teorema
Vol. XXIV/2, 2005, pp. 27-41

Incertidumbre y grados de creencia

Julián Velarde Lombraña

ABSTRACT

The purpose of this paper is twofold: first I defined a certain type of uncertainty that appears in the context of phenomena that are described using exclusively subjective judgments. This type of uncertainty has been traditionally treated inside the framework of the so called "subjective (or Bayesian) probability", and more recently inside the Dempster-Shafer theory of evidence. In the second place, I discuss several formal models on a common basis of uncertainty measures and I propose the one extracted from the theory of Dempster-Shafer as a computational model for the representation and processing of the subjective probability understood as "degree of belief".

RESUMEN

El objetivo de este trabajo es doble: en primer lugar se acota un tipo de incertidumbre; la que aparece en el contexto de fenómenos que son descritos haciendo uso exclusivo de juicios subjetivos. Este tipo de incertidumbre ha sido tratado tradicionalmente dentro del marco de la llamada "probabilidad subjetiva" (o bayesiana), y más recientemente dentro de la teoría de la evidencia de Dempster-Shafer. En segundo lugar, sobre una base común de medidas de incertidumbre, se discuten varios modelos formales, y se propone el extraído de la teoría de Dempster-Shafer como un modelo computacional para la representación y tratamiento de la probabilidad subjetiva entendida como "grado de creencia".

INTRODUCCIÓN

Las nociones de *incertidumbre, grado de creencia* y de *plausibilidad* van ligadas, tanto en su significación como en el desarrollo de métodos para su tratamiento, a la noción de *probabilidad.* Noción ésta última interpretada, ya en los primeros tiempos de su desarrollo (Leibniz y Bernoulli) en varios sentidos.

Boudot (1967), en su estudio del *Ars Conjectandi* de Bernoulli, sigue manteniendo la distinción tradicional de dos sentidos de *probabilidad*/*posibilidad*, poniéndola en correspondencia con la distinción escolástica aplicada a la modalidad: (1) probabilidad (posibilidad) *de re*, llamada también *objetiva* u *ontológica*: la probabilidad (posibilidad) atribuida a los sucesos mismos; y (2) la probabilidad (posibilidad) *de dicto* o *subjetiva* o *epistemoló-*

gica: la probabilidad (posibilidad) atribuida al conocimiento (creencia) que tenemos de los sucesos. Hacking (1975) distingue asimismo entre posibilidad, en la línea de la llamada probabilidad frecuencialista o aleatoria (tipo Von Mises, Borel o Reichenbach) y posibilidad *de dicto*, como ligada a una probabilidad subjetiva o incluso a un estado de creencia. Con todo, esta distinción general —probabilidad subjetiva o epistemológica/probabilidad objetiva o frecuencialista— no es suficientemente precisa, como lo prueban las diversas significaciones que estas expresiones adquieren y las múltiples y confusas interferencias que con otras mantienen en distintas épocas. Así, por ejemplo, Hacking [*opus cit.*] subdistingue tres clases de probabilidad subjetiva: (a) el "personalismo" de De Finetti y Savage; (b) la probabilidad lógica o inductiva desarrollada por Keynes, Carnap y otros; (c) la probabilidad propugnada por los teóricos de la física, como Heisenberg. Pero, en tanto que Boudot [*opus cit.*] sostiene que Bernoulli es el fundador de la probabilidad subjetiva tipo (b), Hacking, al contrario, le sitúa en el grupo (c), reservando para Leibniz el honor de ser el primero en, si no desarrollar al menos, propugnar una lógica de las probabilidades.

La llamada probabilidad subjetiva ha sido, y sigue siendo, susceptible de múltiples interpretaciones, si bien suele haber acuerdo general en señalar como eslabones de su desarrollo a Ramsey (1931), seguido por De Finetti (1974), Good (1950) y Savage (1954). Durante los últimos decenios ha habido múltiples variaciones en las teorías de la probabilidad, si bien todas ellas toman como base para sus modelos formales la teoría axiomática. Por lo que respecta a su interpretación semántica, la probabilidad subjetiva (también llamada frecuentemente bayesiana) es usada en el contexto de fenómenos que son descritos haciendo uso exclusivo de juicios subjetivos. En este contexto, cabe entender la probabilidad como "grado de creencia".

Si los objetos de creencia son los enunciados o proposiciones, entonces una manera posible de tratar la incertidumbre es como una función, Cred., de las proposiciones, *p*, y los sujetos, *s*, al intervalo unidad [0,1], que representa el "grado de creencia" que el sujeto *s* tiene en la proposición *p*:

$$\text{Cred. (s, p)} \rightarrow [0,1]$$

Denotando el grado 0 la no creencia; el grado 1, la creencia plena; y los grados intermedios entre 0 y 1 la creencia parcial.

