Reduce computation in profile empirical likelihood method

Minqiang LI[1], Liang PENG[2]* and Yongcheng QI[3]

[1]*Bloomberg LP, 731 Lexington Avenue, New York, NY 10022, USA*
[2]*School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA*
[3]*Department of Mathematics and Statistics, University of Minnesota–Duluth, Duluth, MN 55812, USA*

Abstract: Since its introduction by Owen (1988, 1990), the empirical likelihood method has been extensively investigated and widely used to construct confidence regions and to test hypotheses in the literature. For a large class of statistics that can be obtained via solving estimating equations, the empirical likelihood function can be formulated from these estimating equations as proposed by Qin and Lawless (1994). If only a small part of parameters is of interest, a profile empirical likelihood method has to be employed to construct confidence regions, which could be computationally costly. In this article the authors propose a jackknife empirical likelihood method to overcome this computational burden. This proposed method is easy to implement and works well in practice. *The Canadian Journal of Statistics* 39: 370–384; 2011 © 2011 Statistical Society of Canada

Résumé: Depuis leur introduction par Owen (1988, 1990), la méthode de vraisemblance empirique a été étudiée de façon exhaustive et elle est beaucoup utilisée dans la littérature pour construire des régions de confiance et confronter des hypothèses. Pour une grande classe de statistiques obtenues en résolvant des équations d'estimation, la fonction de vraisemblance empirique peut être formulée à partir de ces équations d'estimation telles que proposées par Qin et Lawless (1994). Lorsqu'uniquement une petite partie des paramètres sont d'intérêt, une méthode de vraisemblance empirique de profil doit être utilisée pour construire une région de confiance ce qui peut s'avérer très coûteux à évaluer numériquement. Dans cet article, les auteurs proposent une version jackknife de la méthode de vraisemblance empirique pour surmonter les coûts de calculs. Cette méthode est facile à implanter et elle fonctionne bien en pratique. *La revue canadienne de statistique* 39: 370–384; 2011 © 2011 Société statistique du Canada

1. INTRODUCTION

Empirical likelihood method was introduced by Owen (1988, 1990) to construct confidence regions for the mean of a random vector. Like the bootstrap and jackknife methods, the empirical likelihood method is a nonparametric one. Without assuming a family of distributions for the data, the empirical likelihood ratio statistics can be defined to share similar properties as the likelihood ratio for parametric distributions. For instance, the empirical likelihood method produces confidence regions whose shape and orientation are determined entirely by the data. In comparison with the normal approximation method and the bootstrap method for constructing confidence intervals, the empirical likelihood method does not require a pivotal quantity, and it has better small sample performance (see Hall & La Scala, 1990).

* *Author to whom correspondence may be addressed.*
 E-mail: peng@math.gatech.edu

As an effective way for interval estimation and goodness-of-fit test, the empirical likelihood method has been extended and applied in many different fields such as regression models (Chen & Van Keilegom, 2009), quantile estimation (Chen & Hall, 1993), additive risk models (Lu & Qi, 2004), two-sample problems (Zhou & Liang, 2005; Cao & Van Keilegom, 2006; Keziou & Leoni-Aubin, 2008; Ren, 2008), time series models (Hall & Yao, 2003; Chan, Peng & Qi, 2006; Nordman & Lahiri, 2006; Chen & Gao, 2007; Nordman, Sibbertsen & Lahiri, 2007; Guggenberger & Smith, 2008), heavy-tailed models (Lu & Peng, 2002; Peng, 2004; Peng & Qi, 2006a, b), high dimensional data (Chen, Peng & Qin, 2009), and copulas (Chen, Peng & Zhao, 2009).

A common feature in employing the empirical likelihood method is to work with linear constraints such as linear functionals, and a powerful way in formulating the empirical likelihood ratio statistic is via estimating equations as in Qin and Lawless (1994). When the constraints involve nonlinear equations, an important trick is to transform these nonlinear constraints to some linear constraints by introducing some link variables. After this transformation, a profile empirical likelihood method is employed. For example, if one wants to construct an empirical likelihood confidence interval for the variance $\theta = E\{X - EX\}^2$, one can introduce a link variable $\beta = E(X)$ and then formulate the empirical likelihood by using equations $E(X) = \beta$ and $E(X^2) = \theta + \beta^2$. Some other examples include the study of ROC curves (Claeskens et al., 2003), copulas (Chen, Peng & Zhao, 2009), and difference of quantiles (Zhou & Jing, 2003; Shen & He, 2007). Apparently these extra constraints add more computational burden to the empirical likelihood method. Another computational difficulty arises when the profile empirical likelihood method involves a large number of nuisance parameters.

Suppose $X_1, \ldots, X_n$ are independent random vectors with common distribution function F and there is a q-dimensional parameter θ associated with F. Let y^{T} denote the transpose of the vector y and

$$G(x; \theta) = (g_1(x; \theta), \ldots, g_s(x; \theta))^{\mathrm{T}}$$

denote $s(\geq q)$ functionally independent functions, which connect F and θ through the equations $EG(X_1; \theta) = 0$. Write $\theta = (\alpha^{\mathrm{T}}, \beta^{\mathrm{T}})^{\mathrm{T}}$, where α and β are q_1-dimensional and q_2-dimensional parameters, respectively, and $q_1 + q_2 = q$. In order to construct confidence regions for α, Qin and Lawless (1994) proposed to use the following profile empirical likelihood ratio

$$l(\alpha) = 2l_E((\alpha^{\mathrm{T}}, \hat{\beta}^{\mathrm{T}}(\alpha))^{\mathrm{T}}) - 2l_E(\tilde{\theta}), \tag{1}$$

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

$$0 = \frac{1}{n} \sum_{i=1}^{n} \frac{G(X_i; \theta)}{1 + t^{\mathrm{T}}G(X_i; \theta)} \tag{2}$$

$\tilde{\theta} = (\tilde{\alpha}^{\mathrm{T}}, \tilde{\beta}^{\mathrm{T}})^{\mathrm{T}}$ minimizes $l_E(\theta)$ with respect to θ, and $\hat{\beta}(\alpha)$ minimizes $l_E((\alpha^{\mathrm{T}}, \beta^{\mathrm{T}})^{\mathrm{T}})$ with respect to β for fixed α. It has been shown that $l(\alpha_0)$ converges in distribution to $\chi^2_{q_1}$ under some regularity conditions, where α_0 denotes the true value of α. For the second order properties of the empirical likelihood method based on estimating equations, we refer to Chen and Cui (2007).

