Journal of the Operations Research Society of Japan
Vol. 25, No. 3, September 1982

MONOTONE OPTIMAL CONTROL OF ARRIVALS DISTINGUISHED BY REWARD AND SERVICE TIME

Shōichi Nishimura
University of Tsukuba

(Received September 6, 1980; Final April 15, 1982)

Abstract We consider a monotone optimal policy for a discrete time problem of controlling the arriving customers. At each period one customer arrives at a manufacturing factory to order a job distinguished by the reward and the service time with a constant delivery interval. The basic properties of optimal policies are obtained. It is shown that, contrary to intuition, from counterexamples an optimal policy cannot generally be monotone in such cases as finite-horizon problems with and without discounting and an infinite-horizon problem with discounting, while there exists a monotone optimal policy for infinite-horizon problems without discounting.

1. Introduction

In this paper we study the monotonicity of optimal policy for the following discrete time problem of controlling the arrival customers. Suppose that in each period one customer arrives at a manufacturing factory to order a job distinguished by the reward and the service time with the constant delivery interval. Let d be the delivery interval between the acceptance time and the completion time of the job. Let k be the random units of time of job length. Using queueing terminology we say k is the service time. And let r be the random reward received by the manufacturing factory if the customer is accepted. Let us assume that the joint distribution of (k, r) of successive customers are independent and identically distributed. At the decision epoch each job is distinguished by (k, r). Let i be the work backlog (not including the job to be just accepted). Simply we say i is the state of the system. The decision maker of the factory decides whether the arriving customer is to be accepted or rejected from i and (k, r). If the customer is accepted, then the next state is $i + k - 1$ with reward r and if the customer is rejected, then the next state is $i - 1$ for $i \neq 0$ and 0 for $i = 0$. The customer with the service time $k \geq d - i + 1$ has to be rejected.

Ikuta [1] formulated this model as "Order Selection Problem" when the service time k is constant. And he proved the monotonicity of the stationary optimal policy such that the arriving customer is less likely to enter as the state i increases. In our model the decision maker distinguishes not only the reward r but the service time k of arriving customers. The constraint is the completion of the service within the delivery interval. In controlled queueing systems arriving customers are classified only by the reward and their service time is not distinguished. Miller [4] obtained the monotonicity in $M/M/C$ finite capacity queue. Lippman [2] proved the monotonicity of optimal policies for finite-horizon problems with and without discounting allowing an infinite number of customer classes iteratively and extended the monotonic property to the infinite-horizon problems. Lippman and Stidham Jr. [3] considered exponential congestion systems including $M/M/1$ system as a special case and Stidham Jr. [8] considered $GI/M/1$ congestion systems with extensions of batch arrivals, Erlang service-time distribution and others. The cost structure in these models consists of the holding cost and the reward. They proved the monotonicity of optimal policies and compared socially and individually optimal joining policies.

In section 2, we formulate our model as Markov decision processes for finite-horizon problem with a discount factor α. The basic properties of optimal policies are obtained. Under the condition that the state of the system is i and the service time of the arriving customer is k, critical-numbers $v_{n,\alpha}(i, k)$ of the reward are inductively obtained on the horizon length n and are nonnegative. An optimal policy is given as follows: if $r \geq v_{n,\alpha}(i, k)$ then the arriving customer is accepted and if $r < v_{n,\alpha}(i, k)$ then he is rejected. The plausible question is whether for fixed α, n and k, $v_{n,\alpha}(i, k)$ is monotone nondecreasing in i. In other words, the problem is whether the customer who is accepted in the state i, is accepted in the state $i'(i' < i)$. We will answer this question negatively, however, by the first counterexample in Example 1.

