

TEORIA DE JUEGOS COOPERATIVOS VERSUS TEORIA DE LA EVIDENCIA

José Antonio Núñez del Prado, M^a Pilar García Pineda, Antonio Heras Martínez

Universidad Complutense de Madrid

RESUMEN

Aunque conceptualmente diferentes, las formulaciones matemáticas de la teoría de juegos cooperativos, efectuada por VON NEUMANN Y MORGENSTERN en su tratado clásico de 1944, y la teoría de la evidencia, creada por A.P.DEMPSTER y G. SHAFFER, como una generalización de la teoría de la probabilidad a finales de los 70 del siglo pasado, son formalmente iguales. El objetivo de este trabajo es el de mostrar ese paralelismo y el de establecer la posible interpretación de una cualquiera de esas teorías por los métodos y los conceptos de la otra.

INTRODUCCIÓN

La teoría matemática de la evidencia, formulada inicialmente por A.P.DEMPSTER en la década de los 60 del siglo pasado y sistematizada por G. SHAFFER a finales de los 70 de ese mismo siglo, aunque sus orígenes se remontan al siglo XVII cuando se iniciaron los primeros análisis para modelar la incertidumbre de los que surgió como línea principal la teoría de la probabilidad, es una teoría de la creencia no probabilística que los sujetos tienen acerca de algún campo de conocimientos y del modo como operan y razonan con esas creencias.

La teoría de la evidencia, tanto en DEMPSTER – SHAFFER, como en las discusiones del siglo XVII acerca del conocimiento incierto, parte de la crítica de la hipótesis, aceptada en la teoría de la probabilidad clásica y que constituye su fundamento, de la aditividad de las creencias: sean A y B dos proposiciones independientes entre sí acerca del mundo; la propiedad fundamental de la probabilidad es que $P(A \text{ ó } B) = P(A) + P(B)$ pero siempre se ha pensado que la “unión” de dos o más proposiciones podrían reforzarse mutuamente de forma que su “fuerza conjunta” sería más que la suma de la fuerza de cada componente. Esta es la idea básica que se halla detrás de la teoría de la evidencia.

Sea Ω el universo del discurso y 2^Ω el conjunto de todas las proposiciones acerca de Ω . En la teoría de la probabilidad cada $A \in 2^\Omega$ es un subconjunto de Ω que no es más que la “suma” de sus elementos y la probabilidad de A queda determinada por las probabilidades de sus constituyentes $P(A) = \sum P(a_i)$ con $a_i \in A$. Es por esta razón por la que la probabilidad queda definida por una asignación básica de probabilidades a los elementos de Ω $P: \Omega \rightarrow [0, 1]$ con $\sum P(a_i) = 1$, $a_i \in \Omega$. En la teoría de la creencia cada subconjunto $A \in 2^\Omega$ no es la “suma” de sus elementos y ha de ser considerado en sí mismo de forma global; por esto una asignación de creencia pasa de Ω a 2^Ω y se define como una función $m: 2^\Omega \rightarrow [0, 1]$ tal que 1) $m(\emptyset) = 0$ y 2) $\sum m(A) = 1$ $A \in 2^\Omega$. Dada esta asignación básica de creencia se define la creencia total en $A \in 2^\Omega$ como $BEL(A) = \sum m(X)$ con $X < A$. Es decir, la creencia es “como” una probabilidad definida en 2^Ω y no en Ω con la “suma” sobre los elementos de A sustituida por la suma sobre los subconjuntos de A, cambio de perspectiva que pretende reflejar la idea de que todas las proposiciones, sean elementales $a_i \in \Omega$ o compuestas $A \in 2^\Omega$, son consideradas como

entidades propias y en igualdad de condiciones mutuas. Frente a la aditividad de la teoría de las probabilidades las funciones de creencia satisfacen la superaditividad: para cualesquiera dos subconjuntos A, B de Ω se tiene que $BEL(A \cup B) + BEL(A \cap B) \geq BEL(A) + BEL(B)$. En particular si $A \cap B = \emptyset$ se tiene que $BEL(A \cup B) \geq BEL(A) + BEL(B)$ que es una forma de decir que la “unión hace la fuerza”.

