

Jackknife Empirical Likelihood Methods for Risk Measures and Related Quantities

LIANG PENG¹, YONGCHENG QI², RUODU WANG¹ AND JINGPING YANG³

Abstract

The quantification of risks is of importance in insurance. In this paper, we employ the jackknife empirical likelihood methods to construct confidence intervals for the risk measures and some related quantities studied by Jones and Zitikis (2003). A simulation study shows the advantages of the new methods over the normal approximation method and naive bootstrap method.

Keywords: Confidence interval, jackknife empirical likelihood, risk measure.

1 Introduction

In life insurance and finance, quantifying risk is a very important task for pricing an insurance product or managing a financial portfolio. Generally speaking, a risk measure is constructed to be a mapping from a set of risks to some real numbers. Some well-known risk measures include coherent risk measures (Yaari (1987), Artzner (1999)), distortion risk measures, Wang's premium principle and proportional hazards transform risk measures (Wang, Young and Panjer (1997); Wang (1995), (1996), (1998); Wirch and Hardy (1999); Necir and Meraghni (2009)).

For a risk variable X with distribution function F , Jones and Zitikis (2003) defined a large class of risk measures associated with X as

$$R(F) = \int_0^1 F^-(t)\psi(t)dt, \quad (1)$$

where F^- denotes the generalized inverse function of F , and ψ is a nonnegative function chosen for showing the objective opinion about the risk loading. Different choices of ψ result in different risk measures. For example, the proportional hazards transform risk measure has $\psi(t) = r(1-t)^{r-1}$

¹School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA. Email: peng@math.gatech.edu, ruodu.wang@math.gatech.edu

²Department of Mathematics and Statistics, University of Minnesota–Duluth, 1117 University Drive, Duluth, MN 55812. Email: yqi@d.umn.edu

³LMAM, Department of Financial Mathematics, Peking University, Beijing 100871, China. Email: yangjp@math.pku.edu.cn

and Wang's premium principle has $\psi(t) = g'(1 - t)$, where g is an increasing convex function with derivatives over $[0, 1]$; see Jones and Zitikis (2003) for details. Other choices of the function ψ can be found in Jones and Zitikis (2007). Jones and Zitikis (2003) also introduced a related quantity to illustrate the right-tail, left-tail and two-sided deviations, which is defined as

$$r(F) = \frac{R(F)}{\mathbb{E}(X)}. \quad (2)$$

In this paper, we will focus on the statistical inference of the risk measure and its related quantity defined in (1) and (2), respectively.

How to infer $R(F)$ and $r(F)$ plays an important role in the applications of risk measures. Recently, Jones and Zitikis (2003) proposed nonparametric estimation by replacing F^- and $\mathbb{E}(X)$ by the sample quantile function and sample mean respectively, and derived the asymptotic normality. Therefore, confidence intervals for $R(F)$ and $r(F)$ can be constructed via estimating the asymptotic variance. For comparing two risk measures, we refer to Jones and Zitikis (2005). Jones and Zitikis (2007) investigated the nonparametric estimation of the parameter associated with distortion-based risk measures. In order to construct confidence intervals for $R(F)$ and $r(F)$ without estimating the asymptotic variance, we investigate the possibility of applying empirical likelihood method in this paper so as to improve the inference.

Empirical likelihood method is a nonparametric likelihood approach for statistical inference, which has been shown to be powerful in interval estimation and hypothesis testing. We refer to Owen (2001) for an overview on the method. However, it is known that empirical likelihood methods are not effective in dealing with non-linear functionals. Recently, a so-called jackknife empirical likelihood method was proposed by Jing, Yuan and Zhou (2009) to deal with nonlinear functionals. The key idea is to formulate a jackknife sample based on estimating the nonlinear functional and then apply the empirical likelihood method for a mean to the jackknife sample. Since the risk measure $R(F)$ and its related quantity $r(F)$ are non-linear functionals, we propose to employ the jackknife empirical likelihood method to obtain interval estimation for these two quantities.

The paper is organized as follows. In Section 2, the methodologies and main results are presented. A simulation study is given in Section 3. All proofs are put in Section 4. Some conclusions are drawn in Section 5.

2 Methodologies and main results

Put

$$\Psi(t) = \int_0^t \psi(s) ds.$$

Then the risk measure defined in (1) can be written as

$$R = R(F) = \int_{-\infty}^{\infty} (\Psi(1) - \Psi(F(t))) dt.$$

Let X_1, \dots, X_n be independent real-valued random variables with continuous distribution function $F(x)$ and $\mathbb{E}(X_1) \neq 0$. Define the empirical distribution function as

$$F_n(x) = \frac{1}{n} \sum_{j=1}^n \mathbf{1}(X_j \leq x).$$

Then Jones and Zitikis (2003) proposed to estimate $R(F)$ and $r(F)$ by

$$\hat{R}_n = \int_{-\infty}^{\infty} (\Psi(1) - \Psi(F_n(t))) dt, \quad \text{and} \quad \hat{r}_n = \frac{n \int_{-\infty}^{\infty} (\Psi(1) - \Psi(F_n(t))) dt}{\sum_{j=1}^n X_j},$$

respectively, and showed that

$$\sqrt{n}\{\hat{R}_n - R\} \xrightarrow{d} N(0, \sigma_1^2) \quad \text{and} \quad \sqrt{n}\{\hat{r}_n - r(F)\} \xrightarrow{d} N(0, \sigma_2^2) \quad (3)$$

under some regularity conditions, where

$$\sigma_1^2 = Q_F(\Psi, \Psi), \quad \sigma_2^2 = \frac{1}{\mu^2} (Q_F(\Psi, \Psi) - 2r(F)Q_F(\Psi, 1) + (r(F))^2 Q_F(1, 1)) \quad (4)$$

and

$$Q_F(a, b) = \int_{-\infty}^{\infty} (F(x \wedge y) - F(x)F(y)) a(F(x)) b(F(y)) dx dy.$$

Based on (3), confidence intervals for $R(F)$ and $r(F)$ can be obtained via estimating σ_1^2 and σ_2^2 .

