A stochastic approach to some values in cooperative games

Stefano Benati[a,*], Fernando López-Blázquez[b], Justo Puerto[b]

[a] Dipartimento di Sociologia e Ricerca Sociale. Università di Trento.
[b] Instituto Universitario de Investigación Matemática (IMUS). Universidad de Sevilla.

Abstract

Computing values in cooperative games is a hard task because, in general, it involves the evaluation of a non-polynomial number of terms. Sampling theory gives alternative ways to approximately compute these values and, at the same time, measures the error of these estimations. In this paper we present an attempt to estimate values in cooperative games of any number of players using a probabilistic approximation of a cooperative game. We apply this technique to several well-known games and report on the computational results of our approximation.

Keywords: Cooperative Game Theory, Values, Stochastic Games, Statistical sampling.

1. Introduction

Let (N, v) be a cooperative game with set of players $N = \{1, \dots, n\}$ and characteristic function v, $v(\emptyset) = 0$, and let G^N denote the class of N-person cooperative games. The main goal of this class of games is to allocate the grand coalition payoff $v(N)$ among the players in the game. There exist many allocation rules with a number of different properties and any of these rules may be considered as a solution concept of the game.

A well-accepted solution concept in cooperative game theory is the concept of value. A value $\alpha : G^N \to \mathbb{R}^n$ is a n-tuple of real numbers (one per agent in the grand coalition) where the i-th component is the amount allocated to agent i in the game (N, v).

These values are determined by the properties required to them. Linearity is a desirable property that has been required to most of the values considered in the literature. Thus, we will be interested in studying values whose functional form fits into the following formula:

$$\alpha(\boldsymbol{v}_*) = \sum_{S \in \mathcal{U}} a(S) \circ v_*(S), \tag{1}$$

where $\circ$ is the Hadamard product of vectors, $a(S) = (a_1(S), \dots, a_n(S))^t$, $S \in \mathcal{U}$ are real n-vectors depending only on S, $\mathcal{U} = 2^N$ the family of all the coalitions of the game and $\boldsymbol{v}_* = (v_*(S))^t_{S \in \mathcal{U}}$, where $v_*(S) := (v_{1,*}(S, v), \dots, v_{n,*}(S, v))^t$ are n-vectors of real values depending on S, the characteristic function of the game and eventually on each player in the game. There are

* Corresponding author
Email addresses: stefano.benati@unitn.it (Stefano Benati), lopez@us.es (Fernando López-Blázquez), puerto@us.es (Justo Puerto)

many examples of these values although the best known may be the Shapley value, the family of semivalues [4] and least squares values [14].

- The Shapley value of the i-player, for the cooperative game (N, v), is defined as

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{s!\,(n-s-1)!}{n!}(v(S \cup \{i\}) - v(S)), \tag{2}$$

  where $|S| = s$ denotes the cardinal of the set S. The Shapley value corresponds to the form given in (1) with $a_i(S)$ equal to $\frac{s!(n-s-1)!}{n!}$, if $i \notin S$ and zero otherwise; and $v_{i,*}(S) := v(S \cup \{i\}) - v(S)$.

- The semivalues, introduced by Dubey et al. [4], are similar to the Shapley values but with different systems of coefficients:

$$\psi_i(v) = \sum_{S \subseteq N \setminus \{i\}} p_s(v(S \cup \{i\}) - v(S)),$$

  with $p_s \geq 0$, $s = 0, \ldots, n-1$, real numbers such that $\sum_{s=0}^{n-1} \binom{n-1}{s} p_s = 1$. The semivalue ψ corresponds to (1) with $a(S) = p_s$, for $S \subset N \setminus \{i\}$; $a(S) = 0$ otherwise, and $v_{i,*}(S) := v(S \cup \{i\}) - v(S)$.

- The family of least squares values introduced by [3] and [14] also fits to this framework (1). The least squares value for player i is defined as:

$$x_i(v) = \frac{v(N)}{n} + \frac{1}{n\kappa}(n a_i^m(v) - \sum_j a_j^m(v)) \tag{3}$$

  where $a_i^m(v) = \sum_{i \in S \subset N} m(s) v(S), |S| = s$ and $\kappa = \sum_{s=1}^{n-1} m(s) \binom{n-2}{s-1}$. The reader is referred to (19) to check that $x_i(v)$ can expressed in the form of (1). Indeed, $x_i(v) = \frac{v(N)}{n} + \sum\limits_{S \in \mathcal{U}} a_i(S) v_*(S)$, with $a_i(S)$ equal to $\frac{(n-s)m(s)}{n\kappa}$, if $i \in S$ and $-\frac{sm(s)}{n\kappa}$, otherwise; and $v_{i,*}(S) = v(S),\ S \in \mathcal{U}$.

In general, computing any value $\alpha(\boldsymbol{v}_*)$ of a cooperative game is NP-hard, since the number of evaluations needed to obtain it requires the evaluation of some expression for all the coalitions $S \subseteq N$ which is of the order $O(2^n)$. To avoid this inconvenience several authors have proposed different methodologies. Owen [12] proposed a method to compute the Shapley value via a probabilistic approximation of the game by a normal distribution. A different approach, using generating functions, only valid for simple games, is given in Bilbao et al. [1]. A first approach to approximate Shapley values of cooperative games using sampling can be found in Fernndez et al [5], [6]. Later, Keinan et al. [10] apply similar techniques to estimate Shapley values of some neurocontroller games that appear in neuroscience. In [2] the authors rediscover the above methodology and apply it to some standard cooperative games.

In this paper we present a new method for approximating values of cooperative games. Rather than computing the exact values we concentrate on providing good estimations of these values with their corresponding measure of errors. To this end, we use a general probabilistic approximation of a cooperative game. On this approximation, we define a probabilistic value that is seen as an estimator of the actual value of the original game. This estimator is unbiased, consistent and with controlled quadratic error. Our computational experiments in different games with large number of players confirm the viability of our approach.

The paper is organized in four sections. After the introduction, the second section defines the probabilistic approximation of a cooperative game, and proves its main statistical and distributional properties. In Section 3 we present our family of estimators and their properties. Finally, in section four we test the above mentioned estimators on four different well-known classes of cooperative games showing the efficiency of the method.

Through this paper we will assume the usual conventions $\sum_a^b \cdot = 0$ and $\binom{a}{b} = 0$ if $a < b$.

2. The probabilistic approximation of a cooperative game

The goal of this section is to define a probabilistic approximation of a game (N, v) that allows us to estimate its values. The study of probabilistic or stochastic cooperative games is not new. The reader is referred to [8, 9, 15, 16, 17] for different approaches and further analysis on this class of games.

Let $\mathcal{U} = 2^N$ be the family of all the coalitions of the game and let $\mathcal{U}^\tau$ be the τ-fold cartesian product of $\mathcal{U}$. Therefore, any element $\mathbb{S} \in \mathcal{U}^\tau$, is a vector $\mathbb{S} = (S_1, \ldots, S_\tau)^t$ where each $S_i \subseteq N$. We assume further that we are given a probability distribution p on $\mathcal{U}$. Thus, for any $S \in \mathcal{U}$, $p(S)$ is the probability of choosing the coalition S. For simplicity, we will assume that $p(S) > 0$, for all $S \in \mathcal{U}$, although, as we will see later, this condition is not essential. This probability, p, induces on the space $\mathcal{U}^\tau$ the natural product probability.

We introduce a set of random vectors $\widetilde{v}_*(S) := (\widetilde{v}_{1,*}(S), \ldots, \widetilde{v}_{n,*}(S))$ associated with each coalition of the original game. For each coalition S and player $i \in N$, let us define the random variable $\widetilde{v}_{i,*}(S) := \widetilde{v}_{i,*}(S; \tau, p, \mathbb{S})$ as

$$\widetilde{v}_{i,*}(S) = \frac{M_S(\mathbb{S})}{\tau p(S)} v_{i,*}(S) \tag{4}$$

where $M_S := M_S(\mathbb{S}) = \#\{j : S_j \in \mathbb{S} \text{ and } S_j = S\}$ is the number of times that coalition S appears in the random vector of coalitions $\mathbb{S}$. Note that the random vector $\mathbb{M} = (M_S)_{S \in U}$ follows a multinomial distribution with parameters τ (number of elements in the sample $\mathbb{S}$) and $\mathbf{p} = (p(S))_{S \in \mathcal{U}}$ (vector of probabilities of selection for each possible set in the sample $\mathbb{S}$). We denote this fact by $\mathbb{M} \sim \mathcal{M}(\tau;\ \mathbf{p})$. We recall now some properties of multinomial vectors that will be useful in the following:

- $E\mathbb{M} = \tau \mathbf{p}$.
- $Var(\mathbb{M}) = \tau \Lambda$, with $\Lambda = diag(\mathbf{p}) - \mathbf{p}\mathbf{p}^t$.
- $\frac{\mathbb{M}}{\tau} \xrightarrow{a.s.} \mathbf{p}$, as $\tau \longrightarrow \infty$.
- $\sqrt{\tau}\left(\frac{\mathbb{M}}{\tau} - \mathbf{p}\right) \xrightarrow{d} \mathbf{W}$, as $\tau \longrightarrow \infty$, with $\mathbf{W}$ following a 2^n-dimensional normal distribution with mean $\mathbf{0}$ and variance matrix Λ, i.e., $\mathbf{W} \sim \mathcal{N}_{2^n}(\mathbf{0}, \Lambda)$.

