Journal of the Operations Research
Society of Japan
Vol. 30, No. 2, June 1987



# MINIMAX STRATEGIES FOR AVERAGE COST STOCHASTIC GAMES WITH AN APPLICATION TO INVENTORY MODELS

Masami Kurano
*Chiba University*

(Received July 1, 1986; Revised December 1, 1986)

*Abstract* We consider a zero-sum average cost stochastic game with the unbounded lower semi-continuous cost function, and by using the contraction property ([6, 7]) for the average case we give sufficient conditions for which there exists a minimax stationary strategy. Also, we formulate a minimax inventory model as a stochastic game and show that for any $\epsilon > 0$ there exists an $\epsilon$-minimax random $(s, S)$ ordering policy, which is a modification of $(s, S)$ ordering policy, under some weak conditions.

## 1. Introduction and Notation

A zero-sum stochastic game has been investigated by many authors and the existence of equilibrium strategies has been discussed. For example, see [9,11] for the discounted case and [2,13] for the average case.

In this paper we consider an average cost stochastic game with the unbounded lower semi-continuous cost function, and by using the contraction property ([6,7]) for the average case we give sufficient conditions for which there exists a minimax stationary strategy. Also, we apply these results to the inventory model with an unknown demand distribution and show that for any $\varepsilon > 0$ there exists an $\varepsilon$-minimax random $(s,S)$ ordering policy, which is a modification of $(s,S)$ ordering policy, under some weak conditions.

By a Borel set we mean a Borel subset of some complete separable metric space. For a Borel set $X$, $\mathcal{B}_X$ denotes the Borel subsets of $X$. If $X$ is a non-empty Borel set, $B^+(X)$ [ $B_s^+(X)$ ] denotes the set of all non-negative real valued Borel measurable [lower semi-continuous] functions on $X$. The product of the sets $D_1, D_2, \ldots$ will be denoted by $D_1 D_2 \ldots$ .

A zero-sum stochastic game is specified by five objects: $S$, $\{A(x), x \in S\}$,

$B$, $C$, $Q$, where $S$ is any Borel set and denotes the state space, for each $x\varepsilon S$, $A(x)$ is a non-empty Borel subset of a Borel set $A$ such that $\{(x,a): x\varepsilon S, a\varepsilon A(x)\}$ is closed, and denotes the set of actions available to player 1 at state $x$, $B$ is a non-empty Borel set and denotes the set of actions available to player 2, $c\varepsilon B^{+}(SAB)$ is a one-step cost function for player 1 and $Q = Q(\cdot|x,a,b)$ is the law of motion, which is taken to be a stochastic kernel on $\mathcal{B}_S SAB$; i.e., for each $(x,a,b)\,\varepsilon\, SAB$, $Q(\cdot|x,a,b)$ is a probability measure on $\mathcal{B}_S$; and, for each $D\,\varepsilon\,\mathcal{B}_S$, $Q(D|\cdot)\,\varepsilon\, B^{+}(SAB)$.

A strategy of player 1 will be a sequence $\pi = (\pi_0, \pi_1,\ldots)$ such that, for each $t \geq 0$, $\pi_t$ is a stochastic kernel on $\mathcal{B}_A S(ABS)^t$ with $\pi_t(A(x_t)|x_0,a_0,b_0,\ldots, a_{t-1},b_{t-1},x_t) = 1$ for all $(x_0,a_0,b_0,\ldots, a_{t-1},b_{t-1},x_t)\,\varepsilon\, S(ABS)^t$. Let $\Pi$ denote the set of all strategies for player 1. A strategy $\pi = (\pi_0,\pi_1,\ldots)$ is called [analytically measurable] stationary strategy if there is a [analytically measurable] measurable function $f:S \to A$ with $f(x)\,\varepsilon\, A(x)$ for all $x\,\varepsilon\, S$ such that $\pi_t(f(x_t)|x_0,a_0,b_0,\ldots, a_{t-1},b_{t-1},x_t) = 1$ for all $(x_0,a_0,b_0,\ldots, a_{t-1},b_{t-1},x_t)\,\varepsilon\, S(ABS)^t$ and $t \geq 0$. Such a strategy will be denoted by $f^{\infty}$.

A strategy of player 2 is a sequence $\sigma = (\sigma_0, \sigma_1,\ldots)$ such that, for each $t \geq 0$, $\sigma_t$ is a stochastic kernel on $\mathcal{B}_B SA(BSA)^t$. We note that the $t$-th action of player 2 is taken after knowing the action taken by player 1 at the $t$-th time. Let $\Sigma$ denote the set of all strategies for player 2. Stationary strategies of player 2 are defined analogously.

The sample space is the product space $\Omega = S(ABS)^{\infty}$. Let $X_t$, $\Delta_t$ and $\Gamma_t$ be random quantities defined by $X_t(\omega) = x_t$, $\Delta_t(\omega) = a_t$ and $\Gamma_t(\omega) = b_t$ for $\omega = (x_0,a_0,b_0,x_1,a_1,b_1,\ldots)\,\varepsilon\,\Omega$.

Let $H_t = (X_0,\Delta_0,\Gamma_0,\ldots, \Delta_{t-1},\Gamma_{t-1},X_t)$. It is assumed that, for each $\pi = (\pi_0,\pi_1,\ldots)\,\varepsilon\,\Pi$ and $\sigma = (\sigma_0,\sigma_1,\ldots)\,\varepsilon\,\Sigma$, $P(\Delta_t\,\varepsilon\, D_1|H_t) = \pi_t(D_1|H_t)$, $P(\Gamma_t\,\varepsilon\, D_2|H_t,\Delta_t) = \sigma_t(D_2|H_t,\Delta_t)$ and $P(X_{t+1}\,\varepsilon\, D_3|H_{t-1},\Delta_{t-1},\Gamma_{t-1},X_t{=}x,\Delta_t{=}a, \Gamma_t{=}b) = Q(D_3|x,a,b)$ for every $D_1\,\varepsilon\,\mathcal{B}_A$, $D_2\,\varepsilon\,\mathcal{B}_B$ and $D_3\,\varepsilon\,\mathcal{B}_S$.

Then, for each $\pi\,\varepsilon\,\Pi$, $\sigma\,\varepsilon\,\Sigma$ and starting point $x\,\varepsilon\, S$, we can define the probability measure $P^x_{\pi,\sigma}$ on $\Omega$ in an obvious way.

