J. Appl. Prob. **58,** 609–620 (2021)
*doi:*10.1017/jpr.2020.106

ON THE COMPARISON OF SHAPLEY VALUES FOR VARIANCE AND STANDARD DEVIATION GAMES

MARCELLO GALEOTTI,* *University of Florence*
GIOVANNI RABITTI,** *Heriot-Watt University*

Abstract

Motivated by the problem of variance allocation for the sum of dependent random variables, Colini-Baldeschi, Scarsini and Vaccari (2018) recently introduced Shapley values for variance and standard deviation games. These Shapley values constitute a criterion satisfying nice properties useful for allocating the variance and the standard deviation of the sum of dependent random variables. However, since Shapley values are in general computationally demanding, Colini-Baldeschi, Scarsini and Vaccari also formulated a conjecture about the relation of the Shapley values of two games, which they proved for the case of two dependent random variables. In this work we prove that their conjecture holds true in the case of an arbitrary number of independent random variables but, at the same time, we provide counterexamples to the conjecture for the case of three dependent random variables.

Keywords: Covariance matrix; sum of random variables; vector majorization; cooperative games

2010 Mathematics Subject Classification: Primary 91A12
Secondary 62J10

1. Shapley value

The Shapley value is a value attribution method that originated from economic game theory [12]. The intuition is that, considering a team of players together producing a value, the Shapley value attributes this value to individual members of the team. The Shapley value has been successfully applied in many probabilistic and statistical problems, such as collinear regression [6], reliability [8], queuing theory [2], uncertainty quantification [11], inventory [10], multivariate risk analysis [1] and machine learning [7]. See [9] for a survey of applications. The intuition is to regard a set of random variables as the team of players and the statistical/probabilistic index of interest as the produced value.

Formally, consider a set of n players (random variables) $N = \{1, 2, \ldots, n\}$ and any possible subset $J \subseteq N$, called a coalition. The function $\nu : 2^N \mapsto \mathbb{R}$, with the condition that $\nu(\emptyset) = 0$, is called the characteristic function or the value function of the game. The characteristic function

Received 23 April 2020; revision received 6 November 2020.

* Postal address: Department of Statistics, Informatics and Applications, University of Florence, Via delle Pandette 9, 50127 Florence, Italy. Email address: marcello.galeotti@unifi.it
** Postal address: Department of Actuarial Mathematics and Statistics, Heriot-Watt University and Maxwell Institute for Mathematical Sciences, EH14 4AS, Edinburgh, UK. Email address: g.rabitti@hw.ac.uk

of a coalition J, $\nu(J)$, is the value produced by all players in the coalition, for any possible coalition. Then, the Shapley value $\phi_i(\nu)$ is a measure of the value of player i and is given by

$$\phi_i^\nu = \sum_{J \subseteq N \setminus \{i\}} \frac{(n-|J|-1)!\,|J|!}{n!} \left[\nu(J \cup \{i\}) - \nu(J)\right], \tag{1}$$

where $|J|$ is the cardinality of J. It can be recognized that the Shapley value for the ith player is based on the marginal increase in the value function $\nu(J \cup \{i\}) - \nu(J)$ when player i joins coalition J, averaged over all possible coalitions.

The Shapley value can be characterized by several interesting properties:

Efficiency: $\sum_{i=1}^n \phi_i(\nu) = \nu(N)$.

Symmetry: If $\nu(J \cup \{i\}) = \nu(J \cup \{j\})$ for all $J \subseteq N \setminus \{i, j\}$, then $\phi_i(\nu) = \phi_j(\nu)$.

Dummy player: If $\nu(J \cup \{i\}) = \nu(J)$ for all $J \subseteq N$, then $\phi_i(\nu) = 0$.

Linearity: If two value functions ν and μ have respective Shapley values $\phi(\nu)$ and $\phi(\mu)$, then the game with value $\alpha\nu + \beta\mu$ has Shapley value $\alpha\phi(\nu) + \beta\phi(\mu)$ for all $\alpha, \beta \in \mathbb{R}$.

In [12] it is proved that the Shapley value is the unique attribution method satisfying these four properties. Equivalently, the Shapley value (1) can be expressed as

$$\phi_i(\nu) = \frac{1}{n!} \sum_{\psi \in \mathcal{P}(N)} \left(\nu\left(P^\psi(i) \cup \{i\}\right) - \nu\left(P^\psi(i)\right)\right), \tag{2}$$

where $\mathcal{P}(N)$ is the set of all permutations of N and $P^\psi(i)$ is the set of players who precede i in the order determined by the permutation ψ. The representation of Shapley value in (2) based on permutations has been adopted for developing algorithms for the calculation of Shapley values [3,13].

2. Variance and standard deviation games

Colini-Baldeschi, Scarsini and Vaccari [5] recently used Shapley values as the allocation criterion for a portfolio problem.

Consider the random vector $X = (X_1, X_2, \ldots, X_n)$ with finite second moment, and the sum of the random variables $S = \sum_{i=1}^n X_i$. Then, [5] considers the Shapley values using the characteristic functions

$$\nu(J) = \mathrm{Var}(S_J), \qquad \lambda(J) = \sqrt{\mathrm{Var}(S_J)},$$

where the partial sums $S_J = \sum_{i \in J} X_i$ are defined for every $J \subseteq N$. The authors of [5] call ν a variance game and λ a standard deviation game, and characterize the Shapley values for a variance game.

Theorem 1. (Colini-Baldeschi, Scarsini and Vaccari [5].) *For a variance game ν, $\phi_i(\nu) =$ Cov (X_i, S).*

However, an analogous closed-form representation of the Shapley value is not available for a standard deviation game. Thus, comparing Shapley values for variance and for standard deviation games is very difficult. This led the authors of [5] to formulate a conjecture about their comparison, which we describe in the next section.

3. Conjecture and results

Given two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, the vector $\mathbf{x}$ is said to be majorized by $\mathbf{y}$ ($\mathbf{x} \preceq \mathbf{y}$) if $\sum_{i=k}^n x_{(i)} \le \sum_{i=k}^n y_{(i)}$ for all $k \in \{2, \dots, n\}$ and $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$, where $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$ is the increasing rearrangement of $\mathbf{x}$. Using this notion, the conjecture can be formulated as follows.

