Universitat de Barcelona

Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Càlcul del valor de Shapley

Autor: Roger Parra Serret

Director: Dr. Josep Vives
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 23 de gener de 2022

Abstract

The Shapley value is a solution concept used in game theory that involves fairly distributing both gains and costs to several actors working in coalition. In this article, we study its problems and we talk about possible solutions of diferent situations that there exist. We will place much emphasis in some projects like 'An analysis of the Shapley Value and its Uncertainty for the Voting Game' done by Shaheen S. Fatima, Michael Wooldridge and Nicholas R. Jenning, and 'Polynomial calculation of the Shapley value based on sampling' done by Javier Castro, Daniel Gómez and Juan Tejada. Furthermore, we will see another way to compute this value, based on the work mentioned above about random sampling. Finally, we will compare the last two methods that we have presented and we will see an example based on a real situation.

Resum

El valor de Shapley és un concepte de solució que s'utilitza en la teoria de jocs, i que tracta el tema del repartiment just dels guanys entre els membres d'una coalició. En aquest treball, parlem de les diferents dificultats que presenta aquest concepte i d'algunes possibles solucions que se l'hi poden donar en diverses situacions. Farem èmfasi en alguns treballs com 'Anàlisi del valor de Shapley i la seva incertesa per a jocs cooperatius' fet per Shaheen S. Fatima, Michael Wooldridge i Nicholas R. Jenning, i 'Càlcul polinomial del valor de Shapley basat en el mostreig' fet per Javier Castro, Daniel Gómez i Juan Tejada. A més a més, veurem una altra manera de calcular-ne el valor, basat en el treball anteriorment esmentat sobre mostreig aleatori. Finalment, compararem els dos mètodes anteriors i en veurem algún exemple aplicat a situacions reals.

2020 Mathematics Subject Classification. 91A12, 62D05

Agraïments

Vull agrair primer de tot al Dr. Josep Vives per ajudar-me cada vegada que ho necessitava. També m'ha donat suport tant en la correcció, com en continguts que no sabia sobre el tema i sobre Làtex.

També m'agradaria agrair a la meva família per escoltar-me parlar sobre el tema en tot moment i donar-me tant de suport. Han sigut uns mesos molt complicats i exhaustius entre el TFG i la feina, i m'han animat en tot moment, creient que me'n sortiria sense problemes.

Moltes gràcies.

Índex

1 Introducció

1.1 Què és la teoria de Jocs?

La teoria de jocs és una teoria matemàtica que estudia les decisions d'un individu per tal de tenir èxit en un joc o en una situació de conflicte, tenint en compte les decisions que prenen la resta d'agents implicats. Aquesta teoria analitza matemàticament el comportament òptim dels diversos jugadors davant les possibles estratègies aplicables per a la resolució guanyadora del conflicte. La teoria de jocs s'aplica a àmbits com l'economia, la gestió, la biologia, la psicologia o l'estratègia.

1.2 Com apareix?

La teoria de jocs apareix per primera vegada en una carta que escriu James Waldegrave el 1713 en la que dóna una solució d'estratègia mixta a un joc de cartes anomenat Le Her. Ja més tard, el 1838, Antoine-Agustin Cournot, publica *la Recerca dels Principis Matemàtics de la Teoria de la Riquesa* on, en termes generals, s'hi comenta una anàlisi del que seria la Teoria de Jocs. Cournot considera un duopoli, i en presenta una solució que avui dia es considera com una versió restringida de l'equilibri de Nash. Més tard, es podria dir que el fundador de teoria de jocs va ser John Von Neumann, que junt amb Christian Morgenstern, van publicar el 1944 *The Theory of Games and economic behaviour*. Van investigar dos plantejaments diferents de la teoria de jocs. El primer va ser el plantejament estratègic o no cooperatiu. Es van centrar en jocs de dues persones amb objectius totalment oposats. A aquest tipus de jocs se'ls anomena estrictament competitius o de suma zero. El segon va ser el plantejament d'aliances o cooperatiu en el que es busca descobrir la conducta òptima dels jugadors en jocs on hi ha molts participants. Ens centrem en aquesta classe de jocs, els jocs cooperatius.

1.3 El dilema del presoner

Dues persones, que anomenarem jugadors, que no es coneixen entre elles i no tenen cap informació una de l'altra, són detingudes acusades d'algun crim i interrogades en sales diferents. Els dos jugadors tenen dues opcions, culpar a l'altre jugador, o no fer-ho; i depenent de la seva acció, seran sentenciats a més o menys anys de presó segons el que indica la Taula 1.

		Jugador 2	
		Culpar	No Culpar
	Culpar	(10,10)	(0,15)
Jugador 1	No Culpar	(15,0)	(1,1)

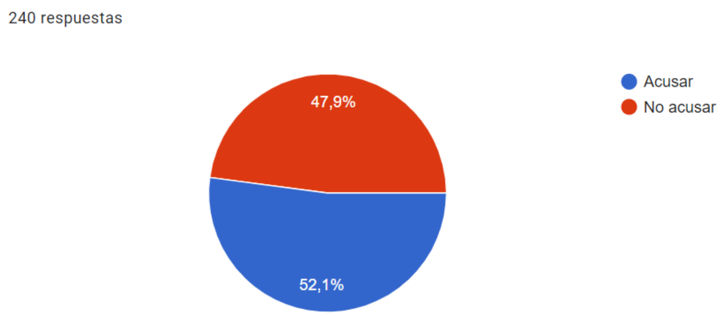
Taula 1: Prisoners Dilemma

Dels vectors que apareixen a la taula, la primera component són els anys que rebria el jugador 1 i la segona component els anys que rebria el jugador 2.

En el cas de que només un dels dos jugadors acusi a l'altre, aquest darrer rebria tota la culpa, i a l'inrevés. Si els dos s'acusen, el jutge no sabria del tot qui és el culpable però veu la possibilitat de que haguéssin col·laborat i posa 10 anys de presó a cadascú. I finalment, si cap dels dos s'acusa, com no sap que decidir perquè no té cap informació decideix només sentenciar-los a 1 any de presó per a cadascú.

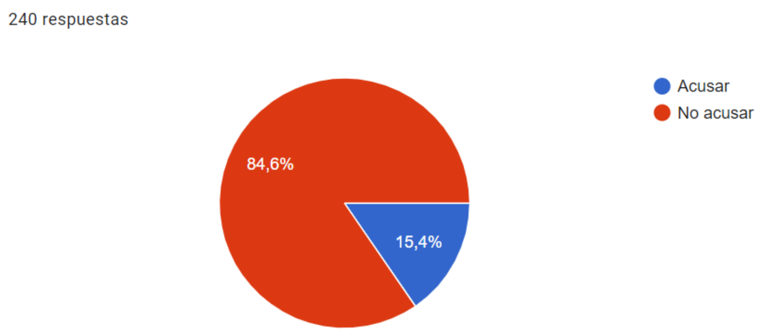
Quina seria la millor solució llavors?

He fet una enquesta a gent del meu entorn, presentant-los la mateixa situació com si ells fossin un dels dos acusats. He obtingut 240 respostes i aquestes han estat bastant repartides.



Tot i així, veiem que el què s'emporta més quantitat de vots, és 'acusar' i és que des del punt de vista de la teoria de jocs, culpar seria la decisió idònia. Per què? Doncs perquè com que no tenim cap informació de l'altre jugador, hem de pensar només en nosaltres. Per tant, si ell ens acusés, tindríem 10 (anys de presó si també l'acusem) < 15 (anys de presó si no l'acusem). En canvi, si ell decidís no acusar-nos, tindríem 0 (anys de presó si l'acusem) < 1 (anys de presó si tampoc l'acusem).

Després de preguntar això, vaig voler veure quins serien els resultats, si es conegués a l'altre acusat. Observem que els resultats obtinguts són els següents:



Com podem veure, la gent aquí ho analitza des d'un punt de vista psicològic ja que, si coneguéssim a l'altre acusat (i ell ens conegués a nosaltres), potser la millor solució seria no acusar pensant que l'altre també sacrificaria 1 any de la seva vida a canvi de no jugar-se'n 15 o 10.

2 Jocs cooperatius amb utilitat transferible

La principal diferència entre els jocs cooperatius i els jocs no cooperatius és que aquests presenten la possibilitat d'establir acords vinculants. Es classifiquen en dos tipus, jocs amb utilitat transferible (TU) o jocs sense utilitat transferible (NTU). Ens centrarem amb els jocs d'utilitat transferible. Amb utilitat transferible ens referim al fet d'assumir que la utilitat pot ser transferida de forma gratuïta d'un jugador a un altre, sent aquesta utilitat comuna i divisible infinitament. Cal conèixer, doncs, quina és la utilitat de cada subgrup de jugadors. En els jocs cooperatius, en general, partim d'una situació on s'han de repartir els beneficis i els costos d'algun projecte entre els diferents jugadors que hi participen. Assumim, de forma general, que els participants són racionals, és a dir, que busquen la màxima utilitat possible independentment del que rebi la resta de jugadors. Amb la teoria de jocs cooperatius amb utilitat transferible l'objectiu serà descobrir com repartir els guanys entre els diferents jugadors implicats, tenint en compte les possibles aliances entre jugadors o la actuació conjunta de tots ells.

Definició 2.1. *Un joc cooperatiu amb utilitat transferible és un parell* (N, v) *on* $N = \{1,...,n\}$ *és el conjunt finit de jugadors que participen en el joc i v:* $2^N \to \mathbb{R}$ *(*2^N *és el conjunt de subconjunts, és a dir, el conjunt de les parts de N) és el que s'anomena funció característica, que assigna a cada subconjunt* $S \subset N$ *(que representen cada una de les possibles coalicions del joc) un nombre real* $v(S)$ $(v(\emptyset) := 0)$.

Així doncs, donada una coalició S, $v(S)$ designa la utilitat o valor que obté la coalició si cooperen els jugadors que la formen, independentment del que faci la resta de jugadors que no pertanyen a la coalició.

Anomenem G a la classe de tots els jocs TU amb conjunt de jugadors finit. Anomenem G(N) a la classe de tots els jocs TU amb conjunt de jugadors N. Siguin $(N, v), (N, v')$ dos jocs TU, veiem que v compleix la propietat de la suma i el producte per escalars. El joc TU, (N, v'') on $v'' = v+v'$ es defineix per $v''(S) := v(S)+v'(S)$ per a cada coalició $S \subset N$. El joc TU, (N, v'') on $v''= \lambda v$ $(\lambda \in \mathbb{R})$ es defineix per $v''(S) := \lambda v(S)$ per a cada coalició $S \subset N$. Si parlem de l'estructura algebràica dels jocs TU diem que el conjunt de tots els jocs TU amb conjunt de jugadors N amb la suma i el producte per escalars $(G(N), +, \cdot; \mathbb{R})$, és un espai vectorial de dimensió $2^n - 1$ on $n = |N|$.

Definició 2.2. *Sigui* $(N, v) \in G(N)$ *un joc TU, direm que* (N, v) *és superadditiu si* $\forall S, T \subset N$ *amb* $S \cap N = \emptyset$*,*

$$v(S \cup T) \geq v(S) + v(T).$$

SG(N) és el nom que rep el conjunt de tots els jocs superadditius de N jugadors.

Els jocs superadditius, per tant, són jocs on als jugadors els hi interessa clarament cooperar entre ells, ja que qualsevol quina sigui la coalició entre diferents jugadors, la utilitat donada per la funció característica a aquesta aliança, serà superior a la utilitat de cada jugador individualment. Podem dir que la teoria de jocs TU té el principal objectiu de buscar una assignació o un conjunt d'assignacions acceptable per tots els jugadors. Una estabilitat.

Definició 2.3. *Donat un joc $(N, v) \in G(N)$, es diu que un vector $x \in \mathbb{R}^N$ és una imputació si satisfà dues condicions:*

1. $x_i \geq v(i)\ \forall i \in N$.
2. $\sum_{i=1}^{n} x_i = v(N)$.

Sigui $(N, v) \in G(N)$, $I(N, v)$ serà el nom que rep el conjunt d'imputacions del joc. Si el joc es superadditiu, $I(N, v) \neq \emptyset$. Les imputacions d'un joc (N, v) són les assignacions de $v(N)$ entre els jugadors, que satisfan la condició de racionalitat individual. És a dir, que compleixen que $x_i \geq v(i)$, $\forall i \in N$; cada jugador obté una utilitat de, com a mínim, el que obtindria sol (el primer punt de la definició anterior).

Definició 2.4. *Sigui $(N, v) \in G(N)$ i siguin $x, y \in I(N, v)$ dues imputacions. Es diu que x domina a y a través de $S \in 2^N \backslash \emptyset$ si:*

1. $x_i > y_i\ \forall\ i \in S$. *(Els jugadors prefereixen el valor que els assigna el vector x abans que el que els assigna el vector y).*
2. $\sum_{i \in S} x_i \leq v(S)$. *(El vector x és factible per S, ja que, la quantitat total proposada per x pels jugadors de S no és superior al que rebrien aquests al aliar-se, $v(S)$)*

Direm que x domina a y si existeix una coalició $S \in 2^N \backslash \emptyset$ tal que x domina a y a través de S.

Definició 2.5. *Definim el nucli o core d'un joc $(N, v) \in G(N)$ com el següent conjunt de $I(N, v)$,*

$$C(N, v) = \{x \in I(N, v) : \textstyle\sum_{i \in S} x_i \geq v(S)\ \forall S \in 2^N\}$$

Proposició 2.6. *Sigui $(N, v) \in G(N)$ un joc TU..*

1. *Si $x \in C(N, v)$, llavors x és no dominada.*
2. *Si, a sobre, el joc és superadditiu, llavors el nucli són les imputacions no dominades.*

Demostració

1. Suposem una imputació $x \in C(N, v)$ i una imputació $y \in I(N, v)$ i una coalició $S \subset N$ no buida tal que y domina a x a través de S . Llavors

$$v(S) \geq \textstyle\sum_{i \in S} y_i > \sum_{i \in S} x_i \geq v(S).$$

Per tant, arribem a una contradicció.

2. Sigui x una imputació no dominada de (N, v), $x \notin C(N, v)$. Llavors existeix $S \subset N$ tal que $\sum_{i \in S} x_i < v(S)$. Definim $y \in \mathbb{R}^N$ com

$$y_i = \begin{cases} x_i + \dfrac{v(S) - \sum_{j \in S} x_j}{|S|}, & \text{si } i \in S \\ v(i) + \dfrac{v(N) - v(S) - \sum_{j \in N \setminus S} v(j)}{|S|}, & \text{si } i \in N \setminus S \end{cases}$$

D'aquesta manera, hem definit una imputació y de forma que domina a x a través de S, per tant, arribem a una contradicció. □

Definició 2.7. *Un joc $(N, v) \in G(N)$ es diu que és simple si:*

1. *$\forall S \in N$, $v(S) = 0$ o $v(S) = 1$.*
2. *$v(N) = 1$.*
3. *(N, v) és monòton, és a dir, $v(S) \leq v(T)$ $\forall S, T \subset N$ amb $S \subset T$.*

$S(N)$ és el nom que rep el conjunt dels jocs simples.

