Contributions to the Chair "Stress Test"

Stéphane Girard
Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France

May 24, 2023

1 Ongoing works

1.1 Estimation of extreme quantiles from a mixture of Pareto distributions (with E. Gobet)

We focus on the situation where the data are drawn from a continuous mixture a Pareto-type distributions with tail-index $\gamma \in [\gamma_1, \gamma_2]$ where $\gamma_1 > 0$. We show that the mixture is still heavy-tailed with tail-index γ_2 and that, in this case, the Hill estimator of the tail-index is strongly biased. A bias correction is introduced thanks to a sharp expansion of the survival distribution of the mixture in the tail. An estimator of extreme quantiles is then derived. Our current work consists in investigating the behavior of this new bias-corrected estimator (i) from the asymptotic point of view, (ii) on simulated data, and (iii) on financial data.

1.2 Estimation of extreme Expected Shortfall with neural networks (with M. Allouche and E. Gobet)

We propose a new parametrization for one-hidden layer neural networks combining eLU and ReLu units able to estimate the Expected Shortfall at extreme levels, starting from heavy-tailed data. We provide an analysis of the uniform error between the extreme log-Expected Shortfall and its neural network approximation. Numerical experiments are conducted on simulated data to compare the performance of our method with other estimators from the literature.

1.3 Reduced-bias estimation of the conditional tail moment at intermediate and extreme levels for heavy-tailed distributions (with M. Allouche and J. El-Methni)

Conditional tail moments are defined as the moments of the loss distribution above a given quantile. Estimating them permits to estimate a number of risk measures such as conditional tail expectation, expected shortfall, conditional value-at-risk or conditional tail variance. For this reason, conditional tail moments play an important role in actuarial and financial risk management applications. Unfortunately, available estimators of extreme conditional tail moments are typically biased and hence may suffer from poor finite-sample performance. We focus on the definition of bias-reduced extreme conditional tail moment estimators for heavy-tailed distributions. The construction of the reduced bias estimators is based on the investigation of the asymptotic proportionality relationship between extreme conditional tail moments and extreme quantiles, as well as of an extrapolation formula established in the heavy-tailed context. We quantify and estimate the bias induced by the use of these relationships and then introduce bias-reduced estimators whose asymptotic properties are rigorously shown. Their finite-sample properties are assessed on a simulation study and illustrated on real data.

1.4 A refined extreme quantile estimator for Weibull tail-distributions (with J. El-Methni)

We address the estimation of extreme quantiles of Weibull tail-distributions. Since such quantiles are asymptotically larger than the sample maximum, their estimation requires extrapolation methods. In the case of Weibull tail-distributions, classical extreme-value estimators are numerically outperformed by estimators dedicated to this set of light-tailed distributions. The latter estimators of extreme quantiles are based on two key quantities: an order statistic to estimate an intermediate quantile and an estimator of the Weibull tail-coefficient used to extrapolate. The common practice is to select the same intermediate sequence for both estimators. We show how an adapted choice of two different intermediate sequences leads to a reduction of the asymptotic bias associated with the resulting refined estimator. The asymptotic normality of the refined estimator is established, and a data-driven method is introduced for the practical selection of the intermediate sequences. Our approach is compared to three estimators of extreme quantiles dedicated to Weibull tail-distributions on simulated data. An illustration on daily wind measures is also provided.

2 Results

2.1 Publications

- Multivariate expectile-based distribution: properties, Bayesian inference, and applications: `arbel:girard:2023`

  Expectiles form a family of risk measures that have recently gained interest over the more common value-at-risk or return levels, primarily due to their capability to be determined by the probabilities of tail values and magnitudes of realisations at once. However, a prevalent and ongoing challenge of expectile inference is the problem of uncertainty quantification, which is especially critical in sensitive applications, such as in medical, environmental or engineering tasks. We address this issue by developing a novel distribution, termed the multivariate expectile based distribution (MED), that possesses an expectile as a closed-form parameter. Desirable properties of the distribution, such as log-concavity, make it an excellent fitting distribution in multivariate applications. Maximum likelihood estimation and Bayesian inference algorithms are described. Simulated examples and applications to expectile and mode estimation illustrate the usefulness of the MED for uncertainty quantification.

