Jackknife Empirical Likelihood for Parametric Copulas

Ruodu Wang*, Liang Peng† and Jingping Yang‡

July 10, 2012

Abstract

For fitting a parametric copula to multivariate data, a popular way is to employ the so-called pseudo maximum likelihood estimation proposed by Genest, Ghoudi and Rivest (1995). Although interval estimation can be obtained via estimating the asymptotic covariance of the pseudo maximum likelihood estimate, we propose a jackknife empirical likelihood method to construct confidence regions for the parameters without estimating any additional quantities such as asymptotic covariance. A simulation study shows the advantages of the new method in case of strong dependence or having more than one parameter involved.

Key-words: Copulas; Empirical likelihood; Interval estimation; Jackknife

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

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

‡LMAM, Department of Financial Mathematics, Center for Statistical Science, Peking University, Beijing, 100871, China. Email address:yangjp@math.pku.edu.cn

1 Introduction

Let $\mathbf{X}_1 = (X_{1,1}, \cdots, X_{1,d})^T, \cdots, \mathbf{X}_n = (X_{n,1}, \cdots, X_{n,d})^T$ be independent random vectors with common distribution function F and continuous marginal distributions $F_1, \cdots, F_d$. Then the copula of $\mathbf{X_1}$ is defined as

$$C(x_1, \cdots, x_d) = F(F_1^-(x_1), \cdots, F_d^-(x_d)) \tag{1}$$

for $0 \le x_1, \cdots, x_d \le 1$, where F_j^- denotes the inverse of F_j. Since the copula is independent of marginals, it becomes a more or less standard tool in modeling dependence in risk management. Many research papers and review papers have appeared in the literature with particular applications in insurance, finance and risk management; see references in Haug, Klüppelberg and Peng (2011).

For fitting a family of parametric copulas $\{C(\cdot;\theta) : \theta \in \Theta \subset R^q\}$ to a data set, a popular semi-parametric estimation is the so-called pseudo maximum likelihood estimation proposed by Genest, Ghoudi and Rivest (1995). That is, $\hat{\theta} = \arg\max \bar{L}(\theta)$, where $\bar{L}(\theta)$ is the pseudo likelihood function for θ defined as

$$\bar{L}(\theta) = \prod_{i=1}^n c(\hat{F}_1(X_{i,1}), \cdots, \hat{F}_d(X_{i,d}); \theta), \tag{2}$$

where $c(\cdot;\theta)$ denotes the density function of the parametric copula family $C(\cdot;\theta)$, and $\hat{F}_j(x) = \frac{1}{n+1}\sum_{i=1}^n I(X_{i,j} \le x)$ for $j = 1, \cdots, d$. Alternatively, the pseudo maximum likelihood estimator can be defined as a root of the score equations

$$\sum_{i=1}^n \mathbf{l}(\hat{F}_1(X_{i,1}), \cdots, \hat{F}_d(X_{i,d}); \theta) = 0, \tag{3}$$

where $\mathbf{l}(x;\theta) = (l_1(x;\theta), \cdots, l_q(x;\theta))$ and $l_j(x;\theta) = \frac{\partial}{\partial\theta_j}\log c(x;\theta)$. The asymptotic distribution of the above pseudo maximum likelihood estimator and a consistent estimator for the asymptotic variance are given in Genest, Ghoudi and Rivest (1995). Since the asymptotic covariance of the pseudo maximum likelihood estimator is complicated and involves the contribution from both the copula and marginals, it is of importance to seek a more efficient way to construct confidence regions for the parameters θ without estimating the asymptotic covariance. In this paper, we investigate the possibility of employing empirical likelihood methods.

Since Owen (1988, 1990) introduced the empirical likelihood method for constructing a confidence interval/region for a mean, it has been extended and applied to many different settings and fields as a powerful interval estimation procedure; see Owen (2001) for more details. A key step in applying the empirical likelihood method is to formulate the nonparametric likelihood function. This is commonly done via estimating equations as proposed by Qin and Lawless (1994). Since the pseudo maximum likelihood estimator is a solution to the score equations (3), one may apply the method in Qin and Lawless (1994) to construct confidence regions for β by defining the empirical likelihood function as

$$\begin{aligned} L_1(\theta) \quad &= \sup\{\textstyle\prod_{i=1}^n (np_i) : p_1 \geq 0, \cdots, p_n \geq 0, \sum_{i=1}^n p_i = 1, \\ &\quad \textstyle\sum_{i=1}^n p_i \mathbf{l}(\hat{F}_1(X_{i,1}), \cdots, \hat{F}_d(X_{i,d}); \theta) = 0\}. \end{aligned}$$

Unfortunately, this likelihood function can not catch the variances of $\hat{F}_j's$ and thus Wilks theorem fails, i.e., $-2\log L_1(\theta)$ does not converge in distribution to a chi-squared limit.

In general, Wilks theorem does not hold when an empirical likelihood method is applied to nonlinear functionals. A common way to deal with nonlinear functionals is to linearize it before employing the empirical likelihood method; see Chen, Peng and Zhao (2009) and Molanes-Lopez, Van Keilegom and Veraverbeke (2009) for constructing confidence intervals for copula at a particular point. However, it remains unknown on how to linearize the score questions (3). Recently, Jing, Yuan and Zhou (2009) proposed a so-called jackknife empirical likelihood method to deal with nonlinear functionals. In this paper, we apply the jackknife empirical likelihood method to construct confidence intervals/regions for the parametric copulas. When the copula is estimated nonparametrically, Peng, Qi and Van Keilegom (2011) proposed a smoothed jackknife empirical likelihood method to construct confidence intervals for a copula at a fixed point.