In section 3, we consider infinite-horizon problems. Let the state of the system be (i, k, r) and actions be acceptance and rejection. Using the technique given by [2], [3] and [6] it is proved that a stationary optimal policy exists and $v_\alpha(i, k) = \lim_{n\to\infty} v_{n,\alpha}(i, k)$ is nonnegative. We also obtain in Example 2 a counterexample such that $v_\alpha(i, k)$ is not nondecreasing for infinite-horizon problem with discounting $0 < \alpha < 1$. These counterexamples come from the variable service time k. The author seems there is a relation between our examples and unsuspected phenomena that the optimal congestion toll cannot be monotonic as given in [3] and [6]. But the natural requirement that there exists a monotone optimal policy for infinite-horizon problem without discounting $\alpha = 1$, is proved in Theorem 3.3 and 3.4. From this result for a fixed k critical-number $v(i, k)$ is monotone nondecreasing in i and for a fixed i critical-number per service time $v(i, k)/k$ is monotone nondecreasing in k as given in Corollary 3.5.

2. Finite Horizon Model

In this section we consider the following discrete time Markov decision process. Suppose that at each period one customer arrives at a factory to order a job distinguished by the reward r and the service time k. The r is a random reward received by a factory if the customer is accepted but there is no rejection cost. Assume that each customer has a constant delivery interval d, which means that the job should be completed before d units of time elapse from his arriving time. Let i be the work backlog at a decision epoch not including the job to be accepted (i is the state of the system). The system is controlled by accepting or rejecting arriving customers upon observing the state i and (k, r).

Let K and R denote the random variables of the service time and the reward respectively. Assume that $K \in \{0, \ldots, d\}$, where $K = 0$ is the event that no customer arrives. Let

(1) $$P(k) = P\{K = k\} \qquad (k = 0, \ldots, d)$$

and to simplify the notation we put

(2) $$F_k(r) = \begin{cases} P\{R \leq r \mid K = k\} & \text{if } p(k) > 0 \\ \begin{cases} 0 & r < 0 \\ 1 & r \geqq 0 \end{cases} & \text{if } p(k) = 0 \text{ or } k = 0. \end{cases}$$

Assume that

(3) $$\begin{aligned} 0 \leq m_k &= E[R^+ \mid K = k] = \int_0^\infty (1 - F_k(r))dr < \infty \\ 0 \leq m &= E[R^+] = \sum_{k=1}^{d} p(k)m_k < \infty, \end{aligned}$$

where R^+ is the positive part of R. Random variables (K, R) of successive customers are independent and identically distributed and their joint distributions are given by (1) and (2).

Suppose that at a decision epoch the state of the system is i ($i = 0, \ldots, d-1$) and the arriving customer is characterized by (k, r). If the customer ($k \leq d - i$) is accepted, then the next state is $i + k - 1$ and if $k \geqq d - i + 1$ then he cannot be accepted from the constraint of the constant delivery interval d. If he is rejected, then the next state is $i - 1$ but if $i = 0$, the next state is also 0. Let $B_{i,k}$ be the set of rewards r which is accepted when the state is i and the service time is k and put $B_i = B_{i,1} \times \ldots \times B_{i,d-i}$, which is an action at the state i.

Let $V_{n,\alpha}(i)$ be the maximal expected α-discounted ($0 < \alpha \leq 1$) return with the initial state i when the horizon length is n. The $V_{n,\alpha}(i)$ satisfy the following recursive equations:

$$(4)\quad \begin{aligned} V_{n,\alpha}(i) &= \sup_{B_i} \{ \sum_{k=1}^{d-i} p(k) \int_{B_{i,k}} (r + \alpha V_{n-1,\alpha}(i + k - 1)) dF_k(r) \\ &\quad + (1 - \sum_{k=1}^{d-i} p(k) \int_{B_{i,k}} dF_k(r))\, \alpha V_{n-1,\alpha}(i - 1)\} \\ &= \alpha V_{n-1,\alpha}(i - 1) \\ &\quad + \sum_{k=1}^{d-i} p(k) \sup_{B_{i,k}} \{ \int_{B_{i,k}} (r - \alpha(V_{n-1,\alpha}(i - 1) - V_{n-1,\alpha}(i+k-1))) dF_k(r)\}, \end{aligned}$$

where $V_{n,\alpha}(-1) = V_{n,\alpha}(0)$.

