Jackknife Empirical Likelihood Intervals for Spearman's Rho

Ruodu Wang[1,2] and Liang Peng[1,2]

Abstract

In connection with copulas, rank correlation such as Kendall's tau and Spearman's rho has been employed in risk management for summarizing dependence among two variables and estimating some parameters in bivariate copulas and elliptical models. In this paper, a jackknife empirical likelihood method is proposed to construct confidence intervals for Spearman's rho without estimating the asymptotic variance. A simulation study confirms the advantages of the proposed method.

Keywords: Empirical Likelihood Method; Jackknife; Spearman's rho.

1 Introduction

Correlation has been used to summarize dependence among variables for a long history and plays an important role in modern finance such as Capital Asset Pricing Model and portfolio selection. Given the fact that copula and elliptical distributions have been heavily employed in risk management, copula-based dependence measures such as Kendall's tau and Spearman's rho are receiving more and more attention. Some pitfalls on using the linear correlation measure in elliptical models are given in Embrechts, McNeil and Straumann (2002). Advantages of using Kendall's tau and Spearman's rho include estimating some parameters in copulas. For example, if (X, Y) is a bivariate meta-Gaussian distribution with copula

$$C_\rho^{Ga}(u_1,u_2)=\int_{-\infty}^{\Phi^{-1}(u_1)}\int_{-\infty}^{\Phi^{-1}(u_2)}\frac{1}{2\pi(1-\rho^2)^{1/2}}\exp\{-\frac{s_1^2-2\rho s_1s_2+s_2^2}{2(1-\rho^2)}\}\mathrm{d}s_1\mathrm{d}s_2$$

[1]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA.

[2]We thank two reviewers for helpful comments. Research was supported by NSA grant H98230-10-1-0170 and NSF grant DMS-1005336.

and continuous marginals, where Φ^{-1} denotes the inverse function of the standard normal distribution function, then the Kendall's tau and Spearman's rho can be written as

$$\rho^\tau = \frac{2}{\pi}\arcsin\rho \quad \text{and} \quad \rho^s = \frac{6}{\pi}\arcsin\frac{\rho}{2}.$$

Therefore ρ can be estimated via estimating ρ^τ and ρ^s. More details can be found in Chapter 5.3 of McNeil, Frey and Embrechts (2005).

Let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be independent random vectors with distribution function H and continuous marginals $F(x) = H(x, \infty)$ and $G(y) = H(\infty, y)$. Then the Kendall's tau and Spearman's rho are defined as

$$\rho^\tau = \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

and

$$\rho^s = 12\mathbb{E}[(F(X_1) - 1/2)(G(Y_1) - 1/2)],$$

respectively. Define

$$F_n(x) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(X_i \le x) \quad \text{and} \quad G_n(x) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}(Y_i \le x).$$

Then the simple nonparametric estimators for ρ^τ and ρ^s are

$$\hat\rho_n^\tau = \frac{2}{n(n-1)}\sum_{1\le i<j\le n}\{\mathbf{1}((X_i - X_j)(Y_i - Y_j) > 0) - \mathbf{1}((X_i - X_j)(Y_i - Y_j) < 0)\}$$

and

$$\hat\rho_n^s = \frac{12}{n}\sum_{i=1}^n \{F_n(X_i) - 1/2\}\{G_n(Y_i) - 1/2\},$$

respectively.

In order to construct confidence intervals for ρ^τ and ρ^s, one can simply use the asymptotic limits of $\sqrt{n}\{\hat\rho_n^\tau - \rho^\tau\}$ and $\sqrt{n}\{\hat\rho_n^s - \rho^s\}$. However, this method requires to estimate the asymptotic variances. As shown in the next section, the asymptotic variance of $\hat\rho_n^s$ is quite complicated and it is hard to estimate it explicitly. Most likely, it involves density estimation and numerical integration. Therefore, bootstrap method is a common way to construct a confidence interval for the Spearman's rho. As an alternative way of constructing confidence intervals, empirical likelihood method introduced in Owen (1988, 1990) is powerful in dealing with linear functionals without estimating any extra quantities such as asymptotic variance. We refer to Owen (2001) for an overview on empirical likelihood method. Since the Kendall's tau and Spearman's rho are non-linear functionals, a direct application of empirical likelihood method fails to obtaining a chi-square limit. Recently, Jing, Yuan and Zhou (2009) proposed a so-called jackknife empirical likelihood method to construct confidence intervals for U-statistics. Since $\hat{\rho}_n^{\tau}$ is a U-statistic, one can directly employ the jackknife empirical likelihood method in Jing, Yuan and Zhou (2009) to construct confidence intervals for the Kendall's tau without estimating the asymptotic variance. In this paper, we employ the jackknife empirical likelihood method to construct confidence intervals for the Spearman's rho and investigate the finite sample behavior of the proposed method.

We organize this paper as follows. Section 2 presents the methodology and asymptotic results. A simulation study and a real data analysis are given in Section 3. All proofs are put in Section 4.

