Improved Estimator of the Conditional Tail Expectation in the case of heavy-tailed losses

Mohamed Laidi [1,*], Abdelaziz Rassoul [2,*], Hamid Ould Rouis [3,*]

[1]*National High School of Technology, Algiers, Algeria; LRDSI laboratory, Blida 1 university, Blida, Algeria*
[2]*High School of Hydraulics, Blida, Algeria*
[3]*University of Blida 1, Blida, Algeria*

Abstract   This paper investigates the extreme-value methodology in the case of a loss distribution with a finite mean but infinite variance. We propose an improved estimator of the conditional tail expectation (CTE) based on the bias-reduced estimators of high quantile for heavy-tailed distributions. The asymptotic normality of the proposed estimator is established and checked in a simulation study. Moreover, we provided some comparison between the new estimator with the known old estimator in terms of bias and mean squared error.The simulation results show significant improvement either in term of bias or in term of RMSE.

1. Introduction and Motivation

Risk management is a subject of concern in finance and actuarial science. Protecting against financial and actuarial risks is essential in order to anticipate financial crises or major insurance claims. For this, there are many tools to quantify and predict risk (risk measures). They make it possible to evaluate a level of danger of a risk, but also to compare different risks between them. Quantification, prevention, comparison and prediction of risk are essential elements of our society.

One of the best known and used risk measure is the Value-at-Risk (or VaR), it is introduced in the 1990s by Morgan [27] (see Jorion [22] for more details). The idea of the VaR is as follows: we fixe a threshold α and calculate a value $VaR(\alpha)$ which will be such that the probability that the catastrophe will occur is smaller than α. $VaR(\alpha)$ can be seen as the amount of extra capital and business needs in order to reduce the probability of going bankrupt to α.

In statistical terms, the VaR of level $\alpha \in]0,1[$ corresponds to the level $(1-\alpha)$ quantile of the distribution function of the losses. This risk measure has several flaws. It gives only one point information to the quantile $q(\alpha)$ and no information beyond this point; it does not take into account the importance of the disaster when it occurs but only its frequency.

To respond to the need for theoretical practical principles, Artzner et al. 1999 [5] introduced the concept of

---

*Correspondence to: Mohamed Laidi (Email: mohamed.laidi@enst.dz). National High School of Technology, Algiers.
Abdelaziz Rassoul (Email: a.rassoul@ensh.dz). High School of Hydraulics, Blida, Algeria.
Hamid Ould Rouis (Email: houldrouis@hotmail.com). University of Blida 1, Blida, Algeria.

coherent risk measure. A risk measure is said to be coherent if it satisfies the four properties: monotony, positive homogeneity, invariance by translation and subadditivity, (for more details see, [5]).

Unfortunately, the VaR risk measure is not subadditive which implies that it is not coherent. For this reason, many authors proposed to replace standard VaR with alternative risk measures such as Conditional Tail Expectation (CTE), also called in the literature on risk measure Tail Value-at-Risk ($TVaR$) or Expected Shortfall (ES) (see, for example, [23], [15], [1], [2], [3], [33], [32], [37] and [38]).

This risk measures gives information of the distribution beyond the $VaR(\alpha)$. Thus, unlike the VaR, it takes into account the information contained in the tail of the distributi which is essential in the finanacial context. The CTE has been studied by many authors such that: [5], [9], [10], [35], [8], [24].

The application of these alternative measures has gained interest growing in literature and industry. In particular, the CTE, due to its properties and its effectiveness in different fields such as finance and actuaries, its use and studies keep increasing (see, e.g., [15], [34], [18], [16] and references therein).

In the literature, the estimation of the CTE and its asymptotic normality under the assumption that the second moment of the loss variable is finite has been established but this assumption is very restrictive in practical problems. For this reason, Necir et al (2010) [28] has extended that in the case of infinite variance. In this paper, we propose an improvement the estimator established by Necir et al. (2010) [28], our considerations are based on the bias-reduced estimators of high quantile for heavy-tailed distributions introduced by Li (2010) [25], also we show its efficiency and its asymptotic normality theoretically, finally we prove the performance of our estimator by some results of simulation study. For this aim, we proceed as follow :

Let X be a loss random variable with cumulative distribution function (cdf) F. F is assumed to be continuous throughout the present paper and defined on the entire real line, with negative loss interpreted as gain. The CTE of the risk X is then defined, for every $t \in (0, 1)$, by the function $\mathbb{C}(t)$ written as follow

$$\mathbb{C}(t) = \mathbf{E}\left(X \mid X > \mathbf{Q}(t)\right), \tag{1}$$

where $\mathbf{Q}(t) = \inf\{x : F(x) \geq t\}$ is the quantile function corresponding to the cdf F. Since F is continuous, we easily check that $\mathbb{C}(t)$ is equal to

$$\mathbb{C}(t) = \frac{1}{1-t} \int_t^1 \mathbf{Q}(s)ds. \tag{2}$$

In this paper we are interesting by the heavy-tailed distribution. A model F is said to be heavy-tailed, with a tail index α $(\alpha > 0)$, if the tail function $\overline{F} = 1 - F \in \mathcal{RV}_{-\alpha}$, where denotes the class of regularly varying functions with index of regular variation equal to α, i.e., non-negative measurable functions g such that, for all $x > 0$, $g(tx)/g(t) \to x^\alpha$, as $t \to \infty$ (see, [17] for more details).