If $F(r)$ is a distribution function with finite mean, we have

$$\max_B \int_B (r - x) dF(r) = \int_x^\infty (r - x) dF(r) = \int_x^\infty (1 - F(r)) dr.$$

Then

$$(5)\qquad B_{i,k} = \{r;\ r \geqq v_{n,\alpha}(i, k)\}$$

attains the supremum in (4), where $v_{n,\alpha}(i, k) = \alpha(V_{n-1,\alpha}(i - 1) - V_{n-1,\alpha}(i + k - 1))$ is the critical reward r under given n, α and (i, k).

We rewrite (4) as

$$(6)\qquad V_{n,\alpha}(i) = \alpha V_{n-1,\alpha}(i - 1) + \sum_{k=1}^{d-i} p(k) \int_{v_{n,\alpha}(i,\ k)}^\infty (1 - F_k(r)) dr.$$

Let us put $V_{0,\alpha}(i) = 0$ ($i = 1, \ldots, d-1$) then from (6) we recursively obtain $V_{n,\alpha}(i)$ with $V_{n,\alpha}(-1) = V_{n,\alpha}(0)$.

Theorem 2.1 $V_{n,\alpha}(i)$ is non-increasing in i, so that $v_{n,\alpha}(i, k)$ is non-negative for $i = 0, .., d - 1, k = 0, ..., d - i$.

Proof: The proof is given by induction on n. From the initial condition $v_{1,\alpha}(i, k) = 0$. Suppose that $V_{n-1,\alpha}(i)$ is non-increasing in i then $v_{n,\alpha}(i, k) \geq 0$. Then from (6) we obtain

$$V_{n,\alpha}(i) - V_{n,\alpha}(i + 1) = \alpha(V_{n-1,\alpha}(i - 1) - V_{n-1,\alpha}(i))$$

$$(7) \qquad + \sum_{k=1}^{d-i-1} p(k) \int_{v_{n,\alpha}(i, k)}^{v_{n,\alpha}(i + 1, k)} (1 - F_k(r))dr$$

$$+ p(d - i) \int_{v_{n,\alpha}(i, d - i)}^{\infty} (1- F_{d-i}(r))dr$$

The third term on the right hand side in (7) is nonnegative. Put the set $A_i = \{k; v_{n,\alpha}(i, k) > v_{n,\alpha}(i + 1, k), k = 1, ..., d - i\}$.

From $\quad 0 \leq 1 - F_k(r) \leq 1, \; v_{n,\alpha}(i, 1) = \alpha\{V_{n-1,\alpha}(i - 1) - V_{n-1,\alpha}(i)\}$

and $\quad v_{n,\alpha}(i, k) - v_{n,\alpha}(i + 1, k) = v_{n,\alpha}(i, 1) - v_{n,\alpha}(i + k, 1)$ we have

$$\begin{aligned} V_{n,\alpha}(i) - V_{n,\alpha}(i + 1) &\geq v_{n,\alpha}(i, 1) \\ &- \sum_{k\epsilon A_i} p(k)(v_{n,\alpha}(i, k) - v_{n,\alpha}(i + 1, k) \\ &= v_{n,\alpha}(i, 1) - \sum_{k\epsilon A_i} p(k)(v_{n,\alpha}(i, 1) - v_{n,\alpha}(i + k, 1)) \\ &\geq (1 - \sum_{k\epsilon A_i} p(k)) \, v_{n,\alpha}(i, 1) \geq 0. \end{aligned}$$

The proof is completed.

From Theorem 2.1. negative reward customers are rejected in optimal policies. The next plausible question is that for fixed α, n and k, $v_{n,\alpha}(i, k)$ is monotone nondecreasing in i. In other words, the problem is whether the customer who is accepted in state i, is accepted in state $i'(i' < i)$. We will answer this question negatively by the first counterexample.