An alternative way to construct confidence intervals is to employ the empirical likelihood method. Since the risk measure R is non-linear, a common technique is to linearize the functional by introducing some link variables before applying the profile empirical likelihood method; see the study for variance, ROC curve (Claeskens et al. (2003)) and copulas (Chen, Peng and Zhao (2009)). Unfortunately it remains unknown on how to linearize R by introducing some link variables. Here we propose to apply the jackknife empirical likelihood method in Jing, Yuan and Zhou (2009). This procedure is easy to implement and is described as follows.

Define $F_{n,i} = \frac{1}{n-1} \sum_{j=1, j \neq i}^n \mathbf{1}(X_j \leq x)$ and $\hat{R}_{n,i} = \int_{-\infty}^{\infty} (\Psi(1) - \Psi(F_{n,i}(t))) dt$ for $i = 1, \dots, n$. Then the jackknife sample is defined as

$$Y_i = n\hat{R}_n - (n-1)\hat{R}_{n,i}, \quad i = 1, \dots, n.$$

Now we apply the empirical likelihood method to the above jackknife sample. That is, we define the jackknife empirical likelihood function for $\theta = R(F)$ as

$$L_1(\theta) = \sup \left\{ \prod_{i=1}^n np_i : p_i \geq 0, \text{ for } i = 1, \dots, n; \sum_{i=1}^n p_i = 1; \sum_{i=1}^n p_i Y_i = \theta \right\}.$$

By Lagrange multiplier technique, we have $p_i = n^{-1} \{1 + \lambda(Y_i - \theta)\}^{-1}$ and $-2 \log L_1(\theta) = 2 \sum_{i=1}^n \log \{1 + \lambda(Y_i - \theta)\}$, where $\lambda = \lambda(\theta)$ satisfies

$$\sum_{i=1}^n \frac{Y_i - \theta}{1 + \lambda(Y_i - \theta)} = 0. \quad (5)$$

The following theorem shows that the Wilks theorem holds for the proposed jackknife empirical likelihood method.

Theorem 1. *Assume that $|\psi(x)| \leq cx^{\alpha-1}(1-x)^{\beta-1}$, $\psi'(x)$ exists and $|\psi'(x)| \leq cx^{\alpha-2}(1-x)^{\beta-2}$ for all $0 < x < 1$ and some constants $\alpha > 1/2$, $\beta > 1/2$ and $c > 0$. Further assume $\mathbb{E}(|X_i|^\gamma) < \infty$ for some γ such that $\gamma > 1/(\alpha - 1/2)$ and $\gamma > 1/(\beta - 1/2)$. Then we have*

$$-2 \log L_1(R_0) \xrightarrow{d} \chi_1^2 \quad \text{as } n \rightarrow \infty,$$

where R_0 denotes the true value of R and χ_1^2 denotes a chi-square distribution with one degree of freedom.

Remark 1. *Some well-known risk measures, such as proportional hazards transform risk measure, Wang's right-tail deviation and Wang's left-tail deviation satisfy the assumptions of Theorem 1; see Jones and Zitikis (2003).*

Based on the above theorem, a confidence interval for R_0 with level b can be obtained as

$$I_b^R = \{R : -2 \log L_1(R) \leq \chi_{1,b}^2\},$$

where $\chi_{1,b}^2$ is the b -th quantile of χ_1^2 .

Next we consider the related quantity $r(F) = R(F)/\mu$ where $\mu = \mathbb{E}(X_1)$. Alternatively, we consider the quantity $R - \theta\mu$ with $\theta = r(F)$. Then one can estimate this quantity by

$$\hat{R}_n - \theta n^{-1} \sum_{i=1}^n X_i = \hat{R}_n - \theta \int_{-\infty}^{\infty} x dF_n(x) = \hat{R}_n - \theta \int_{-\infty}^{\infty} F_n(x) dx.$$

As before, we define the jackknife sample as

$$n \left(\hat{R}_n - \theta \int_{-\infty}^{\infty} x dF_n(x) \right) - (n-1) \left(\hat{R}_{n,i} - \theta \int_{-\infty}^{\infty} x dF_{n,i}(x) \right) = Y_i - \theta X_i$$

for $i = 1, \dots, n$, where Y_i 's are defined as above. So the jackknife empirical likelihood function for $\theta = r(F)$ is defined as

$$L_2(\theta) = \sup \left\{ \prod_{i=1}^n np_i : p_i \geq 0, \text{ for } i = 1, \dots, n; \sum_{i=1}^n p_i = 1; \sum_{i=1}^n p_i(Y_i - \theta X_i) = 0 \right\}.$$

The following theorem shows that the Wilks theorem holds for the proposed jackknife empirical likelihood method for $r(F)$.

Theorem 2. *Assume the conditions of Theorem 1 hold. Further assume $\mathbb{E}X_1 \neq 0$, and $\mathbb{E}(X_1^2) < \infty$. Then*

$$-2 \log L_2(r_0) \xrightarrow{d} \chi_1^2 \quad \text{as } n \rightarrow \infty,$$

where r_0 denotes the true value of $r(F)$.

Based on the above theorem, a confidence interval for r_0 with level b can be obtained as

$$I_b^r = \{r : -2 \log L_2(r) \leq \chi_{1,b}^2\}.$$

3 Simulation study

In this section we examine the finite sample behavior of the proposed jackknife empirical likelihood methods in terms of coverage accuracy and compare with the normal approximation method and the naive bootstrap method. We focus on the proportional hazards transform risk measure with $\psi(s) = a(1-s)^{a-1}$. Since Pareto distribution, log-normal distribution, Weibull distribution and Gamma distribution are widely used in fitting the losses data in insurance (Klugman, Panjer and Willmot (2008)), our simulation study is based on these four distributions.