En cuanto a su interpretación, la frase "grados de creencia" es usada por diferentes autores en diferentes sentidos, principalmente en tres: (a) En sentido subjetivo, como intentando capturar el estado psicológico de un sujeto individual, con independencia de que el estado sea, o no, irracional. (b) Otras veces la teoría es construida como una teoría de la creencia subjetiva "ideal", esto es, no se trata de si tus grados de creencia son, o no, racionales, sino de si ellos "ajustan entre sí". (c) Finalmente, a veces la teoría es construida como una teoría general de la *creencia racional*, en la que podemos asegurar la racionalidad de los grados de creencia particulares (dada la evidencia que el agente tiene a su disposición), así como la racionalidad de su relaciones mutuas.

En todo caso, dejando a un lado la interpretación en cualquiera de los sentidos señalados, la función *Cred.* debe satisfacer los principios del cálculo de probabilidades clásico [Kolmogorov (1933)].

I. Grados de creencia

Shafer (1976) elaboró una teoría de la creencia basada en las ideas formuladas por Dempster (1966, 1967), y como una extensión de la teoría de la probabilidad subjetiva. Shafer parte de una crítica a la probabilidad bayesiana. El teorema de Bayes versa sobre la probabilidad de las causas o hipótesis H sobre un efecto o evidencia E. Así: si H_i (con i = 1, ..., n) denota las hipótesis; $P(H_i)$, nuestra probabilidad subjetiva *a priori* para cada hipótesis; y E, un efecto (una evidencia observada o una proposición verdadera), entonces la regla de Bayes establece cómo readaptar nuestra probabilidad *a priori* $P(H_i)$ a una probabilidad *a posteriori* $P(H_i{:}E)$:

$$P(H_i{:}E) = \frac{P(E{:}H_i)\ .\ P(H_{\cdot i})}{\Sigma_{j=1,\ldots n}\, P(E{:}H_j)\ .\ P(H_j)}$$

Aquí se asume que el espacio de probabilidad forma un conjunto de hipótesis, H_i, mutuamente exclusivas y totalmente exhaustivas sobre el mundo, siendo una, y sólo una, de ellas la exactamente verdadera, y que la medida de probabilidad P asigna un número a cada posible subconjunto E del espacio según el grado de confirmación de que ese subconjunto contiene la hipótesis, aunque desconocida, verdadera. Esto conlleva que si se utiliza nueva evidencia, el valor de probabilidad, generalmente, cambiará (excepto cuando la probabilidad es 0 ó 1), y que hay que recalcular el conjunto de *todas* las probabilidades. La teoría bayesiana exige, pues, para su desarrollo la asunción implícita de que la evidencia puede ser expresada siempre como una certeza, siendo así que la distribución de probabilidad *a priori*, presumiblemente basada en la evidencia, expresa, usualmente, grados de creencia, que no alcanzan la certeza. Shafer propone un esquema en el que una porción de creencia (en lugar de una masa de distribución bayesiana) puede ser asignada a una proposición, pero sin que ello necesariamente implique asignación en grado alguno a su negación. La principal crítica a la teoría bayesiana es que no puede manejar la ignorancia de manera efectiva, ya que si P(A) es la probabilidad de A, entonces la teoría exige la relación P(A) + P(A') = 1 (siendo A' el complemento de A). Y así no hay distinción entre falta de creencia y creencia y descreencia. La teoría de Dempster–Shafer no se adhiere a esta regla de la aditividad de la probabilidad bayesiana, y en su modelo asigna valores de probabilidad a conjuntos de posibilidades, en vez de a eventos singulares. Su teoría de las "funciones de creencia" se basa en dos ideas: (a) la idea de obtener grados de creencia con respecto a una cuestión a partir de las probabilidades subjetivas asignadas a alguna otra cuestión con ella relacionada; y (b) la regla de Dempster para combinar tales grados de creencia cuando están apoyados en puntos de evidencia independientes. El ejemplo típico para ilustrar estas ideas (ya aducido por Hooper en 1689) es la fiabilidad del testimonio: se trata de obtener grados de creencia de los enunciados hechos por los testigos a partir de las probabilidades subjetivas asignadas a la fiabilidad de esos testigos.