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

2 Methodology and Main Results

In order to formulate an empirical likelihood function with $\hat{F}_j's$ taken into account, we follow the idea in Jing, Yuan and Zhou (2009) to construct a jackknife sample first and then apply the empirical likelihood method to the jackknife sample.

Define $\hat{F}_{j,-i}(x) = \frac{1}{n}\sum_{k=1,k\neq i}^n I(X_{k,j} \le x)$ for $j = 1, \cdots, d$ and $i = 1, \cdots, n$ and the jackknife sample $\{\mathbf{Z_i}(\theta) = (Z_{i,1}(\theta), \cdots, Z_{i,q}(\theta))^T\}_{i=1}^n$ as

$$Z_{i,j}(\theta) = \sum_{k=1}^n l_j(\hat{F}_1(X_{k,1}), \cdots, \hat{F}_d(X_{k,d}); \theta) - \sum_{k=1,k\neq i}^n l_j(\hat{F}_{1,-i}(X_{k,1}), \cdots, \hat{F}_{d,-i}(X_{k,d}); \theta)$$

for $i = 1, \cdots, n$ and $j = 1, \cdots, q$. Based on this jackknife sample, we define the jackknife empirical likelihood function as

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

By the Lagrange multiplier technique, we have

$$-2\log L(\theta) = 2\sum_{i=1}^n \log\{1 + \lambda^T \mathbf{Z}_i(\theta)\},$$

where $\lambda = (\lambda_1(\theta), \cdots, \lambda_q(\theta))^T$ satisfies

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

Before showing that the Wilks theorem holds for the above jackknife empirical likelihood method, we list some regularity conditions. Throughout we use θ_0 to denote the true value of θ and define $r(u) = u(1-u)$.

A1) There exist some constants $0 < \alpha_1 < 1/2$ and $M_1 > 0$ such that, uniformly for $0 < u_1, \cdots, u_d < 1$,

$$|l_j(u_1, \cdots, u_d; \theta_0)| \le M_1 \prod_{i=1}^{d} r(u_i)^{-\alpha_1},$$

$$|l_j^{(s)}(u_1, \cdots, u_d; \theta_0)| := |\frac{\partial}{\partial u_s} l_j(u_1, \cdots, u_d; \theta_0)| \le M_1 r(u_s)^{-1} \prod_{i=1}^{d} r(u_i)^{-\alpha_1},$$

$$|l_j^{(sm)}(u_1, \cdots, u_d; \theta_0)| := |\frac{\partial^2}{\partial u_s \partial u_m} l_j(u_1, \cdots, u_d; \theta_0)| \le M_1 r(u_s)^{-1} r(u_m)^{-1} \prod_{i=1}^{d} r(u_i)^{-\alpha_1},$$

and

$$\mathbb{E}[l_j^2(F_1(X_{1,1}), \cdots, F_d(X_{1,d}); \theta_0)] \le M_1$$

for $j = 1, \cdots, q$ and $s, m = 1, \cdots, d$.

A2) For a given $0 < \alpha_2 < 1/2$, there exist some constants $0 < \alpha_3 < 1/2$ and $M_2 > 0$ such that, uniformly for $0 < u_1, \cdots, u_d < 1$

$$\int \cdots \int_{[0,1]^{d-1}} \prod_{i=1}^{d} r(u_i)^{-\alpha_2} c(u_1, \cdots, u_d; \theta_0)\, du_1 \cdots du_{s-1} du_{s+1} \cdots du_d \le M_2 r(u_s)^{-\alpha_3}$$

for $s = 1, \cdots, d$, and

$$\int \cdots \int_{[0,1]^{d-2}} \prod_{i=1}^{d} r(u_i)^{-\alpha_2} c(u_1, \cdots, u_d; \theta_0)\, du_1 \cdots du_{s-1} du_{s+1} \cdots du_{m-1} du_{m+1} \cdots du_d$$

$$\le M_2 r(u_s)^{-\alpha_3} r(u_m)^{-\alpha_3}$$

for $1 \le s < m \le d$.

Remark 1. Commonly used copulas such as Clayton, Frank, Gumbel, Normal and t copulas satisfy A1) and A2).

Theorem 1. Under conditions A1) and A2), we have

$$-2 \log L(\theta_0) \xrightarrow{d} \chi^2(q) \quad \text{as} \quad n \to \infty.$$

Based on the above theorem, an empirical likelihood confidence interval/region for θ_0 with level ξ is $\{\theta : -2 \log L(\theta) > \chi^2_{q,\xi}\}$, where $\chi^2_{q,\xi}$ is the ξ-th quantile of a chi-squared distribution with q degrees of freedom.

3 Simulation and Data Analysis

3.1 Simulation study

In this subsection, we examine the finite behavior of the proposed jackknife empirical likelihood method and compare it with the normal approximation method.

We draw $10,000$ random samples with size $n = 300$ from the Clayton copula $C(u_1, \cdots, u_d; \theta) = (1 - d + u_1^{-\alpha} + \cdots + u_d^{-\alpha})^{-1/\alpha}$, bivariate normal copula $C(u_1, u_2; \theta) = \Phi_\theta(\Phi^{-1}(u_1), \Phi^{-1}(u_2))$, where Φ denotes the standard normal distribution and Φ_θ denotes the standard bivariate normal distribution with correlation θ, and bivariate t-copula with $\theta = (\rho, \nu)$, where $\rho \in (-1, 1)$ and $\nu > 0$.

