Available online at www.sciencedirect.com

SCIENCE DIRECT®

C. R. Acad. Sci. Paris, Ser. I 341 (2005) 307–312

COMPTES RENDUS MATHEMATIQUE

http://france.elsevier.com/direct/CRASS1/

Statistics

Asymptotic normality of the extreme quantile estimator based on the POT method

Jean Diebolt [a], Armelle Guillou [b], Pierre Ribereau [b]

[a] *CNRS, université de Marne-la-Vallée, équipe d'analyse et de mathématiques appliquées, 5, boulevard Descartes, bâtiment Copernic, Champs-sur-Marne, 77454 Marne-la-Vallée cedex 2, France*

[b] *Université Paris VI, laboratoire de statistique théorique et appliquée, boîte 158, 175, rue du Chevaleret, 75013 Paris, France*

Received 6 February 2005; accepted after revision 29 June 2005

Available online 30 August 2005

Presented by Paul Deheuvels

Abstract

The POT (Peaks-Over-Threshold) approach consists in using the generalized Pareto distribution (GPD) to approximate the distribution of excesses over thresholds. In this Note, we propose extreme quantile estimators based on this method. We establish their asymptotic normality under suitable general assumptions. ***To cite this article: J. Diebolt et al., C. R. Acad. Sci. Paris, Ser. I 341 (2005).***

Résumé

Normalité asymptotique de l'estimateur d'un quantile extrême basé sur la méthode POT. La méthode POT (pics au-delà d'un seuil) consiste à utiliser une distribution de Pareto généralisée (GPD) pour approximer la loi des excès au-delà d'un seuil. Dans cette Note, nous proposons des estimateurs de quantiles extrêmes basés sur cette méthode. Nous établissons leurs normalités asymptotiques sous des hypothèses générales. ***Pour citer cet article : J. Diebolt et al., C. R. Acad. Sci. Paris, Ser. I 341 (2005).***

Version française abrégée

Le principe de la méthode POT est d'estimer la distribution des excès au-delà d'un seuil u par une loi GPD dépendant de deux paramètres (γ, σ) après estimation de ces derniers à partir de la loi des excès au-delà de u. En utilisant cette approche, nous pouvons proposer des estimateurs de quantiles extrêmes, qui dépendent des esti-

E-mail addresses: diebolt@math.univ-mlv.fr (J. Diebolt), guillou@ccr.jussieu.fr (A. Guillou).

1631-073X/$ – see front matter © 2005 Académie des sciences. Published by Elsevier SAS. All rights reserved.
doi:10.1016/j.crma.2005.06.032

mateurs de (γ,σ). Nous établissons, sous des hypothèses convenables, en particulier une condition très générale du second ordre (Condition 2.1 ci-dessous), la normalité asymptotique de ces estimateurs quantiles basé sur un seuil aléatoire. Notre démarche s'appuie sur un formalisme développé par Drees [4]. Nous illustrons notre résultat dans le cas où le couple (γ,σ) est estimé par la méthode du maximum de vraisemblance. Un autre exemple d'application, celui des moments pondérés généralisés, est donné dans Diebolt et al. [3].

1. Introduction

Let $X_1,\dots,X_n$ be a sample of n independent and identically distributed (i.i.d.) random variables from some continuous distribution function F. Now the question is how to obtain with such a limited sample a good estimate for a quantile

$$F^{-1}(1-p)=\inf\bigl\{y:\ F(y)\geqslant 1-p\bigr\},$$

where p is small such that the quantile to be estimated is situated on the border of or beyond the range of the data. Estimating such high quantiles is directly linked to the accurate modelling of the tail of the distribution

$$\overline{F}(x):=\mathbb{P}(X>x)$$

for large thresholds x.

From extreme value theory, the behaviour of such extreme quantiles is known to be governed by one crucial parameter (γ) of the underlying distribution, called the extreme value index (EVI). Indeed, as shown by Gnedenko [6], for some $\gamma\in\mathbb{R}$, there exists sequences of constants $(\alpha_n)\in\mathbb{R}$ and $(\sigma_n)\in\mathbb{R}_+$ such that

$$\mathbb{P}\biggl(\frac{X_{n,n}-\alpha_n}{\sigma_n}\leqslant x\biggr)\xrightarrow{d}H_\gamma(x):=\begin{cases}\exp\bigl(-(1+\gamma x)^{-1/\gamma}\bigr) & \text{for } \gamma\neq 0,\\ \exp\bigl(-\exp(-x)\bigr) & \text{for } \gamma=0.\end{cases}\tag{1}$$

In such a case, we say that F is in the maximum domain of attraction of H_γ, denoted by $F\in MDA(H_\gamma)$.

