Journal of Statistical Planning and Inference 139 (2009) 2800–2815

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Second-order refined peaks-over-threshold modelling for heavy-tailed distributions

Jan Beirlant[a,1], Elisabeth Joossens[b], Johan Segers[c,*,1]

[a]Department of Mathematics and Leuven Statistics Research Centre, Katholieke Universiteit Leuven, Celestijnenlaan 200b, B-3001 Heverlee, Belgium
[b]Joint Research Centre, European Commission, Via Fermi 2749, 21027 Ispra (VA), Italy
[c]Institut de statistique, Université catholique de Louvain, Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, Belgium

ARTICLE INFO

Article history:
Received 18 July 2008
Received in revised form
21 October 2008
Accepted 6 January 2009
Available online 20 January 2009

Keywords:
Bias reduction
Hill estimator
Extended Pareto distribution
Extreme value index
Heavy tails
Regular variation
Tail empirical process
Tail probability
Weissman probability estimator

ABSTRACT

Modelling excesses over a high threshold using the Pareto or generalized Pareto distribution (PD/GPD) is the most popular approach in extreme value statistics. This method typically requires high thresholds in order for the (G)PD to fit well and in such a case applies only to a small upper fraction of the data. The extension of the (G)PD proposed in this paper is able to describe the excess distribution for lower thresholds in case of heavy-tailed distributions. This yields a statistical model that can be fitted to a larger portion of the data. Moreover, estimates of tail parameters display stability for a larger range of thresholds. Our findings are supported by asymptotic results, simulations and a case study.

1. Introduction

It is well known that a distribution is in the max-domain of attraction of an extreme value distribution if and only if the distribution of excesses over high thresholds is asymptotically generalized Pareto (GP) (Balkema and de Haan, 1974; Pickands, 1975). This result gave rise to the peaks-over-threshold methodology introduced in Davison and Smith (1990); see also Coles (2001). The method consists of two components: modelling of clusters of high-threshold exceedances with a Poisson process and modelling of excesses associated to the cluster peaks with a GPD. In practice, a way to verify the validity of the model is to check whether the estimates of the GP shape parameter are stable when the model is fitted to excesses over a range of thresholds. The question then arises how to proceed if this threshold stability is not visible for a given data set. From a theoretical point of view, absence of the stability property can be explained by a slow rate of convergence in the Pickands–Balkema–de Haan theorem. In case of heavy-tailed distributions, the same issue arises when fitting a Pareto distribution (PD) to the relative excesses over high, positive thresholds.

A possible solution is to build a more flexible model capable of capturing the deviation between the true excess distribution and the asymptotic model. For heavy-tailed distributions, this deviation can be parametrized using a power series expansion of

* Corresponding author. Tel.: +32 10 474311.
E-mail addresses: jan.beirlant@wis.kuleuven.be (J. Beirlant), elisabeth.joossens@jrc.it (E. Joossens), johan.segers@uclouvain.be (J. Segers).

[1] Supported by IAP research network Grant no. P6/03 of the Belgian government (Belgian Science Policy).

0378-3758/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2009.01.006

the tail function (Hall, 1982), or more generally via second-order regular variation (Geluk and de Haan, 1987; Bingham et al., 1987).

The aim of this paper is to propose such an extension, called the extended Pareto or extended generalized Pareto distribution (EPD/EGPD). A key distinction with other approaches is that although in previous papers the second-order approximation is used for adjusting the inference of the tail index, inference on the tail itself is still based on the GPD; in contrast, in our approach the EP(G)D is fitted directly to the high-threshold excesses. Indeed, as we will show later, even if the (G)PD parameters are estimated in an unbiased way, tail probability estimators may still exhibit asymptotic bias if based upon the (G)PD approximation.

The main advantages of the new model are a reduction of the bias of estimators of tail parameters and a good fit to excesses over a larger range of thresholds. In an actuarial context, the relevance of using more elaborate models has already been discussed for instance in Frigessi et al. (2002) and Cooray and Ananda (2005).

