

# Reduced-bias estimator of the Conditional Tail Expectation of heavy-tailed distributions

El Hadji Deme, Stéphane Girard, Armelle Guillou

► **To cite this version:**

El Hadji Deme, Stéphane Girard, Armelle Guillou. Reduced-bias estimator of the Conditional Tail Expectation of heavy-tailed distributions. M. Hallin et al. Mathematical Statistics and Limit Theorems, Springer, pp.105-123, 2015, 10.1007/978-3-319-12442-1\_7. hal-00823260v2

**HAL Id: hal-00823260**

**<https://hal.inria.fr/hal-00823260v2>**

Submitted on 23 Sep 2013

**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.

# Reduced-bias estimator of the Conditional Tail Expectation of heavy-tailed distributions

El hadji DEME<sup>(1)</sup>, Stéphane GIRARD<sup>(2)</sup> and Armelle GUILLOU<sup>(3)</sup>

<sup>(1)</sup> LERSTAD, Université Gaston Berger de Saint-Louis, Sénégal

<sup>(2)</sup> Team Mistis, Inria Grenoble Rhône-Alpes & Laboratoire Jean Kuntzmann  
655, avenue de l'Europe, Montbonnot, 38334 Saint-Ismier Cedex, France

<sup>(3)</sup> Université de Strasbourg et CNRS, IRMA, UMR 7501,  
7 rue René Descartes, 67084 Strasbourg cedex, France

## Abstract

Several risk measures have been proposed in the literature. In this paper, we focus on the estimation of the Conditional Tail Expectation (CTE). Its asymptotic normality has been first established in the literature under the classical assumption that the second moment of the loss variable is finite, this condition being very restrictive in practical applications. Such a result has been extended by Necir *et al.* (2010) in the case of infinite second moment. In this framework, we propose a reduced-bias estimator of the CTE. We illustrate the efficiency of our approach on a small simulation study and a real data analysis.

**Keywords:** Bias correction, heavy-tailed distribution, conditional tail expectation, kernel estimators.

## 1 Introduction

Different risk measures have been proposed in the literature and used to determine the amount of an asset to be kept in reserve in the financial framework. We refer to Goovaerts *et al.* (1984) for various examples and properties. One of the most popular example in hydrology or climatology is undoubtedly the return period. A frequency analysis in hydrology focuses on the estimation of quantities (e.g., flows or annual rainfall) corresponding to a certain return period. It is closely related to the notion of quantile which has therefore been extensively studied. For a real value random variable  $X$  with  $\mathbb{E}[X] < \infty$ , the quantile of order  $1 - \frac{1}{T}$  expresses the magnitude of the event which is exceeded with a probability equal to  $\frac{1}{T}$ .  $T$  is then called the return period. In an

actuarial context, the Value-at Risk (VaR) is defined as the  $p$ -quantile

$$Q(p) = \inf\{x \geq 0 : F(x) \geq p\}, \quad \text{for } p \in [0, 1],$$

with  $F$  the distribution function of the random variable  $X$ . A second important risk measure, based on the quantile notion, is the Conditional Tail Expectation (CTE) defined by

$$CTE_\alpha[X] = \mathbb{E}(X|X > Q(\alpha)), \quad \text{for } \alpha \in (0, 1).$$

The CTE satisfies all the desirable properties of a coherent risk measure (see Artzner *et al.*, 1999) and it provides a more conservative measure of risk than the VaR for the same level of degree of confidence (see Landsman and Valdez, 2003). For all these reasons, the CTE (sometimes referred to as expected shortfall) is preferable in many applications. It thus continues to receive an increased attention in the actuarial literature (see for instance chapters 2 and 7 in McNeil *et al.*, 2005).

In all the sequel we assume that  $F$  is continuous, which allows us to rewrite the  $CTE_\alpha[X]$  as

$$\mathbb{C}_\alpha[X] = \frac{1}{1-\alpha} \int_\alpha^1 Q(s) ds.$$

Clearly, the CTE is unknown since it depends on  $F$ . Hence, it is desirable to define estimators for this quantity and to study their asymptotic properties. To this aim, suppose that we have at our disposal a sample  $(X_1, \dots, X_n)$  of independent and identically distributed random variables from  $F$  and denote by  $X_{1,n} \leq \dots \leq X_{n,n}$  the order statistics. The asymptotic behaviour of  $\mathbb{C}_{\alpha_n}[X]$  has been studied recently in Pan *et al.* (2013) and Zhu and Li (2012) when  $\alpha_n \rightarrow 1$  as  $n \rightarrow \infty$ . On the contrary, in this paper we consider  $\alpha$  fixed. A natural estimator of  $\mathbb{C}_\alpha[X]$  can then be obtained by

$$\widehat{\mathbb{C}}_{n,\alpha}[X] = \frac{1}{1-\alpha} \int_\alpha^1 Q_n(s) ds, \tag{1}$$

where  $Q_n(s)$  is the empirical quantile function, which is equal to the  $i$ th order statistic  $X_{i,n}$  for all  $s \in ((i-1)/n, i/n]$ , and for all  $i = 1, \dots, n$ . The asymptotic behavior of the estimator  $\widehat{\mathbb{C}}_{n,\alpha}[X]$  has been studied by Brazauskas *et al.* (2008), when  $\mathbb{E}[X^2] < \infty$ . Unfortunately, this condition is quite restrictive. For instance, in the case of Pareto-type distributions, defined as  $1 - F(x) = x^{-\frac{1}{\gamma}} \ell_F(x)$  where  $\ell_F$  is a slowly varying function at infinity satisfying  $\ell_F(\lambda x)/\ell_F(x) \rightarrow 1$  as  $x \rightarrow \infty$  for all  $\lambda > 0$ , this condition of second moment implies that  $\gamma \in (0, 1/2)$ . When  $\gamma \in (1/2, 1)$ , we have  $\mathbb{E}[X^2] = \infty$  but nevertheless the CTE is well-defined and finite since  $\mathbb{E}[X] < \infty$ . Note that, in the case  $\gamma = 1/2$ , the finiteness of the second moment depends on the slowly varying function.

This framework will be the subject of this paper where we assume that

$$1 - F(x) = x^{-1/\gamma} \ell_F(x)$$

where  $\gamma > 0$  is the extreme value index. We focus on the case where  $\gamma \in (1/2, 1)$  and thus  $\mathbb{E}[X^2] = \infty$ , this range of values being excluded in the results of Brazauskas *et al.* (2008). The

estimation of  $\gamma$  has been extensively studied in the literature and the most famous estimator is the Hill (1975) estimator defined as:

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

for an intermediate sequence  $k = k(n)$ , i.e. a sequence such that  $k \rightarrow \infty$  and  $k/n \rightarrow 0$  as  $n \rightarrow \infty$ . More generally, Csörgő *et al.* (1985) extended the Hill estimator into a kernel class of estimators

$$\widehat{\gamma}_{n,k}^K = \frac{1}{k} \sum_{j=1}^k K\left(\frac{j}{k+1}\right) Z_{j,k},$$

where  $K$  is a kernel integrating to one and  $Z_{j,k} = j (\log X_{n-j+1,n} - \log X_{n-j,n})$ . Note that the Hill estimator corresponds to the particular case where  $K(u) = \underline{K}(u) := \mathbb{1}_{\{0 < u < 1\}}$ .

