Contributions to the Chair “Stress Test”

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

May 23, 2024

1 Scientific results

1.1 Publications

- Estimation of extreme quantiles from heavy-tailed distributions with neural networks (with M. Allouche and E. Gobet): `allouche:girard:2024`

  In the context of the PhD thesis of M. Allouche, we proposed 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.

- On the use of a local $\hat{R}$ to improve MCMC convergence diagnostic (with J. Arbel, A. Dutfoy and T. Moins): `moins:arbel:2024`

  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.

- Reduced-bias estimation of the extreme conditional tail expectation for Box-Cox transforms of heavy-tailed distributions (with M. Allouche and J. El-Methni): `allouche:methni:2024`

  Conditional tail expectation (CTE) is a coherent risk measure defined as the mean of the loss distribution above a high quantile. The existence of the CTE as well as the asymptotic properties of associated estimators however require integrability conditions that may be violated when dealing with heavy-tailed distributions. We introduce Box-Cox transforms of the CTE that have two benefits. First, they alleviate these theoretical issues. Second, they enable to recover a number of risk measures such as conditional tail expectation, expected shortfall, conditional value-at-risk or conditional tail variance. The construction of dedicated estimators is based on the investigation of the asymptotic relationship between

Box-Cox transforms of the CTE and quantiles at extreme probability levels, as well as on an extrapolation formula established in the heavy-tailed context. We quantify and estimate the bias induced by the use of these approximations and then introduce reduced-bias estimators whose asymptotic properties are rigorously shown. Their finite-sample properties are assessed on a simulation study and illustrated on real data, highlighting the practical interest of both the bias reduction and the Box-Cox transform.

- A refined extreme quantile estimator for Weibull tail-distributions (with J. El-Methni): `methni:girard:2024`

  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. This analysis is supported by an asymptotic normality result associated with the refined estimator. A data-driven method is introduced for the practical selection of the intermediate sequences and our approach is compared to three estimators of extreme quantiles dedicated to Weibull tail-distributions on simulated data. An illustration on a real data set of daily wind measures is also provided.

1.2 Submitted manuscripts

- Learning out-of-sample Expected Shortfall and Conditional Tail Moments with neural networks. Application to cryptocurrency data (with M. Allouche and E. Gobet): `allouche:girard:2024b`

  We propose new parameterizations for neural networks in order to estimate out-of-sample Expected Shortfall, and even more generally, out-of-sample conditional tail moments, in heavy-tailed settings as functions of confidence levels. The proposed neural network estimator is able to extrapolate in the distribution tails thanks to an extension of the usual extreme-value second-order condition to an arbitrary order. The convergence rate of the uniform error between the log-conditional tail moment and its neural network approximation is established. The finite sample performance of the neural network estimator is compared to bias-reduced extreme-value competitors on simulated data. It is shown that our method outperforms them in difficult heavy-tailed situations where other estimators almost all fail. Finally, the neural network estimator is tested on real data to investigate the behavior of cryptocurrency extreme loss returns.

- On the simulation of extreme events with neural networks (with M. Allouche and E. Gobet): `allouche:girard:2024c`

  This article aims at investigating the use of generative methods based on neural networks to simulate extreme events. Although very popular, these methods are mainly invoked in empirical works. Therefore, providing theoretical guidelines for using such models in extreme values context is of primal importance. To this end, we propose an overview of most recent generative methods dedicated to extremes, giving some theoretical and practical tips on their tail behaviour thanks to both extreme-value and copula tools.

- ANOVEX: ANalysis Of Variability for heavy-tailed EXtremes (with T. Opitz and A. Usseglio-Carleve): `girard:opitz:2024`

(May 23, 2024)

Analysis of variance (ANOVA) is commonly employed to assess differences in the means of independent samples. However, it is unsuitable for evaluating differences in tail behaviour, especially when means do not exist or empirical estimation of moments is inconsistent due to heavy-tailed distributions. Here, we propose an ANOVA-like decomposition to analyse tail variability, allowing for flexible representation of heavy tails through a set of user-defined extreme quantiles, possibly located outside the range of observations. Building on the assumption of regular variation, we introduce a test for significant tail differences among multiple independent samples and derive its asymptotic distribution. We investigate the theoretical behaviour of the test statistics for the case of two samples, each following a Pareto distribution, and explore strategies for setting hyperparameters in the test procedure. To demonstrate the finite-sample performance, we conduct simulations that highlight generally reliable test behaviour for a wide range of situations. The test is applied to identify clusters of financial stock indices with similar extreme log-returns and to detect temporal changes in daily precipitation extremes at rain gauges in Germany.