From the properties of the multinomial vector $\mathbb{M}$ given above, it is not difficult to establish the following properties for the random elements $\widetilde{\mathbf{v}}_{i,*} = (\widetilde{v}_{i,*}(S))_{S\in\mathcal{U}}$. For simplicity throughout this paper we will assume that the characteristic function v (and $\widetilde{v}$) take values in the set of real numbers. The extension of our results to the case of vectorial characteristic functions is straightforward [13].

Theorem 1. *For all $S \in \mathcal{U}$ and $i \in N$,*

(a) $E\widetilde{v}_{i,*}(S) = v_{i,*}(S)$.

(b) $Var(\widetilde{v}_{i,*}(S)) = \dfrac{1-p(S)}{\tau p(S)} v_{i,*}^2(S)$

(c) $Cov(\widetilde{v}_{i,*}(S),\, \widetilde{v}_{i,*}(T)) = -\dfrac{v_{i,*}(S)\, v_{i,*}(T)}{\tau}$

(d) $\widetilde{v}_{i,*}(S) \xrightarrow{a.s.} v_{i,*}(S)$*, as* $\tau \longrightarrow \infty$*.*

(e) Let $\mathbf{v}_{i,*} = (v_{i,*}(S))_{S\in\mathcal{U}}$ *and* $\widetilde{\mathbf{v}}_{i,*} = (\widetilde{v}_{i,*}(S))_{S\in\mathcal{U}}$*, then* $\sqrt{\tau}\,(\widetilde{\mathbf{v}}_{i,*} - \mathbf{v}_{i,*}) \xrightarrow{d} \mathbf{Z} \sim \mathcal{N}_{2^n}(\mathbf{0},\, \mathbf{\Delta})$*, as* $\tau \longrightarrow \infty$*, where*

$$\mathbf{\Delta} := diag\left(\frac{v_{i,*}^2(S)}{p(S)}\right)_{S\in\mathcal{U}} - \mathbf{v}_{i,*}\mathbf{v}_{i,*}^t$$

Proof. Recall that $\mathbb{M} = (M_S)_{S\in\mathcal{U}}$ follows a multinomial distribution, then for all $S \in \mathcal{U}$, $M_S \sim \mathrm{B}inomial(\tau, p(S))$, so $EM_S = \tau p(S)$ and $\mathrm{Var}(M_S) = \tau p(S)\,(1-p(S))$. Also, for $S,\, T \in \mathcal{U}$ with $T \neq S$, $\mathrm{Cov}(M_S, M_T) = -\tau p(S)p(T)$.

(a) For all $S \subset N$,

$$E\widetilde{v}_{i,*}(S) = E\left(\frac{v_{i,*}(S)M_S}{\tau p(S)}\right) = \frac{v_{i,*}(S)}{\tau p(S)} EM_S = v_{i,*}(S).$$

(b) Using standard properties of the variance, we have

$$\mathrm{Var}\,(\widetilde{v}_{i,*}(S)) = \mathrm{Var}\left(\frac{v_{i,*}(S)M_S}{\tau p(S)}\right) = \left(\frac{v_{i,*}(S)}{\tau p(S)}\right)^2 \mathrm{Var}\,(M_S) = \frac{1-p(S)}{\tau p(S)} v_{i,*}^2(S).$$

(c) We have,

$$\begin{aligned}\mathrm{Cov}(\widetilde{v}_{i,*}(S),\, \widetilde{v}_{i,*}(T)) &= \mathrm{Cov}\left(\frac{v_{i,*}(S)M_S}{\tau p(S)},\, \frac{v_{i,*}(T)M_T}{\tau p(T)}\right) = -\frac{v_{i,*}(S)\, v_*(T)}{\tau^2 p(S)p(T)} \mathrm{Cov}\,(M_S,\, M_T) \\ &= -\frac{v_{i,*}(S)\, v_*(T)}{\tau}.\end{aligned}$$

(d) It is straightforward from the fact that $M_S/\tau \xrightarrow{a.s.} p(S)$, as $\tau \uparrow +\infty$, for all $S \in \mathcal{U}$.

(e) Observe that $\widetilde{\mathbf{v}}_{i,*} = \text{diag}\left(\frac{v_{i,*}(S)}{\tau p(S)},\ S \in \mathcal{U}\right)\mathbb{M}$, then the asymptotic normality follows from the asymptotic normality of the multinomial vector $\mathbb{M}$ and the limit covariance matrix follows from *(a)* and *(b)*.

■

In the statistical language, we may interpret $\widetilde{v}_{i,*}(S)$ as an '*estimator*' of the value $v_{i,*}(S)$ based on a sample of size τ. So, property *(a)* in Theorem 1, means that the '*estimator*' $\widetilde{v}_{i,*}(S)$ is unbiased for $v_{i,*}(S)$, for any $S \in \mathcal{U}$. Properties *(b)* and *(c)* allow us to determine variances and covariances between these '*estimators*'. Properties *(d)* and *(e)* (consistency and asymptotic normality) show how the '*estimators*' behave as the sample size increases: they approach the 'true values'. Note that here, we use '*estimator*' rather than estimator, because in the strict statistical sense an estimator must not depend on the estimated value. Anyway, we believe that this statistical interpretation is interesting as we will see in the following sections.

3. Estimation of values in a cooperative game

The main practical difficulty with quantities of the form given in (1) is that they are not easy to compute due to the fact that there are too many terms in the sum (usually 2^{N-1} or 2^N). The aim of this section is to provide approximations to these values based on the probabilistic approximation to the game given in section 2.

Our approach is based on a sample of size τ, $\mathbb{S} = (S_1, \ldots, S_\tau)$ selected according to the probability scheme described in the previous section. The proposed probabilistic approximation to $\alpha(\boldsymbol{v})$ is given by:

$$\widetilde{\alpha(\boldsymbol{v}_*)} := \alpha(\widetilde{\boldsymbol{v}}_*) = \sum_{S\in\mathcal{U}} a(S) \circ \widetilde{v}_*(S) = \sum_{S\in\mathcal{U}} \frac{a(s) \circ v_*(s)}{\tau p(S)} M_S = \left(\frac{1}{\tau}\sum_{S\in\mathbb{S}} \frac{a_i(S) v_{i,*}(S)}{p(S)}\right)_{i\in N}, \quad (5)$$

where $\circ$ is the Hadamard product of vectors and the last equality uses the fact that $M_S = 0$ for all $S \notin \mathbb{S}$. The advantage of using this approximation is that it involves a much smaller number of terms in the summation than the true expression of the value.

Now, we will study some properties of these approximations. In the sequel and for the sake of presentation we will assume that for any vector $x \in \mathbb{R}^n$, its squares $x^2 = (x_1^2, \ldots, x_n^2)^t = x \circ x$, denotes the Hadamard product of x times itself.

Theorem 2. *The approximation* $\widetilde{\alpha(\boldsymbol{v}_*)} = (\widetilde{\alpha_i(\boldsymbol{v}_*)})_{i=1}^n$ *satisfies:*

(a) $E\widetilde{\alpha_i(\boldsymbol{v}_*)} = \alpha_i(\boldsymbol{v}_*),\ \forall\, i \in N.$

(b) $Var(\widetilde{\alpha_i(\boldsymbol{v}_*)}) = \frac{1}{\tau}\left\{\sum_{S\in\mathcal{U}} \frac{a_i^2(S) v_{i,*}^2(S)}{p(S)} - \alpha_i^2(\boldsymbol{v}_*)\right\} := \frac{1}{\tau}\sigma_i^2,\ \forall\, i \in N.$

(c) $\widetilde{\alpha(\boldsymbol{v}_*)} \xrightarrow{a.s.} \alpha(\boldsymbol{v}_*)$, *as* $\tau \longrightarrow \infty$.

(d) $\sqrt{\tau}\left(\widetilde{\alpha_i(\boldsymbol{v}_*)} - \alpha_i(\boldsymbol{v}_*)\right)/\sigma_i \xrightarrow{d} Z \sim \mathcal{N}(0,1),\ \forall\, i \in N.$

Proof. These results are straightforward from Theorem 1 and the explicit expressions for $\alpha(\widetilde{\boldsymbol{v}}_*)$ given in (5). ∎

We will follow the usual convention $0/0 = 0$. With this convention we can assign $p(S) = 0$ in those expressions of S with $a_i(S) = 0$. Doing so, expressions like $a(S) \circ \widetilde{v}_*(S) = M_S v_*(S)/(\tau p(S))$ and similar ones make sense.

