Estimating bivariate t-copulas via Kendall's tau

Liang Peng

Department of Risk Management and Insurance
Robinson College of Business
Georgia State University

Notes

This talk is based on the following paper:

- Liang Peng and Ruodu Wang (2014). Estimating bivariate t-copula via Kendall's tau. Variance. To appear.

Notes

Contents

Notes

Introduction

Copula: For a random vector (X, Y) with continuous marginal distributions F_1 and F_2, its copula is defined as

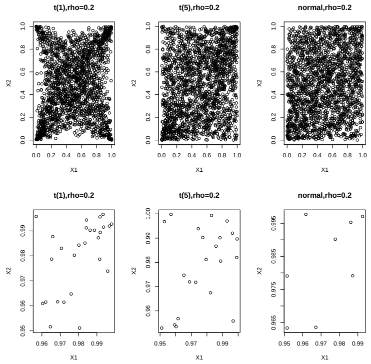
$$C(x,y) = P(F_1(X) \le x, F_2(Y) \le y) \quad \text{for} \quad 0 \le x, y \le 1.$$

t-copula: The t-copula is an elliptical copula defined as

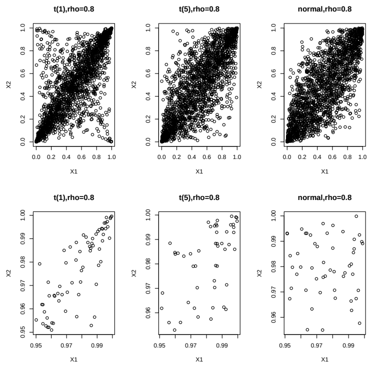
$$C(u,v;\rho,\nu) = \int_{-\infty}^{t_\nu^-(u)} \int_{-\infty}^{t_\nu^-(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \{1 + \frac{x^2-2\rho xy+y^2}{\nu(1-\rho^2)}\}^{-(\nu+2)/2}\, dydx, \quad (1)$$

where $\nu > 0$ is the number of degrees of freedom, $\rho \in [-1,1]$ is the linear correlation coefficient, t_ν is the distribution function of a t-distribution with ν degrees of freedom and t_ν^- denotes the generalized inverse function of t_ν. When $\nu = 1$, the t-copula is also called a Cauchy copula.

Notes



Notes



Notes

Known results: Breymann, Dias and Embrechts (2003) and Mashal, Naldi and Zeevi (2003) showed that empirical fit of the t-copula is better than the Gaussian copula. Some recent applications and generalization of t-copula include: Schloegl and O'Kane (2005) provided formulas for the portfolio loss distribution when t-copula is employed; de Melo and Mendes (2009) priced the options related with retirement funds by using the Gaussian and t copulas;

Notes

Chan and Kroese (2010) used t-copula to model and estimate the probability of a large portfolio loss; Manner and Segers (2011) studied the tails of correlation mixtures of the Gaussian and t copulas; grouped t-copula were given in Chapter 5 of McNeil, Frey and Embrechts (2005); Luo and Shevchenko (2010) and Venter et al. (2007) extended the grouped t-copula; tail dependence for multivariate t-copula and its monotonicity were studied by Chan and Li (2008).

Notes

Estimation: In order to fit the t-copula to a random sample $(X_1, Y_1), \cdots, (X_n, Y_n)$, one has to estimate the unknown parameters ρ and ν first.
Pseudo MLE: Since the distribution of $(F_1(X_i), F_2(Y_i))'s$ is the t-copula, we can use maximum likelihood estimation. However, F_1 and F_2 are unknown. Therefore we estimate them by $F_{n1}(x) = \frac{1}{n+1}\sum_{i=1}^n X_i$ and $F_{n2}(y) = \frac{1}{n+1}\sum_{i=1}^n Y_i$, respectively. Hence, we can apply the MLE to the pseudo data $(F_{n1}(X_i), F_{n2}(Y_i))'s$, which is called pseudo maximum likelihood estimate by Genest, Ghoudi and Rivest (1995).
Although, generally speaking, the pseudo MLE is efficient, its computation becomes a serious issue when applying to t-copulas especially with a large dimension.

