STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, October 2025, pp 1611–1624.
Published online in International Academic Press (www.IAPress.org)

A New Robust Estimation and Hypothesis Testing for Reinsurance Premiums in Big Data Settings

Touil Salah [1], Rassoul Abdelaziz [2], Ould Rouis Hamid [3], Frihi Redouane [3,*]

[1] *LAMDA-RO Laboratory, Department of Mathematics, Hassiba Ben Bouali, Chlef, Algeria*
[2] *GEE Laboratory, National Higher School for Hydraulics, Blida, Algeria*
[3] *LAMDA-RO Laboratory, Department of Mathematics, Blida 1 University, Algeria*

Abstract This research study presents a novel methodology to estimate reassurance premiums in the setting of large datasets, employing the principle of grouping. We present a median-of-means nonparametric estimator that addresses the difficulties posed by huge datasets. We analyze this estimator's consistency and asymptotic normality under specific criteria about the growth rate of subgroups.
Furthermore, we introduce a novel approach to the empirical likelihood method for the median to evaluate excess-of-loss reinsurance. Our proposed method eliminates the need to estimate the estimator's variance structure in advance, which can be difficult and prone to inaccuracies. Numerical simulation analysis is implemented to evaluate the efficacy of our proposed estimator. The results indicate that our estimator is highly resilient in the presence of outliers.

Keywords excess-of-loss reinsurance, median-of-means, big data, empirical likelihood, hypothesis test

AMS 2010 subject classifications 62G35 , 62P05

DOI: 10.19139/soic-2310-5070-968

1. Introduction

An essential problem in actuarial science revolves around pricing insurance risks. The pricing of an insurance risk must accurately capture the level of risk implied by the underlying distribution of the loss random variable. Factors such as the variability of the loss variable, the shape of the distribution, and especially the tail behavior play a significant role in determining the appropriate price. Higher variability and a heavier right-tail distribution require a higher price.

The approach used to determine the price of an insurance risk leads to the concept of a risk measure. A risk measure is a mapping from the set of all loss random variables to the non-negative real numbers that quantify the risk associated with an insurance contract. Artzner (1999) [3] examines the properties of a risk measure must possess to be coherent, building on their previous collaborative work.

Various premium principles have been developed to assess the risk premium accurately. Within the insurance literature, numerous premium calculation principles have been proposed, such as mean, value at risk, variance, and others. In this study, we focus on the Wang premium calculation principle introduced by Wang (1996) [17]. This principle relies on a proportional transformation of the hazard function.

Wang's premium calculation principle satisfies all the desired properties of a premium principle, including sub-additivity and layer additivity. These properties align with the adjusted distribution methods advocated by Venter for

*Correspondence to: Rassoul Abdelaziz (Email: a.rassoul@ensh.dz). National Higher School for Hydraulics, 29 Way of Soumaa, Blida, Algeria.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

no-arbitrage pricing models. In the following, we present the formulation of the PH-premium for regular insurance based on Wang's principle.

The premium for an insured risk X, denoted as the Proportional Hazard (PH) premium, is determined by the continuous distribution function F of the risk and the hazard function $S = 1 - F$. The PH premium is influenced by a parameter $\rho \geq 1$, known as the risk aversion index. In certain actuarial scenarios, such as reinsurance treaties, the focus is estimating a premium for a specified retention level $R > 0$. We denote the reinsurance premium for the high layer $[R, +\infty[$ as $\Pi_{\rho,R}$. This type of problem arises when the insured poses a high level of risk for the insurance company and chooses to transfer a portion of this risk to a reinsurance company. This transfer is necessary because the insurance company may lack sufficient capital to bear the entire risk on its own.

The PH premium is defined as a function of ρ and S by

$$\Pi_{\rho,R} = \int_0^{+\infty} (S(x))^{1/\rho}\, dx \tag{1}$$

Excess-of-loss reinsurance is a prevalent type of reinsurance in which the reinsurer indemnifies only those losses that surpass a designated retention threshold. This category of reinsurance enables the cedent to restrict their risk exposure to a specified threshold. The premium for excess-of-loss reinsurance can be articulated as a result of the additivity property of the proportional hazards premium for losses. For a specified retention level $R > 0$, the risk-adjusted premium for excess-of-loss reinsurance is defined as:

$$\Pi_{\rho,R} = \int_R^{+\infty} (S(x))^{1/\rho}\, dx \tag{2}$$

Now, consider $X_1, X_2, \ldots, X_N$ are independent and identically distributed (i.i.d.) random variables with common distribution function (cdf) F of an insured risk X and $X_{(1)} \leq X_{(2)} \leq .... \leq X_{(N)}$ its order statistics. Further, let $1 \leq k \leq N$ be a sequence of positive integers, such that $k \to \infty$ and $k/N \to 0$ as $N \to \infty$. Integer k represents the number of extremes used in the computation of the tail index estimate. Now, let then the optimal retention level $R_{opt} := F^{\leftarrow}(1 - k/n)$ will be estimated by $R_{opt} := X_{n-k}$, where $X_{n-k,n}$ is the $(k+1)$ largest observation. An empirical estimator of excess-of-loss reinsurance with retention $R = X_{(n-k)}$ is

$$\hat{\Pi}_{\rho,N} = \sum_{i=1}^{k} \left(\frac{i}{N}\right)^{1/\rho} \left(X_{(N-i+1)} - X_{(N-i)}\right). \tag{3}$$

The asymptotic theory for the empirical $\hat{\Pi}_{\rho,N}$ has been known, [13] have developed an asymptotic theory for the excess-of-loss reinsurance estimator, assuming that the underlying i.i.d. random variables $X_1, X_2, ..., X_N$ have finite (r) moments for some $r > 2\rho/(2-\rho)$. Likewise, Centeno et al. (2005) [6] use a bootstrap technique to describe the behavior of a proposed biased estimator for $\Pi_{\rho,0}$ (premium without retention). In this study, we concentrate primarily on the statistical and probabilistic approaches that might be taken to address this matter. The non-parametric confidence intervals that are generated by estimating the asymptotic variance are typically erroneous. This is because the asymptotic variance of $\Pi_{\rho,R}$ is quite complicated. [16] among others, proposed the jackknife empirical likelihood method [12] to improve the inference on $\Pi_{\rho,R}$ which avoids the prior estimation of variance. In this paper, we introduce a simpler method based on the idea of random grouping and the usual empirical likelihood method for the median to study risk measures. Our approach can be classified as one of the so-called divide-conquer methods. More precisely, we divide the data set into several groups, and then obtaining interesting statistics within each group is the first step. In the second step of "conquer", considering robustness, we take the median, instead of mean, of the resulting statistics as our final estimator. It works well, especially in the case of massive data, such as high-frequency data in finance markets. In a world full of big data, we believe that we have developed one effective and robust inference approach to reducing the computational burden arising from an analysis of massive data.

