# Robust estimator of distortion risk premiums for heavy-tailed losses

**Brahim Brahimi and Zoubir Kenioua**

Laboratory of Applied Mathematics, Mohamed Khider University, Biskra, Algeria

**Abstract.** We use the so-called t-Hill tail index estimator proposed by Fabián and Stehlík (2009), rather than Hill's one, to derive a robust estimator for the distortion risk premium of losses. Under the second-order condition of regular variation, we establish its asymptotic normality. By simulation study, we show that this new estimator is more robust than of Necir and Meraghni (2009) both for small and large samples.

**Résumé.** Nous utilisons l'estimateur de l'indice de queue dit t-Hill proposé par Fabián and Stehlík (2009), au lieu de l'estimateur de Hill, pour obtenir un estimateur robuste pour la prime de risque de distorsion des pertes. Sous la condition de second ordre de variation régulière, nous établissons sa normalité asymptotique. Par l'étude de simulation, nous montrons que ce nouvel estimateur est plus robuste que de celui proposé par Necir and Meraghni (2009) pour les petites et les grandes tailles d'échantillon.

## 1. Introduction

In many important applications in Finance, Actuarial Science, Hydrology, Insurance, one of most crucial topics is the determination of the amounts of losses of a heavy-tailed risks. Most of this risk measures, used are special cases of Wang's distortion premium (Wang, 1996), defined as follows

$$\Pi[\psi; F] = \int_0^\infty \psi(1 - F(x))dx, \tag{1}$$

---

*Corresponding author Brahim Brahimi: brah.brahim@gmail.com
Zoubir Kenioua: zoukenioua164@gmail.com

where $\psi : [0,1] \to [0,1]$ is a non-decreasing function called distortion function, such that $\psi(0) = 0$ and $\psi(1) = 1$. The distortion functions $\psi$ are concave, which makes the corresponding distortion premiums $\Pi[\psi; F]$ coherent (Artzner *et al.*, 1999) as proved by Wirch and Hardy (1999). In particular, the proportional-hazards premium (PHP) (see, Rolski *et al.*, 1999, page 82) is defined as follows

$$\Pi[\rho; F] = \int_0^\infty (1 - F(x))^{1/\rho} \, dx, \tag{2}$$

with the concave distortion function $\psi(t) = t^{1/\rho}$ for every $\rho \geq 1$, the parameter $\rho$ is called the risk aversion index. In recent literature, by using the mean excess of loss function Beirlant *et al.* (2001) proposed a semi-parametric estimator for the premium of the largest reinsurance, Necir and Boukhetala (2004) proposed a non-parametric estimator for the net-premium of a high-excess loss layer of heavy-tailed claim amounts and the semi-parametric version was proposed by Necir *et al.* (2007).

Suppose that we have an independent and identically distributed (iid) sample $X_1, ..., X_n$ of size $n$ of an insured risk $X$ with finite mean and a cumulative distribution function (cdf) $F$, note that $X$ is called heavy-tailed, if $1 - F$ is regularly varying with index $-1/\gamma < 0$ (notation: $F \in \mathcal{RV}_{(-1/\gamma)}$), that is

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

Suppose now that $F$ satisfying condition (3) and let denote by $X_{1:n} \leq ... \leq X_{n:n}$ the corresponding order statistics. Also, let $1 < k = k_n$ be the number of extreme observations used in the computation of the tail index. We assume that $k$ satisfies the conditions

$$1 < k < n, \ k \to \infty \text{ and } k/n \to 0 \text{ as } n \to \infty. \tag{4}$$

Since we are concerned with with heavy-tailed losses with infinite second moment, then by following Brahimi *et al.* (2011), we assume that $\gamma \in (1/2, 1)$ and $\rho\gamma \in (0, 1)$, thus we will work with

$$1/2 < \gamma < 1/\rho. \tag{5}$$

Necir and Meraghni (2009) proposed an alternative estimator of (2) and establish its asymptotic normality based on the Weissman's estimator of the high quantile $q_t = F^{\leftarrow}(1 - t)$ defined by

$$\widehat{q}_t = (k/n)^{\widehat{\gamma}^H} X_{n-k:n} t^{-\widehat{\gamma}^H}, \ t \downarrow 0,$$

where $F^{\leftarrow}$ denotes the generalized inverse of $F$ and

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

is the well-known Hill estimator (Hill, 1975) of the tail index $\gamma$. For a fixed aversion parameter $\rho$, their estimator is given by

$$\widehat{\Pi}_{\rho,n}\left(\widehat{\gamma}^H, k\right) := (k/n)^{1/\rho} \frac{X_{n-k:n}}{1 - \widehat{\gamma}^H \rho} + \sum_{i=k+1}^n \left((i/n)^{1/\rho} - ((i-1)/n)^{1/\rho}\right) X_{n-i+1:n}. \tag{7}$$

