Empirical Processes of Extreme Cluster Functionals

Holger Drees, University of Hamburg
joint work with Holger Rootzén, Chalmers University Gothenburg

In the literature on extreme value statistics for time series, quite general results are known about estimators of the marginal tails. Usually it is assumed that the suitably standardized exceedances $X_{i,n} := a_n^{-1}(X_i - u_n)$ converge in distribution to a generalized Pareto distribution, i.e. $P(X_{1,n} > x \mid X_{1,n} > 0) \to (1+\gamma x)_+^{-1/\gamma}$ for some extreme value index $\gamma \in \mathbb{R}$; here $(u_n)_{n\in\mathbb{N}}$ denotes a sequence of thresholds tending to the right endpoint of the range of the marginal d.f. F. Rootzén (2007) analyzed the asymptotic behavior of the tail empirical process

$$e_n(x) := \frac{1}{\sqrt{n\bar{F}(u_n)}} \sum_{i=1}^{n} \big(\mathbb{1}_{\{X_{i,n}>x\}} - \bar{F}(u_n + a_n x)\big), \qquad x \in \mathbb{R},$$

for strong mixing time series and for absolutely regular time series; here $\bar{F}$ denotes the survival function of X_1. From a previous version of that paper, Drees (2000, 2003) derived a weighted approximation for this tail empirical process under absolute regularity and discussed statistical applications, like the analysis of estimators of the extreme value index or extreme quantiles. However, the tail empirical process does not describe the extreme value dependence structure of the times series.

By and large, results on the asymptotic behavior of estimators of the extremal dependence structure under suitable mixing conditions are restricted to estimators of the extremal index and, more general, the distribution of the size of clusters of extreme observations. Unfortunately, these estimators are of very limited value in quantitative risk management. For instance, the distribution of the total sum of losses exceeding a high threshold in a period of given length cannot be described in terms of the cluster size distribution.

We discuss empirical processes which capture more general aspects of the dependence between extreme observations of an absolutely regular stationary time series $(X_i)_{i\in\mathbb{N}}$. To this end, the time series is split into m_n blocks of length r_n, say, and the core of a cluster is defined as the minimal sequence of consecutive standardized observations $X_{i,n}$ which contains all positive values in one block. Now, following an approach by Yun (2000) (see also Segers, 2003), let $\mathcal{F}$ be a family of *cluster functionals*, i.e. measurable functionals $f(Y_{j,n})$ of a block $Y_{j,n} := (X_{i,n})_{(j-1)r_n < i \le jr_n}$ which only depend on the pertaining core.

We give sufficient conditions for the convergence of the empirical process

$$Z_n(f) := \frac{1}{\sqrt{n\bar{F}(u_n)}} \sum_{j=1}^{m_n} \big(f(Y_{j,n}) - E(f(Y_{j,n}))\big), \qquad f \in \mathcal{F},$$

to a Gaussian process with continuous sample paths. The results are obtained by first applying a 'big blocks - small blocks'-technique and then general functional limit theorems for empirical process of i.i.d. observations. Since these general conditions are quite abstract and involved, we also discuss important special cases, like (generalized) tail array sums (including the process of upcrossings over extreme intervals) and functionals describing the distribution function of order statistics in a cluster. In particular, it turns out that the general approach leads to conditions for the convergence of the tail empirical process which are usually easier to verify than the conditions established by Rootzén (2007).

References

Drees, H. (2000). Weighted approximations of tail processes for β-mixing random variables. *Ann. Appl. Probab.* **10**, 1274–1301.

Drees, H. (2003). Extreme quantile estimation for dependent data with applications to finance. *Bernoulli* **9**, 617–657.

Rootzén, H. (2007). Weak convergence of the tail empirical process for stationary sequences. Preprint, Mathematical Statistics, Chalmers University.

Segers, J. (2003). Functionals of clusters of extremes. *Adv. Appl. Probab.* **35**, 1028–1045.

Yun, S. (2000). The distributions of cluster functionals of extreme events in a dth-order Markov chain. *J. Appl. Probab.* **37**, 29–44.