Supongamos que Juan me dice que han encontrado un manuscrito de Russell en la Biblioteca. La credibilidad que me merece (el grado de creencia que asigno a) su testimonio depende de mi "probabilidad subjetiva" de la fiabilidad de Juan. Si mi probabilidad subjetiva de que Juan es fiable es 0,7, entonces mi probabilidad subjetiva de que no es fiable es 0,3, sumando ambas, en tanto que probabilidades, 1. Pero el enunciado de Juan, que debe ser verdadero si Juan es fiable, no necesariamente es falso si él no es fiable. De manera que su testimonio solamente justifica en grado 0,7 de creencia que han encontrado un manuscrito de Russell en la Biblioteca, pero sólo en grado cero (no en grado 0,3) de creencia que no han encontrado un manuscrito de Russell en la Biblioteca. Por lo tanto, este grado cero no significa que estoy seguro de que no han encontrado un manuscrito de Russell en la Biblioteca, como ocurre con las probabilidades; significa, simplemente, que el testimonio de Juan no me ofrece razones para creer que no han encontrado un manuscrito de Russell en la Biblioteca. Los grados 0,7 y 0 constituyen juntos una función de creencia.

La teoría de Shafer busca utilizar las funciones de creencia para valuar numéricamente la masa de evidencia en los problemas prácticos. En el ejemplo anterior partimos de dos cuestiones:

C_1 ¿Es fiable Juan?
C_2 ¿Han encontrado un manuscrito de Russell en la Biblioteca?

Y, partiendo de la probabilidad subjetiva asignada a C_1, derivamos grados de creencia para C_2.

Un único testimonio respecto de una proposición específica (como en el ejemplo) induce una función de creencia relativamente simple (aquella que da un determinado grado de creencia a esa proposición y sus consecuencias, y cero grados de creencia a todas las demás proposiciones. Éste es un caso simple: C_1 y C_2 solamente tienen dos posibles respuestas: sí/no. Pero en casos más complejos tendremos cuestiones C_1 y C_2 con muchas posibles respuestas.

Se parte, entonces, de que cada cuestión se presenta con una lista exhaustiva de respuestas mutuamente exclusivas, sabiendo que exactamente una de ellas es la correcta, aunque no sabemos cuál. El conjunto de respuestas constituye un *esquema*. Tenemos, pues, un esquema S para C_1 y un esquema T para C_2. Sea, ahora: P(s) la probabilidad del elemento *s* de *S*. Dadas esas probabilidades, y dado un subconjunto A de T, buscamos derivar un grado de creencia de A, Cred(A) (nuestro grado de creencia de que A contiene la respuesta correcta a C_2). Una respuesta s a C_1 puede determinar un completo conjunto de respuestas a C_2. Si A es un conjunto de respuestas a C_2, y s desecha todas las respuestas en el complemento de A, i. e., en T – A, entonces s indica que la respuesta a C_2 está en alguna parte de A, de manera que P(s) contribuirá a nuestra creencia en A. Nuestro grado de creencia en A, Cred(A), será la probabilidad total para todas las respuestas s que desechan todas las respuestas en T – A. A la formulación de esto va encaminada la regla de Dempster.

La regla de combinación de Dempster se basa en la idea de independencia probabilista, en tanto que aplicada a cuestiones con respecto a las cuales tenemos probabilidades subjetivas. Así, por ejemplo, puedo usar la regla para combinar evidencia a partir de dos testimonios, si considero la fiabilidad del primer testimonio subjetivamente independiente (antes de tomar en cuenta lo que el testimonio dice) de la fiabilidad del segundo. La regla usa esta independencia subjetiva para evaluar las probabilidades conjuntas de las diversas posibilidades de fiabilidad de ambas. Esto exige, por tanto: (a) independencia de fiabilidad; y (b) ver si hay que desechar algunas posibilidades. Por ejemplo: si Juan me dice que han encontrado un manuscrito de Russell en la Biblioteca y Luis me dice que no han encontrado un manuscrito de Russell en la Biblioteca, entonces ambos testimonios no pueden ser fiables, por ser contradictorios; su fiabilidad ya no es subjetivamente independiente para mí. Cumplidas (a) y (b), se normalizan las probabilidades de las restantes posibilidades por referencia a la unidad (esto es, haciendo su suma igual a 1). Y se usan las probabilidades así normalizadas para obtener nuevos grados de creencia.

La regla de combinación de Dempster busca, pues, aprehender la idea de que: (a) puntos de fiabilidad concordantes se refuerzan mutuamente; (b) puntos en conflicto se destruyen entre sí; y (c) una cadena de razonamientos es más débil que sus eslabones más débiles. Siguiendo con el ejemplo anterior: C_1 (¿es fiable Juan?) y C_2 (¿han encontrado un manuscrito de Russell en la Biblioteca?); siendo S = conjunto de respuestas a C_1 y T = conjunto de respuestas a C_2. Y sea Γ(s) un subconjunto de T.

