



HAL
open science

Confidence intervals for the proportional hazard reinsurance premium for heavy-tailed claims

Abdelkader Amaraoui, Kamal Boukhetala, Jean-François Dupuy

► **To cite this version:**

Abdelkader Amaraoui, Kamal Boukhetala, Jean-François Dupuy. Confidence intervals for the proportional hazard reinsurance premium for heavy-tailed claims. 2024. hal-04440437

HAL Id: hal-04440437

<https://hal.science/hal-04440437>

Preprint submitted on 6 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Confidence intervals for the proportional hazard reinsurance premium for heavy-tailed claims

Abdelkader Ameraoui^a, Kamal Boukhetala^b, Jean-François Dupuy^c

^a*National Higher School of Mathematics P.O. Box 75, Mahelma 16093, Sidi Abdellah, Algiers, Algeria*

^b*Faculty of mathematics P.O. Box 32, Al alia Bab ezzouar, USTHB, Algiers, Algeria*

^c*Univ Rennes, INSA Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France*

Abstract

Estimation of an adequate premium in presence of extreme events is of crucial importance for reinsurance companies. In this paper, we consider the statistical issue of constructing confidence intervals for the proportional hazard premium (PHP) under high-excess loss layer. A straightforward confidence interval is based on the asymptotic normal approximation of the PHP. However, its coverage accuracy can be unsatisfactory. In this paper, we propose two other methods, namely, the likelihood ratio method and the data tilting method. We establish their asymptotic properties. Then, we assess their finite-sample performance via simulations. Finally, we apply the proposed methods on two real data sets (Danish fire losses and Algerian car insurance claims). Our investigations suggest that the data tilting method provides the best results.

Keywords: data tilting ; likelihood ratio method ; Hill estimator ; normal approximation ; tail index

1. Introduction

Extreme events arise in a wide variety of domains : environment (flooding, heat waves), industry (industrial accident, fire), finance (crash, extreme loss) . . . and can cause considerable losses in insurers portfolio. It has therefore become crucial for insurance and reinsurance companies to determine adequate premiums for extreme risks. Several premium calculation principles have been proposed for that purpose. One of the most recent principles is based on a distortion function, that is, an increasing and concave function $g : [0, 1] \rightarrow [0, 1]$ satisfying $g(0) = 0$ and $g(1) = 1$ (Wang, 1996).

Let X be a random risk with distribution function (df) F . The distortion risk measure based on g is defined as

$$\Pi(g) = \int_0^\infty g(1 - F(x)) dx. \quad (1.1)$$

The well-known proportional hazard premium - PHP (Wang, 1995) arises as a special case of (1.1) when $g(x) = x^{1/\rho}$, for $\rho \geq 1$. Several other examples of distortion functions are given in Wang (1996). The PHP of X is thus given by:

$$\Pi_\rho = \int_0^\infty (1 - F(x))^{1/\rho} dx,$$

which is also the distorted expectation of X . The parameter ρ controls the amount of risk loading in the premium. It is called the distortion coefficient, or risk aversion index. The estimation of

Π_ρ has been investigated by several authors, see for example Centeno and Andrade e Silva (2005), Necir and Meraghni (2009), Rassoul (2012a), Deme et al. (2013).

In reinsurance, one is often interested in estimating a premium for a given retention level $R > 0$, that is, a reinsurance premium of the high layer $[R, \infty)$. This issue arises when an insurer decides to transfer part of a risk to a reinsurer, since it may not have sufficient financial resources to bear the total risk. In this case, the reinsurer will not pay the insurer if X is less than or equal to R and will pay $(X - R)$ if X exceeds R (equivalently said, the reinsurer compensates the cedant's loss above the retention level R only). The amount paid by the reinsurer is thus $(X - R)_+$, where $x_+ = \max(0, x)$, and the corresponding PHP for the layer $[R, \infty)$ is defined as the distorted expectation of $(X - R)_+$:

$$\Pi_{\rho,R} = \int_R^\infty (1 - F(x))^{\frac{1}{\rho}} dx.$$

For high-excess loss layers ($R \rightarrow \infty$) of a heavy-tailed insured risk, Necir and Boukhetala (2004), Vandewalle and Beirlant (2006), Necir et al. (2007), Rassoul (2012b), Benkhelifa (2014), Ahmedou et al. (2023) proposed various estimators of $\Pi_{\rho,R}$ and proved their asymptotic normality. These estimates are obtained for a fixed risk aversion index ρ and an optimal retention level $R_{opt} = F^{\leftarrow}(1 - \varepsilon)$, where ε is a small number and $F^{\leftarrow}(s) = \inf\{t > 0, F(t) \geq s\}$ is the generalized inverse of F . As an extension of these contributions, our interest, here, is to construct confidence intervals for $\Pi_{\rho,R}$, in the case of high-excess loss layers.

One obvious first method to construct a confidence interval for $\Pi_{\rho,R}$ is to use the asymptotic normality of an estimate of $\Pi_{\rho,R}$. This approach is described in Section 2.1. Our simulation study in Section 3 will show, however, that the coverage probabilities of the resulting intervals can be quite far from the nominal confidence level. Therefore, in this paper, we investigate two alternative methods, namely a likelihood ratio method and a data tilting method (data tilting is a non-parametric approach which can be seen as a generalization of the empirical likelihood method of Owen (2001)). Both methods have recently proved useful to construct confidence intervals in various contexts. For example, Lu and Peng (2002) and Peng and Qi (2006) use the likelihood ratio method to obtain confidence intervals for the tail index and high quantiles of a heavy-tailed distribution. Empirical likelihood was used by Lu and Peng (2002), Li and Qi (2019), Qi et al. (2023) to construct confidence intervals for the tail index and high quantiles of a heavy-tailed distribution. Data tilting was introduced by Hall and Yao (2003) for constructing confidence regions in regression problems. Peng and Qi (2006) adapted the method to construct confidence intervals for high quantiles of a heavy-tailed distribution. Chan et al. (2007) and Tursunaliyeva and Silvapulle (2016) used data tilting to construct confidence intervals for the value-at-risk (VaR).

Our paper is organized as follows. In Section 2, we briefly describe asymptotic normality-based confidence intervals for $\Pi_{\rho,R}$. Then, we show how the likelihood ratio and data tilting methods can be adapted to construct confidence intervals for $\Pi_{\rho,R}$. We state two results about the asymptotic properties of these methods. In Section 3, the results of a simulation study comparing the performance of the three methods are reported. We also describe two real data applications of the proposed confidence intervals (on a set of Danish fire losses and a set of Algerian car insurance claims). Proofs of the theoretical results are provided in an appendix.

2. Confidence intervals for the premium $\Pi_{\rho,R}$: methodology

2.1. The normal approximation approach

Let X_1, \dots, X_n be a sample of independent and identically distributed risks with common df F . We assume that F is heavy-tailed, that is, for every $x > 0$, the survival function $1 - F$ satisfies:

$$1 - F(x) = x^{-\alpha} \ell(x), \quad (2.2)$$

where $\alpha > 0$ is an unknown parameter (called the tail index) and ℓ is a slowly-varying function at infinity, that is,

$$\lim_{x \rightarrow \infty} \frac{\ell(tx)}{\ell(x)} = 1, \quad \text{for all } t > 0.$$

The survival function $1 - F$ of a heavy-tailed distribution thus satisfies:

$$\lim_{x \rightarrow \infty} \frac{1 - F(tx)}{1 - F(x)} = t^{-\alpha}, \quad \text{for all } t > 0,$$

and is said to be regularly-varying with tail index α .

Let $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$ be the order statistics of the sample X_1, \dots, X_n and $k = k(n)$ be an integer sequence such that $k \rightarrow \infty$ and $k/n \rightarrow 0$ as $n \rightarrow \infty$. Necir et al. (2007) consider the optimal retention level $R_{opt} = F^{\leftarrow}(1 - k/n)$ and propose the following estimator of $\Pi_{\rho,R_{opt}}$:

$$\hat{\Pi}_{\rho, \hat{R}_{opt}} = \left(\frac{k}{n}\right)^{\frac{1}{\rho}} \frac{\rho}{1/\hat{\gamma} - \rho} X_{n-k,n}, \quad (2.3)$$

where $\hat{R}_{opt} = X_{n-k,n}$ and $\hat{\gamma}$ is the Hill estimator (Hill, 1975) of the extreme value index $\gamma = 1/\alpha$, defined by:

$$\hat{\gamma} = \frac{1}{k} \sum_{i=1}^k \log X_{n-i+1,n} - \log X_{n-k,n}. \quad (2.4)$$

Remark 1. In Necir et al. (2007), $\hat{\Pi}_{\rho, \hat{R}_{opt}}$ is obtained by substituting, in $\Pi_{\rho,R} = \int_R^{\infty} (1 - F(x))^{\frac{1}{\rho}} dx$, the retention level $R = R_{opt} = F^{\leftarrow}(1 - k/n)$ by $\hat{R}_{opt} = X_{n-k,n}$ and the tail $1 - F(x)$, $x \rightarrow \infty$, of the heavy-tailed df F by the approximation $1 - \hat{F}(x) = (k/n) X_{n-k,n}^{1/\hat{\gamma}} x^{-1/\hat{\gamma}}$. Integrating over x finally yields (2.3). ■

The asymptotic normality of $\hat{\Pi}_{\rho, \hat{R}_{opt}}$ is established under the so-called second-order regular variation condition (de Haan and Ferreira, 2006). Let $U(t) = (1/(1 - F))^{\leftarrow}(t)$. Then U is said to satisfy the second-order regular variation condition if there exists a function A such that $A(t) \rightarrow 0$ as $t \rightarrow \infty$ and

$$\lim_{t \rightarrow \infty} \frac{U(tx)/U(t) - x^\gamma}{A(t)} = x^\gamma \frac{x^\beta - 1}{\beta} \quad (2.5)$$

for all $x > 0$, where $\beta \leq 0$.

Theorem 2.1. (Necir et al., 2007, 2010) Let $\rho \geq 1$ and assume that (2.5) holds. Let $k = k(n)$ be an integer sequence such that $k \rightarrow \infty$, $k/n \rightarrow 0$ and $\sqrt{k}A(n/k) \rightarrow 0$ as $n \rightarrow \infty$. Then, for $0 < \gamma < 1/\rho$, the following holds as $n \rightarrow \infty$:

$$\frac{(k/n)^{-\frac{1}{\rho}}}{X_{n-k,n}} \sqrt{k}(\hat{\Pi}_{\rho, \hat{R}_{opt}} - \Pi_{\rho, R_{opt}}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2(\rho, \gamma)),$$

where

$$\sigma^2(\rho, \gamma) = \frac{\rho\gamma(\gamma^2 - 2\gamma^3\rho + \gamma^4\rho^2 + 1)}{(1 - \rho\gamma)^2}.$$

Based on this result, a confidence interval with asymptotic confidence level $1 - p$ for $\Pi_{\rho, R_{opt}}$ is:

$$\hat{\Pi}_{\rho, \hat{R}_{opt}} \pm z_{1-p/2} \frac{\sigma(\rho, \hat{\gamma}) X_{n-k,n} (k/n)^{\frac{1}{\rho}}}{\sqrt{k}},$$

where $z_{1-p/2}$ denotes the quantile of order $1 - p/2$ of the standard normal distribution and $\hat{\gamma}$ is the Hill estimator, see Necir et al. (2007). We will see in our simulation study that the coverage accuracy of this interval can be quite far from the nominal confidence level.

Now, we investigate two alternative methods for constructing confidence intervals for $\Pi_{\rho, R}$.

2.2. Likelihood ratio method

First, we provide an interpretation of $\hat{\Pi}_{\rho, \hat{R}_{opt}}$ as a maximum likelihood estimator in a slightly modified form of model (2.2). This observation will motivate our construction of the likelihood ratio confidence interval. Assume that F has the form

$$1 - F(x) = cx^{-\alpha}, \quad \text{for } x > u, \quad (2.6)$$

where u is some high threshold and $c > 0$ is an unknown parameter (Lu and Peng (2002) and Peng and Qi (2006) construct confidence intervals for the tail index and high quantiles of a heavy-tailed distribution, based on this model).

