BIAS CORRECTED GEOMETRIC-TYPE ESTIMATORS

MARGARIDA BRITO, LAURA CAVALCANTE,
AND ANA CRISTINA MOREIRA FREITAS

Abstract. The estimation of the tail index is a central topic in the extreme value analysis. We consider a geometric-type estimator for the tail index and study its asymptotic properties. We propose here two asymptotic equivalent bias corrected geometric-type estimators and establish the corresponding asymptotic behaviour. We also apply the suggested estimators to construct asymptotic confidence intervals for this tail parameter. Some simulations in order to illustrate the finite sample behaviour of the proposed estimators are provided.

1. Introduction

We consider the problem of estimating the Pareto-tail index of a distribution function F, with tail function $\overline{F} = 1 - F \in RV_{-1/\gamma}$, $\gamma > 0$, where RV_α denotes the class of regularly varying functions with index of regular variation equal to α, that is,

$$\lim_{t\to\infty} \frac{\overline{F}\,(tx)}{\overline{F}\,(t)} = x^{-1/\gamma}, \quad \text{for all } x > 0.$$

Equivalently,

$$1 - F\,(x) = x^{-1/\gamma} l\,(x) \quad \text{for } x > 0, \tag{1}$$

where l is a slowly varying function at infinity, that is, l satisfies the condition $l\,(tx)\,/l\,(t) \to 1$ as $t \to \infty$ for all $x > 0$. Denoting by F^{-1} the left continuous inverse of F, $F^{-1}\,(s) = inf\,\{y : F\,(y) \geq s\}$, the condition

2010 *Mathematics Subject Classification.* Primary 62G32; Secondary 62G30, 62E20, 62F12.

Key words and phrases. Bias reduction; extreme value theory; generalised Pareto distribution; heavy tails; tail index estimation.

ACMF is partially supported by FCT grant SFRH/BPD/66174/2009 and LC is supported by FCT grant SFRH/DB/60642/2009. All three authors are supported by FCT project PTDC/MAT/120346/2010. Research funded by the European Regional Development Fund through the programme COMPETE and by the Portuguese Government through the FCT - Fundação para a Ciência e a Tecnologia under the project PEst - C/MAT/UI0144/2011.

(1) is equivalent to the regular variation of the tail function $U(x) = F^{-1}(1-1/x)$, i.e.,

$$U(x) = x^{\gamma} L(x), \tag{2}$$

where L is a slowly varying function at infinity. In this way, the question addressed is the estimation of γ from a finite sample $X_1, \ldots, X_n$.

Let us consider $X_1, X_2, \ldots$ independent and identically distributed (i.i.d.) random variables (r.v.) with distribution function (d.f.) F and let $X_{(1,n)} \leq X_{(2,n)} \leq \cdots \leq X_{(n,n)}$ denote the corresponding order statistics (o.s.) based on the n first observations. We also consider intermediate sequences $k = k_n$ of positive integers ($1 \leq k < n$), that is

$$k \to \infty, \qquad \frac{k}{n} \to 0 \quad \text{as} \quad n \to \infty. \tag{3}$$

One of the most well known estimators for the tail index has been suggested by Hill (1975) and is given by

$$\widehat{H}(k) = \frac{1}{k} \sum_{i=1}^{k} \log X_{(n-i+1,n)} - \log X_{(n-k,n)}.$$

The asymptotic properties of the Hill estimator have been much studied and it is well known that, under certain conditions, $\widehat{H}(k)$ is a strongly consistent estimator (see e.g. Deheuvels et al. (1988)) with asymptotic normal distribution (see e.g. Haeusler and Teugels (1985)).

We recall that the above problem of tail index estimation of a Pareto type distribution is equivalent to the estimation of the exponential tail coefficient. Setting $Z_i = \log X_i$, $i = 1, 2, \ldots$, with X_i as above, we have

$$1 - G(z) = P(Z_1 > z) = r(z)\, e^{-Rz}, \quad z > 0, \tag{4}$$

where $r(z) = l(e^z)$ is a regularly varying function at infinity and $R = 1/\gamma$ is a positive constant, called exponential tail coefficient. Equivalently we have

$$G^{-1}(1-s) = -\frac{1}{R} \log s + \log \widetilde{L}(s), \quad 0 < s < 1,$$

where $\widetilde{L}(s) = L(1/s)$ is a slowly varying function at 0.

The problem of the estimation of the exponential tail coefficient has applications in a variety of domains and an overview of the existing literature is given in Schultze and Steinebach (1996).

We focus this work in the problem of estimating the tail index using a geometric-type estimator of the exponential tail coefficient R, proposed by Brito and Freitas (2003), given by

$$\widehat{R}\,(k) = \sqrt{\frac{\sum_{i=1}^{k}\log^2(n/i) - \frac{1}{k}\left(\sum_{i=1}^{k}\log(n/i)\right)^2}{\sum_{i=1}^{k} Z^2_{(n-i+1,n)} - \frac{1}{k}\left(\sum_{i=1}^{k} Z_{(n-i+1,n)}\right)^2}}. \tag{5}$$

This estimator arises from the study of two estimators based on the least squares method introduced by Schultze and Steinebach (1996). One of these estimators was also introduced by Kratz and Resnick (1996) in an independent but equivalent way. In general, when compared with other tail index estimators, it is reported that the estimators proposed by Schultze and Steinebach have a very good behaviour, performing better in several circumstances.

One of the interesting characteristics of the least squares estimators is the smoothness of the realizations as a function of k. It should be noted that the high variability that some tail estimators present is not a welcome feature, since it makes more difficult the proper selection of the number of upper o.s. involved in the estimation. In this sense, the stability presented in almost all examples can be considered a prominent advantage of the least squares estimators over the classical Hill estimator, which plots often exhibit strong trends and a considerable lack of smoothness resulting in different estimates for neighbouring values of k and an extreme sensibility to the choice of the ideal k-value (see e.g. Csörgő and Viharos (1998)). On the other hand, it can be shown that the asymptotic variance of the geometric-type estimator is twice the asymptotic variance of the Hill estimator. However, given the bias presented by the Hill estimator, the asymptotic variance should not be the only criterion to be considered.

The estimators provided by Schultze and Steinebach were motivated by the fact that $-\log(1-G(z))$, from (4), is approximately linear with slope R, for large z, since $z^{-1}\log r(z) \to 0$ as $z \to \infty$. It is then expected that $-\log(1-G_n(z))$ is also approximately linear for high values of n and z, where G_n denotes the empirical d.f. associated to the random sample $Z_1, \ldots, Z_n$. It was also assumed that $r(z) \equiv c$, $\forall z > 0$, and thus

$$y := -\log(1-G(z)) = Rz - \log c = Rz - d,$$

or equivalentely,

$$z = R^{-1}(y+d) = ay + b,$$

where $a = R^{-1}$, $b = R^{-1}d$ and $d = \log c$.

Denoting by $z_i := z_{(n-i+1,n)}$, $i = 1, \ldots, k \leq n$, the k upper o.s. of the sample $Z_1, \ldots, Z_n$, Schultze and Steinebach approximate $-\log(1-G(z_i))$ by $y_i := -\log(1 - G_n(z_i^-)) = -\log(1-(n-i)/n) = \log(n/i)$, obtaining that y_i is "close" to $Rz_i - d$, or equivalently, z_i is "close" to $ay_i + b$. Following this approach, one of the estimators was obtained by minimizing the function $f_1(a,b) = \sum_{i=1}^{k}(z_i - ay_i - b)^2$ and the other one by minimizing the function $f_2(R,d) = \sum_{i=1}^{k}(y_i - Rz_i + d)^2$, which corresponds to determining the inverse of the slope of the line by minimizing the sum of the distances between the points $\{(z_i, y_i), i = 1, \ldots, k\}$ and the line, measured in horizontal or vertical, respectively.

The $\widehat{R}$ estimator is obtained through a geometrical adaptation of these two perspectives, minimizing the sum of the areas of the rectangles whose sides are the horizontal and vertical segments between the points $\{(z_i, y_i), i = 1, \ldots, k\}$ and the line, in Figure 1, which is equivalent to minimize the function $f(R,d) = \sum_{i=1}^{k}(y_i - Rz_i + d)(R^{-1}y_i + R^{-1}d - z_i)$. In this way both horizontal and vertical distances between the points $\{(z_i, y_i), i = 1, \ldots, k\}$ and the line are minimized.

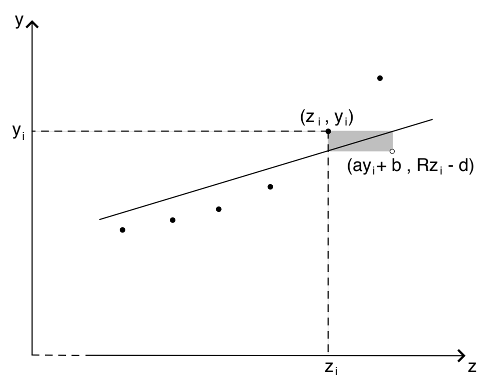

Figure 1. Geometric representation of the rectangles whose areas will be minimized to obtain the estimator of R.

