Online-Workshop on 22nd July 2020 on
"Change-points and Extremes in space and time"

Abstracts

Session: Analysis of Change-points in space and time (10:05 - 11:30)

10:05-10:30 **Alexander Dürre** (Université Libre de Bruxelles)

Title: Robust change-point detection in panel data

In panel data we observe a usually high number N of individuals over a time period T. If T is large, one often assumes stability of the model over time. Different tests have been proposed to check this assumption. They are usually based on a linear or quadratic statistic, which causes unlimited impact of individual observations. A single outlier can therefore spoil the inference. Furthermore, while these methods are optimal under Gaussianity, they often behave poorly under heavy tailed data. We therefore present a new robust change-point test which is based on robustly transformed observations. We derive the asymptotic distribution under short range dependence and N, T going to infinity. A small simulation show its usefulness under heavy tailed distributions.

10:35-11:00 **Annika Bethken** (RU Bochum)

Title: Rank-based change-point tests for long-range dependent time series

We consider change-point tests based on rank statistics to test for structural changes in long-range dependent observations. Under the hypothesis of stationary time series, the asymptotic distributions of the corresponding test statistics are derived. For this, we consider a uniform reduction principle for the empirical process in a two-parameter Skorohod space equipped with a weighted supremum norm. Moreover, special emphasis is laid on an application-oriented approach to the mathematical results by considering self-normalized statistics and an approximation of the distribution of test statistics by subsampling procedures. Theoretical results are motivated by data and accompanied by simulations.

11:05-11:30 **Jana König** (TU Dortmund)

Title: Pointwise Variogram estimation in the presence of outlier blocks

The standard variogram estimator of Matheron is known to be not robust against outliers [3]. Due to this shortcoming several robust variogram estimators have been proposed in the literature [1, 3]. Genton [3], for example, proposed the use of robust scale estimators. This delivers estimators which work fine for isolated outliers [4, 5]. In simulations we observed that such estimators do not perform well in the presence of outlier blocks. An example for outlier blocks could be a large chemical spill in a restricted area. In time series analysis, robust estimates of the autocorrelation function can be constructed using the Multivariate Covariance Determinant (MCD) estimator. This estimator turns out to deliver good results for the estimation of the autocorrelation if here are outlier patches [2]. The estimator simultaneously estimates the autocorrelation for different lags. More precisely, the MCD estimates the variance-covariance matrix of given vectors. In time series analysis, the vectors are chosen such that the autocorrelations are captured in the variance-covariance matrix. The main idea of the MCD is to search for a subset of vectors which minimizes the determinant of the estimated variance-covariance matrix and to use this subset for estimation. Due to the good results obtained by the MCD in time series analysis in case of patchy outliers, we construct robust variogram estimators based on the MCD. We investigate two different approaches to transfer this concept to directional variogram estimation. A direct approach is to build vectors of p subsequent observations such that the variance of the data is captured on the main diagonal and the covariances between different lags on the minor diagonals of the variance-covariance matrix of the vectors. Assuming an intrinsical stationary process, a semivariogram estimate for lag h can be determined for the variance and the covariance for the same time lag. A second idea is to use differences between the data for the lags of interest and to build vectors such that the variogram estimates are directly captured in the estimated variance-covariance matrix. In this approach, we also investigate a different minimization criterion which, instead of minimizing the determinant of the variance covariance matrix, is based on the trace of the matrix. We illustrate the benefits and the shortcomings of the new estimators in a simulation study.

References

[1] Cressie, Hawkins: Robust estimation of the variogram: I. Mathematical Geology 12(2), 115–125 (1980)
[2] Dürre, Fried, Liboshik: Robust estimation of (partial) autocorrelation. WIREs Comput Stat 7, 205–222 (2015)
[3] Genton: Highly robust variogram estimation. Mathematical Geology 30(2), 213–221 (1998)
[4] Kerry, Oliver: Determining the effect of asymmetric data on the variogram. ii. outliers. Computers & Geosciences 33, 1233–1260 (2007)
[5] Lark: A comparison of some robust estimators of the variogram for use in soil survey. European Journal of Soil Science 51, 137–157 (2000)

Session: Analysis of Extremes in space and time (13:00 - 15:00)

13:00-13:25 **Petra Friederichs** (Uni Bonn)

Title: *Definition, description and prediction of extreme weather and climate events*

The talk gives an overview of meteorological extremes and discusses challenges with regard to the definition, description and prediction of weather and climate extremes. The complexity of the topic results from the complexity of the climate system and the very different perspectives approaching weather and climate extremes. Its description and prediction requires a rigorous statistical component, and so the talk takes a look at the role of (extreme value) statistics in it.

The climate system and its subsystems are regarded as a high-dimensional, non-linear, open physico-chemical system far away from thermodynamic equilibrium. As such they are predestined to produce large fluctuations and exhibit very different types of extremes that occur on a variety of scales in space and time. Physical constraints such as energy or momentum conservation shape the characteristics of extremes and their changes in time. However, the definition of extremes is not only based on physical parameters, but extreme weather and climate events are usually related to and thus often defined by their majorsocio-economic impacts and thus their high relevance in nature and society.

13:30-13:55 **Christina Meschede** (TU Dortmund)

Title: *Modeling waiting times of clustered extreme events*

For many applications in the field of extreme value theory, the frequency of the occurrence as well as the return times of extreme events are of great interest. Traditionally, a Poisson process is assumed as model for the occurrence of extreme events, i.e. the waiting times between two successive exceedances are exponentially distributed. In case that the observations are realizations of a strictly stationary process with existing extremal index, Ferro and Segers (2003) showed that extreme events cluster and the waiting times between them are approximately distributed as a random variable following a mixture of an exponential and a dirac measure in zero. Hees et al. (2018) proposed another model for clustered extreme events based on a fractional Poisson process leading to Mittag-Leffler distributed inter-exceedance times.

In this talk, we will introduce a generalized model that includes exponentially, mixed and Mittag-Leffler distributed waiting times as special cases. We suggest the minimum distance method with the Cramér-von Mises distance for estimation of the model parameters.

14:00-14:25 **Tobias Jennessen** (HHU Düsseldorf)

Title: *Method of moments estimators for the extremal index of a stationary time series*

The extremal index θ, a number in the interval [0,1], is known to be a measure of primal importance for analyzing the extremes of a stationary time series. New rank-based estimators for θ are proposed which rely on the construction of approximate samples from the exponential distribution with parameter θ that is then to be fitted via the method of moments.

14:30-14:55 **Katharina Hees** (TU Dortmund)

Title: *Robust estimation of the extremal index of a stationary time series*

A phenomenon observed in many time series is that extreme events occur in clusters. Common examples are extreme weather events like storms, floods or earthquakes. The most common approach to analyze such serially correlated data is to first identify the clusters and then to apply traditional methods from extreme value analysis to the largest peaks in the clusters. The extremal index θ plays an important role in the declustering process. This index can be interpreted asymptotically as the reciprocal of the mean number of exceedances in blocks with at least one exceedance. Another interpretation is that it is the proportion of inter-exceedance times that represent the times between different clusters. Hence, knowledge of the extremal index allows us to decluster the data by sortingthe total n inter-exceedance times by size and assuming the $n\theta$ largest ones to be the inter-cluster times (between clusters) and the others to be the intracluster times (within clusters). In the context for instance of meteorological observation data, we are often confronted with outliers caused by e.g. measurement errors or failures. Several methods for the estimation of the extremal index have been proposed in the literature, but most of them are not robust with respect to outliers. In this talk, we will present a robust estimator of the extremal index and compare it to the existing and well-established methods.