The rest of the paper is organized as follows: In section 2, we present our proposed estimator and its asymptotic properties. Section 3 is devoted to an empirical likelihood approach to testing $\Pi_{\rho,R}$. Section 4 contains some criteria about choices of blocks. Simulations analysis are given in Section 5. The proofs of different results are postponed to section 6.

2. Median-of-means estimate for $\Pi_{\rho,R}$

Initially introduces the median-of-means method [2] to study population means. In this paper, we based this approach to construct a new estimator of the excess-of-loss reinsurance premium $\Pi_{\rho,R}$. To fix the idea, we divide the N observations $X_1, ..., X_N$ into K blocks randomly. Assume that each block contains n data points for simplicity. In block $B_j, j = 1, 2, ..., K$, we note $\tilde{\Pi}^{(j)}_{\rho,n}$ a non-parametric estimator for $\Pi^{(j)}_{\rho,R}$ based on the empirical distribution pertaining to the sample $X_1, \cdots, X_n$ of block j, as follows:

$$\tilde{\Pi}^{(j)}_{\rho,n} := \sum_{i=1}^{k} \left(\frac{i}{n}\right)^{1/\rho} \left(X_{(n-i+1)} - X_{(n-i)}\right). \tag{4}$$

Next, we define the median-of-means estimator of R_v as

$$\tilde{\Pi}^{MoM}_{\rho,N} := Median \left\{\tilde{\Pi}^{(1)}_{\rho,n}, \tilde{\Pi}^{(2)}_{\rho,n}, ..., \tilde{\Pi}^{(K)}_{\rho,n}\right\}. \tag{5}$$

The asymptotic properties of $\tilde{\Pi}^{MoM}_{\rho,N}$ are summarized in the following theorems.

Theorem 1
Assume that $E(|X|^3) < \infty$ and $\sigma^2_\rho(F) > 0$, where,

$$\sigma^2_\rho(F) := \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(F(x \wedge y) - F(x) F(y)\right) a(F(x))\, b(F(y))\, dx\, dy, \tag{6}$$

where $a(\cdot)$ and $b(\cdot)$ are two functions on $[0, 1]$. If F has a strictly positive, continuous density function f, then for any fixed $x > 0$,

$$P\left(\left|\tilde{\Pi}^{MoM}_{\rho,N} - \Pi_{\rho,R}\right| \geq x\right) \leq \frac{C}{(N/K)^{K/5}}, \tag{7}$$

holds for some constant $C := C(x) > 0$ and any positive integer K.

Remark 1
- Note that the constant C at the right hand side of (7) is not uniform in x.
- Theorem 1 directly implies that the convergence of $\tilde{\Pi}^{MoM}_{\rho,N}$ towards to $\Pi_{\rho,R}$ is almost surely by Borel-Cantelli Lemma.

Theorem 2
1. Suppose K is fixed. Let $\Theta_1, \Theta_2, ..., \Theta_K$ be independent and identically distributed standard normal random variables. Then as $N \to \infty$,

$$\frac{\sqrt{N}}{\sigma_\rho(F)} \left(\tilde{\Pi}^{MoM}_{\rho,N} - \Pi_{\rho,R}\right) \xrightarrow{\mathcal{D}} Median\left\{\Theta_1, \Theta_2, ..., \Theta_K\right\}, \tag{8}$$

   where "$\xrightarrow{\mathcal{D}}$" means convergence in distribution.
2. Suppose $N/K^2 \to \infty$ as $K \to \infty$. Then the following asymptotic normality holds,

$$\frac{\sqrt{N}}{\sigma_\rho(F)} \left(\tilde{\Pi}^{MoM}_{\rho,N} - \Pi_{\rho,R}\right) \xrightarrow{\mathcal{D}} \sqrt{\frac{\pi}{2}} \mathcal{N}(0, 1). \tag{9}$$

3. Test by empirical likelihood

In this section, we consider the hypothesis testing problem of whether the excess-of-loss reinsurance premium equals a given value. As commented at the end of the last section, we will not use $\tilde{\Pi}^{MoM}_{\rho,N}$ to construct the test statistic since it involves the unknown $\sigma_\rho(F)$.

Our approach is based on the empirical likelihood (EL) defined by [15].

Since different blocks are disjoint, $\tilde{\Pi}^{(1)}_{\rho,n}, \tilde{\Pi}^{(2)}_{\rho,n}, ..., \tilde{\Pi}^{(K)}_{\rho,n}$ are independent and share the same distribution. So, we can regard them as one sample and apply the EL method.

For each $k = 1, ..., K$, we denote $Z_{n,k} := I\left(\tilde{\Pi}^{(j)}_{\rho,n} \leq \Pi_{\rho,R}\right)$. Obviously, $\mathbb{E}\left(Z_{n,k}\right) = 0.5$ (actually $\mathbb{E}\left(Z_{n,k}\right) = 0.5 = O\left(1/\sqrt{n}\right)$ by Lemma 1 in section 5) and hence the empirical likelihood ratio for $\Pi_{\rho,R}$ is given by

$$\mathcal{R}\left(\Pi_{\rho,R}\right) = \max\left\{\prod_{k=1}^{K} K w_k \mid \sum_{k=1}^{K} w_k Z_{n,k} = 0.5, w_k \geq 0, \sum_{k=1}^{K} w_k = 1\right\}. \tag{10}$$

By the Lagrange multiples method, the maximum point is given by

$$w_k = \frac{1}{K}\frac{1}{1+\lambda\left(Z_{n,k}-0.5\right)}, \tag{11}$$

where $\lambda = \lambda\left(\Pi_{\rho,R}\right)$ satisfies the following equation

$$\frac{1}{K}\sum_{k=1}^{K}\frac{Z_{n,k}}{1+\lambda\left(Z_{n,k}-0.5\right)} = 0. \tag{12}$$

By the use of the same arguments as in [15], we can obtain the following result define by the theorem 3.