The Hill estimator is a pseudo-maximum likelihood estimator based on the exponential approximation of normalized log-spacings $Y_j = j \left( \log X_{j:n} - \log X_{j+1:n} \right)$ for $j = 1, ..., k$. In practice, the Hill estimator depends on the choice of the sample fraction $k$ and is inherently not very robust to large values $Y_j$, which makes the estimator proposed by Necir and Meraghni (2009) sensitive. This constitutes a serious problem in terms of bias and root mean squared error (RMSE). To improve the quality of $\widehat{\Pi}_{\rho,n} \left( \widehat{\gamma}^H, k \right)$, instead of Hill's one, we propose to estimate the tail index $\gamma$ by the so-called t-Hill estimator, proposed by Fabián and Stehlík (2009), given by its harmonic mean

$$\widehat{\gamma} = \widehat{\gamma}(k) := \left( \frac{1}{k} \sum_{j=1}^{k} \frac{X_{n-k:n}}{X_{n-j+1:n}} \right)^{-1} - 1 \tag{8}$$

known as score moment estimation (t-score or t-estimation method). The latter is more robust than the classical Hill estimator $\widehat{\gamma}^H$ defined in Equation (6) (see Stehlík *et al.*, 2012 and the asymptotic normality is given in Theorem 2 of Beran *et al.*, 2014). For other robust estimators for $\gamma$ we referred to Peng and Welsh (2001), Júarez and Schucany (2004), Vandewalle *et al.* (2007) and Kim and Lee (2008). The rest of the paper is organized as follows, in Section 2 we present a construction of a robust estimator of $\Pi[\rho; F]$ in the case of heavy-tailed losses. In Section 3 we establish its asymptotic normality. In Section 4 we carry out a simulation study to illustrate empirical performance and robustness of the estimator. Concluding notes are given in Section 5. Proofs are gathered in Section 6.

Throughout the paper, we use the standard notation $\overset{P}{\to}$ for the convergence in probability and $\mathcal{N}(\mu, \sigma)$ to denote a normal rv with mean $\mu$ and variance $\sigma$.

## 2. Defining the estimator

By using the generalized inverse $F^{\leftarrow}$ and for a fixed risk aversion index $\rho$, we may rewrite (2) into

$$\Pi_{\rho}[X] := - \int_{0}^{1} s^{1/\rho} dF^{\leftarrow}(1-s). \tag{9}$$

The empirical estimator of the risk premium $\Pi_{\rho}[X]$ is obtained by substituting $F^{\leftarrow}$ on the right-hand side of Equation (9) by its empirical counterpart $F_n^{\leftarrow}(s) := \inf\{x \in \mathbb{R} : F_n(x) \geq s\}$, $0 < s \leq 1$, associated to the empirical cdf defined on the real line, defined by $F_n(x) := n^{-1} \# \{X_i \leq x, 1 \leq i \leq n\}$ where $\#A$ denote the cardinality of a set $A$. After straightforward computations, we obtain the formula

$$\Pi_n[X] := \rho^{-1} \int_{0}^{1} s^{1/\rho-1} F_n^{\leftarrow}(1-s) \, ds$$

which may be rewritten, in terms of $X_{1:n}, ..., X_{n:n}$, as an $L$-statistic

$$\Pi_n[X] = \sum_{i=1}^{n} c_{i,n} X_{n-i+1:n}, \tag{10}$$

where

$$c_{i,n} \equiv (i/n)^{1/\rho} - ((i-1)/n)^{1/\rho}. \tag{11}$$

The form (10) is a linear combinations of the order statistics (see, Shorack and Wellner, 1986, page 260). The limit behavior was discussed by many authors: Chernoff *et al.* (1967), Stigler (1974), Mason (1981), Jones and Zitikis (2003) (see its Theorem 3.2 in the case that $X$ is not heavy-tailed) and in Brahimi *et al.* (2011) (in heavy-tailed case).

### 2.1. *Heavy-tailed losses case*

Let $X$ be a nonnegative rv with cdf $F \in \mathcal{RV}_{(-1/\gamma)}$. The condition (3) is equivalent to

$$\lim_{t \to 0} \frac{F^{\leftarrow}(1 - tx)}{F^{\leftarrow}(1 - t)} = x^{-\gamma}, \text{ for every } x > 0 \tag{12}$$

and we say that the function $s \to F^{\leftarrow}(1 - s)$ satisfying condition (12) is regularly varying at zero with the index $(-\gamma) < 0$. The parameter $\gamma$ is called the tail index or extreme value index (EVI). A various tail index estimators have been suggested in the literature, based for instance of the conventional maximum likelihood method, moment estimation, ... (see, e.g. Hill, 1975, Pickands, 1975, Dekkers *et al.*, 1989. Csörgő *et al.*, 1985 and Drees, 1995). For the robustness and bias reduction (see, i.e. Peng and Qi, 2004 and Stehlík *et al.*, 2012). The regular-variation condition itself is not sufficient for establishing asymptotic distributions. To this end, we suppose that cdf $F$ satisfy the well-known by the second-order condition of regular variation with second-order parameter $\tau \leq 0$, that is: there exists a function $t \to a(t)$ with constant sign at infinity and converges to 0 as $t \to \infty$ such that

$$\lim_{t \to \infty} (a(t))^{-1} \left( \frac{\overline{F}(tx)}{\overline{F}(t)} - x^{-1/\gamma} \right) = x^{-1/\gamma} \frac{x^{\tau/\gamma} - 1}{\gamma \tau}, \text{ fore very } x > 0. \tag{13}$$

When $\tau = 0$, the ratio $\dfrac{x^{\tau/\gamma} - 1}{\gamma \tau}$ should be interpreted as $\log x$.

Let $\mathbb{U}$ be the left-continuous inverse of $1/(1 - F)$. Note that $\mathbb{U}(t)$ is defined for $t > 1$, so in terms of $\mathbb{U}$ the condition (13) is equivalent to the following one

$$\lim_{t \to \infty} (A(t))^{-1} \left( \frac{\mathbb{U}(tx)}{\mathbb{U}(t)} - x^{-\gamma} \right) = x^{-\gamma} \frac{x^{\tau} - 1}{\tau}, \text{ for every } x > 0, \tag{14}$$

(see de Haan and Stadtmüller, 1996 or Theorem 3.2.9 in de Haan and Ferreira, 2006).