By substituting $1 - F(x) = cx^{-\alpha}$ in $\Pi_{\rho, R} = \int_R^\infty (1 - F(x))^{\frac{1}{\rho}} dx$ and integrating over x , we obtain the following expression for the premium $\Pi_{\rho, R}$ in model (2.6):

$$\Pi_{\rho, R} = c^{\frac{1}{\rho}} \frac{\rho}{\alpha - \rho} R^{1 - \frac{\alpha}{\rho}}. \quad (2.7)$$

Now, consider the set of left-censored observations $\{(\max(X_i, u), \delta_i); i = 1, \dots, n\}$, where $\delta_i = 1_{\{X_i > u\}}$. Based on these observations, the likelihood of (α, c) in model (2.6) is:

$$L(\alpha, c) = \prod_{i=1}^n (c\alpha X_i^{-\alpha-1})^{\delta_i} (1 - cu^{-\alpha})^{1-\delta_i}. \quad (2.8)$$

Let $u = X_{n-k,n}$, where $k = k(n)$ is an integer sequence such that $k \rightarrow \infty$ and $k/n \rightarrow 0$ as $n \rightarrow \infty$. Then $L(\alpha, c)$ becomes

$$L(\alpha, c) = (c\alpha)^k (1 - cX_{n-k,n}^{-\alpha})^{n-k} \prod_{i=1}^n X_i^{-(\alpha+1)\delta_i},$$

since $\sum_{i=1}^n \delta_i = k$. Let $(\hat{\alpha}, \hat{c})$ denote the maximum likelihood estimator of (α, c) . Equivalently, $(\hat{\alpha}, \hat{c})$ maximizes the log-likelihood

$$\log L(\alpha, c) = k \log \alpha + k \log c - (\alpha + 1) \sum_{i=1}^k \log X_{n-i+1, n} + (n - k) \log(1 - cX_{n-k, n}^{-\alpha}).$$

It is easy to verify that $\hat{\alpha} = \hat{\gamma}^{-1}$ and $\hat{c} = (k/n)X_{n-k, n}^{\hat{\alpha}}$. Substituting $\hat{\alpha}$ and \hat{c} in (2.7), with $\hat{R}_{opt} = X_{n-k, n}$, yields (2.3).

Now, we construct our likelihood ratio-based confidence intervals for $\Pi_{\rho, R_{opt}}$. Motivated by the maximum likelihood interpretation of $\hat{\Pi}_{\rho, \hat{R}_{opt}}$, we propose to maximize $\log L(\alpha, c)$ subject to

$$\alpha > 0, \quad c > 0, \quad \log \Pi_{\rho, R_{opt}} - \frac{1}{\rho} \log c - \log \rho + \log(\alpha - \rho) - \left(1 - \frac{\alpha}{\rho}\right) \log R_{opt} = 0,$$

where the last constraint comes from taking the logarithm in (2.7). Let $\bar{\alpha}(\lambda_n)$ and $\bar{c}(\lambda_n)$ be the maximizers of this constrained log-likelihood, where λ_n denotes Lagrange multiplier. Then it is not difficult to show that $\bar{\alpha}(\lambda_n)$ and $\bar{c}(\lambda_n)$ satisfy

$$\bar{\alpha}(\lambda_n) = \frac{k}{\sum_{i=1}^k \log \left(\frac{X_{n-i+1, n}}{X_{n-k, n}} \right) - \frac{\lambda_n}{\rho} (\Pi_{\rho, R_{opt}} - \log X_{n-k, n})}$$

and

$$\bar{c}(\lambda_n) = \frac{\rho k - \lambda_n}{\rho n - \lambda_n} X_{n-k, n}^{\bar{\alpha}(\lambda_n)},$$

and λ_n satisfies

$$\begin{cases} \bar{\alpha}(\lambda_n) > 0, & \lambda_n < \rho k, \\ \log \Pi_{\rho, R_{opt}} - \frac{1}{\rho} \log \bar{c}(\lambda_n) - \log \rho + \log(\bar{\alpha}(\lambda_n) - \rho) - \left(1 - \frac{\bar{\alpha}(\lambda_n)}{\rho}\right) \log R_{opt} = 0. \end{cases} \quad (2.9)$$

We define the log-likelihood ratio statistic as:

$$\ell(\Pi_{\rho, R_{opt}}) = -2 \{ \log L(\bar{\alpha}(\lambda_n), \bar{c}(\lambda_n)) - \log L(\hat{\alpha}, \hat{c}) \}.$$

The next theorem gives the asymptotic distribution of $\ell(\Pi_{\rho, R_{opt}})$.

Theorem 2.2. *Let $1 \leq \rho < \alpha$ and assume that (2.5) holds. Let $k = k(n)$ be an integer sequence such that $k \rightarrow \infty$, $k/n \rightarrow 0$, $\sqrt{k}/\log n \rightarrow 0$ and $\sqrt{k}A(n/k) \rightarrow 0$ as $n \rightarrow \infty$. Then there exists a unique solution λ_n to (2.9), and the sequence (λ_n) is such that $\ell(\Pi_{\rho, R_{opt}})$ satisfies:*

$$\ell(\Pi_{\rho, R_{opt}}) \xrightarrow{\mathcal{L}} \chi_1^2 \quad \text{as } n \rightarrow \infty.$$

Based on this result, a confidence interval with level $1 - p$ for the true value of $\Pi_{\rho, R_{opt}}$ is:

$$\{ \Pi_{\rho, R_{opt}} : \ell(\Pi_{\rho, R_{opt}}) \leq c_{1-p} \},$$

where c_{1-p} denotes the quantile of order $1 - p$ of the χ_1^2 distribution. The performance of this confidence interval will be assessed via simulations in Section 3.

2.3. Data tilting method

In this section, we adapt the data tilting method to construct new confidence intervals for the reinsurance premium. Data tilting is a non-parametric approach which can be seen as a generalization of the empirical likelihood method (Owen, 2001), where the observations are weighted in order to minimize some distance function. Data tilting was introduced by Hall and Yao (2003) for constructing confidence regions in regression problems. Peng and Qi (2006) adapted the method to construct confidence intervals for high quantiles of a heavy-tailed distribution. Chan et al. (2007) and Tursunalieva and Silvapulle (2016) used it to construct confidence intervals for the VaR.

The method is as follows. Let $q = (q_1, \dots, q_n)$ be a set of weights such that $q_i \geq 0$ and $\sum_{i=1}^n q_i = 1$. First, we find the couple $(\hat{\alpha}(q), \hat{c}(q))$ such that

$$(\hat{\alpha}(q), \hat{c}(q)) = \arg \max_{(\alpha, c)} \sum_{i=1}^n q_i \log \left((c\alpha X_i^{-\alpha-1})^{\delta_i} (1 - cX_{n-k,n}^{-\alpha})^{1-\delta_i} \right),$$

that is, $(\hat{\alpha}(q), \hat{c}(q))$ maximizes a weighted version of the likelihood (2.8), where the contribution of the i -th individual is weighted by q_i . It is easy to see that:

$$\hat{c}(q) = X_{n-k,n}^{\hat{\alpha}(q)} \sum_{i=1}^n q_i \delta_i \quad \text{and} \quad \hat{\alpha}(q) = \frac{\sum_{i=1}^n q_i \delta_i}{\sum_{i=1}^n q_i \delta_i (\log X_i - \log X_{n-k,n})}.$$

Note that $\hat{\alpha}(q)$ and $\hat{c}(q)$ reduce to $\hat{\alpha}$ and \hat{c} when $q_i = 1/n$ for every i .

Now, let $\mathbb{D}(q)$ denote Kullback-Leibler divergence between q and the uniform distribution on the integers $\{1, \dots, n\}$ (which has $q_i = 1/n$ for every $i = 1, \dots, n$). We have $\mathbb{D}(q) = \sum_{i=1}^n q_i \log(nq_i)$. We choose q so as to minimize $\mathbb{D}(q)$ under the constraints $q_i \geq 0$,

$$\sum_{i=1}^n q_i = 1 \quad \text{and} \quad \Pi_{\rho, R_{opt}} = \hat{c}(q)^{\frac{1}{\rho}} \frac{\rho}{\hat{\alpha}(q) - \rho} X_{n-k,n}^{1 - \frac{\hat{\alpha}(q)}{\rho}},$$

or equivalently,

$$\sum_{i=1}^n q_i = 1 \quad \text{and} \quad \Pi_{\rho, R_{opt}} \left(\frac{\hat{\alpha}(q)}{\rho} - 1 \right) - X_{n-k,n} \left(\sum_{i=1}^n q_i \delta_i \right)^{\frac{1}{\rho}} = 0. \quad (2.10)$$

Let

$$A_1(\lambda_1) = 1 - \frac{(n-k)}{n} e^{-1-\lambda_1} \quad \text{and} \quad A_2(\lambda_1) = \frac{A_1(\lambda_1) \Pi_{\rho, R_{opt}}}{\rho \left(\Pi_{\rho, R_{opt}} + X_{n-k,n} (A_1(\lambda_1))^{\frac{1}{\rho}} \right)}.$$

Then, minimizing $\mathbb{D}(q)$ with respect to the q_i , under the constraints stated above, gives:

$$q_i(\lambda_1, \lambda_2) = \begin{cases} \frac{1}{n} e^{-1-\lambda_1} & \text{if } \delta_i = 0 \\ \frac{1}{n} \exp \left(-1 - \lambda_1 - \lambda_2 \left[\frac{\Pi_{\rho, R_{opt}}}{\rho A_2(\lambda_1)} - \frac{\Pi_{\rho, R_{opt}} A_1(\lambda_1) \log(X_i/X_{n-k,n})}{\rho (A_2(\lambda_1))^2} - \frac{X_{n-k,n} (A_1(\lambda_1))^{\frac{1}{\rho}-1}}{\rho} \right] \right) & \text{if } \delta_i = 1, \end{cases}$$

where λ_1 and λ_2 are the Lagrange multipliers, which satisfy (2.10). In what follows, we will note $q_i := q_i(\lambda_1, \lambda_2)$ for notational simplicity. We also denote by $L(\Pi_{\rho, R_{opt}})$ the value $\min_q (2n\mathbb{D}(q))$ achieved under the constraints (2.10). We are now in position to state our second main result.

Theorem 2.3. *Assume that the conditions in Theorem 2.2 hold. Then there exists a solution (λ_1, λ_2) to the constraints (2.10), with $-1 - \log\left(1 + \frac{\sqrt{k}}{n-k}\right) \leq \lambda_1 \leq -1 - \log\left(1 - \frac{\sqrt{k}}{n-k}\right)$, such that under the true value of $\Pi_{\rho, R_{opt}}$,*

$$L(\Pi_{\rho, R_{opt}}) \xrightarrow{\mathcal{L}} \chi_1^2 \quad \text{as } n \rightarrow \infty.$$

Based on this, a confidence interval with level $1 - p$ for the true $\Pi_{\rho, R_{opt}}$ can be constructed as:

$$\{\Pi_{\rho, R_{opt}} : L(\Pi_{\rho, R_{opt}}) \leq c_{1-p}\},$$

where c_{1-p} is the $(1 - p)$ -quantile of the χ_1^2 distribution.

In the next section, we investigate the finite-sample performance of the three types of confidence intervals.

3. A simulation study and real applications

3.1. A simulation study

In this study, we compare the coverage accuracy of the confidence intervals constructed using the normal approximation method, the likelihood ratio method and the data tilting method. Coverage accuracy is evaluated in terms of coverage probabilities and length of the intervals. The simulation scenarios are as follows.

First, we simulate $N = 1000$ samples of size $n = 1000$ from the Pareto distribution with df $F(x) = 1 - x^{-\alpha}$, $x > 0$, and then from the Burr distribution with df $F(x) = 1 - (1 + x^{\beta-\alpha})^{-\alpha/(\beta-\alpha)}$, $x > 0$. We consider various combinations of values for the parameters, namely $\alpha = 2$ and $\rho = 1.2, 1.5$ for Pareto distribution, and $(\alpha, \beta) = (2, 4), (1.5, 3)$ and $\rho = 1.1, 1.2$ for Burr distribution. Recall that ρ is the risk aversion index, which controls the amount of risk loading in the premium. The nominal confidence level of the confidence intervals is set to 0.9.

For each value k of the sample fraction (with $k = 20, 22, 24, \dots, 300$), we calculate the confidence intervals for the three methods, and we compute their coverage probabilities as the proportions of the N intervals which contain the true premium value. These coverage probabilities are plotted against k (see Figures 1 and 2 for Burr distribution and Figure 3 for Pareto distribution).

From these figures, it first appears that the normal approximation method has rather bad performance, with coverage probabilities being far from the nominal confidence level, for every value of k . The likelihood ratio and data tilting methods perform equally on the range $k = 20, \dots, 220$ approximately (except in the case of Burr(2, 4) with $\rho = 1.1$ where both methods perform equally well on the whole range $k = 20, \dots, 300$). The data tilting method is superior when k takes larger values, which indicates that this method is more robust to the bias when a large value of k is used. This may stem from the automatic assignment of the weights q_i in the data tilting methodology. A similar observation was made by Peng and Qi (2006) in their assessment of the data tilting method for constructing confidence intervals for high quantiles of a heavy-tailed distribution.

In a second part, we briefly investigate the finite-sample distribution of $L(\Pi_{\rho, R_{opt}})$ in Theorem 2.3 (in what follows, we focus on the data tilting method since it provides the best results among the three methods). We simulate $N = 1000$ samples of size $n = 1000$ of the Burr(2, 4) and Pareto(1.5) distributions and we calculate $L(\Pi_{\rho, R_{opt}})$ for various values of ρ , with $k = 50, 100$. For each simulation setting, we plot the histogram of the N values of $L(\Pi_{\rho, R_{opt}})$ and we add the density

of the χ_1^2 . Plots for the Burr(2, 4) are given in Figure 4, plots for the Pareto(1.5) are given in Figure 5. Both indicate a correct approximation of the distribution of $L(\Pi_{\rho, R_{opt}})$ by the χ_1^2 .