We draw 10,000 random samples of sizes $n = 300$ and 1000 from the following distributions:

1. Pareto distribution $F_1(x; \theta) = 1 - x^{-\theta}$ for $x \geq 1$;
2. Log-normal distribution $F_2(x; \theta_1, \theta_2) = \Phi((\log x - \theta_1)/\theta_2)$ for $x > 0$, where $\Phi(x)$ denotes the standard normal distribution;
3. Weibull distribution $F_3(x; \theta_1, \theta_2) = 1 - \exp\{-(x/\theta_2)^{\theta_1}\}$ for $x > 0$;
4. Gamma distribution

$$F_4(x; \theta_1, \theta_2) = \int_0^x \frac{\theta_2^{\theta_1}}{\Gamma(\theta_1)} s^{\theta_1-1} \exp\{-\theta_2 s\} ds \quad \text{for } x > 0.$$

For calculating the proposed jackknife empirical likelihood intervals for both $R(F)$ and $r(F)$ (JELCI), we employ the proportional hazards transform risk measure with $\psi(s) = a(1-s)^{a-1}$, $a = 0.55$ and 0.85 , and use the R package 'emplik'. For calculating the confidence intervals for $R(F)$ based on the normal approximation method (NACI), we use the variance estimation in Jones and Zitikis (2003). For computing the naive bootstrap confidence intervals for $r(F)$ (NBCI), we draw 10,000 bootstrap samples with replacement from each random sample X_1, \dots, X_n . Empirical coverage probabilities are reported in Tables 1 and 2 for these three confidence intervals with levels 0.9, 0.95 and 0.99. From these two tables, we conclude that the proposed jackknife empirical likelihood methods give much more accurate coverage probabilities than the other two methods.

4 Proofs

Throughout we put $U_i = F(X_i)$ for $i = 1, \dots, n$, $G_n(t) = n^{-1} \sum_{i=1}^n \mathbf{1}(U_i \leq t)$ and $G_{n,i} = (n-1)^{-1} \sum_{j=1, j \neq i}^n \mathbf{1}(U_j \leq t)$ for $i = 1, \dots, n$. Since F is continuous, U_1, \dots, U_n are independent and uniformly distributed over $(0, 1)$. Without loss of generality we also assume no ties in U_1, \dots, U_n , and let $U_{n,1} < \dots < U_{n,n}$ denote the order statistics of U_1, \dots, U_n . We also use C to denote a generic constant which may be different in different places.

First we list some facts which will be employed in the proofs. We assume $\beta \leq \alpha$ throughout since proofs for the case of $\beta > \alpha$ are exactly the same. Therefore we have $|\psi(x)| \leq cx^{\beta-1}(1-x)^{\beta-1}$ and $|\psi'(x)| \leq cx^{\beta-2}(1-x)^{\beta-2}$ for all $0 < x < 1$. Since $\mathbb{E}|X_1|^\gamma < \infty$ with $\frac{1}{\gamma} + 1 - \beta < \frac{1}{2}$, we have

$$P(|X_1| > x) = o(x^{-\gamma}) \quad \text{as } x \rightarrow \infty, \tag{6}$$

which implies

$$\int_{-\infty}^{\infty} (F(x))^{\beta-1+\delta} (1-F(x))^{\beta-1+\delta} dx \leq 2 + C \int_1^{\infty} x^{-(\beta-1+\delta)\gamma} dx < \infty \tag{7}$$

whenever $\delta \in (\frac{1}{\gamma} + 1 - \beta, \frac{1}{2})$, and

$$\max_{1 \leq j \leq n} |X_j| = \max_{1 \leq j \leq n} |F^-(U_j)| = o_p(n^{1/\gamma}). \quad (8)$$

It follows from the given conditions on ψ that

$$\Psi\left(\frac{1}{n}\right) = O(n^{-\beta}) \quad \text{and} \quad \Psi\left(\frac{1}{n-1}\right) - \Psi\left(\frac{1}{n}\right) = O(n^{-\beta-1}). \quad (9)$$

Lemma 1. *Under the conditions of Theorem 1, we have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (Y_i - R_0) \xrightarrow{d} N(0, \sigma_1^2), \quad (10)$$

where σ_1^2 is given in (4).

Proof. Write

$$\begin{aligned} Y_i &= (n-1) \int_{-\infty}^{\infty} \{\Psi(F_{n,i}(t)) - \Psi(F_n(t))\} dt + \hat{R}_n \\ &= (n-1) \int_0^1 \{\Psi(G_{n,i}(t)) - \Psi(G_n(t))\} dF^-(t) + \hat{R}_n \\ &= (n-1) \int_0^1 \{\Psi(G_{n,i}(t)) - \Psi(G_n(t))\} \mathbf{1}(U_{n,1} \leq t < U_{n,n}) dF^-(t) + \hat{R}_n \\ &= (n-1) \int_0^1 \{\Psi(G_{n,i}(t)) - \Psi(G_n(t))\} \mathbf{1}(U_{n,1} \leq t < U_{n,2}) dF^-(t) \\ &\quad + (n-1) \int_0^1 \{\Psi(G_{n,i}(t)) - \Psi(G_n(t))\} \mathbf{1}(U_{n,2} \leq t < U_{n,n-1}) dF^-(t) \\ &\quad + (n-1) \int_0^1 \{\Psi(G_{n,i}(t)) - \Psi(G_n(t))\} \mathbf{1}(U_{n,n-1} \leq t < U_{n,n}) dF^-(t) + \hat{R}_n \\ &= (n-1) \underbrace{\int_0^1 \{\Psi(G_{n,i}(t)) - \Psi(G_n(t))\} \mathbf{1}(U_{n,1} \leq t < U_{n,2}) dF^-(t)}_{Z_{i,1}} \\ &\quad + (n-1) \underbrace{\int_0^1 \psi(G_n(t)) \{G_{n,i}(t) - G_n(t)\} \mathbf{1}(U_{n,2} \leq t < U_{n,n-1}) dF^-(t)}_{Z_{i,2}} \\ &\quad + \frac{n-1}{2} \underbrace{\int_0^1 \psi'(\xi_{n,i}(t)) \{G_{n,i}(t) - G_n(t)\}^2 \mathbf{1}(U_{n,2} \leq t < U_{n,n-1}) dF^-(t)}_{Z_{i,3}} \\ &\quad + (n-1) \underbrace{\int_0^1 \{\Psi(G_{n,i}(t)) - \Psi(G_n(t))\} \mathbf{1}(U_{n,n-1} \leq t < U_{n,n}) dF^-(t)}_{Z_{i,4}} + \hat{R}_n \\ &= Z_{i,1} + Z_{i,2} + Z_{i,3} + Z_{i,4} + \hat{R}_n, \end{aligned}$$

where

$$\xi_{n,i}(t) = G_n(t) + \theta_i(t)\{G_{n,i}(t) - G_n(t)\} = G_n(t) + \frac{\theta_i(t)}{n-1}\{G_n(t) - \mathbf{1}(U_i \leq t)\}$$

for some $\theta_i(t) \in [0, 1]$.