Now remark that  $\mathbb{C}_\alpha[X]$  can be rewritten as

$$\begin{aligned} \mathbb{C}_\alpha[X] &= \frac{1}{1-\alpha} \int_\alpha^{1-k/n} Q(s) ds + \frac{1}{1-\alpha} \int_0^{k/n} Q(1-s) ds. \\ &=: \mathbb{C}_\alpha^{(1)}[X] + \mathbb{C}_\alpha^{(2)}[X]. \end{aligned}$$

In this spirit, Necir *et al.* (2010) introduced the following estimator of the CTE, which takes into account different asymptotic properties of moderate and high quantiles in the case of Pareto-type distributions:

$$\begin{aligned} \widetilde{\mathbb{C}}_{n,\alpha}[X] &=: \widetilde{\mathbb{C}}_{n,\alpha}^{(1)}[X] + \widetilde{\mathbb{C}}_{n,\alpha}^{(2)}[X] \\ &= \frac{1}{1-\alpha} \sum_{j=1}^{n-k} \left( \left( \frac{j}{n} - \alpha \right)_+ - \left( \frac{j-1}{n} - \alpha \right)_+ \right) X_{j,n} + \frac{k/n}{(1-\alpha)(1-\widehat{\gamma}_{n,k}^H)} X_{n-k,n} \quad (2) \end{aligned}$$

where  $(s - \alpha)_+$  is the classical notation for the positive part of  $(s - \alpha)$ .

The estimator  $\widetilde{\mathbb{C}}_{n,\alpha}^{(1)}[X]$  is obtained similarly to (1) using the well-known properties of the empirical quantile function  $Q_n$  whereas  $\widetilde{\mathbb{C}}_{n,\alpha}^{(2)}[X]$  is obtained using a Weissman estimator of  $Q$ :  $\widehat{Q}(1-s) := X_{n-k,n} \left(\frac{k}{n}\right)^{\widehat{\gamma}_{n,k}^H} s^{-\widehat{\gamma}_{n,k}^H}$ ,  $s \rightarrow 0$  (see Weissman, 1978).

This estimator may suffer from a high bias in finite sample situations, as illustrated on Figure 1 on a Burr distribution with extreme value index  $\gamma = 2/3$ . Besides, it appears that the bias heavily depends on the intermediate sequence, making the choice of  $k$  difficult in practice.

The goal of this paper is twofold. First, we state an asymptotic normality result for  $\widetilde{\mathbb{C}}_{n,\alpha}[X]$  exhibiting the bias term (Section 2) and thus generalizing the one of Necir *et al.* (2010). Second, the precise knowledge of the first order of the bias allows us to propose a reduced-bias approach. The efficiency of our method is illustrated on a small simulation study and a real dataset in Section 3. All the proofs are postponed to Section 4.

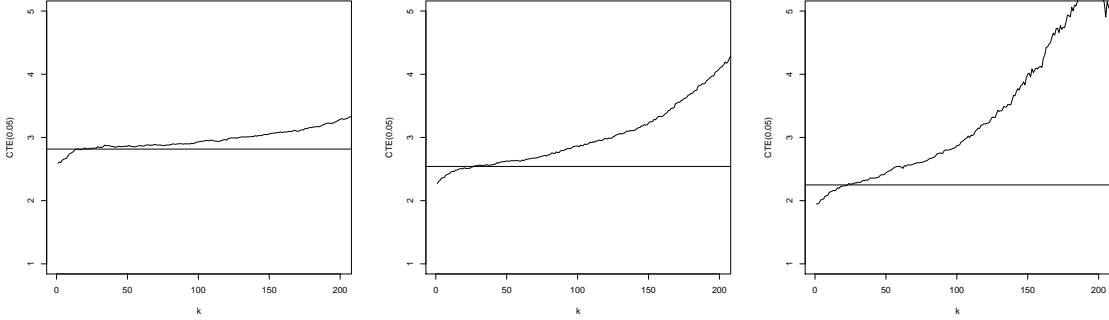


Figure 1: Median of  $\tilde{\mathcal{C}}_{n,0.05}[X]$  as a function of  $k$  based on 500 samples of size 500 from a Burr distribution defined as  $\bar{F}(x) = (1 + x^{-\frac{3\rho}{2}})^{1/\rho}$ . From the left to the right:  $\rho = -1.5$ ,  $\rho = -1$  and  $\rho = -0.75$ . The horizontal line represents the true value of the  $CTE_{0.05}[X]$ .

## 2 Main results

As usual in the extreme value framework, to prove asymptotic normality results, we need a second-order condition on the function  $\mathbb{U}(x) = Q(1 - 1/x)$  such as the following:

**Condition** ( $\mathcal{R}_{\mathbb{U}}$ ). *There exist a function  $A(x) \rightarrow 0$  as  $x \rightarrow \infty$  of constant sign for large values of  $x$  and a second order parameter  $\rho < 0$  such that, for every  $x > 0$ ,*

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

Note that condition ( $\mathcal{R}_{\mathbb{U}}$ ) implies that  $|A|$  is regularly varying with index  $\rho$  (see e.g. Geluk and de Haan, 1987). It is satisfied for most of the classical distribution functions such as the Pareto, Burr and Fréchet ones.

We start to give in Theorem 1, the main expansion of  $\tilde{\mathcal{C}}_{n,\alpha}[X]$  in terms of Brownian bridges, which leads to its asymptotic normality stated in Corollary 1. As it exhibits some bias, we propose a reduced-bias estimator whose expansion is formulated in Theorem 2 and its asymptotic normality is given in Corollary 2.

### 2.1 Asymptotic results for the CTE estimator

**Theorem 1.** *Assume that  $F$  satisfies ( $\mathcal{R}_{\mathbb{U}}$ ) with  $\gamma \in (1/2, 1)$ . Then for any sequence of integer  $k = k(n)$  satisfying  $k \rightarrow \infty$ ,  $k/n \rightarrow 0$  and  $\sqrt{k}A(n/k) = O(1)$  as  $n \rightarrow \infty$ , we have*

$$\frac{n(1-\alpha)}{\sqrt{k}\mathbb{U}(n/k)} \left( \tilde{\mathcal{C}}_{n,\alpha}[X] - \mathcal{C}_\alpha[X] \right) \stackrel{\mathcal{D}}{=} \sqrt{k}A\left(\frac{n}{k}\right) \mathcal{AB}(\gamma, \rho) + \mathbb{W}_{n,1} + \mathbb{W}_{n,2} + \mathbb{W}_{n,3} + o_{\mathbb{P}}(1)$$

where

$$\mathcal{AB}(\gamma, \rho) := \frac{\gamma\rho}{(1-\rho)(\gamma+\rho-1)(1-\gamma)^2}$$

and

$$\left\{ \begin{array}{l} \mathbb{W}_{n,1} := -\frac{\int_0^{1-k/n} \mathbb{B}_n(s) dQ(s)}{\sqrt{k/n} Q(1-k/n)} \\ \mathbb{W}_{n,2} := -\frac{\gamma}{1-\gamma} \sqrt{\frac{n}{k}} \mathbb{B}_n(1-k/n) \\ \mathbb{W}_{n,3} := \frac{\gamma}{(1-\gamma)^2} \sqrt{\frac{n}{k}} \int_0^1 s^{-1} \mathbb{B}_n(1-sk/n) d(s\underline{K}(s)). \end{array} \right.$$

Now, by computing the asymptotic variances of the different processes appearing in Theorem 1, we deduce the following corollary.

**Corollary 1.** *Under the assumptions of Theorem 1, if  $\sqrt{k}A(n/k) \rightarrow \lambda \in \mathbb{R}$ , we have*

$$\frac{n(1-\alpha)}{\sqrt{k}U(n/k)} \left( \tilde{\mathbb{C}}_{n,\alpha}[X] - \mathbb{C}_\alpha[X] \right) \xrightarrow{\mathcal{D}} \mathcal{N}(\lambda \mathcal{AB}(\gamma, \rho), \mathcal{AV}(\gamma))$$

where  $\mathcal{AB}(\gamma, \rho)$  is as above and

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

Since  $\rho < 0$  and  $\gamma \in (1/2, 1)$ , we can easily check that  $\mathcal{AB}(\gamma, \rho)$  is always positive and thus the sign of the function  $A(\cdot)$  determines the sign of the bias of  $\tilde{\mathbb{C}}_{n,\alpha}[X]$ . Note that the asymptotic variance  $\mathcal{AV}(\gamma)$  does not depend on  $\alpha$  and that this result generalizes Theorem 3.1 in Necir *et al.* (2010) in case  $\lambda \neq 0$ . The goal of the next section is to propose a reduced-bias estimator of  $\mathbb{C}_\alpha[X]$ .