Suppose that $X_1, X_2, ..., X_n$ are independent and identically distributed (i.i.d.) random variables with common distribution function F which has regularly varying tails with index $\alpha > 1$, i.e.,

$$\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\alpha}, \text{ for every } x > 0. \tag{3}$$

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., [6]). In this paper, we restrict ourselves to this class of distributions. For more information on the topic and, generally, on extreme value models and their manifold applications, we refer to the monographs by [12], [29] and [31].

In particular, if the index $\alpha \in (1, 2)$, then, condition (3) is equivalent to the statement that F has infinite second moment. In this case, in order to estimate the CTE, let $X_{1:n} < \cdots < X_{n:n}$ denote the order statistics of $X_1, \ldots, X_n$, we set

$$\widetilde{\mathbb{C}}_n(t) = \frac{1}{1-t} \left[\int_t^{1-k/n} \mathbf{Q}_n(s)ds + \int_0^{k/n} \mathbf{Q}_n^w(1-s)ds \right], \tag{4}$$

where $\mathbf{Q}_n(s)$ is the empirical estimation of the quantile function, which is equal to the ith order statistic $X_{i,n}$ for all $s \in ((i-1)/n, i/n)$ and for all $i = 1, ..., n$ and $\mathbf{Q}_n^w(1-s)$ is the estimation of the extreme quantile established

by [36], this estimator is given by

$$\mathbf{Q}_n^w(1-s) = \left(\frac{k}{n}\right)^{1/\widehat{\alpha}_n} X_{n-k,n} s^{-1/\widehat{\alpha}_n}. \tag{5}$$

where $\widehat{\alpha}_n$ is the Hill estimator [21]

$$\widehat{\alpha}_n = \left(\frac{1}{k}\sum_{i=1}^{k}\log X_{n-i+1:n} - \log X_{n-k:n}\right)^{-1} \tag{6}$$

of the tail index $\alpha \in (1,2)$ and $k = k(n)$ is an intermediate integer sequence satisfying the condition

$$k \to \infty \text{ and } k/n \to 0 \text{ as } n \to \infty. \tag{7}$$

After integration, we obtain

$$\widetilde{\mathbb{C}}_n(t) = \frac{1}{1-t}\left[\int_t^{1-k/n}\mathbf{Q}_n(s)ds + \left(\frac{k}{n}\right)\frac{\widehat{\alpha}_n X_{n-k;n}}{(\widehat{\alpha}_n - 1)}\right]. \tag{8}$$

the asymptotic normality of this estimator is found in [28] under some conditions.

We note that, the Hill's estimator $\widehat{\alpha}_n$ plays a pivotal role in statistical inference on distribution tails. This estimator has been extensively studied, improved and even generalized to any real parameter α. Weak consistency of $\widehat{\alpha}_n$ was established by [26] assuming only that the underlying cdf F satisfies condition (3). The asymptotic normality of $\widehat{\alpha}_n$ has been established (see [14]) under the following stricter condition that characterizes Hall's model (see [19] and [20]), where there exist $c > 0$, $d \neq 0$ and $\beta > \alpha > 0$ such that

$$1 - F(x) = cx^{-\alpha} + dx^{-\beta} + o\left(x^{-\beta}\right), \text{ as } x \to \infty, \tag{9}$$

Note that (9), which is a special case of a more general second-order regular variation condition (see [13]), is equivalent to

$$\mathbf{Q}(1-s) = c^{1/\alpha}s^{-1/\alpha}(1 + \alpha^{-1}c^{-\beta/\alpha}ds^{\beta/\alpha-1} + o(1)), \text{ as } s \downarrow 0, \tag{10}$$

The constants α and β are called, respectively, first-order (tail index, shape parameter) and second-order parameters of cdf F. In this paper, we use the bias-reduced estimator of the high quantile $\mathbf{Q}(1-s)$, proposed by [25] who exploited the censored maximum likelihood (CML) based estimators $\widehat{\alpha}$ and $\widehat{\beta}$ of the couple of regular variation parameters (α, β) introduced by [30]. The CML estimators $\widehat{\alpha}$, $\widehat{\beta}$ are defined as the solution of the two equations (under the constraint $\beta > \alpha$):

$$\frac{1}{k}\sum_{i=1}^{k}\frac{1}{G_i(\alpha,\beta)} = 1 \text{ and } \frac{1}{k}\sum_{i=1}^{k}\frac{1}{G_i(\alpha,\beta)}\log\frac{X_{n-i+1,n}}{X_{n-k,n}} = \beta^{-1}, \tag{11}$$

where

$$G_i(\alpha,\beta) = \frac{\alpha}{\beta}\left(1 + \frac{\alpha\beta}{\alpha-\beta}H(\alpha)\right)\left(\frac{X_{n-i+1,n}}{X_{n-k,n}}\right)^{\beta-\alpha} - \frac{\alpha\beta}{\alpha-\beta}H(\alpha), \tag{12}$$

and

$$H(\alpha) = \frac{1}{\alpha} - \frac{1}{k}\sum_{i=1}^{k}\log\frac{X_{n-i+1,n}}{X_{n-k,n}}. \tag{13}$$