When $U_{n,1} \leq t < U_{n,2}$, we have

$$G_n(t) = \frac{1}{n} \quad \text{and} \quad G_{n,i}(t) = \begin{cases} 0 & \text{if } U_i = U_{n,1} \\ \frac{1}{n-1} & \text{else.} \end{cases} \quad (11)$$

Hence, it follows from (8) and (9) that

$$\begin{aligned} \sum_{i=1}^n Z_{i,1} &= (n-1) \int_0^1 \{\Psi(0) - \Psi(\frac{1}{n})\} \mathbf{1}(U_{n,1} \leq t < U_{n,2}) dF^-(t) \\ &\quad + (n-1)^2 \int_0^1 \{\Psi(\frac{1}{n-1}) - \Psi(\frac{1}{n})\} \mathbf{1}(U_{n,1} \leq t < U_{n,2}) dF^-(t) \\ &= -(n-1)\Psi(\frac{1}{n})\{F^-(U_{n,2}) - F^-(U_{n,1})\} \\ &\quad + (n-1)^2\{\Psi(\frac{1}{n-1}) - \Psi(\frac{1}{n})\}\{F^-(U_{n,2}) - F^-(U_{n,1})\} \\ &= O((n-1)n^{-\beta})o_p(n^{1/\gamma}) + O((n-1)^2n^{-1-\beta})o_p(n^{1/\gamma}) \\ &= o_p(n^{1/2-\beta+1/\gamma})\sqrt{n} \\ &= o_p(\sqrt{n}) \end{aligned} \quad (12)$$

since $\frac{1}{2} - \beta + \frac{1}{\gamma} < 0$. Similarly, we can show that

$$\sum_{i=1}^n Z_{i,4} = o_p(\sqrt{n}). \quad (13)$$

Since $\sum_{i=1}^n \{G_{n,i}(t) - G_n(t)\} = 0$, we have

$$\sum_{i=1}^n Z_{i,2} = 0. \quad (14)$$

When $t \geq U_{n,2}$, we have

$$\frac{(n-1)^{-1}\mathbf{1}(U_i \leq t)}{G_n(t)} \leq \frac{1/(n-1)}{2/n} = \frac{n}{2(n-1)},$$

i.e.,

$$\xi_{n,i}(t) \geq G_n(t)\{1 - \frac{n}{2(n-1)}\}$$

uniformly in $t \geq U_{n,2}$. In the same manner, we can show that the equation

$$1 - \xi_{n,i}(t) \geq (1 - G_n(t))\{1 - \frac{n}{2(n-1)}\}$$

holds uniformly in $t < U_{n,n-1}$. Hence, for n large enough,

$$(\xi_{n,i}(t), 1 - \xi_{n,i}(t)) \geq \frac{1}{3}(G_n(t), 1 - G_n(t)) \quad \text{uniformly for } U_{n,2} \leq t < U_{n,n-1} \quad \text{and } 1 \leq i \leq n. \quad (15)$$

Note that

$$\sup_{U_{n,2} \leq t \leq U_{n,n-1}} \frac{G_n(t)}{t} = O_p(1) \quad \text{and} \quad \sup_{U_{n,2} \leq t \leq U_{n,n-1}} \frac{1 - G_n(t)}{1 - t} = O_p(1) \quad (16)$$

(see Page 404 of Shorack and Wellner (1986)). It follows from (15) and (16) that

$$|Z_{i,3}| = O_p \left(n \int_0^1 t^{\beta-2} (1-t)^{\beta-2} \{G_{n,i}(t) - G_n(t)\}^2 \mathbf{1}(U_{n,2} \leq t < U_{n,n-1}) dF^-(t) \right),$$

which coupled with (7) yields

$$\begin{aligned} \sum_{i=1}^n Z_{i,3} &= O_p \left(n \int_0^1 t^{\beta-2} (1-t)^{\beta-2} \sum_{i=1}^n \{G_{n,i}(t) - G_n(t)\}^2 \mathbf{1}(U_{n,2} \leq t < U_{n,n-1}) dF^-(t) \right) \\ &= O_p \left(n \int_0^1 t^{\beta-2} (1-t)^{\beta-2} \frac{n}{(n-1)^2} G_n(t) \{1 - G_n(t)\} \mathbf{1}(U_{n,2} \leq t < U_{n,n-1}) dF^-(t) \right) \\ &= O_p \left(\int_0^1 t^{\beta-1} (1-t)^{\beta-1} \mathbf{1}(U_{n,2} \leq t < U_{n,n-1}) dF^-(t) \right) \\ &= O_p \left(\int_{n^{-1}}^{1-n^{-1}} t^{\beta-1} (1-t)^{\beta-1} dF^-(t) \right) \\ &= O_p \left(n^\delta \int_{n^{-1}}^{1-n^{-1}} t^{\beta-1+\delta} (1-t)^{\beta-1+\delta} dF^-(t) \right) \\ &= O_p \left(n^\delta \int_{-\infty}^{\infty} (F(x))^{\beta-1+\delta} (1-F(x))^{\beta-1+\delta} dx \right) \\ &= O_p(n^\delta) \end{aligned} \quad (17)$$

for any $\delta \in (\frac{1}{\gamma} + 1 - \beta, \frac{1}{2})$. By Jones and Zitikis (2003), we have

$$\sqrt{n}\{\hat{R}_n - R\} \xrightarrow{d} N(0, \sigma_1^2). \quad (18)$$

Hence, the lemma follows from (12), (14), (17), (13) and (18). \square

Lemma 2. *Under the conditions of Theorem 1, we have*

$$\frac{1}{n} \sum_{i=1}^n (Y_i - R)^2 \xrightarrow{p} \sigma_1^2 \quad \text{as } n \rightarrow \infty.$$

Proof. We use the same notations $Z_{i,j}$ as in the proof of Lemma 1. Then, it follows from (11) and (9) that

