



**HAL**  
open science

# On the estimation of the expectation of a heavy-tail distribution

Andrei Novikov

► **To cite this version:**

Andrei Novikov. On the estimation of the expectation of a heavy-tail distribution. 2018. hal-01814734

**HAL Id: hal-01814734**

**<https://hal.science/hal-01814734>**

Preprint submitted on 13 Jun 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the estimation of the expectation of a heavy-tail distribution

Andrei Novikov

Institute of Mathematics, Novosibirsk, Russia.

## Abstract

We use bias-reduced estimators of extreme quantiles from heavy-tail distributions, to build a new estimator of the expectation in the case of infinite second moment. The asymptotic normality of the proposed estimator is established under mild conditions.

**Keywords:** Extreme value theory; Heavy-tail distributions; Hill estimator; Regular variation; Tail index.

## 1 Introduction

Let  $X_1, X_2, \dots$  be independent and identically distributed non-negative random variables with expectation  $\mu < \infty$ , variance  $\sigma^2$  and cumulative distribution function (cdf)  $F$ . Suppose that the tail of  $F$  is regularly varying at infinity with tail index  $(-\alpha) < 0$ , that is

$$\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha}, \quad \text{for any } x > 0, \quad (1)$$

Such heavy-tail distributions include distributions such as Pareto, Burr, Student,  $\alpha$ -stable ( $0 < \alpha < 2$ ), and log-gamma, which are known to be appropriate models for fitting large insurance claims, large fluctuations of prices, log-returns, etc. We are concerned with the construction of a bias-reduced asymptotically normal estimator for the expectation

$$\mu := \int_0^\infty x dF(x),$$

which could be rewritten, in terms of the quantile function (corresponding to the cdf  $F$ )

$$Q(s) := \inf \{x : F(x) \geq s\}, \quad 0 < s < 1,$$

as

$$\mu = \int_0^1 Q(1-s) ds. \quad (2)$$

For a given sample  $X_1, \dots, X_n$ , let

$$Q_n(s) := \inf \{x \in \mathbb{R} : F_n(x) \geq s\}, \quad 0 < s \leq 1,$$

denote the sample quantile function (classical non-parametric estimator of  $Q$ ) associated to the empirical cdf defined on the real line by  $F_n(x) := n^{-1} \sum_{i=1}^n \mathbb{I}(X_i \leq x)$ , with  $\mathbb{I}(\cdot)$  being the indicator function. The natural (unbiased) estimator of  $\mu$  is the sample mean

$$\int_0^1 Q_n(1-s) ds = \frac{1}{n} \sum_{i=1}^n X_i =: \bar{X}_n. \quad (3)$$

From the Central Limit Theorem (CLT), the sequence of r.v.'s  $\{\sqrt{n}(\bar{X}_n - \mu)/\sigma, n \geq 1\}$  converges in distribution to the standard Gaussian r.v., provided that the second-order moment  $\mathbf{E}[X_1^2]$  is finite. This is a very restrictive condition in the context of heavy-tail distributions as the following considerations show. Assume that the r.v.  $X_1$  follows the Pareto law with index  $\alpha > 0$ , that is,  $1 - F(x) = x^{-\alpha}$  for  $x \geq 1$ . When  $\alpha > 1$ , the expectation  $\mu$  exists, but  $\mathbf{E}[X_1^2]$  is only finite for  $\alpha \geq 2$ . Hence, the range  $\alpha \in (1, 2)$  is not covered by the CLT and thus we need to seek another approach to handle this situation. Making use of Weissman's estimator of high quantiles an alternative estimator for  $\mu$  was proposed and its asymptotic normality was established for any  $\alpha \in (1, 2)$ . Let us define the following estimator for  $Q$  :

$$\hat{Q}_n(1-s) := \begin{cases} Q_n^W(1-s) & \text{for } 0 < s < k/n \\ Q_n(1-s) & \text{for } k/n \leq s < 1, \end{cases}$$

where

$$Q_n^W(1-s) := (k/n)^{1/\hat{\alpha}_n^H} X_{n-k,n} s^{-1/\hat{\alpha}_n^H}, \quad s \downarrow 0 \quad (4)$$

is Weissman's estimator of high quantiles, with

$$\hat{\alpha}_n^H := \left( k^{-1} \sum_{i=1}^k \log X_{n-i+1,n} - \log X_{n-k,n} \right)^{-1}, \quad (5)$$

being the well-known Hill estimator of the tail index  $\alpha$ , and  $X_{1,n} \leq \dots \leq X_{n,n}$  denoting the order statistics pertaining to the sample  $X_1, \dots, X_n$ . The number  $k$  represents the number of upper order statistics used in the computation of  $\hat{\alpha}_n^H$ , it is an integer sequence  $k = k_n$  satisfying

$$1 < k < n, \quad k \rightarrow \infty \text{ and } k/n \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (6)$$

By replacing  $Q(1-s)$  by  $\hat{Q}_n(1-t)$  in formula (2), an alternative estimator for  $\mu$  is as follows:

$$\hat{\mu}_n^P = \hat{\mu}_n^P(k) := \int_0^1 \hat{Q}_n(1-s) ds = \int_0^{k/n} Q_n^W(1-s) ds + \int_{k/n}^1 Q_n(1-s) ds,$$

which, by a straightforward calculation, is equal to

$$\hat{\mu}_n^P := \frac{k}{n} \frac{\hat{\alpha}_n^H}{\hat{\alpha}_n^H - 1} X_{n-k,n} + \frac{1}{n} \sum_{i=k+1}^n X_{n-i+1,n}, \quad (7)$$

provided that  $\hat{\alpha}_n^H > 1$ . Moreover, the same author showed that, under suitable regularity assumptions, for any  $\alpha \in (1, 2)$ ,

$$\frac{\sqrt{n}(\hat{\mu}_n^P - \mu)}{\sqrt{k/n}X_{n-k,n}} \xrightarrow{d} \mathcal{N}(0, \sigma^2(\alpha)), \text{ as } n \rightarrow \infty, \quad (8)$$

where

$$\sigma^2(\alpha) := \alpha / (1 - \alpha)^4 (2 - \alpha).$$

Throughout this paper, the standard notations  $\xrightarrow{p}$ ,  $\xrightarrow{d}$  and  $\stackrel{d}{=}$  respectively stand for convergence in probability, convergence in distribution and equality in distribution, while  $\mathcal{N}(a, b^2)$  denotes the normal distribution with expectation  $a$  and variance  $b^2$ .

Actually, this estimator is defined in the more general situation where the r.v.  $X$  is real (not necessarily non-negative) with lower and upper heavy tails. He simultaneously took into account the regular variations of both tails of  $G$  and the balance condition

$$\lim_{t \rightarrow \infty} (1 - F(t)) / (1 - F(t) - F(-t)) = p \in [0, 1].$$

In this paper, we only consider non-negative r.v.'s. Our motivation comes from the actuarial risk theory where insurance losses are represented by such r.v.'s. In this case,  $\hat{\mu}_n^P$  may be interpreted as an estimator of a risk measure called the net premium. Note that in our case, since r.v.  $X$  is non-negative, we have  $F(-x) = 0$  for  $x \geq 0$ , which yields  $p = 1$  in the above balance condition.

Hill's estimator  $\hat{\alpha}_n^H$  plays a pivotal role in statistical inference on distribution tails. This estimator has been thoroughly studied, improved and even generalized to any real parameter  $\alpha$ . Weak consistency of  $\hat{\alpha}_n^H$  was established assuming only that the underlying cdf  $F$  satisfies condition (1). The asymptotic normality of  $\hat{\alpha}_n^H$  has been established under the following stricter condition that characterizes Hall's model.

$$1 - F(x) = cx^{-\alpha} + dx^{-\beta} + o(x^{-\beta}), \text{ as } x \rightarrow \infty, \quad (9)$$

for some  $c > 0$ ,  $d \neq 0$  and  $\beta > \alpha > 0$ . Note that (9), which is a special case of a more general second-order regular variation condition, is equivalent to

$$Q(1 - s) = c^{1/\alpha} s^{-1/\alpha} \left( 1 + \alpha^{-1} c^{-\beta/\alpha} d s^{\beta/\alpha - 1} + o(1) \right), \text{ as } s \downarrow 0. \quad (10)$$

The constants  $\alpha$  and  $\beta$  are called, respectively, first-order (tail index, shape parameter) and second-order parameters of cdf  $F$ .

Extreme value based estimators essentially rely on the number  $k$  of upper order statistics involved in estimate computation. Hill's estimator has, in general, a substantial variance for small values of  $k$  and a considerable bias for large values of  $k$ . Hence, one has to look for a  $k$  value, denoted by  $k^*$ , that balances between these two vices. The choice of this optimal value  $k^*$  represents

a thorny issue in the process of estimating the tail index and related quantities. To solve this problem, several adaptive procedures are available. A theoretical optimal choice of  $k$  is obtained by minimizing the asymptotic mean squared error (RMSE) of  $\hat{\alpha}_n^H$ . Indeed, under condition (9), we have

$$k^* := \left(2^{-1}\alpha\beta^2(\beta-\alpha)^{-3}d^{-2}c^{2\beta/\alpha}\right)^{\frac{\alpha}{2\beta-\alpha}} n^{\frac{2\beta-2\alpha}{2\beta-\alpha}}. \quad (11)$$

Though Peng's estimator  $\hat{\mu}_n^P$  enjoys the asymptotic normality property, it still has a problem due to the fact that, it is based on Weissman's estimator  $Q_n^W$  known to be largely biased. Fortunately, many estimators with reduced biases are proposed in the literature as an alternative to  $Q_n^W$ .

