Zero and non-zero sum risk-sensitive Semi-Markov games

Arnab Bhabak and Subhamay Saha

Department of Mathematics, Indian Institute of Technology Guwahati, Guwahati, Assam, India

ABSTRACT

In this article we consider zero and non-zero sum risk-sensitive average criterion games for semi-Markov processes with a finite state space. For the zero-sum case, under suitable assumptions we show that the game has a value. We also establish the existence of a stationary saddle point equilibrium. For the non-zero sum case, under suitable assumptions we establish the existence of a stationary Nash equilibrium.

1. Introduction

In this paper we consider both zero and non-zero sum risk-sensitive average criterion games for semi-Markov processes. The state space is assumed to be finite and action spaces are Borel. We also assume that the sojourn times are supported on a fixed compact interval. Under general continuity-compactness assumptions and an additional assumption of irreducibility, we show that the zero-sum game admits a value. We also prescribe a saddle point equilibrium which is given by minimizing and maximizing selectors of a pair of optimality equations. For the non-zero sum game problem, under certain additional assumptions we show the existence of a Nash equilibrium. In the non-zero sum case the main step involves showing the existence of solution of a coupled system of equations. In the analysis of both the zero-sum and non-zero sum games, risk sensitive games for discrete-time Markov chains serve as an important intermediate step.

Stochastic games have found applications in diverse areas like economics, computer networks, evolutionary biology and so on. Generally in stochastic control problems, of which stochastic games are a sub-class, the expectation is minimized or maximized. But the obvious practical issue with this approach is that it does not take into account the controller’s attitude toward risk. This motivates the study of risk-sensitive control problems where the expectation of the exponential of the random quantity is considered. Since the pioneering work of Howard and Matheson [1], there has been a lot of work...
on risk-sensitive control of both discrete and continuous time stochastic processes. Risk sensitive games for discrete time Markov chains has been studied by several authors, see for instance [2-4] for zero-sum games and [5, 6] for non-zero sum games. Risk-sensitive games for continuous-time diffusions has been studied in [7-9]. Similarly, risk-sensitive games for continuous-time Markov chains has been studied in [10-12]. The literature on risk-sensitive control of semi-Markov processes is very limited. But, in practical applications assuming that the holding time distribution is memory less can be unreasonable in many situations. Such situations can not be modeled using Markovian dynamics and semi-Markov process is the most natural generalization. In [13], the authors study risk-sensitive control problem for semi-Markov processes on the finite horizon. Risk-sensitive infinite horizon discounted cost problem is considered in [14]. In [15], the authors consider the risk-sensitive average cost criterion for semi-Markov processes. Again, in practical scenarios since there are multiple controllers involved, games appear more naturally as compared to control problems. But to the best of our knowledge, the present paper is the first work on risk-sensitive semi-Markov games.

The rest of the paper is organized as follows. In Section 2, we describe the zero-sum game problem under consideration. In Section 3 we introduce the optimality equations and establish its solution. In Section 4, we describe the non-zero sum game problem. Section 5 establishes the existence of Nash equilibrium for the non-zero sum game. In Section 6 we make some concluding comments.

2. Zero-Sum game model

The risk-sensitive zero-sum semi-Markov game model that we consider here is given by

$$\left( S, A, B, \{A(i) \subset A, B(i) \subset B, i \in S\}, C(i, a, b), \{\rho_{(i, a, b)}(\cdot)\}, \{F_{i, a, b}\}, \left[p_{i, j}(a, b)\right]\right), \quad (1)$$

where,

- $S$ is the state space, which is assumed to be finite and is endowed with the discrete topology.
- The Borel spaces $A$ and $B$ are the action sets for player 1 and 2 respectively. And for each $i \in S$, $A(i) \subset A$, $B(i) \subset B$ are Borel subsets denoting the set of all admissible actions in state $i$ for player 1 and 2 respectively.
- Define $\mathbb{K} = \{(i, a, b) : i \in S, a \in A(i), b \in B(i)\}$ to be the set of admissible state-action pairs. Then $C : \mathbb{K} \rightarrow \mathbb{R}$ is the immediate cost function for player 1 and immediate reward for player 2.
- For each $(i, a, b) \in \mathbb{K}$, the mapping $\rho_{(i, a, b)} : [0, \infty) \rightarrow \mathbb{R}$ denotes the running cost function for player 1 and running reward function for player 2.
- $F_{i, a, b}$ is the sojourn time distribution function for both the players in state $i$ under the actions $a$ and $b$. It is assumed that the sojourn times are positive, so that

$$F_{i, a, b}(0) = 0, \quad (i, a, b) \in \mathbb{K} \quad (2)$$

- Finally, $[p_{i, j}(a, b)]$ is the controlled transition law and satisfies $\sum_{j \in S} p_{i, j}(a, b) = 1$ for every $(i, a, b) \in \mathbb{K}$. 


The game evolves in the following manner. At the initial time \( t=0 \), the process starts at \( X_0 = i_0 \in S \). Suppose player 1 chooses an action \( A_0 = a_0 \in A(i_0) \) and player 2 independently chooses an action \( B_0 = b_0 \in B(i_0) \). As a result player 2 gets an immediate reward \( C(i_0, a_0, b_0) \) from player 1. Player 1 also incurs a holding cost at the rate \( \rho_{(i_0, a_0, b_0)} \). The process stays in state \( i \) for a random amount of time \( S_0 \) whose distribution function is given by \( F_{i_0, a_0, b_0} \) and then jumps to a new state \( X_1 = i_1 \) with probability \( p_{i_0, i_1}(a_0, b_0) \). Immediately after the first transition, players 1 and 2 chooses actions \( A_1 = a_1 \in A(i_1) \) and \( B_1 = b_1 \in B(i_1) \). The same sequence of events as described above repeats itself. Let \( T_n \) to be the time when the \( n \)th transition is completed, then

\[
T_0 = 0 \quad \text{and} \quad T_n = \sum_{i=0}^{n-1} S_i \quad n = 1, 2, \ldots, \tag{3}
\]

where \( S_n, n = 0, 1, 2, \ldots \) denotes the random sojourn times at the \( n \)th state. We denote the number of transitions \( N_t \) in the interval \([0, t] \) by

\[
N_t = \sup\{n \in \mathbb{N} : T_n \leq t\}, \quad t \geq 0. \tag{4}
\]

Let \( \mathcal{H}_n \) be the information available upto time \( T_n \), i.e., \( \mathcal{H}_0 = X_0 \) and for \( n \geq 1, \mathcal{H}_n = \{X_0, A_0, B_0, S_0, \ldots, X_{n-1}, A_{n-1}, B_{n-1}, S_{n-1}, X_n\} \), where for \( n \geq 0, X_n \) is the \( n \)th state, \( A_n \) and \( B_n \) are the actions of player 1 and 2 respectively at the \( n \)th transition time and \( S_n \) is the sojourn time at the \( n \)th state. For \( n \geq 0 \), we also define the admissible history spaces \( H_n \) by \( H_0 = S \) and \( H_n = \mathbb{K} \times (0, \infty) \times H_{n-1} \) for \( n = 1, 2, \ldots \). We endow these spaces with the Borel sigma-algebra. Now we introduce the concept of policies.

**Definition 1.** A randomized history dependent policy or simply a policy for player 1 is a sequence \( \pi^1 = \{\pi^1_n : n \geq 0\} \) of stochastic kernels \( \pi^1_n \) on \( A \) given \( H_n \) such that

\[
\pi^1_n(A(i_n)|h_n) = 1 \quad \forall \ h_n \in H_n, n = 0, 1, \ldots, \tag{5}
\]

A randomized history dependent policy for player 2 can be defined analogously.

Let \( \Phi^1 \) be the set of all stochastic kernels \( \phi^1 \) on \( A \) given \( S \) satisfying \( \phi^1(A(i)|i) = 1 \). A policy \( \pi^1 \) for player 1 is said to be stationary if there exists a stochastic kernel \( \phi^1 \in \Phi^1 \) such that \( \pi^1_n(\cdot|h_n) = \phi^1(\cdot|i_n) \) for all \( h_n = (i_0, a_0, b_0, s_0, \ldots, i_{n-1}, a_{n-1}, b_{n-1}, s_{n-1}, i_n) \in H_n \) and \( n = 0, 1, \ldots, \). We will identify a stationary policy \( \pi^1 \) with \( \phi^1 \). Similarly stationary policies for player 2 can be defined.

For each \( m = 1, 2, \Pi_m \) and \( \Phi^m \) represent the set of all randomized history dependent strategies and the set of all stationary strategies for player \( m \), respectively. We will have the following assumptions on our model.