## 2.2 Reduced-bias method with the least squared approach

From Theorem 1, it is clear that the estimator  $\tilde{\mathbb{C}}_{n,\alpha}[X]$  exhibits a bias due to the use in its construction of the Weissman's estimator which is known to have such a problem. To overcome this issue, we propose to use the exponential regression model introduced in Feuerverger and Hall (1999) and Beirlant *et al.* (1999) to construct a reduced-bias estimator.

More precisely, using  $(\mathcal{R}_U)$ , Feuerverger and Hall (1999) and Beirlant *et al.* (1999, 2002) proposed the following exponential regression model for the log-spacings of order statistics:

$$Z_{j,k} \sim \left( \gamma + A(n/k) \left( \frac{j}{k+1} \right)^{-\rho} \right) + \varepsilon_{j,k}, \quad 1 \leq j \leq k, \quad (3)$$

where  $\varepsilon_{j,k}$  are zero-centered error terms. If we ignore the term  $A(n/k)$  in (3), we retrieve the Hill-type estimator  $\hat{\gamma}_{n,k}^H$  by taking the mean of the left-hand side of (3). By using a least-squares approach, (3) can be further exploited to propose a reduced-bias estimator of  $\gamma$  in which  $\rho$  is substituted by a consistent estimator  $\hat{\rho} = \hat{\rho}_{n,k}$  (see for instance Beirlant *et al.*, 2002) or by a canonical choice, such as  $\rho = -1$  (see e.g. Feuerverger and Hall (1999) or Beirlant *et al.* (1999)).

The least squares estimators of  $\gamma$  and  $A(n/k)$  are then given by

$$\left\{ \begin{array}{l} \hat{\gamma}_{n,k}^{LS}(\hat{\rho}) = \frac{1}{k} \sum_{j=1}^k Z_{j,k} - \frac{\hat{A}_{n,k}^{LS}(\hat{\rho})}{1 - \hat{\rho}}, \\ \hat{A}_{n,k}^{LS}(\hat{\rho}) = \frac{(1 - 2\hat{\rho})(1 - \hat{\rho})^2}{\hat{\rho}^2} \frac{1}{k} \sum_{j=1}^k \left( \left( \frac{j}{k+1} \right)^{-\hat{\rho}} - \frac{1}{1 - \hat{\rho}} \right) Z_{j,k}. \end{array} \right.$$

The main asymptotic properties of  $\hat{\gamma}_{n,k}^{LS}(\hat{\rho})$  and  $\hat{A}_{n,k}^{LS}(\hat{\rho})$  as a function of Brownian bridges have been established in Deme *et al.* (2013, Lemma 5). Note that  $\hat{\gamma}_{n,k}^{LS}(\hat{\rho})$  can be viewed as a kernel estimator

$$\hat{\gamma}_{n,k}^{LS}(\hat{\rho}) = \frac{1}{k} \sum_{j=1}^k K_{\hat{\rho}} \left( \frac{j}{k+1} \right) Z_{j,k},$$

where for  $0 < u \leq 1$ :

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

with  $\underline{K}_{\rho}(u) = ((1 - \rho)/\rho)(u^{-\rho} - 1)\mathbb{1}_{\{0 < u < 1\}}$ .

Now, using the second order refinements of assumption  $(\mathcal{R}_U)$ , we can construct the following asymptotically unbiased estimator of the quantile:

$$\hat{Q}^{LS, \hat{\rho}}(1 - s) = (ns/k)^{-\hat{\gamma}_{n,k}^{LS}(\hat{\rho})} X_{n-k,n} \left( 1 - \hat{\rho}^{-1} \hat{A}_{n,k}^{LS}(\hat{\rho}) \left( 1 - (ns/k)^{-\hat{\rho}} \right) \right),$$

see e.g. Matthys *et al.* (2004).

Thus, in the spirit of (2), we arrive at the following asymptotically unbiased estimator of  $\mathbb{C}_{\alpha}[X]$

$$\begin{aligned} \tilde{\mathbb{C}}_{n,\alpha}^{LS, \hat{\rho}}[X] &:= \frac{1}{1 - \alpha} \sum_{j=1}^{n-k} \left( \binom{j}{n} - \alpha \right)_+ - \left( \binom{j-1}{n} - \alpha \right)_+ X_{j,n} \\ &+ \frac{k/n}{(1 - \alpha)(1 - \hat{\gamma}_{n,k}^{LS}(\hat{\rho}))} \left( 1 - \frac{\hat{A}_{n,k}^{LS}(\hat{\rho})}{\hat{\gamma}_{n,k}^{LS}(\hat{\rho}) + \hat{\rho} - 1} \right) X_{n-k,n}. \end{aligned}$$

Our next goal is to establish, under suitable assumptions, the asymptotic normality of  $\tilde{\mathbb{C}}_{n,\alpha}^{LS, \hat{\rho}}[X]$ . This is done in the following theorem.

**Theorem 2.** *Under the assumptions of Theorem 1, if  $\hat{\rho}$  is a consistent estimator of  $\rho$ , then we have*

$$\frac{n(1 - \alpha)}{\sqrt{k}U(n/k)} \left( \tilde{\mathbb{C}}_{n,\alpha}^{LS, \hat{\rho}}[X] - \mathbb{C}_{\alpha}[X] \right) \stackrel{\mathcal{D}}{=} \mathbb{W}_{n,1} + \mathbb{W}_{n,2} + \mathbb{W}_{n,4} + \mathbb{W}_{n,5} + o_{\mathbb{P}}(1)$$

where  $\mathbb{W}_{n,1}$ ,  $\mathbb{W}_{n,2}$  and  $\mathbb{W}_{n,3}$  are defined in Theorem 1, and

$$\left\{ \begin{array}{l} \mathbb{W}_{n,4} := \frac{\rho\gamma^2}{(\gamma + \rho - 1)(1 - \gamma)^2} \sqrt{\frac{n}{k}} \int_0^1 s^{-1} \mathbb{B}_n(1 - sk/n) d(sK_{\rho}(s)) \\ \mathbb{W}_{n,5} := -\frac{(1 - \gamma)(1 - \rho)}{\gamma + \rho - 1} \mathbb{W}_{n,3}. \end{array} \right.$$

Now, by computing the asymptotic variances of the different processes appearing in Theorem 2, we deduce the following corollary.

**Corollary 2.** *Under the assumptions of Theorem 1, if  $\widehat{\rho}$  is a consistent estimator of  $\rho$ , then we have*

$$\frac{n(1-\alpha)}{\sqrt{k}\mathbb{U}(n/k)} \left( \widetilde{\mathbb{C}}_{n,\alpha}^{\mathbf{LS},\widehat{\rho}}[X] - \mathbb{C}_\alpha[X] \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( 0, \widetilde{\mathcal{AV}}(\gamma, \rho) \right)$$

with

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

As expected, the asymptotic bias of our new estimator of the CTE is equal to zero whereas its asymptotic variance  $\widetilde{\mathcal{AV}}(\gamma, \rho)$  is larger than the one of the original estimator  $\mathcal{AV}(\gamma)$  exhibited in Corollary 1.