Example 1. Put $d = 3$, $p(3) = 0$, $F_1(r) = F_2(r) = 1 - e^{-r}(r > 0)$, then $m_1 = m_2 = 1$. And as the initial condition $V_{0,\alpha}(i) = 0$ $(i = 0, 1, 2)$. From (6) we recursively obtain

$$V_{1,\alpha}(0) = V_{1,\alpha}(1) = p(1) + p(2), \quad V_{1,\alpha}(2) = p(1)$$

$$V_{2,\alpha}(0) = (1 + \alpha)(p(1) + p(2)),$$

$$V_{2,\alpha}(1) = \alpha(p(1) + p(2)) + p(1) + p(2)e^{-\alpha p(2)}$$
$$V_{2,\alpha}(2) = \alpha(p(1) + p(2)) + p(1)e^{-\alpha p(2)}.$$

Then we have

$$v_{3,\alpha}(2, 1) - v_{3,\alpha}(1, 1) = \alpha\{-V_{2,\alpha}(2) + 2V_{2,\alpha}(1) - V_{2,\alpha}(0)\}$$
$$= \alpha\{p(2)(2e^{-\alpha p(2)} - 1) + p(1)(1 - e^{-\alpha p(2)})\} < 0 \quad (0 < \alpha \leq 1),$$

when $p(2)$ is sufficiently close to 1 and $p(1)$ is close to 0. Then in the case of $n = 3$ and $k = 1$ customers whose reward r is $v_{3,\alpha}(2, 1) < r < v_{3,\alpha}(1, 1)$ are rejected at the state $i = 1$ and are accepted at $i = 2$.

Ikuta [1] proved that if K is a constant the critical number of reward $v_{n,\alpha}(i, k)$ is monotone nondecreasing in i by induction. The monotonicity of $v_{n-1,\alpha}(i, k)$ implies the monotonicity of $v_{n,\alpha}(i, k)$. In queueing control problems monotone optimal policies are proved inductively ([2], [3] and [6]) and many other decision problems has this property (see, for example, [1]). This counter example shows that we cannot use the induction to prove the concavity of $V_{n,\alpha}(i)$.

3. Infinite Horizon

In this section we consider optimal policies for infinite-horizon problems. We will prove there exists a monotone stationary optimal policy without discounting ($\alpha = 1$) which maximizes the long-run average expected return per unit time. However, we will obtain a counterexample in Example 2, where this monotonicity of a stationary optimal policy is not satisfied for a discounting problem.

Our original model consists of the infinite action space. We reformulate it to the model with finite actions using the technique (e.g., by Lippman [2], Lippman and Stidham Jr. [3] and Stidham Jr. [6]). They proved the existence of stationary optimal policies for infinite-horizon controlled queueing problems both with and without discounting. Let the state of the system be (i, k, r) at the arriving time of customers where i is the waiting time not including the work seeking admittance, k is the customer's service time and r is the reward. There are two possible actions: accept ($a = 1$) or reject ($a = 0$). Let $V_{n,\alpha}(i, k, r)$ be the maximal expected α-discounted return for n length problem when the initial state is (i, k, r). The functions $V_{n,\alpha}(i, k, r)$ satisfy the following recursive equations:

$$(8) \qquad V_{n+1,\alpha}(i, k, r) = \begin{cases} \max_a \{ar + \alpha V_{n,\alpha}(i + ak - 1)\} & k \le d - i \\ \alpha V_{n,\alpha}(i - 1) & k > d - i \end{cases}$$

$$(9) \qquad V_{n,\alpha}(i) = \sum_{k=1}^{d} p(k) \int_{-\infty}^{\infty} V_{n,\alpha}(i, k, r)\, dF_k(r) \;,$$

where $V_{n,\alpha}(-1) = V_{n,\alpha}(0)$.

It is trivial from definition that $V_{n,\alpha}(i)$ given is the same as (6).

In the reformulated model the action space is finite and the expected return at each period is bounded above by $ER^+ = m < \infty$ by assumption. Using the same logic in Stidham Jr. [6] page 1605, there exists the stationary optimal policy for reformulated model and $\lim_{n\to\infty} V_{n,\alpha}(i, k, r) = V_\alpha(i, k, r)$. And then $\lim_{n\to\infty} V_{n,\alpha}(i) = V_\alpha(i) = \sup_\pi V_\alpha^\pi(i)$, where $V_\alpha^\pi(i)$ and $V_\alpha(i)$ are the expected return function for our original infinite-horizon problem with discounting $(0 < \alpha < 1)$ from a policy π and the optimal policy, respectively. Since $V_{n,\alpha}(i)$ is nonincreasing in i, then $V_\alpha(i)$ is nonincreasing in i. We have