Since this approach is only based on a set of maxima, another method consists in considering the excesses $Y_1,\dots,Y_{N_n}$ over a high threshold u. Here N_n is the number of excesses and $Y_j=X_{i_j}-u>0$ where the i_j's are the i, $1\leqslant i\leqslant n$, such that $X_i>u$. The limiting distribution of these excesses over the threshold u is then the generalized Pareto distribution (GPD) (see Pickands, [7]). This approach is called the GPD approach.

More precisely, denote by $\tau_F\in(0,\infty]$ the right endpoint of F and define the excess distribution function by

$$F_u(x):=\mathbb{P}(X-u\leqslant x\mid X>u)\quad \text{for } 0<u<\tau_F \text{ and } 0<x<\tau_F-u.$$

Then

$$F\in MDA(H_\gamma)\quad\text{iff}\quad \exists\sigma(u)>0\colon \lim_{u\to\tau_F}\sup_{0<x<\tau_F-u}\bigl|F_u(x)-G_{\gamma,\sigma(u)}(x)\bigr|=0,$$

where

$$G_{\gamma,\sigma}(x):=1-\overline{G}_{\gamma,\sigma}(x)=\begin{cases}1-\Bigl(1+\dfrac{\gamma x}{\sigma}\Bigr)^{-1/\gamma} & \text{for } \gamma\neq 0 \text{ and } 1+\dfrac{\gamma x}{\sigma}>0,\\ 1-\exp\Bigl(-\dfrac{x}{\sigma}\Bigr) & \text{for } \gamma=0,\end{cases}$$

denotes the distribution function of the GPD (γ,σ).

Therefore, for large thresholds u one expects that the excess distribution F_u is well approximated by a GPD with shape parameter γ equal to the EVI of F. Schematically, the POT method then works as follows:

- Given a sample $X_1,\dots,X_n$, select a high threshold u. Let N_u be the number of observations $X_{i_1},\dots,X_{i_{N_u}}$ exceeding u and denote the excesses $Y_j=X_{i_j}-u\geqslant 0$.

- Fit a GPD $G_{\gamma,\sigma}$ to the excesses $Y_1, \ldots, Y_{N_u}$ to obtain estimates $\hat{\gamma}$ and $\hat{\sigma}$ for the shape and scale parameters.
- As $\overline{F}(x) = \overline{F}(u) \cdot \overline{F_u}(x-u)$ for $x > u$, one can estimate the tail of F by

$$\widehat{\overline{F}}(x) = \frac{N_u}{n}\left(1 + \hat{\gamma}\,\frac{x-u}{\hat{\sigma}}\right)^{-1/\hat{\gamma}} \quad \text{for } x > u. \tag{2}$$

- Estimates for quantiles $x_p = F^{-1}(1-p) > u$ can be obtained by inverting (2):

$$\hat{x}_p = u + \hat{\sigma}\,\frac{(N_u/np)^{\hat{\gamma}} - 1}{\hat{\gamma}}.$$

In practice we fix u at the $(k+1)$ largest observation $X_{n-k,n}$. A GPD is then fitted to the k excesses $(X_{n-k+1,n} - X_{n-k,n}), \ldots, (X_{n,n} - X_{n-k,n})$. We denote the resulting parameter estimates by $\hat{\gamma}_k^{\mathrm{POT}}$ and $\hat{\sigma}_k^{\mathrm{POT}}$. The corresponding POT quantile estimator is then

$$\hat{x}_{p,k}^{\mathrm{POT}} := X_{n-k,n} + \hat{\sigma}_k^{\mathrm{POT}}\,\frac{(k/(np))^{\hat{\gamma}_k^{\mathrm{POT}}} - 1}{\hat{\gamma}_k^{\mathrm{POT}}} \quad \text{for } p < \frac{k}{n}. \tag{3}$$

The remainder of this Note is the following. In Section 2, we establish the asymptotic normality of this POT estimator under general assumptions. The example where $(\hat{\gamma}_k^{\mathrm{POT}}, \hat{\sigma}_k^{\mathrm{POT}})$ are the maximum likelihood estimators is given in Section 3.

2. Asymptotic normality for the POT quantile estimator

In order to state our main result, we need some formalism, which is essentially based on Drees' [4] results.

If the estimator $\hat{\gamma}$ is based on the (k_n+1) largest observations, then it can be rewritten as $\hat{\gamma} = T_n(Q_n)$ where

$$Q_{n,k_n}(t) := Q_n(t) := F_n^{-1}\left(1 - \frac{k_n}{n}t\right) = X_{n-[k_n t],n}, \quad t \in [0,1],$$

is the empirical tail quantile function, T_n a smooth functional defined on a suitable space and F_n the classical empirical distribution function of the original sample $X_1, \ldots, X_n$.

