Insurance: Mathematics and Economics 53 (2013) 698–703

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Kernel-type estimator of the conditional tail expectation for a heavy-tailed distribution

Abdelaziz Rassoul

GEE Laboratory, National High School, Blida, Algeria

HIGHLIGHTS

- We propose a new estimator of the conditional tail expectation (CTE).
- The new estimator generalizes the old estimator of CTE in heavy-tailed cases.
- A reduced bias estimator of the CTE has been extracted.
- We present the asymptotic normality of the proposed estimator.
- Some results of simulation show the performance of our estimator.

ARTICLE INFO

Article history:
Received February 2013
Received in revised form
July 2013
Accepted 6 September 2013

Keywords:
Risk measure
CTE
Heavy tails
Kernel
Hill estimator
Extreme quantile
Reduced bias

ABSTRACT

In this paper, we are interested in the generalization and improvement of the estimator of the conditional tail expectation (CTE) for a heavy-tailed distribution when the second moment is infinite. It is well known that classical estimators of the CTE are seriously biased under the second-order regular variation framework. To reduce the bias, many authors proposed the use of so-called second-order reduced bias estimators for both first-order and second-order tail parameters. In this work, we have generalized a kernel-type estimator, and we present a number of results on its distributional behavior and compare its performance with the performance of other estimators.

1. Introduction

The conditional tail expectation (CTE) is one of the most important actuarial risk measures; it is risk measure coherent, and it represents the average amount of loss given that the loss exceeds a specified quantile. Hence, the CTE provides a measure of the capital needed due to the exposure to loss, and thus serves as a risk measure. Not surprisingly, therefore, the CTE continues to receive increased attention in the actuarial and financial literature, where we also find numerous generalizations and extensions (see, e.g., Landsman and Valdez, 2003, Hardy and Wirch, 2004, Cai and Li, 2005, Furman and Landsman, 2006, Furman and Zitikis, 2008, and references therein). We next present basic notation and definitions.

Let X be a loss random variable with cumulative distribution function (cdf) F. Usually, the cdf F is assumed to be continuous and defined on the entire real line, with negative loss interpreted as gain. We also make this continuity of F assumption throughout the present paper. The CTE of the risk X is then defined, for every $t \in (0, 1)$, by

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

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

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

Hence, from now on we work with $\mathbb{C}(t)$ and call it the CTE for short.

Suppose that we have independent random variables $X_1, X_2, \ldots,$ each with the cdf F, and let $X_{1,n} < \cdots < X_{n,n}$ denote the order statistics of $X_1, \ldots, X_n$. It is natural to define an empirical estimator of $\mathbb{C}(t)$ by the formula

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

E-mail addresses: rsl_aziz@yahoo.fr, a.rassoul@ensh.dz.

0167-6687/$ – see front matter
http://dx.doi.org/10.1016/j.insmatheco.2013.09.004

where $\mathbb{Q}_n(s)$ is the empirical 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 $i = 1, \ldots, n$. The asymptotic behavior of the estimator $\widehat{\mathbb{C}}_n(t)$ has been studied by Brazauskas et al. (2008) provided that the second moment is finite ($\mathbf{E}[X^2] < \infty$).

This paper deals with the estimation problem of the CTE within the class of heavy-tailed distributions. In mathematical terms, a heavy-tailed distribution of a random variable X is regularly varying at infinity with index $(-1/\gamma) < 0$ if

$$1 - F(x) = x^{-1/\gamma}\mathbb{L}(x), \quad \text{for every } x > 0, \tag{3}$$

where $\mathbb{L}(x)$ is a slowly varying function at infinity. This class includes a number of popular distributions such as Pareto, generalized Pareto, Burr, Fréchet, and Student, 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)). 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 CTE is finite, and, since in that case the results of Brazauskas et al. (2008) cannot be applied, the second moment of X being infinite.

Indeed, recall that, from (1), $\mathbb{C}(t)$ can be rewritten as

$$\mathbb{C}_k(t) = \frac{1}{1-t}\left(\int_t^{1-k/n} \mathbb{Q}(s)ds + \int_0^{k/n} \mathbb{Q}(1-s)ds\right) = \mathbb{C}_{k,1}(t) + \mathbb{C}_{k,2}(t).$$

