Fiducial inference viewed through a possibility-theoretic inferential model lens

Ryan Martin

Department of Statistics, North Carolina State University

www4.stat.ncsu.edu/~rmartin/

NC STATE

1 Introduction & Summary

- Fisher's fiducial argument is well motivated but the specifics are mysterious
- While fiducial inference is commonly referred to as "Fisher's biggest blunder," it's directly tied to many fundamentally important developments in statistics:
  - confidence limits
  - significance testing
  - sufficient statistics & conditional inference
- Fiducial inference also has direct connections to imprecise probability via Dempster
- This paper (arXiv:2303.08630) offers a new connection between fiducial inference and imprecise probability via a possibilistic *inferential model* (IM)
- I consider a special but important class of statistical models
  - includes *location-scale models* as a key special case
  - an agreed-upon fiducial distribution exists and has desirable statistical properties
- Main result in the paper: *for this class of models, the fiducial solution is the "maximal" probability distribution in the credal set determined by the IM*

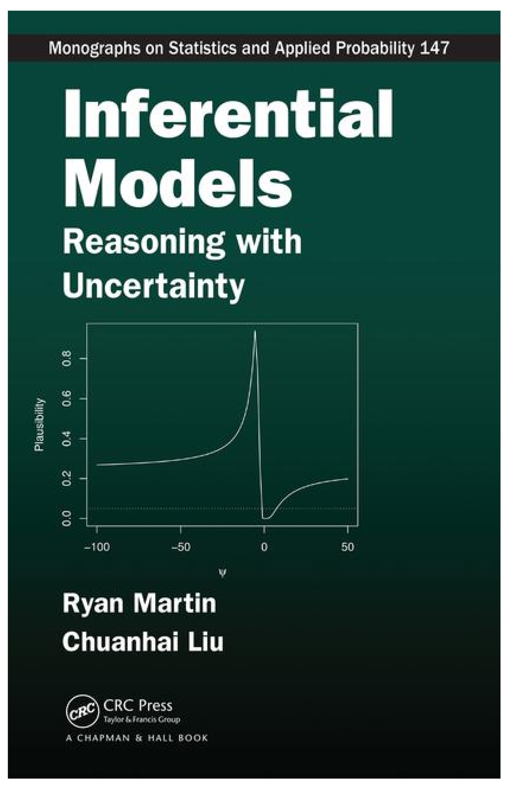
2 Problem setup

- Observable data $X \in \mathbb{X}$, observation is x
- Statistical model $\{\mathsf{P}_\theta : \theta \in \mathbb{T}\}$, uncertain value denoted by Θ
- No prior information about Θ is available, i.e., prior is vacuous
- Goal is reliable *uncertainty quantification* about Θ, given $X = x$
- More specifically, the goal is to assign reliable "degrees of belief" to hypotheses about Θ
- Reliable, prior-free, probabilistic UQ has been called the "Holy Grail of Statistics"

3 Background: inferential models

3.1 High-level

- Goal is to complete the Grail quest: valid, prior-free, (imprecise-)probabilistic uncertainty quantification
- Imprecision isn't part of the model assumptions, but for reliability it's *necessary* that the solution is imprecise
- Book: first rigorous development along these lines
- This IM construction relies on (nested) random sets and some aspects of Dempster–Shafer calculus
- But that's not the only way to do it...
- A number of other subsequent developments:
  - direct constructions (e.g., arXiv:1203.6665)
  - practical applications (e.g., arXiv:1912.00037)
  - computation (arXiv:2103.02659)
  - conformal prediction (arXiv:2112.10234)
  - machine-learning (arXiv:2112.10232)
  - partial prior info (arXiv:2211.14567)



3.2 Possibilistic approach

- Direct and principled IM construction puts possibility theory at the forefront
- Let $L_x(\theta)$ denote the model's likelihood function, and define the *relative likelihood*

$$\eta(x,\theta) = \frac{L_x(\theta)}{\sup_{\vartheta \in \mathbb{T}} L_x(\vartheta)}, \quad \theta \in \mathbb{T}$$