In order to ensure consistency of $T_n(Q_n)$, $(k_n)_{n\in\mathbb{N}}$ will be an intermediate sequence, i.e. $k_n \to \infty$ and $k_n/n \to 0$. However, to obtain a non degenerate limiting distribution, one has to impose stronger restrictions for k_n that depend on the second-order behaviour of the underlying distribution function F. The following expansion of the underlying quantile function, extensively studied by de Haan and Stadtmüller [2], is the most general one.

Condition 2.1. *There exist measurable, locally bounded functions $a, \Phi : (0,1) \to (0,\infty)$ and $\Psi : (0,\infty) \to \mathbb{R}$ such that*

$$\frac{F^{-1}(1-tx) - F^{-1}(1-t)}{a(t)} = \frac{x^{-\gamma} - 1}{\gamma} + \Phi(t)\Psi(x) + R(t,x)$$

for all $t \in (0,1)$ and $x > 0$. (By convention, $(x^{-\gamma}-1)/\gamma := -\log x$ if $\gamma = 0$.) Here

(i) *$\Psi \equiv 0$ and $R(t,x) = \mathrm{o}(1)$ as $t \searrow 0$, for all $x > 0$, or*
(ii) *$x \mapsto \gamma\Psi(x)/(x^{-\gamma}-1)$ is not constant, $\Phi(t) = \mathrm{o}(1)$ and $R(t,x) = \mathrm{o}(\Phi(t))$ as $t \searrow 0$, for all $x > 0$.*

Remark 1. Condition 2.1(i) is equivalent to our basic assumption $F \in MDA(H_\gamma)$ given in (1). In Condition 2.1(ii), Φ is regularly varying at zero with order $-\rho$ for $\rho \leqslant 0$, that is

$$\lim_{t\to 0} \frac{\Phi(tx)}{\Phi(t)} = x^{-\rho}$$

for all $x > 0$. Moreover, according to de Haan and Stadtmüller ([2], Remark 3(ii)), one has

$$\Psi(x) = c_\Psi \times \begin{cases} \overline{G}^{-1}_{\gamma+\rho,1}(x) & \text{if } \rho < 0, \\ -x^{-\gamma}\log x/\gamma & \text{if } \gamma \neq -\rho = 0, \\ \log^2(x) & \text{if } \gamma = -\rho = 0, \end{cases} \tag{4}$$

for some constant $c_\Psi \neq 0$ if the normalizing function a and the function Φ are suitably chosen. This will be assumed in all the sequel and imply a weighted uniform convergence result for the remainder term $R(t, x)$ in a neighbourhood of 0 (see Lemma 2.1 in Drees, [4]).

Now, in order to give our main result, we need some classical assumptions.

Suppose that F is three times differentiable. We denote

$$f(t) := \mathrm{e}^{t(1-\gamma)}U'(\mathrm{e}^t),$$

where U is the tail quantile function defined as $U(t) = F^{-1}(1 - \frac{1}{t})$.

Let V and M be two functions defined as

$$V(t) = \overline{F}^{-1}(\mathrm{e}^{-t}) \quad \text{and} \quad M(t) = \frac{V''(\ln t)}{V'(\ln t)} - \gamma.$$

We assume the first and second-order following conditions

$$\lim_{t\to+\infty} M(t) = 0, \tag{5}$$

M has a constant sign at infinity

and there exists $\rho \leqslant 0$ such that M is regularly varying at infinity with order ρ. (6)

We also suppose that

$$k = k_n \to \infty,\ \frac{k}{n} \to 0,\ \frac{k}{np} \nearrow \infty \text{ such that } \frac{\log(k/(np))}{\sqrt{k}} \to 0, \tag{7}$$

and

$$\lim_{n\to\infty}\left(\log\frac{k}{np}\right)\sup_{v\geqslant\log(n/k)}\left|\frac{f''(v)}{f'(v)} - \beta\right| = 0, \quad \text{where } \beta \leqslant 0. \tag{8}$$

Remark 2. The quotient f''/f' is defined to be zero when f' is zero. For many usual distributions, such as the Uniform, Weibull, Fréchet or standard Normal distributions, the condition (8) is satisfied with $\beta = 0$. However, in the case of the Cauchy distribution, $\beta = -2$.

We are now able to state our main result.