Ahora, si tenemos dos funciones de creencia $Cred_1$ y $Cred_2$ sobre T basadas en puntos independientes de evidencia, i. e., cada una de ellas está basada en un espacio de probabilidad, entonces la regla de Dempster establece cómo formar el espacio de probabilidad producto, eliminando los pares que tienen aplicación sobre conjuntos disjuntos de T, y obtener, entonces, una función de creencia mediante la aplicación Γ de cada par restante a la intersección de los subconjuntos a los que los dos elementos del par son aplicados. La aplicación Γ de S a T establece, para cada elemento $s \in S$, qué elementos $t \in T$ son posibles respuestas a C_2, si s es la respuesta correcta a C_1. En otras palabras: Γ especifica el conjunto D de todos los pares ordenados (s,t) tales que s es compatible con t. D es, pues, una relación de compatiblidad (un subconjunto del producto cartesiano de $S \times T$), que viene referida a la aplicación Γ por :

$$D = \{(s,t) / t \in \Gamma(s)\}$$

y

$$\Gamma(s) = \{t / (s,t) \in D\}$$

II. EL MODELO FORMAL DE LAS FUNCIONES DE CREENCIA

Sea el sistema formal siguiente:

(1) Una colección E de elementos: $e_1, e_2, ..., e_n$
(2) El conjunto potencia de E, esto es, $\wp(E)$
(3) Una familia (campo conjunto) de subconjuntos de E, esto es: $\Phi(E)$.

De manera más precisa: dado un conjunto I, se dice que hay una familia que tiene a I como conjunto de índices (o que hay una familia cuyos índices recorren I) cuando para cada $i \in I$, existe un elemento e_i que depende (es función) de *i*, denotándose dicha familia $(e_i)_{i\in I}$. Dada una familia $(e_i)_{i\in I}$, se dice que es una *familia de elementos del conjunto E*, si para todo $i \in I$ resulta que $e_i \in E$. Y dada una familia $(A_i)_{i\in I}$, se dice que es una *familia de partes del conjunto E* si para todo $i \in I$ resulta que $A_i \subseteq E$. La *familia de partes del conjunto E*, Φ (E), se distingue, a su vez, de *las partes (= subconjuntos) del conjunto E* (o conjunto potencia de E): $\wp(E) = \{A / A \subseteq E\}$), y también de las partes A resultantes de la *partición* ∇ de E: $\nabla(E) = \{A / A \neq \varnothing$, y para todo $e \in E$ se cumple que *e* pertenece a un solo A, y la reunión de las partes A es igual a E$\}$, i.e., *partición* de E es toda familia $(A_i)_{i\in I}$ de partes no vacías de E tales que:

(1) $A_i \cap A_j = \varnothing \ \forall$ i, j $1 \leq i, j \leq n$
(2) $\bigcup_{i\in I} A_i = E$

De las distintas relaciones partes/todo, resulta:

Φ (E) $\subseteq \wp(E)$; $\nabla(E) \subseteq \wp(E)$

II.a. *Espacios medibles*

Sea una familia de subconjuntos de E: Φ (E) $\subseteq$ $\wp$(E), que posee la siguiente estructura algebraica:

(A) Por ser parte de $\wp$(E) es un *conjunto parcialmente ordenado*, ya que:

(I) $\langle \wp(E); \subseteq \rangle$ es un conjunto parcialmente ordenado: en efecto, $\subseteq$ es una *relación* binaria sobre $\wp$(E) que satisface, para todo a, b, c $\in$ $\wp$(E), las tres condiciones siguientes:

1. $a \subseteq a$	reflexividad
2. $a \subseteq b$ & $b \subseteq a \rightarrow a = b$	antisimetría
3. $a \subseteq b$ & $b \subseteq c \rightarrow a \subseteq c$	transitividad

(II) Si $\wp$(E) es un conjunto parcialmente ordenado y Φ (E) $\subseteq$ $\wp$(E), entonces $\langle \Phi (E) ; \subseteq \rangle$ es también un conjunto parcialmente ordenado.