Curiosamente, aunque la teoría de la evidencia fue formulada como una teoría de la creencia en las proposiciones, formalmente es análoga a la formulación de la teoría de los juegos cooperativos. Aquí, el universo del discurso Ω es sustituido por el conjunto de jugadores $\Omega = \{1, 2, 3, \dots, n\}$, el conjunto de todas las proposiciones posibles 2^Ω por el conjunto de todas las coaliciones posibles 2^Ω y la función de creencia BEL por la función característica, normalizada sin pérdida de generalidad, del juego $v: 2^\Omega \rightarrow [0, 1]$ con 1) $v(\emptyset) = 0$, 2) $v(\Omega) = 1$ y alguna forma de superaditividad: si $A \cap B = \emptyset$ se tiene que $v(A \cup B) \geq v(A) + v(B)$ que es igualmente una forma de decir que la “unión hace la fuerza”; una de las más usuales es aceptar que la función característica v la condición de “supermodularidad”: para cualesquiera dos subconjuntos A, B de Ω se tiene que $v(A \cup B) + v(A \cap B) \geq v(A) + v(B)$ que es exactamente la condición que cumplen las funciones de creencia. La idea subyacente es la misma: la coalición $A \in 2^\Omega$ es más que la “suma” de sus jugadores componente y ha de tener mejor pago $v(A)$ que la suma de los pagos individuales $v(A) \geq \sum v(i)$, esto ha de ser superaditiva. Esta similitud de planteamientos hace que ambas teorías puedan ser consideradas equivalentes entre sí.

2.LA REPRESENTACIÓN DE LA IGNORANCIA

Las funciones de creencia permiten representar la ignorancia mejor que la teoría de las probabilidades. Consideremos, por ejemplo, la cuestión ¿Existe Dios? Aquí Ω posee dos y sólo dos posibilidades $\Omega = \{\text{sí}, \text{no}\}$. Desde el punto de vista probabilístico ha de tenerse, por la aditividad de las probabilidades, que $P(\text{sí}) + P(\text{no}) = 1$ y como $P(\text{no}) = 1 - P(\text{sí})$ necesariamente como $P(\text{no}) = P(\text{sí}) = 1/2$. La ignorancia consiste en creer con un 50% de posibilidades que Dios existe y que no existe con otro 50%: en la mitad de los mundos posibles existe Dios y en la otra mitad no. Tal cuantificación no es creíble.

Desde el punto de vista de la teoría de la evidencia la ignorancia puede ser mejor representada por una función de creencia BEL: $2^\Omega \rightarrow [0, 1]$ con BEL(sí) = 0, BEL(no) = 0 y BEL (sí o no) = 1.

