Statistics and Its Interface Volume 8 (2015) 85–91

Kernel-type estimator of the mean for a heavy tailed distribution

Abdelaziz Rassoul

In this paper, we focus on the reduced bias of the mean estimator for a heavy-tailed distribution. It is well known that the classical mean estimator introduced by Peng (2001) is seriously biased under the second order regular variation. To reduce bias, many authors have proposed estimators, for both first and second order parameters of the distribution tail. In this work, we define a kernel type estimator for the mean and we propose a reduced bias estimator. The asymptotic distributional properties of our proposed estimators are derived and we compared their performances with other estimators.

Keywords and phrases: Mean, Heavy tails, Kernel-type estimator, Extreme quantile, Reduced bias.

1. INTRODUCTION

Let X be a positive random variable (rv) with cumulative distribution function (cdf) F. Throughout this paper we shall assume that F is continuous and we note by $\overline{F} = 1 - F$ the tail of the cdf F. The mean of the rv X is then defined by

$$\mu = \mathbf{E}\left(X\right) = \int_0^{+\infty} \overline{F}\left(x\right)dx. \tag{1}$$

Note that we can also rewrite the above integral in terms of the quantile function corresponding to the cdf F, as follows

$$\mu = \int_0^1 \mathbb{Q}(s)ds, \tag{2}$$

where $\mathbb{Q}(t) = \inf\{x : F(x) \geq t\}$ for $t \in [0,1]$.

Now, assume that we have at our disposal a sample of independent and identically distributed (iid) random variables $X_1, X_2, \ldots, X_n$ with the cdf F, and let $X_{1,n} < X_{2,n} < \cdots < X_{n,n}$ be the corresponding order statistics.

One natural candidate for the empirical estimate of right side of (2) is obtained by replacing the true quantile with the sample quantile as follows

$$\widehat{\mu}_n = \int_0^1 \mathbb{Q}_n(s)ds, \tag{3}$$

where $\mathbb{Q}_n(s)$ is the empirical quantile function, which is equal to the i^{th} order statistic $X_{i,n}$ for all $s \in ((i-1)/n, i/n]$, for all $i = 1, \ldots, n$. The asymptotic behavior of the estimator $\widehat{\mu}_n$ has been known by the Central Limit Theorem (CLT), provided that the second moment is finite ($\mathbf{E}[X^2] < \infty$).

In this work we shall be concerned with heavy tailed distributions. In mathematical terms, a heavy-tailed distribution of a random variable X is defined as follows

$$1 - F(x) = x^{-1/\gamma}\mathbb{L}(x), \quad \textit{for every} \quad x > 0, \tag{4}$$

where $\mathbb{L}(x)$ is a slowly varying function at infinity, that is $\lim_{t\to\infty} \mathbb{L}(tx)/\mathbb{L}(t) = 1$ for all $x > 0$. This class includes a number of popular distributions such as Pareto, generalized Pareto, Burr, Fréchet, Student, etc., which are known to be appropriate models for fitting large insurance claims, fluctuations of prices, log-returns, etc. (see, e.g., Beirlant *et al.* (2001) [2]). In the remainder of this paper, we restrict ourselves to this class of distributions. Moreover we focus our paper on the case $\gamma \in (1/2, 1)$ to ensure that the mean is finite and since in that case the results of CTL cannot be applied, because the second moment of X being infinite.

Indeed, recall that from (2), μ can be rewritten as

$$\begin{aligned}\mu &= \int_0^{1-k/n} \mathbb{Q}(s)ds + \int_0^{k/n} \mathbb{Q}(1-s)ds \\ &= \mu_{1,n} + \mu_{2,n}.\end{aligned}$$

Peng (2001) [23] proposed an alternative estimator of μ for a heavy tailed distribution as follows

$$\widetilde{\mu}_{n,k}^H = \int_0^{1-k/n} \mathbb{Q}_n(s)ds + \frac{(k/n)\, X_{n-k,n}}{(1 - \widehat{\gamma}_{n,k}^H)}, \tag{5}$$

where $\widehat{\gamma}_{n,k}^H$ is the Hill estimator [21] of the tail index γ:

$$\widehat{\gamma}_{n,k}^H = \frac{1}{k}\sum_{i=1}^{k} \log X_{n-i+1,n} - \log X_{n-k,n}. \tag{6}$$

Note that to estimate $\mu_{2,n}$ we use a Weissman-type estimator for $\mathbb{Q}$ [26]

$$\widehat{\mathbb{Q}}\left(1-s\right) := X_{n-k,n} \;\; (k/n)^{\widehat{\gamma}_{n,k}^H} \;\; s^{-\widehat{\gamma}_{n,k}^H}, \;\; s \to 0. \tag{7}$$

We note that, Peng (2001) [23] defined his estimator with $\gamma = 1/\alpha$ and he took the general situation where 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 F and the balance condition

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

In this paper, we only consider non-negative rv's. Our motivation comes from the actuarial risk theory where insurance losses are represented by such rv's. In this case, $\widetilde{\mu}^H_{n,k}$ may be interpreted as an estimator of a risk measure called the net premium. Note that in our case, we have $F(x)=0$ for $x<0$, which gives $p=1$ in the above balance condition.

The Hill estimation has been extensively studied in the literature for an intermediate sequence k, such that $k\to\infty$ and $k/n\to 0$ as $n\to\infty$.