Li et al. [25] obtained their bias-reduced estimators $\mathbf{Q}_n^L(1-s)$, of the high quantiles $\mathbf{Q}(1-s)$, by substituting $\left(\widehat{\alpha},\widehat{\beta}\right)$ to (α,β) in (10). That is

$$\widehat{\mathbf{Q}}_n^L(1-s) = \widehat{c}^{1/\widehat{\alpha}}s^{-1/\widehat{\alpha}}(1 + \widehat{\alpha}^{-1}c^{-\widehat{\beta}/\widehat{\alpha}}\widehat{d}s^{\widehat{\beta}/\widehat{\alpha}-1} + o(1)), \text{ as } s \downarrow 0, \tag{14}$$

where

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

The consistency and asymptotic normality of $\widehat{\mathbf{Q}}_n^L(1-s)$ are established by the same authors. By using $\mathbf{Q}_n$ in formula (4), we get, after integration, the new estimator of the CTE as follows:

$$
\overline{\mathbb{C}}_n(t) = \frac{1}{1-t} \left[\int_t^{1-k/n} \mathbf{Q}_n(s) ds + \left(\frac{k}{n} \right) \left(\frac{n\widehat{c}}{k} \right)^{1/\widehat{\alpha}} \left(\frac{\widehat{\alpha}}{\widehat{\alpha}-1} + \frac{\widehat{d}\widehat{c}^{-\widehat{\beta}/\widehat{\alpha}} (k/n)^{\widehat{\beta}/\widehat{\alpha}-1}}{\widehat{\beta}-1} \right) \right],
\tag{16}
$$

provided that $\widehat{\beta} > \widehat{\alpha} > 1$ so that $\overline{\mathbb{C}}_n$ is finite.

The rest of the paper is organized as follows. In Section 2 we study the asymptotic normality of the new CTE estimator. In section 3, we make a simulation study of the new estimator and we show the performance of our estimator and compare it with the old one. We finich our paper by a conclusion given in section 4. The proof of the main result, which is Theorem 1 in Section 2, is postponed to Section 4.

2. Main result and its practical implementation

In the field of the extreme values theory, a function denoted by $\mathbf{U}$ and (sometimes) called tail quantile function, is defined by

$$
\mathbf{U}(t) = F^{\leftarrow}(1-1/t) = \mathbf{Q}(1-1/t),
\tag{17}
$$

where $F^{\leftarrow}$ represented the generalized inverse of the df F. Then, we say that F is heavy-tailed iff $\mathbf{U} \in RV_{1/\alpha}$ (de Haan, 1970 [11]), i.e.

$$
\lim_{t\to\infty} \frac{\mathbf{U}(tx)}{\mathbf{U}(t)} = x^{1/\alpha}, \text{ for any } x > 0.
\tag{18}
$$

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

$$
\mathbf{U}(t) = c^{1/\alpha} t^{1/\alpha} (1 + \alpha^{-1} c^{-\beta/\alpha} d t^{1-\beta/\alpha} + o(1)), \text{ as } t \to \infty.
\tag{19}
$$

This implies that, there is a function $A_1(t)$, which tends to zero as $t \to \infty$ (because $\beta > \alpha$), determines the rate of convergence of $\log(\mathbf{U}(tx)/\mathbf{U}(t))$ to its limit $\alpha^{-1} \log x$, such that:

$$
\lim_{t\to\infty} \frac{\log(\mathbf{U}(tx)/\mathbf{U}(t)) - \alpha^{-1} \log x}{A_1(t)} = \frac{x^{1-\beta/\alpha} - 1}{1 - \beta/\alpha}; \text{ for any } x > 0,
\tag{20}
$$

where

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

Relation (20) is known as the second-order condition of regular variation (see, e.g., [12]). Unfortunately, the second-order regular variation is not sufficient to find asymptotic distributions for the estimators defined by the systems (11) and (15). We strengthen it with a condition, called third-order condition of regular variation and given by (21), that specifies the rate of (20) (see, e.g., [13] or [4]).

$$
\lim_{t\to\infty} \frac{\frac{\log(\mathbf{U}(tx)/\mathbf{U}(t)) - \alpha^{-1} \log x}{A_1(t)} - \frac{x^{1-\beta/\alpha}-1}{1-\beta/\alpha}}{A_2(t)} = D(\alpha, \beta, \rho) \frac{x^{1-\beta/\alpha} - 1}{1 - \beta/\alpha}; \text{ for any } x > 0.
\tag{21}
$$

where $A_2(t) \to 0$ as $t \to \infty$, 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 ρ being a positive constant called third-order parameter. Peng and Qi, [30] established the asymptotic normality of $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{c}$ under the following extract conditions on the sample fraction k, as $n \to \infty$:

$$(i): k^{1/2} |A_1(n/k)| \to \infty, \ (ii): k^{1/2} A_1^2(n/k) \to 0 \text{ and } (iii): k^{1/2} A_1(n/k) A_2(n/k) \to 0. \tag{22}$$

As for $\widehat{d}$, it is asymptotically normal under the assumption $k^{1/2} |A_1(n/k)| / \log(n/k) \to \infty$ added to (ii) and (iii). Note that, from a theoretical point of view, assumptions (7) and (22) are realistic, as the following example shows, indeed, let us choose

$$k = [n^{1-\varepsilon}], 0 < \varepsilon < 1, \tag{23}$$

then it easy to verify that these assumptions hold for any $1/5 < \varepsilon < 1/3$. The notation $[\cdot]$ stands for the integer part of real numbers.

Our main result, namely the asymptotic normality of new estimator of the CTE is formulated in the following theorem.