- Shrinkage for Extreme Partial Least Squares (with J. Arbel and H. Lorenzo): `arbel:girard:2024`

  This work focuses on dimension-reduction techniques for modelling conditional extreme values. Specifically, we investigate the idea that extreme values of a response variable can be explained by nonlinear functions derived from linear projections of an input random vector. In this context, the estimation of projection directions is examined, as approached by the Extreme Partial Least Squares (EPLS) method–an adaptation of the original Partial Least Squares (PLS) method tailored to the extreme-value framework. Further, a novel interpretation of EPLS directions as maximum likelihood estimators is introduced, utilizing the von Mises–Fisher distribution applied to hyperballs. The dimension reduction process is enhanced through the Bayesian paradigm, enabling the incorporation of prior information into the projection direction estimation. The maximum a posteriori estimator is derived in two specific cases, elucidating it as a regularization or shrinkage of the EPLS estimator. We also establish its asymptotic behavior as the sample size approaches infinity. A simulation data study is conducted in order to assess the practical utility of our proposed method. This clearly demonstrates its effectiveness even in moderate data problems within high-dimensional settings. Furthermore, we provide an illustrative example of the method's applicability using French farm income data, highlighting its efficacy in real-world scenarios.

1.3 Conferences

- Organization of a "Computational and statistical methods for extremes in finance" session at the *International Conference on Computational Finance* (ICCF), Amsterdam, april 2024.
- Organization of an "Extreme" session at the "*Journées MAS de la SMAI*", Poitiers, august 2024.
- Chair of the "Extreme and risk" session at the "*Journées de Statistique de la SFdS*", Bordeaux, may 2024.
- Keynote speaker at the "*Journées de Statistique de la SFdS*", Bordeaux, may 2024.
- International conferences (so far):
  SIAM Conference on Uncertainty Quantification `allouche:girard:2024f,allouche:gobet:2024`, ICCF `allouche:girard:2024e,pachebat:girard:2024`, a workshop `allouche:girard:2024d`.

2 PhD theses supervision

One PhD thesis is founded by the Chair:

- 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.

3 Editorial activities

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

4 Ongoing works - Perspectives

4.1 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 cryptocurrency data.

4.2 Bias-reduced simulation of extreme values using Generative Adversarial Networks (with M. Allouche and E. Gobet)

Our goal is to take profit of the bias reduction techniques developped in allouche:girard:2024 to improve the performance of the EVGAN method proposed in the PhD thesis of M. Allouche.

4.3 Simulation of multivariate extreme events with generative models (with E. Gobet and J. Pachebat)

One of the objectives of the PhD thesis of J. Pachebat is to improve the simulation performance of the EVGAN method in high multidimensional settings. The main idea is to leverage multivariate extreme-value theory to better account for the tail dependence.

4.4 Functional Extreme-PLS (with C. Pakzad)

We aim at proposing an extreme dimension reduction method extending the Extreme-PLS approach to the case where the covariate lies in a possibly infinite-dimensional Hilbert space. The ideas are partly borrowed from both Partial Least-Squares and Sliced Inverse Regression techniques. As such, the method will rely on the projection of the covariate onto a subspace and maximize the covariance between its projection and the response conditionally to an extreme event to capture most of the tail information. Moreover, we shall link the covariate and the heavy-tailed response through a non-linear inverse single-index model. Our goal is first to infer the index in this regression framework and second to show the asymptotic consistency of the estimator with convergence rates under the model.

This is the topic of the postdoctoral work of C. Pakzad.

