

Article

Cooperative Stochastic Games with Mean-Variance Preferences

Elena Parilina ^{1,2,3,*}  and Stepan Akimochkin ^{4,†}¹ Department of Mathematical Game Theory and Statistical Decisions, Saint Petersburg State University, 7/9 Universitetskaya nab., 199034 Saint Petersburg, Russia² School of Mathematics and Statistics, Qingdao University, Qingdao 266071, China³ Institute of Applied Mathematics of Shandong, Qingdao 266071, China⁴ Australia and New Zealand Banking Group Limited, 242 Pitt Street, Sydney, NSW 2000, Australia; Stefan.akimochkin@gmail.com

* Correspondence: e.parilina@spbu.ru

† These authors contributed equally to this work.

Abstract: In stochastic games, the player's payoff is a stochastic variable. In most papers, expected payoff is considered as a payoff, which means the risk neutrality of the players. However, there may exist risk-sensitive players who would take into account "risk" measuring their stochastic payoffs. In the paper, we propose a model of stochastic games with mean-variance payoff functions, which is the sum of expectation and standard deviation multiplied by a coefficient characterizing a player's attention to risk. We construct a cooperative version of a stochastic game with mean-variance preferences by defining characteristic function using a maxmin approach. The imputation in a cooperative stochastic game with mean-variance preferences is supposed to be a random vector. We construct the core of a cooperative stochastic game with mean-variance preferences. The paper extends existing models of discrete-time stochastic games and approaches to find cooperative solutions in these games.

Keywords: cooperative stochastic games; mean-variance preferences; stochastic payoff; risk-sensitive payoff; core



Citation: Parilina, E.; Akimochkin, S. Cooperative Stochastic Games with Mean-Variance Preferences. *Mathematics* **2021**, *9*, 230. <https://doi.org/10.3390/math9030230>

Academic Editor: Krzysztof J. Szajowski
Received: 30 November 2020
Accepted: 20 January 2021
Published: 25 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The class of stochastic games was initially introduced by L. Shapley [1]. He considered two-player zero-sum stochastic games with finite state space and finite strategy spaces and proved the existence of optimal stationary strategies when players' payoff function is discounted mathematical expected payoff. This result has been extended to the case of n -player games by Fink [2], Takahashi [3], Sobel [4]. The results in stochastic games of the recent 60 years, mostly for non-cooperative case, are discussed by Neyman and Sorin [5], Solan [6], and Solan and Vieille [7].

Players' payoffs in stochastic games are stochastic variables. In most papers, the expected discounted payoff is considered as a measure of a player's payoff in a stochastic game. In this case, it is supposed that the players are risk neutral. Hence, the risk averse and risk loving attitudes cannot be modeled considering only expectation as a criterion. We propose a model of a stochastic game in which the players' attentions to risk are taken into account. We use a mean-variance preference popular in decision making theory. In particular, this type of preference for random payoffs is used in financial risk management, in a portfolio decision theory proposed by Markowitz [8]. The risk allocation problem in portfolio management based on cooperative game theory approach is also examined by Csóka et al. [9].

We consider n -person stochastic games with finite sets of actions in any state of the game. A mean-variance stochastic game in discrete time with the payoff function, which is a linear combination of expectation and standard deviation of the payoff, is proposed. The goal of the paper is to construct the cooperative version of a mean-variance stochastic game and define the core as a cooperative solution of such a cooperative game. In the

paper, we first define the players' payoffs according to mean-variance preferences in a non-cooperative setting, and second, we define a cooperative game in the form of characteristic function. Then, we determine a solution of a cooperative stochastic game using the method proposed by Suijs and Borm [10], and Suijs et al. in [11]. The contribution of the paper is in proposing the theory of cooperative stochastic games with mean-variances preferences. The research question about the stability of the cooperative solution over time is open and briefly discussed in Section 5.

First, we consider a non-cooperative stochastic game with mean-variance utility function. Second, we define the cooperative version of the game constructing characteristic function using maxmin approach (see Neumann and Morgenstern [12]). This approach supposes that the value of characteristic function for any coalition is the maxmin payoff in zero-sum stochastic game between the coalition and anti-coalition. As the players may have different coefficients before standard deviations in the utility functions, the summation of the utilities does not make sense. When a coalition is formed, the maximal utility of the coalition payoff over the coalition members is used to measure the utility of the coalition payoff. There are other approaches of defining characteristic function in dynamic games proposed by Gromova and Petrosyan [13], Petrosjan and Zaccour [14] and examined by Reddy and Zaccour [15]. However, applications of these approaches need corrections, considering the stochastic nature of payoffs and mean-variance utility functions. Cooperative stochastic games with expected payoff as a utility function are introduced by Petrosjan [16] and Petrosjan and Baranova [17], and then extended by Parilina [18] and Parilina and Tampieri [19] by examining the stability of cooperative solutions in stochastic dynamic settings. As proposed by Suijs et al. [11], we consider the core as a solution of a cooperative game. The conditions of node-consistency of the core in dynamic games played over event trees with stochastic nature are obtained by Parilina and Zaccour [20]. Recently, the sufficient conditions of strong subgame consistency of the core are obtained by Parilina and Petrosyan [21].

Then, we define a solution of a cooperative stochastic game. The idea is briefly described below. Suijs et al. introduce a new class of cooperative games arising from cooperative decision making problems in the stochastic environment [10,11]. They introduce some types of preferences to order stochastic payoffs and model the games with risk neutral, as well as risk averse and risk loving players. Suijs et al. in [11] define the core of the cooperative game with stochastic payoffs and obtain the necessary and sufficient conditions for the non-emptiness of the core. The deterministic equivalent of a stochastic payoff with some properties is introduced by Suijs and Borm [10]. The nucleolus for this type of games is introduced by Suijs et al. [11]. The Shapley-like solutions for the cooperative game with stochastic payoffs are proposed by Timmer et al. [22]. The model of transferable utility game with uncertainty is introduced by Habis and Herings [23]. They propose a solution concept of a weak sequential core for this class of games and show that it is non-empty if all ex post TU-games are convex.

The goal of the paper is to propose a scheme of construction of a cooperative version of a stochastic game, in which not only expectation but also standard deviation of players' payoffs is taken into account. The idea of considering risk-sensitive players recently appears in different variations in some papers on dynamic and differential games. The risk sensitive nonzero-sum continuous time stochastic games are considered by Wei [24], where the existence of Nash equilibrium is proved. The risk-sensitive average payoff criterion for zero-sum (see Bäuerle and Rieder [25]) and for the nonzero-sum discrete-time stochastic games with a denumerable state space (see Wei and Chen [26]), in which an existence of the Nash equilibrium is examined. The risk sensitive Nash equilibria in stochastic games are discussed by Nowak [27]. The class of risk sensitive stochastic games with overlapping generations is considered by Jaśkiewicz and Nowak [28]. Recently, the Nash equilibria for stochastic games of resource extraction in which risk coefficient is included into utility function are found by Asienkiewicz and Balbus [29]. In mean-field games, the model with both expectation and variance in payoff criteria is proposed by Saldi et al. [30].

The paper is organized as follows. The class of mean-variance stochastic games in discrete time is introduced in Section 2. The propositions providing formulae to calculate a variance and covariance of players' payoffs in stochastic games are given in Section 3. In Section 4, we define a cooperative mean-variance stochastic game and an imputation as a cooperative solution of the game. We briefly conclude in Section 5. The proofs of all propositions are given in the appendix.