Sigui doncs (N, v) un joc simple, una coalició s'anomenarà guanyadora si $v(S) = 1$ i perdedora si $v(S) = 0$. W serà el conjunt de totes les coalicions guanyadores. W^m és el conjunt de totes les coalicions guanyadores minimals, és a dir les coalicions guanyadores que no contenen a cap altra coalició guanyadora.

Definició 2.8. *Donat un joc $(N, v) \in S(N)$, es diu que $i \in N$ és un jugador vet de (N, v) si $v(N \setminus \{i\}) = 0$.*

Proposició 2.9. *$(N, v) \in S(N)$. Llavors $C(N, v) \neq \emptyset$, si i només si, el conjunt V dels jugadors vet és no buit. A més, si $C(N, v) \neq \emptyset$ llavors és igual a les imputacions tals que $x_i = 0$, $\forall i \in N \backslash V$.*

Demostració. Agafem $x \in C(N, v)$ i suposem $V = \emptyset$. Llavors $\forall i \in N$,

$$0 = v(N) – v(N \backslash \{\text{i}\}) \geq \textstyle\sum_{j \in N} x_j – \sum_{j \in N \backslash \{i\}} x_j = x_i > 0.$$

Cosa que no és possible. Recíprocament, suposem que $V \neq \emptyset$ i considerem el conjunt no buit $\{x \in I(N, v) : x_i = 0 \ \forall i \in N \backslash \text{V}\}$
És immediat comprovar que aquest conjunt és el nucli. □

Definició 2.10. *Una família de coalicions $F \subset 2^N \backslash \{\emptyset\}$ es denomina equilibrada si existeix una família de nombres reals positius corresponents (coeficients d'equilibri) $\{y_S \colon S \in F\}$ tal que, $\forall i \in N$,*

$$\textstyle\sum_{S \in F, i \in S} y_S = 1.$$

Definició 2.11. *Un joc $(N, v) \in G(N)$ es denomina equilibrat si per cada família equilibrada de coalicions F amb coeficients d'equilibri $\{y_S : S \in F\}$ es té*

$$\textstyle\sum_{S \in F} y_S v(S) \leq v(N).$$

Teorema 2.12. *(Bondareva-Shapley) Sigui $(N, v) \in G(N)$ un joc. Llavors $C(N, v) \neq \emptyset$, si i només si, (N, v) és equilibrat.*

Demostració Sigui $x \in C(N, v)$ i F una família de coalicions equilibrada amb coeficients d'equilibri $\{y_S : S \in F\}$. Llavors,

$$\textstyle\sum_{S \in F} y_S v(S) \leq \sum_{S \in F} \sum_{i \in S} y_S x_i = \sum_{i \in N} [x_i \sum_{S \in F, i \in S} y_S] = \sum_{i \in N} x_i \cdot 1 = v(N).$$

Recíprocament, suposem que v és equilibrat i considerem el següent problema de programació lineal (P):

Volem minimitzar $\sum_{i \in N} x_i$, tenint en compte $\sum_{i \in S} x_i \geq v(S), \forall S \in 2^N \backslash \{\emptyset\}$.

Clarament $C(v) \neq \emptyset$, si i només si, existeix $\bar{x}$ una solució òptima de (P) amb $\sum_{i \in N} \bar{x}_i = v(N)$. El dual de (P) és el següent problema de programació lineal (D):

Volem maximitzar $\sum_{S \in 2^N \backslash \{\emptyset\}} y_S v(S)$, tenint en compte $\sum_{S \in 2^N \backslash \{\emptyset\}} y_S = 1, \forall i \in N$,

$$y_S \geq 0, \forall S \in 2^N \backslash \{\emptyset\}.$$

Clarament, (D) té al menys una solució òptima $\bar{y}$ ja que, el conjunt de solucions factibles és no buit i compacte, i la seva funció objectiu és continua. Definim ara

$$\bar{F} = \{S \subset N : \bar{y}_S > 0\}.$$

Obviament $\bar{F}$ és una família equilibrada amb coeficients d'equilibri $\{\bar{y}_S \in \bar{F}\}$. Com que (D) té una solució òptima llavors (P) també en té una, $\bar{x}$, i a més

$$v(N) \leq \sum_{i \in N} \bar{x}_i = \sum_{S \in 2^N \backslash \{\emptyset\}} \bar{y}_S v(S) \leq v(N). \tag{2.1}$$

Com que v és equilibrat i $\bar{x}$ és una solució òptima de (P), acabem la demostració ja que, per 2.1 tenim $v(N) = \sum_{i \in N} \bar{x}_i$. $\square$

3 El valor de Shapley

3.1 Definició i propietats

Ara parlarem del valor de Shapley, el contingut principal del nostre treball. Lloyd Shapley (1923-2016) va ser un matemàtic i economista americà, professor a la Universitat de California. Les seves contribucions principals van ser a l'economia matemàtica. Va rebre el Premi Nobel d'Economia el 2012 juntament amb *Alvin E. Roth*. Va ser mentor de John F. Nash, qui també va rebre el Nobel de l'Economia l'any 1994.

Entre alguns dels seus mèrits, també hi trobem la participació en l'exèrcit dels Estats Units d'Amèrica el 1943 com a meteoròleg, on, gràcies a la criptografia, va poder desxifrar el codi del clima dels soviètics i se'l va premiar amb l'Estrella de Bronze.

L'any 1953 va introduir l'anomenat valor de Shapley, com a mètode per assignar en cada joc cooperatiu, els guanys de cada coalició, de forma equitativa.

Definició 3.1. *Sigui* $(N, v) \in G(N)$ *un joc:*

1. *Un jugador de* N *s'anomena jugador nul de* (N, v) *si* $\forall S \subset N$, $v(S \cup \{i\}) = v(S)$.

2. *Dos jugadors* $i, j \in N$ *són intercanviables en* (N, v) *si* $\forall S \subset N \backslash \{i, j\}$, $v(S \cup \{i\}) = v(S \cup \{j\})$.

Sigui $f : G(N) \to \mathbb{R}^n$ una aplicació:

- Direm que f satisfà **la propietat d'eficiència** si $\forall (N, v) \in G(N)$, $\sum_{i \in N} f_i(N, v) = v(N)$.

- Direm que f satisfà **la propietat del jugador nul** si $\forall (N, v) \in G(N)$, i per cada jugador nul $i \in N$, $f_i(N, v) = 0$.

- Direm que f satisfà **la propietat de simetria** si $\forall (N, v) \in G(N)$, i per cada jugadors intercanviables $i, j \in N$, $f_i(N, v) = f_j(N, v)$.

- Direm que f satisfà **la propietat d'additivitat** si per tot parell de jocs $(N, v), (N, w) \in G(N)$, $f(N, v + w) = f(N, v) + f(N, w)$.

Teorema 3.2. *Només existeix un únic valor $\phi : G(N) \to \mathbb{R}^n$ que satisfà les propietats d'eficiència, del jugador nul, de simetria i d'additivitat. Aquest valor s'anomena* ***valor de Shapley*** *i ve donat per:*

$$\phi_i(N, v) = \sum_{S \subset N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!}(v(S \cup \{i\}) - v(S)).$$

Demostració
És immediat provar que ϕ satisfà les propietats de jugador nul i additivitat. Per comprovar que satisfà eficiència considerem $(N, v) \in G(N)$ i notem que

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

$$= \sum_{i \in N} \sum_{S \subset N \setminus \{i\}} \frac{s!(n-s-1)!}{n!} v(S \cup \{i\}) - \sum_{i \in N} \sum_{S \subset N \setminus \{i\}} \frac{s!(n-s-1)!}{n!} v(S)$$

$$\underset{*}{=} \sum_{S \subseteq N} s \frac{(s-1)!(n-s)!}{n!} v(S) - \sum_{S \subset N, S \neq N} (n-s) \frac{s!(n-s-1)!}{n!} v(S) \underset{**}{=} v(N).$$

(*) Anem a veure aquesta igualtat. Comencem pel segon sumand, notem que passem de $\sum_{i \in N} \sum_{S \subset N \setminus \{i\}}$ a $\sum_{S \subset N, S \neq N}(n-s)$. Això és degut a que estem comptant les coalicions de N que no contenen a i per a cada $i \in N$. Imaginem un cas de 3 jugadors: per a $i = 1$ tindríem les coalicions $\emptyset, \{2\}, \{3\}, \{2, 3\}$; per a $i = 2$ tindríem les coalicions $\emptyset, \{1\}, \{3\}, \{1, 3\}$; i per a $i = 3$ tindríem les coalicions $\emptyset, \{1\}, \{2\}, \{1, 2\}$; d'aquesta forma podem compendre el $\sum_{S \subset N, S \neq N}$ ja que, hi apareixen totes les coalicions possibles menys N. Ara bé, fixem-nos que cada una d'aquestes coalicions apareix $(n - s)$ vegades, per això hi trobem aquest factor.

Ara, pel que fa al primer sumand, veiem que estem tractant amb $v(S \cup \{i\})$. Si fem el mateix que l'apartat anterior però en el cas de $S' = S \cup \{i\}$, imaginant el cas de 3 jugadors, tindrem: per a $i = 1$ tindríem les coalicions $\{1\}, \{1, 2\}, \{1, 3\}, \{1, 2, 3\}$; per a $i = 2$ tindríem les coalicions $\{2\}, \{1, 2\}, \{2, 3\}, \{1, 2, 3\}$; i per a $i = 3$ tindríem les coalicions $\{3\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}$; per tant veiem que cada coalició es repeteix $s'(= s+1)$ vegades. D'aquí veiem com passem de $= \sum_{i \in N} \sum_{S \subset N \setminus \{i\}}$ a $\sum_{S' \subseteq N, S' \neq \emptyset} s'$ (podem obviar la part $S' \neq \emptyset$ ja que, $v(\emptyset) = 0$) i fent el canvi de la S' i de la s', també veiem com $\frac{s!(n-s-1)!}{n!} v(S \cup \{i\})$ passa a $\frac{(s'-1)!(n-s'+1-1)!}{n!} v(S')$. Ara podem ajuntar factors i, podem anomenar S a S' i s a s' novament sense que els resultats canviïn.

(**) Finalment, per veure la última igualtat simplement caldria sumar i veuríem com només ens queda el cas on $S = N$. Tindríem $\frac{n!0!}{n!} v(N) = v(N)$.

Per provar que ϕ satisfà la propietat de simetria considerem $(N, v) \in G(N), i, j \in N$ simètrics a (N, v) i notem que

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

$$- \sum_{S \subset N \setminus \{j\}} \frac{s!(n-s-1)!}{n!}(v(S \cup \{j\}) - v(S))$$

$$= \sum_{S \subset N \setminus \{i\}, j \in S} \tfrac{s!(n-s-1)!}{n!} (v(S \cup \{i\}) - v(S))$$

$$- \sum_{S \subset N \setminus \{j\}, i \in S} \tfrac{s!(n-s-1)!}{n!} (v(S \cup \{j\}) - v(S))$$

$$= \sum_{S \subset N \setminus \{i,j\}} \tfrac{(s+1)!(n-s-2)!}{n!} (v(S \cup \{i\}) - v(S \cup \{j\})) = 0.$$

Per tant, ϕ satisfà les propietats necessàries. Veiem ara que és l'únic que les satisfà. Per cada coalició no buida $S \subset N$, anem a definir un joc TU $(N, u^S) \in G(N)$ tal que

$$u^S(T) = \begin{cases} 1, & \text{si } S \subset T \\ 0, & \text{en cas contrari} \end{cases}$$

Ara tenim en compte que cada $(N, v) \in G(N)$ es pot veure com un vector

$$(v(S))_{S \in 2^N \setminus \{\emptyset\}} \in \mathbb{R}^{2^n - 1}.$$

Llavors $G(N)$ es pot veure com un espai vectorial de dimensió $2^n - 1$. Veiem que $U(N) = \{(N, u^S) : S \in 2^N \setminus \{\emptyset\}\}$ és una base de tal espai vectorial. Per provar-ho, és suficient provar que $U(N)$ és un conjunt de vectors linealment independents.

Considerem, llavors $\sum_{S \in 2^N \setminus \{\emptyset\}} \alpha_S u^S = 0$ ($\alpha_S \in \mathbb{R}$ $\forall S \in 2^N \setminus \{\emptyset\}$) i suposem que existeix $T \in 2^N \setminus \{\emptyset\}$ amb $\alpha_T \neq 0$. Suposem, sense perdre generalitat, que no existeix $\bar{T} \subset T$ amb $\bar{T} \neq T$, tal que $\alpha_{\bar{T}} \neq 0$. Però llavors $\sum_{S \in 2^N \setminus \{\emptyset\}} \alpha_S u^S(T) = \alpha_T$, cosa que no és possible. Prenem ara f que satisfà eficiència, les propietat de jugador nul i simetria. Clarament, $\forall i \in N$, cada coalició no buida $S \subset N$ i cada $\alpha_S \in \mathbb{R}$ es té

$$f_i(N, \alpha_S u^S) = \begin{cases} \dfrac{\alpha_S}{|S|}, & \text{si } i \in S \\ 0, & \text{en cas contrari} \end{cases}$$

Per acabar, en tenim prou tenint en compte que com f també verifica additivitat, f està unívocament determinat ja que, $U(N)$ és una base de $G(N)$. $\square$

Nota: Podem interpretar el valor de Shapley de la següent manera. Es tracta de la utilitat esperada quan ocurreix el següent:

1. Els jugadors arriben a la coalició en cert moment.

2. Tots els possibles ordres d'arribada són equiprobables.

3. El jugador rep una utilitat igual a la contribució que exerceix en la coalició en el moment en que arriba.

Seguint aquesta idea, podem escriure el valor de Shapley com:

$$\phi_i(N,v) = \frac{1}{n!}\sum_{O\in\pi(N)}(v(B^O(i)\cup\{i\}) - v(B^O(i))),\ \forall i \in N,$$

on $\pi(N)$ és el conjunt d'ordres possibles (permutacions de N) i $B^O(i)$ és el conjunt $\{j \in N : O(j) < O(i)\}$, és a dir, el conjunt dels jugadors que arriben abans que i quan l'ordre d'arribada ve especificat per la permutació O.

Exemple 3.3. Parlarem del joc dels guants. Hi ha 3 jugadors, els dos primers tenen un guant de la mà esquerra i l'altre de la mà dreta. La funció d'utilitat del joc consisteix en $v(S) = 1$ si S conté una parella de guants completa i $v(S) = 0$ en cas contrari.

Tenim doncs el següent valor de Shapley:

O	$Jugador\ 1$	$Jugador\ 2$	$Jugador\ 3$
$(1,2,3)$	0	0	1
$(1,3,2)$	0	0	1
$(2,1,3)$	0	0	1
$(2,3,1)$	0	0	1
$(3,1,2)$	1	0	0
$(3,2,1)$	0	1	0
ϕ	1/6	1/6	2/3

Com hem obtingut els valors anteriors?

