# CAPITULO 1

## TEORÍA EVIDENCIAL

- **La Teoría Evidencial de Dempster y Shafer**
- **Combinación de Evidencias en la Teoría de Dempster y Shafer**
- **Credibilidad, Plausibilidad e Intervalo de Confianza**
- **Casos Particulares de la Teoría Evidencial**
- **Resumen**
- **Textos básicos**

# 1. TEORÍA EVIDENCIAL

A diferencia del modelo de los factores de certidumbre de Shortliffe y Buchanan, el esquema de razonamiento propuesto en su día por Dempster y Shafer sí tiene una fuerte base teórica, hasta el punto que, lo que inicialmente era un modelo de razonamiento propuesto por Dempster, se convirtió en una verdadera teoría tras la formalización de Shafer.

Este esquema de razonamiento es atractivo, entre otras razones, porque:

- permite modelizar de forma sencilla la incertidumbre asociada a evidencias e hipótesis.
- permite considerar conjuntos de hipótesis sin que la confianza depositada en cada uno de ellos tenga que ser distribuida de ningún modo entre cada una de las hipótesis individuales del conjunto.
- permite reflejar de forma elegante la falta de conocimiento tan frecuentemente ligada a los procesos de razonamiento.
- contiene a la teoría de la probabilidad como un caso particular.
- contiene a algunas de las funciones combinatorias de evidencias del modelo de los factores de certidumbre.

Pero... ¿cómo se puede manejar el conocimiento inexacto y la falta de conocimiento en el modelo de Dempster y Shafer?

## 1.1. La Teoría Evidencial de Dempster y Shafer

En primer lugar, y dado un universo de discurso cualquiera, Dempster y Shafer introducen el concepto de *marco de discernimiento*, que definen como "... el conjunto finito de todas las hipótesis que se pueden establecer en el dominio del problema". El marco de discernimiento debe formar un conjunto completo, y por tanto exhaustivo, de hipótesis mútuamente excluyentes.

Por otra parte, el efecto de una determinada evidencia sobre el conjunto global de hipótesis no viene determinado por la contribución de la confianza depositada en las hipótesis individuales. Por el contrario, el efecto de cada evidencia afecta generalmente a un subconjunto de hipótesis del marco de discernimiento. Este planteamiento es coherente con la realidad de casi todos los problemas cotidianos[44]. En problemas reales, lo más normal es que las evidencias permitan discriminar entre grupos de hipótesis

[44] ... recuérdese que ya hemos comentado lo difícil que resulta encontrar evidencias que confirmen a una y sólo una hipótesis.

alternativas, manteniéndose, sin embargo, la incertidumbre entre las hipótesis individuales. Según este planteamiento:

- $\theta$ es el marco de discernimiento
- $A$ es un subconjunto cualquiera del marco
- $h1, ..., hn$ son las hipótesis del marco

Ahora podemos establecer fácilmente el conjunto $\Gamma\theta$ de todos los subconjuntos posibles del marco. En este contexto, la aparición de una determinada evidencia $e$ favorecerá a un determinado subconjunto $A$ de $\theta$, de forma que el grado en que $A$ se vea favorecido se representa por $m(A)$, en donde $m$ es indicativo de la confianza que la evidencia $e$ permite depositar en $A$. $m$ se denomina *función básica de asignación de verosimilitud*, y toma valores en el intervalo cerrado [0, 1]. Al respecto, utilizaremos la siguiente notación:

$$e: A = \{ha, hb, hc\} \rightarrow m(A) = x, \text{ con } x \in [0, 1]$$

El hecho de que la evidencia $e$ apoye al subconjunto $A$ no implica, como ya hemos dicho, que las hipótesis individuales se repartan de forma explícita la confianza depositada en la propia $A$. Esto constituye una diferencia notable con respecto a la teoría clásica de la probabilidad, según la cual, si $h1$, $h2$, $h3$ y $h4$ son las cuatro únicas hipótesis posibles en un dominio dado:

Si $p(h1, h2) = 0.80$

de alguna manera estamos afirmando que:

$p(h1) = 0.40$

$p(h2) = 0.40$

Este mismo ejemplo contemplado bajo la óptica de la teoría evidencial tendría el siguiente tratamiento:

$\theta \;=\; \{ h1, h2, h3, h4 \}$

$\Gamma\theta \;=\; \{ø, (h1), (h2), (h3), (h4), (h1, h2), (h1, h3), (h1, h4), (h2, h3), (h2, h4), (h3, h4), (h1, h2, h3), (h1, h2, h4), (h1, h3, h4), (h2, h3, h4), (h1, h2, h3, h4) \}$

$\Gamma\theta$ contiene a los 16 subconjuntos posibles que se pueden establecer con las cuatro hipótesis iniciales del marco de discernimiento. Nótese que en $\Gamma\theta$ están incluidos el conjunto vacío {ø}, y el propio marco $\{ \theta = (h1, h2, h3, h4) \}$

Si elegimos un subconjunto $A$ del marco, por ejemplo $A = (h1, h2, h3)$, de tal manera que:

$$e: A = (h1, h2, h3) \rightarrow m(A) = 0.75,$$

lo único que se afirma es que, dada la evidencia $e$, la verosimilitud de $A$ es 0.75. Es claro que ninguna de las hipótesis individuales de $A$ se ve afectada por esta asignación de verosimilitud, ya que cualquier otra hipótesis, o cualquier otro conjunto de hipótesis son, en realidad, subconjuntos diferentes de $\Gamma\theta$, independientes -en principio- de la evidencia $e$.

Todo subconjunto del marco de discernimiento para el cual, dada una evidencia $e$, se verifique que $m(A) \neq 0$, se denomina *elemento focal*.

Volviendo por un instante a la función básica de asignación de verosimilitud, Dempster y Shafer definen las siguientes condiciones para $m$:

- $\sum_{A \subset \phi} m(A) = 1$
- $m(\varnothing) = 0$

Ambas condiciones son consecuencia de las restricciones impuestas al marco de discernimiento[45].

Decíamos también que la teoría evidencial proporciona un medio elegante para tratar la falta de conocimiento asociada a los procesos de razonamiento. Supongamos un marco de discernimiento $\theta$ y una evidencia tal que:

$$e: A \subset \theta \ \rightarrow m(A) = s, \text{ con } 0 \leq s \leq 1$$

La primera condición exigida a $m$ establece que $\sum_{A \subset \phi} m(A) = 1$, entonces ...¿qué pasa con el resto de confianza que no ha sido asignada al elemento focal $A$?

Al respecto, Dempster y Shafer postulan que:

Si: $e: A \subset \theta \rightarrow m(A) = s$, con $0 \leq s \leq 1$
Entonces: $m(\theta) = 1 - m(A) = 1 - s$

Esta formulación debe interpretarse del siguiente modo: puesto que la evidencia $e$ supone la asignación de una confianza dada a un determinado elemento focal $A$ del marco, el resto de la confianza no asignada representa "falta de conocimiento" y, por lo tanto, debe ser asignada al propio marco de discernimiento. Con estas premisas podemos reflexionar del siguiente modo:

- La confianza no asignada es ignorancia o falta de conocimiento sobre el grado de importancia de la evidencia en relación al elemento focal considerado. En otras palabras, se sabe que la evidencia apoya al elemento focal en un grado $s$; sin embargo, la confianza no asignada ($1-s$), no sabemos si contribuye o no a $A$ (o a cualquier otro subconjunto del marco.)

45 Recuérdese que el marco debe ser un conjunto completo de hipótesis mutuamente excluyentes.

- La confianza no asignada (1-$s$) debe asignarse al marco ya que, por construcción del esquema, lo que sí sabemos es que la solución está en el marco.

La formulación completa de la aproximación es la siguiente:

$\theta$ = marco de discernimiento = $\{h1,..., hn\}$
$A$ = elemento focal, $A \subset \theta$
$e$ = evidencia referida a $A$
$m(A)$ = medida de asignación básica de verosimilitud de $A$, dado $e$
$e: A \rightarrow m(A) = s$
$m(\theta) = 1 - s$
$m(B) = 0 \qquad \forall B \subset \theta, B \neq \theta, B \neq A$

Si el planteamiento fuese probabilístico, la misma evidencia apoyaría al elemento focal $A$ y al complementario del elemento focal:

$$p(A) = s \rightarrow p(\neg A) = 1 - s$$

y recordemos que éste era uno de los aspectos más débiles de los modelos probabilísticos. Con esta nueva teoría:

Si: $\theta = \{h1, h2, h3, h4\}$
y: $A = \{h1, h2\}$
con: $e: A \rightarrow m(\{h1, h2\}) = s$
entonces: $m(\{h1, h2, h3, h4\}) = 1 - s$

Generalizando un poco estas ideas puede afirmarse, sin equivocarnos demasiado, que el procedimiento según el cual se maneja la falta de información en la teoría evidencial corrige las carencias de los modelos probabilísticos.