Conjecture 1. (Colini-Baldeschi, Scarsini and Vaccari [5].) *For any $n \times n$ covariance matrix Σ, if ν is the corresponding variance game and λ the corresponding standard deviation game, then*

$$\frac{1}{\lambda(N)}\Phi(\lambda) \preceq \frac{1}{\nu(N)}\Phi(\nu),$$

where Φ denotes the vector of the Shapley values.

The authors of [5] showed that the conjecture holds for $n = 2$ dependent random variables and verified it numerically for $n = 3, 4, 5$ independent random variables. We now prove in Section 3.1 that the conjecture holds true for an arbitrary number of independent random variables. However, we show, using two counterexamples in Section 3.2, that the conjecture is not valid for $n = 3$ dependent random variables. From a reviewer's report we learnt about [4], where it is shown how the conjecture in the case of independent random variables can be seen as an application of a far more general theorem, whose proof includes several lemmas and definitions. Although we expect those results to be sound, we think that our alternative proof of the conjecture for independent random variables, being direct and self-contained, can be interesting and useful.

3.1. Proof of the conjecture for independent random variables

Theorem 2. *Assume that the random variables are independent. Then, the conjecture holds for any n.*

Proof. Let ν be the variance game and λ the standard deviation game. We assume the random variables to be independent with variances $\sigma_1^2 \le \sigma_2^2 \le \cdots \le \sigma_n^2$, so that it is easily checked that $\phi_i(\nu) = \sigma_i^2$. Moreover, straightforward computations show that, when $i \le j$, $\phi_i(\lambda) \le \phi_j(\lambda)$ as well.

First of all, we observe that the conjecture holds for $n = 2$ and 3. In fact, the case $n = 2$ is trivial, so that we have to prove the conjecture when $n = 3$. Since the ordering of the vectors $\boldsymbol{\Phi}(\boldsymbol{\nu})$ and $\boldsymbol{\Phi}(\boldsymbol{\lambda})$ by increasing components is the same, and, moreover, we can assume, without loss of generality, that $\mathrm{Var}(X_1 + X_2 + X_3) = \mathrm{SD}\,(X_1 + X_2 + X_3) = 1$, the proof amounts to showing that

$$\phi_1(\lambda) \ge \phi_1(\nu), \qquad \phi_3(\lambda) \le \phi_3(\nu). \tag{3}$$

We start with the first inequality in (3). In fact, after straightforward computations, that inequality is seen to be equivalent to

$$F\,(a, b) = 2 - 2\sqrt{1-a} + 2\sqrt{a} + \sqrt{1-b} - \sqrt{1-a-b} + \sqrt{a+b} - \sqrt{b} - 6a \ge 0$$

in a suitable set of the (a, b) plane, where $\sigma_1^2 = a$, $\sigma_2^2 = b$, and $\sigma_3^2 = c$ satisfy $c = 1 - a - b \ge b \ge a \ge 0$. Namely, our first inequality must be verified in the triangle T with vertices in $(0, 0)$, $(0, \frac{1}{2})$, $(\frac{1}{3}, \frac{1}{3})$. In fact, it is easily computed that $F(0, b) = F(\frac{1}{3}, \frac{1}{3}) = 0$. Moreover, straightforward computations show that $\lim\limits_{a \to 0+} \frac{\partial F}{\partial a}(0, b) = +\infty$ when $b > 0$, while $\frac{\partial^2 F}{\partial a^2} \le -\frac{1}{2}\left[\left(\frac{1}{3}\right)^{-\frac{3}{2}} - \left(\frac{2}{3}\right)^{-\frac{3}{2}}\right] < 0$ throughout T. Hence, for any fixed $\overline{b} \in (0, \frac{1}{2})$, $F\left(a, \overline{b}\right)$ has

https://doi.org/10.1017/jpr.2020.106 Published online by Cambridge University Press

a concave graph when $\left(a, \overline{b}\right) \in T$ and is positive when a lies in some right neighborhood of 0. Then, consider the two oblique sides of T. Replacing b by linear functions of a, the function F is given along these sides by the two functions

$$f_1(a) = F\left(a, \frac{1}{3} - \frac{1}{2}\left(a - \frac{1}{3}\right)\right), \qquad f_2(a) = F(a, a).$$

It is again a matter of standard computations to check that $f_1(0) = f_1\left(\frac{1}{3}\right) = f_2(0) = f_2\left(\frac{1}{3}\right) = 0$, that both f_1' and f_2' are positive in a right neighborhood of 0, and that f_1'' is negative when $0 < a \le \frac{1}{3}$. As to f_2'', instead, straightforward computations show that it changes sign, from negative to positive, only once when $0 < a \le \frac{1}{3}$, whereas $f_2'\left(\frac{1}{3}\right) < 0$. Therefore, both $f_1(a)$ and $f_2(a)$ are positive when $0 < a < \frac{1}{3}$, which leads us to conclude that $F(a, \overline{b}) \ge 0$ when $\left(a, \overline{b}\right) \in T$ and thus $F(a, b) \ge 0$ all over T.

Analogously, using the same symbols, the second inequality in (3) is easily checked to be equivalent to

$$G(a, c) = 2 - 2\sqrt{1-c} + 2\sqrt{c} + \sqrt{1-a} - \sqrt{1-a-c} + \sqrt{a+c} - \sqrt{a} - 6c \le 0$$

in the triangle S of the (a, c) plane with vertices $(0, \frac{1}{2})$, $(0,1)$, $(\frac{1}{3}, \frac{1}{3})$. Then, straightforward computations show that the inequality holds on the sides of S, while $\frac{\partial^2 G}{\partial a^2} > 0$ throughout S, which implies the above inequality.