Notes

Two-step estimation procedure: A more practical method to estimate ρ is through the Kendall's tau, defined as

$$\tau = \mathbb{E}(\text{sign}((X_1 - X_2)(Y_1 - Y_2))) = 4\int_0^1\int_0^1 C(u_1, u_2)\, dC(u_1, u_2) - 1.$$

It is known that τ and ρ have a simple relationship

$$\rho = \sin(\pi\tau/2).$$

By noting this relationship, Lindskog, McNeil and Schmock (2003) proposed to first estimate ρ by

$$\hat{\rho} = \sin(\pi\hat{\tau}/2), \quad \text{where} \quad \hat{\tau} = \frac{2}{n(n-1)} \sum_{1\le i<j\le n} \text{sign}((X_i - X_j)(Y_i - Y_j)),$$

Notes

and then to estimate ν by maximizing the pseudo likelihood function

$$\prod_{i=1}^n c(F_{n1}(X_i), F_{n2}(Y_i); \hat{\rho}, \nu),$$

where $c(u, v; \rho, \nu) = \frac{\partial^2}{\partial u \partial v} C(u, v; \rho, \nu)$ is the density of the t-copula defined in (1). In other words, the estimator $\hat{\nu}$ is defined as a solution to the score equation

$$\sum_{i=1}^n l(\hat{\rho}, \nu; F_{n1}(X_i), F_{n2}(Y_i)) = 0, \tag{2}$$

where $l(\rho, \nu; u, v) = \frac{\partial}{\partial \nu} \log c(u, c; \rho, \nu)$. $\hat{\tau}$ is called the Kendall's tau estimator.

Notes

Asymptotic limit: A recent attempt to derive the asymptotic distribution for the two-step estimator $(\hat{\rho}, \hat{\nu})$ is given by Fantazzini (2010), who employed the techniques for estimating equations. Unfortunately the derived asymptotic distribution in Fantazzini (2010) is not correct since the Kendall's tau estimator is a U-statistic rather than an average of independent observations. Numeric comparisons for the two estimation procedures are given in Dakovic and Czado (2011).

Notes

Methodology

Here we first derive the joint asymptotic limit of the two-step estimator $(\hat{\rho}, \hat{\nu})$ as follows.

Theorem 1. As $n \to \infty$, we have

$$\begin{aligned}&\sqrt{n}\{\hat{\rho}-\rho\}\\&=\cos(\tfrac{\pi\tau}{2})\tfrac{\pi}{\sqrt{n}}\textstyle\sum_{i=1}^{n}4\{C(F_1(X_i),F_2(Y_i))-\mathbb{E}C(F_1(X_1),F_2(Y_1))\}\\&-\cos(\tfrac{\pi\tau}{2})\tfrac{\pi}{\sqrt{n}}\textstyle\sum_{i=1}^{n}2\{F_1(X_i)+F_2(Y_i)-1\}+o_p(1)\end{aligned} \tag{3}$$

and

Notes

$$\begin{aligned}&\sqrt{n}\{\hat{\nu}-\nu\}\\&=-K_\nu^{-1}\{\tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^{n}l(\rho,\nu;F_1(X_i),F_2(Y_i))+K_\rho\sqrt{n}(\hat{\rho}-\rho)\\&+\tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^{n}\int_0^1\int_0^1 l_u(\rho,\nu;u,v)\{I(F_1(X_i)\le u)-u\}c(u,v)\,dudv\\&+\tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^{n}\int_0^1\int_0^1 l_v(\rho,\nu;u,v)\{I(F_2(Y_i)\le v)-v\}c(u,v)\,dudv\}\\&+o_p(1),\end{aligned} \tag{4}$$

where $l_u(\rho,\nu;u,v)=\frac{\partial}{\partial u}l(\rho,\nu;u,v)$, $l_v(\rho,\nu;u,v)=\frac{\partial}{\partial v}l(\rho,\nu;u,v)$, and for $a=\nu,\rho$,

$$K_a=\mathbb{E}\left(\frac{\partial}{\partial a}l(\rho,\nu;F_1(X_1),F_2(Y_1))\right)=\int_0^1\int_0^1\frac{\partial}{\partial a}l(\rho,\nu;u,v)\,dC(u,v).$$