Siguiendo el mismo esquema de presentación que en temas anteriores, consideraremos ahora el modo en que las evidencias aparecen en los problemas del mundo real, y cómo la teoría evidencial trata este tipo de situaciones.

Parece claro que en problemas reales las evidencias no vienen solas. Más aún, distintas evidencias no necesariamente tienen por qué referirse a los mismos elementos focales. Así si $\theta = \{h1, h2, h3, h4\}$ es nuestro marco de discernimiento, o conjunto completo de hipótesis mútuamente excluyentes posibles en nuestro dominio, el conjunto de todos los subconjuntos posibles del marco será:

$\Gamma\theta = \{$ ø, $(h1), (h2), (h3), (h4), (h1, h2), (h1, h3), (h1, h4), (h2, h3),$
$(h2, h4), (h3, h4), (h1, h2, h3), (h1, h2, h4), (h1, h3, h4),$
$(h2, h3, h4), (h1, h2, h3, h4) \}$

Evidentemente, el número de elementos de $\Gamma\theta$ es 2 elevado al número de elementos del marco: $\#\Gamma\theta = 2^{\#\theta}$. En nuestro caso 16.

Sea ahora una evidencia $e1$ tal que:

$$e1 \colon A1 = (h1, h2) \text{ con } m1(h1, h2) = x$$

por lo tanto, $m1(h1, h2, h3, h4) = 1 - x$

Sea ahora una nueva evidencia $e2$ tal que:

$$e2 \colon A2 = (h2, h3, h4) \text{ con } m2(h2, h3, h4) = y$$

por lo tanto, $m2(h1, h2, h3, h4) = 1 - y$

Finalmente, sea una tercera evidencia $e3$ tal que:

$$e3 \colon A3 = (h1) \text{ con } m3(h1) = z$$

por lo tanto, $m3(h1, h2, h3, h4) = 1 - z$

Nótese que cada elemento focal se refiere a cada evidencia concreta. La cuestión estriba ahora en considerar el efecto conjunto de todas las evidencias... pero antes tenemos que saber algo más de la teoría evidencial.

## 1.2. Combinación de Evidencias en la Teoría de Dempster y Shafer

Si dos (o más) fuentes de información proporcionan sendas evidencias (e.g., $e1$ y $e2$), relativas a dos elementos focales (e.g., $A1$ y $A2$) de un mismo marco de discernimiento, las funciones de asignación básica de verosimilitud -i.e., $m1(A1)$ y $m2(A2)$- se combinan para dar una nueva función de asignación básica de verosimilitud -$m12$- que representa el efecto conjunto de ambas evidencias sobre la intersección de los elementos focales correspondientes. De esta forma:

Si $e1 \colon A1 \rightarrow m1(A1) \;\; = x$

Si $e2 \colon A2 \rightarrow m2(A2) \;\; = y$

definimos $m12(C) = m1\,(A1) \times m2\,(A2)$, donde $C = A1 \cap A2$.

Esta expresión puede generalizarse directamente para distintas parejas de elementos focales. Así:

$$m12(C) = \sum_{C = Ai \cap Bj} m1(Ai)\; m2(Bj)$$

Formulación que coincide con la de asignación de probabilidad a la intersección de dos sucesos independientes en la teoría clásica de la probabilidad. Por este motivo se puede afirmar que la teoría evidencial asume implícitamente independencia entre las evidencias. Por otra parte, es claro que la primera condición exigida a la función de asignación básica de verosimilitud se cumple:

$$\sum_{C = Ai \cap Bj} m12(C) = 1$$

Sin embargo, puede ocurrir que distintas evidencias “señalen” a elementos focales muy distintos, tan distintos que no tengan ningún elemento en común, y por lo tanto su intersección sea nula. Este caso introduce una peculiaridad en el modelo:

Si: $e1$: $A1 \rightarrow m1(A1) = x$, $0 < x \leq 1$
Si: $e2$: $A2 \rightarrow m2(A2) = y$, $0 < y \leq 1$
Si: $C = A1 \cap A2 = ø$
Entonces: $m12(C) = m12(ø) = m1(A1) \times m2(A2) = xy \neq 0$

resultado que contradice la segunda condición exigida a la función básica de asignación de verosimilitud, según la cual la solución *tiene* que estar en el marco de discernimiento. ¿Estamos ante una contradicción de la teoría evidencial? Aparentemente sí. Sin embargo estas situaciones ocurren muchas veces en edificios tan bien estructurados como la Física o la Química y, además, tienen una solución muy sencilla... Si aparece una inconsistencia que no debía aparecer ¡hay que eliminarla! Para ello, *por decreto* $m12(ø)=0$, y si hay que corregir algo se corrige.

Este procedimiento se denomina *Normalización* y no es, en absoluto, ni artificioso ni extraño. Concretamente, en la teoría evidencial la normalización de resultados debe conseguir mantener las funciones de asignación básica de verosimilitud entre los límites definidos, lo que supondrá corregir las asignaciones a elementos focales de intersección no nula, de forma que su suma siga siendo la unidad. La nueva expresión para la combinación de evidencias es la siguiente:

$$m12(C) = \frac{\sum_{C = Ai \cap Bj} m1(Ai) \;\; m2(Bj)}{1 - \sum_{Ai \cap Bj = \phi} m1(Ai) \;\; m2(Bj)} = \frac{\sum_{C = Ai \cap Bj} m1(Ai) \;\; m2(Bj)}{\sum_{Ai \cap Bj \neq \phi} m1(Ai) \;\; m2(Bj)}$$

Claramente, la necesidad de normalización aparece cuando las fuentes de información aportan evidencias sobre elementos focales muy diferentes, tan diferentes que no comparten hipótesis individuales.

La expresión $K = \sum_{Ai \cap Bj = \phi} m1(Ai)\, m2(Bj)$ del denominador de la expresión normalizada para $m12(C)$ se denomina *grado de conflicto,* y precisamente es una medida de la compatibilidad existente entre las evidencias que están siendo combinadas. Considerando la expresión para el grado de conflicto, la ecuación de $m12(C)$ se puede reescribir de la siguiente forma:

$$m12(C) = \frac{\sum_{C = Ai \cap Bj} m1(Ai)\, m2(Bj)}{1 - K}$$

expresión en la que el factor $\dfrac{1}{1-K}$ se denomina *factor de normalización*.

Nótese que cuando las distintas evidencias señalan a distintos elementos focales entre los cuales no hay intersecciones nulas, no podemos hablar de evidencias contradictorias. En este caso no hay conflictos, $K$=0, y la expresión normalizada para $m12(C)$ coincide con la expresión sin normalizar. Por el contrario, cuando las evidencias son totalmente contradictorias y todos los elementos focales son disjuntos entre sí, el conflicto es total, $K$=1, y la combinación de evidencias no está definida.

Analizaremos todas estas ideas por medio de un ejemplo en el cual vamos a tratar de obtener la mejor interpretación para una manifestación dada, en un dominio cualquiera, en el que todas las posibles soluciones son $\{h1, h2, h3, h4\}$, y para lo cual disponemos de la siguiente información:

- La evidencia $e1$ apoya a las interpretaciones $h1$ y $h2$ con: $m1(h1, h2) = 0.6$
- La evidencia $e2$ apoya a las interpretaciones $h2$, $h3$, $h4$ con: $m2(h2, h3, h4) = 0.7$
- La evidencia $e3$ apoya a la interpretación $h1$ con: $m3(h1) = 0.8$

El primer paso consiste en formalizar el problema en los términos que la teoría propone:

$\theta = \{h1, h2, h3, h4\}$

$e1$: $A1 = \{h1, h2\}$, $m1(A1) = 0.6$
$m1(\theta) = 0.4$
$m1(B1) = 0.0$, con $B1 \subset \theta$, $B1 \neq A1$, $B1 \neq \theta$

$e2$: $A2 = \{h2, h3, h4\}$, $m2(A2) = 0.7$
$m2(\theta) = 0.3$
$m2(B2) = 0.0$, con $B2 \subset \theta$, $B2 \neq A2$, $B2 \neq \theta$