Theorem 3.1 $V_\alpha(i)$ is nonincreasing in i and satisfies the following equation

$$V_\alpha(i) = \alpha V_\alpha(i - 1) + \sum_{k=1}^{d-i} p(k) \int_{v_\alpha(i,\, k)}^{\infty} (1 - F_k(r))dr \;,$$

where $V_\alpha(-1) = V_\alpha(0)$ and $v_\alpha(i, k) = \alpha(V_\alpha(i - 1) - V_\alpha(i + k - 1))$. Moreover, the stationary critical-number policy that the acception region of reward under (i, k) is $\{r\colon r \ge v_\alpha(i, k)\}$ is optimal among all policies.

The monotonicity of $v_\alpha(i, k)$ in i is not satisfied, in general, which is shown by a counterexample. Put $d = 3$ and $p(2) = 0$, then $V_\alpha(i)$ is given by

$$V_\alpha(0) = \alpha V_\alpha(0) + p(1)\int_0^\infty (1 - F_1(r))dr + p(3)\int_{\alpha(V_\alpha(0) - V_\alpha(2))}^{\infty} (1 - F_3(r))dr$$

$$(10) \qquad V_\alpha(1) = \alpha V_\alpha(0) + p(1)\int_{\alpha(V_\alpha(0) - V_\alpha(1))}^{\infty} (1 - F_1(r))dr$$

$$V_\alpha(2) = \alpha V_\alpha(1) + p(1)\int_{\alpha(V_\alpha(1) - V_\alpha(2))}^{\infty} (1 - F_1(r))dr \;.$$

If $p(3) = 1$ and $p(1) = 0$, then $V_\alpha(2) = \alpha^2 V_\alpha(0)$, $V_\alpha(1) = \alpha V_\alpha(0)$, $V_\alpha(0) > 0$ and $v_\alpha(2, 1) - v_\alpha(1, 1) = \alpha(V_\alpha(1) - V_\alpha(2)) - \alpha(V_\alpha(0) - V_\alpha(1))$
$= -\alpha(1 - \alpha)^2\, V_\alpha(0) < 0$ for $0 < \alpha < 1$.

In this case, however, there exists an optimal monotone policy because of $p(1) = 0$. Let us choose $p(1)$ and $p(3)$ sufficiently close to 0 and 1 respectively under $0 < p(1) + p(3) \leq 1$. We can make a counterexample such that the arriving customer of the service time $k = 1$ is accepted in the state $i = 2$, but is rejected in the state $i = 1$ ($v_\alpha(2, 1) < r < v_\alpha(1, 1)$) as follows:

Example 2. Put $\alpha = 0.5$, $P(R = 0.2 \mid K = 1) = P(R = 1 \mid K = 3) = 1$, $p(0) = 0.06$, $p(1) = 0.04$ and $p(3) = 0.9$. From the elementary calculation, the following $V_{0.5}(i)$ ($i = 0, 1, 2$) satisfies (10) with

$$\int_{0.5(V_{0.5}(0) - V_{0.5}(1))}^{\infty} (1 - F_1(r))dr = 0 : \quad V_{0.5}(0) = 10721.28/9876$$

$$\doteqdot 1.086,\ V_{0.5}(1) = 0.5\ V_{0.5}(0) = 0.543 \quad \text{and} \quad V_{0.5}(2) = 225.52/823 \doteqdot 0.274.$$

There are two customer classes: ($k = 1$, $r = 0.2$) and ($k = 3$, $r = 1$). Using contraction mapping fixed point theorem (see Ross [5] Corollary 6.6), $V_{0.5}(i)$ is the unique solution of (10) for $\alpha = 0.5$.
We obtain $v_{0.5}(1, 1) = 0.5(V_{0.5}(0) - V_{0.5}(1)) \doteqdot 0.2715$ and $v_{0.5}(2, 1) = 0.5(V_{0.5}(1) - V_{0.5}(2)) \doteqdot 0.1345$.
From $0.2715 > 0.2 > 0.1345$ the customer of $k = 1$ is accepted at $i = 2$ but rejected at $i = 1$.

