Journal of Multivariate Analysis 112 (2012) 63–75

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Jackknife empirical likelihood tests for error distributions in regression models

Huijun Feng, Liang Peng*

School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332-0160, United States

ARTICLE INFO

Article history:
Received 4 January 2012
Available online 7 June 2012

AMS 2011 subject classification:
primary 62E20

Keywords:
Goodness-of-fit test
Jackknife empirical Likelihood method
Regression model

ABSTRACT

Regression models are commonly used to model the relationship between responses and covariates. For testing the error distribution, some classical test statistics such as Kolmogorov–Smirnov test and Cramér–von-Mises test suffer from the complicated limiting distribution due to the plug-in estimate for the unknown parameters. Hence some ad hoc procedure such as bootstrap method is needed to obtain critical points. Recently, Khmaladze and Koul (2004) [7] have proposed an asymptotically distribution free test via some Martingale transforms. However, the calculation of such a test becomes quite involved, which usually requires numeric integration when the Cramér–von-Mises type of test is employed. In this paper we propose a novel jackknife empirical likelihood method which is easy to compute and has a chi-square limit so that critical values are ready at hand. A simulation study confirms that the new test has an accurate size and is powerful too.

1. Introduction

Let Y and X denote a univariate response and a d-variate covariate, respectively. For modeling the relationship between Y and X, a widely employed tool is the regression model $Y = m(X; \alpha) + \epsilon$, where m is a known function depending on a q-dimensional unknown parameter α and ϵ is a random error with mean zero. Suppose $\{(X_i^T, Y_i)^T\}_{i=1}^n$ is a random sample from this regression model, i.e.,

$$Y_i = m(X_i; \alpha) + \epsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

where $\epsilon_1, \ldots, \epsilon_n$ are independent and identically distributed random variables with zero mean, $X_1, \ldots, X_n$ are independent and identically distributed random variables and independent of $\epsilon_i's$. A standard way to estimate the unknown parameter α is the least squares estimate

$$\hat{\alpha} = \arg\min_{\alpha} \sum_{i=1}^{n} \{Y_i - m(X_i; \alpha)\}^2,$$

which says that $\hat{\alpha}$ is a solution of the following score equations

$$\sum_{i=1}^{n} \{Y_i - m(X_i; \alpha)\} \frac{\partial}{\partial \alpha} m(X_i; \alpha) = 0. \tag{2}$$

In some applications such as predicting conditional Value-at-Risk in risk management, it is useful to fit a parametric distribution family to the random error ϵ_i so as to improve the accuracy of inference. This results in a corresponding

* Corresponding author.
E-mail addresses: hfeng9@math.gatech.edu (H. Feng), peng@math.gatech.edu (L. Peng).

0047-259X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2012.05.018

parametric distribution family for the conditional distribution of Y_i given X_i. See [4] for more details on regression models. A discussion on using nonlinear regression models for price decisions is given in [2]. A test for association between covariates and response is provided by [1]. Here we are interested in testing whether the distribution of ϵ_i follows from a particular parametric family, i.e., test $H_0 : F_\epsilon \in \mathcal{F} = \{F(\cdot;\beta) : \beta \in \Omega \subset R^s\}$, where F_ϵ denotes the distribution of ϵ_i, and $F(\cdot;\beta)$ denotes a distribution function depending on the parameter β. Obviously one can simply employ some classical goodness-of-fit tests to the estimated errors $\hat\epsilon_i = Y_i - m(X_i;\hat\alpha)$, $i = 1,\ldots,n$. More specifically, one can estimate β first by using the maximum likelihood estimate $\hat\beta$ based on $\hat\epsilon_1,\ldots,\hat\epsilon_n$, and then consider either the Kolmogorov–Smirnov test

$$T_1 = \sup_z \sqrt{n}\left|\frac{1}{n}\sum_{i=1}^n I(\hat\epsilon_i \le z) - F(z;\hat\beta)\right|$$

or the Cramér–von-Mises test

$$\begin{aligned} T_2 &= n\int_{-\infty}^{\infty}\left\{\frac{1}{n}\sum_{i=1}^n I(\hat\epsilon_i \le z) - F(z;\hat\beta)\right\}^2 dF(z;\hat\beta) \\ &= \frac{1}{12n} + \sum_{i=1}^n \left\{\frac{2i-1}{2n} - F(\hat\epsilon_{n,i};\hat\beta)\right\}^2, \end{aligned}$$

where $\hat\epsilon_{n,1} \le \cdots \le \hat\epsilon_{n,n}$ denote the order statistics of $\hat\epsilon_1,\ldots,\hat\epsilon_n$ (see [3]). Due to the plug-in estimators $\hat\alpha$ in $\hat\epsilon_i's$ and $\hat\beta$, the limiting distributions of T_1 and T_2 become quite complicated, which depend on the underlying distribution and thus are no longer distribution free. Hence some ad hoc procedure such as bootstrap method is needed in order to calculate the critical values.

Recently, Khmaladze and Koul [7] have proposed a new goodness-of-fit test via martingale transforms for testing the error distribution. It turns out that the new test statistic is asymptotically distribution free in testing a simple null hypothesis or a composite null hypothesis with a scale distribution family, and hence critical values can be tabulated. Some numeric analyses are given in [7] and [8] for the Kolmogorov–Smirnov type of test. However, when the Cramér–von-Mises type of test is concerned, the calculation of the proposed test in [7] becomes quite complicated, which requires to evaluate some integrals numerically.

In this paper, we propose a novel jackknife empirical likelihood test for testing the error distribution in the regression model (1). It turns out that the asymptotic distribution of the new test has a chi-square limit, and the calculation of the test statistic is quite straightforward and involves no numeric integration. As a powerful tool in interval estimation and hypothesis test, the empirical likelihood method has been applied to many different settings. We refer to [11] for an overview. Some advantages of the empirical likelihood method include that the shape of confidence interval/region is determined by the sample automatically. When the empirical likelihood method is applied to nonlinear functionals directly, the Wilks theorem fails in general, i.e., the limit is no longer a chi-square distribution. To overcome this difficulty, [6] proposed to apply the empirical likelihood method to some jackknife sample constructed from the targeted nonlinear functionals. This is called the jackknife empirical likelihood method. A smoothed jackknife empirical likelihood method is applied to copulas, tail copulas and ROC curves; see [5,13,14].

We organize this paper as follows. Methodology and main asymptotic results are given in Section 2. Section 3 presents a simulation study. Some conclusions are given in Section 4. Proofs are put in Appendix.

2. Methodology

To motivate our new method, we assume $\epsilon_i's$ are observable and β is known for the time being. That is, we want to test $H_0 : F_\epsilon(x) \equiv F(x;\beta)$. This is equivalent to test $H_0 : \int_{-\infty}^{\infty}\{F_\epsilon(x) - F(x;\beta)\}^2 dF(x;\beta) = 0$, which results in the Cramér–von-Mises test when $F_\epsilon(x)$ is replaced by the empirical distribution function based on $\epsilon_1,\ldots,\epsilon_n$.

By noting that $H_0 : \int_{-\infty}^{\infty}\{F_\epsilon(x) - F(x;\beta)\}^2\, dF(x;\beta) = 0$ is equivalent to

$$H_0 : E\int_{-\infty}^{\infty}\{I(\epsilon_1 \vee \epsilon_2 \le x) - 2I(\epsilon_1 \le x)F(x;\beta) + F^2(x;\beta)\}\, dF(x;\beta) = 0,$$

i.e.,

$$H_0 : E\{F^2(\epsilon_1;\beta) - F(\epsilon_1 \vee \epsilon_2;\beta) + 1/3\} = 0,$$

one can directly apply the empirical likelihood method to the above estimating equation based on sample $\{(\epsilon_i, \epsilon_{i+k})^T\}_{i=1}^k$, where $k = [n/2]$. More specifically, by defining the empirical likelihood function as

$$\begin{aligned} L(\beta) = \sup\Bigg\{ &\prod_{i=1}^k (kp_i) : p_1 \ge 0,\ldots,p_k \ge 0, \sum_{i=1}^k p_i = 1, \\ &\sum_{i=1}^k p_i \left(\frac{F^2(\epsilon_i;\beta) + F^2(\epsilon_{i+k};\beta)}{2} - F(\epsilon_i \vee \epsilon_{i+k};\beta) + 1/3\right) = 0\Bigg\}, \end{aligned}$$

it follows from [10] that $-2\log L(\beta)$ converges in distribution to a chi-square limit with one degree of freedom under H_0. Hence, one can use the empirical likelihood ratio test statistic $-2\log L(\beta)$ to test $H_0 : F_\epsilon(x) \equiv F(x;\beta)$. Unfortunately, this test is not powerful since

$$E\int_{-\infty}^{\infty}\{I(\epsilon_1\vee\epsilon_2\le x)-2I(\epsilon_1\le x)F(x;\beta)+F^2(x;\beta)\}\,dF(x;\beta)=O(\delta^2)$$

rather than $O(\delta)$ when $\sup_x|F_\epsilon(x)-F(x;\beta)|=O(\delta)$. To overcome this difficulty, we propose to apply the empirical likelihood method to the following two equations:

$$\begin{cases}E\{F^2(\epsilon_1;\beta)-F(\epsilon_1\vee\epsilon_2;\beta)+1/3\}=0\\ EF(\epsilon_1;\beta)-2EF^3(\epsilon_1;\beta)=0.\end{cases} \tag{3}$$

Note that [9] proposed to employ different estimating equations when $\epsilon_i's$ are observable and β is either known or unknown. We remark that the second equation in (3) can be replaced by some other linear estimating equations. Hence this new method is quite flexible and easy in taking more relevant constraints into account.