***Theorem** 3*
Under the conditions in Theorem 2, we have

$$-2\log\mathcal{R}\left(\Pi_{\rho,R}\right) \xrightarrow{\mathcal{D}} \chi^2_1, \text{ as } K, n \to \infty. \tag{13}$$

Using the result of Theorem 3, the rejection region for the hypothesis with significance level α, with $(0 < \alpha < 1)$

$$H_0 : \Pi_{\rho,R} = \varpi \text{ vs. } H_1 : \Pi_{\rho,R} \neq \varpi \tag{14}$$

can be constructed as

$$\mathcal{R} = \left\{-2\log\mathcal{R}\left(\Pi_{\rho,R}\right) \geq \chi^2_1\left(\beta\right)\right\}, \tag{15}$$

where $\chi^2_1\left(\alpha\right)$ is the upper $\alpha - th$ quantile of χ^2_1.

4. Selection of block

In this section, we discuss the selection of optimal K when applying the EL method in previous sections. From (8), when N is fixed, the small K works better since the median is the middle value for the data set in practice. Furthermore, the median is a robust statistic, which has a breakdown point of 50%. Hence, K should be selected large when the data are contaminated. On the other hand, EL performs not well when K is small. So, we adopt different methods to select K for estimation and inference.

If we are interested in the point estimator (8), We can proceed in the following manner, when the data are not contaminated, we adopt the suggestion by [14], $K = 8\lceil\log(1/\delta)\rceil$ with $(0 < \delta < 1)$, where $\lceil a\rceil$ is the largest integer not greater than a. In practice, δ comes from the uniform distribution on $(0, 1)$. To eliminate the random effect of K, we replicate 500 times and get their mean as our final choice of K. When the data are contaminated, we set $K = \lceil 0.04N\rceil$.

If we are interested in inference (14), we can proceed as follows. We note that the accuracy of the estimator in each block increases as n increases. However, the power of EL increases as K becomes larger. Hence, we propose one information criterion, AAIC, which is analogous to adjusted AIC (AAIC) (Akaike (1973)) [1],

$$AAIC = \frac{1}{K}\sum_{k=1}^{K}\left(\tilde{\Pi}^{(k)}_{\rho,n} - \Pi_{\rho,R}\right)^2 + \frac{m}{K},$$

where $m = \lceil N/K \rceil$. In this paper, the above information criteria are minimized over $K \in [K_{low}, K_{upp}]$. Here we set K_{low} to be 30, which is the usual smallest sample size for EL performing well, and $K_{upp} = \lceil N/S_X \rceil$, where $S_X = S \max\{1, KU_X/3\}$ with KU_X an estimator of the kurtosis. The adjusted factor of 3 is the kurtosis of normal distribution.

Here S is a specific constant, such as 50 or 100. When KU_X is higher, we get larger m to improve the accuracy of each block estimation. This is consistent with the belief that each block should contain more data if the skewness and kurtosis of distribution are bigger. K_{opt} is obtained by minimizing $AAIC$.

5. Simulation study and real data application

5.1. Simulation study

In this section, we investigate the finite sample to show the performance of our proposed methods. The data are drawn from the following two distributions:

1. Gamma distribution:
$$F(x;\alpha,\beta) = \int_0^x \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} dt, x > 0.$$
we set $\alpha = 3, \beta = 1$.
2. Fréchet distribution:
$$F(x;\alpha) = \exp\left\{-x^{-\alpha}\right\}, x > 0.$$
we set $\alpha = 3$.

We replicate 500 Monte Carlo simulations. We present the results of this subsection in three examples.

***Example** 1*
This example is used to estimate $\Pi_{\rho,R}$. We compare the method in Section 2 (MoM, median of $\Pi_{\rho,R}$) given by formula (5), with the traditional method (TM, i.e, the empirical version of $\Pi_{\rho,R}$ defined by formula (3) with using full data) by the Average Square Error (ASE) criterion:

$$ASE = \frac{1}{500} \sum_{j=1}^{500} \left(\tilde{\Pi}_{\rho,n}^{(j)} - \Pi_{\rho,R}\right)^2,$$

where $\tilde{\Pi}_{\rho,n}^{(j)}$ is the estimator of $\Pi_{\rho,R}$ based on the j^{th} sample. To analyze the sensitivity of two methods against outliers, we contaminate each sample by adding $r\%$ of χ^2_{100} observations, where $r \in \{0, 1\}$, and we fix two values of the aversion index ρ, where $\rho \in \{1.1, 1.2\}$. We set $N \in \{600, 1000, ..., 3000\}$ Table 1 and Table 2 presents the results.

We have the following comments.

1. The ASEs of the two methods are small as the sample size increases.
2. When the data are not contaminated, MoM is almost the same as TM. The latter is slightly better. However, when outliers are added, TM does not work since its ASEs are bigger, which implies that TM is very sensitive to outliers. But, MoM has good performance. Hence, our proposed method is better than TM.