We employ the 'copula' package in R to calculate the pseudo maximum likelihood estimator and its asymptotic variance so as to construct a confidence interval/region for θ, denoted by NAM. We also denote the proposed jackknife empirical likelihood method by JELM. For calculating the score equations of the bivariate t-copula, we use the formulas in Dakovic and Czada (2011) with some typos corrected. More specifically, i) the integrals in (7) and (8) have to be divided by 2; ii) x^2 in (8) is x_i^2; iii) the term $\frac{\nu+2}{2\nu}$ in the formula for $\frac{\partial l}{\partial \nu}(u_1, u_2)$ after (11) is $\frac{\nu-2}{2\nu}$. Note that equations (7), (8) and (11) mean those in Dakovic and Czada (2011).

In Tables 1-3 we report coverage probabilities for these two methods with levels 0.9 and 0.95. Note that for the t-copula, the 'copula' package in R does not provide asymptotic covariance. Hence we only report the coverage probabilities for the proposed jackknife empirical likelihood method in this case. From these tables, we observe that (i) the proposed jackknife empirical likelihood method works better than the normal approximation methods for large θ in the Clayton and normal copula (i.e., strong dependence); (ii) results for the cases of $d = 4, \theta = 10, 15$ in Table 1 indicate that the asymptotic variance for the Clayton copula given in the 'copula' package may be problematic when the dimension is large; (iii) the proposed jackknife empirical likelihood method performs well for t-copulas, where the asymptotic variance in the 'copula' package is not available.

3.2 Data analysis

We apply the proposed method to an insurance company data on losses and ALAEs. This particular data set has been analyzed by Frees and Valdez (1998), Klugman and Parsa (1999), Dupuis and Jones (2006), and Peng (2008). Like Klugman and Parsa (1999), we fit the Frank copula

$$C(u,v;\alpha) = -\frac{1}{\alpha}\log\{1 + \frac{(e^{-\alpha u}-1)(e^{-\alpha v}-1)}{e^{-\alpha}-1}\}.$$

Using the 'copula' package in R, we find the pseudo maximum likelihood estimator for α is 2.992 and the confidence intervals based on the normal approximation method are $(2.694, 3.290)$ and $(2.637, 3.348)$ for levels 90% and 95%, respectively. The proposed jackknife empirical likelihood intervals are calculated to be $(2.702, 3.292)$ and $(2.653, 3.352)$ for levels 90% and 95%, respectively, which are slightly skewed to the right than the normal approximation based intervals.

4 Proofs

Lemma 1. Under conditions of Theorem 1, we have

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathbf{Z}_i(\theta_0) \xrightarrow{d} N(0,\Sigma) \quad \text{as} \quad n\to\infty,$$

where $\Sigma = (\sigma_{ij})_{1\le i,j\le q}$,

$$\sigma_{ij} \;=\; \mathbb{E}\left[\left(l_i(\mathbf{T}_1;\theta_0)+\sum_{s=1}^{d}W(i,s)\right)\left(l_j(\mathbf{T}_1;\theta_0)+\sum_{s=1}^{d}W(j,s)\right)\right] < \infty,$$

$\mathbf{T}_1 = (F_1(X_{1,1}),\cdots,F_d(X_{1,d}))^T$ and

$$W(i,s) = \int_0^1\cdots\int_0^1 l_i^{(s)}(u_1,\cdots,u_d;\theta_0)(I(F_s(X_{1,s})\le u_s)-u_s)c(u_1,\cdots,u_d;\theta_0)\,du_1\cdots du_d.$$

Proof. We denote $\mathbf{T}_k = (F_1(X_{k,1}), \cdots, F_d(X_{k,d}))^T$, $\hat{\mathbf{T}}_k = (\hat{F}_1(X_{k,1}), \cdots, \hat{F}_d(X_{k,d}))^T$ and $\hat{\mathbf{T}}_{k,-i} = (\hat{F}_{1,-i}(X_{k,1}), \cdots, \hat{F}_{d,-i}(X_{k,d}))^T$ for $i, k = 1, \cdots, n$. Write

$$\begin{aligned}
& Z_{i,j}(\theta_0) \\
= \;& l_j(\hat{\mathbf{T}}_k;\theta_0) + \sum_{k=1,k\neq i}^{n} \{l_j(\hat{\mathbf{T}}_k;\theta_0) - l_j(\hat{\mathbf{T}}_{k,-i};\theta_0)\} \\
= \;& l_j(\hat{\mathbf{T}}_k;\theta_0) + \sum_{k=1,k\neq i}^{n} \sum_{s=1}^{d} l_j^{(s)}(\hat{\mathbf{T}}_k;\theta_0)\{\hat{F}_s(X_{k,s}) - \hat{F}_{s,-i}(X_{k,s})\} \\
& + \frac{1}{2} \sum_{k=1,k\neq i}^{n} \sum_{s=1}^{d} \sum_{t=1}^{d} l_j^{(st)}(\mathbf{Y}_{k,i};\theta_0)\{\hat{F}_s(X_{k,s}) - \hat{F}_{s,-i}(X_{k,s})\}\{\hat{F}_t(X_{k,t}) - \hat{F}_{t,-i}(X_{k,t})\} \\
= \;& l_j(\hat{\mathbf{T}}_k;\theta_0) + \frac{1}{n} \sum_{k=1,k\neq i}^{n} \sum_{s=1}^{d} l_j^{(s)}(\hat{\mathbf{T}}_k;\theta_0)\{I(X_{i,s} \leq X_{k,s}) - \hat{F}_s(X_{k,s})\} \\
& + \frac{1}{2n^2} \sum_{k=1,k\neq i}^{n} \sum_{s=1}^{d} \sum_{t=1}^{d} l_j^{(st)}(\mathbf{Y}_{k,i};\theta_0) \times \{I(X_{i,s} \leq X_{k,s}) - \hat{F}_s(X_{k,s})\} \\
& \quad \times \{I(X_{i,t} \leq X_{k,t}) - \hat{F}_t(X_{k,t})\} \\
=: \;& I_1(i,j) + I_2(i,j) + I_3(i,j), \qquad (5)
\end{aligned}$$