We shall consider the following average cost criterion:

For any strategies $\pi\,\varepsilon\,\Pi$, $\sigma\,\varepsilon\,\Sigma$ and $x\,\varepsilon\, S$ let

$$(1.1)\qquad \psi(x,\pi,\sigma) = \limsup_{T\to\infty} E^x_{\pi,\sigma}[\Sigma_{t=0}^{T-1} c(X_t,\Delta_t,\Gamma_t)]/T ,$$

where $E^x_{\pi,\sigma}$ is the expectation operator with respect to $P^x_{\pi,\sigma}$.

Let $\psi(x,\pi) = \sup_{\sigma\,\varepsilon\,\Sigma} \psi(x,\pi,\sigma)$. Then for any $\varepsilon \geq 0$, we say that $\pi^*\,\varepsilon\,\Pi$ is $\varepsilon$-minimax if $\psi(x,\pi^*) \leq \psi(x,\pi) + \varepsilon$ for all $x\,\varepsilon\, S$ and $\pi\,\varepsilon\,\Pi$. A 0-minimax strategy is simple called minimax.

In Section 2, we give sufficient conditions for which a minimax stationary strategy exists. In Section 3, a minimax inventory problem is formulated as a stochastic game and it is shown that for any $\varepsilon > 0$, there exists an $\varepsilon$-minimax random $(s,S)$ ordering policy under weak conditions.

## 2. Existence of Minimax Strategy

In this section we shall give sufficient conditions for the existence of a minimax stationary policy.

In order to insure the ergodicity of the process, we introduce the following contraction property ([6,7]).

Condition A. There exist a measure $\gamma$ on $S$ such that $0 < \gamma(S) < 1$ and $Q(D|x,a,b) \geqq \gamma(D)$ for all $D \in \mathcal{B}_S$, $x \in S$, $a \in A(x)$ and $b \in B$.

Under Condition A, we define the map $U$ on $\mathrm{B}^+(S)$ by

(2.1) $$Uu(x) = \inf_{a \in A(x)} \sup_{b \in B} U(x,a,b,u)$$

if this expression exists, where

(2.2) $$U(x,a,b,u) = c(x,a,b) + \int u(y)Q(dy|x,a,b) - \int u(y)\gamma(dy)$$

for each $u \in \mathrm{B}^+(S)$, $x \in S$, $a \in A(x)$ and $b \in B$.

Condition B. The following B1-B2 holds:

B1. $c \in \mathrm{B}_s^+(SAB)$ and $Q(\cdot|x,a,b)$ is weakly continuous in $(x,a,b) \in SAB$, that is, whenever $x_n \to x$, $a_n \to a$ and $b_n \to b$, $Q(\cdot|x_n,a_n,b_n)$ converges weakly to $Q(\cdot|x,a,b)$.

B2. When $x_n \in S \to x \in S$ as $n \to \infty$, for any sequence $\{a_n\}$ with $a_n \in A(x_n)$ for all $n \geqq 1$, there exist a subsequence $\{a_{n_j}\}$ of $\{a_n\}$ and $a \in A(x)$ such that $a_{n_j} \to a$ as $j \to \infty$.

We need the following condition to treat with the unbounded cost.

Condition C. There exists a $\bar{v} \in \mathrm{B}_s^+(S)$ such that the following C1-C3 hold:

C1. $c(x,a,b) \leqq \bar{v}(x)$ for all $x \in S$, $a \in A(x)$ and $b \in B$.

C2. $U\bar{v} \leqq \bar{v}$.

C3. $\int \bar{v}(y)Q(dy|x,a,b)$ is uniformly integrable for $(x,a,b) \in SAB$.

In the next section we shall show that the usual inventory model satisfies Condition B and C.

For any non-empty Borel set $X$, we denote by $\bar{\mathrm{B}}_s^+(X)$ the set of all non-negative real-valued, bounded lower semi-continuous functions on $X$.

**Lemma 2.1.** Suppose that Conditions B and C hold. Then for any $u \in B_s^+(S)$ with $0 \leqq u \leqq \bar{v}$ it holds that (i) $\int u(y)Q(dy|x,a,b) \in B_s^+(SAB)$ and (ii) $\sup_{b \in B} U(x,a,b,u) \in B_s^+(SA)$.

**Proof:** From C3, for any $\varepsilon > 0$ there exists a constant $M$ for which $\int_D \bar{v}(y)D(dy|x,a,b) \leqq \varepsilon/2$ for all $x \in S$, $a \in A(x)$ and $b \in B$, where $D = \{y \in S | \bar{v}(y) \geqq M\}$. Let $u \in B_s^+(S)$ with $0 \leqq u \leqq \bar{v}$. And, for the above $M$, let $u_M(y) = u(y)$ if $u(y) < M$, $= M$ if $u(y) \geqq M$. Then since $u_M \in B_s^+(S)$, it holds from Lemma 4.1 of Maitra [8] that

$$(2.3) \qquad \int u_M(y)Q(dy|x,a,b) \in \bar{B}_s^+(SAB).$$

Also, we obtain

$$(2.4) \qquad \left| \int u(y)Q(dy|x,a,b) - \int u_M(y)Q(dy|x,a,b) \right| \leqq \int_D \bar{v}(y)Q(dy|x,a,b) \leqq \varepsilon/2 \quad \text{for all } (x,a,b) \in SAB.$$

Therefore, by (2.3) and (2.4) it holds that when $(x_n,a_n,b_n) \to (x,a,b)$,

$$\begin{aligned}\liminf_{n\to\infty} &\int u(y)Q(dy|x_n,a_n,b_n,) \\ &\geqq \liminf_{n\to\infty} \int u_M(y)Q(dy|x_n,a_n,b_n) - \varepsilon/2 \\ &\geqq \int u_M(y)Q(dy|x,a,b) - \varepsilon/2 \\ &\geqq \int u(y)Q(dy|x,a,b) - \varepsilon.\end{aligned}$$

As $\varepsilon \to 0$, $\liminf_{n\to\infty} \int u(y)Q(dy|x_n,a_n,b_n) \geqq \int u(y)Q(dy|x,a,b)$, which means (i). Clearly (ii) follows. Q.E.D.

**Lemma 2.2.** Suppose that Conditions A, B and C hold. Then, for any $u \in B_s^+(S)$ with $0 \leqq u \leqq \bar{v}$, $Uu \in B_s^+(S)$.