$e3$: $A3 = \{h1\}$, $m3(A3) = 0.8$
$m3(\theta) = 0.2$
$m3(B3) = 0.0$, con $B3 \subset \theta$, $B3 \neq A3$, $B3 \neq \theta$

Antes de proceder a combinar evidencias (y dado que tal combinación implica la intersección de conjuntos con medidas de asignación básica de verosimilitud relativas a distintas evidencias), al objeto de evitar confusiones utilizaremos la siguiente notación:

$e1$: $A1 = \{h1, h2\}$, $m1(A1) = 0.6$
$m1(\theta_1) = 0.4$

$e2$: $A2 = \{h2, h3, h4\}$, $m2(A2) = 0.7$
$m2(\theta_2) = 0.3$

$e3$: $A3 = \{h1\}$, $m3(A3) = 0.8$
$m1(\theta_3) = 0.2$

Nótese que hemos prescindido de aquellos subconjuntos para los cuales "$mx = 0$", independientemente de la evidencia considerada. Nótese, asimismo, que hemos señalado los respectivos marcos de discernimiento con un subíndice indicativo de la evidencia de que procede la medida de verosimilitud asignada, y que corresponde a la confianza no asignada al elemento focal considerado. Evidentemente:

$$\theta_1 = \theta_2 = \theta_3 = \theta = \{h1, h2, h3, h4\}$$

Procederemos ahora a combinar evidencias:

$e1$ y $e2$:

| | |
|---|---|
| $A1 \cap A2 = \{h2\}$, | $m12\{h2\} = 0.6 \times 0.7 = 0.42$ |
| $A1 \cap \theta_2 = \{h1, h2\}$, | $m12\{h1, h2\} = 0.6 \times 0.3 = 0.18$ |
| $\theta_1 \cap A2 = \{h2, h3, h4\}$, | $m12\{h2, h3, h4\} = 0.4 \times 0.7 = 0.28$ |
| $\theta_1 \cap \theta_2 = \theta$, | $m12\{\theta\} = 0.4 \times 0.3 = 0.12$ |

Obsérvese que las intersecciones deben realizarse sobre todos los subconjuntos para los cuales, en función de cada evidencia, su función básica de asignación de verosimilitud es no nula.

Nótese también que si:

C1 = $A1 \cap A2$, con $m12\{h2\} = 0.42$
C2 = $A1 \cap \theta_2$, con $m12\{h1, h2\} = 0.18$
C3 = $\theta_1 \cap A2$, con $m12\{h2, h3, h4\} = 0.28$
C4 = $\theta_1 \cap \theta_2$, con $m12\{\theta\} = 0.12$

entonces: $\sum_{i=1}^{4} m12(Ci) = 1$

lo cual no hace más que confirmar la primera condición establecida para la función básica de asignación de verosimilitud.

Consideremos ahora la tercera evidencia, $e3$, según la cual:

| | |
|---|---|
| $e3$: $A3 = \{h1\}$, | $m3\{h1\} = 0.8$ |
| | $m3\{\theta\} = 0.2$ |

La combinación de esta nueva evidencia con el resultado de la combinación de las evidencias anteriores produce los siguientes resultados:

$e1$ y $e2$ y $e3$:

| | | |
|---|---|---|
| C1 $\cap A3 = ø$ | con $m123\{ø\}$ | = 0.336 |
| C1 $\cap \theta_3 = \{h2\}$ | con $m123\{h2\}$ | = 0.084 |
| C2 $\cap A3 = \{h1\}$ | con $m123\{h1\}$ | = 0.144 |
| C2 $\cap \theta_3 = \{h1, h2\}$ | con $m123\{h1, h2\}$ | = 0.036 |
| C3 $\cap A3 = ø$ | con $m123\{ø\}$ | = 0.224 |

$C3 \cap \theta_3 = \{h2,h3,h4\}$ con $m123\{h2,h3,h4\} = 0.056$
$C4 \cap A3 = \{h1\}$ con $m123\{h1\} = 0.096$
$C4 \cap \theta_3 = \{\theta\}$ con $m123\{\theta\} = 0.024$