2 Methodology

Define the copula and empirical copula of (X_i, Y_i) as

$$C(x, y) = \mathbb{P}(F(X_1) \le x, G(Y_1) \le y)$$

and

$$C_n(x, y) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}(F_n(X_i) \le x, G_n(Y_i) \le y),$$

respectively. Put

$$C_1(x, y) = \frac{\partial}{\partial x} C(x, y) \quad \text{and} \quad C_2(x, y) = \frac{\partial}{\partial y} C(x, y).$$

Assume that

$$\begin{cases} C_1(x,y) \quad \text{exists and is continuous on the set} \quad \{(x,y): 0 < x < 1, 0 \le y \le 1\}, \\ C_2(x,y) \quad \text{exists and is continuous on the set} \quad \{(x,y): 0 \le x \le 1, 0 < y < 1\}. \end{cases} \tag{1}$$

Then it follows from Proposition 3.1 of Segers (2010) that

$$\sup_{0\le x,y\le 1} |\sqrt{n}\{C_n(x,y) - C(x,y)\} - W(x,y) + C_1(x,y)W(x,1) + C_2(x,y)W(1,y)| = o_p(1), \tag{2}$$

where $W(x,y)$ is a Gaussian process with mean zero and covariance

$$\mathbb{E}[W(x_1,y_1)W(x_2,y_2)] = C(x_1 \wedge x_2, y_1 \wedge y_2) - C(x_1,y_1)C(x_2,y_2). \tag{3}$$

Note that (2) holds via the Skorohod construction. By (2), we have

$$\begin{aligned} \sqrt{n}\{\hat{\rho}_n^s - \rho^s\} &= 12\int_0^1\int_0^1 \sqrt{n}\{C_n(x,y) - C(x,y)\}\mathrm{d}x\mathrm{d}y \\ &\xrightarrow{d} 12\int_0^1\int_0^1\{W(x,y) - C_1(x,y)W(x,1) - C_2(x,y)W(1,y)\}\mathrm{d}x\mathrm{d}y. \end{aligned}$$

Hence, the asymptotic limit depends on the copula $C(x,y)$ and its partial derivatives. In order to avoid estimating the complicated asymptotic variance for constructing confidence intervals for ρ^s, we employ the following jackknife empirical likelihood method.

Define

$$\begin{cases} F_{n,i}(x) = \frac{1}{n-1}\sum_{j=1,j\neq i}^n I(X_j \le x), \\ G_{n,i}(x) = \frac{1}{n-1}\sum_{j=1,j\neq i}^n I(Y_j \le x), \\ \hat{\rho}_{n,i}^s = \frac{12}{n-1}\sum_{j=1,j\neq i}^n \{F_{n,i}(X_j) - 1/2\}\{G_{n,i}(Y_j) - 1/2\} \\ Z_i = n\hat{\rho}_n^s - (n-1)\hat{\rho}_{n,i}^s \end{cases}$$

for $i = 1, \cdots, n$. As in Jing, Yuan and Zhou (2009), a jackknife empirical likelihood function for $\theta = \rho^s$ is defined as

$$L(\theta) = \sup\{\prod_{i=1}^n (np_i) : p_1 \ge 0, \cdots, p_n \ge 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i Z_i = \theta\}.$$

By the Lagrange multiplier technique, we obtain that $p_i = n^{-1}\{1 + \lambda(Z_i - \theta)\}^{-1}$ and $-2\log L(\theta) = 2\sum_{i=1}^n \log\{1 + \lambda(Z_i - \theta)\}$, where $\lambda = \lambda(\theta)$ satisfies

$$\frac{1}{n}\sum_{i=1}^n \frac{Z_i - \theta}{1 + \lambda(Z_i - \theta)} = 0. \tag{4}$$

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

Theorem 1. Assume condition (1) holds. Then $-2\log L(\rho^s)$ converges in distribution to a chi-square distribution with one degree of freedom as $n \to \infty$.

Based on the above theorem, a jackknife empirical likelihood confidence interval for ρ^s with level α can be obtained as

$$I_\alpha = \{\theta : -2\log L(\theta) \le \chi^2_{1,\alpha}\},$$

where $\chi^2_{1,\alpha}$ denotes the α quantile of a chi-square distribution with one degree of freedom.

3 Simulation study and data analysis

3.1 Simulation study

We investigate the finite sample behavior of the proposed jackknife empirical likelihood method and compare it with the normal approximation method in terms of coverage accuracy.