Again, property *(a)* in Theorem 2 shows that our approximations are unbiased and property *(b)* means that as the sample size increases, i.e., $\tau \to \infty$, the approximation approaches (almost surely) the true value $\alpha(\boldsymbol{v}_*)$. The vector $\sigma^2 = (\sigma_1^2, \ldots, \sigma_n^2)^t$, in Theorem 2(b) is related to the precision of our approximations. Let us write

$$\sigma^2 := A_1 - A_2, \text{ where } A_1 := \sum_{S \in \mathcal{U}} \frac{a^2(S) \circ v_*(S)}{p(S)} \text{ and } A_2 := \alpha^2(\boldsymbol{v}_*),$$

and let us define

$$\widetilde{A_1} := \frac{1}{\tau} \sum_{S \in \mathbb{S}} \frac{a^2(S) \circ v_*^2(S)}{p^2(S)}.$$

We have

$$\begin{aligned} E\widetilde{A_1} &:= \frac{1}{\tau} E\left(\sum_{S \in \mathbb{S}} \frac{a^2(S) \circ v_*^2(S)}{p^2(S)} \right) = \frac{1}{\tau} E\left(\sum_{S \in \mathcal{U}} \frac{a^2(S) \circ v_*^2(S)}{p^2(S)} M_S \right) \\ &= E\left(\sum_{S \in \mathcal{U}} \left(\frac{a^2(S) \circ v_*(S)}{p(S)} \right) \circ \widetilde{v}_*(S) \right) = \sum_{S \in \mathcal{U}} \frac{a^2(S) \circ v_*(S)}{p(S)} \circ E\widetilde{v}_*(S) = A_1, \end{aligned}$$

then the approximation $\widetilde{A_1}$ is unbiased for A_1 and as an approximation to σ^2 we propose

$$\widetilde{\sigma^2} := \widetilde{A_1} - \widetilde{\alpha}^2(\boldsymbol{v}_*). \tag{6}$$

In practise, we should determine the sample size τ in such a way that $\widetilde{\alpha(\boldsymbol{v}_*)}$ is 'close enough' to $\alpha(\boldsymbol{v}_*)$. For that purpose, the normal approximation given in property *(d)* of the previous theorem is useful. For instance, suppose that τ will be determined under the condition that

$$P\left(\left| \widetilde{\alpha_i(\boldsymbol{v}_*)} - \alpha_i(\boldsymbol{v}_*) \right| > \varepsilon \right) \leq \beta$$

where $\varepsilon > 0$ and $\beta \in (0, 1)$ are given. Using the normal approximation, we obtain

$$\tau \geq \left(\frac{\sigma_i}{\varepsilon} \quad \Phi^{-1}\left(1 - \frac{\beta}{2}\right) \right)^2 \tag{7}$$

where Φ^{-1} denotes the inverse of the distribution function of a $\mathcal{N}(0, 1)$ random variable.

The formula (7) for the sample size may have the drawback that the value of σ_i^2 is not known. The solution is that we must approximate σ_i^2 using a sample of size τ_1 (usually much smaller that τ). This smaller sample plays here the same role as the pilot samples (or training samples) frequently used in Statistics to obtain brief estimators of quantities of a secondary interest.

These brief estimators are used to estimate some other quantities with more precision. As an approximation to the unknown variance, σ_i^2, we propose here the use of $\widetilde{\sigma_{i,\tau_1}^2}$, obtained from the expression given in (6) by using a pilot or training sample of size τ_1. Finally, the sample size, τ, is calculated by using (7) with σ_i replaced by its approximation $\widetilde{\sigma_{i,\tau_1}}$.

Also, the normal approximation can be used to obtain an (approximated) $(1-\beta)$- confidence interval for $\alpha(\boldsymbol{v}_*)$

$$\widetilde{\alpha(\boldsymbol{v}_*)} \pm \frac{\widetilde{\sigma}}{\sqrt{\tau}} \Phi^{-1}(1-\beta/2) \tag{8}$$

These intervals can be useful to obtain some information about the precision of the procedure used in the approximation.

To illustrate the above developments we include an easy cooperative game where all the elements described a in this section can be followed.

Example 3. *A worked example: Mean value and Shapley value from symmetric voting* Let us consider the game (N, v) where $N = \{1, 2, \ldots, n\}$ and the characteristic function is $v(S) = 1$, if $|S| > n/2$ and 0 otherwise. Suppose, for simplicity, that n is odd. Let us estimate the mean value of the characteristic function of all coalitions containing player i, namely,

$$\alpha_i(\boldsymbol{v}) = \frac{1}{2^{n-1}} \sum_{S:\ i \in S \subset U} v(S). \tag{9}$$

Observe that in this case we have $v_{i,*}(S) = v(S)$, for all $S \subseteq N$ and for all $i \in N$. Then, for any i, it is not difficult to check that $\alpha_i(\boldsymbol{v}) = 1/2$.

For the estimation of (9) we select subsets with uniform probability, $p(S) = 2^{-(n-1)}$, in the family of sets not containing the i-player, i.e., for S such that $i \in S$ and $p(S) = 0$ otherwise. So, the approximation proposed in (5) is

$$\widetilde{\alpha_i(\boldsymbol{v})} := \frac{1}{\tau} \sum_{S \in \mathbb{S}} v(S), \tag{10}$$

where $\mathbb{S} = (S_1, \ldots, S_\tau)$ is the random sample of size τ of the selected subsets. Using (6), we get

$$\widetilde{\sigma_i^2} = \frac{1}{\tau} \sum_{S \in \mathbb{S}} v^2(S) - \widetilde{\alpha}_i^{\,2}. \tag{11}$$

It is an easy exercise to check that $W := \sum_{S \in \mathbb{S}} v(S)$ is *the number of subsets with cardinal* $>$ $n/2$ *among the* τ *selected*. Clearly W follows a binomial distribution with parameters τ and parameter of success $P(|S| > n/2) = 1/2$, that is to say $W \sim Binomial(\tau, 1/2)$.

Using elementary properties of binomial distributions, it is not difficult to obtain directly the properties established in Theorem 2, for instance, $E\widetilde{\alpha_i(\boldsymbol{v})} = 1/2$; $Var(\widetilde{\alpha_i(\boldsymbol{v})}) = \sigma_i^2/\tau$ with $\sigma_i^2 = 1/4$ and also the asymptotic properties $\widetilde{\alpha_i(\boldsymbol{v})} \xrightarrow{a.s.} 1/2$, and $2\sqrt{\tau}\left(\widetilde{\alpha_i(\boldsymbol{v})} - 1/2\right) \xrightarrow{d} \mathcal{N}(0,1)$, as $\tau \longrightarrow \infty$.

We shall also illustrate the approximation of the Shapley value of this game. In this very simple game the exact Shapley value is $\phi_i = 1/n$, for $i = 1, \ldots, n$.

We take $v_{*,i}(S) = v(S\cup\{i\}) - v(S) = 1$ if $S \subset N\setminus\{i\}$ and $|S| = \lfloor n/2 \rfloor$; $v_{i,*}(S) = 0$ otherwise. Under the uniform sampling scheme described above

$$\widetilde{\phi_i(\boldsymbol{v})} = \frac{1}{\tau}F \quad \text{and} \quad \widetilde{\sigma_i^2} = \widetilde{\phi_i(\boldsymbol{v})}(1 - \widetilde{\phi_i(\boldsymbol{v})}),$$

where F=*number of subsets with cardinal equal to* $\lfloor n/2 \rfloor$ *among the* τ *selected.* It is clear that $F \sim \mathit{Binomial}(\tau, 1/n)$, and again it is immediate to check directly that $E\widetilde{\phi_i(\boldsymbol{v})} = \phi_i$, $\mathrm{Var}(\widetilde{\phi_i(\boldsymbol{v})}) = \sigma_i^2/\tau$ with $\sigma_i^2 = (n-1)/n^2$ and the asymptotic properties $\widetilde{\phi_i(\boldsymbol{v})} \xrightarrow{a.s.} 1/n$, and $\sqrt{\tau}\left(\widetilde{\phi_i(\boldsymbol{v})} - 1/n\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{n-1}{n^2}\right)$, as $\tau \longrightarrow \infty$.

4. Applications to some distinguished cooperative games

In this section we will apply the concepts presented in the previous sections to some distinguished cooperative games: the gloves game, the airport game, the voting game and the linear production game. For these games we discuss the estimation of mainly two values that capture the powerfulness and flexibility of our approach: the mean value of the characteristic function of all coalitions containing player i, and the Shapley and the least square value of the i-th player, see (2) and (3).

We define the mean value of the characteristic function of all coalitions containing player i as:

$$\alpha_i(\boldsymbol{v}) = \frac{1}{2^{n-1}} \sum_{S:\ i\in S\subset U} v(S). \tag{12}$$

4.1. Sampling strategies

For the estimation of (12) we select subsets with uniform probability, $p(S) = 2^{-(n-1)}$, in the family of sets containing the i-player, i.e., for S such that $i \in S$ and $p(S) = 0$ otherwise. So, the approximation proposed in (5) is

$$\widetilde{\alpha}_i := \frac{1}{\tau}\sum_{S\in\mathbb{S}} v(S), \tag{13}$$

where $\mathbb{S} = (S_1, \ldots, S_\tau)$ is the random sample of size τ of the selected subsets. Using (6), we get

$$\widetilde{\sigma_i^2} = \frac{1}{\tau}\sum_{S\in\mathbb{S}} v^2(S) - \widetilde{\alpha}_i{}^2. \tag{14}$$

As we have seen in Section 3, our methodology applies to any sampling scheme, e.g., parameters $p(S)$. Nevertheless, to show our results we have resorted to two more sampling strategies: 1) Weighted sampling (WSS) and 2) Stratified sampling (SSS). In the following, we briefly describe these two strategies.

4.1.1. Weighted Sampling strategy

We select random subsets within the family of coalitions in two stages. Firstly, we choose the size s with uniform probability and in the second stage we choose at random a subset of size s with uniform probability within the resulting set of coalitions. Depending on whether we impose that a given player should belong or not to the corresponding set of coalitions it gives rise to slightly different probabilities and estimators.

For the case in which S and $S \cup \{i\}$ are to be considered, we choose $s \in \{0, \ldots, n-1\}$, the size of the subset (possibly empty) that does not contain player-i with uniform probability $(1/n)$ and in the second stage we choose at random a subset of size s with uniform probability $(1/\binom{n-1}{s})$. Then,

$$p(S) = \frac{1}{n\binom{n-1}{s}} = \frac{s!(n-s-1)!}{n!}, \quad \text{for } S \text{ such that } i \notin S, \tag{15}$$

and $p(S) = 0$ otherwise. We will refer to this strategy as (WSS1).