The asymptotic properties of $\widehat{R}(k)$ were investigated in Brito and Freitas (2003). In particular, these authors established the consistency of the estimator and proved that, under general regularity conditions, the distribution of $k^{1/2}\left(\widehat{R}(k) - R\right)$ is asymptotically normal. This estimator also enjoys of certain properties that makes its use specially attractive for the case where R is expected to be small (see e.g. Csörgő and Viharos (1998) and Brito and Freitas (2006)).

In the context of estimating the tail index, we will consider the following geometric-type estimator for γ:

$$\widehat{GT}(k) = \frac{1}{\widehat{R}(k)} = \sqrt{\frac{M_n^{(2)} - \left[M_n^{(1)}\right]^2}{i_n(k)}}\ . \tag{6}$$

where

$$i_n(k) = \frac{1}{k}\sum_{i=1}^{k}\log^2(n/i) - \left(\frac{1}{k}\sum_{i=1}^{k}\log(n/i)\right)^2 \tag{7}$$

and

$$M_n^{(j)}(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\log X_{(n-i+1,n)} - \log X_{(n-k,n)}\right)^j . \tag{8}$$

The asymptotic properties of $\widehat{GT}(k)$ arise naturally from the corresponding properties of $\widehat{R}(k)$ studied in Brito e Freitas (2003).

To deal with the suggested problems, the procedures are formulated under second order conditions. We begin by assuming there exists a positive function a such that, for all $x > 0$,

$$\lim_{t\to\infty}\frac{U(tx) - U(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}. \tag{9}$$

From (2), we can choose $a(t) = \gamma U(t)$. We also suppose that there exists a function $A(t)$, tending to zero as $t \to \infty$, such that

$$\lim_{t\to\infty}\frac{\frac{U(tx)}{U(t)} - x^\gamma}{A(t)} = x^\gamma\frac{x^\rho - 1}{\rho}, \tag{10}$$

for all $x > 0$, where $\rho < 0$ is the shape parameter governing the rate of convergence of $U(tx) - U(t)$ and the function $|A(t)| \in RV_\rho$ (see e.g. Geluk and de Haan (1987)).

In order to obtain information about the upper tail of F, most of the estimators are constructed as functions of the upper k o.s. of a sample of size n (see e.g. Pickands (1975), Dekkers et al. (1989)). When the number of upper o.s. used in the estimation of γ increases, the bias in the estimation becomes larger. This considerable bias that appears in several estimators reveals a difficult problem to go beyond the applications and there are several papers trying to deal with. Once this is such an important research theme, the bias reduction has become popular and received considerable attention in extreme value statistics. Some estimators were built in order to deal with the bias term in an appropriate way (see for example, Peng (1998), Beirlant et al. (1999), Feuerverger and Hall (1999), Gomes et al. (2000), Gomes and Pestana (2007) and Beirlant et al. (2008)). One of the procedures commonly used to deal with this problem was formulated under second order properties of the d.f. and gave rise to the second order reduced-bias estimators.

In Section 2 we study some properties of the geometric-type estimator and, in Section 3, we propose two asymptotic equivalent bias corrected estimators and study their asymptotic behaviour. We use the presented estimators to obtain asymptotic confidence intervals. Proofs are presented in Section 4. A simulation study is provided in Section 5, in order to illustrate and compare the finite sample behaviour of the presented tail index estimators, including the geometric-type and Hill bias corrected estimators.

2. Asymptotic properties of the geometric-type estimator

Here the asymptotic normality of the geometric-type estimator is shown using a method that proves to be very useful for statistical inference. We first derive the asymptotic distributional representation of the geometric-type estimator.

Since $i_n(k) \to 1$ as $n \to \infty$, we begin by considering the asymptotic normality of the following tail index estimator of γ

$$\widetilde{\gamma}\,(k) = \sqrt{M_n^{(2)} - \left[M_n^{(1)}\right]^2}.$$

In the sequel, $\xrightarrow{D}$ and $\overset{D}{=}$ stand, respectively, for convergence and equality in distribution.

Theorem 2.1. *Assume* (9) *holds. For sequences k such that* (3) *holds, we have the following asymptotic distributional representation*

$$\widetilde{\gamma}\,(k) \overset{D}{=} \gamma + \frac{\gamma}{2\sqrt{k}} Q_n - \frac{\gamma}{\sqrt{k}} P_n + \frac{A\left(\frac{n}{k}\right)}{(1-\rho)^2} + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{1}{k}\right),$$

where $P_n = \sqrt{k}\left(\sum_{i=1}^{k} Z_i/k - 1\right)$ *and* $Q_n = \sqrt{k}\left(\sum_{i=1}^{k} Z_i^2/k - 2\right)$, (P_n, Q_n) *is asymptotically normal with mean equal to* $\binom{0}{0}$ *and covariance matrix* $\left(\begin{smallmatrix} 1 & 4 \\ 4 & 20 \end{smallmatrix}\right)$, *and* $\{Z_i\}$ *denote i.i.d. standard exponential r.v..*

Corollary 2.2. *Assume the conditions of Theorem 2.1 hold. If k is such that $\sqrt{k}A(n/k) \to \lambda$ finite, then*

$$\sqrt{k}\,(\widetilde{\gamma}\,(k) - \gamma) \xrightarrow[n\to\infty]{D} N\left(\frac{\lambda}{(1-\rho)^2}, 2\gamma^2\right).$$

Theorem 2.3. *Assume* (9) *holds. For sequences k such that* (3) *holds, we have the following asymptotic distributional representation*

$$\widehat{GT}\,(k) \overset{D}{=} \gamma + \frac{\gamma}{2\sqrt{k}} Q_n - \frac{\gamma}{\sqrt{k}} P_n + \frac{A\left(\frac{n}{k}\right)}{(1-\rho)^2} + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{\log^2 k}{k}\right),$$

where $P_n = \sqrt{k}\left(\sum_{i=1}^{k} Z_i/k - 1\right)$ *and* $Q_n = \sqrt{k}\left(\sum_{i=1}^{k} Z_i^2/k - 2\right)$, (P_n, Q_n) *is asymptotically normal with mean equal to* $\binom{0}{0}$ *and covariance matrix* $\left(\begin{smallmatrix} 1 & 4 \\ 4 & 20 \end{smallmatrix}\right)$, *and* $\{Z_i\}$ *denote i.i.d. standard exponential r.v..*

Corollary 2.4. *Assume the conditions of Theorem 2.3 hold. If* k *is such that* $\sqrt{k}A(n/k) \to \lambda$ *finite, then*

$$\sqrt{k}\left(\widehat{GT}(k) - \gamma\right) \xrightarrow[n\to\infty]{D} N\left(\frac{\lambda}{(1-\rho)^2}, 2\gamma^2\right).$$

3. Bias corrected geometric-type estimation

In this section we improve the geometric-type estimator in the sense of reducing its bias. For this we propose two asymptotic equivalent bias corrected estimators for the tail index, and study the corresponding asymptotic behaviour.

It is convenient to assume that the underlying models belong to Hall's class (Hall (1982)), given by:

$$U(t) = Ct^{\gamma}\left(1 + \frac{A(t)}{\rho}(1 + o(1))\right), \quad \text{as } t \to \infty,$$

where

$$A(t) = \gamma\beta t^{\rho}, \tag{11}$$

with $\gamma > 0$ and $C > 0$, and $\rho < 0$ and $\beta \neq 0$ are, respectively, the shape and scale parameters. This is a very important family with several applications.

3.1. Bias corrected geometric-type estimators.

In order to achieve the improvement of the geometric-type estimator behaviour presented in (6), and following some suggestions in the literature, we derive corrected geometric-type estimators by removing its bias dominant component (see e.g. Caeiro et al. (2005)).

For this we use the asymptotic representation of the geometric-type estimator presented in Theorem 2.3,

$$\widehat{GT}(k) \stackrel{D}{=} \gamma + \frac{\gamma}{2\sqrt{k}}Q_n - \frac{\gamma}{\sqrt{k}}P_n + \frac{A\left(\frac{n}{k}\right)}{(1-\rho)^2} + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{\log^2 k}{k}\right),$$

where the bias dominant component can be written as

$$\frac{A\left(\frac{n}{k}\right)}{(1-\rho)^2} = \frac{\gamma\beta\left(\frac{n}{k}\right)^{\rho}}{(1-\rho)^2}.$$

Thus, removing the bias dominant component directly, we obtain a bias corrected estimator of $\widehat{GT}(k)$ given by

$$\overline{\widehat{GT}}(k) = \widehat{GT}(k)\left(1 - \frac{\beta\left(\frac{n}{k}\right)^{\rho}}{(1-\rho)^2}\right). \tag{12}$$

Considering now the exponential expansion $e^{-x} = 1 - x + o(x)$ as $x \to 0$, we may get the asymptotically equivalent bias corrected estimator

$$\overline{\overline{\widehat{GT}}}(k) = \widehat{GT}(k)\, exp\left\{-\frac{\beta}{(1-\rho)^2}\left(\frac{n}{k}\right)^{\rho}\right\}. \tag{13}$$

We can easily note that the bias dominant component is dependent of the shape ρ and scale β second order parameters. Thus, another challenge of utmost importance to consider is the proper and adequate estimation of the second order parameters, ρ and β, in order to remove the bias dominant component and obtain bias corrected estimators.