Φ (E) posee, además, la estructura algebraica siguiente:

1. Es un *retículo*: $\Re = \langle \Phi (E) ; \cap, \cup \rangle$, en donde $\cup$ y $\cap$ son operaciones binarias sobre Φ (E) que satisfacen, $\forall$a, b, c $\in$ Φ (E), las leyes siguientes:

1a. $a \cup a = a$; $a \cap a = a$	idempotencia
1b. $a \cup b = b \cup a$; $a \cap b = b \cap a$	conmutativa
1c. $a \cup (b \cup c) = (a \cup b) \cup c$; $a \cap (b \cap c) = (a \cap b) \cap c$	asociativa
1d. $a \cap (a \cup b) = a$; $a \cup (a \cap b) = a$	absorción

2. Es un *retículo distributivo*, ya que, $\forall$a, b, c $\in$ Φ (E) , se cumple:

2a. $a \cup (b \cap c) = (a \cup b) \cap (a \cup c)$; $a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$

3. Es, además, un *retículo booleano*: $\Re = \langle \Phi (E) ; \cup, \cap, ' , 0, 1 \rangle$, en donde ' es una operación monaria (*complementación*); 0 y 1 son dos elementos pertenecientes a Φ (E) (llamados, elemento *mínimo* y elemento *máximo*), que satisfacen las siguientes condiciones:

3a. $a \cup 0 = a$; $a \cap 0 = 0$; $a \cup 1 = 1$; $a \cap 1 = a$	Elemento neutro
3b. $a \cup a' = 1$; $a \cap a' = 0$	Ley de complementación

4. Φ (E) es un *álgebra de subconjuntos* cuando

4a. $\varnothing \in \Phi$ (E)
4b. a $\in \Phi$ (E) $\rightarrow$ a' $\in \Phi$ (E)
4c. a, b $\in \Phi$ (E) $\rightarrow$ a $\cup$ b $\in \Phi$ (E)

Los todos así constituidos se denominan *espacios* (o *campos*) *medibles*, y a sus elementos se les denomina conjuntos *medibles.*

II.b. *Funciones de probabilidad*

Un todo (o espacio) de probabilidad es una familia de partes del conjunto E (conjunto de observables): elementos (*puntos muestrales*), e_i, o subconjuntos (*sucesos*), A_i, tal como está definida en las condiciones 1-4, y haciendo corresponder: a E, el suceso cierto; a $\varnothing$, el suceso imposible; y a las operaciones ', $\cup$, $\cap$, respectivamente, las conectivas lógicas *no*, *o* e *y*. La familia Φ (E) así estructurada se llama *álgebra de sucesos*. (Las particiones de conjuntos son también —y, además, específicos— todos de probabilidad).

Una *medida de probabilidad* P es una función numérica sobre un espacio medible Φ (E), i.e., P es una aplicación de los subconjuntos A de Φ (E) en el intervalo [0, 1]. P: Φ (E) $\rightarrow$ [0, 1] tal que,

1. $0 \leq P(A) \leq 1 \ \forall A \in \Phi$ (E) 2.1
2. $P(E) = 1$ 2.2
3. $P(A \cup B) = P(A) + P(B) \ \forall A, B \in \Phi (E) / A \cap B = \varnothing$ 2.3

o en fórmula más general:

Si $\{A_1, A_2, \ldots, A_n\} \subseteq \Phi$ (E) es una colección de conjuntos que son disjuntos (incompatibles) dos a dos, entonces:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) \qquad 2.4$$

La medida de probabilidad posee, pues, unas propiedades que dependen (son función) de la naturaleza (estructura) de los elementos sobre los que se aplica. Se llama *espacio de probabilidad* a la terna < E, Φ (E), P > en la que E es un conjunto, Φ (E) una álgebra de subconjuntos (de sucesos) de E, y P una medida de probabilidad definida sobre ese álgebra.

II.c. *Funciones de creencia*

Sea E un conjunto finito, y $\wp$(E) el conjunto de sus partes. Sea P una aplicación funcional de una aplicación funcional $\wp$(E) en [0,1], tal que:

1. $P(\emptyset) = 0$ 2.5
2. $\Sigma_{A\subseteq \wp(E)} P(A) = 1$ 2.6

P es una función de probabilidad en el conjunto de las partes de E, por lo que la familia probabilizable es el conjunto de las partes de E. Ahora, a partir de P definamos para todo $A \in \wp(E)$ una función, representada como Cred(A), tal que:

$$Cred(A) = \Sigma_{x\subseteq A} P(X) \quad 2.7$$

Se dirá, entonces, que Cred(A) es una *función de credibilidad* en $\wp(E)$.

En un retículo booleano $\wp(A)$, la familia de los $X \subseteq A$ tales que $P(X) > 0$ se denomina *soporte* (elementos *focales*) de Cred(A), y se representa:

$$S_{Cred(A)} = \{X\in \wp(A) \mid P(A) > 0\} \quad 2.8$$

Y una función de credibilidad, Cred(A), con $A\in \wp(E)$, se dice que tiene *soporte simple* si:

$P(E) = 1 - s$, $s\in [0,1]$ 2.9
$P(A) = s$ 2.10
$P(\wedge) = 0, \wedge \neq E, \wedge \neq A, \wedge \subset E$ 2.11

Esto es, en un retículo booleano $\wp(A)$ todos los subconjuntos distintos de E y de A tienen como valores: $P(\wedge) = 0$, y P(E) y P(A) tienen valores complementarios. Y el par < S, P > , en donde S denota el conjunto de todos los elementos focales inducidos por P, es llamado un *cuerpo de evidencia*.