Now we are ready to extend the above idea to test the error distribution in the regression model (1). We consider the cases of simple null hypothesis and composite null hypothesis separately. Throughout we assume that α_0 and β_0 denote the true values of α and β respectively.

2.1. A simple null hypothesis

In this subsection, we are interested in testing $H_0 : F_\epsilon(x)\equiv F(x;\beta_0)$ under model (1).

Put $k=[\frac{n}{2}]$ and define $\epsilon_i(\alpha)=Y_i-m(X_i;\alpha)$, $\tilde\epsilon_i(\alpha)=Y_{k+i}-m(X_{k+i};\alpha)$, $\epsilon_i^*(\alpha)=\max(\epsilon_i(\alpha),\tilde\epsilon_i(\alpha))$ and $h_i(\alpha)=\frac{\partial}{\partial\alpha}\{\epsilon_i^2(\alpha)+\tilde\epsilon_i^2(\alpha)\}$ for $i=1,\ldots,k$. Therefore the least squares estimator $\hat\alpha$ of α is defined as a solution to the equation $\sum_{i=1}^k h_i(\alpha)=0$.

Unfortunately we cannot directly apply the empirical likelihood method to Eq. (3) based on the sample $\{(\epsilon_i(\hat\alpha),\tilde\epsilon_i(\hat\alpha))\}_{i=1}^k$ since this fails to catch the variance of $\hat\alpha$. Generally speaking, the Wilks theorem does not hold when the empirical likelihood method is applied to nonlinear functionals directly. Motivated by the recent jackknife empirical likelihood method in [6], we propose to apply the empirical likelihood method to some jackknife pseudosample. In order to formulate the jackknife sample, it follows from the idea in [6] that one has to estimate α by deleting one observation each time, that is, to solve the equation $\sum_{i=1,i\ne j}^k h_i(\alpha)=0$ for each $j=1,\ldots,k$. When m is a nonlinear function, the above equation does not admit an explicit solution in general. Therefore, the above way of formulating jackknife sample is computationally intensive. Here we propose to apply the approximate jackknife empirical likelihood method in [12] as follows.

Note that

$$\begin{aligned}0&=\sum_{j=1,j\ne i}^k h_j(\alpha)\\&=\sum_{j=1,j\ne i}^k h_j(\alpha)-\sum_{j=1}^k h_j(\hat\alpha)\\&=\sum_{j=1}^k\{h_j(\alpha)-h_j(\hat\alpha)\}-h_i(\alpha)\\&\approx\sum_{j=1}^k\left\{\frac{\partial}{\partial\alpha^T}h_j(\hat\alpha)\right\}\{\alpha-\hat\alpha\}-h_i(\hat\alpha).\end{aligned} \tag{4}$$

Instead of solving $0=\sum_{j=1,j\ne i}^k h_j(\alpha)$, we propose to approximate the solution by

$$\hat\alpha_i=\hat\alpha+\left\{\frac{1}{k}\sum_{j=1}^k\frac{\partial}{\partial\alpha^T}h_j(\hat\alpha)\right\}^{-1}\frac{1}{k}h_i(\hat\alpha).$$

Using Eq. (3), $\hat\alpha$ and $\hat\alpha_i's$, we define the approximate jackknife sample as

$$\begin{aligned}G_1(i)=&\sum_{j=1}^k\left\{\frac{F^2(\epsilon_j(\hat\alpha);\beta_0)+F^2(\tilde\epsilon_j(\hat\alpha);\beta_0)}{2}-F(\epsilon_j^*(\hat\alpha);\beta_0)+1/3\right\}\\&-\sum_{j=1,j\ne i}^k\left\{\frac{F^2(\epsilon_j(\hat\alpha_i);\beta_0)+F^2(\tilde\epsilon_j(\hat\alpha_i);\beta_0)}{2}-F(\epsilon_j^*(\hat\alpha_i);\beta)+1/3\right\}\end{aligned}$$

and

$$G_2(i) = \sum_{j=1}^{k}\{F(\epsilon_j(\hat{\alpha}); \beta_0) + F(\tilde{\epsilon}_j(\hat{\alpha}); \beta_0)\} - \sum_{j=1, j\neq i}^{k}\{F(\epsilon_j(\hat{\alpha}_i); \beta_0) + F(\tilde{\epsilon}_j(\hat{\alpha}_i); \beta_0)\}$$
$$- 2\sum_{j=1}^{k}\{F^3(\epsilon_j(\hat{\alpha}); \beta_0) + F^3(\tilde{\epsilon}_j(\hat{\alpha}); \beta_0)\} + 2\sum_{j=1, j\neq i}^{k}\{F^3(\epsilon_j(\hat{\alpha}_i); \beta_0) + F^3(\tilde{\epsilon}_j(\hat{\alpha}_i); \beta_0)\}$$

for $i = 1, \ldots, k$. Based on the above approximate jackknife sample, we define the jackknife empirical likelihood function as

$$L_n^J = \sup\left\{\prod_{i=1}^{k}(kp_i) : p_1 \geq 0, \ldots, p_k \geq 0, \sum_{i=1}^{k} p_i = 1, \sum_{i=1}^{k} p_i G(i) = 0\right\}$$

where $G(i) = (G_1(i), G_2(i))^T$. By the Lagrange multiplier technique, we have

$$l_n^J := -2\log L_n^J = 2\sum_{i=1}^{k}\log\{1 + \lambda^T G(i)\},$$

where λ satisfies

$$\sum_{i=1}^{k}\frac{G(i)}{1 + \lambda^T G(i)} = 0. \tag{5}$$

Before proving that the Wilks theorem holds for the above jackknife empirical likelihood test, we list some regularity conditions:

- (A1) there are a neighborhood of α_0, say Ω_0 and a function $K(x)$ such that $EK(X_1) < \infty$ with X_1 given in model (1) and
$$\sup_{\alpha\in\Omega_0}\left\{\left|\frac{\partial}{\partial\alpha_i}m(x; \alpha)\right| + \left|\frac{\partial^2}{\partial\alpha_i\partial\alpha_j}m(x; \alpha)\right| + \left|\frac{\partial^3}{\partial\alpha_i\partial\alpha_j\partial\alpha_l}m(x; \alpha)\right|\right\} \leq K(x)$$
for $1 \leq i, j, l \leq q$;
- (A2) $E\frac{\partial}{\partial\alpha^T}h_1(\alpha_0)$ is invertible;
- (A3) $\sup_{y\in\Omega_1}|F''(y; \beta_0)| < \infty$, where Ω_1 denotes the support of ϵ_1.

Theorem 1. *Suppose model* (1) *holds with* $E\epsilon_i = 0$ *and* $E\epsilon_i^{2+\delta_0} < \infty$ *for some* $\delta_0 > 0$. *Further assume conditions* (A1)–(A3) *hold. Then, under* $H_0 : F_\epsilon(x) \equiv F(x; \beta_0)$, *we have* $l_n^J \xrightarrow{d} \chi^2(2)$ *as* $n \to \infty$.

Using Theorem 1, a jackknife empirical likelihood test for testing $H_0 : F_\epsilon(x) \equiv F(x; \beta_0)$ against $H_a : F_\epsilon(x) \not\equiv F(x; \beta_0)$ can be constructed, which rejects H_0 when $l_n^J \geq \chi^2_{2,1-\gamma}$, where γ is the significance level and $\chi^2_{2,1-\gamma}$ denotes the $(1-\gamma)$-th quantile of a chi-square distribution with two degrees of freedom.