Using now the Weissman estimator (Weissman, 1978) of high quantiles $F^{\leftarrow}$ is given by

$$F_n^{\leftarrow(W)}(1 - s) := (k/n)^{\widehat{\gamma}} X_{n-k:n} s^{-\widehat{\gamma}}, \ s \downarrow 0, \tag{15}$$

then formula (9) can be split into

$$\Pi_\rho[X] = -\int_0^{k/n} s^{1/\rho} dF^{\leftarrow}(1 - s) - \int_{k/n}^1 s^{1/\rho} dF^{\leftarrow}(1 - s). \tag{16}$$

By using an integration by part to the second integral yields

$$\Pi_\rho[X] = (k/n)^{1/\rho} F^{\leftarrow}(1 - k/n) - \int_0^{k/n} s^{1/\rho} dF^{\leftarrow}(1 - s) + \rho^{-1} \int_{k/n}^1 s^{1/\rho - 1} F^{\leftarrow}(1 - s) ds$$

$$:= \Pi_\rho^{(1)} + \Pi_\rho^{(2)} + \Pi_\rho^{(3)}.$$

A simple estimator of $\Pi_\rho^{(1)}$ is

$$\Pi_{\rho,n}^{(1)} := (k/n)^{1/\rho} X_{n-k,n}. \tag{17}$$

To estimate $\Pi_\rho^{(2)}$, we note that $\widehat{\gamma}$ is a consistent estimator for $\gamma$ (Stehlík $et\ al.$, 2012) and since $\rho < 1/\gamma$, by substituting $F_n^{\leftarrow(W)}(1-s)$ given in Equation (15) instead of $F^{\leftarrow}(1-s)$ and integrating yield the following estimator

$$\Pi_{\rho,n}^{(2)} := (k/n)^{1/\rho} \frac{\widehat{\gamma}\rho}{1-\widehat{\gamma}\rho} X_{n-k:n}. \tag{18}$$

Finally, by plugging $F_n^{\leftarrow}$ instead of $F^{\leftarrow}$ on second integral of Equation (16) we obtain the estimator

$$\Pi_{\rho,n}^{(3)} := \sum_{i=k+1}^{n} c_{i,n} X_{n-i+1:n}, \tag{19}$$

of $\Pi_\rho^{(3)}$, where $F_n^{\leftarrow}(s) := \inf \{x \in \mathbb{R} : F_n(x) \geq s\}$, $0 < s \leq 1$, denote the sample quantile function associated to the empirical cdf defined on the real line by $F_n(x) := n^{-1} \sum_{i=1}^{n} \mathbb{I}(X_i \leq x)$, with $\mathbb{I}(\cdot)$ being the indicator function and the coefficients $c_{i,n}$ are given in Equation (11). The final form of our estimator

$$\widetilde{\Pi}_{\rho,n} := (k/n)^{1/\rho} \frac{X_{n-k,n}}{1-\widehat{\gamma}\rho} + \sum_{i=k+1}^{n} \left( (i/n)^{1/\rho} - ((i-1)/n)^{1/\rho} \right) X_{n-i+1:n}. \tag{20}$$

In next section we establish the asymptotic normality of our estimator and compared with the estimator proposed by Necir and Meraghni given in Equation (7).

## 3. Asymptotic distribution

We will begin to expose our results as asymptotic representation theorems in the lines of Beran $et\ al.$ (2014). For that purpose, we need to describe the probability theory on which they hold. Indeed, Csörgő $et\ al.$ (1986) have constructed a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ carrying a sequence of independant standard uniform random variables $U_1, U_2, ...$ and a sequence of Brownian bridges $\{\mathbb{B}_n(s); \ 0 \leq s \leq 1\}$ such that for every $0 \leq \zeta < 1/2$,

$$\sup_{1/n \leq s \leq 1-1/n} \frac{n^\zeta |\beta_n(1-s) - \mathbb{B}_n(1-s)|}{s^{1/2-\zeta}} = O_{\mathbb{P}}\left(n^{-\zeta}\right), \tag{21}$$

where, for each $n \geq 1$, $\beta_n$ the quantile process associated with $U_1, ..., U_n$. To define this quantile function, we consider the order statistics $U_{1,n} \leq ... \leq U_{n,n}$ pertaining to $U_1, ..., U_n$ and we introduce the uniform quantile function

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

from which derives the uniform quantile process

$$\beta_n(s) = \sqrt{n}(s - V_n(s)), \text{ for } n \geq 1 \text{ and } s \in [0,1].$$

We places ourselves on that probability space and use the Renyi's representations

$$\{X_i, i \geq 1\} \stackrel{d}{=} \{F^{\leftarrow}(U_{i,n}), i \geq 1\}$$

and

$$\{\{X_{i,n}, 1 \leq i \leq n\}, \ n \geq 1\} \stackrel{d}{=} \{\{F^{\leftarrow}(U_{i,n}), 1 \leq i \leq n\}, \ n \geq 1\},$$

where $F^{\leftarrow}$ denotes the quantile function pertaining to the cdf $F$ and $=^d$ stands for equality in distribution. In the following, we may and do replace the equalities in distribution by simple equalities, so that the weak and strong laws established for on random sequences depending only on $X_1$, $X_2$, etc... In that probability space will hold in any probability space by the virtue of equality in law for stochastic processes.

