BIAS-CORRECTED GEOMETRIC-TYPE ESTIMATORS OF THE TAIL INDEX

MARGARIDA BRITO, LAURA CAVALCANTE, AND ANA CRISTINA MOREIRA FREITAS

ABSTRACT. The estimation of the tail index is a central topic in 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 \( F = 1 - F \in RV_{1/\gamma}, \gamma > 0 \), where \( RV_\alpha \) denotes the class of regularly varying functions with index of regular variation equal to \( \alpha \), that is,

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

Equivalently,

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

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

\[
U(x) = x^\gamma L(x),
\]

where \( L \) is a slowly varying function at infinity. Our main goal here is the estimation of \( \gamma \), namely the construction of bias-corrected asymptotically normal estimators, using geometric-type estimators.

The class of distributions \( F \) verifying condition (1) plays an important role in extreme value statistics. The univariate extreme value theory (EVT) and its applications is introduced and developed in several general texts, such as, for instance in Leadbetter et al. (1983) and Coles (2001). In order to provide the theoretical background of our results in this context, we briefly review some basic concepts and results from classic univariate EVT. First, we recall that two distribution functions \( F_1 \) and \( F_2 \) are of the same type if, for some constants \( a > 0 \) and \( b \), \( F_1(x) = F_2(ax + b) \), for all \( x \). Consider now a sequence of

\textit{2010 Mathematics Subject Classification.} Primary 62G32; Secondary 62G30, 62E20, 62F12.

\textit{Key words and phrases.} Bias reduction; extreme value theory; heavy tails; tail index estimation.
independent and identically distributed (i.i.d.) random variables (r.v.) \(X_1, X_2, \ldots\) 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. For simplicity write \(M_n = X_{(n:n)}\) and assume that there exist sequences \((a_n), a_n > 0\) and \((b_n)\) such that

\[
P(M_n \leq a_n x + b_n) = F^n(a_n x + b_n) \rightarrow G(x),
\]

where \(G\) is a non degenerate d.f.. If \(F\) satisfies the above condition, we say that \(F\) belongs to the domain of attraction for maxima of \(G\) and write \(F \in DA(G)\). According to the extremal types theorem, if \(F \in DA(G)\), then \(G\) must have the same type as one of the following three families

Type I: \(\Lambda(x) = \exp(-e^{-x}), x \in \mathbb{R}\);

Type II: \(\Phi_\alpha(x) = \begin{cases} \exp(-x^{-\alpha}), & x > 0, \alpha > 0, \\ 0, & x \leq 0, \end{cases}\)

Type III: \(\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha), & x < 0, \alpha > 0, \\ 1, & x \geq 0. \end{cases}\)

The three types of distributions are known as the Gumbel, Fréchet and Weibull distributions, respectively.

This result for maxima may be viewed as a counterpart of the Central Limit Theorem for sums. Using the reparametrization of von Mises, the above three types may be written in the following form, called generalized extreme value distribution (GEV):

\[
G_\gamma(x) = \exp \left( - \left( 1 + \gamma \left( \frac{x - \mu}{\sigma} \right) \right)^{-1/\gamma} \right), 1 + \gamma \left( \frac{x - \mu}{\sigma} \right) > 0,
\]

where, for \(\gamma = 0\), \(G(x) = \exp(-\exp(-\frac{x-\mu}{\sigma}))\) and \(\mu\) and \(\sigma > 0\) are location and scale parameters, respectively. The shape parameter \(\gamma\), widely known as the tail index or extreme value index, is a key parameter in extreme value statistics, as it determines the general tail behaviour of the underlying d.f. \(F\). The case where \(\gamma = 0\), corresponds to the Gumbel family and the distributions belonging to its domain of attraction have light or moderate upper tails, such as the exponential, gamma, normal or lognormal distributions. For \(\gamma = -1/\alpha < 0\), corresponding to the Weibull family, the distributions have moderate and short tails, with finite right endpoint, like the uniform and beta distributions. The second type of distributions, where \(\gamma = 1/\alpha > 0\), corresponds to the Fréchet family. Distributions with an upper heavy tail, such as, the Student-t, loggamma, F, Pareto and Burr are well known examples of distributions belonging to the domain of attraction for maxima of this family.

The generalized Pareto distribution (GPD) is the other fundamental class of distributions in univariate extreme value statistics, given by

\[
H_\gamma(x) = 1 - (1 + \gamma x/\sigma)^{-1/\gamma}, 1 + \gamma x/\sigma > 0,
\]

where \(\sigma > 0\) and, for \(\gamma = 0\), the above expression is again interpreted as the limit of \(H_\gamma\) as \(\gamma \to 0\), that is, \(H(x) = 1 - \exp(-x/\sigma)\).
The GPD is related to the GEV, and arises as a limit distribution for excesses over a high threshold $u$. More precisely, there is a function $\sigma()$ such that

$$\lim_{u \to x_0} \sup_{0 \leq x \leq x_0 - u} | P(X - u \leq x \mid X > u) - H_{\gamma, \sigma(u)}(x) | = 0,$$

where $x_0$ denotes the upper endpoint, if and only if $F \in DA(G_\gamma)$. The shape parameter $\gamma$ is then the same for GEV and GPD and the GPD class is also a reparametrization of three different families of distributions, corresponding to the cases $\gamma = 0$, $\gamma < 0$ and $\gamma > 0$, respectively.

Several characterisations of the domains of attraction have been established. For instance, condition (1) is necessary and sufficient for $F \in DA(G_\gamma)$, $\gamma > 0$. Thus, the distributions we consider here, the Pareto-type distributions, belong to the domain of attraction of the heavy-tailed Fréchet distribution with tail index $\gamma$.

The problem of estimating the tail index has received considerable attention and common applications may be found in a big variety of domains such as in geology, hydrology, meteorology, oceanography, economics, telecommunications and sociology. In particular, we developed, in Brito, Cavalcante and Freitas (2015), an important application in the field of seismology, where we estimate tail parameters of the seismic moment distribution. We also mention here that the estimation of the tail index has a relevant application to the estimation of the dimension of strange attractors, including attractors of meteorological models. See for example Lucarini et al. (2012) and Faranda et al. (2012).

Consider now intermediate sequences $k = k_n$ of positive integers ($1 \leq k < n$), that is

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

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

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

The asymptotic properties of the Hill estimator have been much studied and it is well known that, under certain conditions, $\hat{H}$ 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

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

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

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

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

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

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, with a better performance in several situations (see Csörgő and Viharos (1998)). Namely 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 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, whose plots often exhibit strong trends and a considerable lack of smoothness resulting in different estimates for neighbouring values of $k$ and an extreme sensitivity 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 sensitivity to the choice of $k$ 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 equivalently,

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

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 minimising the function $f_1(a,b) = \sum_{i=1}^{k} (z_i - ay_i - b)^2$ and the other one by minimising the function $f_2(R,d) = \sum_{i=1}^{k} (y_i - Rz_i + d)^2$. These choices correspond to determine the inverse of the slope of the line by minimising the sum of the distances between the points $\{(z_i,y_i), i = 1, \ldots, k\}$ and the line, measured horizontally or vertically, respectively.
The \( \hat{R} \) estimator is obtained through a geometrical adaptation of these two perspectives, minimising 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 minimise 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, a function of both the horizontal and the vertical distance between the points \( \{(z_i, y_i), i = 1, \ldots, k\} \) and the line is minimised.

![Figure 1. Geometric representation of the rectangles whose areas will be minimised to obtain the estimator of \( R \).](image)

The asymptotic properties of \( \hat{R} \) 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 \( k^{1/2} \left( \hat{R}(k) - R \right) \) is asymptotically normal. This estimator also enjoys certain properties that make 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 \( \gamma \), we will consider the following geometric-type estimator:

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

where

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

and

\[
M_n^{(j)}(k) = \frac{1}{k} \sum_{i=1}^{k} \left( \log X_{(n-i+1,n)} - \log X_{(n-k,n)} \right)^j.
\]
The asymptotic properties of $\hat{GT}$ arise naturally, through the delta method (see, e.g., Dudewicz and Mishra (1988), Theorem 6.3.18), from the corresponding properties of $\hat{R}$ studied in Brito and Freitas (2003).

Here we assume there exists a positive function $a$ such that, for all $x > 0$,

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

(9) \hspace{1cm} \text{From (2), we can choose } a(t) = \gamma U(t).

The procedures analysed in this work are formulated under second order conditions. We suppose that there exists a function $A(t)$, tending to zero as $t \to \infty$, such that

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

(10) \hspace{1cm} \text{for all } x > 0, \text{ where } \rho < 0 \text{ is the shape parameter governing the rate of convergence of } U(tx)/U(t) \text{ and the function } |A(t)| \in RV_\rho \text{ (see, e.g., Geluk and de Haan (1987)). A very important family satisfying (10) is the Hall’s class (Hall (1982)), defined by:}

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

(11) \hspace{1cm} \text{where}

$$A(t) = \gamma \beta t^\rho,$$

with $\gamma > 0$ and $C > 0$, and $\rho < 0$ and $\beta \neq 0$ are, respectively, the shape and scale parameters.