**Proof:** For any fixed $u \in B_s^+(S)$ with $0 \leqq u \leqq \bar{v}$, let $U(x,a,u) = \sup_{b \in B} U(x,a,b,u)$. Then since $U(x,a,u) \in B_s^+(SA)$, by the definition of $Uu$, for any state sequence $\{x_n\}$ with $x_n \in S \to x \in S$ as $n \to \infty$ and $\varepsilon > 0$ there exists an action sequence $\{a_n\}$ such that

$$Uu(x_n) \geqq U(x_n,a_n,u) - \varepsilon \text{ for all } n \geqq 1.$$

Using the condition B2, there are a subsequence $\{a_{n_j}\}$ of $\{a_n\}$ and $a \in A(x)$ for which $a_{n_j} \to a$ as $j \to \infty$ and

$$\begin{aligned}\liminf_{n\to\infty} Uu(x_n) &\geqq \liminf_{n\to\infty} U(x_n,a_n,u) - \varepsilon \\ &= \lim_{j\to\infty} U(x_{n_j},a_{n_j},u) - \varepsilon \geqq U(x,a,u) - \varepsilon \\ &\geqq Uu(x) - \varepsilon.\end{aligned}$$

As $\varepsilon \to 0$ in the above, $Uu \in B_s^+(S)$ follows. Q.E.D.

We denote by $B(S \to A)$ the set of all Borel measurable functions $f:S \to A$ with $f(x) \in A(x)$ for all $x \in S$ and by $B_a(X \to B)$ the set of all lower semi-analytic functions $h:X \to B$, where $X$ is any Borel set.

**Lemma 2.3.** Suppose that Conditions A, B and C hold. Then, for any $u$, $w \in B_s^+(S)$ with $0 \leq u$, $w \leq \bar{v}$ and $\varepsilon > 0$ there exist $f \in B(S \to A)$ and $h \in B_a(S \to B)$ such that

$$(2.5) \qquad Uu(x) - Uw(x) \leq \int (u(y) - w(y))\bar{Q}(dy|x,f(x),h(x)) + \varepsilon \quad \text{for all } x \in S,$$

where

$$(2.6) \qquad \bar{Q}(dy|x,a,b) = Q(dy|x,a,b) - \gamma(dy).$$

**Proof:** By Lemma 2.1 $U(x,a,w) \in B_s^+(SA)$, so that it holds from the selection theorem ([1,12]) that for any $\varepsilon > 0$ there exist $f \in B(S \to A)$ and $h \in B_a(S \to B)$ such that $U(x,f(x),w) = Uw(x)$ and $U(x,f(x),h(x),u) \geq U(x,f(x),u) - \varepsilon$ for all $x \in S$.

Thus, by the definition of $U$, we have

$$\begin{aligned} Uu(x) - Uw(x) &\leq U(x,f(x),u) - U(x,f(x),w) \\ &\leq U(x,f(x),h(x),u) - U(x,f(x),h(x),w) + \varepsilon, \end{aligned}$$

which implies (2.5). Q.E.D.

**Theorem 2.1.** Suppose that Conditions A, B and C hold. Then there exist a constant $\psi^*$ and a $v \in B_s^+(S)$ with $0 \leq v \leq \bar{v}$ such that

$$(2.7) \qquad v(x) = \inf_{a \in A(x)} \sup_{b \in B} \{c(x,a,b) - \psi^* + \int v(y)Q(dy|x,a,b)\} \quad \text{for all } x \in S,$$

and if

$$(2.8) \qquad \lim_{T\to\infty} E^x_{\pi,\sigma} [v(X_T)]/T = 0 \quad \text{for all } x \in S,\ \pi \in \Pi \text{ and } \sigma \in \Sigma,$$

it holds that

$$(2.9) \qquad \psi^* \leq \psi(x,\pi) \quad \text{for all } x \in S \text{ and } \pi \in \Pi.$$

**Proof:** Let us define the sequence $\{\bar{v}_n\}$ and $\{\underline{v}_n\}$ respectively by $\bar{v}_0 = \bar{v}$, $\underline{v}_0 = 0$, $\bar{v}_{n+1} = U\bar{v}_n$ and $\underline{v}_{n+1} = U\underline{v}_n$ for all $n \geq 1$.

Then, from Lemma 2.2 and the monotonicity of $U$ we have $\bar{v}_0 \geq \bar{v}_n \geq \bar{v}_{n+1} \geq \underline{v}_{n+1} \geq \underline{v}_n \geq 0$ and $\underline{v}_n \in B_s^+(S)$ $(n \geq 1)$.

Now, we show by induction that there exists a constant $M$ such that

(2.10) $\bar{v}_n(x) - \underline{v}_n(x) \leq M\beta^{n-1}$ for all $n \geq 1$,

where $\beta = 1 - \gamma(S)$ and $0 < \beta < 1$.

In fact, from C3 there exists some $M$ such that $\int \bar{v}(y)Q(dy|x,a,b) \leq M$ for all $x \in S$, $a \in A(x)$ and $b \in B$. For any $\varepsilon > 0$ from Lemma 2.3 there exist $f \in B(S \to A)$ and $h \in B_a(S \to B)$ for which

$$\bar{v}_1(x) - \underline{v}_1(x) \leq \int(\bar{v}_0(y) - \underline{v}_0(y))\bar{Q}(dy|x,f(x),h(x)) + \varepsilon$$
$$\leq \int \bar{v}_0(y)Q(dy|x,f(x),h(x)) + \varepsilon \leq M + \varepsilon,$$

so that as $\varepsilon \to 0$ we observe that (2.10) holds for $n = 1$.
Suppose that (2.10) holds for $n$. Similarly, for any $\varepsilon > 0$ there exist $f_n \in B(S \to A)$ and $h_n \in B_a(S \to B)$ such that

$$\bar{v}_{n+1}(x) - \underline{v}_{n+1}(x) \leq \int(\bar{v}_n(y) - \underline{v}_n(y))\bar{Q}(dy|x,f_n(x),h_n(x)) + \varepsilon$$
$$\leq M\beta^{n-1}\bar{Q}(S|x,f_n(x),h_n(x)) + \varepsilon$$
$$= M\beta^n + \varepsilon,$$

which shows that (2.10) holds for $n+1$. Thus, if we let $v = \lim_{n\to\infty} \bar{v}_n$, then $v = \lim_{n\to\infty} \underline{v}_n$ and $v \in B_s^+(S)$. Also, since $\bar{v}_n = U\bar{v}_{n-1} \geq Uv$ and $\underline{v}_n = U\underline{v}_{n-1} \leq Uv$, we get $v \geq Uv$ and $v \leq Uv$, which implies

(2.11) $v = Uv$.

If we let $\psi^* = \int v(y)\gamma(dy)$ in (2.11), (2.11) means (2.7).
For $\psi^*$ and $v \in B_s^+(S)$ as in (2.7), we define

$$\phi(x,a,b) = c(x,a,b) - \psi^* - v(x) + \int v(y)Q(dy|x,a,b)$$

for each $x \in S$, $a \in A(x)$ and $b \in B$.