Theorem 1
Assume that the cdf F satisfies condition (21) with $\alpha \in (1, 2)$ and $\beta/\alpha = \lambda > 1$. Then for any sequence of integers $k = k_n$ satisfying the conditions (7) and (22). Then

$$\frac{\sqrt{n}(\overline{\mathbb{C}}_n(t) - \mathbb{C}(t))(1-t)}{(k/n)^{1/2} (nc/k)^{1/\alpha}} \to_d \mathcal{N}\left(0, \sigma^2(\alpha, \beta)\right) \tag{24}$$

for any fixed $t \in (0, 1)$, where the asymptotic variance $\sigma^2(\alpha, \beta)$ is given by the formula

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

3. Simulation study

In this section, the biased estimator $\tilde{\mathbb{C}}_n(t)$ and the reduced-bias one $\overline{\mathbb{C}}_n(t)$ are compared using simulation study. For this reason, 1000 samples of size $n \in \{250, 500, 1000, 2000\}$ are simulated from a two heavy-tailed distributions :

Fréchet model : defined as : $F(x) = exp(-x^{-\alpha})$ we take $\alpha = 1, 5$ and $\alpha = 1, 75$, respectively, and according to the Hall's model, we find $\beta = 2\alpha, c = 1$, and $d = -1/2$.

Burr model : defined as : $F(x) = 1 - (1 + x^\tau)^{-\lambda}$ with $\lambda = 1, 5$ and $\lambda = 1, 75$, both with $\tau = 1$ and in this case we have $c = 1, \alpha = \tau\lambda, d = -\lambda$ and $\beta = \lambda\tau + \tau$ (For more flexibility, we can take $\tau = \alpha/\lambda$ in the expression of Burr distribution).

In all cases, we assume that α, β, c and d are unknowns, then by the resolving the equations (11) and (12) we calculate the values of $\widehat{\alpha}$ and $\widehat{\beta}$, and resolving the system (15) to calculate the estimators $\widehat{c}$ and $\widehat{d}$. We fix two values of t, for example $t = 0.90$ and $t = 0.95$ to calculate the differents values of the two estimators, the new estimator $\overline{\mathbb{C}}_n$ and the old estimator $\tilde{\mathbb{C}}_n$. We note that, the mean, the bias and root mean squared error (RMSE) of these estimators are estimated over the 1000 replications. We illustrate and compare the bias and the root mean square error (RMSE) of $\tilde{\mathbb{C}}_n(t)$ and $\overline{\mathbb{C}}_n(t)$. Finally, we summarize the results of simulations in the tables (1, 2, 3, and 4). We remark, on one hand, that the bias and the RMSE of the new estimator increase when the sample size increase, and on other hands, the bias and RMSE of our estimator is smaller than the old estimator.
In order to show the influence of the choice of the integer value k to the performance of our estimator $\overline{\mathbb{C}}_n(t)$ and the old estimator $\tilde{\mathbb{C}}_n(t)$, we generate 1000 samples of size 1000 of the parent model (Fréchet and Burr) with two values of index $\alpha = 1.5$ and 1.75 and we picture the two estimators when k varied from 50 to 850. The results are displayed in figures 1 and 2. We observe that the performance of our estimator is clearly better than the old estimator.
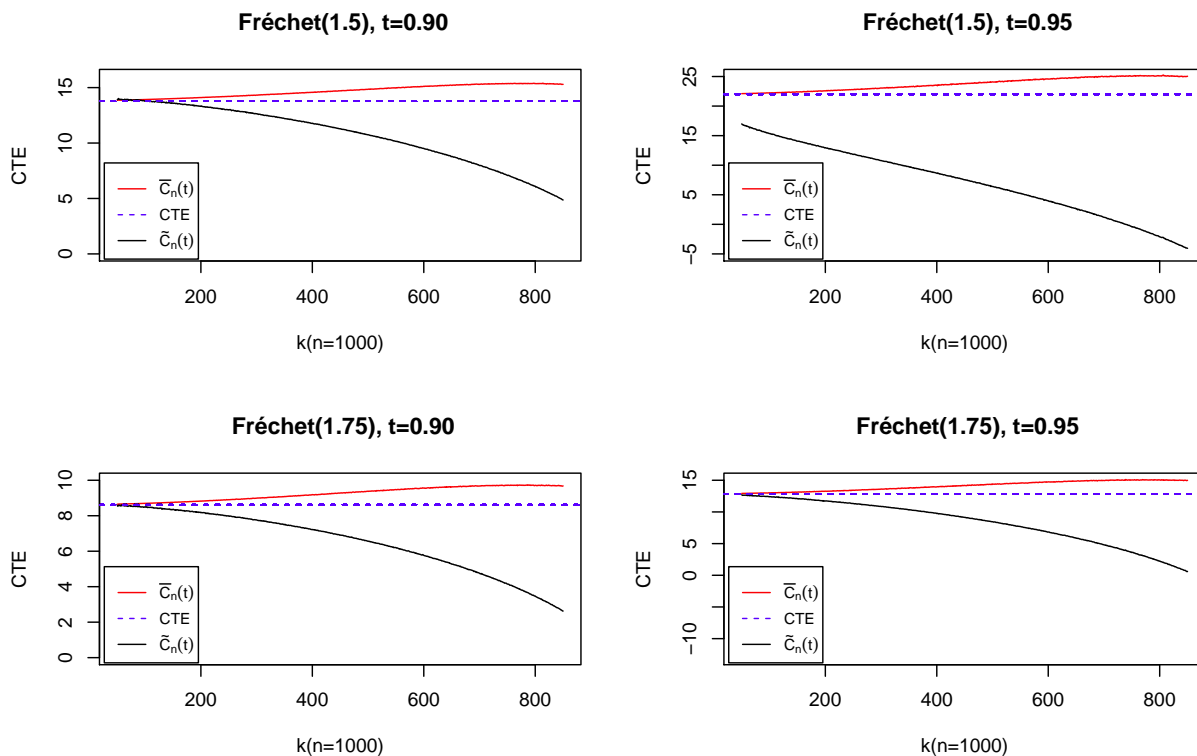
Figure 1. Behavior of the new estimator $\overline{\mathbb{C}}_n$ and old estimator $\widetilde{\mathbb{C}}_n$ of CTE in respect to the variation of k for Fréchet distribution.