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 $\gamma$ increases, the bias in the estimation becomes larger. Bias reduction is a complex problem and received considerable attention in extreme value statistics (see, e.g., Beirlant et al. (2008) and references therein). 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 introduce two asymptotically equivalent bias-corrected estimators and study their asymptotic behaviour. We use the presented estimators to obtain asymptotic confidence intervals. A simulation study is provided in Section 4, 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. The simulations were performed using the statistical programming language R (R Core Team (2015)). Proofs are presented in Section 5. Some conclusions are given in Section 6. The scripts used to perform the simulations are provided in the Appendix.

2. ASYMPOTIC PROPERTIES OF THE GEOMETRIC-TYPE ESTIMATOR

In this section, we derive the asymptotic distributional representation of the geometric-type estimator.
We begin by considering the asymptotic normality of the following tail index estimator of \( \gamma \)

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

In the sequel, \( D \to \) and \( D = \) stand, respectively, for convergence and equality in distribution.

**Theorem 2.1.** Assume (10) holds. For sequences \( k \) such that (3) holds, we have the following asymptotic distributional representation

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

where \( P_n = \sqrt{k} \left( \sum_{i=1}^k Z_i/k - 1 \right) \) 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 \( (0, 0) \) and covariance matrix \( (1, 4; 4, 20) \), 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

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

In the following theorem, we present the asymptotic distributional representation of the geometric-type estimator.

**Theorem 2.3.** Assume (10) holds. For sequences \( k \) such that (3) holds, we have the following asymptotic distributional representation

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

where \( P_n = \sqrt{k} \left( \sum_{i=1}^k Z_i/k - 1 \right) \) 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 \( (0, 0) \) and covariance matrix \( (1, 4; 4, 20) \), 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

\[
\sqrt{k} (\hat{\gamma}_{GT}(k) - \gamma) \xrightarrow{D} N\left(\frac{\lambda}{(1-\rho)^2}, 2\gamma^2\right).
\]

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.
3.1. **Bias-corrected geometric-type estimators.**

Bias-corrected estimators may be constructed in a natural way (see, e.g., Caeiro et al. (2005)) from the asymptotic representation of the geometric-type estimator given in Theorem 2.3,

\[
\hat{GT}(k) = D + \frac{\gamma}{2\sqrt{k}}Q_n - \frac{\gamma}{\sqrt{k}}P_n + \frac{A\left(\frac{n}{k}\right)}{(1-\rho)^2} + o_p\left(\frac{A\left(\frac{n}{k}\right)}{k}\right) + O_p\left(\frac{\log k}{k}\right)
\]

The bias dominant component can be written as

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

and thus, removing this component, we obtain a bias-corrected estimator of \(\hat{GT}\) given by

\[
\overline{\hat{GT}}(k) = \hat{GT}(k) \left(1 - \frac{\beta\left(\frac{n}{k}\right)^\rho}{(1-\rho)^2}\right).
\] (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

\[
\overline{\overline{\hat{GT}}}(k) = \hat{GT}(k) \exp\left\{-\frac{\beta\left(\frac{n}{k}\right)^\rho}{(1-\rho)^2}\right\}.
\] (13)

We can easily note that the bias dominant component is dependent on the shape \(\rho\) and scale \(\beta\) second order parameters. Thus, it is important to consider an adequate estimation of the second order parameters, \(\rho\) and \(\beta\), 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 \(\rho\) (depending on a parameter \(\tau\) whose role is explained below) proposed by Fraga Alves et al. (2003)

\[
\hat{\rho}_n^{(\tau)}(k) = - \frac{3\left(T_n^{(\tau)}(k) - 1\right)}{T_n^{(\tau)}(k) - 3},
\] (14)
where

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

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

\[
\hat{\beta}_{\rho}(k) = \left(\frac{k}{n}\right)^{\hat{\rho}} \frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{k}\right)^{-\hat{\rho}} \frac{1}{k} \sum_{i=1}^{k} U_i - \frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{k}\right)^{-\hat{\rho}} U_i - \frac{1}{k} \sum_{i=1}^{k} \left(\frac{i}{k}\right)^{-2\hat{\rho}} U_i,
\]

where

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

with \(1 \leq i \leq k < n\).

We remark that the class of estimators of \(\rho\) presented above, and consequently also the \(\beta\) estimator, is dependent on a tuning parameter \(\tau \geq 0\). In the literature (see, e.g., Fraga Alves et al. (2003)) 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 \(\rho\) 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 important to ensure that the bias reduction is achieved without increasing the asymptotic variance. For that purpose, the external estimation of \(\rho\) and \(\beta\) should be done at a high level \(k\) (larger than the one used for \(\gamma\)-estimation), denoted here by \(k_h\) (see, e.g., Caeiro et al. (2005) and Gomes et al. (2004)). In fact, the simulation studies presented in the next section show that the estimator of \(\rho\) only stabilizes at high levels of \(k\).

As observed in other studies (see, e.g., Gomes et al. (2007)) the level that seems appropriate to be used in applications is

\[
k_h = \left\lfloor n^{1-\epsilon} \right\rfloor,\]

for some \(\epsilon > 0\) 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 \(\gamma\) is unknown and that \(\widehat{GT}^*\) is one of the estimators \(\widehat{GT}\) or \(\widehat{GT}^*\).
Theorem 3.1. Assume (11) holds. For sequences $k$ such that (3) holds, we have the following asymptotic distributional representation

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

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 $(0,0)$ and covariance matrix $\left( \begin{array}{cc} 1 & 4 \\ 4 & 20 \end{array} \right)$, and $\{Z_i\}$ denote i.i.d. standard exponential r.v.

Remark 3.2. We note that, under (11), we may write the first remainder term of the asymptotic distributional representation of Theorem 2.3 explicitly, i.e.

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

and, from Theorem 3.1, we may write

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

Then, for sequences $k$ such that $k^{\rho-1} \log^2 k = o(n^{\rho})$, we have

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

and

$$
\hat{GT}^* (k) \overset{D}{=} \gamma + \frac{\gamma}{2\sqrt{k}} Q_n - \frac{\gamma}{\sqrt{k}} P_n + o_p \left( \left( \frac{n}{k} \right)^{\rho} \right).
$$

Analogously, we obtain that the distributional representation of $\hat{\gamma}$ of Theorem 2.1 is the same of that of $\hat{GT}$. These computations allow us to conclude that, for the Hall’s model, the bias-corrected geometric-type estimators with adequate sequences $k$, present an asymptotic distributional representation with a smaller remainder term than the original estimator.

Corollary 3.3. Assume the conditions of Theorem 3.1 hold. If we choose $k$ such that $\sqrt{k} A(n/k) \to \lambda$ finite, then

$$
\sqrt{k} \left( \hat{GT}^* (k) - \gamma \right) \overset{D}{\underset{n \to \infty}{\to}} N \left( 0, 2\gamma^2 \right).
$$

Assuming now that $\hat{GT}^{**}$ denotes the version of $\hat{GT}^*$ where the parameters $\rho$ and $\beta$ are estimated externally, we have the following result.
Theorem 3.4. Under the conditions of Theorem 3.1 and assuming consistent estimators for \( \rho \) and \( \beta \) computed at a level that implies \( \hat{\rho} - \rho = o_p(1/\log n) \), we have the following asymptotic distributional representation

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

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 \((0, 0)\) and covariance matrix \(\begin{pmatrix} 1 & 4 \\ 4 & 20 \end{pmatrix}\), and \(\{Z_i\}\) denote i.i.d. standard exponential r.v.

Corollary 3.5. Assume the conditions of Theorem 3.4 hold. If we choose \( k \) such that \( \sqrt{k} A(n/k) \to \lambda \) finite, then

\[
\sqrt{k} \left(\hat{GT}^{**}(k) - \gamma\right) \xrightarrow{D} N \left(0, 2\gamma^2\right).
\]

4. Simulation results

In this section we present some simulations in order to examine the finite sample behaviour of the proposed tail index estimators. We began by generating \( s = 2000 \) independent replicates of sample size \( n = 1000 \) from the Generalised Pareto Distribution 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

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

where \( \theta \) is the value we want to estimate.

The main purpose of the simulations performed in this section is to provide, for nice cases, an illustration of the behaviour of the proposed geometric-type bias-corrected tail index estimators. We compare the new estimators with the Hill ones since the Hill estimator is one of the most used in applications.