Then, it holds from (2.7) that $\sup_{b \in B} \phi(x,a,b) \geq 0$ for all $x \in S$ and $a \in A(x)$, so that using the selection theorem ([1,12]) for any $\varepsilon > 0$ there exists $h \in B_a(SA \to B)$ such that $\phi(x,a,h(x,a)) \geq -\varepsilon$ for all $x \in S$ and $a \in A(x)$. So, for this stationary policy $h^\infty$, we have

$$E^x_{\pi,h^\infty}[\ \phi(X_t,\Delta_t,\Gamma_t)\ ] \geq -\varepsilon$$

for all $\pi \in \Pi$,

which derives

$$E^x_{\pi,h^\infty}[\ \Sigma_{t=0}^{T-1} c(X_t,\Delta_t,\Gamma_t)\ ]/T$$
$$\geq \psi^* + (v(x) - E^x_{\pi,h^\infty}[v(X_T)])/T - \varepsilon.$$

Therefore, as $T\to\infty$ and $\varepsilon\to 0$ in the above, we get $\psi(x,\pi)\geqq\psi^*$. Q.E.D.

The next theorem shows the existence of a minimax stationary strategy in a stochastic game model.

**Theorem 2.2.** Suppose that Conditions A, B and C hold. Then it holds that

(i) there exists $f\in B(S\to A)$ such that

(2.12) $$\sup_{b\in B}\phi(x,f(x),b) = 0 \quad \text{for all } x\in S$$

and

(ii) if (2.8) holds, the stationary strategy $f^\infty$ is minimax.

**Proof:** From the selection theorem ([1,11]), (i) follows. For (ii), from (2.12) it holds that $\phi(x,f(x),b)\leqq 0$ for all $b\in B$, so that by the similar discussion as that of Theorem 2.1 we obtain $\psi(x,f^\infty)\leqq\psi^*$, which implies from (ii) of Theorem 2.1 that $f^\infty$ is minimax. Q.E.D.

## 3. A Minimax Inventory Model

In this section we consider the single-item stochastic inventory model whose demand distributions for each period are assumed to be unknown but are restricted to a class of distributions on $R^+ = (0,\infty)$.

And by transforming this model equivalently to a stochastic game between a decision maker and Nature we shall give a characterization of a minimax ordering policy which minimizes the maximum average expected cost over the infinite planning horizon. Here, the demands in successive periods are assumed to form a sequence of independent random variables whose distributions can change from period to period in a restricted class of distributions and any unfilled demand in a period is backlogged. We note that a reader may refer to Jagannathan [5] for the discounted minimax case.

Let $\mathrm{P}(R^+)$ be the set of all probability measure or, equivalently, distributions on $R^+$. Then it is known that $\mathrm{P}(R^+)$ is a complete separable metric space with respect to the weak topology (for example, see [1]). Let $\mathcal{F}$ be a Borel subset of $\mathrm{P}(R^+)$. Define $S = (-\infty,M]$ and $A = [0,M]$, where $M$ is a capacity of inventory. For each $x\in S$, $A(x) = [0\vee x,M]\subset A$ is the set of actions available to a decision maker (player 1) at state $x$ and denotes the set of inventory after ordering, where $x\vee y = \max\{x,y\}$. And $B = \mathcal{F}$ is the set of actions available to player 2.

Then, the stochastic kernel $Q$ is as follows:

$$Q(D|x,a,F) = P(a-\tilde{x} \in D) \text{ for each } x \in S,\ a \in A(x) \text{ and } F \in \mathcal{F},$$

where $\tilde{x}$ is a random variable with the distribution $F$. For one-step cost, let, for each $x \in S$, $a \in A(x)$ and $F \in \mathcal{F}$,

(3.1) $$c(x,a,F) = K \cdot I_{(0,\infty)}(a-x) + c \cdot (a-x) + L(a,F),$$

where $L(a,F)$ is the expected holding and shortage cost at the inventory $a$ after ordering when the demand distribution is $F$ and $K > 0$ is a set-up cost and $I_D$ is the indicator function of $D$.

We introduce the following conditions to apply the results of Section 2.

Condition D. The following D1-D2 hold.

D1. There exist $\kappa > 0$ and $\delta > 0$ such that

$$\int_0^\infty y^{1+\delta} dF(y) \leqq \kappa \qquad \text{for all } F \in \mathcal{F}.$$

D2. $L(a,F)$ is convex in $a \in A$ for each $F \in \mathcal{F}$ and bounded with $0 \leqq L(a,F) \leqq L$ for some $L$ and all $a \in A$ and $F \in \mathcal{F}$.

Condition E. There is a measure $\gamma$ on $S$ such that $0 < \gamma(S) < 1$ and $Q(D|x,a,F) \geqq \gamma(D)$ for all $D \in \mathcal{B}_S$, $x \in S$, $a \in A(x)$ and $F \in \mathcal{F}$.

Example.

We denote by $N_+(\mu,\sigma^2)$ the normal distribution which is truncated at 0 on the left. For any given $d_i$ $(i=1,2,3,4)$ with $0 < d_1 < d_2$ and $0 < d_3 < d_4$ let

$$\mathcal{F} = \{ N_+(\mu,\sigma^2) \mid d_1 \leqq \mu \leqq d_2,\ d_3 \leqq \sigma^2 \leqq d_4 \}.$$

In this case, D1 holds for $\delta = 2$ and D2 holde for any linear holding and penalty cost functions. Let $f_+(x;\mu,\sigma^2)$ be the density of $N_+(\mu,\sigma^2)$. We observe that

$$Q(D|x,a,N_+(\mu,\sigma^2)) = \int_{a-y \in D} f_+(y;\mu,\sigma^2)dy \text{ for any } D \in \mathcal{B}_S \text{ and } a \in A(x).$$

We define a function $f(y)$ by $f(y) = \min_{d_1 \leqq \mu \leqq d_2, d_3 \leqq \sigma^2 \leqq d_4, a \in A} f_+(a-y;\mu,\sigma^2)$ if $y \leqq 0$, $= 0$ if $0 < y \leqq M$.

Then, it is easily verified that $0 < \gamma(S) < 1$ and

$$Q(D|x,a,N_+(\mu,\sigma^2)) \geqq \gamma(D) \text{ for any } D \in \mathcal{B}_S,\ x \in S,\ a \in A(x) \text{ and } N_+(\mu,\sigma^2) \in \mathcal{F},$$

where $\gamma(D) = \int_D f(y)dy$.