Our main result is the following

**Theorem 1.** *Let $F$ be a df satisfying (14) with $\gamma > 1/2$ and suppose that $F^{\leftarrow}(\cdot)$ is continuously differentiable on $[0, 1)$. Let $k = k_n$ satisfying (4) such that $\sqrt{k}A(n/k) \to 0$ as $n \to \infty$. For any $1 \leq \rho < 1/\gamma$, we have*

$$\frac{\sqrt{n}\left(\widetilde{\Pi}_{\rho,n} - \Pi_\rho\right)}{(k/n)^{1/\rho - 1/2}F^{\leftarrow}(1 - k/n)} \stackrel{d}{\to} \mathcal{N}\left(0, \sigma^2(\gamma, \rho)\right), \ as \ n \to \infty,$$

*where*

$$\sigma^2(\gamma, \rho) = \gamma^2 + \frac{\gamma^2 \rho \left(\rho - 2\rho\gamma^2 + 2\gamma\right)}{(\gamma\rho - 1)^2} + \frac{2\gamma^2}{(\rho + \gamma\rho - 1)(\rho + 2\gamma\rho - 2)}$$
$$+ \frac{2\gamma}{2\gamma - 1} - \frac{2\gamma\rho\left(\rho\gamma^2 - \rho\gamma + 1\right)}{(\gamma\rho - 1)(\rho + \gamma\rho - 1)}.$$

## 4. Simulation study

### 4.1. Performance and comparative study of $\widetilde{\Pi}_{\rho,n}$ and $\widehat{\Pi}_n$

In this simulation study we examine the performance of our estimator $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k)$ given in (20) and compare it with that of $\widehat{\Pi}_{\rho,n}(\widehat{\gamma}^H, k)$ given in (7). Thus we follow the steps below.

**Step 1:** We generate 1000 pseudorandom samples of size $n = 100, 200, 500$ and 1000 from Pareto cdf with $\gamma = 0.6$.

**Step 2:** We estimate the tail index parameter by Hill and t-Hill estimators $\widehat{\gamma}^H(k_1^*)$ and $\widehat{\gamma}(k_2^*)$, respectively given in (6) and (8). We adopt the Reiss and Thomas algorithm (see Reiss and Thomas, 2007, page 137), for choosing the optimal numbers of upper extremes $k_1$ and $k_2$. By this methodology, we define the optimal sample fraction of upper order statistics $k_j^*$ by

$$k_j^* := \arg\min_k \frac{1}{k} \sum_{i=1}^{k} i^\theta \left|\widehat{\gamma}_j(i) - \text{median}\left\{\widehat{\gamma}_j(1), ..., \widehat{\gamma}_j(k)\right\}\right|, j = 1, 2$$

where $\widehat{\gamma}_1 = \widehat{\gamma}^H$ and $\widehat{\gamma}_2 = \widehat{\gamma}$. On the light of our simulation study, we obtained reasonable results by choosing $\theta = 0.3$. For an illustration of the behavior of $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k)$ as a function of $n$, see Figure 1.

**Step 3:** We fix the distortion parameter with respect to Condition (5) by $\rho = 1.12$, then we compute the bias and RMSE of the four estimators $\widehat{\gamma}^H(k_1^*)$, $\widehat{\gamma}(k_2^*)$, $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k_1^*)$ and $\widehat{\Pi}_{\rho,n}(\widehat{\gamma}^H, k_2^*)$. The results are summarized in Table 1. We see that when dealing with large samples our estimator performs better.

| $n$ | $k_1^*$ | $\widehat{\gamma}(k_1^*)$ | | $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k_1^*)$ | | $k_2^*$ | $\widehat{\gamma}^H(k_2^*)$ | | $\widehat{\Pi}_{\rho,n}(\widehat{\gamma}^H, k_2^*)$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | bias | RMSE | bias | RMSE | | bias | RMSE | bias | RMSE |
| 100 | 10 | $-0.0733$ | 0.2511 | 0.3618 | 0.5199 | 17 | $-0.1641$ | 0.2865 | 0.4096 | 0.7332 |
| 200 | 23 | $-0.0571$ | 0.1821 | 0.3562 | 0.5147 | 34 | $-0.0993$ | 0.2350 | 0.3918 | 0.7185 |
| 500 | 62 | $-0.0299$ | 0.1142 | 0.3404 | 0.4820 | 86 | $-0.0301$ | 0.0739 | 0.3639 | 0.6936 |
| 1000 | 129 | $-0.0147$ | 0.0798 | 0.1966 | 0.2687 | 169 | $-0.0181$ | 0.0545 | 0.2827 | 0.5279 |

**Table 1.** $\widehat{\gamma}(k_1^*)$, $\widehat{\gamma}^H(k_2^*)$ $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k_1^*)$ and $\widehat{\Pi}_{\rho,n}(\widehat{\gamma}, k_2^*)$ estimators based on 1000 samples of Pareto-distributed claim amounts with tail index 0.6 and distortion parameter $\rho = 1.12$. The exact value of the premium is 2.0487.

### 4.2. Comparative robustness study

In this subsection we study the sensitivity to outliers of $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k_1^*)$ and compare it with that of $\widehat{\Pi}_{\rho,n}(\widehat{\gamma}^H, k_2^*)$. We consider an $\epsilon$-contaminated model known by mixture of Pareto distributions

$$F_{\gamma_1,\gamma_2,\epsilon}(x) = 1 - (1-\epsilon)x^{-1/\gamma_1} + \epsilon x^{-1/\gamma_2}, \tag{22}$$

where $\gamma_1, \gamma_2 > 0$ and $0 < \epsilon < 0.5$ is the fraction of contamination. Note that for $\epsilon = 0$, $\widehat{\gamma}^H$ and $\widehat{\gamma}$ are asymptotically unbiased. Therefore, for $\epsilon > 0$, the effect of contamination becomes immediately apparent. If $\gamma_1 < \gamma_2$ and $\epsilon > 0$, (22) corresponds to a Pareto distribution contaminated by a longer tailed distribution. For the implementation of mixtures models to the outliers study one refers, for instance, to (Barnett and Lewis, 1995, page 43). In this context, we proceed our study as follows.