Notes

Using the above theorem, we can easily obtain that

$$\sqrt{n}(\hat{\rho}-\rho,\hat{\nu}-\nu)^T\xrightarrow{d}N\left((0,0)^T,\begin{pmatrix}\sigma_1^2 & \sigma_{12}\\ \sigma_{12} & \sigma_2^2\end{pmatrix}\right), \tag{5}$$

where σ_1^2, σ_{12} and σ_2^2 are constants whose values are given in the proof of Theorem 1.

Question: How to construct confidence intervals/regions effectively?

Normal Approximation Method: We seek an alternative way, Empirical Likelihood Method, since the above asymptotic covariance matrix is too complicated.

Notes

Parametric likelihood ratio test

Observations: $X_1, \cdots, X_n$ iid with pdf $f(x; g(\mu))$, where g is a known function, but $\mu = E(X_1)$ is unknown.

Question: test $H_0 : \mu = \mu_0$ against $H_a : \mu \neq \mu_0$

PLRT: Let $\hat{\mu}$ denote the maximum likelihood estimate for μ. Then the likelihood ratio is defined as

$$\lambda = \Pi_{i=1}^n f(X_i; g(\mu_0)) / \Pi_{i=1}^n f(X_i; g(\hat{\mu})).$$

The likelihood ratio test is based on the following
Wilks Theorem. Under H_0, $-2 \log \lambda \xrightarrow{d} \chi^2(1)$ as $n \to \infty$.

Notes

Empirical likelihood method

When we do not fit a class of parametric family to X_i, but still test $H_0 : \mu = \mu_0$ vs $H_a : \mu \neq \mu_0$, a similar approach to the parametric likelihood ratio test was introduced by *Owen* (1988, 1990), which is a nonparametric likelihood ratio test and called empirical likelihood method.

Notes

Define the empirical likelihood ratio function for μ as

$$R(\mu) = \sup\{\prod_{i=1}^n (np_i) \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i X_i = \mu\}.$$

By Lagrange multiplier technique, we have
$p_i = n^{-1}\{1 + \lambda^T(X_i - \mu)\}^{-1}$ and

$$-2 \log R(\mu) = 2 \sum_{i=1}^n \log\{1 + \lambda^T(X_i - \mu)\},$$

where $\lambda = \lambda(\mu)$ satisfies

$$n^{-1} \sum_{i=1}^n \frac{X_i - \mu}{1 + \lambda^T(X_i - \mu)} = 0.$$

Notes

Wilks Theorem: Under H_0,

$$W(\mu_0) := -2\log R(\mu_0) \xrightarrow{d} \chi^2(d) \quad \text{as} \quad n \to \infty,$$

where $\mu \in R^d$.
Confidence interval/region: The above theorem can be employed to construct a confidence interval or region for μ as

$$I_\alpha = \{\mu : W(\mu) \le \chi^2_{d,\alpha}\}.$$

Advantages: i) No need to estimate any additional quantities such as asymptotic variance; ii) the shape of confidence interval/region is determined by the sample automatically; iii) Bartlett correctable

Notes

Estimating equations

A popular way to formulate the empirical likelihood function is via estimating equations.
Observations: $X_1, \cdots, X_n$ iid with common distribution function F and there is a q-dimensional parameter θ associated with F.
Conditions: Let y^T denote the transpose of the vector y and

$$G(x;\theta) = (g_1(x;\theta), \cdots, g_s(x;\theta))^T$$

denote $s(\ge q)$ functionally independent functions, which connect F and θ through the equations $EG(X_1;\theta) = 0$.

Notes

Empirical likelihood function:

$$R(\theta) = \sup\{\prod_{i=1}^n (np_i) : p_i \ge 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i G(X_i;\theta) = 0\}.$$

Wilks Theorem: $-2\log R(\theta_0) \xrightarrow{d} \chi^2(q)$ as $n \to \infty$.