We now turn our attention to the infinite-horizon problem without discounting $\alpha = 1$, in which the objective is to maximize long-run average expected return per unit time. For any bounded function $h(i)$,

$$\sup \{ \sum_{k=1}^{d-i} p(k) \int_{B_{i,k}} (r + h(i+k-1))dF_k(r) + (1 - \sum_{k=1}^{d-i} p(k) \int_{B_{i\ k}} dF_k(r))\ h(i-1)\}$$

is attained by $B_{i,k} = \{r;\ r \geq h(i - 1) - h(i + k - 1)\}$ as was shown in (5) and (6). Then, to prove the existence of a stationary optimal policy it is sufficient that $\mid V_\alpha(i - 1) - V_\alpha(i) \mid$ is bounded for all α and $i = 1, \ldots, d-1$ using Theorem 6.17 and 6.18 in Ross [5] or Theorem 2.11 and Corollary 2.13 in Stidham Jr. and Probhu [7]. We have

$$\mid V_\alpha(i-1) - V_\alpha(i) \mid \leq (1-\alpha)\ V_\alpha(i-1) + \sum_{k=1}^{d-i} p(k) \int_{v_\alpha(i,\ k)}^{\infty} (1 - F_k(r))dr$$

$$\leq (1 - \alpha) \frac{m}{1 - \alpha} + m = 2m \qquad (i = 1, \ldots, d-1)$$

where $m = ER^+$. We have

Theorem 3.2. There exist the nonincreasing bounded function $V(i)$ and the constant g such that for each i

$$V(i) = \lim_{\alpha \to 1^-} (V_\alpha(i) - V_\alpha(0))$$

is well defined and satisfies the functional equation

$$(11) \qquad g + V(i) = V(i-1) + \sum_{k=1}^{d-i} p(k) \int_{v(i,\, k)}^{\infty} (1 - F_k(r))dr.$$

where g is the maximal long-run average expected return per unit time, $V(-1) = V(0)$ and $v(i, k) = V(i-1) - V(i+k-1)$. Moreover, the stationary critical-number policy that the acception region of reward under (i, k) is $\{r: r \geq v(i, k)\}$ is optimal among all policies.

We are now in the position to prove the monotonicity of the optimal stationary policy for infinite-horizon problems without discounting ($\alpha = 1$). We first treat the case of $0 < \sum_{k=1}^{d} p(k) < 1$ because the case of $\sum_{k=1}^{d} p(k) = 1$ is complicated.

Theorem 3.3. Suppose that $0 < \sum_{k=1}^{d} p(k) < 1$, then $v(i, 1)$ is nonnegative and nondecreasing in i, so that $V(i)$ is nonincreasing and concave.

Proof: From Theorem 3.2 $v(i, 1) = V(i-1) - V(i)$ is nonnegative. We will prove $v(i, 1)$ is nondecreasing in i using reduction to absurdity. From (11) we have

$$(12) \qquad g - v(i, 1) = \sum_{k=1}^{d-i} p(k) \int_{v(i,\, k)}^{\infty} (1 - F_k(r))dr$$

and then

$$v(i, 1) - v(i-1, 1) = \sum_{k=1}^{d-i} p(k) \int_{v(i-1,\, k)}^{v(i,\, k)} (1 - F_k(r))dr$$

$$(13) \qquad + p(d-i+1) \int_{v(i-1,\, d-i+1)}^{\infty} (1 - F_{d-i+1}(r))dr$$

$$\geq \sum_{k=1}^{d-i} p(k) \int_{v(i-1,\, k)}^{v(i,\, k)} (1 - F_k(r))dr.$$

Suppose that for some j $(0 \leq j \leq d-1)$

$$(14) \qquad v(d-1, 1) \geqq \ldots \geqq v(j, 1)$$

and

$$(15) \qquad v(j, 1) < v(j-1, 1)$$

Put the set $A = \{k;\ v(j, k) < v(j-1, k),\ p(k) > 0,\ k = 1, \ldots, d-j\}$. Using the equation $v(j-1, k) - v(j, k) = v(j-1, 1) - v(j+k-1, 1)$ and $0 \leq F_k(r) \leq 1$ we have

$$v(j, 1) - v(j-1, 1) \geqq -\sum_{k\varepsilon A} p(k) \int_{v(j, k)}^{v(j-1, k)} (1 - F_k(r))dr$$