2.2. A composite null hypothesis

In this subsection, we are interested in testing $H_0 : F_\epsilon \in \mathcal{F} = \{F(\cdot; \beta) : \beta \in \Omega \subset \mathcal{R}^s\}$ against $H_a : F_\epsilon \notin \mathcal{F}$ under model (1). Define

$$\bar{h}_i(\alpha, \beta) = \frac{\partial}{\partial\beta}\log f(\epsilon_i(\alpha); \beta) + \frac{\partial}{\partial\beta}\log f(\tilde{\epsilon}_i(\alpha); \beta)$$

for $i = 1, \ldots, k$, where $f(x; \beta) = \frac{\partial}{\partial x}F(x; \beta)$. Next we estimate β by solving the score equation $\sum_{i=1}^{k}\bar{h}_i(\hat{\alpha}, \beta) = 0$, and denote the solution by $\hat{\beta}$. Although one may prefer to estimate α and β simultaneously by solving the equations

$$\sum_{i=1}^{k}\frac{\partial}{\partial\alpha}\{\log f(\epsilon_i(\alpha); \beta) + \log f(\tilde{\epsilon}_i(\alpha); \beta)\} = 0 \quad \text{and} \quad \sum_{i=1}^{k}\bar{h}_i(\alpha, \beta) = 0,$$

we propose to estimate them separately, which has less computation in general. In order to formulate the jackknife sample, one needs to solve $\sum_{i=1,i\neq j}^{k}\bar{h}_i(\hat{\alpha}_j,\beta)=0$ for each $j=1,\ldots,k$. Like (4), we have

$$\begin{aligned}0 &= \sum_{i=1,i\neq j}^{k}\bar{h}_i(\hat{\alpha}_j,\beta)\\ &= \sum_{i=1,i\neq j}^{k}\bar{h}_i(\hat{\alpha}_j,\beta)-\sum_{i=1}^{k}\bar{h}_i(\hat{\alpha},\hat{\beta})\\ &= \sum_{i=1,i\neq j}^{k}\bar{h}_i(\hat{\alpha}_j,\beta)-\sum_{i=1}^{k}\bar{h}_i(\hat{\alpha}_j,\hat{\beta})+\sum_{i=1}^{k}\bar{h}_i(\hat{\alpha}_j,\hat{\beta})-\sum_{i=1}^{k}\bar{h}_i(\hat{\alpha},\hat{\beta})\\ &= \sum_{i=1}^{k}(\bar{h}_i(\hat{\alpha}_j,\beta)-\bar{h}_i(\hat{\alpha}_j,\hat{\beta}))+\sum_{i=1}^{k}(\bar{h}_i(\hat{\alpha}_j,\hat{\beta})-\bar{h}_i(\hat{\alpha},\hat{\beta}))-\bar{h}_j(\hat{\alpha}_j,\beta)\\ &\approx \sum_{i=1}^{k}\left\{\frac{\partial}{\partial\beta}\bar{h}_i(\hat{\alpha}_j,\hat{\beta})\right\}(\beta-\hat{\beta})+\sum_{i=1}^{k}\left\{\frac{\partial}{\partial\alpha}\bar{h}_i(\hat{\alpha},\hat{\beta})\right\}(\hat{\alpha}_j-\hat{\alpha})-\bar{h}_j(\hat{\alpha},\hat{\beta}).\end{aligned}$$

Thus, instead of solving $\sum_{i=1,i\neq j}^{k}\bar{h}_i(\hat{\alpha}_j,\beta)=0$, we propose to approximate the solution by

$$\hat{\beta}_j=\hat{\beta}+\left\{\frac{1}{k}\sum_{i=1}^{k}\frac{\partial}{\partial\beta^T}\bar{h}_i(\hat{\alpha},\hat{\beta})\right\}^{-1}\frac{1}{k}\bar{h}_j(\hat{\alpha},\hat{\beta})-\left\{\frac{1}{k}\sum_{i=1}^{k}\frac{\partial}{\partial\beta^T}\bar{h}_i(\hat{\alpha},\hat{\beta})\right\}^{-1}\left\{\frac{1}{k}\sum_{i=1}^{k}\frac{\partial}{\partial\alpha^T}\bar{h}_i(\hat{\alpha},\hat{\beta})\right\}(\hat{\alpha}_j-\hat{\alpha})$$

for $j=1,\ldots,k$.

Based on $\hat{\alpha}$, $\hat{\alpha}_i$, $\hat{\beta}$, $\hat{\beta}_i$ and (3), we formulate the approximate jackknife sample as

$$\begin{aligned}\bar{G}_1(i) &= \sum_{j=1}^{k}\left\{\frac{F^2(\epsilon_j(\hat{\alpha});\hat{\beta})+F^2(\tilde{\epsilon}_j(\hat{\alpha});\hat{\beta})}{2}-F(\epsilon_j^*(\hat{\alpha});\hat{\beta})+1/3\right\}\\ &\quad-\sum_{j=1,j\neq i}^{k}\left\{\frac{F^2(\epsilon_j(\hat{\alpha}_i);\hat{\beta}_i)+F^2(\tilde{\epsilon}_j(\hat{\alpha}_i);\hat{\beta}_i)}{2}-F(\epsilon_j^*(\hat{\alpha}_i);\hat{\beta}_i)+1/3\right\}\end{aligned}$$

and

$$\begin{aligned}\bar{G}_2(i) &= \sum_{j=1}^{k}\{F(\epsilon_j(\hat{\alpha});\hat{\beta})+F(\tilde{\epsilon}_j(\hat{\alpha});\hat{\beta})\}-\sum_{j=1,j\neq i}^{k}\{F(\epsilon_j(\hat{\alpha}_i);\hat{\beta}_i)+F(\tilde{\epsilon}_j(\hat{\alpha}_i);\hat{\beta}_i)\}\\ &\quad-2\sum_{j=1}^{k}\{F^3(\epsilon_j(\hat{\alpha});\hat{\beta})+F^3(\tilde{\epsilon}_j(\hat{\alpha});\hat{\beta})\}+2\sum_{j=1,j\neq i}^{k}\{F^3(\epsilon_j(\hat{\alpha}_i);\hat{\beta}_i)+F^3(\tilde{\epsilon}_j(\hat{\alpha}_i);\hat{\beta}_i)\}\end{aligned}$$

for $i=1,\ldots,k$. Based on the above approximate jackknife sample, the jackknife empirical likelihood function is defined as

$$\bar{L}_n^J=\sup\left\{\prod_{i=1}^{k}(kp_i):p_1\geq 0,\ldots,p_k\geq 0,\sum_{i=1}^{k}p_i=1,\sum_{i=1}^{k}p_i\bar{G}(i)=0\right\}\qquad(6)$$

where $\bar{G}(i)=(\bar{G}_1(i),\bar{G}_2(i))^T$. By the Lagrange multiplier technique, we have

$$\bar{l}_n^J:=-2\log\bar{L}_n^J=2\sum_{i=1}^{k}\log\{1+\bar{\lambda}^T\bar{G}(i)\},$$

where $\bar{\lambda}$ satisfies

$$\sum_{i=1}^{k}\frac{\bar{G}(i)}{1+\bar{\lambda}^T\bar{G}(i)}=0.\qquad(7)$$

Before showing that the Wilks theorem holds for the above jackknife empirical likelihood method, we list some regularity conditions:

- (A4) there are a neighborhood of β_0, say Ω_2, and a function $\bar{K}(\cdot)$ such that $E\bar{K}(\epsilon_1(\alpha_0), \tilde{\epsilon}_1(\alpha_0), X_1, X_{k+1}) < \infty$ and
$$\sup_{\alpha\in\Omega_0,\beta\in\Omega_2}\left\{\left|\frac{\partial}{\partial\theta_i}\bar{h}_1(\alpha,\beta)\right|+\left|\frac{\partial^2}{\partial\theta_i\partial\theta_j}\bar{h}_1(\alpha,\beta)\right|+\left|\frac{\partial^3}{\partial\theta_i\partial\theta_j\partial\theta_l}\bar{h}_1(\alpha,\beta)\right|\right\}\le \bar{K}(\epsilon_1(\alpha_0), \tilde{\epsilon}_1(\alpha_0), X_1, X_{k+1}),$$
where $\theta = (\alpha^T, \beta^T)^T$ and $1 \le i, j, l \le q + s$;
- (A5) $E\frac{\partial}{\partial\beta^T}\bar{h}_1(\alpha_0, \beta_0)$ is invertible;
- (A6) $\sup_{y\in\Omega_3}\sup_{\beta\in\Omega_2}|\frac{\partial^2}{\partial\bar{\theta}^2}F(y;\beta)| < \infty$, where $\bar{\theta} = (y, \beta^T)^T$ and Ω_3 denotes the support of ϵ_i which is independent of β.

Theorem 2. *Suppose model* (1) *holds with* $E\epsilon_i = 0$ *and* $E\epsilon_i^{2+\delta_0} < \infty$ *for some* $\delta_0 > 0$*. Further assume* (A1)–(A2) *and* (A4)–(A6) *hold. Then* $\bar{l}_n^J \xrightarrow{d} \chi^2(2)$ *as* $n \to \infty$.

As before, Theorem 2 can be employed to test $H_0 : F_\epsilon \in \mathcal{F}$ against $H_a : F_\epsilon \notin \mathcal{F}$.