We remark that the geometric-type estimator has a lower bias dominant component than the Hill estimator when evaluated at the same threshold, i.e. for the same k.

Estimation of the second order parameters.

Here, we consider the class of estimators of the parameter ρ (depending on τ) proposed by Fraga Alves et al. (2003)

$$\widehat{\rho}_n^{(\tau)}(k) = -\left|\frac{3\left(T_n^{(\tau)}(k) - 1\right)}{T_n^{(\tau)}(k) - 3}\right|, \tag{14}$$

where

$$T_n^{(\tau)}(k) = \begin{cases} \dfrac{\left(M_n^{(1)}(k)\right)^{\tau} - \left(M_n^{(2)}(k)/2\right)^{\tau/2}}{\left(M_n^{(2)}(k)/2\right)^{\tau/2} - \left(M_n^{(3)}(k)/6\right)^{\tau/3}}, & \text{if } \tau > 0 \\[2ex] \dfrac{\log\left(M_n^{(1)}(k)\right) - \frac{1}{2}\log\left(M_n^{(2)}(k)/2\right)}{\frac{1}{2}\log\left(M_n^{(2)}(k)/2\right) - \frac{1}{3}\log\left(M_n^{(3)}(k)/6\right)}, & \text{if } \tau = 0, \end{cases}$$

with M_n^j as in (8), and the β estimator obtained in Gomes and Martins (2002)

(15)
$$\widehat{\beta}_{\widehat{\rho}}(k) = \left(\frac{k}{n}\right)^{\widehat{\rho}} \frac{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\right)\frac{1}{k}\sum_{i=1}^{k}U_i - \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}U_i}{\left(\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}\right)\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}}U_i - \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-2\widehat{\rho}}U_i},$$

where

$$U_i = i\left(\log\frac{X_{(n-i+1,n)}}{X_{(n-i,n)}}\right),$$

with $1 \le i \le k < n$.

We remark that the class of estimators of ρ presented above, and consequently also the β estimators, is dependent on a tuning parameter $\tau \ge 0$. In the literature it has been suggested the use of the tuning parameter $\tau = 0$ when $\rho \in [-1, 0)$ and $\tau = 1$ when $\rho \in (-\infty, -1)$. This parameter must be chosen appropriately in order to provide a higher stability for the estimator of ρ and as such, a graphical study supporting this choice must always be seen as a relevant tool.

Choice of the k_h level to be used in the second order parameters estimation.

It is known that the external estimation of ρ and β at a larger k value than the one used for γ-estimation has clear advantages, allowing the bias reduction without increasing the asymptotic variance (see e.g. Caeiro et al. (2005)).

Through some simulation studies presented in the next chapter we can notice that the estimator of ρ only stabilizes at high levels of k, which justifies the suggestion given in some works that ρ must be estimated at a high level k_h (see e.g. Caeiro and Gomes (2008) and Gomes et al. (2004)). Moreover, the number k_h of the top observations to be considered for the estimation of ρ and β should be such as to ensure that $\widehat{\rho} - \rho = o_p(1/\log n)$.

In the lines of other studies, and among some suggestions (see e.g. Gomes et al. (2007)), the level that seemed to be the most appropriate to consider in illustrations is

(16)
$$k_h = \left\lfloor n^{1-\epsilon} \right\rfloor, \text{ for some } \epsilon > 0 \text{ small},$$

where $\lfloor x \rfloor$ denotes the integer part of x.

3.2. **Asymptotic properties of geometric-type bias corrected estimators.**

We begin by assuming that only the tail index parameter γ is unknown and that $\widehat{GT}^{*}(k)$ is one of the estimators $\overline{\widehat{GT}}(k)$ or $\overline{\overline{\widehat{GT}}}(k)$.

Theorem 3.1. *Assume* (10) *holds. For sequences k such that* (3) *holds, and $A(t)$ as in* (11)*, we have the following asymptotic distributional representation*

$$\widehat{GT}^{*}(k) \overset{D}{=} \gamma + \frac{\gamma}{2\sqrt{k}} Q_n - \frac{\gamma}{\sqrt{k}} P_n + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{log^2 k}{k}\right),$$

where $P_n = \sqrt{k}\left(\sum_{i=1}^{k} Z_i/k - 1\right)$ and $Q_n = \sqrt{k}\left(\sum_{i=1}^{k} Z_i^2/k - 2\right)$, (P_n, Q_n) is asymptotically normal with mean equal to $\binom{0}{0}$ and covariance matrix $\left(\begin{smallmatrix}1 & 4\\ 4 & 20\end{smallmatrix}\right)$, and $\{Z_i\}$ denote i.i.d. standard exponential r.v..

Corollary 3.2. *Assume the conditions of Theorem 3.1 hold. If we choose k such that $\sqrt{k}A(n/k) \to \lambda$ finite, then*

$$\sqrt{k}\left(\widehat{GT}^{*}(k) - \gamma\right) \xrightarrow[n\to\infty]{D} N\left(0, 2\gamma^2\right).$$

Assuming now that $\widehat{GT}^{**}(k)$ denotes the version of $\widehat{GT}^{*}(k)$ where the parameters ρ and β are estimated externally, we have the following result

Theorem 3.3. *Under the conditions of Theorem 3.1 and assuming consistent estimators for ρ and β computed at a level that implies $\widehat{\rho} - \rho = o_p(1/\log n)$, we have the following asymptotic distributional representation*

$$\widehat{GT}^{**}(k) \overset{D}{=} \gamma + \frac{\gamma}{2\sqrt{k}} Q_n - \frac{\gamma}{\sqrt{k}} P_n + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{log^2 k}{k}\right),$$

where $P_n = \sqrt{k}\left(\sum_{i=1}^{k} Z_i/k - 1\right)$ and $Q_n = \sqrt{k}\left(\sum_{i=1}^{k} Z_i^2/k - 2\right)$, (P_n, Q_n) is asymptotically normal with mean equal to $\binom{0}{0}$ and covariance matrix $\left(\begin{smallmatrix}1 & 4\\ 4 & 20\end{smallmatrix}\right)$, and $\{Z_i\}$ denote i.i.d. standard exponential r.v..

Corollary 3.4. *Assume the conditions of Theorem 3.3 hold. If we choose k such that $\sqrt{k}A(n/k) \to \lambda$ finite, then*

$$\sqrt{k}\left(\widehat{GT}^{**}(k) - \gamma\right) \xrightarrow[n\to\infty]{D} N\left(0, 2\gamma^2\right).$$

4. Proofs

For the proof of Theorem 2.1 we need the following Lemma.

Lemma 4.1 (Dekkers and de Haan (1993), Lemma 3.1)**.** *Let* $Y_{(1,n)} \leq Y_{(2,n)} \leq \cdots \leq Y_{(n,n)}$ *denote the o.s. based on the* n *first observations of the sequence* $Y_1, \ldots, Y_n$ *of i.i.d. r.v. with common d.f.* $1 - 1/x$ $(x > 1)$*. Let* k *be such that* (3) *holds. For* $\gamma > 0$*, define*

$$T_n = \sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{k} \log Y_{(n-i+1,n)} - \log Y_{(n-k,n)} - 1 \right\},$$

$$V_n = \sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^{k} \left(\log Y_{(n-i+1,n)} - \log Y_{(n-k,n)} \right)^2 - 2 \right\}.$$

Then (T_n, V_n) *is asymptotically normal with mean equal to* $\binom{0}{0}$ *and covariance matrix* $\left(\begin{smallmatrix} 1 & 4 \\ 4 & 20 \end{smallmatrix}\right)$.

Proof of Theorem 2.1. Note that the condition (10) is equivalent to

$$\lim_{t \to \infty} \frac{\log U(tx) - \log U(t) - \gamma \log x}{A(t)} = \frac{x^{\rho} - 1}{\rho}.$$

Consequently we have

$$\log U(tx) - \log U(t) = \gamma \log x + A(t) \frac{x^{\rho} - 1}{\rho} (1 + o(1))$$

and

$$(\log U(tx) - \log U(t))^2 = (\gamma \log x)^2 + 2\gamma \frac{x^{\rho} - 1}{\rho} (\log x) A(t) + o(A(t)),$$

as $t \to \infty$.

Let us consider the variables presented in Lemma 4.1.

Since $\left(X_{(1,n)}, X_{(2,n)}, \cdots, X_{(n,n)}\right) \stackrel{D}{=} \left(U\left(Y_{(1,n)}\right), U\left(Y_{(2,n)}\right), \cdots, U\left(Y_{(n,n)}\right)\right)$, without loss of generality we can write $\left(X_{(i,n)}\right) = \left(U\left(Y_{(i,n)}\right)\right)$.