In the case in which is not required that i should be in S, we choose the size of the subset $s \in \{1, \ldots, n-1\}$ (excluding $\emptyset$ and N) with uniform probability $(n-1)^{-1}$. Secondly, we choose a subset of size s with uniform probability $\binom{n}{s}^{-1}$. Therefore:

$$p(S) = p(s) = \frac{1}{(n-1)\binom{n}{s}}, \quad \text{for } S \text{ such that } i \in S, \tag{16}$$

and $p(S) = 0$ otherwise. We will refer to this strategy as (WSS2).

4.1.2. Stratified sampling strategy

This strategy selects independent random samples of size τ_s within the family of coalitions of size s. If we require, for instance, that player i belongs to each chosen coalition and we draw uniformly within that collection of coalitions, then

$$p(S) = \frac{1}{\binom{n-1}{s-1}}, \quad \text{if } i \in S \text{ and } |S| = s,$$

and 0 otherwise. We will refer to this strategy as (SSS).

4.2. Estimation of the Shapley Value

As we mentioned in the introduction, the Shapley value may be the most important value for cooperative games. For that reason, we devote this section to derive specific implementations of the stochastic approximation presented above to the Shapley value. For the estimation of the Shapley value associated to the i-player, see (2), we will use two different approaches. The first one is based in its original expression (2) and the second one is based on the least squares formulas (3). The choice of the estimation motivates different sampling strategies.

4.2.1. Computing the Shapley value from the standard formula (2)

In order to apply our methodology to estimate the Shapley value using the original expressions (2), we can apply WSS1.

Then, the approximations given in (5) and (6) yield respectively,

$$\widetilde{\phi}_i = \frac{1}{\tau}\sum_{S\in\mathbb{S}} v_{*,i}(S) \quad \text{and} \quad \widetilde{\sigma_i^2} = \frac{1}{\tau}\sum_{S\in\mathbb{S}} v_{*,i}^2(S) - \widetilde{\phi}_i^{\,2}. \tag{17}$$

4.2.2. Computing the Shapley Value from the Least Squares formula

To perform the estimation of Shapley value from (3), we first show how to adapt the Least Squares formula (3) to be applicable within our framework to estimate this value. Then, we will compare two implementations of the stochastic approximation, relying on different interpretations of the Least Squares formula that lead to two different sampling strategies: 1) a weighted sampling; and 2) a stratified sampling.

The Least Squares value of a cooperative game is [14]:

$$x_i(v) = \frac{v(N)}{n} + \frac{1}{n\kappa}(\underbrace{na_i^m(v)}_{1^{st}\text{ term}} - \underbrace{\sum_j a_j^m(v)}_{2^{nd}\text{ term}}), \tag{18}$$

where $a_i^m(v) = \sum_{i\in S\subset N} m(s)v(S), |S| = s$ and $\kappa = \sum_{s=1}^{n-1} m(s)\binom{n-2}{s-1}$. Formula (18) can be rewritten as a weighted sum in the form of (1) (in which $\mathcal{U}$ are all non empty and proper subsets of N):

$$x_i(v) = \frac{v(N)}{n} + \sum_{S\in\mathcal{U}} a_i(S)v(S). \tag{19}$$

In equation (19) the weights $a_i(S)$ are determined as follows: the values $v(S)$ appear in (18) for all $S \subset N$. If $i \in S$, the term $v(S)$ appears in the first term with weight $n\,m(s)$, then in the second term with weight $s\,m(s)$ (one for each $j \in S$). If $i \notin S$, $v(S)$ appears only in the second term with weight $s\,m(s)$. So that:

$$a_i(S) = \begin{cases} \frac{(n-s)m(s)}{n\kappa}, & i \in S \\ -\frac{sm(s)}{n\kappa}, & i \notin S. \end{cases} \tag{20}$$

The Shapley value is then obtained when:

$$m(s) = \frac{1}{n-1}\binom{n-2}{s-1}^{-1} \text{ and } \kappa = 1. \tag{21}$$

Therefore, to get the stochastic approximation we must replace in the above expression the deterministic value $v(S)$ in (19) by the corresponding stochastic terms:

$$\widetilde{v}(S) = \frac{M_S(\mathbb{S})}{\tau p(S)} v(S). \tag{22}$$

Let $\mathbb{S} = (S_1, \ldots, S_\tau)$ be a random sample of $\mathcal{U}$ of size τ, the stochastic approximation of the Shapley value is:

$$\tilde{x}_i(v) = \frac{v(N)}{n} + \frac{1}{\tau} \sum_{S \in \mathcal{S}^\tau} \frac{a_i(s)}{p(S)} v(S), \quad \forall i \in N. \tag{23}$$

Next, we discuss two different sampling procedures to be applied to the above approximation of the Shapley value.

The first one is WSS2. The reason is because we observe that always $v(\emptyset) = 0$ and that from expression (20) $a_i(N) = 0$. Therefore, these two sets, namely $\emptyset$ and the grand coalition N, can be excluded from the sampling scheme provided that we estimate via (23).

Therefore, from equation (20), taking advantage that:

$$\binom{n-2}{s-1}^{-1} \binom{n}{s} = \frac{n(n-1)}{s(n-s)} \tag{24}$$

we have that:

$$\frac{a_i(s)}{p(S)} = \begin{cases} \frac{(n-1)}{s}, & i \in S \\ -\frac{n-1}{n-s}, & i \notin S. \end{cases} \tag{25}$$

These values are used in formula (23) to calculate the estimated $\tilde{x}_i(\boldsymbol{v})$. Moreover, they are used to calculate estimated variance using (6):

$$\widetilde{\sigma_i^2} = \widetilde{A_1} - \widetilde{A_2}$$

where

$$\widetilde{A_1} := \frac{1}{\tau} \sum_{S \in \mathbb{S}} \frac{a^2(S) v^2(S)}{p^2(S)}, \tag{26}$$

$$\widetilde{A_2} := (\tilde{x}_i(v) - \frac{v(N)}{n})^2. \tag{27}$$

Note that the estimated variance in (6) considers values that do not contain the deterministic value $\frac{v(N)}{n}$.

The second approach is to apply the *Stratified sampling* (SSS). In doing that we must decompose the summation (18) into separate terms $a_i^m(v)$ and observe that:

$$a_i^m(v) = \sum_{i \in S \subset N} m(s) v(S) = \sum_{s=1}^{n-1} \underbrace{\sum_{\substack{S \subset N \\ |S|=s \\ i \in S}} m(s) v(S)}_{a_{i,s}^m(v)}. \tag{28}$$

Assuming that sample sets of $\mathcal{U}_{i,s}$ of cardinal τ_s are drawn with uniform probability, we have $\frac{1}{p(S)} = \binom{n-1}{s-1}$ and as $m(s) = \frac{1}{n-1}\binom{n-2}{s-1}^{-1}$ results:

$$\widetilde{a^m_{i,s}(v)} = a^m_{i,s}(\tilde{v}) = \frac{m(s)}{p(S)\tau_s}\sum_{S\in\mathbb{S}_{i,s}} v(S) = \frac{1}{\tau_s(n-s)}\sum_{S\in\mathbb{S}_{i,s}} v(S)$$

We can use $a^m_{i,s}(\tilde{v})$ to estimate $a^m_{i,s}(v)$, then use the estimated value to calculate estimated $a^m_i(\tilde{v})$, and from them calculate the estimated $\tilde{x}_i$.

Regarding the standard error computation, we can estimate $Var[a^m_{i,s}(\tilde{v})]$, and then taking advantage of the deterministic linear combinations:

$$Var[a^m_i(\tilde{v})] = \sum_s Var[a^m_{i,s}(\tilde{v})] \tag{29}$$

and then:

$$Var[\tilde{x}_i] = \left(\frac{n-1}{n}\right)^2 Var[a^m_i(\tilde{v})] + \left(\frac{1}{n}\right)^2 \sum_{j\neq i} Var[a^m_j(\tilde{v})]. \tag{30}$$

We compare later the two stochastic approximations, that is, the one with weighted and the one with stratified sampling, letting them evaluate the same sample size τ in different important classes of games.

In the following, we present our application to four classes of games. In the first two cases that we analyze no approximation is needed at all because the quantities of interest can be easily calculated exactly. Anyway, we will use the procedure described in Section 3 to illustrate how our approximations work in these simple cases. In the last example, we will concentrate only in the Shapley and the least square value. There, we apply our methodology to a case in which it is harder (and almost impossible for moderate or large n) to evaluate the required value. Nevertheless, we will be able to use our approach to estimate the Shapley value of large size Voting and Linear Production games.

To better illustrate the possibilities and flexibility of our stochastic approximations we have followed different sampling mechanisms in the applications. In the first two games, the Shapley value has been estimated using the original expression (2) and thus, we have applied WSS1. In the last two games, the estimation has been done using the Least squares formulas. Hence in this cases we have resorted to WSS2 and SSS to compare experimentally their performance.

Note that the computation of most values in cooperative game theory involve, in general, the calculation of 2^{n-1} terms. With our approximation, we only need τ terms. As a measure of the gain thus obtained, we propose the sampling ratio, $f = \tau/2^{n-1}$, which is also reported in the following examples.