We draw $10,000$ random samples of sample size $n = 100, 300$ from a bivariate normal distribution with correlation ρ and marginals being the standard normal distribution. In this case, the Spearman's rho equals $\frac{6}{\pi}\arcsin(\rho/2)$. We calculate the jackknife empirical likelihood interval I_α at levels $\alpha = 0.9, 0.95, 0.99$ for $\rho = 0, \pm 0.2, \pm 0.8$, which correspond to $\rho^s = 0, \pm 0.1913, \pm 0.7859$, respectively. For constructing a confidence interval based on the asymptotic limit of $\hat{\rho}^s_n$, we employ the percentile bootstrap confidence interval. More specifically, we draw $1,000$ bootstrap samples of size n from each original sample. Based on each bootstrap sample, we calculate the Spearman's rho estimator. Therefore we obtained $1,000$ bootstrapped Spearman's rho estimators denoted by $\hat{\rho}^{s*}_{n,1}, \cdots, \hat{\rho}^{s*}_{n,1000}$. Let c_1 and c_2 denote the $[1000(1-\alpha)/2] - th$ and $[1000(1+\alpha)/2] - th$ largest order statistics of $\{\hat{\rho}^{s*}_{n,i} - \hat{\rho}^s_n\}_{i=1}^{1000}$. Hence, the percentile bootstrap confidence interval for ρ^s with level α is

$$I^B_\alpha = (\hat{\rho}^s_n - c_2, \quad \hat{\rho}^s_n - c_1).$$

The empirical coverage probabilities and average interval lengths for both I_α and I_α^B are reported in Tables 1 and 2, which show that i) the proposed jackknife empirical likelihood method produces much more accurate confidence intervals than the percentile bootstrap method in most cases, specially for $n = 100$; ii) the interval lengths of the jackknife empirical likelihood method are slightly longer.

3.2 Data analysis

Next, we apply the proposed method to the Danish fire insurance claims. This data set is available at *www.ma.hw.ac.uk/∼mcneil/*, which consists of loss to buildings, loss to contents and loss to profits. As described there, the data were collected at the Copenhagen Reinsurance Company and comprise 2167 fire losses over the period 1980 to 1990. They have been adjusted for inflation to reflect 1985 values and are expressed in millions of Danish Kroner. Here we consider the first two variables: loss to building and loss to contents; see Figure 1 below. For computing I_α^B, we draw $1,000$ bootstrap samples as before. We find that $\hat{\rho}_n^s = 0.1411$, $I_{0.9}^B = (0.0959, \quad 0.1866)$, $I_{0.95}^B = (0.0897, \quad 0.1942)$, $I_{0.9} = (0.0962, \quad 0.1862)$ and $I_{0.95} = (0.0882, \quad 0.1952)$, which show that the proposed jackknife empirical likelihood method produces similar interval length as the bootstrap method. Both intervals indicate the Spearman's rho is positive, which means that the loss to contents is positively correlated with the loss to profits.

4 Proofs

Before proving Theorem 1, we show the following two lemmas.

Lemma 1. Under conditions of Theorem 1, we have

$$\begin{aligned}\sqrt{n}\{\tfrac{1}{n}\sum_{i=1}^n Z_i - \rho^s\} &\xrightarrow{d} 12\int_0^1\int_0^1\{W(x,y) - C_1(x,y)W(x,1) - C_2(x,y)W(1,y)\}\mathrm{d}x\mathrm{d}y \\ &\stackrel{d}{=} N(0,\sigma^2)\end{aligned}$$

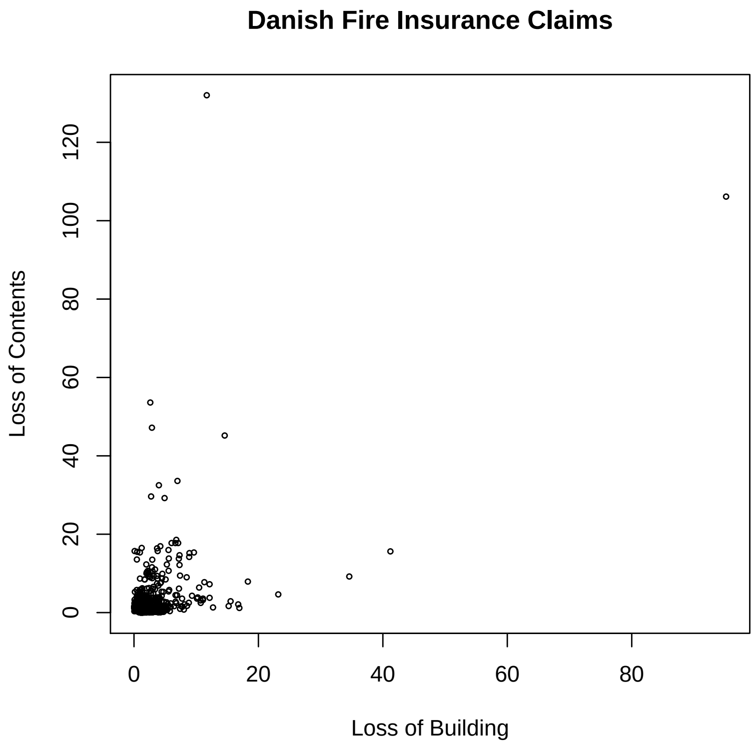

Figure 1: Scatterplot of the Danish fire insurance data.