2. Stochastic Games with Mean-Variance Preferences

2.1. Stochastic Game

Define a finite stochastic game G using the following notations. The finite set of players is $N = \{1, \dots, n\}$, and the finite set of states is $S = \{1, 2, \dots, t\}$. In any state $s \in S$, an n -person normal-form game $\langle N, \{A_i^s\}_{i \in N}, \{K_i^s\}_{i \in N} \rangle$ is defined, where A_i^s is the finite set of Player i 's actions in state s , $K_i^s(a_1^s, \dots, a_n^s) = K_i^s(a^s) \in \mathbb{R}$ is a payoff function of Player i in state s , $i \in N$, $s \in S$. The action profile $a^s = (a_1^s, \dots, a_n^s)$ belongs to the set $A^s = \prod_{i \in N} A_i^s$.

The transition map π is $T \rightarrow \Delta(S)$ function, where $T = \{(s, a^s) \mid s \in S; a^s \in A^s\}$ and $\Delta(S)$ is a probability distribution on S . For each pair (s, a^s) , mapping π identifies the vector of probabilities $(\pi(1 \mid s; a^s), \dots, \pi(s' \mid s; a^s), \dots, \pi(t \mid s; a^s))$, where $\pi(s' \mid s; a^s)$ is the probability that state s' will be realized at the next stage if at the present stage (state s) action-profile a^s has been realized, $\pi(s' \mid s; a^s) \geq 0$ and $\sum_{s' \in S} \pi(s' \mid s; a^s) = 1$ for any $a^s \in A^s$ and $s \in S$.

Let vector $p^0 = (p_1^0, \dots, p_t^0)$ be a vector of initial distribution on S , and p_s^0 is the probability that state s will be realized at the first stage of the game, $p_s^0 \geq 0$ for any $s \in S$, and $\sum_{s \in S} p_s^0 = 1$.

Let all players have the same discount factor $\delta \in (0, 1)$ for their payoffs.

Definition 1. The stochastic game G is defined by

$$\langle N, S, \pi, p^0, \delta \rangle. \quad (1)$$

Definition 2. Stochastic subgame G^s is a finite stochastic game (1) with initial distribution vector p^0 , such that $p_s^0 = 1$. Subgame G^s is a stochastic game starting from state $s \in S$.

The implementation of a stochastic game can be briefly described in the following way. First, the chance moves and chooses the initial state of the game according to vector p^0 . Let the initial state be $s \in S$. Second, the first stage of game G is realised, i.e., the normal-form game $\langle N, \{A_i^s\}_{i \in N}, \{K_i^s\}_{i \in N} \rangle$ is played. Players simultaneously choose their actions from the action sets $\{A_i^s\}_{i \in N}$. Let Player i 's action in state s be $a_i^s \in A_i^s$, $s \in S$. Therefore, an action profile $a^s = (a_1^s, \dots, a_n^s) \in \prod_{i \in N} A_i^s$ is implemented. Player i obtains her payoff $K_i^s(a^s)$, $i \in N$. Then, the transition to the next stage is realized. The probability that state s' will appear at the next stage is equal to $\pi(s' \mid s; a^s)$, $s' \in S$. After that, the game proceeds in the described way.

Assume that players know all the parameters of the game. We consider a class of stationary strategies. Let Player i 's strategy in game G be a mapping $\eta_i(\cdot)$, such that $\eta_i(s) = a_i^s \in A_i^s$, $s \in S$. Denote the set of stationary strategies of Player i as H_i . The stationary strategy defines an action at any stage depending only on which state is realized at the stage, i.e., $\eta_i : s \mapsto a_i^s \in A_i^s$, $s \in S$, and not depending on the stage number and the history of the stage. A vector of stationary strategies is a strategy profile $\eta = (\eta_1, \dots, \eta_n) \in H_1 \times \dots \times H_n$ in stochastic game G . Obviously, η is a strategy profile in any subgame G^s .

2.2. Mean-Variance Payoff Functions

Players' payoffs in stochastic games are random variables and they are functions of stationary strategy profile η . Denote Player i 's payoff in stochastic game G as $\bar{\zeta}_i = \bar{\zeta}_i(\eta)$ and

Player i 's payoff in stochastic subgame G^s as $\xi_i^s = \bar{\xi}_i^s(\eta)$ for any $s \in S$. Random variables $\bar{\xi}_i$ and $\xi_i = (\xi_i^1, \dots, \xi_i^t)'$ satisfy the following equation:

$$\bar{\xi}_i(\eta) = p^0 \xi_i(\eta).$$

To find a player's "optimal" strategy, it is necessary to determine how to compare random payoffs. The most common way is to measure a random payoff by its mathematical expectation, assuming the risk neutrality of the players. Strictly speaking, ξ_i is preferred to ψ_i or $\xi_i \succ_E \psi_i$, if and only if $E(\xi_i) > E(\psi_i)$, where $E(\xi)$ is the mathematical expectation of random variable ξ . This preference is complete and transitive but it does not allow one to model different attentions to risk.

Example 1. One may compare random payoffs using the principle of stochastic domination. Let $\xi_i, \psi_i \in L^1(\mathbb{R})$ be random payoffs of Player i and $F_{\xi_i}(\cdot)$ and $F_{\psi_i}(\cdot)$ be cumulative distribution functions of ξ_i and ψ_i , respectively. It is said that ξ_i stochastically dominates ψ_i and write $\xi_i \succsim_F \psi_i$, if and only if, for all $x \in \mathbb{R}$ inequality $F_{\xi_i}(x) \leq F_{\psi_i}(x)$ holds. This preference is incomplete, therefore, random payoffs can be incomparable with this preference. Another disadvantage of this preference is its computational complexity for a class of stochastic games. Thus, its application to the theory of stochastic games is questionable.

We use the method of ordering random payoffs proposed in [10,11]. Let every player i have a utility function u_i of a random payoff such that $u_i : L^1(\mathbb{R}) \rightarrow \mathbb{R}$. For random payoffs $\xi, \psi \in L^1(\mathbb{R})$, payoff ξ is preferred or equivalent to payoff ψ by Player i if and only if $u_i(\xi) \geq u_i(\psi)$. In [10] the preferences $(\succsim_i)_{i \in N}$ are considered for the cooperative games with stochastic payoffs and they are such that for any player $i \in N$ there exists a utility function $u_i : L^1(\mathbb{R}) \rightarrow \mathbb{R}$ satisfying properties:

1. For all $\xi \in L^1(\mathbb{R})$: $\xi \sim_i u_i(\xi)$.
2. For all $\xi, \psi \in L^1(\mathbb{R})$: $\xi \succsim_i \psi$ if and only if $u_i(\xi) \geq u_i(\psi)$.
3. For all $d \in \mathbb{R}$: $u_i(d) = d$.
4. For all $\xi \in L^1(\mathbb{R})$: $u_i(\xi - u_i(\xi)) = 0$.
5. For all $\xi \in L^1(\mathbb{R})$ and all $d, d' \in \mathbb{R}$ with $d < d'$: $u_i(d + \xi) < u_i(d' + \xi)$.

Expression $\xi \sim_i u_i(\xi)$ means that the random payoff ξ is equivalent to $u_i(\xi)$, and $\xi \succsim_i \psi$ means that random payoff ξ is preferred or equivalent to random payoff ψ for Player i . In [10], the function $u_i(\cdot)$ is called a deterministic equivalent of the preference. In our work, we call function $u_i(\xi)$ the utility function of a random payoff ξ . Obviously, the utility function $u_i(\xi_i) = E(\xi_i)$ satisfies properties 1–5 given above.