### 3 Finite sample behavior

#### 3.1 A small simulation study

In this section, the biased estimator  $\widetilde{\mathbb{C}}_{n,\alpha}[X]$  and the reduced-bias one  $\widetilde{\mathbb{C}}_{n,\alpha}^{\mathbf{LS},-1}[X]$  are compared on a small simulation study. To this aim, 500 samples of size 500 are simulated from a Burr distribution defined as:  $\overline{F}(x) = (1 + x^{-\frac{3}{2}\rho})^{1/\rho}$ . The associated extreme value index is  $\gamma = 2/3$  and  $\rho$  is the second order parameter. Three values for  $\alpha \in \{0.05, 0.10, 0.20\}$  are used and different values of  $\rho \in \{-0.75, -1, -1.5\}$  are considered to assess its impact. The median and median squared error (MSE) of these estimators are estimated over the 500 replications. The results are displayed in Figure 2 and Figure 3. It appears on Figure 2 that the closer  $\rho$  is to 0, the more important is the bias of  $\widetilde{\mathbb{C}}_{n,\alpha}[X]$  whatever the value of  $\alpha$  is. The effect of the bias correction on the MSE is illustrated on Figure 3. We can observe that the MSE of the reduced-bias estimator  $\widetilde{\mathbb{C}}_{n,\alpha}^{\mathbf{LS},-1}[X]$  is almost constant with respect to  $k$ , especially when the bias of  $\widetilde{\mathbb{C}}_{n,\alpha}[X]$  is strong, *i.e.* when  $\rho$  is close to 0.

#### 3.2 Real data analysis

Our real dataset concerns a Norwegian fire insurance portfolio from 1972 until 1992. Together with the year of occurrence, the data contain the value ( $\times 1000$  Krone) of the claims. A priority of 500 units was in force. These data were of some concern in that the number of claims had risen systematically with a maximum in 1988 as illustrated in Figure 4(a). We concentrate here on the year 1976 where the average claim size per year reached a peak as was the case in 1988. The sample size is  $n = 207$ . Figure 4(b) shows the histogram corresponding to this year 1976. From



Figure 4(c) we can observe the difficulty to find a stable part in the plot of the Hill estimator  $\hat{\gamma}_{n,k}^H$  as a function of  $k$ , due to the bias of this estimator. We can apply our methodology to this real dataset as the extreme value index (or at least its estimator) is in the interval  $(1/2, 1)$  whatever the value of  $k$  is. Figure 4(d)-(f) shows the biased estimator  $\tilde{\mathbb{C}}_{n,\alpha}[X]$  (dashed line) and the reduced-bias one  $\tilde{\mathbb{C}}_{n,\alpha}^{\text{LS},-1}[X]$  (full line) for three different values of  $\alpha$ : 0.05, 0.10 and 0.20. The reduced-bias estimator  $\tilde{\mathbb{C}}_{n,\alpha}^{\text{LS},-1}[X]$  is almost constant for a large range of values of  $k$  which makes the choice of  $k$  easier in practice.

## 4 Proofs

Let  $Y_1, \dots, Y_n$  be independent and identically distributed random variables from the unit Pareto distribution  $G$ , defined as  $G(y) = 1 - y^{-1}$ ,  $y \geq 1$ . For each  $n \geq 1$ , let  $Y_{1,n} \leq \dots \leq Y_{n,n}$  be the order statistics pertaining to  $Y_1, \dots, Y_n$ . Clearly  $X_{j,n} \stackrel{\mathcal{D}}{=} \mathbb{U}(Y_{j,n})$ ,  $j = 1, \dots, n$ . In order to use the results from Csörgő *et al.* (1986), a probability space  $(\Omega, \mathbb{A}, \mathbb{P})$  is constructed carrying a sequence  $\xi_1, \xi_2, \dots$  of independent random variables uniformly distributed on  $(0, 1)$  and a sequence of Brownian bridges  $\mathbb{B}_n(s)$ ,  $0 \leq s \leq 1$ ,  $n = 1, 2, \dots$ . The resulting empirical quantile is denoted by

$$\beta_n(t) = \sqrt{n} (t - \mathbb{V}_n(t))$$

where  $\mathbb{V}_n(s) = \xi_{j,n}$ ,  $\frac{j-1}{n} < s \leq \frac{j}{n}$ ,  $j = 1, \dots, n$  and  $\mathbb{V}_n(0) = 0$ .

The following lemma gives an asymptotic expansion for the second random term appearing in (2).

**Lemma 1.** *Under the assumptions of Theorem 1, we have*

$$\frac{n(1-\alpha)}{\sqrt{k}\mathbb{U}(n/k)} \left( \tilde{\mathbb{C}}_{n,\alpha}^{(2)}[X] - \mathbb{C}_\alpha^{(2)}[X] \right) \stackrel{\mathcal{D}}{=} \sqrt{k}A \left( \frac{n}{k} \right) \mathcal{AB}(\gamma, \rho) + \mathbb{W}_{n,2} + \mathbb{W}_{n,3} + o_{\mathbb{P}}(1).$$

**Proof of Lemma 1.** Note that  $\tilde{\mathbb{C}}_{n,\alpha}^{(2)}[X]$  can be rewritten as follows

$$(1-\alpha)\tilde{\mathbb{C}}_{n,\alpha}^{(2)}[X] = \frac{k/n}{1 - \hat{\gamma}_{n,k}^H} \mathbb{U}(Y_{n-k,n}).$$

As a consequence, the following expansion holds:

$$\frac{n(1-\alpha)}{\sqrt{k}\mathbb{U}(n/k)} \left( \tilde{\mathbb{C}}_{n,\alpha}^{(2)}[X] - \mathbb{C}_\alpha^{(2)}[X] \right) = \sum_{j=1}^4 T_{n,j},$$

where

$$\begin{aligned} T_{n,1} &:= \frac{\sqrt{k}}{1 - \hat{\gamma}_{n,k}^H} \left[ \frac{\mathbb{U}(Y_{n-k,n})}{\mathbb{U}(n/k)} - \left( \frac{k}{n} Y_{n-k,n} \right)^\gamma \right], \\ T_{n,2} &:= \frac{\sqrt{k}}{1 - \hat{\gamma}_{n,k}^H} \left[ \left( \frac{k}{n} Y_{n-k,n} \right)^\gamma - 1 \right], \\ T_{n,3} &:= \frac{1}{(1 - \hat{\gamma}_{n,k}^H)(1 - \gamma)} \sqrt{k} (\hat{\gamma}_{n,k}^H - \gamma), \\ T_{n,4} &:= \frac{n}{\sqrt{k}\mathbb{U}(n/k)} \left[ \frac{k/n}{1 - \gamma} \mathbb{U}(n/k) - (1 - \alpha)\mathbb{C}_\alpha^{(2)}[X] \right]. \end{aligned}$$

We study each term separately.

Term  $T_{n,1}$ . According to de Haan and Ferreira (2006, Theorem 2.3.9), for any  $\delta > 0$ , we have

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

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

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

$$T_{n,1} = o_{\mathbb{P}}(1). \quad (4)$$

Term  $T_{n,2}$ . The equality  $Y_{n-k,n} \stackrel{\mathcal{D}}{=} (1 - \xi_{n-k,n})^{-1}$  yields

$$\begin{aligned} \sqrt{k} \left[ \left(\frac{k}{n}Y_{n-k,n}\right)^\gamma - 1 \right] &\stackrel{\mathcal{D}}{=} \sqrt{k} \left( \left(\frac{n}{k}(1 - \xi_{n-k,n})\right)^{-\gamma} - 1 \right) \\ &= -\gamma\sqrt{k} \left(\frac{n}{k}(1 - \xi_{n-k,n}) - 1\right) (1 + o_{\mathbb{P}}(1)) \quad \text{by a Taylor expansion} \\ &= -\gamma\sqrt{\frac{n}{k}}\beta_n \left(1 - \frac{k}{n}\right) (1 + o_{\mathbb{P}}(1)) \\ &= -\gamma\sqrt{\frac{n}{k}} \left( \mathbb{B}_n \left(1 - \frac{k}{n}\right) + O_{\mathbb{P}}(n^{-\nu}) \left(\frac{k}{n}\right)^{1/2-\nu} \right) (1 + o_{\mathbb{P}}(1)), \end{aligned}$$

for  $0 \leq \nu < 1/2$ , by Csörgő *et al.* (1986). Thus, using again that  $\widehat{\gamma}_{n,k}^H \xrightarrow{\mathbb{P}} \gamma$ , it follows that

$$T_{n,2} \stackrel{\mathcal{D}}{=} -\frac{\gamma}{1-\gamma} \sqrt{\frac{n}{k}} \mathbb{B}_n \left(1 - \frac{k}{n}\right) (1 + o_{\mathbb{P}}(1)) = \mathbb{W}_{n,2} + o_{\mathbb{P}}(1). \quad (5)$$