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^n Z_{i,1}^2 &= \frac{(n-1)^2}{n} \int_0^1 \int_0^1 \left\{ \Psi(0) - \Psi\left(\frac{1}{n}\right) \right\}^2 \mathbf{1}(U_{n,1} \leq t_1, t_2 < U_{n,2}) dF^-(t_1) dF^-(t_2) \\
&\quad + \frac{(n-1)^3}{n} \int_0^1 \int_0^1 \left\{ \Psi\left(\frac{1}{n-1}\right) - \Psi\left(\frac{1}{n}\right) \right\}^2 \mathbf{1}(U_{n,1} \leq t_1, t_2 < U_{n,2}) dF^-(t_1) dF^-(t_2) \\
&= O\left(\frac{(n-1)^2}{n} n^{-2\beta}\right) o_p(n^{2/\gamma}) + O\left(\frac{(n-1)^3}{n} n^{-2-2\beta}\right) o_p(n^{2/\gamma}) \\
&= o_p(1).
\end{aligned} \tag{19}$$

Similarly,

$$\frac{1}{n} \sum_{i=1}^n Z_{i,4}^2 = o_p(1). \tag{20}$$

It is easy to check that

$$\begin{aligned}
&\frac{1}{n} \sum_{i=1}^n Z_{i,2}^2 \\
&= \frac{(n-1)^2}{n} \int_0^1 \int_0^1 \psi(G_n(t_1)) \psi(G_n(t_2)) \sum_{i=1}^n \{G_{n,i}(t_1) - G_n(t_1)\} \{G_{n,i}(t_2) - G_n(t_2)\} \\
&\quad \times \mathbf{1}(U_{n,2} \leq t_1, t_2 < U_{n,n-1}) dF^-(t_2) dF^-(t_1) \\
&= \int_0^1 \int_0^1 \psi(G_n(t_1)) \psi(G_n(t_2)) \{G_n(t_1 \wedge t_2) - G_n(t_1) G_n(t_2)\} \mathbf{1}(U_{n,2} \leq t_1, t_2 < U_{n,n-1}) dF^-(t_2) dF^-(t_1) \\
&= 2 \underbrace{\int_0^1 \int_0^{t_1} \psi(G_n(t_1)) \psi(G_n(t_2)) G_n(t_2) \{1 - G_n(t_1)\} \mathbf{1}(U_{n,1} \leq t_1, t_2 < U_{n,n-1}) dF^-(t_2) dF^-(t_1)}_{I_1}.
\end{aligned}$$

By (16), we have

$$\sup_{U_{n,2} \leq t_1, t_2 < U_{n,n-1}} \psi(G_n(t_1)) \psi(G_n(t_2)) G_n(t_2) \{1 - G_n(t_1)\} = O_p\left(t_1^{\beta-1} (1-t_1)^{\beta-1} t_2^{\beta-1} (1-t_2)^{\beta-1} t_2 (1-t_1)\right).$$

Similar to the proof of (7), we can show that

$$\begin{aligned}
&\int_0^1 \int_0^{t_1} t_1^{\beta-1} (1-t_1)^{\beta-1} t_2^{\beta-1} (1-t_2)^{\beta-1} t_2 (1-t_1) dF^-(t_2) dF^-(t_1) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^x F(x)^{\beta-1} (1-F(x))^{\beta-1} F(y)^{\beta-1} (1-F(y))^{\beta-1} F(y) (1-F(x)) dy dx \\
&< \infty.
\end{aligned}$$

By the Glivenko-Cantelli theorem, $\sup_{0 < t < 1} |G_n(t) - t| \rightarrow 0$ almost surely. It then follows from the dominated convergence theorem that

$$I_1 \xrightarrow{p} \int_0^1 \int_0^{t_1} \psi(t_1) \psi(t_2) t_2 (1-t_1) dF^-(t_2) dF^-(t_1).$$

Hence

$$\frac{1}{n} \sum_{i=1}^n Z_{i,2}^2 \xrightarrow{p} \int_0^1 \int_0^1 \psi(t_1) \psi(t_2) \{t_1 \wedge t_2 - t_1 t_2\} dF^-(t_2) dF^-(t_1) = \sigma_1^2. \quad (21)$$

Note that

$$\begin{aligned} & \sum_{i=1}^n \{G_{n,i}(t_1) - G_n(t_1)\}^2 \{G_{n,i}(t_2) - G_n(t_2)\}^2 \\ &= \sum_{i=1}^n \left\{ \frac{G_n(t_1)}{n-1} - \frac{\mathbf{1}(U_i \leq t_1)}{n-1} \right\}^2 \left\{ \frac{G_n(t_2)}{n-1} - \frac{\mathbf{1}(U_i \leq t_2)}{n-1} \right\}^2 \\ &= \frac{n}{(n-1)^4} \{ -3G_n^2(t_1)G_n^2(t_2) + G_n^2(t_1)G_n(t_2) + G_n(t_1)G_n^2(t_2) + 4G_n(t_1)G_n(t_2)G_n(t_1 \wedge t_2) \\ &\quad - 2G_n(t_1)G_n(t_1 \wedge t_2) - 2G_n(t_2)G_n(t_1 \wedge t_2) + G_n(t_1 \wedge t_2) \} \\ &= \frac{n}{(n-1)^4} \{ \underbrace{3G_n(t_1)G_n(t_2)(G_n(t_1 \wedge t_2) - G_n(t_1)G_n(t_2))}_{II_1} - \underbrace{G_n(t_1)(G_n(t_1 \wedge t_2) - G_n(t_1)G_n(t_2))}_{II_2} \\ &\quad - \underbrace{G_n(t_2)(G_n(t_1 \wedge t_2) - G_n(t_1)G_n(t_2))}_{II_3} + \underbrace{(1 - G_n(t_1))(1 - G_n(t_2))G_n(t_1 \wedge t_2)}_{II_4} \} \\ &= \frac{n}{(n-1)^4} \{ II_1 - II_2 - II_3 + II_4 \}. \end{aligned}$$

