Hacettepe Journal of Mathematics and Statistics
Volume 44 (3) (2015), 735 – 746

An improved estimator of the distortion risk measure for heavy-tailed claims

Rassoul Abdelaziz*

Abstract

The main aim of this paper is to propose an alternative estimate of the distortion risk measure for heavy-tailed claims. Our approach is based on the result of Balkema and de Haan (1974) [3], and Pickands (1975) [22] for approximating the tail of the distribution by a generalized Pareto distribution. The asymptotic normality of the new estimator is established, and its performance illustrated by some results of simulation who shows the advantages of the new estimator over the estimator based on the classical extreme-value theory.

Keywords: Premium principle, Distortion risk measure, POT method, Extremes values theory, Generalized Pareto distribution, Loss distribution.

Received 11/02/2014 : Accepted 18/06/2014 Doi : 10.15672/HJMS.2014207471

1. Introduction

A number of risks measures found in finance and insurance literature are special cases of the distortion risk measure, defined by

$$H[F,g] = \int_0^{+\infty} g\left(\overline{F}(x)\right) dx. \tag{1.1}$$

where $X \geq 0$ is a loss random variable with cumulative distribution function (cdf) F and the de-cumulative distribution function (ddf) $\overline{F} = 1 - F$, which is also known as survival function. The distortion function $g : [0,1] \to [0,1]$ is assumed to be an increasing function such that $g(0) = 0$ and $g(1) = 1$.
Dhaene et al. (2012) [9] show that, when the distortion function g is right continuous on $[0,1)$, the formula (1.1) may be rewritten as follows

$$H[F,g] = \int_0^1 \mathbb{Q}(1-s)dg(s), \tag{1.2}$$

*Laboratoire de Génie de l'eau et Environement (GEE)
Ecole Nationale Supérieure d'Hydraulique, Blida, Algérie,
Email: `a.rassoul@ensh.dz`

where $\mathbb{Q}$ is the quantile function corresponding the cdf F, that is

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

The risk measure $H[F, g]$, which can also be viewed as a premium calculation principle, has manifested in the econometric literature, particularly in Yaari's (1987) [31] dual theory of choice under risk, and has been introduced into actuarial literature by Wang (1996) [28]. A number of risk measures of this form have been discussed by Wirch and Hardy (1999) [30].
In Artzner (1999) [1] and Artzner et al. (1999) [2] a risk measure satisfying the four axioms of subadditivity, monotonicity, positive homogeneity and translation invariance is called Coherent, and also demonstrated that the risk measure $H[F, g]$ is coherent when g is concave. Note that the class of concave distortion risk measures is only a subset of the class of coherent risk measures.
Many special cases that have arisen in the finance and insurance literature are such:

- VaR: $g(x) = 1_{[1-q,1]}$ for some $q \in]0, 1[$
- Tail-VaR: $g(x) = \min\{\frac{x}{1-q}, 1\}$ for some $q \in (0, 1)$
- Proportional Hazard Transform: $g(x) = x^{1/\rho}$ for some $\rho > 1$
- Dual-Power Transform: $g(x) = 1 - (1 - x)^{\rho}$ for some $\rho > 1$
- Gini principle: $g(x) = (1 + \rho)x - \rho x^2$, with $0 < \rho \leq 1$.
- Lookback distortion: $g(x) = x^{\rho}(1 - \rho \ln(x))$, with $0 < \rho \leq 1$.

Detailed studies of distortion risk measures, also known as Wang's risk measures, can be found in, for example, Wang (1996) [28], Wang and Young (1998) [29], Hürlimann (1998) [12], and Hua and Joe, (2012) [13].

A number of authors have tackled the distortion risk measure from the statistical inferential point of view. A short survey and classification of papers in the area follows:

- Light-tailed distributions
  - Classical-type asymptotic results
  - Asymptotic results aimed at variance reduction
- Heavy-tailed distributions
  - Fisher-Tippett-Gnedenko type extreme-value methods
  - Pickands-Balkema-de Haan type Peack Over Threshold methods

Jones and Zitikis (2003) [16] noticed that the empirical counterpart of $H[F, g]$ is a linear combination of order statistics, commonly known as L-statistic. This opens up a fruitful venue for developing statistical inferential results, which have been actively investigated by a number of researchers. Specifically, let $X_1, ..., X_n$ be independent copies of X; and let $X_{1,n}, ..., X_{n,n}$ be the corresponding ascending order statistics. The empirical estimator of the risk premium $H[F, g]$ is obtained by substituting the quantile $\mathbb{Q}$ on the right-hand side of equation (1.2) by its empirical counterpart

$$\mathbb{Q}_n(s) := inf\{x : F_n(x) \geq s\} := F_n^{-1}(s),$$

on the real line, defined by

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{X_i \leq x\}},$$

with $1_{\{.\}}$ being the indicator function. After straightforward computation, we obtain the formula

$$\widehat{H}_n[F_n, g] = \int_0^1 \widehat{Q}_n(1 - s)dg(s),$$

where $\widehat{Q}_n(1-s)$ is an empirical estimator of the quantile function, given by the formula

$$\widehat{Q}_n(1-s) := X_{n-k+1,n}, \text{ where } \frac{k-1}{n} < s \leq \frac{k}{n},$$

Then, the empirical estimator of $H[F,g]$ is given by the formula

$$\widehat{H}_n[F,g] = \sum_{i=1}^{n} \left(g\left(\frac{i}{n}\right) - g\left(\frac{i-1}{n}\right) \right) X_{n-i+1,n}.$$