3. “PROBABILIDADES” SUPERIORES

DEMPSTER y SHAFFER introducen para cada función de creencia BEL: $2^\Omega \rightarrow [0, 1]$ la función P^* : $2^\Omega \rightarrow [0, 1]$ definida por $P^*(A) = 1 - BEL(\neg A)$, $A \in 2^\Omega$ y $\neg A = \Omega - A$ el complemento de A respecto del universo del discurso Ω que representa igualmente un grado superior de creencia en A: BEL($\neg A$) es la creencia en “no A” y $1 - BEL(\neg A)$ es una medida de la posibilidad de A. Puede mostrarse que siempre se tiene que $BEL(A) \leq P^*(A)$. Entre estas dos funciones BEL y P^* DEMPSTER y SHAFER consideran el conjunto $M(P)$ de todas las probabilidades $P: 2^\Omega \rightarrow [0, 1]$ entre esas dos funciones: $BEL(A) \leq P(A) \leq P^*(A)$, que es la forma original en la que aparecieron las funciones de creencia. La motivación tras ello es la siguiente: desde el punto de vista de la teoría de las probabilidades, cuando se tiene un conjunto de proposiciones inciertas en la mayoría de las ocasiones no se sabe que probabilidades asociarlas pero sí se sabe que esas probabilidades forman un conjunto $M(P)$; la función $\inf P(A)$ es una medida de la mínima creencia que se está dispuesto a asignar a A y $\sup P(A)$ es la máxima creencia que se está dispuesto a asociar a A: la asignación $\inf(A)$ constituye una función de creencia BEL asociada a dicho conjunto de probabilidades $M(P)$ y $\sup(A)$ constituye la asignación de “probabilidad superior” P^* asociada a BEL(A) asociada a BEL. En la teoría de juegos cooperativos, curiosamente, aparece una situación similar. Si (Ω, v) es un juego supermodular $v(\Omega) = 1 \geq \sum v(i)$ y el exceso $v(\Omega) - \sum v(i)$ representa la ganancia extra que obtienen los jugadores por sus coaliciones. Este exceso hay que repartirlo adecuadamente entre dichos jugadores y un tal reparto, denominado IMPUTACIÓN, consiste en una función $X: 2^\Omega \rightarrow [0, 1]$ con $X(i) = x_i$ tal que 1) $x_i \geq v(i)$ ya que es un extra añadido y 2) $\sum x_i = 1$, que define así una probabilidad sobre Ω . De todas estas posibles imputaciones, la teoría de juegos busca aquellas que sean NO-DOMINADAS, es decir aquellas que no puedan ser mejoradas por un mejor reparto de los beneficios denominadas el NÚCLEO del juego (Ω, v) . Un teorema de caracterización muestra que el núcleo lo forman aquellas y sólo aquellas imputaciones tales que $X(A) = \sum x_i \geq \sum v(i)$, $i \in A \in 2^\Omega$, es decir por las probabilidades que mayorizan el valor del juego v. Se tiene así que el núcleo es el conjunto $M(P)$ de todas las probabilidades $P: 2^\Omega \rightarrow [0, 1]$ entre esas dos funciones: $v(A) \leq P(A) \leq v^*(A)$ como

en el caso de las funciones de creencia. Ahora se hace posible mostrar un teorema de la teoría de juegos cooperativos que afirma que el núcleo es vacío para los juegos cooperativos de suma constante, es decir para aquellos juegos en los que $v(A) = 1 - v(-A)$; estos juegos supermodulares de suma constante son precisamente aquellos en los que v se convierte en una distribución de probabilidad sobre el conjunto de jugadores y por lo tanto $M(P)$ ha de coincidir con la única probabilidad v que define el juego: los juegos cooperativos de suma constante son los espacios de probabilidad clásicos.

4. ESTADÍSTICA EN LOS JUEGOS COOPERATIVOS

La interpretación de la función característica de los juegos cooperativos como una función de creencia permite introducir nociones «probabilísticas» en la teoría de juegos:

4.1.- Función de distribución inferior y superior de un juego: Si v es la función característica del juego y X es cualquier función definida en Ω se denomina función de distribución inferior de X $F_*(x) = v(X \leq x)$ en analogía con la teoría de la probabilidad estándar. Igualmente se define la función de distribución superior como $F^*(x) = v(X \geq x)$. Se puede mostrar que $F_*(x) \leq F^*(x)$.

4.2.- Esperanza inferior y superior de una «variable aleatoria x »:

$$E_*(X) = \int_{-\infty}^{\infty} x dF_*(v)$$

$$E^*(X) = \int_{-\infty}^{\infty} x dF^*(v)$$

Los intercambios de las *estrellas* en las respectivas funciones de distribución son necesarios para que $E_*(X) \leq E^*(X)$.

Al igual que en el caso probabilístico estándar, los conceptos de valor esperado inferior y superior generalizan los conceptos de función característica y probabilidad superior en el sentido de que si I_T es la función indicador del subconjunto $T \subseteq \Omega$ se tienen que $E_*(X) = v(T)$ y $E^*(X) = P^*(T) = 1 - v(-T)$.