Remark 1. Theorems 1 and 2 still hold when estimators for α and β are replaced by solving some other estimating equations.

3. A simulation study

In this section, we investigate the finite sample behavior of the proposed jackknife empirical likelihood test and compare it with the Cramér–von-Mises test. Since the test in [7] is hard to implement for the type of Cramér–von-Mises test and only applicable to testing a simple null hypothesis or a composite null hypothesis with a scale distribution family, we do not compare our new test with it.

Consider the model $Y_i = \exp(\alpha X_i) + \epsilon_i$ in Section 7 of [7] with $\alpha = 0.25$ and $X_i \sim Uniform(2, 4)$. First consider the case of large sample size by drawing 10,000 random samples of size $n = 200$ and 500 from the above model with either

$$F_\epsilon(x) = \left(1-\frac{\delta}{\sqrt{n}}\right)N(0,1)+\frac{\delta}{\sqrt{n}}t(\nu) \tag{8}$$

or

$$F_\epsilon(x) = \left(1-\frac{\delta}{\sqrt{n}}\right)t(\nu)+\frac{\delta}{\sqrt{n}}N(0,1) \tag{9}$$

for $\delta = 0, 0.5, 1, 2, 3$.

The aim is to test either $H_0 : \epsilon_i \sim N(0, 1)$ or $H_0 : \epsilon_i \sim t(3)$ or $H_0 : \epsilon_i \sim t(8)$ or $H_0 : F_\epsilon \in \mathcal{F}^n = \{N(0, \sigma^2) : \sigma > 0\}$ or $H_0 : F_\epsilon \in \mathcal{F}^t = \{t(\nu) : \nu > 2\}$. In the case of composite null hypothesis, β equals either σ or ν and $\hat{\beta}$ is the corresponding moment estimator based on the estimated errors $\hat{\epsilon}_i's$. For computing the power of the Cramér–von-Mises test, a parametric bootstrap method with repetition 1,000 is employed to obtain the critical values. More specifically, we generate 1,000 random samples with size n from F_ϵ in the case of simple null hypothesis or $F_\epsilon(;\hat{\beta})$ in the case of composite hypothesis. Denote them by $\{\epsilon_i^{*(j)}\}_{i=1}^n$ for $j = 1, \ldots, 1000$. For each $j = 1, \ldots, 1000$, we further generate a bootstrap sample

$$Y_i^{*(j)} = m(X_i; \hat{\alpha}) + \epsilon_i^{*(j)} \quad \text{for } i = 1, \ldots, n.$$

Based on $\{(X_i, Y_i^{*(j)})^T\}_{i=1}^n$ for each $j = 1, \ldots, 1000$, we compute the corresponding least squares estimator for α, the moment estimator for β in the case of composite null hypothesis and estimated errors, say $\hat{\alpha}^{*(j)}$, $\hat{\beta}^{*(j)}$, $\{\hat{\epsilon}_i^{*(j)}\}_{i=1}^n$. Using these bootstrap quantities, we obtained 1000 bootstrapped Cramér–von-Mises test statistics, which give the critical values. Note that [8] employed the naive bootstrap method, i.e., resampling from the estimated errors nonparametrically, for obtaining critical values for the Kolmogorov–Smirnov test. Since we are testing a parametric distribution family for ϵ_i, it prefers to employing the parametric bootstrap method.

The empirical sizes and powers of the proposed jackknife empirical likelihood method and the Cramér–von-Mises test are reported in Tables 1–4. From these tables, we observe that (i) results for $\delta = 0$ show that the size of the proposed jackknife empirical likelihood method is close to the nominal level and its accuracy is improved when the sample size becomes large; (ii) results for $\delta = 0.5, 1, 2, 3$ show that the proposed jackknife empirical likelihood method is more powerful than the Cramér–von-Mises test for most cases, especially for the simple null hypothesis; (iii) both tests almost have no power for testing $H_0 : F_\epsilon \in \mathcal{F}^n$ when δ is not large.

Second we consider the case of small sample size by drawing 10,000 random samples with size $n = 50$ and 100 from the above model. It turns out that the size of the proposed jackknife empirical likelihood method is larger than the nominal level for $n = 50$. Hence we propose the following bootstrap calibration method, where more details on calibration for empirical likelihood methods can be found in [11].

Table 1
Powers of the proposed jackknife empirical likelihood test (JEL) and the Cramér–von-Mises test (CM) are reported for the case of $n = 200$ and $\nu = 3$. Define $\mathcal{F}^n = \{N(0, \sigma^2) : \sigma > 0\}$ and $\mathcal{F}^t = \{t(\nu) : \nu > 2\}$.

δ	H_0	JEL Level 5%	CM Level 5%	JEL Level 10%	CM Level 10%
0	$N(0, 1)$	0.0500	0.0479	0.0999	0.0944
	$t(3)$	0.0592	0.0541	0.1128	0.1031
	$\mathcal{F}^n$	0.0597	0.0512	0.1115	0.1007
	$\mathcal{F}^t$	0.0584	0.0451	0.1095	0.0974
0.5	$N(0, 1)$	0.0760	0.0667	0.1405	0.1297
	$t(3)$	0.0814	0.0497	0.1423	0.1006
	$\mathcal{F}^n$	0.0651	0.0490	0.1170	0.0996
	$\mathcal{F}^t$	0.0663	0.0504	0.1210	0.1063
1	$N(0, 1)$	0.1542	0.1384	0.2426	0.2193
	$t(3)$	0.1247	0.0531	0.1989	0.1164
	$\mathcal{F}^n$	0.0620	0.0541	0.1157	0.1039
	$\mathcal{F}^t$	0.1007	0.0547	0.1702	0.1227
2	$N(0, 1)$	0.4158	0.3793	0.5404	0.5129
	$t(3)$	0.3238	0.1067	0.4340	0.2424
	$\mathcal{F}^n$	0.0609	0.0604	0.1104	0.1096
	$\mathcal{F}^t$	0.2496	0.1289	0.3559	0.2601
3	$N(0, 1)$	0.6677	0.6397	0.7709	0.7466
	$t(3)$	0.6092	0.2737	0.7181	0.4959
	$\mathcal{F}^n$	0.0819	0.0877	0.1464	0.1498
	$\mathcal{F}^t$	0.4995	0.3133	0.6227	0.5199
$\sqrt{n}$	$\mathcal{F}^n$	0.9580	0.9649	0.9752	0.9817

Table 2
Powers of the proposed jackknife empirical likelihood test (JEL) and the Cramér–von-Mises test (CM) are reported for the case of $n = 200$ and $\nu = 8$. Define $\mathcal{F}^n = \{N(0, \sigma^2) : \sigma > 0\}$ and $\mathcal{F}^t = \{t(\nu) : \nu > 2\}$.

δ	H_0	JEL Level 5%	CM Level 5% Level 5%	JEL Level 10% Level 10%	CM Level 10% Level 10%
0	$N(0, 1)$	0.0527	0.0478	0.1044	0.0969
	$t(8)$	0.0550	0.0538	0.1062	0.1040
	$\mathcal{F}^n$	0.0632	0.0551	0.1185	0.1011
	$\mathcal{F}^t$	0.0548	0.0494	0.1043	0.0934
0.5	$N(0, 1)$	0.0819	0.0724	0.1445	0.1357
	$t(8)$	0.0784	0.0671	0.1363	0.1276
	$\mathcal{F}^n$	0.0628	0.0505	0.1152	0.0998
	$\mathcal{F}^t$	0.0717	0.0632	0.1335	0.1198
1	$N(0, 1)$	0.1593	0.1394	0.2555	0.2292
	$t(8)$	0.1289	0.1216	0.2391	0.2100
	$\mathcal{F}^n$	0.0593	0.0499	0.1169	0.1028
	$\mathcal{F}^t$	0.1299	0.1204	0.2158	0.2016
2	$N(0, 1)$	0.5104	0.4574	0.6399	0.5908
	$t(8)$	0.4636	0.3894	0.5907	0.5283
	$\mathcal{F}^n$	0.0624	0.0508	0.1169	0.1014
	$\mathcal{F}^t$	0.4043	0.3809	0.5350	0.5158
3	$N(0, 1)$	0.8368	0.7973	0.9040	0.8786
	$t(8)$	0.8042	0.7508	0.8810	0.8469
	$\mathcal{F}^n$	0.0593	0.0527	0.1140	0.1033
	$\mathcal{F}^t$	0.7598	0.7399	0.8480	0.8369
$\sqrt{n}$	$\mathcal{F}^n$	0.2728	0.2528	0.3721	0.3655

Draw 1,000 resamples from $\{(\epsilon_i(\hat{\alpha}), \tilde{\epsilon}_i(\hat{\alpha}))\}_{i=1}^k$ with size $k = [n/2]$, say $\{(\epsilon_i^{*(b)}(\hat{\alpha}), \tilde{\epsilon}_i^{*(b)}(\hat{\alpha}))\}_{i=1}^k$ for $b = 1, \ldots, 1{,}000$. For each resample $\{(\epsilon_i^{*(b)}(\hat{\alpha}), \tilde{\epsilon}_i^{*(b)}(\hat{\alpha}))\}_{i=1}^k$, we use the model (1) to generate a resample

$$Y_i^{*(b)} = m(X_i; \hat{\alpha}) + \epsilon_i^{*(b)}, \qquad Y_{k+i}^{*(b)} = m(X_{k+i}; \hat{\alpha}) + \tilde{\epsilon}_i^{*(b)}$$

for $i = 1, \ldots, k$. Next based on $\{(X_i, Y_i^{*(b)})\}_{i=1}^{2k}$, we re-estimate the parameters and calculate the jackknife empirical likelihood function, which results in 1,000 jackknife empirical likelihood functions. Therefore, the bootstrap calibrated jackknife empirical likelihood test is computed by obtaining critical values from the computed 1,000 jackknife empirical likelihood functions instead of the chi-square distribution with two degrees of freedom. Denote this bootstrap calibrated jackknife empirical likelihood test as BCJEL.