Posem-nos en el cas del jugador 1. Aquest, només té contribució marginal igual a 1 quan entra a la coalició en segona posició i el jugador 3 ja hi és, perquè es la única manera de que ell sigui el que provoca que la parella es completi. Quan entra i la parella ja està completa, o quan al entrar ell, no completa la parella, diem que la seva contribució marginal és 0. El mateix passa amb el jugador 2.

En canvi, els jugador 3 té més possibilitats de completar la parella. Com és l'únic jugador que té un guant de la mà dreta, mentre al entrar a una coalició, ja hi sigui algún dels altres dos jugadors, ell serà qui completi la parella i per tant tindrà contribució marginal igual a 1. En canvi, sent el primer en entrar a la coalició, és la única manera que no sigui ell el que completi la parella de guants, i per tant, tindrà contribució marginal igual a 0.

Exemple 3.4. Parlament d'Espanya. Eleccions març 2012.

Tenim $N = \{PP, PSOE, IU\}$ on PP va aconseguir 50 escons, PSOE va aconseguir 47 escons i IU va aconseguir 12 escons (el total són 109 escons). Per poder governar, han de tenir mínim 3/5 parts del total dels escons (3/5 de $109 = 65,4 \approx 66$). El problema el modelem tal que $v(S) = 1$ si la coalició S té majoria per governar i $v(S) = 0$ si no en té. Quedaria així:

S	$\emptyset$	$\{PP\}$	$\{PSOE\}$	$\{IU\}$
$v(S)$	0	0	0	0

S	$\{PP, PSOE\}$	$\{PP, IU\}$	$\{PSOE, IU\}$	$\{PP, PSOE, IU\}$
$v(S)$	1	0	0	1

Llavors anem a calcular les contribucions marginals de cada partit per cada ordre possible i el seu respectiu valor de Shapley:

O	PP	$PSOE$	IU
$(PP, PSOE, IU)$	0	1	0
$(PP, IU, PSOE)$	0	1	0
$(PSOE, PP, IU)$	1	0	0
$(PSOE, IU, PP)$	1	0	0
$(IU, PP, PSOE)$	0	1	0
$(IU, PSOE, PP)$	1	0	0
ϕ	1/2	1/2	0

En aquest cas veiem, que la coalició serà guanyadora, si i només si, PP i PSOE hi pertanyen. Per això, IU no té cap influència en el joc. Dels dos protagonistes, per molt que el PP tingui més escons, si està sol no aconsegueix res; sempre necessita al PSOE.

Nota: $\forall (N, v) \in SG(N)$, el valor de Shapley és una imputació.

Definició 3.5. *Sigui* $(N, v) \in G(N)$ *un joc, direm que el joc és convex si,* $\forall i \in N; S, T \subset N \backslash \{i\}$ *amb* $S \subset T$*,* $v(T \cup \{i\}) - v(T) \geq v(S \cup \{i\})–v(S)$.

Tot joc convex és superadditiu.

Teorema 3.6. *Sigui* $(N, v) \in G(N)$ *un joc convex. Llavors el valor de Shapley del joc pertany al nucli.*

Demostració

Es veu directament del fet que el valor de Shapley és la mitjana artmètica dels vectors de les distribucions marginals.

3.2 Desavantatges

El problema principal que presenta el valor de Shapley és que, com molts investigadors han comentat al llarg dels anys, la manera de calcular el valor només es pot determinar després de calcular totes les possibles coalicions. Què implica això?

Primerament, una immensa quantitat de temps de càlcul. La majoria de vegades, només una solució aproximada és factible. Recordem que per a n jugadors, hi ha

$2^n - 1$ possibles coalicions. Això requereix molta estona de càlcul. I si a això li sumem la possibilitat de que la funció característica no sigui clara, hauríem de jugar amb instàncies aleatòries, fet que augmentaria les variances de les estimacions dels valors de Shapley.

Un altre problema que sol tenir el valor de Shapley és que es sol malinterpretar. Les definicions d'aquest són molt curoses i cal tenir-les clares a l'hora d'aplicar-lo, però no serveix de res saber com es calcula si després no sabem interpretar-lo i donar una resposta a les nostres preguntes.

També veiem, que el valor de Shapley és un valor simple sense model de predicció. Això significa que no serveix quan vols aplicar canvis en les dades d'entrada. Per exemple: "Si canviés de feina a una on guanyés una quantitat y, el valor de Shapley es veuria afectat en un augment de x". Això no ho podem fer.

Per últim, també podríem veure com un problema el fet que sempre necessitem accés a les dades si volem calcular-lo per a unes noves dades inicials. No podem calcular el valor de Shapley fent servir un tipus de funció de predicció.

4 Anàlisi del valor de Shapley i la seva incertesa per a jocs cooperatius

4.1 El problema del càlcul del valor de Shapley

Per abordar els problemes esmentats, alguns investigadors han desenvolupat alguns algoritmes per poder solucionar-los.

Shaheen S. Fatima, Michael Wooldridge i Nicholas R. Jennings van idear un algoritme d'aproximació al valor, basat en la aleatorització, amb una complexitat temporal lineal amb el nombre de jugadors i amb un error d'aproximació baix.

Com hem comentat abans, el valor de Shapley ens dóna una única solució pels jocs cooperatius i s'utilitza per avaluar quina participació té cada jugador dins del joc. Tot i així, hi ha una part d'incertesa associada a aquest valor, que permet una dimensió adicional per poder avaluar la participació de cada jugador al joc. Per tant, els jugadors ja no volen només saber quin es el valor de Shapley, sinó també quina és aquesta incertesa. Per tant, el nostre objectiu seria esbrinar el valor de Shapley, la incertesa i estudiar-ne la relació entre les dues coses i el joc.

Ara bé, com que el problema de determinar el valor de Shapley és P-complet (en la teoria de la complexitat computacional, la classe de complexitat P-complet és un conjunt de problemes de decisió de gran utilitat per identificar problemes que poden ser resolts eficientment en màquines paral·leles), necessitem presentar un nou mètode polinomial temporal aleatori per determinar l'aproximació del valor de Shapley. Amb aquest mètode, comparem aquest valor i el correlacionem amb la seva incertesa de forma que permeti als agents comparar jocs en la base del seu valor de Shapley i la seva incertesa. Aquest estudi mostra que aquesta incertesa primer augmenta amb el valor de Shapley i després disminueix. Això significa que la incertesa està en el seu mínim quan el valor està en el seu màxim i que els agents no sempre han de comprometre el valor de Shapley per tal de reduir la incertesa.

Definició 4.1. *Sigui Ω el conjunt $N \setminus \{i\}$ i $f_i : \Omega \to 2^{N\setminus\{i\}}$ una variable aleatòria que pren valors en el conjunt de tots els subconjunts de $N \setminus \{i\}$, que és interpretada com la tria aleatòria que fa una coalició on s'uneix el jugador i, i on la seva funció de probabilitat és*

$$g\{f_i = S\} = \tfrac{|S|!(n-|S|-1)!}{n!}.$$

Donat un joc (N, v), la incertesa (β_i) d'un jugador $i \in N$ és la variancia de la seva contribució marginal, és a dir, $\beta_i(N, v) = Var(\Delta_i v \circ f_i)$ on $\Delta_i v = v(S \cup \{i\}) - v(S)$.

(Fent servir aquesta nomenclatura, el valor de Shapley d'un jugador $i \in N$ abans comentat, es pot escriure com l'esperança de la seva contribució marginal, és a dir, $\phi_i(N, v) = E(\Delta_i v \circ f_i)$)

Podem veure com $\sum_{S\subseteq N\setminus\{i\}} \frac{|S|!(n-|S|-1)!}{n!} = 1.$

$$\textstyle\sum_{S\subseteq N\setminus\{i\}} g\{f_i = S\} = \sum_{S\subseteq N\setminus\{i\}} \frac{|S|!(n-|S|-1)!}{n!} = \sum_{S\subseteq N} \binom{n}{s} \frac{s!(n-s)!}{n+1!}$$

$$= \sum_{s=0}^{n} \frac{n!}{(n-s)!s!} \frac{s!(n-s)!}{n+1!} = \sum_{s=0}^{n} \frac{1}{n+1} = 1.$$

4.2 Aplicació a jocs de votació ponderada

Definició 4.2. *Sigui $(N, v) \in G(N)$ un joc definit per una imputació $\{w \in \mathbb{R}^n, q \in \mathbb{R}_+\}$ tal que:*

$$v(S) = \begin{cases} 1, \ si \displaystyle\sum_{j \in S} w_j \geq q \\ 0, \ en\ cas\ contrari \end{cases}$$

Aquest joc es representa per $\{N,w,q\}$ o $\{q;w_1,...,w_n\}$ ($n=|N|$) i s'anomena joc de votació ponderada.

Aplicarem ara els càlculs del valor de Shapley a jocs de votació ponderada.

Exemple 4.3. El joc del parlament
Sigui w_i el nombre de vots que té el jugador $i \in N$ i la q el nombre de vots necessaris per guanyar el joc. La contribució marginal de cada jugador és o bé 0 o bé 1.

El problema de determinar el valor de Shapley en un joc de votació ponderada és P-complet. Això fa que calcular aquest valor sigui bastant difícil de trobar en general. Per tal de solucionar-ho, es van proposar dos mètodes: *La simulació de Montecarlo* i el mètode de *funcions generalitzades*. El primer d'aquests fa servir el que en anglès s'anomena *swings* (el *swing* d'un jugador $i \in N$, és un parell de coalicions $(S, S \cup \{i\})$ tal que S és una coalició perdedora i $S \cup \{i\}$ és una coalició guanyadora) de cada jugador com a variable aleatòria amb una distribució donada i defineix el valor de Shapley en termes d'aquesta variable aleatòria. En canvi, el segon mètode no dóna una aproximació del valor de Shapley sinó que és un mètode exacte. El problema és que necesita una gran quantitat d'emmagatzematge i de càlculs i només es pot aplicar a jocs amb quotes i pesos enters.

El nostre mètode és semblant al de *la simulació de Montecarlo* ja que també buscarem una aproximació. Però la diferència és que nosaltres agafarem els pesos com a variables aleatòries. D'aquesta forma evitarem necessitar tants càlculs i emmagatzematge.

Exemple 4.4. Tots els jugadors tenen les mateixes ponderacions.
Agafem un joc de votació ponderada, $\{q; j, ..., j\}$ sobre un parlament amb m partits, cada partit amb j seients. Si $j \geq q$ els jugadors no necesitarien formar cap coalició. En canvi, si $q = mj$ (m és el nombre de jugadors) només seria viable una coalició formada per tots els jugadors. Per tant, la quota satisfà la condició;

$$(j + 1) \leq q \leq j(m - 1).$$

La majoria és decisiva, per tant, $v(S) = 1$ si $w(S) = \sum_{i \in S} w_i = jd \geq q$, on d és el nombre de jugadors que formen S; i $v(S) = 0$ en cas contrari.

Agafem un jugador qualsevol. Aquest jugador es pot unir a una coalició, sent el jugador número i que s'hi uneix ($1 \leq i \leq m$). Per tant, la contribució marginal del jugador serà 1, si i només si, entra a la coalició com a jugador número $[q/j]$. En qualsevol dels altres casos la seva contribució marginal serà 0. Per tant podem calcular el valor de Shapley:

$$\phi_i = 1/m, \quad \forall i \in N.$$

I la incertesa de cada un d'aquests serà la variancia de les seves contribucions marginals:

$$\beta_i = (1 - \phi_i)\phi_i^2 + (1 - \phi_i)^2\phi_i = (1 - \phi_i)\phi_i, \, \forall i \in N.$$

Ara ja ho podem relacionar.

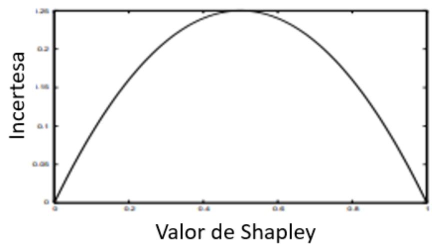


Figura 1: Valor de Shapley vs Incertesa

Veiem a la imatge (el valor de Shapley es mou en l'interval $[0, 1]$) que, com hem dit a l'inici, primer creix i després decreix.

Exemple 4.5. Escollir el capità
Anem a proposar ara un joc on els pesos no són tots iguals, sinó que tots valen 1, menys un que val $k > 1$. El joc l'escribim així $\{q; 1, ...1, k\}$. Consisteix en m jugadors d'un equip i el seu entrenador. Volen escollir qui serà el capità del equip i per tal de fer-ho han de votar; tots tenen un vot menys l'entrenador que en té k. Si ho veiem en coalicions, $v(S) = 1$ si $\sum_{i \in S} w_s \geq q$ i $v(S) = 0$ si $\sum_{i \in S} w_s < q$; és a dir, el nombre de vots necesaris perque algú sigui escollit capità és q. Hem de considerar dos casos: $q \leq m$ o $q > m$.

1. Comencem pel primer cas dels dos, $q \leq m$. Veiem que $k + 1 \leq q \leq m + k - 1$ ja que, si $q \leq k$, llavors l'entrenador pot triar ell sol qui serà el capità i no té sentit que votin els jugadors. Mentre que si $q \geq m + k$, ni que formessin una coalició entre tots els jugadors i l'entrenador arribarien a la quota necessària. També veiem que $2 \leq k \leq q - 1$. Les contribucions marginals possibles pels nostres jugadors i l'entrenador són o bé 1 o bé 0.

Comencem per l'entrenador, qui es pot afegir a una coalició com a i-èsima persona amb $1 \leq i \leq m+1$. La seva contribució marginal serà 1, si i només si, s'uneix com a i-èsima persona però $k-j-1 \leq i \leq q$; és a dir, que no tinguessin majoria sense ell, pero amb ell sí que la tinguessin. Això passarà en k casos. Per tant el valor de Shapley del entrenador, definit per ϕ_e, és:

$$\phi_e = k/(m+1).$$

Ara considerem qualsevol dels jugadors, on la seva contribució marginal serà 1 en dos casos: si s'uneix a una coalició que conté ja l'entrenador i ho fa com a $(q-k+1)$-èsim jugador, que serien $(q-k)(m-1)!$ possibilitats de les $(m-1)!$ possibles permutacions; i l'altre opció és que s'uneixi a una coalició que ja contingui $(q-1)$ jugadors, és a dir, que s'uneixi com a q-èsim membre, i de permutacions complint aquesta condició n'hi ha $(m-q+1)(m-1)!$. El seu valor de Shapley, definit per ϕ_j, serà doncs:

$$\phi_j = (q-k+m-q+1)(m-1)!/(m+1)! = (m-k+1)/m(m+1).$$

Fent servir ara la definició de incertesa, siguin β_e i β_j les seves respectives incerteses, tenim:

$$\beta_e = (1-k/(m+1))(k/(m+1)).$$

$$\beta_j = (1-(m-k+1)/m(m+1))((m-k+1)/m(m+1)).$$

2. Anem ara al cas $q > m$. Els intervals de q i de la k són els mateixos que en l'apartat 1. Anem a parlar del cas del entrenador, qui pot afegir-se a una coalició com a i-èsima persona amb $1 \leq i \leq m+1$, tot i que novament la seva contribució marginal serà 1, si i només si, s'uneix com a i-èsima persona però $k-j-1 \leq i \leq q$, cosa que passarà en k casos. Novament:

$$\phi_e = k/(m+1).$$

En el cas dels jugadors, com $q > m$, ara ja només hi ha un cas en el que la contribució marginal d'aquests jugadors serà 1: quan s'uneixi a una coalició que ja conté al entrenador com a $(q-k+1)$-èssim membre. Són $(q-1)(m-1)!$ combinacions de les $(m+1)!$ permutacions possibles. Per tant, tindrem:

$$\phi_j = (q-k)/m(m+1).$$

Per tant, finalment tindrem:

$$\beta_e = (1-k/(m+1))(k/(m+1)).$$

$$\beta_j = (1-(q-k)/m(m+1))((q-k)/m(m+1)).$$

Com que la relació entre el valor de Shapley i la seva incertesa és la mateixa que en l'exemple anterior, el gràfic d'abans també ens serveix: la incertesa té el seu valor més alt quan el valor de Shapley és 1/2. Creix el principi a mesura que avança ϕ i disminueix després.