Notes

Profile empirical likelihood method

Suppose we are only interested in a part of θ. Then like the parametric profile likelihood ratio test, we have the profile empirical likelihood method.

Observations: $X_1, \cdots, X_n$ iid with common distribution function F and there is a q-dimensional parameter θ associated with F. Write $\theta = (\alpha^T, \beta^T)^T$, where α and β are q_1-dimensional and q_2-dimensional parameters, respectively, and $q_1 + q_2 = q$. Now we are interested in α.

Conditions: Let y^T denote the transpose of the vector y and

$$G(x; \theta) = (g_1(x; \theta), \cdots, g_s(x; \theta))^T$$

denote $s(\geq q)$ functionally independent functions, which connect F and θ through the equations $EG(X_1; \theta) = 0$.

Notes

Profile empirical likelihood ratio:

$$l(\alpha) = 2l_E((\alpha^T, \hat{\beta}^T(\alpha))^T) - 2l_E(\tilde{\theta}),$$

where $l_E(\theta) = \sum_{i=1}^n \log\{1 + \lambda^T G(X_i; \theta)\}$, $\lambda = \lambda(\theta)$ is the solution of the following equation

$$0 = \frac{1}{n} \sum_{i=1}^n \frac{G(X_i; \theta)}{1 + \lambda^T G(X_i; \theta)},$$

$\tilde{\theta} = (\tilde{\alpha}^T, \tilde{\beta}^T)^T$ minimizes $l_E(\theta)$ with respect to θ, and $\hat{\beta}(\alpha)$ minimizes $l_E((\alpha^T, \beta^T)^T)$ with respect to β for fixed α.

Wilks Theorem: $l(\alpha_0) \xrightarrow{d} \chi^2(q_1)$ as $n \to \infty$, where α_0 denotes the true value of α.

Notes

Jackknife empirical likelihood method

Empirical likelihood method has difficulties in dealing with nonlinear functionals.

Example: Covariance. Suppose $(X_1, Y_1), \cdots, (X_n, Y_n)$ are iid with covariance $\sigma_{12} = E\{(X_1 - E(X_1))(Y_1 - E(Y_1))\}$, and we are interested in testing $H_0 : \sigma_{12} = \sigma_0$ against $H_a : \sigma_{12} \neq \sigma_0$.

Method 1: Define the empirical likelihood function

$$R(\sigma_{12}) = \sup\{\prod_{i=1}^n (np_i) : p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i\{X_i - \sum_{j=1}^n p_j X_j\}\{Y_i - \sum_{j=1}^n p_j Y_j\} = \sigma_{12}\}.$$

In this way, the above minimization is too complicated due to no formula for $p_i's$.

Notes

Method 2: Define the empirical likelihood function

$$R(\sigma_{12}) = \sup\{\prod_{i=1}^n (np_i) : p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i\{X_i - n^{-1}\sum_{j=1}^n X_j\}\{Y_i - n^{-1}\sum_{j=1}^n Y_j\} = \sigma_{12}\}.$$

Then $-2\log R(\sigma_0)$ can not converge in distribution to a chi-squared distribution since the above procedure fails to catch the variances contribution made by $n^{-1}\sum_{i=1}^n X_i$ and $n^{-1}\sum_{i=1}^n Y_i$. As a matter of fact, the limit is a weighted sum of two independent chi-squared random variables.

Notes

Method 3: Define the empirical likelihood function as

$$R(\mu_1, \mu_2\sigma_{12}) = \sup\{\prod_{i=1}^n (np_i) : p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i X_i = \mu_1, \sum_{i=1}^n p_i Y_i = \mu_2, \sum_{i=1}^n p_i (X_i - \mu_1)(Y_i - \mu_2) = \sigma_{12}\}$$

and the profile empirical likelihood function as

$$R_P(\sigma_{12}) = \max_{\mu_1,\mu_2} R(\mu_1, \mu_2, \sigma_{12}).$$

Wilks theorem: $-2\log R_P(\sigma_0) \xrightarrow{d} \chi^2(1)$ as $n \to \infty$.
Computional issues: After introducing the link variable μ, the computation is increased.