Ejemplo:

Sea E = {a, b, c}. Por tanto,
$\wp(E)$ = { {a,b,c} {a,b} {a,c} {b,c} {a} {b} {c} Ø }
Sean las siguientes probabilidades (se comprueba que se cumple 2.6):
P({a,b,c}) = 0; P({a,b}) = 0,18; P({a,c}) = 0,52; P({b,c}) = 0; P({a}) = 0,03; P({b}) = 0,07; P({c}) = 0,20; P(Ø) = 0.
Sea A = {b,c}, entonces:
Cred(A) = P({b,c}) + P({b}) + P({c}) = 0,27.
El soporte de Cred(A) es {b} y {c}.

Las propiedades principales de las funciones Cred. son:

1. $Cred(\emptyset) = 0$, ya que $P(\emptyset) = 0$ y $\Sigma_{\emptyset} P(\emptyset) = 0$ 2.12.
2. $Cred(E) = 1$, ya que, por 2.7, $Cred(E) = \Sigma_{x\subseteq E} P(X) = 1$ 2.13

3. $\forall$ A, B $\in \wp(E)$:
$\mathrm{Cred}(A \cup B) \geq \mathrm{Cred}(A) + \mathrm{Cred}(B) - \mathrm{Cred}(A \cap B)$ 2.14

(sin olvidar que debe cumplirse: $\mathrm{Cred}(A) + \mathrm{Cred}(B) - \mathrm{Cred}(A \cap B) \leq 1$). De manera general, 2.14 deviene:

$\mathrm{Cred}(A_1 \cup A_2 \cup ... \cup A_n) \geq \Sigma_{I:\emptyset \neq I \subseteq \{1,2,\ldots,n\}} (-1)^{(\mathrm{Card.}\ I)+1} . \mathrm{Cred}(\cap Ai)_{i \in I}$ 2.15

4. $\forall$ A, B $\in \wp(E)$:
$A \subseteq B \Rightarrow \mathrm{Cred}(A) \leq \mathrm{Cred}(B)$ 2.16

Esta propiedad (monotonía de la inclusión) conlleva otras importantes propiedades, y en ella se apoya la idea de ampliar la noción de Cred a los conjuntos difusos.

5. A toda función P de $\wp(E)$ en [0,1] se puede hacer corresponder una, y solamente una, función Cred. de $\wp(E)$ en [0,1]; esto es:

$\forall$ A $\in \wp(E)$:
$P(A) = \Sigma_{X \subseteq A} (-1)^{(\mathrm{Card.}\ A\ -\ \mathrm{Cad.}\ X)} . \mathrm{Cred}(X)$ 2.17

De esta manera, se establece una biyección entre las funciones de probabilidad, P, y las funciones de creencia, Cred.

II.d. *Combinación de funciones de creencia: la regla de Dempster*

Sean dos funciones de creencia, Cred_1 y Cred_2, que satisfacen 2.5 y 2.6, sobre el mismo espacio de probabilidad $\wp(E)$. La regla de Dempster establece cómo obtener una función de credibilidad, Cred, como resultado de combinar Cred_1 y Cred_2 , en los pasos siguientes:

(1) Formar el espacio de probabilidad producto, a partir de los soportes de Cred_1 y Cred_2. Esto es, la regla agrega dos cuerpos de evidencia, $P_1(A_i)$ y $P_2(B_j)$, definidos sobre el mismo esquema de discernimiento en un solo cuerpo de evidencia P(D):

$\forall$ D $\in \wp(E)$:
$P(D) = \Sigma_{Ai \cup Bj = D} P_1(A_i) . P_2(B_j)$ 2.18

En donde $A_i \in \wp(E)$ y $B_j \in \wp(E)$.