Moreover, we can prove (see Appendix A) that, for any set of n independent random variables, $\phi_1(\lambda) \ge \sigma_1^2$, which implies

$$\sum_{i=2}^{n} \phi_i(\lambda) \le \sum_{i=2}^{n} \sigma_i^2.$$

Now, assume $n \ge 4$. Consider the set of $n-1$ random variables $\widetilde{N} = (Y, X_3, \ldots, X_n)$, where $Y = X_1 + X_2$. Hence, $\widetilde{N}$ is a set of $n - 1 \ge 3$ independent random variables and we can apply the induction hypothesis. Assume, without loss of generality, that $\nu(N) = \lambda(N) = \nu(\widetilde{N}) = \lambda(\widetilde{N}) = 1$. Denoting by $\widetilde{\phi}_k$ the Shapley values relative to the set $\widetilde{N}$, we observe that, for any $k \ge 3$, $\phi_k(\nu) = \widetilde{\phi}_k(\nu) = \sigma_k^2$.

The next step is to prove that, for any $k \ge 3$, $\phi_k(\lambda) \le \widetilde{\phi}_k(\lambda)$. Given a permutation ψ of the elements of $\widetilde{N}$, denote by Z the sum of the random variables different from Y preceding X_k in ψ (where possibly $Z = 0$) and by $a^2 \ge 0$ the variance of Z. When we pass from the computation of $\widetilde{\phi}_k(\lambda)$ to that of $\phi_k(\lambda)$, to each previous permutation ψ correspond n new permutations. More precisely, consider the function

$$\Phi : \{\text{permutations of } (X_1, X_2, X_3, \ldots, X_n)\} \to \{\text{permutations of } (Y, X_3, \ldots, X_n)\}$$

defined by replacing, in each permutation, X_1 with Y and removing X_2. So, taking Z as above and letting $0 \le h \le n-3$ be the number of addends in Z, it follows that there exist $(h+1)!\,(n-h-2)!$ permutations of $(Y, X_3, \ldots, X_n)$ corresponding to the quantity $\sqrt{a^2 + \sigma_Y^2 + \sigma_k^2} - \sqrt{a^2 + \sigma_Y^2}$ in the computation of $\widetilde{\phi}_k(\lambda)$. Then, for each such permutation, say ψ', there are $h+1$ permutations in $\Phi^{-1}\left(\psi'\right)$ corresponding to the above quantity in the computation of $\phi_k(\lambda)$ and $n-h-1$ permutations corresponding to $\sqrt{a^2 + \sigma_1^2 + \sigma_k^2}$-$\sqrt{a^2 + \sigma_1^2}$. Similarly, there are $h!\,(n-h-1)!$ permutations of

https://doi.org/10.1017/jpr.2020.106 Published online by Cambridge University Press

$(Y, X_3, \ldots, X_n)$ corresponding to the quantity $\sqrt{a^2+\sigma_k^2}-\sqrt{a^2}$ in the computation of $\widetilde{\phi}_k(\lambda)$. For each such permutation, say ψ'', there are $n-h-1$ permutations in $\Phi^{-1}\left(\psi''\right)$ corresponding to the above quantity in the computation of $\phi_k(\lambda)$ and $h+1$ permutations corresponding to $\sqrt{a^2+\sigma_2^2+\sigma_k^2}-\sqrt{a^2+\sigma_2^2}$. Therefore, having multiplied by $\frac{n}{n}$ the expression of $\widetilde{\phi}_k(\lambda)$, the difference with $\phi_k(\lambda)$ is constituted by replacing $(h+1)!\,(n-h-1)!$ times the quantity

$$F\left(a^2, \sigma_Y^2, \sigma_k^2\right)=\sqrt{a^2+\sigma_Y^2+\sigma_k^2}-\sqrt{a^2+\sigma_Y^2}+\sqrt{a^2+\sigma_k^2}-\sqrt{a^2},$$

where $\sigma_Y^2=\sigma_1^2+\sigma_2^2$, by the quantity

$$G\left(a^2, \sigma_1^2, \sigma_2^2, \sigma_k^2\right)=\sqrt{a^2+(\sigma_1)^2+\sigma_k^2}-\sqrt{a^2+(\sigma_1)^2}+\sqrt{a^2+(\sigma_2)^2+\sigma_k^2}-\sqrt{a^2+(\sigma_2)^2}.$$

Finally, it follows from straightforward computations, which we omit for the sake of synthesis, that $F \geq G$ whenever $a^2+\sigma_1^2+\sigma_2^2+\sigma_k^2 \leq 1$, where either $a=0$ or $\sigma_1^2 \leq \sigma_2^2 \leq \min\left(a^2, \sigma_k^2\right)$. This proves that

$$\phi_k(\lambda) \leq \widetilde{\phi}_k(\lambda) \qquad \text{when } k \geq 3. \tag{4}$$

Now, our first step will be to prove that

$$\sum_{i=3}^{n} \phi_i(\lambda) \leq \sum_{i=3}^{n} \sigma_i^2. \tag{5}$$

Denoting in the following each σ_i^2 as $a_i \geq 0$, we observe that the above inequality holds both when $a_1+a_2 \leq a_3$, by the induction hypothesis, as a consequence of (4), and when $a_1=a_2$, since in such a case $\phi_2(\lambda)=\phi_1(\lambda) \geq a_1=a_2$, as proved in Appendix A.

Then, either $a_1=a_2$ or $a_1+a_2 \leq a_3$ hold, implying (5), or else we can determine $\alpha > 0$, $\beta \geq 0$ such that $a_1+\alpha=a_2-\beta$, $\alpha=(n-1)\,\beta$. Therefore, we consider a path $\mathbf{a}(x)=(a_1(x), a_2(x), a_3(x), \ldots, a_n(x))$, $0 \leq x \leq \overline{x}$, $\overline{x}>1$, such that $a_1(0)=a_1+\alpha$ and $a_i(0)=a_i-\beta$, $i=2, 3, \ldots, n$, while $a_1(x)=a_1(0)-\alpha x$ and $a_i(x)=a_i(0)+\beta x$, $i=2, 3, \ldots, n$, until $a_1\left(\overline{x}\right)+a_2\left(\overline{x}\right)=a_3\left(\overline{x}\right)$ for some $\overline{x}>1$.