In Tables 5 and 6 we report the empirical sizes and powers of the proposed jackknife empirical likelihood method, its bootstrap calibrated version and the Cramér–von-Mises test. From these two tables we observe that (i) the size of the jackknife empirical likelihood test is larger than the nominal level for $n = 50$, but gets more accurate when $n = 100$;

Table 3
Powers of the proposed jackknife empirical likelihood test (JEL) and the Cramér–von-Mises test (CM) are reported for the case of $n = 500$ and $\nu = 3$. Define $\mathcal{F}^n = \{N(0, \sigma^2) : \sigma > 0\}$ and $\mathcal{F}^t = \{t(\nu) : \nu > 2\}$.

δ	H_0	JEL Level 5%	CM Level 5%	JEL Level 10%	CM Level 10%
0	$N(0, 1)$	0.0505	0.0544	0.1061	0.1018
	$t(3)$	0.0532	0.0517	0.1040	0.0994
	$\mathcal{F}^n$	0.0545	0.0518	0.1023	0.1036
	$\mathcal{F}^t$	0.0561	0.0483	0.1054	0.0991
0.5	$N(0, 1)$	0.0751	0.0692	0.1370	0.1296
	$t(3)$	0.0675	0.0487	0.1458	0.1018
	$\mathcal{F}^n$	0.0528	0.0505	0.1027	0.1009
	$\mathcal{F}^t$	0.0615	0.0510	0.1183	0.1016
1	$N(0, 1)$	0.1543	0.1400	0.2491	0.2284
	$t(3)$	0.1252	0.0589	0.1913	0.1234
	$\mathcal{F}^n$	0.0527	0.0490	0.1021	0.0967
	$\mathcal{F}^t$	0.0921	0.0650	0.1636	0.1304
2	$N(0, 1)$	0.4609	0.4191	0.5903	0.5483
	$t(3)$	0.2982	0.1173	0.4189	0.2567
	$\mathcal{F}^n$	0.0496	0.0556	0.1005	0.1061
	$\mathcal{F}^t$	0.2378	0.1249	0.3450	0.2525
3	$N(0, 1)$	0.7846	0.7387	0.8616	0.8361
	$t(3)$	0.5919	0.3000	0.7065	0.5217
	$\mathcal{F}^n$	0.0570	0.0592	0.1117	0.1095
	$\mathcal{F}^t$	0.4695	0.3151	0.5948	0.5180
$\sqrt{n}$	$\mathcal{F}^n$	0.9999	1	1	1

Table 4
Powers of the proposed jackknife empirical likelihood test (JEL) and the Cramér–von-Mises test (CM) are reported for the case of $n = 500$ and $\nu = 8$. Define $\mathcal{F}^n = \{N(0, \sigma^2) : \sigma > 0\}$ and $\mathcal{F}^t = \{t(\nu) : \nu > 2\}$.

δ	H_0	JEL Level 5%	CM Level 5%	JEL Level 10%	CM Level 10%
0	$N(0, 1)$	0.0507	0.0503	0.1015	0.0999
	$t(8)$	0.0541	0.0485	0.1040	0.1010
	$\mathcal{F}^n$	0.0551	0.0527	0.1088	0.1032
	$\mathcal{F}^t$	0.0511	0.0503	0.0984	0.0982
0.5	$N(0, 1)$	0.0776	0.0708	0.1441	0.1355
	$t(8)$	0.0779	0.0686	0.1347	0.1303
	$\mathcal{F}^n$	0.0527	0.0510	0.1045	0.0985
	$\mathcal{F}^t$	0.0691	0.0642	0.1282	0.1250
1	$N(0, 1)$	0.1584	0.1433	0.2484	0.2340
	$t(8)$	0.1474	0.1232	0.2381	0.2078
	$\mathcal{F}^n$	0.0553	0.0536	0.1048	0.1024
	$\mathcal{F}^t$	0.1266	0.1275	0.2134	0.2086
2	$N(0, 1)$	0.5191	0.4777	0.6462	0.6103
	$t(8)$	0.4508	0.4014	0.5821	0.5322
	$\mathcal{F}^n$	0.0573	0.0552	0.1068	0.1026
	$\mathcal{F}^t$	0.3926	0.3897	0.5230	0.5232
3	$N(0, 1)$	0.8720	0.8316	0.9279	0.9041
	$t(8)$	0.8209	0.7696	0.8949	0.8615
	$\mathcal{F}^n$	0.0544	0.0514	0.1010	0.0995
	$\mathcal{F}^t$	0.7559	0.7572	0.8411	0.8533
$\sqrt{n}$	$\mathcal{F}^n$	0.5899	0.5452	0.6937	0.6699

(ii) the size of the bootstrap calibrated jackknife empirical likelihood test is comparable with that of the Cramér–von-Mises test; (iii) for testing t distributions, the bootstrap calibrated jackknife empirical likelihood test is more powerful than the Cramér–von-Mises test for the simple null hypothesis, but less powerful for the composite null hypothesis; (iv) both the bootstrap calibrated jackknife empirical likelihood test and the Cramér–von-Mises test perform similar for testing normal distributions; v) for sample size $n = 100$, the jackknife empirical likelihood test has a reasonable size and is most powerful.

4. Conclusions

We propose some jackknife empirical likelihood methods to test whether the error distribution in a regression model belongs to a particular parametric family. Unlike classical goodness-of-fit tests, the new tests always have a chi-square limit and so no ad hoc techniques such as bootstrap method are needed to obtain critical values. Also the calculation is quite straightforward and involves no numeric integration unlike the method in [7]. When the sample size is small ($n = 50$), the

Table 5
Powers of the proposed jackknife empirical likelihood test (JEL), its bootstrap calibrated version (BCJEL) and the Cramér–von-Mises test (CM) are reported for the case of $n = 50$ and $\nu = 8$. Define $\mathcal{F}^n = \{N(0, \sigma^2) : \sigma > 0\}$ and $\mathcal{F}^t = \{t(\nu) : \nu > 2\}$.

δ	H_0	JEL Level 5%	BCJEL Level 5%	CM Level 5%	JEL Level 10%	BCJEL Level 10%	CM Level 10%
0	$N(0, 1)$	0.0836	0.0442	0.0510	0.1437	0.0901	0.1000
	$t(8)$	0.0877	0.0423	0.0506	0.1450	0.0893	0.0981
	$\mathcal{F}^n$	0.1133	0.0460	0.0539	0.1743	0.0934	0.1039
	$\mathcal{F}^t$	0.0816	0.0430	0.0376	0.1380	0.0895	0.0802
0.5	$N(0, 1)$	0.1327	0.0786	0.0671	0.1999	0.1368	0.1319
	$t(8)$	0.1273	0.0733	0.0644	0.1899	0.1263	0.1218
	$\mathcal{F}^n$	0.1118	0.0458	0.0476	0.1730	0.0914	0.0916
	$\mathcal{F}^t$	0.1036	0.0586	0.0580	0.1666	0.1072	0.1104
1	$N(0, 1)$	0.2289	0.1430	0.1335	0.3153	0.2308	0.2174
	$t(8)$	0.2118	0.1386	0.1171	0.2952	0.2136	0.1977
	$\mathcal{F}^n$	0.1170	0.0468	0.0509	0.1782	0.0930	0.0995
	$\mathcal{F}^t$	0.1619	0.0975	0.1031	0.2294	0.1620	0.1775
2	$N(0, 1)$	0.5123	0.3665	0.3669	0.6150	0.5038	0.4957
	$t(8)$	0.5128	0.3685	0.3329	0.6146	0.5027	0.4691
	$\mathcal{F}^n$	0.1107	0.0515	0.0425	0.1729	0.0879	0.1044
	$\mathcal{F}^t$	0.4047	0.2723	0.3058	0.5123	0.3929	0.4302
3	$N(0, 1)$	0.6505	0.4868	0.5327	0.7474	0.6349	0.6659
	$t(8)$	0.7566	0.5973	0.5724	0.8348	0.7374	0.7105
	$\mathcal{F}^n$	0.1101	0.0399	0.0530	0.1772	0.0838	0.1047
	$\mathcal{F}^t$	0.6576	0.4966	0.5395	0.7558	0.6370	0.6576

Table 6
Powers of the proposed jackknife empirical likelihood test (JEL), its bootstrap calibrated version (BCJEL) and the Cramér–von-Mises test (CM) are reported for the case of $n = 100$ and $\nu = 8$. Define $\mathcal{F}^n = \{N(0, \sigma^2) : \sigma > 0\}$ and $\mathcal{F}^t = \{t(\nu) : \nu > 2\}$.