That is, Condition E holds for this $\mathcal{F}$.

Lemma 3.1. Suppose that Conditions D and E hold. Then, Conditions A, B and C in Section 2 are satisfied in a stochastic game defined above.

Proof: For any integer m and real number $\beta'$ such that

$0 < \beta' \leqq \gamma(S) - c\cdot\kappa\cdot\{K+L+c\cdot(M+m)\}^{-1}$, let define a function $\bar{v}$ on $S$ by

$$\begin{aligned}\bar{v}(x) &= (K + L + c\cdot(M+m))/\beta' \quad \text{if } x \in (-m,M], \\ &= (K + L + c\cdot(M+j+1))/\beta' \quad \text{if } x \in (-j-1,-j] \quad \text{for } j \geqq m.\end{aligned}$$

Then, it holds that $U(x,a,F,v) \leqq \bar{v}(x)$ for all $x \in S$, $a \in A(x)$ and $F \in \mathscr{F}$, where $U(x,a,F,v)$ is defined in (2.2).

In fact, for example, when $x \in (-m,M]$, we have

$$\begin{aligned}U(x,a,F,\bar{v}) &= c(x,a,F) + \int\bar{v}(y)Q(dy|x,a,F) \\ &\leqq K + L + c\cdot(M+m) + \{(1-\gamma(S))(K+L+c\cdot(M+m)) + c\kappa\}/\beta' \\ &\leqq \bar{v}(x),\end{aligned}$$

where $\bar{Q}$ is defined in (2.6). Thus we get $U\bar{v} \leqq \bar{v}$.

Also, it is easily verified that other conditions in Conditions A, B and C hold. Q.E.D.

Before stating the theorem, we give the following lemma.

**Lemma 3.2.** Suppose that $g(x,\lambda)$ is $K$-convex in $x \in R^+$ for each $\lambda \in \Gamma$. Then, $\sup_{\lambda\in\Gamma} g(x,\lambda)$ is $K$-convex in $x \in R^+$.

**Proof:** Let $g(x) = \sup_{\lambda\in\Gamma} g(x,\lambda)$.

For any $\varepsilon > 0$ and $x \in S$, $g(x) \leqq g(x,\lambda) + \varepsilon$ for some $\lambda \in \Gamma$. Thus,

$$\begin{aligned}&K + g(x+d) - g(x) - d\{(g(x) - g(x-e))/e\} \\ &= K + g(x+d) + dg(x-e)/e - (1+d/e)g(x) \\ &\geqq K + g(x+d,\lambda) + dg(x-e,\lambda)/e - (1+d/e)g(x,\lambda) - (1+d/e)\varepsilon \\ &\geqq -(1+d/e)\varepsilon \text{ from the hypothesis of } K\text{-convexity.}\end{aligned}$$

As $\varepsilon \to 0$ in the above, we have

$$K + g(x+d) - g(x) - d\{(g(x) - g(x-e))/e\} \geqq 0$$

$$\text{for all } x \in S,\ d > 0 \text{ and } e > 0,$$

which implies $K$-convexity of $g$. Q.E.D.

**Theorem 3.1.** Under Conditions D and E, a minimax $(s,S)$ ordering policy exists.

**Proof:** By Theorem 2.1, there exist a constant $\psi^*$ and $v \in B_s^+(S)$ such that

$$\begin{aligned}v(x) = \inf_{a\in x} \sup_{F\in\mathscr{F}} \{&K\cdot I_{(0,\infty)}(x-a) + c\cdot(a-x) \\ &+ L(a,F) - \psi^* + \int v(a-y)dF(y)\}.\end{aligned}$$

Now, we show that $v$ is $K$-convex. For the operator $U$ defined in (2.1), let $u_0 = 0$ and $u_n = Uu_{n-1}$ for $n \geqq 1$. First, we show by induction that $u_n$ is

$K$-convex for all $n \geq 0$. If we define $G(x,a,F,u) = c\cdot a + L(a,F) + \int u(a-y)Q(dy|x,a,F)$ for each $x \in S$, $a \in A(x)$, $F \in \mathcal{F}$ and $u \in B_s^+(S)$, we can write

$$U(x,a,F,u) = -c\cdot x + \min\{G(x,x,F,u),\ K + G(x,a,F,u)I_{(x,M]}(a)\} - \int u(y)\gamma(dy).$$

From the results of Iglehart [3,4], $G(x,a,F,u_n)$ is $K$-convex in $a \in A$ if $u_n$ is $K$-convex.

Since $\sup_{F \in \mathcal{F}} G(x,a,F,u_n)$ is $K$-convex in $a \in A$ for Lemma 3.2, by using the results of Iglehart again it holds that $u_{n+1} = Uu_n$ is $K$-convex. Therefore, since $v = \lim_{n\to\infty} u_n$ by the similar discussion as Theorem 2.1, $v$ is $K$-convex. By Theorem 2.2, the minimax stationary strategy $f^\infty$ exists. Since $\sup_{F \in \mathcal{F}} G(x,a,F,v)$ is $K$-convex in $a \in A$, we can prove, by the same way as used in Iglehart [3,4], that $f^\infty$ is an $(s,S)$ ordering policy. Q.E.D.

We say that $\pi = (\pi_0, \pi_1, \ldots) \in \Pi$ is a random $(s,S)$ ordering policy if there exist $\varepsilon_1 (0 < \varepsilon_1 < 1)$ and a map $f: S \to A$ satisfying that $f(x) = S_1$, if $x \leq s_1$, $=x$ if $x > S_1$ for some $s_1 < S_1$ such that $\pi_t$ selects the action $\Delta_t = f(X_t)$ with probability $1-\varepsilon_1$ and the action $X_t \vee S_1$ with probability $\varepsilon_1$.

Then we can state the main theorem.

**Theorem 3.2.** Suppose that Condition D holds and $L(a,F)$ is linear in $F \in P(R^+)$ for each $a \in A$.

Then for any $\varepsilon > 0$ there exists a random $(s,S)$ ordering policy which is $\varepsilon$-minimax.

In order to prove Theorem 3.2, we shall introduce a subsidiary stochastic game for which Condition E holds.

Let $\Phi \in P(R^+)$ be such that $\Phi$ has density $\phi(x)$ with $\phi(x) = (2M)^{-1}$ if $M \leq x \leq 3M$ and $= 0$ otherwise. For this $\Phi$ and $\varepsilon_1$ $(0 < \varepsilon_1 < 1)$, put $\mathcal{F}_{\varepsilon_1} = \{\varepsilon_1\Phi + (1-\varepsilon_1)F : F \in \mathcal{F}\}$.