For recent literature on statistical inference for distortion premiums, we refer to Jones and Zitikis (2003) [16], Jones and Zitikis (2007) [17], Centeno and Andrade (2005) [8], Furman and Zitikis (2008) [10], Brazauskas et al. (2008) [6], Greselin et al. (2009) [11], Necir et al. (2010) [20], Joseph H. T. Kim. (2010) [18], Peng et al. (2012) [21] and the references therein.
The asymptotic normality of the estimator $\widehat{H}_n[F,g]$ is established by Jones and Zitikis (2003) [16] as follows

$$\sqrt{n}\left(\widehat{H}_n[F_n,g] - H[F,g]\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,\sigma^2\right),$$

in particular, if g is differentiable, we have

$$\sigma^2 := \int_0^1 \int_0^1 (\min\{F(t),F(s)\} - F(t)F(s))\, g(1-F(t))\, g(1-F(s))\, dtds,$$

by provided that the second moment are finite, that is $E(X^2) < \infty$. This is a very restrictive condition in the context of heavy-tailed distributions as the following considerations show. Assume that the rv X_1 follows the Fréchet law with index $\gamma > 0$, that is, $1 - F(x) = exp\left\{-x^{-1/\gamma}\right\}$ for $x > 1$. When $\gamma \in (0.5, 1]$, the mean exist, but the second moment $E\left(X^2\right)$ is infinite. Hence, the range is not covered by the CLT and thus, another approach to handle this situation is needed. Making use of the results of Balkema and de Haan (1974) [3], and Pickands (1975) [22] to approximate the tail of the distribution by the Generalized Pareto Distribution (GPD), this result is know by the Peack Over Threshold method (POT) to propose a alternative estimator for the distortion risk premiums. Moreover, under suitable assumptions we established its asymptotic normality, and we presente some results of simulation to illustrate the performance of our estimator applying to the proportional hazard premium PHP. Empirical studies have shown that Financial and actuarial data exhibit heavy tails or Pareto like distributions. The class of regularly varying cdf's is a major subclass of heavy-tailed distributions, it includes distributions such as Pareto, Burr, Student, Lévy-stable, and loggamma, which are known to be appropriate models for fitting large insurance claims, large fluctuations of prices, log-returns, etc. (see, e.g., Beirlant et al., 2001 [4]; Reiss and Thomas, 2007 [23] and Rolski et al., 1999 [25]).

Note that throughout this paper, the standard notations $\xrightarrow{\mathbb{P}}$, $\xrightarrow{\mathcal{D}}$ and $\stackrel{d}{=}$ 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 $N_2(\mu,\Sigma)$ denote the bivariate normal distribution with mean vector μ and matrix of variance-covariance Σ.

The paper is organized as follows. In section 2, we introduice the differents notions and definitions of the used tools and the mains assumptions. In sections 3 we introduice the new estimator of $H_{g,n}$, and presente the main result about the limiting behavior of the proposed estimator. Some results of simulation and illustration are given in section 4. The Proofs of the mains results are postponed until section 5.

2. Main assumptions, notations, and the POT method

Distortion functions. We assume that the distortion function g is regularly varying at infinity, with index of regular variation $r \in [0,1]$, that is,

$$(2.1) \qquad g(x) = x^r \ell(x),$$

where ℓ is a slowly varying function, that is, $\ell(tx)/\ell(x) \to 1$ when $x \to \infty$ for any $t > 0$. For further properties of these functions, we refer to, for example, Resnick (1987) [24], Seneta (1976) [26]. Examples of such distortion functions are:

- VaR: $r = 0$ and $\ell(x) = 1_{[1-q,1]}(x)$
- Tail-VaR: $r = 1$ and $\ell(x) = 1/(1-q)$
- Proportional Hazard Transform: $r = 1/\rho$ and $\ell(x) = 1$
- Dual-Power Transform: $r = 1$ and $\ell(x) = \rho - \frac{\rho(\rho-1)}{2}x + o(x)$
- Gini Principle: $r = 1$ and $\ell(x) = 1 + \rho - \rho x$.
- Lookback distortion: $r = \rho$ and $\ell(x) = (1 - \rho \ln(x))$.

Distribution functions. We deal only with losses X that are heavy tailed. More specifically, we work within the class of of regularly varying cdf's. Namely, the survival function or the tail of cdf F is said to be with regulary varying at infinity, that is

$$(2.2) \qquad \overline{F}(x) = cx^{-1/\xi}\left(1 + x^{-\delta}\mathbb{L}(x)\right) \text{ when } x \to \infty,$$

for $\xi \in (0,1)$, $\delta > 0$ and some real constant c, where $\mathbb{L}$ a slowly varying function.

The POT method. Let $X_1, ..., X_n$ be independent and identically distributed random variables, each with the same cdf F, and let u_n be some a large number, 'high level,' which we later let tend to infinity when $n \to \infty$. With the notation

$$\overline{F}_{u_n}(y) = \mathbf{P}[X_1 - u_n > y \mid X_1 > u_n],$$

we have that

$$\overline{F}_{u_n}(y) = \frac{\overline{F}(u_n + y)}{\overline{F}(u_n)},$$

and thus

$$(2.3) \qquad \overline{F}_{u_n}(y) = \left(1 + \frac{y}{u_n}\right)^{-1/\xi}\left[\frac{1 + (u_n + y)^{-\delta}\mathbb{L}(u_n + y)}{1 + u_n^{-\delta}\mathbb{L}(u_n)}\right].$$