The computational complexity in using the profile empirical likelihood method comes from computing $l_E((\alpha^{\mathrm{T}}, \hat{\beta}^{\mathrm{T}}(\alpha))^{\mathrm{T}})$. When the nuisance parameter β is known, one can simply replace $\hat{\beta}^{\mathrm{T}}(\alpha)$ and $\tilde{\beta}$ in (1) by the true value of β so that the computation is reduced significantly. In order to avoid computing $l_E((\alpha^{\mathrm{T}}, \hat{\beta}^{\mathrm{T}}(\alpha))^{\mathrm{T}})$, one may choose to replace $\hat{\beta}^{\mathrm{T}}(\alpha)$ by some other different estimator, for example, solving q_2 equations of $n^{-1} \sum_{i=1}^{n} G(X_i; \theta) = 0$. Although this reduces computation especially when one can find an explicit estimator of β in terms of the sample and α, Wilks' theorem does not hold, that is, $l(\alpha_0)$ does not converge in distribution to $\chi^2_{q_1}$. Instead,

$l(\alpha_0)$ generally converges in distribution to a weighted sum of independent chi-square random variables; see Hjort, McKeague and Van Keilegom (2009). Since the weights in the limit have to be estimated, this empirical likelihood method does not preserve the important properties of the standard empirical likelihood method: self-studentization, automatically determined shape of confidence region, and Bartlett correction. Therefore, it is of importance to develop an empirical likelihood method which has a chi-square limit and is computationally efficient than the profile empirical likelihood method especially when the number of nuisance parameters is large. Moreover, when some estimating equations involve U-statistics, the profile empirical likelihood method is extremely complicated.

Motivated by the recent study on using jackknife empirical likelihood method to deal with nonlinear constraints in U-statistics (Jing, Yuan & Zhou, 2009), we propose a jackknife empirical likelihood method to construct confidence regions for the interesting parameter α with the nuisance parameter β being simply replaced by some estimator. The jackknife empirical likelihood ratio statistic is obtained by applying the standard empirical likelihood method to the jackknife pseudo sample. The jackknife method was originally used to estimate the variance of a statistic and to construct bias-corrected estimators of parameters. See, for example, Shao and Tu (1995) for details. The proposed jackknife empirical likelihood method allows us to compute the nuisance parameters simply through a subset of estimating equations and yet still retains the attractive chi-square limiting distribution for the empirical likelihood ratio.

We organize this article as follows. In Section 2, the new methodology and main results are given. Section 3 presents a simulation study. Section 4 presents two case studies of financial applications. To save space, an outline of proofs is given in the Appendix and detailed proofs can be found in the technique report, which is available at http://people.math.gatech.edu/~peng/.

2. METHODOLOGY AND MAIN RESULTS

As in the introduction, let $G(x; \theta) = (g_1(x; \theta), \ldots, g_s(x; \theta))^{\mathrm{T}}$ denote $s(\geq q)$ functionally independent functions with $EG(X_1; \theta_0) = 0$, where $\theta_0 = (\alpha_0^{\mathrm{T}}, \beta_0^{\mathrm{T}})^{\mathrm{T}}$ denotes the true value of $\theta = (\alpha^{\mathrm{T}}, \beta^{\mathrm{T}})^{\mathrm{T}}$, and α and β are q_1-dimensional and q_2-dimensional parameters, respectively. Note that only the parameter α is of interest under consideration. To remove the nuisance parameter β, we propose to first estimate it from some q_2 estimating equations, and then work with the remaining $s - q_2$ equations, where β is replaced by the obtained estimator. The details are as follows:

Define

$$G_a(x; \alpha, \beta) = (g_1(x; \alpha, \beta), \ldots, g_{s-q_2}(x; \alpha, \beta))^{\mathrm{T}}$$

and

$$G_b(x; \alpha, \beta) = (g_{s-q_2+1}(x; \alpha, \beta), \ldots, g_s(x; \alpha, \beta))^{\mathrm{T}}$$

Without loss of generality, we solve the last q_2 equations of $n^{-1} \sum_{i=1}^n G(X_i; \theta)$ to get an estimator for β. That is, $\tilde{\beta}(\alpha; \mathbf{X})$ is the solution to

$$\frac{1}{n} \sum_{i=1}^n G_b(X_i; \alpha, \beta) = 0 \tag{3}$$

with respect to β for each fixed α, where $\mathbf{X} = (X_1, \ldots, X_n)^{\mathrm{T}}$. Obviously, the best choice of the q_2 equations is to have explicit formula for $\tilde{\beta}(\alpha; \mathbf{X})$, if possible. Set

$$T_n(\alpha) = \frac{1}{n} \sum_{i=1}^{n} G_a(X_i; \alpha, \tilde{\beta}(\alpha, \mathbf{X}))$$

and let $\tilde{\beta}(\alpha; \mathbf{X}_{-i})$ denote the solution to the equations

$$\frac{1}{n-1} \sum_{j=1, j\neq i}^{n} G_b(X_j; \alpha, \beta) = 0 \tag{4}$$

with respect to β for each fixed α, where $\mathbf{X}_{-i} = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)^{\mathrm{T}}$. Similarly, define

$$T_{n,-i}(\alpha) = \frac{1}{n-1} \sum_{j=1, j\neq i}^{n} G_a(X_j; \alpha, \tilde{\beta}(\alpha; \mathbf{X}_{-i})).$$

Then the jackknife pseudo sample is defined as

$$Y_i(\alpha) = (Y_{i,1}(\alpha), \ldots, Y_{i,s-q_2}(\alpha))^{\mathrm{T}} = nT_n(\alpha) - (n-1)T_{n,-i}(\alpha) \quad \text{for} \quad i = 1, \ldots, n$$

As in Tukey (1958), one expects that $Y_i's$ are approximately independent. This motivates us to apply the standard empirical likelihood method to the jackknife sample $Y_1(\alpha), \ldots, Y_n(\alpha)$ for constructing empirical likelihood confidence regions for α. Hence we define the jackknife empirical likelihood function as

$$L^J(\alpha) = \sup\left\{ \prod_{i=1}^{n} (np_i) : \sum_{i=1}^{n} p_i = 1, \quad \sum_{i=1}^{n} p_i Y_i(\alpha) = 0, \quad p_1 \geq 0, \ldots, p_n \geq 0 \right\}$$

It follows from the Lagrange multiplier technique that the above maximization is achieved at $p_i = n^{-1}\{1 + \lambda^{\mathrm{T}} Y_i(\alpha)\}^{-1}$ and the log empirical likelihood ratio $\ell^J(\alpha) = -\log L^J(\alpha)$ is given by

$$\ell^J(\alpha) = \sum_{i=1}^{n} \log\{1 + \lambda^{\mathrm{T}} Y_i(\alpha)\} \tag{5}$$

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

$$\frac{1}{n} \sum_{i=1}^{n} \frac{Y_i(\alpha)}{1 + \lambda^{\mathrm{T}} Y_i(\alpha)} = 0. \tag{6}$$