In this paper, we use the bias-reduced estimator of the high quantile  $Q(1-s)$ , recently proposed by who exploited the censored maximum likelihood (CML) based estimators  $(\hat{\alpha}, \hat{\beta})$  of the couple of regular variation parameters  $(\alpha, \beta)$ . The CML estimators  $(\hat{\alpha}, \hat{\beta})$  are defined as the solution of the two equations (under the constraint  $\beta > \hat{\alpha}_n^H$ )

$$\frac{1}{k} \sum_{i=1}^k \frac{1}{G_i(\alpha, \beta)} = 1 \quad \text{and} \quad \frac{1}{k} \sum_{i=1}^k \frac{1}{G_i(\alpha, \beta)} \log \frac{X_{n-i+1,n}}{X_{n-k,n}} = \beta^{-1}, \quad (12)$$

where

$$G_i(\alpha, \beta) = \frac{\alpha}{\beta} \left(1 + \frac{\alpha\beta}{\alpha-\beta} H(\alpha)\right) \left(\frac{X_{n-i+1,n}}{X_{n-k,n}}\right)^{\beta-\alpha} - \frac{\alpha\beta}{\alpha-\beta} H(\alpha), \quad (13)$$

with

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

The bias-reduced estimators  $Q_n^{LPY}(1-s)$ , of the high quantiles  $Q(1-s)$ , are obtained by substituting  $(\hat{\alpha}, \hat{\beta})$  to  $(\alpha, \beta)$  in (10). That is

$$Q_n^{LPY}(1-s) := \hat{c}^{1/\hat{\alpha}} s^{-1/\hat{\alpha}} \left(1 + \hat{\alpha}^{-1} \hat{c}^{-\hat{\beta}/\hat{\alpha}} \hat{d} s^{\hat{\beta}/\hat{\alpha}-1}\right), \quad s \downarrow 0, \quad (14)$$

where

$$\begin{cases} \hat{c} = \frac{\hat{\alpha}\hat{\beta}}{\hat{\alpha}-\hat{\beta}} \frac{k}{n} X_{n-k,n}^{\hat{\alpha}} \left(\frac{1}{\hat{\beta}} - \frac{1}{k} \sum_{i=1}^k \log \frac{X_{n-i+1,n}}{X_{n-k,n}}\right), \\ \hat{d} = \frac{\hat{\alpha}\hat{\beta}}{\hat{\beta}-\hat{\alpha}} \frac{k}{n} X_{n-k,n}^{\hat{\beta}} \left(\frac{1}{\hat{\alpha}} - \frac{1}{k} \sum_{i=1}^k \log \frac{X_{n-i+1,n}}{X_{n-k,n}}\right). \end{cases} \quad (15)$$

The consistency and asymptotic normality of  $Q_n^{LPY}(1-s)$  are established by the same authors. Now we can define another estimator for the quantile function  $Q$  as follows:

$$\tilde{Q}_n(1-s) = \begin{cases} Q_n^{LPY}(1-s) & \text{for } 0 < s < k/n \\ Q_n(1-s) & \text{for } k/n \leq s < 1. \end{cases}$$

By replacing  $Q$  by  $\tilde{Q}_n$ , in formula (2), we get

$$\hat{\mu}_n = \hat{\mu}_n(k) := \int_0^1 \tilde{Q}_n(1-s) ds = \int_0^{k/n} Q_n^{LPY}(1-s) ds + \int_{k/n}^1 Q_n(1-s) ds. \quad (16)$$

An elementary integral calculation leads to a new bias-reduced estimator for  $\mu$  defined by the following formula:

$$\hat{\mu}_n := (k/n) (n\hat{c}/k)^{1/\hat{\alpha}} \left( \frac{\hat{\alpha}}{\hat{\alpha}-1} + \frac{\hat{d}\hat{c}^{-\hat{\beta}/\hat{\alpha}}(k/n)^{\hat{\beta}/\hat{\alpha}-1}}{\hat{\beta}-1} \right) + \frac{1}{n} \sum_{i=k+1}^n X_{n-i+1,n}, \quad (17)$$

provided that  $\hat{\beta} > \hat{\alpha} > 1$  so that  $\hat{\mu}_n$  be finite.

The rest of the paper is organized as follows. In Section 2, we briefly discuss the third order-condition of regular variation before establishing the asymptotic normality of  $\hat{\mu}_n$ . Some concluding remarks notes made in Section 4. Finally, some of the main results used in Section 3 are gathered in the Appendix.

## 2 Main results

In the theory of extremes, a function, denoted by  $U$  and (sometimes) called tail quantile function, is used quite often. It is defined by

$$U(t) := (1/(1-F))^{-1}(t) = Q(1-1/t), \quad 1 < t < \infty.$$

In terms of this function, Hall's conditions (9) and (10) are equivalent to

$$U(t) = c^{1/\alpha} t^{1/\alpha} \left( 1 + \alpha^{-1} c^{-\beta/\alpha} dt^{1-\beta/\alpha} + o(1) \right), \quad t \rightarrow \infty. \quad (18)$$

This implies that

$$\lim_{t \rightarrow \infty} \frac{\log [U(tx)/U(t)] - \alpha^{-1} \log x}{A_1(t)} = \frac{x^{1-\beta/\alpha} - 1}{1 - \beta/\alpha}, \quad \text{for any } x > 0, \quad (19)$$

where

$$A_1(t) := d\alpha^{-1} (1 - \beta/\alpha) c^{-\beta/\alpha} t^{1-\beta/\alpha}.$$

The function  $A_1(t)$ , which tends to zero as  $t \rightarrow \infty$  (because  $\beta > \alpha$ ), determines the rate of convergence of  $\log [U(tx)/U(t)]$  to its limit  $\alpha^{-1} \log x$ . Relation (19) is known as the second-order condition of regular variation.

Unfortunately, the second-order regular variation is not sufficient to find asymptotic distributions for the estimators defined by the systems (12) and (15). We strengthen it into a condition, called third-order condition of regular variation and given by (20), that specifies the rate of (19).

$$\lim_{t \rightarrow \infty} \frac{\frac{\log [U(tx)/U(t)] - \alpha^{-1} \log x}{A_1(t)} - \frac{x^{1-\beta/\alpha} - 1}{1 - \beta/\alpha}}{A_2(t)} = D(\alpha, \beta, \rho), \quad (20)$$

where  $A_2(t) \rightarrow 0$  as  $t \rightarrow \infty$ , with constant sign near infinity and

$$D(\alpha, \beta, \rho) := \frac{1}{\rho} \left( \frac{x^{1-\beta/\alpha+\rho} - 1}{1 - \beta/\alpha + \rho} - \frac{x^{1-\beta/\alpha} - 1}{1 - \beta/\alpha} \right),$$

with  $\rho$  being a positive constant called third-order parameter. The asymptotic normality of  $\hat{\alpha}$ ,  $\hat{\beta}$  and  $\hat{c}$  was established under the following extract conditions on the sample fraction  $k$ , as  $n \rightarrow \infty$ ,

$$(i) \sqrt{k} |A_1(n/k)| \rightarrow \infty, \quad (ii) \sqrt{k} A_1^2(n/k) \rightarrow 0, \quad (iii) \sqrt{k} A_1(n/k) A_2(n/k) \rightarrow 0. \quad (21)$$

As for  $\hat{d}$ , it is asymptotically normal under the assumption  $(\sqrt{k} |A_1(n/k)|) / \log(n/k) \rightarrow \infty$  added to (ii) and (iii).

**Example 1** Consider the Fréchet cdf with shape parameter  $\alpha > 0$

$$F(x) = \exp(-x^{-\alpha}), \quad x > 0. \quad (22)$$

The corresponding tail quantile function is defined by  $U(t) = (-\log(1 - 1/t))^{-1/\alpha}$ , for  $t > 1$ . Applying Taylor's expansion (to the third order) to  $U$  and identifying with (18), yield  $\beta = 2\alpha$ ,  $c = 1$  and  $d = -1/2$ . The condition (20) holds for  $A_1(t) = t^{-1}/2\alpha$ ,  $A_2(t) = (\alpha - 3)t^{-2}/12\alpha^2$  and  $\rho = 3\alpha$ .

Note that, from a theoretical point of view, assumptions (6) and (21) are realistic, as the following example shows. Indeed, let us choose  $k = [n^\epsilon]$ ,  $0 < \epsilon < 1$ , then it easy to verify that these assumptions hold for any  $2/3 < \epsilon < 4/5$ . The notation  $[\cdot]$  stands for the integer part of real numbers.