Then,

$$\begin{aligned} M_n^{(1)} &= \frac{1}{k}\sum_{i=1}^{k} \log X_{(n-i+1,n)} - \log X_{(n-k,n)} \\ &= \frac{1}{k}\sum_{i=1}^{k} \log U\left(\frac{Y_{(n-i+1,n)}}{Y_{(n-k,n)}} Y_{(n-k,n)}\right) - \log U\left(Y_{(n-k,n)}\right) \\ &= \gamma + \frac{\gamma}{\sqrt{k}} T_n + \frac{A\left(Y_{(n-k,n)}\right)}{1-\rho} + o_p\left(A\left(Y_{(n-k,n)}\right)\right), \end{aligned}$$

since

$$\frac{1}{k}\sum_{i=1}^{k}\left(\frac{\left(\frac{Y_{(n-i+1,n)}}{Y_{(n-k,n)}}\right)^{\rho} - 1}{\rho}\right) \overset{D}{=} \frac{1}{k}\sum_{i=1}^{k}\left(\frac{Y_i^{\rho}-1}{\rho}\right),$$

which tends to $E\left(\frac{Y_1^{\rho}-1}{\rho}\right) = \frac{1}{1-\rho}$.

We have also

$$\begin{aligned} M_n^{(2)} &= \frac{1}{k}\sum_{i=1}^{k}\left[\log X_{(n-i+1,n)} - \log X_{(n-k,n)}\right]^2 \\ &= 2\gamma^2 + \frac{\gamma^2}{\sqrt{k}} V_n + A\left(Y_{(n-k,n)}\right)\frac{2\gamma\,(2-\rho)}{(1-\rho)^2} + o_p\left(A\left(Y_{(n-k,n)}\right)\right), \end{aligned}$$

since

$$\frac{1}{k}\sum_{i=1}^{k}\left(\frac{\left(\frac{Y_{(n-i+1,n)}}{Y_{(n-k,n)}}\right)^{\rho} - 1}{\rho}\log\frac{Y_{(n-i+1,n)}}{Y_{(n-k,n)}}\right) \overset{D}{=} \frac{1}{k}\sum_{i=1}^{k}\left(\frac{Y_i^{\rho}-1}{\rho}\log Y_i\right),$$

which tends to $E\left(\frac{Y_1^{\rho}-1}{\rho}\log Y_1\right) = \frac{2-\rho}{(1-\rho)^2}$.

Considering $h(x) = x^2$ and the Taylor expansion of $h\left(M_n^{(1)}\right)$ around $h(\gamma)$ we obtain

$$\left[M_n^{(1)}\right]^2 = \gamma^2 + \frac{2\gamma^2}{\sqrt{k}} T_n + \frac{2\gamma}{1-\rho} A\left(Y_{(n-k,n)}\right) + o_p\left(A\left(Y_{(n-k,n)}\right)\right) + O_p\left(\frac{1}{k}\right).$$

Using the above representations we obtain

$$\widetilde{\gamma}^2(k) = M_n^{(2)} - \left[M_n^{(1)}\right]^2$$
$$= \gamma^2 + \frac{\gamma^2}{\sqrt{k}}(V_n - 2T_n) + d\ A\left(Y_{(n-k,n)}\right) + o_p\left(A\left(Y_{(n-k,n)}\right)\right) + O_p\left(\frac{1}{k}\right),$$

where $d = 2\gamma/(1-\rho)^2$.

Since $A(t) \in RV_\rho$, then $A(tx) = x^\rho A(t)(1+o(1))$.

Noting that $(k/n)\,Y_{(n-k,n)} = 1 + o_P(1)$, we have

$$A\left(Y_{(n-k,n)}\right) = A\left(\frac{n}{k}(1+o_P(1))\right)$$
$$= A\left(\frac{n}{k}\right) + o_P\left(A\left(\frac{n}{k}\right)\right).$$

Therefore, we may write

$$\widetilde{\gamma}^2(k) = \gamma^2 + \frac{\gamma^2}{\sqrt{k}}(V_n - 2T_n) + d\ A\left(\frac{n}{k}\right) + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{1}{k}\right).$$

Considering $g(x) = \sqrt{x}$ and the Taylor expansion of $g\left(\widetilde{\gamma}^2(k)\right)$ around $g\left(\gamma^2\right)$ we obtain

(17)

$$\widetilde{\gamma}(k) = \gamma + \frac{1}{2\gamma}\left[\frac{\gamma^2}{\sqrt{k}}(V_n - 2T_n) + d\ A\left(\frac{n}{k}\right) + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{1}{k}\right)\right]$$
$$= \gamma + \frac{\gamma}{2\sqrt{k}}V_n - \frac{\gamma}{\sqrt{k}}T_n + \frac{A\left(\frac{n}{k}\right)}{(1-\rho)^2} + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{1}{k}\right).$$

Recall that $\log\left(Y_{(n-i+1,n)}/Y_{(n-k,n)}\right)$ are exponential standard r.v., $Exp(1)$. Using Lemma 4.1, from (17) we can write

$$\widetilde{\gamma}(k) \stackrel{D}{=} \gamma + \frac{\gamma}{2\sqrt{k}}Q_n - \frac{\gamma}{\sqrt{k}}P_n + \frac{A\left(\frac{n}{k}\right)}{(1-\rho)^2} + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{1}{k}\right),$$

where $P_n = \sqrt{k}\left(\sum_{i=1}^k Z_i/k - 1\right)$, $Q_n = \sqrt{k}\left(\sum_{i=1}^k Z_i^2/k - 2\right)$, with Z_i i.i.d. exponential standard r.v., are jointly asymptotic normal.

This completes the proof. □

Proof of Corollary 2.2. As $(P_n, Q_n) \xrightarrow{D} N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 4 \\ 4 & 20 \end{pmatrix}\right]$,

$$V\left(\sqrt{k}\left(\widetilde{\gamma}(k) - \gamma\right)\right) = V\left(\frac{\gamma}{2}Q_n\right) + V(\gamma P_n) - 2Cov\left(\frac{\gamma}{2}Q_n, \gamma P_n\right) \xrightarrow[n\to\infty]{} 2\gamma^2.$$

The result follows from the proof of Theorem 2.1. □

For proving Theorem 2.3 we use the following auxiliary Lemma.

Lemma 4.2 (Brito and Freitas (2003), Lemma 2)**.** *Let k be a sequence of positive integers such that $1 \leq k \leq n$. For the sequence $i_n(k)$ defined in* (7) *we have*

$$i_n(k) = 1 + O\left(\frac{log^2 k}{k}\right).$$

Proof of Theorem 2.3. We recall that

$$\widehat{GT}(k) = \frac{\widetilde{\gamma}(k)}{\sqrt{i_n(k)}}.$$

Note now that we can write

$$\sqrt{k}\left(\widehat{GT}(k) - \gamma\right) = \sqrt{k}\left(\widetilde{\gamma}(k) - \gamma\right) + \sqrt{k}\widetilde{\gamma}(k)\left(\frac{1}{\sqrt{i_n(k)}} - 1\right).$$

As $\widetilde{\gamma}(k) \xrightarrow[n\to\infty]{P} \gamma$, from Lemma 4.2 we get

$$\sqrt{k}\widetilde{\gamma}(k)\left(\frac{1}{\sqrt{i_n(k)}} - 1\right) = O_p\left(\frac{log^2 k}{\sqrt{k}}\right),$$

So, from the proof of Theorem 2.1 and Lemma 4.2, we have

$$\widehat{GT}(k) = \gamma + \frac{\gamma}{2\sqrt{k}}V_n - \frac{\gamma}{\sqrt{k}}T_n + \frac{A\left(\frac{n}{k}\right)}{(1-\rho)^2} + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{log^2 k}{k}\right),$$

where T_n and V_n are the same as in proof of Theorem 2.1 and the result follows. □

Proof of Corollary 2.4. By Theorem 2.3, the result is established in a similar way to the proof of Corollary 2.2. □

Proof of Theorem 3.1. Recall that $\widehat{GT}(k) \xrightarrow{P} \gamma$ as $n \to \infty$. If all parameters are known, except the tail index γ, we get

$$\begin{aligned}
\overline{\widehat{GT}}(k) &= \widehat{GT}(k)\left(1 - \frac{\beta\left(\frac{n}{k}\right)^{\rho}}{(1-\rho)^2}\right) \\
&= \widehat{GT}(k) - \frac{A\left(\frac{n}{k}\right)}{(1-\rho)^2}\left(1 + o_p(1)\right) \\
&= \gamma + \frac{\gamma}{2\sqrt{k}}V_n - \frac{\gamma}{\sqrt{k}}T_n + o_p\left(A\left(\frac{n}{k}\right)\right) + O_p\left(\frac{log^2 k}{k}\right).
\end{aligned}$$

With an easy calculation, we also have

$$\begin{aligned}\overline{\overline{\widehat{GT}}}(k) &= \widehat{GT}(k)\, exp\left(-\frac{\beta\left(\frac{n}{k}\right)^{\rho}}{(1-\rho)^2}\right)\\ &= \widehat{GT}(k)\left[1-\frac{A\left(\frac{n}{k}\right)}{\gamma(1-\rho)^2}+o_p\left(\frac{A\left(\frac{n}{k}\right)}{\gamma(1-\rho)^2}\right)\right]\\ &= \gamma+\frac{\gamma}{2\sqrt{k}}V_n-\frac{\gamma}{\sqrt{k}}T_n+o_p\left(A\left(\frac{n}{k}\right)\right)+O_p\left(\frac{log^2 k}{k}\right).\end{aligned}$$

