On the other hand, if $m > m'$, then it is sufficient to add $m - m'$ trivial values $p_{m'+1} = \cdots = p_m = 0$.

Consider that there are m persons with abilities' $\{p_1, \cdots, p_m\}$ for m jobs during N time periods. We treat this problem is in state $(N, \Phi; p_1, \cdots, p_m)$. We also indicate this problem as $P_{N,\Phi}(p_1, \cdots, p_m)$. When a state of this problem is $(N, \Phi; p_1, \cdots, p_m)$ and there are n jobs at the initial time period, then we use the notation $P_{N;p_1,\cdots,p_m}(n)$ for this subproblem. When there are n jobs with $x_{(1)}, \cdots, x_{(n)}$, we also consider a subproblem $P_{N,\Phi;p_1,\cdots,p_m}(n; x_{(1)}, \cdots, x_{(n)})$. The objective of this problem is to maximize the total expected reward by assigning m persons with $\{p_1, \cdots, p_m\}$ to m appearing jobs.

For the problems $P_{N,\Phi;p_1,\cdots,p_m}$, $P_{N,\Phi;p_1,\cdots,p_m}(n)$ and $P_{N,\Phi;p_1,\cdots,p_m}(n; x_{(1)}, \cdots, x_{(n)})$, let $v_{N,\Phi;p_1,\cdots,p_m}$, $v_{N,\Phi;p_1,\cdots,p_m}(n)$ and $v_{N,\Phi;p_1,\cdots,p_m}(n; x_{(1)}, \cdots, x_{(n)})$ be total expected rewards obtainable under the optimal policy, respectively. As in Ross (1970), the principle of optimality implies the following optimality equations.

$$v_{N,\Phi;p_1,\cdots,p_m} = \sum_{n=0}^{m} v_{N,\Phi;p_1,\cdots,p_m}(n)p_N(n) \tag{21}$$

$$v_{N,\Phi;p_1,\cdots,p_m}(n) = \mathrm{E}\left[v_{N,\Phi;p_1,\cdots,p_m}(n;X_{(1)},\cdots,X_{(n)})\right] \tag{22}$$

$$v_{N,\Phi;p_1,\cdots,p_m}\left(n;x_{(1)},\cdots,x_{(n)}\right) = \max_{\{\bar{p}_1,\cdots,\bar{p}_n\}\subset\{p_1,\cdots,p_m\}} \max_{\sigma\in\mathcal{S}_n} \left\{ \sum_{j=1}^{n} \bar{p}_{\sigma(j)}x_{(j)} + v_{N-1,T(\overline{\Phi},\boldsymbol{x});p_1^*,\cdots,p_{m-n}^*} \right\}, \tag{23}$$

where $\{p_1^*,\cdots,p_{m-n}^*\}$ are the remaining $m-n$ persons except to n persons, $\{\bar{p}_1,\cdots,\bar{p}_n\}$, out of m p's, $\{p_1,\cdots,p_m\}$ $(p_1^*\geq\cdots\geq p_{m-n}^*, \bar{p}_1\geq\cdots\geq\bar{p}_n)$. The optimal policy and the total expected reward obtainable under this policy are determined by threshold values as usual. These threshold values are the same ones obtained for the optimal selection problem in the last section, i.e., $\{d_{N,\Phi,m}^i\}$ and $\{d_{N,\Phi,m}^i(n)\}$. The proof of the following two theorems are complicated, and we only sketch the outlines of the proofs.

Theorem 5 *For the sequential stochastic assignment problem $P_{N,\Phi;p_1,\cdots,p_m}$, the optimal policy is as follows. "When the problem is in state $(N,\ \Phi;p_1,\cdots,p_m)$, n jobs appear with $x_{(1)},\cdots,x_{(n)}$. Let $\{b_j\}_{j=1,2,\cdots,m}$ be a rearranged sequence of two sequences $\{x_{(i)}\}_{i=1,\cdots,n}$ and $\{d_{N-1,\Phi,m-n}^i\}_{i=1,\cdots,m-n}$ from the largest to the smallest. If $b_j = x_{(i)}$, $(j=1,\cdots,m, i=1,\cdots,n)$ then it is optimal to assign the j-th p_j to a job with $x_{(i)}$. If $b_j = d_{N-1,\Phi,m-n}^i$, $(j=1,\cdots,m, i=1,\cdots,m-n)$ then it is optimal not to assign this p_j in this time."*