Table 1. Comparison between the new estimator $\overline{\mathbb{C}}_n$ and old estimator $\widetilde{\mathbb{C}}_n$ of CTE in terms of bias and RMSE respecting to the variation of sample size n based on Fréchet distribution.

Fréchet distribution, $\alpha = 1.5$								
t	n	CTE	$\bar{\mathbb{C}}_n(t)$	bias	RMSE	$\tilde{\mathbb{C}}_n(t)$	bias	RMSE
0.9	250	13.793	14.457	−0.664	0.66415	11.048	2.745	2.7454
	500		14.409	−0.616	0.61578	11.418	2.375	2.3755
	1000		14.287	−0.494	0.49345	11.837	1.956	1.9559
	2000		14.14	−0.347	0.34733	11.852	1.941	1.941
0.95	250	21.984	22.78	−0.796	0.79685	16.115	−5.869	5.8683
	500		22.703	−0.719	0.71965	16.473	−5.511	5.5105
	1000		22.674	−0.69	0.69043	16.944	−5.04	5.0392
	2000		22.67	−0.686	0.68661	17.247	−4.737	4.7364

Table 2. Comparison between the new estimator $\overline{\mathbb{C}}_n$ and old estimator $\widetilde{\mathbb{C}}_n$ of CTE in terms of bias and RMSE respecting to the variation of sample size n based on Fréchet distribution.

			Fréchet distribution, $\alpha = 1.75$					
t	n	CTE	$\overline{\mathbb{C}}_n(t)$	bias	RMSE	$\widetilde{\mathbb{C}}_n(t)$	bias	RMSE
0.9	250	8.6207	8.6977	-0.077	0.076993	8.2478	0.3729	0.37295
	500		8.6771	-0.0564	0.05637	8.3136	0.3071	0.30711
	1000		8.6586	-0.0379	0.037901	8.5085	0.1122	0.11218
	2000		8.6504	-0.0297	0.029695	8.5374	0.0833	0.083322
0.95	250	12.866	12.963	-0.097	0.096298	12.185	0.681	0.68126
	500		12.956	-0.09	0.089431	12.218	0.648	0.64858
	1000		12.937	-0.071	0.070963	12.624	0.242	0.24248
	2000		12.917	-0.051	0.050889	12.766	0.100	0.10069

Table 3. Comparison between the new estimator $\overline{\mathbb{C}}_n$ and old estimator $\widetilde{\mathbb{C}}_n$ of CTE in terms of bias and RMSE respecting to the variation of sample size n based on Burr distribution.

			Burr distribution, $\alpha = 1.5$					
t	n	CTE	$\overline{\mathbb{C}}_n(t)$	bias	RMSE	$\widetilde{\mathbb{C}}_n(t)$	bias	RMSE
0.9	250	13.676	14.365	-0.689	0.68807	14.909	-1.233	1.2321
	500		14.226	-0.55	0.54908	14.658	-0.982	0.98132
	1000		14.124	-0.448	0.44795	14.506	-0.83	0.82922
	2000		14.026	-0.35	0.34997	14.491	-0.815	0.81428
0.95	250	21.891	23.239	-1.348	1.3474	24.05	-2.159	2.1587
	500		22.996	-1.105	1.1044	23.874	-1.983	1.9827
	1000		22.753	-0.862	0.86155	23.709	-1.818	1.8173
	2000		22.602	-0.711	0.71095	23.64	-1.749	1.7489

Table 4. Comparison between the new estimator $\overline{\mathbb{C}}_n$ and old estimator $\widetilde{\mathbb{C}}_n$ of CTE in terms of bias and RMSE respecting to the variation of sample size n based on Burr distribution.