First, we consider $\gamma_1 = 0.6$, $\gamma_2 = 2$ to have the contaminated model and let $\rho = 1.12$. Then we consider four contamination scenarios according to $\epsilon = 5\%, 10\%, 15\%, 25\%$.

For each value $\epsilon$, we generate 1000 samples of size $n = 100$, 200 and 1000 from the model (22). Finally, we compare the $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k_1^*)$ and $\widehat{\Pi}_{\rho,n}(\widehat{\gamma}^H, k_2^*)$ estimators with this true value, by computing for each estimator, the appropriate bias and RMSE and summarize the results in Table 2.
As expected, the estimator $\widehat{\Pi}_{\rho,n}(\widehat{\gamma}^H, k_2^*)$ as well as $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k_1^*)$ turn out to be more sensitive to this type of contaminations. For example, in 0% contamination for $n = 200$, the couple (bias, RMSE) for $\widehat{\Pi}_{\rho,n}(\widehat{\gamma}^H, k_2^*)$ take the values $(0.3918, 0.7185)$, while for 15% contamination
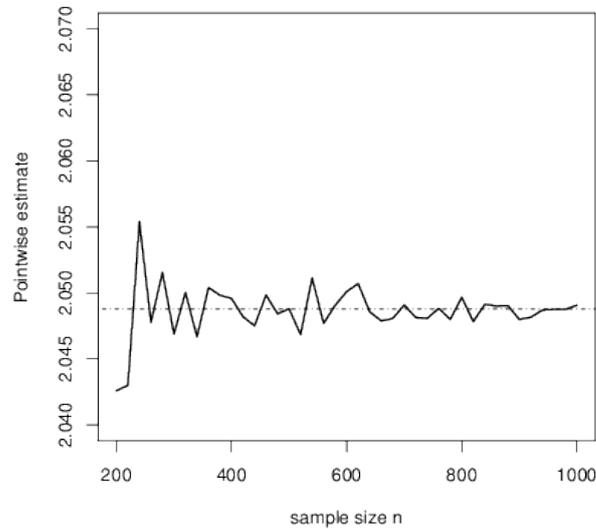
**Fig. 1.** Plot of the premium estimator $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k_1)$ as a function of the sample size $n$ based on 200 Pareto-distributed claims with tail index 0.6. The distortion parameter is 1.12. The horizontal doted line represents the true value of the premium 2.0487.

the bias and the RMSE are given by the couple $(-1.4355, 2.3107)$. We may conclude that the bias and RMSE of $\widehat{\Pi}_{\rho,n}(\widehat{\gamma}^H, k_2^*)$ estimator are more sensitive (or note robust) to outliers. However, for 0% contamination the (bias, RMSE) of $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k_1^*)$ is $(0.3562, 0.5147)$, while for 15% contamination is $(0.4508, 0.6870)$. Both bias and RMSE of $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k_1^*)$ estimation are note sensitive to outliers. Then we may conclude that is the better estimator.

## 5. Concluding notes

We showed that the new estimator of premium based on t-Hill estimator is more robust and performs better than the one based on Hill estimator proposed by Necir and Meraghni (2009). Our estimator $\widetilde{\Pi}_{\rho,n}(\widehat{\gamma}, k)$ is based on Weissman's estimation of high quantiles, so we would lead to improve our result to use one of several bias-reduced estimators have been proposed (see for example Matthys and Beirlant, 2003).

**Acknowledgement.** We thank the editor and the referee for their constructive and useful comments that led to a much improved paper.

## 6. Proofs

To establish the asymptotic normality of $\widetilde{\Pi}_{\rho,n}$ we need the asymptotic approximation of $\widehat{\gamma}$ with the same sequence of Brownian bridges as $\widetilde{\Pi}_{\rho,n}$, for this reason we give the following results.

| $n$ | % contamination | $\widetilde{\Pi}_{\rho,n}\left(\widehat{\gamma}, k_1^*\right)$ bias | RMSE | $\widehat{\Pi}_{\rho,n}\left(\widehat{\gamma}^H, k_2^*\right)$ bias | RMSE |
|---|---|---|---|---|---|
| 100 | 5 | 0.4043 | 0.6664 | −0.4286 | 1.4727 |
| | 10 | 0.4389 | 0.6862 | −0.7291 | 1.9123 |
| | 15 | 0.4598 | 0.7464 | −1.2786 | 2.1247 |
| | 25 | 1.0578 | 1.1305 | −1.2103 | 2.1828 |
| 200 | 5 | 0.3831 | 0.5532 | −0.4713 | 1.5118 |
| | 10 | 0.3964 | 0.5675 | −1.2496 | 2.1907 |
| | 15 | 0.4508 | 0.6870 | −1.4355 | 2.3107 |
| | 25 | 0.9470 | 1.0197 | −1.7366 | 2.3274 |
| 1000 | 5 | 0.2124 | 0.3211 | −0.3794 | 2.1222 |
| | 10 | 0.2329 | 0.3349 | −1.0662 | 2.3978 |
| | 15 | 0.2931 | 0.3749 | −1.2501 | 2.0355 |
| | 25 | 0.8124 | 0.9291 | −1.5238 | 2.3596 |

**Table 2.** $\widetilde{\Pi}_{\rho,n}\left(\widehat{\gamma}, k_1^*\right)$ and $\widehat{\Pi}_{\rho,n}\left(\widehat{\gamma}^H, k_2^*\right)$ are based on 1000 samples of mixture of Pareto distributions with tail index 0.6, $\epsilon = 5\%, 10\%, 15\%, 25\%$ and distortion parameter $\rho = 1.12$. The exact value of the premium is 2.0487.