δ	H_0	JEL Level 5%	BCJEL Level 5%	CM Level 5%	JEL Level 10%	BCJEL Level 10%	CM Level 10%
0	$N(0, 1)$	0.0615	0.0446	0.0518	0.1153	0.0905	0.1024
	$t(8)$	0.0618	0.0434	0.0480	0.1108	0.0861	0.1006
	$\mathcal{F}^n$	0.0709	0.0393	0.0487	0.1302	0.0819	0.0990
	$\mathcal{F}^t$	0.0620	0.0431	0.0393	0.1159	0.0934	0.0820
0.5	$N(0, 1)$	0.0934	0.0694	0.0716	0.1586	0.1279	0.1331
	$t(8)$	0.0931	0.0702	0.0611	0.1562	0.1308	0.1178
	$\mathcal{F}^n$	0.0739	0.0408	0.0494	0.1341	0.0865	0.1000
	$\mathcal{F}^t$	0.0696	0.0522	0.0542	0.1251	0.0999	0.1054
1	$N(0, 1)$	0.1748	0.1424	0.1426	0.2634	0.2246	0.2361
	$t(8)$	0.1649	0.1335	0.1224	0.2542	0.2142	0.2041
	$\mathcal{F}^n$	0.0738	0.0428	0.0533	0.1335	0.0839	0.1032
	$\mathcal{F}^t$	0.1135	0.0868	0.0989	0.1838	0.1517	0.1670
2	$N(0, 1)$	0.4978	0.4370	0.4276	0.6211	0.5703	0.5642
	$t(8)$	0.4729	0.4068	0.3741	0.5965	0.5442	0.5078
	$\mathcal{F}^n$	0.0717	0.0392	0.0539	0.1317	0.0827	0.1037
	$\mathcal{F}^t$	0.3392	0.2794	0.3086	0.4570	0.3990	0.4359
3	$N(0, 1)$	0.7753	0.7153	0.7165	0.8662	0.8264	0.8221
	$t(8)$	0.7948	0.7374	0.6990	0.8727	0.8374	0.8161
	$\mathcal{F}^n$	0.0751	0.0383	0.0548	0.1299	0.0810	0.1037
	$\mathcal{F}^t$	0.6630	0.5852	0.6317	0.7731	0.7169	0.7500

sizes of the jackknife empirical likelihood tests are larger than the nominal level and a bootstrap calibration is proposed to improve the size. A simulation study confirms that the sizes of the new methods are reasonably accurate for sample size larger than 100 and powerful too.

Acknowledgments

We thank an associate editor and two reviewers for helpful comments. Peng's research was supported by NSF grant DMS-1005336 and HKSAR-RGC-GRF 400410.

Appendix. Proofs

Before proving theorems, we need some lemmas.

Lemma 1. *Under conditions of Theorem 1, we have as $n \to \infty$*

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k} G_1(i) = \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\left\{\frac{F^2(\epsilon_i(\alpha_0);\beta_0)+F^2(\tilde{\epsilon}_i(\alpha_0);\beta_0)}{2} - F(\epsilon_i(\alpha_0)\vee\tilde{\epsilon}_i(\alpha_0);\beta_0)+\frac{1}{3}\right\}+o_p(1)$$
$$=: W_{k1}+o_p(1)$$

and

$$\begin{aligned}\frac{1}{\sqrt{k}}\sum_{i=1}^{k} G_2(i) &= \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F(\epsilon_i(\alpha_0);\beta_0)+F(\tilde{\epsilon}_i(\alpha_0);\beta_0)-2F^3(\epsilon_i(\alpha_0);\beta_0)-2F^3(\tilde{\epsilon}_i(\alpha_0);\beta_0)\}\\ &\quad + E\left\{2F'(\epsilon_1(\alpha_0);\beta_0)-12F^2(\epsilon_1(\alpha_0);\beta_0)F'(\epsilon_1(\alpha_0);\beta_0)\right\}\\ &\quad \times E\left\{\frac{\partial}{\partial\alpha^T}m(X_1;\alpha_0)\right\}\left\{E\frac{\partial}{\partial\alpha^T}h_1(\alpha_0)\right\}^{-1}\frac{1}{\sqrt{k}}\sum_{i=1}^{k}h_i(\alpha_0)+o_p(1)\\ &=: W_{k2}+o_p(1).\end{aligned}$$

Proof. For simplicity we write $F(x)$, $m(x)$, ϵ_i and $\tilde{\epsilon}_i$ instead of $F(x;\beta_0)$, $m(x;\alpha_0)$, $\epsilon_i(\alpha_0)$ and $\tilde{\epsilon}_i(\alpha_0)$, respectively. So

$$\begin{aligned}F(\epsilon_j^*(\hat{\alpha}))-F(\epsilon_j^*(\hat{\alpha}_i)) &= F(\epsilon_j(\hat{\alpha}))-F(\epsilon_j(\hat{\alpha}_i))+\{F(\tilde{\epsilon}_j(\hat{\alpha}))-F(\epsilon_j(\hat{\alpha}))-F(\tilde{\epsilon}_j(\hat{\alpha}_i))+F(\epsilon_j(\hat{\alpha}_i))\}I(\epsilon_j(\hat{\alpha})\le\tilde{\epsilon}_j(\hat{\alpha}))\\ &\quad +\{F(\tilde{\epsilon}_j(\hat{\alpha}_i))-F(\epsilon_j(\hat{\alpha}_i))\}\{I(\epsilon_j(\hat{\alpha})\le\tilde{\epsilon}_j(\hat{\alpha}))-I(\epsilon_j(\hat{\alpha}_i)\le\tilde{\epsilon}_j(\hat{\alpha}_i))\}\\ &= I_1(j,i)+I_2(j,i)+I_3(j,i).\end{aligned}$$

Since $\max_{1\le i\le k}|\hat{\alpha}-\hat{\alpha}_i| = O_p(k^{-\delta})$ for some $\delta > 1/2$, by conditions (A1)–(A3), there are some $\delta' \in (1/2,\delta)$ and some $M > 0$ such that

$$\begin{aligned}\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\sum_{j\ne i} I_3(j,i) &= O_p\left(\left|\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\sum_{j\ne i}\{F(\tilde{\epsilon}_j(\hat{\alpha}_i))-F(\epsilon_j(\hat{\alpha}_i))\}\right.\right.\\ &\quad\left.\left.\times I\left(|\epsilon_j(\hat{\alpha}_i)-\tilde{\epsilon}_j(\hat{\alpha}_i)|\le M\left\|\left(\frac{1}{k}\sum_{l=1}^{k}\frac{\partial}{\partial\alpha^T}h_l(\hat{\alpha})\right)^{-1}h_i(\hat{\alpha})\right\|k^{-1}\right)\right|\right)\\ &= O_p\left(\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\sum_{j\ne i}M^2\left\|\left(\frac{1}{k}\sum_{l=1}^{k}\frac{\partial}{\partial\alpha^T}h_l(\hat{\alpha})\right)^{-1}h_i(\hat{\alpha})\right\|k^{-1}I(|\epsilon_j(\hat{\alpha})-\tilde{\epsilon}_j(\hat{\alpha})|\le k^{-\delta'})\right)\\ &= o_p(1).\end{aligned} \tag{10}$$

Since $\sum_{i=1}^{k}h_i(\hat{\alpha})=0$, it follows from Taylor expansions that

$$\begin{aligned}\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\sum_{j\ne i} I_1(j,i) &= \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\sum_{j\ne i}F'(\epsilon_j(\hat{\alpha}))\left\{\frac{\partial}{\partial\alpha^T}m(X_j;\hat{\alpha})\right\}\left\{\frac{1}{k}\sum_{l=1}^{k}\frac{\partial}{\partial\alpha^T}h_l(\hat{\alpha})\right\}^{-1}\frac{h_i(\hat{\alpha})}{k}+o_p(1)\\ &= -\frac{1}{\sqrt{k}}\sum_{j=1}^{k}F'(\epsilon_j(\hat{\alpha}))\left\{\frac{\partial}{\partial\alpha^T}m(X_j;\hat{\alpha})\right\}\left\{\frac{1}{k}\sum_{l=1}^{k}\frac{\partial}{\partial\alpha^T}h_l(\hat{\alpha})\right\}^{-1}\frac{h_j(\hat{\alpha})}{k}+o_p(1)\\ &= o_p(1).\end{aligned}$$

