A Simulation Study on the Performance of Extreme-Value Index Estimators and Proposed Robustifying Modifications

Tsourti Zoi and Panaretos Ioannis

Abstract

This article is concerned with the systematic study of the extreme values that can be observed in every field of our lives. Although the most commonly-used statistical measure of any phenomenon is the 'arithmetic mean', the occurrence of a single extreme can have much more drastic impacts. This alone suggests the importance of studying extremes.

There are more than one approaches to analyze extremes in statistics, the most popular one being the semi-parametric approach. The essence of semi-parametric technique is that the extreme behaviour of any phenomenon under study is described by a single index (the so-called extreme-value index γ) and can be estimated by using only the largest observations. A series of (semi-parametric) estimators for the extreme value index γ have been proposed in the literature (Hill, Pickands and Moment among others). For almost all of these estimators consistency and asymptotic normality have been proven. Still, one of the most serious objections one could raise against these methods is their sensitivity towards the choice of number of upper order statistics used in the estimation (k). The proper choice of k is the 'Achilles' Hill' of the semi-parametric approach.

In the present article we suggest some modifications of the standard estimators, based on smoothing and robustifying procedures, in an effort to lessen the sensitivity of extreme-value index estimators to the choice of k and, accordingly, diminish the problem of choosing k.

The performance of these modified estimators as well as the performance of the standards ones is evaluated via an extensive simulation study. In the simulation study, samples of different sizes and distributions with several different types of tails are taken into consideration.

The results of this study reveal that there is not a uniformly best estimator. On the contrary, indications about the area where the true value of extreme-value index lies can be proven to be really useful to the proper choice of estimator as well as to whether the

application of a smoothing (averaging) procedure is a good idea or not. More particularly, the performance of estimators on a specific data-set depends on the distribution of the data. In cases of long-tailed data (with an infinite upper end-point) Moment and Moment-Ratio estimators seem to estimate more satisfactorily the non-negative extreme-value index γ. However, when it comes to upper-bounded distributions (characterized by a negative value of γ) Peng's and Moment estimators are more preferable. As long as the impact of smoothing (averaging) procedures is concerned, we deduced that it is effective (improving the performance of standard estimators) in cases where the true value of extreme-value index is non-positive. Particularly, mean-averaging procedures improve greatly the performance of Pickands estimator (in case of zero γ), while median-averaging of Moment and Peng's also leads to improved estimators (for γ negative).