Contributions to extreme-value analysis

Stéphane Girard

INRIA Rhône-Alpes & LJK (team MISTIS).

655, avenue de l'Europe, Montbonnot. 38334 Saint-Ismier Cedex, France

Stephane.Girard@inria.fr

Abstract: This report summarizes my contributions to extreme-value statistics. I worked on the estimation of the extreme-value index, Weibull-tail coefficient and end-point estimation. I also studied conditional extremes, that is the situation where the extreme-value behavior depends on a covariate.

Contributions

Extreme value theory is a branch of statistics dealing with the extreme deviations from the bulk of probability distributions. More specifically, it focuses on the limiting distributions for the minimum or the maximum of a large collection of random observations from the same arbitrary (unknown) distribution. Let $x_1 < \cdots < x_n$ denote n ordered observations from a random variable X representing some quantity of interest. A p_n-quantile of X is the value q_{p_n} such that the probability that X is greater than q_{p_n} is p_n, i.e. $P(X > q_{p_n}) = p_n$. When $p_n < 1/n$, such a quantile is said to be extreme since it is usually greater than the maximum observation x_n. To estimate such extreme quantiles requires therefore specific methods to extrapolate information beyond the observed values of X. Those methods are based on Extreme value theory. This kind of issues appeared in hydrology. One objective was to assess risk for highly unusual events, such as 100-year floods, starting from flows measured over 50 years.

The decay of the survival function $P(X > x) = 1 - F(x)$, where F denotes the cumulative distribution function associated to X, is driven by a real parameter called the extreme-value index γ. I have proposed several estimators for this parameter, see [1, 2, 3, 4]. When this parameter is positive, the survival function is said to be heavy-tailed, when this parameter is negative, the survival function vanishes above its right end point. Some estimation methods for the right end point have been proposed in [5, 6]. If this parameter is zero, then the survival function decreases to zero at an exponential rate. An important part of our work is dedicated to the study of such distributions. For instance, in reliability, the distributions of interest are included in a semi-parametric family whose tails are decreasing exponentially fast. These so-called Weibull tail-distributions encompass a variety of light-tailed distributions, such as Weibull, Gaussian, gamma and logistic distributions. Let us recall that a cumulative distribution function F has a Weibull tail if it satisfies the following property: There exists $\theta > 0$ such that for all $\lambda > 0$,

$$\lim_{y \to \infty} \frac{\log(1 - F(\lambda y))}{\log(1 - F(y))} = \lambda^{1/\theta}.$$

Dedicated methods have been proposed to estimate the Weibull tail-coefficient θ since the relevant information is only contained in the extreme upper part of the sample. More specifically, the estimators I proposed are based on the log-spacings between the upper order statistics [7, 8, 9, 10, 11]. See also [12, 13, 14, 15] for the estimation of the associated extreme quantiles. These methods can also be seen as an improvement of the ET method [16, 17, 18]. Of course, the choice of a tail model is an important issue, see [19] for the introduction of a goodness-of-fit test and [20] for its implementation.

I also addressed the estimation of extreme level curves. This problem is equivalent to estimating quantiles when covariate information is available and when their order converges to one as the sample size increases. We show that, under some conditions, these so-called "extreme conditional quantiles" can still be estimated through a kernel estimator of the conditional survival function. Sufficient conditions on the rate of convergence of their order to one are provided to obtain asymptotically Gaussian distributed estimators. Making use of this result, some estimators of the extreme-value parameters are introduced and extreme conditional quantiles estimators are deduced [21, 22, 23, 24, 25, 26, 27, 28, 29, 30]. Finally, the tail copula is widely used to describe the dependence in the tail of multivariate distributions. In some situations such as risk management, the dependence structure may be linked with some covariate. The tail copula thus depends on this covariate and is referred to as the conditional tail copula. The aim of [31] is to propose a nonparametric estimator of the conditional tail copula and to establish its asymptotic normality.