Exemple 4.6. Anem ara a comentar l'opció on hi ha alguns jugadors amb ponderació gran i la resta amb ponderació petita. Dels primers en tindrem pm i dels segons $(1-p)m$ amb $p \in \mathbb{R}$, $0 \leq p \leq 1$. La quota en aquests casos serà q complint, $(k+1) \leq q \leq pmk + (1-p)m - 1$ i pel que fa a k, $2 \leq k \leq (q-1)$. Podríem pensar una situació en una empresa de m treballadors on han de triar quina opció de reforma els agrada més entre algunes opcions possibles. Entre tots aquests, n'hi ha pm que com la seva posició en l'empresa és més gran, tenen k vots; i els $(1-p)m$ restants, fa poc que han arribat a l'empresa i només tenen 1 vot $(0 \leq p \leq 1)$. Per tant, com a la resta d'exemples, tindrem $v(S) = 1$ si S té més de q vots, o 0 altrament.
Aquí és on farem servir un **Mètode Aleatoritzat del valor de Shapley**, i per tal de trobar-lo el que farem serà considerar una mostra de l'empresa, és a dir, de la població de jugadors, una gran coalició de tamany M. Sigui $\hat{p}$ la proporció de la mostra que té k vots. Independentment de com estigui distribuïda la població, la proporció de jugadors amb "vot gran"en una mostra de tamany M es distribueix segons una normal amb $\mu = p$, i $\sigma^2 = p(1-p)/M$:

$$\hat{p} \sim N(p, p(1-p)/M).$$

A partir, d'això obtenim el valor de Shapley de la següent manera: la contribució marginal en aquesta mostra aleatòria és 1 si el pes de la mostra (en el moment en que s'afageix) és inferior a q però és més gran que $(q-k)$. Altrament és 0. Sabem que la ponderació mitjà d'aquesta mostra és $\hat{p}Mk + (1-\hat{p})M$. Sigui a la proporció de treballadors amb vot gran, necessaris perquè la mitja de la mostra aleatòria sigui $(q-k)$ (seria amb $a = (q-k-M)/(M(k-1))$). Sigui b la proporció de treballadors amb vot gran, necessaris perquè la mitja de la mostra aleatòria sigui $(q-\epsilon)$ (seria amb $b = (q-\epsilon-M)/(M(k-1))$). La contribució marginal esperada d'un treballador amb vot gran de la mostra aleatòria és l'àrea de la curva definida per la distribució normal esmentada anteriorment entre els límits a i b:

$$\Delta_g^M = \tfrac{1}{x\sqrt{2\pi\sigma^2}} \int_a^b e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Per tant, definint el valor de Shapley d'un treballador amb vot gran per ϕ_g, aquest valor és:

$$\phi_g = \tfrac{1}{m} \textstyle\sum_{M=1}^m \Delta_g^M.$$

I definint la seva incertesa per β_g, aquesta és:

$$\beta_g = \tfrac{1}{m} \textstyle\sum_{M=1}^m (\Delta_g^M - \phi_g)^2.$$

Anem ara al cas dels treballadors amb vot petit. La seva contribució marginal, com sempre serà 1 si al afegir-se a la mostra és inferior a q, però més gran o igual a $(q-1)$; altrament és 0. Sabem que la ponderació mitjana de la mostra és $\hat{p}Mk + (1-\hat{p})M$. Sigui c la proporció de treballadors amb vot gran perquè la ponderació mitjana sigui $(q-1)$ (seria amb $c = (q-1-M)/(M(k-1))$). Sigui també d la proporció de treballadors amb vot gran perquè la ponderació mitjana sigui $(q-\epsilon)$ (seria amb $d = (q-\epsilon-M)/(M(k-1))$). La contribució marginal del treballador amb vot petit serà doncs l'àrea de la curva definida per la distribució normal d'abans però ara, amb límits c i d:

$$\Delta_p^M = \frac{1}{x\sqrt{2\pi\sigma^2}} \int_c^d e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Per tant, definint el valor de Shapley d'un treballador amb vot petit per ϕ_p, aquest valor és:

$$\phi_p = \frac{1}{m} \sum_{M=1}^{m} \Delta_p^M.$$

I definint la seva incertesa per β_p, aquesta és:

$$\beta_p = \frac{1}{m} \sum_{M=1}^{m} (\Delta_p^M - \phi_p)^2.$$

Per acabar l'exemple, anem a comparar el valor de Shapley i la seva Incertesa en cada un dels casos, si variem les ponderacions (k), la quota (q) o el nombre de treballadors o jugadors (m). Cal variar k, p, m seguint:

1. Cap jugador pot guanyar sol. $(k < q)$

2. La quota és més petita que el nombre total de possibles vots. $(q < mpk + (1-p)m)$

Per això veiem les sis imatges següents:

Considerem la primera, amb $m = 200, q = 200$, que és un gràfic dels treballadors amb vot gran on variem les ponderacions. Per cada valor de k, la figura ensenya el valor de Shapley i la seva incertesa per $p \in (0'1, 0'9)$. La segona és per treballadors amb vot petit.

Si mirem la tercera, amb $m = 200, j = 5$, és un gràfic dels treballadors amb vot gran on variem la quota. Per cada valor de q, la figura ensenya el valor de Shapley i la seva incertesa $p \in (0'1, 0'9)$. La quarta és per treballadors amb vot petit.

Finalment en la cinquena, amb $q = 25, j = 5$, és un gràfic dels treballadors amb vot gran on variem el nombre de treballadors (jugadors). Per cada valor de m, la figura ensenya el valor de Shapley i la seva incertesa $p \in (0'1, 0'9)$. La sisena és per treballadors amb vot petit.

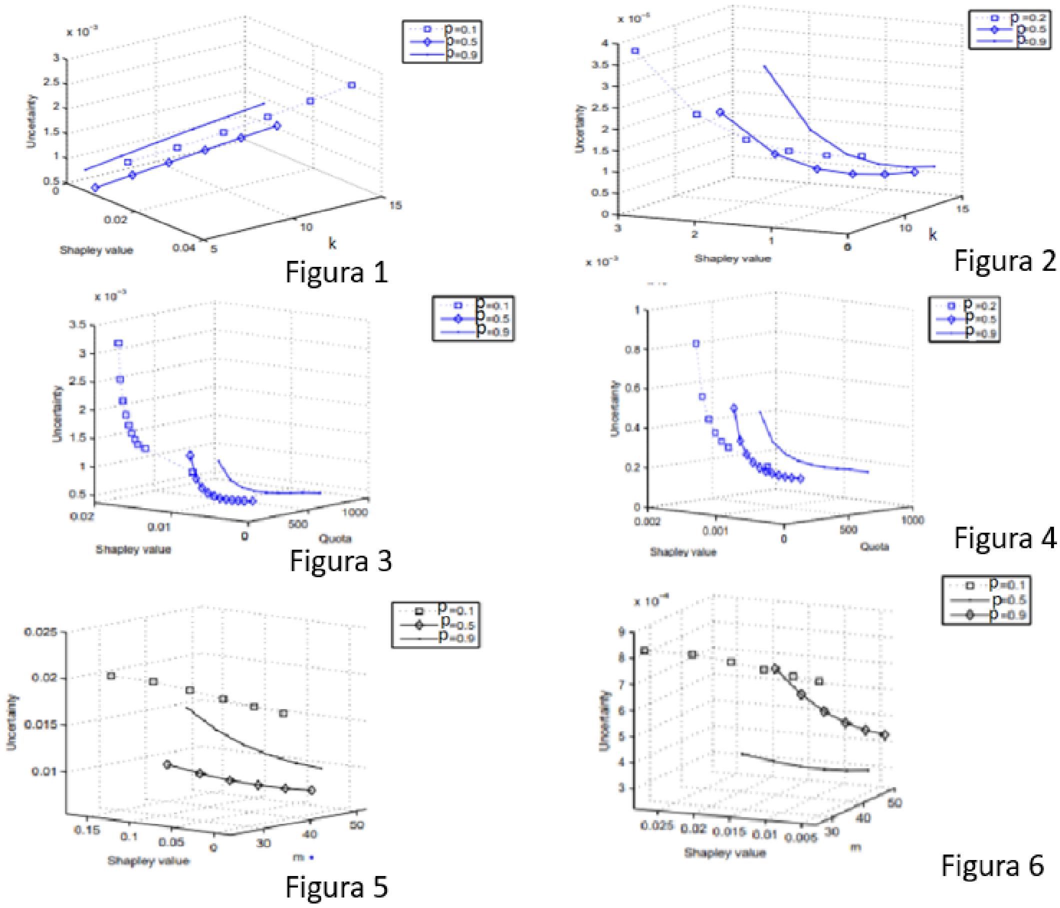


Figura 2: Valor de Shapley vs Incertesa

Veiem doncs que amb dos tipus de jugadors, variant k, p, q i m i satisfent les condicions 1 i 2, la incertesa augmenta amb el valor de Shapley.

Cal notar que els exemples de dalt són amb un mínim de 20 treballadors i amb més d'un treballador de cada tipus. Això implica, que el valor de Shapley és inferior a 0,5 i la relació amb la incertesa actua com a la primera figura de l'exemple anterior.

Exemple 4.7. Més de dos tipus

Considerem ara més de dos tipus de jugadors amb w_i el pes del jugagor $i \in N$. Sigui m el nombre de jugadors i q la quota, el joc serà de la forma $\{q; w_1, w_2, ..., w_m\}$. Considerem una població de jugadors en la que la ponderació de cada jugador segueix una distribució normal estandaritzada - $N(0,1)$. Abans, no hem necessitat que la població seguís una distribució normal ja que només teníem dos tipus de jugadors. Per tant, tot i que l'exemple d'abans és vàlid per qualsevol tipus de distribució de la població, els resultats d'aquesta secció seran només per quan segueix una distribució normal. Com que aquesta distribució permet ponderacions negatives, la convertim en $N(4,1)$. Com hem dit a la Definició 4.1, el valor de Shapley és l'esperança de la seva contribució marginal a alguna coalició al atzar. Per tant, per tal de determinar aquest valor per a la població esmentada anteriorment ($N(4,1)$), farem servir la següent regla de la Teoria de les Mostres que serveix per a la distribució normal.

Per a una distribució normal amb mitjana μ i amb desviació típica σ, tenim que la suma de les ponderacions dels m jugadors a la mostra té distribució $N(m\mu, m\sigma^2)$. Per tant, per la nostra distribució, la suma de les ponderacions dels jugadors en una mostra aleatòria de tamany m vindrà donada per $N(4m, m)$ Fem servir aquesta regla per determinar el valor de Shapley:

Apliquem novament el Mètode Aleatoritzat del valor de Shapley en aquest apartat. Per un jugador i amb ponderació w_i, anomenarem ϕ_i al seu valor de Shapley i β_i la seva incertesa. Sigui M el tamany d'aquesta mostra aleatòria en la qual la distribució de les ponderacions individuals dels jugadors segueix una $N(4,1)$. La contribució marginal del jugador i a aquesta mostra aleatòria és 1, si la ponderació total dels M jugadors de la mostra és més gran o igual a $q - w_i$ però inferior a q. Altrament, la seva contribució marginal és 0. Per tant, l'esperança de la distribució marginal del jugador i, representada per Δ_i^M, a la coalició de la mostra, és l'àrea sota la curva definida per $N(4M, M)$ a l'interval $[q - w_i, q - \epsilon]$, és a dir,

$$\Delta_i^M = \frac{1}{x\sqrt{2\pi\sigma}} \int_{q-\epsilon}^{q-w_i} e^{-\frac{(x-4M)^2}{2M}} dx.$$

i el seu valor de Shapley és:

$$\phi_i = \frac{1}{n} \sum_{M=1}^{m} \Delta_i^M.$$

La complexitat d'aquest mètode és $O(m)$. A més, les dues fonts d'incertesa són M i ϵ. Aquesta incertesa disminueix amb M i augmenta amb ϵ. Es pot calcular com:

$$\beta_i = \frac{1}{n} \sum_{M=1}^{m} (\phi_i - \Delta_i^M)^2.$$

Pel cas de més de dos tipus de jugadors, definim les condicions a partir de q i w_i (amb $1 \leq i \leq m$) següents:

1. Cap jugador pot guanyar sol. ($w_i < q$ per $1 \leq i \leq m$)

2. El nombre de jugadors necessaris per guanyar és menys de m.

Fem servir les equacions d'abans variem els paràmetres $q, w_i(1 \leq i \leq m)$ i m complint aquestes condicions, i determinem la relació del valor de Shapley i la incertesa.

Si mirem la primera imatge, per cada q, les ponderacions varien entre 1 i $q-1$. La incertesa primer augmenta amb el valor de Shapley i després decreix. A les següents figures posem $m = 50$ i $m = 100$ respectivament. La incertesa d'un jugador primer augmenta amb el valor de Shapley i després disminueix.

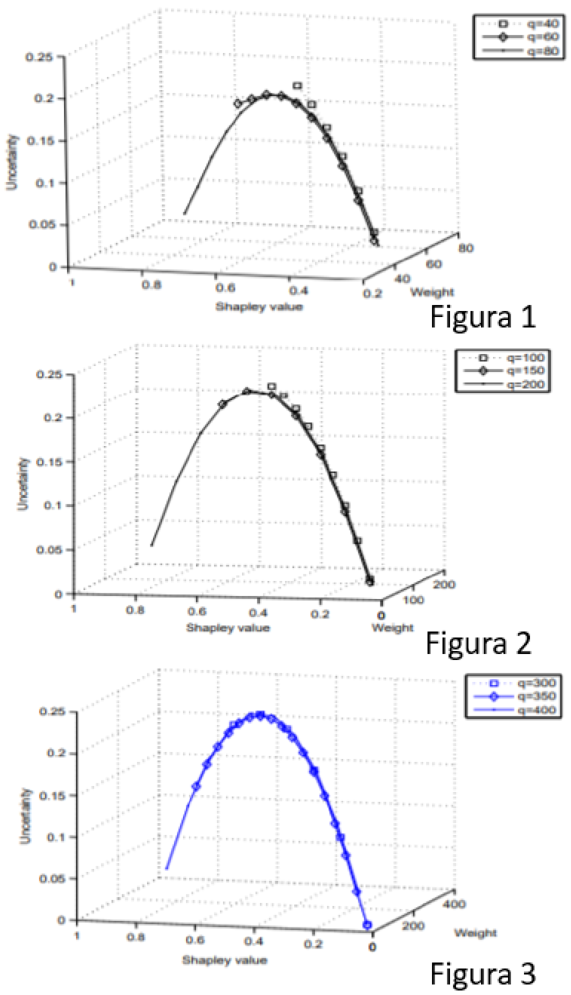


Figura 3: Valor de Shapley vs Incertesa

Resumint, hem vist que la incertesa està al seu mínim, quan el valor de Shapley està al seu màxim i els agents no sempre han de comprometre el seu valor per tal de reduir la incertesa. Això és degut a que si el valor de Shapley és entre $0,5$ i 1, un increment del valor està associat amb un decreixement de la incertesa.