Upon recalling the definition of the generalised Pareto distribution, we have that, for all parameter values $\beta > 0$ and $\xi > 0$,

$$(2.4) \qquad \mathbb{G}_{\xi,\beta}(y) = 1 - \left(1 + \xi\frac{y}{\beta}\right)^{-1/\xi}, \quad 0 \le y < \infty.$$

We see that, the right-hand side of equation (2.3) is a perturbed version of $\mathbb{G}_{\xi,\beta_n}(y)$, with the notation $\beta_n = u_n\xi$. Balkema and de Haan (1974) [3], and Pickands (1975) [22] have shown that F_{u_n} is approximated by a generalized Pareto distribution GPD function $\mathbb{G}_{\xi,\beta_n}$ with shape parameter $\xi \in \mathbb{R}$ and scale parameter $\beta = \beta(u_n)$, in the following sense:

$$(2.5) \qquad \sup_{y>0} |F_{u_n}(y) - \mathbb{G}_{\xi,\beta}(y)| = O(u_n^{-\delta}\mathbb{L}(u_n)),$$

where, for any $\delta > 0$, we have $u_n^{-\delta}\mathbb{L}(u_n) \to 0$ when $u_n \to \infty$.

Approximation (2.5) suggests to define an estimator of $\overline{F}_{u_n}(y)$ as follows:

$$(2.6) \qquad \widehat{\overline{F}}_{u_n}(y) = \overline{\mathbb{G}}_{\widehat{\xi}_n,\widehat{\beta}_n}(y),$$

for appropriate estimates $\widehat{\xi}_n$ and $\widehat{\beta}_n$ of ξ and β, respectively. Note that β will be estimated separately, i.e. $\beta = \xi u_n$ will not be used. The reason for this is to achieve greater flexibility in the parameter fitting, compensating for the underlying distribution not being an exact GPD. Theorem 3.2 in Smith (1987) [27] gives us the asymptotic distribution of the tail parameters $\left(\widehat{\xi}_n, \widehat{\beta}_n\right)$ as follows

$$\sqrt{np_n}\begin{pmatrix} \widehat{\beta}_n/\beta - 1 \\ \widehat{\xi}_n - \xi \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}_2\left(0, \mathbf{\Sigma}^{-1}\right) \text{ when } n \to \infty, \tag{2.7}$$

provided that $\sqrt{np_n}u_n^{-\delta}\mathbb{L}(u_n) \to 0$ when $n \to \infty$ and the function $x \mapsto x^{-\delta}\mathbb{L}(x)$ is non-increasing for all sufficiently large x, where

$$\mathbf{\Sigma}^{-1} = (1+\xi)\begin{pmatrix} 2 & -1 \\ -1 & 1+\xi \end{pmatrix}. \tag{2.8}$$

We note that when $\sqrt{np_n}u_n^{-\delta}\mathbb{L}(u_n) \nrightarrow 0$, then the limiting distribution in (2.7) is biased.

Next we define an estimator of $\overline{F}(u_n)$. For this, let $N \equiv N_n(u_n)$ be defined by

$$N = \#\left\{X_i : \ X_i > u_n : 1 \le i \le n\ \right\},$$

which is the number of those X_i's that exceed u_n. Since N follows the binomial distribution $\mathcal{B}(p_n, n)$ with the parameter $p_n = \mathbf{P}[X_1 > u_n]$, which is equal to $\overline{F}(u_n)$, we have a natural estimator of $\overline{F}(u_n)$ defined by

$$\widehat{p}_n = \frac{N}{n}.$$

From the definition of $\overline{F}_{u_n}(y)$ we have $\overline{F}(u_n+y) = \overline{F}(u_n)\overline{F}_{u_n}(y)$. Hence, with the above defined estimators for $\overline{F}_{u_n}(y)$ and $\overline{F}(u_n)$, we have the following estimator of $\overline{F}(u_n+y)$:

$$\begin{aligned} \widehat{\overline{F}}(u_n+y) &= \widehat{\overline{F}}(u_n)\widehat{\overline{F}}_{u_n}(y) \\ &= \widehat{p}_n\overline{\mathbb{G}}_{\widehat{\xi}_n,\widehat{\beta}_n}(y). \end{aligned} \tag{2.9}$$

We shall use $\widehat{\overline{F}}(u_n+y)$ to construct an estimator for the distortion risk measure $H\left[F,g\right]$ and then show in a simulation study that in this way constructed empirical distortion risk measure outperforms the one constructed using Fisher-Tippett-Gnedenko type extreme-value methods.

3. The new estimator and the main result

We start constructing a POT-based estimator of $H\left[F,g\right]$ using the following lemma.

3.1. Lemma. *Assume that F and g satisfying (2.2) and (2.1) respectively, and u_n be some large level. Then, when $n \to \infty$, we have that*

$$H_n[F,g] = \int_0^{u_n} g\left(\overline{F}(x)\right)dx + (p_n)^r\,\frac{\beta}{r-\xi} + r_n \tag{3.1}$$

with the remainder term

$$r_n = O\big(u_n^{1-r/\xi-\delta}\big),$$

which converges to 0 *when $n \to \infty$ because $1 - r/\xi - \delta < 0$.*

The proof of the lemma 3.1 is relegated to Section 5. With p_n, β and ξ on the right-hand side of equation (3.1) replaced by their estimators, we obtain an estimator of $H\left[F,g\right]$, defined as follows:

$$\widehat{H}_n\left[F,g\right] = \int_0^{u_n} g\left(\overline{F}_n(x)\right)dx + \left(\widehat{p}_n\right)^r\frac{\widehat{\beta}_n}{r-\widehat{\xi}_n}. \tag{3.2}$$

The asymptotic normality of $\widehat{H}_n[F,g]$ is established in the following theorem.