More generally, Csörgő et al. (1985) [8] extended the Hill estimator into a kernel class of estimators

$$\widehat{\gamma}^K_{n,k}=\frac{1}{k}\sum_{i=1}^{k}K\left(\frac{i}{k+1}\right)Z_{i,k}, \tag{8}$$

where K is a kernel integrating to one and

$$Z_{i,k}=i\left(\log X_{n-i+1,n}-\log X_{n-i,n}\right),\ 1\le i\le k<n.$$

Note that the Hill estimator corresponds to the particular case where $K=\underline{K}:=1_{(0,1)}$.

In this spirit, we propose a kernel-type estimator for the mean. Thus, μ can be estimated by

$$\widetilde{\mu}^K_{n,k}=\int_0^{1-k/n}\mathbb{Q}_n(s)ds+\frac{(k/n)\,X_{n-k,n}}{(1-\widehat{\gamma}^K_{n,k})}. \tag{9}$$

Asymptotic normality for $\widetilde{\mu}^K_{n,k}$ is obviously related to the one of $\widehat{\gamma}^K_{n,k}$. As usual in the extreme value framework, to prove such type of results, we need a second-order condition on the tail quantile function $\mathbb{U}$, defined as

$$\mathbb{U}(z)=\inf\left\{y:F(y)\ge 1-1/z\right\},\ z>1. \tag{10}$$

We say that the function $\mathbb{U}$ satisfies the second-order regular variation condition with second-order parameter $\rho\le 0$ if there exists a function $A(t)$ which does not change its sign in a neighbourhood of infinity with $\lim_{t\longrightarrow\infty}A(t)=0$, such that, for every $x>0$,

$$\lim_{t\to\infty}\frac{\log\mathbb{U}(tx)-\log\mathbb{U}(t)-\gamma\log(x)}{A(t)}=\frac{x^\rho-1}{\rho}, \tag{11}$$

when $\rho=0$, then the ratio on the right-hand side of equation (11) should be interpreted as $\log x$. As an example of heavy-tailed distributions satisfying the second order condition, we have the so called and frequently used Hall's model [20] which is a class of cdf's, such that

$$\mathbb{U}(t)=ct^\gamma\left(1+dA(t)/\rho+o\left(t^\rho\right)\right)\ \text{as}\ \ t\to\infty, \tag{12}$$

where $\gamma>0$, $\rho\le 0$, $c>0$, and $d\in\mathbb{R}^*$. For example, if we consider the special case

$$1-F(x)=cx^{-1/\gamma}\big(1+dx^{\rho/\gamma}+o\big(x^{\rho/\gamma}\big)\big),$$

with $c>0$, $d\in\mathbb{R}$ and $\rho<0$, (11) holds and we may choose $A(t)=d\gamma\rho t^\rho$.

This sub-class of heavy-tailed distributions contains the Pareto, Burr, Fréchet and t-Student, cdf's usually used, in insurance mathematics, as models for dangerous risks. For statistical inference concerning the second-order parameter ρ, we refer, for example, to Peng and Qi (2004) [24], Gomes *et al.* (2007) [16], Gomes and Pestana (2007) [15]. In the sequel, we assume that equation (11) holds with $\rho<0$.

The remainder of this paper is organized as follows. In Section 2, we study the asymptotic properties of the general kernel estimator of the mean $\widetilde{\mu}^K_{n,k}$. This result illustrates the fact that this estimator can exhibit severe bias in many situations. To solve this problem a reduced-bias approach is also proposed. The efficiency of our method is shown on a small simulation study in Section 3. All proofs are deferred to section 4.

Note that throughout this paper, the standard notations $\overset{\mathbb{P}}{\to}$, $\overset{d}{\to}$ and $\overset{d}{=}$ respectively stand for convergence in probability, convergence in distribution and equality in distribution, while $\mathcal{N}(a,b^2)$ denotes the normal distribution with mean a and variance b^2.

2. MAINS RESULTS

For study of the asymptotic normality of the estimator of $\widetilde{\mu}^K_{n,k}$, we need some results and classical assumptions about the kernel:

Condition $(\mathcal{K})$: Let K be a function defined on $(0,1]$

- CK1. $K(s)\ge 0$ whenever $0<s\le 1$ and $K(1)=0$;
- CK2. $K(\cdot)$ is differentiable, nonincreasing and right continuous on $(0,1]$;
- CK3. K and K' are bounded;
- CK4. $\int_0^1 K(u)du=1$;
- CK5. $\int_0^1 u^{-1/2}K(u)du<\infty$.

2.1 Asymptotic results for the mean estimator

Theorem 2.1. *Assume that F satisfies (11) with $\gamma\in(1/2,1)$. If further $(\mathcal{K})$ holds and the sequence k satisfies $k\to\infty$, $k/n\to 0$ and if $\sqrt{k}A(n/k)\to\lambda\in\mathbb{R}$, as $n\to\infty$, we have*

$$\frac{\sqrt{k}}{(k/n)\,\mathbb{U}(n/k)}\left(\widetilde{\mu}^K_{n,k}-\mu\right)\overset{d}{\to}\mathcal{N}\left(\lambda\mathcal{A}\mathcal{B}_K\left(\gamma,\rho\right),\mathcal{A}\mathcal{C}_K\left(\gamma,\rho\right)\right),$$

where

$$\mathcal{AB}_K(\gamma,\rho) = \left(\frac{1}{(\gamma-1)(\gamma+\rho-1)} + \frac{1}{(1-\gamma)^2}\int_0^1 \frac{K(s)}{s^\rho}ds\right);$$

and

$$\mathcal{AC}_K(\gamma,\rho) = \frac{\gamma^2}{(1-\gamma)^2(2\gamma-1)} + \frac{\gamma^2}{(1-\gamma)^4}\int_0^1 K^2(s)\,ds.$$

Note that the result of Theorem 2.1 generalizes Theorem 1 in Peng (2001) [23] in case $\lambda \neq 0$ and when we use a general kernel instead of K.