Example 2. One may consider a preference \succsim_α such that for any $\xi, \psi \in L^1(\mathbb{R})$ the relation $\xi \succsim_\alpha \psi$ holds if α -quantile of c.d.f. F_ξ of a random variable ξ is not less than the α -quantile of c.d.f. F_ψ , i.e., $u(\xi) = u_\alpha^\xi = \sup\{t | F_\xi(t) < \alpha\}$. This utility function also satisfies properties 1–5 given above.

Let Player i 's utility function of random payoff ξ_i be

$$u_i(\xi_i) = E(\xi_i) + b_i (Var(\xi_i))^{\frac{1}{2}}, \quad (2)$$

where $Var(\xi_i)$ is the variance of ξ_i , parameter $b_i \in \mathbb{R}$ shows player's attention to risk. Note that if $b_i = 0$, it is understood as Player i is risk neutral. If $b_i > 0$, then Player i is risk loving, and if $b_i < 0$, then Player i is a risk averse person. Denote the preference corresponding to utility function (2) as \succsim_{b_i} . The utility function (2) satisfies properties 1–5.

Preference \succsim_{b_i} is complete but not implied by \succsim_F . The preference \succsim_{b_i} is attractive for applications because of its computation simplicity and possibility of modeling attention to risk. If $b_i = 0$ for any $i \in N$, utility function (2) becomes $u_i(\xi_i) = E(\xi_i)$, which coincides with preference \succsim_E .

2.3. Stochastic Games with Mean-Variance Preferences

We define a stochastic game in which any player $i \in N$ has a utility function $u_i: L^1(\mathbb{R}) \rightarrow \mathbb{R}$ of form (2).

Definition 3. Stochastic game G_b with mean-variance preferences is defined by tuple

$$G_b = \langle N, S, \pi, p^0, \delta, b \rangle, \quad (3)$$

where $b = (b_1, \dots, b_n) \in \mathbb{R}^n$ is a profile of parameters determining players' utility functions (2) for any $i \in N$.

By G_b^s , $s \in S$, we define the subgame of stochastic game (3) with $p^0 = (p_1^0, \dots, p_t^0)$ such that $p_s^0 = 1$.

In the next section, we present the formulae to calculate the utilities of random payoffs in stochastic game (3) and any subgame G_b^s . We also consider the case when players form a coalition and calculate its summarized stochastic payoff.

3. Expectation and Variance of Random Payoffs in Stochastic Games

Consider a strategy profile $\eta = (\eta_1, \dots, \eta_n)$, where η_i is a stationary strategy of Player i , such that $\eta_i(s) = a_i^s \in A_i^s$ for all $s \in S$, $i \in N$. First, find expectation $E(\xi_i^s)$ of Player i 's payoff in subgame G^s :

$$E(\xi_i^s(\eta)) = K_i^s(a^s) + \delta \sum_{s' \in S} \pi(s'|s; a^s) E(\xi_i^{s'}(\eta))$$

or in a vector form:

$$E(\xi_i(\eta)) = K_i(a) + \delta \Pi(\eta) E(\xi_i(\eta)), \quad (4)$$

where $E(\xi_i(\eta)) = (E(\xi_i^1(\eta)), \dots, E(\xi_i^t(\eta)))'$, $K_i(a) = (K_i^1(a^1), \dots, K_i^t(a^t))'$. Equation (4) is equivalent to the following one:

$$E(\xi_i(\eta)) = (\mathbb{I} - \delta \Pi(\eta))^{-1} K_i(a), \quad (5)$$

where \mathbb{I} is $t \times t$ identity matrix and $\Pi(\eta)$ is a matrix of transition probabilities in stochastic game (1), which is a function of strategy profile η taking a form of

$$\Pi(\eta) = \begin{pmatrix} \pi(1|1; a^1) & \dots & \pi(s|1; a^1) & \dots & \pi(t|1; a^1) \\ \vdots & & \ddots & & \vdots \\ \pi(1|t; a^t) & \dots & \pi(s|t; a^t) & \dots & \pi(t|t; a^t) \end{pmatrix}.$$

Matrix $(\mathbb{I} - \delta \Pi(\eta))^{-1}$ is denoted as $\hat{\Pi}(\eta)$ and it always exists for $\delta \in (0, 1)$.

Remark 1. All the eigenvalues of stochastic matrix $\Pi(\eta)$ are from the interval $[-1, 1]$. For the existence of matrix $\hat{\Pi}(\eta)$, it is necessary and sufficient that the determinant of matrix $(\Pi(\eta) - \mathbb{I}/\delta)$ be not equal to zero. Therefore, matrix $(\Pi(\eta) - \mathbb{I}/\delta)$ must not have an eigenvalue equal to $1/\delta$. The last condition is satisfied because $1/\delta > 1$. Therefore, it cannot be an eigenvalue of stochastic matrix $\Pi(\eta)$.

An expectation of Player i 's payoff in stochastic game G equals

$$E(\bar{\xi}_i(\eta)) = p^0 E(\xi_i(\eta)) = p^0 \hat{\Pi}(\eta) K_i(a). \quad (6)$$

The variance $\text{Var}(\bar{\xi}_i(\eta))$ of Player i 's payoff $\bar{\xi}_i(\eta)$ in game G can be found using the following proposition.

Proposition 1. Let in stochastic game G or G_b , Player i 's strategy be $\eta_i \in H_i$. The variance of Player i 's payoff is

$$\text{Var}(\bar{\xi}_i(\eta)) = (p^0 \tilde{\Pi}(\eta) \circ K_i(a))(2\hat{\Pi}(\eta) - \mathbb{I})K_i - (p^0 \hat{\Pi}(\eta)K_i(a))^2, \quad (7)$$

where operation $A \circ B$ is the Hadamard product (Let A and B be $m \times n$ matrices with entries in \mathbb{R} . The Hadamard product of A and B is an $m \times n$ matrix defined by $[A \circ B]_{i,j} = [A]_{i,j}[B]_{i,j}$ for all $1 \leq i \leq m, 1 \leq j \leq n$.) of matrices A and B of the same size, and $\tilde{\Pi}(\eta) = (\mathbb{I} - \delta^2 \Pi(\eta))^{-1}$.

Proof. See Appendix A.1. \square

Corollary 1. The variance $\text{Var}(\xi_i^s(\eta))$ of Player i 's payoff in subgame G^s , $s \in S$, can be found by formula (7) substituting $p^0 = (p_1^0, \dots, p_t^0)$ s. t. $p_s^0 = 1$ and $p_{s'}^0 = 0$ for any $s' \neq s, s' \in S$.

For further calculations, in particular, for the construction of a cooperative form of stochastic game G_b and subgames G_b^s , it is required to determine the random payoff of any coalition $C \subset N$, which is denoted as $\bar{\xi}_C$ and equals

$$\bar{\xi}_C(\eta) = \sum_{i \in C} \bar{\xi}_i(\eta),$$

and for subgame G^s :

$$\xi_C^s(\eta) = \sum_{i \in C} \xi_i^s(\eta).$$

We should notice that players' payoffs ξ_i^s and ξ_j^s are not independent variables for $i \neq j, i, j \in C$. The following proposition gives the formulae to calculate an expectation and variance of random payoff $\bar{\xi}_C(\eta)$ of any coalition C .