Similarly we have

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\sum_{j\ne i} I_2(j,i) = o_p(1).$$

Therefore,

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\sum_{j\ne i}\{F(\epsilon_j^*(\hat{\alpha}))-F(\epsilon_j^*(\hat{\alpha}_i))\} = o_p(1). \tag{11}$$

Using $\sum_{i=1}^{k}h_i(\hat{\alpha})=0$ again, we have

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\sum_{j\ne i}\{F^2(\epsilon_j(\hat{\alpha}))+F^2(\tilde{\epsilon}_j(\hat{\alpha}))-F^2(\epsilon_j(\hat{\alpha}_i))-F^2(\tilde{\epsilon}_j(\hat{\alpha}_i))\}$$

$$= \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \sum_{j \neq i} 2F(\epsilon_j(\hat{\alpha}))F'(\epsilon_j(\hat{\alpha})) \left\{ \frac{\partial}{\partial \alpha^T} m(X_j; \hat{\alpha}) \right\} \left\{ \frac{1}{k} \sum_{l=1}^{k} \frac{\partial}{\partial \alpha^T} h_l(\hat{\alpha}) \right\}^{-1} \frac{h_i(\hat{\alpha})}{k}$$

$$+ \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \sum_{j \neq i} 2F(\tilde{\epsilon}_j(\hat{\alpha}))F'(\tilde{\epsilon}_j(\hat{\alpha})) \left\{ \frac{\partial}{\partial \alpha^T} m(X_{k+j}; \hat{\alpha}) \right\} \left\{ \frac{1}{k} \sum_{l=1}^{k} \frac{\partial}{\partial \alpha^T} h_l(\hat{\alpha}) \right\}^{-1} \frac{h_i(\hat{\alpha})}{k} + o_p(1)$$

$$= -\frac{1}{\sqrt{k}} \sum_{j=1}^{k} 2F(\epsilon_j(\hat{\alpha}))F'(\epsilon_j(\hat{\alpha})) \left\{ \frac{\partial}{\partial \alpha^T} m(X_j; \hat{\alpha}) \right\} \left\{ \frac{1}{k} \sum_{l=1}^{k} \frac{\partial}{\partial \alpha^T} h_l(\hat{\alpha}) \right\}^{-1} \frac{h_j(\hat{\alpha})}{k}$$

$$+ \frac{1}{\sqrt{k}} \sum_{j=1}^{k} 2F(\tilde{\epsilon}_j(\hat{\alpha}))F'(\tilde{\epsilon}_j(\hat{\alpha})) \left\{ \frac{\partial}{\partial \alpha^T} m(X_{k+j}; \hat{\alpha}) \right\} \left\{ \frac{1}{k} \sum_{l=1}^{k} \frac{\partial}{\partial \alpha^T} h_l(\hat{\alpha}) \right\}^{-1} \frac{h_j(\hat{\alpha})}{k} + o_p(1)$$

$$= o_p(1). \tag{12}$$

Thus, it follows from (11) and (12) that

$$\frac{1}{\sqrt{k}} \sum_{i=1}^{k} G_1(i) = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \left\{ \frac{F^2(\epsilon_i(\hat{\alpha})) + F^2(\tilde{\epsilon}_i(\hat{\alpha}))}{2} - F(\epsilon_i^*(\hat{\alpha})) + 1/3 \right\} + o_p(1). \tag{13}$$

Similar to (12), we can show that

$$\begin{aligned} \frac{1}{\sqrt{k}} \sum_{i=1}^{k} G_2(i) &= \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \{F(\epsilon_i(\hat{\alpha})) + F(\tilde{\epsilon}_i(\hat{\alpha})) - 2F^3(\epsilon_i(\hat{\alpha})) - 2F^3(\tilde{\epsilon}_i(\hat{\alpha}))\} \\ &\quad + \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \sum_{j \neq i} \{F(\epsilon_j(\hat{\alpha})) + F(\tilde{\epsilon}_j(\hat{\alpha})) - F(\epsilon_j(\hat{\alpha}_i)) - F(\tilde{\epsilon}_j(\hat{\alpha}_i))\} \\ &\quad - \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \sum_{j \neq i} \{F^3(\epsilon_j(\hat{\alpha})) + F^3(\tilde{\epsilon}_j(\hat{\alpha})) - F^3(\epsilon_j(\hat{\alpha}_i)) - F^3(\tilde{\epsilon}_j(\hat{\alpha}_i))\} \\ &= \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \{F(\epsilon_i(\hat{\alpha})) + F(\tilde{\epsilon}_i(\hat{\alpha})) - 2F^3(\epsilon_i(\hat{\alpha})) - 2F^3(\tilde{\epsilon}_i(\hat{\alpha}))\} + o_p(1). \end{aligned} \tag{14}$$

It is easy to show that

$$\sqrt{k}\{\hat{\alpha} - \alpha_0\} = -\left\{ E \frac{\partial}{\partial \alpha^T} h_1(\alpha_0) \right\}^{-1} \frac{1}{\sqrt{k}} \sum_{i=1}^{k} h_i(\alpha_0) + o_p(1). \tag{15}$$

Like the proof of (10), we have

$$\begin{aligned} \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \{F(\epsilon_i^*(\hat{\alpha})) - 2/3\} &= \frac{1}{\sqrt{k}} \sum_{i=1}^{k} F(\epsilon_i(\hat{\alpha})) - \frac{2}{3}\sqrt{k} + \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \{F(\tilde{\epsilon}_i(\hat{\alpha})) - F(\epsilon_i(\hat{\alpha}))\} I(\epsilon_i(\hat{\alpha}) < \tilde{\epsilon}_i(\hat{\alpha})) \\ &= \frac{1}{\sqrt{k}} \sum_{i=1}^{k} F(\epsilon_i(\hat{\alpha})) - \frac{2}{3}\sqrt{k} + \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \{F(\tilde{\epsilon}_i(\hat{\alpha})) - F(\epsilon_i(\hat{\alpha}))\} I(\epsilon_i < \tilde{\epsilon}_i) + o_p(1) \\ &= \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \left\{ F(\epsilon_i) - F'(\epsilon_i) \frac{\partial}{\partial \alpha^T} m(X_i; \alpha_0)(\hat{\alpha} - \alpha_0) \right\} - \frac{2}{3}\sqrt{k} \\ &\quad + \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \left\{ F(\tilde{\epsilon}_i) - F'(\tilde{\epsilon}) \frac{\partial}{\partial \alpha^T} m(X_{k+i}; \alpha_0)(\hat{\alpha} - \alpha_0) - F(\epsilon_i) \right. \\ &\quad \left. + F'(\epsilon_i) \frac{\partial}{\partial \alpha^T} m(X_i; \alpha_0)(\hat{\alpha} - \alpha_0) \right\} I(\epsilon_i < \tilde{\epsilon}_i) + o_p(1) \\ &= \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \{F(\epsilon_i \vee \tilde{\epsilon}_i) - 2/3\} - E \left\{ F'(\epsilon_1) \frac{\partial}{\partial \alpha^T} m(X_1; \alpha_0) I(\epsilon_1 > \tilde{\epsilon}_1) \right. \\ &\quad \left. + F'(\tilde{\epsilon}_1) \frac{\partial}{\partial \alpha^T} m(X_{k+1}; \alpha_0) I(\epsilon_1 \leq \tilde{\epsilon}_1) \right\} \sqrt{k}(\hat{\alpha} - \alpha_0) + o_p(1) \end{aligned}$$

$$= \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F(\epsilon_i \vee \tilde{\epsilon}_i) - 2/3\} - 2E\{F(\epsilon_1)F'(\epsilon_1)\}E \times \left\{\frac{\partial}{\partial \alpha^T} m(X_1;\alpha_0)\right\}\sqrt{k}(\hat{\alpha}-\alpha_0) + o_p(1). \quad (16)$$

It is easy to verify that

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F(\epsilon_i(\hat{\alpha})) - 1/2\} = \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F(\epsilon_i) - 1/2\} - E\{F'(\epsilon_1)\}E\left\{\frac{\partial}{\partial \alpha^T} m(X_1;\alpha_0)\right\}\sqrt{k}(\hat{\alpha}-\alpha_0) + o_p(1), \quad (17)$$