Our main result, namely the asymptotic normality of the bias-reduced estimator  $\hat{\mu}_n$ , is formulated in the last of the following four theorems. In Theorem 1, we give an approximation of  $\hat{\alpha}$  in terms of Brownian bridges, which leads to its asymptotic normality stated in Theorem 2. We do the same thing to  $\hat{\mu}_n$  in Theorem 3. But, this does not meet our needs to achieve the major object of this paper. Then, we need to approximate both  $\hat{\alpha}$  and  $\hat{\mu}_n$  by linear functional of the same sequence of standard Brownian bridges  $B_n(s)$ .

**Theorem 1** Assume that the third order condition (20) holds with  $\beta/\alpha =: \lambda > 1$  and let  $k = k_n$  be an integer sequence satisfying (6) and (21). Then there exists a sequence of Brownian bridges  $\{B_n(s), 0 \leq s \leq 1\}$  such that

$$\sqrt{k}(\hat{\alpha} - \alpha) = \alpha(\eta_1 W_{1n} + \eta_2 W_{2n} + \eta_3 W_{3n}) + o_p(1), \quad \text{as } n \rightarrow \infty$$

where  $W_{1n}$ ,  $W_{2n}$  and  $W_{3n}$  are sequences of centered Gaussian r.v.'s defined by

$$\begin{aligned} W_{1n} &:= \sqrt{n/k} B_n(1 - k/n) - \sqrt{n/k} \int_0^1 s^{-1} B_n(1 - ks/n) ds, \\ W_{2n} &:= (\lambda^{-1} - 1) \sqrt{n/k} B_n(1 - k/n) + (\lambda - 1) \sqrt{n/k} \int_0^1 s^{\lambda-2} B_n(1 - ks/n) ds, \\ W_{3n} &:= (1 - \lambda) \sqrt{n/k} \int_0^1 s^{\lambda-2} (\log s) B_n(1 - ks/n) ds \\ &\quad + \lambda^{-2} \sqrt{n/k} B_n(1 - k/n) - \sqrt{n/k} \int_0^1 s^{\lambda-2} B_n(1 - ks/n) ds, \end{aligned}$$

and

$$\eta_1 := \frac{\lambda^4}{(\lambda - 1)^4}, \quad \eta_2 := \frac{\lambda^2(2\lambda - 1)(3\lambda - 1)}{(\lambda - 1)^5}, \quad \eta_3 := \frac{\lambda^3(2\lambda - 1)^2}{(\lambda - 1)^4}.$$

**Theorem 2** Under the assumptions of Theorem 1, we have

$$\sqrt{k}(\hat{\alpha} - \alpha) \xrightarrow{d} \mathcal{N}\left(0, \alpha^2 \beta^4 / (\alpha - \beta)^4\right), \quad \text{as } n \rightarrow \infty. \quad (23)$$

**Theorem 3** Under the assumptions of Theorem 1, we have, as  $n \rightarrow \infty$

$$\frac{\sqrt{n}}{\sqrt{k/n}(nc/k)^{1/\alpha}} \{\hat{\mu}_n - \mu\} = -\frac{\alpha}{(\alpha - 1)^2} \{\eta_1 W_{1n} + \eta_2 W_{2n} + \eta_3 W_{3n}\} + W_{4n} + o_p(1),$$

where  $W_{1n}$ ,  $W_{2n}$  and  $W_{3n}$  are those of Theorem 1 and

$$W_{4n} := -\frac{\int_{k/n}^1 B_n(1 - s) dQ(1 - s)}{\sqrt{k/n}(nc/k)^{1/\alpha}}.$$

**Theorem 4** Under the assumptions of Theorem 1, we have

$$\frac{\sqrt{n}}{\sqrt{k/n}(nc/k)^{1/\alpha}} \{\hat{\mu}_n - \mu\} \xrightarrow{d} \mathcal{N}\left(0, \sigma^2(\alpha, \beta)\right), \quad \text{as } n \rightarrow \infty, \quad (24)$$

where

$$\sigma^2(\alpha, \beta) := \frac{\alpha^2 \beta^4}{(\alpha - 1)^4 (\alpha - \beta)^4} + \frac{2}{2 - \alpha} + \frac{2\alpha \beta^2}{(\alpha - 1)^2 (\alpha - \beta)^2}. \quad (25)$$

The following corollary to Theorem 4 provides a straightforward practical way to build confidence intervals for  $\mu$ .

**Corollary 1** Under the assumptions of Theorem 1, we have

$$\frac{\sqrt{n}}{\sqrt{k/n\sigma}(\hat{\alpha}, \hat{\beta})(n\hat{c}/k)^{1/\hat{\alpha}}} \{\hat{\mu}_n - \mu\} \xrightarrow{d} \mathcal{N}(0, 1), \quad \text{as } n \rightarrow \infty,$$

where  $\hat{\alpha}$ ,  $\hat{\beta}$  and  $\hat{c}$  are the estimates of  $\alpha$ ,  $\beta$  and  $c$  given in (12) and (15) respectively.

### 3 Proofs

#### 3.1 Proof of Theorem 1.

First recall that  $\lambda = \beta/\alpha > 1$ . Then, we have, as  $n \rightarrow \infty$

$$\hat{\alpha} - \alpha = \alpha \{ \eta_1 (S_1 - 1) + \eta_2 (S_2 - \lambda^{-1}) + \eta_3 (S_3 - \lambda^{-2}) \} + o_p \left( k^{-1/2} \right), \quad (26)$$

where

$$S_1 := \frac{1}{k} \sum_{i=1}^k \log \frac{Y_{n-i+1,n}}{Y_{n-k,n}}, \quad S_2 := \frac{1}{k} \sum_{i=1}^k \left( \frac{Y_{n-i+1,n}}{Y_{n-k,n}} \right)^{1-\lambda}, \quad (27)$$

and

$$S_3 := \frac{1}{k} \sum_{i=1}^k \left( \frac{Y_{n-i+1,n}}{Y_{n-k,n}} \right)^{1-\lambda} \log \frac{Y_{n-i+1,n}}{Y_{n-k,n}}, \quad (28)$$

with  $Y_{1,n} \leq \dots \leq Y_{n,n}$  being the order statistics pertaining to a sample  $Y_1, \dots, Y_n$  of independent and identically distributed non-negative random variables, defined on the same probability space as the  $X'_i$ 's, with cdf

$$G(y) = 1 - y^{-1}, \quad \text{for } y > 1. \quad (29)$$

A probability space  $(\Omega, \mathcal{A}, \mathbf{P})$  may be constructed, carrying an infinite sequence  $U_1, U_2, \dots$  of independent  $(0, 1)$ -uniform r.v.'s and a sequence of Brownian bridges  $\{B_n(s), 0 \leq s \leq 1\}$ ,  $n = 1, 2, \dots$ , having, amongst others, the property stated in Lemma 1. Let  $U_{1,n} \leq \dots \leq U_{n,n}$  denote the order statistics pertaining to  $U_1, \dots, U_n$  and define the empirical quantile function  $V_n(s)$  as

$$V_n(s) = U_{i,n} \quad \text{for } (i-1)/n < s \leq i/n, \quad i = 1, \dots, n, \quad \text{and } V_n(0) = U_{1,n}.$$

**Lemma 1** *On this probability space, for every  $0 \leq \tau < 1/2$ , we have, as  $n \rightarrow \infty$*

$$\sup_{1/n \leq s \leq 1-1/n} \frac{|\sqrt{n}(s - V_n(s)) - B_n(s)|}{(s(1-s))^{1/2-\tau}} = O_p(n^{-\tau}). \quad (30)$$

Without loss of generality, we assume that

$$Y_i = G^{-1}(U_i) = (1 - U_i)^{-1}, \quad i = 1, \dots, n,$$

and

$$Y_{i,n} = G^{-1}(U_{i,n}) = (1 - U_{i,n})^{-1}, \quad i = 1, \dots, n,$$

where  $G^{-1}$  denotes the quantile function pertaining to cdf  $G$  given by formula (29). Then, this allows us to write

$$Y_{n-i+1,n} = (1 - V_n(1-s))^{-1}, \quad \text{for } \frac{i-1}{n} < s \leq \frac{i}{n}, \quad i = 1, \dots, n.$$

Making use of the previous representation of the order statistics  $Y_{n-i+1,n}$ , we may rewrite the three statistics in (27) and (28) into

$$\begin{aligned} S_1 &= \frac{n}{k} \int_0^{k/n} \log \left( \frac{1 - V_n(1-s)}{1 - U_{n-k,n}} \right)^{-1} ds, \\ S_2 &= \frac{n}{k} \int_0^{k/n} \left( \frac{1 - V_n(1-s)}{1 - U_{n-k,n}} \right)^{-1+\lambda} ds, \end{aligned}$$

and

$$S_3 = \frac{n}{k} \int_0^{k/n} \left( \frac{1 - V_n(1-s)}{1 - U_{n-k,n}} \right)^{-1+\lambda} \log \left( \frac{1 - V_n(1-s)}{1 - U_{n-k,n}} \right)^{-1} ds.$$