Theorem 6 *The total expected reward $v_{N,\Phi;p_1,\cdots,p_m}$ and $v_{N,\Phi;p_1,\cdots,p_m}(n)$ obtainable under the optimal policy for problem in state $(N,\ \Phi;p_1,\cdots,p_m)$ are as follows.*

$$v_{N,\Phi;p_1,\cdots,p_m} = \sum_{i=1}^{m} p_i d_{N,\Phi,m}^i, \qquad v_{N,\Phi;p_1,\cdots,p_m}(n) = \sum_{i=1}^{m} p_i d_{N,\Phi,m}^i(n). \tag{24}$$

In order to prove Theorems 5 and 6, we employ the induction principle on N. When $N=1$, Theorem 5 are obvious since

$$\max_{\sigma\in\mathcal{S}_n}\left\{\sum_{j=1}^{n}\bar{p}_{\sigma(j)}x_{(j)}\right\} = \sum_{j=1}^{n} p_j x_{(j)}, \tag{25}$$

which is known as Hardy's lemma. On the other hand,

$$v_{1,\Phi;p_1,\cdots,p_m} = \sum_{i=1}^{m} p_i E\left[X_{(i)}\right] = \sum_{i=1}^{m} p_i d_{1,\Phi,m}^i$$

implies Theorem 6. We assume that they are true for any value less than $n-1$.

Proof of Theorem 5

Since the induction assumptions imply $v_{N-1;p_1^*,\cdots,p_{m-n}^*} = \sum_{i=1}^{m-n} p_i^* d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}$, we have

$$\max_{\sigma\in\mathcal{S}_n}\left\{\sum_{j=1}^{n}\bar{p}_{\sigma(j)}x_{(j)} + v_{N-1,T(\overline{\Phi},\boldsymbol{x});p_1^*,\cdots,p_{m-n}^*}\right\}$$
$$= \max_{\sigma\in\mathcal{S}_n}\left\{\sum_{j=1}^{n}\bar{p}_{\sigma(j)}x_{(j)} + \sum_{i=1}^{m-n} p_i^* d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}\right\}$$
$$= \max_{\sigma\in\mathcal{S}_n}\left\{\sum_{j=1}^{n}\bar{p}_{\sigma(j)}x_{(j)}\right\} + \sum_{i=1}^{m-n} p_i^* d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}. \quad (26)$$

Equation (25) yields

$$\max_{\sigma\in\mathcal{S}_n}\left\{\sum_{j=1}^{n}\bar{p}_{\sigma(j)}x_{(j)}\right\} + \sum_{i=1}^{m-n} p_i^* d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} = \sum_{j=1}^{n}\bar{p}_j x_{(j)} + \sum_{i=1}^{m-n} p_i^* d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}.$$

If an individual decides to assign n $\bar{p}$'s ($\{\bar{p}_1,\cdots,\bar{p}_n\}\subset\{p_1,\cdots,p_m\}$), it is optimal to assign the j-th largest $\bar{p}_j$ to a job with $x_{(j)}$. This implies

$$v_{N,\Phi;p_1,\cdots,p_m}(n;x_{(1)},\cdots,x_{(n)})$$
$$= \max_{\{\bar{p}_1,\cdots,\bar{p}_n\}\subset\{p_1,\cdots,p_m\}}\max_{\sigma\in\mathcal{S}_n}\left\{\sum_{j=1}^{n}\bar{p}_{\sigma(j)}x_{(j)} + v_{N-1,T(\overline{\Phi},\boldsymbol{x});p_1^*,\cdots,p_{m-n}^*}\right\}$$
$$= \max_{\{\bar{p}_1,\cdots,\bar{p}_n\}\subset\{p_1,\cdots,p_m\}}\left\{\sum_{j=1}^{n}\bar{p}_j x_{(j)} + \sum_{i=1}^{m-n} p_i^* d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}\right\}. \quad (27)$$

In this point, the problem is how to assign the n $\{x_{(i)}\}_{i=1,\cdots,n}$ and the $m-n$ $\{d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}\}_{i=1,\cdots,m-n}$ to m p's, i.e., $\{p_1,\cdots,p_m\}$. Let $\{b_j\}_{j=1,2,\cdots,m}$ be a rearranged sequence of two sequences $\{x_{(i)}\}_{i=1,\cdots,n}$ and $\{d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}\}_{i=1,\cdots,m-n}$ from the largest to the smallest, then Equation (27) equals to $\max_{\sigma\in\mathcal{S}_n}\left\{\sum_{i=1}^{m}p_i b_{\sigma(i)}\right\}$. Since $b_1\ge b_2\ge\cdots\ge b_m$, Equation (25) implies

$$\max_{\sigma\in\mathcal{S}_n}\left\{\sum_{i=1}^{m}p_i b_{\sigma(i)}\right\} = \sum_{i=1}^{m}p_i b_i.$$

If $b_j = x_{(i)}$ $(j=1,\cdots,m, i=1,\cdots,n)$, then it is optimal to assign $x_{(i)}$ to the j-th p_j. If $b_j = d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}$ $(j=1,\cdots,m,\ i=1,\cdots,m-n)$, then it is optimal not to assign at this time period. □

Proof of Theorem 6.