La primera condición exigida para la función básica de asignación de verosimilitud sigue cumpliéndose (i.e., la suma de los $m123$ de las correspondientes intersecciones sigue siendo 1); sin embargo, la segunda condición se vulnera en dos ocasiones, por lo que hay que normalizar las expresiones. En este caso, y recordando que el grado de conflicto se calcula según la expresión:

$$K = \sum_{Ai \cap Bj = \varphi} m1(Ai)\ \ m2(Bj)\text{, resulta } K = 0.336 + 0.224 = 0.560$$

Llevando este resultado a la ecuación que nos permite calcular la medida de asignación básica de verosimilitud normalizada (o lo que es lo mismo, dividiendo los resultados anteriores por 1 - $K$), obtenemos los siguientes resultados:

$m123\{h2\} = 0.084/0.44 = 0.191$
$m123\{h1\} = 0.144/0.44 = 0.327$
$m123\{h1, h2\} = 0.036/0.44 = 0.082$
$m123\{h2,h3,h4\} = 0.056/0.44 = 0.127$
$m123\{h1\} = 0.096/0.44 = 0.218$
$m123\{\theta\} = 0.024/0.44 = 0.055$

Agrupando ahora los términos correspondientes obtenemos:

$m123\{h1\} = 0.545$
$m123\{h2\} = 0.191$
$m123\{h1, h2\} = 0.082$
$m123\{h2,h3,h4\} = 0.127$
$m123\{\theta\} = 0.055$

Este resultado final indica que, aunque ninguna evidencia por separado permite afirmar nada sobre ninguna hipótesis individual del marco, la concurrencia de evidencias permite que dos hipótesis concretas -$h1$ y $h2$- se destaquen del resto. Por otra parte, la combinación de evidencias hace que la confianza todavía no asignada disminuya drásticamente en relación a la obtenida cuando consideramos evidencias individuales.

## 1.3. Credibilidad, Plausibilidad e Intervalo de Confianza

La teoría evidencial permite el seguimiento de la evolución dinámica de la confianza depositada en los subconjuntos del marco de discernimiento a medida que aparecen nuevas evidencias. Para ello se definen dos nuevas medidas, la *credibilidad* y la *plausibilidad* que, respectivamente, son un indicador de la mínima y de la máxima confianza que podemos depositar en un elemento focal dado.

Formalmente la credibilidad se define según la ecuación:

$$Cr(A) = \sum_{B \subset A} m(B)$$

en donde $A$ es el elemento focal considerado, subconjunto del marco de discernimiento. La credibilidad es una medida de las contribuciones que todos los subconjuntos de $A$ ejercen sobre el propio $A$. Consideremos el ejemplo del apartado anterior, y calculemos la evolución de la credibilidad en el elemento focal $\{h1, h2\}$, a medida que van apareciendo las evidencias $e1$, $e2$ y $e3$.

Con $e1$, la credibilidad de $\{h1, h2\}$ coincide exactamente con la medida de asignación básica de verosimilitud, ya que -aparte del propio elemento focal considerado- no existe ningún subconjunto de $\{h1, h2\}$ con valores de $m$ distintos de 0. Así:

$$e1\text{: } \{h1, h2\} \rightarrow m1\{h1, h2\} = Cr(h1, h2) = 0.6$$

Consideremos ahora la segunda evidencia combinada con la primera, y analicemos el mismo elemento focal. En este caso obteníamos las siguientes intersecciones:

| | | | |
|---|---|---|---|
| $(h2)$ | $\rightarrow$ | $m12\ (h2)$ | $= 0.42$ |
| $(h1, h2)$ | $\rightarrow$ | $m12(h1, h2)$ | $= 0.18$ |
| $(h2, h3, h4)$ | $\rightarrow$ | $m12(h2, h3, h4)$ | $= 0.28$ |
| $(h1, h2, h3, h4)$ | $\rightarrow$ | $m12(h1, h2, h3, h4)$ | $= 0.12$ |

por lo que, de acuerdo con la expresión propuesta:

$$Cr(h1, h2) = m12(h2) + m12(h1, h2) = 0.42 + 0.18 = 0.60$$

Este resultado coincide con el anterior -una única evidencia-, por lo que la credibilidad asociada al elemento focal considerado no se ha visto modificada tras la aparición de $e2$. Consideremos, finalmente, el efecto conjunto de las tres evidencias, que nos producía los siguientes resultados:

| | | | |
|---|---|---|---|
| $(h1)$ | $\rightarrow$ | $m123(h1)$ | $= 0.545$ |
| $(h2)$ | $\rightarrow$ | $m123(h2)$ | $= 0.191$ |
| $(h1, h2)$ | $\rightarrow$ | $m123(h1, h2)$ | $= 0.082$ |
| $(h2, h3, h4)$ | $\rightarrow$ | $m123(h2, h3, h4)$ | $= 0.127$ |
| $(h1, h2, h3, h4)$ | $\rightarrow$ | $m123(h1, h2, h3, h4)$ | $= 0.055$ |