as $n \to \infty$, where

$$\sigma^2 = \mathbb{E}[(12 \int_0^1 \int_0^1 \{W(x,y) - C_1(x,y)W(x,1) - C_2(x,y)W(1,y)\} \mathrm{d}x \mathrm{d}y)^2].$$

Proof. For $i = 1, \cdots, n$, write

$$
\begin{aligned}
& Z_i - \rho^s \\
= \;& n\hat{\rho}_n^s - (n-1)\hat{\rho}_{n,i}^s - \rho^s \\
= \;& 12\sum_{j=1}^{n}(F_n(X_j) - 1/2)(G_n(Y_j) - 1/2) - 12\sum_{j\neq i}(F_{n,i}(X_j) - 1/2)(G_{n,i}(Y_j) - 1/2) - \rho^s \\
= \;& 12\sum_{j\neq i}[(F_n(X_j) - 1/2)(G_n(Y_j) - 1/2) - (F_{n,i}(X_j) - 1/2)(G_{n,i}(Y_j) - 1/2)] \\
& +(12(F_n(X_i) - 1/2)(G_n(Y_i) - 1/2) - \rho^s) \\
= \;& 12\left(\sum_{j\neq i}(F_n(X_j) - F_{n,i}(X_j))(G_n(Y_j) - 1/2) + \sum_{j\neq i}(G_n(Y_j) - G_{n,i}(Y_j))(F_n(X_j) - 1/2)\right. \\
& \left. + \sum_{j\neq i}(F_n(X_j) - F_{n,i}(X_j))(G_n(Y_j) - G_{n,i}(Y_j))\right) + (12(F_n(X_i) - 1/2)(G_n(Y_i) - 1/2) - \rho^s) \\
= \;& 12\left(\sum_{j\neq i}(F_n(X_j) - F_{n,i}(X_j))(G_n(Y_j) - 1/2) + \sum_{j\neq i}(G_n(Y_j) - G_{n,i}(Y_j))(F_n(X_j) - 1/2)\right) \\
& +O(1/n) + (12(F_n(X_i) - 1/2)(G_n(Y_i) - 1/2) - \rho^s) \\
= \;& 12\{\underbrace{\sum_{j=1}^{n}(F_n(X_j) - F_{n,i}(X_j))(G_n(Y_j) - 1/2)}_{V_{1,i}} + \underbrace{\sum_{j=1}^{n}(G_n(Y_j) - G_{n,i}(Y_j))(F_n(X_j) - 1/2)}_{V_{2,i}} \\
& + \underbrace{((F_n(X_i) - 1/2)(G_n(Y_i) - 1/2) - \rho^s/12)}_{V_{3,i}}\} + O(1/n).
\end{aligned}
$$

Thus

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(Z_i - \rho^s) = \frac{12}{\sqrt{n}}\sum_{i=1}^{n}(V_{i,1} + V_{i,2} + V_{i,3}) + O(1/\sqrt{n}).
$$

We already have

$$
\begin{aligned}
\frac{12}{\sqrt{n}}\sum_{i=1}^{n} V_{3,i} &= \sqrt{n}(\hat{\rho}_n^s - \rho^s) \\
&\xrightarrow{d} 12\int_0^1\int_0^1 \{W(x,y) - C_1(x,y)W(x,1) - C_2(x,y)W(1,y)\}\mathrm{d}x\mathrm{d}y.
\end{aligned}
$$

It is easy to check that

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{i=1}^{n} V_{1,i} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\sum_{j=1}^{n}(F_n(X_j)-F_{n,i}(X_j))(G_n(Y_j)-1/2) \\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(G_n(Y_j)-1/2)\sum_{i=1}^{n}(F_n(X_j)-F_{n,i}(X_j)) \\
&= \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(G_n(Y_j)-1/2)\times 0 \\
&= 0.
\end{aligned}$$

Similarly

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} V_{2,i} = 0.$$

Thus it follows from the above equations that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(Z_i-\rho^s) \xrightarrow{d} 12\int_0^1\int_0^1\{W(x,y)-C_1(x,y)W(x,1)-C_2(x,y)W(1,y)\}\mathrm{d}x\mathrm{d}y.$$

□

Lemma 2. Under conditions of Theorem 1, we have

$$\frac{1}{n}\sum_{i=1}^{n}(Z_i-\rho^s)^2 \xrightarrow{p} \sigma^2 \quad \text{as} \quad n\to\infty.$$