4.3. The Gloves game

Suppose a game (N, v) with set of players $N = \{1, \dots, n_1,\ n_1 + 1, \dots, n_1 + n_2\}$, where the first n_1 elements represent left hand gloves and the remaining n_2 are right hand gloves. Let $n = n_1 + n_2$. For any subset $S \subseteq N$, we consider $v(S) = \min\{|S_{\text{left}}|, |S_{\text{right}}|\}$, where

$$S_{\text{left}} = S \cap \{1, \dots, n_1\} \text{ and } S_{\text{right}} = S \cap \{n_1 + 1, \dots, n_1 + n_2\}.$$

It can be checked that, for $i = 1, \dots, n_1$ the mean value defined in (12) is

$$\alpha_i(v) = \sum_{k=0}^{\min\{n_1, n_2\}} k p_k$$

where

$$p_k = \frac{1}{2^{n-1}} \left\{ \binom{n_2}{k} \sum_{j=k+1}^{n_1} \binom{n_1 - 1}{j - 1} + \binom{n_1 - 1}{k - 1} \sum_{j=k+1}^{n_2} \binom{n_2}{j} + \binom{n_1 - 1}{k - 1} \binom{n_2}{k} \right\}, \tag{31}$$

for $k = 0, \dots, \min\{n_1, n_2\}$. (A similar formula for $\alpha_i(v)$ with $i = n_1 + 1, \dots, n$ can be obtained by interchanging n_1 and n_2 in (31)).

Under the uniform sampling, in the set of coalitions containing the player i, the approximations given in (13) and (14) result in

$$\widetilde{\alpha}_i = \frac{1}{\tau} \sum_{k=0}^{\min\{n_1, n_2\}} k W_k \quad \text{and} \quad \widetilde{\sigma_i^2} = \frac{1}{\tau} \sum_{k=0}^{\min\{n_1, n_2\}} k^2 W_k - \widetilde{\alpha}_i^2,$$

with W_k =*number of subsets with* $\min\{|S_{\text{left}}|, |S_{\text{right}}|\} = k$ *among the* τ *selected.* Note that $\boldsymbol{W} = \left(W_0, \dots, W_{\min\{n_1, n_2\}}\right)$ is a random vector following a multinomial $\mathcal{M}\left(\tau, p_1, \dots, p_{\min\{n_1, n_2\}}\right)$ distribution. . .

For instance, in the case $n_1 = 30$ and $n_2 = 15$ we illustrate our approximation with a sample size $\tau = 1000$ for the mean value of players $i = 1$ and $i = 31$. The results are presented in the following table

Player	Exact (α_i)	$\widetilde{\alpha}_i$	$\widetilde{\sigma_i^2}$	95%-confidence interval
1	7.4925	7.482	3.6957	(7.4782, 7.4858)
31	7.9812	7.949	3.1444	(7.9455, 7.9525)

These results show that the estimation is rather accurate even for a very small sampling fraction $f = 1000/2^{44} = 5,68 \times e - 11$.

The Shapley value

We also consider the Shapley value of this game. For $i = 1, \dots, n_1$, and S such that $i \notin S$, we consider

$$v_{i,*}(S) = v(S \cup \{i\}) - v(S) = \begin{cases} 1, & |S_{\text{left}}| < |S_{\text{right}}| \\ 0, & |S_{\text{left}}| \geq |S_{\text{right}}| \end{cases}$$

and the number of coalitions of size $s = |S| = \{1, \ldots, n-1\}$ with $i \notin S$ and $|S_{\text{left}}| < |S_{\text{right}}|$ is $\sum_{t<s/2} \binom{n_1-1}{t}\binom{n_2}{s-t}$, then the corresponding Shapley value is

$$\phi_i(v) = \sum_{s=1}^{n-1} \frac{s!(n-s-1)!}{n!} \sum_{t<s/2} \binom{n_1-1}{t}\binom{n_2}{s-t}.$$

Under the sampling strategy WSS1 described in (15), the approximations given in (17) are

$$\widetilde{\phi}_i = \frac{1}{\tau} U_i \quad \text{and} \quad \widetilde{\sigma_i^2} = \widetilde{\phi}_i(1-\widetilde{\phi}_i),$$

with $U_i =$*number of coalitions sampled for which* $|S_{left}| < |S_{right}|$ (recall that the sampling is performed in the set of coalitions not containing player i). We have that U_i is a binomial random variable, namely $U_i \sim Binomial(\tau, p_i)$, where $p_i = \frac{1}{n}\sum_{s=1}^{n-1}\sum_{t<s/2} \frac{\binom{n_1-1}{t}\binom{n_2}{s-t}}{\binom{n-1}{s}}$. Again, the properties stated in Theorem 2 can be checked directly and for instance, we have

$$E\widetilde{\phi}_i = p_i = \phi_i \quad \text{and} \quad Var(\widetilde{\phi}_i) = \frac{1}{\tau}\phi_i(1-\phi_i).$$

As a numerical example, suppose again that $n_1 = 30$ and $n_2 = 15$. Firstly, we run a pilot sample of size 500 and we obtained $\widetilde{\sigma_1^2} = 0.0529$. This approximation to the variance was used in formula (7) along with the values $\epsilon = 0.01$ and $\beta = 0.05$, so that we obtain a sample size $\tau = 2033$ ($f = 2033/2^{44} = 1.15e-10$) and for this sample size we report the results presented in the following table

Player	Exact (ϕ_i)	$\widetilde{\phi}_i$	$\widetilde{\sigma_i^2}$	95%-confidence interval
1	0.04475	0.0497	0.0472	(0.04949, 0.04991)
31	0.91050	0.9134	0.0787	(0.91363, 0.91417)

As compared with the previous estimates for the mean value, doubling the sample size, we get an improvement of the results of the standard errors of one order of magnitude.

4.4. The Airport game

Let (N, v) be an airport game, with $N = \{1, \ldots, n\}$, see [11]. The characteristic function is defined as follows. Consider integers $0 = n_0 < n_1 < \cdots < n_k = n$ and $0 = c_0 < c_1 < \cdots < c_k$. For $i \in N$, define $c(i) = c_r$, if $i \in A_r; = \{n_{r-1}+1, \ldots, n_r\}$. For $S \subset N$ the characteristic function is $v(S) = \max\{c(i) : i \in S\}$.

Elementary combinatorial arguments show that

$$\#\{S \subset N : i \in S \text{ and } v(S) = c(j)\} = \begin{cases} 2^{n_r - 1}, & \text{if } j = r \\ \\ 2^{n_j-1} - 2^{n_{j-1}-1} -, & \text{if } j > r \end{cases}$$

then, the mean value, (12), is

$$\alpha_i(v) = \sum_{j=r}^{k} c(j) p_j^{(i)}, \tag{32}$$

with

$$p_j^{(i)} = \begin{cases} \dfrac{1}{2^{n-n_r}}, & \text{if } j = r \\[2ex] \dfrac{1}{2^{n-n_j}} - \dfrac{1}{2^{n-n_{j-1}}}, & \text{if } j > r. \end{cases}$$

Sampling τ coalitions with uniform probability in the set of coalitions containing the i-player, and using the approximations given in (13) and (14), we obtain

$$\widetilde{\alpha}_i = \frac{1}{\tau} \sum_{j=r}^{k} c(j) W_j^{(i)} \quad \text{and} \quad \widetilde{\sigma_i^2} = \frac{1}{\tau} \sum_{j=r}^{k} c^2(j) W_j^{(i)} - \widetilde{\alpha}_i^2$$

with $W_j^{(i)}$ =*number of selected coalitions with* $v(S) = c(j)$ *among the* τ *selected.* Note that $\boldsymbol{W}^{(i)} = \left(W_r^{(i)}, \ldots, W_k^{(i)}\right) \sim \mathcal{M}\left(\tau; p_r^{(i)}, \ldots, p_k^{(i)}\right)$. It is well known, that under multinomial sampling the proportion $W_j^{(i)}/\tau$ is unbiased of its expected value, $p_j^{(i)}$, then $\widetilde{\alpha}_i$ is unbiased for (32), in agreement with Theorem 2(a).

As a numerical example, we consider $n = 45$ players, $c_r = r$, for $r = 1, \ldots, 10$ and n_r given in the following table:

n_1	n_2	n_3	n_4	n_5	n_6	n_7	n_8	n_9	n_{10}
10	20	28	34	38	40	42	43	44	45

We run a training sample of size $\tau_0 = 500$, obtaining an estimate of the variance of the mean value of the first player, $\widetilde{\sigma_1^2} = 1.1184$. This value along with $\epsilon = 0.05$ and $\beta = 0.05$ were used to determine the sample size $\tau = 14036$ ($f = 7.98e - 10$), see (7). For this sample size we obtain the results shown in the next table

Player	Exact (α_i)	$\widetilde{\alpha}_i$	$\widetilde{\sigma_i^2}$	95%-confidence interval
1	9.0854	9.0968	1.3322	$(9.0966, 9.0969)$
11	9.0854	9.0649	1.3528	$(9.0647, 9.0651)$
21	9.0854	9.0754	1.3277	$(9.0753, 9.0756)$
29	9.0854	9.0948	1.3251	$(9.0947, 9.095)$
35	9.0859	9.0726	1.352	$(9.0724, 9.0728)$
39	9.0938	9.089	1.2751	$(9.0888, 9.0891)$
41	9.125	9.1225	1.1196	$(9.1224, 9.1227)$
43	9.25	9.239	0.6977	$(9.2389, 9.2391)$
44	9.5	9.497	0.25	$(9.4969, 9.4971)$
45	10.0	10.0	0.0	$(10.0, 10.0)$

Estimation of the Shapley value of the player $i \in A_r = \{n_{r-1}+1, \ldots, n_r\}$.

