Modeling Clusters of Extreme Values in Time Series

Natalia M. Markovich

Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia nat.markovich@gmail.com, markovic@ipu.rssi.ru

The study of clusters of extreme values of a time series (exceedances over a sufficiently high threshold) is of fundamental interest in many applied fields including climate research, insurance and telecommunications. Usually clustering of exceedances corresponds to consecutive large losses occurring in a short period of time. Our main result states that limit distributions of cluster and inter-cluster sizes for a stationary sequence under specific mixing condition have geometric forms. The cluster size implies the number of consecutive exceedances of the time series over a threshold and the inter-cluster size the number of consecutive observations running under the threshold. In Ferro & Segers (2003) the inter-cluster size normalized by the tail function is proved to be exponentially distributed. A geometric model of the limiting cluster size distribution was presented in Robinson & Tawn (2000) without rigorous proof. The inter-cluster size is also geometric distributed if clusters of exceedances are independent (Santos & Fraga Alves 2012). Recent results of the author will be presented. It is shown that geometric limit distributions of both cluster and inter-cluster sizes can be represented by a level of a sufficiently high quantile of the underlying process used as the threshold and a so-called extremal index. The latter serves as a dependence measure. Its reciprocal has a simple interpretation as the limiting mean cluster size (Leadbetter et al. 1983). The presented models are in a good agreement with cluster and inter-cluster distributions of autoregressive maximum and moving maxima processes and with real telecommunication data concerning packet traffic rates of Skype and Internet television peer-to-peer video applications. The suggested geometric distributions give rise to further modeling of distributions of return intervals (the time intervals between clusters) and durations of clusters widely used in seismology and climatology. These distributions are shown to be bounded by sums of stable and exponentially distributed components. Further development of the author's inferences concerns the score function of the inter-cluster size normalized by the tail function and its relation to the extremal index and the score function of the underlying process. The latter can be applied in numerous fields.

Key Words: Cluster of exceedances over a threshold, extremal index, distributions of cluster and inter-cluster sizes, score function