In case of heavy-tailed distributions, it is more convenient to work with relative excesses X/u rather than absolute excesses $X-u$. Under the domain of attraction condition the limit distribution of X/u given $X>u$ for $u\to\infty$ is the PD. The EPD and EGPD presented here are related through the same affine transformation that links these relative and absolute excesses. Building on the theory of generalized regular variation of second order in de Haan and Stadtmüller (1996), it is also possible to construct an extension of the GPD with comparable merits applicable to distributions in all max-domains of attraction. However, parameter estimation in this more general setting is numerically quite involved (Beirlant et al., 2002b): the model contains one additional parameter and the upper endpoint of the distribution depends in a complicated way on the parameters, which complicates both theory and computations.

Bias-reduction methods have already been proposed in, amongst others, Feuerverger and Hall (1999), Gomes et al. (2000), Beirlant et al. (1999, 2002a) and Gomes and Martins (2002, 2004). These methods focus on the distribution of log-spacings of high order statistics. Moreover, *ad hoc* construction methods for asymptotically unbiased estimators of the extreme value index were introduced in Peng (1998), Drees (1996) and Segers (2005). In contrast, next to providing bias-reduced tail index estimators, our model can be fitted directly to the excesses over a high threshold. The fitted model can then be used to estimate any tail-related risk measure, such as tail probabilities, tail quantiles (or value-at-risk), etc.

In the same spirit as in this paper, a mixture model with two Pareto components was proposed in Peng and Qi (2004). The advantage of our model is that it also incorporates the popular GPD. From our experience, this connection can assist in judging the quality of the GPD fit; see for instance the case study in Example 5.3.

The paper is structured as follows. The next section provides the definition of the E(G)PD, which is shown to yield a more accurate approximation to the distribution of absolute and relative excesses for a wide class of heavy-tailed distributions. Estimators of the EPD parameters are derived in Section 3 using the linearized score equations, and their asymptotic normality is formally stated. In Section 4, we compare the asymptotic distribution and the finite-sample behavior of the estimators of the extreme value index following from PD, GPD and EPD modelling. To illustrate how to apply the methodology to the estimation of general tail-related risk measures, we elaborate in Section 5 on tail probability estimation with theoretical results and a practical case. The appendices, finally, contain the statement and proof of an auxiliary result on a certain tail empirical process followed by the proofs of the main theorems.

2. The extended (generalized) Pareto distribution

Definition 2.1. The *EPD* with parameter vector (γ,δ,τ) in the range $\tau<0<\gamma$ and $\delta>\max(-1,1/\tau)$ is defined by its distribution function

$$G_{\gamma,\delta,\tau}(y)=\begin{cases}1-\{y(1+\delta-\delta y^{\tau})\}^{-1/\gamma} & \text{if } y>1,\\ 0 & \text{if } y\leqslant 1.\end{cases}$$

The *EGPD* is defined by its distribution function

$$H_{\gamma,\delta,\tau}(x)=G_{\gamma,\delta,\tau}(1+x),\quad x\in\mathbb{R}.$$

The ordinary PD with shape parameter $\alpha>0$ is a member of the EPD family: take $\gamma=1/\alpha$ and $\delta=0$ (arbitrary τ). The generalized Pareto distribution (GPD) with positive shape parameter $\gamma>0$ and scale parameter $\sigma>0$ is a member of the EGPD family: take $\tau=-1$ and $\delta=\gamma/\sigma-1$. Finally, the distribution of the random variable Y is EPD(γ,δ,τ) if and only if the distribution of $Y-1$ is EGPD(γ,δ,τ).

We will use the E(G)PD to model tails of heavy-tailed distributions that satisfy a certain second-order condition, to be described next. For a distribution function F, write $\overline{F}=1-F$. Recall that a positive, measurable function f defined in some right neighborhood of infinity is *regularly varying* with index $\beta\in\mathbb{R}$ if $\lim_{u\to\infty}f(ux)/f(u)=x^{\beta}$ for all $x\in(0,\infty)$; notation $f\in\mathscr{R}_{\beta}$. The following definition describes a subset of the class of distribution functions F for which $\overline{F}\in\mathscr{R}_{-1/\gamma}$, $\gamma>0$. Note that the latter is precisely the class of distributions in the max-domain of attraction of the Fréchet distribution with shape parameter $1/\gamma$.