Given a coalition $S \subset N \setminus \{i\}$, let us define $\ell(S) = \max\{j : S \cap A_j \neq \emptyset\}$ (for convenience, $\max \emptyset = 0$, so that $\ell(\emptyset) = 0$). Then,

$$v_{i,*}(S) = v(S \cup \{i\}) - v(S) = \begin{cases} 0, & \text{if } \ell(S) \geq r \\ \\ c(r) - c(\ell(S)), & \text{if } \ell(S) < r \end{cases}$$

and for $0 \leq j < r$

$$\#\{S : S \subset N \setminus \{i\}, |S| = s \text{ and } \ell(S) = j\} = \binom{n_j}{s} - \binom{n_{j-1}}{s}, \; s \geq 0,$$

(for convenience, $n_{-1} = -1$) then, after some algebra, the exact Shapley value is

$$\phi_i := \phi_i(v) = \sum_{S \subseteq N - \{i\}} \frac{s!\,(n-s-1)!}{n!} v_{i,*}(S) = \sum_{s=0}^{n-1} \frac{s!\,(n-s-1)!}{n!} \sum_{j=0}^{r-1} \sum_{\substack{S \subset N - \{i\} \\ |S| = s \\ \ell(S) = j}} (c(r) - c(j))$$

$$= \sum_{s=0}^{n-1} \frac{s!\,(n-s-1)!}{n!} \sum_{j=0}^{r-1} (c(r) - c(j)) \left\{ \binom{n_j}{s} - \binom{n_{j-1}}{s} \right\}$$

$$= \sum_{s=0}^{n-1} \frac{s!\,(n-s-1)!}{n!} \sum_{j=0}^{r-1} \binom{n_j}{s} (c(j+1) - c(j)).$$

For the estimation of ϕ_i, for $i \in A_r$, we can take advantage of the fact that $v_{i,*}(S) = 0$ if $\ell(S) \geq r$, so that instead of the sampling scheme given in (15), we propose the following probability on the set of coalitions not containing player i,

$$p(S) = \begin{cases} \dfrac{1}{(n_{r-1}+1)\binom{n_{r-1}}{s}}, & \text{if } \ell(S) \leq r-1 \text{ and } |S| = s \in \{0, \ldots, n_{r-1}\} \\ \\ 0, & \text{otherwise} \end{cases} \tag{33}$$

which can be seen as a two stages sampling such that in the first stage the size of the coalition (not containig player i), s, is selected with equal probability among $\{0, \ldots, n_{r-1}\}$ and in the second stage a coalition is selected with uniform probability within the set of coalitions of size s with $\ell(S) \leq r-1$. Define the random variables $W_{s,j}$ =*number of coalitions with $\ell(S) = j$ and $|S| = s$ among the τ selected.* Then, the approximations given in (5) and (6) yield, respectively

$$\widetilde{\phi}_i = \frac{(n_{r-1}+1)!}{\tau\, n!} \sum_{s=0}^{n_{r-1}} \frac{(n-s-1)!}{(n_{r-1}-s)!} \sum_{j=0}^{r-1} (c(r) - c(j))\, W_{s,j},$$

$$\widetilde{\sigma_i^2} = \frac{1}{\tau} \left(\frac{(n_{r-1}+1)!}{n!} \right)^2 \sum_{s=0}^{n_{r-1}} \left(\frac{(n-s-1)!}{(n_{r-1}-s)!} \right)^2 \sum_{j=0}^{r-1} (c(r) - c(j))^2\, W_{s,j} - \widetilde{\phi}_i^2.$$

Similarly as in the previous example, we run a training sample of size $\tau_0 = 1000$ and obtained $\widetilde{\sigma^2_{45}} = 2.341$. For $\beta = 0.01$ and $\epsilon = 0.01$, we determined the sample size $\tau = 278581$ ($f = 1.58e - 8$). For this sample size we obtained the following results

Player	Exact (ϕ_i)	$\widetilde{\phi_i}$	$\widetilde{\sigma_i^2}$	99%-confidence interval
1	0.0222	0.0222	0.0	$(0.0222, 0.0222)$
11	0.0508	0.0509	0.0195	$(0.0509, 0.0509)$
21	0.0908	0.0915	0.0918	$(0.0915, 0.0915)$
29	0.1496	0.1492	0.244	$(0.1492, 0.1493)$
35	0.2405	0.2402	0.5169	$(0.2402, 0.2402)$
39	0.3834	0.3842	0.952	$(0.3841, 0.3842)$
41	0.5834	0.5827	1.5167	$(0.5826, 0.5827)$
43	0.9167	0.918	2.4164	$(0.918, 0.918)$
44	1.4167	1.4116	3.3844	$(1.4116, 1.4116)$
45	2.4167	2.4142	4.2718	$(2.4142, 2.4142)$

4.5. The Simple Voting Game

The Simple Voting Game is defined as $G := (q, w_i, i = 1, \ldots, n)$ where q is the quota and w_i is the weight given to the $i - th$ voter (player). The implementation of WSS2 follows from the straightforward calculation of (19). When we apply stratified sampling (SSS), it can be readily seen that we have two indexes, k_1 and k_2 such that:

- For all S such that $|S| < k_1$, $v(S) = 0$;
- For all S such that $|S| > k_2$, $v(S) = 1$.

Therefore, since $v(S) = 1$ for all $|S| > k_2$, we can replace estimated values $a^m_{i,s}(\tilde{v})$ with exact values $a^m_{i,s}(v) = 1$ to obtain:

$$a^m_i(\tilde{v}) = (n - 1 - k_2)/(n - 1) + \sum_{s=k_1}^{k_2} a^m_{i,s}(\tilde{v})$$

Furthermore, estimation can be improved observing that the Shapley value of different players should be equal whenever the associated voting weights are the same. So one can improve the estimated Shapley value $\tilde{\phi}$ projecting it on the feasible set x by solving the quadratic program:

$$\min_{x \in \mathbb{R}^n} \|x - \phi_i(\tilde{v})\| : \sum_{i=1}^{n} x_i = 1 \text{ and } x_i = x_j \text{ whenever } w_i = w_j.$$

The first experiment considers a small size game with 17 players. Let

$$G = \{\text{quota} = 45;\ w = [11, 11, 9, 9, 8, 8, 5, 5, 4, 4, 3, 3, 3, 1, 1, 1, 1]\},$$

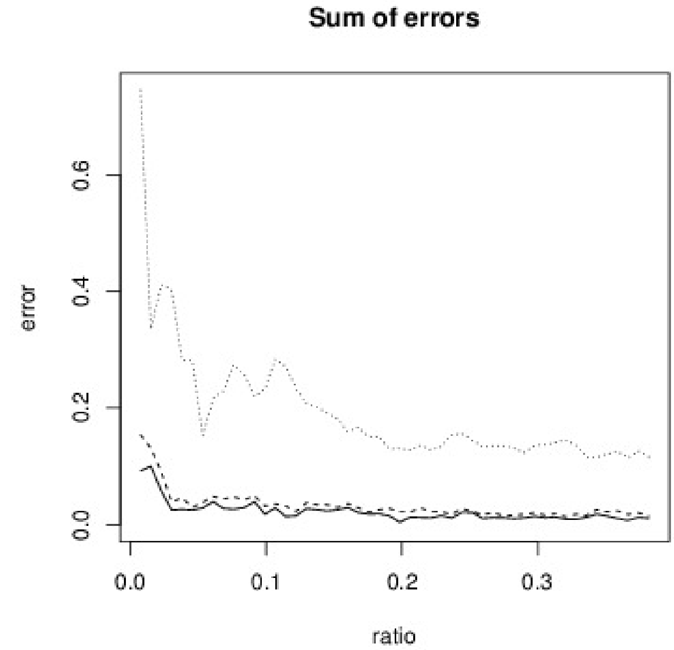

Figure 1: Sum of approximation errors with respect to v evaluation in the Simple Voting Game (expressed as ratio). Dotted line: Weighted sampling, Broken line: Stratified sampling, Continuous line: Stratified sampling and quadratic correction.

here we can calculate the exact Shapley value ϕ_i for all $i = 1, \ldots, n$, and then we can compare it to the approximations given the two sampling schemes. Their performance is compared using as an approximation index the overall summation of absolute differences among true and estimated values:

$$\sum_{i=1}^{n} |\phi_i - \phi_i(\tilde{v})|.$$

Controlling for f, the results are reported in Fig 1. It can be seen that the approximation error of the stratified sampling SSS is much better than the weighted sampling WSS2. Moreover, the improvement through the quadratic correction is substantial. This can also be viewed in Figure 2, in which Figure 1 is zoomed to appreciate the quadratic improvement. On average, after having used the quadratic projection, the approximation error decreased more than 30%. This behavior motivates that in our tables we only report data from the SSS estimation since they are much better that those corresponding to WSS2.

In Table 1, the computational results of the approximation calculated with $\tau = 5000$ ($f = 0.076$) are reported. Column exact is the real Shapley value, column SSS reports the Stratified Sampling approximation, column SSS + QP reports the approximation when the quadratic correction is applied and column SD reports the standard error of SSS. There it can be seen how close the approximations are to the actual Shapley value, to the point that the resulting standard error are sufficiently small to detect the correct power ranking of players.

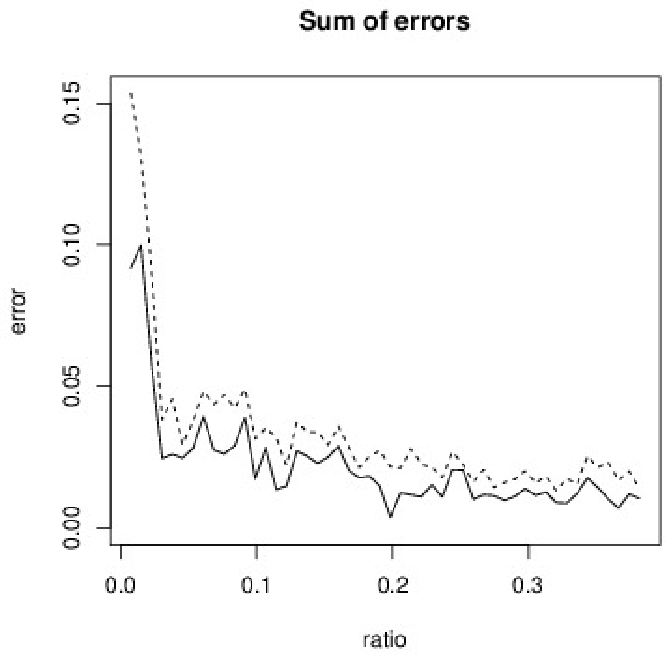

Figure 2: Sum of approximation errors with respect to v evaluation in the small Simple voting game (expressed as ratio). Zooming Figure 1. Broken line: Weighted sampling, Continuous line: Stratified sampling and quadratic correction.