Before we present our main results, we first list the assumptions we need. For this purpose, denote

$$\frac{\partial y}{\partial x} = \left(\frac{\partial y_i}{\partial x_j} \right)_{1 \leq i \leq m, 1 \leq j \leq n}$$

for $y = (y_1, \ldots, y_m)^{\mathrm{T}}$ and $x = (x_1, \ldots, x_n)^{\mathrm{T}}$, and define

$$\hat{\alpha} = \arg\min \ell^J(\alpha), \quad \Sigma = (E\{g_k(X_1; \alpha_0, \beta_0) g_l(X_1; \alpha_0, \beta_0)\})_{1 \le k, l \le s} = (\sigma_{k,l})_{1 \le k, l \le s},$$

$$\Sigma_1 = E\Big\{ \tfrac{\partial}{\partial \beta} G_b(X_1; \alpha_0, \beta_0) \Big\}, \quad \Sigma_2 = E\Big\{ \tfrac{\partial}{\partial \alpha} G_b(X_1; \alpha_0, \beta_0) \Big\},$$

$$\Sigma_3 = E\Big\{ \tfrac{\partial}{\partial \alpha} G_a(X_1; \alpha_0, \beta_0) \Big\} - E\Big\{ \tfrac{\partial}{\partial \beta} G_a(X_1; \alpha_0, \beta_0) \Big\} \Sigma_1^{-1} \Sigma_2$$

and $\Sigma^* = (\sigma^*_{k,l})_{1 \le k, l \le s - q_2}$ where

$$\begin{aligned}
\sigma^*_{k,l} = {} & \sigma_{k,l} - E\Big\{ \tfrac{\partial}{\partial \beta} g_k(X_1; \alpha_0, \beta_0) \Big\} \Sigma_1^{-1} (\sigma_{s-q_2+1,l}, \ldots, \sigma_{s,l})^{\mathrm{T}} \\
& - E\Big\{ \tfrac{\partial}{\partial \beta} g_l(X_1; \alpha_0, \beta_0) \Big\} \Sigma_1^{-1} (\sigma_{s-q_2+1,k}, \ldots, \sigma_{s,k})^{\mathrm{T}} \\
& + E\Big\{ \tfrac{\partial}{\partial \beta} g_k(X_1; \alpha_0, \beta_0) \Big\} \Sigma_1^{-1} (\sigma_{ij})_{s-q_2+1 \le i, j \le s} \Sigma_1^{-1} E\Big\{ \tfrac{\partial}{\partial \beta^{\mathrm{T}}} g_l(X_1; \alpha_0, \beta_0) \Big\}
\end{aligned}$$

Some regularity conditions are as follows:

- (A1) There is a neighborhood of α_0 and β_0, say $\Omega_{\alpha_0} \times \Omega_{\beta_0}$, such that $G_b(x; \alpha, \beta)$ are continuous function of $\alpha \in \Omega_{\alpha_0}$ and $\beta \in \Omega_{\beta_0}$ for all x, and $\sup_{\alpha \in \Omega_{\alpha_0}, \beta \in \Omega_{\beta_0}} \|G_b(x; \alpha, \beta)\|^3 \le K(x)$ for some function K satisfying that $EK(X_1) < \infty$;
- (A2) For each $\alpha \in \Omega_{\alpha_0}$, there is a function $\beta(\alpha) \in \Omega_{\beta_0}$ such that $EG_b(X_1; \alpha, \beta(\alpha)) = 0$;
- (A3) $\Big\| \tfrac{\partial}{\partial \beta} G_b(x; \alpha, \beta) \Big\|^3$, $\Big\| \tfrac{\partial^2}{\partial \beta^{\mathrm{T}} \partial \beta} g_l(x; \alpha, \beta) \Big\|^3$, and $\Big| \tfrac{\partial^3}{\partial \beta_i \partial \beta_j \partial \beta_m} g_l(x; \alpha, \beta) \Big|$ for $l = s - q_2 + 1, \ldots, s$, $i, j, m = 1, \ldots, q_2$ are bounded by $K(x)$ uniformly in $\alpha \in \Omega_{\alpha_0}$ and $\beta \in \Omega_{\beta_0}$;
- (A4) $\beta(\alpha)$ defined in (A2) has continuous first derivatives;
- (A5) $\Big\| \tfrac{\partial}{\partial \alpha} G_b(x; \alpha, \beta) \Big\|^3$ and $\Big| \tfrac{\partial^3}{\partial \alpha_i \partial \beta_j \partial \beta_m} g_l(x; \alpha, \beta) \Big|$ for $l = s - q_2 + 1, \ldots, s$, $i = 1, \ldots, q_1$, $j, m = 1, \ldots, q_2$ are bounded by $K(x)$ uniformly in $\alpha \in \Omega_{\alpha_0}$ and $\beta \in \Omega_{\beta_0}$;
- (A6) Assume Σ_1 is invertible, Σ^* is positive definite and Σ_3 has rank q_1;
- (A7) $\Big\| \tfrac{\partial}{\partial \beta} G_a(x; \alpha, \beta) \Big\|$, $\|G_a(x; \alpha, \beta)\|^3$, $\Big| \tfrac{\partial^3}{\partial \alpha_k \partial \beta_j \partial \beta_m} g_l(x; \alpha, \beta) \Big|$, and $\Big| \tfrac{\partial^3}{\partial \beta_i \partial \beta_j \partial \beta_m} g_l(x; \alpha, \beta) \Big|$ for $i, j, m = 1, \ldots, q_2$, $k = 1, \ldots, q_1$, and $l = 1, \ldots, s - q_2$ are bounded by $K(x)$ uniformly in $\alpha \in \Omega_{\alpha_0}$ and $\beta \in \Omega_{\beta_0}$.

The following two propositions show the existence and some properties of $\tilde{\beta}(\alpha; \mathbf{X})$, $\tilde{\beta}(\alpha_0; \mathbf{X}_{-i})$, and $\hat{\alpha}$.

Proposition 1. *(i) Under conditions (A1) and (A2), with probability one, there exist solutions $\tilde{\beta}(\alpha; \mathbf{X}) \in \Omega_{\beta_0}$ and $\tilde{\beta}(\alpha; \mathbf{X}_{-i}) \in \Omega_{\beta_0}$ to (3) and (4), respectively, such that*

$$\tilde{\beta}(\alpha; \mathbf{X}) - \beta(\alpha) = o(1) \quad \text{and} \quad \max_{1 \le i \le n} \|\tilde{\beta}(\alpha; \mathbf{X}_{-i}) - \beta(\alpha)\| = o(1) \tag{7}$$

almost surely for each $\alpha \in \Omega_{\alpha_0}$. (ii) Under conditions (A1)–(A3), we have

$$\tilde{\beta}(\alpha_0; \mathbf{X}) - \beta_0 + \Sigma_1^{-1} \frac{1}{n} \sum_{i=1}^{n} G_b(X_i; \alpha_0, \beta_0) = O_p(n^{-1}), \tag{8}$$

$$\max_{1 \le i \le n} \left\| \tilde{\beta}(\alpha_0; \mathbf{X}_{-i}) - \beta_0 + \Sigma_1^{-1} \frac{1}{n-1} \sum_{j=1, j \ne i}^{n} G_b(X_j; \alpha_0, \beta_0) \right\| = O_p(n^{-1}) \tag{9}$$

$$\max_{1 \le i \le n} \left\| \tilde{\beta}(\alpha_0; \mathbf{X}) - \tilde{\beta}(\alpha_0; \mathbf{X}_{-i}) + \Sigma_{n1}^{-1} D_n(i) \right\| = o_p(n^{-3/2}) \tag{10}$$

where

$$\Sigma_{n1} = \frac{1}{n} \sum_{j=1}^{n} \frac{\partial}{\partial \beta} G_b(X_j; \alpha_0, \tilde{\beta}(\alpha_0; \mathbf{X})), \quad D_n(i) = (D_{n,s-q_2+1}(i), \dots, D_{n,s}(i))^{\mathrm{T}}$$