4.3 Conclusions

Tot i que el valor de Shapley, aconseguit a partir de les aproximacions, és una solució única que ens dóna la informació de quant ha aportat un jugador en el tema que s'estigui tractant, també té una certa quantitat d'incertesa associada. Per tant, ens interessa saber quina és aquesta incertesa i no només quin és el valor de Shapley. Si obtenim uns resultats bons, però la incertesa és alta, seria com si no haguéssim obtingut informació.

En aquest apartat hem vist com calcular els dos valors de forma aproximada i determinar-ne la relació entre ells. Com hem dit, hem pogut observar, que de forma general, la incertesa primer creix amb el valor de Shapley, i després decreix.

5 Càlcul polinomial del valor de Shapley basat en el mostreig

5.1 Introducció

Seguint amb altres estudis sobre el valor de Shapley, Javier Castro, Daniel Gómez i Juan Tejada, van publicar un article l'abril del 2018 sobre aquest. L'objectiu del seu treball, va ser desenvolupar un algoritme eficient que pogués estimar el valor de Shapley per una gran classe de jocs.

Consisteix en un mètode polinomial basat en la teoria de mostres que pot ser utilitzat per estimar el valor de Shapley per a jocs cooperatius. El mostreig, és un procés que consisteix en extreure un grup representatiu d'individus, d'una població particular. S'utilitza en circumstàncies on no és pràctic obtenir informació de cada membre de la població com és el cas quan tractem del valor de Shapley.

Abans d'entrar en l'algoritme, hi ha alguns conceptes que ens presenten i que caldria esmentar, com per exemple una manera alternativa de calcular el valor de Shapley. Sigui N el nombre de jugadors, definim $O : \{1, ..., n\} \to \{1, ..., n\}$ com la permutació que assigna a cada posició k, el jugador $O(k)$. Denotem per $\pi(N)$ le sèrie de totes les possibles permutacions d'N jugadors ($\pi(N)$ te cardinalitat $n!$). Donada una permutació O, denotem per $Pre^i(O)$ la sèrie de predecessors del jugador i en l'ordre O ($Pre^i(O) = \{O(1), ..., O(k-1)\}$ si $i = O(k)$). Per tant el valor de Shapley es pot expressar de la següent manera:

$$\phi_i = \sum_{0\in\pi(N)} \tfrac{1}{n!}(v(Pre^i(O) \cup \{i\}) - v(Pre^i(O))),\ i = 1, ..., n.$$

Podem generalitzar la definició del valor de Shapley al cas de jugadors amb diferents ponderacions.

Definició 5.1. *Sigui $w = (w_1, ..., w_n)$ un vector amb les ponderacions de cada jugador $i \in N$ amb $w_i > 0$ i O una permutació, una funció de probabilitat que representa la probabilitat dels diferents ordres de $\pi(N)$ es defineix com:*

$$P_w(O) = \textstyle\prod_{k=1}^{n}(w_{O(k)}/\sum_{l=1}^{k} w_{O(l)}).$$

A partir d'aquí, definim el valor de Shapley ponderat com:

$$\phi_i = \sum_{0\in\pi(N)} P_w(O)(v(Pre^i(O) \cup \{i\}) - v(Pre^i(O)))\ i = 1, ..., n.$$

Nota. En la definició clàssica del valor de Shapley, se suposa que tots els diferents ordres tenen la mateixa probabilitat.

En el procés de mostreig, denotarem una població finita per $P = (a_1, ...a_p)$ on a_i és la unitat de mostreig. La característica observada per a cada unitat de mostreig es denotarà per $x(a_i) = (x(a_i)_1, ..., x(a_i)_r)$. $\theta = (\theta_1, ...\theta_r)$ serà un vector de paràmetres que volem estimar i la seva estimació serà $\hat{\theta} = (\hat{\theta_1}, ..., \hat{\theta_r})$.

En aquest estudi, es considerarà una mostra M com a element de $\underbrace{P \times ... \times P}_{m\ vegades}$, és a dir, una mostra amb reposició.

5.2 L'estimació del valor de Shapley

L'objectiu d'aquesta secció és obtenir una estimació del valor de Shapley i algunes propietats d'aquest, utilitzant un únic procés de mostreig per tots els jugadors $i \in N$. Anomenarem aquest procés **Mètode de mostreig aleatori** i el definirem de la següent manera:

1. La població del procés P serà la sèrie de tots els possibles ordres de N jugadors, $P = \pi(N)$.

2. El vector de paràmetres que estudiarem serà $\phi = (\phi_1, ..., \phi_n)$ on n és el nombre de jugadors.

3. Les característiques observades per cada unitat de mostreig $O \in \pi(N)$, són les contribucions marginals dels jugadors en l'ordre O, és a dir,

$$x(O) = (x(O)_1, ..., x(O)_n) \text{ on } x(O)_i = v(Pre^i(O) \cup \{i\}) - v(Pre^i(O)).$$

4. L'estimació del paràmetre ϕ, $\hat{\phi}$, serà la mitjana de les contribucions marginals de la mostra M, és a dir

$$\hat{\phi} = (\hat{\phi}_1, ..., \hat{\phi}_n) \text{ on } \hat{\phi}_i = \frac{1}{m} \sum_{O \in M} x(O)_i.$$

5. El procés de selecció per determinar la mostra, agafarà qualsevol ordre $O \in \pi(N)$ amb probabilitat $\frac{1}{n!}$.

Algoritme de càlcul

Inici del bucle
    Determinem m
    $Contador := 0$ i $\hat{\phi}_i := 0$
    Inici del bucle
        Agafem $O \in \pi(N)$ amb probabilitat $\frac{1}{n!}, \forall i \in N$
        Inici del bucle
            Calculem $Pre^i(O)$
            Calculem $x(O)_i := v(Pre^i(O) \cup \{i\} - v(Pre^i(O))$
            $\hat{\phi}_i : \hat{\phi}_i + x(O)_i$
        Final del bucle
        $Contador := Contador + 1$
    Final del bucle
$\hat{\phi}_i := \frac{\hat{\phi}_i}{m} \; \forall i \in N$
Final del bucle

Exemple 5.2. Anem a veure un exemple, fent servir el mètode de mostreig aleatori.

Considerem un joc (N, v) de 3 jugadors tal que v segueix la següent taula:

S	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$v(S)$	0	1	3	0	5	7	4	10

Hi ha 6 possibles ordres (3!). Anem a calcular les contribucions marginals de cada jugador per cada un d'aqueststs ordres.

O	$x^1(O)$	$x^2(O)$	$x^3(O)$
$(1,2,3)$	1	4	5
$(1,3,2)$	1	3	6
$(2,1,3)$	2	3	5
$(2,3,1)$	6	3	1
$(3,1,2)$	7	3	0
$(3,2,1)$	6	4	0

El valor de Shapley el calculem fent la mitjana de cada una d'aquestes columnes. Per tant $\phi(v) = (\frac{23}{6}, \frac{20}{6}, \frac{17}{6})$

Veiem que compleix la propietat d'eficiència:

$$\tfrac{23}{6} + \tfrac{20}{6} + \tfrac{17}{6} = \tfrac{60}{6} = 10 = v(\{1,2,3\}).$$

Anem ara a utilitzar el mètode per trobar el valor de Shapley aproximat que hem definit. Triarem $m = 3$, és a dir, agafarem 3 dels 6 posssibles ordres que hem comentat, per exemple, $(1,2,3), (1,3,2)$ i $(3,2,1)$. La taula que defineix les contribucions marginals és:

O	$x^1(O)$	$x^2(O)$	$x^3(O)$
$(1,2,3)$	1	4	5
$(1,3,2)$	1	3	6
$(3,2,1)$	6	4	0

Per tant, tenim que la nostra aproximació serà, $\hat{\phi}(v) = (\frac{8}{3}, \frac{11}{3}, \frac{11}{3})$.

Veiem que clarament difereix molt respecte al valor de Shapley real. Això és degut a la baixa quantitat d'ordres que fem servir. Per això, en casos on el càlcul del valor de Shapley és senzill, no caldria usar-lo. En canvi, en els exemples que veurem a continuació, és molt pràctic.

Podem veure que també compleix la propietat d'eficiència:

$$\tfrac{8}{3} + \tfrac{11}{3} + \tfrac{11}{3} = \tfrac{30}{3} = 10 = v(\{1,2,3\}).$$

Veiem abans algunes propietats:

Proposició 5.3. *L'estimador $\hat{\phi}_i$ és no esbiaixat i la seva variança ve donada per*

$$Var[\hat{\phi}_i] = \tfrac{\sigma^2}{m} \text{ on } \sigma^2 = \textstyle\sum_{O \in \pi(N)} (x(O)_i - \phi_i)^2 \frac{1}{n!}$$

Demostració
La demostració és directa tenint en compte que, l'estimador $\hat{\phi}_i$ és una mitjana de la mostra i ϕ_i és la mitjana de la població. □

Corol.lari 5.4. *L'estimador $\hat{\phi}_i$ és consistent en probabilitat, és a dir,*

$$\lim_{x\to\infty} P(\left|\hat{\phi}_i - \phi_i\right| > \epsilon) = 0, \forall \epsilon > 0.$$

Corol.lari 5.5. *Si un jugador nul per un joc $(N, v) \in G(N)$ $(\forall S \subset N, v(S \cup \{i\}) = v(S))$, llavors $\hat{\phi}_i = \phi_i$.*

Demostració
La demostració és directa tenint en compte que el valor de $x(O)_i$ és constant $\forall O \in \pi(N)$. □

Proposició 5.6. *L'estimador $\hat{\phi}_i$ és eficient en la imputació, és a dir,*

$$\sum_{i=1}^{|N|} \hat{\phi}_i = v(N).$$

Demostració
La demostració és directa tenint en compte que la suma de les contribucions marginals en qualsevol ordre és $v(N)$. □

Per acabar la secció, determinarem el tamany de la mostra m. Per tal d'aconseguir-ho, volem garantir que l'error de l'estimació del procés és inferior a e amb probabilitat major a $1 - \alpha$. Seguint el teorema central del límit, obtenim que l'estimador $\hat{\phi}_i \sim N(\phi_i, \sigma^2/m)$ i per tant, si $m \geq Z^2_{\alpha/2}\sigma^2/e^2$, llavors $P(\left|\hat{\phi}_i - \phi_i\right| \leq e) \geq 1 - \alpha$, amb $Z_{\alpha/2}$ sent el valor tal que $P(Z \geq Z_{\alpha/2}) = \alpha/2$ i $Z \sim N(0, 1)$.

Tenint en compte que σ^2 és desconegut, és necessari donar una cota superior d'aquest valor per tal de determinar el tamany de la mostra. Per acotar σ^2, definim el màxim i el mínim que pot assolir la variable $\hat{\phi}_i$, és a dir, $x^i_{max} = max_{S\subset N\setminus\{i\}} x(S)$ i $x^i_{min} = min_{S\subset N\setminus\{i\}} x(S)$. Ara, podem observar que per qualsevol variable aleatòria acotada per x^i_{min} i x^i_{max}, la màxima variança s'assoleix quan aquesta variable pren els dos extrems amb la mateixa probabilitat $1/2$, per tant es compleix la següent desigualtat:

$$\sigma^2 \leq \frac{1}{2}(x^i_{max} - \frac{x^i_{max}+x^i_{min}}{2})^2 + \frac{1}{2}(x^i_{min} - \frac{x^i_{max}+x^i_{min}}{2})^2 = \frac{(x^i_{max}-x^i_{min})^2}{4}.$$

Nota. Observem que és possible calcular el valor de Shapley ponderat amb vector de ponderacions w de forma semblant, modificant les probabilitats dels diferents ordres de la següent manera:

$$P_w(O) = \prod_{k=1}^{n}(w_{O(k)}/\sum_{l=1}^{k} w_{O(l)}) \text{ per cada } O \in \pi(N).$$

5.3 Casos particulars

Ara el que farem és aproximar el valor de Shapley per alguns jocs coneguts en els quals el valor pot ser calculat de forma polinomial temporal per qualsevol coalició. Per mesurar l'error de l'aproximació de la mostra farem servir alguns exemples pels quals sabem el valor de Shapley.

Agafarem $\alpha = 0,01$ per calcular la cota d'error teòric (e_{th}). Recordem que aquest error representa un valor que satisfà: $P(\left|\hat{\phi}_i - \phi_i\right| \leq e_{th}) \geq 1 - \alpha = 0.99$.

Exemple 5.7. Joc simètric

Sigui $(N, v) \in G(N)$ un joc on $N = \{1, ..., 1000\}$ tal que $\forall S \subset N$,

$$v(S) = \begin{cases} 1, & \text{si } |S| > 500 \\ 0, & \text{altrament} \end{cases}$$

Tenint en compte que els jugadors d'aquest joc són simètrics, és molt fàcil veure que el valor de Shapley per a cada un d'ells és $0,001$. Per veure-ho només cal entendre que la contribució marginal d'un jugador serà 1 només si és el participant número 501 de la coalició; i altrament serà 0. Això passarà una de cada mil vegades i per tant, tenim que el valor de Shapley per a cada jugador és $\frac{1}{1000}$.