Como se intuye puede demostrarse que el núcleo del juego $(\Omega, \wp(\Omega), v)$ puede caracterizarse como el conjunto de probabilidades P de Ω tales que $E^*(X) \leq E_P(X) \leq E^*(X)$.

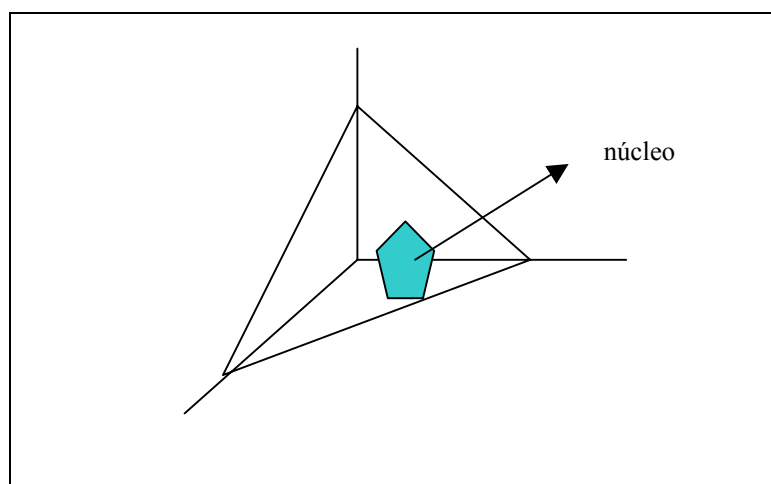
4.3.-Condicionización. Si $(\Omega, \wp(\Omega), v)$ es un juego cooperativo y B es una coalición se obtienen un nuevo juego cooperativo restringido a B $(B, \wp(B), v_B)$ en el cual v_B se define

$$v_B(A) = \frac{v(A \cap B)}{v(B)}$$

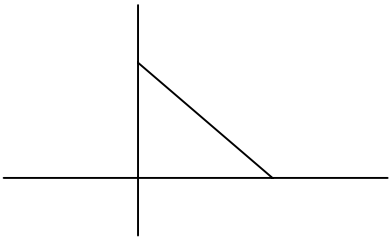
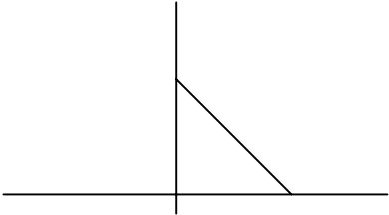
como en el caso probabilístico estándar.

5. SOLUCIONES UNÍVOVAS DEL JUEGO: VALOR DE SHAPLEY

El «core» o núcleo de un juego cooperativo $(\Omega, \wp(\Omega), v)$ representa todas las posibles redistribuciones entre los jugadores de los beneficios conseguidos por su cooperación. Como hemos visto dicho núcleo es un conjunto convexo compacto del simplex de probabilidades de Ω que contiene una cantidad infinita de elementos.



¿Cuál elegir de todos ellos?. La teoría de juegos cooperativos ha considerado diversas elecciones particulares y concretas del núcleo $C(v)$ tales como el valor de SHAPLEY. Si bien desde el punto de vista de la teoría de juegos, tales elecciones parecen adecuadas, desde el punto de vista de la teoría de la evidencia no resultan tan razonables. Esta fricción entre ambos enfoques, matemáticamente análogos, puede ilustrarse por medio del siguiente ejemplo.