and

$$\begin{aligned}
D_{n,l}(i) &= \tfrac{1}{2} n^{-1} G_b^{\mathrm{T}}(X_i; \alpha_0, \beta_0) \Sigma_1^{-1} E\left\{ \frac{\partial^2}{\partial \beta^{\mathrm{T}} \partial \beta} g_l(X_1; \alpha_0, \beta_0) \right\} \Sigma_1^{-1} n^{-1} G_b(X_i; \alpha_0, \beta_0) \\
&\quad + n^{-1} g_l(X_i; \alpha_0, \beta_0) - n^{-1} \left\{ \frac{\partial}{\partial \beta} g_l(X_i; \alpha_0, \beta_0) \right\} \Sigma_1^{-1} \frac{1}{n-1} \sum_{j=1}^{n} G_b(X_j; \alpha_0, \beta_0) \\
&\quad + n^{-1} \left\{ \frac{\partial}{\partial \beta} g_l(X_i; \alpha_0, \beta_0) \right\} \Sigma_1^{-1} \frac{1}{n-1} G_b(X_i; \alpha_0, \beta_0)
\end{aligned}$$

(iii) Under conditions (A1)–(A5), we have

$$\begin{cases} \max_{1 \le i \le n} \left\| \frac{\partial}{\partial \alpha} \tilde{\beta}(\alpha_0; \mathbf{X}_{-i}) + \Sigma_1^{-1} \Sigma_2 \right\| = O_p(n^{-1/2}), \\ \left\| \frac{\partial}{\partial \alpha} \tilde{\beta}(\alpha_0; \mathbf{X}) + \Sigma_1^{-1} \Sigma_2 \right\| = O_p(n^{-1/2}) \end{cases} \tag{11}$$

and

$$\max_{1 \le i \le n} \left\| \frac{\partial}{\partial \alpha} \tilde{\beta}(\alpha_0; \mathbf{X}) - \frac{\partial}{\partial \alpha} \tilde{\beta}(\alpha_0; \mathbf{X}_{-i}) + \Sigma_1^{-1} A_i \right\| = o_p(n^{-1}) \tag{12}$$

where $A_i = (a_{s-q_2+1}(i), \dots, a_s(i))^{\mathrm{T}}$ and

$$\begin{aligned}
a_l^{\mathrm{T}}(i) &= -\tfrac{1}{n} G_b^{\mathrm{T}}(X_i; \alpha_0, \beta_0) \Sigma_1^{-1} E\left\{ \frac{\partial^2}{\partial \beta^{\mathrm{T}} \partial \alpha} g_l(X_1; \alpha_0, \beta_0) \right\} + \tfrac{1}{n} \frac{\partial}{\partial \alpha} g_l(X_i; \alpha_0, \beta_0) \\
&\quad + \tfrac{1}{n} G_b^{\mathrm{T}}(X_i; \alpha_0, \beta_0) \Sigma_1^{-1} E\left\{ \frac{\partial^2}{\partial \beta^{\mathrm{T}} \partial \beta} g_l(X_1; \alpha_0, \beta_0) \right\} \Sigma_1^{-1} \Sigma_2 \\
&\quad - \tfrac{1}{n} \frac{\partial}{\partial \beta} g_l(X_i; \alpha_0, \beta_0) \Sigma_1^{-1} \Sigma_2
\end{aligned}$$

Proposition 2. *Under (A1)–(A7), with probability tending to one, $\ell^J(\alpha)$ attains its minimum value at some point $\hat{\alpha}$ in the interior of the ball $\|\alpha - \alpha_0\| \le n^{-1/3}$, and $\hat{\alpha}$ and $\hat{\lambda} = \lambda(\hat{\alpha})$ satisfy*

$$Q_{1n}(\hat{\alpha}, \hat{\lambda}) = 0 \quad \text{and} \quad Q_{2n}(\hat{\alpha}, \hat{\lambda}) = 0$$

where

$$Q_{1n}(\alpha, \lambda) = \frac{1}{n} \sum_{i=1}^{n} \frac{Y_i(\alpha)}{1 + \lambda^{\mathrm{T}} Y_i(\alpha)}$$

and

$$Q_{2n}(\alpha, \lambda) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{1 + \lambda^{\mathrm{T}} Y_i(\alpha)} \left\{ \frac{\partial}{\partial \alpha} Y_i(\alpha) \right\}^{\mathrm{T}} \lambda$$

Next we show that Wilks' theorem holds for the proposed jackknife empirical likelihood method.

Theorem 1. *Under conditions (A1)–(A7), we have* $LR(\alpha_0) := 2\ell^J(\alpha_0) - 2\ell^J(\hat{\alpha})$ *converges in distribution to* $\chi^2_{q_1}$, *where* $\hat{\alpha}$, $\tilde{\beta}(\alpha; \mathbf{X})$ *and* $\tilde{\beta}(\alpha; \mathbf{X}_{-i})$ *are given in Propositions 1 and 2.*

Based on Theorem 1, an asymptotically accurate $100\gamma\%$ confidence region for α is given by

$$I_\gamma = \{\alpha : LR(\alpha) \le \chi^2_{q_1}(\gamma)\}$$

where $\chi^2_{q_1}(\gamma)$ is the γ quantile of a χ^2 distribution with q_1 degrees of freedom.

Remark 1. When $s = q$, we have $\ell^J(\hat{\alpha}) = 0$. Moreover, when $\tilde{\beta}(\alpha; \mathbf{X})$ has an explicit formula in terms of the sample $\mathbf{X}$ and α, the computation of the proposed jackknife empirical likelihood method is only slightly heavier than the standard empirical likelihood method. Indeed, the software R package for the empirical likelihood method can be employed as in our simulation study.

Remark 2. If one is interested in a part of $\alpha = (\tilde{\alpha}_1^{\mathrm{T}}, \tilde{\alpha}_2^{\mathrm{T}})^{\mathrm{T}}$, say $\tilde{\alpha}_1$, then we can show that $2\min_{\tilde{\alpha}_2} \ell^J(\alpha) - 2\ell^J(\hat{\alpha})$ converges in distribution to a chi-square limit with the degrees of freedom being the length of $\tilde{\alpha}_1$. This method may be called jackknife profile empirical likelihood method, which is appealing when a part of nuisance parameters can be solved explicitly.

Remark 3. When equations (3) are independent of α, then regularity conditions involving the partial derivatives of G_b with respect to α can be removed.