It follows from (16) that

$$\sup_{U_{n,2} \leq t_1, t_2 \leq U_{n,n-1}} \frac{|G_n(t_1 \wedge t_2) - G_n(t_1)G_n(t_2)|}{t_1 \wedge t_2 - t_1 t_2} = \sup_{U_{n,2} \leq t_1, t_2 \leq U_{n,n-1}} \frac{|G_n(t_1 \wedge t_2)(1 - G_n(t_1 \vee t_2))|}{t_1 \wedge t_2(1 - t_1 \vee t_2)} = O_p(1),$$

which, coupled with (15) and (16), yields that

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n Z_{i,3}^2 \\ &= O_p \left(\frac{(n-1)^2}{4n} \int_0^1 \int_0^1 t_1^{\beta-2} (1-t_1)^{\beta-2} t_2^{\beta-2} (1-t_2)^{\beta-2} \right. \\ &\quad \times \sum_{i=1}^n \{G_{n,i}(t_1) - G_n(t_1)\}^2 \{G_{n,i}(t_2) - G_n(t_2)\}^2 \mathbf{1}(U_{n,2} \leq t_1, t_2 < U_{n,n-1}) dF^-(t_2) dF^-(t_1) \Big) \\ &= O_p \left(n^{-2} \int_0^1 \int_0^1 t_1^{\beta-2} (1-t_1)^{\beta-2} t_2^{\beta-2} (1-t_2)^{\beta-2} \right. \\ &\quad \times \{ II_1 - II_2 - II_3 + II_4 \} \mathbf{1}(U_{n,2} \leq t_1, t_2 < U_{n,n-1}) dF^-(t_2) dF^-(t_1) \Big). \end{aligned}$$

From the above equation we can get that

$$\begin{aligned}
& \frac{1}{n} \sum_{i=1}^n Z_{i,3}^2 \\
&= O_p \left(\underbrace{n^{-2} \int_{U_{n,2}}^{U_{n,n-1}} \int_{U_{n,2}}^{U_{n,n-1}} t_1^{\beta-2} (1-t_1)^{\beta-2} t_2^{\beta-2} (1-t_2)^{\beta-2} t_1 t_2 (t_1 \wedge t_2 - t_1 t_2) dF^-(t_2) dF^-(t_1)}_{III_1} \right) \\
&+ O_p \left(\underbrace{n^{-2} \int_{U_{n,2}}^{U_{n,n-1}} \int_{U_{n,2}}^{U_{n,n-1}} t_1^{\beta-2} (1-t_1)^{\beta-2} t_2^{\beta-2} (1-t_2)^{\beta-2} t_1 (t_1 \wedge t_2 - t_1 t_2) dF^-(t_2) dF^-(t_1)}_{III_2} \right) \\
&+ O_p \left(\underbrace{n^{-2} \int_{U_{n,2}}^{U_{n,n-1}} \int_{U_{n,2}}^{U_{n,n-1}} t_1^{\beta-2} (1-t_1)^{\beta-2} t_2^{\beta-2} (1-t_2)^{\beta-2} t_2 (t_1 \wedge t_2 - t_1 t_2) dF^-(t_2) dF^-(t_1)}_{III_3} \right) \\
&+ O_p \left(\underbrace{n^{-2} \int_{U_{n,2}}^{U_{n,n-1}} \int_{U_{n,2}}^{U_{n,n-1}} t_1^{\beta-2} (1-t_1)^{\beta-2} t_2^{\beta-2} (1-t_2)^{\beta-2} (1-t_1)(1-t_2) (t_1 \wedge t_2) dF^-(t_2) dF^-(t_1)}_{III_4} \right) \\
&= O_p(III_1) + O_p(III_2) + O_p(III_3) + O_p(III_4).
\end{aligned}$$

It is easy to check from (7) that for every $\delta \in (\frac{1}{\gamma} + 1 - \beta, \frac{1}{2})$

$$\begin{aligned}
III_2 + III_3 &= 2n^{-2} \int_{U_{n,2}}^{U_{n,n-1}} \int_{U_{n,2}}^{t_1} t_1^{\beta-2} (1-t_1)^{\beta-2} t_2^{\beta-2} (1-t_2)^{\beta-2} (t_1 + t_2) t_2 (1-t_1) dF^-(t_2) dF^-(t_1) \\
&\leq 4n^{-2} \int_{U_{n,2}}^{U_{n,n-1}} t_1^{\beta-1} (1-t_1)^{\beta-1} \int_{U_{n,2}}^{t_1} t_2^{\beta-1} (1-t_2)^{\beta-2} dF^-(t_2) dF^-(t_1) \\
&= n^{-2} \int_0^1 t_1^{\beta-1} (1-t_1)^{\beta-1} O(U_{n,2}^{-\delta} + (1-t_1)^{-1-\delta}) dF^-(t_1) \\
&= O(n^{-2} U_{n,2}^{-\delta}) \int_0^1 t_1^{\beta-1} (1-t_1)^{\beta-1} dF^-(t_1) + O(n^{-2}) \int_0^1 t_1^{\beta-1} (1-t_1)^{\beta-2-\delta} dF^-(t_1) \\
&= O(n^{-2} U_{n,2}^{-\delta}) (U_{n,2}^{-\delta} + (1-U_{n,n-1})^{-\delta}) + O(n^{-2}) (U_{n,2}^{-\delta} + (1-U_{n,n-1})^{-1-2\delta}) \\
&= O_p(n^{-2+2\delta} + n^{-1+2\delta}) \\
&= o_p(1).
\end{aligned}$$

Similarly, we can show that

$$III_1 = o_p(1) \quad \text{and} \quad III_4 = o_p(1).$$

Hence,

$$\frac{1}{n} \sum_{i=1}^n Z_{i,3}^2 = o_p(1). \quad (22)$$

Since $\hat{R}_n \xrightarrow{p} R$, we have

$$\frac{1}{n} \sum_{i=1}^n (\hat{R}_n - R)^2 = o_p(1). \quad (23)$$

It follows from (19), (20), (22) and (23) that

$$\frac{1}{n} \sum_{i=1}^n \{Z_{i,1} + Z_{i,3} + Z_{i,4} + \hat{R}_n - R\}^2 = O\left(\frac{1}{n} \sum_{i=1}^n \{Z_{i,1}^2 + Z_{i,3}^2 + Z_{i,4}^2 + (\hat{R}_n - R)^2\}\right) = o_p(1). \quad (24)$$

Note that

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n Z_{i,2} \{Z_{i,1} + Z_{i,3} + Z_{i,4} + \hat{R}_n - R\} &\leq \sqrt{\frac{1}{n} \sum_{i=1}^n Z_{i,2}^2} \sqrt{\frac{1}{n} \sum_{i=1}^n \{Z_{i,1} + Z_{i,3} + Z_{i,4} + \hat{R}_n - R\}^2} \\ &= o_p(1). \end{aligned} \quad (25)$$