where

$$\mathbf{Y}_{k,i} = \beta_k \hat{\mathbf{T}}_k + (1-\beta_k)\hat{\mathbf{T}}_{k,-i}$$

and $\beta_k \in [0,1]$ depending on i and j. Since

$$\sup_{1\leq i\leq n} \frac{F_s(X_{i,s})}{\hat{F}_s(X_{i,s})} = O_p(1) \quad \text{and} \quad \sup_{1\leq i\leq n} \frac{1-F_s(X_{i,s})}{1-\hat{F}_s(X_{i,s})} = O_p(1) \qquad (6)$$

(see (4) in Page 415 of Shorack and Wellner (1986)), it follows from A1) that

$$\begin{aligned}
& \tfrac{1}{\sqrt{n}} \textstyle\sum_{i=1}^n I_2(i,j) \\
= \;& n^{-3/2} \textstyle\sum_{i=1}^n \sum_{k=1}^n \sum_{s=1}^d l_j^{(s)}(\hat{\mathbf{T}}_k;\theta_0)\{I(X_{i,s} \leq X_{k,s}) - \hat{F}_s(X_{k,s})\} \\
& -n^{-3/2} \textstyle\sum_{i=1}^n \sum_{s=1}^d l_j^{(s)}(\hat{\mathbf{T}}_i;\theta_0)\{1 - \hat{F}_s(X_{i,s})\} \\
= \;& -n^{-3/2} \textstyle\sum_{k=1}^n \sum_{s=1}^d l_j^{(s)}(\hat{\mathbf{T}}_k;\theta_0)\{1 - 2\hat{F}_s(X_{k,s})\} \\
= \;& O_p(n^{-3/2} \textstyle\sum_{i=1}^n \sum_{s=1}^d r(\hat{F}_s(X_{i,s}))^{-1} \prod_{t=1}^d r(\hat{F}_t(X_{i,t}))^{-\alpha_1}) \\
= \;& O_p(n^{-3/2} \textstyle\sum_{i=1}^n \sum_{s=1}^d r(F_s(X_{i,s}))^{-1} \prod_{t=1}^d r(F_t(X_{i,t}))^{-\alpha_1}).
\end{aligned} \qquad (7)$$

By A2) and choosing $\delta > 1$ and $\delta\alpha_3 < 1/2$, where α_3 is given in A2), we have for any $\epsilon > 0$

$$
\begin{aligned}
& \mathbb{P}(n^{-3/2}\textstyle\sum_{i=1}^n r(F_s(X_{i,s}))^{-1}\prod_{t=1}^d r(F_t(X_{i,t}))^{-\alpha_1} > \epsilon) \\
\leq\ & \mathbb{P}(n^{-3/2}\textstyle\sum_{i=1}^n I(n^{-\delta} \leq F_s(X_{i,s}) \leq 1-n^{-\delta}) r(F_s(X_{i,s}))^{-1}\prod_{t=1}^d r(F_t(X_{i,t}))^{-\alpha_1} > \epsilon) \\
& +\mathbb{P}(\min_{1\leq i\leq n} F_s(X_{i,s}) < n^{-\delta}) + \mathbb{P}(\max_{1\leq i\leq n} F_s(X_{i,s}) > 1-n^{-\delta}) \\
\leq\ & (n^{3/2}\epsilon)^{-1}\textstyle\sum_{i=1}^n \mathbb{E}[I(n^{-\delta} \leq F_s(X_{i,s}) \leq 1-n^{-\delta}) r(F_s(X_{i,s}))^{-1}\prod_{t=1}^d r(F_t(X_{i,t}))^{-\alpha_1}] \\
& +o(1) \\
\leq\ & M_2 n^{-1/2}\epsilon^{-1}\mathbb{E}[I(n^{-\delta} \leq F_s(X_{1,s}) \leq 1-n^{-\delta}) r(F_s(X_{1,s}))^{-1-\alpha_3}] + o(1) \\
\leq\ & M_2 n^{-1/2+\delta\alpha_3}\epsilon^{-1} + o(1) \\
=\ & o(1).
\end{aligned}
\tag{8}
$$

Therefore, it follows from (7) and (8) that

$$
\frac{1}{\sqrt{n}}\sum_{i=1}^n I_2(i,j) = o_p(1) \quad \text{for} \quad j = 1, \cdots, q. \tag{9}
$$