***Example** 2*
This example is for statistical inference on $\Pi_{\rho,R}$. We set $K_{low} = 30$, choose sample sizes $N \in \{2000, 3000, ..., 8000\}$ for $S = 50$, and $N \in \{3000, 4000, ..., 8000\}$ for $S = 100$. The nominal significance level is 0.05. We compared our proposed method with the normal approximation method. However, the normal approximation method performs badly even when we use the known $\sigma^2(F)$, not its estimator. Hence, we report only the empirical size and power of our proposed method. Furthermore, we also report ASE and average K (AK)

Table 1. ASE for Gamma distribution in Example 1

ρ	1.1				1.2			
r	0%		1%		0%		1%	
N	TM	MoM	TM	MoM	TM	MoM	TM	MoM
600	0.0124	0.0133	0.1304	0.0074	0.0137	0.0144	0.1415	0.0087
1000	0.0122	0.0130	0.1305	0.0073	0.0136	0.0143	0.1416	0.0085
1400	0.0121	0.0127	0.1303	0.0071	0.0134	0.0142	0.1418	0.0082
1800	0.0121	0.0125	0.1303	0.0069	0.0133	0.0140	0.1416	0.0081
2200	0.0120	0.0125	0.1302	0.0068	0.0131	0.0138	0.1415	0.0080
2600	0.0119	0.0124	0.1302	0.0068	0.0131	0.0137	0.1414	0.0078
3000	0.0119	0.0124	0.1301	0.0067	0.0130	0.0137	0.1414	0.0078

Table 2. ASE for Fréchet distribution in Example 1

ρ	1.1				1.2			
r	0%		1%		0%		1%	
N	TM	MoM	TM	MoM	TM	MoM	TM	MoM
600	0.0154	0.0166	0.2021	0.0079	0.0187	0.0195	0.1751	0.0085
1000	0.0151	0.0162	0.2012	0.0076	0.0183	0.0193	0.1745	0.0083
1400	0.0149	0.0157	0.2007	0.0074	0.0179	0.0192	0.1739	0.0082
1800	0.0148	0.0156	0.2001	0.0072	0.0177	0.0191	0.1734	0.0080
2200	0.0147	0.0156	0.1992	0.0071	0.0174	0.0189	0.1731	0.0079
2600	0.0145	0.0155	0.1991	0.0071	0.0171	0.0189	0.1731	0.0079
3000	0.0145	0.0154	0.1990	0.0070	0.0170	0.0188	0.1728	0.0078

in the empirical size. For the power, we consider $\Pi_{\rho,R} + \theta$ with $\theta \in \{0.1, 0.3, 0.5\}$ as the alternative hypothesis. The simulations are displayed in Tables 3-4. The results with $S = 100$ are better than $S = 50$ since its ASE is slightly smaller. The size of the proposed test is closer to 0.05 as N increases.

Table 3. Empirical size, AK and ASE for Gamma distribution in Example 2

	ρ	1.1			1.2		
S	N	Size	AK	ASE	Size	AK	ASE
50	2000	0.061	33.84	0.0058	0.067	36.71	0.0085
	3000	0.062	34.45	0.0055	0.065	38.52	0.0084
	4000	0.059	36.78	0.0053	0.063	39.41	0.0082
	5000	0.058	39.53	0.0052	0.061	40.28	0.0080
	6000	0.054	41.78	0.0049	0.059	41.84	0.0079
	7000	0.053	43.54	0.0047	0.055	43.74	0.0077
	8000	0.052	46.97	0.0045	0.053	49.75	0.0075
100	3000	0.059	35.78	0.0053	0.061	43.74	0.0082
	4000	0.058	38.47	0.0048	0.059	45.89	0.0081
	5000	0.057	41.47	0.0045	0.057	47.83	0.0079
	6000	0.052	44.27	0.0040	0.054	51.14	0.0077
	7000	0.052	49.19	0.0037	0.051	54.95	0.0075
	8000	0.051	55.87	0.0036	0.049	59.47	0.0073

Table 4. Empirical size, AK and ASE for Fréchet distribution in Example 2

	ρ	1.1			1.2		
S	N	Size	AK	ASE	Size	AK	ASE
50	2000	0.069	35.74	0.0069	0.065	37.84	0.0073
	3000	0.065	38.04	0.0066	0.062	38.86	0.0071
	4000	0.060	40.58	0.0065	0.061	39.05	0.0071
	5000	0.058	43.96	0.0063	0.057	39.89	0.0069
	6000	0.054	45.79	0.0062	0.054	41.15	0.0068
	7000	0.053	48.09	0.0061	0.053	45.52	0.0068
	8000	0.051	52.71	0.0061	0.052	50.41	0.0067
100	3000	0.053	37.58	0.0065	0.060	41.52	0.0070
	4000	0.055	42.49	0.0063	0.059	45.48	0.0068
	5000	0.057	46.38	0.0061	0.055	49.74	0.0068
	6000	0.054	50.49	0.0059	0.053	53.07	0.0067
	7000	0.052	54.46	0.0058	0.050	57.17	0.0065
	8000	0.050	58.19	0.0058	0.048	60.89	0.0064

Table 5. Empirical power for Gamma distribution in Example 2

	ρ	1.1			1.2		
S	N	0.1	0.3	0.5	0.1	0.3	0.5
50	2000	0.357	0.695	1.00	0.415	0.747	1.00
	3000	0.374	0.754	1.00	0.452	0.812	1.00
	4000	0.392	0.825	1.00	0.524	0.893	1.00
	5000	0.412	0.892	1.00	0.597	0.945	1.00
	6000	0.439	0.947	1.00	0.662	0.995	1.00
	7000	0.468	0.989	1.00	0.723	1.00	1.00
	8000	0.507	1.000	1.00	0.804	1.00	1.00
100	3000	0.310	0.857	1.00	0.652	0.818	1.00
	4000	0.331	0.912	1.00	0.729	0.892	1.00
	5000	0.371	0.962	1.00	0.806	0.953	1.00
	6000	0.418	0.999	1.00	0.882	0.985	1.00
	7000	0.451	1.000	1.00	0.926	1.00	1.00
	8000	0.487	1.000	1.00	0.973	1.00	1.00

***Example** 3*
In this example, we consider the testing problem for $\Pi_{\rho,R}$. We set the sample size as $\left\{10^4, 10^5, 10^6\right\}$. For convenience of calculations, we fix $K \in \{30, 60\}$. We report ASE and empirical size. The other settings are the same as those in Example 2. From Table 7, the results with $K = 30$ are better than $K = 60$ Since smaller K produces a large sample size of each block, the block estimator is more accurate. On the other hand, the block size $K = 30$ is enough to make EL perform satisfactorily.