Notes

Jackknife empirical likelihood for covariance: Put

$$\hat{\sigma}_n = \frac{1}{n}\sum_{i=1}^n \{X_i - n^{-1}\sum_{j=1}^n X_j\}\{Y_i - n^{-1}\sum_{j=1}^n Y_j\}$$

and

$$\hat{\sigma}_{n,l} = \frac{1}{n-1}\sum_{i\neq l}\{X_i - \frac{1}{n-1}\sum_{j\neq l} X_j\}\{Y_i - \frac{1}{n-1}\sum_{j\neq l} Y_j\}.$$

Then define the jacknife sample as

$$Z_l = n\hat{\sigma}_n - (n-1)\hat{\sigma}_{n,l}$$

for $l = 1, \cdots, n$, and define the jackknife empirical likelihood function as

$$R(\sigma_{12}) = \sup\{\prod_{i=1}^n (np_i); p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i Z_i = \sigma_{12}\}.$$

Notes

Wilks Theorem: $-2\log R(\sigma_0) \xrightarrow{d} \chi^2(1)$ as $n \to \infty$.
Computation: In R, one can employ the package 'emplik'.

Notes

Interval estimation for ν. A direct application of empirical likelihood method fails to catch the contribution made by the first step estimation for ρ. Here we consider the jackknife empirical likelihood. In order to construct a jackknife sample as in Jing, Yuan and Zhou (2009), we first define for $i = 1, \cdots, n$ $\hat{\rho}_i = \sin(\pi\hat{\tau}_i/2)$,

$$\hat{\tau}_i = \frac{2}{(n-1)(n-2)} \sum_{1\le j<l\le n, j\ne i, l\ne i} \mathrm{sign}((X_j - X_l)(Y_j - Y_l)),$$

$$F_{n1,i}(x) = \frac{1}{n}\sum_{j\ne i} I(X_j \le x), \quad F_{n2,i}(y) = \frac{1}{n}\sum_{j\ne i} I(Y_j \le y),$$

and then define the jackknife sample as

$$Z_i(\nu) = \sum_{j=1}^n l(\hat{\rho}, \nu; F_{n1}(X_j), F_{n2}(Y_j)) - \sum_{j\ne i} l(\hat{\rho}_i, \nu; F_{n1,i}(X_j), F_{n2,i}(Y_j))$$

for $i = 1, \cdots, n$.

Notes

Based on this jackknife sample, the jackknife empirical likelihood function for ν is defined as

$$L_1(\nu) = \sup\{\prod_{i=1}^n (np_i) : p_1 \ge 0, \cdots, p_n \ge 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i Z_i(\nu) = 0\}.$$

By the Lagrange multiplier technique, we have

$$l_1(\nu) := -2\log L_1(\nu) = 2\sum_{i=1}^n \log\{1 + 2\lambda_1 Z_i(\nu)\},$$

where $\lambda_1 = \lambda_1(\nu)$ satisfies

$$\sum_{i=1}^n \frac{Z_i(\nu)}{1 + \lambda_1 Z_i(\nu)} = 0.$$

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

Notes

Theorem 2. As $n \to \infty$, $l_1(\nu_0)$ converges in distribution to a chi-square limit with one degree of freedom, where ν_0 denotes the true value of ν.
Based on the above theorem, one can construct a confidence interval with level α for ν_0 without estimating the asymptotic variance as

$$I_1(\alpha) = \{\nu : l_1(\nu) \le \chi^2_{1,\alpha}\},$$

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

Notes

Interval estimation for (ρ, ν). Since the Kendall's tau estimator is not a linear functional, one can not apply the empirical likelihood method directly to construct a confidence region for (ρ, ν). Here we employ the jackknife empirical likelihood method by defining the jackknife empirical likelihood function as

$$\begin{aligned} L_2(\rho,\nu) &= \sup\{\textstyle\prod_{i=1}^n (np_i) : p_1 \ge 0, \cdots, p_n \ge 0, \sum_{i=1}^n p_i = 1, \\ &\textstyle\sum_{i=1}^n p_i Z_i(\nu) = 0, \sum_{i=1}^n p_i(n\hat{\rho} - (n-1)\hat{\rho}_i) = \rho\}. \end{aligned}$$

