ST790 — Fall 2022
Imprecise-Probabilistic Foundations of Statistics

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

Week 05a

This lecture

- Belief functions recap
- Connection to possibility measures
- Credal set contents
- Dempster's rule of combination
- Examples
- Generalization of Bayes rule

Recap

- $\underline{\Pi} : 2^{\mathbb{X}} \to [0, 1]$ is a *belief function* if
  - $\underline{\Pi}(\varnothing) = 0$
  - $\underline{\Pi}(\mathbb{X}) = 1$
  - it's ∞-monotone
- Dual $\overline{\Pi}$ is a *plausibility function*
- For finite frame $\mathbb{X}$, there exists a *basic probability assignment* m, i.e., a probability mass function on $2^{\mathbb{X}}$, with

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

- Set A with $m(A) > 0$ are called *focal elements*

Connection to possibility theory

- A necessity measure is a belief function
- How can we tell if a belief function a necessity measure?
- *Fact:* $\mathcal{X}$ has a maxitive capacity iff $\mathcal{X}$ is *nested*, i.e.,

$$\mathcal{X}(\omega) \subseteq \mathcal{X}(\omega') \text{ or } \mathcal{X}(\omega) \supseteq \mathcal{X}(\omega') \quad \text{for all } (\omega, \omega')$$

- Roughly, possibility measures $\Longleftrightarrow$ nested random sets
- So, belief = necessity iff focal elements are nested
  - (finite $\mathbb{X}$) belief function determines a random set $\mathcal{X}$
  - focal elements are the realizations of $\mathcal{X}$
- In the belief function literature, "nested focal elements" is often referred to as *consonance*[1][2]

---

[1] Consonance = "no conflict" = "evidence points in a single direction"
[2] See Figure 2.7 in Cuzzolin for an illustration

Remarks on consonance

- Consonance ("no conflict") is a strong assumption
- Can't be justified in many contexts
- Shafer criticizes Shackle on the grounds that it's too optimistic to "ban the appearance of conflict"
- Shafer describes one general class of problems where consonance makes sense:
  - *inferential evidence*
  - "the evidence for a cause that is provided by an effect"
- That's what we're dealing with in statistics!
- BTW, there's interesting work on *consonant approximations* general belief functions[3]

[3] e.g., Dubois & Prade (*IJAR* 1990)

Credal set contents

- Shafer doesn't interpret belief as a lower prob
- But there's still a credal set, so what's in it?
- $\mathscr{C}$(Bel) $\rightarrow$
- Edges parallel to sides of the triangle
- *Shapley's theorem:* $\leq |\mathbb{X}|!$ vertices, given by (3.12)
- Section 4.3 of Nguyen

3.1 The multiple semantics of belief functions 67

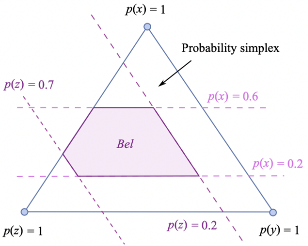


Fig. 3.3: A belief function is a credal set with boundaries determined by lower and upper bounds (3.10) on probability values.

Extremal probabilities of credal sets associated with belief functions Although the set $\mathcal{P}[Bel]$ (3.10) is a polytope in the simplex $\mathcal{P}$ of all probabilities we can define on Θ, not all credal sets there 'are' belief functions. Credal sets associated with BFs have vertices of a very specific type (see Fig. 3.3). The latter are all the distributions P^π induced by a permutation $\pi = \{x_{\pi(1)}, \ldots, x_{\pi(|\Theta|)}\}$ of the singletons of $\Theta = \{x_1, \ldots, x_n\}$ of the form [245, 339]

$$P^\pi[Bel](x_{\pi(i)}) = \sum_{A \ni x_{\pi(i)};\, A \not\ni x_{\pi(j)}\ \forall j<i} m(A). \qquad (3.12)$$

Such an extremal probability (3.12) assigns to a singleton element put in position $\pi(i)$ by the permutation π the mass of all the focal elements containing it, but not containing any elements preceding it in the permutation order [1879].