Proof. Write

$$\begin{aligned}
\sigma^2 &= \mathbb{E}[(12\int_0^1\int_0^1\{W(x,y)-C_1(x,y)W(x,1)-C_2(x,y)W(1,y)\}\mathrm{d}x\mathrm{d}y)^2] \\
&= 144\mathbb{E}[(\underbrace{\iint W(x,y)\mathrm{d}x\mathrm{d}y}_{A_1}-\underbrace{\iint C_1(x,y)W(x,1)\mathrm{d}x\mathrm{d}y}_{A_2}-\underbrace{\iint C_2(x,y)W(1,y)\mathrm{d}x\mathrm{d}y}_{A_3})^2] \\
&= 144\mathbb{E}[A_1^2+A_2^2+A_3^2-2A_1A_2-2A_1A_3+2A_2A_3],
\end{aligned}$$

where by convention we use $\int = \int_0^1$. Using (3) we have

$$\begin{aligned}
\mathbb{E}(A_1^2) &= \iiiint (C(x_1\wedge x_2, y_1\wedge y_2)-C(x_1,y_1)C(x_2,y_2))\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}y_2 \\
&= 4\iint C(x,y)(1-x)(1-y)\mathrm{d}x\mathrm{d}y-\left(\iint C(x,y)\mathrm{d}x\mathrm{d}y\right)^2,
\end{aligned} \tag{5}$$

$$\mathbb{E}(A_2^2) = \iiiint C_1(x_1,y_1)C_1(x_2,y_2)(x_1\wedge x_2-x_1x_2)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}y_2, \tag{6}$$

and

$$\mathbb{E}(A_3^2) = \iiiint C_2(x_1, y_1)C_2(x_2, y_2)(y_1 \wedge y_2 - y_1y_2)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}y_2. \tag{7}$$

By integration by parts, we have

$$\begin{aligned}
& \textstyle\iiint C_1(x_1, y_1)C(x_1, y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 \\
= & \textstyle\iint C(1, y_1)C(1, y_2)\mathrm{d}y_1\mathrm{d}y_2 - \iiint C_1(x_1, y_2)C(x_1, y_1)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 \\
= & \textstyle\frac{1}{4} - \iiint C_1(x_1, y_2)C(x_1, y_1)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2,
\end{aligned}$$

which implies that

$$\iiint C_1(x_1, y_1)C(x_1, y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 = \frac{1}{8}. \tag{8}$$

Similarly,

$$\begin{aligned}
& \iiint x_1C_1(x_1, y_1)C(x_1, y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 \\
= & \iint C(1, y_1)C(1, y_2)\mathrm{d}y_1\mathrm{d}y_2 - \iiint (x_1C_1(x_1, y_1)C(x_1, y_2) - C(x_1, y_1)C(x_1, y_2))\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2,
\end{aligned}$$

which implies that

$$2\iiint x_1C_1(x_1, y_1)C(x_1, y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 + \iiint C(x_1, y_1)C(x_1, y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 = \frac{1}{4}. \tag{9}$$

It follows from (8), (9) and (3) that

$$\begin{aligned}
\mathbb{E}(A_1A_2) = & \textstyle\iiint C_1(x_1, y_1)(C(x_1 \wedge x_2, y_2) - x_1C(x_2, y_2))\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}y_2 \\
= & \textstyle\iiint C_1(x_1, y_1)C(x_1, y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 - \iiint x_1C_1(x_1, y_1)C(x_1, y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 \\
& \textstyle - \iiint C(x_1, y_1)C(x_1, y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 + \left(\iint C(x, y)\mathrm{d}x\mathrm{d}y\right)^2 \\
= & \textstyle -\frac{1}{8} + \iiint x_1C_1(x_1, y_1)C(x_1, y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 + \left(\iint C(x, y)\mathrm{d}x\mathrm{d}y\right)^2.
\end{aligned} \tag{10}$$

Using the same arguments, we can show that

$$\mathbb{E}(A_1A_3) = -\frac{1}{8} + \iiint y_1C_2(x_1, y_1)C(x_2, y_1)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_1 + \left(\iint C(x, y)\mathrm{d}x\mathrm{d}y\right)^2, \tag{11}$$

and

$$\mathbb{E}(A_2A_3) = \iiiint C_1(x_1, y_1)C_2(x_2, y_2)(C(x_1, y_2) - x_1y_2)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}y_2. \tag{12}$$

Note that

$$\begin{aligned} \tfrac{1}{n}\sum_{i=1}^{n}(Z_i-\rho^s)^2 &= \tfrac{144}{n}\sum_{i=1}^{n}(V_{1,i}+V_{2,i}+V_{3,i}+O(1/n))^2 \\ &= \tfrac{144}{n}\sum_{i=1}^{n}(V_{1,i}+V_{2,i}+V_{3,i})^2+O(1/n) \end{aligned} \tag{13}$$

since $V_{1,i}$, $V_{2,i}$ and $V_{3,i}$ are uniformly bounded for $i=1,\cdots,n$. A straightforward calculation shows that