- Next define the IM's possibility contour

$$\pi_x(\theta) = \mathsf{P}_\theta\{\eta(X,\theta) \le \eta(x,\theta)\}, \quad \theta \in \mathbb{T}$$

- This possibility contour determines the IM's necessity/possibility output

$$\overline{\Pi}_x(A) = \sup_{\theta \in A} \pi_x(\theta) \quad \text{and} \quad \underline{\Pi}_x(A) = 1 - \overline{\Pi}_x(A^c), \quad A \subseteq \mathbb{T}$$

- Data-driven uncertainty quantification via $(\underline{\Pi}_x, \overline{\Pi}_x)$
  - has certain reliability/validity properties (see below)
  - can be characterized as a x-slice of a (complexity-reduced) consonant outer approximation of the imprecise joint distribution of (X, Θ); see arXiv:2211.14567

3.3 Key properties

- Focus here on statistical reliability, but there are coherence properties too
- For this case with vacuous prior information, the following *validity* property is trivial

$$\sup_{\Theta \in \mathbb{T}} \mathsf{P}_\Theta\{\pi_X(\Theta) \le \alpha\} \le \alpha, \quad \alpha \in [0,1]$$

- Important and familiar statistical consequences:
  - the testing rule "reject H if $\overline{\Pi}_x(H) \le \alpha$" controls Type I error rate at level α
  - the set $\{\theta : \pi_x(\theta) > \alpha\}$ is a $100(1-\alpha)$% confidence seet
- Even more can be said in terms of reliability (arXiv:2304.05740)

4 Invariant statistical models

- Physical models often describe explicitly how the various parameter affect the observable data
- Common examples are *location* and *location-scale models*
- Generalization involves *groups* $\mathscr{G}$ of transformations acting on $\mathbb{X}$ (and on $\mathbb{T}$)
- Basically, $\{\mathsf{P}_\theta : \theta \in \mathbb{T}\}$ is assumed *invariant* wrt to transformations $g \in \mathscr{G}$

$$\mathsf{P}_\theta(X \in \cdot) = \mathsf{P}_{g\theta}(gX \in \cdot), \quad \theta \in \mathbb{T}$$

- There are a number of (mild) technical conditions in the paper that I'm skipping here
- For simplicity, it's common to link $\mathbb{T} = \mathscr{G}$ and treat parameter as a group element
- Key points:
  - this structure translates into structure in the likelihood...
  - fiducial solution agrees with the default-prior Bayes solution (based on right Haar prior)
  - both give exact frequentist confidence limits

5 Fiducial and IM connections

- IM returns a possibility measure, with a corresponding credal set

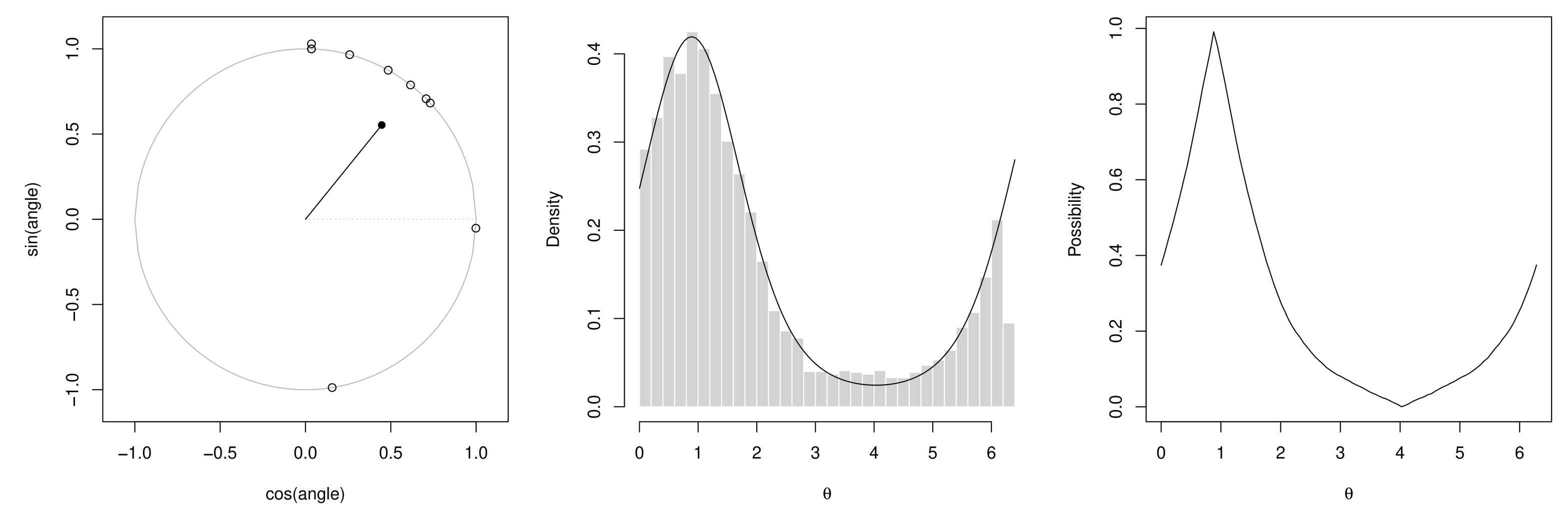
$$\mathscr{C}(\overline{\Pi}_x) = \{\mathsf{Q}_x \in \text{probs}(\mathbb{T}) : \mathsf{Q}_x(\cdot) \le \overline{\Pi}_x(\cdot)\}$$