Hence, denoting by $\phi_i(\lambda)(x)$ the Shapley values for the standard deviation game corresponding to $\mathrm{Var}(X_i)=a_i(x)$, $i=1, \ldots, n$,

$$\sum_{i=3}^{n} \phi_i(\lambda)(0) \leq \sum_{i=3}^{n} a_i(0), \qquad \sum_{i=3}^{n} \phi_i(\lambda)(\overline{x}) \leq \sum_{i=3}^{n} a_i(\overline{x}).$$

Let $\phi_i(\lambda)(x)=f_i(x)$, $i=3, \ldots, n$. What we want to prove is that

$$\sum_{i=3}^{n} f_i''(x)=\sum_{i=3}^{n}\left(f_i(x)-a_i(x)\right)'' \geq 0, \qquad 0 \leq x \leq \overline{x},$$

which clearly implies

$$\sum_{i=3}^{n} \phi_i(\lambda)(x) \leq \sum_{i=3}^{n} a_i(x), \qquad 0 \leq x \leq \overline{x}.$$

https://doi.org/10.1017/jpr.2020.106 Published online by Cambridge University Press

In order to do this, recall the definition of the Shapley value, i.e.

$$\phi_i(\lambda)(x) = f_i(x) = \frac{1}{n!} \sum \left(\lambda \left(P^{\lambda}(i) \cup \{i\} \right) - \lambda \left(P^{\lambda}(i) \right) \right),$$

and call $f_{ij}(x)$ the addend of $f_i(x)$ corresponding to j terms preceding $a_i(x)$ in a permutation $P^{\lambda}(i)$, $j = 0, 1, \ldots, n-1$. Consider the functions

$$g_1(x) = \sum_{i=3}^{n} \left(f_{i0}(x) + f_{i1}(x) \right), \qquad g_j(x) = \sum_{i=3}^{n} f_{ij}(x), \quad j = 2, \ldots, n-1.$$

Then, it can be shown that $g_j''(x) > 0$ as $0 \le x \le \overline{x}$ (see Appendix B). In this way, (5) is proved.

We have to consider, now, $k \geq 4$; i.e. we have to prove

$$\sum_{i=k}^{n} \phi_i(\lambda) \le \sum_{i=k}^{n} \sigma_i^2, \qquad k = 4, \ldots, n. \tag{6}$$

Let us start from $k = 4$. Again setting $\sigma_i^2 = a_i$, it follows from the same previous arguments that the inequality holds if $a_1 = a_2 = a_3$ or $a_1 + a_2 \le a_4$. Suppose neither one is the case. Then we can find α, β, γ such that

$$\begin{aligned} a_1 + \alpha = a_2 + \beta = a_3 - \gamma, &\qquad \alpha > \gamma > 0, \\ \alpha = s\gamma, &\qquad s > 0, \\ \beta = h\gamma, & \\ s + h = n - 2, & \end{aligned}$$

the extreme cases being given by $h = s = \frac{n-2}{2}$, when $a_2 = a_3$, and $h = -1$, when $a_2 = a_1$. We consider, as above, a path joining the two n-tuples where (6) holds, but inverting, for our convenience, the direction. Thus, $a(0)$ corresponds to $a_1(0) + a_2(0) = a_4(0)$, while $a(\overline{x})$ will correspond to $a_1(\overline{x}) = a_2(\overline{x}) = a_3(\overline{x})$, so that $a(x)$ is defined by

$$a_1(x) = a_1(0) + \alpha x, \qquad a_2(x) = a_2(0) + \beta x, \qquad a_i(x) = a_i(0) - \gamma x, \quad i = 3, \ldots, n.$$

Using the previous notation, we denote the functions $\phi_i(\lambda)(x)$ as $f_i(x)$, $i = 1, 2, \ldots, n$. Hence, from what we have seen, $f_1(x) + f_2(x) \geq a_1(x) + a_2(x)$ along the whole path. First of all, we can show that, for $x \in [0, \overline{x}]$,

$$f_3'(x) \le f_4'(x) \le \cdots \le f_n'(x), \tag{7}$$

while we recall that

$$\sum_{i=3}^{n} a_i(x) = \sum_{i=3}^{n} a_i(0) - (n-2)\,\gamma x.$$

In fact, compare, for example, $f_3'(x)$ and $f_4'(x)$: clearly they are equal, as are $f_3(x)$ and $f_4(x)$, when $a_3 = a_4$. So, change a_3 into $a_3 - z$ and a_4 into $a_4 + z$, for any small $z > 0$. Assume, by contradiction, that, for some $\widehat{x} \in (0, \overline{x})$,

$$\frac{\partial f_3'(\widehat{x}, 0)}{\partial z} - \frac{\partial f_4'(\widehat{x}, 0)}{\partial z} > 0.$$

https://doi.org/10.1017/jpr.2020.106 Published online by Cambridge University Press

Then, for a sufficiently small $h > 0$, letting $x \in (\widehat{x}, \widehat{x} + h)$ and setting $z = x$ as well, $f_3'(x) > f_4'(x)$. But, since $f_3\ (\widehat{x}) = f_4\ (\widehat{x})$, this implies that $f_3\ (\widehat{x} + h) > f_4\ (\widehat{x} + h)$, which leads to a contradiction, since we have observed that $\sigma_i^2 < \sigma_j^2$, $i < j$, implies $\phi_i(\lambda) \le \phi_j(\lambda)$.

Hence, moving, say, from $\overline{a} = \frac{1}{2}(a_3(x) + a_4(x))$ and setting $a_3(x) = \overline{a} - \overline{z}$, $a_4(x) = \overline{a} + \overline{z}$ at each $z \in (0, \overline{z})$ the above steps can be repeated, observing that $\phi_j(\lambda) - \phi_j(\lambda)$ increases with $\sigma_j^2 - \sigma_i^2$. Thus, the inequality $f_3'(x) \le f_4'(x)$ follows and so, by the same arguments, do the others.