Now we consider a subsidiary inventory model $G(\mathcal{F}_{\varepsilon_1})$ in which the set of actions available to player 2 is $\mathcal{F}_{\varepsilon_1}$ but the state space and the set of actions available to a decision maker (player 1) at state $x$ are respectively $S=(-\infty,M]$ and $A(x)=[0 \vee x, M]$.

Notice that the sample space of $G(\mathcal{F}_{\varepsilon_1})$ is $\Omega' = S(A\mathcal{F}_{\varepsilon_1} S)^\infty$. In $G(\mathcal{F}_{\varepsilon_1})$, we denote respectively by $X_t'$, $\Delta_t'$ and $\Gamma_t'$ the state and the actions at the $t$-th time taken by players 1 and 2 $(t \geq 0)$. Also, in $G(\mathcal{F}_{\varepsilon_1})$ let $\Pi'$ and $\Sigma'$ be respectively the classes of strategies for players 1 and 2 and $\psi'(x,\pi',\sigma')$ the average cost defined by (1.1) for any $x \in S$, $\pi' \in \Pi'$ and $\sigma' \in \Sigma'$. In the proof of Theorem 3.2 given later it is shown that Condition E holds for $G(\mathcal{F}_{\varepsilon_1})$, so that applying Theorem 3.1 under Condition D there exists a minimax $(s,S)$ ordering policy for $G(\mathcal{F}_{\varepsilon_1})$.

To investigate the relation between $\Pi$ ($\Sigma$) and $\Pi'$ ($\Sigma'$), we introduce the following transformation.

Let $\{Y_t\}$ be a sequence of independent random variables such that for each $t \geqq 0$ $Y_t$ is uniformly distributed on $(0,1)$.

For any $t \geqq 0$ and the random quantity $H'_t = (X'_0, \Delta'_0, \Gamma'_0, \ldots, \Delta'_{t-1}, \Gamma'_{t-1}, X'_t)$ $\varepsilon\, S(A\, \mathcal{F}_{\varepsilon_1} S)^t$, we define a random quantity $\tilde{H}_t = (\tilde{X}_0, \tilde{\Delta}_0, \tilde{\Gamma}_0, \ldots, \tilde{\Delta}_{t-1}, \tilde{\Gamma}_{t-1}, \tilde{X}_t)$ $\varepsilon\, S(A\, \mathcal{F} S)^t$ by

$$\tilde{X}_0 = X'_0,\ \tilde{\Gamma}_j = (\Gamma'_j - \varepsilon_1 \Phi)/(1-\varepsilon_1),\ \tilde{\Delta}_j = \Delta'_j$$

and $\tilde{X}_{j+1} = \tilde{\Delta}_j - \tilde{\Gamma}_j^{-1}(Y_j)$ for each $j \geqq 0$,

where for any $F \in \mathrm{P}(R^+)$ $F^{-1}$ is a left continuous inverse and $F^{-1}(t) = \inf\{x\colon F(x) \geqq t\}$.

We note that $\tilde{\Gamma}_j \in \mathcal{F}$ because $\Gamma'_j \in \mathcal{F}_{\varepsilon_1}$.

Using the above transformation, from $\pi = (\pi_0, \pi_1, \ldots) \in \Pi$ and $\sigma = (\sigma_0, \sigma_1, \ldots) \in \Sigma$ we construct $\pi' = (\pi'_0, \pi'_1, \ldots) \in \Pi'$ and $\sigma' = (\sigma'_0, \sigma'_1, \ldots) \in \Sigma'$ by

$\pi'_t(\cdot|H'_t) = \pi_t(\cdot|\tilde{H}_t)$ and

$\sigma'_t(D|H'_t, \Delta'_t) = \mathrm{Prob}(\varepsilon_1 \Phi + (1-\varepsilon_1)\tilde{F} \in D)$ for any Borel subset $D$ of $\mathcal{F}_{\varepsilon_1}$ and $t \geqq 0$,

where $\tilde{F}$ is distributed with $\sigma_t(\cdot|\tilde{H}_t, \tilde{\Delta}_t)$.

To make the above definition possible, we only need to show that $\pi_t(A(X'_t)|\tilde{H}_t) = 1$ for all $t \geqq 0$. In fact, since $X'_t = \Delta'_{t-1} - W'_{t-1}$ and $\tilde{X}_t = \Delta'_{t-1} - W_{t-1}$ and $W'_{t-1}$ and $W_{t-1}$ are respectively distributed with $\Gamma'_{t-1} = \varepsilon_1 \Phi + (1-\varepsilon_1)\tilde{\Gamma}_{t-1}$ and $\tilde{\Gamma}_{t-1}$, it holds from the property of $\Phi$ that $\mathrm{Prob}(X'_t \leqq \mathrm{Max}\{\tilde{X}_t, 0\}) = 1$. Thus $\mathrm{Prob}(A(X'_t) \supset A(\tilde{X}_t)) = 1$ so that by $\pi_t(A(\tilde{X}_t)|\tilde{H}_t) = 1$ we get $\pi_t(A(X'_t)|\tilde{H}_t) = 1$ for all $t \geqq 0$.

For convenience, we say $\pi' \in \Pi'$ ($\sigma' \in \Sigma'$) a strategy constructed from $\pi \in \Pi$ ($\sigma \in \Sigma$) using the random transformation ($\rho$).

Conversely, we try to construct $\pi \in \Pi$ and $\sigma \in \Sigma$ from $\pi' \in \Pi'$ and $\sigma' \in \Sigma'$.

Let $\{\eta_t\}$ and $\{Z_t\}$ be sequences of independent random variables with $\mathrm{Prob}(\eta_t = 1) = 1 - \mathrm{Prob}(\eta_t = 0) = \varepsilon_1$ and $Z_t$ is distributed with $\Phi$ for all $t \geqq 0$.

For any $t \geqq 0$ and the random quantity $H_t = (X_0, \Delta_0, \Gamma_0, \ldots, \Delta_{t-1}, \Gamma_{t-1}, X_t)$ $\varepsilon\, S(A\, \mathcal{F} S)^t$, we define a random quantity $\tilde{H}'_t = (\tilde{X}'_0, \tilde{\Delta}'_0, \tilde{\Gamma}'_0, \ldots, \tilde{\Delta}'_{t-1}, \tilde{\Gamma}'_{t-1}, X'_t)$ $\varepsilon$ $S(A\, \mathcal{F}_{\varepsilon_1} S)^t$ by

$$\tilde{X}'_0 = X_0,\ \tilde{\Delta}'_j = \Delta_j,\ \tilde{\Gamma}'_j = \varepsilon_1 \Phi + (1-\varepsilon_1)\Gamma_j \text{ and}$$

$$\tilde{X}'_{j+1} = \tilde{\Delta}_j - z_j \text{ if } \eta_j = 1, = X_{j+1} \text{ if } \eta_j = 0$$
for each $j \geq 0$.