□

Proof of Corollary 3.2. From the proof of Theorem 3.1 we have

$$\sqrt{k}\left(\widehat{GT}^*(k)-\gamma\right)=\frac{\gamma}{2}V_n-\gamma T_n+\sqrt{k}o_p\left(A\left(\frac{n}{k}\right)\right)+O_p\left(\frac{log^2 k}{\sqrt{k}}\right).$$

Since $\sqrt{k}A(n/k)\to\lambda$ as $n\to\infty$,

$$\sqrt{k}\left(\widehat{GT}^*(k)-\gamma\right)=\frac{\gamma}{2}V_n-\gamma T_n+o_p(1).$$

It remains to compute the values of the asymptotic variance and mean.

$$E\left[\sqrt{k}\left(\widehat{GT}^*(k)-\gamma\right)\right]=\frac{\gamma}{2}E(V_n)-\gamma E(T_n)\xrightarrow[n\to\infty]{}0,$$

$$V\left[\sqrt{k}\left(\widehat{GT}^*(k)-\gamma\right)\right]=\frac{\gamma^2}{4}V(V_n)+\gamma^2V(T_n)-2Cov\left(\frac{\gamma}{2}V_n,\gamma T_n\right)\xrightarrow[n\to\infty]{}2\gamma^2.$$

□

Proof of Theorem 3.3. If ρ and β are estimated consistently, we can use the Taylor's expansion for bivariate functions and get

$$\begin{aligned}\frac{\widehat{\beta}}{(1-\widehat{\rho})^2}\left(\frac{n}{k}\right)^{\widehat{\rho}} &= \frac{\beta}{(1-\rho)^2}\left(\frac{n}{k}\right)^{\rho}+\left(\widehat{\beta}-\beta\right)\frac{1}{(1-\rho)^2}\left(\frac{n}{k}\right)^{\rho}(1+o_p(1))\\ &\quad+\frac{\beta}{(1-\rho)^2}(\widehat{\rho}-\rho)\left(\frac{n}{k}\right)^{\rho}\left(\frac{2}{1-\rho}+\log\left(\frac{n}{k}\right)\right)(1+o_p(1))\\ &= \frac{A(n/k)}{\gamma(1-\rho)^2}\left(\frac{\widehat{\beta}}{\beta}+\frac{2(\widehat{\rho}-\rho)}{1-\rho}+(\widehat{\rho}-\rho)\log\left(\frac{n}{k}\right)\right)(1+o_p(1)),\end{aligned}$$

where $\widehat{\beta}$ and $\widehat{\rho}$ are the estimators of β and ρ, respectively.

Therefore we have

$$\widehat{GT}(k)\left(1-\frac{\widehat{\beta}\left(\frac{n}{k}\right)^{\widehat{\rho}}}{(1-\widehat{\rho})^2}\right)=\widehat{GT}(k)-\frac{A\left(\frac{n}{k}\right)}{(1-\rho)^2}+o_p\left(A\left(\frac{n}{k}\right)\right)$$

$$=\gamma+\frac{\gamma}{\sqrt{k}}\left(\frac{V_n}{2}-T_n\right)+o_p\left(A\left(\frac{n}{k}\right)\right)+O_p\left(\frac{log^2k}{k}\right)$$

and

$$\widehat{GT}(k)\,exp\left(-\frac{\widehat{\beta}\left(\frac{n}{k}\right)^{\widehat{\rho}}}{(1-\widehat{\rho})^2}\right)=\gamma+\frac{\gamma}{\sqrt{k}}\left(\frac{V_n}{2}-T_n\right)+o_p\left(A\left(\frac{n}{k}\right)\right)+O_p\left(\frac{log^2k}{k}\right),$$

since $\widehat{\rho}$ and $\widehat{\beta}$ are consistent estimators of ρ and β computed at a level such that $\widehat{\rho}-\rho=o_p(1/\log n)$. The result follows. □

Proof of Corollary 3.4. The result follows using the same approach as in the proof of Corollary 3.2. □

5. Simulation results

In this section we present some simulations in order to examine the finite sample behaviour of the proposed tail index estimators. We have generated s=2000 independent replicates of sample size 1000 from the Generalised Pareto Distribution (GPD) with d.f.

$$F(x)=1-(1+\gamma x)^{-1/\gamma},\quad x\geq 0,\quad \gamma=1,$$

and from the Burr distribution with d.f.

$$F(x)=1-\left(1+x^{-\rho/\gamma}\right)^{1/\rho},\quad x\geq 0,\quad \gamma=1\text{ and }\rho=-2.$$

Remark that $\beta=1$ for both families, and for GPD $\rho=-\gamma$.

The results were compared using mean values of the estimates and through relative root mean square error (RRMSE), with the expression

$$\widehat{RRMSE}\left(\widehat{\theta}\right)=\frac{\sqrt{\frac{1}{s}\sum_{i=1}^{s}\left(\widehat{\theta}_i-\theta\right)^2}}{\theta},$$

where θ is the value we want to estimate.

The main purpose of the simulations performed in this section is to provide a general insight into the distributional behaviour of the new geometric-type bias corrected tail index estimators proposed, (12) and (13). Once the evaluation of their behaviour encompasses the comparison with similar corrections of Hill estimator, we start by presenting in Figures 2 and 3 the behaviour of both original estimators for the chosen distributions.

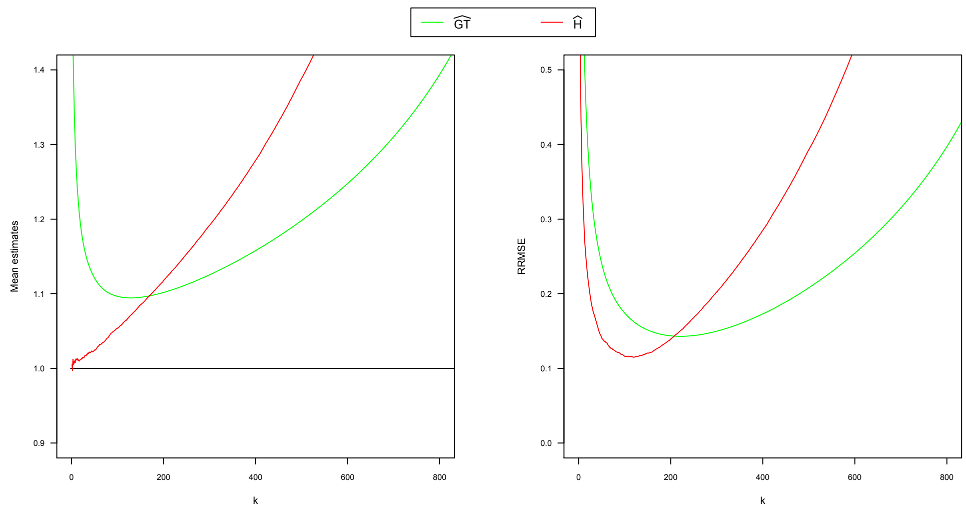

Figure 2. Mean estimates (left) and RRMSE (right) of $\widehat{GT}$ and $\widehat{H}$, for a sample size n=1000 (and 2000 replicates), as a function of k, from a GPD given by $F(x) = 1 - (1 + \gamma x)^{-1/\gamma}$, $x \geq 0$ with $\gamma = 1$.

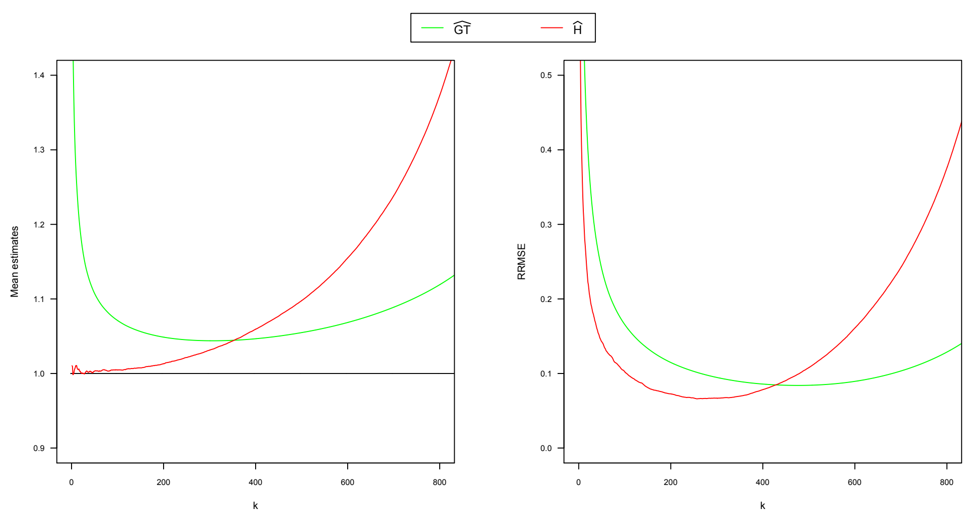

Figure 3. Mean estimates (left) and RRMSE (right) of $\widehat{GT}$ and $\widehat{H}$, for a sample size n=1000 (and 2000 replicates), as a function of k, from a Burr distribution given by $F(x) = 1 - \left(1 + x^{-\rho/\gamma}\right)^{1/\rho}$, $x \geq 0$, with $\gamma = 1$ and $\rho = -2$.