(16)
$$\geqq -\sum_{k\varepsilon A} p(k)\ (v(j-1, k) - v(j, k))$$
$$= -\sum_{k\varepsilon A} p(k)\ (v(j-1, 1) - v(j+k-1, 1)$$
$$= -v(j-1, 1) \sum_{k\varepsilon A} p(k) + \sum_{k\varepsilon A} p(k)\ v(j+k-1, 1).$$

From (14) we have

(17) $$v(j, 1) - v(j-1, 1) \geqq (v(j, 1) - v(j-1, 1)) \sum_{k\varepsilon A} p(k)$$
$$> v(j, 1) - v(j-1, 1),$$

where the last inequality is derived from $0 < \sum_{k=1}^{d} p(k) < 1$ and (15). This contradiction in (17) comes from (15), then $v(i, 1)$ is nondecreasing in i and the proof is completed.

Next we treat the case $\sum_{k=1}^{d} p(k) = 1$.

Theorem 3.4. Suppose $\sum_{k=1}^{d} p(k) = 1$, then there exists a nonnegative and nondecreasing function $v(i, 1)$ satisfying (12). That is, there exists a non-increasing and concave function $V(i)$ satisfying (11).

Proof: Let $U(i)$ and $u(i, 1) = U(i-1) - U(i)$ satisfy (11) or (12). Suppose that for some j $(0 \leq j \leq d-1)$

(18) $$u(d-1, 1) \geqq \ldots \geqq u(j, 1)$$

and

(19) $$u(j, 1) < u(j-1, 1).$$

If one of inequalities (13), (16) and (17) is strict inequality, then we can derive the contradiction as in the proof of Theorem 3.3.
Then it is necessary that

(20) $$\sum_{k=1}^{d-j} p(k) = 1$$

and

$$(21)\qquad F_k(u(j-1,\ k)) = F_k(u(j,\ k)) = 0 \quad \text{for} \quad p(k) > 0$$

and using the monotonicity of $u(i,\ 1)$ in (18) and $u(j+k,\ 1) = u(j,\ 1)$ for $k \in A$ in (16)

$$(22)\qquad u(j+k,\ 1) = u(j,\ 1) \quad \text{for} \quad 1 \le k \le \max\{k\colon p(k) > 0\}.$$

Now put $V(i)$ and $v(i,\ k)$ as follows:

$$V(d-1) = U(d-1)$$

$$(23)\qquad v(i,\ 1) = \begin{cases} u(i,\ 1) & j \le i \le d-1 \\ u(j,\ 1) & 1 \le i < j \\ 0 & i = 0 \end{cases}$$

$$v(i,\ k) = V(i-1) - V(i+k-1) \qquad (i=1,\ \ldots,\ d-1,\ k=1,\ \ldots,\ d-i)$$

$$v(0,\ k) = V(0) - V(k-1) \qquad (k=1,\ \ldots,\ d)$$

From the definition $v(i,\ 1)$ is nonnegative and nondecreasing in i because $u(i,\ 1)$ is nondecreasing in $j \le i \le d-1$ from (18) and $u(j,\ 1)$ is nonnegative. The proof will be completed if we prove

$$(24)\qquad g = v(i,\ 1) + \sum_{k=1}^{d-i} p(k) \int_{v(i,\ k)}^{\infty} (1 - F_k(r))dr$$

for $i = 0,\ \ldots,\ j-1$. In the cases of $i = 1,\ \ldots,\ j-1$ we have $v(i,\ 1) = u(j,\ 1)$, $v(i,\ k) = ku(j,\ 1) = u(j,\ k)$ then

$$(25)\qquad \begin{aligned} g &= u(j,\ 1) + \sum_{k=1}^{d-j} p(k) \int_{u(j,\ k)}^{\infty} (1 - F_k(r))dr \\ &= v(i,\ 1) + \sum_{k=1}^{d-j} p(k) \int_{v(i,\ k)}^{\infty} (1 - F_k(r))dr \quad (i = 1,\ \ldots,\ j-1). \end{aligned}$$