Term  $T_{n,3}$ . According to Theorem 1 in Deme *et al* (2013) and by the consistency in probability of  $\widehat{\gamma}_{n,k}^H$ , we have

$$\begin{aligned} T_{n,3} &\stackrel{\mathcal{D}}{=} \frac{1}{(1-\gamma)^2} \left\{ \frac{\sqrt{k}A(n/k)}{1-\rho} + \gamma\sqrt{\frac{n}{k}} \int_0^1 s^{-1} \mathbb{B}_n \left(1 - s\frac{k}{n}\right) d(s\underline{K}(s)) \right\} + o_{\mathbb{P}}(1) \\ &= \frac{1}{(1-\rho)(1-\gamma)^2} \sqrt{k}A(n/k) + \mathbb{W}_{n,3} + o_{\mathbb{P}}(1). \end{aligned} \quad (6)$$

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

$$T_{n,4} = \sqrt{k} \left\{ \frac{1}{1-\gamma} - \int_1^\infty x^{-2} \frac{\mathbb{U}(nx/k)}{\mathbb{U}(n/k)} dx \right\} = -\sqrt{k} \int_1^\infty x^{-2} \left( \frac{\mathbb{U}(nx/k)}{\mathbb{U}(n/k)} - x^\gamma \right) dx.$$

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

$$\begin{aligned} T_{n,4} &= -\sqrt{k}A_0\left(\frac{n}{k}\right) \int_1^\infty x^{\gamma-2} \frac{x^\rho - 1}{\rho} dx (1 + o(1)) \\ &= \sqrt{k}A\left(\frac{n}{k}\right) \frac{1}{(1-\gamma)(\gamma+\rho-1)} (1 + o(1)). \end{aligned} \quad (7)$$

Combining (4)-(7), Lemma 1 follows. ■

**Proof of Theorem 1.** Combining Lemma 1 with statement (4.3) in Necir *et al.* (2010), we get

$$\frac{n(1-\alpha)}{\sqrt{k}\mathbb{U}(n/k)} \left( \tilde{\mathbb{C}}_{n,\alpha}[X] - \mathbb{C}_\alpha[X] \right) \stackrel{\mathcal{D}}{=} \sqrt{k}A \left( \frac{n}{k} \right) \mathcal{AB}(\gamma, \rho) + \mathbb{W}_{n,1} + \mathbb{W}_{n,2} + \mathbb{W}_{n,3} + o_{\mathbb{P}}(1).$$

Theorem 1 is thus established.  $\blacksquare$

**Proof of Corollary 1.** From Theorem 1, we only have to compute the asymptotic variance of the limiting process. The computations are tedious but quite direct. We only give below the main arguments, i.e.

$$\begin{aligned} \mathbb{E}W_{n,1}^2 &= \frac{\int_0^{1-k/n} (1-t) \left( \int_0^t s dQ(s) \right) dQ(t)}{k/n Q^2(1-k/n)} + \frac{\int_0^{1-k/n} t \left( \int_t^{1-k/n} (1-s) dQ(s) \right) dQ(t)}{k/n Q^2(1-k/n)} \\ &= \frac{\int_{k/n}^1 u \left( \int_u^1 dQ(1-v) \right) dQ(1-u)}{k/n Q^2(1-k/n)} - \frac{\int_{k/n}^1 u \left( \int_u^1 v dQ(1-v) \right) dQ(1-u)}{k/n Q^2(1-k/n)} \\ &\quad + \frac{\int_{k/n}^1 \left( \int_{k/n}^u v dQ(1-v) \right) dQ(1-u)}{k/n Q^2(1-k/n)} - \frac{\int_{k/n}^1 u \left( \int_{k/n}^u v dQ(1-v) \right) dQ(1-u)}{k/n Q^2(1-k/n)} \\ &=: Q_{1,n} + Q_{2,n} + Q_{3,n} + Q_{4,n}. \end{aligned}$$

Recall now that  $Q(1-s) = s^{-\gamma} \ell(s)$  with  $\ell$  a slowly varying function at 0. By integration by parts and using Lemma 6 in Deme *et al.* (2013), it follows that

$$Q_{1,n} = \frac{1}{2} \left[ 1 + \frac{\int_{k/n}^1 Q^2(1-u) du}{k/n Q^2(1-k/n)} \right] \rightarrow \frac{\gamma}{2\gamma-1}.$$

Now remark that  $d \left( \int_u^1 v dQ(1-v) \right) = -u dQ(1-u)$  which implies that

$$Q_{2,n} = -\frac{1}{2} \frac{k}{n} \left[ \frac{\int_{k/n}^1 v dQ(1-v)}{k/n Q(1-k/n)} \right]^2 = o(1) \tag{8}$$

this last result coming from the fact that, according to Proposition 1.3.6 in Bingham *et al.* (1987): for all  $\varepsilon > 0$ ,  $x^{-\varepsilon} \ell(x) \rightarrow \infty$  as  $x \rightarrow 0$ . Thus, choosing  $0 < \varepsilon < \gamma - \frac{1}{2}$  entails

$$\begin{aligned} 0 \leq s \left( \frac{\int_s^1 t d(Q(1-t))}{sQ(1-s)} \right)^2 &= s \left( 1 + \frac{\int_s^1 t^{-\gamma} \ell(t) dt}{s^{1-\gamma} \ell(s)} \right)^2 \\ &\leq s (1 + C s^{\gamma-1-\varepsilon})^2 = O \left( s^{1+2[\gamma-1-\varepsilon]} \right) = o(1) \end{aligned}$$

where  $C$  is a suitable constant. Consequently,  $Q_{2,n} \rightarrow 0$ . The two others terms,  $Q_{3,n}$  and  $Q_{4,n}$ , can be treated similarly, leading to

$$\begin{aligned} Q_{3,n} = Q_{1,n} &\rightarrow \frac{\gamma}{2\gamma-1} \\ Q_{4,n} = Q_{2,n} &\rightarrow 0. \end{aligned}$$

Finally,

$$\mathbb{E}W_{n,1}^2 \rightarrow \frac{2\gamma}{2\gamma-1},$$

and direct computations now lead to

$$\begin{aligned}
\mathbb{E}W_{n,2}^2 &\longrightarrow \frac{\gamma^2}{(1-\gamma)^2} \\
\mathbb{E}W_{n,3}^2 &\longrightarrow \frac{\gamma^2}{(1-\gamma)^4} \text{ by Corollary 1 in Deme et al. (2013)} \\
\mathbb{E}(W_{n,1}W_{n,2}) &\longrightarrow \frac{\gamma}{1-\gamma} \text{ by (8)} \\
\mathbb{E}(W_{n,1}W_{n,3}) &= 0 \\
\mathbb{E}(W_{n,2}W_{n,3}) &= 0.
\end{aligned}$$