In view of these results, $\widetilde{\mu}_{n,k}^K$ is an estimator of μ with an asymptotic bias given by

$$(k/n)\,\mathbb{U}(n/k)A(n/k)\mathcal{AB}_K(\gamma,\rho).$$

For a specific kernel, the asymptotic bias and variance can be computed. In the following remark we take the case $K = \underline{K}$.

Remark 2.2. *In the special case where $K = \underline{K}$, Theorem 2.1 is equivalent to*

$$\frac{\sqrt{k}(\widetilde{\mu}_{n,k}^K - \mu)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathcal{N}\left(\lambda\frac{\gamma\rho}{(1-\rho)(\gamma+\rho-1)(1-\gamma)^2}, \sigma_\gamma^2\right), \tag{13}$$

where the asymptotic variance σ_γ^2 is given by the formula

$$\sigma_\gamma^2 = \frac{\gamma^4}{(1-\gamma)^4(2\gamma-1)}.$$

Now, we want to propose a reduced-bias estimator of the mean μ.

2.2 Bias-correction for the mean estimator

Recall that, from Theorem 2.1,

$$\widetilde{\mu}_{n,k}^K - (k/n)\,\mathbb{U}(n/k)A(n/k)\mathcal{AB}_K(\gamma,\rho),$$

is an asymptotically unbiased estimator for μ. Note that γ, ρ, $\mathbb{U}(n/k)$ and $A(n/k)$ are unknown quantities that we have to estimate. Under the condition (11), from Feuerverger and Hall (1999) [11] and Beirlant et al. (1999, 2002) [1][3] proposed that, for k not too large the scaled logarithmic spacings of order statistics

$$Z_{i,k} := i(logX_{n-i+1,n} - logX_{n-i,n}),\ i = 1, ..., k,$$

can be modelled by the following generalised regression models:

$$Z_{i,k} \sim \left(\gamma + A(\frac{n}{k})\left(\frac{i}{k+1}\right)^{-\rho}\right) + \epsilon_{i,k},\ 1 \le i \le k, \tag{14}$$

where $\epsilon_{i,k}$ are zero-centered error terms. We note that, if we ignore the term $A(n/k)$ in (14), we obtain the Hill estimator [21] $\widehat{\gamma}_{n,k}^H$ by taking the mean of the left-hand side of (14). By using a least-squares approach, formula (14) can be further exploited to propose a reduced-bias estimator for γ in which ρ is substituted by a consistent estimator $\widehat{\rho} = \widehat{\rho}(n,k)$ (see for instance Beirlant et al. (2002) [3] and Fraga Alves et al. (2003) [12] or by a canonical choice, such as $\rho = -1$ (see e.g. Feuerverger and Hall (1999) [11] or Beirlant et al. (1999) [1]).

The least squares estimators for γ and $A(n/k)$ are then given by

$$\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}) = \frac{1}{k}\sum_{i=1}^k Z_{i,k} - \frac{\widehat{A}_{n,k}^{L.S}(\widehat{\rho})}{1-\widehat{\rho}};$$

and

$$\widehat{A}_{n,k}^{L.S}(\widehat{\rho}) = \frac{(1-2\widehat{\rho})(1-\widehat{\rho})^2}{\widehat{\rho}^2}\frac{1}{k}\sum_{i=1}^k\left(\left(\frac{i}{k+1}\right)^{-\widehat{\rho}} - \frac{1}{1-\widehat{\rho}}\right)Z_{i,k}.$$

Note that $\widehat{\gamma}_{n,k}^{L.S}(\rho)$ can be viewed as the kernel estimator $\widehat{\gamma}_{n,k}^{K_\rho}$, where for $0 < u \le 1$:

$$K_\rho(u) := \frac{1-\rho}{\rho}\underline{K}(u) + \left(1 - \frac{1-\rho}{\rho}\right)\underline{K}_\rho(u), \tag{15}$$

with $\underline{K}(u) = 1_{(0,1)}$ and $\underline{K}_\rho(u) = (\frac{1-\rho}{\rho})(u^{-\rho}-1)1_{(0,1)}$, both kernels satisfying condition $(\mathcal{K})$. On the contrary $\underline{K}_\rho$ does not satisfy statement (CK1) in $(\mathcal{K})$. We refer to Gomes and Martins (2004) [13] and Gomes et al. (2007) [16] for other techniques of bias reduction based on the estimation of the second order parameter.

We are now able to obtain a reduced-bias estimator for the mean μ from equation (11) and using the above estimators for the different unknown quantities:

$$\begin{aligned}\widehat{\mu}_{n,k}^{K,\widehat{\rho}} &= \widetilde{\mu}_{n,k}^K - \left(\frac{k}{n}\right)X_{n-k,n}\widehat{A}_{n,k}^{L.S}(\widehat{\rho})\mathcal{AB}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right)\\ &= \left(\frac{k}{n}\right)X_{n-k,n}\Bigg(\frac{1}{1-\widehat{\gamma}_{n,k}^K}\\ &\quad - \widehat{A}_{n,k}^{L.S}(\widehat{\rho})\mathcal{AB}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right)\Bigg)\\ &\quad + \int_0^{1-k/n}\mathbb{Q}_n(s)ds.\end{aligned}$$

The asymptotic normality of $\widehat{\mu}_{n,k}^{K,\widehat{\rho}}$ is established in the following theorem.

