Master 2 Internship
Censoring and Extreme Value Theory

Subject:

The problem of censoring, where only partial information on a random variable is available, typically that it exceeds a given threshold, is common in many scientific disciplines. For instance, when studying advanced age mortality, one often has that some individuals of a birth cohort are still alive at the time of follow-up, meaning that only a lower bound for their actual lifetime is available. In insurance, the settlement of major claim cases might take a long time, which implies that at the evaluation time of a portfolio of insurance contracts, some of the claims will not be completely settled. For such cases, the exact total claim amount is not available, but one knows that it will be at least the amount that had been paid up to the time of evaluation. In general, censoring leads thus to the situation where not all cases are fully observed, or, otherwise stated, to incomplete data. The presence of censoring is thus a major challenge when interest is in inferring about the upper tail of the distribution, since such investigations involve estimation based on the largest observations in a dataset, but due to the censoring, some of these will be incomplete.

Extreme value theory in the presence of censoring has been already studied in extreme value statistics, see, e.g., Beirlant et al. (2007), Einmahl et al. (2008), Gomes & Neves (2011), or Worms & Worms (2014). Recently, Bladt & Rodionov (2024) developed convergence theory for extreme level Kaplan-Meier integrals. This new approach has been used in Bladt et al. (2024) in order to propose bias-corrected estimator of the extreme value index.

Missions

This Master 2 internship will be divided into several parts:
- Familiarization with extreme value theory;
- Detailed study of the paper Bladt et al. (2024);
- Quantile estimation: theory, simulations, real data analysis (if time).

Candidate profile

No prerequisites in extreme value theory are required (this knowledge can be acquired in part 1 of the Master 2 internship), but a solid background in statistics and probability, particularly in stochastic processes, is crucial. A good knowledge of R is also required.

How to apply

The candidates should send a CV and grade transcripts (each year of the bachelor) to Armelle Guillou (armelle.guillou@math.unistra.fr).

A PhD may be proposed as a continuation of this internship.

References

[1] Beirlant, J., Guillou, A., Dierckx, G. & Fils-Villetard, A. (2007). Estimation of the extreme value index and extreme quantiles under random censoring, Extremes, 10, 151-174.

[2] Bladt, M., Goegebeur, Y. & Guillou, A. (2024). Asymptotically unbiased estimator of the extreme value index under random censoring, https://hal.science/hal-04786783v1.

[3] Bladt, M. & Rodionov, I. (2024). Censored extreme value estimation, https://arxiv.org/abs/2312.10499.

[4] Einmahl, J.H.J., Fils-Villetard, A. & Guillou, A. (2008). Statistics of extremes under random Censoring, Bernoulli, 14, 207-227.

[5] Gomes, M.I. & Neves, M.M. (2011). Estimation of the extreme value index for randomly censored data, Biometrical Letters, 48, 1-22.

[6] Worms, J. & Worms, R. (2014). New estimators of the extreme value index under random right censoring, for heavy-tailed distributions, Extremes, 17, 337-358.