ST790 — Fall 2022

Imprecise-Probabilistic Foundations of Statistics

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

Week 08b

This lecture

- Quick recap of the statistical setup
- Towards valid (& efficient) imprecise-probabilistic inference
- Dempster's framework
- Examples/applications
- Discussion

Recap

- Problem setup:
  - uncertain variables: $Y \in \mathbb{Y}$ and $\Theta \in \mathbb{T}$
  - data Y is observable, Θ is to be inferred
  - joint imprecise probability $\overline{\mathsf{P}}_{Y,\Theta}$
- Common scenario:
  - Θ is an unknown parameter in model for Y
  - precise $\mathsf{P}_{Y|\Theta}$ and imprecise prior $\overline{\mathsf{P}}_{\Theta}$
- Inferential model (IM): $y \mapsto (\underline{\Pi}_y, \overline{\Pi}_y)$
- Use IM output for *probabilistic reasoning*
- IM-based probabilistic reasoning ought to be reliable

Recap, cont.

- Towards this, the IM is *valid* if

$$\overline{\mathsf{P}}_{Y,\Theta}\{\overline{\Pi}_Y(A) \le \alpha,\, \Theta \in A\} \le \alpha, \quad \text{all } (A, \alpha)$$

- Equivalent condition in terms of lower probability

$$\overline{\mathsf{P}}_{Y,\Theta}\{\underline{\Pi}_Y(A) > 1 - \alpha,\, \Theta \notin A\} \le \alpha, \quad \text{all } (A, \alpha)$$

- Validity implies control on the probability of erroneous inferences, hence probabilistic reasoning is "reliable"
- Special case of a vacuous prior for Θ:

$$\sup_{\theta \in A} \mathsf{P}_{Y|\theta}\{\overline{\Pi}_Y(A) \le \alpha\} \le \alpha$$
$$\sup_{\theta \notin A} \mathsf{P}_{Y|\theta}\{\underline{\Pi}_Y(A) \ge 1 - \alpha\} \le \alpha \quad \text{all } (A, \alpha)$$

Recap, cont.

- *False confidence theorem:* A precise IM[1] (e.g., default-prior Bayes, fiducial, etc) isn't valid in the sense above
- That is, for any α there exists A such that

$$\sup_{\theta \notin A} \mathsf{P}_{Y|\theta}\{\Pi_Y(A) \geq 1 - \alpha\} > \alpha$$

- So, for precise IMs, even in the "no-prior" case, there *always exists* assertions afflicted by false confidence
- One way to avoid this risk is to require that the IM be valid
  - IM needs to be imprecise
  - but not all imprecise IMs are valid

[1] IM with lower and upper outputs equal: $\underline{\Pi}_y = \overline{\Pi}_y = \Pi_y$, say

Towards a valid IM

- There are a number of approaches that we might consider
  - Dempster's formulation
  - Generalized Bayes
  - ...
- Dempster's approach is appealing
  - extends/generalizes Fisher's fiducial argument
  - Bayes's solution is a special case
- So, I want to start with Dempster's construction
- However:
  - Dempster's IM isn't valid in the sense above
  - not an error, Dempster wasn't trying for validity

Dempster's IM

- Build a joint space $\mathbb{Y} \times \mathbb{T}$, the frame/*state-space model*
- Dempster proposes quantifying uncertainty about (Y, Θ) using random sets[2] in the frame $\mathbb{Y} \times \mathbb{T}$
- For example, prior info about Θ gets *extended* to $\mathbb{Y} \times \mathbb{T}$ through the construction of *cylinders*

$$\text{random set for } \Theta \to \underbrace{\mathbb{Y} \times (\text{random set for } \Theta)}_{\text{random set for } (Y, \Theta)}$$

- We'll end up with several random sets in $\mathbb{Y} \times \mathbb{T}$ and be interested in their intersection, i.e., Dempster's rule

[2] Recall: random sets determine special belief functions

Dempster's IM, cont.

- Dempster's original formulation (1960s) considers random sets that are set-valued mappings of random variables
- Same idea here:
  - $\mathcal{T}^{\text{model}} = \mathcal{T}^{\text{model}}(U)$ encodes the model
  - $\mathcal{T}^{\text{prior}} = \mathcal{T}^{\text{prior}}(V)$ encodes the prior
  - $\mathcal{T}^{\text{obs}} = \mathcal{T}^{\text{obs}}(W)$ encodes the observations y
- Then interest is in the intersection

$$\mathcal{T}_y = \underbrace{\mathcal{T}^{\text{model}} \cap \mathcal{T}^{\text{prior}} \cap \mathcal{T}^{\text{obs}}}_{\text{determined by } U, V, W}$$

- Dempster focuses on the vacuous prior case, $\mathcal{T}^{\text{prior}} \equiv \mathbb{Y} \times \mathbb{T}$, so this term can be ignored in the above expression

Dempster's IM, cont.

- What's $\mathcal{T}^{\text{model}}(U)$?
- Think about how you might generate data from model $\mathsf{P}_{Y|\Theta}$
  - draw a random seed $U \sim \mathsf{P}_U$
  - plug U into a (model, Θ)-dependent function
  - output Y
- Write this algorithm[3] as $Y = a(\Theta, U)$
- Then define

$$\mathcal{T}^{\text{model}}(u) = \{(Y, \Theta) \in \mathbb{Y} \times \mathbb{T} : Y = a(\Theta, u)\},$$

  all those data–parameter pairs compatible with given u
- $\mathcal{T}^{\text{model}}(U)$, with $U \sim \mathsf{P}_U$, is a random set

[3] a.k.a. "association" or "data-generating equation"

Dempster's IM, cont.