By A1), (6) and noting that

$$
\begin{aligned}
& \textstyle\sum_{i=1}^n (I(X_{i,s} \leq X_{k,s}) - \hat{F}_s(X_{k,s}))^2 \\
=\ & (n+1)\hat{F}_s(X_{k,s})(1-\hat{F}_s(X_{k,s})) - \hat{F}_s^2(X_{k,s}) \\
\leq\ & (n+1) r(\hat{F}_s(X_{k,s})),
\end{aligned}
$$

we have

$$
\begin{aligned}
& |n^{-5/2}\textstyle\sum_{i=1}^n\sum_{k=1,k\neq i}^n l_j^{(st)}(\mathbf{Y}_{k,i};\theta_0)(I(X_{i,s} \leq X_{k,s}) - \hat{F}_s(X_{k,s})) \\
& \quad \times(I(X_{i,t} \leq X_{k,t}) - \hat{F}_t(X_{k,t}))| \\
=\ & O_p(n^{-5/2}\textstyle\sum_{i=1}^n\sum_{k=1}^n r(F_s(X_{k,s}))^{-1} r(F_t(X_{k,t}))^{-1}\prod_{m=1}^d r(F_m(X_{k,m}))^{-\alpha_1} \\
& \times\{(I(X_{i,s} \leq X_{k,s}) - \hat{F}_s(X_{k,s}))^2 + (I(X_{i,t} \leq X_{k,t}) - \hat{F}_t(X_{k,t}))^2\}) \\
=\ & O_p(n^{-5/2}\textstyle\sum_{k=1}^n r(F_s(X_{k,s}))^{-1} r(F_t(X_{k,t}))^{-1}\prod_{m=1}^d r(F_m(X_{k,m}))^{-\alpha_1} \\
& \times(n+1)\{r(\hat{F}_s(X_{k,s})) + r(-\hat{F}_t(X_{k,t}))\}) \\
=\ & O_p(n^{-3/2}\textstyle\sum_{k=1}^n r(F_t(X_{k,t}))^{-1}\prod_{m=1}^d r(F_m(X_{k,m}))^{-\alpha_1}) \\
& +O_p(n^{-3/2}\textstyle\sum_{k=1}^n r(F_s(X_{k,s}))^{-1}\prod_{m=1}^d r(F_m(X_{k,m}))^{-\alpha_1})
\end{aligned}
\tag{10}
$$

for $s, t = 1, \cdots, q$. Like the proof of (8), we have

$$n^{-3/2} \sum_{k=1}^{n} r(F_t(X_{k,t}))^{-1} \prod_{m=1}^{d} r(F_m(X_{k,m}))^{-\alpha_1} = o_p(1)$$

for $t = 1, \cdots, d$, i.e.,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} I_3(i,j) = o_p(1) \quad \text{for} \quad j = 1, \cdots, q. \tag{11}$$

Write

$$\begin{aligned} I_1(i,j) &= l_j(\mathbf{T}_i;\theta_0) + \textstyle\sum_{s=1}^{d} l_j^{(s)}(\mathbf{T}_i;\theta_0)\{\hat{F}_s(X_{i,s}) - F_s(X_{i,s})\} \\ &\quad + \tfrac{1}{2} \textstyle\sum_{s=1}^{d} \sum_{t=1}^{d} l_j^{(st)}(\mathbf{Y}_i^*;\theta_0)\{\hat{F}_s(X_{i,s}) - F_s(X_{i,s})\}\{\hat{F}_t(X_{i,t}) - F_t(X_{i,t})\} \\ &=: II_1(i,j) + II_2(i,j) + II_3(i,j), \end{aligned}$$

where

$$\mathbf{Y}_i^* = \beta_i^* \hat{\mathbf{T}}_i + (1 - \beta_i^*)\mathbf{T}_i$$

and $\beta_i^* \in [0,1]$.

Since

$$\max_{1 \le i \le n} \left| \frac{\sqrt{n}\{\hat{F}_s(X_{i,s}) - F_s(X_{i,s})\}}{F_s^{1/2}(X_{i,s})(1 - F_s(X_{i,s}))^{1/2}} \right| = O_p(\log n) \tag{12}$$

for $s = 1, \cdots, d$ (see Mason (1981)), using the same arguments in proving (8), we can show that

$$\frac{1}{n} \sum_{i=1}^{n} II_3(i,j) = o_p(1) \quad \text{for} \quad j = 1, \cdots, q. \tag{13}$$

It is easy to check that

$$\begin{aligned} &\mathbb{E}(\{\hat{F}_{s,-i}(X_{i,s}) - \tfrac{n-1}{n} F_s(X_{i,s})\}\{\hat{F}_{s,-k}(X_{k,s}) - \tfrac{n-1}{n} F_s(X_{k,s})\} | X_i, X_k) \\ = &\ \tfrac{n-2}{n^2}\{F_s(X_{i,s} \wedge X_{k,s}) - F_s(X_{i,s})F_s(X_{k,s})\} \end{aligned} \tag{14}$$

for $i \neq k$. Put

$$W_1(i,j,s) = l_j^{(s)}(\mathbf{T}_i;\theta_0)\{\hat{F}_s(X_{i,s}) - F_s(X_{i,s})\},$$

$$W_2(i,j,s) = l_j^{(s)}(\mathbf{T}_i;\theta_0)\{\hat{F}_{s,-i}(X_{i,s}) - \frac{n-1}{n} F_s(X_{i,s})\}$$

and

$$W_3(i,j,s) = \int_0^1 \cdots \int_0^1 l_j^{(s)}(u_1,\cdots,u_d;\theta_0)\{I(F_s(X_{i,s}) \le u_s) - u_s\} \times c(u_1,\cdots,u_d;\theta_0)\,du_1\cdots du_d.$$