Next, we show that, as  $n \rightarrow \infty$

$$\begin{aligned} \sqrt{k} (S_1 - 1) &= W_{1n} + o_p(1), \\ \sqrt{k} (S_2 - \lambda^{-1}) &= W_{2n} + o_p(1), \end{aligned}$$

and

$$\sqrt{k} (S_3 - \lambda^{-2}) = W_{3n} + o_p(1),$$

where  $W_{1n}, W_{2n}$  and  $W_{3n}$  are the Gaussian r.v.'s defined in Theorem 1. We will only consider the asymptotic distribution of  $S_3$ . The proofs for  $S_1$  and  $S_2$  use similar arguments. By letting  $f(x) = x^{\lambda-1} \log x$ , the statistic  $S_3$  becomes

$$S_3 = -(n/k) \int_0^{k/n} f \left( \frac{1 - V_n(1-s)}{1 - U_{n-k,n}} \right) ds.$$

An application of standard calculus gives  $\int_0^1 f(s) ds = -\lambda^{-2}$ . Therefore

$$S_3 - \lambda^{-2} = -(n/k) \int_0^{k/n} \left[ f \left( \frac{1 - V_n(1-s)}{1 - U_{n-k,n}} \right) - f \left( \frac{s}{k/n} \right) \right] ds.$$

We divide the integral above in two parts, then we study the asymptotic behavior of each integral. Observe that

$$\begin{aligned} S_3 - \lambda^{-2} &= -(n/k) \int_0^{1/n} \left[ f \left( \frac{1 - V_n(1-s)}{1 - U_{n-k,n}} \right) - f \left( \frac{s}{k/n} \right) \right] ds \\ &\quad - (n/k) \int_{1/n}^{k/n} \left[ f \left( \frac{1 - V_n(1-s)}{1 - U_{n-k,n}} \right) - f \left( \frac{s}{k/n} \right) \right] ds \\ &=: -\Delta_n - \Omega_n. \end{aligned}$$

Next, we show that  $\sqrt{k}\Delta_n$  converges to 0 in probability. Indeed, we have  $1 - V_n(1-s) = 1 - U_{n,n}$ , for  $0 < s \leq 1/n$ , it follows that

$$\begin{aligned} \Delta_n &= (n/k) \int_0^{1/n} \left[ f \left( \frac{1 - U_{n,n}}{1 - U_{n-k,n}} \right) - f \left( \frac{s}{k/n} \right) \right] ds \\ &= k^{-1} f \left( \frac{1 - U_{n,n}}{1 - U_{n-k,n}} \right) - \int_0^{1/k} f(s) ds. \end{aligned}$$

An elementary calculation gives  $\int_0^{1/k} f(s) ds = \lambda^{-1} k^{-\lambda} (\log k^{-1} - \lambda^{-1})$ , and we have  $n(1 - U_{n,n-k})/k \xrightarrow{P} 1$ , as  $n \rightarrow \infty$ , therefore

$$\Delta_n = \{1 + o_p(1)\} k^{-\lambda} \log k^{-1} - \lambda^{-1} k^{-\lambda} (\log k^{-1} - \lambda^{-1}).$$

Since  $\lambda > 1$  and  $k \rightarrow \infty$ , then  $k^{-\lambda+1/2} \rightarrow 0$  and  $k^{-\lambda+1/2} \log k^{-1} \rightarrow 0$ , it follows that  $\sqrt{k}\Delta_n \xrightarrow{P} 0$  as  $n \rightarrow \infty$ . Consider now the second term  $\Omega_n$  which may be rewritten into

$$\begin{aligned} \Omega_n &= (n/k) \int_{1/n}^{k/n} \left[ f\left(\frac{1 - V_n(1-s)}{1 - U_{n-k,n}}\right) - f\left(\frac{s}{1 - U_{n-k,n}}\right) \right] ds \\ &\quad + (n/k) \int_{1/n}^{k/n} \left[ f\left(\frac{s}{1 - U_{n-k,n}}\right) - f\left(\frac{s}{k/n}\right) \right] ds \\ &=: \Omega_{n1} + \Omega_{n2}. \end{aligned}$$

Making use of Taylor's expansion of  $f$ , we get

$$f\left(\frac{1 - V_n(1-s)}{1 - U_{n-k,n}}\right) - f\left(\frac{s}{1 - U_{n-k,n}}\right) = f'\left(\frac{\varphi_n(s)}{1 - U_{n-k,n}}\right) \left(\frac{1 - V_n(1-s)}{1 - U_{n-k,n}} - \frac{s}{1 - U_{n-k,n}}\right),$$

and

$$f\left(\frac{s}{1 - U_{n-k,n}}\right) - f\left(\frac{s}{k/n}\right) = f'(s\psi_n) \left(\frac{s}{1 - U_{n-k,n}} - \frac{s}{k/n}\right),$$

where

$$\min\{1 - V_n(1-s), s\} < \varphi_n(s) < \max\{1 - V_n(1-s), s\} \quad (31)$$

and

$$\min\left\{\frac{1}{1 - U_{n-k,n}}, \frac{1}{k/n}\right\} < \psi_n < \max\left\{\frac{1}{1 - U_{n-k,n}}, \frac{1}{k/n}\right\}. \quad (32)$$

Observe now that  $\Omega_{n1}$  and  $\Omega_{n2}$  may be rewritten into

$$\Omega_{n1} = (n/k) \int_{1/n}^{k/n} f'\left(\frac{s}{1 - U_{n-k,n}}\right) \left[\frac{1 - V_n(1-s)}{1 - U_{n-k,n}} - \frac{s}{1 - U_{n-k,n}}\right] ds + \Omega_{n1}^*,$$

and

$$\Omega_{n2} = \{1 + o_p(1)\} (n/k) \int_{1/n}^{k/n} f'\left(\frac{s}{k/n}\right) \left[\frac{s}{1 - U_{n-k,n}} - \frac{s}{k/n}\right] ds + \Omega_{n2}^*,$$

where

$$\begin{aligned} \Omega_{n1}^* &:= (n/k) \int_{1/n}^{k/n} \left[ f'\left(\frac{\varphi_n(s)}{1 - U_{n-k,n}}\right) - f'\left(\frac{s}{1 - U_{n-k,n}}\right) \right] \\ &\quad \times \left[ \frac{1 - V_n(1-s)}{1 - U_{n-k,n}} - \frac{s}{1 - U_{n-k,n}} \right] ds, \end{aligned}$$

and

$$\begin{aligned}\Omega_{n2}^* &:= \{1 + o_p(1)\} (n/k) \int_{1/n}^{k/n} \left[ f' \left( \frac{s\psi_n}{k/n} \right) - f' \left( \frac{s}{k/n} \right) \right] \\ &\quad \times \left[ \frac{s}{1 - U_{n-k,n}} - \frac{s}{k/n} \right] ds.\end{aligned}$$

From Lemma 3 (see the Appendix), both  $\sqrt{k}\Omega_{n1}^*$  and  $\sqrt{k}\Omega_{n2}^*$  converge to 0 in probability. Since  $n(1 - U_{n-k,n})/k \xrightarrow{P} 1$ , then

$$\Omega_{n1} = \{1 + o_p(1)\} \int_{1/n}^{k/n} f' \left( \frac{s}{k/n} \right) [1 - s - V_n(1 - s)] ds + o_p(k^{-1/2}),$$

and

$$\Omega_{n2} = -\{1 + o_p(1)\} \frac{k/n - (1 - U_{n-k,n})}{(k/n)^2} \int_{1/n}^{k/n} \left( \frac{s}{k/n} \right) f' \left( \frac{s}{k/n} \right) ds + o_p(k^{-1/2}).$$

The derivative of function  $f$  equals  $f'(x) = (\lambda - 1)x^{\lambda-2} \log x + x^{\lambda-2}$ , then

$$\begin{aligned}\Omega_{n1} &= (\lambda - 1) (n/k) \int_{1/k}^1 t^{\lambda-2} (\log t) [1 - V_n(1 - kt/n) - kt/n] dt \\ &\quad + (n/k) \int_{1/k}^1 t^{\lambda-2} [1 - V_n(1 - kt/n) - kt/n] dt + o_p(k^{-1/2}),\end{aligned}$$

and

$$\begin{aligned}\Omega_{n2} &= (\lambda - 1) (n/k) [k/n - (1 - U_{n-k,n})] \int_{1/k}^1 t^{\lambda-1} \log t dt \\ &\quad + (n/k) [k/n - (1 - U_{n-k,n})] \int_{1/k}^1 t^{\lambda-1} dt + o_p(k^{-1/2}) \\ &= \lambda^{-2} (n/k) [k/n - (1 - U_{n-k,n})] + o_p(k^{-1/2}).\end{aligned}$$

Fix  $0 < \tau < 1/2$ , then using approximation (30), in Lemma 1, yields

$$\begin{aligned}\sqrt{k}\Omega_{n1} &= (\lambda - 1) \sqrt{n/k} \int_{1/k}^1 t^{\lambda-2} (\log t) B_n(1 - kt/n) dt \\ &\quad + \sqrt{n/k} \int_{1/k}^1 t^{\lambda-2} B_n(1 - kt/n) dt + \sqrt{k}\tilde{\Omega}_{n1}(\tau) + o_p(1),\end{aligned}$$

and

$$\sqrt{k}\Omega_{n2} = -\lambda^{-2} \sqrt{n/k} B_n(1 - k/n) + \sqrt{k}\tilde{\Omega}_{n2}(\tau) + o_p(1),$$

where

$$\begin{aligned}\sqrt{k}\tilde{\Omega}_{n1}(\tau) &= (\lambda - 1) O_p(n^{-\tau}) (k/n)^{1/2-\tau} (n/k)^{1/2} \int_0^1 t^{\lambda-2+(1/2-\tau)} |\log t| dt \\ &\quad + O_p(n^{-\tau}) (k/n)^{1/2-\tau} \sqrt{n/k} \int_0^1 t^{\lambda-2+(1/2-\tau)} dt,\end{aligned}$$

and

$$\sqrt{k}\tilde{\Omega}_{n2}(\tau) = \lambda^{-2}O_p(n^{-\tau})\sqrt{n/k}(k/n)^{1/2-\tau}.$$