Therefore, the lemma follows from (19)–(25). \square

Proof of Theorem 1. First we observe by using (7) that for any $\delta \in (\frac{1}{\gamma} + 1 - \beta, \frac{1}{2})$

$$\begin{aligned} \max_{1 \leq i \leq n} |Z_{i,2}| &\leq \int_0^1 \psi(G_n(t)) \mathbf{1}(U_{n,2} \leq t < U_{n,n-1}) dF^-(t) \\ &= O_p\left(\int_{U_{n,2}}^{U_{n,n-1}} t^{\beta-1} (1-t)^{\beta-1} dF^-(t)\right) \\ &= O_p(U_{n,2}^{-\delta} + (1 - U_{n,n-1})^{-\delta}) \\ &= o_p(n^{1/2}). \end{aligned} \quad (26)$$

Similarly we can show that

$$\max_{1 \leq i \leq n} |Z_{i,j}| = o_p(n^{1/2}) \quad \text{for } j = 1, 3, 4.$$

Hence, $\max_{1 \leq i \leq n} |Y_i| = o_p(n^{1/2})$. By the standard arguments in the empirical likelihood method (see Chapter 11 of Owen (2001)), it follows from Lemmas 1 and 2 that

$$-2 \log L_1(R) = \frac{\{\sum_{i=1}^n (Y_i - R)\}^2}{\sum_{i=1}^n (Y_i - R)^2} + o_p(1) \xrightarrow{d} \chi^2(1).$$

\square

In order to prove Theorem 2, we need the following lemmas.

Lemma 3. *Under the conditions of Theorem 2, we have*

$$\sqrt{n} \left(\hat{R}_n - \frac{R(F)}{\mu} \frac{1}{n} \sum_{i=1}^n X_i \right) \xrightarrow{d} N(0, \bar{\sigma}^2) \quad \text{as } n \rightarrow \infty,$$

where

$$\begin{aligned} \bar{\sigma}^2 &= \int_0^1 \int_0^1 \psi(t_1) \psi(t_2) (t_1 \wedge t_2 - t_1 t_2) dF^-(t_1) dF^-(t_2) + \frac{R^2(F)}{\mu^2} \mathbb{E}(X_1 - \mu)^2 \\ &\quad + 2 \frac{R(F)}{\mu} \int_0^1 \int_0^1 \psi(t_1) (t_1 \wedge t_2 - t_1 t_2) dF^-(t_1) dF^-(t_2). \end{aligned}$$

Proof. It is known that there exists a Brownian bridge W such that

$$\sup_{0 \leq t \leq 1} \frac{\sqrt{n}(G_n(t) - t) - W(t)}{t^\delta(1-t)^\delta} = o_p(1) \quad (27)$$

for any $\delta \in (0, 1/2)$ (see Chapter 4 of Csorgo and Horvath (1993)). Then we have

$$\begin{aligned} \sqrt{n}\left\{\hat{R}_n - \frac{R(F)}{\mu} \frac{1}{n} \sum_{i=1}^n X_i\right\} &= \sqrt{n} \int_0^1 \{\Psi(t) - \Psi(G_n(t))\} dF^-(t) + \frac{R(F)}{\mu} \sqrt{n} \int_0^1 \{t - G_n(t)\} dF^-(t) \\ &\xrightarrow{d} - \int_0^1 \psi(t)W(t) dF^-(t) - \frac{R(F)}{\mu} \int_0^1 W(t) dF^-(t). \end{aligned} \quad \square$$

Lemma 4. *Under the conditions of Theorem 2, we have*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \left(Y_i - \frac{R(F)}{\mu} X_i \right) \xrightarrow{d} N(0, \bar{\sigma}^2) \quad \text{as } n \rightarrow \infty.$$

Proof. It can be shown in a way similar to the proof of Lemma 1. □

Lemma 5. *Under the conditions of Theorem 2, we have*

$$\frac{1}{n} \sum_{i=1}^n \left(Y_i - \frac{R(F)}{\mu} X_i \right)^2 \xrightarrow{p} \bar{\sigma}^2 \quad \text{as } n \rightarrow \infty.$$

Proof. It can be proved in a similar way to the proof of Lemma 2. □

Proof of Theorem 2. This can be done in a way similar to the proof of Theorem 1. □

5 Conclusions

This paper proposes jackknife empirical likelihood methods to construct confidence intervals for the risk measures and some related quantities studied by Jones and Zitikis (2003). Unlike the normal approximation method, the new methods do not need to estimate the asymptotic variance explicitly and are easy to implement by employing the R package 'emplik'. A simulation study shows that the jackknife empirical likelihood confidence intervals are more accurate than the normal approximation based confidence intervals.

Acknowledgments

Peng's research was supported by NSA grant H98230-10-1-0170, Qi's research was supported by NSA grant H98230-10-1-0161, and Yang's research was partly supported by the National Basic Research Program (973 Program) of China (2007CB814905) and National Natural Science Foundation of China (Grants No. 10871008).

References

- [1] P. Artzner (1999). Application of coherent risk measures to capital requirements in insurance. *North American Actuarial Journal* 3, 11–29.
- [2] J. Chen, L. Peng and Y. Zhao (2009). Empirical likelihood based confidence intervals for copulas. *J. Multivariate Anal.* 100, 137 - 151.
- [3] G. Claeskens, B. Jing, L. Peng and W. Zhou (2003). Empirical likelihood confidence regions for comparison distributions and ROC curves. *The Canadian Journal of Statistics.* 31(2), 173 - 190.
- [4] M. Csorgo and L. Horvath (1993). Weighted Approximations in Probability and Statistics. *Wiley, Chichester.*
- [5] B. Jing, J. Yuan and W. Zhou (2009). Jackknife empirical likelihood. *Journal of American Statistical Association* 104, 1224–1232.
- [6] B.L. Jones and R. Zitikis (2003). Empirical estimation of risk measures and related quantities. *North American Actuarial Journal* 7, 44–54.
- [7] B.L. Jones and R. Zitikis (2005). Testing for the order of risk measures: an application of L-statistics in an actuarial science. *Metron LXIII*, 193–211.
- [8] B.L. Jones and R. Zitikis (2007). Risk measures, distortion parameters, and their empirical estimation. *Insurance: Mathematics and Economics* 41, 279–297.
- [9] S.A. Klugman, H.H. Panjer and G.E. Willmot (2008). Loss Distributions: From Data to Decisions. *Wiley.*