Since

$$W_1(i,j,s) = \frac{n}{n+1}W_2(i,j,s) + l_j^{(s)}(\mathbf{T}_i;\theta_0)\{\frac{1}{n+1} - \frac{2}{n+1}F_s(X_{i,s})\},$$

it follows from the same arguments in proving (8) that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n W_1(i,j,s) = \frac{1}{\sqrt{n}}\sum_{i=1}^n W_2(i,j,s) + o_p(1) \tag{15}$$

for $j = 1,\cdots,q$ and $s = 1,\cdots,d$. By (12), we have

$$\begin{aligned} & \max_{1\le i\le n}|\tfrac{\sqrt{n}\{\hat{F}_{s,-i}(X_{i,s}) - \frac{n-1}{n}F_s(X_{i,s})\}}{F_s^{1/2}(X_{i,s})(1-F_s(X_{i,s}))^{1/2}}| \\ \le\ & \max_{1\le i\le n}\tfrac{n+1}{n}|\tfrac{\sqrt{n}\{\hat{F}_s(X_{i,s}) - F_s(X_{i,s})\}}{F_s^{1/2}(X_{i,s})(1-F_s(X_{i,s}))^{1/2}}| \\ & + \max_{1\le i\le n}\{\sqrt{n}F_s^{1/2}(X_{i,s})(1-F_s(X_{i,s}))^{1/2}\}^{-1} \\ =\ & O_p(\log n) \end{aligned} \tag{16}$$

for $s = 1,\cdots,d$. Using (16) and the same arguments in proving (8), we have

$$\frac{1}{n}\sum_{i=1}^n W_2^2(i,j,s) = o_p(1) \quad \text{for} \quad j = 1,\cdots,q, s = 1,\cdots,d. \tag{17}$$

By (14) and (17), we have

$$\begin{aligned} & \mathbb{E}\{\tfrac{1}{n}\textstyle\sum_{i,k=1,i\ne k}^n W_2(i,j,s)W_2(k,j,s)\} \\ =\ & \mathbb{E}(\mathbb{E}\{\tfrac{1}{n}\textstyle\sum_{i,k=1,i\ne k}^n W_2(i,j,s)W_2(k,j,s)|X_i,X_k\}) \\ =\ & \mathbb{E}\{\tfrac{n-2}{n^3}\textstyle\sum_{i,k=1,i\ne k}^n l_j^{(s)}(\mathbf{T}_i;\theta_0)l_j^{(s)}(\mathbf{T}_k;\theta_0)(F_s(X_{i,s})\wedge F_s(X_{k,s}) - F_s(X_{i,s})F_s(X_{k,s}))\} \\ =\ & \textstyle\int_0^1\cdots\int_0^1 l_j^{(s)}(u_1,\cdots,u_d;\theta_0)l_j^{(s)}(v_1,\cdots,v_d;\theta_0)(u_s\wedge v_s - u_sv_s)\times \\ & c(u_1,\cdots,u_d;\theta_0)c(v_1,\cdots,v_d;\theta)\,du_1\cdots du_d dv_1\cdots dv_d + o(1), \end{aligned} \tag{18}$$

$$\begin{aligned}
&\mathbb{E}\{\tfrac{1}{n}\textstyle\sum_{i=1}^n\sum_{k=1}^n W_2(i,j,s)W_3(k,j,s)\}\\
=\;& \mathbb{E}(\mathbb{E}\{\tfrac{1}{n}\textstyle\sum_{i=1}^n\sum_{k=1}^n W_2(i,j,s)W_3(k,j,s)|X_i,X_k\})\\
=\;& \mathbb{E}(\mathbb{E}\{\tfrac{1}{n}\textstyle\sum_{i,k=1,i\neq k}^n W_2(i,j,s)W_3(k,j,s)|X_i,X_k\})\\
=\;& \mathbb{E}\{\tfrac{1}{n^2}\textstyle\sum_{i,k=1,i\neq k}^n l_j^{(s)}(\mathbf{T}_i;\theta_0)(I(X_{k,s}\le X_{i,s})-F_s(X_{i,s}))W_3(k,j,s)\}\\
=\;& \textstyle\int_0^1\cdots\int_0^1 l_j^{(s)}(u_1,\cdots,u_d;\theta_0)l_j^{(s)}(v_1,\cdots,v_d;\theta_0)(u_s\wedge v_s-u_sv_s)\times\\
& c(u_1,\cdots,u_d;\theta_0)c(v_1,\cdots,v_d;\theta_0)\,du_1\cdots du_d dv_1\cdots dv_d+o(1)
\end{aligned}\tag{19}$$

and

$$\begin{aligned}
&\mathbb{E}\{\tfrac{1}{n}\textstyle\sum_{i=1}^n\sum_{k=1}^n W_3(i,j,s)W_3(k,j,s)\}\\
=\;& \mathbb{E}\{\tfrac{1}{n}\textstyle\sum_{i=1}^n W_3^2(i,j,s)\}\\
=\;& \textstyle\int_0^1\cdots\int_0^1 l_j^{(s)}(u_1,\cdots,u_d;\theta_0)l_j^{(s)}(v_1,\cdots,v_d;\theta_0)(u_s\wedge v_s-u_sv_s)\times\\
& c(u_1,\cdots,u_d;\theta_0)c(v_1,\cdots,v_d;\theta_0)\,du_1\cdots du_d dv_1\cdots dv_d
\end{aligned}\tag{20}$$

for $j=1,\cdots,q$ and $s=1,\cdots,d$. Hence, it follows from (17)–(20) that for any $\epsilon>0$