Cuzzolin (2021), p. 67

Credal set contents, cont.

- Any $\mathrm{P} \in \mathscr{C}(\mathrm{Bel})$ is an *allocation of probability*
- Intuitively: just allocate the mass $m(A)$ to the points x in A
- Formally: $\alpha : \mathbb{X} \times 2^{\mathbb{X}} \to [0,1]$ is an *allocation* (of m) if

$$\sum_{x \in A} \alpha(x, A) = m(A), \quad A \subseteq \mathbb{X}$$

- e.g., uniform allocation (assoc w/ Shapley value) is

$$\alpha(x, A) \equiv \frac{m(A)}{|A|}, \quad x \in A, \quad A \subseteq \mathbb{X}$$

Theorem.

$\mathscr{C}(\mathrm{Bel}_m) = \{\mathrm{P}_\alpha : \alpha \text{ an allocation of } m\}$, where P_α has mass fun

$$p_\alpha(x) = \sum_{A : A \ni x} \alpha(x, A), \quad x \in \mathbb{X}$$

Example

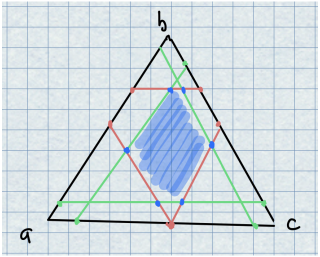
Example from Week 02b slides:

A	$\varnothing$	$\{a\}$	$\{b\}$	$\{c\}$	$\{a,b\}$	$\{a,c\}$	$\{b,c\}$	$\{a,b,c\}$
$m(A)$	0.0	0.1	0.1	0.2	0.3	0.0	0.2	0.1

Find P corresponding to the uniform allocation α:

$$p_\alpha(x) = \sum_{A:A\ni x} \frac{m(A)}{|A|}, \quad x \in \{a,b,c\}$$

$$p_\alpha(a) = 0.1 + \frac{0.3}{2} + \frac{0.1}{3} = 0.283$$
$$p_\alpha(b) = 0.1 + \frac{0.3}{2} + \frac{0.2}{2} + \frac{0.1}{3} = 0.383$$
$$p_\alpha(c) = 0.2 + \frac{0.2}{2} + \frac{0.1}{3} = 0.333$$



Remarks on allocations

- Roughly, allocations $\approx$ mixtures
- For belief funs induced by $\mathcal{X}$ on general $\mathbb{X}$[4]
  - $\mathcal{X}(\omega)$ set-valued map on $(\Omega, \mathscr{B}(\Omega), \mu)$
  - belief function $\underline{\Pi}(A) = \mu\{\omega : \mathcal{X}(\omega) \subseteq A\}$
  - $\mathsf{P} \in \mathscr{C}(\underline{\Pi})$ iff there exists probability measures Q_ω supported on $\mathcal{X}(\omega)$ such that

$$\mathsf{P}(A) = \int_\Omega \mathsf{Q}_\omega(A)\, \mu(d\omega)$$

- Same is basically true for general belief functions,[5] but too complicated for us here...

[4] Wasserman (*Ann. Stat* 1990) "Prior envelopes..."

[5] Shafer (*Ann. Prob* 1979) "Allocations of probability"

Dempster's rule of combination

- Two "independent" belief functions $\underline{\Pi}_1$ and $\underline{\Pi}_2$ on $\mathbb{X}$, determined by mass functions m_1 and m_2, resp.
- How to combine?

Dempster's rule of combination.