In the multivariate context, we focus on extreme geometric quantiles [32]. Their asymptotics are established, both in direction and magnitude, under suitable integrability conditions, when the norm of the associated index vector tends to one.

Applications of extreme-value theory are found in hydrology [33, 34, 35] and more generally in risk estimation [36, 37, 38].

References

[1] E. Deme, L. Gardes, and S. Girard. On the estimation of the second order parameter for heavy-tailed distributions. *REVSTAT - Statistical Journal*, 11:277–299, 2013.

[2] J. Diebolt, M. El-Aroui, M. Garrido, and S. Girard. Quasi-conjugate Bayes estimates for GPD parameters and application to heavy tails modelling. *Extremes*, 8:57–78, 2005.

[3] L. Gardes and S. Girard. Asymptotic properties of a Pickands type estimator of the extreme value index. In Louis R. Velle, editor, *Focus on probability theory*, pages 133–149. Nova Science, New-York, 2006.

[4] L. Gardes and S. Girard. Asymptotic distribution of a Pickands type estimator of the extreme value index. *Comptes-Rendus de l'Académie des Sciences, Série I*, 341:53–58, 2005.

[5] S. Girard, A. Guillou, and G. Stupfler. Estimating an endpoint with high order moments in the Weibull domain of attraction. *Statistics and Probability Letters*, 82:2136–2144, 2012.

[6] S. Girard, A. Guillou, and G. Stupfler. Estimating an endpoint with high order moments. *Test*, 21:697–729, 2012.

[7] L. Gardes, S. Girard, and A. Guillou. Weibull tail-distributions revisited: a new look at some tail estimators. *Journal of Statistical Planning and Inference*, 141(1):429–444, 2011.

[8] L. Gardes and S. Girard. Estimation of the Weibull tail-coefficient with linear combination of upper order statistics. *Journal of Statistical Planning and Inference*, 138:1416–1427, 2008.

[9] J. Diebolt, L. Gardes, S. Girard, and A. Guillou. Bias-reduced estimators of the Weibull tail-coefficient. *Test*, 17:311–331, 2008.

[10] L. Gardes and S. Girard. Comparison of Weibull tail-coefficient estimators. *REVSTAT - Statistical Journal*, 4(2):373–188, 2006.

[11] S. Girard. A Hill type estimate of the Weibull tail-coefficient. *Communication in Statistics - Theory and Methods*, 33(2):205–234, 2004.

[12] L. Gardes and S. Girard. Estimation de quantiles extrêmes pour les lois à queue de type Weibull : une synthèse bibliographique. *Journal de la Société Française de Statistique*, 154:98–118, 2013.

[13] J. El Methni, L. Gardes, S. Girard, and A. Guillou. Estimation of extreme quantiles from heavy and light tailed distributions. *Journal of Statistical Planning and Inference*, 142(10):2735–2747, 2012.

[14] J. Diebolt, L. Gardes, S. Girard, and A. Guillou. Bias-reduced extreme quantiles estimators of Weibull-tail distributions. *Journal of Statistical Planning and Inference*, 138:1389–1401, 2008.

[15] L. Gardes and S. Girard. Estimating extreme quantiles of Weibull tail-distributions. *Communication in Statistics - Theory and Methods*, 34:1065–1080, 2005.

[16] J. Diebolt and S. Girard. A note on the asymptotic normality of the ET method for extreme quantile estimation. *Statistics and Probability Letters*, 62(4):397–406, 2003.

[17] J. Diebolt, M. Garrido, and S. Girard. Asymptotic normality of the ET method for extreme quantile estimation. Application to the ET test. *Comptes-Rendus de l'Académie des Sciences, Série I*, 337:213–218, 2003.

[18] S. Girard and J. Diebolt. Consistency of the ET method and smooth variations. *Comptes-Rendus de l'Académie des Sciences, Série I*, 329:821–826, 1999.