**Assumption 1.**

(i) For each \( i \in S \), the set \( A(i) \) and \( B(i) \) are compact subsets of \( A \) and \( B \).

(ii) For each \( i, j \in S, (a, b) \rightarrow C(i, a, b) \) and \( (a, b) \rightarrow p_{ij}(a, b) \) are continuous in \( (a, b) \in A(i) \times B(i) \).

(iii) The family \( \{F_{i, a, b}\} \) is supported on a compact interval and is weakly continuous, that is, there exists \( B > 0 \) such that

\[
F_{i, a, b}(B) = 1, \quad (i, a, b) \in \mathbb{K}, \tag{5}
\]
and for each \( i \in S \) and \( u \) bounded measurable, \((a, b) \mapsto \int_{0}^{B} u(s) dF_{i,a,b}(s)\) is continuous in \((a, b) \in A(i) \times B(i)\).

(iv) For every \( i \in S \), the mapping \((a, b, s) \mapsto \rho_{(i,a,b)}(s)\) is continuous in \((a, b, s) \in A(i) \times B(i) \times [0,B]\).

Since the spaces \( A(i) \) and \( B(i) \) are compact and the state space is finite, so it follows by Assumption 1 that,

\[
M_{\rho} := \sup_{(i,a,b) \in A(i) \times B(i)} \left| \rho_{(i,a,b)}(s) \right| < \infty. \tag{6}
\]

Given the initial state \( X_{0} = i \) and a pair of policies \((\pi^{1}, \pi^{2})\), the distribution of \(\{(X_{n}, A_{n}, B_{n}, S_{n})\}\) is uniquely determined by the Tulcea theorem [16]. We denote such a distribution by \(\mathbb{P}^{\pi^{1}, \pi^{2}}\), and \(\mathbb{E}^{\pi^{1}, \pi^{2}}\) be the corresponding expectation operator. The following Markov relations are satisfied almost surely under each distribution \(\mathbb{P}^{\pi^{1}, \pi^{2}}\): For each \( i, j \in S \), \( C \) Borel subset of \( A, D \) Borel subset of \( B \) and \( n \in \mathbb{N}\),

\[
\begin{align*}
\mathbb{P}^{\pi^{1}, \pi^{2}}[X_{0} = i] &= 1, \\
\mathbb{P}^{\pi^{1}, \pi^{2}}[A_{n} \in C, B_{n} \in D | \mathcal{H}_{n}] &= \pi^{1}_{n}(C|\mathcal{H}_{n}) \pi^{2}_{n}(D|\mathcal{H}_{n}), \\
\mathbb{P}^{\pi^{1}, \pi^{2}}[S_{n} \leq t| \mathcal{H}_{n}, A_{n}, B_{n}] &= F_{X_{n}, A_{n}, B_{n}}(t), \\
\mathbb{P}^{\pi^{1}, \pi^{2}}[X_{n+1} = j| \mathcal{H}_{n}, A_{n}, B_{n}, S_{n}] &= p_{X_{n},j}(A_{n}, B_{n}).
\end{align*} \tag{7}
\]

Now we describe the evaluation criterion for our game. The total cost incurred by player 1 and the total reward gained by player 2 up to time \( t > 0 \) is given by:

\[
C_{t} = \sum_{k=0}^{N_{i} - 1} \left[ C(X_{k}, A_{k}, B_{k}) + \int_{0}^{S_{k}} \rho_{(X_{k}, A_{k}, B_{k})}(r) dr \right] + C(X_{N_{i}}, A_{N_{i}}, B_{N_{i}}) \\
+ \int_{0}^{t-T_{N_{i}}} \rho_{(X_{N_{i}}, A_{N_{i}}, B_{N_{i}})}(r) dr. \tag{8}
\]

For risk-sensitivity parameter \( \theta > 0 \) and a policy pair \((\pi^{1}, \pi^{2})\) define,

\[
J_{0}(i, \pi^{1}, \pi^{2}) := \lim_{t \to \infty} \frac{1}{\theta t} \log \left[ \mathbb{E}^{\pi^{1}, \pi^{2}}_{i}(e^{\theta C_{t}}) \right]. \tag{9}
\]

We further make the following definitions.

\[
L(i, \theta) = \sup_{\pi^{2} \in \Pi_{2}} \inf_{\pi^{1} \in \Pi_{1}} J_{0}(i, \pi^{1}, \pi^{2}), \\
U(i, \theta) = \inf_{\pi^{1} \in \Pi_{1}} \sup_{\pi^{2} \in \Pi_{2}} J_{0}(i, \pi^{1}, \pi^{2}),
\]

where \( J_{0}(i, \pi^{1}, \pi^{2}) \) is defined in (9). \( L(\cdot) \) is called the lower value of the game and \( U(\cdot) \) is called the upper value of the game.

**Definition 2.** If \( U(i, \theta) = L(i, \theta) \) for all \( i \in S \), then we say that the game has a value. If the game has a value, then the common function is referred to as the value function of the game and will be denoted by \( V(\cdot) \).
Here player 1 is interested in minimizing \( J_0(i, \pi^1, \pi^2) \) over \( \pi^1 \in \Pi_1 \) for each \( \pi^2 \in \Pi_2 \), and player 2 wants to maximize \( J_0(i, \pi^1, \pi^2) \) over \( \pi^2 \in \Pi_2 \) for each \( \pi^1 \in \Pi_1 \). This motivates the following definition.

**Definition 3.** Suppose that the value of the game exists. A policy \( \pi^{*1} \in \Pi_1 \) is said to be optimal for player 1, if for any \( i \in S \),

\[
V(i, \theta) = \sup_{\pi^2 \in \Pi_2} J_0(i, \pi^{*1}, \pi^2), \quad \forall i \in S.
\]

Similarly, for player 2 a policy \( \pi^{*2} \in \Pi_2 \) is said to be optimal, if for any \( i \in S \),

\[
V(i, \theta) = \inf_{\pi^1 \in \Pi_1} J_0(i, \pi^1, \pi^{*2}), \quad \forall i \in S.
\]

If \( \pi^{*m} \in \Pi_m \) is optimal for player \( m(m = 1, 2) \), then \((\pi^{*1}, \pi^{*2}) \in \Pi_1 \times \Pi_2 \) is called a saddle point equilibrium.

### 3. Analysis of Zero-Sum game

For \( i \in S \), let \( \mathcal{P}(A(i)) \) and \( \mathcal{P}(B(i)) \) denote the set of all probability measures on \( A(i) \) and \( B(i) \) respectively. The analysis of the zero-sum game crucially depends on the following equation.

\[
e^{\theta h(i)} = \sup_{\phi \in \mathcal{P}(B(i))} \inf_{\psi \in \mathcal{P}(A(i))} \left[ \int_{A(i)} \int_{B(i)} \psi(da) \phi(db) e^{\theta \psi C(i,a,b)} e^{\theta \psi C(i,a,b)} \int_0^B e^{\theta \psi C(i,a,b)} dF_{i,a,b}(s) \right] \times \sum_{j \in S} e^{\theta h(j)} p_j(a,b), \quad i \in S.
\]

(10)

where \( g \) is a real number and \( h(.) \) is a real function defined on the state space \( S \). Using Assumption 1 and Fan’s minimax theorem [17], equation (10) can also be written as:

\[
e^{\theta h(i)} = \inf_{\psi \in \mathcal{P}(A(i))} \sup_{\phi \in \mathcal{P}(B(i))} \left[ \int_{A(i)} \int_{B(i)} \psi(da) \phi(db) e^{\theta \psi C(i,a,b)} e^{\theta \psi C(i,a,b)} \int_0^B e^{\theta \psi C(i,a,b)} dF_{i,a,b}(s) \right] \times \sum_{j \in S} e^{\theta h(j)} p_j(a,b), \quad i \in S.
\]

(11)

The importance of the above equations is illustrated by the next theorem.

**Theorem 1.** Suppose that equation (10) and hence equation (11) is satisfied by a pair \((g, h(.)\). Under Assumption 1, it follows that the game has a value and is given by \( g = V(i, \theta) \). Further if \( \phi^{*1} \in \Phi^1 \) is the outer minimizing selector of the right hand side of (11) and if \( \phi^{*2} \in \Phi^2 \) is the outer maximizing selector of the right hand side of (10), then \((\phi^{*1}, \phi^{*2}) \) is a saddle point equilibrium.