$$\begin{aligned}
&\mathbb{P}(|\tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^n(W_2(i,j,s)-W_3(i,j,s))|>\epsilon)\\
=\;& \mathbb{P}(\tfrac{1}{n}\textstyle\sum_{i=1}^n W_2^2(i,j,s)+\tfrac{1}{n}\sum_{i,k=1,i\neq k}^n W_2(i,j,s)W_2(k,j,s)\\
& -\tfrac{2}{n}\textstyle\sum_{i=1}^n\sum_{k=1}^n W_2(i,j,s)W_3(k,j,s)+\tfrac{1}{n}\sum_{i=1}^n\sum_{k=1}^n W_3(i,j,s)W_3(k,j,s)>\epsilon^2)\\
\le\;& \mathbb{P}(\tfrac{1}{n}\textstyle\sum_{i=1}^n W_2^2(i,j,s)>\epsilon^2/2)\\
& +\tfrac{2}{\epsilon^2}\mathbb{E}\{\tfrac{1}{n}\textstyle\sum_{i,k=1,i\neq k}^n W_2(i,j,s)W_2(k,j,s)-\tfrac{2}{n}\sum_{i=1}^n\sum_{k=1}^n W_2(i,j,s)W_3(k,j,s)\\
& +\tfrac{1}{n}\textstyle\sum_{i=1}^n\sum_{k=1}^n W_3(i,j,s)W_3(k,j,s)\}\\
=\;& o(1).
\end{aligned}\tag{21}$$

By (13), (15) and (21), we have

$$\begin{aligned}
&\tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^n I_1(i,j)\\
=\;& \tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^n l_j(\mathbf{T}_i;\theta_0)+\tfrac{1}{\sqrt{n}}\sum_{i=1}^n\sum_{s=1}^d\int_0^1\cdots\int_0^1 l_j^{(s)}(u_1,\cdots,u_d;\theta_0)\\
& \times(I(F_s(X_{i,s})\le u_s)-u_s)c(u_1,\cdots,u_d;\theta_0)\,du_1\cdots du_d+o_p(1)
\end{aligned}\tag{22}$$

for $j = 1, \cdots, q$. Note that by A1) and A2),

$$\begin{aligned}
& \mathbb{E}\left[\int_0^1 \cdots \int_0^1 l_j^{(s)}(u_1, \cdots, u_d; \theta_0)(I(F_s(X_{i,s}) \leq u_s) - u_s)c(u_1, \cdots, u_d; \theta_0)\, du_1 \cdots du_d\right]^2 \\
= & \int_0^1 \cdots \int_0^1 l_j^{(s)}(u_1, \cdots, u_d; \theta_0)c(u_1, \cdots, u_d; \theta_0)l_j^{(s)}(v_1, \cdots, v_d; \theta_0)c(v_1, \cdots, v_d; \theta_0) \\
& \times(\min\{u_s, v_s\} - u_s v_s)\, du_1 \cdots du_d dv_1 \cdots dv_d \\
< & \ \infty.
\end{aligned}$$

Hence, the lemma follows from (9), (11), (22) and the central limit theorem.

Lemma 2. Under conditions of Theorem 1, we have

$$\frac{1}{n}\sum_{i=1}^n \mathbf{Z}_i(\theta_0)\mathbf{Z}_i^T(\theta_0) \xrightarrow{p} \Sigma \quad \text{as} \quad n \to \infty,$$

where Σ is defined in Lemma 1.

Proof. Using the same notation in the proof of Lemma 1, we can show that for fixed $j, m = 1, \cdots, q$,

$$\frac{1}{n}\sum_{i=1}^n I_1(i,j)I_1(i,m) = \mathbb{E}[l_j(\mathbf{T}_1; \theta_0)l_m(\mathbf{T}_1; \theta_0)] + o_p(1),$$

$$\frac{1}{n}\sum_{i=1}^n I_3(i,j)\{I_1(i,m) + I_2(i,m)\} = o_p(1), \quad \frac{1}{n}\sum_{i=1}^n I_3(i,j)I_3(i,m) = o_p(1),$$

$$\frac{1}{n}\sum_{i=1}^n I_1(i,j)I_2(i,m) = \mathbb{E}\left[l_j(\mathbf{T}_1; \theta_0)\sum_{s=1}^d W(m,s)\right] + o_p(1),$$

and

$$\frac{1}{n}\sum_{i=1}^n I_2(i,j)I_2(i,m) = \mathbb{E}\left[\sum_{s=1}^d\sum_{t=1}^d W(j,s)W(m,t)\right] + o_p(1),$$

which implies that

$$\frac{1}{n}\sum_{i=1}^n Z_{i,j}(\theta_0)Z_{i,m}(\theta_0) \xrightarrow{p} \sigma_{jm} \quad \text{for} \quad j, m = 1, \cdots, q,$$

i.e., the lemma holds.

Lemma 3. Under conditions of Theorem 1, we have for $j = 1, \cdots, q$,

$$\max_{1\leq i\leq n} |Z_{i,j}(\theta_0)| = o_p(n^{1/2}).$$

Proof. We shall use the same notation in the proof of Lemma 1. For any $M > 0$, we have

$$\begin{aligned} & \mathbb{P}\left(\max_{1\le i\le n} |I_2(i,j)| \ge n^{1/2}M\right) \\ \le \ & \mathbb{P}\left(\max_{1\le i\le n} \frac{1}{n}\sum_{k=1,k\ne i}^{n}\sum_{s=1}^{d} |l_j^{(s)}(\hat{\mathbf{T}}_k;\theta_0)| \ge n^{1/2}M\right) \\ \le \ & \mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n}\sum_{s=1}^{d} |l_j^{(s)}(\hat{\mathbf{T}}_k;\theta_0)| \ge n^{1/2}M\right). \end{aligned}$$