For  $\lambda > 1$ ,  $\int_0^1 t^{\lambda-2+(1/2-\tau)} |\log t| dt = (\lambda - 1/2 - \tau)^{-2}$  and  $\int_0^1 t^{\lambda-2+(1/2-\tau)} dt = (\lambda - 1/2 - \tau)^{-1}$  are finite integrals. Then both quantities  $\sqrt{k}\tilde{\Omega}_{n1}$  and  $\sqrt{k}\tilde{\Omega}_{n2}$  are equal to  $O_p(k^{-\tau})$  for all large  $n$ , which tends in probability to 0 as  $n \rightarrow \infty$ . Recall that up to now we have showed that

$$\begin{aligned}\sqrt{k}\Omega_{n1} &= (\lambda - 1)\sqrt{n/k}\int_{1/k}^1 t^{\lambda-2}(\log t)B_n(1 - kt/n)dt \\ &\quad + \sqrt{n/k}\int_{1/k}^1 t^{\lambda-2}B_n(1 - kt/n)dt + o_p(1),\end{aligned}$$

and

$$\sqrt{k}\Omega_{n2} = \lambda^{-2}\sqrt{n/k}B_n(1 - k/n) + o_p(1).$$

It remains to prove that

$$\begin{aligned}I_n &:= (\lambda - 1)\sqrt{n/k}\int_0^{1/k} t^{\lambda-2}(\log t)B_n(1 - kt/n)dt \\ &\quad + \sqrt{n/k}\int_0^{1/k} t^{\lambda-2}B_n(1 - kt/n)dt,\end{aligned}$$

converges, in probability, to 0. Indeed, since  $E|B_n(1 - ks/n)| \leq \sqrt{ks/n}$ , then

$$\mathbf{E}|I_n| \leq (\lambda - 1)\int_0^{1/k} t^{\lambda-2+1/2}(|\log t| + 1)dt.$$

Since

$$\int_0^{1/k} t^{\lambda-2+1/2}(|\log t| + 1)dt = \frac{2}{(2\lambda - 1)^2}k^{-\lambda+\frac{1}{2}}(2\lambda - \log k + 2\lambda \log k + 1),$$

which tends to 0 as  $n \rightarrow \infty$ , then  $I_n$  converges to 0 in probability. This completes the proof of Theorem 1.  $\square$

### 3.2 Proof of Theorem 2.

To establish the asymptotic normality of  $\hat{\alpha}$ , given in (23), we proceed by similar arguments as for  $\hat{\mu}_n$  in the proof of Theorem 4.  $\square$

### 3.3 Proof of Theorem 3

Let us divide the integral (2), in two parts, as follows:

$$\mu = \mu_{1,n}(k) + \mu_{2,n}(k),$$

where

$$\mu_{1,n}(k) := \int_0^{k/n} Q(1-s) ds \text{ and } \mu_{2,n}(k) := \int_{k/n}^1 Q(1-s) ds.$$

Recall that, in Section 1 formula (16), we have defined estimator  $\hat{\mu}_n$  of  $\mu$  by

$$\hat{\mu}_n = \hat{\mu}_{1,n}(k) + \hat{\mu}_{2,n}(k),$$

where

$$\hat{\mu}_{1,n}(k) := (k/n)(n\hat{c}/k)^{1/\hat{\alpha}} \left\{ \frac{\hat{\alpha}}{\hat{\alpha}-1} + \frac{\hat{d}\hat{c}^{-\hat{\beta}/\hat{\alpha}}(k/n)^{\hat{\beta}/\hat{\alpha}-1}}{\hat{\beta}-1} \right\}$$

and

$$\hat{\mu}_{2,n}(k) := \frac{1}{n} \sum_{i=k+1}^n X_{n-i+1,n}.$$

To simplify notations, let us set

$$Z_{ni} := \frac{\sqrt{n}}{\sqrt{k/n}(nc/k)^{1/\alpha}} \{\hat{\mu}_{i,n}(k) - \mu_{i,n}(k)\}, \quad i = 1, 2. \quad (33)$$

First, we consider  $Z_{n1}$ . It is easy to verify that, as  $n \rightarrow \infty$

$$\mu_{1,n}(k) = \{1 + o_p(1)\} \frac{k}{n} (nc/k)^{1/\alpha} \frac{\alpha}{\alpha-1},$$

and, under the condition (6), we have

$$\hat{\mu}_{1,n}(k) = \{1 + o_p(1)\} \frac{k}{n} \frac{\hat{\alpha}}{\hat{\alpha}-1} (n\hat{c}/k)^{1/\hat{\alpha}}.$$

It follows that

$$\hat{\mu}_{1,n}(k) - \mu_{1,n}(k) = \{1 + o_p(1)\} \frac{k}{n} \left\{ \frac{\hat{\alpha}}{\hat{\alpha}-1} (n\hat{c}/k)^{1/\hat{\alpha}} - \frac{\alpha}{\alpha-1} (nc/k)^{1/\alpha} \right\}.$$

Let us write  $Z_{n1} = T_{1n} + T_{2n}$ , where

$$T_{1n} := \{1 + o_p(1)\} \sqrt{k} \left\{ \frac{\hat{\alpha}}{\hat{\alpha}-1} - \frac{\alpha}{\alpha-1} \right\},$$

and

$$T_{2n} := \{1 + o_p(1)\} \sqrt{k} \left\{ \frac{(n\hat{c}/k)^{1/\hat{\alpha}}}{(nc/k)^{1/\alpha}} - 1 \right\}.$$

We begin by showing that  $T_{2n} \xrightarrow{P} 0$ , as  $n \rightarrow \infty$ . First observe that  $T_{2n}$  may be rewritten into

$$T_{2n} = \{1 + o_p(1)\} \sqrt{k} \left\{ (nc/k)^{1/\hat{\alpha}-1/\alpha} (\hat{c}/c)^{1/\hat{\alpha}-1/\alpha} - 1 \right\}.$$

Assumptions (i) and (ii) of Theorem 1 imply that  $k^{1/2}/\log(n/k) \rightarrow \infty$ . Also, the asymptotic normality of  $\hat{\alpha}$  gives  $\hat{\alpha} - \alpha = O_p(k^{-1/2})$ . Therefore  $(1/\hat{\alpha} - 1/\alpha) \log(nc/k) \xrightarrow{P} 0$ , this implies that  $(nc/k)^{1/\hat{\alpha}-1/\alpha} \xrightarrow{P} 1$ , as  $n \rightarrow \infty$ . On the other hand, we have

$$\hat{c}/c - 1 = \alpha^{-1} (\hat{\alpha} - \alpha) \log \frac{n}{k} + o_p\left(k^{-1/2} \log \frac{n}{k}\right).$$

Since  $\widehat{c}$  is a consistent estimator of  $c$ , then Taylor's expansion gives

$$(\widehat{c}/c)^{1/\widehat{\alpha}-1/\alpha} - 1 = \alpha^{-1} (1 + o_p(1)) (\widehat{\alpha} - \alpha) (\widehat{c}/c - 1), \text{ as } n \rightarrow \infty.$$

It suffices now to show that  $\sqrt{k} \left( (\widehat{c}/c)^{1/\widehat{\alpha}-1/\alpha} - 1 \right)$  converges to 0 in probability. Indeed, again by using the fact that  $\widehat{\alpha} - \alpha = O_p(k^{-1/2})$ , yields

$$\sqrt{k} \left( (\widehat{c}/c)^{1/\widehat{\alpha}-1/\alpha} - 1 \right) = O_p(1) \left( k^{-1/2} \log \frac{n}{k} + o_p \left( \frac{\log \frac{n}{k}}{\sqrt{k}} \right) \right),$$

which tends in probability to 0, because we already have  $\sqrt{k}/\log(n/k) \rightarrow \infty$ . Now, we consider the term  $T_{1n}$ . Since  $\widehat{\alpha}$  is a consistent estimator of  $\alpha$ , then it is easy to show that

$$T_{1n} = -\frac{1 + o_p(1)}{(\alpha - 1)^2} \sqrt{k} (\widehat{\alpha} - \alpha), \text{ as } n \rightarrow \infty.$$

From 1, we infer that

$$T_{1n} = -\frac{(1 + o_p(1))\alpha}{(\alpha - 1)^2} (\eta_1 W_{1n} + \eta_2 W_{2n} + \eta_3 W_{3n}), \text{ as } n \rightarrow \infty.$$