3. SIMULATION STUDY

Suppose the random vector (X, Y) has marginal distribution functions F_1, F_2 and copula $C(x, y) = P(F_1(X) \le x, F_2(Y) \le y)$. In fitting a parametric family to the copula, a useful quantity is the Spearman's rho defined as $\rho_s = 12E\{F_1(X)F_2(Y)\} - 3$. For example, if one employs the Gaussian copula

$$C(u, v; \theta) = \int_{-\infty}^{\Phi^-(u)} \int_{-\infty}^{\Phi^-(v)} \frac{1}{2\pi\sqrt{1-\theta^2}} \exp\left\{-\frac{s^2 - 2\theta st + t^2}{2(1-\theta^2)}\right\} \mathrm{d}s \mathrm{d}t$$

or the t copula

$$C(u, v; \theta) = \int_{-\infty}^{t_4^-(u)} \int_{-\infty}^{t_4^-(v)} \frac{1}{2\pi\sqrt{1-\theta^2}} \left\{1 + \frac{s^2 - 2\theta st + t^2}{4(1-\theta^2)}\right\}^{-3} \mathrm{d}s \mathrm{d}t$$

where $\theta \in [-1, 1]$ and Φ^- and t_ν^- denote the inverse function of the standard normal distribution function and t distribution function with degrees of freedom ν, respectively, then $\rho_s = 6\pi^{-1}\arc\sin(\theta/2)$. Hence, Spearman's rho is of importance in fitting a parametric copula. Here, we consider constructing a confidence interval for the Spearman's rho by fitting either the Gaussian copula or the t copula and modeling marginals by t distributions. In this case, a profile empirical likelihood method can be employed to construct a confidence interval for ρ_s by considering the following estimating equations

$$\begin{cases} \rho_s = 12E\{F_1(X; \nu_1)F_2(Y; \nu_2)\} - 3 \\ 0 = EX^2 - \frac{\nu_1}{\nu_1 - 2} = EY^2 - \frac{\nu_2}{\nu_2 - 2} \end{cases}$$

where F_1 and F_2 have distributions t_{ν_1} and t_{ν_2}, respectively. On the other hand, the proposed jackknife empirical likelihood method can be employed to the above estimating equations as well.

First we draw 10,000 random samples with sample size $n = 100$ and 300 from the Gaussian copula and t copula by using the package "copula" in R and transform the marginals to have either normal distributions or t distributions. For computing the coverage probabilities of the proposed jackknife empirical likelihood method, we employ the package "emplik" in the software R. For computing the coverage probabilities of the profile empirical likelihood method, we use the package "emplik" to obtain the likelihood ratio as a function of nuisance parameters and then use the package "nlm" to find the minimum. These coverage probabilities are reported in Tables 1 and 2. From these tables, we observe that (i) the proposed jackknife empirical likelihood method performs much better than the profile empirical likelihood method when $n = 100$; and (ii) both methods perform well when $n = 300$ although the jackknife empirical likelihood method is slightly better.

Second, we calculate the average interval lengths for both methods by drawing 1,000 random samples from the above models. Tables 3 and 4 show that the interval length of the proposed jackknife empirical likelihood method is slightly bigger than that of the profile empirical likelihood method.

Third, we draw 10,000 random samples from a three-dimensional normal copula and t copula with marginal distributions t_7, t_8, t_9, and $\theta = (\theta_1, \theta_2, \theta_2)^{\mathrm{T}}$. Then we apply the proposed jackknife empirical likelihood method and the profile empirical likelihood method to construct confidence regions for the three Spearman's rho. Coverage probabilities are reported in Tables 5 and 6, which

TABLE 1: Empirical coverage probabilities for the proposed jackknife empirical likelihood confidence interval (JELCI) and the profile empirical likelihood confidence interval (PELCI) with nominal levels 0.9 and 0.95 for the two-dimensional Gaussian copula and marginal distributions t_7 and t_8.

	Level 0.9 JELCI	Level 0.9 PELCI	Level 0.95 JELCI	Level 0.95 PELCI
$(n, \theta) = (100, 0.25)$	0.8683	0.8001	0.9184	0.8438
$(n, \theta) = (100, 0.5)$	0.8573	0.7970	0.9114	0.8416
$(n, \theta) = (100, 0.75)$	0.8563	0.8066	0.9076	0.8505
$(n, \theta) = (300, 0.25)$	0.8961	0.8950	0.9455	0.9436
$(n, \theta) = (300, 0.5)$	0.8969	0.8960	0.9451	0.9437
$(n, \theta) = (300, 0.75)$	0.8961	0.8945	0.9453	0.9433

TABLE 2: Empirical coverage probabilities for the proposed jackknife empirical likelihood confidence interval (JELCI) and the profile empirical likelihood confidence interval (PELCI) with nominal levels 0.9 and 0.95 for the two-dimensional t copula and marginal distributions t_7 and t_8.

	Level 0.9 JELCI	Level 0.9 PELCI	Level 0.95 JELCI	Level 0.95 PELCI
$(n, \theta) = (100, 0.25)$	0.8679	0.8054	0.9221	0.8518
$(n, \theta) = (100, 0.5)$	0.8605	0.8055	0.9149	0.8531
$(n, \theta) = (100, 0.75)$	0.8575	0.8123	0.9109	0.8612
$(n, \theta) = (300, 0.25)$	0.8948	0.8923	0.9439	0.9420
$(n, \theta) = (300, 0.5)$	0.8963	0.8937	0.9464	0.9439
$(n, \theta) = (300, 0.75)$	0.8935	0.8928	0.9447	0.9434

TABLE 3: Empirical interval lengths for the proposed jackknife empirical likelihood confidence interval (JELCI) and the profile empirical likelihood confidence interval (PELCI) with nominal levels 0.9 and 0.95 for the two-dimensional Gaussian copula and marginal distributions t_7 and t_8.

	Level 0.9 JELCI	Level 0.9 PELCI	Level 0.95 JELCI	Level 0.95 PELCI
$(n, \theta) = (100, 0.25)$	0.949	0.938	1.131	1.117
$(n, \theta) = (100, 0.5)$	1.036	1.011	1.233	1.201
$(n, \theta) = (100, 0.75)$	1.130	0.986	1.348	1.174
$(n, \theta) = (300, 0.25)$	0.548	0.547	0.654	0.652
$(n, \theta) = (300, 0.5)$	0.593	0.590	0.706	0.703
$(n, \theta) = (300, 0.75)$	0.634	0.632	0.755	0.752

TABLE 4: Empirical interval lengths for the proposed jackknife empirical likelihood confidence interval (JELCI) and the profile empirical likelihood confidence interval (PELCI) with nominal levels 0.9 and 0.95 for the two-dimensional t copula and marginal distributions t_7 and t_8.

	Level 0.9 JELCI	Level 0.9 PELCI	Level 0.95 JELCI	Level 0.95 PELCI
$(n, \theta) = (100, 0.25)$	0.956	0.941	1.140	1.121
$(n, \theta) = (100, 0.5)$	1.038	1.015	1.236	1.207
$(n, \theta) = (100, 0.75)$	1.118	0.976	1.334	1.164
$(n, \theta) = (300, 0.25)$	0.553	0.551	0.659	0.657
$(n, \theta) = (300, 0.5)$	0.593	0.590	0.707	0.703
$(n, \theta) = (300, 0.75)$	0.631	0.629	0.752	0.749

TABLE 5: Empirical coverage probabilities for the proposed jackknife empirical likelihood confidence interval (JELCI) and the profile empirical likelihood confidence interval (PELCI) with nominal levels 0.9 and 0.95 for the three-dimensional Gaussian copula and marginal distributions t_7, t_8, and t_9.