Theorem 2.3. *Under the same assumptions of Theorem 2.1, and if $\widehat{\rho}$ is a consistent estimator for $\rho < 0$, then, we have*

$$\frac{\sqrt{k}(\widehat{\mu}_{n,k}^{K,\widehat{\rho}} - \mu)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathcal{N}\left(0, \widetilde{\mathcal{AC}}_K(\gamma,\rho)\right),$$

Table 1. Comparison of the estimator $\widetilde{\mu}_{n,k}^{K}$ and the reduced-bias estimator $\widehat{\mu}_{n,k}^{L,S,\widehat{\rho}}$ for 500 samples of size 400 and 1,000 of Fréchet distribution with $\gamma = 2/3$ and $\rho = -1$ (The true value is 2.5553)

N	400				1,000			
k	25	75	110	160	50	125	200	250
$\widetilde{\mu}_{n,k}^{K}$	2.6602	2.7415	2.8075	2.9973	2.5163	2.6353	2.8207	3.1258
$\widetilde{\mu}_{n,k}^{L,S,\widehat{\rho}}$	2.6459	2.6525	2.6357	2.6202	2.4951	2.5254	2.5293	2.5516
bias1	0.1049	0.1862	0.2522	0.4421	-0.0389	0.0800	0.2655	0.5706
bias2	0.0906	0.0973	0.0804	0.0650	-0.0601	-0.0299	-0.0259	-0.0037
mse1	0.0110	0.0347	0.0636	0.1954	0.0015	0.0064	0.0705	0.3255
mse2	0.0082	0.0095	0.0065	0.0042	0.0036	0.00089	0.00067	0.000013

where

$$\widetilde{\mathcal{AC}}_K(\gamma,\rho) = \mathcal{AC}_K(\gamma,\rho) + \frac{\gamma^2}{\rho^2}(1-2\rho)(1-\rho)^2\,\mathcal{AB}_K^2(\gamma,\rho) + \frac{2\gamma^2(1-2\rho)(1-\rho)}{\rho^2(1-\gamma)^2} \times \left(1-(1-\rho)\int_0^1 \frac{K(s)}{s^\rho}ds\right)\mathcal{AB}_K(\gamma,\rho).$$

We observe that, the estimator $\widehat{\mu}_{n,k}^{K,\widehat{\rho}}$ has a null asymptotic bias which was not the case for $\widetilde{\mu}_{n,k}^{K}$ equation (13), but with an increase of the asymptotic variance.

Remark 2.4. *In the special case $K = \underline{K}$, the result of the Theorem 2.3 becomes*

$$\frac{\sqrt{k}\left(\widehat{\mu}_{n,k}^{K,\widehat{\rho}}-\mu\right)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathcal{N}\left(0, \frac{\gamma^4(\gamma-\rho)^2}{(2\gamma-1)(\gamma+\rho-1)^2(1-\gamma)^4}\right).$$

Now, in the special case where $K = K_\rho$, as mentioned in (15), the estimator $\widehat{\gamma}_{n,k}^{K_\rho}$ coincides with $\widehat{\gamma}_{n,k}^{L.S}(\rho)$.

The aim of the next corollary is to establish the asymptotic normality of the resulting mean estimator $\widehat{\mu}_{n,k}^{K_\rho,\widehat{\rho}}$, denoted by $\widehat{\mu}_{n,k}^{L.S,\widehat{\rho}}$, when the least squares approach is adopted.

Corollary 2.5. *Under the same assumptions as in Theorem 2.3 and in the special case where $K = K_\rho$, we have*

$$\frac{\sqrt{k}\left(\widehat{\mu}_{n,k}^{L.S,\widehat{\rho}}-\mu\right)}{(k/n)\mathbb{U}(n/k)} \xrightarrow{d} \mathcal{N}\left(0, \widetilde{\mathcal{AC}}_{K_\rho}(\gamma,\rho)\right),$$

where

$$\widetilde{\mathcal{AC}}_{K_\rho}(\gamma,\rho) = \frac{\gamma^2(1-\rho)^2}{\rho^2(1-\gamma)^4} + \frac{\gamma^2}{(2\gamma-1)(1-\gamma)^2} + \frac{\gamma^2(1-2\rho)(1-\rho)(\gamma\rho+2\rho+\gamma-\rho-1)}{\rho^2(1-\gamma)^3(\gamma+\rho-1)^2}.$$

3. SIMULATION STUDY

In this section, we carry out a simulation study (by means of the statistical software **R**, see Ihaka and Gentleman, 1996) [22] to illustrate the performance of the biased estimator $\widetilde{\mu}_{n,k}^{K}$ with the kernel $K = \underline{K} = 1_{(0,1)}$ and the reduced-bias estimator $\widetilde{\mu}_{n,k}^{L,S,\widehat{\rho}}$, we compare the two estimators in terms of the bias and the mean squared error (mse) at different values of the fraction k, satisfies the condition: $k \to \infty$ and $k/n \to 0$ as $n \to \infty$. Through its application to sets of samples taken from Fréchet distributions with tail of distribution $\overline{F}(x) = 1 - \exp(-x^{-1/\gamma})$, $x > 0$ (with tail index $\gamma = 2/3$ and $\gamma = 3/4$) and the second order parameter $\rho = -1$, we generate 500 independent replicates of sizes 400 and 1,000 from the selected parent distribution. For each simulated sample, we obtain an estimate of the estimators $\widetilde{\mu}_{n,k}^{K}$ and $\widetilde{\mu}_{n,k}^{L,S,\widehat{\rho}}$. The overall estimated mean is then taken as the empirical mean of the values in the 500 repetitions. We not that, the notations bias1 and bias2 represents the bias of the estimators $\widetilde{\mu}_{n,k}^{K}$ and $\widetilde{\mu}_{n,k}^{L,S,\widehat{\rho}}$ respectively, and the notations mse1 and mse2 represents the mean squared error of the estimators $\widetilde{\mu}_{n,k}^{K}$ and $\widetilde{\mu}_{n,k}^{L,S,\widehat{\rho}}$ respectively. To this end, we summarize the results in Table 1 for $\gamma = 2/3$ and in Table 2 for $\gamma = 3/4$.