$$\begin{aligned} &\tfrac{1}{n}\sum_{i=1}^{n}V_{1,i}^2 \\ =\ &\tfrac{1}{n}\sum_{i=1}^{n}\Big(\sum_{j=1}^{n}(F_n(X_j)-F_{n,i}(X_j))(G_n(Y_j)-1/2)\Big)^2 \\ =\ &\tfrac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}(F_n(X_j)-F_{n,i}(X_j))(G_n(Y_j)-1/2)(F_n(X_k)-F_{n,i}(X_k))(G_n(Y_k)-1/2) \\ =\ &\tfrac{1}{n}\sum_{j=1}^{n}\sum_{k=1}^{n}(G_n(Y_j)-1/2)(G_n(Y_k)-1/2)\sum_{i=1}^{n}(F_n(X_j)-F_{n,i}(X_j))(F_n(X_k)-F_{n,i}(X_k)) \\ =\ &\tfrac{1}{(n-1)^2}\sum_{j=1}^{n}\sum_{k=1}^{n}(G_n(Y_j)-1/2)(G_n(Y_k)-1/2)(F_n(X_j\wedge X_k)-F_n(X_j)F_n(X_k)) \\ \xrightarrow{p}\ &\iiiint(y_1-\tfrac{1}{2})(y_2-\tfrac{1}{2})(x_1\wedge x_2-x_1x_2)C(\mathrm{d}x_1,\mathrm{d}y_1)C(\mathrm{d}x_2,\mathrm{d}y_2) \\ =\ &\iint\tfrac{1}{4}(x_1\wedge x_2-x_1x_2)\mathrm{d}x_1\mathrm{d}x_2-\iiint(x_1\wedge x_2-x_1x_2)C_1(x_1,y_1)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_1 \\ &+\iiiint C_1(x_1,y_1)C_1(x_2,y_2)(x_1\wedge x_2-x_1x_2)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}y_2 \\ =\ &\tfrac{1}{48}+\tfrac{1}{2}\iint C(x,y)\mathrm{d}x\mathrm{d}y-\iint xC(x,y)\mathrm{d}x\mathrm{d}y+\mathbb{E}(A_2^2) \end{aligned} \tag{14}$$

as $n\to\infty$. Similarly, we can show that

$$\frac{1}{n}\sum_{i=1}^{n}V_{2,i}^2\xrightarrow{p}\frac{1}{48}+\frac{1}{2}\iint C(x,y)\mathrm{d}x\mathrm{d}y-\iint yC(x,y)\mathrm{d}x\mathrm{d}y+\mathbb{E}(A_3^2), \tag{15}$$

$$\begin{aligned} &\tfrac{1}{n}\sum_{i=1}^{n}V_{3,i}^2 \\ \xrightarrow{p}\ &\iint(x-\tfrac{1}{2})^2(y-\tfrac{1}{2})^2C(\mathrm{d}x,\mathrm{d}y)-(\rho^s/12)^2 \\ =\ &\iint(x-\tfrac{1}{2})(y-\tfrac{1}{2})C(x,y)\mathrm{d}x\mathrm{d}y-\tfrac{1}{48}-\big(\iint C(x,y)\mathrm{d}x\mathrm{d}y-\tfrac{1}{4}\big)^2 \\ =\ &4\iint(x-1)(y-1)C(x,y)\mathrm{d}x\mathrm{d}y+\iint 2(x+y)C(x,y)\mathrm{d}x\mathrm{d}y-3\iint C(x,y)\mathrm{d}x\mathrm{d}y \\ &-\tfrac{1}{48}-\big(\iint C(x,y)\mathrm{d}x\mathrm{d}y\big)^2+\tfrac{1}{2}\iint C(x,y)\mathrm{d}x\mathrm{d}y-\tfrac{1}{16} \\ =\ &\mathbb{E}(A_1^2)+\iint 2(x+y)C(x,y)\mathrm{d}x\mathrm{d}y-\tfrac{5}{2}\iint C(x,y)\mathrm{d}x\mathrm{d}y-\tfrac{1}{12}, \end{aligned} \tag{16}$$

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} V_{1,i}V_{2,i} \\
\xrightarrow{p}\quad & \iiiint (x_2-\tfrac{1}{2})(y_1-\tfrac{1}{2})(C(x_1,y_2)-x_1y_2)C(\mathrm{d}x_1,\mathrm{d}y_1)C(\mathrm{d}x_2,\mathrm{d}y_2) \\
=\quad & \iint \tfrac{1}{4}(C(x_1,y_2)-x_1y_2)\mathrm{d}x_1\mathrm{d}y_2 - \iiint \tfrac{1}{2}C_2(x_2,y_2)(C(x_1,y_2)-x_1y_2)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_2 \\
& -\iiint \tfrac{1}{2}C_1(x_1,y_1)(C(x_2,y_1)-x_2y_1)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 \\
& +\iiiint C_1(x_1,y_1)C_2(x_2,y_2)(C(x_1,y_2)-x_1y_2)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}y_1\mathrm{d}y_2 \\
=\quad & \tfrac{3}{16}-\tfrac{1}{4}\iint C(x,y)\mathrm{d}x\mathrm{d}y - \tfrac{1}{2}\iint C_2(x_2,y_2)C(x_1,y_2)dx_1\mathrm{d}x_2\mathrm{d}y_2 \\
& -\tfrac{1}{2}\iint C_1(x_1,y_1)C(x_2,y_1)dx_1\mathrm{d}y_1\mathrm{d}y_2 + \mathbb{E}(A_2A_3),
\end{aligned} \tag{17}
$$