Orthogonal sum of $\underline{\Pi}_1$ and $\underline{\Pi}_2$, denoted by $\underline{\Pi}_1 \oplus \underline{\Pi}_2$ is given by

$$(\underline{\Pi}_1 \oplus \underline{\Pi}_2)(A) = \frac{1}{1-\kappa} \sum_{A_1, A_2 : \varnothing \neq A_1 \cap A_2 \subseteq A} m_1(A_1)\, m_2(A_2), \quad A \subseteq \mathbb{X},$$

where κ denotes the "degree of conflict"[a]

$$\kappa = \sum_{A_1, A_2 : A_1 \cap A_2 = \varnothing} m_1(A_1)\, m_2(A_2)$$

[a] $\underline{\Pi}_1$ and $\underline{\Pi}_2$ aren't *combinable* if $\kappa = 1$

Dempster's rule, cont.

- Frame $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$
- Two mass functions

$$m_1(A) = \begin{cases} 0.7 & A = \{\theta_1\} \\ 0.3 & A = \{\theta_1, \theta_2\} \end{cases}$$

$$m_2(A) = \begin{cases} 0.6 & A = \{\theta_2, \theta_3, \theta_4\} \\ 0.4 & A = \Theta \end{cases}$$

- $\kappa = 0.7 \times 0.6 = 0.42$
- Combined mass function

$$m(A) = \begin{cases} 0.483 & A = \{\theta_1\} \\ 0.207 & A = \{\theta_1, \theta_2\} \\ 0.310 & A = \{\theta_2\} \end{cases}$$

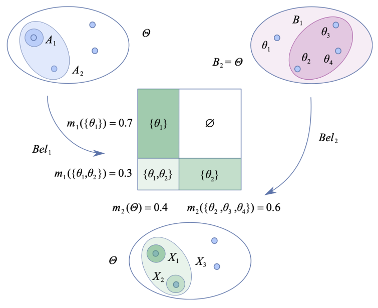


Fig. 2.3: Example of Dempster combination, and its graphical representation.

Cuzzolin (2021), p. 39

Dempster's rule, cont.

- Dempster's rule is
  - commutative, i.e., $\underline{\Pi}_1 \oplus \underline{\Pi}_2 = \underline{\Pi}_2 \oplus \underline{\Pi}_1$
  - associative, i.e., $(\underline{\Pi}_1 \oplus \underline{\Pi}_2) \oplus \underline{\Pi}_3 = \underline{\Pi}_1 \oplus (\underline{\Pi}_2 \oplus \underline{\Pi}_3)$
- This makes belief function theory quite appealing:
  - any relevant piece of evidence gets encoded as a belief function
  - if they're judged to be independent, then just combine them using Dempster's rule as above
- Similar to probability:
  - if I have independent marginal distributions
  - then combine them by multiplying

Dempster's rule, cont.

- Dempster's rule also generalizes Bayes's rule
- Compare "updating" and "combining"
- Start with two belief functions
  - one is general, $\underline{\Pi}$
  - other is extreme, $\underline{\Pi}_B(A) = 1(A \supseteq B)$
- Second piece of evidence: "B is true"
- Combine via Dempster's rule:

$$\underline{\Pi}(A \mid B) := (\underline{\Pi} \oplus \underline{\Pi}_B)(A) = \frac{\underline{\Pi}(A \cup B^c) - \underline{\Pi}(B^c)}{1 - \underline{\Pi}(B^c)}$$

- In terms of plausibility:

$$\overline{\Pi}(A \mid B) = \frac{\overline{\Pi}(A \cap B)}{\overline{\Pi}(B)}$$

Remarks about DS

- Powerful framework, commonly used in AI
- I'll show some more interesting/practical examples later
- Some potential project ideas:[6]
  - variations on the DS framework
  - efficient computations
  - ...
- While individual belief functions are coherent, some issues can arise when they're combined via Dempster's rule...

[6]You can find some details in Cuzzolin's book

Next lecture

- Choquet integration

- Probability via expected values (previsions)
- Lower probability via lower previsions
- Coherence
- ...