It follows that

$$Z_{n1} = -\frac{\alpha}{(\alpha - 1)^2} \{ \eta_1 W_{1n} + \eta_2 W_{2n} + \eta_3 W_{3n} \} + o_p(1). \quad (34)$$

Let us now consider the asymptotic distribution of  $Z_{n2}$ , in (33). It is shown that

$$Z_{n2} = -\frac{\int_{k/n}^1 B_n(1-s) dQ(1-s)}{\sqrt{k/n} Q(1-k/n)} + o_p(1).$$

On the other hand, from (10), we have  $Q(1-k/n) \sim (nc/k)^{1/\alpha}$ , as  $n \rightarrow \infty$ , it follows that

$$Z_{n2} = W_{4n} + o_p(1). \quad (35)$$

Combining (34) and (35) achieves the proof of Theorem 3.  $\square$

### 3.4 Proof of Theorem 4.

Now, we investigate the asymptotic normality of  $\widehat{\mu}_n$  given in(24). Since  $W_{in}$ ,  $i = 1, \dots, 4$  are sequences of centred Gaussian r.v.'s, then

$$\frac{\sqrt{n}}{\sqrt{k/n} (nc/k)^{1/\alpha}} \{ \widehat{\mu}_n - \mu \} \xrightarrow{d} \mathcal{N}(0, \Gamma \Sigma \Gamma^t), \text{ as } n \rightarrow \infty,$$

where

$$\Gamma := \left( -\frac{\alpha}{(\alpha - 1)^2} \eta_1, -\frac{\alpha}{(\alpha - 1)^2} \eta_2, -\frac{\alpha}{(\alpha - 1)^2} \eta_3, 1 \right),$$

$\Gamma^t$  is the transpose of  $\Gamma$  and  $\Sigma$  is the variance-covariance matrix of the vector  $(W_{1n}, \dots, W_{4n})$  defined by

$$\Sigma = \begin{bmatrix} 1 & \frac{1}{\lambda^2} - \frac{1}{\lambda} & \frac{2}{\lambda^3} - \frac{1}{\lambda^2} & -1 \\ \frac{1}{\lambda^2} - \frac{1}{\lambda} & \frac{1}{2\lambda-1} - \frac{1}{\lambda^2} & \frac{1}{(2\lambda-1)^2} - \frac{1}{\lambda^3} & \frac{\lambda-1}{\lambda} \\ \frac{2}{\lambda^3} - \frac{1}{\lambda^2} & \frac{1}{(2\lambda-1)^2} - \frac{1}{\lambda^3} & \frac{2}{(2\lambda-1)^3} - \frac{1}{\lambda^4} & -\frac{1}{\lambda^2} \\ -1 & \frac{\lambda-1}{\lambda} & -\frac{1}{\lambda^2} & \frac{2}{2-\alpha} \end{bmatrix}.$$

Note that the elements of  $\Sigma$  were obtained after tedious computations of the limits of the expectations  $\mathbf{E}[W_{in}W_{jn}]$  for  $i, j = 1, 4$  ( $i \leq j$ ). Finally, a standard calculation of the product  $\Gamma\Sigma\Gamma^t$  yields

$$\Gamma\Sigma\Gamma^t = \frac{\alpha^2\beta^4}{(\alpha-1)^4(\alpha-\beta)^4} + \frac{2}{2-\alpha} + \frac{2\alpha\beta^2}{(\alpha-1)^2(\alpha-\beta)^2},$$

which is denoted by  $\sigma^2(\alpha, \beta)$ . This completes the proof of Theorem 4.  $\square$

## 4 Conclusion

The main objective of this paper was to propose a bias-reduced estimator for the expectation of a heavy-tail distribution. This was achieved on the basis of the bias-reduction of the first and second order parameter estimators of regularly varying distributions and the corresponding high quantiles estimators. In addition, the newly introduced estimator is asymptotically normal, making confidence intervals easily constructible.

## References

- [1] B. Abdous, A.L. Fougères, and K. Ghoudi. Extreme behaviour for bivariate elliptical distributions. *Canadian Journal of Statistics*, 33(3):317–334, 2005.
- [2] Y. Aragon, A. Daouia, and C. Thomas-Agnan. Nonparametric frontier estimation: A conditional quantile-based approach. *Econometric Theory*, 21(2):358–389, 2005.
- [3] T.G. Bali and D. Weinbaum. A conditional extreme value volatility estimator based on high-frequency returns. *Journal of Economic Dynamics and Control*, 31(2):361–397, 2007.
- [4] J. Beirlant, T. De Wet, and Y. Goegebeur. Nonparametric estimation of extreme conditional quantiles. *Journal of statistical computation and simulation*, 74(8):567–580, 2004.
- [5] J. Beirlant and Y. Goegebeur. Local polynomial maximum likelihood estimation for pareto-type distributions. *Journal of Multivariate Analysis*, 89(1):97–118, 2004.
- [6] V.E. Berezkin, G.K. Kamenev, and A.V. Lotov. Hybrid adaptive methods for approximating a nonconvex multidimensional pareto frontier. *Computational Mathematics and Mathematical Physics*, 46(11):1918–1931, 2006.

- [7] H. Bi, G. Wan, and N.D. Turvey. Estimating the self-thinning boundary line as a density-dependent stochastic biomass frontier. *Ecology*, 81(6):1477–1483, 2000.
- [8] G. Bouchard, S. Girard, A. Iouditski, and A. Nazin. Some linear programming methods for frontier estimation. *Applied Stochastic Models in Business and Industry*, 21(2):175–185, 2005.
- [9] G. Bouchard, S. Girard, AB Iouditski, and AV Nazin. Nonparametric frontier estimation by linear programming. *Automation and Remote Control*, 65(1):58–64, 2004.
- [10] H.N.E. Byström. Managing extreme risks in tranquil and volatile markets using conditional extreme value theory. *International Review of Financial Analysis*, 13(2):133–152, 2004.
- [11] H.N.E. Byström. Extreme value theory and extremely large electricity price changes. *International Review of Economics & Finance*, 14(1):41–55, 2005.
- [12] C. Cazals, J.P. Florens, and L. Simar. Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106(1):1–25, 2002.
- [13] V. Chavez-Demoulin and A.C. Davison. Generalized additive modelling of sample extremes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(1):207–222, 2005.
- [14] A. Daouia. Asymptotic representation theory for nonstandard conditional quantiles. *Journal of Nonparametric Statistics*, 17(2):253–268, 2005.
- [15] A. Daouia, J.P. Florens, and L. Simar. Functional convergence of quantile-type frontiers with application to parametric approximations. *Journal of Statistical Planning and Inference*, 138(3):708–725, 2008.
- [16] A. Daouia, J.P. Florens, and L. Simar. Frontier estimation and extreme value theory. *Bernoulli*, 16(4):1039–1063, 2010.
- [17] A. Daouia, J.P. Florens, and L. Simar. Regularization of nonparametric frontier estimators. *Journal of Econometrics*, 2012.
- [18] A. Daouia, L. Gardes, and S. Girard. Nadarayas estimates for large quantiles and free disposal support curves. *Exploring Research Frontiers in Contemporary Statistics and Econometrics*, pages 1–22, 2012.
- [19] A. Daouia, L. Gardes, and S. Girard. On kernel smoothing for extremal quantile regression. *Bernoulli*, 19:2557–2589, 2013.
- [20] A. Daouia, L. Gardes, S. Girard, and A. Lekina. Kernel estimators of extreme level curves. *Test*, 20(2):311–333, 2011.

- [21] A. Daouia and I. Gijbels. Robustness and inference in nonparametric partial frontier modeling. *Journal of Econometrics*, 161(2):147–165, 2011.
- [22] A. Daouia and I. Gijbels. Estimating frontier cost models using extremiles. *Exploring Research Frontiers in Contemporary Statistics and Econometrics*, pages 65–81, 2012.
- [23] A. Daouia, S. Girard, and A. Guillou. A  $\gamma$ -moment approach to monotonic boundaries estimation: with applications in econometric and nuclear fields. *Journal of Econometrics*, 178:727–740, 2014.
- [24] A. Daouia and A. Ruiz-Gazen. Robust nonparametric frontier estimators: qualitative robustness and influence function. *Statistica Sinica*, 16(4):1233, 2006.
- [25] A. Daouia and L. Simar. Nonparametric efficiency analysis: A multivariate conditional quantile approach. *Journal of Econometrics*, 140(2):375–400, 2007.
- [26] B. Das and S.I. Resnick. Conditioning on an extreme component: Model consistency with regular variation on cones. *Bernoulli*, 17(1):226–252, 2011.
- [27] B. Das and S.I. Resnick. Detecting a conditional extreme value model. *Extremes*, 14(1):29–61, 2011.
- [28] AC Davison and NI Ramesh. Local likelihood smoothing of sample extremes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(1):191–208, 2000.
- [29] A.L.M. Dekkers and L. de Haan. On the estimation of extreme-value index and large quantiles estimation. *Ann. Statist.*, 17:1795–1832, 1989.
- [30] A. Delaigle and I. Gijbels. Estimation of boundary and discontinuity points in deconvolution problems. *Statistica Sinica*, 16(3):773, 2006.
- [31] G. Dierckx, Y. Goegebeur, and A. Guillou. Local robust and asymptotically unbiased estimation of conditional pareto-type tails. *Test*, 23(2):330–355, 2014.
- [32] J.P. Florens and L. Simar. Parametric approximations of nonparametric frontiers. *Journal of Econometrics*, 124(1):91–116, 2005.
- [33] L. Gardes and S. Girard. A moving window approach for nonparametric estimation of the conditional tail index. *Journal of Multivariate Analysis*, 99(10):2368–2388, 2008.
- [34] L. Gardes and S. Girard. Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels. *Extremes*, 13(2):177–204, 2010.