We formulate the CTE estimator for a heavy-tailed distribution satisfying (3) as follows:

$$\widetilde{\mathbb{C}}_{n,k}(t) = \frac{1}{(1-t)}\int_t^{1-k/n} \mathbb{Q}_n(s)ds + \frac{(k/n)\,X_{n-k,n}}{(1-t)\left(1-\widehat{\gamma}^H_{n,k}\right)}, \tag{4}$$

where $\widehat{\gamma}^H_{n,k}$ is the Hill estimator (Hill, 1975) of the tail index γ:

$$\widehat{\gamma}^H_{n,k} = \frac{1}{k}\sum_{i=1}^{k} i\left(\log X_{n-i+1,n} - \log X_{n-i,n}\right). \tag{5}$$

Note that to estimate $\mathbb{C}_{k,2}(t)$ we use a Weissman-type (Weissman, 1978) estimator for $\mathbb{Q}$:

$$\widehat{\mathbb{Q}}\,(1-s) := X_{n-k,n}\ (k/n)^{\widehat{\gamma}^H_{n,k}}\ s^{-\widehat{\gamma}^H_{n,k}}, \quad s \to 0. \tag{6}$$

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

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

$$\widehat{\gamma}^K_{n,k} = \frac{1}{k}\sum_{i=1}^{k} K\left(\frac{i}{k+1}\right) Z_{i,k}, \tag{7}$$

where K is a kernel integrating to 1 and $Z_{i,k} = i(\log X_{n-i+1,n} - \log X_{n-i,n})$.

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 CTE. Thus, the CTE can be estimated by

$$\widetilde{\mathbb{C}}^K_{n,k}(t) = \frac{1}{(1-t)}\int_t^{1-k/n} \mathbb{Q}_n(s)ds + \frac{(k/n)\,X_{n-k,n}}{(1-t)\left(1-\widehat{\gamma}^K_{n,k}\right)}. \tag{8}$$

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

$$\mathbb{U}(z) = \inf\{y : F(y) \geq 1 - 1/z\}, \quad z > 1. \tag{9}$$

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

$$\lim_{t\to\infty} \frac{\log \mathbb{U}(tx) - \log \mathbb{U}(t) - \gamma \log(x)}{A(t)} = \frac{x^\rho - 1}{\rho}; \tag{10}$$

when $\rho = 0$, the ratio on the right-hand side of Eq. (10) 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, which is a class of cdfs, such that

$$\mathbb{U}\,(t) = ct^\gamma\,(1 + dA(t)/\rho + o\,(t^\rho)) \quad \text{as } t \to \infty, \tag{11}$$

where $\gamma > 0$, $\rho \leq 0$, $c > 0$, and $d \in \mathbb{R}^*$.

This subclass of heavy-tailed distributions contains the Pareto, Burr, Fréchet, and Student-t cdfs 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), Gomes et al. (2005) and Gomes and Pestana (2007).

The remainder of this paper is organized as follows. In Section 2, we study the asymptotic properties of the general kernel estimator of the CTE $\widetilde{\mathbb{C}}^K_{n,k}$. 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. The proofs of the different results are postponed to Section 4.

Note that, throughout this paper, the standard notations $\xrightarrow{\mathbb{P}}$, $\xrightarrow{d}$, and $\overset{d}{=}$, respectively, stand for convergence in probability, convergence in distribution, and equality in distribution, $\mathcal{N}(a, b^2)$ denotes the normal distribution with mean a and variance b^2, and $\{\mathbf{B}_n\,(s)\,, 0 < s < 1\}_{n\geq 1}$ is a sequence of standard Brownian bridges.

2. Main results

For studying the asymptotic normality of the estimator of $\widetilde{\mathbb{C}}^K_{n,k}$, we need some results and classical assumptions about the kernel.

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

CK1. $K(s) \geq 0$ whenever $0 < s \leq 1$ and $K(1) = K'(1) = 0$.
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 CTE estimator

Theorem 2.1. *Assume that F satisfies (10) with $\gamma \in (1/2, 1)$. If, furthermore, $(\mathcal{K})$ holds and the sequence k satisfies $k \to \infty$, $k/n \to 0$ and $\sqrt{k}A(n/k) \to 0$ as $n \to \infty$, we have*

$$\frac{(1-t)\sqrt{k}}{(k/n)\,\mathbb{U}(n/k)}\left(\widetilde{\mathbb{C}}^K_{n,k}(t) - \mathbb{C}\,(t)\right) \overset{d}{=} \sqrt{k}A\left(\frac{n}{k}\right)\mathcal{A}\mathcal{B}_K\,(\gamma,\rho) + \mathbb{W}_{1,n} + \mathbb{W}_{2,n}\,(K) + \mathbb{W}_{3,n} + o_{\mathbb{P}}\,(1)\,,$$