Si evaluamos ahora la credibilidad de $(h1, h2)$ obtenemos:

$$Cr(h1, h2) = m123(h1) + m123(h2) + m123(h1, h2) = 0.545 + 0.191 + 0.082 = 0.818$$

Nuestra credibilidad en el elemento focal considerado (es decir, la mínima confianza que podemos depositar en él) ha aumentado tras considerar las tres evidencias conjuntamente.

La otra medida importante, la plausibilidad, es un indicador de la máxima confianza que podemos depositar en un elemento focal dado, y se calcula considerando también las contribuciones de otros subconjuntos con intersección no nula, de acuerdo con la expresión:

$$Pl(A) \;=\; \sum_{B \cap A \neq \phi} m(B)$$

en donde *A* es el elemento focal considerado, subconjunto del marco de discernimiento.

La plausibilidad no sólo tiene en cuenta los subconjuntos del propio elemento focal, sino también todas las contribuciones de todos aquellos subconjuntos que tienen algo que ver con dicho elemento focal. Volviendo a nuestro ejemplo:

$e1$: $Pl(h1, h2) = m1(h1, h2) + m1(h1, h2, h3, h4) = 0.6 + 0.4 = 1.0$

$e1$ y $e2$: $Pl(h1, h2) = m12(h2) + m12(h1,h2) + m12(h2, h3, h4) + m12(h1, h2, h3, h4) = 1.0$

$e1$ y $e2$ y $e3$: $Pl(h1, h2) = m123(h1) + m123(h2) + m123(h1, h2) + m123(h2, h3, h4) + m123(h1, h2, h3, h4) = 1.0$

En este caso la plausibilidad del elemento focal considerado no varía, y su valor es siempre la unidad. Ello indica que, siendo optimistas, la confianza que podemos depositar en todas o alguna de las hipótesis de dicho elemento focal es completa (lo cual, por cierto, se refleja en los valores del resultado final del ejemplo). En otros casos, sin embargo, ambas medidas -*Cr* y *Pl*- varían durante el proceso inferencial[46].

Una tercera medida importante es el llamado *Intervalo de Confianza*, que se construye, para cada elemento focal, a partir de la credibilidad y de la plausibilidad. Así, en cada nivel del proceso de razonamiento, el intervalo de confianza es el segmento del espacio numérico [0,1] que tiene como valor mínimo el valor de la credibilidad del elemento focal, y como valor máximo el correspondiente valor de la plausibilidad. De este modo, y recurriendo de nuevo a nuestro ejemplo, los intervalos de confianza del elemento focal ($h1$, $h2$) varían del siguiente modo:

$e1$: IC($h1$, $h2$) = [0.600, 1.000]
$e1$ y $e2$: IC($h1$, $h2$) = [0.600, 1.000]
$e1$ y $e2$ y $e3$: IC($h1$, $h2$) = [0.818, 1.000]

Observamos aquí que, al considerar conjuntamente las tres evidencias, el intervalo de confianza se hace más estrecho, a la vez que se aproxima a la unidad. Conceptualmente, el intervalo de confianza representa la incertidumbre asociada al elemento focal considerado. Además, se puede demostrar que si *A* es un elemento focal:

$$0 \leq Cr(A) \leq 1$$

---

[46] Analícese, por ejemplo, el caso del elemento focal ($h1$)

$0 \leq Pl(A) \leq 1$
$Cr(A) \leq Pl(A)$
$Cr(A) \leq \text{Prob}(A) \leq Pl(A)$

en donde Prob($A$) es la probabilidad estadística.