Then we repeat the experiment to the large size voting game G, with $n = 40$ players. Game players' weight are described in the second column of Table 2, the quota is 102. In this case we cannot calculate the exact Shapley value and the approximation quality is evaluated through the standard error (SD). All the results are reported in Table 2. We run the stochastic approximation with stratified sampling SSS and $\tau = 5000$, which implies a sampling fraction $f = 9.09e-09$. As it can be seen, the standard error is so small that the approximation seems a reliable estimate of the true and *unknown* Shapley value.

4.6. The Linear Production Game

Linear production games were introduced by Owen [12]. In these games each agent $i \in N$ owns a resource bundle $b_i \in \mathbb{R}^q_+$. The resources can be used to produce m types of goods according to some technological matrix $A \in \mathbb{R}^{q \times m}$ which can be sold at prices $p_1, \ldots, p_m$. When a coalition S is formed its members pool their resources so as to maximize the market value of their products. This class of games has become very popular and important among game theoretician by its implications in the cost sharing field and its equivalence to the class of all balanced games. In spite of that, it is well-known the difficulty to compute values of these games.

player	weight	Exact	SSS	SSS+QP	SD	95% interval
1	11	0.1323	0.1329	0.1313	0.0057	(0.1218 , 0.1440)
2	11	0.1323	0.1298	0.1313	0.0056	(0.1187 , 0.1408)
3	9	0.1052	0.1038	0.1074	0.0060	(0.0921 , 0.1156)
4	9	0.1052	0.1109	0.1074	0.0058	(0.0995 , 0.1223)
5	8	0.0924	0.0918	0.0957	0.0058	(0.0806 , 0.1031)
6	8	0.0924	0.0996	0.0957	0.0057	(0.0884 , 0.1108)
7	5	0.0557	0.0569	0.0557	0.0059	(0.0454 , 0.0683)
8	5	0.0557	0.0544	0.0557	0.0059	(0.0429 , 0.0660)
9	4	0.0440	0.0384	0.0406	0.0059	(0.0268 , 0.0500)
10	4	0.0440	0.0428	0.0406	0.0059	(0.0312 , 0.0544)
11	3	0.0327	0.0335	0.0334	0.0058	(0.0221 , 0.0448)
12	3	0.0327	0.0371	0.0334	0.0058	(0.0258 , 0.0484)
13	3	0.0327	0.0295	0.0334	0.0058	(0.0180 , 0.0409)
14	1	0.0107	0.0055	0.0097	0.0059	(-0.0060 , 0.0170)
15	1	0.0107	0.0090	0.0097	0.0057	(-0.0023 , 0.0202)
16	1	0.0107	0.0125	0.0097	0.0057	(0.0013 , 0.0237)
17	1	0.0107	0.0117	0.0097	0.0058	(0.0004 , 0.0230)

Table 1: Exact and approximate value of the Shapley Value to the Small Simple Voting Game $n = 17$ and $\tau = 5000$.

The characteristic function is given by:

$$v(S) = \max\big\{ \sum_{j=1}^{m} p_j x_j : \; Ax \leq \sum_{i \in S} b_i, \; x = (x_1, \ldots, x_m) \geq 0 \big\}.$$

Our goal in this section is to give a procedure to obtain good estimates of the Shapley value of this class of games of any size.

The first experiment considers $n = 15$ players, $m = 12$ products, $q = 5$ production constraints. All data for A, B, p are random (0-1)-uniform values. The implementation of weighted sampling WSS2 follows from the straightforward calculation of (19). When we apply stratified sampling (SSS), it can be readily seen that coefficients $a^m_{i,s}(v)$ are the sum of $\binom{n}{s}$ terms, so that, for small or large values of s, $a^m_{i,s}(v)$ can be conveniently calculated rather than estimated. In our implementation, this is done for $s = 2$ and for $s = n-1$. This implies some gain in the final precision.

For $n = 15$ it is still possible to compute the exact Shapley value of this game since the number of linear programs that have to be solved, for the evaluation of each ϕ terms, is bounded above by 2^{12}. We run the weighted and stratified sampling, WSS2 and SSS, respectively, with different sampling fraction f and we calculate the sum of the absolute value of the errors (the difference between the actual and the estimated Shapley value). The comparison is reported in Figure 3, where it can be seen that the stratified sampling SSS strongly outperforms WSS2, as observed in the voting game.

In Figure 4 the same figure is zoomed to observe how the stratified sampling error decreases.

It can be seen that for sampling fraction values greater than 20% the gains of further samples are modest as the approximation is quite close to the real Shapley value.

In Table 3 we report the comparison between the exact and the approximate Shapley value for $\tau = 10000$, considering interval estimate too. As it can be seen, the approximation error is one order of magnitude smaller than the value.

The quality of the approximation depends on the standard error, that in turn depends on the sample size τ and not on the population size 2^n. In the following example, we report the stratified sampling approximation, SSS, for a large production game in which this property is more evident.

Random data are generated as before, but the number of players is now $n = 40$. We run two simulations, one with $\tau = 10000$ and another one with $\tau = 360000$, giving rise to sample sampling fractions, respectively, of $f = 1.82e - 08$ and $f = 6.55e - 07$. Running the algorithm in R with the freeware (non-commercial) linear solver lpSolve, it results in one minute of computation for $\tau = 360000$ and negligible times for $\tau = 10000$. Results are reported in Table 4. When we are interested to determine the Shapley value, we can see that $\tau = 10000$ is sufficient. Here, the Shapley values are of unitary order, while the standard errors are of order 10^{-2}, namely two orders of magnitude lower.

If we were interested in determining the players' ranking with respect to the Shapley value, we can see that the difference between players' Shapley values is of order 10^{-2}. Therefore we must increase τ to decrease the standard error. As it can be seen, with $\tau = 360,000$, the standard error is of order 10^{-3}, small enough to determine estimation intervals that do not overlap with each other so that these estimates can clearly determine meaningful players' ranking.

5. Conclusions

We proposed a methodology to calculate values of cooperative games that relies on the concept of stochastic approximation.

The idea is to replace the exponential number of terms of the value formula summation with just a sample of them. Applying probability concepts to the sample and to the reduced sum, we can prove that the expectation of the estimate is the actual value of the game and that its standard error keeps under control the difference between the estimate and the actual value. We proved the viability of our approach calculating values of different games with different sampling strategies, and we found that the sample size needs not to be large to obtain good approximations. Successful applications depend on the sampling strategy and the game characteristic function. Therefore, future research should consider different sampling strategies, such as different stratifications, and different classes of games, as for instance network games. Moreover, while the theoretical properties of many cooperative games are often well-understood, we are somewhat missing their economic applications. To date, one reason to this could have been the time complexity to calculate values, which has prevented the analysis of games whose number of players is more than minimal. We hope that the stochastic approximation will be the technique to foster new empiric analysis.

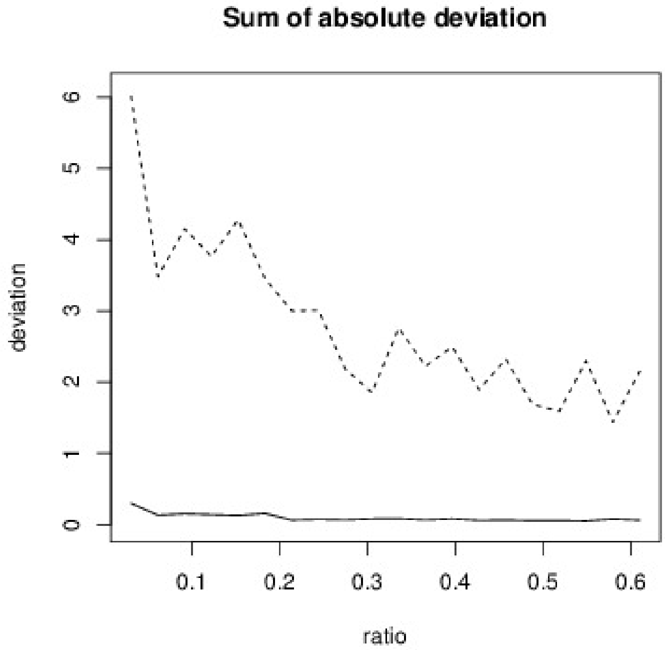

Figure 3: Sum of approximation errors with respect to v evaluation (expressed as ratio) in the small Linear production game $n = 12$: Broken line: Weighted sampling, Continuous line: Stratified sampling.

Acknowledgements

This research has been partially financed by the Spanish Ministry of Science and Technology projects FQM-331, MTM2016-74983-C02-01.

References

[1] Bilbao J.M., Fernández J.R., Jiménez A., López J.J. (2000) "Generating functions for computing power indices efficiently". TOP 8:191-213.

[2] Castro J., Gómez D., Tejada J. (2009) "Polynomial calculation of the Shapley value based on sampling". Computer and Operations Research 36:1726-1730.