For $i = 1,\ \ldots,\ j-1$ the equation (24) is proved. Moreover, from (21) $F_k(v(1,\ k)) = F_k(u(j,\ k)) = 0$ then for $i = 1$ in (25) we have

$$(26)\qquad \begin{aligned} g &= v(1,\ 1) + \sum_{k=1}^{d-j} p(k) \int_{v(1,\ k)}^{\infty} (1 - F_k(r))dr \\ &= v(1,\ 1) + \sum_{k=1}^{d-j} p(k) \int_{0}^{\infty} (1 - F_k(r))dr - \sum_{k=1}^{d-j} p(k)\ v(1,\ k). \end{aligned}$$

In the case of $i = 0$, we have $v(0, k) = V(0) - V(k - 1) + V(k) - V(k) = V(0) - V(k) - V(0) + V(1) = v(1, k) - v(1, 1) \leq v(1, k)$ and $F_k(v(0, k)) = F_k(v(1, k)) = F_k(u(j, k)) = 0$ then from (26)

$$v(0, 1) + \sum_{k=1}^{d-j} p(k) \int_{v(0, k)}^{\infty} (1 - F_k(r))dr$$

$$= \sum_{k=1}^{d-j} p(k) \int_{v(1, k) - v(1, 1)}^{\infty} (1 - F_k(r))dr$$

$$= \sum_{k=1}^{d-j} p(k) \int_{0}^{\infty} (1 - F_k(r))dr + v(1, 1) - \sum_{k=1}^{d-j} p(k)\ v(1, k) = g.$$

The proof is completed.

If there exists j satisfying (18) and (19), every customer arriving in state j is accepted by the stationary optimal policy because of (20) and (21). The transition probability from state j to state $j - 1$ is 0 and states $i = 0, \ldots, j - 1$ are transient as the stationary Markov chain. For a stationary optimal policy satisfying (18) and (19) there exists a stationary optimal monotone policy such that every customer arriving in state $i = 0, \ldots, j$ is accepted.

Corollary 3.5 There exists a critical-number monotone stationary optimal policy for infinite-horizon without discounting $\alpha = 1$ such that the acception region of reward r under (i, k) is $B_{i,k} = \{r;\ r \geq v(i, k)\}$. For a fixed k, $v(i, k)$ is nondecreasing in i and for a fixed i the critical-number of reward per unit service time $v(i, k)/k$ is nondecreasing in $1 \leq k \leq d - i$.

Proof: The proof is immediately obtained from that $V(i)$ is nonincreasing and concave.

In our model the constant delivery interval is assumed and then the FIFS queue discipline is adopted. We may consider its variations, for example, a variable delivery interval model or a multi-stage production model. Especially in the latter case the decision maker should form the job scheduling at the decision epock. It seems to the author that it is necessary to study composite problems of scheduling and decision.

Acknowledgements

The author would like to express his sincere thanks to Prof. S. Ikuta for his discussions. He is indebted to the referees for their valuable comments and improvements.

References

[1] S. Ikuta, "A Sequential Selection Process and its Applications," to appear.

[2] S. A. Lippman, "Applying a New Device in the Optimization of Exponential Queueing Systems," *Opns. Res.* 23, 687-710 (1975).

[3] S. A. Lippman and S. Stidham, JR., "Individual versus Social Optimization in Exponential Congestion Systems," *Opns. Res.* 25, 233-247 (1977).

[4] B. Miller, "A Queing Reward System with Several Customer Classes," *Management Sci.* 16, 234-245 (1969).

[5] S. Ross, *Applied Probability Models with Optimization Applications.* Holden-Day, San Francisco, 1970.

[6] S. Stidham. JR., "Socially and Individually Optimal Control of Arrivals to a GI/M/1 Queue," *Management Sci.* 24, 1598-1610 (1978).

[7] S. Stidham. JR. and N. U. Prabhu, "Optimal Control of Queueing Theory," *Lecture Notes in Economics and Math. Systems*, Vol. 98, 263-294, Springer, Berlin and New York, 1974.

Shōichi NISHIMURA: Socio Economic Planning, University of Tsukuba Sakuramura, Niiharigun, Ibaraki 305, Japan.