**Proposition 1.** *Assume that the second order condition (14) holds with $\gamma > 1/2$ and let $k = k_n$ be an integer sequence satisfying (4) and $\sqrt{k}A\left(n/k\right) \to 0$. Then, there exists a sequence of Brownian bridges $\{\mathbb{B}_n\left(s\right), \ 0 \le s \le 1\}$ such that*

$$\sqrt{k}\left(\widehat{\gamma} - \gamma\right) = \gamma\left(\gamma + 1\right)^2 \int_0^1 s^{\gamma-1}\mathbb{B}_n\left(s\right) ds + o_p\left(1\right),$$

*leading to*

$$\sqrt{k}\left(\widehat{\gamma} - \gamma\right) \overset{d}{\to} \mathcal{N}\left(0, \frac{\gamma^2\left(1 + \gamma\right)^2}{\left(1 + 2\gamma\right)}\right), \ as \ n \to \infty,$$

*Proof.* Our proofs are conducted in the probability space described in Section 3. Then we are entitled to write

$$S_k := \frac{1}{k}\sum_{j=1}^{k} \frac{X_{n-k:n}}{X_{n-j+1:n}}$$

then $\widehat{\gamma} = S_k^{-1} - 1$. Here we suppose that $\sqrt{k}A\left(n/k\right) \to \lambda = 0$. By using similar arguments and techniques used in Brahimi *et al.* (2013) and by considering the approximation given in (21) (see, Csörgő *et al.*, 1986) it follows that

$$\sqrt{k}\left(S_k - \frac{1}{\gamma + 1}\right) = -\gamma \int_0^1 s^{\gamma-1}\mathbb{B}_n\left(s\right) dt + o_p\left(1\right).$$

Using now the map $g\left(x\right) = 1/x - 1$, since $g\left(1/\left(\gamma + 1\right)\right) = \gamma$ and applying the delta method yields

$$\sqrt{k}\left(\widehat{\gamma} - \gamma\right) = \left(\gamma + 1\right)^2 \gamma \int_0^1 s^{\gamma-1}\mathbb{B}_n\left(s\right) ds + o_p\left(1\right).$$

It is clear that $\sqrt{k}\left(\widehat{\gamma} - \gamma\right)$ is a Gaussian rv with mean 0 and variance $\frac{\gamma^2(1+\gamma)^2}{(1+2\gamma)}$. This completes the proof of Proposition 1.                                                                                                       □

### 6.1. *Proof of Theorem 1*

By using Proposition 1 and from Necir *et al.* (2007) we show that under the assumptions of Theorem 1, we use the same sequence of Brownian bridges $\{\mathbb{B}_n(s), 0 \le s \le 1\}$ such that, for all large $n$

$$\frac{\left(\Pi_{\rho,n}^{(1)} - \Pi_{\rho}^{(1)}\right)}{(k/n)^{1/\rho}F^{\leftarrow}(1 - k/n)} = -\gamma(n/k)^{1/2}\mathbb{B}_n(1 - k/n) + o_p(1).$$

Let $Y_1, Y_2....$ be iid rv's with cdf $1 - 1/y, y > 1$, and let $Y_{1,n} \le Y_{2,n} \le ... \le Y_{n,n}$ be the associated order statistics. Note that

$$\{X_{n-i+1:n}, 1 \le i \le n\} \overset{d}{=} \{\mathbb{U}(Y_{n-i+1:n}), 1 \le i \le n\}$$

Then, for $\rho \ge 1$ and without loss of generality, we may rewrite the statistic $\Pi_{\rho,n}^{(2)}$ as follows

$$\Pi_{\rho,n}^{(2)} = (k/n)^{1/\rho}\frac{\widehat{\gamma}\rho}{1 - \widehat{\gamma}\rho}\mathbb{U}(Y_{n-k:n}),$$

where $\widehat{\gamma}$ is the t-Hill estimator of $\gamma$. So we have

$$\frac{\sqrt{k}\left(\Pi_{\rho,n}^{(2)} - \Pi_{\rho}^{(2)}\right)}{(k/n)^{1/\rho}\mathbb{U}(n/k)} := \sum_{j=1}^{4}\Delta_{jn},$$

where

$$\Delta_{1n} := \sqrt{k}\frac{\widehat{\gamma}\rho}{1 - \widehat{\gamma}\rho}\left(\frac{\mathbb{U}(Y_{n-k:n})}{\mathbb{U}(n/k)} - \left(\frac{Y_{n-k:n}}{n/k}\right)^{\gamma}\right),$$

$$\Delta_{2n} := \sqrt{k}\frac{\widehat{\gamma}\rho}{1 - \widehat{\gamma}\rho}\left(\left(\frac{Y_{n-k:n}}{n/k}\right)^{\gamma} - 1\right),$$

$$\Delta_{3n} := \sqrt{k}\left(\frac{\widehat{\gamma}\rho}{1 - \widehat{\gamma}\rho} - \frac{\gamma\rho}{1 - \gamma\rho}\right)$$

and

$$\Delta_{4n} := \sqrt{k}(k/n)^{-1/\rho}\left(\frac{\frac{(k/n)^{1/\rho}\rho}{(1/\gamma - \rho)}\mathbb{U}(n/k) - \Pi_{\rho}^{(2)}}{\mathbb{U}(n/k)}\right).$$

As showed in Necir *et al.* (2007), we have : $\Delta_{1n} \to 0$ and $\Delta_{4n} \to 0$ as $n \to \infty$.

