ST790 — Fall 2022
Imprecise-Probabilistic Foundations of Statistics

Ryan Martin
North Carolina State University
www4.stat.ncsu.edu/~rmartin

Week 04b

This lecture

- Belief functions — origins & perspectives
- Definition and basic properties
- Connections to random sets & possibility measures
- Examples
- ...

Introduction

- *Belief function theory* might be the largest "IP community"[1]
- *BELIEF 2022* conference in October[2]
- Probably the first formal IP theory (1960s)
- Synonymous with *Dempster–Shafer theory*[3]
- Commonly used in CS, ENGR, AI, ...
- There are some fundamental differences between "belief functions" and "imprecise probabilities"
- Dempster & Shafer had different perspectives
- Our focus will be on Shafer's...

[1] BFAS, `https://bfasociety.org`
[2] `http://hebergement.universite-paris-saclay.fr/belief2022/`
[3] `https://en.wikipedia.org/wiki/Dempster-Shafer_theory`

Origins

- I previously mentioned Art Dempster's formulation
- Motivated by statistical inference, "first to understand and then to replace"[4] Fisher's fiducial argument
- Closely related to random sets, or *multivalued maps*
- Dempster's view is closer to imprecise probability:
  - imprecision is unavoidable in some problems
  - leads to *bounds* on (subjective) probabilities
  - need a calculus[5] for inference & reasoning
- So, roughly, Dempster views a "belief function" through a credal-set lens, i.e., bounds on (subjective) probabilities

[4] In Dempster's foreword to Shafer's 1976 book

[5] e.g., "Dempster–Shafer calculus for statisticians" (*IJAR* 2008)

Origins, cont.

- Glenn Shafer's approach was quite different, though the math is more-or-less the same as Dempster's
- S developed D's ideas into a general & powerful framework
- In the spirit of Kolmogorov, Shafer developed an axiom-based system, *A Mathematical Theory of Evidence*
- Shafer's view is different from Dempster's/IP's:
  - "evidence $\to$ degrees-of-belief" is imprecise
  - degrees of belief need not be related to probabilities
  - Dempster's calculus is a central piece
- So, roughly, Shafer views a "belief function" as its own thing irrespective of any probability bounds

Belief functions

Definition.

A functional $\underline{\Pi} : 2^{\mathbb{X}} \to [0,1]$ is a *belief function* if

1. $\underline{\Pi}(\varnothing) = 0$
2. $\underline{\Pi}(\mathbb{X}) = 1$
3. $\underline{\Pi}$ is ∞-monotone

The dual, $\overline{\Pi}$, is called a *plausibility function*

- For good reason, Shafer's book focuses exclusively on the *finite-$\mathbb{X}$ case*, as does much of the literature
- I'll do the same (for the most part)
- Common in the literature to see $\underline{\Pi}(\cdot)$ and $\overline{\Pi}(\cdot)$ denoted as $\mathrm{bel}(\cdot)$ and $\mathrm{pl}(\cdot)$, respectively[6]

[6] Because they're not interpreted as "lower" and "upper" probabilities

Belief functions, cont.

- Shafer's approach is similar to Kolmogorov's
  - purposely offers no justification for his axioms, no claims that one would be "irrational" to refute them
  - focus is on developing the subsequent math
- His axioms are similar to Kolmogorov's too — just swap countable-additivity for ∞-monotonicity
- Contains probability & possibility as special cases
- Choquet: equivalent to a random set (when $\mathbb{X}$ is finite)
- Imprecise-probabilistic consequences (if you wish):
  - no-sure-loss
  - coherence
- So far, nothing new to us...

Belief functions, cont.

- Set/recall some notation and terminology[7]
  - $\mathbb{X}$ is called the *frame of discernment* or *frame* (finite)
  - a *basic probability assignment* is just a probability mass function $m : 2^{\mathbb{X}} \to [0, 1]$
  - subsets A with $m(A) > 0$ are called *focal elements*[8]
- For finite frames, Choquet's theorem implies a one-to-one correspondence between $\underline{\Pi}$'s and m's, i.e.,

$$\underline{\Pi}(A) = \sum_{B \in 2^{\mathbb{X}} : B \subseteq A} m(B), \quad A \in 2^{\mathbb{X}}$$

- Similarly, $\overline{\Pi}(A) = \sum_{B \in 2^{\mathbb{X}} : B \cap A \neq \varnothing} m(B)$

[7] I'll mostly follow notation and terminology in Cuzzolin's near-encyclopedic book, *The Geometry of Uncertainty*, 2021

[8] If m is understood as the mass function of a random set $\mathcal{X}$ on finite $\mathbb{X}$, then the collection of focal elements is just the support of $\mathcal{X}$

Example[9]

- For fixed $S \in 2^{\mathbb{X}}$ and $s \in [0,1]$, a *simple support function* is

$$\underline{\Pi}(A) = \begin{cases} 0 & \text{if } A = \varnothing \\ s & \text{if } A \supseteq S \text{ but } A \neq \mathbb{X} \\ 1 & \text{if } A = \mathbb{X} \end{cases}$$

- Corresponds to a random set $\mathcal{X}$ with mass function m...
- Interpretation:
  - Shafer: "corresponds to a body of evidence whose precise and full effect is to support the subset S to the degree s"
  - Me: "I'm $100s$% sure S is true"
- In statistics:
  - simple, easily elicitable prior, provides valuable info
  - what do we do with it, how do we combine with...?

[9] See Example 2 on page 36 of Cuzzolin...

DS calculus: preview

- Consider two simple support functions $\underline{\Pi}_1$ and $\underline{\Pi}_2$ on $\mathbb{X}$
  - determined by pairs (S_1, s_1) and (S_2, s_2)
  - corresponding random sets $\mathcal{X}_1$ and $\mathcal{X}_2$
- Suppose, in my judgment, the bodies of evidence leading to $\underline{\Pi}_1$ and to $\underline{\Pi}_2$ are "independent"
- *How to combine independent bodies of evidence?*
- The DS calculus says, roughly:
  - treat $\mathcal{X}_1$ and $\mathcal{X}_2$ as independent in the usual prob sense
  - interpret $\mathcal{X}_1 \cap X_2 \subseteq A$ as "support for A"
  - remove "conflict cases" where $\mathcal{X}_1 \cap \mathcal{X}_2 = \varnothing$

DS calculus: preview, cont.

- Assume $S_1 \cap S_2 \neq \varnothing$, "no conflict"
- Combined belief function is given by

$$(\underline{\Pi}_1 \oplus \underline{\Pi}_2)(A) = \mathrm{P}(\underbrace{\mathcal{X}_1 \cap \mathcal{X}_2 \subseteq A}_{\text{support for } A} \mid \underbrace{\mathcal{X}_1 \cap \mathcal{X}_2 \neq \varnothing}_{\text{no conflict}})$$

$$= \cdots$$

$$= \begin{cases} 0 & A \not\supseteq S_1 \cap S_2 \\ s_1 s_2 & A \supseteq S_1 \cap S_2,\ A \not\supseteq S_1, S_2 \\ s_1 & A \supseteq S_1,\ A \not\supseteq S_2 \\ s_2 & A \supseteq S_2,\ A \not\supseteq S_1 \\ 1 - (1 - s_1)(1 - s_2) & A \supseteq S_1 \cup S_2,\ A \neq \mathbb{X} \\ 1 & A = \mathbb{X} \end{cases}$$

- If $S_1 \cap S_2 = \varnothing$, then there's "conflict" — common!

Next lecture

- More belief functions
- Dempster's rule of combination
- Examples
- Credal set-related properties
- ...