La taula següent ens mostra la cota de l'error teòric per aquesta estimació. Per calcular la cota d'aquest error teòric hem fet servir $\sigma^2_{\phi_i} \leq (x^i_{max} - x^i_{min})^2/4 = 1/4$.

m	1000	10000	100000	1000000	10000000	100000000
e_{th}	0.0407	0.0129	0.00407	0.00129	0.000407	0.000129

Anem a mostrar el càlcul del primer error teòric. Sigui $m = 10^3$, $\sigma^2_{\phi_i} \leq \frac{1}{4}$ i $\alpha = 0,01$. Seguint que $P(Z \geq Z_{\alpha/2}) = 0.005$ on $Z \sim N(0,1)$ trobem que $Z_{\alpha/2} = 2.575$ Per tant, seguint la nostra fórmula, $m \geq Z^2_{\alpha/2}\sigma^2/e^2_{th}$, trobem $e^2_{th} \geq 0.0407143$

Exemple 5.8. Joc no simètric

Sigui $(N, v) \in G(N)$ un joc sobre la votació dels Estats Units on $N = \{1, ..., 51\}$ i les seves respectives ponderacions són:

$$w = \{w_1, ..., w_{51}\}=\{45, 41, 27, 26, 26, 25, 21, 17, 17, 14, 13, 13, 12, 12, 12, 11, \underbrace{10, ..., 10}_{4\, vegades},$$
$$\underbrace{9, ..., 9}_{4\, vegades}, 8, 8, \underbrace{7, ..., 7}_{4\, vegades}, \underbrace{6, ..., 6}_{4\, vegades}, 5, \underbrace{4, ..., 4}_{9\, vegades}, \underbrace{3, ..., 3}_{7\, vegades}\}.$$

Per estimar el valor de Shapley, Owen utilitza l'extensió multilineal. La taula següent (figura 1) mostra el valor de Shapley real, i l'aproximació donada per l'extensió multilineal i el procés de mostreig. Pel que fa a l'error teòric, l'hem obtingut tenint en compte $\sigma^2_{\phi_i} \leq (x^i_{max} - x^i_{min})^2/4 = 1/4$ i per tant, coincideix amb la taula anterior.

Player	True Shapley	Multilinear	Sampling
1	0.08831	0.08852	0.088257
2	0.07973	0.07976	0.079723
3	0.05096	0.05113	0.050952
4	0.04898	0.04915	0.048959
5	0.04898	0.04915	0.048976
6	0.04700	0.04716	0.047010
7	0.03917	0.03928	0.039199
8	0.03147	0.03157	0.031455
9	0.03147	0.03157	0.031420
10	0.02577	0.02586	0.025759
11	0.02388	0.02396	0.023847
12	0.02388	0.02396	0.023859
13	0.02200	0.02207	0.022037
14	0.02200	0.02207	0.021991
15	0.02200	0.02207	0.022018
16	0.02013	0.02019	0.020128
17	0.01827	0.01833	0.018281
18	0.01827	0.01833	0.018278
19	0.01827	0.01833	0.018277
20	0.01827	0.01833	0.018262
21	0.01641	0.01647	0.016452
22	0.01641	0.01647	0.016417
23	0.01641	0.01647	0.016424
24	0.01641	0.01647	0.016401
25	0.01456	0.01461	0.014572
26	0.01456	0.01461	0.014577
27	0.01272	0.01276	0.012716
28	0.01272	0.01276	0.012716
29	0.01272	0.01276	0.012703
30	0.01272	0.01276	0.012736
31	0.01088	0.01092	0.010905
32	0.01088	0.01092	0.010901
33	0.01088	0.01092	0.010871
34	0.01088	0.01092	0.010888
35	0.009053	0.009078	0.0090567
36	0.007230	0.007243	0.0072156
37	0.007230	0.007243	0.0072248
38	0.007230	0.007243	0.0072389
39	0.007230	0.007243	0.0072208
40	0.007230	0.007243	0.0072332
41	0.007230	0.007243	0.0072324
42	0.007230	0.007243	0.0072160
43	0.007230	0.007243	0.0072478
44	0.007230	0.007243	0.0072206
45	0.005412	0.005431	0.0054173
46	0.005412	0.005431	0.0054107
47	0.005412	0.005431	0.0054221
48	0.005412	0.005431	0.0054045
49	0.005412	0.005431	0.0054231
50	0.005412	0.005431	0.0054140
51	0.005412	0.005431	0.0054084

Figura 4: Taula Valor de Shapley

m	1000	10000	100000	1000000	10000000	100000000
e_{th}	0.0407	0.0129	0.00407	0.00129	0.000407	0.000129

Uns altres tipus d'errors que podríem calcular en aquest cas per cada m és l'error mitjà i el màxim error. Calcularíem el valor absolut de la diferència entre el valor de Shapley real i el valor de Shapley obtingut pel mètode de mostreig aleatori per cada un dels 51 (en aquest cas) jugadors. Pel que fa a l'error màxim, buscaríem entre aquestes 51 diferències, quina es la màxima; i en quant al error mitjà, sumaríem totes les diferències i dividiríem el resultat entre el nombre de jugadors (51). Per exemple, en el cas de $m = 10^8$ de la taula, tindríem que l'error màxim és $e_{max} = 0,000115$ i el mitjà és $e_{avg} = 0,00003$.

És important veure que, per mostres de tamany gran com la nostra, l'estimació obtinguda de la mostra és millor que la que dóna Owen, ja que el màxim error i l'error mitjà de l'extensió multilineal és superiors en els dos casos.

Exemple 5.9. El joc de l'aeroport

Sigui (N, c) un joc on $N = \{1, ..., 100\}$ definit tal que $\forall S \subset N, c(S) = max_{i \in S} c_i$, on:

$$c = \{c_1, ..., c_{100}\} = \{\underbrace{1, ..., 1}_{8\, vegades}, \underbrace{2, ..., 2}_{12\, vegades},$$

$$\underbrace{3, ..., 3}_{6\, vegades}, \underbrace{4, ..., 4}_{14\, vegades}, \underbrace{5, ..., 5}_{8\, vegades}, \underbrace{6, ..., 6}_{9\, vegades}, \underbrace{7, ..., 7}_{13\, vegades}, \underbrace{8, ..., 8}_{10\, vegades}, \underbrace{9, ..., 9}_{10\, vegades}, \underbrace{10, ..., 10}_{10\, vegades}\}.$$

Les dues taules d'aquest apartat tracten el valor de Shapley i l'error respectivament. La primera (figura 2), mostra el valor de Shapley real i l'aproximació donada pel procés de mostreig i l'extensió multilineal ($m = 10^8$), fent servir $\sigma^2_{\phi_i} \leq (x^i_{max} - x^i_{min})^2/4 = \frac{10^4}{4} = 25$.

Player	Shapley	Sampling	Player	Shapley	Sampling
1	0.01	0.01000240	26	0.033369565	0.03336215
2	0.01	0.01002122	27	0.046883079	0.04681264
3	0.01	0.00999404	28	0.046883079	0.04693904
4	0.01	0.01001061	29	0.046883079	0.04686664
5	0.01	0.00998937	30	0.046883079	0.04683868
6	0.01	0.01000476	31	0.046883079	0.04688131
7	0.01	0.01000974	32	0.046883079	0.04694949
8	0.01	0.00998920	33	0.046883079	0.04686780
9	0.020869565	0.02086825	34	0.046883079	0.04687685
10	0.020869565	0.02086269	35	0.046883079	0.04681076
11	0.020869565	0.02085620	36	0.046883079	0.04693215
12	0.020869565	0.02087864	37	0.046883079	0.04683531
13	0.020869565	0.02089870	38	0.046883079	0.04687293
14	0.020869565	0.02086929	39	0.046883079	0.04688354
15	0.020869565	0.02089350	40	0.046883079	0.04686851
16	0.020869565	0.02087001	41	0.063549745	0.06363301
17	0.020869565	0.02086078	42	0.063549745	0.06354695
18	0.020869565	0.02085292	43	0.063549745	0.06358871
19	0.020869565	0.02085290	44	0.063549745	0.06351779
20	0.020869565	0.02087947	45	0.063549745	0.06346493
21	0.033369565	0.03335218	46	0.063549745	0.06350608
22	0.033369565	0.03335009	47	0.063549745	0.06354124
23	0.033369565	0.03339154	48	0.063549745	0.06356064
24	0.033369565	0.03329275	49	0.082780515	0.08280608
25	0.033369565	0.03338934	50	0.082780515	0.08276462
51	0.082780515	0.08284313	76	0.139369662	0.13940729
52	0.082780515	0.08268617	77	0.139369662	0.13934942
53	0.082780515	0.08276607	78	0.139369662	0.13951629
54	0.082780515	0.08285130	79	0.139369662	0.13936278
55	0.082780515	0.08290920	80	0.139369662	0.13935631
56	0.082780515	0.08279981	81	0.189369662	0.18945158
57	0.082780515	0.08277814	82	0.189369662	0.18946746
58	0.106036329	0.10594986	83	0.189369662	0.18939991
59	0.106036329	0.10600076	84	0.189369662	0.18936684
60	0.106036329	0.10592028	85	0.189369662	0.18945528
61	0.106036329	0.10605632	86	0.189369662	0.18939632
62	0.106036329	0.10598094	87	0.189369662	0.18929814
63	0.106036329	0.10627204	88	0.189369662	0.18936075
64	0.106036329	0.10611809	89	0.189369662	0.18925926
65	0.106036329	0.10614854	90	0.189369662	0.18938667
66	0.106036329	0.10605967	91	0.289369662	0.28941387
67	0.106036329	0.10604222	92	0.289369662	0.28922169
68	0.106036329	0.10618083	93	0.289369662	0.28929850
69	0.106036329	0.10598174	94	0.289369662	0.28949639
70	0.106036329	0.10612261	95	0.289369662	0.28946205
71	0.139369662	0.13937247	96	0.289369662	0.28920360
72	0.139369662	0.13923064	97	0.289369662	0.28943995
73	0.139369662	0.13934868	98	0.289369662	0.28930091
74	0.139369662	0.13927767	99	0.289369662	0.28926759
75	0.139369662	0.13934098	100	0.289369662	0.28925455

Figura 5: Taula Valor de Shapley

m	1000	10000	100000	1000000	10000000	100000000
e_{th}	0.4073	0.1288	0.04073	0.01288	0.004073	0.001288

En aquest exemple, com hem vist en el cas anterior, quan comparem amb el valor de Shapley real podem treure les dades dels errors màxim i mitjà. Per tant, tindríem que l'error màxim és $e_{max} = 0,00064$ i el mitjà és $e_{avg} = 0,00015$.

Exemple 5.10. Joc de les sabates

Sigui (N, v) un joc de sabates on $N = \{1, ..., 100\}$. La funció característica es defineix com segueix, $\forall S \subset N$, $v(S) = min\{|S_{esquerra}|, |S_{dreta}|\}$, on $|S_{esquerra}|$ és el nombre de sabates del peu esquerra i $|S_{dreta}|$ és el nombre de sabates del peu dret. En aquest exemple $|S_{esquerra}| = |S_{dreta}| = 50$. Pel procés de mostreig, hem suposat, sense perdre generalitat, que els primers 50 jugadors pertanyen al grup de les sabates del peu esquerra i els altres 50 corresponen al grup de les sabates del peu dret.

Es pot veure fàcilment que el valor de Shapley per a cada jugador és 1/2. Ho podem entendre com, si un jugador entra a una coalició, la seva contribució marginal serà 1 si hi ha algún jugador sense parella del peu contrari i 0 si no n'hi ha. Com hi ha el mateix nombre de jugadors, de cada tipus, la mitad de les vegades, podrem formar una parella, i la mitad de les vegades no.

Per veure l'error teòric tornem a tenir $\sigma^2_{\phi_i} \leq (x^i_{max} - x^i_{min})^2/4 = \frac{1}{4}$.

m	1000	10000	100000	1000000	10000000	100000000
e_{th}	0.0407	0.0129	0.00407	0.00129	0.000407	0.000129

Exemple 5.11. L'arbre d'abast mínim

Sigui $G = (N \cup \{0\}, E)$ un graf on $N = \{1, ...100\}$, i el cost associat a l'aresta és:

$$C_{ij} = \begin{cases} 1, & \text{si}\, i = j + 1, i = j - 1, i = 1 \wedge j = 100, i = 100 \wedge j = 1, \\ 101, & \text{si}\, i = 0\, \text{o}\, j = 0, \\ \infty, & \text{altrament} \end{cases}$$

Un joc de l'arbre d'abast mínim, (N, c) és un joc de costos, on donada una coalició $S \subset N$, $c(S)$ és la suma del cost de l'aresta de l'arbre d'abast mínim, és a dir, $c(S)$=\{L'arbre d'abast mínim del graph $G|_{S \cup \{0\}}$\}, que és el graf restringit a les arestes de S i el node 0. El valor de Shapley per aquest joc és 2, $\forall i \in N$, ja que si miressim totes les contribucions marginals possibles de l'aresta i, només provocaria un cost de 101, una de cada 100 vegades, i la resta provocaría un cost de 1. Tenim doncs $(101 + 99 \cdot 1)/2 = 2$ com volíem veure.

Pel que fa al error teòric, el calculem fent servir que $\sigma^2_{\phi_i} \leq (x^i_{max} - x^i_{min})^2/4 = \frac{10^4}{4} = (101 - (-100))^2/4 = 10100,25$.

m	1000	10000	100000	1000000	10000000	100000000
e_{th}	$8,186$	$2,588$	$0,818$	$0,258$	$0,082$	$0,026$

5.4 Conclusió

Com en l'apartat anterior i com hem vist des del principi del treball, el problema del càlcul del valor de Shapley, exceptuant casos senzills, sempre recau en la seva complexitat. En aquest cas, es proposa una solució aproximada basada en temps polinomial. El procés de mostreig vist aquí és una eina important per veure-ho. Aquest mètode ens dóna una estimació del valor amb algunes propietats importants com el ser no esbiaixat, la consistència, el no error per a jugadors nuls i que és possible calcular l'error teòric a partir de la probabilitat.

Recordem també, que la única restricció de l'algoritme és el fet que la utilitat de cada coalició s'ha de calcular polinomialment, però és menys restrictiva que la de l'aproximació que apareix a la literatura del valor de Shapley.

6 Mostreig aleatori estructurat

6.1 Introducció

El novembre de 2016, a partir del treball de Javier Castro, Daniel Gómez i Juan Tejada; Tjeerd van Campen, Herbert Hamers, Bart Husslage i Roy Lindelauf van idear un nou mètode de càlcul del valor de Shapley anomenat *Mostreig aleatori estructurat* basat en el mètode del mostreig aleatori.

Aquest mètode està basat en la idea de que cada jugador apareix en cada posició el mateix nombre de vegades. Per tal d'aconseguir el resultat, es realitzen canvis en els ordres seleccionats aleatòriament, posant els jugadors en l'ordre que els toca. Les contribucions marginals dels jugadors en els nous ordres són utilitzades per calcular el valor de Shapley.

Imaginem un joc de 3 jugadors. Agafem $m = 6$, és a dir, 6 ordres aleatòriament i els dividirem en 3 grups de 2. Per calcular les contribucions marginals del jugador 1, el que farem es canviar el jugador 1 per colocar-lo a les posicions que toquen en cada un dels ordres. Trindrem: dos ordres que tinguin al jugador 1 en primera posició, dos que el tinguin en segona posició, i dos que el tinguin en la tercera posició.

Per exemple:

Grup	O_1	O_2	O_3
1	$-$	1	$-$
1	1	$-$	$-$
2	$-$	$-$	1
2	1	$-$	$-$
3	$-$	$-$	1
3	$-$	1	$-$

Ho canviaríem a:

Grup	O_1	O_2	O_3
1	1	σ_1	$-$
1	1	$-$	$-$
2	$-$	1	σ_2
2	σ_2	1	$-$
3	$-$	$-$	1
3	$-$	σ_3	1

A partir d'aquí, agafarem les contribucions marginals i en calcularíem el valor de Shapley per aquest jugador. Repetiriem el mateix amb la resta de jugadors.