Moreover, we observe that $\sum_{i=4}^n f_i''(x)$ can change sign at most once in $(0, \overline{x})$, from positive to negative, since it can be computed that $\sum_{i=4}^n f_i''(x) = 0$ implies $\sum_{i=4}^n f_i'''(x) < 0$. Finally, we recall that, at $x = \overline{x}$, $f_3(\overline{x}) \ge a_3(\overline{x})$.

Hence, assume, by contradiction, that there exists a subinterval $\left[p, q\right]$ of $[0, \overline{x}]$ such that

$$\sum_{i=4}^n f_i(x) = \sum_{i=4}^n a_i(x) \text{ when } x = p, q, \qquad \sum_{i=4}^n f_i(x) > \sum_{i=4}^n a_i(x) \text{ when } p < x < q.$$

Then, when $p < x < q$, $f_3(x) < a_3(x)$; otherwise, $f_1(x) + f_2(x) + f_3(x) \ge a_1(x) + a_2(x) + a_3(x)$ would imply $\sum_{i=4}^n f_i(x) \le \sum\limits_{i=4}^n a_i(x)$. In particular, we will have $f_3(x) < a_3(x)$ when $x \in \left[r, q\right)$, where r is such that $\sum_{i=4}^n f_i'(r) = -(n-3)\gamma$ and $\sum_{i=4}^n f_i'(x) < -(n-3)\gamma$ when $x \in (r, q)$. Hence, because of (7), $f_3'(x) < -\gamma$ for $x \in (r, q)$. Also, since $\sum_{i=4}^n f_i''(x) \le 0$ in $[r, \overline{x}]$, it follows that $\sum_{i=4}^n f_i'(x) < -(n-3)\gamma$ in $[r, \overline{x}]$, so that $f_3'(x) < -\gamma$ in $[r, \overline{x}]$ as well, implying $f_3(x) < a_3(x)$ in $[r, \overline{x}]$ and thus leading to a contradiction, as $f_3(\overline{x}) \ge a_3(\overline{x})$.

In fact, by recurrent arguments, the other cases $k > 4$ are proved analogously. This concludes the proof of the conjecture in the case of independent random variables. □

3.2. The conjecture does not hold for $n > 2$ dependent random variables

We will show that the conjecture does not hold in the case of three dependent random variables. Before providing numerical counterexamples, we consider a more general setting. Namely, let X_1, X_2, X_3 be random variables such that $\text{Var}(X_1 + X_2 + X_3) = 1$ with $\text{Cov}(X_i, X_1 + X_2 + X_3) = \frac{1}{3}$ (i.e. $\phi_i(\nu) = \frac{1}{3}$) for $i = 1, 2, 3$ and, moreover, $\sigma_1 < \sigma_2 < \sigma_3$. Then, we will prove that $\phi_1(\lambda) < \phi_2(\lambda) < \phi_3(\lambda)$.

First of all, recalling $\text{Var}(X_1 + X_2 + X_3) = 1$, it is easily calculated that

$$6\phi_i(\lambda) = 2 - 2\sqrt{\text{Var}(X_j + X_k)} + 2\sigma_i + \sqrt{\text{Var}(X_i + X_j)} - \sigma_j + \sqrt{\text{Var}(X_i + X_k)} - \sigma_k$$

where $i \ne j \ne k$. Hence, let us compare $\phi_1(\lambda)$ and $\phi_2(\lambda)$ (the other case is analogous). Again, by straightforward computations it can be checked that $\phi_1(\lambda) < \phi_2(\lambda)$ is equivalent to $\sigma_1 + \sqrt{\text{Var}(X_1 + X_3)} < \sigma_2 + \sqrt{\text{Var}(X_2 + X_3)}$. Squaring and observing that $\phi_1(\nu) = \phi_2(\nu)$ is equivalent to $\sigma_1^2 + \rho_{13}\sigma_1\sigma_3 = \sigma_2^2 + \rho_{23}\sigma_2\sigma_3$, we are eventually led to show that $\sigma_1\sqrt{\text{Var}(X_1 + X_3)} < \sigma_2\sqrt{\text{Var}(X_2 + X_3)}$, i.e.

$$\frac{\text{Var}(X_1 + X_3)}{\text{Var}(X_2 + X_3)} < \frac{\sigma_2^2}{\sigma_1^2}.$$

Since $\sigma_1^2 + \rho_{13}\sigma_1\sigma_3 = \sigma_2^2 + \rho_{23}\sigma_2\sigma_3$ implies $\rho_{13}\sigma_1\sigma_3 = \sigma_2^2 - \sigma_1^2 + \rho_{23}\sigma_2\sigma_3$, the above inequality is easily seen to be equivalent to

$$\frac{\sigma_2^2 - \sigma_1^2}{\text{Var}(X_2 + X_3)} < \frac{\sigma_2^2 - \sigma_1^2}{\sigma_1^2},$$

https://doi.org/10.1017/jpr.2020.106 Published online by Cambridge University Press

implying $\sigma_1^2 < \text{Var}(X_2 + X_3)$. Then, assume, by contradiction, that $\sigma_1^2 \geq \text{Var}(X_2 + X_3)$, i.e. $\sigma_1^2 \geq \sigma_2^2 + \rho_{23}\sigma_2\sigma_3 + \rho_{23}\sigma_2\sigma_3 + \sigma_3^2 = \sigma_1^2 + \rho_{13}\sigma_1\sigma_3 + \rho_{23}\sigma_2\sigma_3 + \sigma_3^2$. Hence, we should have $0 \geq \rho_{13}\sigma_1\sigma_3 + \rho_{23}\sigma_2\sigma_3 + \sigma_3^2 = \text{Cov}\,(X_3, X_1 + X_2 + X_3)$. But we have assumed that $\text{Cov}\,(X_3, X_1 + X_2 + X_3) = \phi_3(\nu) = \frac{1}{3}$, so we are led to a contradiction.