- Well-known characterization: $\mathsf{Q}_x \in \mathscr{C}(\overline{\Pi}_x) \iff \mathsf{Q}_x\{\pi_x(\Theta) \le \alpha\} \le \alpha$
- A *maximal element* $\mathsf{Q}_x^\star$ satisfies $\mathsf{Q}_x^\star\{\pi_x(\Theta) \le \alpha\} = \alpha$

Theorem. *Under the invariant statistical model setting, the IM's credal set has a maximal element and it corresponds to the fiducial (and default-prior Bayes) solution*

6 Example

Directional data (left) with uncertain mean angle Θ, fiducial density (middle), IM contour (right)



7 A new fiducial argument?

- Theorem says that the fiducial solution is the maximal element of the IM's credal set
- This assumes an invariant model, so that there's an agreed-upon fiducial solution that works
- But the IM solution always works, invariant model or otherwise
- *Can we just define the "fiducial solution" as a/the maximal element of the IM credal set?*
- Note: I'm not advocating for a fiducial solution, since it can't meet my validity requirements
- But this does raise some interesting questions:
  - sound definition of "maximal element"
  - computation of this new fiducial solution
  - comparison to existing solutions (e.g., Hannig's generalized fiducial inference)
  - statistical properties

8 False confidence theorem, revisited

- That no fiducial (or default-prior Bayes) solution can meet the validity requirement is a consequence of the *false confidence theorem* (arXiv:1706.08565)
- Roughly, the false confidence theorem says that, however one defines the probability Q_x, there exists $(\alpha, A) \in [0,1] \times 2^{\mathbb{T}}$ such that $A \not\ni \Theta$ and $\mathsf{P}_\Theta\{\mathsf{Q}_X(A) > 1-\alpha\} > \alpha$
- Currently, no characterization of the problematic A's is available
- *Conjecture:* the problematic A's correspond to hypotheses about non-linear functionals of Θ
- New reason why fiducial generally doesn't work: even in invariant models, the best probabilistic approximation of the marginal IM isn't the marginal fiducial distribution
- With the invariant model structure in this paper, I proved something relevant to this conjecture, namely, that hypotheses concerning linear functionals are safe from false confidence

9 Conclusion

- New imprecise-probability-centric connection between fiducial and IMs
- Offers new understanding of Fisher's fiducial inference
- Maybe even a new approach to constructing fiducial solutions (if one insists)
- New support for IMs
  - IMs "work" in various senses, no assumptions about invariance, etc, and offer much more than the fiducial or Bayes solutions
  - but the fiducial solution in invariant problems is very good, so it'd be an issue if there were no connections, if the IM solution didn't agree