We conclude that in most cases the mean estimators gives rise to a general better performance of the $\widetilde{\mu}_{n,k}^{L,S,\widehat{\rho}}$ relatively to the estimator $\widetilde{\mu}_{n,k}^{K}$, when both are considered at levels k, especially for difficult distributions such as the Fréchet with index $\gamma \in (1/2, 1)$ when the second moment are infinite, we remark that all the biases decrease when n increase, then notes that, the bias and the mean squared error (mse) of $\widetilde{\mu}_{n,k}^{L,S,\widehat{\rho}}$ are smaller than the corresponding ones of the $\widetilde{\mu}_{n,k}^{K}$ estimator. We note that similar results can be obtained with other distributions satisfies (11) with $\gamma \in (1/2, 1)$.

4. PROOFS

Let $Y_1, ..., Y_n$ be iid. rv's from the unit Pareto distribution G, defined as $G(y) = 1 - 1/y, y > 1$. For each $n \geq 1$, let $Y_{1,n} \leq ... \leq Y_{n,n}$ be the order statistics pertaining to $Y_1, ..., Y_n$. Clearly $X_{j,n} \overset{d}{=} \mathbb{U}(Y_{j,n})$, $j = 1, ..., n$. In order to use results from Csörgő et al. (1986) [6], a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ is constructed carrying an infinite sequence $\zeta_1, \zeta_2, ..., \zeta_n$ of independent random variables uniformly distributed on $(0, 1)$ and a sequence of Brownian

Table 2. Comparison of the estimator $\widetilde{\mu}_{n,k}^{K}$ and the reduced-bias estimator $\widehat{\mu}_{n,k}^{L,S,\widehat{\rho}}$ for 500 samples of size 400 and 1,000 of Fréchet distribution with $\gamma = 3/4$ and $\rho = -1$ (The true value is 3.2636)

N	400				1,000			
k	25	75	110	160	50	125	200	250
$\widetilde{\mu}_{n,k}^{K}$	3.3384	3.7606	4.4488	4.9647	3.1850	3.4628	3.9343	4.9052
$\widetilde{\mu}_{n,k}^{L,S,\widehat{\rho}}$	3.2703	3.2545	3.1454	3.0733	3.1429	3.2356	3.2442	3.2684
bias1	0.0748	0.4969	1.1852	1.7011	-0.0786	0.1992	0.6706	1.6415
bias2	0.0067	-0.0092	-0.1182	-0.1903	-0.1207	-0.0281	-0.0194	0.0047
mse1	0.0056	0.2469	1.4046	2.8937	0.0062	0.0397	0.4497	2.6947
mse2	0.00004	0.00008	0.01398	0.03623	0.0146	0.00079	0.00038	0.00002

bridges $\{\mathbf{B}_n(s), 0 \leq s \leq 1\}$. The resulting empirical quantile is denoted by

$$\beta_n(t) = \sqrt{n}\left(t - \mathbb{V}_n(s)\right),\ 0 \leq t \leq 1,$$

where

$$\mathbb{V}_n(s) = \zeta_{j,n},\ \frac{j-1}{n} \leq s \leq \frac{j}{n},\ j = 1, 2, ..., n$$

and $\mathbb{V}_n(0) = \zeta_{1,n}$, where $\zeta_{1,n} \leq \zeta_{2,n} \leq ... \leq \zeta_{n,n}$ denote the order statistics based on the sample $\zeta_1, \zeta_2, ..., \zeta_n$.

Proof of the Theorem 2.1. Let us rewrite

$$\left(\widetilde{\mu}_{n,k}^{K} - \mu\right) = A_{n,1} + A_{n,2}, \tag{16}$$

where

$$A_{n,1} = \int_0^{1-k/n} \left(\mathbb{Q}_n(s) - \mathbb{Q}(s)\right) ds,$$

and

$$A_{n,2} = \frac{k/n}{1 - \widehat{\gamma}_{n,k}^{K}} X_{n-k,n} - \int_{1-k/n}^{1} \mathbb{Q}(s) ds.$$

We shall show below that there are Brownian bridges $\mathbf{B}_n$ such that

$$\frac{\sqrt{k} A_{n,1}}{(k/n)\mathbb{U}(n/k)} = -\frac{\int_0^{1-k/n} \mathbf{B}_n(s) d\mathbb{Q}(s)}{(k/n)^{1/2}\mathbb{U}(n/k)} + o_{\mathbf{P}}(1) = \mathbb{W}_{n,3} \tag{17}$$

and

$$\frac{\sqrt{k} A_{n,2}}{(k/n)\,\mathbb{U}(n/k)} = \sqrt{k} A(n/k) \mathcal{AB}_K(\gamma, \rho) + \mathbb{W}_{1,n} + \mathbb{W}_{2,n}(K), \tag{18}$$

where

$$\mathbb{W}_{1,n} = -\frac{\gamma}{1-\gamma}\sqrt{\frac{n}{k}} \mathbf{B}_n\left(1 - \frac{k}{n}\right)\left(1 + o_{\mathbb{P}}(1)\right);$$

$$\mathbb{W}_{2,n}(K) = \frac{\gamma}{(1-\gamma)^2}\sqrt{\frac{n}{k}}\left\{\int_0^1 \frac{1}{s}\mathbf{B}_n\left(1 - \frac{sk}{n}\right) d\left(sK(s)\right)\right\}.$$

The proof of statement (17) is similar to that of Theorem 1 in Peng (2001) [23] and Theorem 2 in Necir *et al.* (2010) [25], though some adjustments are needed since we are now concerned with the mean. We therefore present main blocks of the proof together with pinpointed references to Necir *et al.* (2010) [25] for specific technical details.