Finally, as in Peng and Qi (2006), we investigate the shape of $L(\Pi_{\rho, R_{opt}})$ in the data tilting method. We simulate a sample of size $n = 1000$ from the Burr($\alpha, 2\alpha$) distribution (then, from Pareto(α)), and we compute the data tilting function $L(\Pi_{\rho, R_{opt}})$ at all $\hat{\Pi}_{\rho, \hat{R}_{opt}} \pm \hat{\Pi}_{\rho, \hat{R}_{opt}} \times i/1000$, $i = 0, 1, \dots, 1000$, where $\hat{\Pi}_{\rho, \hat{R}_{opt}}$ denotes Necir et al. (2007) estimator of $\Pi_{\rho, R_{opt}}$ (see Section 2.1). The data tilting function $L(\Pi_{\rho, R_{opt}})$ is plotted for various $1 \leq \rho$ and for $k = 50, 100$; see Figure 6 (for Burr) and Figure 7 (for Pareto). These plots reveal that the function is convex, which implies that the resulting confidence region is indeed an interval. We observe that the confidence intervals are wider when $k = 50$. This may be due to an increasing variability caused by the smaller number of observations above the threshold $X_{n-k, n}$. We also observe that as ρ increases, the confidence intervals shift to the right on the x -axis (with an upper bound which may be quite large). This, again, is natural, since as risk aversion increases, one expects the premium to be larger.

Overall, our short simulation study indicates that the data tilting method should be preferred. Its coverage accuracy is generally close to the nominal level, and it has a smaller sensitivity to the choice of the sample fraction k than the likelihood ratio method. A short comparison of the length of the intervals provided by the three methods will be reported in the next section.

3.2. Application to Danish fire data

In this section, we illustrate our confidence intervals on the set of Danish fire data, which consists of $n = 2167$ fire insurance loss over one million Danish Krone, recorded between 1980 and 1990 (the loss include damage to buildings, furnishings, personal properties and loss of profits). These data have been analyzed by several authors (e.g., McNeil, 1997; Resnick, 1997; Peng and Qi, 2006) and are now available in the R package `evir` (Pfaff and McNeil, 2018). Using various methods (statistical tests and graphical devices), McNeil (1997) and Resnick (1997) have shown that the data can be considered as heavy-tailed. A Kolmogorov-Smirnov test for a heavy-tailed distribution was further developed by Koning and Peng (2008), and confirmed that a heavy-tailed distribution is appropriate for these data. Consensus values for the tail index estimate range between 1.2 and 1.4, as suggested by the Hill plot, see Figure 8. Here, we take $\hat{\alpha} = 1.3$. We consider various values for the risk aversion index ρ , with $1 \leq \rho < \hat{\alpha}$ (namely, $\rho = 1, 1.1, 1.2$).

For each ρ and each of the normal approximation, likelihood ratio and data tilting method, we plot the confidence intervals of level 0.90 for $\Pi_{\rho, R}$ against the sample fraction k (with $k = 20, 22, \dots, 250$), see Figure 9. On the same figure, we also plot the length of the intervals against the sample fraction. From these plots, we observe that the normal approximation based interval has consistently larger length than the intervals constructed from the likelihood ratio and data tilting methods. The intervals based on these latter two methods have roughly similar length, with slightly smaller length for the data tilting intervals when ρ increases. On Figure 9, we also observe that the level of the intervals starts to decrease around $k = 190$. However, when k increases, one may expect that the level of the premium increases. This pattern may be due to the bias increasing when a large value of k is used. One can check on the Hill plot of Figure 8 that after a stable phase in the estimation of α , the bias starts to increase around $k = 190$. For these reasons, we suggest to consider values of k smaller than 190 for the purpose of calculating confidence intervals for the premium, on this data set.

3.3. Application to Algerian car insurance claims

The data considered in this section consist of $n = 1009$ third-party liability insurance claims, where each claim corresponds to a car accident having occurred in Northern Algeria. The data were communicated to us by an Algerian insurance company. First, we investigate the heavy-tail hypothesis for these data. We use the Kolmogorov-Smirnov goodness-of-fit test proposed by Koning and Peng (2008). For every $k = 1, 2, \dots$ the test statistic is defined as:

$$\sup_{r>1} \sqrt{k} |1 - G_k(r) - r^{-1/\hat{\gamma}}| \quad (3.11)$$

where $1 - G_k(r) = \frac{1}{k} \sum_i 1_{\{X_i > rX_{n-k,n}\}}$ and $\hat{\gamma}$ is the Hill estimator (2.4). We calculate (3.11) for k ranging from 10 to 800, and we plot the values against k (see Figure 10). From Koning and Peng (2008), the limiting critical value with level 0.95 is 1.076. This value is represented as the horizontal line on Figure 10. Koning and Peng (2008) explain that "*Ideally, when the heavy tailed hypothesis is true, test should not reject the null hypothesis for small values of k , and reject it for large values of k since the critical values are obtained by ignoring the bias*". Deciding whether k is small or large is somewhat subjective, but Figure 10 seems to be consistent with the ideal pattern described by Koning and Peng (2008), which suggests that the heavy-tailed hypothesis holds for the Algerian car claim data. Indeed, for every k less than 400, the hypothesis of a heavy-tailed distribution is consistently accepted. When k is larger than 400, the hypothesis is consistently rejected. Next, we calculate and plot the Hill estimator $\hat{\alpha} = 1/\hat{\gamma}$ of α (see Figure 11). Note how the bias starts to dominate when $k \geq 400$, which indicates that we should consider values of k smaller than 400 for purpose of inference on α . For k less than 200 (approximately), Hill estimator is quite unstable. Therefore, we suggest to consider values of k in the neighborhood of 200 to estimate α , and we retain the value $\hat{\alpha} = 1.45$.

Now, we consider premium estimation for various values of $\rho < \hat{\alpha}$, namely, $\rho = 1.1, 1.2$. On Figure 12, we plot the confidence intervals of level 0.90 for $\Pi_{\rho,R}$, and their length, against the sample fraction k (with $k = 100, 102, \dots, 400$). Again, we observe that the normal approximation method yields larger intervals, while the likelihood ratio and data tilting methods, here, provide similar intervals.

4. Discussion

In this paper, we investigated several methods for constructing premium confidence intervals, namely, a normal approximation based method, a likelihood ratio method and a data tilting method. Our investigations suggest that for finite sample sizes, the normal approximation method performs rather poorly, both in terms of coverage probabilities and length of the intervals. On the other hand, based on our simulations, it appears that the likelihood ratio and data tilting methods provide good results, with a slight superiority for the data tilting method. Now, several problems may deserve further attention. One of these is the construction of premium confidence intervals based on censored data. Another interesting problem arises when the claim amounts are observed together with some additional information (in the form of covariates). In this case, it becomes of interest to estimate a conditional premium. All these problems are the topic for our future research.

Appendix : proofs

Proof of Theorem 2.2: We first prove three technical lemmas (in what follows, we will note R instead of R_{opt} in order to simplify the notations).

Lemma 1. Let $g : \mathbb{R} \rightarrow \mathbb{R}$ be defined as

$$g(\lambda) = \Pi_{\rho,R} \left(\frac{\hat{\alpha}/\rho}{1 - \frac{\lambda\hat{\alpha}}{\rho k}(\Pi_{\rho,R} - \log X_{n-k,n})} - 1 \right) - X_{n-k,n} \left(\frac{\rho k - \lambda}{\rho n - \lambda} \right)^{\frac{1}{\rho}}.$$

Then there exists a unique λ such that $g(\lambda) = 0$.

Proof. The derivative of g with respect to λ is

$$g'(\lambda) = \Pi_{\rho,R} \frac{\hat{\alpha}^{2\frac{1}{k}}(\Pi_{\rho,R} - \log X_{n-k,n})}{\rho^2 \left(1 - \frac{\lambda\hat{\alpha}}{\rho k}(\Pi_{\rho,R} - \log X_{n-k,n}) \right)^2} + X_{n-k,n} \left(\frac{\rho k - \lambda}{\rho n - \lambda} \right)^{\frac{1}{\rho}-1} \frac{(n-k)}{(\rho n - \lambda)^2}. \quad (4.12)$$

The quantity

$$\bar{c}(\lambda) = \frac{\rho k - \lambda}{\rho n - \lambda} X_{n-k,n}^{\bar{\alpha}(\lambda)}$$

must be positive, therefore, the constraint $\rho k - \lambda > 0$ (or equivalently, $\lambda < \rho k$) must apply. Under this constraint, the second term in the sum (4.12) is strictly positive. Now, consider the case where $\Pi_{\rho,R} - \log X_{n-k,n} > 0$ (the case where $\Pi_{\rho,R} - \log X_{n-k,n} \leq 0$ can be treated using similar arguments). Then $g'(\lambda) > 0$ and g is strictly increasing.

It is easy to see that

$$\bar{\alpha}(\lambda) = \frac{\hat{\alpha}}{1 - \frac{\lambda\hat{\alpha}}{\rho k}(\Pi_{\rho,R} - \log X_{n-k,n})},$$

and since $\bar{\alpha}(\lambda)$ must be positive, the constraint

$$\lambda < \frac{\rho k}{\hat{\alpha}(\Pi_{\rho,R} - \log X_{n-k,n})}$$

must also apply. Therefore, we must have

$$\lambda \in (-\infty, m),$$

where

$$m = \min \left(\rho k, \frac{\rho k}{\hat{\alpha}(\Pi_{\rho,R} - \log X_{n-k,n})} \right) > 0.$$

We have $g(-\infty) = -\Pi_{\rho,R} - X_{n-k,n} < 0$. If $m = \rho k / \{\hat{\alpha}(\Pi_{\rho,R} - \log X_{n-k,n})\}$, then $g(m-) = +\infty$. If $m = \rho k$, that is, if $\rho k \leq \rho k / \{\hat{\alpha}(\Pi_{\rho,R} - \log X_{n-k,n})\}$, it is easy to show that:

$$1 < \frac{1}{1 - \hat{\alpha}(\Pi_{\rho,R} - \log X_{n-k,n})},$$

and thus,

$$0 < \frac{\hat{\alpha}/\rho}{1 - \hat{\alpha}(\Pi_{\rho,R} - \log X_{n-k,n})} - 1.$$

This, finally, implies that if $m = \rho k$,

$$g(m) = \Pi_{\rho,R} \left(\frac{\hat{\alpha}/\rho}{1 - \hat{\alpha}(\Pi_{\rho,R} - \log X_{n-k,n})} - 1 \right) > 0.$$

In summary, g is continuous, strictly increasing, $g(-\infty) < 0$ and $g(m) > 0$, with $m > 0$. Therefore, there exists a unique $\lambda \in (-\infty, m)$ such that $g(\lambda) = 0$. ■

Lemma 2. *Let*

$$b_n = \frac{k\rho^2}{\Pi_{\rho,R}(\Pi_{\rho,R} - \log X_{n-k,n})\hat{\alpha}^2},$$

where $k = k(n)$ satisfies $k \rightarrow \infty$, $k/n \rightarrow 0$ and $\sqrt{k}/\log n \rightarrow 0$ as $n \rightarrow \infty$, and let λ_n denote the solution to (2.9). Then, as $n \rightarrow \infty$, $\mathbb{P}(|\lambda_n| < b_n) \rightarrow 1$ and $|\lambda_n|/\sqrt{k} \xrightarrow{\mathbb{P}} 0$.

Proof. If λ_n is the solution to (2.9), then $g(\lambda_n) = 0$ (existence of λ_n is ensured by Lemma 1). We wish to show that $\mathbb{P}(|\lambda_n| < b_n) \rightarrow 1$ as $n \rightarrow \infty$. We have $\mathbb{P}(|\lambda_n| < b_n) = \mathbb{P}(-b_n < \lambda_n < b_n) = \mathbb{P}(g(-b_n) < 0 < g(b_n))$ since g is strictly increasing. To prove that $\mathbb{P}(g(-b_n) < 0 < g(b_n)) \rightarrow 1$, we prove that $\mathbb{P}(g(b_n) > 0) \rightarrow 1$ and $\mathbb{P}(g(-b_n) < 0) \rightarrow 1$.