We start by presenting the behaviour of the original geometric-type and Hill estimators, which is illustrated in Figures 2 and 3.
Figure 2. Mean estimates (left) and RRMSE (right) of $\hat{GT}$ and $\hat{H}$, 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 $\hat{GT}$ and $\hat{H}$, for a sample size $n = 1000$ (and 2000 replicates), as a function of $k$, from a Burr distribution given by $F(x) = 1 - (1 + x^{\rho/\gamma})^{1/\rho}$, $x \geq 0$, with $\gamma = 1$ and $\rho = -2$. 
To illustrate the behaviour of the corrected estimators, we consider the suitable estimators of the parameter $\rho$ proposed by Fraga Alves et al. (2003), in (14), and the $\beta$ estimator obtained in Gomes and Martins (2002), in (15). Firstly we need to choose the tuning parameter $\tau$, in which we will support the estimation of the second order parameters $\rho$ and $\beta$. The use of $\tau = 0$ when $\rho \in [-1,0)$ and $\tau = 1$ when $\rho \in (-\infty,-1)$ is recommended for some other estimators (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 ($\rho = -2$). For our estimator, the behaviour of $\hat{\rho}_\tau$ for the values of the control parameter $\tau \in \{0, 0.5, 1\}$ for both distributions is shown in Figure 4. This figure confirms the prevalent choice of $\tau = 0$ for GPD but suggests that the choice of $\tau = 0.5$ instead of $\tau = 1$ seems to be more suitable for Burr distribution, leading to better estimates of $\beta$ and $\rho$.

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

We also remark that the estimator of $\rho$ presents a high variation for most $k$ values, stabilizing only at very high levels of $k$, for which the estimates get closer to the true value of the parameter. This fact confirms that the estimation of $\rho$ at a high level is favourable and highly recommended. For exploring the results we consider in (16) $\epsilon = 0.005$ and $\epsilon = 0.001$, i.e. we use the following $k_h$ levels:

\begin{align*}
&k_{h1} = \left\lfloor n^{0.995} \right\rfloor \text{ and } k_{h2} = \left\lfloor n^{0.999} \right\rfloor. \n\end{align*}

To give an idea about the behaviour of $\hat{\beta}$ according to the choice of $\tau$ and the level $k_h$, we present in Figure 5 the estimates of $\beta$ computed with $\hat{\rho}_\tau(k_{h1})$ and $\hat{\rho}_\tau(k_{h2})$, $\tau \in \{0, 0.5, 1\}$, for both distributions. This figure indicates that the estimates of $\beta$ present a better
behaviour for larger values of $k$. The adequate estimation of the second order parameters is crucial in order to get better estimates of the tail index using corrected estimators. Following what seems to be more appropriate graphically, we chose to estimate $\rho$ and $\beta$ using $\tau = 0$ for GPD and $\tau = 0.5$ for Burr distribution, both computed at the same level $k_{h_1}$ or $k_{h_2}$.

![Figure 5](image.png)

**Figure 5.** Mean estimates of $\hat{\beta}_{\tau}(k_{h_1})$ and $\hat{\beta}_{\tau}(k_{h_2})$, $\tau = \{0, 0.5, 1\}$, 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 $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x \geq 0$ and $\rho = -2$, both with $\gamma = 1$ ($\beta = 1$).

From Figures 6 and 7, in which the behaviour of the geometric-type estimator is presented as well as 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 show a very good behaviour. We also note that, for sample size $n = 1000$, the performance of the corrected estimators are slightly better when we calculate the second order parameters using the level $k_{h_2}$ instead of using the $k_{h_1}$ level.
Figure 6. Mean estimates (left) and RRMSE (right) of $\hat{G}_T$, $\bar{G}_T$ and $\overline{G}_T$, with $\hat{\rho}$ and $\hat{\beta}$ computed at the levels $k_{h1} = \lceil n^{0.995} \rceil$ and $k_{h2} = \lceil n^{0.999} \rceil$, 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 $\hat{G}_T$, $\bar{G}_T$ and $\overline{G}_T$, with $\hat{\rho}$ and $\hat{\beta}$ computed at the levels $k_{h1} = \lceil n^{0.995} \rceil$ and $k_{h2} = \lceil n^{0.999} \rceil$, for a sample size $n = 1000$ (and 2000 replicates), as a function of $k$, from a Burr distribution given by $F(x) = 1 - (1 + x - \rho/\gamma)^{1/\rho}$, $x \geq 0$, with $\gamma = 1$, $\rho = -2$ ($\beta = 1$; $\tau = 0.5$).
Using the asymptotic normality of the geometric-type estimator, we may construct asymptotic confidence intervals for the tail index, with \((1 - \alpha)\)-level, in the usual way:

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

The confidence bounds for the corresponding geometric-type bias-corrected estimators are similar to the previous ones. The 95% confidence bounds for the geometric-type bias-corrected estimators are reported in Tables 1 and 2. For each distribution, we present three values of \(k\) and for the corresponding bias-corrected estimators are reported in Tables 1 and 2. For each distribution, 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 \(\hat{\rho}\) and \(\hat{\beta}\) 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 \(\gamma = 1\) \((\rho = -1, \beta = 1; \tau = 0)\).

<table>
<thead>
<tr>
<th>(k)</th>
<th>(\hat{GT})</th>
<th>(\hat{GT}<em>{\hat{\rho}(k</em>{h1}), \hat{\beta}(k_{h1})})</th>
<th>(\hat{GT}<em>{\hat{\rho}(k</em>{h2}), \hat{\beta}(k_{h2})})</th>
<th>(\hat{GT}<em>{\hat{\rho}(k</em>{h1}), \hat{\beta}(k_{h1})})</th>
<th>(\hat{GT}<em>{\hat{\rho}(k</em>{h2}), \hat{\beta}(k_{h2})})</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>1.25 ± 0.180</td>
<td>1.013 ± 0.162</td>
<td>1.002 ± 0.160</td>
<td>1.018 ± 0.163</td>
<td>1.008 ± 0.161</td>
</tr>
<tr>
<td>500</td>
<td>1.198 ± 0.149</td>
<td>1.011 ± 0.125</td>
<td>0.997 ± 0.124</td>
<td>1.025 ± 0.127</td>
<td>1.013 ± 0.126</td>
</tr>
<tr>
<td>700</td>
<td>1.310 ± 0.137</td>
<td>1.037 ± 0.109</td>
<td>1.020 ± 0.107</td>
<td>1.063 ± 0.111</td>
<td>1.050 ± 0.110</td>
</tr>
</tbody>
</table>

**Table 2.** Confidence bounds \((\alpha = 0.05)\) using the geometric-type estimator and the corresponding bias-corrected estimators, with \(\hat{\rho}\) and \(\hat{\beta}\) 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 \(F(x) = 1 - (1 + x^{-\rho}/\gamma)^{1/\rho}, \ x \geq 0\) with \(\gamma = 1\) and \(\rho = -2\) \((\beta = 1; \tau = 0.5)\).

<table>
<thead>
<tr>
<th>(k)</th>
<th>(\hat{GT})</th>
<th>(\hat{GT}<em>{\hat{\rho}(k</em>{h1}), \hat{\beta}(k_{h1})})</th>
<th>(\hat{GT}<em>{\hat{\rho}(k</em>{h2}), \hat{\beta}(k_{h2})})</th>
<th>(\hat{GT}<em>{\hat{\rho}(k</em>{h1}), \hat{\beta}(k_{h1})})</th>
<th>(\hat{GT}<em>{\hat{\rho}(k</em>{h2}), \hat{\beta}(k_{h2})})</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
<td>1.044 ± 0.167</td>
<td>1.026 ± 0.164</td>
<td>1.022 ± 0.164</td>
<td>1.026 ± 0.164</td>
<td>1.022 ± 0.164</td>
</tr>
<tr>
<td>500</td>
<td>1.055 ± 0.131</td>
<td>1.012 ± 0.125</td>
<td>1.005 ± 0.125</td>
<td>1.013 ± 0.126</td>
<td>1.006 ± 0.125</td>
</tr>
<tr>
<td>700</td>
<td>1.089 ± 0.114</td>
<td>1.011 ± 0.106</td>
<td>1.000 ± 0.105</td>
<td>1.014 ± 0.106</td>
<td>1.004 ± 0.105</td>
</tr>
</tbody>
</table>

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

In order to have an idea of the good behaviour of the geometric-type bias-corrected estimators, we compared them with the corresponding Hill bias-corrected estimators (Caeiro et al. (2005)), given by

\[
\tilde{H}(k) = \hat{H}(k) \left( 1 - \frac{\hat{\beta} \left( \frac{n}{k} \right) \hat{\rho}}{1 - \hat{\rho}} \right)
\]

and

\[
\bar{H}(k) = \hat{H}(k) \exp \left\{ -\frac{\hat{\beta} \left( \frac{n}{k} \right) \hat{\rho}}{1 - \hat{\rho}} \right\},
\]
where \( \hat{\rho} \) and \( \hat{\beta} \) are the estimators of the shape and scale parameters, respectively.

We present here the results for sample sizes \( n = 300 \) and \( n = 1000 \) with a number of replications \( s = 2000 \). For larger number of replications, the simulation results were very similar and so they are not reported here.

From Figures 8, 9, 10 and 11 we observe that, for the considered GPD and Burr distribution, both the geometric-type and the Hill bias-corrected estimators present a good performance. Particularly, we note that, in general, the geometric-type estimator has a better performance for intermediate \( k \) values, while the best behaviour of the Hill estimator is observed for small values of \( k \). Estimates given by the geometric-type estimator are quite far from the target value for small values of \( k \), while, for the Hill estimator, the same is observed for large \( k \) values. Moreover, both Hill corrected estimators present a better performance when the second order parameters are based on level \( k_{h1} \). For the geometric-type estimators, the results are slightly better for the level \( k_{h2} \), with the exception of the Burr distribution with \( n = 300 \).