- [35] L. Gardes and S. Girard. Functional kernel estimators of large conditional quantiles. *Electronic Journal of Statistics*, 6:1715–1744, 2012.
- [36] L. Gardes and S. Girard. Nonparametric estimation of the conditional tail copula. *Journal of Multivariate Analysis*, 137:1–16, 2015.
- [37] L. Gardes and S. Girard. On the estimation of the functional Weibull tail-coefficient. *Journal of Multivariate Analysis*, 146:29–45, 2016.
- [38] L. Gardes, S. Girard, and A. Lekina. Functional nonparametric estimation of conditional extreme quantiles. *Journal of Multivariate Analysis*, 101(2):419–433, 2010.
- [39] J. Geffroy, S. Girard, and P. Jacob. Asymptotic normality of the  $l_1$ -error of a boundary estimator. *Journal of Nonparametric Statistics*, 18(1):21–31, 2006.
- [40] A. Ghorbel and A. Trabelsi. Predictive performance of conditional extreme value theory in value-at-risk estimation. *International Journal of Monetary Economics and Finance*, 1(2):121–148, 2008.
- [41] I. Gijbels, E. Mammen, B.U. Park, and L. Simar. On estimation of monotone and concave frontier functions. *Journal of the American Statistical Association*, 94(445):220–228, 1999.
- [42] S. Girard. On the asymptotic normality of the  $l_1$ -error for haar series estimates of poisson point processes boundaries. *Statistics & Probability Letters*, 66(1):81–90, 2004.
- [43] S. Girard, A. Guillou, and G. Stupfler. Frontier estimation with kernel regression on high order moments. *Journal of Multivariate Analysis*, 116:172–189, 2013.
- [44] S. Girard, A. Guillou, and G. Stupfler. Uniform strong consistency of a frontier estimator using kernel regression on high order moments. *ESAIM: Probability and Statistics*, 18:642–666, 2014.
- [45] S. Girard, A. Iouditski, and A.V. Nazin.  $l_1$ -optimal nonparametric frontier estimation via linear programming. *Automation and Remote Control*, 66(12):2000–2018, 2005.
- [46] S. Girard and P. Jacob. Extreme values and haar series estimates of point process boundaries. *Scandinavian Journal of Statistics*, 30(2):369–384, 2003.
- [47] S. Girard and P. Jacob. Projection estimates of point processes boundaries. *Journal of Statistical Planning and Inference*, 116(1):1–15, 2003.
- [48] S. Girard and P. Jacob. Extreme values and kernel estimates of point processes boundaries. *ESAIM: Probability and Statistics*, 8(1):150–168, 2004.

- [49] S. Girard and P. Jacob. Frontier estimation via kernel regression on high power-transformed data. *Journal of Multivariate Analysis*, 99(3):403–420, 2008.
- [50] S. Girard and P. Jacob. A note on extreme values and kernel estimators of sample boundaries. *Statistics & Probability Letters*, 78(12):1634–1638, 2008.
- [51] S. Girard and P. Jacob. Frontier estimation with local polynomials and high power-transformed data. *Journal of Multivariate Analysis*, 100(8):1691–1705, 2009.
- [52] S. Girard and L. Menneteau. Central limit theorems for smoothed extreme value estimates of poisson point processes boundaries. *Journal of Statistical Planning and Inference*, 135(2):433–460, 2005.
- [53] S. Girard and L. Menneteau. Smoothed extreme value estimators of non-uniform point processes boundaries with application to star-shaped supports estimation. *Communications in Statistics Theory and Methods*, 37(6):881–897, 2008.
- [54] Y. Goegebeur, A. Guillou, and A. Schorgen. Nonparametric regression estimation of conditional tails: the random covariate case. *Statistics*, 23:1–24, 2013.
- [55] W.H. Greene. Maximum likelihood estimation of econometric frontier functions. *Journal of Econometrics*, 13(1):27–56, 1980.
- [56] P. Hall, L. Peng, and C. Rau. Local likelihood tracking of fault lines and boundaries. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):569–582, 2001.
- [57] P. Hall and L. Simar. Estimating a changepoint, boundary, or frontier in the presence of observation error. *Journal of the American Statistical Association*, 97(458):523–534, 2002.
- [58] P. Hall and N. Tajvidi. Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science*, pages 153–167, 2000.
- [59] P. Jacob and C. Suquet. Estimating the edge of a poisson process by orthogonal series. *Journal of Statistical Planning and Inference*, 46(2):215–234, 1995.
- [60] P. Jacob and C. Suquet. Regression and edge estimation. *Statistics & Probability Letters*, 27(1):11–15, 1996.
- [61] S.O. Jeong and L. Simar. Linearly interpolated fdh efficiency score for nonconvex frontiers. *Journal of Multivariate Analysis*, 97(10):2141–2161, 2006.
- [62] A. Kneip, L. Simar, and P.W. Wilson. Asymptotics for dea estimators in nonparametric frontier models. Technical report, Discussion paper, 2003.

- [63] R. Koenker and K. Hallock. Quantile regression: An introduction. *Journal of Economic Perspectives*, 15(4):43–56, 2001.
- [64] A. P. Korostelev and A. B. Tsybakov. Asymptotic efficiency in estimation of a convex set. *Problemy Peredachi Informatsii*, 30(4):33–44, 1994.
- [65] Aleksandr Petrovich Korostelev and Alexandre B Tsybakov. *Minimax linewise algorithm for image reconstruction*. Springer, 1993.
- [66] S.C. Kumbhakar and C.A.K. Lovell. *Stochastic frontier analysis*. Cambridge University Press, 2003.
- [67] V. Marimoutou, B. Raggad, and A. Trabelsi. Extreme value theory and value at risk: application to oil market. *Energy Economics*, 31(4):519–530, 2009.
- [68] C. Martins-Filho and F. Yao. A smooth nonparametric conditional quantile frontier estimator. *Journal of Econometrics*, 143(2):317–333, 2008.
- [69] J. El Methni, L. Gardes, and S. Girard. Nonparametric estimation of extreme risks from conditional heavy-tailed distributions. *Scandinavian Journal of Statistics*, 41:988–1012, 2014.
- [70] J. El Methni, L. Gardes, and S. Girard. Kernel estimation of extreme regression risk measures. *Electronic journal of statistics*, 12(1):359–398, 2018.
- [71] A. Novikov. Linear aggregation methods applied to boundary estimators. <http://hal.archives-ouvertes.fr/hal-01074697>, 2014.
- [72] A. Novikov. Nonlinear aggregation methods for boundary estimators. <http://hal.archives-ouvertes.fr/hal-01075368>, 2014.
- [73] A. Novikov. Asymptotic distribution of convex-hull estimators. <http://hal.archives-ouvertes.fr/hal-01110490>, 2015.
- [74] A. Novikov. Boundary of a gaussian process. <http://hal.archives-ouvertes.fr/hal-01108160>, 2015.
- [75] B.U. Park and L. Simar. Efficient semiparametric estimation in a stochastic frontier model. *Journal of the American Statistical Association*, 89(427):929–936, 1994.
- [76] L. Peng. Bias-corrected estimators for monotone and concave frontier functions. *Journal of Statistical Planning and Inference*, 119(2):263–275, 2004.
- [77] S. Rao. Frontier estimation as a particular case of conditional extreme value analysis. <http://hal.archives-ouvertes.fr/hal-00776107>, 2013.

- [78] S. Rao. Linear aggregation of conditional extreme-value index estimators. <http://hal.archives-ouvertes.fr/hal-00771647>, 2013.
- [79] S. Rao. Linear aggregation of frontier estimators. <http://hal.archives-ouvertes.fr/hal-00798394>, 2013.
- [80] S. Rao. Nonlinear aggregation of conditional extreme-value index estimators. <http://hal.archives-ouvertes.fr/hal-00776108>, 2013.
- [81] S. Rao. Nonlinear aggregation of frontier estimators. *Journal of Parametric and Non-Parametric Statistics*, 1, 2013. <http://jpanps.altervista.org/>.
- [82] S. Rao. A review on conditional extreme value analysis. <http://hal.archives-ouvertes.fr/hal-00770546>, 2013.
- [83] S. Rao. Estimation of the proportional hazard premium under random censoring. <http://hal.archives-ouvertes.fr/hal-01795436>, 2018.
- [84] O. Rosen and A. Cohen. Extreme percentile regression. In *Statistical Theory and Computational Aspects of Smoothing: Proceedings of the COMPSTAT94 satellite meeting held in Semmering, Austria*, pages 27–28, 1994.
- [85] P. Schmidt and T.F. Lin. Simple tests of alternative specifications in stochastic frontier models. *Journal of Econometrics*, 24(3):349–361, 1984.
- [86] J.K. Sengupta. Stochastic data envelopment analysis: a new approach. *Applied Economics Letters*, 5(5):287–290, 1998.
- [87] L. Simar. Aspects of statistical analysis in dea-type frontier models. *Journal of Productivity Analysis*, 7(2):177–185, 1996.
- [88] L. Simar. Detecting outliers in frontier models: A simple approach. *Journal of Productivity Analysis*, 20(3):391–424, 2003.
- [89] L. Simar and P.W. Wilson. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management science*, 44(1):49–61, 1998.
- [90] L. Simar and P.W. Wilson. Of course we can bootstrap dea scores! but does it mean anything? logic trumps wishful thinking. *Journal of Productivity Analysis*, 11(1):93–97, 1999.
- [91] L. Simar and P.W. Wilson. A general methodology for bootstrapping in non-parametric frontier models. *Journal of Applied Statistics*, 27(6):779–802, 2000.

