Fiducial inference viewed through a possibility-theoretic inferential model lens

Ryan Martin[1]
North Carolina State University
Researchers.One
www4.stat.ncsu.edu/~rgmarti3

ISIPTA'23
Oviedo, Spain
July 11th 2023

[1]Research partially supported by the NSF, SES–2051225

Introduction

> *Statisticians want numerical measures of the degree to which data support hypotheses.* —Hacking

- Most options consider only precise probabilities, e.g.,
  - Bayesian
  - fiducial
- Fiducial is often (unfairly) called "Fisher's biggest blunder"
- Despite its failures, fiducial still had a major impact:
  - Neyman's confidence sets
  - Dempster's theory $\rightsquigarrow$ imprecise probability

Intro, cont.

- Also: there are cases where fiducial works great
- So, if another method is to work well, then it shouldn't be totally different from fiducial
- In particular, my *inferential model* (IM) framework
  - takes the form of a data-dependent possibility distribution
  - is provably valid in general, i.e., it always works
- Connection between IMs and fiducial?
- Main result: in a class of problems where fiducial works, its "posterior" is a maximal member of the IM's credal set

Problem setup

- Data $X \in \mathbb{X}$, observed value x
- Model $\{\mathsf{P}_\theta : \theta \in \mathbb{T}\}$, e.g., $\mathsf{Bin}(n, \theta)$
- Uncertain true value is Θ, generic values denoted by θ
- Assume *no prior info* about Θ is available[2]
- Data + model $\Longrightarrow$ likelihood: $\theta \mapsto L_x(\theta)$

- **Goal:** *reliably quantify uncertainty about* Θ, *given* $X = x$

[2] This was the case Fisher and many other statisticians consider

Inferential models (IMs)

- IMs deal with imprecise probabilities[3]
- Vacuous prior info about $\Theta \Rightarrow$ IM's possibility contour is

$$\pi_x(\theta) = \mathsf{P}_\theta\{R(X, \theta) \leq R(x, \theta)\}$$

  where $R(x, \theta) = L_x(\theta) / \sup_\vartheta L_x(\vartheta)$, relative likelihood
- Upper probability[4] defined via optimization

$$\overline{\Pi}_x(A) = \sup_{\theta \in A} \pi_x(\theta), \quad A \subseteq \mathbb{T}$$

- Magnitudes of $\overline{\Pi}_x(\cdot)$ used for drawing inferences

---

[3] Modern version is possibility-theoretic (M., `arXiv:2211.14567`)

[4] Lower probability $\underline{\Pi}_x$ defined via conjugacy

IMs, cont.

Validity theorem.

The IM above is (strongly) valid, i.e.,

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

- Basic validity result,[5] familiar for p-values
- Important consequences wrt reliable possibilistic UQ
  - tests have exact Type I error rate control[6]
  - exact confidence sets, i.e., if $C_\alpha(x) = \{\theta : \pi_x(\theta) > \alpha\}$, then

$$\sup_{\Theta} \mathsf{P}_{\Theta}\{C_\alpha(X) \not\ni \Theta\} \leq \alpha, \quad \alpha \in [0,1]$$

[5] Generalizations to "partial prior info" in M., arXiv:2211.14567

[6] "Uniform" error control in Cella and M., arXiv:2304.05740

Invariant statistical models

- Fiducial argument works when the model has more structure
- I mean a *group invariance* structure:
  - group $\mathcal{G}$ of transformations acting on $\mathbb{X}$ (and on $\mathbb{T}$)
  - roughly, model is invariant wrt to $\mathcal{G}$ if

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

  - e.g., location-scale models & affine transformations
- The following results are "well-known"
  - the fiducial and default-prior Bayes solutions are the same
  - Bayes solution uses the right invariant Haar prior
  - both solutions give exact confidence sets

IM and fiducial connection

- IM's credal set: $\mathscr{C}(\overline{\Pi}_x) = \{\mathsf{Q}_x \in \text{probs}(\mathbb{T}) : \mathsf{Q}_x(\cdot) \leq \overline{\Pi}_x(\cdot)\}$
- Well-known characterization

$$\mathsf{Q}_x \in \mathscr{C}(\overline{\Pi}_x) \iff \mathsf{Q}_x\{\pi_x(\Theta) \leq \alpha\} \leq \alpha, \quad \alpha \in [0,1]$$

- "Maximal" member $\mathsf{Q}_x^\star$ satisfies

$$\mathsf{Q}_x^\star\{\pi_x(\Theta) \leq \alpha\} = \alpha, \quad \alpha \in [0,1]$$

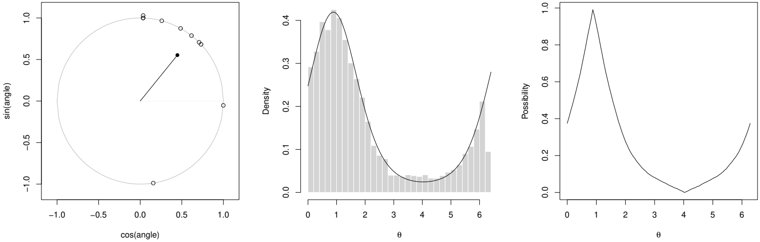
Theorem (M., `arXiv:2303.08630`).

For an invariant statistical model as described in the paper:

- IM's credal set has a maximal member
- it corresponds to the fiducial distribution
- in particular, the IM and fiducial confidence sets are the same

Example

- Directional data, i.e., angles or points on a circle
- Simple von Mises model with unknown mean direction Θ
- Group structure: Θ plays the role of a rotation
- Plots show data,[7] fiducial density, and IM contour[8]



[7] From Example 1 in Mardia's *Directional Statistics*
[8] Details in M., `arXiv:2303.08630`

Conclusion

- New connection between fiducial and IMs
- Explains why fiducial sometimes works
  - IM always works
  - fiducial works when it agrees with the IM
- Also explains why fiducial doesn't always work
  - marginalization via different calculus
  - fiducial–IM connection isn't preserved under marginalization
- More details in the paper, we can talk during poster session

The end — *Thanks!*

https://wordpress-courses2223.wolfware.ncsu.edu/st-790-001-fall-2022/

Valid and efficient imprecise-probabilistic inference with partial priors, II. General framework

Ryan Martin*

https://arxiv.org/abs/2211.14567