			Burr distribution, $\alpha = 1.75$					
t	n	CTE	$\overline{\mathbb{C}}_n(t)$	bias	RMSE	$\widetilde{\mathbb{C}}_n(t)$	bias	RMSE
0.9	250	8.5455	9.0332	-0.4877	0.48776	9.3583	-0.8128	0.81281
	500		8.9188	-0.3733	0.37334	9.0502	-0.5047	0.5047
	1000		8.8353	-0.2898	0.28979	8.9942	-0.4487	0.44873
	2000		8.7708	-0.2253	0.22533	8.9115	-0.366	0.36607
0.95	250	12.811	13.697	-0.886	0.88641	14.386	-1.575	1.5751
	500		13.556	-0.745	0.74521	14.082	-1.271	1.2709
	1000		13.381	-0.57	0.56966	13.788	-0.977	0.97676
	2000		13.246	-0.435	0.4355	13.596	-0.785	0.78538
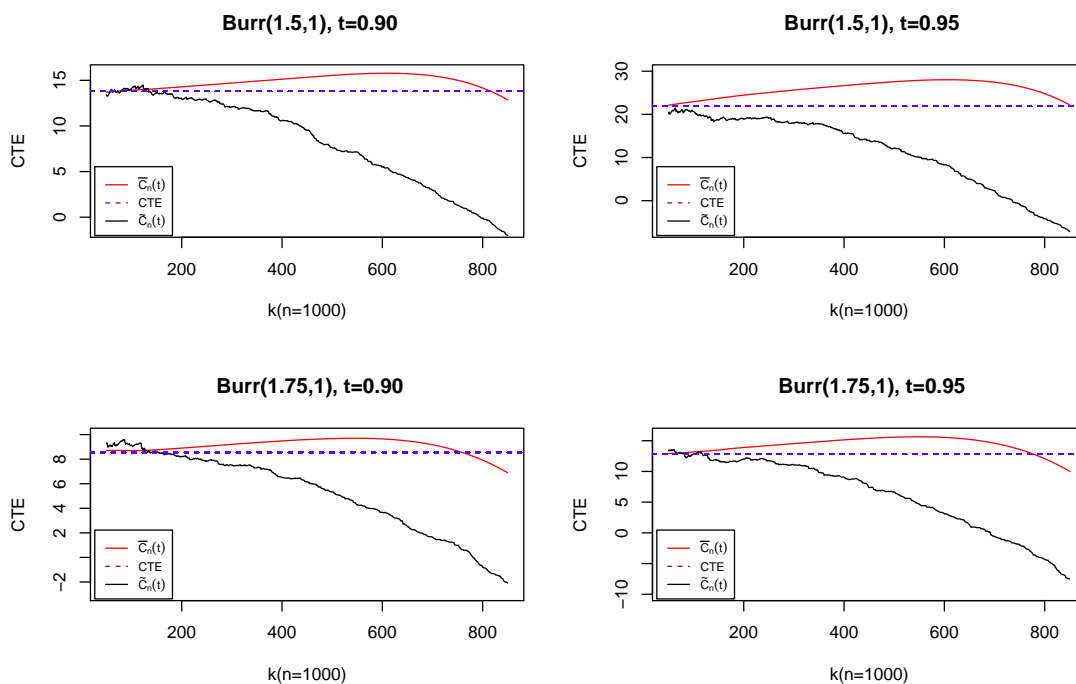
Figure 2. Behavior of the new estimator $\overline{\mathbb{C}}_n$ and old estimator $\widetilde{\mathbb{C}}_n$ of CTE in respect to the variation of k for Burr distribution.

The analysis and simulation as shown in tables 1, 2, 3, 4, and figures 1, 2 indicate that the new CTE estimator is more performant than the estimator contructed by Necir et al (2010) [28].

4. Conclusion

From the research that has undertaken, it is possible to conclude that the estimation of the CTE by the use of extreme quantile established by Li (2010) [25] gave us a more efficient estimate compared to that introduced by Necir et al (2010) [28]. The results of the simulations show significant improvement either in term of bias or in term of RMSE. The next stage of our research will be the construction of a confidence interval of our CTE estimator with the covering probability.

Appendix: Proof of Theorem 1

We start the proof of Theorem 1 with the decomposition

$$(1-t)\,\mathbb{C}(t) = C_{1.n}(s) + C_{2,n}(s),$$

where

$$C_{1.n}(s) = \int_t^{1-k/n} \mathbf{Q}\,(s)\,ds \text{ and } C_{2.n}(s) = \int_0^{k/n} \mathbf{Q}\,(1-s)\,ds.$$

Also, we have

$$(1-t)\,\overline{\mathbb{C}}(t) = \overline{C}_{1.n}(s) + \overline{C}_{2,n}(s),$$

where

$$\overline{C}_{1.n}(s) = \int_t^{1-k/n} \mathbf{Q}_n(s)ds,$$

and

$$\overline{C}_{2.n}(s) = \left(\frac{k}{n}\right) \left(\frac{n\widehat{c}}{k}\right)^{1/\widehat{\alpha}} \left(\frac{\widehat{\alpha}}{\widehat{\alpha}-1} + \frac{\widehat{d}\widehat{c}^{-\widehat{\beta}/\widehat{\alpha}} (k/n)^{\widehat{\beta}/\widehat{\alpha}-1}}{\widehat{\beta}-1}\right).$$

First, it is easy to verify that, as $n \to \infty$

$$C_{2.n}(s) = (1 + o_p(1)) \left(\frac{k}{n}\right) \left(\frac{nc}{k}\right)^{1/\alpha} \left(\frac{\alpha}{\alpha-1}\right),$$

and, under the condition (6), we have

$$\overline{C}_{2.n}(s) = (1 + o_p(1)) \left(\frac{k}{n}\right) \left(\frac{n\widehat{c}}{k}\right)^{1/\widehat{\alpha}} \left(\frac{\widehat{\alpha}}{\widehat{\alpha}-1}\right).$$