(2) A partir de P(D) obtenemos una nueva función de creencia, Cred, según 2.7,

$$\text{Cred}(A) = \Sigma_{x \subseteq A} P(D) \qquad 2.19$$

Ejemplo:

Sean dos funciones de probabilidad P_1 y P_2 sobre el mismo espacio probabilístico $\wp(E)$. Y sea:

$E = \{a, b, c\}$
$\wp(E) = \{ \{a,b,c\} \{a,b\} \{a,c\} \{b,c\} \{a\} \{b\} \{c\} Ø \}$
$P_1(E)$ la distribución de probabilidad P_1 en $\wp(E)$
$P_2(E)$ la distribución de probabilidad P_2 en $\wp(E)$
$P(E)$ resultado de $P_1 . P_2$
$\text{Cred}_1(E)$ función de creencia inducida por $P_1(E)$
$\text{Cred}_2(E)$ función de creencia inducida por $P_2(E)$
$\text{Cred}(E)$ resultado de combinar Cred_1 con Cred_2.

Tal como se indica en la tabla siguiente:

$\wp(E)$	$P_1(E)$	$P_2(E)$	$\text{Cred}_1(E)$	$\text{Cred}_{2(E)}$	$P(E) = P_1 . P_2$	$\text{Cred}(E) = \text{Cred}_1 o \text{Cred}_2$
Ø	0	0	0	0	0	0
{a}	0,03	0,05	0,03	0,05	0,0015	0,0015
{b}	0,07	0,05	0,07	0,05	0,0035	0,0035
{c}	0,10	0,10	0,10	0,10	0,0100	0,0100
{a,b}	0,20	0,30	0,30	0,40	0,1150	0,1200
{a,c}	0,25	0,20	0,38	0,35	0,1215	0,1330
{b,c}	0,35	0,25	0,52	0,40	0,1945	0,2080
{a,b,c}	0	0,05	1	1	0,5540	1

Propiedades relativas a la combinación de las funciones de creencia son:

$$\text{Cred}(A) + \text{Cred}(A') \leq 1 \qquad 2.20$$

En efecto, si por 2.7 tenemos $\text{Cred}(A') = \Sigma_{x \subseteq A'} P(X)$, entonces:

$$\text{Cred}(A) + \text{Cred}(A') = \Sigma_{x \subseteq A} P(X) + \Sigma_{x \subseteq A'} P(X).$$

Por tanto, la unión del conjunto de los subconjuntos de A y el conjunto de los subconjuntos de A' da lugar a un subconjunto de $\wp$(E). De ahí que, si Cred(E) = 1, entonces Cred. de un subconjunto suyo sea ≤ 1.

II.e. *Funciones de plausibilidad*

Una función de plausibilidad Pl. sobre el espacio $\wp$(E), en el que está definida una función de creencia Cred., viene definida así:

$$Pl(A) = 1 - Cred(A') \qquad 2.21$$

Con las propiedades siguientes: Por 2.7 la la fórmula anterior se convierte en:

$$Pl(A) = 1 - \Sigma_{x\subseteq A'} P(X) \qquad 2.22$$

Y, por tanto,

$$Pl(A) = \Sigma_{B:B\cap A\neq\varnothing} P(B) \qquad 2.23$$

Y por 2.15:

$$Pl(A_1 \cap A_2 \cap \ldots \cap A_n) = \Sigma_{I:\varnothing\neq I\subseteq\{1,2,\ldots,n\}} (-1)^{(Card.\ I)+1} . Pl(\cup Ai)_{i\in I} \qquad 2.24$$

Por lo tanto, 2.17 se convierte en:

$$P(A) = \Sigma_{X\subseteq A} (-1)^{(Card.\ A\ -\ Cad.\ X)} . 1 - Pl(A')) \qquad 2.25$$

Y, finalmente, 2.20 se convierte en:

$$Pl(A) + Pl(A') \geq 1 \qquad 2.26$$

Ejemplo sobre E = {a, b, c} en la tabla siguiente:

$\wp$(E)	P	Cred	Pl
∅	0	0	0
a	0,2	0,2	1
b	0	0	0,5
c	0	0	0,5
a, b	0,3	0,5	1
a, c	0,3	0,5	1
b, c	0	0	0,8
E	0,2	1	1

De la comparación entre las funciones de credibilidad (o confianza), probabilidad y plausibilidad —y sus propiedades— resulta:

$\forall$ A, B $\in \wp(E)$:
Cred(A $\cup$ B) $\geq$ Cred(A) + Cred(B) - Cred(A $\cap$ B) 2.27
P(A $\cup$ B) = P(A) + P(B) - P(A $\cap$ B) 2.28
Pl(A $\cup$ B) $\leq$ Pl(A) + Pl(B) - Pl(A $\cap$ B) 2.29

Y con respecto a la complementación:

Cred(A) + Cred(A') $\leq$ 1 2.30
P(A) + P(A') = 1 2.31
Pl(A) + Pl(A') $\geq$ 1 2.32