	Level 0.9 JELCI	Level 0.9 PELCI	Level 0.95 JELCI	Level 0.95 PELCI
$(n, \theta_1, \theta_2, \theta_3) = (100, 0.25, 0.5, 0.25)$	0.8216	0.7306	0.8799	0.7746
$(n, \theta_1, \theta_2, \theta_3) = (100, 0.5, 0.5, 0.5)$	0.8164	0.7363	0.8740	0.7794
$(n, \theta_1, \theta_2, \theta_3) = (100, 0.75, 0.5, 0.75)$	0.7908	0.7524	0.8497	0.7958
$(n, \theta_1, \theta_2, \theta_3) = (300, 0.25, 0.5, 0.25)$	0.8934	0.8907	0.9434	0.9384
$(n, \theta_1, \theta_2, \theta_3) = (300, 0.5, 0.5, 0.5)$	0.8893	0.8872	0.9429	0.9388
$(n, \theta_1, \theta_2, \theta_3) = (300, 0.75, 0.5, 0.75)$	0.8867	0.8859	0.9406	0.9390

show that the proposed jackknife empirical likelihood method performs better than the profile empirical likelihood.

In conclusion, the proposed jackknife empirical likelihood method provides good coverage accuracy and the computation is much simpler than the profile empirical likelihood method. Moreover, the package "emplik" in the software R is ready to use for the proposed method.

TABLE 6: Empirical coverage probabilities for the proposed jackknife empirical likelihood confidence interval (JELCI) and the profile empirical likelihood confidence interval (PELCI) with nominal levels 0.9 and 0.95 for the three-dimensional t copula and marginal distributions t_7, t_8, and t_9.

	Level 0.9 JELCI	Level 0.9 PELCI	Level 0.95 JELCI	Level 0.95 PELCI
$(n, \theta_1, \theta_2, \theta_3) = (100, 0.25, 0.5, 0.25)$	0.8160	0.7341	0.8733	0.7784
$(n, \theta_1, \theta_2, \theta_3) = (100, 0.5, 0.5, 0.5)$	0.8025	0.7348	0.8648	0.7835
$(n, \theta_1, \theta_2, \theta_3) = (100, 0.75, 0.5, 0.75)$	0.7860	0.7443	0.8475	0.7920
$(n, \theta_1, \theta_2, \theta_3) = (300, 0.25, 0.5, 0.25)$	0.8867	0.8819	0.9390	0.9341
$(n, \theta_1, \theta_2, \theta_3) = (300, 0.5, 0.5, 0.5)$	0.8858	0.8834	0.9372	0.9340
$(n, \theta_1, \theta_2, \theta_3) = (300, 0.75, 0.5, 0.75)$	0.8842	0.8857	0.9334	0.9352

4. CASE STUDIES

4.1. Testing the Drift Parameter in the Variance Gamma Model

The class of variance gamma (VG) distributions was introduced by Madan and Seneta (1987) as an alternative model for stock returns beyond the usual normal distribution assumption. It has so far been used extensively by financial economists especially in pricing financial derivatives, see Madan and Milne (1991) and Madan, Carr and Chang (1998) for applications, and Seneta (2007) for a historical account of the development.

The VG process Z_t is a time-changed Lévy process where the subordinator is a Gamma process. It is parameterized by three parameters: the drift parameter μ, the volatility parameter σ, and the subordinator variance parameter ν. More specifically, let S_t be a gamma process subordinator with a unit mean rate and a variance rate ν where $\nu > 0$. Let W_t be a standard Brownian motion. Then the VG process is defined as $dZ_t = \mu dS_t + \sigma dW_{S_t}$. That is, the calendar time t is now replaced with the time change S_t. Let $X \equiv Z_{t+\delta} - Z_t$ be the increment of Z_t with interval δ. The characteristic function of X is $Ee^{iuX} = e^{\Psi(u)\delta}$, where the characteristic exponent $\Psi(u)$ is given by

$$\Psi(u) = -\frac{1}{\nu} \log \left\{ 1 + \frac{u^2 \sigma^2 \nu}{2} - i\mu\nu u \right\}$$

Given a sample $\mathbf{X}$ of increments X with sample size n, we are interested in the hypothesis $H_0 : \mu = 0$. This amounts to asking whether it is sufficient to model the data in interest using a martingale process. For example, it is interesting to know whether one needs to introduce a drift parameter for the log change of US/Japan exchange rate process or not. The parameter of interest is thus μ, and the two nuisance parameters are (σ, ν). To employ the jackknife empirical likelihood estimation method in this article, we need the estimating equations. One approach would be to just use the mean equation alone. However, this ignores all higher moments which are the very reason why financial economists use VG process as an alternative to Brownian motion with drift. Therefore, below we use a different approach based on higher-order moments. The lowest-order moments or central moments would seem to be natural choices. However, the odd moments or central moments are all proportional to μ and thus all zero under the null $H_0 : \mu = 0$. Thus, we use only the even moments with orders 2, 4, and 6.

By differentiating the characteristic function of the VG process increment, the raw even moments $m_j(\mu, \nu, \sigma) := EX^j$ can be computed easily and estimating equations are obtained by equating the empirical moments to the raw moments. Then we apply the jackknife empirical likelihood method to the last equation by solving the first two equations for ν and σ.

We use four different financial time series, namely, the S&P 500 index (SPX), the CBOE volatility index (VIX), the effective federal funds rate (FFR), and the exchange rate between the British Pound and US Dollar (GBP/USD). These four time series are very important in four distinctive large financial sectors, that is, the equity market, the financial derivative market, the fixed-income market, and the foreign exchange market. The sample periods are January 1950 to October 2009 for the SPX, January 1990 to October 2009 for the VIX, July 1954 to September 2009 for the FFR, and January 1971 to July 2009 for the GBP/USD.

Plot of the SPX shows a few significant events, including the amazing run of the SPX before year 1987, the market crash in 1987, the internet boom in late 1990s, the IT bubble burst in 2001, and most recently the subprime mortgage crisis. Plot of the VIX has been largely stable, except for the abrupt surge during the recent market turmoil. The FFR has had quite a few surges, especially during the hyper-inflation period in the late 1970s and early 1980s, and has recently reached unprecedented low levels due to the Fed's effort to boost the economy. The most noticeable feature on the GBP/USD subplot is the free fall and fast appreciation of the British Pound before and after the Plaza Accord in 1985. The recent global financial crisis has also weakened the British Pound due to the flight to quality by international investors.

Our samples are the weekly log change processes of the above four financial time series. Plots of the samples confirm the common notion that financial data often deviate a lot from the normal distribution assumption, see for example, Blattberg & Gonedes (1974) or Kon (1984). Mandelbrot (1963) traces awareness of this non-normality as far back as year 1915. This nonnormality is especially true for the times series of the log change of the effective federal funds rate, where we see a huge kurtosis.