By the induction principle, we will prove that $d^i_{N,\Phi,m}$ is the expectation (under the optimal policy) to which the i-th p_i is assigned in the problem in state $(N, \Phi; p_1, \cdots, p_m)$ $(i \leq m)$. From this, it will be shown obviously that the total expected reward is given by Equation (24). Similarly, $d^i_{N,\Phi,m}(n)$ is also the expectation to which the i-th p_i is assigned in the same problem when n jobs appear for decisions.

We note that $d^i_{N,\Phi,m}(n) = U_n(d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}, d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}|1,\infty)$. Since

$$\begin{aligned}\alpha_{1,n} &= U_n(d_1,\infty|1,\infty)\\ &= \int_0^{d_1} d_1 h_{n,1}\left(x_{(1)}\right) f\left(x_{(1)}\right) dx_{(1)} + \int_{d_1}^{\infty} x_{(1)} h_{n,1}\left(x_{(1)}\right) f\left(x_{(1)}\right) dx_{(1)}, \end{aligned} \tag{28}$$

then

$$\begin{aligned} d^1_{N,\Phi,m} &= U_n\left(d^1_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n},\infty \middle| 1,\infty\right)\\ &= \int_0^{d^1_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}} d^1_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} h_{n,1}\left(x_{(1)}\right) f\left(x_{(1)}\right) dx_{(1)}\\ &\quad + \int_{d^1_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}}^{\infty} x_{(1)} h_{n,1}\left(x_{(1)}\right) f\left(x_{(1)}\right) dx_{(1)}. \end{aligned} \tag{29}$$

This value is an expectation to which the largest p_1 is assigned under the optimal policy. This comes from the fact that; if $x_{(1)} > d^1_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}$, then it is optimal to assign p_1, and if, otherwise, it is optimal to decide not to assign p_1 at this time period. This result follows by conditioning on the initial $x_{(1)}$ recalling that p_1 is used if and only if this value lie in the interval $(d^1_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n},\infty)$. The second term of Equation (28) corresponds to the first case, and the first term is the expectation to which the p_i is assigned for the future decisions under the optimal policy.

For the general cases, since

$$\begin{aligned}\alpha_{i,n} &= U_n(d_i,d_{i-1}|1,\infty)\\ &= \int_0^{d_i} d_i h_{n,1}\left(x_{(1)}\right) f\left(x_{(1)}\right) dx_{(1)} + \int_{d_i}^{d_{i-1}} x_{(1)} h_{n,1}\left(x_{(1)}\right) f\left(x_{(1)}\right) dx_{(1)}\\ &\quad + \int_{d_{i-1}}^{y} U_n\left(d_{i-1},d_{i-2}\left|2,x_{(1)}\right.\right) f\left(x_{(1)}\right) dx_{(1)}, \end{aligned} \tag{30}$$

we investigate about these three terms. Equation (30) implies

$$\begin{aligned} d^i_{N,\Phi,m}(n) &= U_n\left(d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}, d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} \middle| 1,\infty\right)\\ &= \int_0^{d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}} d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} h_{n,1}\left(x_{(1)}\right) f\left(x_{(1)}\right) dx_{(1)}\\ &\quad + \int_{d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}}^{d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}} x_{(1)} h_{n,1}\left(x_{(1)}\right) f\left(x_{(1)}\right) dx_{(1)}\\ &\quad + \int_{d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}}^{y} U_n\left(d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}, d^{i-2}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} \middle| 2,x_{(1)}\right) f\left(x_{(1)}\right) dx_{(1)}. \end{aligned} \tag{31}$$

If $x_{(1)} < d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}$, then not only p_i but also $p_1, \cdots, p_{i-1}$ are not assigned at this time period. At the next instant, p_i will be also the i-th largest among the remaining $m-n$ p's. This implies that the expectation (under the optimal policy) which will be assigned to p_i is $d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}$ from the future decisions. This case corresponds to the first term of Equation (31).