First, note that:

$$\begin{aligned} \frac{1 - \frac{b_n}{\rho k}}{1 - \frac{b_n}{\rho n}} &= \frac{1}{1 - \frac{k}{n}\rho \frac{1}{\Pi_{\rho,R}(\Pi_{\rho,R} - \log X_{n-k,n})\hat{\alpha}^2}} - \frac{\rho}{\Pi_{\rho,R}(\Pi_{\rho,R} - \log X_{n-k,n})\hat{\alpha}^2 - \frac{k}{n}\rho} \\ &= 1 + o_{\mathbb{P}}(1). \end{aligned}$$

Therefore,

$$\begin{aligned} X_{n-k,n} \left(\frac{\rho k - b_n}{\rho n - b_n} \right)^{\frac{1}{\rho}} &= X_{n-k,n} \left(\frac{k}{n} \right)^{\frac{1}{\rho}} \left(\frac{1 - \frac{b_n}{\rho k}}{1 - \frac{b_n}{\rho n}} \right)^{\frac{1}{\rho}} \\ &= X_{n-k,n} \left(\frac{k}{n} \right)^{\frac{1}{\rho}} (1 + o_{\mathbb{P}}(1)) \\ &= \hat{\Pi}_{\rho, \hat{R}_{opt}} \left(\frac{\hat{\alpha}}{\rho} - 1 \right) + o_{\mathbb{P}}(1) \\ &= \Pi_{\rho,R} \left(\frac{\alpha}{\rho} - 1 \right) + o_{\mathbb{P}}(1), \end{aligned} \tag{4.13}$$

since $\hat{\Pi}_{\rho, \hat{R}_{opt}} - \Pi_{\rho,R} = o_{\mathbb{P}}(1)$ (Necir et al., 2007) and $\hat{\alpha} - \alpha = o_{\mathbb{P}}(1)$ (Mason, 1982). Now, using (4.13) and Taylor's expansion in

$$g(b_n) = \frac{\Pi_{\rho,R}\hat{\alpha}}{\rho} \left(\frac{1}{1 - \frac{b_n\hat{\alpha}}{\rho k}(\Pi_{\rho,R} - \log X_{n-k,n})} \right) - \Pi_{\rho,R} - X_{n-k,n} \left(\frac{\rho k - b_n}{\rho n - b_n} \right)^{\frac{1}{\rho}},$$

we obtain:

$$\begin{aligned} g(b_n) &= \frac{\Pi_{\rho,R}\hat{\alpha}}{\rho} \left(1 + \frac{b_n\hat{\alpha}}{\rho k}(\Pi_{\rho,R} - \log X_{n-k,n})(1 + o_{\mathbb{P}}(1)) \right) - \Pi_{\rho,R} - \Pi_{\rho,R} \left(\frac{\alpha}{\rho} - 1 \right) + o_{\mathbb{P}}(1) \\ &= \frac{\Pi_{\rho,R}\hat{\alpha}}{\rho} + \frac{\Pi_{\rho,R}b_n\hat{\alpha}^2}{\rho^2 k}(\Pi_{\rho,R} - \log X_{n-k,n})(1 + o_{\mathbb{P}}(1)) - \frac{\Pi_{\rho,R}\alpha}{\rho} + o_{\mathbb{P}}(1) \\ &= 1 + o_{\mathbb{P}}(1). \end{aligned}$$

Therefore, $\mathbb{P}(g(b_n) > 0) \rightarrow 1$ as $n \rightarrow \infty$. Similarly, we can show that $g(-b_n) = -1 + o_{\mathbb{P}}(1)$, hence $\mathbb{P}(g(-b_n) < 0) \rightarrow 1$ as $n \rightarrow \infty$, which concludes the first part of the proof.

Next, we have:

$$\begin{aligned}\mathbb{P}(|\lambda_n| < b_n) &= \mathbb{P}\left(\frac{|\lambda_n|}{\sqrt{k}} < \sqrt{k} \frac{\rho^2}{\Pi_{\rho,R}(\Pi_{\rho,R} - \log X_{n-k,n})\hat{\alpha}^2}\right) \\ &= \mathbb{P}\left(\frac{|\lambda_n|}{\sqrt{k}} < \frac{\rho^2}{\Pi_{\rho,R}\hat{\alpha}^2} \frac{1}{\frac{\Pi_{\rho,R}}{\sqrt{k}} - \frac{\log X_{n-k,n}}{\sqrt{k}}}\right).\end{aligned}$$

From Theorem 2 of Necir et al. (2007), if $\sqrt{k}/\log n \rightarrow 0$, then $\log X_{n-k,n}/\sqrt{k} \xrightarrow{\mathbb{P}} \infty$. Therefore,

$$\frac{\rho^2}{\Pi_{\rho,R}\hat{\alpha}^2} \frac{1}{\frac{\Pi_{\rho,R}}{\sqrt{k}} - \frac{\log X_{n-k,n}}{\sqrt{k}}} \xrightarrow{\mathbb{P}} 0$$

and since $\mathbb{P}(|\lambda_n| < b_n) \rightarrow 1$, we conclude that $|\lambda_n|/\sqrt{k} \xrightarrow{\mathbb{P}} 0$. ■

Lemma 3. *Let λ_n denote the solution to (2.9), and let $k = k(n)$ be such that $k \rightarrow \infty$, $k/n \rightarrow 0$ and $\sqrt{k}/\log n \rightarrow 0$ as $n \rightarrow \infty$. Then, as $n \rightarrow \infty$, $\bar{\alpha}(\lambda_n)/\hat{\alpha} \xrightarrow{\mathbb{P}} 1$.*

Proof. First, using Taylor's expansion, we obtain

$$\begin{aligned}g(\lambda_n) &= \frac{\Pi_{\rho,R}\hat{\alpha}}{\rho} + \frac{\Pi_{\rho,R}\hat{\alpha}^2}{\rho^2 k} \lambda_n (\Pi_{\rho,R} - \log X_{n-k,n})(1 + o_{\mathbb{P}}(1)) - \frac{\Pi_{\rho,R}\alpha}{\rho} + o_{\mathbb{P}}(1) \\ &= \frac{\Pi_{\rho,R}}{\rho} (\hat{\alpha} - \alpha) + \frac{\Pi_{\rho,R}\hat{\alpha}^2}{\rho^2 k} \lambda_n (\Pi_{\rho,R} - \log X_{n-k,n})(1 + o_{\mathbb{P}}(1)).\end{aligned}$$

If λ_n is the solution to (2.9), then $g(\lambda_n) = 0$, and thus,

$$\frac{\lambda_n}{k} = -\rho(\hat{\alpha} - \alpha) \frac{1}{(\Pi_{\rho,R} - \log X_{n-k,n})\hat{\alpha}^2} (1 + o_{\mathbb{P}}(1)).$$

Now, we have:

$$\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} - 1 = \frac{1}{1 - \frac{\lambda_n \hat{\alpha}}{\rho k} (\Pi_{\rho,R} - \log X_{n-k,n})} - 1,$$

and by Taylor's expansion,

$$\begin{aligned}\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} - 1 &= \frac{\lambda_n \hat{\alpha}}{\rho k} (\Pi_{\rho,R} - \log X_{n-k,n})(1 + o_{\mathbb{P}}(1)) \\ &= -\rho(\hat{\alpha} - \alpha) \frac{1}{(\Pi_{\rho,R} - \log X_{n-k,n})\hat{\alpha}^2} \frac{\hat{\alpha}}{\rho} (\Pi_{\rho,R} - \log X_{n-k,n})(1 + o_{\mathbb{P}}(1)) \\ &= -\frac{(\hat{\alpha} - \alpha)}{\alpha} (1 + o_{\mathbb{P}}(1)).\end{aligned}$$

Therefore, by consistency of $\hat{\alpha}$, we obtain that $\bar{\alpha}(\lambda_n)/\hat{\alpha} - 1 \xrightarrow{\mathbb{P}} 0$, which concludes the proof. ■

We are now in position to prove Theorem 2.2. First, recall that

$$\log L(\alpha, c) = k \log \alpha + k \log c - (\alpha + 1) \sum_{i=1}^k \log X_{n-i+1,n} + (n - k) \log(1 - cX_{n-k,n}^{-\alpha}).$$

Using the fact that $\hat{\alpha}^{-1} = k^{-1} \sum_{i=1}^k \log X_{n-i+1,n} - \log X_{n-k,n}$, we can re-write $\log L(\alpha, c)$ as

$$\log L(\alpha, c) = k \log \alpha - k(\alpha + 1)\hat{\alpha}^{-1} + k \log(cX_{n-k,n}^{-\alpha}) - k \log X_{n-k,n} + (n - k) \log(1 - cX_{n-k,n}^{-\alpha}),$$

and thus,

$$\begin{aligned} \ell(\Pi_{\rho,R}) &= -2 \{ \log L(\bar{\alpha}(\lambda_n), \bar{c}(\lambda_n)) - \log L(\hat{\alpha}, \hat{c}) \} \\ &= -2 \left\{ k \log \left(\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} \right) - k(\bar{\alpha}(\lambda_n) - \hat{\alpha})\hat{\alpha}^{-1} + k \log \left(\frac{\bar{c}(\lambda_n)X_{n-k,n}^{-\bar{\alpha}(\lambda_n)}}{\hat{c}X_{n-k,n}^{-\hat{\alpha}}} \right) \right. \\ &\quad \left. + (n - k) \log \left(\frac{1 - \bar{c}(\lambda_n)X_{n-k,n}^{-\bar{\alpha}(\lambda_n)}}{1 - \hat{c}X_{n-k,n}^{-\hat{\alpha}}} \right) \right\}. \end{aligned}$$

Now, using $\hat{c} = (k/n)X_{n-k,n}^{\hat{\alpha}}$ and $\bar{c}(\lambda_n) = \frac{\rho k - \lambda_n}{\rho n - \lambda_n} X_{n-k,n}^{\bar{\alpha}(\lambda_n)}$, we can re-write $\ell(\Pi_{\rho,R})$ as

$$\ell(\Pi_{\rho,R}) = -2 \left\{ k \log \left(\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} \right) - k \left(\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} - 1 \right) + k \log \left(\frac{\rho k - \lambda_n}{\rho n - \lambda_n} \cdot \frac{n}{k} \right) + (n - k) \log \left(\frac{\rho n - \rho k}{\rho n - \lambda_n} \cdot \frac{n}{n - k} \right) \right\},$$

and straightforward calculations yield:

$$\ell(\Pi_{\rho,R}) = -2k \left\{ \log \left(\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} \right) - \left(\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} - 1 \right) \right\} - 2k \log \left(1 - \frac{\lambda_n}{\rho k} \right) + 2n \log \left(1 - \frac{\lambda_n}{\rho n} \right).$$

By Lemma 2, $|\lambda_n|/\sqrt{k} \xrightarrow{\mathbb{P}} 0$ (which also implies $|\lambda_n|/k \xrightarrow{\mathbb{P}} 0$ and $|\lambda_n|/n \xrightarrow{\mathbb{P}} 0$), and by Lemma 3, $\bar{\alpha}(\lambda_n)/\hat{\alpha} \xrightarrow{\mathbb{P}} 1$. Thus, by Taylor's expansions, we obtain

$$\begin{aligned} \ell(\Pi_{\rho,R}) &= -2k \left\{ \log \left(\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} \right) - \left(\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} - 1 \right) \right\} + O_{\mathbb{P}} \left(\frac{\lambda_n^2}{k} \right) \\ &= \left(\sqrt{k} \left(\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} - 1 \right) \right)^2 + o_{\mathbb{P}}(1). \end{aligned}$$

Now, according to the proof of Lemma 3, we have

$$\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} - 1 = -\frac{(\hat{\alpha} - \alpha)}{\alpha} (1 + o_{\mathbb{P}}(1)),$$

therefore,

$$\left(\sqrt{k} \left(\frac{\bar{\alpha}(\lambda_n)}{\hat{\alpha}} - 1 \right) \right)^2 = \left(\sqrt{k} \left(\frac{\hat{\gamma} - \gamma}{\gamma} \right) \right)^2 (1 + o_{\mathbb{P}}(1)),$$

Finally, if (2.5) holds and $\sqrt{k}A(n/k) \rightarrow 0$ as $n \rightarrow \infty$, $\sqrt{k}(\hat{\gamma} - \gamma)/\gamma \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ (see de Haan and Peng, 1998). Therefore,

$$\ell(\Pi_{\rho,R}) \xrightarrow{\mathcal{L}} \chi_1^2.$$

■

Proof of Theorem 2.3: The proof follows the same lines as Peng and Qi (2006), therefore, we provide an outline of the proof and emphasize the parts that are specific to our problem.