- A refined Weissman estimator for extreme quantiles: `allouche:elmethni:2023`

  Weissman extrapolation methodology for estimating extreme quantiles from heavy-tailed distributions is based on two estimators: an order statistic to estimate an intermediate quantile and an estimator of the tail-index. The common practice is to select the same intermediate sequence for both estimators. In this work, we show how an adapted choice of two different intermediate sequences leads to a reduction of the asymptotic bias associated with the resulting refined Weissman estimator. The asymptotic normality of the latter estimator is established and a data-driven method is introduced for the practical selection of the intermediate sequences. Our approach is compared to Weissman estimator and to six bias reduced estimators of extreme quantiles on a large scale simulation study. It appears that the refined Weissman estimator outperforms its competitors in a wide variety of situations, especially in the challenging high bias cases. Finally, an illustration on an actuarial real data set is provided.

- Extreme Partial Least-Squares regression `bousebata:enjolras:2023`

  We propose a new approach, called Extreme-PLS, for dimension reduction in regression and adapted to distribution tails. The objective is to find linear combinations of predictors that best explain the extreme values of the response variable in a non-linear inverse regression

model. The asymptotic normality of the Extreme-PLS estimator is established in the single-index framework and under mild assumptions. The performance of the method is assessed on simulated data. A statistical analysis of French farm income data, considering extreme cereal yields, is provided as an illustration.

2.2 Submitted manuscripts

- Estimation of extreme quantiles from heavy-tailed distributions with neural networks (with M. Allouche and E. Gobet): `allouche:girard:2023`

  In the context of the PhD thesis of M. Allouche, we propose a new parametrization for one-hidden layer eLU neural networks able to estimate extreme quantiles, starting from heavy-tailed data. We provide an analysis of the uniform error between the extreme log-quantile and its neural network approximation. Numerical experiments are conducted on simulated data to compare the performance of our method with other estimators from the literature. The method is then extended to the estimation of conditional extreme quantiles by combining multi-layers ReLU neural networks with the previous one layer eLU neural network. An illustration is provided on rainfall data in the Cévennes-Vivarais region.

- Reparameterization of extreme value framework for improved Bayesian workflow (with J. Arbel, A. Dutfoy and T. Moins): `moins:arbel:2023`

  Using Bayesian methods for extreme value analysis offers an alternative to frequentist ones, with several advantages such as easily dealing with parametric uncertainty or studying irregular models. However, computation can be challenging and the efficiency of algorithms can be altered by poor modelling choices, and among them the parameterization is crucial. We focus on the Poisson process characterization of univariate extremes and outline two key benefits of an orthogonal parameterization. First, Markov chain Monte Carlo convergence is improved when applied on orthogonal parameters. This analysis relies on convergence diagnostics computed on several simulations. Second, orthogonalization also helps deriving Jeffreys and penalized complexity priors, and establishing posterior propriety thereof. Our framework is applied to return level estimation of Garonne flow data (France).

- On the use of a local $\hat{R}$ to improve MCMC convergence diagnostic (with J. Arbel, A. Dutfoy and T. Moins): `moins:arbel:2023b`

  Diagnosing convergence of Markov chain Monte Carlo is crucial and remains an essentially unsolved problem. Among the most popular methods, the potential scale reduction factor, commonly named $\hat{R}$, is an indicator that monitors the convergence of output chains to a target distribution, based on a comparison of the between-and within-variances. Several improvements have been suggested since its introduction in the 90s. Here, we aim at better understanding the $\hat{R}$ behaviour by proposing a localized version that focuses on quantiles of the target distribution. This new version relies on key theoretical properties of the associated population value. It naturally leads to proposing a new indicator $\hat{R}^{\infty}$, which is shown to allow both for localizing the Markov chain Monte Carlo convergence in different quantiles of the target distribution, and at the same time for handling some convergence issues not detected by other $\hat{R}$ versions.

2.3 Conferences

- Member of Scientific Program Committee of the 16th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2023), HTW Berlin, University of Applied Sciences, Berlin, Germany, from 16 to 18 December 2023. Organizer of an invited session "*Machine learning for extremes*" (4 invited talks to be confirmed).

- Chair of one session of the EVA conference, Milan, june 2023.
- International conferences (so far):
  EVA girard:opitz:2023,arbel:dutfoy:2023,allouche:girard:2023c,elmethni:girard:2023b,
  ERCIM elmethni:girard:2023, ATI workshop allouche:girard:2023b.