5.2. Application to Norwegian fire insurance dataset

We conclude this section with a brief illustration of the estimation procedure using a fire insurance dataset analyzed by [4]. The dataset contains the sizes of $9,181$ fire insurance claims from a Norwegian insurance company, spanning the years 1972 to 1992. These claim amounts have been adjusted for inflation using the Norwegian Consumer Price Index (CPI) and are expressed in thousands of Norwegian kroner (NKR). The dataset is available, for example, in the R package CASdatasets, which can be downloaded from the following link:

Table 6. Empirical power for Fréchet distribution in Example 2

	ρ	1.1			1.2		
S	N	0.1	0.3	0.5	0.1	0.3	0.5
50	2000	0.342	0.675	1.00	0.474	0.847	1.00
	3000	0.378	0.715	1.00	0.541	0.907	1.00
	4000	0.401	0.783	1.00	0.592	0.962	1.00
	5000	0.424	0.852	1.00	0.638	1.00	1.00
	6000	0.448	0.925	1.00	0.749	1.00	1.00
	7000	0.473	0.988	1.00	0.793	1.00	1.00
	8000	0.517	1.000	1.00	0.958	1.00	1.00
100	3000	0.321	0.848	1.00	0.682	0.878	1.00
	4000	0.351	0.897	1.00	0.761	0.947	1.00
	5000	0.384	0.951	1.00	0.842	0.989	1.00
	6000	0.421	0.988	1.00	0.874	1.00	1.00
	7000	0.457	1.000	1.00	0.932	1.00	1.00
	8000	0.491	1.000	1.00	0.963	1.00	1.00

Table 7. Empirical size for Gamma and Fréchet distribution in Example 3

			K=30		K=60	
Distribution	ρ	N	ASE	Size	ASE	Size
Gamma	1.1	10^4	0.0068	0.057	0.0074	0.062
		10^5	0.0067	0.056	0.0073	0.060
		10^6	0.0066	0.053	0.0072	0.057
	1.2	10^4	0.0076	0.061	0.0082	0.059
		10^5	0.0075	0.059	0.0081	0.057
		10^6	0.0074	0.058	0.0081	0.055
Fréchet	1.1	10^4	0.0066	0.063	0.0075	0.062
		10^5	0.0065	0.059	0.0073	0.059
		10^6	0.0065	0.057	0.0071	0.056
	1.2	10^4	0.0073	0.065	0.0078	0.061
		10^5	0.0072	0.061	0.0075	0.056
		10^6	0.0072	0.060	0.0074	0.054

http://dutangc.perso.math.cnrs.fr/RRepository/pub/. For the period from 1985 to 1992, the annual number of claims is consistent, so we focus on this time frame in our application.
The parameters of the Norwegian fire insurance dataset are provided in Table 8 We model the Norwegian fire insurance dataset using appropriate distributions, and the fitting results are presented in Table 9 Based on the results in Table 9, we conclude that the log-normal distribution provides a good fit for the data, with parameters $shape = 0.87$ and $scale = 1962.99$.

Table 8. Parameters of description of the Norwegian fire insurance dataset

Min	1st Qu.	Median	Mean	sd	3rd Qu.	Max
0.5	0.7	1.020	2.217	7.760	1.8	465.365

The results of the estimation of the reinsurance premium for three values of distortion parameters 1.1, 1.2, and 1.3, are presented in Table 10. From Table 10, we deduce that the new estimator (MoM) is very robust compared with the traditional estimator (TM), which is very sensitive to the outliers data.

Table 9. Results of the fitting Norwegian fire insurance dataset

	Weibull	log-Normal	gamma
Kolmogorov-Smirnov statistic	0.261276	**0.124491**	0.200603
Akaike's Information Criterion	158927.9	**152859.3**	159471.2

Table 10. Results of estimation of reinsurance premium based on the Norwegian fire insurance dataset

ρ	1.1	1.2	1.3
Theoretical PHT	1.6355	1.9382	2.2603
TM	1.7431	2.4021	3.2225
MoM	1.6122	1.9628	2.2539

6. Proofs

In this section, we show the theorems 1-3; for this reason, we need the following Berry-Essen bound, which is due to [13].

Lemma 1
Assume that $E(|X|^3) < \infty$ and $\sigma_\rho^2(F) > 0$ which is defined in (6). If F has a strictly positive, continuous density function f on $[-\eta, \eta]$ for some $\eta > 0$, then there exists a constant $C > 0$ such that

$$\sup_{x\in\mathbb{R}} \left\{ \mathbb{P} \left| \left(\frac{\sqrt{N}}{\sigma_\rho(F)} \left(\tilde{\Pi}_{\rho,N} - \Pi_{\rho,R}\right) \leq x \right) - \Phi(x) \right| \right\} \leq \frac{C}{\sqrt{N}} \tag{16}$$

where Φ is the cumulative function of a standard normal random variable.

Proof of theorem 1
Define the random variables

$$\eta_{n,j} := \frac{\sqrt{n}}{\sigma_\rho(F)} \left(\tilde{\Pi}^{(j)}_{\rho,n} - \Pi_{\rho,R}\right), j = 1, ..., K \tag{17}$$

From 16, we have

$$\sup_{x\in\mathbb{R}} \left\{ \left| (\eta_{n,j} \leq x) - \Phi(x) \right| \right\} \leq \frac{C}{\sqrt{n}}$$

for each $j = 1, ..., K$. Setting $x = \sqrt{n}z/\sigma_\rho(F)$, we get

$$\mathbb{P}\left(\left(\tilde{\Pi}^{(j)}_{\rho,n} - \Pi_{\rho,R}\right) \geq z \right) \leq \frac{C}{\sqrt{n}} + 1 - \Phi\left(\frac{\sqrt{n}z}{\sigma(F)}\right)$$

for all $z > 0$. Use the elementary inequality

$$1 - \Phi\left(\frac{\sqrt{n}z}{\sigma_\rho(F)}\right) \leq e^{-\frac{nz^2}{2\sigma_\rho^2(F)}},$$

which is $o(1/\sqrt{n})$ for large n and fixed $z > 0$. Thus

$$\mathbb{P}\left(\tilde{\Pi}^{(j)}_{\rho,n} - \Pi_{\rho,R} \geq z\right) \leq \frac{C}{2\sqrt{n}},$$

and similarly, we have

$$\mathbb{P}\left(\tilde{\Pi}^{(j)}_{\rho,n} - \Pi_{\rho,R} \leq -z\right) \leq \frac{C}{2\sqrt{n}},$$

where C is a constant depending on z but not n. As a consequence, we have

$$\mathbb{P}\left(\left|\tilde{\Pi}_{\rho,n}^{(j)}-\Pi_{\rho,R}\right|\geq z\right)\leq\frac{C}{\sqrt{n}}. \tag{18}$$