We may also note that, in general, the corrected geometric-type estimators present, as a function of \( k \), a smoother behaviour than the Hill estimators. In particular, the geometric estimators are more stable and less sensitive to the choice of the number of largest observations taken into account for the estimation than the Hill ones.

Figure 8. Mean estimates (left) and RRMSE (right) of \( \hat{GT} \), \( \hat{GT} \), \( \hat{H} \) and \( \hat{H} \), with \( \hat{\rho} \) and \( \hat{\beta} \) computed at the levels \( k_{h1} = \lfloor n^{0.995} \rfloor \) and \( k_{h2} = \lfloor n^{0.999} \rfloor \), for a sample size \( n = 300 \) (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 9. Mean estimates (left) and RRMSE (right) of $\hat{GT}$, $\hat{GT}$, $\hat{H}$ and $\hat{H}$, with $\hat{\rho}$ and $\hat{\beta}$ computed at the levels $k_{h1} = \lceil n^{0.995} \rceil$ and $k_{h2} = \lceil n^{0.999} \rceil$, 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 10. Mean estimates (left) and RRMSE (right) of $\hat{GT}$, $\hat{GT}$, $\hat{H}$ and $\hat{H}$, with $\hat{\rho}$ and $\hat{\beta}$ computed at the levels $k_{h1} = \lceil n^{0.995} \rceil$ and $k_{h2} = \lceil n^{0.999} \rceil$, for a sample size $n = 300$ (and 2000 replicates), as a function of $k$, from a Burr distribution given by $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x \geq 0$, with $\gamma = 1$ and $\rho = -2$ ($\beta = 1$; $\tau = 0.5$).
Figure 11. Mean estimates (left) and RRMSE (right) of $\hat{GT}, \hat{H}$, with $\hat{\rho}$ and $\hat{\beta}$ 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 $F(x) = 1 - (1 + x^{-\rho/\gamma})^{1/\rho}$, $x \geq 0$, with $\gamma = 1$ and $\rho = -2$ ($\beta = 1; \tau = 0.5$).

As expected, all the estimators present a better behaviour for the largest sample size $n = 1000$.

Overall, we may conclude that the performance of the geometric-type estimator is improved by bias correction and the corresponding corrected versions present a very good behaviour.

5. Proofs

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

Lemma 5.1 (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 \(\left(\frac{0}{0}\right)\) and covariance matrix \(\left(\frac{1}{2} \frac{4}{20}\right)\).

**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 that

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

and so

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

as \(t \to \infty\).

Let us consider the variables presented in Lemma 5.1. Using the probability integral transform, we have

\[
\left( X_1, X_2, \ldots, X_n \right) \overset{D}{=} \left( U(Y_1), U(Y_2), \ldots, U(Y_n) \right), \quad n \geq 1,
\]

and so, we can write, without loss of generality,

\[
(X_{(i,n)}) = \left( U(Y_{(i,n)}) \right).
\]

Then,

\[
M_n^{(1)} = \frac{1}{k} \sum_{i=1}^{k} \log X_{(n-i+1,n)} - \log X_{(n-k,n)}
\]

\[
= \frac{1}{k} \sum_{i=1}^{k} \log U\left( \frac{Y_{(n-i+1,n)}}{Y_{(n-k,n)}} \right) - \log U\left( Y_{(n-k,n)} \right)
\]

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

since

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

which tends to

\[
E\left( \frac{Y_1^\rho - 1}{\rho} \right) = \frac{1}{1 - \rho}.
\]

We have also

\[
M_n^{(2)} = \frac{1}{k} \sum_{i=1}^{k} \left[ \log X_{(n-i+1,n)} - \log X_{(n-k,n)} \right]^2
\]

\[
= 2\gamma^2 + \frac{\gamma^2}{\sqrt{k}} V_n + A\left( Y_{(n-k,n)} \right) \frac{2\gamma (2 - \rho)}{(1 - \rho)^2} + o_p\left( A\left( Y_{(n-k,n)} \right) \right),
\]
since
\[
\frac{1}{k} \sum_{i=1}^{k} \left( \frac{Y_{i}^{n_i+1,n}}{Y_{i}^{n_i,k,n}} \right)^{\rho} - 1 = o \left( \frac{1}{\rho} \log Y_i \right),
\]
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 \( \gamma \), we obtain
\[
\left[ M_n^{(1)} \right]^2 = \gamma^2 + \frac{2\gamma^2}{\sqrt{k}} T_n + d A \left( Y_{n-k,n} \right) + o_{p} \left( A \left( Y_{n-k,n} \right) \right) + o_{p} \left( \frac{1}{k} \right),
\]
Using the above representations we obtain that
\[
\tilde{\gamma}^2 (k) = M_n^{(2)} - \left[ M_n^{(1)} \right]^2
= \gamma^2 + \frac{\gamma^2}{\sqrt{k}} (V_n - 2T_n) + d A \left( Y_{n-k,n} \right) + o_{p} \left( A \left( Y_{n-k,n} \right) \right) + o_{p} \left( \frac{1}{k} \right),
\]
where \( d = 2\gamma/(1-\rho)^2 \).

Since \( A(t) \in RV_{\rho} \), then \( A(tx) = x^\rho A(t)(1 + o(1)) \). Noting now that \((k/n)Y_{n-k,n} = 1 + o_{p}(1)\), we have
\[
A \left( Y_{n-k,n} \right) = A \left( \frac{n}{k} (1 + o(1)) \right)
= A \left( \frac{n}{k} \right) + o_{p} \left( A \left( \frac{n}{k} \right) \right).
\]
Therefore, we may write
\[
\tilde{\gamma}^2 (k) = \gamma^2 + \frac{\gamma^2}{\sqrt{k}} (V_n - 2T_n) + d A \left( \frac{n}{k} \right) + o_{p} \left( A \left( \frac{n}{k} \right) \right) + o_{p} \left( \frac{1}{k} \right).
\]

Considering \( g(x) = \sqrt{x} \) and the Taylor expansion of \( g \left( \tilde{\gamma}^2 (k) \right) \) around \( \gamma^2 \), we obtain
\[
\tilde{\gamma} (k) = \gamma + \frac{1}{2\gamma} \left[ \frac{\gamma^2}{\sqrt{k}} (V_n - 2T_n) + d A \left( \frac{n}{k} \right) + o_{p} \left( A \left( \frac{n}{k} \right) \right) + o_{p} \left( \frac{1}{k} \right) \right]
\]
\[
= \gamma + \frac{\gamma}{2\sqrt{k}} V_n - \frac{\gamma}{\sqrt{k}} T_n + A \left( \frac{n}{k} \right) + o_{p} \left( A \left( \frac{n}{k} \right) \right) + o_{p} \left( \frac{1}{k} \right).
\]

Now, we recall that \( \log \left( Y_{n-i+1,n}/Y_{n-k,n} \right) \) are exponential standard r.v., \( \text{Exp} \(1) \). Using Lemma 5.1, from (18) we can write
\[
\tilde{\gamma} (k) = D \gamma + \frac{\gamma}{2\sqrt{k}} Q_n - \frac{\gamma}{\sqrt{k}} P_n + A \left( \frac{n}{k} \right) + o_{p} \left( A \left( \frac{n}{k} \right) \right) + o_{p} \left( \frac{1}{k} \right).
\]
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 asymptotically normal. This completes the proof. \( \square \)

**Proof of Corollary 2.2.** The result follows from the proof of Theorem 2.1 noting that, since \( (P_n, Q_n) \xrightarrow{D} N \left[ \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & 4 \\ 4 & 20 \end{array} \right) \right] \), then, in particular,

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

\( \square \)

For proving Theorem 2.3 we use the following auxiliary lemma.