For $i = 1, \dots, n$, let $\delta_i = 1_{\{X_i > X_{n-k,n}\}}$ and let $q_{(i)}$ denote the concomitant of $X_{n-i+1,n}$, that is,

$$q_{(i)} = \frac{1}{n} \exp \left(-1 - \lambda_1 - \lambda_2 \left[\frac{\Pi_{\rho, R_{opt}}}{\rho A_2(\lambda_1)} - \frac{\Pi_{\rho, R_{opt}} A_1(\lambda_1) \log(X_{n-i+1,n}/X_{n-k,n})}{\rho (A_2(\lambda_1))^2} - \frac{X_{n-k,n} (A_1(\lambda_1))^{\frac{1}{\rho}-1}}{\rho} \right] \right)$$

for $i = 1, \dots, k$ and $q_{(i)} = \frac{1}{n} e^{-1-\lambda_1}$ for $i = k+1, \dots, n$. Since

$$1 = \sum_{i=1}^n q_i = \sum_{i=1}^n \delta_i q_i + \sum_{i=1}^n (1 - \delta_i) q_i,$$

and $\sum_{i=1}^n \delta_i q_i = \sum_{i=1}^k q_{(i)}$, we have

$$\begin{aligned} \sum_{i=1}^k q_{(i)} &= 1 - \frac{(n-k)}{n} e^{-1-\lambda_1} \\ &= A_1(\lambda_1). \end{aligned}$$

Now, consider the second constraint in (2.10):

$$\Pi_{\rho, R_{opt}} \left(\frac{\hat{\alpha}(q)}{\rho} - 1 \right) - X_{n-k,n} \left(\sum_{i=1}^n q_i \delta_i \right)^{\frac{1}{\rho}} = 0.$$

This constraint can be rewritten as:

$$\hat{\alpha}(q) = \frac{\rho \left(\Pi_{\rho, R_{opt}} + X_{n-k,n} (A_1(\lambda_1))^{\frac{1}{\rho}} \right)}{\Pi_{\rho, R_{opt}}}.$$

Using this, we can write:

$$\begin{aligned} A_2(\lambda_1) &= \frac{A_1(\lambda_1)}{\hat{\alpha}(q)} \\ &= \frac{\sum_{i=1}^n q_i \delta_i}{\hat{\alpha}(q)} \\ &= \sum_{i=1}^n q_i \delta_i (\log X_i - \log X_{n-k,n}) \\ &= \sum_{i=1}^k q_{(i)} \log(X_{n-i+1,n}/X_{n-k,n}), \end{aligned}$$

where the second to third line follows from the expression of $\hat{\alpha}(q)$. Therefore, the constraints in (2.10) are equivalent to:

$$A_1(\lambda_1) = \sum_{i=1}^k q_{(i)} \quad \text{and} \quad A_2(\lambda_1) = \sum_{i=1}^k q_{(i)} Z_i,$$

where $Z_i := \log(X_{n-i+1,n}/X_{n-k,n})$, or also, to:

$$\sum_{i=1}^k q(i) = A_1(\lambda_1) \quad \text{and} \quad \frac{\sum_{i=1}^k q(i) Z_i}{\sum_{i=1}^k q(i)} = \frac{A_2(\lambda_1)}{A_1(\lambda_1)}. \quad (4.14)$$

Now, by using the expression of $q(i)$, the second equality in (4.14) can be rewritten as:

$$\frac{\sum_{i=1}^k \exp\left(\lambda_2 \frac{\Pi_{\rho, R_{opt}} A_1(\lambda_1) Z_i}{\rho(A_2(\lambda_1))^2}\right) Z_i}{\sum_{i=1}^k \exp\left(\lambda_2 \frac{\Pi_{\rho, R_{opt}} A_1(\lambda_1) Z_i}{\rho(A_2(\lambda_1))^2}\right)} = \frac{\Pi_{\rho, R_{opt}}}{\rho\left(\Pi_{\rho, R_{opt}} + X_{n-k,n}(A_1(\lambda_1))^{\frac{1}{\rho}}\right)}. \quad (4.15)$$

The existence of a solution (λ_1, λ_2) to (2.10) will be established by proving that : *i*) for every λ_1 , there exists $\lambda_2(\lambda_1)$ such that $(\lambda_1, \lambda_2(\lambda_1))$ satisfies (4.15), *ii*) there exists some λ_1 such that $(\lambda_1, \lambda_2(\lambda_1))$ also satisfies $\sum_{i=1}^k q(i) = A_1(\lambda_1)$.

Proof of *i*): let

$$f(\lambda) = \frac{\sum_{i=1}^k Z_i \exp(\lambda Z_i)}{\sum_{i=1}^k \exp(\lambda Z_i)}.$$

One can check that $f(\lambda) \xrightarrow{\lambda \rightarrow -\infty} Z_k$, $f(\lambda) \xrightarrow{\lambda \rightarrow +\infty} Z_1$ and f is increasing and continuous on $[Z_k, Z_1]$. Therefore, f is invertible on $[Z_k, Z_1]$, with inverse f^{-1} . Now, let $\lambda_2(\lambda_1)$ be defined (uniquely, for a given λ_1) as:

$$\lambda_2(\lambda_1) = f^{-1}\left(\frac{\Pi_{\rho, R_{opt}}}{\rho\left(\Pi_{\rho, R_{opt}} + X_{n-k,n}(A_1(\lambda_1))^{\frac{1}{\rho}}\right)}\right) \frac{\rho(A_2(\lambda_1))^2}{\Pi_{\rho, R_{opt}} A_1(\lambda_1)}. \quad (4.16)$$

It is not difficult to see that $\lambda_2(\lambda_1)$ satisfies (4.15), and thus, $(\lambda_1, \lambda_2(\lambda_1))$ is solution to the constraint $\frac{\sum_{i=1}^k q(i) Z_i}{\sum_{i=1}^k q(i)} = \frac{A_2(\lambda_1)}{A_1(\lambda_1)}$.

Thereafter, it will be convenient, for notational simplicity, to re-express $\lambda_2(\lambda_1)$ as:

$$\lambda_2(\lambda_1) = \tilde{\lambda} \times \frac{\rho(A_2(\lambda_1))^2}{\Pi_{\rho, R_{opt}} A_1(\lambda_1)},$$

with $\tilde{\lambda} \equiv \tilde{\lambda}(\lambda_1) = f^{-1}\left(\frac{\Pi_{\rho, R_{opt}}}{\rho(\Pi_{\rho, R_{opt}} + X_{n-k,n}(A_1(\lambda_1))^{\frac{1}{\rho}})}\right)$.

Proof of *ii*): first, by Taylor's expansion, we have, as in Peng and Qi (2006), $f(\pm k^{-1/4}) - f(0) = \pm k^{-1/4} f'(0)(1 + o_{\mathbb{P}}(1))$, with $f(0) = \sum_{i=1}^k Z_i / \sum_{i=1}^k 1 = \hat{\gamma}$ and

$$f'(0) = \frac{1}{k} \sum_{i=1}^k Z_i^2 - \hat{\gamma}^2 = \gamma^2(1 + O_{\mathbb{P}}(k^{-1/2})). \quad (4.17)$$

Using this, we can show that $\mathbb{P}(-k^{-1/4} < \tilde{\lambda} < k^{-1/4}) \rightarrow 1$ as $n \rightarrow \infty$. A second Taylor's expansion yields

$$f(\tilde{\lambda}) - \hat{\gamma} = \tilde{\lambda} f'(0)(1 + O_{\mathbb{P}}(k^{-1/4})) = \tilde{\lambda} \gamma^2(1 + O_{\mathbb{P}}(k^{-1/4})), \quad (4.18)$$

which implies $\tilde{\lambda} = O_{\mathbb{P}}(k^{-1/2})$. Finally, using Taylor's expansion, we have

$$\sum_{i=1}^k \exp(\tilde{\lambda} Z_i) = k \left(1 + O_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right) \right). \quad (4.19)$$

Now, let

$$h(\lambda_1) = \sum_{i=1}^k q(i) - A_1(\lambda_1).$$

By plugging $\lambda_2(\lambda_1)$ into $q(i)$, we can re-express $h(\lambda_1)$ as:

$$h(\lambda_1) = \frac{1}{n} e^{-1-\lambda_1} \exp \left(-\tilde{\lambda} \frac{\rho(A_2(\lambda_1))^2}{\Pi_{\rho, R_{opt}} A_1(\lambda_1)} \left[\frac{\Pi_{\rho, R_{opt}}}{\rho A_2(\lambda_1)} - \frac{X_{n-k, n}(A_1(\lambda_1))^{\frac{1}{\rho}-1}}{\rho} \right] \right) \sum_{i=1}^k \exp(\tilde{\lambda} Z_i) - A_1(\lambda_1).$$

Let $\lambda'_1 = -1 - \log \left(1 + \frac{\sqrt{k}}{n-k} \right)$ and $\lambda''_1 = -1 - \log \left(1 - \frac{\sqrt{k}}{n-k} \right)$. Then it is easy to see that

$$e^{-1-\lambda'_1} = 1 + \frac{\sqrt{k}}{n-k} \quad \text{and} \quad A_1(\lambda'_1) = \frac{k}{n} \left(1 - \frac{1}{\sqrt{k}} \right). \quad (4.20)$$

We also have

$$A_2(\lambda'_1) = \gamma \frac{k}{n} \left(1 - \frac{1}{\sqrt{k}} + o_{\mathbb{P}} \left(\frac{1}{\sqrt{k}} \right) \right), \quad (4.21)$$

which follows from (2.3), from the fact that $\frac{k}{n} \left(1 - \frac{1}{\sqrt{k}} \right) \leq A_1(\lambda) \leq \frac{k}{n} \left(1 + \frac{1}{\sqrt{k}} \right)$ for every $\lambda'_1 \leq \lambda \leq \lambda''_1$, and from the following result for the Hill estimate: $\sqrt{k}(\hat{\gamma} - \gamma) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma^2)$.

Now, it follows from (4.19), (4.20) and (4.21) that $h(\lambda'_1) = \frac{k}{n} \frac{1}{\sqrt{k}} (1 + o_{\mathbb{P}}(1))$. Similarly, $h(\lambda''_1) = -\frac{k}{n} \frac{1}{\sqrt{k}} (1 + o_{\mathbb{P}}(1))$. Since $\lambda'_1 \leq \lambda_1 \leq \lambda''_1$ and $h(\lambda'_1)$ and $h(\lambda''_1)$ can be made arbitrarily close to 0 by letting $n \rightarrow \infty$, λ_1 satisfies $h(\lambda_1) = 0$, or equivalently, $\sum_{i=1}^k q(i) = A_1(\lambda_1)$, with $\lambda_2(\lambda_1)$ defined in (4.16). That is, we have shown that there exists a solution $(\lambda_1, \lambda_2(\lambda_1))$ of (4.14), or equivalently, of (2.10).

Now, let $\hat{\lambda}_1$ be such that $h(\hat{\lambda}_1) = 0$. By using similar arguments as Peng and Qi (2006), we can show that

$$1 + \hat{\lambda}_1 = O_{\mathbb{P}} \left(\frac{\sqrt{k}}{n} \right). \quad (4.22)$$

Moreover, from (4.18), we can show that

$$\tilde{\lambda}^2 = \left(\frac{\hat{\gamma} - \gamma}{\gamma^2} \right)^2 (1 + O_{\mathbb{P}}(k^{-1/4})). \quad (4.23)$$

As in Peng and Qi (2006), we have

$$L(\Pi_{\rho, R_{opt}}) = (1 + o_{\mathbb{P}}(1)) \left(n(1 + \hat{\lambda}_1)^2 + \tilde{\lambda}^2 \sum_{i=1}^k (Z_i - \hat{\gamma})^2 \right).$$

It follows from (4.17), (4.22) and (4.23) that

$$\begin{aligned} L(\Pi_{\rho, R_{opt}}) &= (1 + o_{\mathbb{P}}(1))O_{\mathbb{P}}(k/n) + (1 + o_{\mathbb{P}}(1))k\gamma^2(1 + O_{\mathbb{P}}(k^{-1/2}))\tilde{\lambda}^2 \\ &= o_{\mathbb{P}}(1) + (1 + o_{\mathbb{P}}(1)) \left(\sqrt{k} \left(\frac{\hat{\gamma} - \gamma}{\gamma} \right) \right)^2. \end{aligned}$$

Under the conditions of Theorem 2.2, $\sqrt{k}(\hat{\gamma} - \gamma)/\gamma \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$. Therefore,

$$L(\Pi_{\rho, R_{opt}}) \xrightarrow{\mathcal{L}} \chi_1^2.$$

■

References

- Ahmedou, S., Deme, E.H., Fofana, S. 2023. An improved estimator of distortion risk premiums under dependence insured risks with heavy-tailed marginals. *Far East Journal of Theoretical Statistics* 67(3), 243-277.
- Benkhelifa, L. 2014. Kernel-type estimator of the reinsurance premium for heavy-tailed loss distributions. *Insurance: Mathematics and Economics* 59, 65-70.
- Centeno, M.L., Andrade e Silva, J. 2005. Applying the Proportional Hazard Premium calculation principle. *Astin Bulletin* 35(2), 409-425.
- Chan, N.H., Deng, S.-J., Peng, L., Xia, Z. 2007. Interval estimation of value-at-risk based on GARCH models with heavy-tailed innovations. *Journal of Econometrics* 137(2), 556-576.
- de Haan, L., Ferreira, A. 2006. *Extreme Value Theory: An Introduction*. Springer, New York.
- de Haan, L., Peng, L. 1998. Comparison of tail index estimators. *Statistica Neerlandica* 52(1), 60-70.
- Deme, E. H., Girard, S., Guillou, A. 2013. Reduced-bias estimator of the Proportional Hazard Premium for heavy-tailed distributions. *Insurance: Mathematics and Economics* 52(3), 550-559.
- Hall, P., Yao, Q. 2003. Data tilting for time series. *Journal of the Royal Statistical Society Series B* 65(2), 425-442.
- Hill, B.M. 1975. A simple approach to inference about the tail of a distribution. *Annals of Statistics* 3, 1136-1174.
- Koning, A.J., Peng, L., 2008. Goodness-of-fit tests for a heavy tailed distribution. *Journal of Statistical Planning and Inference* 138(12), 3960-3981.
- Li, Y., Qi, Y. 2019. Adjusted empirical likelihood method for the tail index of a heavy-tailed distribution. *Statistics & Probability Letters* 152, 50-58.
- Lu, J.C., Peng, L. 2002. Likelihood Based Confidence Intervals for the Tail Index. *Extremes* 5, 337-352.