Notes

Theorem 3. As $n \to \infty$, $-2 \log L_2(\rho_0, \nu_0)$ converges in distribution to a chi-square limit with two degrees of freedom, where $(\rho_0, \nu_0)^T$ denotes the true value of $(\rho, \nu)^T$.
Based on the above theorem, one can construct a confidence region with level α for $(\rho_0, \nu_0)^T$ without estimating the asymptotic variance as

$$I_2(\alpha) = \{(\rho, \nu) : -2 \log L_2(\rho, \nu) \le \chi^2_{2,\alpha}\},$$

where $\chi^2_{2,\alpha}$ denotes the α-th quantile of a chi-square limit with two degrees of freedom.

Notes

Simulation

We investigate the finite sample behavior of the proposed jackknife empirical likelihood method for constructing confidence intervals for ν and compare it with the parametric bootstrap method in terms of coverage probability.
We employ the R packge 'copula' to draw $1,000$ random samples with size $n = 200, 500$ from the t-copula with $\rho = 0.1, 0.5, 0.9$ and $\nu = 3, 8$. For computing the confidence interval based on normal approximation, we use parametric bootstrap method to obtain the critical values by resampling $1,000$ samples with size n from the t-copula with parameters $\hat{\rho}$ and $\hat{\nu}$. The R package 'emplik' is employed to compute the coverage probability of the proposed jackknife empirical likelihood method.

Notes

(n, ρ, ν)	JELM Level 90%	NAM Level 90%	JELM Level 95%	NAM Level 95%
$(200, 0.1, 3)$	0.886	0.813	0.935	0.844
$(200, 0.5, 3)$	0.849	0.771	0.908	0.802
$(200, 0.9, 3)$	0.878	0.826	0.928	0.849
$(200, 0.1, 8)$	0.831	0.600	0.909	0.615
$(200, 0.5, 8)$	0.815	0.594	0.886	0.611
$(200, 0.9, 8)$	0.837	0.664	0.902	0.680
$(500, 0.1, 3)$	0.871	0.825	0.923	0.853
$(500, 0.5, 3)$	0.874	0.838	0.933	0.870
$(500, 0.9, 3)$	0.876	0.844	0.932	0.869
$(500, 0.1, 8)$	0.871	0.728	0.939	0.760
$(500, 0.5, 8)$	0.862	0.747	0.920	0.769
$(500, 0.9, 8)$	0.892	0.774	0.942	0.797

Notes

Data Analysis

First we fit the bivariate t-copula to the data set on 3283 daily log-returns of equity for two major Dutch banks, ING and ABN AMRO Bank, over the period 1991–2003, giving $\hat{\rho} = 0.682$ and $\hat{\nu} = 2.617$. The empirical likelihood ratio function $l_1(\nu)$ is plotted against ν below from 1.501 to 3.5 with step 0.001, which shows that the proposed jackknife empirical likelihood intervals for ν are $(2.280, 3.042)$ for level 0.9 and $(2.246, 3.129)$ for level 0.95. The normal-approximation-based intervals for ν are $(2.257, 2.910)$ for level 0.9 and $(2.195, 2.962)$ for level 0.95. As we see, the intervals based on the jackknife empirical likelihood method are slightly longer and more skewed to the right than those based on the normal approximation method.

Notes

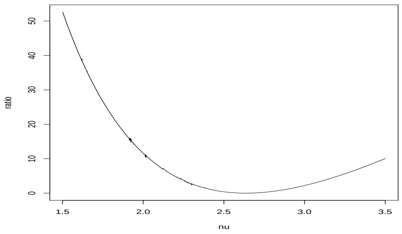


Figure: Equity. The empirical likelihood ratio $l_1(\nu)$ is plotted against ν from 1.501 to 3.5 with step 0.001 for the daily log-returns of equity for two major Dutch banks (ING and ABN AMRO Bank) over the period 1991–2003.