In order to prove Theorem 1, we need the following auxiliary lemma.
Lemma 1. Suppose Assumption 1 holds. Then the following holds:

(i) Given $\alpha \in (0, 1)$, there exists an integer $r_\alpha > 0$ such that, for every $(i, a, b) \in K$, the inequality $\int_0^B e^{-r_\alpha} dF_{i,a,b}(s) \leq \alpha$ holds for every $r \geq r_\alpha$.

(ii) For each $\alpha \in (0, 1)$, $t \geq 0$ and $n \in \mathbb{N}$, $\mathbb{P}_{i}^{\pi_1, \pi_2}[N_t \geq n] \leq \alpha^n e^{r_\alpha t}$ for all $i \in S$ and $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$, where $r_\alpha$ is as in part (ii). Thus,

$$\mathbb{P}_{i}^{\pi_1, \pi_2}[N_t < \infty] = 1. \quad (12)$$

Proof. The proof is a simple generalization of Lemma 4.1 in [15]. \hfill \Box

Proposition 1. Let $(g, h(.))$ be a solution of equation (10) and hence of equation (11). Under Assumption 1, the following are true.

For each $i \in S$, $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$ and $t > 0$:

$$e^{\theta \bar{h}(i)} \geq \mathbb{E}_{i}^{\phi^{1}, \pi_2}\left[ e^{\frac{1}{2} \left( \sum_{k=1}^{N_t} C(X_{i,k}, a_k, b_k) + \int_0^t \varphi^{0_2}(s) ds - gT_{i+1} + h(X_{i+1}) \right) - \frac{1}{2} \sum_{a,b} \mathbb{P}(B(i)) \right], \quad (13)$$

and also we have,

$$e^{\theta \bar{h}(i)} \leq \mathbb{E}_{i}^{\pi_1, \phi^{2}}\left[ e^{\frac{1}{2} \left( \sum_{k=1}^{N_t} C(X_{i,k}, a_k, b_k) + \int_0^t \varphi^{0_2}(s) ds - gT_{i+1} + h(X_{i+1}) \right) - \frac{1}{2} \sum_{a,b} \mathbb{P}(B(i)) \right], \quad (14)$$

where $\phi^{1}$ and $\phi^{2}$ are as in Theorem 1.

Proof. From (11) we have for any $\varphi \in \mathcal{P}(B(i))$

$$e^{\theta \bar{h}(i)} \geq \int_{\mathcal{A}(i)} \int_{B(i)} \varphi^{1}(da) \varphi(db) e^{\theta C(i,a,b)} \int_0^B e^{\theta \left( \int_0^t \varphi^{0_2}(s) ds - gT_{i+1} + h(X_{i+1}) \right) - \frac{1}{2} \sum_{a,b} \mathbb{P}(B(i))} dF_{i,a,b}(s) \times \sum_{j \in S} e^{\theta \bar{h}(j)} p_{ij}(a,b), \quad i \in S. \quad (15)$$

Thus for any $\pi_2 \in \Pi_2$ we have,

$$e^{\theta \bar{h}(i)} \geq \mathbb{E}_{i}^{\phi^{1}, \pi_2}\left[ e^{\theta \left( \sum_{k=1}^{N_t} C(X_{i,k}, a_k, b_k) + \int_0^t \varphi^{0_2}(s) ds - gT_{i+1} + h(X_{i+1}) \right) - \frac{1}{2} \sum_{a,b} \mathbb{P}(B(i))} \right], \quad i \in S. \quad (16)$$

More generally, via equations (7) it follows that for every $n \in \mathbb{N}$,

$$e^{\theta \bar{h}(n)} \geq \mathbb{E}_{i}^{\phi^{1}, \pi_2}\left[ e^{\theta \left( \sum_{k=1}^{N_t} C(X_{i,k}, a_k, b_k) + \int_0^t \varphi^{0_2}(s) ds - gT_{i+1} + h(X_{i+1}) \right) - \frac{1}{2} \sum_{a,b} \mathbb{P}(B(i))} \right]_{\mathcal{H}_n}, \quad i \in S. \quad (17)$$

We prove by induction that for every non-negative integer $n$,

$$e^{\theta \bar{h}(i)} \geq \mathbb{E}_{i}^{\phi^{1}, \pi_2}\left[ e^{\theta \left( \sum_{k=1}^{N_t} C(X_{i,k}, a_k, b_k) + \int_0^t \varphi^{0_2}(s) ds - gT_{i+1} + h(X_{i+1}) \right) - \frac{1}{2} \sum_{a,b} \mathbb{P}(B(i))} 1_{[N_t \leq n]} \right] \quad (18)$$

$$+ \mathbb{E}_{i}^{\phi^{1}, \pi_2}\left[ e^{\theta \left( \sum_{k=1}^{N_t} C(X_{i,k}, a_k, b_k) + \int_0^t \varphi^{0_2}(s) ds - gT_{i+1} + h(X_{i+1}) \right) - \frac{1}{2} \sum_{a,b} \mathbb{P}(B(i))} 1_{[N_t > n]} \right].$$

To show this, from (15) we get,
Thus the induction argument is complete. Then the Monotone convergence theorem, where (16) was used to deduce the first inequality, whereas the fact that the random variables $1_{[N_i = 0]}$ and $1_{[N_i = n]}$ are measurable was used in the last step. Since $T_{n+2} = T_{n+1} + S_{n+1}$, by (3) it follows that

$$
E_{i}^{\phi_{i}, \pi} \left[ \sum_{k=0}^{n} \left( C(X_k, A_k, B_k) + \int_{0}^{s_k} \rho(X_k, A_k, B_k)(t) dt \right) - gT_{n+1} + h(X_{n+1}) \right] 1_{[N_i > n]}
$$

so, together with the induction hypothesis it follows that (17) is also valid for $n + 1$. Thus the induction argument is complete. Then the Monotone convergence theorem, together with (12) gives,

$$
\lim_{n \to \infty} E_{i}^{\phi_{i}, \pi} \left[ \sum_{k=0}^{n} \left( C(X_k, A_k, B_k) + \int_{0}^{s_k} \rho(X_k, A_k, B_k)(t) dt \right) - gT_{n+1} + h(X_{n+1}) \right] 1_{[N_i \leq n]} = E_{i}^{\phi_{i}, \pi} \left[ \sum_{k=0}^{N} \left( C(X_k, A_k, B_k) + \int_{0}^{s_k} \rho(X_k, A_k, B_k)(t) dt \right) - gT_{n+1} + h(X_{n+1}) \right] .
$$
Now using Assumption 1 and Lemma 1 we get that
\[ \mathbb{E}_t^{\theta t^2} \left[ e^{\theta \sum_{k=0}^n (C(X_k, A_k, B_k) + \int_0^{S_k} \rho(X_k, A_k, B_k)(r)dr) - gT_{n+1}} \right] \to 0 \quad \text{as } n \to \infty. \]

Now taking \( n \to \infty \) on both the sides of (17) and using the last convergence and (18) we get the desired inequality (13).

The other inequality (14) also follows analogously starting from (10).

**Proof of Theorem 1.** We have, \( T_{N_t} \leq t < T_{N_{t+1}} = T_{N_t} + S_{N_t} \), for every \( t > 0 \), and thus
\[ 0 \leq t - T_{N_t} \leq S_{N_t} \leq B \quad \text{and} \quad T_{N_{t+1}} - t \leq S_{N_t} \leq B. \] (19)

Now from (8) we have
\[ \sum_{k=0}^{N_t} C(X_k, A_k, B_k) + \int_0^{S_k} \rho(X_k, A_k, B_k)(r)dr - gT_{N_{t+1}} = (C_t - tg) + \int_{t-T_{N_t}}^{S_{N_t}} \rho(X_{N_{t+1}}, A_{N_{t+1}}, B_{N_{t+1}})(r)dr - (T_{N_{t+1}} - t)g. \]

and together with the equality (6) and (19) it follows that
\[ \sum_{k=0}^{N_t} C(X_k, A_k, B_k) + \int_0^{S_k} \rho(X_k, A_k, B_k)(r)dr - gT_{N_{t+1}} - (C_t - tg) \leq B(M_p + |g|). \] (20)