Next, we show that $\Delta_{2n} + \Delta_{3n}$ is asymptotically normal. Assume, without loss of generality, that the rv's $Y_1, Y_2, ..., Y_n$ are defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ which carries the sequence $(U_n)_{n \ge 1}$ in such a way that $Y_n = (1 - U_n)^{-1}$ for $n = 1, 2, ...$ and $Y_{i,n} = (1 - U_{i,n})^{-1}$,

$i = 1, ..., n,$. Then, this allows us to write $Y_{n-i+1,n} = (1 - V_n (1 - s))^{-1}$, for $\dfrac{i-1}{n} < s \le \dfrac{i}{n}$, $i = 1, ..., n$. From Necir *et al.* (2007) we have For $\Delta_{2n}$

$$\Delta_{2n} = - (n/k)^{1/2} \frac{\rho \gamma^2}{1 - \gamma \rho} (1 + o_p (1)) \mathbb{B}_n (1 - k/n) .$$

For $\Delta_{3n}$ and by using the map $h(\theta) = \rho / \left(\frac{1}{\theta} - \rho\right)$ and applying the delta method yields:

$$\Delta_{3n} = \frac{\rho}{(\gamma \rho - 1)^2} \sqrt{k} (\gamma - \widehat{\gamma}) .$$

From Proposition 1 we get

$$\Delta_{3n} = \frac{\gamma \rho (\gamma + 1)^2}{(\gamma \rho - 1)^2} \int_0^1 s^{\gamma - 1} \mathbb{B}_n (s) \, ds + o_p (1) .$$

Finally we have

$$\frac{\sqrt{k} \left(\Pi_{\rho,n}^{(2)} - \Pi_\rho^{(2)}\right)}{(k/n)^{1/\rho} F^{\leftarrow} (1 - k/n)} = \frac{\gamma \rho (\gamma + 1)^2}{(\gamma \rho - 1)^2} \int_0^1 s^{\gamma - 1} \mathbb{B}_n (s) \, ds - (n/k)^{1/2} \frac{\rho \gamma^2}{1 - \gamma \rho} \mathbb{B}_n (1 - k/n) + o_p (1) .$$

From Necir *et al.* (2007) we have

$$\frac{\sqrt{k} \left(\Pi_{\rho,n}^{(3)} - \Pi_\rho^{(3)}\right)}{(k/n)^{1/\rho} F^{\leftarrow} (1 - k/n)} = \frac{\rho^{-1} \int_{k/n}^1 s^{1/\rho - 1} \mathbb{B}_n (1 - s) F^{\leftarrow\prime} (1 - s) \, ds}{F^{\leftarrow} (1 - k/n) (k/n)^{1/\rho - 1/2}} + o_p (1) .$$

Then

$$\frac{\sqrt{n} \left(\widetilde{\Pi}_{\rho,n} - \Pi_\rho\right)}{(k/n)^{1/\rho - 1/2} F^{\leftarrow} (1 - k/n)} = \Lambda (\gamma, \rho) + o_p (1)$$

where

$$\Lambda (\gamma, \rho) := W_{n1} + W_{n2} + W_{n3} + o_p (1)$$

and

$$W_{n1} := - (n/k)^{1/2} \gamma \mathbb{B}_n (1 - k/n) ,$$

$$W_{n2} := \frac{\gamma \rho (\gamma + 1)^2}{(\rho \gamma - 1)^2} \int_0^1 s^{\gamma - 1} \mathbb{B}_n (s) \, ds - (n/k)^{1/2} \frac{\gamma^2 \rho}{1 - \gamma \rho} \mathbb{B}_n (1 - k/n) ,$$

$$W_{n3} := \frac{\rho^{-1} \int_{k/n}^1 s^{1/\rho - 1} \mathbb{B}_n (1 - s) F^{\leftarrow\prime} (1 - s) \, ds}{F^{\leftarrow} (1 - k/n) (k/n)^{1/\rho - 1/2}} .$$

It is clear that $\Lambda (\gamma, \rho)$ is a Gaussian rv with mean 0 and variance

$$E (\Lambda (\gamma, \rho))^2 = E \left(W_{n1}^2\right) + E \left(W_{n2}^2\right) + E \left(W_{n3}^2\right) + 2E (W_{n1} W_{n2}) + 2E (W_{n1} W_{n3}) + 2E (W_{n2} W_{n3}) .$$

An elementary calculation gives, we get

$$
\begin{aligned}
E\left(W_{n1}^2\right) &= \gamma^2 + o\left(1\right), \\
E\left(W_{n2}^2\right) &= \frac{\gamma^2\rho^2}{(1-\gamma\rho)^2} + \frac{2\gamma}{2\gamma-1} + o\left(1\right), \\
E\left(W_{n3}^2\right) &= \frac{2\gamma^2}{(\rho+\gamma\rho-1)(\rho+2\gamma\rho-2)} + o\left(1\right), \\
E\left(W_{n1}W_{n2}\right) &= \frac{\gamma^3\rho}{1-\gamma\rho} + o\left(1\right), \\
E\left(W_{n1}W_{n3}\right) &= \frac{\gamma\rho}{\rho+\gamma\rho-1} + o\left(1\right)
\end{aligned}
$$

and

$$
E\left(W_{n2}W_{n3}\right) = -\frac{\gamma^3\rho^2}{(\gamma\rho-1)(\rho+\gamma\rho-1)} + o\left(1\right).
$$

The proof of Theorem 1 is completed by combining all the preceding results. □

## References

Artzner, P., Delbaen, F., Eber, J.M., Heath, D., 1999. Coherent measures of risk. *Mathematical Finance.* **9**, 203-228.