It follows that

$$\overline{C}_{2.n}(s) - C_{2.n}(s) = (1 + o_p(1)) \left(\frac{k}{n}\right) \left[\frac{\widehat{\alpha}}{\widehat{\alpha}-1} \left(\frac{n\widehat{c}}{k}\right)^{1/\widehat{\alpha}} - \frac{\alpha}{\alpha-1} \left(\frac{nc}{k}\right)^{1/\alpha}\right], \qquad (25)$$

let us write $\overline{C}_{2.n}(s) - C_{2.n}(s) = A_{n,1}(s) + A_{n,2}(s)$, where

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

We begin by showing that $A_{n,2} \to 0$, as $n \to \infty$. First observe that $A_{n,2}$ may be rewritten into

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

Assumptions (i) and (ii), imply that $\sqrt{k}/\log(n/k) \to \infty$. Also, from Theorem 1 of [30], the asymptotic normality of $\widehat{\alpha}$ gives $\widehat{\alpha} - \alpha = O(1/\sqrt{k})$. Therefore

$$(1/\widehat{\alpha} - 1/\alpha) \log(nc/k) \overset{P}{\to} 0,$$

this implies that

$$(nc/k)^{1/\widehat{\alpha}-1/\alpha} \overset{P}{\to} 1 \text{ as } n \to \infty.$$

On the other hand, from equation (4.7) in [25], we have

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

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

$$(\widehat{c}/c)^{1/\widehat{\alpha}-1/\alpha} - 1 = \alpha^{-1} (\widehat{\alpha} - \alpha) \log(n/k) (\widehat{c}/c - 1), \text{ as } n \to \infty.$$

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 $\widehat{\alpha} - \alpha = O_p(1/\sqrt{k})$, yield

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

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

$$A_{n,1} = -\left(1 + o_p\left(1\right)\right) \frac{\sqrt{k}}{\left(\alpha - 1\right)^2}\left(\widehat{\alpha} - \alpha\right).$$

From theorem 2.1 in [7], we infer that

$$\frac{\sqrt{n}\left(1 - t\right)}{\left(k/n\right)^{1/2}\left(nc/k\right)^{1/\alpha}}\left(\overline{C}_{2,n}(s) - C_{2,n}(s)\right) = -\frac{\alpha}{\left(\alpha - 1\right)^2}\left[\eta_1 W_1 + \eta_2 W_2 + \eta_3 W_3\right] + o_{\mathbf{P}}(1),$$

where

$$W_1 = \sqrt{n/k}B_n\left(1 - k/n\right) - \sqrt{n/k}\int_0^1 s^{-1}B_n(1 - ks/n)ds,$$

$$W_2 = \left(\lambda^{-1} - 1\right)\sqrt{n/k}B_n\left(1 - k/n\right) + \left(\lambda - 1\right)\sqrt{n/k}\int_0^1 s^{\lambda-2}B_n(1 - ks/n)ds,$$

$$W_3 = \left(1 - \lambda\right)\sqrt{n/k}\int_0^1 s^{\lambda-2}(\log s)B_n(1 - ks/n)ds$$
$$+ \lambda^{-2}\sqrt{n/k}B_n\left(1 - k/n\right) - \sqrt{n/k}\int_0^1 s^{\lambda-2}B_n(1 - ks/n)ds,$$

and

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

From the Proof of statement (4.2) in [28], we shall show below that there are Brownian bridges B_n such that

$$\frac{\sqrt{n}\left(1 - t\right)}{\left(k/n\right)^{1/2}\mathbf{Q}(1 - k/n)}\left(\overline{C}_{1,n}(s) - C_{1,n}(s)\right) = -\frac{\int_{k/n}^1 B_n(1 - s)d\mathbf{Q}(1 - s)}{\left(k/n\right)^{1/2}\mathbf{Q}(1 - k/n)} + o_{\mathbf{P}}(1) \qquad (26)$$

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

$$\frac{\sqrt{n}\left(1 - t\right)}{\left(k/n\right)^{1/2}\left(nc/k\right)^{1/\alpha}}\left(\overline{C}_{1,n}(s) - C_{1,n}(s)\right) = W_4 + o_{\mathbf{P}}(1),$$

where

$$W_4 = -\frac{\int_{k/n}^1 B_n(1 - s)d\mathbf{Q}(1 - s)}{\left(k/n\right)^{1/2}\mathbf{Q}(1 - k/n)},$$

Finally, we have

$$\frac{\sqrt{n}\left(1 - t\right)}{\left(k/n\right)^{1/2}\left(nc/k\right)^{1/\alpha}}\left(\overline{\mathbb{C}}_n(t) - \mathbb{C}(t)\right) = -\frac{\alpha}{\left(\alpha - 1\right)^2}\left[\eta_1 W_1 + \eta_2 W_2 + \eta_3 W_3\right] + W_4 + o_{\mathbf{P}}(1).$$

All done, after a standard calculation we obtain

$$\frac{\sqrt{n}\left(1 - t\right)}{\left(k/n\right)^{1/2}\left(nc/k\right)^{1/\alpha}}\left(\mathbb{C}_n(t) - \mathbb{C}(t)\right) \to_d N\left(0, \sigma^2\left(\alpha, \beta\right)\right).$$