To illustrate the behaviour of the corrected estimators we consider the suitable estimators of the parameter ρ proposed by Fraga Alves et al. (2003), in (14), and the β estimator obtained in Gomes and Martins (2002), in (15). Firstly we need to choose the tuning parameter τ, in which we will support the estimation of the second order parameters ρ and β. To achieve this, we draw in Figure 4 the behaviour of $\widehat{\rho}_\tau$ for the values of the control parameter $\tau \in \{0, 0.5, 1\}$ for both distributions and analyse the variations that it causes in their behaviour.

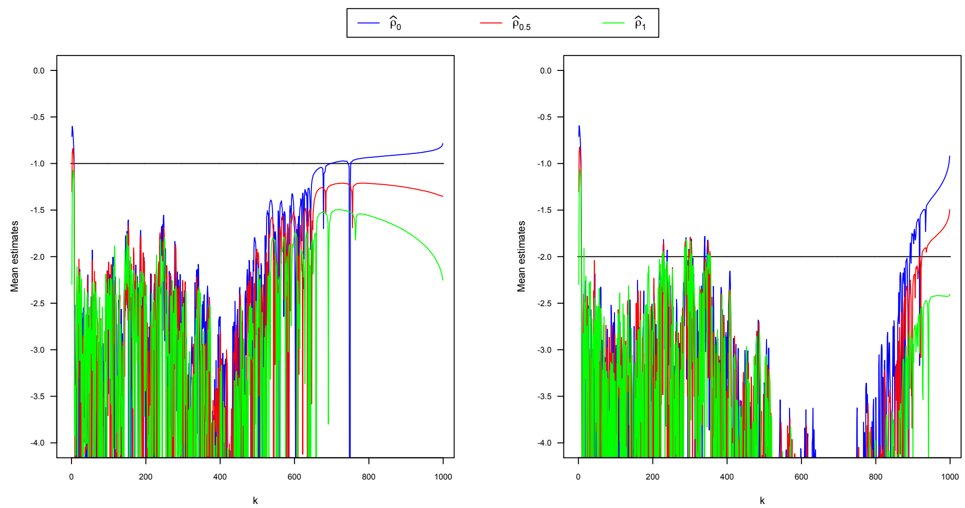

Figure 4. Mean estimates of $\widehat{\rho}_\tau$, $\tau = \{0, 0.5, 1\}$, for GPD (left) and Burr (right) distributions. GPD given by $F(x) = 1 - (1 + \gamma x)^{-1/\gamma}$, $x \geq 0$ ($\rho = -1$), and Burr distribution given by $F(x) = 1 - \left(1 + x^{-\rho/\gamma}\right)^{1/\rho}$, $x \geq 0$ and $\rho = -2$, both with $\gamma = 1$ ($\beta = 1$).

It is suggested in some works the use of $\tau = 0$ when $\rho \in [-1, 0)$ and $\tau = 1$ when $\rho \in (-\infty, -1)$ (see e.g. Fraga Alves et al. (2003)). This leads to the choice of $\tau = 0$ for the GPD ($\rho = -1$) and $\tau = 1$ for Burr distribution ($\rho = -2$). The Figure 4 confirms the prevalent choice of $\tau = 0$ for GPD but suggests that perhaps the choice of $\tau = 0.5$ instead of $\tau = 1$ seems to be more suitable for Burr distribution, leading to better estimates of β and ρ.

We also remark that the estimator of ρ presents a high variation in the majority of k values, stabilizing only at very high levels of k, for which the estimates gets closer to the true value of the parameter. This fact reaffirm that estimation of ρ at a high level is favourable and highly recommended.

For exploring the results we consider in (16) $\epsilon = 0.005$ and $\epsilon = 0.001$, ie, we use the following k_h levels:

$$k_{h1} = \left\lfloor n^{0.995} \right\rfloor \quad \text{and} \quad k_{h2} = \left\lfloor n^{0.999} \right\rfloor . \tag{18}$$

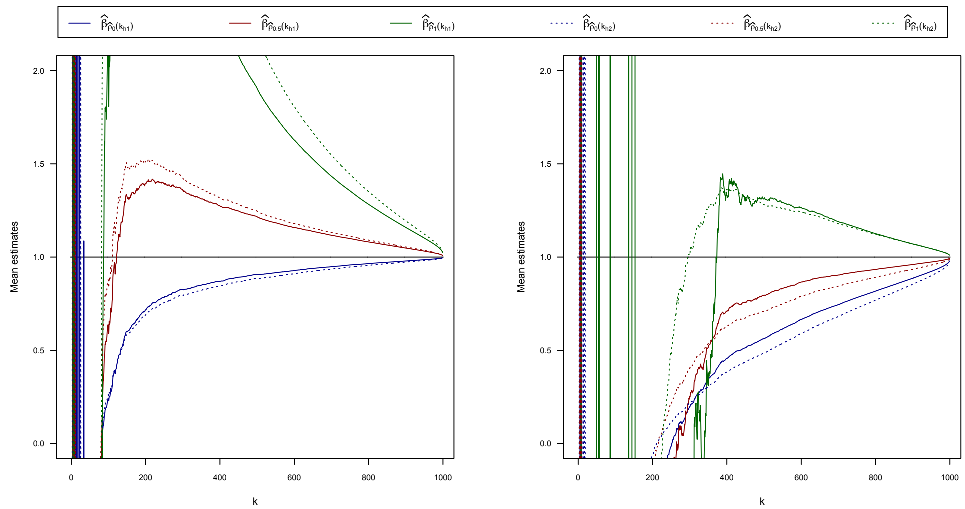

Figure 5. Mean estimates of $\widehat{\beta}_{\widehat{\rho}_\tau(k_{h1})}$ and $\widehat{\beta}_{\widehat{\rho}_\tau(k_{h2})}$, $\tau = \{0, 0.5, 1\}$, for GPD (left) and Burr (right) distributions. GPD given by $F(x) = 1 - (1+\gamma x)^{-1/\gamma}$, $x \geq 0$ ($\rho = -1$), and Burr distribution given by $F(x) = 1 - \left(1 + x^{-\rho/\gamma}\right)^{1/\rho}$, $x \geq 0$ and $\rho = -2$, both with $\gamma = 1$ ($\beta = 1$).

To give an idea about the behaviour of $\widehat{\beta}$ according to the choice of τ and the level k_h, we present in Figure 5 the estimates of β computed with $\widehat{\rho}_\tau(k_{h1})$ and $\widehat{\rho}_\tau(k_{h2})$, $\tau \in \{0, 0.5, 1\}$, for both distributions. One aspect that stands out in this figure is that estimates of β are more favourable the higher the k value used for its calculation.

Following what seems to be graphically more propitious, we chose to estimate ρ and β using $\tau = 0$ for GPD and $\tau = 0.5$ for Burr distribution, both computed at the same level k_{h1} or k_{h2}. The correct estimation of these parameters is crucial in order to get better estimates of the tail index using corrected estimators.

Now we have the necessary tools to estimate the tail index using the bias corrected tail index estimators. In this way, the illustrations that follow contain a graphical representation of the behaviour of the estimators corrected according to the choices on τ made for each distribution.

From the asymptotic normality we construct confidence intervals for the tail index, with $(1-\alpha)$-level, in the usual way:

$$I_{\widehat{GT}}(k,\alpha) = \left\{ \gamma : \frac{1}{\sqrt{2}\gamma}\, k^{1/2} |\widehat{GT} - \gamma| \leq \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \right\}.$$

The confidence bounds for the corresponding geometric-type bias corrected estimators are similar to the previous ones.

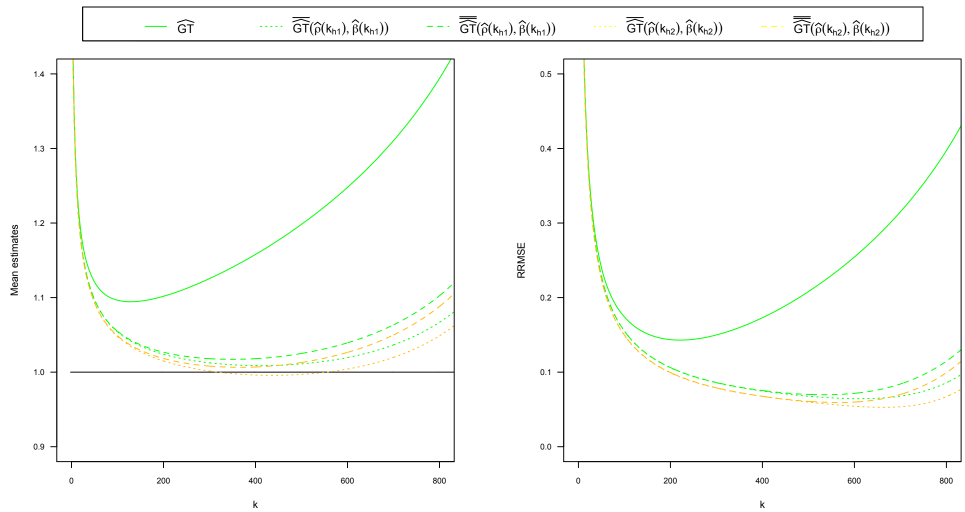

Figure 6. Mean estimates (left) and RRMSE (right) of $\widehat{GT}$, $\overline{\widehat{GT}}$ and $\overline{\overline{\widehat{GT}}}$, with $\widehat{\rho}$ and $\widehat{\beta}$ computed at the levels $k_{h1} = \lfloor n^{0.995} \rfloor$ and $k_{h2} = \lfloor n^{0.999} \rfloor$, for a sample size n=1000 (and 2000 replicates), as a function of k, from a GPD given by $F(x) = 1 - (1 + \gamma x)^{-1/\gamma}$, $x \geq 0$ with $\gamma = 1$ ($\rho = -1$, $\beta = 1$; $\tau = 0$).