Combining all these results, Corollary 1 follows. ■

**Proof of Theorem 2.** We use the following decomposition

$$\frac{n(1-\alpha)}{\sqrt{k}\mathbb{U}(n/k)} \left( \tilde{\mathbb{C}}_{n,\alpha}^{\text{LS},\hat{\rho}}[X] - \mathbb{C}_\alpha[X] \right) = \sum_{i=1}^7 S_{n,i}$$

where

$$\begin{aligned}
S_{n,1} &= \frac{n(1-\alpha)}{\sqrt{k}\mathbb{U}(n/k)} \left( \tilde{\mathbb{C}}_{n,\alpha}^{(1)}[X] - \mathbb{C}_\alpha^{(1)}[X] \right) \\
S_{n,2} &= \frac{1}{1-\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho})} \left( 1 - \frac{\hat{A}_{n,k}^{\text{LS}}(\hat{\rho})}{\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho}) + \hat{\rho} - 1} \right) \sqrt{k} \left[ \frac{\mathbb{U}(Y_{n-k,n})}{\mathbb{U}(n/k)} - \left( \frac{k}{n} Y_{n-k,n} \right)^\gamma \right] \\
S_{n,3} &= \frac{1}{1-\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho})} \left( 1 - \frac{\hat{A}_{n,k}^{\text{LS}}(\hat{\rho})}{\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho}) + \hat{\rho} - 1} \right) \sqrt{k} \left[ \left( \frac{k}{n} Y_{n-k,n} \right)^\gamma - 1 \right] \\
S_{n,4} &= \frac{1}{(1-\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho}))(1-\gamma)} \sqrt{k} (\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho}) - \gamma) \\
S_{n,5} &= \sqrt{k} A(n/k) \left[ \frac{1}{(1-\gamma)(\gamma + \rho - 1)} - \frac{1}{(1-\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho})) (\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho}) + \hat{\rho} - 1)} \right] \\
S_{n,6} &= -\frac{1}{(1-\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho})) (\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho}) + \hat{\rho} - 1)} \sqrt{k} \left( \hat{A}_{n,k}^{\text{LS}}(\hat{\rho}) - A(n/k) \right) \\
S_{n,7} &= \frac{n}{\sqrt{k}\mathbb{U}(n/k)} \left[ \frac{k/n}{1-\gamma} \left( 1 - \frac{A(n/k)}{\gamma + \rho - 1} \right) \mathbb{U}(n/k) - (1-\alpha) \mathbb{C}_\alpha^{(2)}[X] \right].
\end{aligned}$$

Now, we are going to study separately the terms  $S_{n,1}, \dots, S_{n,7}$ .

Term  $S_{n,1}$ . Statement (4.3) in Necir *et al.* (2010) leads to

$$S_{n,1} = \mathbb{W}_{n,1} + o_{\mathbb{P}}(1). \tag{9}$$

Term  $S_{n,2}$ . Note that

$$S_{n,2} = \frac{1-\hat{\gamma}_{n,k}^H}{1-\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho})} \left( 1 - \frac{\hat{A}_{n,k}^{\text{LS}}(\hat{\rho})}{\hat{\gamma}_{n,k}^{\text{LS}}(\hat{\rho}) + \hat{\rho} - 1} \right) T_{n,1}$$

where  $T_{n,1}$  is defined in the proof of Lemma 1. Thus combining Lemma 5 in Deme *et al.* (2013) with the consistency of  $\widehat{\boldsymbol{\rho}}$  and (4), we obtain that

$$S_{n,2} = o_{\mathbb{P}}(1). \quad (10)$$

Term  $S_{n,3}$ . Similarly, we observe that  $S_{n,3} = T_{n,2}(1 + o_{\mathbb{P}}(1))$  where  $T_{n,2}$  is defined in the proof of Lemma 1. Thus according to (5), we have

$$S_{n,3} \stackrel{\mathcal{D}}{=} \mathbb{W}_{n,2} + o_{\mathbb{P}}(1). \quad (11)$$

Term  $S_{n,4}$ . Combining Lemma 5 in Deme *et al.* (2013) with the consistency of  $\widehat{\gamma}_{n,k}^{\mathbf{LS}}(\widehat{\boldsymbol{\rho}})$ , we infer that

$$S_{n,4} \stackrel{\mathcal{D}}{=} \frac{\gamma + \rho - 1}{\rho\gamma} \mathbb{W}_{n,4} + o_{\mathbb{P}}(1). \quad (12)$$

Term  $S_{n,5}$ . Under the assumption that  $\sqrt{k}A(n/k) = O(1)$  and by the consistency of  $\widehat{\boldsymbol{\rho}}$  and  $\widehat{\gamma}_{n,k}^{\mathbf{LS}}(\widehat{\boldsymbol{\rho}})$  we have

$$S_{n,5} = o_{\mathbb{P}}(1). \quad (13)$$

Term  $S_{n,6}$ . Using Lemma 5 in Deme *et al.* (2013), we get

$$\begin{aligned} S_{n,6} &\stackrel{\mathcal{D}}{=} -\frac{\gamma(1-\rho)}{(1-\gamma)(\gamma+\rho-1)} \sqrt{\frac{n}{k}} \int_0^1 s^{-1} \mathbb{B}_n \left(1 - \frac{sk}{n}\right) d(s(\underline{K}(s) - K_{\rho}(s))) + o_{\mathbb{P}}(1) \\ &= -\frac{(1-\rho)(1-\gamma)}{\gamma+\rho-1} \left( \mathbb{W}_{n,3} - \frac{\gamma+\rho-1}{\rho\gamma} \mathbb{W}_{n,4} \right) + o_{\mathbb{P}}(1) \\ &= \mathbb{W}_{n,5} + \frac{(1-\rho)(1-\gamma)}{\gamma\rho} \mathbb{W}_{n,4} + o_{\mathbb{P}}(1). \end{aligned} \quad (14)$$

Term  $S_{n,7}$ . Remark that

$$S_{n,7} = -\frac{\sqrt{k}A(n/k)}{(1-\gamma)(\gamma+\rho-1)} + T_{n,4},$$

where  $T_{n,4}$  is defined in the proof of Lemma 1. Thus using (7) and the assumption that  $\sqrt{k}A(n/k) = O(1)$ , we deduce that

$$S_{n,7} = o_{\mathbb{P}}(1). \quad (15)$$

Combining (9)-(15), Theorem 2 follows. ■

**Proof of Corollary 2.** From Theorem 2, we only have to compute the asymptotic variance of the limiting process. As in Corollary 1, the computations are quite direct and the desired asymptotic

variance can be obtained by noticing that

$$\begin{aligned}
\mathbb{E}W_{n,5}^2 &\longrightarrow \frac{\gamma^2(1-\rho)^2}{(1-\gamma)^2(\gamma+\rho-1)^2} \\
\mathbb{E}(W_{n,1}W_{n,4}) &= 0 \\
\mathbb{E}(W_{n,1}W_{n,5}) &= 0 \\
\mathbb{E}W_{n,4}^2 &= \frac{\gamma^4(1-\rho)^2}{(1-\gamma)^4(\gamma+\rho-1)^2} \\
\mathbb{E}(W_{n,2}W_{n,5}) &= 0 \\
\mathbb{E}(W_{n,2}W_{n,4}) &= 0 \\
\mathbb{E}(W_{n,4}W_{n,5}) &= -\frac{\rho\gamma^3(1-\rho)}{(1-\gamma)^3(\gamma+\rho-1)^2}.
\end{aligned}$$

■

## Acknowledgements

The first author acknowledges support from AIRES-Sud (AIRES-Sud is a program from the French Ministry of Foreign and European Affairs, implemented by the "Institut de Recherche pour le Développement", IRD-DSF) and from the "Ministère de la Recherche Scientifique" of Sénégal.