We claim that

$$\left|\tilde{\Pi}_{\rho,N}^{MoM}-\Pi_{\rho,R}\right|\leq Median\left\{\left|\tilde{\Pi}_{\rho,n}^{(j)}-\Pi_{\rho,R}\right|,j=1,2,...,K\right\}. \tag{19}$$

in fact,

$$\tilde{\Pi}_{\rho,N}^{MoM}-\Pi_{\rho,R}=Median\left\{\tilde{\Pi}_{\rho,n}^{(j)}-\Pi_{\rho,R},j=1,2,...,K\right\}.$$

Note that

$$Median\left\{\tilde{\Pi}_{\rho,n}^{(j)}-\Pi_{\rho,R},j=1,2,...,K\right\}\leq Median\left\{\left|\tilde{\Pi}_{\rho,n}^{(j)}-\Pi_{\rho,R}\right|,j=1,2,...,K\right\}$$

which implies that

$$\tilde{\Pi}_{\rho,N}^{MoM}-\Pi_{\rho,R}\leq Median\left\{\left|\tilde{\Pi}_{\rho,n}^{(j)}-\Pi_{\rho,R}\right|,j=1,2,...,K\right\}.$$

Similarly, we can also prove

$$-\left(\tilde{\Pi}_{\rho,N}^{MoM}-\Pi_{\rho,R}\right)\leq Median\left\{\left|\tilde{\Pi}_{\rho,n}^{(j)}-\Pi_{\rho,R}\right|,j=1,2,...,K\right\}.$$

This prove (19). Consequently, we have:

$$\begin{aligned}\mathbb{P}\left(\left|\tilde{\Pi}_{\rho,N}^{MoM}-\Pi_{\rho,R}\right|\geq z\right) &\leq \mathbb{P}\left(Median\left\{\left|\tilde{\Pi}_{\rho,n}^{(j)}-\Pi_{\rho,R}\right|,j=1,2,...,K\right\}\geq z\right)\\ &: =\mathbb{P}\left(E\right).\end{aligned}$$

Define the Bernoulli random variables

$$\theta_j=I\left(\left|\tilde{\Pi}_{\rho,n}^{(j)}-\Pi_{\rho,R}\right|\geq z\right),j=1,2,...,K$$

$E\left(\theta_j\right)\leq C/\sqrt{n}$. Obviously, the event E happens if and only if $\sum_{j=1}^{K}\theta_j$ is larger than $K/2$. Thus,

$$\mathbb{P}\left(E\right)=\mathbb{P}\left(\sum_{j=1}^{K}\theta_j\geq\frac{K}{2}\right)\leq e^{-KE(\theta_1)}\left(\frac{2eKC}{K\sqrt{n}}\right)^{K/2}\leq\frac{C}{n^{K/5}},$$

where we have used Chernoff's inequality in the last step. This ends the proof of Theorem 1. □

For any fixed x, define the independent and identically distributed Bernoulli random variables

$$\xi_{n,j}\left(x\right):=I\left(\eta_{n,j}\leq x\right),j=1,2,...,K$$

and set $p_n(x)=\mathbb{P}(\eta_{n,j}\leq x)$. From Lemma 1,

$$\left|p_n(x)-\Phi(x)\right|=O\left(1/\sqrt{n}\right)$$

for all real x. The following lemma gives the central limit theorem for partial sums of $\xi_{n,j}\left(x\right)$.

***Lemma** 2*
Suppose $n/K\to\infty$ as $K\to\infty$. Then we have

$$\sqrt{K}\left(\frac{1}{K}\sum_{j=1}^{K}\xi_{n,j}\left(x\right)-\Phi(x)\right)\overset{\mathcal{D}}{\to}\mathcal{N}\left(0,\Phi(x)\left(1-\Phi(x)\right)\right)$$

for any fixed real x. Particularly, as $K\to\infty$,

$$\sqrt{K}\left(\frac{1}{K}\sum_{j=1}^{K}\xi_{n,j}\left(x/\sqrt{K}\right)-\frac{1}{2}-\frac{x}{\sqrt{2\pi K}}\Phi(x)\right)\overset{\mathcal{D}}{\to}\mathcal{N}\left(0,1/4\right). \tag{20}$$

Proof
By independence, for any real t, and $i^2 = -1$, we have

$$\mathbb{E}\left(e^{it\sqrt{K}\left(\frac{1}{K}\sum_{j=1}^{K}\xi_{n,j}(x)-\Phi(x)\right)}\right) = \left(\mathbb{E}\left(e^{it\frac{1}{\sqrt{K}}|\xi_{n,j}(x)-\Phi(x)|}\right)\right)^K,$$

and by the Taylor's expansion

$$\begin{aligned}\mathbb{E}\left(e^{it\frac{1}{\sqrt{K}}(\xi_{n,j}(x)-\Phi(x))}\right) &= p_n e^{it\frac{1}{\sqrt{K}}(1-\Phi(x))} + (1-p_n)\, e^{-it\frac{1}{\sqrt{K}}\Phi(x)} \\ &= 1 + it\frac{p_n}{\sqrt{K}}(1-\Phi(x)) - it\frac{(1-p_n)}{\sqrt{K}}\Phi(x) \\ &\quad -\frac{p_n}{2K}\left[t(1-\Phi(x))\right]^2 - \frac{(1-p_n)}{2K}\left[t\Phi(x)\right]^2 + o\left(K^{-1}\right) \\ &= 1 - \frac{p_n}{2K}\left[\Phi(x)(1-\Phi(x))\right] + o\left(K^{-1}\right). \qquad (21)\end{aligned}$$

where we have used the fact that $|p_n - \Phi(x)| = O(1/\sqrt{n})$, $n/K \to \infty$, $K \to \infty$ and

$$\left|\frac{p_n}{\sqrt{K}}(1-\Phi(x)) + \frac{1-p_n}{\sqrt{K}}\Phi(x)\right| = \left|\frac{p_n - \Phi(x)}{\sqrt{K}}\right| = o(1/K)$$