- Mason, D.M. 1982. Laws of large numbers for sums of extreme values. *The Annals of Probability* 10(3), 754-764.
- McNeil, A.J. 1997. Estimating the tails of loss severity distributions using extreme value theory. *Astin Bulletin* 27(1), 117-137.
- Necir, A., Boukhetala, K. 2004. Estimating the risk adjusted premium of the largest reinsurance covers. In: Antoch, Jaromir (Ed.), *Proceeding of Computational Statistics*. Physica-Verlag, Springer, 1577-1584.
- Necir, A., Brahim, B., Meraghni, D. 2010. Erratum to: 'Statistical estimate of the proportional hazard premium of loss'. *Scandinavian Actuarial Journal* 3, 246-247.
- Necir, A., Meraghni, D. 2009. Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts. *Insurance: Mathematics and Economics* 45(1), 49-58.
- Necir, A., Meraghni, D., Meddi, F. 2007. Statistical estimate of the proportional hazard premium of loss. *Scandinavian Actuarial Journal* 3, 147-161.
- Owen, A.B. 2001. *Empirical Likelihood*. Chapman and Hall/CRC.
- Peng, L., Qi, Y., 2006. Confidence regions for high quantiles of a heavy tailed distribution. *The Annals of Statistics* 34(4), 1964-1986.
- Pfaff, B., McNeil, A. 2018. *evir: Extreme Values in R*. R package version 1.7-4. <https://CRAN.R-project.org/package=evir>
- Qi, Y., Xie, M., Yang, J. 2023. Inference of high quantiles of a heavy-tailed distribution from block data. *Statistics* 57(4), 918-940.
- Rassoul, A. 2012a. Estimating of the proportional hazard premium for heavy-tailed claim amounts with the POT method. *REVSTAT - Statistical Journal* 10(3), 335-349.
- Rassoul, A. 2012b. Reduced bias estimation of the reinsurance premium of loss distribution. *Journal of Statistics Applications & Probability* 1(2), 147-155.
- Resnick, S.I. 1997. Discussion of the Danish Data on Large Fire Insurance Losses. *Astin Bulletin* 27(1), 139-151.
- Tursunaliyeva, A., Silvapulle, P. 2016. Nonparametric estimation of operational value-at-risk (Op-VaR). *Insurance: Mathematics and Economics* 69, 194-201.
- Vandewalle, B., Beirlant, J., 2006. On univariate extreme value statistics and the estimation of reinsurance premiums. *Insurance: Mathematics and Economics* 38(3), 441-459.
- Wang, S. 1995. Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics* 17(1), 43-54.
- Wang, S. 1996. Premium calculation by transforming the layer premium density. *Astin Bulletin* 26(1), 71-92.

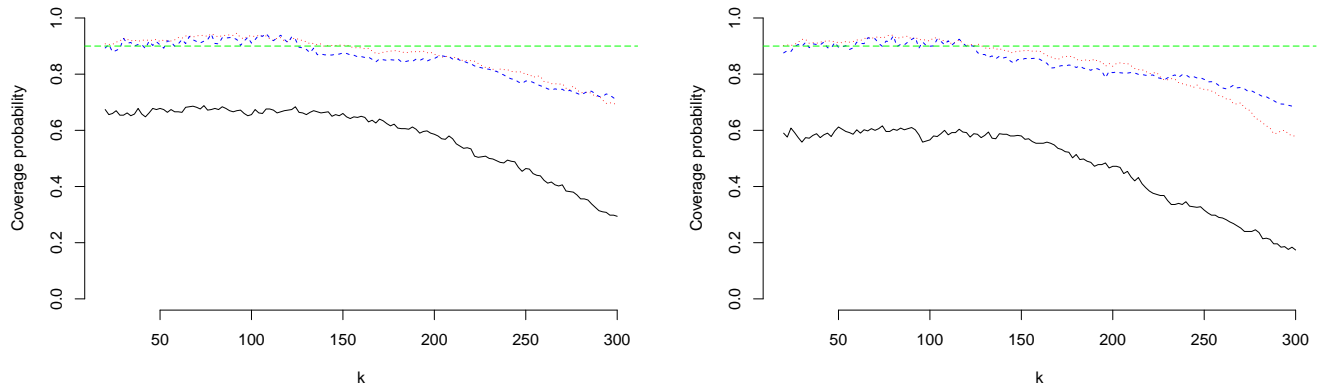


Figure 1: Coverage probabilities for Burr(2,4), plotted against $k = 20, 22, \dots, 300$ (based on $N = 1000$ samples of size $n = 1000$), with $\rho = 1.1$ (left) and $\rho = 1.2$ (right). Red dotted line: likelihood ratio method, black line: normal approximation method, blue dashed line: data tilting method. Horizontal green line is the nominal confidence level (0.9).

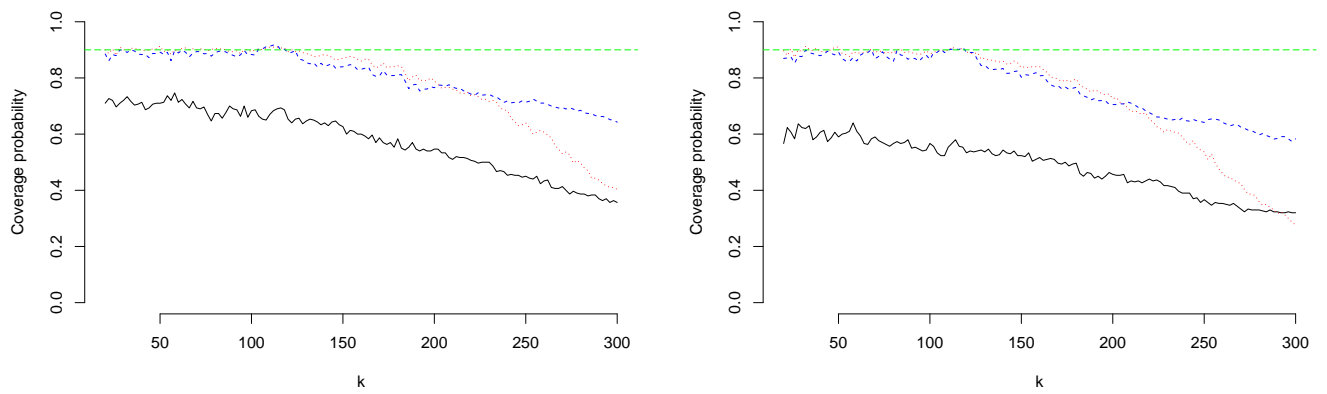


Figure 2: Coverage probabilities for Burr(1.5,3), plotted against $k = 20, 22, \dots, 300$ (based on $N = 1000$ samples of size $n = 1000$), with $\rho = 1.1$ (left) and $\rho = 1.2$ (right). Red dotted line: likelihood ratio method, black line: normal approximation method, blue dashed line: data tilting method. Horizontal green line is the nominal confidence level (0.9).

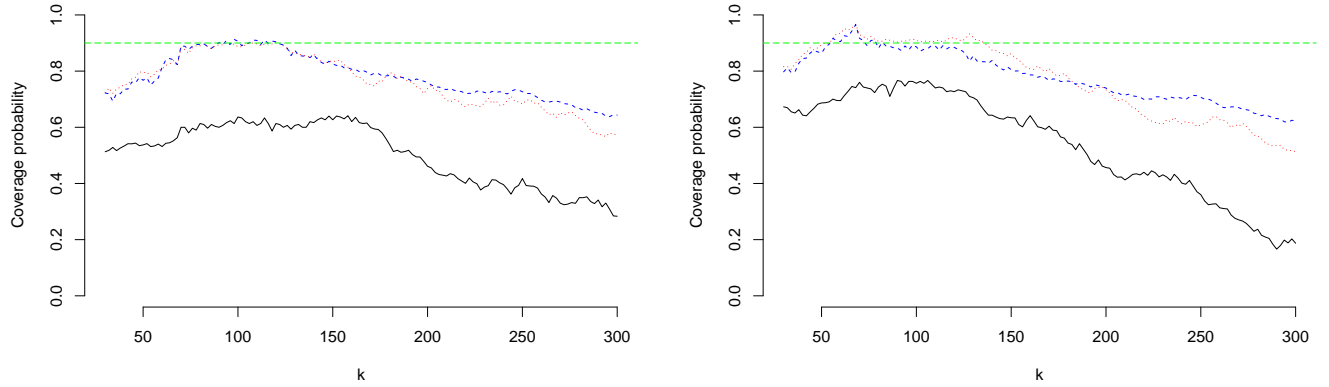


Figure 3: Coverage probabilities for Pareto(2), plotted against $k = 20, 22, \dots, 300$ (based on $N = 1000$ samples of size $n = 1000$), with $\rho = 1.2$ (left) and $\rho = 1.5$ (right). Red dotted line: likelihood ratio method, black line: normal approximation method, blue dashed line: data tilting method. Horizontal green line is the nominal confidence level (0.9).

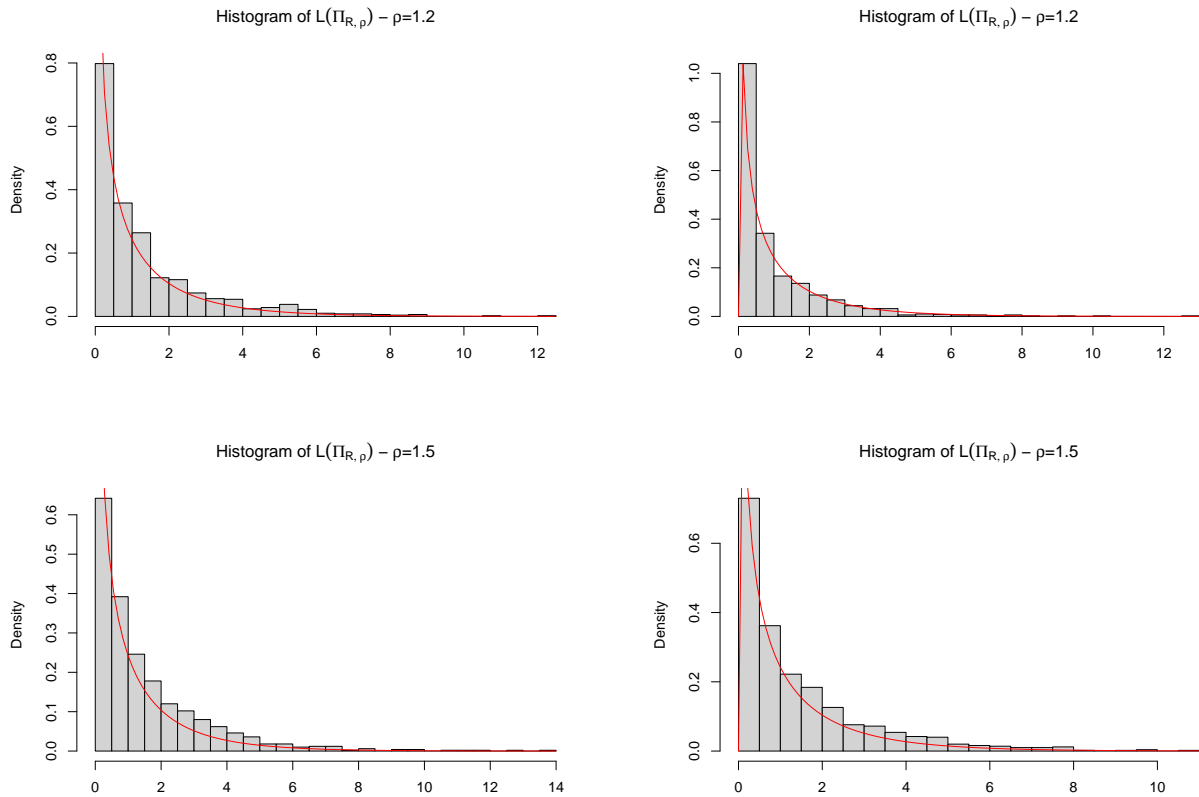


Figure 4: Histograms of the N realizations of $L(\Pi_{\rho, R_{opt}})$, with samples distributed as Burr(2, 4) (left: $k = 50$, right: $k = 100$). Red line : density of the χ_1^2 .