4.5 Measure of extreme inequality in wealth distribution (with J. El-Methni)

In 1912, Italian economist Corrado Gini introduced the first measure of inequality, the Gini coefficient. This statistical index is used to account for the variability of a variable within a population. It thus makes it possible to quantify the level of inequality in the distribution of wealth in a population. The Gini coefficient is a number ranging from 0 to 1, where 0 means perfect equality (everyone has the same share of wealth) and 1 represents perfect inequality (one person has all the wealth). However, it turns out that, if the focus is on the distribution of wealth among the highest income earners, the Gini coefficient underestimates inequality. Indeed, if the variable of interest follows a heavy-tailed distribution (with possibly infinite variance), then it has the same Gini index as a light-tailed distribution. The Gini coefficient is therefore robust to extreme observations and fails to discriminate between two different distributions of wealth for extreme values. The objectives are to define and theoretically study a new coefficient for measuring inequalities in the distribution of wealth, dedicated to extreme values. This will enable to take better account of the behavior of the distribution tail and no longer underestimate inequalities.

A PhD thesis co-advised with J. El-Methni should start on these research questions in fall 2024.

References

`allouche:girard:2024` ....
M. Allouche, S. Girard, and E. Gobet.
Estimation of extreme quantiles from heavy-tailed distributions with neural networks.
Statistics and Computing, 34:12, 2024.

`allouche:girard:2024e` ....
M. Allouche, S. Girard, and E. Gobet.
Learning of extreme expected shortfall with neural networks. Application to cryptocurrency data.
In *International Conference on Computational Finance*, Amsterdam, Netherlands, April 2024.

`allouche:girard:2024d` ....
M. Allouche, S. Girard, and E. Gobet.
Learning out-of-sample Expected Shortfall and Conditional Tail Moments with neural networks. Application to cryptocurrency data.
In *Stochastic Numerics and Statistical Learning: Theory and Applications Workshop*, Thuwal, Saudi Arabia, May 2024.

`allouche:girard:2024b` ....
M. Allouche, S. Girard, and E. Gobet.
Learning out-of-sample expected shortfall and conditional tail moments with neural networks. application to cryptocurrency data.
`hal-03751980`, 2024.

`allouche:girard:2024c` ....
M. Allouche, S. Girard, and E. Gobet.
On the simulation of extreme events with neural networks.
`hal-04416809`, 2024.

`allouche:girard:2024f` ....
M. Allouche, S. Girard, E. Gobet, and J. Pachebat.
Neural networks based learning applied to extreme statistics and sampling rare events.
In *SIAM Conference on Uncertainty Quantification*, Trieste, Italy, March 2024.

`allouche:gobet:2024` ....
M. Allouche, E. Gobet, and S. Girard.
Estimation of extreme risk measures with neural networks.
In *SIAM Conference on Uncertainty Quantification*, Trieste, Italy, March 2024.

`allouche:methni:2024` ....
M. Allouche, J. El Methni, and S. Girard.
Reduced-bias estimation of the extreme conditional tail expectation for Box-Cox transforms of heavy-tailed distributions.
Journal of Statistical Planning and Inference, 2024.
to appear.

`arbel:girard:2024` ....
J. Arbel, S. Girard, and H. Lorenzo.
Shrinkage for Extreme Partial Least Squares.
`hal-04251783`, 2024.

`girard:opitz:2024` ....
S. Girard, T. Opitz, and A. Usseglio-Carleve.
ANOVEX: ANalysis of Variability for heavy-tailed EXtremes.
`hal-04200300`, 2024.

`methni:girard:2024` ....
J. El Methni and S. Girard.
A refined extreme quantile estimator for Weibull tail-distributions.
REVSTAT - Statistical Journal, 2024.
to appear.

`moins:arbel:2024` ....
T. Moins, J. Arbel, A. Dutfoy, and S. Girard.
On the use of a local $\hat{R}$ to improve MCMC convergence diagnostic.
Bayesian Analysis, 2024.
to appear.

`pachebat:girard:2024` ....
J. Pachebat, S. Girard, and E. Gobet.
Simulation of multivariate extreme events with generative models.
In *International Conference on Computational Finance*, Amsterdam, Netherlands, April 2024.