Según esto, es posible, en caso de ignorancia, que se dé: Cred(A) = Cred(A') = 0; y Pl(A) = Pl(A') = 1. Esto es, dos enunciados opuestos pueden aparecer, a la vez, como plausibles sin ser en grado alguno creíbles. Pero con este modelo de las funciones de creencia es posible distinguir entre la falta de creencia y la descreencia, ya que (falta de creencia) Cred(A) = 0 no implica que Cred(A') = 1, lo que implicaría, a su vez que Pl(A) = 0. Mientras que (descreencia) Cred(A') = 1 implica que Cred(A) = 0.
Y, finalmente:

$\forall$ A $\in \wp(E)$:
0 $\leq$ P(A) $\leq$ Cred(A) $\leq$ Pl(A) $\leq$ 1 2.33

Esto es, la plausibilidad de un enunciado es siempre igual o mayor que su credibilidad.

II.f. *Posibilidad y necesidad*

Cuando, en un retículo booleano de las partes de un conjunto E: $\wp(E)$, en el que está definida una función de credibilidad, los elementos focales de $\wp(E)$, como definidos en 2.8, forman una "estructura nidificada", las funciones de credibilidad y plausibilidad , como definidas en 2.7 y 2.21, se reducen, respectivamente, a una medida de necesidad y a una medida de posibilidad.
Los elementos A_i (i=1,...,n) $\in \wp(E)$ *focales* forman una estructura nidificada: $< A_i ; \wp(E) ; \subseteq >$ si pueden ser ordenados así:

$A_1 \subseteq A_2 \subseteq ... \subseteq A_n$

Se dice, entonces, que los elementos focales A_i son *consonantes*, en cuyo caso se cumple:

$\forall$ A, B $\in \wp$(E) :
Cred(A $\cap$ B) = mín. { Cred(A), Cred(B) } 2.34
Pl(A $\cup$ B) = máx. { Pl(A), Pl(B) } 2.35

Y las funciones de credibilidad / plausibilidad que cumplen, respectivamente, 2.34 y 2.35 son llamadas, respectivamente, medidas de necesidad N y de posibilidad Π, por lo que 2.21 se convierte en:

$\forall$ A, B $\in \wp$(E) :
N(A) = 1 - Π(A')

Esto es, la necesidad de que A es verdadero corresponde a la imposibilidad de que A es falso, viniendo medida la imposibilidad por el complemento para 1 de la medida de posibilidad.

III. CONCLUSIONES

La probabilidad subjetiva, en tanto que se aplica a fenómenos descritos haciendo uso de juicios subjetivos, es susceptible de ser interpretada (entendida) como grado de creencia. Si los objetos de creencia son los enunciados o proposiciones, cabe tratar la incertidumbre subjetiva generada por esas proposiciones como una función (función de credibilidad) de las proposiciones *p* y de los sujetos *s* en el intervalo unidad.

El modelo formal propuesto busca cómo computar: (1º) las funciones de creencia simples; y (2º) la combinación de funciones, lo que permite cuantificar la amplitud de evidencia que conlleva un mensaje complejo y no susceptible de descomposición en componentes independientes.

Departamento de Filosofía
Universidad de Oviedo
c/ Tte. Alfonso Martínez s/n, E-33071 Oviedo
E-mail: velarde@uniovi.es

REFERENCIAS BIBLIOGRAFICAS

BOUDOT, P. M. (1967), "Probabilité et logique de l'argumentation selon Jacques Bernoulli", en *Les Études Philosophiques*, 28, pp. 265-288.
DE FINETTI, B. (1974), *Theory of Probabiliy*, Wiley, Nueva York.

DEMPSTER, A. P.(1966), "New methods for reasoning towards posterior distributions based on sample data", en *Annals of Mathematical Statistics*, 37, pp. 355-374.

— (1967) "Upper and lower probabilities induced by a multivalued mapping", en *Annals of Mathematical Statistics*, 38, pp. 325-339.

GOOD, I. J. (1950), *Probability and the Weighing of Evidence*, Hafner's, Nueva York.

HACKING, I. (1975), *The Emergence of Probability*, Cambridge University Press, Cambridge.

KOLMOGOROV, A. N. (1956), *Foundations of the Theory of Probability*, (primera edición alemana, 1933); segunda edición Chelsea, Nueva York.

RAMSEY, P. (1931), *The Foundations of Mathematics and Other Essays*, Humanities Press, Nueva York.

SHAFER, G. (1976), *A Mathematical Theory of Evidence*, Princeton University Press, Princeton.

— (1990), "Perspectives on the theory and practice of belief functions", en *International Journal of Approximate Reasoning*, 4, pp. 323-362.

SAVAGE, L. J.(1954), *The Foundations of Statistics*, Wiley, Nueva York.