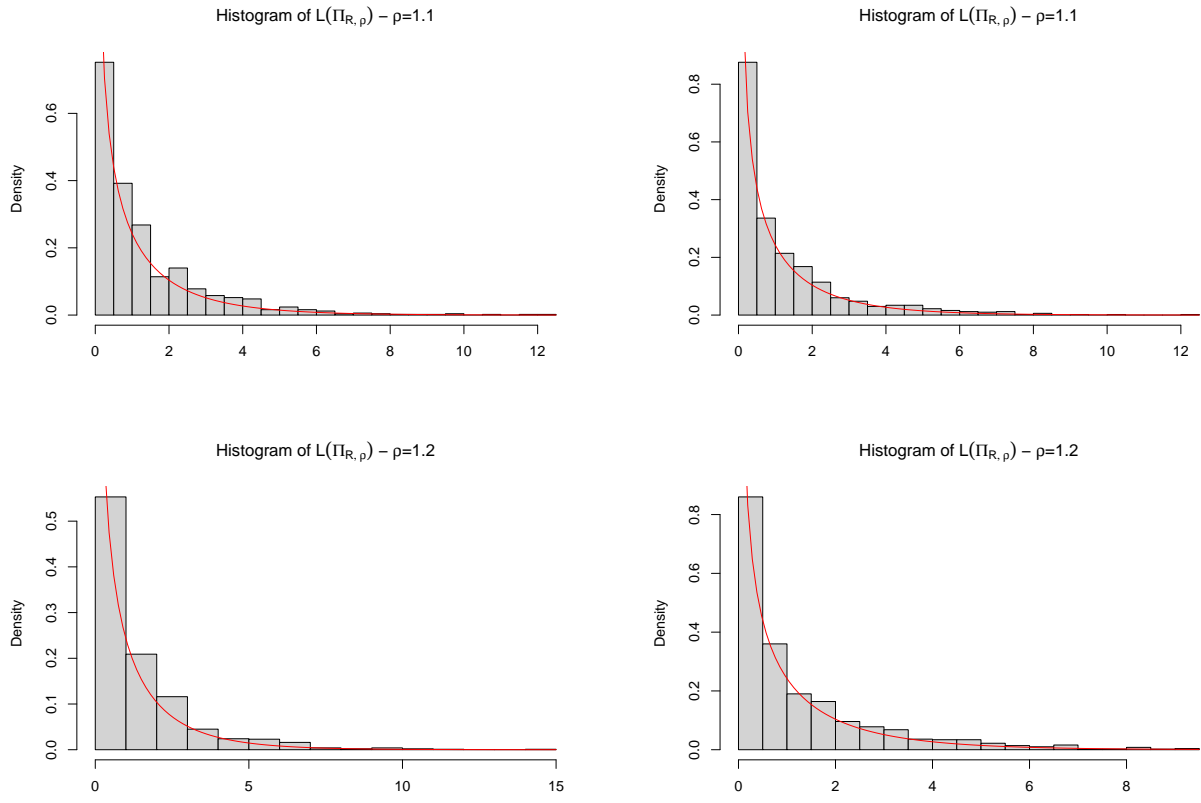


Figure 5: Histograms of the N realizations of $L(\Pi_{\rho,R_{opt}})$, with samples distributed as Pareto(1.5) (left: $k = 50$, right: $k = 100$). Red line : density of the χ_1^2 .

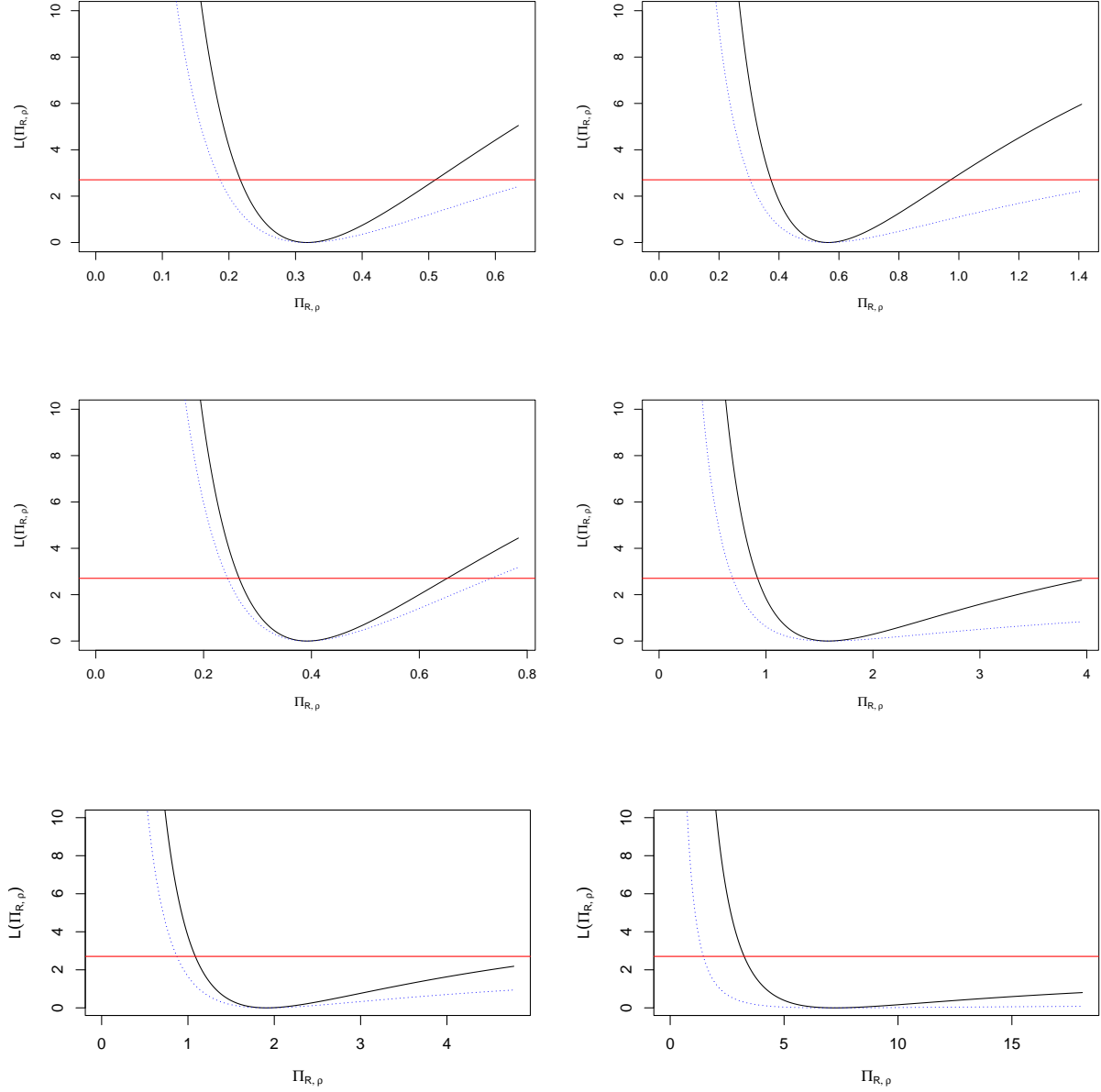


Figure 6: **Burr distribution.** Left column: $(\alpha, \beta) = (2, 4)$ and $\rho = 1$ (top), $\rho = 1.2$ (middle), $\rho = 1.5$ (bottom). Right column: $(\alpha, \beta) = (1.5, 3)$ and $\rho = 1$ (top), $\rho = 1.1$ (middle), $\rho = 1.2$ (bottom). Plots of the function $\mathbb{L}(\Pi_{R,\rho})$ are given for $k = 50$ (blue dotted line) and $k = 100$ (black line). Red horizontal line: χ_1^2 -quantile of order 0.9.

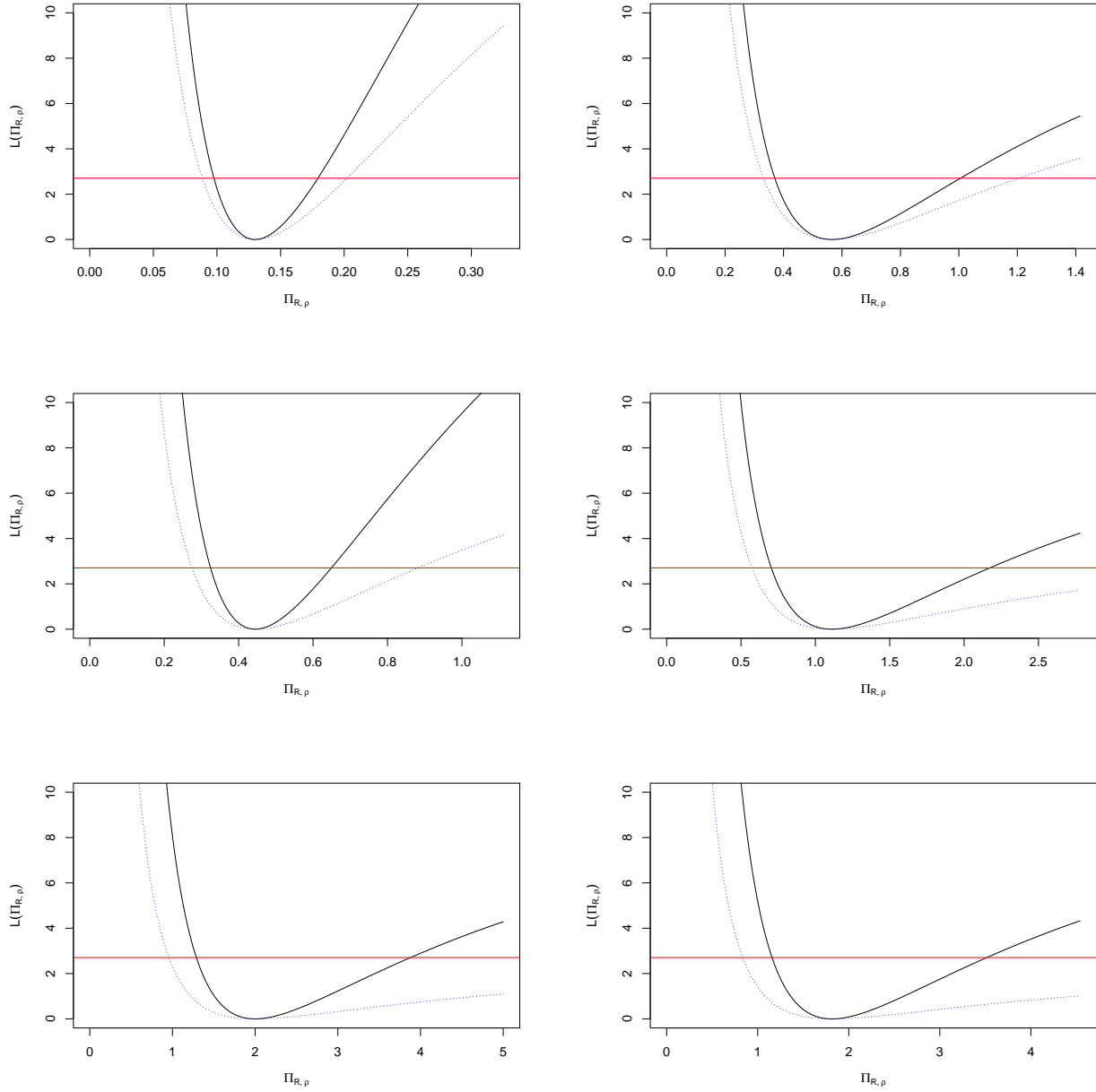


Figure 7: **Pareto distribution.** Left column: $\alpha = 3$ and $\rho = 1.2$ (top), $\rho = 1.5$ (middle), $\rho = 2$ (bottom). Right column: $\alpha = 1.5$ and $\rho = 1$ (top), $\rho = 1.1$ (middle), $\rho = 1.2$ (bottom). Plots of the function $\mathbb{L}(\Pi_{R,\rho})$ are given for $k = 50$ (blue dotted line) and $k = 100$ (black line). Red horizontal line: χ^2_1 -quantile of order 0.9.

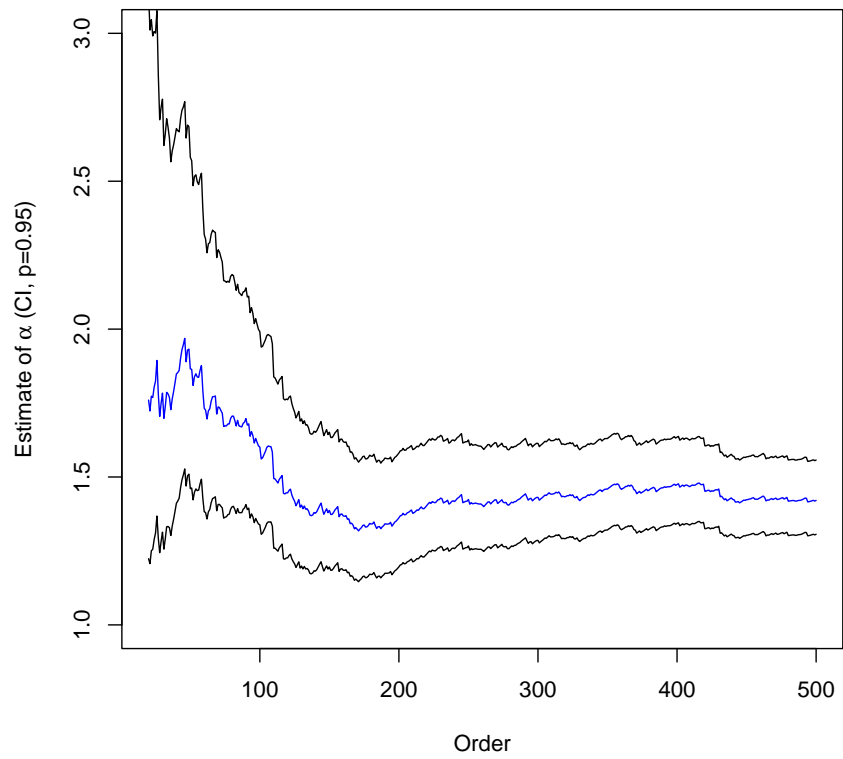


Figure 8: Hill estimate of the tail index α (with 95%-confidence intervals) for the Danish fire data.