Acknowledgment

REFERENCES

1. C. Acerbi. Coherent representations of subjective risk-aversion. *Risk measures for the 21st century*, pages 147–207, 2004.
2. C. Acerbi and D. Tasche. Expected shortfall: a natural coherent alternative to value at risk. *Economic notes*, 31(2):379–388, 2002.
3. C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, 2002.
4. M. F. Alves, M. I. Gomes, L. de Haan, and C. Neves. A note on second order conditions in extreme value theory: linking general and heavy tail conditions. *REVSTAT Statistical Journal*, 5(3):285–304, 2007.
5. P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.
6. J. Beirlant and G. Matthys. Extreme quantile estimation for heavy-tailed distributions. *Rapport technique, Department of Mathematics, KU Leuven*, page 23, 2001.
7. B. Brahimi, D. Meraghni, A. Necir, and D. Yahia. A bias-reduced estimator for the mean of a heavy-tailed distribution with an infinite second moment. *Journal of Statistical Planning and Inference*, 143(6):1064–1081, 2013.
8. V. Brazauskas, B. L. Jones, M. L. Puri, and R. Zitikis. Estimating conditional tail expectation with actuarial applications in view. *Journal of Statistical Planning and Inference*, 138(11):3590–3604, 2008.
9. J. Cai and K. S. Tan. Optimal retention for a stop-loss reinsurance under the var and cte risk measures. *ASTIN Bulletin: The Journal of the IAA*, 37(1):93–112, 2007.
10. J.-J. Cai, J. H. Einmahl, L. Haan, and C. Zhou. Estimation of the marginal expected shortfall: the mean when a related variable is extreme. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 77(2):417–442, 2015.
11. L. de Haan. De (1970) on regular variation and its application to the weak convergence of sample extremes. *Mathematical Centre Tracts*, 32, 1970.
12. L. de Haan and A. Ferreira. Extreme value theory: An introduction, 2006.
13. L. de Haan and U. Stadtmüller. Generalized regular variation of second order. *Journal of the Australian Mathematical Society (Series A)*, 61(03):381–395, 1996.
14. L. d. De Haan and L. Peng. Comparison of tail index estimators. *Statistica Neerlandica*, 52(1):60–70, 1998.
15. J. Dhaene, L. Henrard, Z. Landsman, A. Vandendorpe, and S. Vanduffel. Some results on the cte-based capital allocation rule. *Insurance: Mathematics and Economics*, 42(2):855–863, 2008.
16. A. El Attar, M. El Hachlougi, and Z. E. A. Guennoun. Optimality of reinsurance treaties under a mean-ruin probability criterion. *Statistics, Optimization & Information Computing*, 7(2):383–393, 2019.
17. B. Gnedenko. Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of mathematics*, pages 423–453, 1943.
18. O. Hakim. Pot approach for estimation of extreme risk measures of eur/usd returns. *Statistics, Optimization & Information Computing*, 6(2):240–247, 2018.
19. P. Hall. On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 37–42, 1982.
20. P. Hall, A. Welsh, et al. Adaptive estimates of parameters of regular variation. *The Annals of Statistics*, 13(1):331–341, 1985.
21. B. M. Hill et al. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, 3(5):1163–1174, 1975.
22. P. Jorion. *Value at risk: the new benchmark for controlling market risk*. Irwin Professional Pub., 1997.
23. R. Kaas, M. Goovaerts, J. Dhaene, and M. Denuit. *Modern actuarial risk theory: using R*, volume 128. Springer Science & Business Media, 2008.
24. Z. M. Landsman and E. A. Valdez. Tail conditional expectations for elliptical distributions. *North American Actuarial Journal*, 7(4):55–71, 2003.
25. D. Li, L. Peng, and J. Yang. Bias reduction for high quantiles. *Journal of Statistical Planning and Inference*, 140(9):2433–2441, 2010.
26. D. M. Mason. Laws of large numbers for sums of extreme values. *The Annals of Probability*, pages 754–764, 1982.
27. J. Morgan. Creditmetrics-technical document. *JP Morgan, New York*, 1997.
28. A. Necir, A. Rassoul, and R. Zitikis. Estimating the conditional tail expectation in the case of heavy-tailed losses. *Journal of Probability and Statistics*, 2010.
29. L. Peng. Asymptotically unbiased estimators for the extreme-value index. *Statistics & Probability Letters*, 38(2):107–115, 1998.
30. L. Peng and Y. Qi. Estimating the first-and second-order parameters of a heavy-tailed distribution. *Australian & New Zealand Journal of Statistics*, 46(2):305–312, 2004.
31. S. I. Resnick. *Heavy-tail phenomena: probabilistic and statistical modeling*. Springer Science & Business Media, 2007.
32. R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, 26(7):1443–1471, 2002.
33. R. T. Rockafellar, S. Uryasev, et al. Optimization of conditional value-at-risk. *Journal of risk*, 2:21–42, 2000.
34. K. S. Tan, C. Weng, and Y. Zhang. Optimality of general reinsurance contracts under cte risk measure. *Insurance: Mathematics and Economics*, 49(2):175–187, 2011.
35. D. Tasche. Expected shortfall and beyond. *Journal of Banking & Finance*, 26(7):1519–1533, 2002.
36. I. Weissman. Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815, 1978.

37. J. L. Wirch and M. R. Hardy. A synthesis of risk measures for capital adequacy. *Insurance: mathematics and Economics*, 25(3):337–347, 1999.
38. Y. Yamai, T. Yoshiba, et al. Comparative analyses of expected shortfall and value-at-risk: their estimation error, decomposition, and optimization. *Monetary and economic studies*, 20(1):87–121, 2002.