**Lemma 5.2** (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.** Recalling that

\[
\hat{G}T(k) = \frac{\bar{\gamma}(k)}{\sqrt{i_n(k)}},
\]

we can write

\[
\sqrt{k} \left( \frac{\hat{G}T(k) - \gamma}{\sqrt{i_n(k)}} \right) = \sqrt{k} (\bar{\gamma}(k) - \gamma) + \sqrt{k} \bar{\gamma}(k) \left( \frac{1}{\sqrt{i_n(k)}} - 1 \right).
\]

As \( \bar{\gamma}(k) \xrightarrow{n \to \infty} \gamma \), from Lemma 5.2 we get

\[
\sqrt{k} \bar{\gamma}(k) \left( \frac{1}{\sqrt{i_n(k)}} - 1 \right) = O_p \left( \frac{\log^2 k}{\sqrt{k}} \right).
\]

So, from the proof of Theorem 2.1 and Lemma 5.2, we have

\[
\hat{G}T(k) = \gamma + \frac{\gamma}{2 \sqrt{k}} V_n - \gamma \frac{T_n}{\sqrt{k}} + A \left( \frac{n}{k} \right) + o_p \left( A \left( \frac{n}{k} \right) \right) + O_p \left( \frac{\log^2 k}{k} \right),
\]

where \( T_n \) and \( V_n \) are the same as in proof of Theorem 2.1. \( \square \)

**Proof of Corollary 2.4.** The result follows from the proof of Theorem 2.3 noting that, as \( (P_n, Q_n) \xrightarrow{D} N \left[ \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \left( \begin{array}{cc} 1 & 4 \\ 4 & 20 \end{array} \right) \right] \), then, in particular,

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

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

$$\hat{GT}(k) = \hat{GT}(k) \left( 1 - \frac{\beta \left( \frac{n}{k} \right)}{(1 - \rho)^2} \right)$$

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

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

With an easy calculation, we also have

$$\hat{GT}(k) = \hat{GT}(k) \exp \left( - \frac{\beta \left( \frac{n}{k} \right)}{(1 - \rho)^2} \right)$$

$$= \hat{GT}(k) \left[ 1 - \frac{A \left( \frac{n}{k} \right)}{\gamma (1 - \rho)^2} + o_p \left( \frac{A \left( \frac{n}{k} \right)}{\gamma (1 - \rho)^2} \right) \right]$$

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

□

Proof of Corollary 3.3. From the proof of Theorem 3.1 we have

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

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

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

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

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

$$V \left[ \sqrt{k} \left( \hat{GT}^* (k) - \gamma \right) \right] = \frac{\gamma^2}{4} V \left( V_n \right) + \gamma^2 V \left( T_n \right) - 2 \text{Cov} \left( \frac{\gamma}{2} V_n, \gamma T_n \right) \xrightarrow{n \to \infty} 2 \gamma^2.$$
Proof of Theorem 3.4. If \( \rho \) and \( \beta \) are estimated consistently, we can use the Taylor’s expansion for bivariate functions and get

\[
\frac{\hat{\beta}}{(1-\hat{\beta})^2} \left( \frac{n}{k} \right)^{\hat{\beta}} = \frac{\beta}{(1-\rho)^2} \left( \frac{n}{k} \right)^{\rho} + \left( \hat{\beta} - \beta \right) \frac{1}{(1-\rho)^2} \left( \frac{n}{k} \right)^{\rho} (1 + o_p(1))
\]

\[
+ \frac{\beta}{(1-\rho)^2} (\hat{\rho} - \rho) \left( \frac{n}{k} \right)^{\rho} \left( \frac{2}{1-\rho} + \log \left( \frac{n}{k} \right) \right) (1 + o_p(1))
\]

\[
= \frac{A(n/k)}{\gamma (1-\rho)^2} \left( \hat{\beta} + \frac{2(\hat{\rho} - \rho)}{1-\rho} + (\hat{\rho} - \rho) \log \left( \frac{n}{k} \right) \right) (1 + o_p(1)),
\]

where \( \hat{\beta} \) and \( \hat{\rho} \) are the estimators of \( \beta \) and \( \rho \), respectively. Therefore we have

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

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

and

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

since \( \hat{\rho} \) and \( \hat{\beta} \) are consistent estimators of \( \rho \) and \( \beta \) computed at a level such that \( \hat{\rho} - \rho = o_p(1/\log n) \).

Proof of Corollary 3.5. The result follows using the same approach as in the proof of Corollary 3.3.

6. Conclusions

Most of the classic estimators of the tail Pareto-type index, such as Hill-type estimators, show a moderate or severe bias, presenting also in some cases an irregular behaviour, especially for not large sample sizes. The geometric-type estimators of the tail index present, in general, a smooth behaviour, but may also suffer from small sample bias.

In this paper we have introduced two bias-corrected estimators of \( \gamma \). These estimators are based on an asymptotic distributional representation of the geometric-type estimator, derived under common second order conditions. We also derived asymptotic distributional representations of the new two bias-corrected geometric-type estimators and established their asymptotic normality. From these results we see that the asymptotic variance is not increased by the bias reduction.

The finite sample behaviour of the new bias-corrected estimators was analised and compared with other estimators through a small scale simulation study, for the GPD and Burr distributions. Beside the original geometric-type estimator, we considered two bias-corrected Hill estimators. The reduced versions of both estimators show a good
performance, but we may notice that, in general, the bias-corrected geometric-type estimators present a smoother path as a function of $k$, being less sensitive to the choice of the $k$ upper observations used in the estimation than the reduced-bias Hill estimators, especially for small samples.

Overall, we may conclude that the empirical results show the improvement obtained with the bias correction and the adequacy of the new estimators for the considered distributions. The better performance of the bias-corrected comparatively to the original geometric-type estimators is reflected in the obtained confidence bounds, with smaller amplitudes and very similar bound values for different choices of the number of o.s. $k$.

Finally, we note that the simulation study may be reproduced and extended using the R-script in the Appendix.

Acknowledgements

ACMF was partially supported by FCT grant SFRH/BPD/66174/2009 and FCT project FAPESP/19805/2014. LC was partially supported by FCT grant SFRH/BD/60642/2009. All three authors were partially supported by FCT project PTDC/MAT/120346/2010, which is funded by national and European structural funds through the programs FEDER and COMPETE. All authors were partially supported by CMUP (UID/MAT/00144/2010), which is funded by FCT (Portugal) with national (MEC) and European structural funds through the programs FEDER, under the partnership agreement PT2020.

The authors warmly thank the referees for their useful suggestions, that greatly improved the presentation of the paper, in particular the suggestions to include a conclusions section and an appendix with the used scripts.

References


value distribution parameters as dynamical indicators of stability, *International Journal of
Bifurcation and Chaos*, 22, 1250276.
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
Haeusler, E. and Teugels, J.L. (1985). On asymptotic normality of Hill’s estimator for the
for singular measures, *Chaos*, 22, 023135.
3*, 119–131.
R Core Team (2015). *R: A language and environment for statistical computing*, Vienna, Aus-
## SIMULATION FUNCTION

Simulation = function (n, s, f) {

# Load libraries
library(actuar) # to qburf function
library(VGAM) # to qgpd function

# Range of order statistics
k1=1 # minimum
k2=n-1 # maxima

d=k2-k1+1

# High level to compute second order parameters
k1difKh1=floor(n^(0.995))
k1difKh2=floor(n^(0.999))

# Vectors of uniform distributed random values
u=rep(1,n)
u=as.vector(u)
u1=rep(1,n)
u1=as.vector(u1)

# Sample vector
x=rep(1,n)
x=as.vector(x)

# Matrices to save the estimated values for each k and for s replicates
Gsvezes=array(0, dim=c(s,d)) # Geometric estimator
GRVKh1svezes=array(0, dim=c(s,d)) # Reduced Bias Geometric estimator (kh1)
GRVEKh1svezes=array(0, dim=c(s,d)) # Reduced Bias Geometric Exponential estimator (kh1)
GRVKh2svezes=array(0, dim=c(s,d)) # Reduced Bias Geometric estimator (kh2)
GRVEKh2svezes=array(0, dim=c(s,d)) # Reduced Bias Geometric Exponential estimator (kh2)
Hsvezes=array(0, dim=c(s,d)) #to Hill estimator
HRVKh1svezes=array(0, dim=c(s,d)) # Reduced Bias Hill estimator (kh1)
HRVEKh1svezes=array(0, dim=c(s,d)) # Reduced Bias Hill Exponential estimator (kh1)
HRVKh2svezes=array(0, dim=c(s,d)) # Reduced Bias Hill estimator (kh2)
HRVEKh2svezes=array(0, dim=c(s,d)) # Reduced Bias Hill Exponential estimator (kh2)

# Process to compute estimated values for s replicated samples
for (j in 1:s) {

  # Seed for random generation
  if (f=="GPD") {
    set.seed(2000+j) #GPD
  } else if (f=="BURR") {
    set.seed(8000+j) #Burr
  }

  # Generate uniform distributed random values
  u[n]=(runif(1, min=0, max=1))^{(1/n)}
  for (i in 2:n) {
    u1[i]=runif(1, min=0, max=1)
    u[n-i+1]=u[n-i+2]^{(1/(n-i+1))}
  }

  # Generate the (Burr or GPD) sample from u
  if (f=="GPD") {
    #GPD parameters: Location=1, Scale=1 and Gamma=1
    x=qgpd(u, 0, 1, 1)
  } else if (f=="BURR") {
    #Burr parameters: Gamma=1 and Rho=-2 --> a=-1/Rho=0.5, b=-Rho/Gamma=2 and s=1
    x=qburr(u, 0.5, 2, 0, 1)
  }