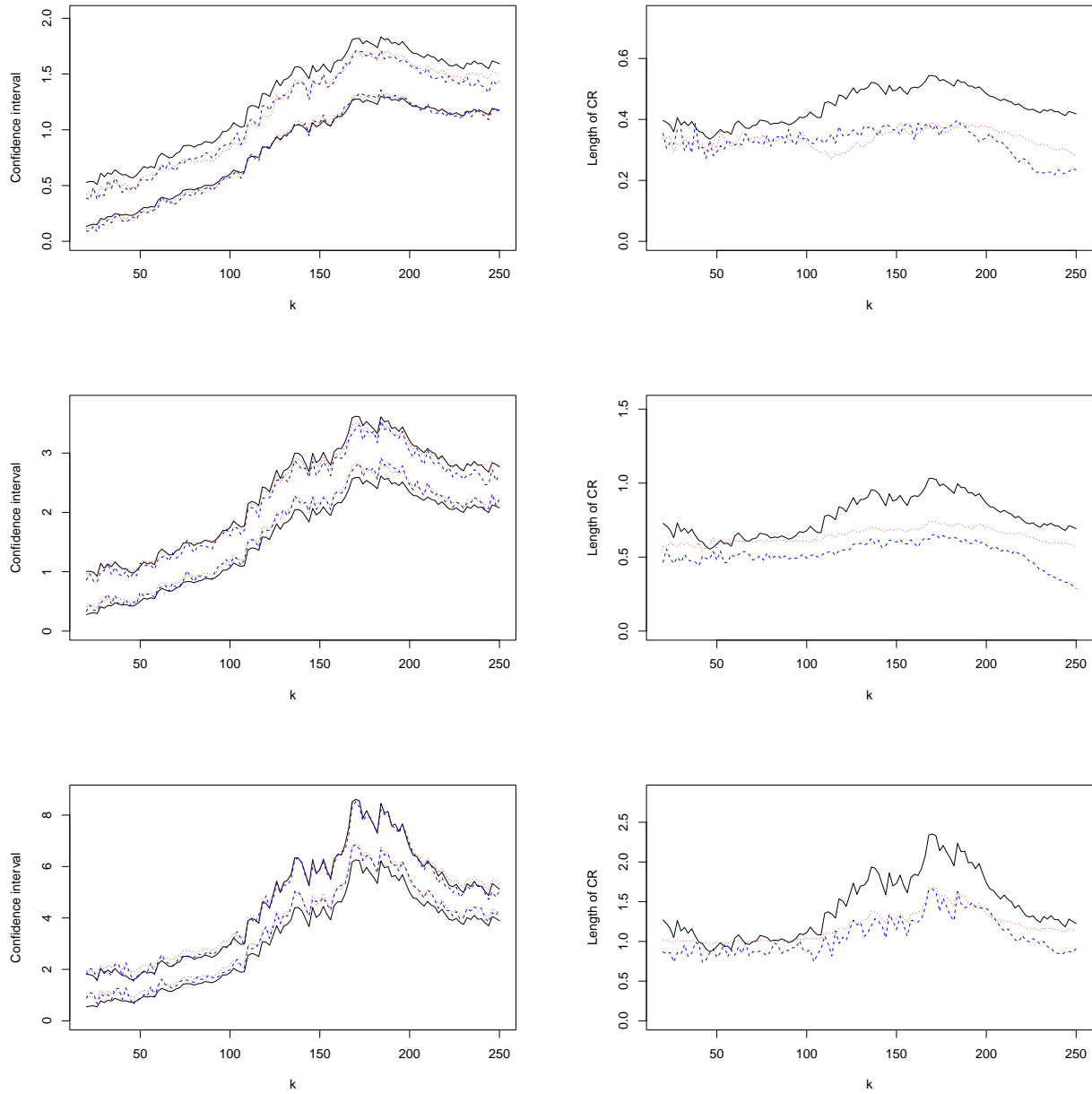


Figure 9: **Danish fire data.** 90%-confidence intervals for $\Pi_{R,\rho}$ (left) and their length (right), calculated for $\rho = 1.0$ (top), $\rho = 1.1$ (middle), $\rho = 1.2$ (bottom), and plotted against $k = 20, 22, \dots, 250$. Red dotted line: likelihood ratio method, black line: normal approximation method, blue dashed line: data tilting method.

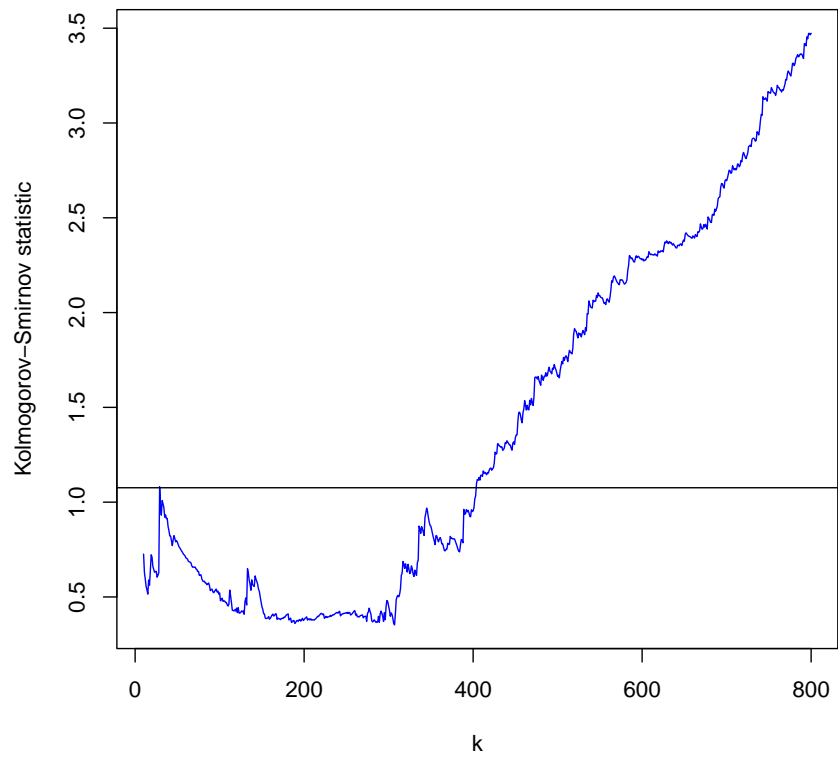


Figure 10: Kolmogorov-Smirnov test of the heavy-tail hypothesis for the Algerian car claim data.

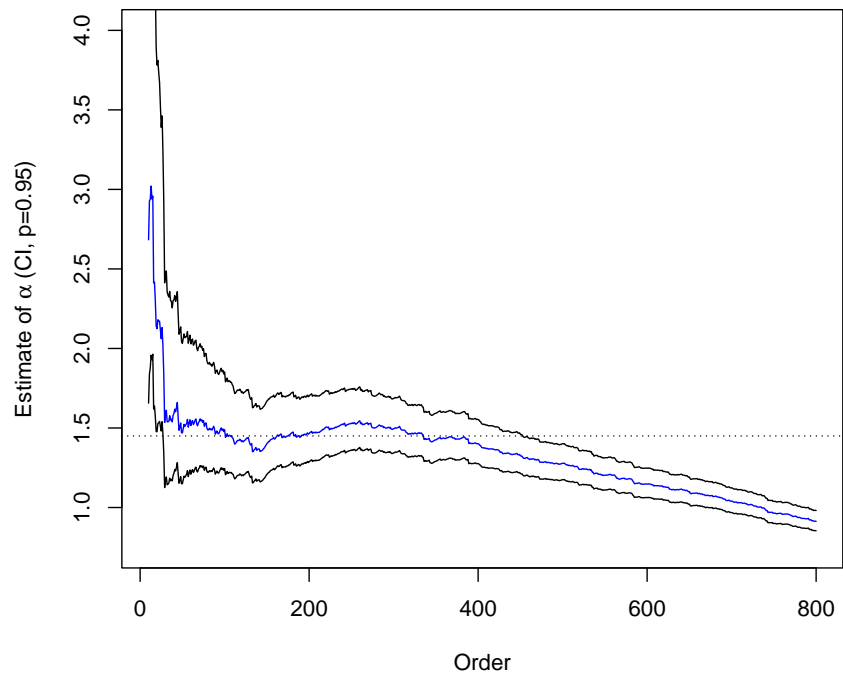


Figure 11: Hill estimate of the tail index α (with 95%-confidence intervals) for the Algerian car claim data (the dotted line indicates $\hat{\alpha} = 1.45$).

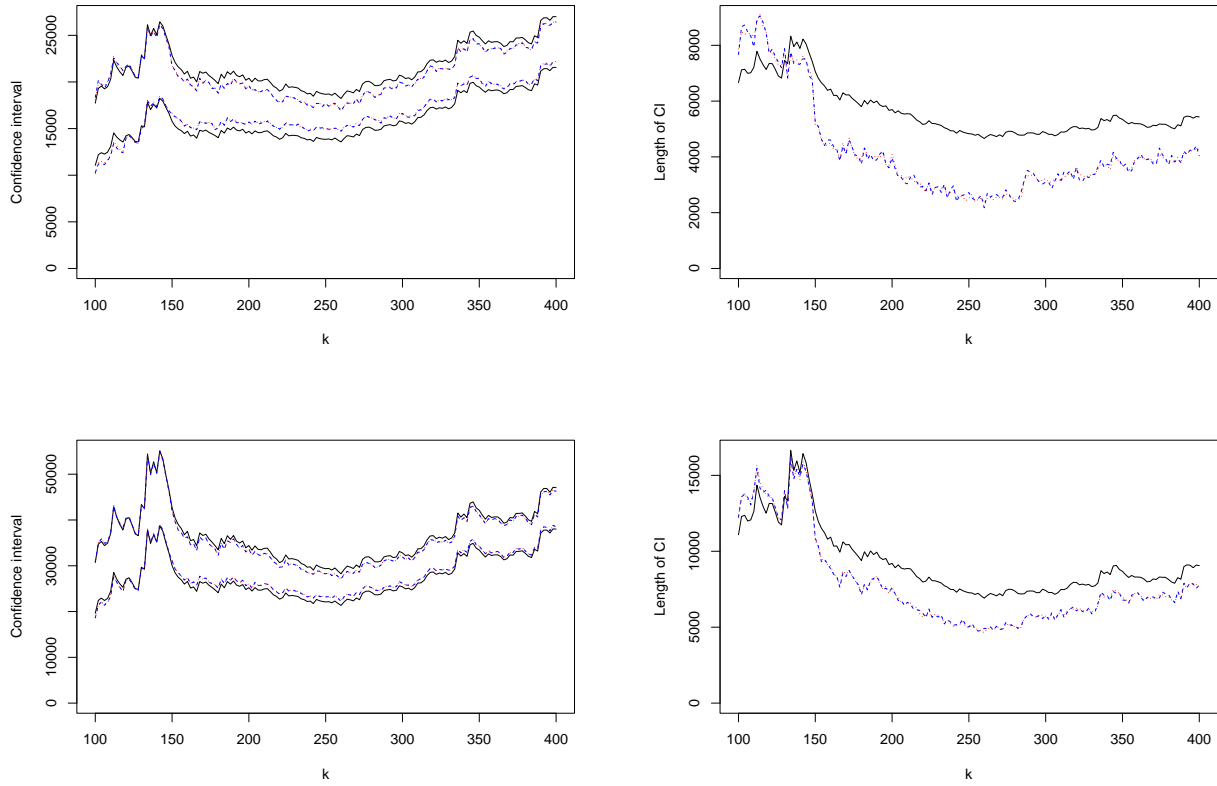


Figure 12: **Algerian car claim data.** 90%-confidence intervals for $\Pi_{R,\rho}$ (left) and their length (right), calculated for $\rho = 1.1$ (1st row), $\rho = 1.2$ (2nd row), and plotted against $k = 100, 102, \dots, 400$. Red dotted line: likelihood ratio method, black line: normal approximation method, blue dashed line: data tilting method.