Now the first conclusion of this lemma follows easily from (21).
For the second part, we observe that the above calculations still hold if we replace x with $x/\sqrt{K}$ and note the fact fact

$$\Phi\left(x/\sqrt{K}\right) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}}\int_0^{x/\sqrt{K}} e^{-u^2/2}du = \frac{1}{2} + \frac{x}{\sqrt{2\pi K}} + o\left(K^{-1/2}\right).$$

Now, the proof is complete, according to Slutsky's Theorem. □

Proof of Theorem 2

1. This follows immediately from Lemma 1 and the continuous mapping theorem since the Median function is continuous.
2. First, we observe that

$$\begin{aligned}\frac{\sqrt{N}}{\sigma(F)}\left(\tilde{\Pi}^{MoM}_{\rho,N} - \Pi_{\rho,R}\right) &= \sqrt{K}\frac{\sqrt{n}}{\sigma_\rho(F)}\left(\tilde{\Pi}^{MoM}_{\rho,N} - \Pi_{\rho,R}\right) \\ &= \sqrt{K}Median\left\{\eta_{n,j}, j=1,...,K\right\}. \qquad (22)\end{aligned}$$

We first assume K is odd and for any real x, we have

$$\begin{aligned}&\mathbb{P}\left(\sqrt{K}Median\left\{\eta_{n,j}, j=1,...,K\right\} \le x\right) \\ =\ &\mathbb{P}\left(\sum_{j=1}^{K} I\left(\eta_{n,j} \le \frac{x}{\sqrt{K}}\right) \ge \frac{K+1}{2}\right) \\ =\ &\mathbb{P}\left(\sum_{j=1}^{K}\xi_{n,j}\left(\frac{x}{\sqrt{K}}\right) \ge \frac{K+1}{2}\right) \\ =\ &\mathbb{P}\left(\sqrt{K}\left\{\frac{1}{K}\sum_{j=1}^{K}\xi_{n,j}\left(\frac{x}{\sqrt{K}}\right) - \frac{1}{2} - \frac{x}{\sqrt{2\pi K}}\right\} \ge -\frac{x}{\sqrt{2\pi}} + O\left(\frac{1}{\sqrt{K}}\right)\right)\end{aligned}$$

which tends to $\Phi\left(\sqrt{2/\pi x}\right)$ by the above lemma 2. If K is even, then

$$\mathbb{P}\left(\sqrt{K}Median\left\{\eta_{n,j}, j=1,...,K\right\} \le x\right) \ge \mathbb{P}\left(\sum_{j=1}^{K} I\left(\eta_{n,j} \le \frac{x}{\sqrt{K}}\right) \ge \frac{K}{2}+1\right)$$

and

$$\mathbb{P}\left(\sqrt{K}Median\left\{\eta_{n,j}, j=1,...,K\right\} \le x\right) \le \mathbb{P}\left(\sum_{j=1}^{K} I\left(\eta_{n,j} \le \frac{x}{\sqrt{K}}\right) \ge \frac{K}{2}\right)$$

The right-hand sides of the above two inequalities tend to $\Phi\left(\sqrt{2/\pi x}\right)$ as $K \to \infty$. Now, we complete the whole proof of Theorem 2.

□

Proof of Theorem 3
Recall that

$$Z_{n,k} = I\left(\tilde{\Pi}^{(k)}_{\rho,n} - \Pi_{\rho,R}\right), \text{ for } k = 1, 2, ..., K.$$

Write (11) as

$$f(\lambda) = \frac{1}{K}\sum_{j=1}^{K} \frac{Z_{n,k}-0.5}{1+\lambda\left(Z_{n,k}-0.5\right)} = 0. \tag{23}$$

Write

$$U_{n,k} = \lambda\left(Z_{n,k}-0.5\right).$$

By (23), we can easily get

$$\bar{Z}_{n,k} - 0.5 = \lambda\overline{S}, \tag{24}$$

where

$$\overline{S} = \frac{1}{K}\sum_{k=1}^{K}\frac{\left(Z_{n,k}-0.5\right)^2}{1+U_{n,k}}$$

and

$$\bar{Z}_{n,k} = \frac{1}{K}\sum_{k=1}^{K} Z_{n,k}.$$

Note

$$S = \frac{1}{K}\sum_{k=1}^{K}\left(Z_{n,k}-0.5\right)^2 = 0.25,$$

and

$$Z_K = \max_{1\le k\le K}\left|Z_{n,k}-0.5\right| = 0.5.$$

Combining the constraint condition $\omega_i > 0$, we can derive that $1 + U_{n,k} > 0$, and

$$\lambda S \le \lambda\overline{S}\left(1+\max_{1\le k\le K} U_{n,k}\right) \le \lambda\overline{S}\left(1+\lambda Z_K\right) = \left(\bar{Z}_{n,k}-0.5\right)\left(1+\lambda Z_K\right)$$

The last equality follows by (24). Hence,

$$\lambda\left[S - \left(\bar{Z}_{n,k}-0.5\right)Z_K\right] \le \bar{Z}_{n,k} - 0.5.$$

Furthermore, together with Lemma 2, $\bar{Z}_{n,k}-0.5=O_p\left(1/\sqrt{K}\right)$. We have

$$\lambda\left[0.25-O_p\left(1/\sqrt{K}\right)\right]=O_p\left(1/\sqrt{K}\right).$$

So $\lambda=O_p\left(1/\sqrt{K}\right)$. In addition, we have

$$\max_{1\leq k\leq K}|U_{n,k}|=O_p\left(1/\sqrt{K}\right)=o_p(1).$$

Expanding (23) gives

$$\begin{aligned} 0 &= \frac{1}{K}\sum_{k=1}^{K}\frac{(Z_{n,k}-0.5)}{1+U_{n,k}} \\ &= \frac{1}{K}\sum_{k=1}^{K}(Z_{n,k}-0.5)\left(1-U_{n,k}+\frac{U_{n,k}^2}{1+U_{n,k}}\right) \\ &= \left(\bar{Z}_{n,k}-0.5\right)-\lambda S+\frac{1}{K}\sum_{k=1}^{K}\frac{(Z_{n,k}-0.5)}{1+U_{n,k}}U_{n,k}^2 \end{aligned} \tag{25}$$