3.2. Theorem. *Let F be a distribution function fulfilling (2.2) with $\xi \in (0.5,1)$ and the distortion function g is differentiable and regularly varying at infinity with index $0 \le r \le 1$. Suppose that $\mathbb{L}$ is locally bounded in $[x_0,+\infty)$ for $x_0 \ge 0$ and $x \to x^{-\delta}\mathbb{L}(x)$ is non-increasing near infinity, for some $\delta > 0$. For any $u_n = O(n^{\alpha\xi})$ with $\alpha \in (0,1)$, we have*

$$\frac{\sqrt{n}}{\gamma_n \sigma_n}\left(\widehat{H}_n[F,g] - H[F,g]\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)\,, \text{ as } n \to \infty,$$

where

$$\sigma_n^2 := 1 + \frac{\theta_1^2}{\gamma_n^2} p_n(1-p_n) + \frac{2(1+\xi)\theta_2^2}{p_n\gamma_n^2} + \frac{(1+\xi)^2\theta_3^2}{p_n\gamma_n^2} - \frac{(1+\xi)\theta_2\theta_3}{p_n\gamma_n^2}.$$

and

$$\gamma_n^2 = \mathbf{Var}\left[\int_0^{u_n} g'\left(\overline{F}(x)\right)\mathbf{1}(X \le x)\,dx\right],$$

with

$$\theta_1 = \frac{\beta g'(p_n)}{r-\xi}, \theta_2 = \frac{\beta g(p_n)}{r-\xi}, \theta_3 = \frac{\beta g(p_n)}{(r-\xi)^2},$$

and $\beta = u_n\xi$.

4. Simulation Study

To illustrate the result of the Theorem 3.2 , we carry out a simulation study (by means of the statistical software **R**, see Ihaka and Gentleman, 1996) [14], in this study we are interesting by a popular risks measure named Proportional Hazard Premium (PHP) where the distortion function is given by $g(x) = x^{1/\rho}$ with $\rho > 1$, to illustrate the performance of our estimation and its comparison with the parametric estimator, through its application to sets of samples taken from two distinct Pareto distributions $\overline{F}(x) = x^{-1/\xi}, x \ge 1$ (with tail index $\xi = 2/3$ and $\xi = 3/4$), we are interesting by the PHP risk measure, that is, the distortion function is given by $g(x) = x^{1/\rho}$ with the distortion paramater $\rho > 1$, in this case the esimator of the PHP is given by

$$\widehat{H}_{\rho,n} = \int_0^{u_n}\left(\frac{1}{n}\sum_{j=1}^{n}\mathbf{1}_{(X_j \ge x)}\right)^{1/\rho} dx + (\widehat{p}_n)^{1/\rho}\,\frac{\rho\widehat{\beta}_n}{1-\rho\widehat{\xi}_n}.$$

In the first part, we evaluate the root mean squared error (rmse), the accuracy of the confidence intervals via and their lengths (length) and the coverage probabilities (cprob), the confidence level $1-\zeta$ is fixed at 0.95, we generate 200 independent replicates of sizes 500, 1000 and 2000 from the selected parent distribution for $\xi = 2/3$. For each simulated sample, we obtain an estimate of the estimators premium H_ρ for two distinct aversion index values $\rho = 1.1$ and $\rho = 1.2$. In each case we compute, by averaging over all samples, the confidence bounds and the coverage probability and length of the corresponding confidence interval. Note that lcb and ucb stand respectively for lower confidence bound and upper confidence bound.
To this end. We summarize the results in Table 1 for $\xi = 2/3$, $\rho = 1.1$, and Table 2 for $\xi = 2/3$,$\rho = 1.2$.

In this second part, we generate 200 independent replicate of size 1000 from the selected parent distribution $\overline{F}(x) = x^{-1/\xi}, x \ge 1$ (with tail index $\xi = 2/3$ and $\xi = 3/4$) and estimate the PHP for two distinct aversion index values $\rho = 1.1$ and $\rho = 1.2$. We interesting by the comparison of our estimator $\widehat{H}_{\rho,n}$ with the old estimator constructed

Table 1. Point estimates and 95%-confidence intervals for H,based on200 samples of Pareto-distributed rv's with tail index $\xi = 2/3$ and $\rho = 1.1$.

$\rho = 1.1$			$H = 3.75$			
n	$\widehat{H}_{\rho,n}$	rmse	lcb	ucb	cprob	length
500	3.312	0.561	2.23	4.39	0.54	2.168
1000	4.037	0.286	3.139	4.934	0.71	1.793
2000	3.765	0.050	3.189	4.342	0.82	1.153

Table 2. Point estimates and 95%-confidence intervals for H,based on 200 samples of Pareto-distributed rv's with tail index $\xi = 2/3$ and $\rho = 1.2$.

$\rho = 1.2$			$H = 5$			
n	$\widehat{H}_{\rho,n}$	rmse	lcb	ucb	cprob	length
500	5.194	0.835	2.852	7.537	0.640	4.683
1000	5.069	0.355	3.444	6.696	0.815	3.252
2000	5.028	0.311	3.617	6.439	0.890	2.822

by the extreme values methods by (Necir and Meraghni 2009 [19]) and noted $\widetilde{H}_{\rho,n}$, this comparaison is in terms the **bias** and the mean squared error **(MSE).** We summarize the results in Table (3)

Table 3. Analog between the new estimator and the old estimator of the premium hazard proportional for two tail index and two risk aversions index

ξ	2/3		3/4	
ρ	1.1	1.2	1.1	1.2
H_ρ	3.75	5	5.714	10
$\widehat{H}_{\rho,n}$	3.752	5.071	5.815	10.036
bias	0.002	0.071	0.101	0.037
MSE	0.0998	0.256	0.340	1.796
$\widetilde{H}_{\rho,n}$	4.042	5.280	6.050	8.718
bias	0.292	0.280	0.336	-1.283
MSE	0.116	0.299	0.457	2.048

From these results, we observe that the new estimator has smaller bias and mean squared error than the old estimator in most cases, the new estimator performs worse, which may be explained by the Theorem 3.2.