Using (14) we get that \( e^{-2\theta||h||} \leq \mathbb{E}_t^{\pi^1, \theta^2} \left[ e^{\theta \sum_{k=0}^n (C(X_k, A_k, B_k) + \int_0^{S_k} \rho(X_k, A_k, B_k)(r)dr) - gT_{n+1}} \right]. \) Using (3), we have
\[ e^{-2\theta||h||} \leq \mathbb{E}_t^{\pi^1, \theta^2} \left[ e^{\theta \sum_{k=0}^n (C(X_k, A_k, B_k) + \int_0^{S_k} \rho(X_k, A_k, B_k)(r)dr) - gT_{n+1}} \right], \]
so that \( e^{-2\theta||h|| - \theta B(M_p + |g|) + \theta tg} \leq \mathbb{E}_t^{\pi^1, \theta^2} \left[ e^{\theta C_1} \right]. \) Taking logarithm on both sides, dividing by \( \theta t \) and then taking limit \( t \to \infty \) we get,
\[ \theta \leq J_\theta(i, \pi^1, \phi^2), \quad \forall i \in S. \]

For the other inequality consider inequality (13). Then proceeding similarly as above we have the following inequality,
\[ e^{2\theta||h|| + \theta B(M_p + |g|) + \theta tg} \geq \mathbb{E}_t^{\phi^1, \pi^2} \left[ e^{\theta C_1} \right]. \]

Again taking logarithm on both sides, dividing by \( \theta t \) and then taking limit \( t \to \infty \) we get,
\[ \theta \geq J_\theta(i, \phi^1, \pi^2), \quad \forall i \in S. \]

Since \((\pi^1, \pi^2)\) was arbitrary, we get
\[ g \leq \inf_{\pi^1 \in \Pi_1} J_\theta(i, \pi^1, \phi^2) \leq L(i, \theta) \leq U(i, \theta) \leq \sup_{\pi^2 \in \Pi_2} J_\theta(i, \phi^1, \pi^2) \leq g. \]

Hence we have the desired conclusions.

In view of Theorem 1, in order to establish the existence of the value of the game and saddle point equilibrium, it suffices to show the existence of solution of the optimality equation (10). For that we impose one more assumption on our model.
Assumption 2. Under each stationary policy, the embedded discrete-time Markov chain \( \{X_n\} \) is irreducible.

In order to establish the existence of solution of (10), we first consider risk-sensitive average criterion game problem for the discrete time process \( \{X_n\} \). For that we consider policies \( (\pi^1, \pi^2) \), where for each positive integer \( n \), the kernels \( (\pi^1_n, \pi^2_n) \) depends only on \( X_0, A_0, B_0, X_1, \ldots, X_{n-1}, A_{n-1}, B_{n-1}, X_n \). Given a bounded continuous function \( D \) on \( \mathbb{K} \), define the discrete-time average at \( h \) policies and consider the value function

\[
V_{h,D}(i, \pi^1, \pi^2) := \lim_{n \to \infty} \sup \frac{1}{\theta n} \ln \left( \mathbb{E}^{\pi^1, \pi^2}_{i} \left[ e^{\theta \sum_{k=0}^{n-1} D(X_k, A_k, B_k)} \right] \right)
\]

and \( \theta \)-optimal discrete time average value function, if it exists, is given by

\[
V_{h,D}^*(i) := \inf_{\pi^1} \sup_{\pi^2} V_{h,D}(i, \pi^1, \pi^2) = \sup_{\pi^1} \inf_{\pi^2} V_{h,D}(i, \pi^1, \pi^2)
\]

It is easy to see that the value function \( V_{h,D}^*(\cdot) \) satisfies the following.

\[
V_{h,D}^*(\cdot) \leq V_{h,D_1}^*(\cdot) \quad \text{if} \quad D \leq D_1 \quad \text{and} \quad V_{h,c+D}^*(\cdot) = c + V_{h,D}^*(\cdot)
\]

where \( c \in \mathbb{R} \). Since \( D \leq D + \|D - D_1\| \) it follows that \( V_{h,D}^*(\cdot) \leq V_{h,D_1}^*(\cdot) + \|D - D_1\| \).

Similarly, by interchanging the roles of \( D \) and \( D_1 \), this yields that

\[
\|V_{h,D}^*(\cdot) - V_{h,D_1}^*(\cdot)\| \leq \|D - D_1\|.
\]

Observing that \( V_{h,0}^* = 0 \), the monotonicity property in (23) yields that, for bounded continuous functions \( D, D_1 \),

\[
V_{h,D}^* \leq 0 \leq V_{h,D_1}^*, \quad \text{when} \quad D \leq 0 \leq D_1.
\]

We have the following theorem.

Theorem 2. Under Assumptions 1 and 2, we have the following:

(i) For each bounded continuous function \( D \) on \( \mathbb{K} \) there exist \( \mu_D \in \mathbb{R} \) and \( h_D : S \to \mathbb{R} \) such that

\[
e^{\theta |\mu_D + h_D(i)|} = \sup_{\varphi \in \mathcal{P}(B(i))} \inf_{\psi \in \mathcal{P}(A(i))} \left[ \int_{A(i)} \int_{B(i)} \psi(da) \varphi(db) e^{\theta D(i, a, b)} \sum_{j \in S} p_{i,j}(a, b) e^{\theta h(j)} \right]
\]

and \( \mu_D = V_{h,D}^*(i), \quad \text{for} \quad i \in S. \)

(ii) For bounded continuous functions \( D, D_1 \),

\[
|\mu_D - \mu_{D_1}| \leq \|D - D_1\|.
\]

Proof. The proof of (i) follows by putting together arguments and results from the existing literature on risk-sensitive control of discrete-time Markov chains. We just outline the steps and cite appropriate references.
Step 1: Using a standard contraction argument similar to the proof of Theorem 3.1(a) of [6] it can be shown that for each \( \beta \in (0,1) \) there exists function \( V_\beta(\cdot) \) on \( S \) satisfying

\[
e^{\theta V_\beta(i)} = \sup_{\psi \in \mathcal{P}(B(i))} \inf_{\phi \in \mathcal{P}(A(i))} \left[ \int_{A(i)} \int_{B(i)} \psi(da) \phi(db) e^{\theta D(i,a,b)} \sum_{j \in S} p_{i,j}(a,b) e^{\theta \beta V_\beta(j)} \right]
\]

\[
= \inf_{\psi \in \mathcal{P}(A(i))} \sup_{\phi \in \mathcal{P}(B(i))} \left[ \int_{A(i)} \int_{B(i)} \psi(da) \phi(db) e^{\theta D(i,a,b)} \sum_{j \in S} p_{i,j}(a,b) e^{\theta \beta V_\beta(j)} \right], \quad i \in S.
\]

(28)

Also it is true that \( \| V_\beta \| \leq \| D \| \frac{1}{1-\beta} \).

Step 2: Fix a sequence \( \beta_n \uparrow 1 \). For \( n \geq 1 \), define

\[
z_{\beta_n} = \sup_{i \in S} V_{\beta_n}(i), \quad w_{\beta_n}(i) = V_{\beta_n}(i) - z_{\beta_n}, \quad g_{\beta_n} = (1 - \beta_n) z_{\beta_n}.
\]

From (28) and (29) we get

\[
e^{\theta w_{\beta_n}(i) + \theta g_{\beta_n}} = \sup_{\psi \in \mathcal{P}(B(i))} \inf_{\phi \in \mathcal{P}(A(i))} \left[ \int_{A(i)} \int_{B(i)} \psi(da) \phi(db) e^{\theta D(i,a,b)} \sum_{j \in S} p_{i,j}(a,b) e^{\theta \beta_n w_{\beta_n}(j)} \right]
\]

\[
= \inf_{\psi \in \mathcal{P}(A(i))} \sup_{\phi \in \mathcal{P}(B(i))} \left[ \int_{A(i)} \int_{B(i)} \psi(da) \phi(db) e^{\theta D(i,a,b)} \sum_{j \in S} p_{i,j}(a,b) e^{\theta \beta_n w_{\beta_n}(j)} \right], \quad i \in S.
\]

(31)

Now arguing as in Proposition 3.1 in [6], it can be shown that there exists a subsequence of \( \beta_n \), which we relabel as \( \beta_n \) and function \( h_D(i) \) and constant \( \mu_D \) such that \( h_D(i) = \lim_{n \to \infty} w_{\beta_n}(i) \) and \( \mu_D = \lim_{n \to \infty} g_{\beta_n} \).

Step 3: Now taking limit in (30) and using Step 2 we get (26).

Step 4: The fact that \( \mu_D = V_{\beta_D}(i) \) follows as in Lemma 2.3 in [4].

The proof of (ii) is straightforward from part (i) and (24). \( \square \)

Lemma 2. Suppose that Assumption 1 is valid and for each \( g \in \mathbb{R} \) define the function \( D_g : \mathbb{K} \to \mathbb{R} \) by

\[
D_g(i,a,b) = C(i,a,b) + \frac{1}{\theta} \ln \left( \int_0^B e^{\theta \left( \int_0^t \rho_{i,a,b}(r) dr - g \right)} dF_{i,a,b}(s) \right).
\]

(32)

The following assertions hold.