Table 7 reports the test statistic $LR(0)$ for the log changes of the four financial time series considered. The statistic is asymptotically χ_1^2 distributed, and the asymptotic P-values are reported. The skewness and kurtosis for each time series are also included in the table, as well as the result of a naive t-test. As we would expect, the statistic $LR(0)$ strongly rejects $H_0 : \mu = 0$ for the log changes of the S&P 500 index with a P-value of only 0.0002. However, the statistic cannot reject $H_0 : \mu = 0$ for the log changes of the CBOE volatility index, and the exchange rate between British Pound and US Dollar. These are the same conclusions one gets from the naive t-test assuming that the sample is drawn from a normal distribution with unknown variance. However, the test on the log changes of FFR is interesting. A naive t-test suggests that $\mu = 0$, while the test statistic $LR(0)$ rejects $\mu = 0$ at more than 99% level. This apparent discrepancy might be explained by the fact that we have used higher-order moments in the estimating equations and that the empirical distribution of log changes of FFR is very different from the normal distribution with a very large kurtosis. It also highlights the advantage of using empirical likelihood estimation over the naive t-test which assumes a normal distribution.

TABLE 7: Jackknife empirical likelihood ratio test for $H_0 : \mu = 0$ in the variance gamma model for financial time series. The samples are the weekly log changes of the SPX, VIX, FFR, and GBP/USD time series.

Log change of	n	Skewness	Kurtosis	t-Test	$LR(0)$	P-value
SPX	3059	−0.54	7.77	3.5451	14.3046	0.0002
VIX	1018	0.39	4.82	0.1005	1.1303	0.2877
FFR	2881	−0.32	55.83	−0.2819	7.8141	0.0052
GBP/USD	1987	−0.43	7.10	−0.6286	0.8917	0.3450

The skewness and kurtosis of the samples are also reported as well as the naive t-statistic assuming that the samples are drawn from a normal distribution with unknown variances.

4.2. Testing Whether a Normal Tempered Stable Process Is Normal Inverse Gaussian

A broader class of models that includes the VG process as a special case is the so-called *normal tempered stable* process. See Cont and Tankov (2004). This class of models is obtained by time changing an independent Brownian motion with drift by a tempered stable subordinator S_t. The Lévy measure of S_t with index A and parameter ν is given by

$$\rho_S(x) = \frac{1}{\Gamma(1-A)} \left(\frac{1-A}{\nu} \right)^{1-A} \frac{e^{-(1-A)x/\nu}}{x^{A+1}} 1_{x>0}$$

where $\nu > 0$ and $0 \leq A < 1$ are two constants. The increment X of the time-changed process Y_t again has a closed-form characteristic function with characteristic exponent as follows

$$\Psi(u) = \frac{1-A}{\nu A} \left\{ 1 - \left(1 + \frac{\nu(u^2\sigma^2/2 - i\mu u)}{1-A} \right)^A \right\}$$

Two important special cases are the VG process (where $A = 0$) and the normal inverse Gaussian process (where $A = 1/2$). See Barndorff-Nielsen (1997) and Rydberg (1997) for early studies on the normal inverse Gaussian process.

We are now interested in the null hypothesis $H_0 : A = 1/2$, that is, whether the data implies a normal inverse Gaussian process. Thus, the parameter of interest is A and the nuisance parameters are now (μ, ν, σ). To employ the jackknife empirical likelihood estimation method, we construct estimating equations from the raw moments $m_j(A, \mu, \nu, \sigma) := EX^j$. It turns out that if $\mu = 0$ in the normal tempered stable process, then the lowest four moments are not functions of A. Also, if $\mu = 0$, then all odd moments are zero. Thus, to reliably estimate A using moments when μ is small, we have to include a moment m_j with $j \geq 6$. We choose to use the moments m_1, m_2, m_4, and m_6 which can be calculated straightforwardly. Then the jackknife empirical likelihood method is applied to the last equation by using the first three equations to compute $\tilde{\mu}(A; \mathbf{X})$, $\tilde{\nu}(A; \mathbf{X})$, and $\tilde{\sigma}(A; \mathbf{X})$.

We apply the above procedure to the same four time series as before. The results are shown in Table 8. As we see, the test statistic $LR(1/2)$ cannot reject $H_0 : A = 1/2$ for the log changes of the S&P 500 index, the CBOE VIX index, and the exchange rate between the British Pound and US Dollar. However, it strongly rejects $A = 1/2$ for the log change time series of the effective federal funds rate.

TABLE 8: Jackknife empirical likelihood ratio test for $H_0 : A = 1/2$ in a normal tempered stable process for financial time series.

Log change of	n	$LR(1/2)$	P-value
SPX	3059	1.0944	0.2955
VIX	1018	1.2903	0.2560
FFR	2881	10.1324	0.0015
GBP/USD	1987	0.2629	0.6082

The weekly log changes of the SPX, VIX, FFR, and GBP/USD are studied.

ACKNOWLEDGEMENTS

APPENDIX: OUTLINE OF PROOFS

Proof of Proposition 1. It follows from Taylor expansions.

Proof of Proposition 2. Using Proposition 1, we can show that

$$\max_{1 \leq i \leq n} |Y_{i,l}(\alpha)| = o_p(n^{1/2}) \tag{13}$$

uniformly in $\|\alpha - \alpha_0\| \leq n^{-1/3}$ for $l = 1, \ldots, s-q_2$

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\alpha_0) \xrightarrow{d} (W_1, \ldots, W_{s-q_2})^{\mathrm{T}} \tag{14}$$

and

$$\frac{1}{n} \sum_{i=1}^{n} Y_i(\alpha_0) Y_i^{\mathrm{T}}(\alpha_0) \xrightarrow{p} (E(W_l W_k))_{1 \leq k, l \leq s-q_2} \tag{15}$$

where for $l = 1, \ldots, s-q_2$,

$$W_l = Z_l - E\left\{ \frac{\partial}{\partial \beta} g_l(X_1; \alpha_0, \beta_0) \right\} \Sigma_1^{-1} (Z_{s-q_2+1}, \ldots, Z_s)^{\mathrm{T}}$$

and $(Z_1, \ldots, Z_s)^{\mathrm{T}} \sim N(0, \Sigma)$. Hence, the proposition follows from (13) to (15) and similar arguments in Owen (1990).

Proof of Theorem 1. By Proposition 1, we can show that

$$\frac{1}{n} \sum_{i=1}^{n} \frac{\partial}{\partial \alpha} Y_i(\alpha_0) \xrightarrow{p} \Sigma_3$$

Put $V = \{\Sigma_3^T (\Sigma^*)^{-1} \Sigma_3\}^{-1}$. Similar to the proof of Theorem 1 of Qin and Lawless (1994), we can show that

$$\sqrt{n}\{\hat{\alpha} - \alpha_0\} = -V \Sigma_3^T (\Sigma^*)^{-1} \sqrt{n} Q_{1n}(\alpha_0, 0) + o_p(1)$$

and

$$\sqrt{n}\hat{\lambda} = (\Sigma^*)^{-1} \{I - \Sigma_3 V \Sigma_3^{\mathrm{T}} (\Sigma^*)^{-1}\} \sqrt{n} Q_{1n}(\alpha_0, 0) + o_p(1)$$

So,

$$l^J(\hat{\alpha}) = \frac{n}{2} Q_{1n}^{\mathrm{T}}(\alpha_0, 0)(\Sigma^*)^{-1} \{I - \Sigma_3 V \Sigma_3^{\mathrm{T}} (\Sigma^*)^{-1}\} Q_{1n}(\alpha_0, 0) + o_p(1)$$