## References

- [1] Artzner, P., Delbaen, F., Eber, J-M., Heath, D. (1999). Coherent measures of risk, *Mathematical Finance*, **9**, 203-228.
- [2] Beirlant, J., Dierckx, G., Goegebeur, M., Matthys, G. (1999). Tail index estimation and an exponential regression model, *Extremes*, **2**, 177-200.
- [3] Beirlant, J., Dierckx, G., Guillou, A., Starica, C. (2002). On exponential representations of log-spacings of extreme order statistics, *Extremes*, **5**, 157-180.
- [4] Bingham, N.H., Goldie, C.M., Teugels, J.L. (1987). *Regular variation*, Cambridge.
- [5] Brazauskas, V., Jones, B., Puri, M., Zitikis, R. (2008). Estimating conditional tail expectation with actuarial applications in view, *Journal of Statistical Planning and Inference*, **138**, 3590-3604.
- [6] Csörgő, M., Csörgő, S., Horváth, L., Mason, D.M. (1986). Weighted empirical and quantile processes, *Annals of Probability*, **14**, 31-85.
- [7] Csörgő, S., Deheuvels, P., Mason, D.M. (1985). Kernel estimates of the tail index of a distribution, *Annals of Statistics*, **13**, 1050-1077.

- [8] de Haan, L., Ferreira, A. (2006). *Extreme value theory: an introduction*, Springer.
- [9] Deme, E., Girard, S., Guillou, A. (2013). Reduced-bias estimator of the Proportional Hazard Premium for heavy-tailed distributions, *Insurance Mathematic & Economics*, **52**, 550-559.
- [10] Feuerverger, A., Hall, P. (1999). Estimating a tail exponent by modelling departure from a Pareto distribution, *Annals of Statistics*, **27**, 760-781.
- [11] Geluk, J.L., de Haan, L. (1987). *Regular variation, extensions and Tauberian theorems*, CWI tract 40, Center for Mathematics and Computer Science, P.O. Box 4079, 1009 AB Amsterdam, The Netherlands.
- [12] Goovaerts, M.J., de Vlyder, F., Haezendonck, J. (1984). *Insurance premiums, theory and applications*, North Holland, Amsterdam.
- [13] Hill, B. M. (1975). A simple approach to inference about the tail of a distribution, *Annals of Statistics*, **3**, 1136–1174.
- [14] Landsman, Z., Valdez, E. (2003). Tail conditional expectations for elliptical distributions, *North American Actuarial Journal*, **7**, 55-71.
- [15] Matthys, G., Delafosse, E., Guillou, A., Beirlant, J. (2004). Estimating catastrophic quantile levels for heavy-tailed distributions, *Insurance Mathematic & Economics*, **34**, 517-537.
- [16] McNeil, A.J., Frey, R., Embrechts, P. (2005). *Quantitative risk management: concepts, techniques, and tools*, Princeton University Press.
- [17] Necir, A., Rassoul, A., Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses, *Journal of Probability and Statistics*, **ID 596839**, 17 pp.
- [18] Pan, X., Leng, X., Hu, T. (2013). Second-order version of Karamata's theorem with applications, *Statistics and Probability Letters*, **83**, 1397-1403.
- [19] Weissman, I., (1978). Estimation of parameters and larges quantiles based on the  $k$  largest observations, *Journal of American Statistical Association*, **73**, 812–815.
- [20] Zhu, L., Li, H. (2012). Asymptotic analysis of multivariate tail conditional expectations, *North American Actuarial Journal*, **16**, 350-363.

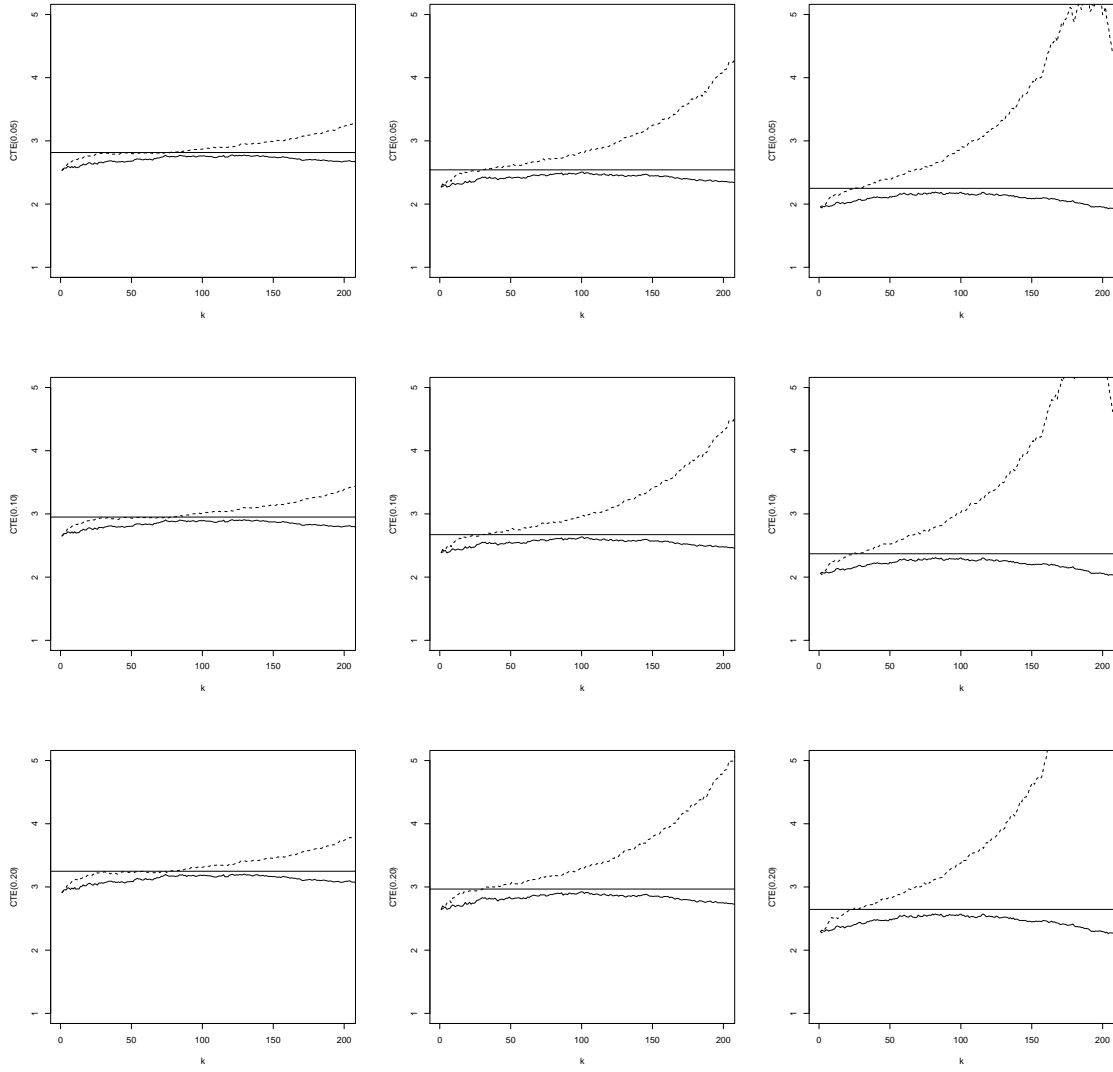


Figure 2: Median of  $\tilde{\mathcal{C}}_{n,\alpha}[X]$  (dotted line) and  $\tilde{\mathcal{C}}_{n,\alpha}^{LS,-1}[X]$  (full line) as a function of  $k$  based on 500 samples of size 500 for  $\alpha = 0.05$  (top),  $\alpha = 0.10$  (middle) and  $\alpha = 0.20$  (bottom) from a Burr distribution defined as  $\bar{F}(x) = (1 + x^{-\frac{3\rho}{2}})^{1/\rho}$ . From the left to the right:  $\rho = -1.5$ ,  $\rho = -1$  and  $\rho = -0.75$ . The horizontal line represents the true value of the  $CTE_\alpha[X]$ .