(i) \( D_g \) is bounded continuous on \( \mathbb{K} \) for each \( g \in \mathbb{R} \).

(ii) \( \| D_g - D_g' \| \leq B |g - g_1|, g, g_1 \in \mathbb{R} \).

(iii) There exist \( g^- \geq 0 \) such that \( D_{g^-} \leq 0 \).

(iv) \( D_{g^+} \geq 0 \) for some \( g^+ \leq 0 \).

Proof. The proof is a straightforward generalization of Lemma 6.1 in [15]. \( \square \)

We finally have the existence theorem.
Theorem 3. (Existence of solutions) Under Assumptions 1 and 2, there exists \( g \in \mathbb{R} \) and \( h : S \to \mathbb{R} \) such that the optimality equation (10) is satisfied.

**Proof.** For each \( g \) consider \( D_g \) given by (32). Combining Lemma 2 and Theorem 2 we get that \( \mu_{D_g} \) is continuous in \( g \). So again using Lemma 2 and intermediate value property we get the existence of a \( g \) such that \( \mu_{D_g} = 0 \). Hence we have the result from Theorem 2.

\[ \square \]

4. Non-zero sum game model

In the non-zero sum game model we assume that there is no immediate cost and individual players have their own running cost functions. For \( m = 1, 2 \), we denote the running cost function for player \( m \) by \( \rho^m \). Here the evolution of the game is similar, except for the fact that upon taking their individual actions both players incur a holding cost up to the next transition. The definition of the policies is same as the zero-sum case. Thus, the total cost up to a positive time \( t \) for player 1 is given by:

\[
C_1^t = \sum_{k=0}^{N(t)-1} \int_0^{S_k} \rho_{(X_k, A_k, B_k)}^1(r) \, dr + \int_0^{t-S_{N(t)-1}} \rho_{(X_{N(t)}, A_{N(t)}, B_{N(t)})}^1(r) \, dr,
\]

while for player 2 it is given by:

\[
C_2^t = \sum_{k=0}^{N(t)-1} \int_0^{S_k} \rho_{(X_k, A_k, B_k)}^2(r) \, dr + \int_0^{t-S_{N(t)-1}} \rho_{(X_{N(t)}, A_{N(t)}, B_{N(t)})}^2(r) \, dr.
\]

Here the objective of each player is to minimize their own average costs.

**Definition 4.** Fix a pair of policies \((\pi^1, \pi^2) \in \Pi_1 \times \Pi_2\). For \( m = 1, 2 \), define the value function for player 1 as

\[
V_1^0(\pi^2) = \inf_{\pi_1} J_1^0(i, \pi_1, \pi^2),
\]

where \( J_1^0 \) is given by (9), with \( C_t \) replaced by \( C_1^t \). Similarly, the value function for player 2 is given by \( V_2^0(\pi^1) = \inf_{\pi_2} J_2^0(i, \pi_1, \pi^2) \), where \( J_2^0 \) is given by (9), with \( C_t \) replaced by \( C_2^t \).

**Definition 5.** (Nash equilibrium) A pair of policies \((\pi^1, \pi^2) \in \Pi_1 \times \Pi_2\) is called a Nash equilibrium for the non-zero sum game if

\[
J_1^0(i, \pi^1, \pi^2) \leq J_1^0(i, \pi_1, \pi^2) \quad \text{and} \quad J_2^0(i, \pi^1, \pi^2) \leq J_2^0(i, \pi^1, \pi^2),
\]

for all \( i \in S \) and \((\pi^1, \pi^2) \in \Pi_1 \times \Pi_2\).

5. Analysis of non-zero sum game

We wish to establish the existence of Nash equilibrium for the non-zero sum game. To that end we, just like in the zero-sum case first consider a discrete time non-zero sum game given by the embedded Markov chain. Given two bounded continuous functions \( D_1 \) and \( D_2 \) on \( \mathbb{R} \), we define for \( i \in S \), under \((\pi^1, \pi^2)\), the discrete-time cost functional for player \( m, m = 1, 2 \) by...
\[ V_{0,D_m}(i, \pi^1, \pi^2) := \limsup_{n \to \infty} \frac{1}{\theta n} \ln \left( \mathbb{E}_{i} \right. \right. \left. \left. \sum_{k=0}^{\pi_n-1} D_m(X_k, A_k, B_k) \right) \right). \quad (35) \]

We have the following discrete-time theorem.

**Theorem 4.** Suppose that Assumptions 1 and 2 are satisfied. Fix a pair of stationary strategies \((\phi^1, \phi^2)\). Then there exist functions \(y_{\phi^1}^i\) and \(y_{\phi^2}^i\) on \(S\) and constants \(\mu_{\phi^1}^i\) and \(\mu_{\phi^2}^i\) such that the following are true.

\(\text{(i)}\) \[ e^{\theta y_{\phi^2}^i(i) + \theta y_{\phi^1}^i} = \inf_{\varphi \in \mathcal{P}(A(i))} \left\{ \int_{A(i)} \int_{B(i)} e^{\theta D_1(i,a,b)} \times \sum_{j \in S} e^{\theta y_{\phi^2}^j(j)} p_{i,j}(a,b) \phi^2(db|i) \psi(da) \right\} \quad \forall i \in S, \]

and \(\mu_{\phi^2}^i = \inf_{\pi^1} V_{0,D_1}(i, \pi^1, \pi^2)\).

\(\text{(ii)}\) \[ e^{\theta y_{\phi^1}^i(i) + \theta y_{\phi^1}^i} = \inf_{\varphi \in \mathcal{P}(B(i))} \left\{ \int_{A(i)} \int_{B(i)} e^{\theta D_1(i,a,b)} \times \sum_{j \in S} e^{\theta y_{\phi^1}^j(j)} p_{i,j}(a,b) \phi^1(da|i) \right\} \quad \forall i \in S, \]

and \(\mu_{\phi^1}^i = \inf_{\pi^2} V_{0,D_1}(i, \pi^1, \pi^2)\).

**Proof.** The proof again follows by putting together arguments and result from the existing literature. So like in the zero-sum case we outline the steps and cite appropriate references.

**Step 1:** Let \(x \in (0,1)\). Then again using a contraction argument similar to the proof of Theorem 3.1(a) of [6] the following can be shown.

\(\text{(a)}\) For each fixed \(\phi^2 \in \Phi^2\), there exists a function \(w_{\phi^2,x}^i\) such that

\[ e^{\theta w_{\phi^2,x}^i} = \inf_{\psi \in \mathcal{P}(A(i))} \left[ \int_{A(i)} \int_{B(i)} e^{\theta D_1(i,a,b)} \times \sum_{j \in S} e^{\theta w_{\phi^2,x}^j(j)} p_{i,j}(a,b) \phi^2(db|i) \psi(da) \right], \quad (38) \]

for all \(i \in S\).

\(\text{(b)}\) For each fixed \(\phi^1 \in \Phi^1\), there exists a function \(w_{\phi^1,x}^i\) on \(S\) such that

\[ e^{\theta w_{\phi^1,x}^i} = \inf_{\varphi \in \mathcal{P}(B(i))} \left[ \int_{A(i)} \int_{B(i)} e^{\theta D_1(i,a,b)} \times \sum_{j \in S} e^{\theta w_{\phi^1,x}^j(j)} p_{i,j}(a,b) \phi^1(da|i) \right], \quad (39) \]

for all \(i \in S\).

**Step 2:** Fix an arbitrary sequence \(\{x_n\} \in (0,1)\) satisfying \(x_n \uparrow 1\), as \(n \to \infty\). For each \(n \geq 1\) set

\[ \gamma_{\phi^2}^x = \sup_{i \in S} w_{\phi^2,x_n}^i(i), \quad \gamma_{\phi^1}^x = \sup_{i \in S} w_{\phi^1,x_n}^i(i), \]

\[ \mu_{\phi^2}^x = (1 - x_n) \gamma_{\phi^2}^x, \quad \mu_{\phi^1}^x = (1 - x_n) \gamma_{\phi^1}^x, \]

\[ \gamma_{\phi^2}^x(i) = w_{\phi^2,x_n}^i(i) - \gamma_{\phi^2}^x, \quad \gamma_{\phi^1}^x(i) = w_{\phi^1,x_n}^i(i) - \gamma_{\phi^1}^x. \]
Now arguing as in Proposition 3.1 in [6], it can be shown that there exists functions $y^{\phi_1}$ and $y^{\phi_2}$ and constants $\mu^{\phi_1}$ and $\mu^{\phi_2}$ such that along a subsequence $y^{\phi_m}(i) = \lim_{n \to \infty} y^{\phi_m}(i)$ and $\mu^{\phi_m} = \lim_{n \to \infty} \mu^{\phi_m}$, for $m = 1, 2$. From Step 1, using boundedness of $D_m, m = 1, 2$, it is easy to see that $\mu^{\phi_m}$ is bounded and hence existence of a convergent subsequence is trivial. It follows from the definition that $y^{\phi_m}(i) \leq 0$ for each $i$. The finiteness of the state space and irreducibility ensures the lower bound.

