

Distributed Inference for Tail Empirical and Quantile Processes

Chen Zhou^{1,2}

¹Erasmus University Rotterdam

²De Nederlandsche Bank

Joint work with Liujun Chen and Deyuan Li (Fudan University)

EVA 2021

Distributed inference in extreme value statistics

- ▶ Distributed inference: data are stored in a distributed way
- ▶ Divide-and-Conquer algorithm: “average” in the central machine
- ▶ Oracle estimator: imaginary estimator using all observations
- ▶ **Oracle property:**
Does the DC estimator possess the same asymptotic theory as the oracle estimator?

Distributed inference in extreme value statistics

- ▶ Distributed inference: data are stored in a distributed way
- ▶ Divide-and-Conquer algorithm: “average” in the central machine
- ▶ Oracle estimator: imaginary estimator using all observations
- ▶ **Oracle property:**
Does the DC estimator possess the same asymptotic theory as the oracle estimator?
- ▶ **For extreme value statistics, it is not obvious!**
- ▶ Chen et al (2021) shows the oracle property for the Hill estimator under certain (necessary and) sufficient conditions
Check Liujun Chen’s talk, also in this conference!

This talk: a tool for checking the oracle property

- ▶ Can we have a set of tools for checking the oracle property for (most) extreme value estimators?

This talk: a tool for checking the oracle property

- ▶ Can we have a set of tools for checking the oracle property for (most) extreme value estimators?
- ▶ In extreme value statistics, the key tools for establishing asymptotic theory are the tail empirical and quantile processes
- ▶ The goal: link the tail empirical and quantile processes for the oracle sample and distributed samples (keeping weights)
- ▶ Major difficulty: if the number of machines is large

This talk: a tool for checking the oracle property

- ▶ Can we have a set of tools for checking the oracle property for (most) extreme value estimators?
- ▶ In extreme value statistics, the key tools for establishing asymptotic theory are the tail empirical and quantile processes
- ▶ The goal: link the tail empirical and quantile processes for the oracle sample and distributed samples (keeping weights)
- ▶ Major difficulty: if the number of machines is large
- ▶ Relate the weighted tail empirical and quantile processes for the oracle sample and those for the distributed samples:
 - ▶ *Their random limits are “averaging”*
 - ▶ We provide proper uniform inequalities such that the “averaging” property can be inherited to estimators.

See you in my talk!