  # True value of the tail index (Gamma)
  truevalue=1

  # Compute the estimated values and save them in the matrix for each replicate
  #Geometric estimator
  gamma.G095=Geometric(x, k=k1:k2)
  Gsvezes[j,]=gamma.G095$estimate
  #Hill estimator
  gamma.H095=Hill1(x, k=k1:k2)
  Hsvezes[j,]=gamma.H095$estimate
  #Reduced Bias Geometric estimator (kh1)
  gamma.GRV095k1difKh1=GeometricredviesK1dif(x, k=k1:k2, k1difKh1, fun=f)
  GRVx1difKh1svezes[j,]=gamma.GRV095k1difKh1$estimate
  #Reduced Bias Exponential Geometric estimator (kh1)
  gamma.GRVE095k1difKh1=GeometricredviesexpK1dif(x, k=k1:k2, k1difKh1, fun=f)
  GRVEk1difKh1svezes[j,]=gamma.GRVE095k1difKh1$estimate
  #Reduced Bias Hill estimator (kh1)
  gamma.HRV095k1difKh1=Hill1redviesK1dif(x, k=k1:k2, k1difKh1, fun=f)
  HRVx1difKh1svezes[j,]=gamma.HRV095k1difKh1$estimate
  #Reduced Bias Geometric estimator (kh2)
  gamma.GRV095k1difKh2=GeometricredviesK1dif(x, k=k1:k2, k1difKh2, fun=f)
  GRVx1difKh2svezes[j,]=gamma.GRV095k1difKh2$estimate
  #Reduced Bias Exponential Geometric estimator (kh2)
  gamma.GRVE095k1difKh2=GeometricredviesexpK1dif(x, k=k1:k2, k1difKh2, fun=f)
  GRVEk1difKh2svezes[j,]=gamma.GRVE095k1difKh2$estimate
  #Reduced Bias Hill estimator (kh2)
  gamma.HRV095k1difKh2=Hill1redviesK1dif(x, k=k1:k2, k1difKh2, fun=f)
  HRVx1difKh2svezes[j,]=gamma.HRV095k1difKh2$estimate
# Reduced Bias Exponential Hill estimator (kh2)

\[
\gamma_{HRVEk1difKh2} = \text{HRVestimexpK1dif}(x, k=k1:k2, \text{k1difKh2, fun=f})
\]

\[
HRVEk1difKh2svezes[i,j] = \gamma_{HRVE095k1difKh2}\text{estimate}
\]

# Vectors of the mean estimates of the s replicates

# Geometric estimator

\[
\text{VectormediaG} = \text{rep}(0,d)
\]

\[
\text{VectormediaG} = \text{as.vector(VectormediaG)}
\]

for (i in 1:k2) {VectormediaG[i]=mean(Gsvezes[,i], na.rm=TRUE)}

# Hill estimator

\[
\text{VectormediaH} = \text{rep}(0,d)
\]

\[
\text{VectormediaH} = \text{as.vector(VectormediaH)}
\]

for (i in 1:k2) {VectormediaH[i]=mean(Hsvezes[,i], na.rm=TRUE)}

# Reduced Bias Geometric estimator (kh1)

\[
\text{VectormediaGRVk1difKh1} = \text{rep}(0,d)
\]

\[
\text{ VectormediaGRVk1difKh1} = \text{as.vector(VectormediaGRVk1difKh1)}
\]

for (i in 1:k2) {VectormediaGRVk1difKh1[i]=mean(GRVk1difKh1svezes[,i], na.rm=TRUE)}

# Reduced Bias Exponential Geometric estimator (kh1)

\[
\text{VectormediaGRVEk1difKh1} = \text{rep}(0,d)
\]

\[
\text{ VectormediaGRVEk1difKh1} = \text{as.vector(VectormediaGRVEk1difKh1)}
\]

for (i in 1:k2) {VectormediaGRVEk1difKh1[i]=mean(GRVEk1difKh1svezes[,i], na.rm=TRUE)}

# Reduced Bias Hill estimator (kh1)

\[
\text{VectormediaHRVk1difKh1} = \text{rep}(0,d)
\]

\[
\text{VectormediaHRVk1difKh1} = \text{as.vector(VectormediaHRVk1difKh1)}
\]

for (i in 1:k2) {VectormediaHRVk1difKh1[i]=mean(HRVk1difKh1svezes[,i], na.rm=TRUE)}

# Reduced Bias Exponential Hill estimator (kh1)

\[
\text{VectormediaHRVEk1difKh1} = \text{rep}(0,d)
\]

\[
\text{VectormediaHRVEk1difKh1} = \text{as.vector(VectormediaHRVEk1difKh1)}
\]

for (i in 1:k2) {VectormediaHRVEk1difKh1[i]=mean(HRVEk1difKh1svezes[,i], na.rm=TRUE)}

# Vectors of the relative mean squared error of the s replicates

# Geometric estimator

\[
\text{rmseGrel} = \text{rep}(0,d)
\]

\[
\text{rmseGrel} = \text{as.vector(rmseGrel)}
\]

for (i in 1:k2) {rmseGrel[i]=sqrt(mean((Gsvezes[,i]-truevalue)^2, na.rm=TRUE))/truevalue}

# Hill estimator

\[
\text{rmseHrel} = \text{rep}(0,d)
\]

\[
\text{rmseHrel} = \text{as.vector(rmseHrel)}
\]
for (i in 1:k2) {rmseHrel[i]=sqrt(mean((Hsvezes[,i]-truevalue)^2, na.rm=TRUE))/truevalue}
#Reduced Bias Geometric estimator (kh1)
rmseGRVrelk1difKh1=rep(0,d)
rmseGRVrelk1difKh1=as.vector(rmseGRVrelk1difKh1)
for (i in 1:k2) {rmseGRVrelk1difKh1[i]=sqrt(mean((GRVk1difKh1svezes[,i]-truevalue)^2, na.rm=TRUE))/truevalue}
#Reduced Bias Exponential Geometric estimator (kh1)
rmseGRVErelk1difKh1=rep(0,d)
rmseGRVErelk1difKh1=as.vector(rmseGRVErelk1difKh1)
for (i in 1:k2) {rmseGRVErelk1difKh1[i]=sqrt(mean((GRVEk1difKh1svezes[,i]-truevalue)^2, na.rm=TRUE))/truevalue}
#Reduced Bias Hill estimator (kh1)
rmseHRVrelk1difKh1=rep(0,d)
rmseHRVrelk1difKh1=as.vector(rmseHRVrelk1difKh1)
for (i in 1:k2) {rmseHRVrelk1difKh1[i]=sqrt(mean((HRVk1difKh1svezes[,i]-truevalue)^2, na.rm=TRUE))/truevalue}
#Reduced Bias Exponential Hill estimator (kh1)
rmseHRVErelk1difKh1=rep(0,d)
rmseHRVErelk1difKh1=as.vector(rmseHRVErelk1difKh1)
for (i in 1:k2) {rmseHRVErelk1difKh1[i]=sqrt(mean((HRVEk1difKh1svezes[,i]-truevalue)^2, na.rm=TRUE))/truevalue}
#Reduced Bias Geometric estimator (kh2)
rmseGRVrelk1difKh2=rep(0,d)
rmseGRVrelk1difKh2=as.vector(rmseGRVrelk1difKh2)
for (i in 1:k2) {rmseGRVrelk1difKh2[i]=sqrt(mean((GRVk1difKh2svezes[,i]-truevalue)^2, na.rm=TRUE))/truevalue}
#Reduced Bias Exponential Geometric estimator (kh2)
rmseGRVErelk1difKh2=rep(0,d)
rmseGRVErelk1difKh2=as.vector(rmseGRVErelk1difKh2)
for (i in 1:k2) {rmseGRVErelk1difKh2[i]=sqrt(mean((GRVEk1difKh2svezes[,i]-truevalue)^2, na.rm=TRUE))/truevalue}
#Reduced Bias Hill estimator (kh2)
rmseHRVrelk1difKh2=rep(0,d)
rmseHRVrelk1difKh2=as.vector(rmseHRVrelk1difKh2)
for (i in 1:k2) {rmseHRVrelk1difKh2[i]=sqrt(mean((HRVk1difKh2svezes[,i]-truevalue)^2, na.rm=TRUE))/truevalue}
#Reduced Bias Exponential Hill estimator (kh2)
rmseHRVErelk1difKh2=rep(0,d)
rmseHRVErelk1difKh2=as.vector(rmseHRVErelk1difKh2)
for (i in 1:k2) {rmseHRVErelk1difKh2[i]=sqrt(mean((HRVEk1difKh2svezes[,i]-truevalue)^2, na.rm=TRUE))/truevalue}

