ST790 — Fall 2022
Imprecise-Probabilistic Foundations of Statistics

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

Week 07b

This lecture

- Quick recap of imprecise prob stuff we covered
- Brief list of things we didn't cover
- Transition into "applications"
- Start with *statistical inference*
- .....

Recap

- Initial motivation:
  - precise prob can't accommodate all kinds of uncertainty, e.g., epistemic uncertainty about models
  - too strong an assumption in real applications
  - perhaps some need for imprecision in statistical inference...?
- De Finetti's notion of *coherence* and its generalization
- Capacities, K-monotone, K-alternating
- Credal sets (of probability distributions)
  - no sure loss iff credal set is non-empty
  - coherent iff capacity equals envelope of its credal set

Recap, cont.

- Different models:
  - random sets
  - possibility measures
  - belief functions
  - lower previsions
- Properties, characterization of credal sets
- Specifically for possibility measures:
  - extension principle
  - (imprecise-)probability-to-possibility transform
- Deeper dive into coherence
- Updating rules:
  - Dempster's rule
  - Generalized Bayes rule

Some stuff we didn't cover

- History of these developments
- Other kinds of imprecise prob models:[1]
  - interval probabilities
  - fuzzy sets
  - p-boxes
  - non-numerical...[2]
- Lots more details on the models we discussed
  - random sets
  - belief functions
  - lower previsions / sets of desirable gambles
- Dependence in imprecise probability
  - judgments about independence, exchangeability, etc
  - combine marginal specifications into joint specifications accounting for dependence

[1] some of these in, e.g., Ch. 4 of *Intro to IP*

[2] e.g., `https://researchers.one/articles/22.09.00004`

Transitioning

- Two general types of logic:
  - *deductive* (e.g., population to sample)
  - *inductive* (e.g., sample to population)
- Theory of probability (precise or imprecise) is deductive
  - set some axioms
  - logic & math to derive consequences
- Applications of prob (precise or imprecise) are inductive
  - less structured, maybe subjective
  - driven by basic principles and priorities
- Easy to think that "deductive $\gg$ inductive," but...

> *In deductive reasoning all knowledge attainable is already latent in the postulates... In inductive reasoning we are performing part of the process by which new knowledge is created.* —Fisher

Statistical inference

- Isn't statistical inference settled? *NO!*
- Bayesian vs. frequentist, *two-theory problem*[3] is serious
- Not just an impractical detail for philosophers to debate
- I think imprecise probability holds the key to resolving this

> *Two contending philosophical parties, the Bayesians and the frequentists, have been vying for supremacy over the past two-and-a-half centuries... Unlike most philosophical arguments, this one has important practical consequences. The two philosophies represent competing visions of how science progresses and how mathematical thinking assists in that progress.* —Efron

[3] There are more than two theories, just two "dominant" ones

Statistical inference, cont.

- Problem setup:
  - uncertain variables: $Y \in \mathbb{Y}$ and $\Theta \in \mathbb{T}$
  - data Y is observable, Θ is to be inferred
  - e.g., Θ is an unknown parameter in model for Y
- Notation:
  - joint imprecise probability $\overline{\mathrm{P}}_{Y,\Theta}$
  - e.g., precise $\mathrm{P}_{Y|\Theta}$ and imprecise prior $\overline{\mathrm{P}}_{\Theta}$
  - assume all imprecise probabilities are coherent
- Covers the "Bayesian" and "frequentist" cases, and more

frequentist		Bayesian
every prior	$\leftarrow$ some priors $\rightarrow$	one prior

Statistical inference, cont.

- "Inference" can mean various things
- My perspective here:
  - goal is *uncertainty quantification*
  - mapping $(y, \overline{\mathsf{P}}, \ldots) \mapsto (\underline{\Pi}_y, \overline{\Pi}_y)$, imprecise prob[4] on $\mathbb{T}$
  - if $Y = y$ is observed, quantify uncertainty about Θ via $\overline{\Pi}_y$
- This mapping is what I call an *inferential model* (IM),[5] i.e., a model for how data gets converted to inference/UQ
- Bayes, fiducial, Dempster, generalized Bayes, ... are IMs
- Relevant questions:
  - what do we want to do with the IM output?
  - how do we interpret the IM output?
  - what properties do we want the IM output to satisfy?

[4]Technically, could be a lower/upper prevision...

[5]Original focus was too narrow, I know better now...

Inferential models

- *What do we want to do the IM output?*
- There are lots of things we could do:
  - estimators, confidence regions
  - hypothesis tests and other decision procedures
- These are important but, to me, secondary
- My priority is that the IM facilitates a form of probabilistic reasoning, i.e., for any "assertion" $A \subseteq \mathbb{T}$ about Θ,

$$\overline{\Pi}_y(A) \text{ is small } \implies \text{ infer } A^c$$
$$\underline{\Pi}_y(A) \text{ is large } \implies \text{ infer } A$$

- Note that this is *for any assertion A*

IMs, cont.