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F(\tilde{\epsilon}_i(\hat{\alpha})) - 1/2\} = \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F(\tilde{\epsilon}_i) - 1/2\} - E\{F'(\epsilon_1)\}E\left\{\frac{\partial}{\partial \alpha^T} m(X_1;\alpha_0)\right\}\sqrt{k}(\hat{\alpha}-\alpha_0) + o_p(1), \quad (18)$$

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F^2(\epsilon_i(\hat{\alpha})) - 1/3\} = \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F^2(\epsilon_i) - 1/3\} - 2E\{F(\epsilon_1)F'(\epsilon_1)\}E\left\{\frac{\partial}{\partial \alpha^T} m(X_1;\alpha_0)\right\}\sqrt{k}(\hat{\alpha}-\alpha_0) + o_p(1), \quad (19)$$

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F^2(\tilde{\epsilon}_i(\hat{\alpha})) - 1/3\} = \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F^2(\tilde{\epsilon}_i) - 1/3\} - 2E\{F(\epsilon_1)F'(\epsilon_1)\}E\left\{\frac{\partial}{\partial \alpha^T} m(X_1;\alpha_0)\right\}\sqrt{k}(\hat{\alpha}-\alpha_0) + o_p(1), \quad (20)$$

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F^3(\epsilon_i(\hat{\alpha})) - 1/4\} = \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F^3(\epsilon_i) - 1/4\} - 3E\{F^2(\epsilon_1)F'(\epsilon_1)\}E\left\{\frac{\partial}{\partial \alpha^T} m(X_1;\alpha_0)\right\}\sqrt{k}(\hat{\alpha}-\alpha_0) + o_p(1) \quad (21)$$

and

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F^3(\tilde{\epsilon}_i(\hat{\alpha})) - 1/4\} = \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\{F^3(\tilde{\epsilon}_i) - 1/4\} - 3E\{F^2(\epsilon_1)F'(\epsilon_1)\}E \times \left\{\frac{\partial}{\partial \alpha^T} m(X_1;\alpha_0)\right\}\sqrt{k}(\hat{\alpha}-\alpha_0) + o_p(1). \quad (22)$$

Hence the lemma follows from (13)–(22). □

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

$$\frac{1}{k}\sum_{i=1}^{k} G_j(i)G_l(i) \xrightarrow{p} \lim_{n\to\infty} E(W_{kj}W_{kl})$$

for $j, l = 1, 2$ *as* $n \to \infty$.

Proof. Put

$$A_{i1} = \sum_{j\neq i}\frac{F^2(\epsilon_j(\hat{\alpha})) + F^2(\tilde{\epsilon}_j(\hat{\alpha})) - F^2(\epsilon_j(\hat{\alpha}_i)) - F^2(\tilde{\epsilon}_j(\hat{\alpha}_i))}{2}$$

and

$$A_{i2} = \sum_{j\neq i}\{F(\epsilon_j^*(\hat{\alpha})) - F(\epsilon_j^*(\hat{\alpha}_i))\}.$$

Like the proof of (10), we have

$$\begin{aligned} A_{i2} &= \sum_{j\neq i}\{F(\epsilon_j(\hat{\alpha})) - F(\epsilon_j(\hat{\alpha}_i))\}I(\epsilon_j(\hat{\alpha}) > \tilde{\epsilon}_j(\hat{\alpha})) + \sum_{j\neq i}\{F(\tilde{\epsilon}_j(\hat{\alpha})) - F(\tilde{\epsilon}_j(\hat{\alpha}_i))\}I(\epsilon_j(\hat{\alpha}) \le \tilde{\epsilon}_j(\hat{\alpha})) + o_p(1) \\ &= \sum_{j\neq i}F'(\epsilon_j)\left\{\frac{\partial}{\partial \alpha^T} m(X_j)\right\}\left\{\frac{1}{k}\sum_{l=1}^{k}\frac{\partial}{\partial \alpha^T}h_l(\alpha_0)\right\}^{-1} h_i(\alpha_0)k^{-1}I(\epsilon_j > \tilde{\epsilon}_j) \\ &\quad + \sum_{j\neq i}F'(\tilde{\epsilon}_j)\left\{\frac{\partial}{\partial \alpha^T} m(X_{k+j})\right\}\left\{\frac{1}{k}\sum_{l=1}^{k}\frac{\partial}{\partial \alpha^T}h_l(\alpha_0)\right\}^{-1} h_i(\alpha_0)k^{-1}I(\epsilon_j \le \tilde{\epsilon}_j) + o_p(1) \\ &= 2E\{F(\epsilon_1)F'(\epsilon_1)\}E\left\{\frac{\partial}{\partial \alpha^T} m(X_1)\right\}\left\{E\frac{\partial}{\partial \alpha^T}h_1(\alpha_0)\right\}^{-1} h_i(\alpha_0) + o_p(1). \end{aligned} \quad (23)$$

It is easy to check that

$$A_{i1} = 2E\{F(\epsilon_1)F'(\epsilon_1)\}E\left\{\frac{\partial}{\partial\alpha^T}m(X_1)\right\}\left\{E\frac{\partial}{\partial\alpha^T}h_1(\alpha_0)\right\}^{-1}h_i(\alpha_0) + o_p(1). \tag{24}$$

Thus, it follows from (23) and (24) that

$$\begin{aligned}\frac{1}{k}\sum_{i=1}^{k}G_1^2(i) &= \frac{1}{k}\sum_{i=1}^{k}(A_{i1}-A_{i2})^2 + \frac{1}{k}\sum_{i=1}^{k}\left(\frac{F^2(\epsilon_i(\hat\alpha)) + F^2(\tilde\epsilon_i(\hat\alpha))}{2} - F(\epsilon_i^*(\hat\alpha)) + \frac{1}{3}\right)^2 \\ &\quad + \frac{2}{k}\sum_{i=1}^{k}(A_{i1}-A_{i2})\left(\frac{F^2(\epsilon_i(\hat\alpha)) + F^2(\tilde\epsilon_i(\hat\alpha))}{2} - F(\epsilon_i^*(\hat\alpha)) + \frac{1}{3}\right) \\ &= \frac{1}{k}\sum_{i=1}^{k}\left(\frac{F^2(\epsilon_i) + F^2(\tilde\epsilon_i)}{2} - F(\epsilon_i \vee \tilde\epsilon_i) + \frac{1}{3}\right)^2 + o_p(1) \\ &= \lim_{n\to\infty} W_{k1}^2 + o_p(1).\end{aligned}$$

The rest can be shown in a similar way. □

Proof of Theorem 1. It follows from Lemmas 1 and 2 and some standard arguments in the empirical likelihood method (see [11, Chapter 11]). □

Proof of Theorem 2. This can be shown in a similar way to the proof of Theorem 1 although some more tedious expansions are needed. □

References

[1] D. Basso, L. Finos, L. Salmaso, A new association test in linear models with application to experimental designs, J. Appl. Statist. Sci. 15 (2007) 99–109.
[2] L. Corain, V.B. Melas, L. Salmaso, A discussion on nonlinear models for price decisions in rating-based product preference models, Comm. Statist. Simulation Comput. 38 (2009) 1178–1201.
[3] R.B. D'Agostino, M.A. Stephens, Goodness-of-Fit Techniques, Marcel Dekker, Inc., 1986.
[4] A.C. Davison, Statistical Models, Cambridge University Press, 2008.
[5] Y. Gong, L. Peng, Y. Qi, Smoothed jackknife empirical likelihood method for ROC curve, J. Multivariate Anal. 101 (2010) 1520–1531.
[6] B.Y. Jing, J.Q. Yuan, W. Zhou, Jackknife empirical likelihood, J. Amer. Statist. Assoc. 104 (2009) 1224–1232.
[7] E.V. Khmaladze, H.L. Koul, Martingale transformations goodness-of-fit tests in regression models, Ann. Statist. 32 (2004) 995–1034.
[8] H. Koul, L. Sakhanenko, Goodness-of-fit testing in regression: a finite sample comparison of bootstrap methodology and Khmaladze transformation, Statist. Probab. Lett. 74 (2005) 290–302.
[9] M. Li, L. Peng, Empirical Likelihood Test via Estimating Equations, J. Statist. Plann. Inference 141 (2011) 2428–2439.
[10] A. Owen, Empirical likelihood ratio confidence intervals for single functional, Biometrika 75 (1988) 237–249.
[11] A. Owen, Empirical Likelihood, Chapman & Hall/CRC, 2001.
[12] L. Peng, Approximate jackknife empirical likelihood method for estimating equations, Canad. J. Statist. 40 (2012) 110–123.
[13] L. Peng, Y. Qi, Smoothed jackknife empirical likelihood method for tail copulas, Test 19 (2010) 514–536.
[14] L. Peng, Y. Qi, I. Van Keilegom, Jackknife empirical likelihood method for copulas. Test, (2011) (in press).