6.2 Algoritme del mostreig aleatori estructurat per a n jugadors

Sigui (N, v) un joc de n jugadors, l'algoritme de càlcul del valor de Shapley, seguint el mètode aleatori estructurat és el següent:

1. S'escull un subconjunt $\pi_r(N) \subset \pi(N)$ de r ordres del $n!$ possibles tal que r sigui múltiple de n, és a dir, $\exists t \in \mathbb{N}$ complint $r = tn$.
2. Es divideix el conjunt π_r en n grups de tamany t.
3. Per cada jugador $i \in N$:
   (a) Canviem el jugador i amb el jugador a la posició j per a cada un dels t ordres en el grup j, on $j \in \{1, ..., n\}$, i a aquest conjunt resultant l'anomenem π'_r.
   (b) Calculem les contribucions marginals $x^i(O)$ del jugador i per cada un dels nous ordres $O \in \pi'_r$.
   (c) Aproximem el valor de Shapley del jugador i.

Exemple 6.1. Recordem l'exemple de l'apartat anterior, el joc de 3 jugadors on

S	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$v(S)$	0	1	3	0	5	7	4	10

Agafem $r = 6$ i per exemple, els ordres següents: $(2, 1, 3), (3, 2, 1), (3, 1, 2), (1, 2, 3), (1, 3, 2)$ i $(2, 3, 1)$. (Recordem que podríem repetir ja que la mostra l'agafem amb reposició, però per veure-ho més clar, agafem 6 ordres diferents)

Per tant, per calcular el valor de Shapley del jugador 1 trindrem les següents contribucions marginals:

Grup	*Ordre incial*	*Ordre canviat*	$x^1(O)$
1	$(2, 1, 3)$	$(1, 2, 3)$	1
1	$(3, 2, 1)$	$(1, 2, 3)$	1
2	$(3, 1, 2)$	$(3, 1, 2)$	7
2	$(1, 2, 3)$	$(2, 1, 3)$	2
3	$(1, 3, 2)$	$(2, 3, 1)$	6
3	$(2, 3, 1)$	$(2, 3, 1)$	6

Repetim el procés per els jugadors 2 i 3:

Grup	*Ordre incial*	*Ordre canviat*	$x^2(O)$
1	$(2, 1, 3)$	$(2, 1, 3)$	3
1	$(3, 2, 1)$	$(2, 3, 1)$	3
2	$(3, 1, 2)$	$(3, 2, 1)$	4
2	$(1, 2, 3)$	$(1, 2, 3)$	4
3	$(1, 3, 2)$	$(1, 3, 2)$	3
3	$(2, 3, 1)$	$(1, 3, 2)$	3

Grup	*Ordre incial*	*Ordre canviat*	$x^3(O)$
1	$(2,1,3)$	$(3,1,2)$	0
1	$(3,2,1)$	$(3,2,1)$	0
2	$(3,1,2)$	$(1,3,2)$	6
2	$(1,2,3)$	$(1,3,2)$	6
3	$(1,3,2)$	$(1,2,3)$	5
3	$(2,3,1)$	$(2,1,3)$	5

Per tant el valor de Shapley aproximat serà $\hat{\phi}(v) = (\frac{23}{6}, \frac{20}{6}, \frac{22}{6})$.

El primer detall que podem observar en aquest càlcul del valor de Shapley és que, mentre el mètode de mostreig aleatori compleix la propietat d'eficiència, el mètode de mostreig aleatori estructurat no la compleix.

$$\tfrac{23}{6} + \tfrac{20}{6} + \tfrac{22}{6} = \tfrac{65}{6} = 10,8\hat{3} > 10 = v(\{1,2,3\}).$$

Aquesta falta d'eficiència és deguda a que el mètode de mostreig aleatori estructurat només considera les contribucions marginals d'un sol jugador per cada ordre de mostreig.

L'altra cosa que podem veure és que ambdós mètodes requereixen el mateix nombre de càlculs de contribucions marginals. Tot i així, el mètode de mostreig aleatori estructurat necesita també un canvi en l'ordre dels jugadors, per tant, el temps de càlcul és molt superior.

Tot i així, com veurem a continuació, aquest mètode és molt millor que l'anterior en termes de calcular el valor de Shapley.

6.3 Comparació de l'error entre el mètode de mostreig aleatori i el mètode de mostreig aleatori estructurat

La idea que Tjeerd van Campen, Herbert Hamers, Bart Husslage i Roy Lindelauf van plantejar en aquest apartat és comparar el valor de Shapley calculat pels dos mètodes en diferents jocs d'una certa quantitat de jugadors, juntament amb el valor de Shapley real. Primer de tot expliquem una mica els tipus d'errors amb els que treballarem i com es mesuren i després farem dos tipus d'anàlisis, un que depèn del nombre d'ordres de les mostres dels dos mètodes i l'altre que es basa en el nombre de jugadors dels jocs cooperatius.

Definició 6.2. *Donat un joc (N, v), on $N = \{1, 2, ..., n\}$ i sigui $S \subseteq N$ un conjunt no buit de jugadors, un joc d'unanimitat u_S es defineix com*

$$u_S(T) = \begin{cases} 1, & si\ S \subseteq T \\ 0, & altrament \end{cases}$$

per cada possible coalició $T \subseteq N$. S'interpreta com que la utilitat només pot ser 1 quan tots els jugadors de S estan involucrats en la coalició.

Definició 6.3. *Sigui (N, v) un joc representat com una combinació lineal de jocs d'unanimitat, és a dir, $v(T) = \sum_{S \subseteq N, S \neq \emptyset} c_S u_S(T)$ on c_S és el coeficient del joc d'unanimitat corresponent. Aquest joc s'anomena "sum-of-unanimity-game"o SOUG i el valor de Shapley per un jugador i es pot calcular com $\phi_i(v) = \sum_{S: i \in S} \frac{c_S}{|S|}$*

Nota: Tot joc cooperatiu por ser expressat com a $SOUG$ i aquesta representació porporciona una forma eficient de calcular el valor de Shapley. Tot i així, per tal de transformar un joc cooperatiu en un SOUG, tots els $2^n - 1$ valors $v(S)$ de les coalicions del joc (N, v) s'han de considerar. Per tant, aquest procediment és útil només quan el joc $SOUG$ ja ve donat.

Exemple 6.4. Anem a calcular el valor de Shapley per un joc de tipus $SOUG$. Imaginem els següents jocs d'unanimitat per a 3 persones $u_{\{1\}}$, $u_{\{1,2\}}$ i $u_{\{2,3\}}$ amb les següents utilitats per cada coalició i per a cada un dels jocs:

S	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$u_{\{1\}}$	0	1	0	0	1	1	0	1
$u_{\{1,2\}}$	0	0	0	0	1	0	0	1
$u_{\{2,3\}}$	0	0	0	0	0	0	1	1

Siguin $c_{\{1\}} = 2$, $c_{\{1,2\}} = -3$ $c_{\{2,3\}} = 5$ i tots els altres coeficients igual a 0. Llavors pel corresponent joc de 3 persones de tipus $SOUG$ (N, v) tenim:

$$v(S) = 2 \cdot u_{\{1\}}(S) - 3 \cdot u_{\{1,2\}}(S) + 5 \cdot u_{\{2,3\}}(S).$$

I la següent taula:

S	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$v(S)$	0	2	0	0	-1	2	5	4

Ara podríem calcular el valor de Shapley amb les contribucions marginals com sempre, però la definició anterior ens mostra una altra manera de calcular-lo,

$$\phi_1(v) = \frac{c_{\{1\}}}{|\{1\}|} + \frac{c_{\{1,2\}}}{|\{1,2\}|} + \frac{c_{\{1,3\}}}{|\{1,3\}|} + \frac{c_{\{1,2,3\}}}{|\{1,2,3\}|} = \frac{2}{1} + \frac{-3}{2} + \frac{0}{2} + \frac{0}{3} = \frac{1}{2}.$$
$$\phi_2(v) = \frac{c_{\{2\}}}{|\{2\}|} + \frac{c_{\{1,2\}}}{|\{1,2\}|} + \frac{c_{\{2,3\}}}{|\{2,3\}|} + \frac{c_{\{1,2,3\}}}{|\{1,2,3\}|} = \frac{0}{1} + \frac{-3}{2} + \frac{5}{2} + \frac{0}{3} = 1.$$
$$\phi_3(v) = \frac{c_{\{3\}}}{|\{1\}|} + \frac{c_{\{1,3\}}}{|\{1,3\}|} + \frac{c_{\{2,3\}}}{|\{2,3\}|} + \frac{c_{\{1,2,3\}}}{|\{1,2,3\}|} = \frac{0}{1} + \frac{0}{2} + \frac{5}{2} + \frac{0}{3} = \frac{5}{2}.$$

Per tant tenim $\phi(v) = (\frac{1}{2}, 1, \frac{5}{2})$.

La idea ara, és aplicar els dos mètodes a jocs del tipus $SOUG$. Es generen una certa quantitat de jocs de tipus $SOUG$ aleatòriament, generant de forma també aleatòria, el nombre de jocs d'unanimitat i els coeficients; per tant, els jocs cooperatius a partir dels quals generem els jocs $SOUG$ també es generen de forma aleatòria. D'aquesta forma, com els jocs són coneguts, ens permetem una forma eficient de calcular el valor de Shapley.

Definició 6.5. *Sigui j el nombre de jocs $SOUG$ generats i n el nombre de jugadors dels jocs cooperatius. Sigui ϕ_i amb $i \in N$ el valor de Shapley del jugador i i $\hat{\phi}_i$ la seva aproximació a partir d'algún mètode de càlcul, anomenem $AAAE$ ("Average average absolute error"), a la mitjana de la mitjana de l'error absolut de cada joc i el calculem de la següent forma:*

$$AAAE = \frac{1}{j} \sum_{k=1}^{j} \left(\frac{1}{n} \sum_{i=1}^{n} \left| \hat{\phi}_i(v_k) - \phi_i(v_k) \right| \right).$$

Definició 6.6. *Sigui j el nombre de jocs $SOUG$ generats i n el nombre de jugadors dels jocs cooperatius. Sigui ϕ_i amb $i \in N$ el valor de Shapley del jugador i i $\hat{\phi}_i$ la seva aproximació a partir d'algún mètode de càlcul, anomenem $AAPE$ ("Average average percentage error"), a la mitjana de la mitjana de l'error relatiu de cada joc i el calculem de la següent forma:*

$$AAPE = \frac{1}{j} \sum_{k=1}^{j} \left(\frac{1}{n} \sum_{i=1}^{n} \frac{\left| \hat{\phi}_i(v_k) - \phi_i(v_k) \right|)}{|\phi_i(v_k)|} \right..$$

En el nostre cas, agafarem $j = 50$, és a dir, 50 jocs generats aletòriament de tipus SOUG. El procés a seguir en aquest estudi serà doncs:

1. Generar aleatòriament els 50 jocs $SOUG$ i normalitzar el valor de la coalició formada per tots els jugadors per cada joc.
2. Calcular el valor de Shapley per cada jugador de cada un dels 50 jocs.
3. Utilitzar el mètode del mostreig aleatori per aproximar el valor de Shapley per cada jugador dels 50 jocs, i calcular-ne el $AAAE$ i el $AAPE$.
4. Utilitzar el mètode del mostreig aleatori estructurat per aproximar el valor de Shapley per cada jugador dels 50 jocs, i calcular-ne el $AAAE$ i el $AAPE$.

Nota: Pot ser que algun jugador tingui un valor de Shapley molt pròxim a 0 i això pot provocar problemes amb el $AAPE$, per això, els valors de Shapley que siguin inferiors al $0,1\%$ de $v(N)$ no els tindrem en compte pel càlcul del $AAPE$.

6.3.1 Nombre d'ordres

En aquesta secció, l'objectiu és veure si, com pensem, quan més gran sigui la nostra mostra d'ordres, més bona serà la nostra aproximació utilitzant els mètodes. Per tant, veurem com evolucionen els errors a mesura que el nombre d'ordres de la mostra augmenta. S'agafen els 50 jocs de 100 jugadors cada un i es construeix a partir de 100 jocs d'unanimitat, amb una coalició aleatòria $S \subset N$. Els coeficients dels jocs d'unanimitat seran nombres enters entre -10 i 10. L'estudi realitzat varia entre 500 i 5000 ordres tot i que n'hi ha 100! (menys d'un $0,1\%$).

La següent figura ens mostra els dos tipus d'errors que hem definit (AAAE, AAPE) pels dos mètodes que estem comparant. Podem observar com va variant cada un dels errors a mesura que el nombre d'ordres va augmentant de 500 a 5000.

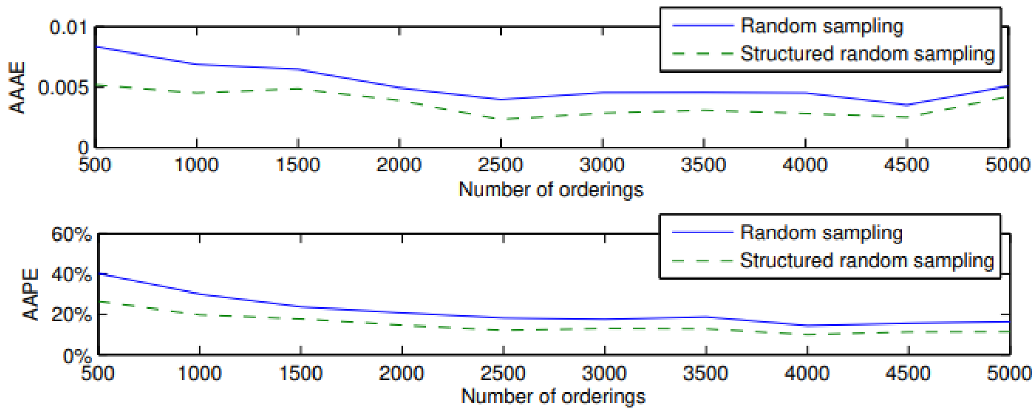


Figura 6: Anàlisi dels errors en els mètodes d'aproximació del valor de Shapley

Podem veure que els $AAAE$s pels dos mètodes són molt petits. En canvi els $AAPE$s són certament més grans. Podem veure, per exemple, que quan utilitzem 5000 ordres, el $AAPE$ és d'un 11% pel mètode de mostreig aleatori estructurat i és d'un 16% pel mètode de mostreig aleatori. Per tant, clarament quants més ordres utilitzem, millor és l'aproximació (tot i que hem utilitzat relativament pocs ordres i hem d'entendre que aquests resultats podrien ser molt més prometedors). Tot i així, podem veure com el mètode de mostreig aleatori estructurat millora clarament l'altre mètode d'aproximació.