**Step 3:** First we rewrite equations (38) and (39) in terms of the quantities defined in Step 2. Then taking limit $n \to \infty$ and using Step 2, we obtain equations (36) and (37) respectively.

**Step 4:** The interpretations of $\mu^{\phi_1}$ and $\mu^{\phi_2}$ follows by similar arguments as in Theorem 4.1 of [6].

In order to establish the existence of a Nash equilibrium we need the following additional assumption.

**Assumption 3.** Fix a state $i^* \in S$. Define $\tau^* = \inf\{n \geq 1 : X_n = i^*\}$. We assume that there exist constants $R$ and $M$ such that

$$\sup_{\phi^1 \in \Phi^1} \sup_{\phi^2 \in \Phi^2} \sup_{i \in S} \mathbb{E}^{\phi^1, \phi^2}[R^{\tau^*}] \leq M.$$ 

For this $R$, we further assume that $\theta$ is such that

$$e^{2BM_\theta} \leq R,$$

where $M_\rho = \max\{M_{\rho^1}, M_{\rho^2}\}$ where $M_{\rho^i}$ as in (6) with $\rho$ replaced by $\rho^i$.

By Proposition 3 in [5], it follows that aperiodicity of the embedded discrete time Markov chain $\{X_n\}$ under each pair of stationary policies is a sufficient condition for the first part of Assumption 3. The remaining assumptions in the cited proposition follows by finiteness of state space, continuity of $p_{ij}(\cdot)$ and Assumption 2. Next we obtain the following theorem as a consequence of the previous theorem.

**Theorem 5.** Assume that Assumptions 1, 2 and 3 hold. Fix $(\phi^1, \phi^2) \in \Phi^1 \times \Phi^2$. Then there exist constants $g^{\phi_1}, g^{\phi_2}$, real valued functions $h^{\phi_1}, h^{\phi_2}$ on $S$ with $h^{\phi_1}(i^*) = h^{\phi_2}(i^*) = 0$, such that the following are true.

(i)

$$e^{0h^{\phi_2}(i)} = \inf_{\psi \in \mathcal{P}(A(i))} \left\{ \int_A \int_{B(i)} \int_0^B e^\theta \left[ \int_0^t \int_{A(i)} p_{i,a,b}(s) dt \right] dF_{i,a,b}(s) \right\} \times \sum_{j \in S} e^{0h^{\phi_2}(i)} p_{i,j}(a,b) \phi^2(db) |i\psi(da), \forall i \in S. \tag{40}$$

(ii)

$$e^{0h^{\phi_1}(i)} = \inf_{\phi \in \mathcal{P}(B(i))} \left\{ \int_A \int_{B(i)} \int_0^B e^\theta \left[ \int_0^t \int_{B(i)} p_{i,a,b}(s) dt \right] dF_{i,a,b}(s) \right\} \times \sum_{j \in S} e^{0h^{\phi_1}(i)} p_{i,j}(a,b) \phi^1(da) |i\phi(db), \forall i \in S. \tag{41}$$

(iii)

$$g^{\phi_2} = \inf_{\pi^1 \in \Pi_1} I^1_\theta(i^*, \pi^1, \phi^2) \text{ for all } i \text{ and } g^{\phi_1} = \inf_{\pi^2 \in \Pi_2} I^2_\theta(i^*, \phi^1, \pi^2) \text{ for all } i.$$


(iv) For \((i, a, b) \in \mathbb{K}\), let \(D^{x_2}_1(i, a, b) = \frac{1}{\theta} \ln(\int_0^B e^{\theta \int_0^t r^i_{u^1(a, b)}(t) dt - \theta x^2_2} dF_{i, a, b}(s))\) and \(D^{x_1}_2(i, a, b) = \frac{1}{\theta} \ln(\int_0^B e^{\theta \int_0^t r^i_{u^2(a, b)}(t) dt - \theta x^1_2} dF_{i, a, b}(s))\). Then \(h^{x_1}\) and \(h^{x_2}\) have the following representations.

\[
\begin{align*}
  h^{x_1}(i) &= \inf_{\phi^1 \in \Phi^1} \frac{1}{\theta} \ln \mathbb{E}_i \left[ e^{\theta \sum_{k=0}^{i-1} D^{x_1}_1 (X_k, A_k, B_k)} \right], \quad \forall i \in S \setminus \{i^*\}. \\
  h^{x_2}(i) &= \inf_{\phi^2 \in \Phi^2} \frac{1}{\theta} \ln \mathbb{E}_i \left[ e^{\theta \sum_{k=0}^{i-1} D^{x_2}_2 (X_k, A_k, B_k)} \right], \quad \forall i \in S \setminus \{i^*\}.
\end{align*}
\]

**Proof.** The proof of (i) and (ii) follows from Theorem 4 by a similar trick as in Theorem 3 of the zero-sum game section. Proof of (iii) follows by arguments similar to Theorem 1. Finally, the proof of (iv) follows by arguments similar to Lemma 8.1 in [6]. Note that in the proofs of (i), (ii) and (iii) we do not require Assumption 3. It is in the proof of (iv) where we need Assumption 3. \(\square\)

Now, fix any \((\phi^1, \phi^2) \in \Phi^1 \times \Phi^2\). Define

\[
\begin{align*}
\Delta(\phi^2) &= \{ \phi^{x_2} \in \Phi^2 : \text{for each } i \in S, \quad \int_{A(i)} \int_{B(i)} \int_0^B e^{\theta \int_0^t r^i_{u^2(a, b)}(t) dt - \theta x^2_2} dF_{i, a, b}(s) \\
&\quad \times \sum_{j \in S} e^{\theta \phi^{x_2}(j) p_{u^2}(a, b) \phi^{x_2}(db)(i) \phi^{x_2}(da)(i)} = \inf_{\psi \in \mathbb{P}(A(i))} \left\{ \int_{A(i)} \int_{B(i)} \int_0^B e^{\theta \int_0^t r^i_{u^2(a, b)}(t) dt - \theta x^2_2} dF_{i, a, b}(s) \\
&\quad \times \sum_{j \in S} e^{\theta \phi^{x_2}(j) p_{u^2}(a, b) \phi^{x_2}(db)(i) \psi(da)(i)} \right\} \} \}
\end{align*}
\]

and

\[
\begin{align*}
\Delta(\phi^1) &= \left\{ \phi^{x_2} \in \Phi^2 : \text{for each } i \in S, \quad \int_{A(i)} \int_{B(i)} \int_0^B e^{\theta \int_0^t r^i_{u^1(a, b)}(t) dt - \theta x^1_2} dF_{i, a, b}(s) \\
&\quad \times \sum_{j \in S} e^{\theta \phi^{x_1}(j) p_{u^1}(a, b) \phi^{x_2}(db)(i) \phi^{x_1}(da)(i)} = \inf_{\psi \in \mathbb{P}(B(i))} \left\{ \int_{A(i)} \int_{B(i)} \int_0^B e^{\theta \int_0^t r^i_{u^1(a, b)}(t) dt - \theta x^1_2} dF_{i, a, b}(s) \\
&\quad \times \sum_{j \in S} e^{\theta \phi^{x_1}(j) p_{u^1}(a, b) \phi^{x_2}(db)(i) \psi(da)(i)} \right\} \} \}.
\end{align*}
\]

It follows from our assumptions that the sets \(\Delta(\phi^2)\) and \(\Delta(\phi^1)\) are non-empty.

**Lemma 3.** Suppose that Assumptions 1 and 2 is true. For each \((\phi^1, \phi^2) \in \Phi^1 \times \Phi^2\), \(\Delta(\phi^2) \times \Delta(\phi^1)\) is convex and compact with respect to the weak topology.