As we have observed, $\phi_2(\lambda) < \phi_3(\lambda)$ is proved in the same way. In fact, $\phi_2(\lambda) < \phi_3(\lambda)$ is easily checked to be equivalent to

$$\sigma_1 + \sqrt{\text{Var}(X_1 + X_3)} < \sigma_2 + \sqrt{\text{Var}(X_2 + X_3)}. \tag{8}$$

Repeating the above steps relative to the comparison of $\phi_1(\lambda)$ and $\phi_2(\lambda)$, since $\phi_2(\nu) = \phi_3(\nu)$ implies $\sigma_2^2 + \rho_{12}\sigma_1\sigma_2 = \sigma_3^2 + \rho_{13}\sigma_1\sigma_3$, it follows that (8) is equivalent to

$$\frac{\text{Var}(X_1 + X_2)}{\text{Var}(X_1 + X_3)} < \frac{\sigma_3^2}{\sigma_2^2},$$

i.e., recalling again that $\sigma_2^2 + \rho_{12}\sigma_1\sigma_2 = \sigma_3^2 + \rho_{13}\sigma_1\sigma_3$, after simple computations,

$$\frac{\sigma_3^2 - \sigma_2^2}{\text{Var}(X_1 + X_3)} < \frac{\sigma_3^2 - \sigma_2^2}{\sigma_2^2}.$$

Then, assume, by contradiction, that $\sigma_2^2 \geq \text{Var}(X_1 + X_3)$. Hence, $\sigma_2^2 \geq \sigma_1^2 + \sigma_3^2 + 2\rho_{13}\sigma_1\sigma_3 = \sigma_1^2 + \sigma_2^2 + \rho_{12}\sigma_1\sigma_2 + \rho_{13}\sigma_1\sigma_3$, i.e.

$$0 \geq \sigma_1^2 + \rho_{12}\sigma_1\sigma_2 + \rho_{13}\sigma_1\sigma_3 = \text{Cov}\,(X_1, X_1 + X_2 + X_3) = \phi_1(\nu),$$

and we are led to a contradiction, having hypothesised $\phi_1(\nu) = \frac{1}{3}$. Thus, $\phi_1(\lambda) < \phi_2(\lambda) < \phi_3(\lambda)$, which, since $\phi_1(\lambda) + \phi_2(\lambda) + \phi_3(\lambda) = 1$, implies $\phi_2(\lambda) + \phi_3(\lambda) > \frac{2}{3} = \phi_2(\nu) + \phi_3(\nu)$, disproving the conjecture.

Example 1. We provide a numerical example of the above description:

$$\sigma_1^2 = \frac{1}{6}, \quad \sigma_2^2 = \frac{2}{3}, \quad \sigma_3^2 = \frac{5}{6}, \quad \rho_{12} = \frac{1}{2} \quad \rho_{13} = 0, \quad \rho_{23} = -\frac{3}{2\sqrt{5}}.$$

In particular, in such a case, $\text{Var}(X_2 + X_3) = \frac{1}{2}$ and $\text{Var}(X_1 + X_3) = 1$, while the covariance matrix is given by

$$\sum = \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & 0 \\ \frac{1}{6} & \frac{2}{3} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{5}{6} \end{pmatrix}.$$

Example 2. This is another example with all non-negative covariances:

$$\sigma_1^2 = \frac{1}{6}, \quad \sigma_2^2 = \frac{1}{5}, \quad \sigma_3^2 = \frac{3}{10}, \quad \rho_{12} = \frac{0.8}{\sqrt{1.2}}, \quad \rho_{13} = \frac{0.2}{\sqrt{1.8}}, \quad \rho_{23} = 0.$$

The covariance matrix is

$$\sum = \begin{pmatrix} \frac{1}{6} & \frac{2}{15} & \frac{1}{30} \\ \frac{2}{15} & \frac{1}{5} & 0 \\ \frac{1}{30} & 0 & \frac{3}{10} \end{pmatrix}.$$

https://doi.org/10.1017/jpr.2020.106 Published online by Cambridge University Press

4. Future research

In Section 3.1 we have shown that the conjecture holds true for any number of random variables in the particular case of a diagonal covariance matrix. We present another case in which the conjecture is valid for any number of random variables.

Theorem 3. *Assume that all the correlation coefficients of the covariance matrix are unitary, i.e.* $\rho_{ij} \equiv 1$ *for all* $i \neq j$*. Then, the normalized Shapley values for the variance and the standard deviation games coincide. In particular, the conjecture holds true.*

Proof. Under the theorem assumption, the standard deviation value function can be written as

$$\lambda(J) = \left\{ \mathrm{Var}\left[\sum_{j \in J} X_j \right] \right\}^{1/2} = \left\{ \sum_{j \in T} \sigma_j^2 + 2 \sum_{k,j \in T} \sigma_k \sigma_j \right\}^{1/2}$$

$$= \left\{ \left(\sum_{j \in T} \sigma_j \right)^2 \right\}^{1/2} = \sum_{j \in T} \sigma_j .$$

Hence, it easily follows that $\lambda(N) = \sum_{j \in N} \sigma_j$ is the sum of all the standard deviations and that the marginal value increase simplifies to $\lambda(J \cup i) - \lambda(J) = \sigma_i$. This implies that the normalized Shapley values for the standard deviation games becomes

$$\frac{1}{\lambda(N)} \phi_i(\lambda) = \frac{1}{\lambda(N)} \sum_{T \subseteq N \setminus \{i\}} \frac{|J|!\,(N - |J| - 1)!}{N!} (\lambda(J \cup i) - \lambda(J)) = \frac{\sigma_i}{\sum_{j \in N} \sigma_j} \tag{9}$$

for all $i = 1, 2, \ldots, n$. If we multiply the numerator and the denominator in (9) by $\sum_{j \in N} \sigma_j$, we find

$$\frac{1}{\lambda(N)} \phi_i(\lambda) = \frac{\sigma_i \left(\sum_{j \in N} \sigma_j \right)}{\left(\sum_{j \in N} \sigma_j \right)^2} = \frac{\sigma_i^2 + \sum_{j \neq i} \sigma_j}{\left(\sum_{j \in N} \sigma_j \right)^2} = \frac{\mathrm{Cov}(X_i, S)}{\mathrm{Var}(S)} = \frac{1}{\nu(N)} \phi_i(\nu),$$

where ν is the variance value function. □

In conclusion, we believe that the approach we have followed to prove the conjecture in the independent case for general n can be used to check whether the conjecture of [5] holds for some specific dependence structures.