- [92] L. Simar and P.W. Wilson. Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productivity Analysis*, 13(1):49–78, 2000.
- [93] L. Simar and P.W. Wilson. Testing restrictions in nonparametric efficiency models. *Communications in Statistics-Simulation and Computation*, 30(1):159–184, 2001.
- [94] L. Simar and P.W. Wilson. Statistical inference in nonparametric frontier models: recent developments and perspectives. *The measurement of productive efficiency and productivity growth*, pages 421–521, 2008.
- [95] L. Simar and P.W. Wilson. Inferences from cross-sectional, stochastic frontier models. *Econometric Reviews*, 29(1):62–98, 2009.
- [96] L. Simar and P.W. Wilson. Performance of the bootstrap for dea estimators and iterating the principle. *Handbook on data envelopment analysis*, pages 241–271, 2011.
- [97] L. Simar and V. Zelenyuk. Stochastic fdh/dea estimators for frontier analysis. *Journal of Productivity Analysis*, 36(1):1–20, 2011.
- [98] C. Wang, H. Zhuang, Z. Fang, and T. Lu. A model of conditional var of high frequency extreme value based on generalized extreme value distribution. *Journal of Systems & Management*, 3:003, 2008.
- [99] H. Wang and C.L. Tsai. Tail index regression. *Journal of the American Statistical Association*, 104(487):1233–1240, 2009.
- [100] H.J. Wang, D. Li, and X. He. Estimation of high conditional quantiles for heavy-tailed distributions. *Journal of the American Statistical Association*, 107:1453–1464, 2012.

## A Appendix: Auxiliary results

**Lemma 2** *Let  $k = k_n$  be a sequence of integers satisfying (6) and  $f(x) = x^{\lambda-1} \log x$ ,  $\lambda > 1$ . Then, uniformly on  $s \in [1/n, k/n]$ , we have*

$$f' \left( \frac{\varphi_n(s)}{1 - U_{n-k,n}} \right) - f' \left( \frac{s}{1 - U_{n-k,n}} \right) = o_p(1) \left( \frac{s}{k/n} \right)^{\lambda-2} \log \frac{s}{k/n}, \text{ as } n \rightarrow \infty.$$

**Proof.** We have  $n(1 - U_{n-k,n})/k \xrightarrow{p} 0$ , as  $n \rightarrow \infty$ , then

$$f' \left( \frac{\varphi_n(s)}{1 - U_{n-k,n}} \right) - f' \left( \frac{s}{1 - U_{n-k,n}} \right) = (1 + o_p(1)) \left[ f' \left( \frac{\varphi_n(s)}{k/n} \right) - f' \left( \frac{s}{k/n} \right) \right].$$

A straightforward calculation of the derivative of  $f$  yields

$$f' \left( \frac{\varphi_n(s)}{k/n} \right) - f' \left( \frac{s}{k/n} \right) = \left( \frac{s}{k/n} \right)^{\lambda-2} \left[ (\lambda-1) \left( \left( \frac{\varphi_n(s)}{s} \right)^{\lambda-2} - 1 \right) \log \frac{\varphi_n(s)}{k/n} + (\lambda-1) \log \frac{\varphi_n(s)}{s} + \left( \frac{\varphi_n(s)}{s} \right)^{\lambda-2} - 1 \right]. \quad (36)$$

Observe now, that inequalities (31) imply

$$\min \left\{ \frac{1-s-V_n(1-s)}{s}, 0 \right\} < \frac{\varphi_n(s)}{s} - 1 < \max \left\{ \frac{1-s-V_n(1-s)}{s}, 0 \right\}.$$

We have

$$\sup_{1/n \leq s < 1} \frac{|1-s-V_n(1-s)|}{s} \xrightarrow{p} 0 \text{ as } n \rightarrow \infty,$$

it follows that

$$\sup_{1/n \leq s \leq k/n} \left| \frac{\varphi_n(s)}{s} - 1 \right| \xrightarrow{p} 0 \text{ as } n \rightarrow \infty. \quad (37)$$

On the other hand, we infer that

$$\sup_{1/n \leq s \leq k/n} \left| \frac{s}{\varphi_n(s)} - 1 \right| \xrightarrow{p} 0 \text{ as } n \rightarrow \infty. \quad (38)$$

By applying the mean value theorem to the functions  $x \rightarrow \log x$  and  $x \rightarrow x^{\lambda-1}$  respectively, then by using (37) and (38), we show readily that, as  $n \rightarrow \infty$

$$\sup_{1/n \leq s \leq k/n} \left| \log \frac{\varphi_n(s)}{s} \right| \xrightarrow{p} 0 \text{ and } \sup_{1/n \leq s \leq k/n} \left| \left( \frac{\varphi_n(s)}{s} \right)^{\lambda-1} - 1 \right| \xrightarrow{p} 0. \quad (39)$$

Note that the first result of (39) implies that

$$\sup_{1/n \leq s \leq k/n} \left| \log \frac{\varphi_n(s)}{k/n} - \log \frac{s}{k/n} \right| \xrightarrow{p} 0, \text{ as } n \rightarrow \infty. \quad (40)$$

By using equations (39) and (40) together, we show that, uniformly in  $s \in [1/n, k/n]$ , the right-hand side of equation (36), is equal to  $o_p(1) \left( \frac{s}{k/n} \right)^{\lambda-2} \log \frac{s}{k/n}$ .  $\square$

**Lemma 3** *We have  $\sqrt{k}\Omega_{n1}^* \xrightarrow{p} 0$  and  $\sqrt{k}\Omega_{n2}^* \xrightarrow{p} 0$ , as  $n \rightarrow \infty$ .*

*Proof.* We only show the first result. The second one is obtained by similar arguments. Recall that

$$\Omega_{n1}^* := - (n/k) \int_{1/n}^{k/n} \left[ f' \left( \frac{\varphi_n(s)}{1-U_{n-k,n}} \right) - f' \left( \frac{s}{1-U_{n-k,n}} \right) \right] \times \left[ \frac{1-V_n(1-s)}{1-U_{n-k,n}} - \frac{s}{1-U_{n-k,n}} \right] ds.$$

Using Lemma 2, we get, as  $n \rightarrow \infty$

$$\Omega_{n1}^* = o_p(1) (n/k) \int_{1/n}^{k/n} \left( \frac{s}{k/n} \right)^{\lambda-2} \left| \log \frac{s}{k/n} \right| \left| \frac{1 - V_n(1-s)}{k/n} - \frac{s}{k/n} \right| ds.$$

By a change of variables, we get

$$\Omega_{n1}^* = o_p(1) (n/k) \int_{1/k}^1 s^{\lambda-2} |\log s| |1 - V_n(1 - ks/n) - ks/n| ds.$$

Making use of approximation(30) yields

$$\sqrt{k}\Omega_{n1}^* = o_p(1) (n/k)^{1/2} \int_{1/k}^1 s^{\lambda-2} |\log s| \left( |B_n(1 - ks/n)| + (ks/n)^{1/2-\tau} O_p(n^{-\tau}) \right) ds.$$

In other words

$$\begin{aligned} \sqrt{k}\Omega_{n1}^* &= o_p(1) (n/k)^{1/2} \int_{1/k}^1 s^{\lambda-2} |\log s| |B_n(1 - ks/n)| ds. \\ &\quad + o_p(k^{-\tau}) \int_{1/k}^1 s^{\lambda-3/2-\tau} |\log s| ds. \end{aligned}$$

The expectation of the first term of right-hand side of the previous equation is less than or equal to

$$o_p(1) (n/k)^{1/2} \int_0^1 s^{\lambda-2} |\log s| \mathbf{E}[|B_n(1 - ks/n)|] ds.$$

Using the fact that  $E|B_n(1 - ks/n)| \leq (ks/n)^{1/2}$ , we show that the previous quantity is less than or equal to  $o_p(1) \int_0^1 s^{\lambda-3/2} |\log s| ds$ . Since both integrals  $\int_0^1 s^{\lambda-3/2-\tau} |\log s| ds$  and  $\int_0^1 s^{\lambda-3/2-\tau} |\log s| ds$  are finite, then  $\sqrt{k}\Omega_{n1}^* = o_p(1)$ , as  $n \rightarrow \infty$ .  $\square$