And for any $(s,S)$ ordering strategy $\pi' = (\pi'_0, \pi'_1, \ldots) \in \Pi'$ and any strategy $\sigma' = (\sigma'_0, \sigma'_1, \ldots) \in \Sigma'$, we construct $\pi = (\pi_0, \pi_1, \ldots) \in \Pi$ and $\sigma = (\sigma_0, \sigma_1, \ldots) \in \Sigma$ by

$$\pi_t(\cdot|H_t) = \pi'_t(\cdot|\tilde{H}'_t) \text{ and}$$
$$\sigma_t(D|H_t) = \sigma'_t(D'|\tilde{H}'_t, \tilde{\Delta}'_t) \text{ for each } t \quad 0 \text{ and any Borel subset } D \text{ of } \mathcal{F},$$
where $D' = \{\varepsilon_1\Phi + (1-\varepsilon_1)F : F \in D\}$.

We say $\pi \in \Pi$ ($\sigma \in \Sigma$) a strategy constructed from $\pi' \in \Pi'$ ($\sigma' \in \Sigma'$) using the random transformation ($\nu$).
In this case, since $\pi'$ is an $(s,S)$ ordering policy, $\pi$ becomes a random $(s,S)$ ordering policy.

Lemma 3.3. Suppose that Conditions in Theorem 3.2 hold. Then for any $\varepsilon > 0$ there exists $\varepsilon_1 > 0$ satisfying the following: For any $\pi \in \Pi$, there is $\pi' \in \Pi'$ such that for any $\sigma \in \Sigma$ there exists $\sigma' \in \Sigma'$ for which

$$\text{(3.2)} \qquad |\, \psi(x,\pi,\sigma) - \psi'(x,\pi',\sigma') \,| < \varepsilon/2$$

and conversely for any $\sigma' \in \Sigma'$ there exists $\sigma \in \Sigma$ satisfying (3.2).

Proof: For any given $\pi \in \Pi$ and $\varepsilon_1 > 0$ let $\pi' \in \Pi'$ be a strategy constructed from $\pi$ using the random transformation ($\rho$). Then when $\varepsilon_1$ is sufficiently small, we will show that this $\pi' \in \Pi'$ is the desired strategy.

For any $\sigma \in \Sigma$, let $\sigma' \in \Sigma'$ be a strategy constructed from $\sigma \in \Sigma$ using the random transformation ($\rho$). Then, by the method of construction we observe that $P^x_{\pi',\sigma'}(\tilde{H}_t \in D) = P^x_{\pi,\sigma}(H_t \in D)$ for all $t \geq 0$ and any Borel subset $D$ of $S(A\,\mathcal{F}\,S)^t$, so that

$$\text{(3.3)} \qquad E^x_{\pi,\sigma}[c(X_t,\Delta_t,\Gamma_t)] = E^x_{\pi',\sigma'}[c(\tilde{X}_t,\tilde{\Delta}_t,\tilde{\Gamma}_t)].$$

From the property of $\Phi$ we can assume that $L(a,\Phi) \leq L'$ for all $a \in A$ and some $L'$. Thus we get, by the linearity of $L$,

$$\text{(3.4)} \qquad |L(a,F) - L(a,\varepsilon_1\Phi + (1-\varepsilon_1)F)| \leq \varepsilon_1(L + L').$$

Also, by the definition we have

$$\text{(3.5)} \qquad E^x_{\pi',\sigma'}[KI_{(0,\infty)}(\Delta'_t - \tilde{X}_t) - KI_{(0,\infty)}(\Delta'_t - X'_t)] \leq 2\varepsilon_1 K$$

and

$$\text{(3.6)} \qquad |E^x_{\pi',\sigma'}[c\cdot(\tilde{X}_t - X'_t)]| \leq \varepsilon_1(3M + \kappa).$$

Thus, we have

$$|E^x_{\pi',\sigma'}[c(\tilde{X}_t,\tilde{\Delta}_t,\tilde{\Gamma}_t) - c(X'_t,\Delta'_t,\Gamma'_t)|$$
$$= |E^x_{\pi',\sigma'}[c(\tilde{X}_t,\Delta'_t,\tilde{\Gamma}_t) - c(X'_t,\Delta'_t,\Gamma'_t)]|, \text{ from the definition of } \tilde{\Delta}_t,$$
$$\leq |E^x_{\pi',\sigma'}[L(\Delta'_t,\tilde{\Gamma}_t) - L(\Delta'_t,\Gamma'_t)]| + |E^x_{\pi',\sigma'}[c\cdot(\tilde{X}_t - X'_t)]|$$
$$+ |E^x_{\pi',\sigma'}[KI_{(0,\infty)}(\Delta'_t - \tilde{X}_t) - KI_{(0,\infty)}(\Delta'_t - X'_t)]|, \text{ from (3.1)},$$
$$\leq \varepsilon_1(L + L') + 2\varepsilon_1 K + \varepsilon_1(3M + K), \text{ from (3.4) - (3.6)}.$$

Therefore, for any $\varepsilon > 0$ there exists $\varepsilon_1 > 0$ such that

$$(3.7)\qquad |E^x_{\pi',\sigma'}[c(\tilde{X}_t,\tilde{\Delta}_t,\tilde{\Gamma}_t)] - E^x_{\pi',\sigma'}[c(X'_t,\Delta'_t,\Gamma'_t)] \leq \varepsilon/2.$$

By (3.3) and (3.7), we get $|\psi(s,\pi,\sigma) - \psi'(x,\pi',\sigma')| \leq \varepsilon/2$.

Conversely, for any $\sigma' \in \Sigma'$, let $\sigma \in \Sigma$ be a strategy constructed from $\sigma' \in \Sigma'$ using the random transformation ($\nu$). Then, similarly as the above discussion we can prove that for any $\varepsilon > 0$ there is $\varepsilon_1 > 0$ such that $|\psi(x,\pi,\sigma) - \psi'(x,\pi',\sigma')| \leq \varepsilon/2$, which completes the proof. Q.E.D.