If $d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} \le x_{(1)} < d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}$, then $p_1, \cdots, p_{i-1}$ are not assigned at this time period, but the i-th p_i is assigned to this job. This case corresponds to the second term of Equation (31).

If $d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} \le x_{(1)}$, then p_i is not assigned to a job with $x_{(1)}$. Since $d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} \le x_{(1)}$, one of the $p_1, \cdots, p_{i-1}$ is assigned to this job under the optimal policy. In this case, an individual decides for a job with $x_{(2)}$. There will consider the following three cases, i.e., (1) $x_{(2)} < d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}$, (2) $d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} \le x_{(2)} < d^{i-2}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}$ and (3) $d^{i-2}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n} \le x_{(2)}$. These cases are similar classification for $x_{(1)}$. The third term of Equation (31) corresponds to this case.

This result follows by conditioning on the initial $x_{(1)}$ recalling that p_i is used if and only if this value lies in the interval $(d^i_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}, d^{i-1}_{N-1,T(\overline{\Phi},\boldsymbol{x}),m-n}]$. The expectation to which the p_i is assigned under the optimal policy is, therefore, equal to Equation (31), and this completes the proofs. □

Remark 2 *If we assume $p_1 = \cdots = p_k = 1$ and $p_{k+1} = \cdots = p_m = 0$ in Theorem 6, then we obtain a similar result in Theorem 4. From this fact, the sequential stochastic assignment problem is a generalization of an optimal selection problem in the last section. As is shown in Theorem 5, $d^i_{N,\Phi,k}(m)$ means a threshold value for the optimal policy. Theorems 5, 6 and Lemma 6 indicate the properties of the optimal policy. For example, $d^i_{N,\Phi,k}(m)$ is decreasing with respect to Φ. If a state corresponds to an economic condition, as information about this state becomes smaller, the economic condition improves. From this fact, we say that; as information about the state improves to better, the threshold value increases. We also consider that; as the number of random variables increases, the threshold value increases. It is, therefore, natural that $d^i_{N,\Phi,k}(m)$ is increasing with respect to m. This can be shown by using the induction principle.*

Remark 3 *In this paper, we only treat a case where the total number of jobs is previously known, but it is possible to show similar results when this number is not known previously. We can obtain a proof for this case by a method similar to one used in Nakai (1986a).*

Acknowledgment The author would like to thank the referees for their helpful comments.

References

Albright, S. C.: A Markov-Decision-Chain Approach to a Stochastic Assignment Problem, *Operations Research*, **22** (1974), 61–64.

Brown, M. and Solomon, H.: Optimal Issuing Policies under Stochastic Field Lives, *Journal of Applied Probability*, **10** (1973), 761–768.

Derman C., Lieberman, G. J. and Ross, S. M.: A Sequential Stochastic Assignment Problem, *Management Science*, **18** (1972), 349–355.

Nakai, T.: The Problem of Optimal Stopping in a Partially Observable Markov Chain, *Journal of Optimization Theory and Applications*, **45** (1985), 425–442.

Nakai, T.: Optimal Selection Problem with a Random Number of Applicants per Period, *Operations Research*, **34** (1986), 478–485.

Nakai, T.: A Sequential Stochastic Assignment Problem in a Partially Observable Markov Chain, *Mathematics of Operations Research*, **11** (1986), 230–240.

Nakai, T.: A Stochastic Ordering and Related Sequential Decision Problems, *Journal of Information & Optimization Sciences*, **11** (1990), 49–65.

Nakai, T.: A Partially Observable Decision Problem under a Shifted Likelihood Ratio Ordering, *Mathematical and Computer Modelling*, **22** (1995), 237–246.

Nakai, T.: An Optimal Selection Problem on a Partially Observable Markov Chain, *Stochastic Modelling in Innovative Manufacturing* (Eds. Christer, A. H., Osaki, S. and Thomas, L. C.), Lecture Notes in Economics and Mathematical Systems, **445**, Springer-Verlag, Berlin (1996), 140–154.

Pinedo, M. L. and Ross, S. M.: Scheduling Jobs Subject to Nonhomogeneous Poisson Shocks, *Management Science*, **26** (1980), 1250–1257.

Ross, S. M.: *Applied Probability Models with Optimization Applications*, Holden-Day, San Francisco, California (1970).

Ross, S. M.: *Stochastic Processes*, John-Wiley and Sons, New York, New York (1983).

Wilks, S. S.: *Mathematical Statistics*, John-Wiley & Sons, New York (1962).

Received June 28, 1998