Proposition 2. The expectation of payoff $\bar{\xi}_C(\eta)$ of coalition C in stochastic game G or G_b is

$$E(\bar{\xi}_C(\eta)) = p^0 \hat{\Pi}(\eta) \sum_{i \in C} K_i(a). \quad (8)$$

The variance of payoff $\bar{\xi}_C(\eta)$ of coalition C in stochastic game G or G_b is

$$\text{Var}(\bar{\xi}_C(\eta)) = \sum_{i \in C} \text{Var}(\bar{\xi}_i(\eta)) + 2 \sum_{\substack{l < m \\ l, m \in C}} \text{Cov}(\bar{\xi}_l(\eta), \bar{\xi}_m(\eta)), \quad (9)$$

where variance $\text{Var}(\bar{\xi}_i(\eta))$ of Player i 's payoff is defined by formula (7), and $\text{Cov}(\bar{\xi}_l(\eta), \bar{\xi}_m(\eta))$ is a covariance of Players l and m ' payoffs, which is equal to

$$\begin{aligned} \text{Cov}(\bar{\xi}_l(\eta), \bar{\xi}_m(\eta)) &= (p^0 \tilde{\Pi}(\eta) \circ K_l(a))(\hat{\Pi}(\eta) - 0.5\mathbb{I})K_m(a) \\ &\quad + (p^0 \tilde{\Pi}(\eta) \circ K_m(a))(\hat{\Pi}(\eta) - 0.5\mathbb{I})K_l(a) - (p^0 \hat{\Pi}(\eta)K_l(a))(p^0 \hat{\Pi}(\eta)K_m(a)). \end{aligned} \quad (10)$$

Proof. See Appendix A.2. \square

Corollary 2. The variance of payoff $\xi_C^s(\eta)$ of coalition C in subgame G^s or G_b^s , $s \in S$, can be found by formula (9) substituting $p^0 = (p_1^0, \dots, p_t^0)$ s. t. $p_s^0 = 1$ and $p_{s'}^0 = 0$ for any $s' \in S, s' \neq s$.

Equation (6) and Proposition 1 provide formulae to calculate the utilities of players' payoffs in stochastic games with mean-variance preferences for any strategy profile η . Proposition 2 provides formulae to calculate the utilities of coalition payoffs in stochastic games with mean-variance preferences.

4. Cooperative Stochastic Games with Mean-Variance Preferences

In this section, we construct a cooperative version of a stochastic game G_b and subgame G_b^s , starting from state $s \in S$. Therefore, we determine characteristic function $v(\cdot) : 2^N \rightarrow \mathbb{R}$ satisfying two properties:

1. $v(\emptyset) = 0$,
2. (superadditivity) For any disjoint coalitions $C, C' \subset N$, $C \cap C' = \emptyset$, inequality $v(C) + v(C') \leq v(C \cup C')$ holds.

The value $v(C)$ can be interpreted as a “worth” or “power” of coalition C when its members play together as a unit. We use “maxmin” approach to determine superadditive characteristic function [12].

The characteristic function for subgame G_b^s is denoted as $v^s(\cdot)$ and defined as

$$v^s(C) = \max_{\eta_C} \min_{\eta_{N \setminus C}} \max_{i \in C} u_i(\xi_C^s(\eta_C, \eta_{N \setminus C})), \quad (11)$$

for any coalition $C \subset N$. Here, $\eta_C \in \prod_{i \in C} H_i$ is a stationary strategy of coalition C and $\eta_{N \setminus C} \in \prod_{i \in N \setminus C} H_i$ is a stationary strategy of coalition $N \setminus C$. Strategies $(\eta_C, \eta_{N \setminus C})$ form a strategy profile in any subgame G_b^s and game G_b . In Equation (11) and further calculations, the utility function of any player i is of form (2), as required in the definition of stochastic game G_b . An expression in the right-handed side of (11); we find maxmin of the maximal value of utility among players from C . Maxmin value is such that coalition C can guarantee by strategy η_C when coalition $N \setminus C$ plays against coalition C using strategy $\eta_{N \setminus C}$.

Characteristic function for game G_b is denoted by $\bar{v}(\cdot)$ and it can be defined for coalition C using the following expression:

$$\bar{v}(C) = \max_{\eta_C} \min_{\eta_{N \setminus C}} \max_{i \in C} u_i(\bar{\xi}_C(\eta_C, \eta_{N \setminus C})). \quad (12)$$

This function is equal to the maximal value of utility function of random payoff $\bar{\xi}_C(\eta_C, \eta_{N \setminus C})$ (among the players from coalition C), which coalition C can guarantee by itself using stationary strategy $\eta_C \in \prod_{i \in C} H_i$ if coalition $N \setminus C$ plays against C in zero-sum game using strategy $\eta_{N \setminus C} \in \prod_{i \in N \setminus C} H_i$.

Cooperation in stochastic game G_b assumes that players form the grand coalition N and play together to maximize the maximal utility of their total payoff $\bar{\xi}_N$. We call the stationary strategy profile η^* in game G_b a cooperative one if

$$\eta^* = \arg \max_{\eta \in \prod_{i \in N} H_i} \max_{i \in N} u_i(\bar{\xi}_N(\eta)). \quad (13)$$

Therefore, the value of characteristic function $\bar{v}(N)$ can be rewritten as

$$\bar{v}(N) = \max_{i \in N} u_i(\bar{\xi}_N(\eta^*)). \quad (14)$$

The value $v^s(N)$ of the characteristic function $v^s(N)$ in cooperative subgame G_b^s is

$$v^s(N) = \max_{\eta \in \prod_{i \in N} H_i} \max_{i \in N} u_i(\xi_N^s(\eta)). \quad (15)$$

Definition 4. Cooperative stochastic game with mean-variance preferences is a tuple $\langle N, \bar{v} \rangle$ with characteristic function $\bar{v} : 2^N \rightarrow \mathbb{R}$ defined by (12) and (14), and $\bar{v}(\emptyset) = 0$.

Definition 5. Cooperative stochastic subgame with mean-variance preferences is a tuple $\langle N, v^s \rangle$ with characteristic function $v^s : 2^N \rightarrow \mathbb{R}$ defined by (11) and (15), and $v^s(\emptyset) = 0$.

In cooperation, players with mean-variance preferences realize a cooperative stationary strategy profile η^* given by (13).

The main problem of cooperative game theory is to find a “fair” imputation of coalition N 's total payoff. In a classical cooperative game theory, values of characteristic function are non-random variables and there is a variety of solution concepts to determine, in some sense, “fair” imputation.

The concept of a random payoff imputation is proposed in [11], in which it is supposed that the components of the imputation are random variables. We give a definition of an imputation of a random payoff $\bar{\xi}_N(\eta^*)$.

Definition 6 (see [11]). An imputation of random payoff $\bar{\xi}_N(\eta^*)$ is a vector $\varphi = (\varphi_1, \dots, \varphi_n)$, such that $\varphi_i = d_i + r_i(\bar{\xi}_N(\eta^*) - E(\bar{\xi}_N(\eta^*)))$ and $d \in \mathbb{R}^n, r \in \mathbb{R}^n$ satisfy the following conditions:

1. $\sum_{i \in N} d_i = E(\bar{\xi}_N(\eta^*)),$
2. $\sum_{i \in N} r_i = 1$ and $r_i \geq 0$ for any $i \in N$.

According to Definition 6, the payoff $\bar{\xi}_N(\eta^*)$ of coalition N in game G_b is a random vector characterized by an allocation of expected payoff $E(\bar{\xi}_N(\eta^*))$ and allocation of random variable $(\bar{\xi}_N(\eta^*) - E(\bar{\xi}_N(\eta^*)))$, which can be interpreted as a risk of random payoff $\bar{\xi}_N(\eta^*)$. Denote the set of imputations in cooperative game G_b as $\Phi(N)$.