Notes

Second we fit the t copula to the nonzero losses to building and content in the Danish fire insurance claims. This data set is available at *www.ma.hw.ac.uk/∼mcneil/*, which comprises 2167 fire losses over the period 1980 to 1990. We find that $\hat{\rho} = 0.134$ and $\hat{\nu} = 9.474$. The proposed jackknife empirical likelihood intervals for ν are $(6.830, 16.285)$ and $(6.415, 17.785)$ for levels 0.9 and 0.95 respectively, and the normal-approximation-based intervals for ν are $(0.978, 12.719)$ and $(-2.242, 13.070)$ for levels 0.9 and 0.95 respectively. The above negative value is due to some large values of the bootstrapped estimators of ν. It is clear that the proposed jackknife empirical likelihood intervals are shorter and more skewed to the right than the normal approximation based intervals.

Notes

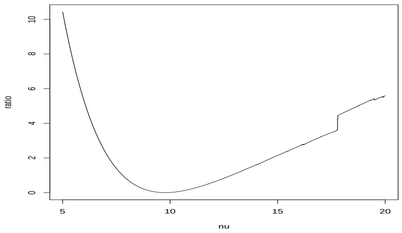


Figure: Danish fire losses. The empirical likelihood ratio $l_1(\nu)$ is plotted against ν from 5.005 to 20 with step 0.005 for the nonzero losses to building and content in the Danish fire insurance claims.

Notes

Proofs

Proof of Theorem 1. Define

$$g(x,y) = \mathbb{E}\mathrm{sign}((x-X_1)(y-Y_1)) - \tau$$
$$= 4\{C(F_1(x),F_2(y)) - \mathbb{E}C(F_1(X_1),F_2(Y_1))\} -2\{F_1(x)-\tfrac{1}{2}\} - 2\{F_2(y)-\tfrac{1}{2}\},$$

$$\psi(x_1,y_1,x_2,y_2) = \mathrm{sign}((x_1-x_2)(y_1-y_2)) - g(x_1,y_1) - g(x_2,y_2).$$

It follows from the Hoeffding decomposition and results in Hoeffding (1948) that

$$\sqrt{n}\{\hat{\tau}-\tau\} = \frac{2}{\sqrt{n}}\sum_{i=1}^n g(X_i,Y_i) + \frac{2\sqrt{n}}{n(n-1)}\sum_{1\le i<j\le n}\psi(X_i,Y_i,X_j,Y_j) = \frac{2}{\sqrt{n}}\sum_{i=1}^n g(X_i,Y_i) + o_p(1), \quad (6)$$

which implies (3).

Notes

By the Taylor expansion, we have

$$\begin{aligned}
0 &= \frac{1}{\sqrt{n}}\sum_{i=1}^n l(\hat{\rho},\hat{\nu};F_{n1}(X_i),F_{n2}(Y_i)) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n l(\rho,\nu;F_{n1}(X_i),F_{n2}(Y_i)) \\
&\quad +\frac{1}{\sqrt{n}}\sum_{i=1}^n\{\frac{\partial}{\partial\rho}l(\rho,\nu;F_{n1}(X_i),F_{n2}(Y_i))\}(\hat{\rho}-\rho) \\
&\quad +\frac{1}{\sqrt{n}}\sum_{i=1}^n\{\frac{\partial}{\partial\nu}l(\rho,\nu;F_{n1}(X_i),F_{n2}(Y_i))\}(\hat{\nu}-\nu)+o_p(1) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n l(\rho,\nu;F_1(X_i),F_2(Y_i)) \\
&\quad +\frac{1}{\sqrt{n}}\sum_{i=1}^n l_u(\rho,\nu;F_1(X_i),F_2(Y_i))\{F_{n1}(X_i)-F_1(X_i)\} \\
&\quad +\frac{1}{\sqrt{n}}\sum_{i=1}^n l_v(\rho,\nu;F_1(X_i),F_2(Y_i))\{F_{n2}(Y_i)-F_2(Y_i)\} \\
&\quad +\frac{1}{n}\sum_{i=1}^n\{\frac{\partial}{\partial\rho}l(\rho,\nu;F_1(X_i),F_2(Y_i))\}\sqrt{n}(\hat{\rho}-\rho) \\
&\quad +\frac{1}{n}\sum_{i=1}^n\{\frac{\partial}{\partial\nu}l(\rho,\nu;F_1(X_i),F_2(Y_i))\}\sqrt{n}(\hat{\nu}-\nu)+o_p(1),
\end{aligned} \quad (7)$$

which implies (4). More details can be found in Wang, Peng and Yang (2013).