Theorem 2.2. *Suppose that F is three times differentiable and that assumptions* (5)–(8) *and those ensuring the asymptotic normality of*

$$\sqrt{k}\begin{pmatrix} \hat{\gamma}_k^{\mathrm{POT}} - \gamma \\ \dfrac{\hat{\sigma}_k^{\mathrm{POT}}}{a(k/n)} - 1 \\ \dfrac{X_{n-k,n} - F^{-1}(1-k/n)}{a(k/n)} \end{pmatrix} = \begin{pmatrix} G_k^{(1)} \\ G_k^{(2)} \\ G_k^{(3)} \end{pmatrix} \longrightarrow^d N(B, \Gamma) \tag{9}$$

are satisfied. Then, under the additional assumption

$$\sqrt{k}\Phi\left(\frac{k}{n}\right) \longrightarrow \lambda \in [0,\infty), \tag{10}$$

we have

- *under Condition* 2.1(ii), *if* $\gamma < 0$:
$$\frac{\sqrt{k}(\hat{x}_{p,k}^{\mathrm{POT}} - x_p)}{\hat{\sigma}_k^{\mathrm{POT}} \int_1^{k/np} s^{\hat{\gamma}_k^{\mathrm{POT}}-1} \log s \, \mathrm{d}s} = G_k^{(1)} - \gamma G_k^{(2)} + \gamma^2 G_k^{(3)} + c_\Psi \frac{\gamma^2}{\gamma+\rho} \lambda \mathbb{1}_{\{\rho<0\}} + \mathrm{o}_{\mathbb{P}}(1), \tag{11}$$
- *under Condition* 2.1(ii), *if* $\gamma \geqslant 0$, *by choosing*
$$\Phi(t) = \frac{M(1/t)}{\widetilde{c_\Psi}} \cdot \begin{cases} \frac{1}{\rho} & \text{if } \rho < 0, \\ 1 & \text{if } \gamma \neq -\rho = 0, \\ \frac{1}{2} & \text{if } \gamma = -\rho = 0, \end{cases} \tag{12}$$
$$\frac{\sqrt{k}(\hat{x}_{p,k}^{\mathrm{POT}} - x_p)}{\hat{\sigma}_k^{\mathrm{POT}} \int_1^{k/np} s^{\hat{\gamma}_k^{\mathrm{POT}}-1} \log s \, \mathrm{d}s} = G_k^{(1)} - \widetilde{c_\Psi} \lambda \mathbb{1}_{\{\gamma \geqslant 0, \beta = 0\}} \cdot \begin{cases} \rho & \text{if } \rho < 0, \\ 1 & \text{if } \gamma \neq -\rho = 0 \\ 2 & \text{if } \gamma = -\rho = 0, \end{cases} + \mathrm{o}_{\mathbb{P}}(1), \tag{13}$$
where $\widetilde{c_\Psi} = \pm c_\Psi$ *in order to ensure that* Φ *takes its values in* $(0,\infty)$.

In these two cases, the limiting distribution is a normal distribution whose variance depends on the components of Γ.

Proof of Theorem 2.2. We rewrite the difference $\hat{x}_{p,k}^{\mathrm{POT}} - x_p$ as follows:

$$\hat{x}_{p,k}^{\mathrm{POT}} - x_p = \left(\left\{\frac{(k/(np))^{\hat{\gamma}_k^{\mathrm{POT}}} - 1}{\hat{\gamma}_k^{\mathrm{POT}}} - \frac{(k/(np))^{\gamma} - 1}{\gamma}\right\}\hat{\sigma}_k^{\mathrm{POT}}\right) + \left(\frac{(k/(np))^{\gamma} - 1}{\gamma}\left\{\hat{\sigma}_k^{\mathrm{POT}} - a\left(\frac{k}{n}\right)\right\}\right)$$
$$+ \left(a\left(\frac{k}{n}\right)\frac{G_k^{(3)}}{\sqrt{k}}\right) + \left(U\left(\frac{n}{k}\right) - U\left(\frac{1}{p}\right) + a\left(\frac{k}{n}\right)\frac{(k/(np))^{\gamma} - 1}{\gamma}\right).$$

The three first terms in brackets can be treated as in de Haan and Rootzén [1]. For the last term, if $\gamma < 0$ we have to use Lemma 2.1 in Drees [4], while if $\gamma \geqslant 0$, we can prove, by a Taylor expansion, that Condition 2.1(ii) is satisfied for $a(t) := \frac{1}{t}U'(\frac{1}{t})[1 + cM(\frac{1}{t})]$ with c a suitable constant and $\Phi(\cdot)$ satisfying (12). Then, following again the lines of proof of de Haan and Rootzén [1], Theorem 2.2 follows. See Diebolt et al. [3] for more details. □