We use the statement from [11] about necessary and sufficient conditions that an imputation belongs to the core of cooperative stochastic game G_b with mean-variance preferences as a definition of the core.

Definition 7. An imputation φ belongs to the core of the cooperative game G_b , if and only if the inequality

$$\sum_{i \in C} \{d_i + r_i(u_i(\bar{\xi}_N(\eta^*)) - E(\bar{\xi}_N(\eta^*)))\} \geq \bar{v}(C)$$

holds for any coalition $C \subset N$.

To illustrate the idea of defining the core in cooperative stochastic game with mean-variance preferences, we provide a numerical example demonstrating calculations step by step.

Example 3. Consider two-person stochastic game with two states $\langle N, A_1^s, A_2^s, K_1^s, K_2^s \rangle, s = 1, 2$, with players' payoffs:

$$s = 1 : \begin{pmatrix} (3, 5) & (3, 10) \\ (10, 4) & (8, 6) \end{pmatrix}, \quad s = 2 : \begin{pmatrix} (2, 4) & (2, 10) \\ (7, 3) & (5, 4) \end{pmatrix}$$

Transition probabilities are defined in matrices:

$$s = 1 : \begin{pmatrix} (0.7, 0.3) & (0.4, 0.6) \\ (0.4, 0.6) & (0.3, 0.7) \end{pmatrix}, \quad s = 2 : \begin{pmatrix} (0.9, 0.1) & (0.4, 0.6) \\ (0.2, 0.8) & (0.3, 0.7) \end{pmatrix},$$

where element (i, j) of the first (second) matrix consists of transition probabilities from states 1 (2) to the states 1, 2 respectively under condition that Player 1 chooses strategy i and Player 2 chooses strategy j in state 1 (2).

Let discount factor δ be 0.9 and vector of initial distribution p^0 on state set $\{1, 2\}$ be $(0.5, 0.5)$. Suppose that both players are risk averse and $b_1 = -2, b_2 = -1$.

Using condition (13), we obtain $\bar{\eta} = (\bar{\eta}_1, \bar{\eta}_2)$, where $\bar{\eta}_1 = (0, 0), \bar{\eta}_2 = (1, 1)$, i.e., player 1 chooses action 2 in both states, and player 2 chooses action 1 in both states. The maximal value of the maximal utility function is equal to the following one:

$$\bar{v}(N) = \max_{i \in N} u_i(\bar{\xi}_N(\bar{\eta})) = 94.1709.$$

Therefore, imputation defined by (d, r) is in the core of game G_b according to Definition 7 if and only if the following inequality takes place for any $C \subset N$:

$$\sum_{i \in C} (d_i + r_i(u_i(\bar{\xi}_N(\bar{\eta})) - E(\bar{\xi}_N(\bar{\eta}))) \geq \bar{v}(C).$$

If $C = \{1\}$, the inequality is

$$d_1 + r_1(u_1(\bar{\xi}_N(\bar{\eta})) - E(\bar{\xi}_N(\bar{\eta}))) \geq \bar{v}(\{1\}),$$

and for $C = \{2\}$:

$$d_2 + r_2(u_2(\bar{\xi}_N(\bar{\eta})) - E(\bar{\xi}_N(\bar{\eta}))) \geq \bar{v}(\{2\}).$$

Taking into account Definition 6 of the imputation, the following conditions hold:

$$\begin{aligned} d_1 + d_2 &= E(\bar{\xi}_N(\bar{\eta})) = 128.2, \\ r_1 + r_2 &= 1. \end{aligned}$$

The following system defines the core of the cooperative stochastic game with mean-variance preferences:

$$\begin{cases} d_1 + r_1(60.1419 - 128.2000) \geq 10.9955, \\ d_2 + r_2(94.1709 - 128.2000) \geq 25.4971, \\ d_1 + d_2 = 128.2000, \\ r_1 + r_2 = 1, \end{cases}$$

or in equivalent form:

$$\begin{cases} d_1 \in [10.9955 + 68.0581r_1, 68.6739 + 34.0291r_1], \\ d_2 = 128.2000 - d_1, \\ r_1 \in [0, 1], \quad r_2 = 1 - r_1. \end{cases} \quad (16)$$

The region of (r_1, d_1) , $r_1 \in [0, 1]$, satisfying the first inequality of the latter system, is depicted in Figure 1.

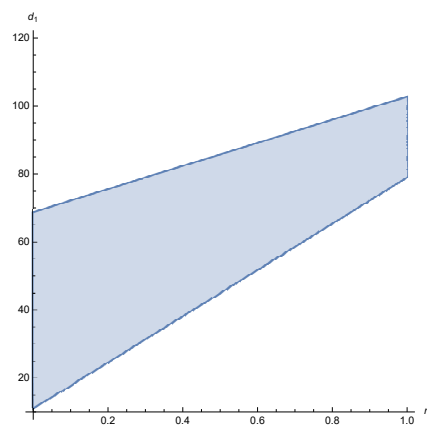


Figure 1. The region of (r_1, d_1) defining the core in Example 3.

We can easily notice that the core is non-empty and contains an infinite number of imputations. The imputation of the core is a vector

$$(d_1 + r_1(\bar{\xi}_N(\eta^*) - 128.2), d_2 + r_2(\bar{\xi}_N(\eta^*) - 128.2)),$$

where $d_1, d_2, r_1, r_2 \in \mathbb{R}$ and satisfy (16).

5. Conclusions

We have constructed a model of cooperative stochastic games with mean-variance preferences. The utility function for mean-variance preference takes into account players' attention to risk. The risk sensitivity of a player is modeled by a linear combination of mathematical expectation and standard deviation of the payoff. We have provided formulas to calculate player's and coalition utilities for any stationary strategy profile. We have defined a cooperative stochastic game with mean-variance preferences in the form of characteristic function determined by a maxmin approach. A particular cooperative solution proposed for the game is the core. The natural research question arising in the area of dynamic games is the sustainability and time-consistency of the cooperative solution over time. The problem of stability of a cooperative agreement in stochastic games with mean preferences is discussed by Parilina and Tampieri [19], Parilina [18], Parilina and Petrosyan [21], where conditions of stable cooperation are examined. The other research question is how to determine another cooperative solution like the Shapley value, nucleolus, etc., for stochastic games with mean-variance preferences and how to sustain them over time.

Author Contributions: Conceptualization, E.P. and S.A.; methodology, E.P. and S.A.; software, E.P. and S.A.; validation, E.P. and S.A.; writing—original draft preparation, E.P. and S.A.; writing—review and editing, E.P. and S.A.; project administration, E.P. and S.A. All authors have read and agreed to the published version of the manuscript.

Funding: The work of the first author was supported by the Shandong Province “Double-Hundred Talent Plan” (No. WST2017009).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Appendix A.1. Proof of Proposition 1

In the proof, we omit the symbol of strategy-profile η and action-profile a to simplify the description of calculations. We use the following notations: $K_i^s = K_i^s(a)$, $E(\bar{\xi}_i) = E(\bar{\xi}_i(\eta))$, $\Pi = \Pi(\eta)$, $\hat{\Pi} = \hat{\Pi}(\eta)$, $\pi(l|m; a^m) = \pi(l|m)$. To prove Equation (7), we use the definition of the variance:

$$\text{Var}(\bar{\xi}_i) = E(\bar{\xi}_i^2) - E^2(\bar{\xi}_i).$$

Note some useful properties of Hadamard product $d \circ h$ of vectors $d, h \in \mathbb{R}^t$ used to prove Proposition 1:

1. $d \circ h = h \circ d$ for all $d, h \in \mathbb{R}^t$;
2. $(d \circ h)g = (d \circ g)h$ for all $d, h, g \in \mathbb{R}^t$;
3. $d \circ (h + g) = d \circ h + d \circ g$ for all $d, h, g \in \mathbb{R}^t$;
4. $(dL \circ h) = (d \circ Lh)$ for all $d, h \in \mathbb{R}^t$ and for any matrix $L = [L_{ij}]_{i=1, \dots, t, j=1, \dots, t}$, $L_{ij} \in \mathbb{R}$.