[19] J. Diebolt, M. Garrido, and S. Girard. A goodness-of-fit test for the distribution tail. In M. Ahsanulah and S. Kirmani, editors, *Extreme Value Distributions*, pages 95–109. Nova Science, New-York, 2007.

[20] J. Diebolt, J. Ecarnot, M. Garrido, S. Girard, and D. Lagrange. Le logiciel Extremes, un outil pour l'étude des queues de distribution. *La revue de Modulad*, 30:53–60, 2003.

[21] L. Gardes and S. Girard. On the estimation of the functional Weibull tail-coefficient. *Journal of Multivariate Analysis*, 2015. to appear.

[22] A. Daouia, L. Gardes, and S. Girard. On kernel smoothing for extremal quantile regression. *Bernoulli*, 19:2557–2589, 2013.

[23] L. Gardes and S. Girard. Functional kernel estimators of large conditional quantiles. *Electronic Journal of Statistics*, 6:1715–1744, 2012.

[24] J. Carreau and S. Girard. Spatial extreme quantile estimation using a weighted log-likelihood approach. *Journal de la Société Française de Statistique*, 152(3):66–83, 2011.

[25] A. Daouia, L. Gardes, S. Girard, and A. Lekina. Kernel estimators of extreme level curves. *Test*, 20(14):311–333, 2011.

[26] L. Gardes, S. Girard, and A. Lekina. Functional nonparametric estimation of conditional extreme quantiles. *Journal of Multivariate Analysis*, 101:419–433, 2010.

[27] L. Gardes and S. Girard. A moving window approach for nonparametric estimation of the conditional tail index. *Journal of Multivariate Analysis*, 99:2368–2388, 2008.

[28] S. Girard and S. Louhichi. On the strong consistency of the kernel estimator of extreme conditional quantiles. In E. Ould-Said, editor, *Recent Advances in Statistical Methodology and its Application*. Springer, 2015. to appear.

[29] A. Daouia, L. Gardes, and S. Girard. Nadaraya's estimates for large quantiles and free disposal support curves. In I. Van Keilegom and P. Wilson, editors, *Exploring research frontiers in contemporary statistics and econometrics*, pages 1–22. Springer, 2012.

[30] L. Gardes and S. Girard. Functional kernel estimators of conditional extreme quantiles. In F. Ferraty, editor, *Recent advances in functional data analysis and related topics*, pages 135–140. Springer, Physica-Verlag, 2011.

[31] L. Gardes and S. Girard. Nonparametric estimation of the conditional tail copula. *Journal of Multivariate Analysis*, 2015. to appear.

[32] S. Girard and A. G. Stupfler. Intriguing properties of extreme geometric quantiles. *REVSTAT - Statistical Journal*, 2015. to appear.

[33] J. Carreau, D. Ceresetti, E. Ursu, S. Anquetin, J.D. Creutin, L. Gardes, S. Girard, and G. Molinié. Evaluation of classical spatial-analysis schemes of extreme rainfall. *Natural Hazards and Earth System Sciences*, 12:3229–3240, 2012.

[34] B. Barroca, P. Bernardara, S. Girard, and G. Mazo. Considering hazard estimation uncertain in urban resilience strategies. *Natural Hazards and Earth System Sciences*, 15:25–34, 2015.

[35] L. Gardes and S. Girard. Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels. *Extremes*, 13(2):177–204, 2010.

[36] E. Deme, S. Girard, and A. Guillou. Reduced-bias estimators of the conditional tail expectation for heavy-tailed distributions. In M. Hallin et al., editor, *Mathematical Statistics and Limit Theorems*, pages 105–123. Springer, 2015.

[37] J. El Methni, L. Gardes, and S. Girard. Nonparametric estimation of extreme risks from conditional heavy-tailed distributions. *Scandinavian Journal of Statistics*, 41:988–1012, 2014.

[38] E. Deme, S. Girard, and A. Guillou. Reduced-bias estimator of the proportional hazard premium for heavy-tailed distributions. *Insurance: Mathematics and Economics*, 22:550–559, 2013.