- [10] A. Necir and D. Meraghni (2009). Empirical estimation of the proportional hazard premium for heavy-tailed claim amounts. *Insurance: Mathematics and Economics* 45, 49–58.
- [11] A. Owen (2001). Empirical Likelihood. *Chapman & Hall/CRC*.
- [12] G.R. Shorack and J.A. Wellner (1986). Empirical Processes with Applications to Statistics. *New York: Wiley*.
- [13] S. Wang (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics* 17, 43–54.
- [14] S. Wang (1996). Premium calculation by transforming the layer premium density. *ASTIN Bulletin* 26, 71–92.
- [15] S. Wang (1998). An actuarial index of the right-tail risk. *North American Actuarial Journal* 2, 88–101.
- [16] S. Wang, V.R. Young and H.H. Panjer (1997). Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics* 21, 173–183.
- [17] J.L. Wirch and M.R. Hardy (1999). A synthesis of risk measures for capital adequacy. *Insurance: Mathematics and Economics* 25, 337–347.
- [18] M.E. Yaari (1987). The dual theory of choice under risk. *Econometrica* 55, 95–115.

Table 1: Coverage probabilities for $R(F)$ are reported for the intervals based on the proposed jackknife empirical likelihood method (JELCI) and the normal approximation method (NACI).

(n, a, F)	JELCI	NACI	JELCI	NACI	JELCI	NACI
	level 0.9	level 0.9	level 0.95	level 0.95	level 0.99	level 0.99
$(300, 0.55, F_1(; 4))$	0.6089	0.4279	0.6965	0.4857	0.8208	0.5879
$(300, 0.85, F_1(; 4))$	0.8602	0.8431	0.9200	0.8992	0.9771	0.9572
$(1000, 0.55, F_1(; 4))$	0.6269	0.4516	0.7101	0.5130	0.8385	0.6188
$(1000, 0.85, F_1(; 4))$	0.8716	0.8628	0.9276	0.9185	0.9793	0.9696
$(300, 0.55, F_2(; 0, 1))$	0.6708	0.5266	0.7535	0.5894	0.8651	0.6852
$(300, 0.85, F_2(; 0, 1))$	0.8518	0.8399	0.9139	0.8943	0.9750	0.9532
$(1000, 0.55, F_2(; 0, 1))$	0.7132	0.5906	0.7901	0.6557	0.8931	0.7577
$(1000, 0.85, F_2(; 0, 1))$	0.8698	0.8628	0.9301	0.9170	0.9818	0.9719
$(300, 0.55, F_3(; 4, 1))$	0.8934	0.8698	0.9450	0.9291	0.9863	0.9808
$(300, 0.85, F_3(; 4, 1))$	0.8996	0.8975	0.9518	0.9499	0.9877	0.9870
$(1000, 0.55, F_3(; 4, 1))$	0.8939	0.8816	0.9468	0.9396	0.9872	0.9853
$(1000, 0.85, F_3(; 4, 1))$	0.8997	0.8990	0.9516	0.9510	0.9896	0.9888
$(300, 0.55, F_4(; 4, 1))$	0.8646	0.8120	0.9214	0.8751	0.9749	0.9458
$(300, 0.85, F_4(; 4, 1))$	0.8950	0.8917	0.9480	0.9428	0.9893	0.9869
$(1000, 0.55, F_4(; 4, 1))$	0.8777	0.8458	0.9285	0.9077	0.9821	0.9672
$(1000, 0.85, F_4(; 4, 1))$	0.9006	0.9002	0.9494	0.9487	0.9906	0.9880

Table 2: Coverage probabilities for $r(F)$ are reported for the intervals based on the proposed jackknife empirical likelihood method (JELCI) and the naive bootstrap method (NBCI).

(n, a, F)	JELCI	NBCI	JELCI	NBCI	JELCI	NBCI
	level 0.9	level 0.9	level 0.95	level 0.95	level 0.99	level 0.99
$(300, 0.55, F_1(; 4))$	0.5080	0.3755	0.5849	0.4215	0.7123	0.5005
$(300, 0.85, F_1(; 4))$	0.7350	0.6880	0.8074	0.7498	0.9063	0.8238
$(1000, 0.55, F_1(; 4))$	0.5566	0.4226	0.6363	0.4727	0.7625	0.5594
$(1000, 0.85, F_1(; 4))$	0.7871	0.7488	0.8566	0.8075	0.9408	0.8802
$(300, 0.55, F_2(; 0, 1))$	0.5187	0.4103	0.5934	0.4617	0.7054	0.5470
$(300, 0.85, F_2(; 0, 1))$	0.7014	0.6377	0.7715	0.7051	0.8701	0.7980
$(1000, 0.55, F_2(; 0, 1))$	0.6217	0.5229	0.6943	0.5885	0.8077	0.6820
$(1000, 0.85, F_2(; 0, 1))$	0.7818	0.7449	0.8523	0.8074	0.9328	0.8856
$(300, 0.55, F_3(; 4, 1))$	0.8597	0.8405	0.9222	0.8995	0.9768	0.9636
$(300, 0.85, F_3(; 4, 1))$	0.8930	0.8820	0.9461	0.9348	0.9881	0.9809
$(1000, 0.55, F_3(; 4, 1))$	0.8745	0.8682	0.9311	0.9232	0.9832	0.9781
$(1000, 0.85, F_3(; 4, 1))$	0.8990	0.8945	0.9516	0.9465	0.9904	0.9871
$(300, 0.55, F_4(; 4, 1))$	0.7711	0.7190	0.8424	0.7913	0.9277	0.8797
$(300, 0.85, F_4(; 4, 1))$	0.8565	0.8396	0.9188	0.8975	0.9748	0.9601
$(1000, 0.55, F_4(; 4, 1))$	0.8115	0.7766	0.8763	0.8465	0.9520	0.9245
$(1000, 0.85, F_4(; 4, 1))$	0.8775	0.8695	0.9339	0.9238	0.9837	0.9749