Determine an expectation of the squared payoff to Player i in stochastic game G as follows:

$$E(\bar{\xi}_i^2) = W_\infty = \lim_{k \rightarrow \infty} W_k, \quad (\text{A1})$$

where W_k is the expectation of the squared payoff of the i th player in stochastic game G under condition that the game ends at stage k . The value of W_k can be obtained by the following equation:

$$W_k = \sum_{j_1=1}^t \sum_{j_2=1}^t \dots \sum_{j_k=1}^t \left(p_{j_1}^0 \pi(j_2|j_1) \dots \pi(j_k|j_{k-1}) \left(\sum_{l=1}^k \delta^{l-1} K_i^{j_l} \right)^2 \right)$$

or another form of the same equation:

$$W_k = \sum_{j_1=1}^t p_{j_1}^0 \sum_{j_2=1}^t \pi(j_2|j_1) \dots \sum_{j_k=1}^t \pi(j_k|j_{k-1}) \left(\sum_{l=1}^k \delta^{l-1} K_i^{j_l} \right)^2.$$

We calculate W_k recurrently and establish the general form of W_k using the method of mathematical induction. It is obvious that

$$W_1 = p^0(K_i \circ K_i).$$

The following computations are needed to determine W_2 :

$$\begin{aligned} W_2 &= \sum_{j_1=1}^t p_{j_1}^0 \sum_{j_2=1}^t \pi(j_2|j_1) (K_i^{j_1} + \delta K_i^{j_2})^2 = \sum_{j_1=1}^t p_{j_1}^0 \sum_{j_2=1}^t \pi(j_2|j_1) (K_i^{j_1})^2 + \\ &+ \delta^2 \sum_{j_1=1}^t p_{j_1}^0 \sum_{j_2=1}^t \pi(j_2|j_1) (K_i^{j_2})^2 + 2\delta \sum_{j_1=1}^t p_{j_1}^0 \sum_{j_2=1}^t \pi(j_2|j_1) K_i^{j_1} K_i^{j_2}. \end{aligned}$$

We obtain

$$\begin{aligned} W_2 &= \sum_{j_1=1}^t p_{j_1}^0 (K_i^{j_1})^2 \sum_{j_2=1}^t \pi(j_2|j_1) + \delta^2 \sum_{j_1=1}^t p_{j_1}^0 \Pi_{j_1} (K_i \circ K_i) + 2\delta \sum_{j_1=1}^t p_{j_1}^0 K_i^{j_1} \sum_{j_2=1}^t \pi(j_2|j_1) K_i^{j_2} = \\ &= \sum_{j_1=1}^t p_{j_1}^0 (K_i^{j_1})^2 + \delta^2 p^0 \Pi (K_i \circ K_i) + 2\delta \sum_{j_1=1}^t p_{j_1}^0 K_i^{j_1} \Pi_{j_1} K_i = \\ &= p^0(K_i \circ K_i) + \delta^2 p^0 \Pi (K_i \circ K_i) + 2\delta (p^0 \circ K_i) \Pi K_i, \end{aligned}$$

where Π_j is the j th row of matrix Π . We should notice that

$$W_2 = W_1 + \delta \{ \delta p^0 \Pi (K_i \circ K_i) + 2(p^0 \circ K_i) \Pi K_i \}.$$

Making similar calculations, we obtain an expression for the value of W_3 :

$$W_3 = W_2 + \delta^2 \{ \delta^2 p^0 \Pi^2 (K_i \circ K_i) + 2(p^0 \circ K_i) \Pi^2 K_i + 2\delta (p^0 \Pi \circ K_i) \Pi K_i \}.$$

Suppose the following formula is true for any $k = 1, 2, \dots, n-1$:

$$W_{k+1} = W_k + \delta^k \{ \delta^k p^0 \Pi^k (K_i \circ K_i) + 2 \sum_{j=1}^k \delta^{k-j} (p^0 \Pi^{k-j} \circ K_i) \Pi^j K_i \}, \quad (\text{A2})$$

and we prove it using the method of mathematical induction for $k = n$. Let Equation (A2) hold for $k = n-1$. The value of W_{n+1} is equal to the following one:

$$W_{n+1} = \sum_{j_1=1}^t p_{j_1}^0 \sum_{j_2=1}^t \pi(j_2|j_1) \dots \sum_{j_{n+1}=1}^t \pi(j_{n+1}|j_n) \left(\sum_{l=1}^{n+1} \delta^{l-1} K_i^{j_l} \right)^2.$$

Taking into account that

$$\left(\sum_{l=1}^{n+1} \delta^{l-1} K_i^{j_l}\right)^2 = \left(\sum_{l=1}^n \delta^{l-1} K_i^{j_l}\right)^2 + (\delta^n K_i^{j_{n+1}})^2 + 2\delta^n K_i^{j_{n+1}} \sum_{l=1}^n \delta^{l-1} K_i^{j_l},$$

we can write an expression:

$$\begin{aligned} W_{n+1} &= W_n + \sum_{j_1=1}^t p_{j_1}^0 \sum_{j_2=1}^t \dots \sum_{j_{n+1}=1}^t \pi(j_{n+1}|j_n) (\delta^n K_i^{j_{n+1}})^2 + \\ &+ \sum_{j_1=1}^t p_{j_1}^0 \sum_{j_2=1}^t \dots \sum_{j_{n+1}=1}^t \pi(j_{n+1}|j_n) \delta^{2n} (K_i^{j_{n+1}})^2 + \\ &+ 2\delta^n \sum_{j_1=1}^t p_{j_1}^0 \sum_{j_2=1}^t \dots \sum_{j_{n+1}=1}^t \pi(j_{n+1}|j_n) K_i^{j_{n+1}} \sum_{l=1}^n \delta^{l-1} K_i^{j_l} = \\ &= W_n + \delta^n \left\{ \delta^n p^0 \Pi^n (K_i \circ K_i) \right. \\ &\left. + 2\delta^n \sum_{l=1}^n \sum_{j_1=1}^t \dots \sum_{j_{n+1}=1}^t \left(p_{j_1}^0 \pi(j_2|j_1) \dots \pi(j_l|j_{l-1}) \delta^{l-1} K_i^{j_l} \times \pi(j_{l+1}|j_l) \dots \pi(j_{n+1}|j_n) K_i^{j_{n+1}} \right) \right\} \end{aligned}$$

Taking into account Equation (A2) for $k = n$, which is true by an assumption of mathematical induction, we obtain an expression for W_{n+1} :

$$\begin{aligned} W_{n+1} &= W_n + \delta^n \left\{ \delta^n p^0 \Pi^n (K_i \circ K_i) + 2 \sum_{l=1}^n \left(\delta^{n-l} (p^0 \Pi^{l-1} \circ K_i) \Pi^{n-l+1} K_i \right) \right\} \\ &= W_n + \delta^n \left\{ \delta^n p^0 \Pi^n (K_i \circ K_i) + 2 \sum_{l=1}^n \delta^{n-l} (p^0 \Pi^{k-l} \circ K_i) \Pi^l K_i \right\}. \end{aligned}$$