# List of outputs
out <- list(estimateG = VectormediaG, estimateH = VectormediaH, estimateGRVKh1 = VectormediaGRVk1difKh1, estimateGRVEKh1 = VectormediaGRVEk1difKh1, estimateHRVKh1 = VectormediaHRVk1difKh1, estimateHRVEKh1 = VectormediaHRVEk1difKh1, estimateGRVKh2 = VectormediaGRVk1difKh2, estimateGRVEKh2 = VectormediaGRVEk1difKh2, estimateHRVKh2 = VectormediaHRVk1difKh2, estimateHRVEKh2 = VectormediaHRVEk1difKh2, rmseG = rmseGrel1, rmseH = rmseHrel1, rmseGRVKh1 = rmseGRVrelk1difKh1, rmseGRVEKh1 = rmseGRVrelk1difKh1, rmseHRVKh1 = rmseHRVrelk1difKh1, rmseHRVEKh1 = rmseHRVrelk1difKh1, rmseGRVKh2 = rmseGRVrelk1difKh2, rmseGRVEKh2 = rmseGRVrelk1difKh2, rmseHRVKh2 = rmseHRVrelk1difKh2, rmseHRVEKh2 = rmseHRVrelk1difKh2, fun = f)

class <- "EVI"
out <- structure(out, class = class)
### ESTIMATOR FUNCTIONS

#### GEOMETRIC ESTIMATOR

```r
Geometric = function (data, k = k1:k2)
{
  X <- sort(data[data > 0], decreasing = TRUE)
  n <- length(X)
  if (!all(k < n)) {
    k <- k[k < n]
    warning("Only those k for which X_{n-k:n} is positive are retained. ",
            call. = FALSE)
  }

  l <- log(X[1:(max(k) + 1)])
  s1 <- cumsum(l[1:max(k)])
  H <- s1[k]/k - l[k + 1]
  s2 <- cumsum((l[1:max(k)])^2)
  o <- log(n/(1:(max(k)+1)))
  m1 <- cumsum(o[1:max(k)])
  m2 <- cumsum((o[1:max(k)])^2)
  R1 <- m2[k] - (m1^2)[k]/k
  R2 <- s2[k] - (s1^2)[k]/k
  Geometric <- 1/sqrt(R1/R2)

  out <- list(n = n, k = k, estimate = Geometric)
  class <- "EVI"

  out <- structure(out, class = class)
}
```

#### HILL ESTIMATOR

```r
Hill1 = function (data, k = k1:k2)
{
  X <- sort(data[data > 0], decreasing = TRUE)
  n <- length(X)
  if (!all(k < n)) {
    k <- k[k < n]
    warning("Only those k for which X_{n-k:n} is positive are retained. ",
            call. = FALSE)
  }

  s1 <- cumsum(log(X[1:max(k)]))[k]
```
```r
H <- s1/k - log(X[k + 1])
out <- list(n = n, k = k, estimate = H)

class <- "EVI"
out <- structure(out, class = class)
```

## REDUCED BIAS GEOMETRIC ESTIMATOR

```r
GeometricredviesKdif = function (data, k = k1:k2, k1dif, fun = "GPD") {
    X <- sort(data[data > 0], decreasing = TRUE)
    n <- length(X)
    if (!all(k < n)) {
        k <- k[k < n]
        warning("Only those k for which X_{n-k:n} is positive are retained.",
                call. = FALSE)
    }
    # Rho estimates (Fraga Alves et al. (2003))
    # We consider tau=0, tau=0.5 and tau=1
    tau1=0
    numeradorT1=log(M1)-(1/2)*log(M2/2)
    denominadorT1=(1/2)*log(M2/2) - (1/3)*log(M3/6)
    T1=numeradorT1/denominadorT1
    rho1=-abs((3*T1-3)/(T1-3))
    tau2=0.5
    numeradorT2=M1^(tau2)-(M2/2)^(tau2/2)
    denominadorT2=(M2/2)^(tau2/2) - (M3/6)^(tau2/3)
    T2=numeradorT2/denominadorT2
    rho2=-abs((3*T2-3)/(T2-3))
    tau3=1
    numeradorT3=M1^(tau3)-(M2/2)^(tau3/2)
    denominadorT3=(M2/2)^(tau3/2) - (M3/6)^(tau3/3)
    T3=numeradorT3/denominadorT3
    rho3=-abs((3*T3-3)/(T3-3))

    # Beta estimates (Gomes and Martins (2002))
    # at level k1dif
    if (fun=="GPD") {
        RO <- rho1
    } else if (fun=="BURR") {
        RO <- rho2
    }
    num=seq(max(k))
```
m = num^{(-RO[k1dif])}
T1 = \text{cumsum}(m[1:max(k)])[k]
1 <- \log(X[1:(\text{max}(k) + 1)])
s1 <- \text{cumsum}(1[1:max(k)])
H <- s1[k]/k - 1[k + 1]
s2 <- \text{cumsum}((1[1:max(k)])^{-2})
o <- \log(n/(1:(\text{max}(k)+1)))
m1 <- \text{cumsum}(o[1:max(k)])
m2 <- \text{cumsum}((o[1:max(k)])^{-2})
R1 <- m2[k] - (m1^{-2})[k]/k
R2 <- s2[k] - (s1^{-2})[k]/k
aa = (k^{-RO[k1dif]-1}) \ast T1[k]
l1 = 1[1:max(k)]
l2 = 1[2:(\text{max}(k)+1)]
novol = num \ast (l1 - l2)
novos1 = \text{cumsum}(novol[1:max(k)])
ee = novos1/k
novo2l = novol \ast m[k]
U1 = \text{cumsum}(novo2l[1:max(k)])
ff = (k^{-RO[k1dif]-1}) \ast U1[k]
g = m^{-2}
novo3l = novol \ast g[k]
V1 = \text{cumsum}(novo3l[1:max(k)])
uu = (k^{-((2*RO[k1dif]-2)-1)}) \ast V1[k]
B = ((k/n)^{RO[k1dif]}) \ast ((aa*ee-ff)/(aa*ff-uu))
R3 <- sqrt(R2)*(1-(B[k1dif]/(1-RO[k1dif])^2)*((n/k)^{RO[k1dif]}))
Geometricredvies <- R3/sqrt(R1)
out <- \text{list}(n = n, k = k, estimate = Geometricredvies)
class <- "EVI"
out <- \text{structure}(out, class = class)

## REDUCED BIAS EXPONENTIAL GEOMETRIC ESTIMATOR

GeometricredviesexpK1dif = function (data, k = k1:k2, k1dif, fun = "GPD")
{
  X <- sort(data[data > 0], decreasing = TRUE)
n <- length(X)
if (!all(k < n)) {
  k <- k[k < n]
  warning("Only those k for which X_{n-k:n} is positive are retained.",
call. = FALSE)
}
# Rho estimates (Fraga Alves et al. (2003))
# We consider tau=0, tau=0.5 and tau=1
1 <- \log(X[1:(\text{max}(k) + 1)])
s1 <- \text{cumsum}(1[1:max(k)])[k]
s2 <- \text{cumsum}((1[1:max(k)])^{-2})[k]
s3 <- \text{cumsum}((1[1:max(k)])^{-3})[k]
M1 <- s1/k - 1[k + 1]
M2 <- s2/k - 2 * 1[k + 1] * s1/k + (1[k + 1])^{-2}
M3 <- s3/k - (1[k + 1])^3 - 3 * 1[k + 1] * s2/k + 3 * s1/k * (1[k + 1])^2

tau1=0
numeradorT1=log(M1)-(1/2)*log(M2/2)
denominadorT1=(1/2)*log(M2/2) - (1/3)*log(M3/6)
T1=numeradorT1/denominadorT1
rho1=-abs((3*T1-3)/(T1-3))
tau2=0.5
numeradorT2=M1^((tau2)-(M2/2)^((tau2/2))
denominadorT2=(M2/2)^((tau2/2)) - (M3/6)^((tau2/3))
T2=numeradorT2/denominadorT2
rho2=-abs((3*T2-3)/(T2-3))
tau3=1
numeradorT3=M1^((tau3)-(M2/2)^((tau3/2))
denominadorT3=(M2/2)^((tau3/2)) - (M3/6)^((tau3/3))
T3=numeradorT3/denominadorT3
rho3=-abs((3*T3-3)/(T3-3))

# Beta estimates (Gomes and Martins (2002))
# at level k1dif
if (fun=="GPD") {
  RO <- rho1
}
else if (fun=="BURR") {
  RO <- rho2
}

num=seq(max(k))
m=num^(-RO[k1dif])
T1=cumsum(m[1:max(k)])[k]
1 <- 1/2 * log(X[1:(max(k) + 1)])
s1 <- cumsum(1[1:max(k)])
H <- s1[k]/k - 1[k + 1]
s2 <- cumsum(1[1:max(k)]^2)
o <- 1/2 * log(n/(1:(max(k)+1)))
m1 <- cumsum(o[1:max(k)])
m2 <- cumsum(o[1:max(k)]^2)
R1 <- m2[k] - (m1^2)[k]/k
R2 <- s2[k] - (s1^2)[k]/k
aa=(k^((RO[k1dif]-1)))*T1[k]
l1=1[1:max(k)]
l2=1[2:(max(k)+1)]
novol=num*(11-12)
novos1=cumsum(novol[1:max(k)])
ee=novos1/k
novo2l=novol*m[k]
U1=cumsum(novo2l[1:max(k)])
ff=(k^((RO[k1dif]-1)))*U1[k]
g=m^2
novo3l=novol*g[k]
V1=cumsum(novo3l[1:max(k)])
uu=(k^((2*RO[k1dif]-1)))*V1[k]
b=((k/n)^((RO[k1dif])))^((aa*ee-ff)/(aa*ff-uu))