A. Appendix

In order to show that $\sum_{i=2}^{n} \phi_i(\lambda) \leq \sum_{i=2}^{n} \sigma_i^2$, we prove that $\phi_1(\lambda) \geq \sigma_1^2$. First of all, set $\sigma_i^2 = a_i$. Then define

$$f(a_1, \ldots, a_n) = (n-1)!\sqrt{a_1} + \sum_{k=2}^{n} (k-1)!\,(n-k)! \left(\left\{ a_1 + \sum_{j_1 < \cdots < j_{k-1}} a_{j_r} \right\}^{1/2} - \left\{ \sum_{j_1 < \cdots < j_{k-1}} a_{j_r} \right\}^{1/2} \right),$$

where $2 \leq j_1 < \cdots < j_{k-1} \leq n$, and $g(a_1) = n!\,a_1$. Thus, it is clear that $\phi_1(\lambda) \geq \sigma_1^2$ is equivalent to $f(a_1, \ldots, a_n) \geq g(a_1)$ whenever $0 \leq a_1 \leq a_2 \leq \cdots \leq a_n$ and $a_1 + a_2 + \cdots + a_n = 1$.

In fact, it is easily computed that, whatever the $a_2, \ldots, a_n$ satisfying the previous conditions, $f(0, a_2, \ldots, a_n) = g(0)$, $f\left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right) = g\left(\frac{1}{n}\right)$, and finally $f(a_1, \ldots, a_n) > g(a_1)$ when $0 < a_1 \le \frac{1}{n^2}$.

So, assume, by contradiction, that $f(a'_1, \ldots, a'_n) = g(a'_1)$ for some vector $\mathbf{a'} = (a'_1, \ldots, a'_n)$ satisfying the above conditions with $\frac{1}{n^2} < a'_1 < \frac{1}{n}$, and, moreover, that there exist vectors $\mathbf{a}$, still satisfying the previous conditions, with $\|\mathbf{a} - \mathbf{a}'\| < \varepsilon$ for a small enough $\varepsilon > 0$ and $a'_1 < a_1$, such that $f(\mathbf{a}) < g(a_1)$.

Hence, since $f\left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right) = g\left(\frac{1}{n}\right)$, there must also exist a vector $\mathbf{a}''$, with $a'_1 < a''_1 \le \frac{1}{n}$, such that $f(\mathbf{a}'') = g(a''_1)$, while, for suitable values of $\mathbf{a}$ close to $\mathbf{a}''$, satisfying the above conditions with $a_1 < a''_1$, it again holds that $f(\mathbf{a}) < g(a_1)$.

Let us start, precisely, from such an $\mathbf{a}''$. It is easily computed that, for $a_1 > 0$, $\frac{\partial f}{\partial a_1} > 0$, while, for $k = 2, \ldots, n$, $\frac{\partial f}{\partial a_k} < 0$. Hence, for a suitably small $x > 0$ and suitable non-negative values $b_2, \ldots, b_n$ such that $b_2 + \cdots + b_n = 1$, the vector $\mathbf{a} = (a''_1 - x, a''_2 + b_2 x, \ldots, a''_n + b_n x)$ satisfies $f(\mathbf{a}) < g(a_1)$. Therefore, it is possible to construct a path $a(x)$, $0 \le x \le \overline{x}$, joining a'' to some a' defined as above, such that $f(\mathbf{a}(x)) < g(a_1(x))$ when $0 < x < \overline{x}$, $f(\mathbf{a}(0)) = g(a_1(0))$, $f(\mathbf{a}(\overline{x})) = g(a_1(\overline{x}))$. In fact, we can expect, at most, such a path to be piecewise linear, since, moving, say, linearly from $\mathbf{a}''$, some $a_i(x)$, $2 \le i < n$, might reach an initially higher a_{i+1} before the path reaches $\mathbf{a}'$. However, a piecewise linear path can be approximated, as well as we want, by a smooth one. Hence, after inverting, for our convenience, the direction of the path, i.e. letting it go from $\mathbf{a}'$ to $\mathbf{a}''$, we can write $\mathbf{a}(x) = (a'_1 + x, a'_2 - \beta_2(x), \ldots, a'_n - \beta_n(x))$, $0 \le x \le \overline{x}$, with $0 \le \beta_2(x), \ldots, \beta_n(x)$, $\beta_2(x) + \cdots + \beta_n(x) = x$, $\beta'_2(x), \ldots, \beta'_n(x) \ge 0$, $\beta''_2(x), \ldots, \beta''_{n-1}(x) \le 0$, so that, setting $\widetilde{f}(x) = f(\mathbf{a}(x))$, $\widetilde{g}(x) = g(a_1(x)) = n!\,(a'_1 + x)$, $\widetilde{f}(0) = \widetilde{g}(0)$, $\widetilde{f}(\overline{x}) = \widetilde{g}(\overline{x})$, $\widetilde{f}(x) < \widetilde{g}(x)$ when $0 < x < \overline{x}$. Moreover, as $\widetilde{g}'(x) = n!$, it follows that $\widetilde{f}'(x) < n!$ when $0 < x < \varepsilon$ and $\widetilde{f}'(x) > n!$ when $x' - \varepsilon < x < \overline{x}$ for a suitably small $\varepsilon > 0$. However, straightforward computations show that $\widetilde{f}''(x) < 0$ when $0 < x < \overline{x}$, leading to a contradiction. In fact, since $\beta_n(x) = x - \beta_2(x) - \cdots - \beta_{n-1}(x)$, $\beta''_n(x) = -\beta''_2(x) - \cdots - \beta''_{n-1}(x)$. Hence, in particular, in the expression of $\widetilde{f}''(x)$, to each term of the type

$$\frac{1}{2}\beta''_n(x)\left(\frac{1}{\sqrt{\sum_r a_{j_r} + a_h}} - \frac{1}{\sqrt{\sum_r a_{j_r} + a_h + a_1}}\right)$$

corresponds a sum

$$\frac{1}{2}\left[\sum_r \beta''_{j_r}(x)\left(\frac{1}{\sqrt{\sum_r a_{j_r} + a_n}} - \frac{1}{\sqrt{\sum_r a_{j_r} + a_n + a_1}}\right)\right.$$
$$\left. + \sum_{2 \le h \ne j_r, n} \beta''_h(x)\left(\frac{1}{\sqrt{\sum_r a_{j_r} + a_h}} - \frac{1}{\sqrt{\sum_r a_{j_r} + a_h + a_1}}\right)\right].$$

https://doi.org/10.1017/jpr.2020.106 Published online by Cambridge University Press