3. Example

Our key point is Corollary 2.1 in Drees [4]. According to this result, if we define

$$D_n := \begin{cases} F^{-1}\left(1 - \frac{k_n}{n}\right) - a\left(\frac{k_n}{n}\right)\left\{\frac{\mathbb{1}_{\{\gamma\neq 0\}}}{\gamma} + \frac{c_\Psi}{\gamma+\rho}\Phi\left(\frac{k_n}{n}\right) \cdot \mathbb{1}_{\{\gamma \neq -\rho > 0\}}\right\} & \text{if } \gamma \geqslant -\frac{1}{2}, \\ X_{n,n} & \text{if } \gamma < -\frac{1}{2}, \end{cases}$$

then, under Condition 2.1(ii) with a, Φ and Ψ satisfying Lemma 2.1 in Drees [4] and the additional assumption (10), we have

$$\sqrt{k_n}\left(\frac{Q_n - D_n}{a(k_n/n)} - \overline{G}_{\gamma,1}^{-1} - \frac{1}{\gamma}\mathbb{1}_{\{\gamma\neq 0\}}\right) \longrightarrow \mathbb{V}_\gamma + \lambda\left(\Psi + \frac{c_\Psi}{\gamma+\rho}\mathbb{1}_{\{\gamma\neq -\rho>0\}}\right) \tag{14}$$

weakly in a suitable space, where $\mathbb{V}_\gamma := t^{-(\gamma+1)}\mathbb{W}(t)$, $t \in [0,1]$, with $\mathbb{W}$ a standard Brownian motion.

Now, we use Theorem 2.1 in Drees et al. [5] which gives the asymptotic normality of the maximum likelihood estimators under Condition 2.1(ii) for $\gamma > -1/2$ and the additional assumption (10). Combining these results with (14), we deduce that the convergence (9) holds. Hence, under the assumptions of our Theorem 2.2, for $\gamma > -1/2$, (11) and (13) are satisfied and the limiting distribution is a $N(\lambda C, \sigma^2)$ distribution, where

$$\sigma^2 = (1+\gamma)^2 \mathbb{1}_{\{\gamma \geqslant 0\}} + (1 + 4\gamma + 5\gamma^2 + 2\gamma^3 + 4\gamma^4)\mathbb{1}_{\{\gamma<0\}},$$

and

$$\begin{aligned} C = {} & \left[c_\Psi \frac{\rho(\gamma+1)}{(1-\rho)(1+\gamma-\rho)} - \widetilde{c_\Psi}\rho \mathbb{1}_{\{\beta=0\}}\right]\mathbb{1}_{\{\gamma\geqslant 0,\rho<0\}} + [c_\Psi - \widetilde{c_\Psi}\mathbb{1}_{\{\beta=0\}}]\mathbb{1}_{\{\gamma>0,\rho=0\}} \\ & + [2c_\Psi - 2\widetilde{c_\Psi}\mathbb{1}_{\{\beta=0\}}]\mathbb{1}_{\{\gamma=0,\rho=0\}} + \left[c_\Psi \rho^2 \frac{1+3\gamma+2\gamma^2}{(1-\rho)(\gamma+\rho)(1+\gamma-\rho)}\right]\mathbb{1}_{\{\gamma<0,\rho<0\}}. \end{aligned}$$

Another example of application has been proposed in Diebolt et al. [3]. It concerns the case where $(\hat{\gamma}_k^{\mathrm{POT}}, \hat{\sigma}_k^{\mathrm{POT}})$ are the generalized probability-weighted moments estimators. Also a simulation study is provided in order to compare these two POT quantile estimators to classical estimators: the Weissman and the moment ones.

References

[1] L. de Haan, H. Rootzén, On the estimation of high quantiles, J. Statist. Plann. Inference 35 (1993) 1–13.
[2] L. de Haan, U. Stadtmüller, Generalized regular variation of second order, J. Austral. Math. Soc. Ser. A 61 (1996) 381–395.
[3] J. Diebolt, A. Guillou, P. Ribereau, Asymptotic normality of extreme quantile estimators based on the peaks-over-threshold approach, submitted for publication, 2005.
[4] H. Drees, On smooth statistical tail functionals, Scand. J. Statist. 25 (1) (1998) 187–210.
[5] H. Drees, A. Ferreira, L. de Haan, On maximum likelihood estimation of the extreme value index, Ann. Appl. Probab. 14 (3) (2004) 1179–1201.
[6] B.V. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, Ann. Math. 44 (1943) 423–453.
[7] J. Pickands III, Statistical inference using extreme order statistics, Ann. Statist. 3 (1975) 119–131.