[3] Dragan I. (2006) "The Least Squares Value and the Shapley Value for Cooperative TU Games". TOP 4:61-73.

[4] Dubey P., Neyman A., Weber J. (1981) "Value Theory without Efficiency". Mathematics of Operations Research 4:99-131.

[5] Fernández F.R., Mayor J.A., Puerto J, Zafra M.J.(2004) "Sampling theory and Shapley value". Proceeding of the VI Spanish Meeting of Game Theory. Elche.

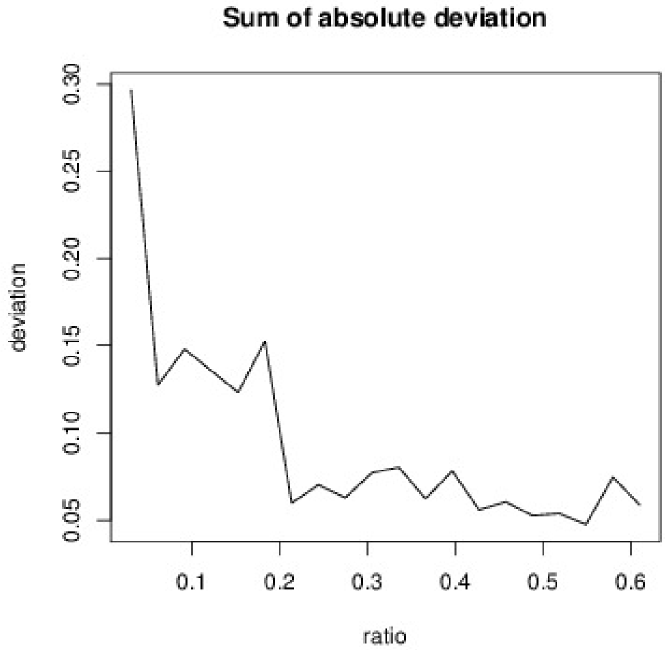

Figure 4: Sum of approximation errors with respect to v evaluation (expressed as ratio) in the small Linear production game $n = 12$: Continuous line: Stratified sampling.

[6] Fernández F.R., Puerto J. (2005) "Sampling Theory and Cooperative Games". Worhshop on Stochastic Method in Game Theory, Centro Mejorana. Erice.

[7] Fernández F.R., Puerto J. (2006) *Teoria de Juegos Multiobjetivos*. Imagraf. Malaga.

[8] Fernández F. R., Puerto J. and Zafra M. J., (2002) "Cores of Stochastic cooperative games", International Game Theory Review 4: 265-280.

[9] Granot, D. (1977) "Cooperative games in stochastic function form", Management Science 23: 621-630.

[10] Keinan A., Sandbank B., Hilgetag C.C., Meilijson I., Ruppin E. (2006) "Axiomatic Scalable Neurocontroller Analysis via the Shapley Value". Artificial Life 12:333-352.

[11] Littlechild S., Owen G. (1973) "A simple expression for the Shapley value in a special case". Management Science 20:370-372.

[12] Owen G. (1995) *Game Theory*. Academic Press.California.

[13] Puerto J., Fernndez F.R., Hinojosa Y. (2008) "Partially ordered cooperative games: extended core and Shapley value". Annals of Operations Research, 158:143-159, 2008.

[14] Ruiz L.M., Valenciano F., Zarzuelo J.M. (1998) "The Family of Least Squares Values for Transferable Utility Games". Games and Economic Behavior 24:109-130.

[15] Timmer J., (2006) "The compromise value for cooperative games with random payoffs", Mathematical Methods of Operations Research, 64:95-106.

[16] Timmer J., Borm P., and Tijs S., (2003) "On three Shapley-like solutions for cooperative games with random payoffs", International Journal of Game Theory 32: 595-613.

[17] Timmer J., Borm P., and Tijs, S. (2005) "Convexity in stochastic cooperative situations", International Game Theory Review 7: 25-42.

player	weight	SS	SS+QP	SD
1	12	0.05529	0.05830	0.00478
2	12	0.06072	0.05830	0.00479
3	12	0.05796	0.05830	0.00492
4	12	0.05535	0.05830	0.00505
5	12	0.06218	0.05830	0.00488
6	8	0.03706	0.03984	0.00503
7	8	0.04310	0.03984	0.00484
8	8	0.03919	0.03984	0.00508
9	8	0.03732	0.03984	0.00488
10	8	0.04124	0.03984	0.00500
11	8	0.04045	0.03984	0.00496
12	8	0.04203	0.03984	0.00491
13	8	0.03881	0.03984	0.00512
14	8	0.03933	0.03984	0.00512
15	5	0.02843	0.02502	0.00495
16	5	0.02278	0.02502	0.00505
17	5	0.02225	0.02502	0.00505
18	5	0.01889	0.02502	0.00504
19	5	0.02971	0.02502	0.00508
20	5	0.02466	0.02502	0.00498
21	5	0.02349	0.02502	0.00490
22	5	0.02843	0.02502	0.00493
23	5	0.02395	0.02502	0.00498
24	5	0.02988	0.02502	0.00519
25	5	0.02281	0.02502	0.00505
26	1	0.00594	0.00498	0.00499
27	1	0.00648	0.00498	0.00503
28	1	0.00753	0.00498	0.00503
29	1	0.00098	0.00498	0.00492
30	1	0.00766	0.00498	0.00495
31	1	0.00109	0.00498	0.00492
32	1	0.00444	0.00498	0.00501
33	1	0.00206	0.00498	0.00498
34	1	0.00115	0.00498	0.00479
35	1	0.00460	0.00498	0.00488
36	1	0.00757	0.00498	0.00509
37	1	0.00770	0.00498	0.00488
38	1	0.00384	0.00498	0.00505
39	1	0.00876	0.00498	0.00499
40	1	0.00492	0.00498	0.00501

Table 2: Approximate value of the Shapley Value to the Large Voting Game

Player	Shapley (exact)	Shapley (est.)	Standard Error	95% Interval
1	0.65720	0.65480	0.01171	(0.63186, 0.67774)
2	0.76555	0.76888	0.01178	(0.74579, 0.79197)
3	0.89288	0.88795	0.01112	(0.86616, 0.90975)
4	0.65340	0.65763	0.01068	(0.63670, 0.67857)
5	0.77422	0.77437	0.01105	(0.75271, 0.79603)
6	0.55946	0.55941	0.01145	(0.53696, 0.58186)
7	1.26183	1.26547	0.00872	(1.24837, 1.28256)
8	1.20149	1.20770	0.01034	(1.18743, 1.22797)
9	0.79972	0.79589	0.01016	(0.77598, 0.81580)
10	1.02142	1.01826	0.01107	(0.99656, 1.03995)
11	0.83565	0.84304	0.01137	(0.82075, 0.86532)
12	0.98868	0.99305	0.01027	(0.97293, 1.01317)
13	0.88628	0.88933	0.01139	(0.86700, 0.91166)
14	0.66389	0.65436	0.01056	(0.63367, 0.67506)
15	0.82248	0.81404	0.01136	(0.79178, 0.83630)

Table 3: Exact and Approximate value of the Shapley Value to the Small Linear Production Game with SSS and $\tau = 10000$

player	Shapley-$\tau 1$	Standard Error-$\tau 1$	Shapley-$\tau 2$	Standard Error-$\tau 2$
1	0.7483	0.0357	0.73390	0.00444
2	0.6847	0.0356	0.70440	0.00444
3	1.0573	0.0347	1.05924	0.00433
4	0.7285	0.0353	0.72841	0.00439
5	0.7155	0.0356	0.70573	0.00442
6	0.6124	0.0358	0.61362	0.00441
7	1.1820	0.0348	1.21112	0.00433
8	1.2689	0.0345	1.26939	0.00430
9	0.7121	0.0356	0.70259	0.00438
10	0.8738	0.0356	0.87206	0.00443
11	0.9221	0.0352	0.93731	0.00439
12	0.8161	0.0356	0.80896	0.00442
13	0.8658	0.0356	0.87550	0.00442
14	0.4787	0.0343	0.44231	0.00432
15	0.9798	0.0355	0.97659	0.00438
16	0.3927	0.0346	0.38779	0.00428
17	0.3125	0.0339	0.28079	0.00420
18	1.0755	0.0355	1.08252	0.00437
19	0.9878	0.0353	1.02532	0.00438
20	1.2855	0.0350	1.28415	0.00436
21	0.8044	0.0360	0.80236	0.00444
22	0.5266	0.0345	0.51503	0.00434
23	1.1500	0.0350	1.14763	0.00435
24	0.6329	0.0358	0.64333	0.00441
25	0.8860	0.0355	0.88368	0.00438
26	1.1240	0.0347	1.14455	0.00434
27	1.2671	0.0353	1.27387	0.00432
28	0.9598	0.0357	0.95851	0.00445
29	1.0255	0.0354	1.03748	0.00442
30	1.1549	0.0347	1.17000	0.00433
31	0.9550	0.0358	0.97016	0.00443
32	1.0790	0.0357	1.08852	0.00441
33	1.1030	0.0351	1.07281	0.00440
34	1.1603	0.0353	1.16440	0.00436
35	0.2951	0.0339	0.27580	0.00421
36	1.4138	0.0340	1.40480	0.00426
37	0.5357	0.0353	0.55060	0.00438
38	0.4943	0.0346	0.47902	0.00432
39	0.5537	0.0356	0.56209	0.00441
40	1.1802	0.0348	1.15496	0.00435

Table 4: Approximate value of the Shapley Value for the Large Linear Production Game with SSS and $\tau = 10000$ (Shapley-$\tau 1$) and $\tau 2 = 360000$ (Shapley-$\tau 2$)