<p>Sea $\Omega = \{1,2\}$ un juego con dos jugadores y sea v la función característica definida por:</p> $v : \wp(\Omega) \rightarrow [0,1]$ $\emptyset \rightarrow v(\emptyset) = 0$ $\{1\} \rightarrow v\{1\} = 0$ $\{2\} \rightarrow v\{2\} = 0$ $\{1,2\} \rightarrow v\{1,2\} = 1$ <p>Ningún jugador consigue nada por sí mismo pero cooperando entre ellos logran una unidad de beneficio. Se ve que el núcleo de este juego consiste de todas las probabilidades sobre Ω, es decir $M(P)$ es el simplex $[0, 1]$.</p>  <p>El valor de SHAPLEY de $(\{1,2\}, \wp\{1,2\}v)$ es, como siempre, el centro de gravedad del núcleo y en este caso dicho centro de gravedad es la probabilidad $P(1/2, 1/2)$: esto es, los jugadores se reparten el beneficio por igual.</p>	<p>Consideremos la cuestión histórica ¿existe Dios?. Hay dos respuestas posibles: SI o NO. Sea $\Omega = \{si, no\}$ y la función de creencia mínima:</p> $v : \wp(\Omega) \rightarrow [0,1]$ $\emptyset \rightarrow v(\emptyset) = 0$ $\{si\} \rightarrow BEL\{si\} = 0$ $\{no\} \rightarrow BEL\{no\} = 0$ $\{si, no\} \rightarrow BEL\{si, no\} = 1$ <p>ya que todos estamos dispuestos a creer que existe o que no existe, pero salvo por creencias personales, no sabemos en absoluto si <i>sí</i> o si <i>no</i>. Como antes el conjunto de probabilidades P tales que $BEL \leq P \leq P^*$ es el simplex $[0,1]$.</p>  <p>El valor de SHAPLEY sería igualmente la probabilidad $P(1/2, 1/2)$ que equivale a creer que existe en un 50% de las veces y en otro 50% que no lo que no parece una redistribución de la creencia total $BEL\{si, no\} = 1$ muy razonable.</p>
--	--

Claramente la redistribución de la creencia a partes iguales sólo puede depender de forma subjetiva de la persona a la que se le plantea la cuestión. De igual modo, en el ejemplo del juego cooperativo, la redistribución del beneficio dependerá del trabajo real llevado a cabo por cada jugador en la coalición y no sólo por el beneficio total que dicha coalición reporta. La cuestión no es valadí como el siguiente caso *legal* demuestra:

EL CASO DEL PADRE A CARA O CRUZ. Un hombre fue acusado en un caso de paternidad sobre la base de un marcador genético cuya frecuencia en la población adulta es del 1% y que se trasmite con probabilidad 1 de padres a hijos. Tanto el presunto padre como el niño causante del litigio poseían el citado marcador, por lo que el fiscal del caso planteó la conveniencia de obtener LA PROBABILIDAD DE QUE EL ACUSADO FUERA EL PADRE DADO QUE EL NIÑO TENÍA EL MARCADOR. Si representamos el suceso “el acusado es el padre” por A y el hecho cierto de que “el niño tienen el marcador” por B el fiscal aplicó el teorema de Bayes y calculó

$$P(A/B) = \frac{P(B/A)P(A)}{P(B/A)P(A) + P(B/A^c)P(A^c)}$$

De lo anteriormente dicho sabemos que $P(B/A) = 1$ y $P(B/A^c) = 0.01 = 1\%$. Por tanto sólo se necesita conocer los valores de $P(A)$ y $P(A^c)$, sustituir y calcular la probabilidad de $P(A/B)$. El fiscal estimó que ambas eran 0.5, valores que trataban de reflejar el desconocimiento que de la posible paternidad se tenía y puesto que podía ser o no el padre, lo lógico, pensó el fiscal, parecía signar igual probabilidad a ambos supuestos. El resultado de esa redistribución de la creencia, que corresponde al valor de SHAPLEY, no pudo ser más concluyente en contra del acusado porque $P(A/B)$ resulto ser aproximadamente 0.99. El defensor recurrió esa redistribución de la creencia y basó su recurso precisamente en la asignación, considerada por el fiscal, del valor de SHAPLEY a la función de creencia mínima:

$$v: \wp(\{A = \text{el acusado es el padre}, A^c = \text{el acusado no es el padre}\}) \rightarrow [0,1]$$