Barnett, V. and Lewis, T., 1995. *Outliers in statistical data.* Third ed. John Willey & Sons. New York.

Beirlant, J., Matthys, G. and Dierckx, G., 2001. Heavy-tailed distributions and rating. Astin Bulletin 31, 37 58.

Beran, J.; Schell, D.; Stehlík, M., 2014. The harmonic moment tail index estimator: asymptotic distribution and robustness. *Ann. Inst. Statist. Math.* **66**, 193-220.

Brahimi, B., Meraghni, D., Necir, A. and Yahia, D., 2013. A bias-reduced estimator for the mean of a heavy-tailed distribution with an infinite second moment. *J. Statist. Plann. Inference.* **143**(6), 1064-1081.

Brahimi, B., Meraghni, D., Necir, A. and Zitikis, R., 2011. Estimating the distortion parameter of the proportional-hazard premium for heavy-tailed losses. *Insurance Math. Econom.* **49**(3), 325-334.

Chernoff, H. Gastwirth J.L. andJohns, M.V. 1997. Asymptotic distribution of linear combinations of functions of order statistics with applications to estimation, *Annals of Mathematical Statistics.* **38**, 52-72.

Csörgő, M., Csörgő, S., Horváth, L. and Mason, D.M., 1986. Weighted empirical and quantile processes. *Ann. Probab.* **14**(1), 31-85.

Csörgő, S., Deheuvels, P. and Mason, D., 1985. Kernel estimates of the tail index of a distribution. *Ann. Statist.* **13** (3), 1050-1077.

Dekkers, A.L.M., Einmahl, J.H.J. and de Haan, L., 1989. A moment estimator for the index of an extreme value distribution. *Ann. Statist.* **17** (4), 1833-1855.

de Haan, L. and Ferreira, A., 2006. *Extreme Value Theory: An Introduction.* Springer Series in Operations Research and Financial Engineering. Springer, New York.

de Haan, L. and Stadtmüller, U., 1996. Generalized regular variation of second order. J. Australian Math. Soc. (Series A) **61**, 381-395.

Drees, H., 1995. Refined Pickands estimators of the extreme value index. *Ann. Statist.* **23**(6), 2059-2080.

Fabián, Z., 2001. Induced cores and their use in robust parametric estimation, *Communication in Statistics, Theory Methods.* **30**, 537-556.

Fabiá, Z. and Stehlík, M., 2009. On robust and distribution sensitive Hill like method. IFAS Research Paper Series. 2009-43

Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. *Ann. Statist.* **3**(5), 1163-1174.

Jones, B.L., Zitikis, R., 2003. Empirical estimation of risk measures and related quantities. *North American Actuarial Journal.* 7, 44-54.

Júarez, S.F. and Schucany, W.R., 2004. Robust and efficient estimation for the generalized Pareto distribution. *Extremes*, **7**, 237-251.

Kim, M. and Lee, S., 2008. Estimation of a tail index based on minimum density power divergence. *Journal of Multivariate Analysis*, **99**, 2453-2471.

Peng, L., Qi, Y., 2004. Estimating the first- and second-order parameters of a heavy-tailed distribution. *Aust. N. Z. J. Stat.* **46**(2), 305-312.

Peng, L. and Welsh, A.H., 2001. Robust estimation of the generalized Pareto distribution. *Extremes*, **4**, 53-65.

Pickands, J., 1975. Statistical inference using extreme order statistics. *Ann. Statist.* **3**, 119-131.

Mason, D.M., 1981. Asymptotic normality of linear combinations of order statistics with a smooth score function, *Annals of Statistics.* **9**, 899-908.

Matthys, G. and Beirlant, J., 2003. Estimating the extreme value index and high quantiles with exponential regression models. *Statist. Sinica.* **13**(3), 853-880.

Necir, A. and Boukhetala, K., 2004. Estimating the risk-adjusted premium for the largest claims reinsurance covers. COMPSTAT 2004. *Proceedings in Computational Statistics*, 1577-1584.

Necir, A., Meraghni, D. and Meddi, F., 2007. Statistical estimate of the proportional hazard premium of loss. *Scand. Actuar. J.* **3**, 147-161.

Necir, A., Meraghni, D., 2009. Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts. *Insurance Math. Econom.* **45**, 49-58.

Reiss, R.-D., Thomas, M., 2007. Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields, 3rd ed. Birkhäuser Verlag, Basel, Boston, Berlin.

Rolski, T., Schimidli, H., Schmidt, V. and Teugels, J.L., 1999. *Stochastic Processes for Insurance and Finance.* John Wiley & Sons, Chichester.

Shorack, G., R. and Wellner, J., A., 1986. *Empirical Processes with Applications to Statistics.* Wiley.

Stehlík, M., Fabi án, Z. and Střelec, L. 2012. Small sample robust testing for normality against Pareto tails. *Comm. Statist. Simulation Comput.* **41**(7), 1167-1194.

Stigler, S. 1974. Linear functions of order statistics with smooth weight functions. *Annals of Statistics.* **2**, 676-693.

Vandewalle, B., Beirlant, J., Christmann, A. and Hubert, M., 2007. A robust estimator for the tail index of Pareto-typ e distributions . Computational Statistics & Data Analysis,

**51**, 6252-6268.

Wang, S. S., 1996. Premium calculation by transforming the layer premium density. *ASTIN Bulletin.* **26** , 71-92.

Weissman, I., 1978. Estimation of parameters and large quantiles based on the $k$ largest observations. Journal of American Statistical Association **73**, 812-815.

Wirch, J. L., and Hardy, M. R., 1999. A synthesis of risk measures for capital adequacy. Insurance: Mathematics and Economics. **25**(3), 337-347.