Claramente,

Si $Cr(A) = 0$ y $Pl(A) = 1$, entonces la ignorancia sobre $A$ es total.
Si $Cr(A) = Pl(A) = 1$, entonces $A$ es absolutamente cierto.
Si $Cr(A) = Pl(A) = 0$, entonces $A$ es absolutamente falso.

## 1.4. Casos Particulares de la Teoría Evidencial

La teoría evidencial contiene, bajo ciertos supuestos y en ciertas situaciones, al modelo de los factores de certidumbre de Shortliffe y Buchanan.

Analicemos el caso de dos evidencias independientes que apoyen al mismo elemento focal $A$ de un determinado marco de discernimiento. En este caso, el tratamiento que le daríamos a nuestro problema sería el siguiente:

$e1$: $A1$, $m1(A1) = \text{s}1$, $m1(\theta_1) = 1 - s1$
$e2$: $A2$, $m2(A2) = \text{s}2$, $m2(\theta_2) = 1 - s2$

con $A1 = A2$ y $\theta_1 = \theta_2$

Combinando ambas evidencias obtenemos el siguiente resultado:

$e1$ y $e2$:

$A1 \cap A2 = A$, $m12(A) = s1 \times s2$
$A1 \cap \theta_2 = A$, $m12(A) = s1 \times (1 - s2) = s1 - s1 \times s2$
$\theta_1 \cap A2 = A$, $m12(A) = (1 - s1) \times s2 = s2 - s1 \times s2$
$\theta_1 \cap \theta_2 = \theta$, $m12(\theta) = (1 - s1) \times (1 - s2)$

Agrupando ahora las expresiones para $A$ obtenemos:

$$m12(A) = s1\, s2 + s1 - s1\, s2 + s2 - s1\, s2 = s1 + s2 - s1\, s2$$

pero este resultado es exactamente el que obtendríamos si aplicáramos el modelo de Shortliffe y Buchanan en el caso de dos evidencias independientes que apoyan a la misma hipótesis, cada una de las cuales con su correspondiente factor de certidumbre. Así:

Si: $e1$ Entonces $A$ con $\text{CF}(A,e1) = s1$, $s1 > 0$

Si: $e2$ Entonces $A$ con $\text{CF}(A,e2) = s2$, $s2 > 0$

el resultado que obtendríamos para CF($A$, $e1$ y $e2$) = $s1 + s2 - s1\ s2$

En otros casos, sin embargo, la teoría evidencial proporciona resultados diferentes de los que obtendríamos aplicando el modelo de los factores de certidumbre. Más concretamente, la teoría evidencial funciona igual que el modelo de Shortliffe y Buchanan en aquellos casos en los que este modelo funciona bien, y lo mejora notablemente cuando el modelo de los factores de certidumbre presenta carencias. Al respecto, los autores de este texto dejan como ejercicio para sus sufridos lectores el análisis del siguiente problema: ¿Qué resultados se obtendrían aplicando la teoría evidencial y el modelo de Shortliffe y Buchanan cuando dos evidencias $e1$ y $e2$ apoyan respectivamente a $A$ y a $\neg A$?

## 1.5. Resumen

La teoría evidencial de Dempster y Shafer presenta un esqueleto formal para tratar convenientemente tanto el conocimiento inexacto como la falta de conocimiento. La teoría trabaja sobre conjuntos de hipótesis, sin que la confianza aportada por una determinada evidencia tenga que distribuirse entre las hipótesis individuales del conjunto considerado. Cuando existen evidencias que apoyan a grupos de hipótesis contradictorias, el modelo normaliza los resultados correspondientes. Para ello se define un índice -grado de conflicto-, que cuantifica las incompatibilidades encontradas. El problema de la incertidumbre se trata a partir del llamado intervalo de confianza, que se construye a partir de dos medidas, la credibilidad y la plausibilidad, indicativas respectivamente de la mínima y la máxima confianza que, en un momento dado, podemos depositar en un conjunto determinado de hipótesis. El intervalo de confianza evoluciona dinámicamente a medida que van apareciendo nuevas evidencias. Un resultado particularmente interesante de la teoría evidencial es que contiene al modelo de Shortliffe y Buchanan en aquellos casos en los que dicho modelo funciona bien, y sin embargo lo mejora en aquellos casos en los que dicho modelo presenta deficiencias.

## 1.6. Textos básicos


- Shafer, “A Mathematical Theory of Evidence”, Princeton University Press, eds.,1976.