Taking into account Equation (A1) and expression for W_k , we obtain

$$W_\infty = \lim_{k \rightarrow \infty} (W_{k-1} + \Omega_k) = \lim_{k \rightarrow \infty} (W_{k-2} + \Omega_{k-1} + \Omega_k) = \dots = \lim_{k \rightarrow \infty} (W_1 + \sum_{j=1}^k \Omega_j) = W_1 + \lim_{k \rightarrow \infty} \sum_{j=1}^k \Omega_j,$$

where

$$\begin{aligned} \Omega_j &= \delta^j \{ \delta^j p^0 \Pi^j (K_i \circ K_i) + 2 \sum_{l=1}^j \delta^{j-l} (p^0 \Pi^{l-1} \circ K_i) \Pi^l K_i \} \\ &= p^0 (\delta^2 \Pi)^j (K_i \circ K_i) + 2 \delta^j \sum_{l=1}^j (p^0 (\delta \Pi)^{j-l} \circ K_i) \Pi^l K_i. \end{aligned}$$

To find the sum $\sum_{j=1}^k \Omega_j$, we first calculate

$$\sum_{j=1}^k p^0 (\delta^2 \Pi)^j (K_i \circ K_i) = p^0 (\tilde{\Pi} - \mathbb{I})(K_i \circ K_i),$$

where $\tilde{\Pi} = (\mathbb{I} - \delta^2 \Pi)^{-1}$.

Then, we calculate the second sum:

$$\begin{aligned} 2 \sum_{j=1}^k \delta^j \sum_{l=1}^j (p^0 (\delta \Pi)^{j-l} \circ K_i) \Pi^l K_i &= 2 \sum_{n=0}^{\infty} \sum_{s=1}^{\infty} \delta^n (p^0 (\delta \Pi)^n \circ K_i) (\delta \Pi)^s K_i \\ &= 2 (p^0 \tilde{\Pi} \circ K_i) (\hat{\Pi} - \mathbb{I}) K_i. \end{aligned}$$

Taking into account the last two expressions, we obtain

$$\begin{aligned} W_\infty &= p^0(K_i \circ K_i) + p^0(\tilde{\Pi} - \mathbb{I})(K_i \circ K_i) + 2(p^0\tilde{\Pi} \circ K_i)(\hat{\Pi} - \mathbb{I})K_i \\ &= p^0\tilde{\Pi}(K_i \circ K_i) + 2(p^0\tilde{\Pi} \circ K_i)(\hat{\Pi} - \mathbb{I})K_i. \end{aligned}$$

Therefore, we obtain the following expression for $E(\bar{\xi}_i^2)$:

$$E(\bar{\xi}_i^2) = p^0\tilde{\Pi}(K_i \circ K_i) + 2(p^0\tilde{\Pi} \circ K_i)(\hat{\Pi} - \mathbb{I})K_i. \quad (\text{A3})$$

Since $p^0\tilde{\Pi}(K_i \circ K_i)$ is equal to $(p^0\tilde{\Pi} \circ K_i)K_i$ by the property of Hadamard product, Equation (A3) can be rewritten as follows

$$E(\bar{\xi}_i^2) = (p^0\tilde{\Pi} \circ K_i)(2\hat{\Pi} - \mathbb{I})K_i.$$

There is a final formula for the variance of Player i 's payoff:

$$\text{Var}(\bar{\xi}_i) = (p^0\tilde{\Pi} \circ K_i)(2\hat{\Pi} - \mathbb{I})K_i - (p^0\hat{\Pi}K_i)^2.$$

Appendix A.2. Proof of Proposition 2

As in Appendix A.1, we omit the symbols of strategy profile in notations. So, we have $K_i^s = K_i^s(a^s)$, $E(\bar{\xi}_i) = E(\bar{\xi}_i(\eta))$, $\Pi = \Pi(\eta)$, $\hat{\Pi} = \hat{\Pi}(\eta)$, $\xi = \xi(\eta)$, $\pi(l|m; a^m) = \pi(l|m)$.

Equation (8) takes place because the expectation of payoff ξ_C is equal to the sum of expectations of the payoffs ξ_i of the players from coalition C and it follows from (5).

To prove Equation (9), we take into account that $\bar{\xi}_C = \sum_{i \in C} \bar{\xi}_i$ and use the following formula for computation of the variance of coalition C 's payoff in stochastic game G :

$$\text{Var}(\bar{\xi}_C) = \sum_{i \in C} \text{Var}(\bar{\xi}_i) + 2 \sum_{\substack{l < m, \\ l, m \in C}} \text{cov}(\bar{\xi}_l, \bar{\xi}_m). \quad (\text{A4})$$

The variance of $\bar{\xi}_i$ can be calculated by Equation (7), so we should obtain the formula to calculate covariance of the l th and m th players' payoffs. Covariance $\text{cov}(\bar{\xi}_l, \bar{\xi}_m)$ can be derived as follows:

$$\text{cov}(\bar{\xi}_l, \bar{\xi}_m) = E(\bar{\xi}_l \bar{\xi}_m) - E(\bar{\xi}_l)E(\bar{\xi}_m).$$

The way of determining covariance $\text{cov}(\bar{\xi}_l, \bar{\xi}_m)$ is similar to the way of determining $E(\bar{\xi}_i^2)$ in Appendix A.1. That is why we will not perform all computations in detail.

The expectation of multiplication of Player l and m 's payoffs is

$$E(\bar{\xi}_l \bar{\xi}_m) = W'_\infty = \lim_{k \rightarrow \infty} W'_k, \quad (\text{A5})$$

where W'_k is the expectation of payoffs multiplication under condition that the game ends at stage k . The value of W'_k can be expressed in the following form:

$$W'_k = \sum_{j_1=1}^t p_{j_1}^0 \sum_{j_2=1}^t \pi(j_2|j_1) \dots \sum_{j_k=1}^t \pi(j_k|j_{k-1}) \left(\sum_{r=1}^k \delta^{r-1} K_l^{j_r} \right) \left(\sum_{r=1}^k \delta^{r-1} K_m^{j_r} \right).$$

It is obvious that

$$W'_1 = p^0(K_l \circ K_m).$$

Expression for W'_2 is

$$W'_2 = W'_1 + \delta \{ \delta p^0 \Pi(K_l \circ K_m) + (p^0 \circ K_l) \Pi K_m + (p^0 \circ K_m) \Pi K_l \}.$$

Making similar computation as in the previous proof, we can write the recurrent equation for W'_k :

$$W'_{k+1} = W'_k + \delta^k \left\{ \delta^k p^0 \Pi^k (K_l \circ K_m) + \sum_{j=1}^k \delta^{k-j} \{ (p^0 \Pi^{k-j} \circ K_l) \Pi^j K_m + (p^0 \Pi^{k-j} \circ K_m) \Pi^j K_l \} \right\}.$$

With respect to (A5), we obtain

$$\begin{aligned} E(\bar{\xi}_l \bar{\xi}_m) &= \sum_{k=0}^{\infty} p^0 (\delta^2 \Pi)^k (K_l \circ K_m) \\ &+ \sum_{k=1}^{\infty} \delta^k \left\{ \sum_{j=1}^k (p^0 (\delta \Pi)^{k-j} \circ K_l) \Pi^j K_m + \sum_{j=1}^k (p^0 (\delta \Pi)^{k-j} \circ K_m) \Pi^j K_l \right\}. \end{aligned} \quad (\text{A6})$$