- Value y of Y is observed
- Encode as $\mathcal{T}^{\text{obs}} \equiv \{(Y, \Theta) \in \mathbb{Y} \times \mathbb{T} : Y = y\}$, constant
- Then $\mathcal{T}_y = \mathcal{T}^{\text{model}} \cap \mathcal{T}^{\text{obs}}$ is determined by U,

$$\mathcal{T}_y(U) = \{(y, \Theta) \in \mathbb{Y} \times \mathbb{T} : y = a(\Theta, U)\}$$

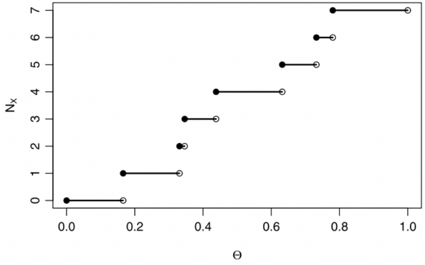
- Project $\mathcal{T}_y(U)$ in $\mathbb{Y} \times \mathbb{T}$ down to a random set in $\mathbb{T}$
- Dempster's rule gives the y-dependent IM with

$$\begin{aligned}
\underline{\Pi}_y(A) &= \mathsf{P}_U\{\mathcal{T}_y(U) \subseteq A \mid \mathcal{T}_y(U) \neq \varnothing\} \\
\overline{\Pi}_y(A) &= \mathsf{P}_U\{\mathcal{T}_y(U) \cap A \neq \varnothing \mid \mathcal{T}_y(U) \neq \varnothing\} \\
&= 1 - \underline{\Pi}_y(A^c)
\end{aligned}$$

Example

- Suppose $Y \mid \Theta \sim \text{Bin}(n, \Theta)$
  - individual Bernoulli's: $Y_i = 1(U_i \leq \Theta)$, $U_i \overset{\text{iid}}{\sim} \text{Unif}(0, 1)$
  - key relationship: $U_{(Y)} \leq \Theta < U_{(Y+1)}$
- Post-conflict-removal, the basic random set is

$$\mathcal{T}^{\text{model}}(U) = \{(Y, \Theta) : U_{(Y)} \leq \Theta < U_{(Y+1)}\}$$

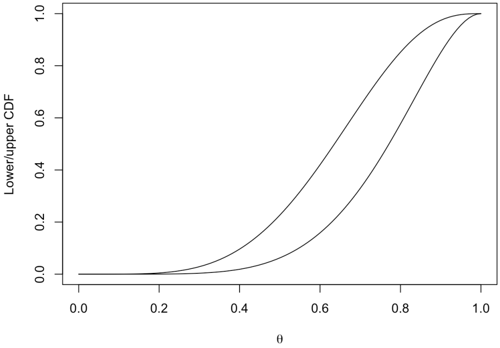


Plot of $\mathcal{T}^{\text{model}}(U)$, where "$N_X$" is our Y; from M. et al (*Stat Sci* 2010)

Example, cont.

- Summaries of Dempster's IM
- Lower and upper CDFs

$$\theta \mapsto \begin{cases} \underline{\Pi}_y([0,\theta]) = \mathsf{P}_U\{U_{(y+1)} \le \theta\} \\ \overline{\Pi}_y([0,\theta]) = \mathsf{P}_U\{U_{(y)} \le \theta\} \end{cases}$$



Properties

- Conditioning[4] out conflict $\implies$ sufficient statistics
- Extension of Fisher's fiducial argument:
  - for those cases when Fisher's argument can be applied, the solution agrees with Dempster's
  - Dempster's solution can be applied when Fisher's can't
- Generalizes Bayes's theorem:
  - if prior is precise, then $\underline{\Pi}_y = \overline{\Pi}_y$ and equals Bayes's
  - in this case, $\mathcal{T}_y(U)$ is a singleton
  - conditioning on $\mathcal{T}_y(U) \neq \varnothing \iff$ conditioning on $a(\Theta, U) = y$

[4] Conditioning is generally more powerful than sufficiency

Dempster's IM isn't valid

Theorem (M. 2022)

If an IM suffers from sure loss in the sense that, e.g.,

$$\sup_y \overline{\Pi}_y(A) < \underline{\mathsf{P}}_\Theta(A) \quad \text{for some } A,$$

then it's not valid.

- The result in the paper is stronger than above[5]
- Neat connection between statistical & behavioral properties
- Since Dempster's rule can incur sure loss,[6] the above theorem implies Dempster's IM isn't valid
- Also follows from the false confidence theorem since, in some examples, Dempster's IM is precise

[5] It can also be explained better than it is in the current version

[6] Gong & Meng's *three prisoner example*

Discussion

- Dempster's framework is intuitive and powerful
- Spurred much of the development in imprecise probability
- Generalizes both Bayesian and fiducial inference
- Lots of applications:
  - many outside of statistics (see Cuzzolin's book)
  - signal-plus-background (Edlefsen et al, *AoAS* 2009)
  - ...
- Computation can be a challenge:
  - conditional distributions of random sets
  - very recent progress has been made on this[7]

[7] Jacob et al (*JASA* 2021, with discussion)

Discussion, cont.

- Dempster's IM not being valid isn't an error/flaw, his formulation was based more on logic
- Failure to avoid sure loss is what pushed Walley and his followers away from Dempster's theory
- Doesn't meet my validity criteria either
- So we need to keep looking for a valid IM construction...

Next lecture

- Generalized Bayes
- Properties
- Walley's formulation for vacuous priors
- IM efficiency considerations