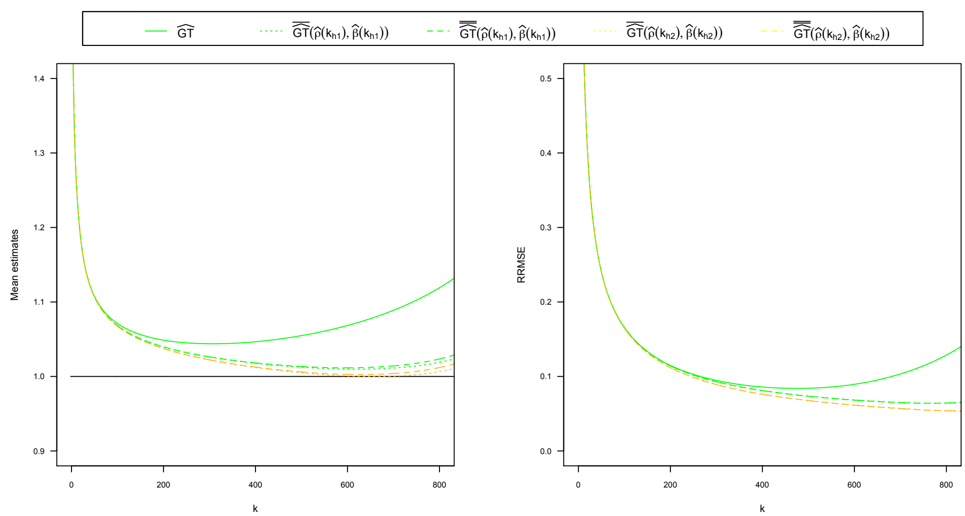

Figure 7. Mean estimates (left) and RRMSE (right) of $\widehat{GT}$, $\overline{\widehat{GT}}$ and $\overline{\overline{\widehat{GT}}}$, with $\widehat{\rho}$ and $\widehat{\beta}$ computed at the levels $k_{h1} = \lfloor n^{0.995} \rfloor$ and $k_{h2} = \lfloor n^{0.999} \rfloor$, for a sample size n=1000 (and 2000 replicates), as a function of k, from a Burr distribution given by $F(x) = 1 - \left(1 + x^{-\rho/\gamma}\right)^{1/\rho}$, $x \geq 0$, with $\gamma = 1$, $\rho = -2$ ($\beta = 1$; $\tau = 0.5$).

From the Figures 6 and 7, in which the geometric-type estimator is confronted with its new corrected versions, we observe that using both

GPD and Burr distribution, the performance of the geometric-type estimator was improved by bias correction and the resulting geometric-type bias corrected estimators shows a very good behaviour.

We also note that the performance of the corrected estimators are slightly better when we calculate the second order parameters using the level k_{h2} instead of using the k_{h1} level. The corresponding 95% confidence bounds of the geometric-type estimator and of the corresponding bias corrected estimators are reported in Tables 1 and 2. We present three values of k for the illustration of the influence of the choice of k.

Table 1. Confidence bounds ($\alpha = 0.05$) using the geometric-type estimator and the corresponding bias corrected estimators, with $\widehat{\rho}$ and $\widehat{\beta}$ computed at the levels $k_{h1} = \lfloor n^{0.995} \rfloor$ and $k_{h2} = \lfloor n^{0.999} \rfloor$, for a sample size n=1000 (and 2000 replicates), as a function of k. GPD $F(x) = 1 - (1+\gamma x)^{-1/\gamma}$, $x \geq 0$ with $\gamma = 1$ ($\rho = -1$, $\beta = 1$; $\tau = 0$).

k	$\widehat{GT}$	$\overline{\overline{GT}}_{\widehat{\rho}(k_{h1}),\widehat{\beta}(k_{h1})}$	$\overline{\overline{GT}}_{\widehat{\rho}(k_{h2}),\widehat{\beta}(k_{h2})}$	$\overline{\overline{\overline{GT}}}_{\widehat{\rho}(k_{h1}),\widehat{\beta}(k_{h1})}$	$\overline{\overline{\overline{GT}}}_{\widehat{\rho}(k_{h2}),\widehat{\beta}(k_{h2})}$
300	1.125 ± 0.180	1.013 ± 0.162	1.002 ± 0.160	1.018 ± 0.163	1.008 ± 0.161
500	1.198 ± 0.149	1.011 ± 0.125	0.997 ± 0.124	1.025 ± 0.127	1.013 ± 0.126
700	1.310 ± 0.137	1.037 ± 0.109	1.020 ± 0.107	1.063 ± 0.111	1.050 ± 0.110

Table 2. Confidence bounds ($\alpha = 0.05$) using the geometric-type estimator and the corresponding bias corrected estimators, with $\widehat{\rho}$ and $\widehat{\beta}$ computed at the levels $k_{h1} = \lfloor n^{0.995} \rfloor$ and $k_{h2} = \lfloor n^{0.999} \rfloor$, for a sample size n=1000 (and 2000 replicates), as a function of k. Burr distribution $F(x) = 1 - \left(1 + x^{-\rho/\gamma}\right)^{1/\rho}$, $x \geq 0$, with $\gamma = 1$ and $\rho = -2$ ($\beta = 1$; $\tau = 0.5$).

k	$\widehat{GT}$	$\overline{\overline{GT}}_{\widehat{\rho}(k_{h1}),\widehat{\beta}(k_{h1})}$	$\overline{\overline{GT}}_{\widehat{\rho}(k_{h2}),\widehat{\beta}(k_{h2})}$	$\overline{\overline{\overline{GT}}}_{\widehat{\rho}(k_{h1}),\widehat{\beta}(k_{h1})}$	$\overline{\overline{\overline{GT}}}_{\widehat{\rho}(k_{h2}),\widehat{\beta}(k_{h2})}$
300	1.044 ± 0.167	1.026 ± 0.164	1.022 ± 0.164	1.026 ± 0.164	1.022 ± 0.164
500	1.055 ± 0.131	1.012 ± 0.125	1.005 ± 0.125	1.013 ± 0.126	1.006 ± 0.125
700	1.089 ± 0.114	1.011 ± 0.106	1.000 ± 0.105	1.014 ± 0.106	1.004 ± 0.105

We note that when using corrected estimators, the amplitude of the asymptotic confidence intervals is smaller.

In order to have an idea of the good behaviour of the geometric-type bias corrected estimators, we compare them with the corresponding

Hill bias corrected estimators (see e.g. Caeiro et al. (2005)), given by

$$\overline{\widehat{H}}(k) = \widehat{H}(k)\left(1 - \frac{\widehat{\beta}\left(\frac{n}{k}\right)^{\widehat{\rho}}}{1-\widehat{\rho}}\right)$$

and

$$\overline{\overline{\widehat{H}}}(k) = \widehat{H}(k)\, exp\left\{-\frac{\widehat{\beta}}{1-\widehat{\rho}}\left(\frac{n}{k}\right)^{\widehat{\rho}}\right\},$$

where $\widehat{\rho}$ and $\widehat{\beta}$ are the estimators of the shape and scale parameters, respectively.

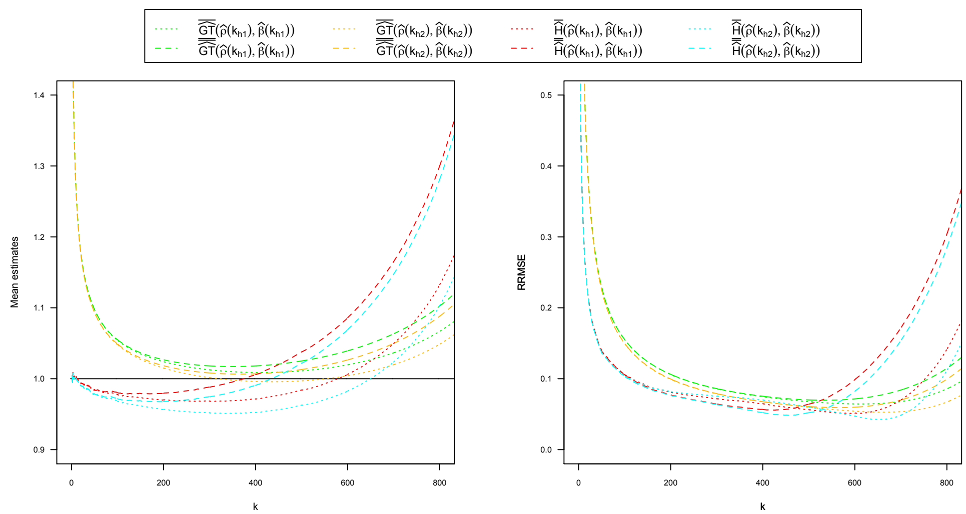

Figure 8. Mean estimates (left) and RRMSE (right) of $\overline{\widehat{GT}}$, $\overline{\overline{\widehat{GT}}}$, $\overline{\widehat{H}}$ and $\overline{\overline{\widehat{H}}}$, with $\widehat{\rho}$ and $\widehat{\beta}$ computed at the levels $k_{h1} = \lfloor n^{0.995} \rfloor$ and $k_{h2} = \lfloor n^{0.999} \rfloor$, for a sample size n=1000 (and 2000 replicates), as a function of k, from a GPD given by $F(x) = 1 - (1+\gamma x)^{-1/\gamma}$, $x \geq 0$ with $\gamma = 1$ ($\rho = -1$, $\beta = 1$; $\tau = 0$).