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} V_{1,i}V_{3,i} \\
=\quad & \tfrac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{n}(F_n(X_j)-F_{n,i}(X_j))[(F_n(X_i)-\tfrac{1}{2})(G_n(Y_i)-\tfrac{1}{2})-\rho^2/12](G_n(Y_j)-\tfrac{1}{2}) \\
=\quad & \tfrac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}[\mathbf{1}(X_i\le X_j)-F_n(X_j)][(F_n(X_i)-\tfrac{1}{2})(G_n(Y_i)-\tfrac{1}{2})-\rho^2/12](G_n(Y_j)-\tfrac{1}{2}) \\
\xrightarrow{p}\quad & \iiiint_{x_2=x_1}^{1}(y_2-\tfrac{1}{2})C(\mathrm{d}x_2,\mathrm{d}y_2)[(x_1-\tfrac{1}{2})(y_1-\tfrac{1}{2})-\rho^s/12]C(\mathrm{d}x_1,\mathrm{d}y_1) \\
=\quad & \iiint (y_2-\tfrac{1}{2})(1-C_2(x_1,y_2))\mathrm{d}y_2[(x_1-\tfrac{1}{2})(y_1-\tfrac{1}{2})-\rho^s/12]C(\mathrm{d}x_1,\mathrm{d}y_1) \\
=\quad & \iiint -(y_2-\tfrac{1}{2})C_2(x_1,y_2)\mathrm{d}y_2(x_1-\tfrac{1}{2})(y_1-\tfrac{1}{2})C(\mathrm{d}x_1,\mathrm{d}y_1)-(\rho^s/12)^2 \\
=\quad & \iint \left(\int C(x_1,y_2)\mathrm{d}y_2-\tfrac{1}{2}x_1\right)(x_1-\tfrac{1}{2})(y_1-\tfrac{1}{2})C(\mathrm{d}x_1,\mathrm{d}y_1)-(\rho^s/12)^2 \\
=\quad & \tfrac{1}{2}\iint x_1C(x_1,y_2))\mathrm{d}x_1\mathrm{d}y_2-\tfrac{1}{4}\iint C(x,y)\mathrm{d}x\mathrm{d}y-\tfrac{1}{12}+\tfrac{1}{16}+\iint x_1C_1(x_1,y_1)(x_1-\tfrac{1}{2})\mathrm{d}x_1\mathrm{d}y_1 \\
& -\iiint C_1(x_1,y_1)C(x_1,y_2)(x_1-\tfrac{1}{2})\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2-(\rho^s/12)^2 \\
=\quad & \tfrac{1}{2}\iint xC(x,y)\mathrm{d}x\mathrm{d}y-\tfrac{1}{48}+\tfrac{1}{4}-\iint xC(x,y)\mathrm{d}x\mathrm{d}y-\tfrac{1}{8}-\iiint x_1C_1(x_1,y_1)C(x_1,y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}x_2 \\
& +\tfrac{1}{2}\iiint C_1(x_1,y_1)C(x_1,y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2-\left(\iint C(x,y)\mathrm{d}x\mathrm{d}y-\tfrac{1}{4}\right)^2 \\
=\quad & \tfrac{1}{24}-\tfrac{1}{2}\iint xC(x,y)\mathrm{d}x\mathrm{d}y-\iiint x_1C_1(x_1,y_1)C(x_1,y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}x_2 \\
& +\tfrac{1}{2}\iiint C_1(x_1,y_1)C(x_1,y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2+\tfrac{1}{2}\iint C(x,y)\mathrm{d}x\mathrm{d}y-\left(\iint C(x,y)\mathrm{d}x\mathrm{d}y\right)^2 \\
=\quad & -\tfrac{1}{12}-\tfrac{1}{2}\iint xC(x,y)\mathrm{d}x\mathrm{d}y+\tfrac{1}{2}\iiint C_1(x_1,y_1)C(x_1,y_2)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 \\
& +\tfrac{1}{2}\iint C(x,y)\mathrm{d}x\mathrm{d}y-\mathbb{E}(A_1A_2),
\end{aligned} \tag{18}
$$

and

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n} V_{2,i}V_{3,i} \xrightarrow{p} & -\frac{1}{12} - \frac{1}{2}\iint yC(x,y)\mathrm{d}x\mathrm{d}y + \frac{1}{2}\iiint C_2(x_1,y_1)C(x_2,y_1)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}y_2 \\ & +\frac{1}{2}\iint C(x,y)\mathrm{d}x\mathrm{d}y - \mathbb{E}(A_1A_3).\end{aligned} \tag{19}$$