Lemma 3.4. Suppose that Conditions in theorem 3.2 hold. Then for any $\varepsilon > 0$ there exist $\varepsilon_1 > 0$ satisfying the following:

For any $(s,S)$ ordering policy $\pi' \in \Pi'$, there exists a random $(s,S)$ ordering policy $\pi \in \Pi$ such that for any $\sigma \in \Sigma$ there is $\sigma' \in \Sigma'$ for which (3.2) holds and conversely for any $\sigma' \in \Sigma'$ there existe $\sigma \in \Sigma$ satisfying (3.2).

Proof: For any $(s,S)$ ordering policy $\pi' \in \Pi'$, we construct a random $(s,S)$ ordering policy $\pi$ from $\pi'$ using the random transformation ($\nu$). Then, similarly as the proof of Lemma 3.3 we can prove that this $\pi$ has the desired property. Q.E.D.

PROOF OF THEOREM 3.2. We try to approximate the inventory game model by a subsidiary inventory model $G(\mathcal{F}\varepsilon_1)$. For any $\varepsilon > 0$, let $\varepsilon_1$ be such that Lemma 3.3 and 3.14 hold. In $G(\mathcal{F}\varepsilon_1)$, if we define $\gamma(\cdot)$ by $\gamma(D) = \varepsilon_1\mu(D \cap [-2M,-M])/2M$ for $D \in \mathcal{B}_S$, we observe that for $x \in S$, $a \in A(x)$ and $F' \in \mathcal{F}\varepsilon_1$,

$$\bar{Q}(D|x,a,F') = Q(D|x,a,F') - \gamma(D)$$
$$\geq \varepsilon_1 \int_D (\phi(a-y) - (2M)^{-1} I_{[-2M,-M]}(y))d\mu$$
$$= \varepsilon_1/2 > 0,$$

where $\mu$ is the Lebesque measure.

This means that Condition E holds for $G(\mathcal{F}_{\varepsilon_1})$.

Therefore, by Theorem 3.1 there exists a minimax $(s,S)$ ordering policy $f^\infty \in \Pi'$ for which

$$(3.8)\qquad \inf_{\pi' \in \Pi'} \sup_{\sigma' \in \Sigma'} \psi'(x,\pi',\sigma') = \sup_{\sigma' \in \Sigma'} \psi'(x,f^\infty,\sigma').$$

Applying Lemma 3.4, there exists a random $(s,S)$ ordering policy $\pi^* \in \Pi$ for which the properties in Lemma 3.4 hold.
For this $\pi^*$, we have

$$\sup_{\sigma \in \Sigma} \psi(x,\pi^*,\sigma) \leq \sup_{\sigma' \in \Sigma'} \psi'(x,f^\infty,\sigma') + \varepsilon/2,$$
from Lemma 3.4,
$$= \inf_{\pi' \in \Pi'} \sup_{\sigma' \in \Sigma'} \psi'(x,\pi',\sigma') + \varepsilon/2, \text{ from (3.8)},$$
$$\leq \inf_{\pi \in \Pi} \sup_{\sigma \in \Sigma} \psi(x,\pi,\sigma) + \varepsilon,$$
from Lemma 3.3,

which implies that the random $(s,S)$ ordering policy $\pi^*$ is $\varepsilon$-minimax. Q.E.D.

Remark: Let $\mathcal{F}(\mu,\sigma^2)$ be the class of distribution functions $F$ on $R^+$ such that $\int x dF(x) = \mu$ and $\int x^2 dF(x) = \mu^2 + \sigma^2$ where $\mu$ and $\sigma^2$ are finite constants. We suppose that the holding and penalty cost functions are both linear. Then, since $\mathcal{F}(\mu,\sigma^2)$ is a Borel set and Condition D is satisfied, it holds from Theorem 3.2 that for any $\varepsilon > 0$ an $\varepsilon$-minimax random $(s,S)$ ordering policy exists for $\mathcal{F} = \mathcal{F}(\mu,\sigma^2)$.

We note that Nakagami [10] has studied the inventory problem with the unbounded lower semi-continuous cost function and by using weighted supremum norms and the Banach contraction principle derived the optimal inventory equation for the discounted case.

## Acknowledgement

The author wishes to express his thanks to referees for their very helpful comments which led to an improvement of the presentation of the material.

## References

[1] Bertsekas, D.P. and Shreve, S.D.: *Stochastic Optimal Control - The Discrete Time Case*, Academic Press, 1978.

[2] Hoffman, A.J. and Karp, R.M.: On non-terminating stochastic games, *Management Sci.*, 12 (1966), 359-370.

[3] Iglehart, D.L.: Dynamic Programming and Stationary Analysis of Inventory Problem, Chap. 1 in H. Scarf, D. Gilford, and M. Shelly (eds.), *Multistage Model and Techniques*, Stanford University Press, Stanford, 1963.

[4] Iglehart, D.L: Optimality of $(s,S)$ Inventory Policy in the Infinite Horizon Dynamic Inventory Problem, *Management Sci.*, 9 (1963), 159-267.

[5] Jagannathan, R.: Minimax Ordering for the Infinite stage Dynamic Inventory Problem, *Management Sci.*, 24 (1978), 1138-1149.

[6] Kurano, M.: Semi-Markov decision processes and their applications in replacement model, *J. Oper. Res. Soc. Japan*, 28 (1985), 18-29.

[7] Kurano, M.: Markov decision processes with a Borel measurable cost function - the average case, *Math. Oper. Res.*, 11 (1986), 309-320.

[8] Maitra, A.: Discounted dynamic programming on compact metric space, *Sankhya Ser.* A, 30 (1968), 211-216.

[9] Maitra, A. and Parthasarathy, T.: On stochastic games, *J. Opt. Theory and Appl.*, 5 (1970), 289-300.

[10] Nakagami, J.: A contraction principle for the optimal inventory equation, *Technical Reports of Mathematical Sciences no.*1, Chiba University, Chiba, 260 Japan 1985.

[11] Nowak, A.S.: Existence of equilibrium stationary strategies in discounted noncooperative stochastic games with uncountable state space, *J. Opt. Theory and Appl.*, 45 (1985), 591-602.

[12] Shreve, S.E. and Bertsekas, D.P.: Alternative theoretical frameworks for finite horizon discrete time stochastic optimal control, *SIAM J. Contr. Optimiz.*, 16 (1978), 953-978.

[13] Tanaka, K., Iwase, S. and Wakuta, K.: On Markov games with the expected average reward criterion, *Sci. Rep. Niigata Univ., Ser.* A. No.13, (1976), 31-41.

Masami KURANO: Department of Mathematics, Faculty of Education, Chiba University, Yayoi-cho, Chiba, 260, Japan.