**Proof.** We first show that \(\Delta(\phi^2)\) is convex. For that let \(\tilde{\phi}^1, \tilde{\psi}^1 \in \Delta(\phi^2)\) and \(\lambda \in [0, 1]\), define: \(\phi^1_\lambda(.|i) = \lambda \tilde{\phi}^1_\lambda(.|i) + (1 - \lambda) \tilde{\psi}^1_\lambda(.|i)\) for all \(i \in S\). By writing down the expression of \(\phi^1_\lambda\) one easily gets that \(\phi^1_\lambda \in \Delta(\phi^2)\). Thus \(\Delta(\phi^2)\) is convex. By analogous argument \(\Delta(\phi^1)\) is also convex, which together implies that \(\Delta(\phi^2) \times \Delta(\phi^1)\) is convex.
By the compactness of $\Phi^1 \times \Phi^2$ and the fact that $\Delta(\phi^2) \times \Delta(\phi^1)$ is a subset of $\Phi^1 \times \Phi^2$, its enough to show that $\Delta(\phi^2) \times \Delta(\phi^1)$ is a closed subset. First we show that $\Delta(\phi^2)$ is a closed subset of the compact space $\Phi^1$. Let $\{\phi_n^*\} \subset \Delta(\phi^2)$ be an arbitrary sequence converging to $\phi^* \in \Phi^1$, and $G(i,a) := \int_{B(i)} \int_0^B e^{\theta \left[ \int_{0}^{\rho_{i,a,b}(t)} dt - g^2(t) \right]} dF_{i,a,b}(s) \times \sum_{j \in S} e^{\theta h^2(j)} p_{i,j}(a,b) \phi^2_j(db|i)$ for $i \in S$ and $a \in A(i)$. By Assumption 1, we have that for each $i \in S$, $G(i,.)$ is a bounded continuous function on $A(i)$. Thus by definition of weak topology we obtain
\[
\int_{A(i)} G(i,a) \phi^* n_a(da|i) \to \int_{A(i)} G(i,a) \phi^* (da|i).
\]
as $n \to \infty$. Since $\{\phi_n^*\} \subset \Delta(\phi^2)$
\[
\int_{A(i)} G(i,a) \phi^* n_a(da|i) = \inf_{\mu \in \mathcal{P}(A(i))} \left\{ \int_{A(i)} G(i,a) \mu(da) \right\}
\]
for all $n = 1, 2, \ldots$. Hence, we have $\phi^* \in \Delta(\phi^2)$. Thus, $\Delta(\phi^2)$ is closed. Similarly, $\Delta(\phi^1)$ is closed. So combining we get $\Delta(\phi^2) \times \Delta(\phi^1)$ is convex and compact.

**Lemma 4.** Suppose that Assumptions 1, 2 and 3 hold. For each $i \in S$, the functions $\phi^1 \to h^{\phi^1}(i)$ and $\phi^2 \to h^{\phi^2}(i)$ are continuous in $\phi^1$ in $\Phi^1$ and $\phi^2$ in $\Phi^2$ respectively. Continuity also holds for the functions $\phi^1 \to g^{\phi^1}$ and $\phi^2 \to g^{\phi^2}$.

**Proof.** By (iii) of Theorem 5, we have $|g^{\phi^1}| \leq M_\rho$ and $|g^{\phi^2}| \leq M_\rho$. We also have $\|D_{\phi^1}^1\| \leq 2BM_\rho$ and $\|D_{\phi^2}^2\| \leq 2BM_\rho$. Thus by Assumption 3, we have $h^{\phi^1}(i) \leq \frac{1}{\rho} \ln M$ for $m = 1, 2$ and for all $i \in S$. Now Assumption 3 also implies that $\sup_{\rho^1 \in \Phi^1} \sup_{\rho^2 \in \Phi^2} \sup_{i \in S} E_i^{\rho^1, \rho^2} \tau^* \leq K$, for some $K$. So by Jensen's inequality we have $h^{\phi^1}(i) \geq -2KBM_\rho$ for $m = 1, 2$ and for all $i \in S$. Now suppose $\phi_n^2 \to \phi^2$. Let us consider subsequences $\{g^{\phi_{n_j}^2}\}$, $\{h^{\phi_{n_j}^1}(i)\}$. We will get a further subsequence such that $g^{\phi_{n_j}^2} \to g^*$ for some constant $g^*$ and $h^{\phi_{n_j}^1}(j) \to u(j)$ for all $j \in S$ for some function $u$ on $S$. We have,
\[
e^{\theta h^{\phi_{n_j}^2}(i)} = \inf_{\psi \in \mathcal{P}(A(i))} \left\{ \int_{A(i)} \int_{B(i)} \int_0^B e^{\theta \left[ \int_{0}^{\rho_{i,a,b}(t)} dt - g^{\phi_{n_j}^2}(t) \right]} dF_{i,a,b}(s) \times \sum_{j \in S} e^{\theta h^{\phi_{n_j}^1}(j)} p_{i,j}(a,b) \phi^2_{n_j}(db|i) \psi(da) \right\} \quad \forall i \in S.
\]
Now by our assumptions, definition of weak convergence and extended Fatou's lemma (Lemma 8.3.7 in [18]), we obtain by taking limit $l \to \infty$ in the above equation,
\[
e^{\theta u(i)} = \inf_{\psi \in \mathcal{P}(A(i))} \left\{ \int_{A(i)} \int_{B(i)} \int_0^B e^{\theta \left[ \int_{0}^{\rho_{i,a,b}(t)} dt - g^*(t) \right]} dF_{i,a,b}(s) \times \sum_{j \in S} e^{\theta u(j)} p_{i,j}(a,b) \phi^2(db|i) \psi(da) \right\} \quad \forall i \in S.
\]
Thus again arguing as in Theorem 5, we will get that $g^{*} = \inf_{i \in \Pi_{\Omega}} J^{1}_{i}(i, \pi^{1}, \phi^{2}) = g^{\phi^{2}}$ and $u(i) = \inf_{\phi^{1} \in \Phi^{1}} \frac{1}{\Omega} \sum_{i=0}^{i^{*}} D^{\phi^{1}, \phi^{2}}(X_{i}, A_{i}, B_{i}) = h^{\phi^{2}}(i)$ over $i \in S \setminus \{i^{*}\}$. Since every subsequence has a further subsequence which converges to the same limit, we are done.

Now we state the main theorem of this section.

**Theorem 6.** Suppose that Assumptions 1, 2 and 3 hold. There exists constants $g^{*1}, g^{*2}, \text{real valued functions } y^{*1}, y^{*2} \text{ on } S \text{ and a pair of stationary policies } (\phi^{*1}, \phi^{*2}) \in \Phi^{1} \times \Phi^{2} \text{ such that}

$$e^{\theta y^{*1}(i)} = \inf_{\phi \in \mathcal{P}(A_{i})(i)} \left\{ \int_{A_{i}} \int_{B(i)} \int_{0}^{B} e^{\theta \int_{0}^{t} \rho^{i}_{A_{i}, B(i)}(t) \, dt - g^{*1}} \, dF_{i,a,b}(s) \times \sum_{j \in S} e^{\theta y^{*1}(j)} p_{i,j}(a, b) \phi^{i}(db|i) \psi(da) \right\}$$

$$= \int_{A_{i}} \int_{B(i)} \frac{1}{\Omega} \sum_{i=0}^{i^{*}} D^{\phi^{1}, \phi^{2}}(X_{i}, A_{i}, B_{i}) = h^{\phi^{2}}(i) \quad \forall i \in S \setminus \{i^{*}\}.$$  

(44)

and

$$e^{\theta y^{*2}(i)} = \inf_{\phi \in \mathcal{P}(B_{i})(i)} \left\{ \int_{A_{i}} \int_{B(i)} \int_{0}^{B} e^{\theta \int_{0}^{t} \rho^{i}_{A_{i}, B(i)}(t) \, dt - g^{*2}} \, dF_{i,a,b}(s) \times \sum_{j \in S} e^{\theta y^{*2}(j)} p_{i,j}(a, b) \varphi(db|i) \phi^{i}(da|i),$$

(45)

for all $i \in S$. Moreover, the pair of policies $(\phi^{*1}, \phi^{*2}) \in \Phi^{1} \times \Phi^{2}$ is a Nash-equilibrium and we have $J_{m}(i, \phi^{*1}, \phi^{*2}) = g^{*m}$ for all $i \in S$ and $m = 1, 2$.