Now, we show the statement (18), we have

$$A_{n,2} = \frac{(k/n)}{1 - \widehat{\gamma}_{n,k}^{K}} \mathbb{U}(Y_{n-k,n}) - \int_0^{k/n} \mathbb{U}(1/s) ds,$$

so

$$\begin{aligned}
\frac{\sqrt{k} A_{n,2}}{(k/n)\mathbb{U}(n/k)} &= \sqrt{k}\frac{1}{1 - \widehat{\gamma}_{n,k}^{K}}\left(\frac{\mathbb{U}(Y_{n-k,n})}{\mathbb{U}(n/k)} - \left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}\right) \\
&+ \sqrt{k}\frac{1}{1 - \widehat{\gamma}_{n,k}^{K}}\left(\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma} - 1\right) \\
&+ \frac{1}{(1 - \widehat{\gamma}_{n,k}^{K})(1-\gamma)}\sqrt{k}\left(\widehat{\gamma}_{n,k}^{K} - \gamma\right) \\
&+ \sqrt{k}\left(\frac{1}{1-\gamma} - \frac{\int_1^{\infty} s^{-2}\mathbb{U}(ns/k) ds}{\mathbb{U}(n/k)}\right) \\
&= T_{n,1} + T_{n,2} + T_{n,3} + T_{n,4}.
\end{aligned}$$

We study each term separately.

For $T_{n,1}$: According to de Haan and Ferreira (2006, p. 60 and Theorem 2.3.9, p. 48) [19], for any $\delta > 0$, we have

$$\begin{aligned}
&\frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - \left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma} \\
&= A_0\left(\frac{n}{k}\right)\left\{\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}\frac{\left(\frac{k}{n}Y_{n-k,n}\right)^{\rho} - 1}{\rho}\right\} \\
&+ A_0\left(\frac{n}{k}\right)\left\{o_{\mathbb{P}}(1)\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma+\rho\pm\delta}\right\},
\end{aligned}$$

where $A_0(t) \sim A(t)$ as $t \to \infty$.

Thus, since $(k/n)Y_{n-k,n} = 1 + o_{\mathbb{P}}(1)$ and $\widehat{\gamma}_{n,k}^{K} \xrightarrow{\mathbb{P}} \gamma$, it readily follows that

$$T_{n,1} = o_{\mathbb{P}}(1).$$

Term $T_{n,2}$**:** The equality $Y_{n-k,n} \overset{d}{=} (1-\zeta_{n-k,n})^{-1}$, yields

$$\begin{aligned}
&\sqrt{k}\left(\left(\frac{k}{n}Y_{n-k,n}\right)^{\gamma}-1\right) \\
&\overset{d}{=} \sqrt{k}\left([(n/k)(1-\zeta_{n-k,n})]^{-\gamma}-1\right) \\
&= -\gamma\sqrt{k}\left([(n/k)(1-\zeta_{n-k,n})]-1\right)(1+o_{\mathbb{P}}(1)) \\
&= -\gamma\sqrt{\frac{n}{k}}\mathbf{B}_n\left(1-\frac{k}{n}\right)(1+o_{\mathbb{P}}(1)) \\
&= -\gamma\sqrt{\frac{n}{k}}\left(\mathbf{B}_n\left(1-\frac{k}{n}\right)\right. \\
&\quad \left.+O_{\mathbb{P}}\left(n^{-v}\right)\left(\frac{k}{n}\right)^{1/2-v}\right)(1+o_{\mathbb{P}}(1)),
\end{aligned}$$

for $0 < v \leq 1/2$, by Csörgő et al. (1986) [7]. Thus, using again that $\widehat{\gamma}_{n,k}^{K} \overset{\mathbb{P}}{\to} \gamma$, it follows that

$$T_{n,2} \overset{d}{=} -\frac{\gamma}{1-\gamma}\sqrt{\frac{n}{k}}\mathbf{B}_n\left(1-\frac{k}{n}\right)(1+o_{\mathbb{P}}(1)) = \mathbb{W}_{1,n}.$$

Term $T_{n,3}$**:** By the consistency in probability of $\widehat{\gamma}_{n,k}^{K}$, we have

$$\begin{aligned}
&\sqrt{k}\left(\widehat{\gamma}_{n,k}^{K}-\gamma-A\left(\frac{n}{k}\right)\int_0^1\frac{K(s)}{s^{\rho}}ds\right) \\
&\overset{d}{=} \gamma\sqrt{\frac{n}{k}}\int_0^1\frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d(sK(s))+o_{\mathbb{P}}(1),
\end{aligned}$$

then, we obtain

$$\begin{aligned}
T_{n,3} &\overset{d}{=} \frac{1}{(1-\gamma)^2}\left\{\sqrt{k}A\left(\frac{n}{k}\right)\int_0^1\frac{K(s)}{s^{\rho}}ds\right\} \\
&+\frac{\gamma}{(1-\gamma)^2}\left\{\sqrt{\frac{n}{k}}\int_0^1\frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d(sK(s))\right\} \\
&+o_{\mathbb{P}}(1) \\
&= \frac{\sqrt{k}}{(1-\gamma)^2}A\left(\frac{n}{k}\right)\int_0^1\frac{K(s)}{s^{\rho}}ds+\mathbb{W}_{2,n}(K)+o_{\mathbb{P}}(1).
\end{aligned}$$