From Figures 8 and 9, we observe that using GPD and Burr distribution, both the geometric-type and the Hill bias corrected estimators present a good performance. Particularly, we note that for GPD the geometric-type estimator has a better posture for intermediate k-values, while the best behaviour of Hill estimator takes place at low values of k. In the case of Burr distribution, a greater distance from the target value is notable at low k-values for the geometric-type estimators, whereas for the Hill estimators the same it is visible for high k-values.

The Hill estimator exhibits in general a lower RRMSE than the geometric-type estimator, which can be understood considering that the asymptotic variance of the Hill estimator is half of the one of the geometric-type estimator.

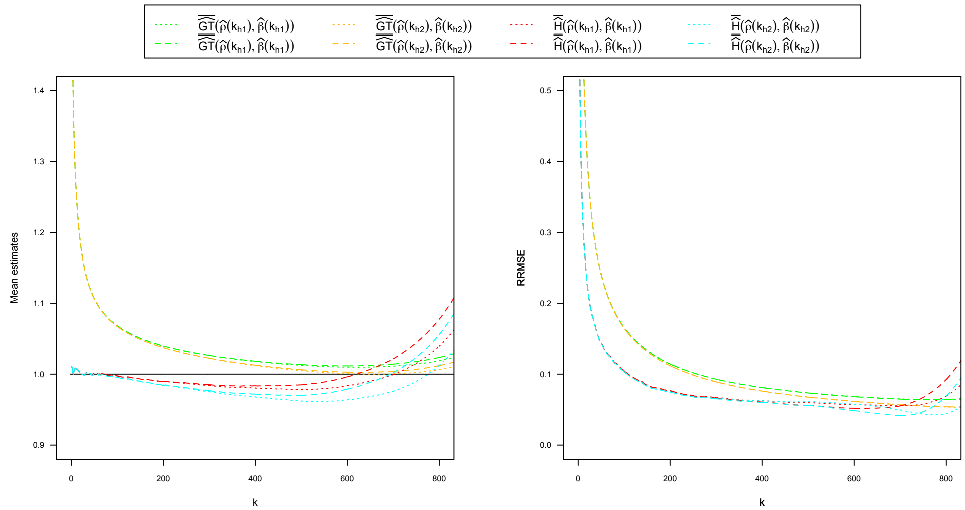

Figure 9. Mean estimates (left) and RRMSE (right) of $\overline{\widehat{GT}}$, $\overline{\overline{\widehat{GT}}}$, $\overline{\widehat{H}}$ and $\overline{\overline{\widehat{H}}}$, with $\widehat{\rho}$ and $\widehat{\beta}$ computed at the levels $k_{h1} = \lfloor n^{0.995} \rfloor$ and $k_{h2} = \lfloor n^{0.999} \rfloor$, for a sample size n=1000 (and 2000 replicates), as a function of k, from a Burr distribution given by $F(x) = 1 - \left(1 + x^{-\rho/\gamma}\right)^{1/\rho}$, $x \geq 0$, with $\gamma = 1$ and $\rho = -2$ ($\beta = 1$; $\tau = 0.5$).

In addition, for GPD and for large k, the estimates based on $\overline{\widehat{H}}$ clearly show far better results than those conducted with $\overline{\overline{\widehat{H}}}$. Unlike what happens with the corrected geometric-type estimators, the corrected Hill ones have the best estimates when the second order parameters are computed using the level k_{h1} instead of using the k_{h2} level, except for very high k-values in which prevails the use of k_{h2}.

We may conclude that the behaviour of the geometric-type estimator is improved by bias correction. The corrected versions show a good performance and for some cases it is even highlighted.

References

Beirlant, J., Dierckx, G., Goegebeur Y. and Matthys, G. (1999). Tail index estimation and an exponential regression model, *Extremes*, **2**, 177–200.

Beirlant, J., Dierckx, G. and Guillou, A. (2005). Estimation of the extreme-value index and generalised quantile plots, *Bernoulli*, **11**, 949–970.

Beirlant, J., Figueiredo, F., Gomes, M.I. and Vandewalle, B. (2008). Improved reduced-bias tail index and quantile estimators, *J. Statist. Plann. and Inference*, **138**, 1851–1870.

Brito, M. and Freitas, A.C.M. (2003). Limiting behaviour of a geometric estimator for tail indices, *Insur. Math. Econ.*, **33**, 221–226.

Brito, M. and Freitas, A.C.M. (2006). Weak convergence of a bootstrap geometric-type estimator with applications to risk theory, *Insur. Math. Econ.* **38**, 571–584.

Caeiro, F. and Gomes, M. I. (2008). Minimum-variance reduced-bias tail index and high quantile estimation, *Revstat* **6**, 1–20.

Caeiro, F., Gomes, M.I. and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator, *Revstat*, **3**, 113–136.

Csörgő, S. and Viharos, L. (1998). *Estimating the tail index.* In "Asymptotic Methods in Probability and Statistics" (B. Szyszkowicz, Eds.), North Holland, Amesterdam, 833–881.

Deheuvels, P., Haeusler, E. and Mason, D.M. (1988). Almost sure convergence of the Hill estimator, *Math. Proc. Camb. Phil. Soc.*, **104**, 371–381.

Dekkers, A.L.M. and de Haan, L. (1993). Optimal choice of sample fraction in extreme-value estimation, *J. Multivar. Anal.*, **47**, 173–195.

Dekkers, A.L.M., Einmahl, J.H.J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution, *Ann. Stat.*, **17**, 1833–1855.

Feuerverger, A. and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution, *Ann. Statist.* **27**, 760–781.

Fraga Alves, M.I., Gomes, M.I. and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter, *Portugaliae Math.*, **60**, 193–213.

Geluk, J. and de Haan, L. (1987). *Regular Variation, Extensions and Tauberian Theorems*, CWI Tract 40, Center of Mathematics and Computer Science, Amsterdam, Netherlands.

Gomes, M. I., Caeiro, F. and Figueiredo, F. (2004). Bias reduction of a tail index estimator trough an external estimation of the second order parameter, Statistics **38**, 497–510.

Gomes, M.I. and Martins, M.J. (2002). "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter, *Extremes*, **5**, 5–31.

Gomes, M.I., Martins, M.J. and Neves, M. (2000). Alternatives to a semi-parametric estimator of parameters of rare events – the jackknife methodology, *Extremes*, **3**, 207–229.

Gomes, M.I, Martins, M.J. and Neves, M. (2007). Improving second order reduced bias extreme value index estimation, *Revstat* **5**, 177–207.

Gomes, M.I. and Pestana, D. (2007). A simple second order reduced bias tail index estimator, *J. Statist. Comput. and Simulation* **77**, 487–504.

Haeusler, E. and Teugels, J.L. (1985). On asymptotic normality of Hill's estimator for the exponent of regular variation, *Ann. Statist.* **13**, 743–756.

Hall, P. (1982). On some simple estimates of an exponent of regular variation, *J. Royal Statist. Soc.*, **B44**, 37–42.

Hill, B.M. (1975). A simple approach to inference about the tail of a distribution, *Ann. Stat.*, **3**, 1163–1174.

Kratz, M. and Resnick, S. (1996). The qq-estimator of the index of regular variation, *Commun. Statist. Stochastic Models* **12**, 699–724.

Peng, L. (1998). Asymptotically unbiased estimator for the extreme-value-index, *Statistics and Probability Letters* **38**, 107–115.

Pickands, J. (1975). Statistical inference using extreme order statistics, *Ann. Stat.*, **3**, 119–13.

Schultze, J. and Steinebach, J. (1996). On least squares estimates of an exponential tail coefficient, *Statist. Decisions* **14**, 353–372.

University of Porto, Portugal
Current address: Department of Mathematics, University of Porto, Portugal
E-mail address: mabrito@fc.up.pt

University of Porto, Portugal
Current address: Department of Mathematics, University of Porto, Portugal
E-mail address: laucavalcante@fc.up.pt

University of Porto, Portugal
Current address: Department of Mathematics, University of Porto, Portugal
E-mail address: amoreira@fep.up.pt