5. Proofs

The following propositions are instrumental for the proof of Theorem 3.2.

5.1. Proposition. *Let F be a distribution function fulfilling (2.2) with $\xi \in (0,1)$, $\delta > 0, r \in [0,1]$ and some real c. Suppose that $\mathbb{L}$ is locally bounded in $[x_0, +\infty)$ for $x_0 \geq 0$.*

Then for n large enough, for any $u_n = O\left(n^{\alpha\xi}\right)$, $\alpha \in (0,1)$, we have that

$$p_n = cn^{-\alpha}(1+o(1)), \tag{5.1}$$

$$\gamma_n^2 = O\left(n^{2\alpha(\xi-r+1)}\right), \tag{5.2}$$

and

$$\sqrt{np_n}u_n^{-\delta}\mathbb{L}\left(u_n\right) = O\left(n^{-\alpha/2-\alpha\xi\delta+1/2}\right). \tag{5.3}$$

Proof of the proposition 5.1. We will now prove the result (5.1), let $\overline{F}\left(x\right) = cx^{-1/\xi}\left(1+x^{-\delta}\mathbb{L}\left(x\right)\right)$. Then for n large enough, we have

$$\begin{aligned} p_n &= \mathbf{P}\left(X > u_n\right) = \overline{F}\left(u_n\right) \\ &= cu_n^{-1/\xi}\left(1+u_n^{-\delta}\mathbb{L}\left(u_n\right)\right), \end{aligned}$$

with $u_n = O\left(n^{\alpha\xi}\right)$, then we obtain the statement (5.1). The result (5.3) are straightforward from the result (5.1). We shall next prove statement (5.2). Note that the quantity γ_n^2 defined the formulation of the theorem is equal to $\mathbf{Var}[Z]$, where

$$Z = \int_0^{u_n} g'\left(\overline{F}(x)\right)\mathbf{1}\left(X \le x\right)dx.$$

Since $\overline{F}(x) = x^{-1/\xi}O(1)$, $g(x) = x^r O(1)$ and $u_n = n^{\alpha\xi}O(1)$, we have that

$$\begin{aligned} \mathbf{E}\left[Z\right] &= \int_0^{u_n} g'\left(\overline{F}(x)\right)F(x)dx \\ &= \int_0^{u_n} g'\left(\overline{F}(x)\right)dx - \int_0^{u_n} g'\left(\overline{F}(x)\right)\overline{F}(x)dx \\ &= n^{\alpha(1+\xi-r)}O(1). \end{aligned}$$

Furthermore,

$$\begin{aligned} \mathbf{E}\left[Z^2\right] &= \int_0^{u_n}\int_0^{u_n} g'\left(\overline{F}(x)\right)g'\left(\overline{F}\left(y\right)\right)\min\left(F(x),F\left(y\right)\right)dxdy \\ &= \int_0^{u_n} g'\left(\overline{F}(x)\right)\left(\int_0^{x} g'\left(\overline{F}\left(y\right)\right)F\left(y\right)dy\right)dx \\ &\quad + \int_0^{u_n} g'\left(\overline{F}(x)\right)\left(\int_x^{u_n} g'\left(\overline{F}\left(y\right)\right)F\left(y\right)dy\right)dx \\ &= 2n^{2\alpha(\xi-r+1)}O(1). \end{aligned}$$

Consequently, statement (5.2) holds. □

Proof of Lemma 3.1. We start with the elementary equation

$$H_{g,n} = \int_0^{u_n} g\left(\overline{F}(x)\right)dx + \int_{u_n}^{\infty} g\left(\overline{F}(x)\right)dx.$$

Hence, the remainder term r_n noted in the formulation of the Lemma 3.1 is

$$r_n = \int_{u_n}^{\infty} g\left(\overline{F}(x)\right)dx - \left(p_n\right)^r\frac{\beta}{r-\xi}.$$

Next we express the integral in the definition of r_n as follows:

$$\begin{aligned} \int_{u_n}^{\infty} g\left(\overline{F}(x)\right)dx &= \int_0^{\infty} g\left(\overline{F}\left(s+u_n\right)\right)ds \\ &= \int_0^{\infty} g\left(p_n\overline{F}_{u_n}(s)\right)ds. \end{aligned}$$

Since $\overline{F}_{u_n}(s) = \overline{F}\,(u_n+s)\,/\overline{F}\,(u_n)$, we have that

$$\overline{F}_{u_n}(s) = \left(1+\frac{\xi}{\beta}s\right)^{-1/\xi} \frac{1+(u_n+s)^{-\delta}\mathbb{L}\,(u_n+s)}{1+u_n^{-\delta}\mathbb{L}\,(u_n)}.$$

Consequently,

$$\int_{u_n}^{\infty} g\left(\overline{F}(x)\right) dx = (p_n)^r \,\frac{\beta}{r-\xi}\left(\frac{1+(u_n+s)^{-\delta}\mathbb{L}\,(u_n+s)}{1+u_n^{-\delta}\mathbb{L}\,(u_n)}\right)^r.$$