Simplifying the summands of the last expression of Equation (A6), we can rewrite it as follows

$$E(\bar{\xi}_l \bar{\xi}_m) = p^0 \tilde{\Pi} (K_l \circ K_m) + (p^0 \tilde{\Pi} \circ K_l) (\hat{\Pi} - \mathbb{I}) K_m + (p^0 \tilde{\Pi} \circ K_m) (\hat{\Pi} - \mathbb{I}) K_l. \quad (\text{A7})$$

Since $p^0 \hat{\Pi} (K_l \circ K_m)$ is equal to $(p^0 \hat{\Pi} \circ K_m) K_l$ and to $(p^0 \hat{\Pi} \circ K_l) K_m$ by the property of Hadamard product, Equation (A7) can be rewritten as follows

$$E(\bar{\xi}_l \bar{\xi}_m) = (p^0 \tilde{\Pi} \circ K_l) (\hat{\Pi} - 0.5 \mathbb{I}) K_m + (p^0 \tilde{\Pi} \circ K_m) (\hat{\Pi} - 0.5 \mathbb{I}) K_l. \quad (\text{A8})$$

Following Equations (7), (A4) and (A8), the covariance of payoffs of Players l and m in stochastic game G can be found by formula:

$$\text{cov}(\bar{\xi}_l, \bar{\xi}_m) = (p^0 \tilde{\Pi} \circ K_l) (\hat{\Pi} - 0.5 \mathbb{I}) K_m + (p^0 \tilde{\Pi} \circ K_m) (\hat{\Pi} - 0.5 \mathbb{I}) K_l - (p^0 \hat{\Pi} K_l) (p^0 \hat{\Pi} K_m).$$

So, we prove that formula (9) is true, where $\text{cov}(\bar{\xi}_l, \bar{\xi}_m)$ can be determined by (10).

References

- Shapley, L.S. Stochastic games. *Proc. Natl. Acad. Sci. USA* **1953**, *39*, 1095–1100. [\[CrossRef\]](#)
- Fink, A.M. Equilibrium in a stochastic n-person game. *J. Sci. Hiroshima Univ.* **1964**, *28*, 89–93. [\[CrossRef\]](#)
- Takahashi, M. Equilibrium points of stochastic, noncooperative n-person games. *J. Sci. Hiroshima Univ.* **1964**, *28*, 95–99. [\[CrossRef\]](#)
- Sobel, M.J. Non-cooperative stochastic games. *Ann. Math. Statist.* **1971**, *42*, 1930–1935. [\[CrossRef\]](#)
- Neyman, A.; Sorin, S. *Stochastic Games and Applications*; Nato Science Series C; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2003; Volume 570.
- Solan, E. Stochastic games. In *Encyclopedia of Database Systems*; Springer: Berlin, Germany, 2009.
- Solan, E.; Vieille, N. Stochastic games. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 13743–13746. [\[CrossRef\]](#) [\[PubMed\]](#)
- Markowitz, H.M. Portfolio Selection. *J. Financ.* **1952**, *7*, 77–91.
- Csóka, P.; Herings, P. J.-J.; Kóczy, L.A. Stable allocations of risk. *Games Econ. Behav.* **2009**, *67*, 266–276. [\[CrossRef\]](#)
- Suijs, J.; Borm, P. Stochastic cooperative games: Superadditivity, convexity, and certainty equivalents. *Games Econ. Behav.* **1999**, *27*, 331–345. [\[CrossRef\]](#)
- Suijs, J.; Borm, P.; De Waegenaere, A.; Tijs, S. Cooperative Games with Stochastic Payoffs. *Eur. J. Oper. Res.* **1999**, *113*, 193–205. [\[CrossRef\]](#)
- von Neumann, J.; Morgenstern, O. *Theory of Games and Economic Behavior*; Princeton University Press: Princeton, NJ, USA, 1944.
- Gromova, E.V.; Petrosyan, L.A. On an approach to constructing a characteristic function in cooperative differential games. *Automat. Remote Control* **2017**, *78*, 1680–1692. [\[CrossRef\]](#)
- Petrosjan, L.; Zaccour, G. Time-consistent Shapley value allocation of pollution cost reduction. *J. Econ. Dyn. Control* **2003**, *27*, 381–398. [\[CrossRef\]](#)
- Reddy, P.V.; Zaccour, G. A friendly computable characteristic function. *Math. Soc. Sci.* **2016**, *82*, 18–25. [\[CrossRef\]](#)
- Petrosjan, L.A. Cooperative Stochastic Games. In *Advances in Dynamic Games: Application to Economics, Engineering and Environmental Management*; Annals of the International Society of Dynamic Games; Haurie, A., Muto, S., Petrosjan, L.A., Raghavan, T.E.S., Eds.; Birkhäuser: Basel, Switzerland, 2006; pp. 139–146.
- Petrosjan, L.A.; Baranova, E.M. Cooperative Stochastic Games in Stationary Strategies. *Game Theory Appl.* **2006**, *11*, 7–17.
- Parilina, E.M. Stable cooperation in stochastic games. *Automat. Remote Control* **2015**, *76*, 1111–1122. [\[CrossRef\]](#)
- Parilina, E.M.; Tampieri, A. Stability and cooperative solution in stochastic games. *Theory Decision* **2018**, *80*, 601–625. [\[CrossRef\]](#)

20. Parilina, E.; Zaccour, G. Node-consistent core for games played over event trees. *Automatica* **2015**, *53*, 304–311. [[CrossRef](#)]
21. Parilina, E.; Petrosyan, L. On a Simplified Method of Defining Characteristic Function in Stochastic Games. *Mathematics* **2020**, *8*, 1135. [[CrossRef](#)]
22. Timmer, J.; Borm, P.; Tijs, S. On three Shapley-like solutions for cooperative games with random payoffs. *Int. J. Game Theory* **2003**, *32*, 595–613.
23. Habis, H.; Herings, P. J.-J. Transferable utility games with uncertainty. *J. Econ. Theory* **2011**, *146*, 2126–2139. [[CrossRef](#)]
24. Wei, Q. Nonzero-sum risk-sensitive finite-horizon continuous-time stochastic games. *Stat. Probabil. Lett.* **2019**, *147*, 96–104. [[CrossRef](#)]
25. Bäuerle, N.; Rieder, U. Zero-sum risk-sensitive stochastic games. *Stoch. Process. Appl.* **2017**, *127*, 622–642. [[CrossRef](#)]
26. Wei, Q.; Chen, X. Risk-Sensitive Average Equilibria for Discrete-Time Stochastic Games. *Dyn. Games Appl.* **2019**, *9*, 521–549. [[CrossRef](#)]
27. Nowak, A.S. Notes on Risk-Sensitive Nash Equilibria. In *Advances in Dynamic Games. Annals of the International Society of Dynamic Games*; Nowak, A.S., Szajowski, K., Eds.; Birkhäuser: Boston, MA, USA, 2005; Volume 7.
28. Jaśkiewicz, A.; Nowak, A.S. Stationary Markov perfect equilibria in risk sensitive stochastic overlapping generations models. *J. Econ. Theory* **2014**, *151*, 411–447. [[CrossRef](#)]
29. Asienkiewicz, H.; Balbus, Ł. Existence of Nash equilibria in stochastic games of resource extraction with risk-sensitive players. *TOP* **2019**, *27*, 502–518. [[CrossRef](#)]
30. Saldi, N.; Basar, T.; Raginsky, M. Approximate markov-nash equilibria for discrete-time risk-sensitive mean-field games. *Math. Oper. Res.* **2020**, *45*, 1596–1620. [[CrossRef](#)]