R3 <- sqrt(R2)*exp(-((RO[k1dif]-1)*((n/k)^RO[k1dif]))
Geometricredviesexp <- R3/sqrt(R1)
out <- list(n = n, k = k, estimate = Geometricredviesexp)

class <- "EVI"
out <- structure(out, class = class)
}

## REDUCED BIAS HILL ESTIMATOR (Caeiro et al. (2005))
Hill1redviesK1dif=function (data, k = k1:k2, k1dif, fun = "GPD") {
  X <- sort(data[data > 0], decreasing = TRUE)
  n <- length(X)
  if (!all(k < n)) {
    k <- k[k < n]
    warning("Only those k for which X_{n-k:n} is positive are retained. ",
        call. = FALSE)
  }

  # Rho estimates (Fraga Alves et al. (2003))
  # We consider tau=0, tau=0.5 and tau=1
  l <- log(X[1:(max(k) + 1)])
  s1 <- cumsum(l[1:max(k)])[k]
  s2 <- cumsum((l[1:max(k)])^2)[k]
  s3 <- cumsum((l[1:max(k)])^3)[k]
  M1 <- s1/k - l[k + 1]
  M2 <- s2/k - 2 * l[k + 1] * s1/k + (l[k + 1])^2
  M3 <- s3/k - (l[k + 1])^3 - 3 * l[k + 1] * s2/k + 3 * s1/k * (l[k + 1])^2
  tau1=0
  numeradorT1=log(M1)-(1/2)*log(M2/2)
  denominadorT1=(1/2)*log(M2/2) - (1/3)*log(M3/6)
  T1=numeradorT1/denominadorT1
  rho1=-abs((3*T1-3)/(T1-3))
  tau2=0.5
  numeradorT2=M1^"(tau2)-(M2/2)"^"(tau2/2)
  denominadorT2=(M2/2)^"(tau2/2)" - (M3/6)^"(tau2/3)
  T2=numeradorT2/denominadorT2
  rho2=-abs((3*T2-3)/(T2-3))
  tau3=1
  numeradorT3=M1^"(tau3)-(M2/2)"^"(tau3/2)
  denominadorT3=(M2/2)^"(tau3/2)" - (M3/6)^"(tau3/3)
  T3=numeradorT3/denominadorT3
  rho3=-abs((3*T3-3)/(T3-3))

  # Beta estimates (Gomes and Martins (2002))
  # at level k1dif
  if (fun=="GPD") {
    RO <- rho1
  } else if (fun=="Burr") {
    RO <- rho2
  }

  num=seq(max(k))
m=num^"(-RO[k1dif])"
  T1=cumsum(m[1:max(k)])[k]
1 <- log(X[1:(max(k) + 1)])
s1 <- cumsum(l[1:max(k)])
H <- s1/k - log(k)
s2 <- cumsum((l[1:max(k)])^2)
o <- log(n/(1:(max(k)+1)))
m1 <- cumsum(o[1:max(k)])
m2 <- cumsum((o[1:max(k)])^2)
R1 <- m2[k] - (m1^2)[k]/k
R2 <- s2[k] - (s1^2)[k]/k
aa=(k^(RO[k1dif]-1))*T1[k]
l1=l[1:max(k)]
l2=l[2:(max(k)+1)]
novol=num*(l1-l2)
novos1=cumsum(novol[1:max(k)])
e=nnovos1/k
novol2l=novol*m[k]
U1=cumsum(novol2l[1:max(k)])
ff=(k^(RO[k1dif]-1))*U1[k]
gm=2
novo3l=novol*g[k]
V1=cumsum(novo3l[1:max(k)])

B=((k/n)^(RO[k1dif]))*((aa*ee-ff)/(aa*ff-uu))
s1 <- cumsum(log(X[1:max(k)]))[k]
H <- s1/k - log(X[k + 1])
Hredvies <- (s1/k - log(X[k + 1]))*(1-(B[k1dif]/(1-RO[k1dif]))*((n/k)^(RO[k1dif])))

out <- list(n = n, k = k, estimate = Hredvies)
class <- "EVI"
out <- structure(out, class = class)

## REDUCED BIAS EXPONENTIAL HILL ESTIMATOR (Caeiro et al. (2005))

Hill1redviesexpK1dif = function (data, k = k1:k2, k1dif, fun = "GPD")
{
  X <- sort(data[data > 0], decreasing = TRUE)
  n <- length(X)
  if (!all(k < n)) {
    k <- k[k < n]
    warning("Only those k for which X_{n-k:n} is positive are retained.",
            call. = FALSE)
  }

  # Rho estimates (Fraga Alves et al. (2003))
  # We consider tau=0, tau=0.5 and tau=1
  1 <- log(X[1:(max(k) + 1)])
s1 <- cumsum(l[1:max(k)])[k]
s2 <- cumsum((l[1:max(k)])^2)[k]
s3 <- cumsum((l[1:max(k)])^3)[k]
M1 <- s1/k - 1[k + 1]
M2 <- s2/k - 2 * 1[k + 1] * s1/k + (1[k + 1])^2
M3 <- s3/k - (1[k + 1])^3 - 3 * 1[k + 1] * s2/k + 3 * s1/k * (1[k + 1])^2
}
tau1=0
numeradorT1=log(M1)-(1/2)*log(M2/2)
denominadorT1=(1/2)*log(M2/2) - (1/3)*log(M3/6)
T1=numeradorT1/denominadorT1
rho1=-abs((3*T1-3)/(T1-3))

tau2=0.5
numeradorT2=M1^"(tau2)-(M2/2)^"(tau2/2)
denominadorT2=(M2/2)^"(tau2/2) - (M3/6)^"(tau3/3)
T2=numeradorT2/denominadorT2
rho2=-abs((3*T2-3)/(T2-3))

tau3=1
numeradorT3=M1^"(tau3)-(M2/2)^"(tau3/2)
denominadorT3=(M2/2)^"(tau3/2) - (M3/6)^"(tau3/3)
T3=numeradorT3/denominadorT3
rho3=-abs((3*T3-3)/(T3-3))

# Beta estimates (Gomes and Martins (2002))
# at level k1dif
if (fun="GPD") {
  RO <- rho1
}
else if (fun="BURR") {
  RO <- rho2
}

num=seq(max(k))
m=num^-RO[k1dif]
T1=cumsum(m[1:max(k)])[k]
l <- log(X[1:(max(k) + 1)])
s1 <- cumsum(l[1:max(k)])
H <- s1[k]/k - 1[k + 1]
s2 <- cumsum(l[1:max(k)]^2)
o <- 1log(n/(1:(max(k)+1)))
m1 <- cumsum(o[1:max(k)])
m2 <- cumsum(o[1:max(k)]^2)
R1 <- m1[k] - (m1^"2)[k]/k
R2 <- s2[k] - (s1^"2)[k]/k
aa=(k^"(RO[k1dif]-1)))*T1[k]
l1=1[1:max(k)]
l2=1[2:(max(k)+1)]
novol=num*(11-12)
novos1=cumsum(novol[1:max(k)])
ee=novos1/k

novol2=aa*ee/(aa*ff-uu)

B=((k/n)^"(RO[k1dif]))*(((aa*ee-ff))/(aa*ff-uu))

Hredviesexp <- (s1/k - log(X[k + 1]))*exp(-(B[k1dif]/(1-RO[k1dif])))*((n/k)^"RO[k1dif])"
out <- list(n = n, k = k, estimate = Hredviesexp)

class <- "EVI"
out <- structure(out, class = class)

##############################################################################
##############################
## RUNNING THE SIMULATION ##
##############################

## SETTING INPUTS
# Sample size
n=300

# Number of independent replicates
s=2000

# Distribution function (Uncomment the chosen function)
f="GPD"
#f="BURR"
Sim=Simulation(n, s, f)

## OUTPUTS OF THE SIMULATION

# Geometric-type estimates
Sim$estimateG
Sim$estimateGRVKh1
Sim$estimateGRVKh2
Sim$estimateGRVEKh1
Sim$estimateGRVEKh2

# Hill-type estimates
Sim$estimateH
Sim$estimateHRVKh1
Sim$estimateHRVKh2
Sim$estimateHRVEKh1
Sim$estimateHRVEKh2

# RRMSE for Geometric-type estimates
Sim$rrmseG
Sim$rrmseGRVKh1
Sim$rrmseGRVKh2
Sim$rrmseGRVEKh1
Sim$rrmseGRVEKh2

# RRMSE for Hill-type estimates
Sim$rrmseH
Margarida Brito, Centro de Matemática & Departamento de Matemática, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal
E-mail address: mabrito@fc.up.pt

Laura Cavalcante, Centro de Matemática da Universidade do Porto, Rua do Campo Alegre 687, 4169-007 Porto, Portugal
E-mail address: lauracavalcante@gmail.com

Corresponding Author: Ana Cristina Moreira Freitas, Centro de Matemática & Faculdade de Economia da Universidade do Porto, Rua Dr. Roberto Frias, 4200-464 Porto, Portugal
E-mail address: amoreira@fep.up.pt