The final term in (25) above has a norm bounded by

$$\frac{1}{K}\sum_{k=1}^{K}\lambda^2\frac{|Z_{n,k}-0.5|^3}{1+U_{n,k}}=O\left(1\right)O_p\left(1/K\right)O_p\left(1\right)=o_p\left(1/\sqrt{K}\right).$$

Hence

$$\lambda=S^{-1}\left(\bar{Z}_{n,k}-0.5\right)+\beta=4\left(\bar{Z}_{n,k}-0.5\right)+\beta$$

with $\beta=o_p\left(1/\sqrt{K}\right)$. By (25) and using Taylor expansion, we can find

$$\log\left(1+U_{n,k}\right)=U_{n,k}-\frac{1}{2}U_{n,k}^2+\eta_k$$

holds for some finite $B>0, 1\leq k\leq K$,

$$P\left(|\eta_k|\leq B\left|U_{n,k}\right|^3\right)\to 1,\ \text{as } K\to\infty \text{ and } m\to\infty.$$

Now, direct calculation yields that

$$\begin{aligned} -2\log\mathcal{R}\left(\Pi_{\rho,R}\right) &= 2\sum_{k=1}^{K}\log\left(1+U_{n,k}\right) \\ &= 2\sum_{k=1}^{K}\left(U_{n,k}-\frac{1}{2}U_{n,k}^2+\eta_k\right) \\ &= 4K\left(\bar{Z}_{n,k}-0.5\right)^2-4K\beta^2+2\sum_{k=1}^{K}\eta_k \end{aligned}$$

By Lemma 2, we have

$$4K\left(\bar{Z}_{n,k}-0.5\right)^2\xrightarrow{\mathcal{D}}\chi_1^2.$$

Note that:

$$4K\beta^2=4Ko_p\left(1/K\right)=o_p\left(1\right),$$

and

$$\left|\sum_{k=1}^{K}\eta_k\right|\leq B\left|\lambda\right|^3\sum_{k=1}^{K}\left|Z_{n,k}-0.5\right|^3=O_p\left(1/\sqrt{K^3}\right)O\left(1\right)=o_p\left(1\right).$$

This completes the proof. □

7. Conclusion

In this paper, we propose a new and robust non-parametric estimator for the excess-of-loss reinsurance premium based on a grouping strategy. The asymptotic properties, including consistency and asymptotic normality of the proposed estimator, are obtained. Due to the complexity of the variance term in the normal approximation of the proposed estimator, we construct a new test for the excess-of-loss reinsurance premium based on the empirical likelihood method for the median. Numerical simulations confirm that our newly proposed estimator is quite robust regarding outliers.

Acknowledgement

The authors would like to thank the reviewers, the Associate Editor, and the Managing Editor for their very valuable comments and suggestions, which led to an improved presentation of the paper.

REFERENCES

1. H. Akaike, *Information theory and an extension of the maximum likelihood principle. In B. N. Petrov, F. Csaki (Eds.)*, Proceedings of the 2nd International Symposium on Information Theory, 267–281, 1973.
2. Y. Alon Matias and M. Szegedy, *The space complexity of approximating the frequency moments*, Journal of Computer and System Sciences, 58, 137–147, 2002.
3. P. Artzner, F. Delbaen, J.M. Eber, and D. Heath, *Coherent measures of risk*, Mathematical finance, 9(3), 203-228, 1999.
4. J. Beirlant, Goegebeur, Segers and J. Teugels, *Statistics of extremes: theory and applications*, John Wiley, 2006
5. D. Bhati, S.K. Kattumannil, and N. Sreelakshmi, *Jackknife empirical likelihood-based inference for probability weighted moments*, J. Korean Stat. Soc. https://doi.org/10.1007/s42952-020-00062-9, 2020.
6. M.L. Centeno, and J. Andrade e Silva, *Applying the proportional hazard premium calculation principle*, Astin Bulletin 35, 409–425, 2005.
7. J. Chen, A.M. Variyath, and B. Abraham, *Adjusted empirical likelihood and its properties*, Journal of Computational and Graphical Statistics, 17, 426–443, 2008.
8. H.A David, and H.N. Nagaraja, *Order Statistics*, New York: John Wiley and Sons, 2003.
9. J.A. Greenwood, J.M. Landwehr, N. C. Matalas, and J.R. Wallis, *Probability weighted moments*, Water Resource Research, 15, 1055–1064, 1979.
10. W. Hoeffding, *A class of statistics with asymptotically normal distribution*, The annals of mathematical statistics, 19, 293–325, 1948.
11. J.R.M. Hosking, J.R. Wallis, and E.F. Wood, E. F. *Estimation of the generalized extreme-value distribution by the method of probability-weighted moments*, Technometrics, 27, 251–261, 1985.
12. B.Y. Jing, J. Yuan, and W. Zhou, *Jackknife empirical likelihood*, Journal of the American Statistical Association, 104, 1224–1232, 2009.
13. B.L. Jnes, and R. Zitikis, *Empirical estimation of risk measures and related quantities*, North American Actuarial Journal 7, 44–54, 2003.
14. G. Lugosi and S. Mendelson, *Sub-Gaussian estimators of the mean of a random vector*, Ann. Statist., 47 , 783–794, 2019.
15. A. Owen, *Empirical likelihood ratio confidence regions*, The Annals of Statistics, 18 (1990), 90–120, 1990.
16. L. Peng, Y. Qi, R. Wang and J. Yang, *Jackknife empirical likelihood method for some risk measures and related quantiles*, Insurance: Mathematics and Economics, 51, 142–150, 2012.
17. S. Wang, *Premium calculation by transforming the layer premium density*, ASTIN Bulletin: The Journal of the IAA, 26(1), 71-92, 1996.
18. Y. Zhao, X, Meng, and H. Yang, *Jackknife empirical likelihood inference for the mean absolute deviation*, Computational Statistics & Data Analysis, 91, 92–101, 2015.