Meeting 4 (lecture 4): Other probabilistic accounts for beliefs

A quick look at (an)other theory for understanding beliefs – part II

Arthur Dempster: A generalization of Bayesian Inference. *Journal of the Royal Statistical Society. Series B.* 1968, Vol. 30 (2) : 205-247.
Glenn Shafer: *Mathematical Theory of Evidence*. Princeton University Press, 1976,

constitute the grounds for *Dempster-Shafer theory of belief functions*

"(…) belief functions is a mathematical theory of how to combine degrees of rational belief derived from different evidential sources."

[Nance D. (2019). Belief functions and burdens of proof. *Law, Probability and Risk* 18: 53-76].

"The Dempster-Shafer theory, also known as the theory of belief functions, is a generalization of the Bayesian theory of subjective probability. Whereas the Bayesian theory requires probabilities for each question of interest, belief functions allow us to base degrees of belief for one question on probabilities for a related questions"

[Shafer G.: *Dempster-Shafer Theory*. www.glennshafer.com/assets/downloads/articles/article-48.pdf] :

Consider an event A

The axioms of probability as a measure state that $P(A) + P(\neg A) = 1$

Now, consider what is referred to as *epistemic uncertainty*.

There is evidence (knowledge) that supports belief in A ($\text{supp}(A)$) to a certain amount, where support – like probability – is measured on a scale from 0 to 1.

The evidence also supports belief in $\neg A$ to a certain amount ($\text{supp}(\neg A)$).

However,

$$\text{supp}(A) + \text{supp}(\neg A) \text{ is not by necessity equal to 1}$$

One may say that a portion of the total support provided by the evidence is withheld or *uncommitted* as between A and $\neg A$.

Example

Let A stand for the statement that a marketing campaign has increased the sales of a certain product in Sweden.

From marketing research it is found that the sales of the product in Stockholm in August 2021 has increased compared to August 2020, while the sales in Malmö in August 2021 has slightly decreased compared to August 2020.

The evidence (marketing research results) may then lead to that the support of A is 0.6 while the support of $\neg A$ is 0.2. Such supports could follow from considering that Stockholm has about 3 times higher population than Malmö, but these two communities cannot be said to represent fully the population of buyers in Sweden.

Hence, there is uncommitted support of 0.2 as between A and $\neg A$. This amount of support is therefore – at this stage – on the disjunction A or $\neg A$ ($\text{supp}(A \cup \neg A) = 0.2$).

Support is now transformed to belief so that the belief in one single event equals the support of that event, while the belief in a disjunction is the sum of the supports of the individual events of the disjunction plus the uncommitted support (of that disjunction.

$$\mathrm{Bel}(A) = \mathrm{supp}(A) = 0{,}6$$

$$\mathrm{Bel}(\neg A) = \mathrm{supp}(\neg A) = 0{,}2$$

$$\begin{aligned}\mathrm{Bel}(A \text{ or } \neg A) &= \mathrm{supp}(A) + \mathrm{supp}(\neg A) + \mathrm{supp}(A \cup \neg A) \\ &= 0{,}6 + 0{,}2 + 0{,}2 = 1\end{aligned}$$

These three values are referred to as the *belief function* .

Note that $\mathrm{Bel}(A \text{ or } \neg A) = 1 = P(A \cup \neg A)$, but $\mathrm{Bel}(A) + \mathrm{Bel}(\neg A) < 1$

The construction can be pictured as

When the beliefs are such that $\text{Bel}(A) + \text{Bel}(\neg A) = 1$ always, the beliefs are referred to as *Bayesian beliefs*. Hence, one may say that belief functions are generalizations of subjective probabilities

Plausibility

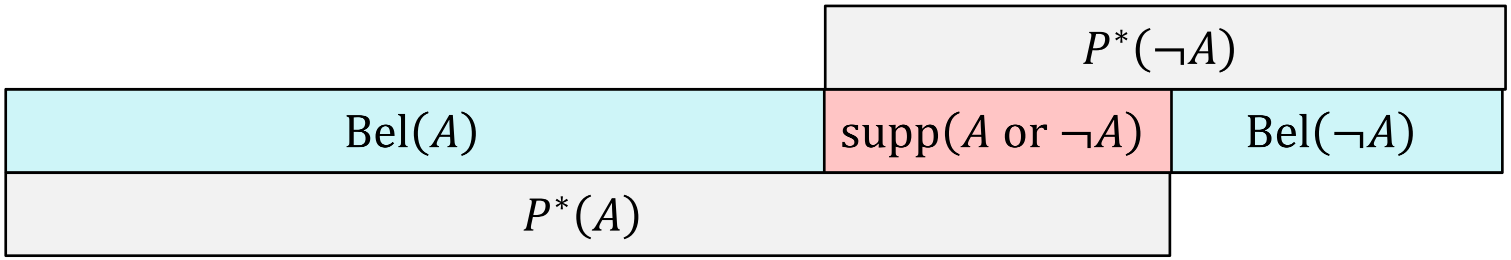
The *plausibility* (or *upper probability*) of an event A is the maximum potential belief in A:

$$P^*(A) = 1 - \text{Bel}(\neg A)$$

and the plausibility of $\neg A$ is analogously

$$P^*(\neg A) = 1 - \text{Bel}(A)$$

Graphically



Potential decision rules:

- Choose A if $$BR = \frac{\text{Bel}(A)}{\text{Bel}(\neg A)} > C$$
- Choose A if $$PR = \frac{P^*(A)}{P^*(\neg A)} > C$$
- Choose A if $$BR_2 = \frac{\text{Bel}(A) + \theta}{\text{Bel}(\neg A) + \theta} > C \qquad \theta = \frac{1}{2} \cdot \text{supp}(A \text{ or } \neg A)$$

Recall the example from Meeting 3

A =“The seized crowbar belongs to Mr Johnson”
B = “The seized crowbar is blue”
C = “Witness says: ‘Mr Johnson’s crowbar is unpainted’”
C' = “Mr Johnson’s crowbar is unpainted”

What is the evidence?

We know that the seized crowbar is blue and we know what witness said that Mr Johnson’s crowbar is unpainted.

$$\Longrightarrow \ \text{Bel}(B) = 1, \text{Bel}(C) = 1$$

$$BR_2 = \frac{\text{Bel}(A) + \theta}{\text{Bel}(\neg A) + \theta} > C$$

Assume we would apply decision rule 3 with C = 1.2
(20% exceedance of the equal stands)

$$\theta = \frac{1}{2} \cdot \text{supp}(A \text{ or } \neg A)$$

$\text{supp}(A|B, C)$?

A ="The seized crowbar belongs to Mr Johnson"
B = "The seized crowbar is blue"
C = "Witness says: 'Mr Johnson's crowbar is unpainted'"
C' = "Mr Johnson's crowbar is unpainted"

We cannot forget that there are initial reasons to believe that *A* is true.
Assume $\text{supp}(A) = 0.5$.

Note that this does not imply that is $\text{supp}(\neg A)$ should also be 0.5. Assume the uncommitted support is 0.2 so that $\text{supp}(\neg A) = 0.3$. Hence, $\text{supp}(A \text{ or } \neg A) = 0.2$

Since $\text{Bel}(A) = \text{supp}(A)$ and $\text{Bel}(\neg A) = \text{supp}(\neg A)$, choosing decision rule 3 we get

$$BR_2 = \frac{\text{Bel}(A) + \theta}{\text{Bel}(\neg A) + \theta} = \frac{0.5 + 0.1}{0.3 + 0.1} = 1.5 > \text{C} = 1.2$$

$\Rightarrow$ Choose *A* !

A ="The seized crowbar belongs to Mr Johnson"
B = "The seized crowbar is blue"
C = "Witness says: 'Mr Johnson's crowbar is unpainted'"
C' = "Mr Johnson's crowbar is unpainted"

Then, given B and C we might commit more support to $\neg A$ without affecting the support of A. Let's say that we add 0.2 to the support of $\neg A$, which means that $\text{supp}(A|B,C)$ is still 0.5, while $\text{supp}(\neg A|B,C) = 0.5 \neq 0.3$

This updates the values plugged in to decision rule 3:

$$BR_2 = \frac{\text{Bel}(A|B,C) + \theta}{\text{Bel}(\neg A|B,C) + \theta} = \frac{0.5 + 0}{0.5 + 0} = 1 < C$$

$\Rightarrow$ Choose $\neg A$!

Note that we could go further analysing what would happen if we take C' into consideration, but since this is not an observable event, it cannot be used for updating.

Application to decisions of courts

Criminal law

Let H_p = "The defendant is guilty"
H_d = "The defendant is not guilty"

Depending on the country's judicial system, there may be different standards of proof.

In the Western World criminal law it is common to have
"beyond reasonable doubt"
as standard of proof

For many and historically, this means that the probability of H_p must be very high, but no common threshold is defined. Some would say 0.95, 0.98, 0.99,…

Would it work with

$$BR = \frac{\mathrm{Bel}(A)}{\mathrm{Bel}(\neg A)} > C \ ? \qquad PR = \frac{P^*(A)}{P^*(\neg A)} > C \ ? \qquad BR_2 = \frac{\mathrm{Bel}(A) + \theta}{\mathrm{Bel}(\neg A) + \theta} > C \ ?$$

Civil law

Let H_p = "The plaintiff is right"
H_d = "The respondent is right"

In the Western World civil law it is common to have as standard of proof "preponderance of evidence" or "balance of probabilities"

For many and historically this means that the probability of H_p must be proven higher than the probability of H_d

Would it work with

$$BR = \frac{\mathrm{Bel}(A)}{\mathrm{Bel}(\neg A)} > C \ ? \qquad PR = \frac{P^*(A)}{P^*(\neg A)} > C \ ? \qquad BR_2 = \frac{\mathrm{Bel}(A) + \theta}{\mathrm{Bel}(\neg A) + \theta} > C \ ?$$