Since, when $h \neq n$,

$$\frac{1}{\sqrt{\sum_r a_{j_r} + a_h}} - \frac{1}{\sqrt{\sum_r a_{j_r} + a_h + a_1}} \geq \frac{1}{\sqrt{\sum_r a_{j_r} + a_h}} - \frac{1}{\sqrt{\sum_r a_{j_r} + a_h + a_1}},$$

the sum of the above two quantities is ≤ 0.

B. Appendix

Let us consider, first, $g_1(x)$. For any fixed $x \in [0, \overline{x}]$, set $a_i(x) = a_i$. Observe that, whenever $3 \leq l < m \leq n$,

$$\frac{1}{2}\left[2\frac{\beta^2}{(a_l + a_m)^{\frac{3}{2}}} - \frac{\beta^2}{4\,(a_m)^{\frac{3}{2}}} - \frac{\beta^2}{4\,(a_l)^{\frac{3}{2}}}\right] = \frac{\beta^2}{4}\,(\widehat{a})^{-\frac{3}{2}}\left(\sqrt{2} - 1\right)$$

for some $a_l \leq \widehat{a} \leq a_m$, so that $\widehat{a} \geq a_2 \geq a_1$. Hence, straightforward computations lead us to conclude that $g_1''(x) \geq 0$ certainly holds if

$$(n-1)\left(1 + \frac{1}{2\sqrt{2}}\right) < \left((n-1)^2 + n - 2\right)\left(1 - \frac{1}{2\sqrt{2}}\right)$$

when $n \geq 4$. In fact, the *worst* case is precisely $n = 4$, when it is easily checked that

$$3\left(1 + \frac{1}{2\sqrt{2}}\right) < 11\left(1 - \frac{1}{2\sqrt{2}}\right).$$

As to $g_j(x)$, when $2 \leq j \leq n-2$, it can be shown that, in order to prove $g_j''(x) \leq 0$, it is sufficient to check the inequality

$$\begin{aligned}(j+1)^{-\frac{3}{2}}&\left[\binom{n-2}{j-1}(n-j-1)^2 + \binom{n-2}{j}(j+1)^2\right]\\ &\leq (j)^{-\frac{3}{2}}\left[\binom{n-2}{j-1}(n-j)^2 + \binom{n-2}{j}(j)^2\right],\end{aligned}$$

which follows through straightforward steps. Finally, $f_{i,n-1}(x) = \frac{1}{n}(1 - \sqrt{1 - a_i\,(x)})$, so that it is obvious that $f''_{i,n-1}(x) \geq 0$.

Acknowledgement

Marcello Galeotti is a member of INDAM (National Institute of High Mathematics) affiliated to the group of Mathematical Analysis and Probability (GNAMPA). Giovanni Rabitti would like to thank the Fondazione Invernizzi for the support received when he was postdoc at Bocconi University.

References

[1] Abbasi, B. and Hosseinifard, S. Z. (2013). Tail conditional expectation for multivariate distributions: A game theory approach. *Statist. Prob. Lett.* **83**, 2228–2235.

[2] Anily, S. and Haviv, M. (2010). Cooperation in service systems. *Operat. Res.* **58**, 660–673.

[3] Castro, J., Gómez, D. and Tejada, J. (2009). Polynomial calculation of the Shapley value based on sampling. *Comput. Operat. Res.* **36**, 1726–1730.

[4] CHEN, Z., HU, Z. AND TANG, Q. (2020). Allocation inequality in cost sharing problem. *J. Math. Econom.* **91**, 111–120.

[5] COLINI-BALDESCHI, R., SCARSINI, M. AND VACCARI, S. (2018). Variance allocation and Shapley value. *Methodology Comput. Appl. Prob.* **20**, 919–933.

[6] LIPOVETSKY, S. AND CONKLIN, M. (2001). Analysis of regression in game theory approach. *Appl. Stoch. Models Bus. Ind.* **17**, 319–330.

[7] LUNDBERG, S. M. *et al.* (2020). From local explanations to global understanding with explainable AI for trees. *Nature Mach. Intel.* **2**, 56–67.

[8] MARICHAL, J.-L. AND MATHONET, P. (2013). On the extensions of Barlow–Proschan importance index and system signature to dependent lifetimes. *J. Multivar. Anal.* **115**, 48–56.

[9] MORETTI, S. AND PATRONE, F. (2008). Transversality of the Shapley value. *TOP* **16**, 1.

[10] MÜLLER, A., SCARSINI, M. AND SHAKED, M. (2002). The newsvendor game has a nonempty core. *Games Econom. Behav.* **38**, 118–126.

[11] OWEN, A. B. (2014). Sobol' indices and Shapley value. *SIAM/ASA J. Uncert. Quant.* **2**, 245–251.

[12] SHAPLEY, L. S. (1953). A value for *n*-person games. In *Contributions to the Theory of Games*, eds H. W. Kuhn and A. W. Tucker, Princeton University Press, pp. 307–317.

[13] SONG, E., NELSON, B. L. AND STAUM, J. (2016). Shapley effects for global sensitivity analysis: Theory and computation. *SIAM/ASA J. Uncert. Quant.* **4**, 1060–1083.