Since function $\mathbb{L}$ is locally bounded in $[x_0,\infty)$ for $x_0 \geq 0$ and $x^{-\delta}\mathbb{L}(x)$ is non-increasing near infinity, then for all large n, we have that

$$u_n^{r/\xi}\int_{u_n}^{\infty} x^{-r/\xi-\delta}\mathbb{L}(x)dx = O(u_n^{-\delta}).$$

Consequently, for all large n,

$$\int_{u_n}^{\infty} g\left(\overline{F}(x)\right) dx = \frac{\beta}{\xi}\int_{1}^{\infty} g\left(p_n\,(z)^{-1/\xi}\right) dz\left(1-u_n^{-\delta}\mathbb{L}\,(u_n)+O(u_n^{-\delta}\mathbb{L}\,(u_n)\right).$$

This implies that $r_n = O(u_n^{1-r/\xi-\delta})$ and concludes the proof of Lemma 3.1. □

Proof of Theorem 3.2. We write

$$\sqrt{n}\big(\widehat{H}_{g,n}-H_g\big) = A_n + B_n,$$

where

$$A_n = \sqrt{n}\int_0^{u_n}\left(g\left(\overline{F}_n\,(x)\right) - g\left(\overline{F}(x)\right)\right)dx$$

and

$$B_n = \sqrt{n}\left((\widehat{p}_n)^r\,\frac{\widehat{\beta}_n}{r-\widehat{\xi}_n} - \int_{u_n}^{\infty} g\left(\overline{F}(x)\right)dx\right).$$

Using Lemma 3.1 and the fact that $\sqrt{n}\;u_n^{1-r/\xi-\delta} \to 0$, as $n\to\infty$, we have that

$$\begin{aligned}
B_n &= \sqrt{n}\left((\widehat{p}_n)^r\,\frac{\widehat{\beta}_n}{r-\widehat{\xi}_n} - \int_{u_n}^{\infty} g\left(\overline{F}(x)\right)dx\right)\\
&= \sqrt{n}\left(B_{n,1} + O\big(u_n^{1-r/\xi-\delta}\big)\right)\\
&= \sqrt{n}B_{n,1} + o(1),
\end{aligned}$$

where

$$\begin{aligned}
B_{n,1} &= (\widehat{p}_n)^r\,\frac{\widehat{\beta}_n}{r-\widehat{\xi}_n} - (p_n)^r\,\frac{\beta}{r-\xi}\\
&= \frac{\widehat{\beta}_n}{r-\widehat{\xi}_n}\,(\widehat{p}_n^r - p_n^r) + \frac{(p_n)^r\,\beta}{(r-\widehat{\xi}_n)}(\widehat{\beta}_n/\beta - 1) + \frac{(p_n)^r}{(r-\widehat{\xi}_n)(r-\xi)}(\widehat{\xi}_n - \xi).
\end{aligned}$$

By Smith (1987) [27], we have that

$$\widehat{\beta}_n/\beta - 1 = O_{\mathbf{P}}\left(u_n^{-\delta}\mathbb{L}\,(u_n)\right) \tag{5.4}$$

and

$$\widehat{\xi}_n - \xi = O_{\mathbf{P}}\left(u_n^{-\delta}\mathbb{L}\,(u_n)\right). \tag{5.5}$$

Furthermore, by the CLT, we have that

$$\widehat{p}_n - p_n = O_{\mathbf{P}}(\sqrt{p_n/n}). \tag{5.6}$$

Consequently, we have that

$$B_{n,1} = \theta_1 \left(1 + o_{\mathbf{P}}(1)\right) \sqrt{n}(\widehat{p}_n - p_n) + \theta_2 \left(1 + o_{\mathbf{P}}(1)\right) \sqrt{n}(\widehat{\beta}_n/\beta - 1) \\ + \theta_3 \left(1 + o_{\mathbf{P}}(1)\right) \sqrt{n}(\widehat{\xi}_n - \xi),$$

where

$$\theta_1 = \frac{\widehat{\beta}_n}{r - \widehat{\xi}_n}, \quad \theta_2 = \frac{(p_n)^r \beta}{(r - \widehat{\xi}_n)}, \quad \theta_3 = \frac{(p_n)^r}{(r - \xi)^2}.$$

We now examine A_n, and start with the equations

$$A_n = \frac{\sqrt{n}}{\gamma_n} \int_0^{u_n} \left(g\left(\overline{F}_n(x)\right) - g\left(\overline{F}(x)\right)\right) dx$$

$$= \frac{\sqrt{n}}{\gamma_n} \int_0^{u_n} \left(\overline{F}_n(x) - \overline{F}(x)\right) g'\left(\overline{F}(x)\right) dx + o_{\mathbf{P}}(1). \tag{5.7}$$

Continuing with (5.7), we have that

$$\begin{aligned} A_n &= -\frac{\sqrt{n}}{\gamma_n} \int_0^{u_n} \left(F_n(x) - F(x)\right) g'\left(\overline{F}(x)\right) dx + o_{\mathbf{P}}(1) \\ &= -\frac{\sqrt{n}}{\gamma_n} \int_0^{u_n} \left(\frac{1}{n} \sum \mathbf{1}\left(X_i \le x\right) - F(x)\right) g'\left(\overline{F}\left(x\right)\right) dx + o_{\mathbf{P}}(1) \\ &= -\frac{\sqrt{n}}{\gamma_n} \left(\frac{1}{n} \sum \int_0^{u_n} \mathbf{1}\left(X_i \le x\right) g'\left(\overline{F}(x)\right) dx - \int_0^{u_n} F(x) g'\left(\overline{F}(x)\right) dx\right) + o_{\mathbf{P}}(1) \\ &= -\frac{\sqrt{n}}{\gamma_n} \left(\overline{Z} - \mathbf{E}\left[Z_1\right]\right) + o_{\mathbf{P}}(1), \end{aligned}$$

where $\overline{Z}$ is the arithmetic average of the n random variables

$$Z_i := \int_0^{u_n} g'\left(\overline{F}(x)\right) \mathbf{1}\left(X_i \le x\right) dx.$$

Note that the quantity γ_n^2 defined in the formulation of the Theorem 3.2 is equal to $\mathbf{Var}[Z_1]$.

