Valid uncertainty quantification about a model[1]

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

ISIPTA 2019
Ghent, Belgium
July 4th, 2019

[1] https://www.researchers.one/article/2018-08-21

Setup

- Consider a classical statistical inference problem:
  - observable data Y;
  - statistical model $Y \sim \mathsf{P}_{Y|\theta}$ depending on $\theta \in \Theta$;
  - goal is to quantify uncertainty about θ based on $Y = y$.
- For statistical inference to be *valid* (in a sense), uncertainty must be quantified as a non-additive belief.[2]
- But non-additivity alone isn't enough, some care is needed in the construction.
- How to construct an inferential model that's valid?

[2] M. (2019) "False confidence, non-additive beliefs, and valid statistical inference," based on my *BELIEF 2018* lecture; on *Researchers.One* and *IJAR*.

Valid inferential models

- Express statistical model via an "association"

$$Y = a(\theta, U), \quad U \sim \mathsf{P}_U.$$

- We don't observe U, but can use a suitable *random set*, $\mathcal{S} \sim \mathsf{P}_{\mathcal{S}}$, to predict/guess its value.
- Push the random set through (y, assoc) to Θ,

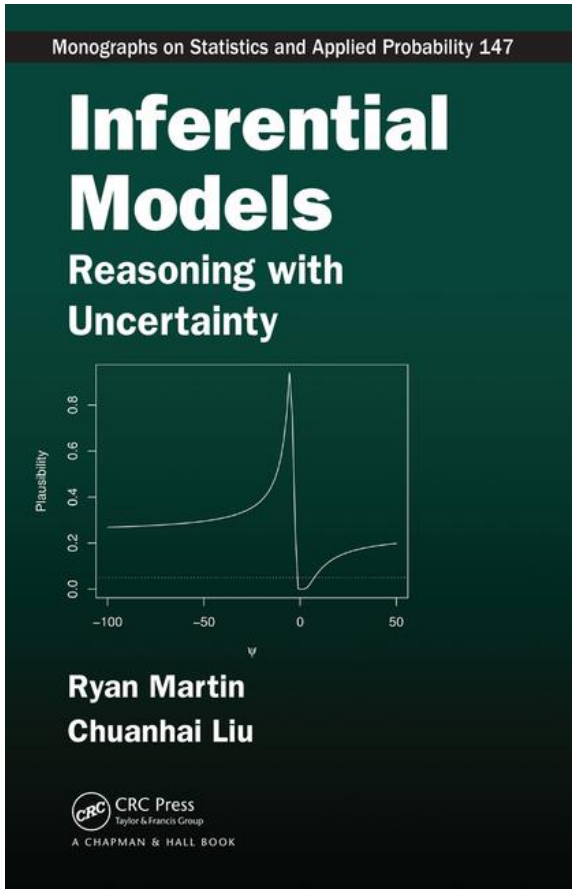
$$\Theta_y(\mathcal{S}) = \bigcup_{u \in \mathcal{S}} \{\vartheta : y = a(\vartheta, u)\}.$$

- Distribution of $\mathcal{S}$ induces y-dependent non-additive beliefs:

$$\begin{aligned} \text{bel}_y(A) &= \mathsf{P}_{\mathcal{S}}\{\Theta_y(\mathcal{S}) \subseteq A\} \\ \text{pl}_y(A) &= 1 - \text{bel}_y(A^c), \qquad A \subseteq \Theta. \end{aligned}$$

Commercial

Still looking for the perfect summer-time read...?



This paper

- Stuff described above takes the statistical model as given.
- *What if the statistical model itself is also uncertain?*
- Express the "parameter" as $\theta = (M, \theta_M)$, where
  - M is a model index
  - θ_M is a model-specific parameter
- Then θ_M is a nuisance parameter.
- Dealing with model uncertainty is like marginal inference...

This paper, cont.

- My approach handles marginal inference by manipulating the association, to isolate the interest parameter.
- General details in paper and poster.
- Here I'll just make an analogy to linear regression:
  - re-express data as (suff stat, residuals)
  - if model is given, *ignore the residuals*
  - if model is uncertain, *use the residuals*
- After marginalization is complete, proceed with the same random set business to get a valid inferential model for M.

One more commercial

- Existing peer review system is detrimental in various ways.
- More on this later...
- Successful reform requires new ideas, and action.
- H. Crane and I developed a new open-access publication platform, featuring an *author-driven* peer review process.

RESEARCHERS.ONE

- For details, check us out at

www.researchers.one

www.twitter.com/@ResearchersOne

The end

Thanks!

rgmarti3@ncsu.edu
www4.stat.ncsu.edu/~rmartin

RESEARCHERS.ONE

International Journal of Approximate Reasoning
Volume 113, October 2019, Pages 39-73

False confidence, non-additive beliefs, and valid statistical inference

Ryan Martin ✉