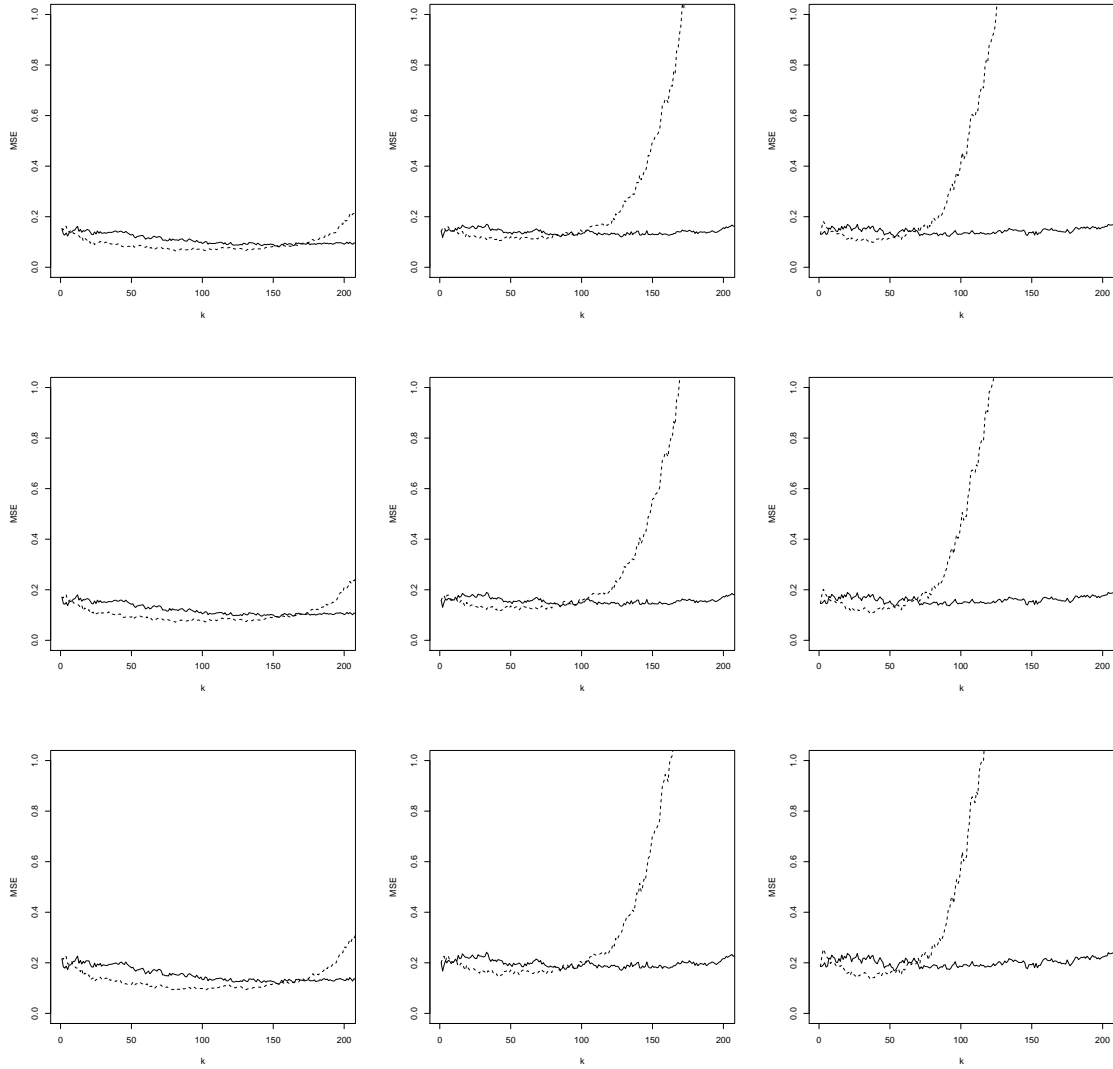


Figure 3: MSE of  $\tilde{C}_{n,\alpha}[X]$  (dotted line) and  $\tilde{C}_{n,\alpha}^{LS,-1}[X]$  (full line) as a function of  $k$  based on 500 samples of size 500 for  $\alpha = 0.05$  (top),  $\alpha = 0.10$  (middle) and  $\alpha = 0.20$  (bottom) from a Burr distribution defined as  $\bar{F}(x) = (1 + x^{-\frac{3\rho}{2}})^{1/\rho}$ . From the left to the right:  $\rho = -1.5$ ,  $\rho = -1$  and  $\rho = -0.75$ .

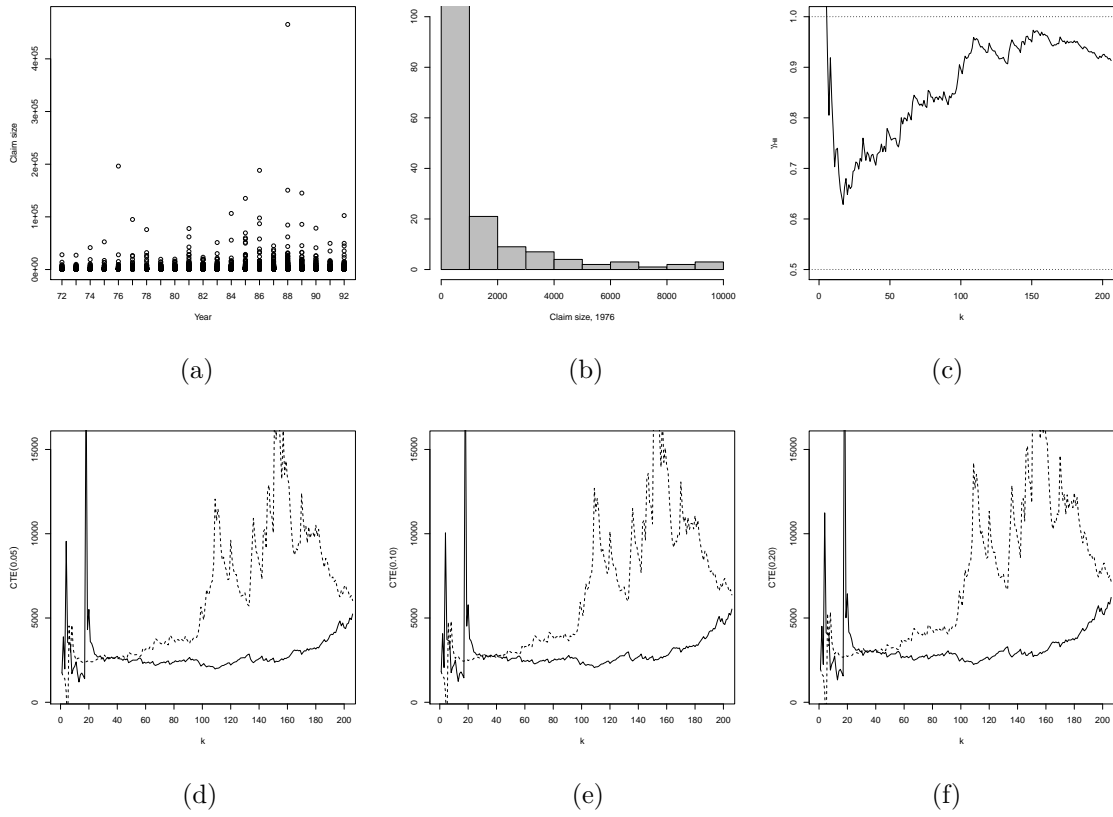


Figure 4: (a) Time plot for the Norwegian fire insurance data; (b) Histogram of the claim size for the year 1976; (c) Hill estimator as a function of  $k$  for the year 1976; Biased estimator  $\tilde{C}_{n,\alpha}[X]$  (dotted line) and reduced-bias one  $\tilde{C}_{n,\alpha}^{LS,-1}[X]$  (full line) as a function of  $k$  for  $\alpha = 0.05$  (d),  $\alpha = 0.10$  (e) and  $\alpha = 0.20$  (f).