Term $T_{n,4}$**:** A change of variables and an integration by parts yield

$$\begin{aligned}
T_{n,4} &= \sqrt{k}\left(\frac{1}{1-\gamma}-\frac{\int_1^{\infty}s^{-2}\mathbb{U}(ns/k)ds}{\mathbb{U}(n/k)}\right) \\
&= -\sqrt{k}\left(\int_1^{\infty}s^{-2}\left(\frac{\mathbb{U}(ns/k)}{\mathbb{U}(n/k)}-s^{\gamma}\right)ds\right).
\end{aligned}$$

Thus, Theorem 2.3.9 in de Haan and Ferreira (2006) [19] entails that, for $\gamma \in (1/2, 1)$

$$\begin{aligned}
T_{n,4} &= -\sqrt{k}A_0\left(\frac{n}{k}\right)\int_1^{\infty}s^{\gamma-2}\frac{s^{\rho}-1}{\rho}ds(1+o_{\mathbb{P}}(1)) \\
&= \sqrt{k}A\left(\frac{n}{k}\right)\frac{1}{(\gamma-1)(\gamma+\rho-1)}(1+o_{\mathbb{P}}(1)).
\end{aligned}$$

Finally, combining statement (17) and statement (18), we obtain

$$\begin{aligned}
&\sqrt{k}\frac{(\widetilde{\mu}_{n,k}^{K}-\mu)}{(k/n)\,\mathbb{U}(n/k)} \\
&\overset{d}{=} \sqrt{k}A\left(\frac{n}{k}\right)\mathcal{AB}_K(\gamma,\rho)+\mathbb{W}_{1,n}+\mathbb{W}_{2,n}(K)+\mathbb{W}_{3,n} \\
&+o_{\mathbb{P}}(1).
\end{aligned}$$

The sum

$$\sqrt{k}A(\frac{n}{k})\mathcal{AB}_K(\gamma,\rho)+\mathbb{W}_{1,n}+\mathbb{W}_{2,n}(K)+\mathbb{W}_{3,n}$$

is a Gaussian random variable. To calculate its expectation and asymptotic variance, the computations are tedious but quite direct. the classical Sultsky's lemma completes the proof of Theorem 2.1. ☐

Proof of the Theorem 2.3. According to Theorem 2.1 and equation (13), we have

$$\frac{\sqrt{k}(\widehat{\mu}_{n,k}^{K,\widehat{\rho}}-\mu)}{(k/n)\,\mathbb{U}(n/k)} \overset{d}{=} \mathbb{W}_{1,n}+\mathbb{W}_{2,n}(K)+\mathbb{W}_{3,n}+\mathbb{W}_{4,n}+o_{\mathbb{P}}(1),$$

where

$$\begin{aligned}
\mathbb{W}_{4,n} &= \sqrt{k}\left(A(n/k)\mathcal{AB}_K(\gamma,\rho)\right) \\
&\quad -\sqrt{k}\left(\widehat{A}_{n,k}^{L.S}(\widehat{\rho})\mathcal{AB}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right)\frac{X_{n-k,n}}{\mathbb{U}(n/k)}\right) \\
&= -\mathcal{AB}_K(\gamma,\rho)\sqrt{k}\left(\widehat{A}_{n,k}^{L.S}(\widehat{\rho})-A(n/k)\right) \\
&\quad -\sqrt{k}\widehat{A}_{n,k}^{L.S}(\widehat{\rho})\left(\mathcal{AB}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right)-\mathcal{AB}_K(\gamma,\rho)\right) \\
&\quad -\sqrt{k}\widehat{A}_{n,k}^{L.S}(\widehat{\rho})\mathcal{AB}_K\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho}),\widehat{\rho}\right)\left(\frac{X_{n-k,n}}{\mathbb{U}(n/k)}-1\right) \\
&\overset{d}{=} -\mathcal{AB}_K(\gamma,\rho)\gamma(1-\rho) \\
&\quad \times\sqrt{\frac{n}{k}}\left\{\int_0^1\frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d\left(s\left(\underline{K}(s)-K_{\rho}(s)\right)\right)\right\} \\
&\quad +o_{\mathbb{P}}(1).
\end{aligned}$$

By the result of Lemma 5 of Deme et al. (2013) [9], for any consistent estimator $\widehat{\rho}$ of ρ, we have

(19)
$$\begin{aligned}
&\sqrt{k}\left(\widehat{\gamma}_{n,k}^{L.S}(\widehat{\rho})-\gamma\right) \overset{d}{=} \gamma\sqrt{\frac{n}{k}}\int_0^1\frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d\left(sK_{\rho}(s)\right) \\
&\quad +o_{\mathbb{P}}(1),
\end{aligned}$$

and

(20)
$$\begin{aligned}
&\sqrt{k}\left(\widehat{A}_{n,k}^{L.S}(\widehat{\rho})-A\left(\frac{n}{k}\right)\right) \\
&\overset{d}{=} \gamma(1-\rho)\sqrt{\frac{n}{k}}\int_0^1\frac{1}{s}\mathbf{B}_n\left(1-\frac{sk}{n}\right)d\left(s\left(\underline{K}(s)-K_{\rho}(s)\right)\right) \\
&+o_{\mathbb{P}}(1),
\end{aligned}$$

and by using the consistency and the inequality

$$|\frac{e^x-1}{x}-1|\le e^{|x|}-1 \ for\ all\ x\in\mathbb{R}.$$

Moreover, direct computations lead to the desired asymptotic variance which ends the proof of Theorem 2.3. □