Hence by the same arguments in (7) and (8) we have

$$n^{-3/2}\sum_{k=1}^{n}\sum_{s=1}^{d} |l_j^{(s)}(\hat{\mathbf{T}}_k;\theta_0)| = o_p(1),$$

i.e., $\mathbb{P}\left(\max_{1\le i\le n} |I_2(i,j)| \ge n^{1/2}M\right) = o(1)$, which implies that

$$\max_{1\le i\le n} |I_2(i,j)| = o_p(n^{1/2}). \tag{23}$$

Note that in (10) and (11), we actually showed

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} |I_3(i,j)| = o_p(1),$$

which implies

$$\max_{1\le i\le n} |I_3(i,j)| = o_p(n^{1/2}). \tag{24}$$

Similarly, we have

$$\max_{1\le i\le n} |II_2(i,j)| = o_p(n^{1/2}) \quad \text{and} \quad \max_{1\le i\le n} |II_3(i,j)| = o_p(n^{1/2}). \tag{25}$$

Since $\mathbb{E}[l_j^2(\mathbf{T}_1;\theta_0)] < \infty$, we have $n\mathbb{P}(l_j^2(\mathbf{T}_1;\theta_0) \ge n) = o(1)$, i.e.,

$$\max_{1\le i\le n} |II_1(i,j)| = o_p(n^{1/2}). \tag{26}$$

Hence the lemma follows from (23) to (26).

Proof of Theorem 1. It follows from Lemmas 1-3 and the standard arguments in the empirical likelihood method for a mean vector (see Owen (1990)).

Acknowledgment. Peng's research was supported by NSF Grant DMS-1005336. Yang's research was partly supported by the National Basic Research Program (973 Program) of China (2007CB814905) and the National Natural Science Foundation of China (Grants No. 10871008)

References

[1] CHEN, J., PENG, L. AND ZHAO, Y. (2009). Empirical likelihood based confidence intervals for copulas. *J. Multi. Anal. 100, 137 – 151.*

[2] DAKOVIC, R. AND CZADO, C. (2011). Comparing point and interval estimates in the bivariate t-copula model with applocation to financial data. *Stat. Papers, to appear.*

[3] DUPUIS, D. AND JONES, B.L. (2006). Multivariate extreme value theory and its usefulness in understanding risk. *North American Actuarial Journal 10(4), 1 – 27.*

[4] FERMANIAN, J.D., RADULOVIC, D. AND WEGKAMP, M.(2004). Weak convergence of empirical copula processes. *Bernoulli 10, 847–860.*

[5] FREES, E.W. AND VALDEZ, E.A. (1998). Understanding relationships using copulas. *North American Actuarial Journal 2, 1 - 25.*

[6] GENEST, C., GHOUDI, K. AND RIVEST, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika 82, 543–552.*

[7] HAUG, S. KLÜPPELBERG, C. AND PENG, L. (2011). Statistical models and methods for dependence in insurance data. *Journal of Korean Statistical Society 40, 125–139.*

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

[9] Klugman, S.A. and Parsa, R. (1999). Fitting bivariate loss distributions with copulas. *Insurance: Mathematics and Economics 24, 139 - 148.*

[10] Mason, D.M. (1981). Bounds for weighted empirical distribution functions, *Ann. Probab. 9, 881–884.*

[11] Molanes-Lopez, E., Van Keilegom, I. and Veraverbeke, N. (2009). Empirical likelihood for non-smooth criterion functions. *Scand. J. Statist., 36, 413–432.*

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

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

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

[15] Peng, L. (2008). Estimating the probability of a rare event via elliptical copulas. *North American Actuarial Journal 12(2), 116–128.*

[16] Peng, L., Qi, Y. and Van Keilegom, I. (2011). Jackknife empirical likelihood methods for copulas. *Test, to appear.*

[17] Qin,J. and Lawless, J.F. (1994). Empirical likelihood and general estimating equations. *Ann. Statist. 22, 300–325.*

[18] Shorack, G.R. and Wellner, J.A. (1986). Empirical Processes With Applications to Statistics. *John Wiley & Sons, New York.*

Table 1: Empirical coverage probabilities are reported for Clayton copulas with dimension $d = 2, 4$.

(d, θ)	JELM Level 0.9	NAM Level 0.9	JELM Levle 0.95	NAM Level 0.95
(2,0.2)	0.8846	0.8875	0.9363	0.9417
(2,1)	0.8902	0.8950	0.9430	0.9448
(2,10)	0.9114	0.9162	0.9563	0.9566
(2,15)	0.9184	0.9160	0.9628	0.9582
(4,0.2)	0.8750	0.8734	0.9336	0.9331
(4,1)	0.8767	0.8791	0.9295	0.9294
(4,10)	0.9167	0.9418	0.9573	0.9703
(4,15)	0.9211	0.9519	0.9604	0.9781

Table 2: Empirical coverage probabilities are reported for the bivariate normal copula.

θ	JELM Level 0.9	NAM Level 0.9	JELM Levle 0.95	NAM Level 0.95
0.2	0.8847	0.8851	0.9438	0.9434
0.5	0.8864	0.8750	0.9411	0.9314
0.8	0.8880	0.8818	0.9393	0.9331

Table 3: Empirical coverage probabilities are reported for the bivariate t copula.

$\theta = (\rho, \nu)$	JELM Level 0.9	JELM Level 0.95
(0.2, 3)	0.8853	0.9404
(0.5,3)	0.8874	0.9385
(0.8,3)	0.8945	0.9476
(0.2,8)	0.8808	0.9352
(0.5,8)	0.8861	0.9412
(0.8,8)	0.8878	0.9415