Next, we shall show that

$$\frac{\sqrt{n}}{\gamma_n} \left(\overline{Z} - \mathbf{E}\left[Z_1\right]\right) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

when $n \to \infty$. We shall next employ the Lindeberg-Feller Theorem. For this, we write:

$$\begin{aligned} \frac{\sqrt{n}}{\gamma_n} \left(\overline{Z} - \mathbf{E}\left[Z_1\right]\right) &= \sum_{k=1}^{n} \frac{\int_0^{u_n} g'\left(\overline{F}(x)\right) \mathbf{1}\left(X_k \le x\right) dx - \mathbf{E}\left[Z_1\right]}{\gamma_n \sqrt{n}} \\ &\equiv \sum_{k=1}^{n} \xi_{k,n}, \end{aligned}$$

where $\mathbf{E}\left(\xi_{k,n}\right) = 0$, $\mathbf{E}\left(\xi_{k,n}^2\right) = 1/n$, and $\sum_{k=1}^{n} \mathbf{E}\left(\xi_{k,n}^2\right) = 1$ for all $n \ge 1$. Furthermore, for all $\alpha \in (0,1), \xi \in (0,1)$ and $\epsilon > 0$, where $u_n = O\left(n^{\alpha\xi}\right)$ was used. This means that

$$\begin{aligned} \sum_{k=1}^{n} \mathbf{E}\left[\left|\xi_{k,n}\right|^2 1 \left|\xi_{k,n}\right| > \epsilon\right] &= \frac{1}{\gamma_n^2} \mathbf{E}\left[\left[Z_k - \mathbf{E}\left[Z_1\right]\right]^2 1 \left|Z_k - \mathbf{E}\left[Z_1\right]\right| > \epsilon \gamma_n \sqrt{n}\right] \\ &\le \frac{u_n^2}{\gamma_n^2} \mathbf{P}\left[\left|Z_k - \mathbf{E}\left[Z_1\right]\right| > \epsilon \gamma_n \sqrt{n}\right] \\ &\le \frac{u_n^2}{\gamma_n^4 \epsilon^2 n}. \end{aligned}$$

We have, from (5.2) with $u_n = n^{\alpha\xi}$, that

$$\frac{u_n^2}{\gamma_n^4 n} = n^{\alpha(4(r-1)-2\xi)-1} O(1)$$

As $\alpha\left(4\left(r-1\right)-2\xi\right)-1<0$, we conclut that

$$\sum_{k=1}^{n} \mathbf{E}\left[\left|\xi_{k,n}\right|^2; \left|\xi_{k,n}\right| > \epsilon\right] \to 0 \text{ as } n \to \infty.$$

Finally, we obtain that

$$\frac{\sqrt{n}}{\gamma_n}\left(\widehat{H}_{\rho,n} - H_\rho\right) \to -\frac{\sqrt{n}}{\gamma_n}\left(\overline{Z} - \mathbf{E}\left[Z_1\right]\right) + \theta_1 \frac{\sqrt{p_n\left(1-p_n\right)}}{\gamma_n} \frac{\sqrt{n}}{\sqrt{p_n\left(1-p_n\right)}}\left(\widehat{p}_n - p_n\right)$$
$$+ \frac{\theta_2}{\sqrt{p_n}\gamma_n}\sqrt{np_n}\left(\widehat{\beta}_n/\beta - 1\right) + \frac{\theta_3}{\sqrt{p_n}\gamma_n}\sqrt{np_n}\left(\widehat{\xi}_n - \xi\right) + o_{\mathbf{P}}(1),$$

From **Lemma A-2** of Johansson 2003 [15], under the assumptions of Theorem 3.2, for any real numbers, t_1, t_2, t_3 and t_4, we have

$$\mathbf{E}\left[\exp\left\{it_1\frac{\sqrt{n}}{\gamma_n}\left(\overline{Z} - \mathbf{E}\left[Z_1\right]\right) + i\sqrt{np_n}\left(t_2, t_3\right)\left(\begin{array}{c}\widehat{\beta}_n/\beta - 1\\ \widehat{\xi}_n - \xi\end{array}\right) + it_4\frac{\sqrt{n}\left(\widehat{p}_n - p_n\right)}{\sqrt{p_n\left(1-p_n\right)}}\right\}\right]$$
$$\to \exp\left\{-\frac{t_1^2}{2} - \frac{1}{2}\left(t_2, t_3\right)\mathbf{\Sigma}^{-1}\left(\begin{array}{c}t_2\\ t_3\end{array}\right) - \frac{t_4^2}{2}\right\}\left(1 + o_{\mathbf{P}}\left(1\right)\right).$$

as $n \to \infty$, where $\mathbf{\Sigma}^{-1}$ is that in (2.8), $\gamma_n^2 = Var(Z_1)$ and $i^2 = -1$. It follows that, with this result that

$$\frac{\sqrt{n}}{\gamma_n\sigma_n}\left(\widehat{H}_{\rho,n} - H_\rho\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,1\right), \text{ as } n \to \infty,$$

where

$$\sigma_n^2 = 1 + \frac{\theta_1^2}{\gamma_n^2}p_n\left(1-p_n\right) + \frac{2\left(1+\xi\right)\theta_2^2}{p_n\gamma_n^2} + \frac{\left(1+\xi\right)^2\theta_3^2}{p_n\gamma_n^2} - \frac{\left(1+\xi\right)\theta_2\theta_3}{p_n\gamma_n^2}.$$

This complet the proof of Theorem 3.2. □

Acknowledgement

The author are very grateful to the referees for a careful reading of the paper that led to significant improvements of the earlier draft.