Prooof of Corollary 2.5. Recall that K_ρ does not satisfy condition $(\mathcal{K})$ but it can be rewritten as formula (15) with both K and K_ρ satisfying $(\mathcal{K})$. So, following the lines of proof of Theorem 2.3 and using in particular the equation (19) and equation (20), Corollary 2.5 follows. □

ACKNOWLEDGEMENTS

The author are very grateful to the referees and the Associate Editor for their careful reading of the paper and their comments that led to significant improvements of the initial draft. The research was supported by the National Research Program (PNR) of Algeria.

Received 24 September 2013

REFERENCES

[1] Beirlant, J., Dierckx, G., Goegebeur, M. and Matthys, G. (1999). Tail index estimation and an exponential regression model, *Extremes*, **2**, 177–200. MR1771132

[2] Beirlant, J., Matthys, G. and Dierckx, G. (2001). Heavy-tailed distributions and rating. *Astin Bulletin*, **31**, 37–58. MR1945631

[3] Beirlant, J., Dierckx, G., Guillou, A. and Starica, C. (2002). On exponential representations of log-spacings of extreme order statistics, *Extremes*, **5**, 157–180. MR1965977

[4] Beirlant, J. and Teugels, J. (1987). Asymptotic normality of Hill's estimator. Extreme value theory (Oberwolfach, 1987), 148–155, Lecture Notes in Statist., **51**, Springer, New York, 1989.

[5] Bingham, N. H., Goldie, C. M. and Teugels, J. L. (1987). *Regular Variation*. Cambridge University Press, Cambridge. MR0898871

[6] Csörgö, M., Csörgö, S., Horveath, L. and Mason, D. (1986a). Weighted empirical and quantile processes. *Ann. Probab.*, **14**, 31–85. MR0815960

[7] Csörgö, S., Horveath, L. and Mason, D. (1986b). What portion of the sample partial sum asymptotically stable or normal? *Probab. Theory Related Fields*, **72**, 1–16.

[8] Csörgö, M., Deuhevels, P. and Mason, D. (1985). Kernel estimates of the tail index of a distribution. *Annals of Statistics*, **13**, 1050–1077. MR0803758

[9] Deme, E. H., Girard, S. and Guillou, A. (2013). Reduced-bias estimator of the Proportional Hazard Premium for heavy-tailed distributions. *Insurance: Mathematics and Economics*, **52(3)**, 550–559. MR3054747

[10] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin. MR1458613

[11] Feuerverger, A. and Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution, *Annals of Statistics*, **27**, 760–781.

[12] Fraga Alves, M. I., de Haan, L. and Lin, T. (2003). Estimation of the parameter controlling the speed of convergence in extreme value theory. *Math. Methods Statist.*, **12**, 155–176. MR2025356

[13] Gomes, M. I. and Martins, M. J. (2004). Bias reduction and explicit semi-parametric estimation of the tail index. *Journal of Statistical Planning and Inference*, **124**, 361–378.

[14] Gomes, M. I. and Figueiredo, F. (2006). Bias reduction in risk modelling: semi-parametric quantile estimation. *Test*, **15(2)**, 375–396.

[15] Gomes, M. I. and Pestana, S. (2007). A simple second-order reduced bias tail index estimator. *J. Stat. Comput. Simul.*, **77**, 487–504.

[16] Gomes, M. I., Martins, M. J. and Neves, M. (2007). Improving second order reduced bias extreme value index estimation. *RevStat,* **5(2)**, 177–207. MR2365939

[17] Groeneboom, P., Lopuhaà, H. P. and de Wolf, P. P. (2003). Kernel-type estimators for the extreme value index, *Annals of Statistics*, **31**, 1956–1995. MR2036396

[18] Hall, P. and Welsh, A. H. (1985). Adaptative estimates of parameters of regular variation. *Annals of Statistics*, **13**, 331–341.

[19] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer, New York.

[20] Hall, P. (1982). On some simple estimates of an exponent of regular variation. *J. Roy. Statist. Soc.*, **44(1)**, 37–42. MR0655370

[21] Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, **3**, 1163–1174. MR0378204

[22] Ihaka, R. and Gentleman, R. (1996). **R**: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5(3)**, 299–314.

[23] Peng, L. (2001). Estimating the mean of a heavy-tailed distribution. *Stat. Probab. Lett.*, **52**, 255–264. MR1838213

[24] Liang, L. and Qi, Y. (2004). Estimating the first and second order parameters of a heavy tailed distribution. *Australian & New Zealand Journal of Statistics*. **46(2)**, 305–312.

[25] Necir, A., Rassoul, A. and Zitikis, R. (2010). Estimating the Conditional Tail Expectation in the case Heavy-Tailed Losses. *Journal of Probability and Statistics,* 10.1155/596839. MR2609004

[26] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *J. Amer. Statist. Assoc.*, **73**, 812–815. MR0521329

Abdelaziz Rassoul
GEE Laboratory
National High School of Hydraulics
Blida
Algeria
E-mail address: a.rassoul@gmail.ensh