Notes

The values of σ_1^2, σ_{12} and σ_2^2 can be calculated straightforward by using the Law of Large Numbers, which are

$$\sigma_1^2 = \cos^2(\tfrac{\pi\tau}{2})\pi^2\{8\int_0^1\int_0^1\{2C^2(u,v)-2(u+v)C(u,v)+uv\}\, dC(u,v)+\tfrac{5}{3}-\tau^2+2\tau\},$$

$$\sigma_2^2 = K_\nu^{-2}(K^2+R_1+R_2+2R_3+2R_4+2R_5+K_\rho^2\sigma_1^2+2K_\rho(L_1+L_2+L_3)),$$

$$\sigma_{12}^2 = -K_\nu^{-1}(K_\rho\sigma_1^2+L_1+L_2+L_3),$$

where

Notes

$$K^2 = \int_0^1 \int_0^1 l(\rho, \nu; u, v)^2 \, dC(u, v),$$

$$R_1 = \int_0^1 \int_0^1 \int_0^1 \int_0^1 l_u(\rho, \nu; u_1, v_1) l_u(\rho, \nu; u_2, v_2)(u_1 \wedge u_2 - u_1 u_2) \, dC(u_1, v_1) dC(u_2, v_2),$$

$$R_2 = \int_0^1 \int_0^1 \int_0^1 \int_0^1 l_v(\rho, \nu; u_1, v_1) l_v(\rho, \nu; u_2, v_2)(v_1 \wedge v_2 - v_1 v_2) \, dC(u_1, v_1) dC(u_2, v_2),$$

$$R_3 = \int_0^1 \int_0^1 \int_0^1 \int_0^1 l_u(\rho, \nu; u_1, v_1) l_v(\rho, \nu; u_2, v_2)(C(u_1, v_2) - u_1 v_2) \, dC(u_1, v_1) dC(u_2, v_2),$$

Notes

$$R_4 = \int_0^1 \int_0^1 \int_0^1 \int_0^1 l_u(\rho, \nu; u_1, v_1) l(\rho, \nu; u_2, v_2)(I(u_2 \le u_1) - u_1) \, dC(u_1, v_1) dC(u_2, v_2),$$

$$R_5 = \int_0^1 \int_0^1 \int_0^1 \int_0^1 l_v(\rho, \nu; u_1, v_1) l(\rho, \nu; u_2, v_2)(I(v_2 \le v_1) - v_1) \, dC(u_1, v_1) dC(u_2, v_2),$$

$$L_1 = \cos(\frac{\pi\tau}{2})\pi \int_0^1 \int_0^1 l(\rho, \nu; u, v)\{4C(u, v) - 2u - 2v\} \, dC(u, v),$$

$$L_2 = \cos(\tfrac{\pi\tau}{2})\pi \int_0^1 \int_0^1 \int_0^1 \int_0^1 l_u(\rho, \nu; u_1, v_1)\{4C(u_2, v_2) - 2u_2 - 2v_2\} \times \{I(u_2 \le u_1) - u_1\} \, dC(u_1, v_1) dC(u_2, v_2),$$

and

$$L_3 = \cos(\tfrac{\pi\tau}{2})\pi \int_0^1 \int_0^1 \int_0^1 \int_0^1 l_v(\rho, \nu; u_1, v_1)\{4C(u_2, v_2) - 2u_2 - 2v_2\} \times \{I(v_2 \le v_1) - v_1\} \, dC(u_1, v_1) dC(u_2, v_2).$$

Notes

THANK YOU

Notes