**Proof.** Let $2^{\Omega_{1} \times \Omega_{2}}$ be the power set of $\Phi^{1} \times \Phi^{2}$ and define the multi function $\Psi : \Phi^{1} \times \Phi^{2} \rightarrow 2^{\Omega_{1} \times \Omega_{2}}$ by $\Psi((\phi^{1}, \phi^{2})) = \Delta(\phi^{2}) \times \Delta(\phi^{1})$. Next we show that $\Psi$ has a closed graph. Let $\{(\phi^{1}_{n}, \phi^{2}_{n})\} \subset \Phi^{1} \times \Phi^{2}$ and $\{(\bar{\phi}^{1}_{n}, \bar{\phi}^{2}_{n})\} \subset \Phi^{1} \times \Phi^{2}$ be arbitrary sequences with $\{(\phi^{1}_{n}, \phi^{2}_{n})\} \in \Psi((\phi^{1}_{n}, \phi^{2}_{n}))$ and $\{(\bar{\phi}^{1}_{n}, \bar{\phi}^{2}_{n})\}$ converges to $(\bar{\phi}^{1}, \bar{\phi}^{2})$ and $(\bar{\phi}^{1}, \bar{\phi}^{2})$, respectively. Then by the definition of $\Delta(\phi^{2}_{n})$, we have

$$\inf_{\phi \in \mathcal{P}(A_{i})(i)} \left\{ \int_{A_{i}} \int_{B(i)} \int_{0}^{B} e^{\theta \int_{0}^{t} \rho^{i}_{A_{i}, B(i)}(t) \, dt - g^{*1}} \, dF_{i,a,b}(s) \times \sum_{j \in S} e^{\theta y^{*1}(j)} p_{i,j}(a, b) \phi^{i}(db|i) \psi(da) \right\}$$

$$= \int_{A_{i}} \int_{B(i)} \frac{1}{\Omega} \sum_{i=0}^{i^{*}} D^{\phi^{1}, \phi^{2}}(X_{i}, A_{i}, B_{i}) = h^{\phi^{2}}(i) \quad \forall i \in S \setminus \{i^{*}\}.$$  

(46)

Now using our assumptions, Lemma 4 and extended Fatou’s lemma (Lemma 8.3.7 in [18]) we obtain by taking limit $n \rightarrow \infty$ in (46),
for all \(i \in S\), which implies \(\tilde{\phi}^{s^{1}} \in \Delta(\hat{\phi}^{2})\). Using similar arguments as above, we can also show that \(\tilde{\phi}^{s^{2}} \in \Delta(\hat{\phi}^{1})\). Hence, the multi function \(\Psi\) has a closed graph. Therefore by Fan’s fixed point theorem \([17]\) we have the existence of \(( \phi^{s^{1}}, \phi^{s^{2}} ) \in \Phi^{1} \times \Phi^{2}\) such that \(((\phi^{s^{1}}, \phi^{s^{2}})) \in \Delta(\phi^{s^{1}}) \times \Delta(\phi^{s^{2}})\). Now using Theorem 5 we obtain solution to the coupled system of equations (44) and (45).

Now for the Nash equilibrium part, it follows from (44) and arguments similar to Theorem 1, that

\[
g^{s^{1}} = J_{1}(i, \phi^{s^{1}}, \phi^{s^{2}}) = V^{1}_{0}(\phi^{s^{2}})
\]

Analogously, starting from (45) it can be shown that

\[
g^{s^{2}} = J_{2}(i, \phi^{s^{1}}, \phi^{s^{2}}) = V^{2}_{0}(\phi^{s^{1}})
\]

Hence we are done.

\[\square\]

6. Conclusion

In this paper we have studied both zero-sum and non-zero sum risk-sensitive average criterion games for semi-Markov process. Here we assume that the state space is finite and the sojourn time distributions are supported within a fixed compact interval. So it remains an open problem to extend the setting to more general state space and sojourn time distributions. Note that such a problem is also open for the control case as well, because in \([15]\) where the control problem is studied similar assumptions are made and crucially used in the analysis. Also, in the non-zero sum case we assume that the immediate cost is zero. The reason behind that is, in presence of immediate cost, in the expression for cost accumulated upto time \(t\), the component corresponding to immediate cost involves a summation whose upper limit is a random variable equal to the number of transitions upto time \(t\). So in presence of that we are unable to establish uniform(over stationary policies) bounds as in the first two lines of the proof of Lemma 4. This lemma is crucially used in proving the existence of Nash equilibrium.

Funding

The research of the second named author is supported by the Mathematical Research Impact Centric Support (MATRICS) grant, File No: MTR/2020/000350, by the Science and Engineering Research Board (SERB), Department of Science and Technology (DST), Government of India.
References

[1] Howard, R. A., Matheson, J. E. (1972). Risk-sensitive Markov decision processes. *Manag. Sci.* 18(7):356–369. DOI: 10.1287/mnsc.18.7.356.

[2] Basu, A., Ghosh, M. K. (2014). Zero-sum risk-sensitive stochastic games on a countable state space. *Stochastic Process. Appl.* 124(1):961–983. DOI: 10.1016/j.spa.2013.09.009.

[3] Bäuerle, N., Rieder, U. (2017). Zero-sum risk-sensitive stochastic games. *Stochastic Process. Appl.* 127(2):622–642. DOI: 10.1016/j.spa.2016.06.020.

[4] Cavazos-Cadena, R., Hernández-Hernández, D. (2019). The vanishing discount approach in a class of zero-sum finite games with risk-sensitive average criterion. *SIAM J. Control Optim.* 57(1):219–240. DOI: 10.1137/18M1165104.

[5] Basu, A., Ghosh, M. K. (2018). Nonzero-sum risk-sensitive stochastic games on a countable state space. *Math. OR*. 43(2):516–532. DOI: 10.1287/moor.2017.0870.

[6] Wei, Q., Chen, X. (2019). Risk-sensitive average equilibria for discrete-time stochastic games. *Dyn. Games Appl.* 9(2):521–549. DOI: 10.1007/s13235-018-0267-5.

[7] Biswas, A., Saha, S. (2020). Zero-sum stochastic differential games with risk-sensitive cost. *Appl. Math. Optim.* 81(1):113–140. DOI: 10.1007/s00245-018-9479-8.

[8] Ghosh, M. K., Kumar, K. S., Pal, C., Pradhan, S. (2021). Nonzero-sum risk-sensitive stochastic differential games with discounted costs. *Stoch. Anal. Appl.* 39(2):306–326. DOI: 10.1080/07362994.2020.1796707.

[9] Ghosh, M. K., Pradhan, S. (2020). Zero-sum risk-sensitive stochastic differential games with reflecting diffusions in the orthant. *ESAIM COCV*. 26:114. DOI: 10.1051/cocv/2020029.

[10] Ghosh, M. K., Kumar, K. S., Pal, C. (2016). Zero-sum risk-sensitive stochastic games for continuous time Markov chains. *Stoch. Anal. Appl.* 34(5):835–851. DOI: 10.1080/07362994.2016.1180995.

[11] Wei, Q. (2018). Zero-sum games for continuous-time Markov jump processes with risk-sensitive finite-horizon cost criterion. *Oper. Res. Lett.* 46(1):69–75. DOI: 10.1016/j.orl.2017.11.008.

[12] Wei, Q. (2019). Nonzero-sum risk-sensitive finite-horizon continuous-time stochastic games. *Statist. Probab. Lett.* 147:96–104. DOI: 10.1016/j.spl.2018.12.004.

[13] Huang, Y., Lian, Z., Guo, X. (2018). Risk-sensitive semi-Markov decision processes with general utilities and multiple criteria. *Adv. Appl. Probab.* 50(3):783–804. DOI: 10.1017/apr.2018.36.

[14] Bhabak, A., Saha, S. (2021). Risk-sensitive semi-Markov decision problems with discounted cost and general utilities. arXiv:2101.04510.

[15] Chávez-Rodríguez, S., Cavazos-Cadena, R., Cruz-Suárez, H. (2016). Controlled semi-Markov chains with risk-sensitive average cost criterion. *J. Optim. Theory Appl.* 170(2):670–686. DOI: 10.1007/s10957-016-0916-z.

[16] Arapostathis, A., Borkar, V. S., Fernández-Gaucherand, E., Ghosh, M. K., Marcus, S. I. (1993). Discrete-time controlled Markov processes with average cost criterion: a survey. *SIAM J. Control Optim.* 31(2):282–344. DOI: 10.1137/0331018.

[17] Fan, K. (1952). Fixed-point and minimax theorems in locally convex topological linear spaces. *Proc. Natl. Acad. Sci. USA.* 38(2):121–126. DOI: 10.1073/pnas.38.2.121.