3 PhD theses supervision

Two PhD theses are founded by the Chair:

- Michael Allouche (co-advised with Emmanuel Gobet, Ecole Polytechnique) "*Contributions to generative modeling and dictionary learning: Theory and application*", Institut Polytechnique de Paris, defended on December 9th, 2022.
- Jean Pachebat (co-advised with Emmanuel Gobet, Ecole Polytechnique) "*How AI models can deal with extreme values? Application to risk assessment*", Institut Polytechnique de Paris, started on February 1st, 2023.

One PhD thesis is connected to the Chair topics:

- Théo Moins (co-advised with Anne Dutfoy, EDF R&D) "*Bayesian assessment of extrapolation limits in extreme-value statistics*", started on October 1st, 2020.

4 Editorial activities

- Associate Editor of *Revstat* since 2019.
- Member of the Advisory Board of *Dependence Modeling* since 2015.

5 Fundings

S. Girard and E. Gobet got a funding of 3,500 euros from the MIAI Grenoble Alpes institute (Multidisciplinary Institute in Artificial Intelligence) for the PhD project of J. Pachebat "*How AI models can deal with extreme values? Application to risk assessment*".

References

allouche:elmethni:2023 .....
M. Allouche, J. El-methni, and S. Girard.
A refined Weissman estimator for extreme quantiles.
Extremes, 2023.
https://doi.org/10.1007/s10687-022-00452-8, to appear.

allouche:girard:2023c .....
M. Allouche, S. Girard, and E. Gobet.
Estimation of extreme quantiles from heavy-tailed distributions with neural networks.
In *13th International Conference on Extreme Value Analysis*, Milan, Italy, june 2023.

allouche:girard:2023 .....
M. Allouche, S. Girard, and E. Gobet.
Estimation of extreme quantiles from heavy-tailed distributions with neural networks.
https://hal.archives-ouvertes.fr/hal-03751980, 2023.

allouche:girard:2023b .....
M. Allouche, S. Girard, and E. Gobet.
Generative modeling of extremes with neural networks.
In *Accelerating Generative Models and Nonconvex Optimisation Workshop*, London, United Kingdom, March 2023. The Alan Turing Institute.

arbel:dutfoy:2023 .....
J. Arbel, A. Dutfoy, S. Girard, and T. Moins.
Reparameterization of extreme value framework for improved bayesian workflow.
In *13th International Conference on Extreme Value Analysis*, Milan, Italy, june 2023.

arbel:girard:2023 .....
J. Arbel, S. Girard, H. Nguyen, and A. Usseglio-Carleve.
Multivariate expectile-based distribution: properties, Bayesian inference, and applications.
Journal of Statistical Planning and Inference, 225:146–170, 2023.

bousebata:enjolras:2023 .....
M. Bousebata, G. Enjolras, and S. Girard.
Extreme Partial Least-Squares.
Journal of Multivariate Analysis, 194:105101, 2023.

elmethni:girard:2023b .....
J. El-Methni and S. Girard.
A refined extreme quantiles estimator for weibull tail-distributions.
In *13th International Conference on Extreme Value Analysis*, Milan, Italy, june 2023.

elmethni:girard:2023 .....
J. El-Methni and S. Girard.
A refined extreme quantiles estimator for weibull tail-distributions.
In *16th International Conference of the ERCIM WG on Computational and Methodological Statistics*, Berlin, Germany, December 2023.

girard:optiz:2023 .....
S. Girard, T. Opitz, A. Usseglio-Carleve, and C. Yan.
Analysis of variability in extremes with application in clustering of extreme events.
In *13th International Conference on Extreme Value Analysis*, Milan, Italy, june 2023.

moins:arbel:2023b .....
T. Moins, J. Arbel, A. Dutfoy, and S. Girard.
On the use of a local $\hat{R}$ to improve MCMC convergence diagnostic.
https://hal.inria.fr/hal-03600407, 2023.

moins:arbel:2023 .....
T. Moins, J. Arbel, S. Girard, and A. Dutfoy.
Reparameterization of extreme value framework for improved Bayesian workflow.
https://hal.archives-ouvertes.fr/hal-03806159, 2023.