Similarly, we can show that $l^J(\alpha_0) = (n/2)Q_{1n}^{\mathrm{T}}(\alpha_0, 0)(\Sigma^*)^{-1}Q_{1n}(\alpha_0, 0) + o_p(1)$. Therefore,

$$
LR(\alpha_0) = \left\{ (\Sigma^*)^{-1/2} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} Y_i(\alpha_0) \right\}^T \{ (\Sigma^*)^{-1/2} \Sigma_3 V \Sigma_3^{\mathrm{T}} (\Sigma^*)^{-1/2} \} \left\{ (\Sigma^*)^{-1/2} \frac{1}{n} \sum_{i=1}^{n} Y_i(\alpha_0) \right\}
$$
$$
+ o_p(1) \xrightarrow{d} \chi^2_{q_1}
$$

BIBLIOGRAPHY

O. E. Barndorff-Nielsen (1997). Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2, 41–68.

R. C. Blattberg & N. J. Gonedes (1974). A comparison of the stable and student distributions as statistical models for stock prices. *Journal of Business*, 47, 244–280.

R. Cao & I. Van Keilegom (2006). Empirical likelihood tests for two-sample problems via nonparametric density estimation. *The Canadian Journal of Statistics*, 34, 61–77.

N. H. Chan, L. Peng & Y. Qi (2006). Quantile inference for near-integrated autoregressive time series with infinite variance. *Statistica Sinica*, 16, 15–28.

S. Chen & H. Cui (2007). On the second order properties of empirical likelihood with moment restrictions. *Journal of Econometrics*, 141, 492–516.

S. Chen & J. Gao (2007). An adaptive empirical likelihood test for parametric time series regression models. *Journal of Econometrics*, 141, 950–972.

S. Chen & P. Hall (1993). Smoothed empirical likelihood confidence intervals for quantiles. *The Annals of Statistics*, 21, 1166–1181.

S. Chen & I. Van Keilegom (2009). A review on empirical likelihood methods for regression. *Test*, 18, 415–447.

J. Chen, L. Peng & Y. Zhao (2009). Empirical likelihood based confidence intervals for copulas. *Journal of Multivariate Analysis*, 100, 137–151.

S. X. Chen, L. Peng & Y. Qin (2009). Empirical likelihood methods for high dimension. *Biometrika*, 96, 711–722.

G. Claeskens, B. Y. Jing, L. Peng & W. Zhou (2003). Empirical likelihood confidence regions for comparison distributions and ROC curves. *The Canadian Journal of Statistics*, 31, 173–190.

R. Cont & P. Tankov (2004). "*Financial Modeling with Jump Processes,*" Chapman & Hall/CRC, Boca Raton.

P. Guggenberger & R. J. Smith (2008). Generalized empirical likelihood tests in time series models with potential identification failure. *Journal of Econometrics*, 142, 134–161.

P. Hall & B. La Scala (1990). Methodology and algorithms of empirical likelihood. *International Statistical Review*, 58, 109–127.

P. Hall & Q. Yao (2003). Data tilting for time series. *Journal of the Royal Statistical Society Series B*, 65, 425–442.

N. L. Hjort, I. W. McKeague & I. Van Keilegom (2009). Extending the scope of empirical likelihood. *The Annals of Statistics*, 37, 10179–11115.

B. Y. Jing, J. Q. Yuan & W. Zhou (2009). Jackknife empirical likelihood. *Journal of the American Statistical Association*, 104, 1224–1232.

A. Keziou & S. Leoni-Aubin (2008). On empirical likelihood for semiparametric two-sample density ratio models. *Journal of Statistical Planning and Inference*, 138, 915–928.

S. J. Kon (1984). Models of stock returns—A comparison. *Journal of Finance*, 39, 147–165.

J. Lu & L. Peng (2002). Likelihood based confidence intervals for the tail index. *Extremes*, 5, 337–352.

X. Lu & Y. Qi (2004). Empirical likelihood for the additive risk model. *Probability and Mathematical Statistics*, 24, 419–431.

D. B. Madan & F. Milne (1991). Option pricing with VG martingale components. *Mathematical Finance*, 1(4), 39–55.

D. B. Madan & E. Seneta (1987). Chebyshev polynomial approximations and characteristic function estimation. *Journal of the Royal Statistical Society Series B*, 49, 163–169.

D. B. Madan, P. P. Carr & E. C. Chang (1998). The variance Gamma process and option pricing. *European Finance Review*, 2, 79–105.

B. Mandelbrot (1963). The variation of certain speculative prices. *Journal of Business*, 36(4), 394–419.

D. J. Nordman & S. N. Lahiri (2006). A frequency domain empirical likelihood for short- and long-range dependence. *The Annals of Statistics*, 34, 3019–3050.

D. J. Nordman, P. Sibbertsen & S. N. Lahiri (2007). Empirical likelihood confidence intervals for the mean of a long-range dependent process. *Journal of Time Series Analysis*, 28, 576–599.

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

A. Owen (1990). Empirical likelihood ratio confidence regions. *The Annals of Statistics*, 18, 90–120.

L. Peng (2004). Empirical likelihood confidence interval for a mean with a heavy tailed distribution. *The Annals of Statistics*, 32, 1192–1214.

L. Peng & Y. Qi (2006a). A new calibration method of constructing empirical likelihood-based confidence intervals for the tail index. *Australian & New Zealand Journal of Statistics*, 48, 59–66.

L. Peng & Y. Qi (2006b). Confidence regions for high quantiles of a heavy tailed distribution. *The Annals of Statistics*, 34, 1964–1986.

J. Qin & J. F. Lawless (1994). Empirical likelihood and general estimating equations. *The Annals of Statistics*, 22, 300–325.

J. Ren (2008). Weighted empirical likelihood in some two-sample semiparametric models with various types of censored data. *The Annals of Statistics*, 36, 147–166.

T. H. Rydberg (1997). The normal inverse Gaussian Levy process: Simulation and approximation. *Communications in Statistics: Stochastic Models*, 13, 887–910.

E. Seneta (2007). The Early Years of the Variance-Gamma Process. In: "*Advances in Mathematical Finance,*" M. C. Fu, R. A. Jarrow, J.-Y. J. Yen, and R. J. Elliott, editors. Birkhauser, Boston, pp. 3–19.

J. Shao & D. Tu (1995). "*The Jackknife and Bootstrap,*" Springer, New York.

J. Shen & S. He (2007). Empirical likelihood for the difference of quantiles under censorship. *Statistical Papers*, 48, 437–457.

J. W. Tukey (1958). Bias and confidence in not-quite large samples. *The Annals of Statistics*, 29, 614.

W. Zhou & B. Y. Jing (2003). Smoothed empirical likelihood confidence intervals for the difference of quantiles. *Statistica Sinica*, 13, 83–95.

Y. Zhou & H. Liang (2005). Empirical-likelihood-based semiparametric inference for the treatment effect in the two-sample problem with censoring. *Biometrika*, 92, 271–282.