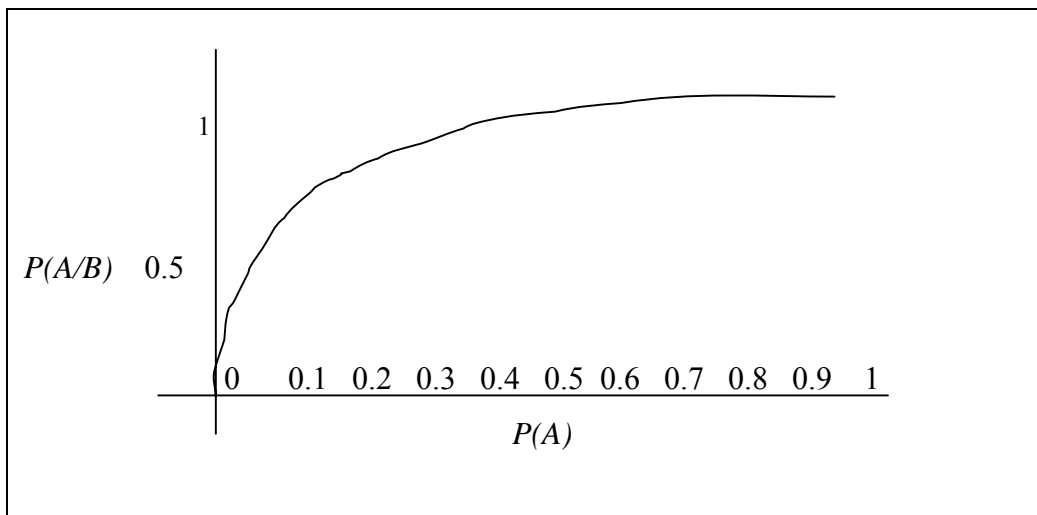
$$\emptyset \rightarrow v(\emptyset) = 0$$

$$\{A\} \rightarrow BEL\{A\} = 0$$

$$\{A^c\} \rightarrow BEL\{A^c\} = 0$$

$$\{A, A^c\} \rightarrow BEL\{A, A^c\} = 1$$

El defensor mostró que llevada a sus últimas consecuencias, semejante asignación de probabilidades equivalía a declarar padre a cualquier adulto por el procedimiento de CARA o CRUZ. Una vez más, prosiguió el defensor, se confundía ignorancia con probabilidad. Para rematar su discurso obtuvo $P(A/B)$ para distintos valores de $P(A)$ que representamos en la siguiente gráfica.



Se ve que para valores bajos de $P(A)$, entre 0 y 0.1, $P(A/B)$ da valores bajos que difícilmente condenan a cualquiera. Puesto que $BEL(A) = \inf \{P(A): P \text{ probabilidad del simplex } [0,1]\} = 0$ es la creencia adecuada para el suceso $A = \text{el acusado es el padre}$, cualquier redistribución de la creencia total $BEL\{A, A^c\} = 1$ en alguna probabilidad del simplex $[0,1]$, como el valor de SHAPLEY, puede ser erróneo y puramente subjetivo.

BIBLIOGRAFIA

C.G.G AITKEN y D.A STONEY: THE USO OF STATISTICS IN FORENSIC SCIENCES (Ellis Horwood 1991)

A. P. DEMPSTER: “Upper and lower probabilities induced by a multivalued mapping” (Annals of Mathematical Statistics, 38 pág. 325-339, 1967).

S.F FIENBERG, ed.: THE EVOLVING ROLE OF STATISTICS ASSESSMENTS AS EVIDENCE IN THE COURTS (Springer-Verlag, 1989)

F. MONTES SUAY: LEY Y PROBABILIDAD (Las matemáticas y sus aplicaciones en el mundo social y económico, U.I.M.P. 2003)

J. ROSENMÜLLER: THE THEORY OF GAMES AND MARKETS (North-Holland, Amsterdam 1981)

G. SHAFER: A MATHEMATICAL THEORY OF EVIDENCE (Princeton University Press, 1976).