- *How do we interpret the IM output?*
  - bounds on probabilities
  - bounds on buying/selling prices for gambles
  - degrees of belief
  - ...
- *What properties do we want IMs to satisfy?*
  - maybe depends on our desired interpretation
  - e.g., coherence
  - ...
  - validity and efficiency
- To me, these two are difficult to separate
  - I don't know what to believe, so I don't just trust any IM
  - if the IM is reliable, then I'm willing to trust it[6]

[6] "Objectively subjective," cf. Lewis's *Principal Principle*

Validity

- The kind of *reliability* that I'm after is with respect to the "probabilistic reasoning" step above
- An IM $y \mapsto (\underline{\Pi}_y, \overline{\Pi}_y)$ is *valid* (relative to $\overline{\mathrm{P}}$) iff

$$\overline{\mathrm{P}}_{Y,\Theta}\{\underbrace{\overline{\Pi}_Y(A) \leq \alpha,\, \Theta \in A}_{\text{inference might be wrong}}\} \leq \alpha, \quad \text{all } (A, \alpha) \in 2^{\mathbb{T}} \times [0,1]$$

- Duality between $\underline{\Pi}_y$ and $\overline{\Pi}_y$ leads to an equivalent formula...
- When model is precise and prior is vacuous:[7]

$$\sup_{\theta \in A} \mathrm{P}_{Y|\theta}\{\overline{\Pi}_Y(A) \leq \alpha\} \leq \alpha, \quad \text{all } (A, \alpha) \in 2^{\mathbb{T}} \times [0,1]$$

[7] This is the case I focused on in my earlier work

Validity, cont.

Definition.

An IM $y \mapsto (\underline{\Pi}_y, \overline{\Pi}_y)$ is *valid* (relative to $\overline{\mathsf{P}}$) iff

$$\overline{\mathsf{P}}_{Y,\Theta}\{\overline{\Pi}_Y(A) \leq \alpha, \ \Theta \in A\} \leq \alpha, \quad \text{all } (A, \alpha) \in 2^{\mathbb{T}} \times [0, 1]$$

- Idea: control the $\overline{\mathsf{P}}$-probability of erroneous inference, assertion-wise and threshold-wise
- Not (necessarily) a frequentist notion!
- Questions:
  - *Why "for all A"?*
    Restrict to one or a few A's defeats the purpose
  - *Why "for all α"?*
    I can be flexible, but what range of α's is wide enough?
  - *Why the same α inside and out?*
    Could put "$f(\alpha)$" inside, but useless if I don't know f

Validity, cont.

- Validity is strong and weak at the same time
  - strong in the sense that it implies statistical procedures have error rate control guarantees[8]
  - weak[9] in the sense it's achieved by a vacuous IM
- Challenge is to have "just enough" of imprecision[10]
- For example:
  - generalized Bayes is valid[11]
  - but not efficient — recall tendency to dilate
- Other familiar IM constructions are not valid...

- More on construction of valid & efficient IMs later

[8] Theorem 1, `https://researchers.one/articles/21.05.00001`

[9] Stronger notions of validity are *possible* (pun intended)

[10] That's what *efficiency* is intended to help with, more later

[11] Corollary 3, same link as above

False confidence

- *Question:* Is imprecision necessary for validity?
- If we specifically consider an imprecise model like above, then answer seems clearly *Yes*
- What about if we don't explicitly mention imprecision?[12]
  - suppose that $Y \sim \mathsf{P}_{Y|\theta}$, where θ is unknown
  - Bayes/fiducial/etc IM is Π_Y, a probability on $\mathbb{T}$
- There are reasons to doubt that validity holds in this case...
- Can actually prove that it fails

[12] Remember, "no prior" really means "every prior," so we can't escape imprecision just by pretending it's not there!

False confidence, cont.

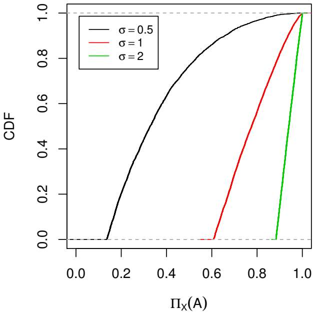
False confidence theorem.

A precise IM $y \mapsto \Pi_y$ is not valid (wrt vacuous prior), i.e., for any $\alpha \in [0,1]$, there exists $A \subset \mathbb{T}$ such that

$$\sup_{\theta \notin A} \mathsf{P}_{Y|\theta}\{\underbrace{\Pi_Y(A) > 1-\alpha}_{\text{confidence in } A \not\ni \theta}\} > \alpha.$$

(Balch, M., & Ferson 2019, `arXiv:1706.08565`)

- Satellite collision example
  - $A = \{\text{non-collision}\}$
  - then $\Pi_Y(A)$ as a random variable, with a CDF $\longrightarrow$
  - truth: *on collision course*
  - different noise levels, σ
- *False confidence:* $\Pi_Y(A)$ is almost always large!



False confidence, cont.

- Take-away: *validity fails without imprecision*
- Remarks:[13]
  - not all A's are afflicted with false confidence
  - don't know which A's are afflicted, hence the risk
  - dangerous, users can do anything with MCMC output
  - problem doesn't go away asymptotically
- Two options:
  - figure out which A's are afflicted and warn users
  - use a valid/imprecise IM
- Not all imprecise IMs are valid...

[13] There are some (loose?) connections between false confidence and incoherence, but this is still work-in-progress

Next lecture

- IM constructions:
  - Dempster's formulation
  - generalized Bayes
- Validity?
- Examples
- .....