References

[1] Artzner, P. *Application of coherent capital requirements*, North American Actuarial Journal **3**(2), 11-25, 1999.

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

[3] Balkema, A. and de Haan L. *Residual Lifetime at Great Age*, Annals of Probability **2**, 792-804, 1974.

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

[5] Bingham, N., Goldie, C. and Teugels, J. *Regular Variation*, Encyclopedia of Mathematics and its Applications **27**, Cambridge University Press, 1987.

[6] Brazauskas, V., Jones, B. L., Puri, M. L. and Zitikis, R. *Estimating conditional tail expectation with actuarial applications in view*, J. Statist. Plann. Inference **138**, 3590-3604, 2008.

[7] Cebrian, A. C., Denuit, M. and Lambert, P. *Generalized Pareto Fitto the Society of Actuaries Large Claims Database*, N. Am. Actuar. J. **7**, 18-36, 2003.

[8] Centeno, M. L. and Andrade e Silva, J. *Applying the proportional hazard premium calculation principle*, Astin Bulletin **35**, 00-00, 2005.
[9] Dhaene, J., Kukush, A., Linders, D. and Tang, Q. *Remarks on quantiles and distortion risk measures*, European Actuarial Journal **2**(2), 319-328, 2012.
[10] Furman, E. and Zitikis, R. *Weighted risk capital allocations*, Insurance: Mathematics and Economics **43**, 263-269, 2008.
[11] Greselin, R., Puri, ML. and Zitikis, R. *L-functions, processes, and statistics in measuring economic inequality and actuarial risks*, Statistics and Its Interface **2** (2), 227-245, 2009.
[12] Hürlimann W. *Inequalities for Look Back Option Strategies and Exchange Risk Modelling*. Paper presented at the First Euro-Japanese Workshop on Stochastic Modelling in Insurance, Finance, Production and Reliability, Brussels, 1998.
[13] Hua, L. and Joe, H. *Tail comonotonicity: Properties, constructions, and asymptotic additivity of risk measures*, Insurance: Mathematics and Economics **51** (2), 492-503, 2012.
[14] Ihaka, R. and Gentleman, R. *R: A language for data analysis and graphics*, J. Comput. Graph. Statist. **5**, 299-314, 1996.
[15] Johansson, J. *Estimating the mean of heavy-tailed distributions*, *Extremes* 6, 91-131, 2003.
[16] Jones, B. L. and Zitikis, R. *Empirical estimation of risk measures and related quantities*, North American Actuarial Journal, **7** (4), 44-54, 2003.
[17] Jones, B. L. and Zitikis, R. *Risk measures, distortion parameters, and their empirical estimation*, Insurance: Mathematics and Economics **41** (2), 279-297, 2007.
[18] Joseph H. T. Kim. *Bias correction for estimated distortion risk measure using the bootstrap*, Insurance: Mathematics and Economics **47** (2), 198-205, 2010.
[19] Necir, A. and Meraghni, D. *Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts*, Insurance: Mathematics and Economics **45**, 49-58, 2009.
[20] Necir, A., Rassoul, A. and Zitikis, R. *Estimating the conditional tail expectation in the case of heavy-tailed losses*, Journal of Probability and Statistics, 2010.
[21] Peng, L., Qib, Y., Wang, R. and Yang, J. *Jackknife empirical likelihood method for some risk measures and related quantities*, Insurance: Mathematics & Economics, **51**, 142-150, 2012.
[22] Pickands, J. *Statistical Inference Using Extreme Order Statistics*, Annals of Statistics **3**, 119-131, 1975.
[23] Reiss, R.D. and Thomas, M. *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, Birkhauser, Basel, 2007.
[24] Resnick, S. *Extreme Values, Regular Variation and Point Processes*, Springer-Verlag, 1987.
[25] Rolski, T., Schimidli, H., Schmidt, V. and Teugels, J. L. *StochasticProcesses for Insurance and Finance*, John Wiley & Sons, Chichester, 1999.
[26] Seneta, E. *Regularly Varying Functions*, Lecture Notes in Mathematics, Springer-Verlag, **508**, 1976.
[27] Smith, R. *Estimating Tails of Probability Distributions*, The Annals of Statistics, **15**, 1174-1207, 1987.
[28] Wang, S. S. *Premium Calculation by Transforming the LayerPremium Density*, ASTIN Bulletin **26**, 71-92, 1996.
[29] Wang, S. and Young, V.R. *Ordering risks. Expected utility theory versus Yaari' dual theory of risk*, Insurance: Mathematics & Economics. **22**, 145-161, 1998.
[30] Wirch, J.L. and Hardy, M.R. *A synthesis of risk measures for capital adequacy*, Insurance: Math. Econom. **25**, 337-347, 1999.
[31] Yaari, M. E. *The Dual Theory of Choice Under Risk*, Econometrica **55**, 95-115, 1987.