Multi-resolution Theory for Approximating Infinite-p-Zero-n: Transitional Inference, Individualized Predictions, and a World Without Bias-Variance Tradeoff

Xiao-Li Meng

Discussion

Ruggero Bellio
University of Udine, Italy

22 Sept. 2022

1. General comments

Very interesting talk (and paper), thought-provoking, with three important topics

- Transitional inference ⇒ using data from an approximating population to draw inferences on a new individual
- **Multi-Resolution** theory ⇒ information increasing without bounds and how to mathematically capture this
- What happens in a **deterministic world** ⇒ some surprising results

1. General comments

Very interesting talk (and paper), thought-provoking, with three important topics

- Transitional inference ⇒ using data from an approximating population to draw inferences on a new individual
- **Multi-Resolution** theory ⇒ information increasing without bounds and how to mathematically capture this
- What happens in a **deterministic world** ⇒ some surprising results

1. General comments

Very interesting talk (and paper), thought-provoking, with three important topics

- Transitional inference $\Rightarrow$ using data from an approximating population to draw inferences on a new individual
- **Multi-Resolution** theory $\Rightarrow$ information increasing without bounds and how to mathematically capture this
- What happens in a **deterministic world** $\Rightarrow$ some surprising results

1. General comments

Very interesting talk (and paper), thought-provoking, with three important topics

- Transitional inference ⇒ using data from an approximating population to draw inferences on a new individual
- **Multi-Resolution** theory ⇒ information increasing without bounds and how to mathematically capture this
- What happens in a **deterministic world** ⇒ some surprising results

1. General comments

Very interesting talk (and paper), thought-provoking, with three important topics

- Transitional inference $\Rightarrow$ using data from an approximating population to draw inferences on a new individual
- **Multi-Resolution** theory $\Rightarrow$ information increasing without bounds and how to mathematically capture this
- What happens in a **deterministic world** $\Rightarrow$ some surprising results

1. General comments

Very interesting talk (and paper), thought-provoking, with three important topics

- Transitional inference $\Rightarrow$ using data from an approximating population to draw inferences on a new individual
- **Multi-Resolution** theory $\Rightarrow$ information increasing without bounds and how to mathematically capture this
- What happens in a **deterministic world** $\Rightarrow$ some surprising results

1. General comments

Very interesting talk (and paper), thought-provoking, with three important topics

- Transitional inference $\Rightarrow$ using data from an approximating population to draw inferences on a new individual
- **Multi-Resolution** theory $\Rightarrow$ information increasing without bounds and how to mathematically capture this
- What happens in a **deterministic world** $\Rightarrow$ some surprising results

Some important points

In a **deterministic world** there are some counter-intuitive results

- Multi-resolution effects may matter for multiple-descent phenomena
- For linear models, there is an enlightening result

$$\text{Avg. pred. error} = [\tau^2 + A(r)]\,\frac{(n+1)\,(n-2)}{n\,(n-r-2)}$$

  which may not explode if $\tau^2 = 0$ when $r/n \to 1$ so that optimal R may be close to n
- Implications of over-fitting are not always obvious, but possibility for a black-box procedure to resist it

Some important points

In a **deterministic world** there are some counter-intuitive results

- Multi-resolution effects may matter for multiple-descent phenomena
- For linear models, there is an enlightening result

$$\text{Avg. pred. error} = [\tau^2 + A(r)] \, \frac{(n+1)\,(n-2)}{n\,(n-r-2)}$$

  which may not explode if $\tau^2 = 0$ when $r/n \to 1$ so that optimal R may be close to n

- Implications of over-fitting are not always obvious, but possibility for a black-box procedure to resist it

Some important points

In a **deterministic world** there are some counter-intuitive results

- Multi-resolution effects may matter for multiple-descent phenomena
- For linear models, there is an enlightening result

$$\text{Avg. pred. error} = [\tau^2 + A(r)]\,\frac{(n+1)\,(n-2)}{n\,(n-r-2)}$$

  which may not explode if $\tau^2 = 0$ when $r/n \to 1$ so that optimal R may be close to n

- Implications of over-fitting are not always obvious, but possibility for a black-box procedure to resist it

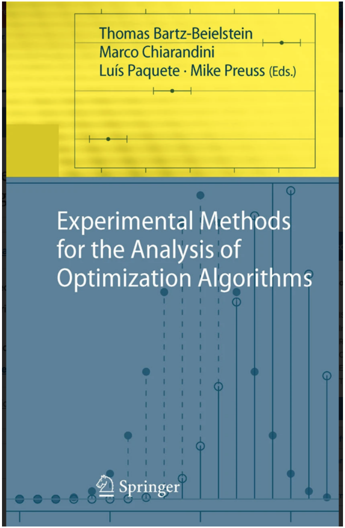
Some important points

Open points from §5.4 of the JASA paper with Xinran Li

- Role of regularization methods
- Importance of systematic investigations of the deterministic MR framework for complex ML models

2. A couple of *personal* notes

Examples of deterministic world



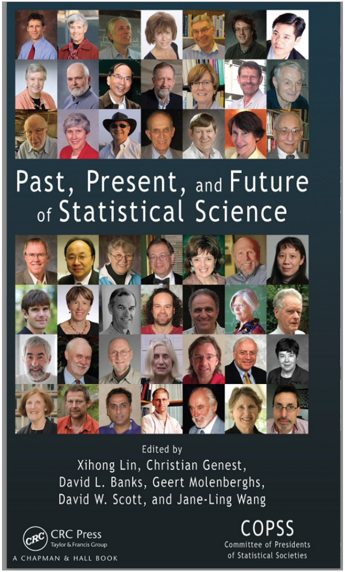
Relevance vs Robustness

This is very central in statistics. Another recent example, already covered in this conference:

Conformal prediction (Vovk et al. 2009, Annals of Statistics; Lei and Wasserman, 2014, JRSS B) strives for valid *distribution-free* sequential prediction, but loses (full) *conditional validity*, since we cannot find $C_n(x)$ such as

$$P\{Y_{n+1} \in C_n(x) | X_{n+1} = x\} \geq 1-\alpha \qquad \text{for all } P \text{ and almost all } x$$

3. The COPSS article



The COPSS article

45

A trio of inference problems that could win you a Nobel Prize in statistics (if you help fund it)

Xiao-Li Meng
Department of Statistics
Harvard University, Cambridge, MA

Statistical inference is a field full of problems whose solutions require the same intellectual force needed to win a Nobel Prize in other scientific fields. Multi-resolution inference is the oldest of the trio. But emerging applications such as individualized medicine have challenged us to the limit: infer estimands with resolution levels that far exceed those of any feasible estimator. Multi-phase inference is another reality because (big) data are almost never collected, processed, and analyzed in a single phase. The newest of the trio is multi-source inference, which aims to extract information in data coming from very different sources, some of which were never intended for inference purposes. All of these challenges call for an expanded paradigm with greater emphases on qualitative consistency and relative optimality than do our current inference paradigms.

Three questions

The part devoted to multi-resolution inference ended with three questions about

- Covariate ordering
- Trade-off between n and R
- Appropriate R for testing

The first two points are already (to some extent) addressed in the JASA paper.

Further treatment in the forthcoming article "*Statistical paradoxes and paradoxes in big data (II): Multi-resolution inference, Simpson's paradox, and individualized treatments*"
cited in the JASA paper?