6.3.2 Nombre de jugadors

En aquesta segon anàlisi, calcularem també els errors com a l'apartat anterior, però ara variarem el nombre de jugadors dels jocs cooperatius. Agafarem també 50 jocs SOUG generats aleatòriament, amb els paràmetres utilitzats per construir aquests jocs, agafats també aleatòriament de la mateixa manera que a l'apartat anterior. En aquest cas, agafarem 5000 ordres i variarem el nombre de jugadors de 10 a 100. A més a més, el valor normalitzat de la coalició formada per tots els jugadors en cada joc de n jugadors es calcula per $0,01n$ permetent d'aquesta manera, el càlcul dels errors absoluts pels diferents mètodes d'aproximació per a diferents nombres de jugadors

La següent figura ens mostra els dos tipus d'errors que hem definit (AAAE, AAPE) pels dos mètodes que estem comparant. Podem observar com va variant cada un dels errors a mesura que el nombre de jugadors va augmentant de 10 a 100.

Els dos mètodes, com podem veure, tenen els dos tipus d'errors bastant baixos. Tot i així, veiem que augmenten a mesura que augmenta el nombre de jugadors. Possiblement, aquest fet, es deu al fet que el nombre d'ordres està fixat ja que, quan augmentem el nombre de jugadors, el nombre d'ordres possibles també augmenta molt. Per tant, quants més ordres utilitzéssim, més acurats serien els resultats.

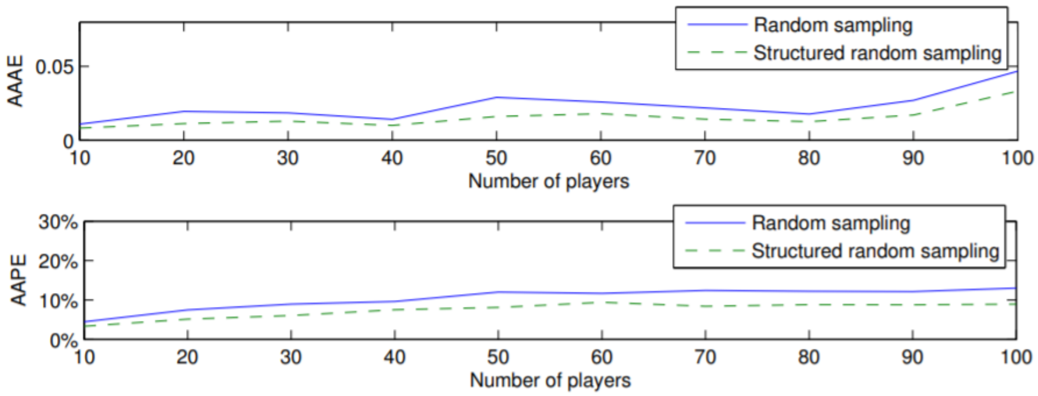


Figura 7: Anàlisi dels errors en els mètodes d'aproximació del valor de Shapley

Tot i així, tampoc augmenta de forma desmesurada, només una mica. Per exemple el *AAPE* és d'un 3% pel mètode de mostreig aleatori estructurat i és d'un 4% pel mètode de mostreig aleatori quan el nombre de jugadors és 10; i és d'un 9% pel mètode de mostreig aleatori estructurat i és d'un 13% pel mètode de mostreig aleatori quan el nombre de jugadors és 100. Novament, podem veure com el mètode de mostreig aleatori estructurat és millor que l'altre mètode d'aproximació.

6.4 Exemple del mètode aplicat a un cas real: La xarxa informàtica d'Al Qaeda del atac al WTC del 9/11

El dia 11 de setembre del 2001, el grup terrorista Al Qaeda va segrestar quatre avions als Estats Units d'Amèrica. Dos d'aquests avions van estrellar-se contra les dues Torres Bessones, uns edificis del "World Trade Center"de Nova York, el tercer va volar cap al Pentagon i el quart a Pensilvania. Van morir 2977 persones i moltes més van quedar ferides.

Basant-se en els recursos públics disponibles, Krebs va fer un mapa de tota la xarxa d'informació que va fer servir el grup terrorista durant l'atac al WTC. Inclou la xarxa de 19 pilots i segrestadors que van ser responsables del atac i s'estén a 69 membres més que van ser-ne còmplices.

En la figura, els tamanys dels vèrtexs, és proporcional al nombre de coneccions de cada membre.

Lindelauf et al. (2013) va calcular-ne el valor de Shapley del joc de connexions de la xarxa dels 19 pilots i segrestadors. Calcular el valor de Shapley de l'extensió de 69 membres no es podia aconseguir amb el seu mètode degut als 69! ordres possibles a considerar-se. Afortunadament, amb el mètode de mostreig aleatori estructurat podem calcular-ne una aproximació, així es pot fer una classificació basant-se en les connexions, de qui va ser més responsable de l'incident.

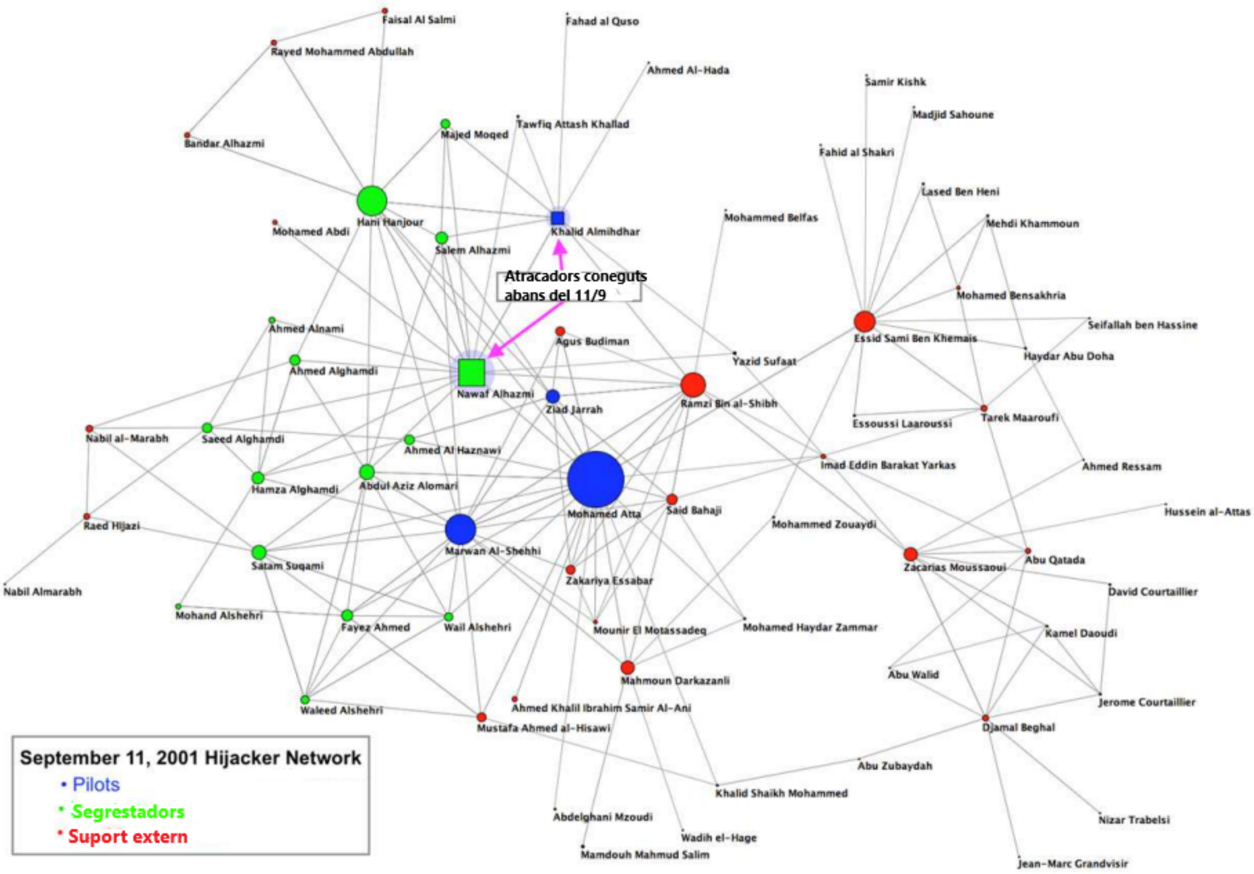


Figura 8: Xarxa terrorista d'Al Qaeda

$v(S)$ es defineix de la següent manera:

$$v(S) = \begin{cases} \sum_{i \in S} w_i, & \text{si } S \text{ està connectada} \\ 0, & \text{en cas contrari} \end{cases}$$

on w_i són els punts obtinguts, depenent de l'acció que realitzés el jugador i amb algún jugador amb el que tingués connexió. Algunes de les accions possibles són: assistir a reunions, accions radicals, afiliacions, assistència a camp d'entrenament terroristes, ajuda en atacs anteriors, etc...

Comptem com a valor de la gran coalició $v(N) = 1$ on $N = \{1, ..., 69\}$. Tot i així, veurem que la suma dels 10 primers membres, ja es superior a 1, per tant més de la mitad de jugadors tindran un valor de Shapley negatiu. Podem veure el top 15 a la següent taula, i es considera a aquests terroristes com als principals culpables del famós atemptat.

Ranking	Nom	Valor de Shapley aprox.
1	Mohamed Atta	0.1137
2	Essid Sami Ben Khemais	0.1111
3	Hani Hanjour	0.1107
4	Djamal Beghal	0.1070
5	Khalid Almihdhar	0.1069
6	Mahmoun Darkazanli	0.1067
7	Zacarias Moussaoui	0.1009
8	Nawaf Alhazmi	0.0995
9	Ramzi Bin al-Shibh	0.0985
10	Raed Hijazi	0.0949
11	Hamza Alghamdi	0.0090
12	Fayez Ahmed	0.0088
13	Marwan Al-Shehhi	0.0046
14	Satam Suqami	0.0038
15	Saeed Alghamdi	0.0037

Figura 9: Top 15 Terroristes

Aquest ranking s'ha obtingut fent servir el mètode de mostreig aleatori estructurat utilitzant 10000 ordres amb un temps de càlcul 5 min (amb ordinador). S'ha repetit el procés utilitzant 500000 i el temps de càlcul ha estat d'unes 4 hores, però això no ha tingut cap efecte en quant al top 15. Recordem que 500000 ordres segueix sense ser una gran part del total d'ordres ja que, aquest és de 69!. Per això, veiem que es necessita una quantitat diminuta del total per obtenir uns resultats decents.

En aquest top 10, podem trobar a Mohamed Atta (pilot que va anar al nord de WTC), Hani Hanjour (pilor que va anar al Pentagon), Khalid Almihdhar i Nawaf Alhazmi (segrestadors del vol que va anar al Pentagon) i Zacarias Moussaoui (arrestat dies abans del atemptat). També hi podem trobar gent que no eren ni pilots ni segrestadors, com Essid Sami Ben Khemais, un cap d'operacions d'Al Qaeda a Itàlia i Djamal Beghal.

6.5 Conclusió

El valor de Shapley és un dels conceptes de solucion més prominents en la teoria dels jocs cooperatius. Tot i així, en general, la seva complexitat es enorme quan més gran es el nombre de jugadors. Hi ha molts jocs on és interessant calcular el valor de Shapley però normalment el nombre de jugadors és molt elevat. Per això basant-se en el mètode de mostreig aleatori, s'ha desenvolupat un mètode certament millorat anomenat mètode de mostreig aleatori estructurat.

A partir d'aquí, hem vist com els errors entre el valor de Shapley real i els valors obtinguts en els mètodes esmentats era inferior en el nou mètode, sent més prominent quan el nombre d'ordres és més petit o el nombre de jugadors augmenta. Els temps de càlcul en aquest mètode eren relativament superiors, però degut als resultats, ens ho podem permetre.

Hem aplicat aquest mètode a un joc de connexions per l'atac terrorista d'Al Qaeda del 11 de setembre del 2001 i la seva xarxa d'informacions, per tal d'identificar quins van ser els homes clau en aquell atemptat.

7 Conclusions

El valor de Shapley, com hem repetit nombroses vegades durant aquest treball, és molt útil però a la vegada molt complex. Calcular-lo és extremadament difícil i requereix de molt de temps quan el nombre de jugadors dels jocs cooperatius són elevats. En aquesta memòria, hem vist una diversitat de formes de calcular-lo; algunes de les quals, tot i que siguin només aproximacions, permeten solucionar el problema que teníem del temps de càlcul, i ens donen una solució força aproximada del valor.

La primera forma que hem vist és un mètode semblant a la *simulació de Montecarlo*, però la diferència és que en el nostre ens centrem en buscar una aproximació del valor, agafant els pesos com a variables aleatòries. Aquest mètode ens presenta una certa incertesa que podem trobar al buscar l'aproximació del valor de Shapley, i hi veiem una forma de calcular-la i com relacionar els dos valors trobats (valor de Shapley i incertesa).

Després hem vist el *Mètode de mostreig aleatori*, que consisteix en calcular una aproximació del valor de Shapley agafant una mostra de tots els ordres possibles, i calcular les contribucions marginals dels jugadors a partir d'aquests ordres, i fer-ne la mitjana per trobar aquesta aproximació. També veiem una forma de calcular-ne una cota de l'error teòric.

Finalment, hem mostrat un mètode basat en aquest anterior, el *Mètode de mostreig aleatori estructurat*. Ara, la mostra d'ordres la tranformem de forma que cada jugador aparegui en cada posició el mateix nombre de vegades. D'aquesta manera, al haver de calcular-ne l'aproximació de cada jugador per separat, perdem la propietat d'eficiència, però veiem que els errors respecte al valor de Shapley exacte, són inferiors als del mètode de mostreig aleatori. S'ha caculat els errors AAAE i AAPE per una certa quantitat de jocs, augmentant primer el nombre d'ordres de la mostra i després el nombre de jugadors, comparant-lo amb el mètode de mostreig aleatori; i veient, d'aquesta manera, com el mètode de mostreig aleatori estructurat és millor en quant a termes d'aproximació. També és una mica més lent, però arribem a la conclusió de que val la pena.

Referències

[1] Julio González-Díaz, Ignacio García-Jurado, M. Gloria Fistras-Janeiro: *An Introductory Course on Mathematical Game Theory*, 2010

[2] Carles Pallarola, Josep Maria Izquierdo, Jesús Marín, Francisco Javier Martínez, Marina Nuñez i Neus Ybern: *Jocs cooperatius i apicacions econòmiques*, 1999

[3] *Wighted Voting Games*

https://personal.utdallas.edu/ chandra/documents/6311/ weight-games.pdf

[4] *Jocs amb utilitats transferibles*

http://eio.usc.es/eipc1/BASE/BASEMASTER/ FORMULARIOS-PHP-DPTO/MATERIALES/Mat_G101145171_JuegosTU.pdf

[5] Shaheen S. Fatima, Michael Wooldridge i Nicholas R. Jennings, *An analysis of the Shapley Value and its Uncertainty for the Voting Game* http://www.cs.ox.ac.uk/people/michael.wooldridge/pubs/amec05b.pdf

[6] Javier Castro, Daniel Gómez i Juan Tejada, *Polynomial calculation of the Shapley value based on sampling*

https://www.academia.edu/7555845/Polynomial_calculation_of_the_Shapley_value_based_on_sampling

[7] Tjeerd van Campen, Herbert Hamers, Bart Husslage i Roy Lindelauf, *Mostreig aleatori estructurat*

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2874859