Hence, it follows from (5)–(7), (10)–(19) that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}(V_{1,i}+V_{2,i}+V_{3,i})^2 &= \frac{1}{n}\sum_{i=1}^{n}(V_{1,i}^2+V_{2,i}^2+V_{3,i}^2+2V_{1,i}V_{2,i}+2V_{1,i}V_{3,i}+2V_{2,i}V_{3,i}) \\ &\xrightarrow{p} \mathbb{E}(A_1^2)+\mathbb{E}(A_2^2)+\mathbb{E}(A_3^2)-2\mathbb{E}(A_1A_2)-2\mathbb{E}(A_1A_3)+\mathbb{E}(A_2A_3),\end{aligned}$$

i.e.,

$$\frac{1}{n}\sum_{i=1}^{n}(Z_i-\rho^s)^2 \xrightarrow{p} \sigma^2.$$

□

Proof of Theorem 1. Since $V_{1,i}$, $V_{2,i}$ and $V_{3,i}$ defined in the proof of Lemma 1 are uniformly bounded for $i = 1, \cdots, n$, we have $\sup_{1\le i\le n}|Z_i|$ is bounded. Hence, using the standard arguments in the empirical likelihood method (see Chapter 11 of Owen (2001)), Lemmas 1 and 2, we obtain that

$$-2\log L(\rho^s) = \frac{\{\sum_{i=1}^{n}(Z_i-\rho^s)\}^2}{\sum_{i=1}^{n}(Z_i-\rho^s)^2} + o_p(1) \xrightarrow{d} \chi^2(1).$$

□

References

[1] Embrechts, P., McNeil, A.J. and Straumann, D. (2002). Correlation and dependency in risk management: properties and pitfalls. In *Risk Management: Value at Risk and Beyond (ed. M. Dempster), pp. 176–223. Cambridge University Press.*

[2] Jing, B, Yuan, J. and Zhou, W. (2009). Jackknife empirical likelihood. *J. Amer. Statist. Assoc.* 104, 1224–1232.

[3] McNeil, A.J., Frey, R. and Embrechts, P. (2005). Quantitative Risk Management: Concepts, Techniques and Tools. *Princeton University Press.*

[4] Owen, A. (1988). Empirical likelihood ratio confidence intervals for single functional. *Biometrika* 75, 237–249.

[5] Owen, A. (1990). Empirical likelihood ratio confidence regions. *Ann. Statist.* 18, 90–120.

[6] Owen, A. (2001). *Empirical Likelihood.* Chapman & Hall/CRC.

[7] Segers, J. (2010). Weak convergence of empirical copula processes under nonrestrictive smoothness assumptions. http://arxiv.org/abs/1012.2133

Table 1: Coverage probabilities for the intervals I_α and I_α^B at levels $\alpha = 0.9, 0.95, 0.99$ are reported for $n = 100, 300$ and $\rho = 0, \pm 0.2, \pm 0.8$.

(n, ρ)	$I_{0.9}$	$I_{0.9}^B$	$I_{0.95}$	$I_{0.95}^B$	$I_{0.99}$	$I_{0.99}^B$
(100, 0)	0.9024	0.8874	0.9524	0.9352	0.9898	0.9794
(100, 0.2)	0.9016	0.8867	0.9524	0.9349	0.9900	0.9791
(100, −0.2)	0.9003	0.8858	0.9513	0.9347	0.9896	0.9773
(100, 0.8)	0.9013	0.8876	0.9473	0.9264	0.9850	0.9624
(100, −0.8)	0.8926	0.8691	0.9390	0.9105	0.9818	0.9509
(300, 0)	0.9055	0.8999	0.9530	0.9476	0.9915	0.9864
(300, 0.2)	0.9035	0.8996	0.9513	0.9440	0.9906	0.9852
(300, −0.2)	0.9073	0.9017	0.9529	0.9467	0.9908	0.9860
(300, 0.8)	0.9037	0.8957	0.9529	0.9393	0.9900	0.9776
(300, −0.8)	0.9008	0.8920	0.9505	0.9377	0.9899	0.9782

Table 2: Average interval lengths for I_α and I_α^B at levels $\alpha = 0.9, 0.95, 0.99$ are reported for $n = 100, 300$ and $\rho = 0, \pm 0.2, \pm 0.8$.

(n, ρ)	$I_{0.9}$	$I_{0.9}^B$	$I_{0.95}$	$I_{0.95}^B$	$I_{0.99}$	$I_{0.99}^B$
$(100, 0)$	0.337	0.332	0.403	0.394	0.529	0.515
$(100, 0.2)$	0.327	0.322	0.391	0.383	0.515	0.501
$(100, -0.2)$	0.327	0.322	0.390	0.382	0.515	0.499
$(100, 0.8)$	0.148	0.148	0.177	0.177	0.235	0.236
$(100, -0.8)$	0.147	0.147	0.176	0.175	0.234	0.231
$(300, 0)$	0.192	0.190	0.229	0.227	0.302	0.298
$(300, 0.2)$	0.186	0.185	0.222	0.220	0.293	0.289
$(300, -0.2)$	0.186	0.185	0.222	0.220	0.293	0.288
$(300, 0.8)$	0.083	0.082	0.099	0.098	0.130	0.129
$(300, -0.8)$	0.083	0.082	0.099	0.098	0.130	0.128