



Mini Review

Volume 7 Issue 2 - June 2018
 DOI: 10.19080/BBOAJ.2018.07.555708

Biostat Biometrics Open Acc J

Copyright © All rights are reserved by Yongcheng Qi

Jackknife Empirical Likelihood Methods



Yongcheng Qi*

Department of Mathematics and Statistics, University of Minnesota Duluth, USA

Submission: February 15, 2018; **Published:** June 25, 2018

*Corresponding author: Yongcheng Qi, Department of Mathematics and Statistics, University of Minnesota Duluth, 1117 University Drive, Duluth, MN 55812, USA; Email: yqi@d.umn.edu

Abstract

Since Jing et al. [1] propose the jackknife empirical likelihood method for U-statistics, the method has been developed and applied to inference problems in statistical theory and other areas such as biostatistics, medical statistics and insurance. In this short review paper, we give an introduction on one sample and two-sample jackknife empirical methods and smoothed jackknife empirical likelihood methods and present a brief literature review on applications of these methods.

Keywords: Empirical likelihood; Jackknife empirical likelihood; Smoothed jackknife empirical; Likelihood; Wilks' theorem; confidence interval; Medical statistics; Nonparametric; Probability; Normal-approximation; Computing; Nonlinear statistics; U-statistics; Independent random vectors; Chi-square distribution; Pseudo-sample; Density functions; Asymptotic covariances; Dimensional means; Heavy computational; Errors-in-variables; Lagrange multipliers

Abbreviations: ELM: Empirical Likelihood Method; JEL: Jackknife Empirical Likelihood; ROC: Receiver Operating Characteristic

Introduction

The study of empirical likelihood method (ELM) dates back to Thomas & Grunkemeier [2] who introduce a nonparametric likelihood ratio test statistic for constructing confidence intervals for the survival probability for right censored data. The method has drawn much attention and has been extensively investigated and widely used to construct confidence regions and to test hypotheses after the work by Owen [3,4] for the mean vectors and functionals. The empirical likelihood is a non-parametric method, but it possesses some good properties of the parametric likelihood method, and it has many advantages over some classical and modern methods, such as the normal-approximation-based method and the bootstrap method. Computing the empirical likelihood ratio involves optimizing the likelihood function over n parameters, where n is the sample size. For linear functionals or estimating functions, the optimization problem can be easily solved by using the method of Lagrange multipliers. The major problem of extending the ELM to nonlinear functionals or nonlinear statistics is the difficulty of nonlinear optimizations. The research in this direction has been greatly stimulated since Jing et al. [1] propose a jackknife empirical likelihood (JEL) method for U-statistics. The main idea of the JEL method is to apply the ELM to the jackknife samples, and the parameters of interest become the means of jackknife samples. For jackknife method, see, e.g., Shao & Tu [5].

Jackknife empirical likelihood method

Throughout, for each $X = (x_1, \dots, x_n) \in \mathbb{R}^n$, define $X_{(-i)}$ as the $(n-1)$ - dimensional sub- Vector of X when the i -th component x_i

is removed from X for $1 \leq i \leq n$, that is,

$$X_{(-i)} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n).$$

One-sample JEL. Suppose X_1, \dots, X_n are independent random vectors with common distribution function F and a d -dimensional parameter θ of interest is associated with F . Suppose X_1, \dots, X_n are independent random vectors with common distribution function F . write $X = (X_1, \dots, X_n)$. If there exists a function $T_n(X)$, $X \in \mathbb{R}^n$ such that

$$\sqrt{n} (T_n(X) - \theta) \xrightarrow{d} N(0, \Sigma), \tag{2.1}$$

for some positive definite matrix Σ . The so-called jackknife pseudo-sample $\{t_1, \dots, t_n\}$ is defined as

$$t_i = nT_n(X) - (n-1)T_{n-1}(X_{(-i)}), \quad i = 1, \dots, n.$$

As in Tukey [6], one expects that t_i 's are approximately independent. This motivates us to apply the standard empirical likelihood method to the jackknife sample $\{t_1, \dots, t_n\}$ for constructing empirical likelihood confidence regions for θ . The jackknife empirical likelihood function is defined as

$$L^j(\theta) = \sup \left\{ \prod_{i=1}^n (np_i) : \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i t_i = \theta, p_i \geq 0, \dots, p_n \geq 0 \right\}.$$

It follows from the Lagrange multiplier technique that the above maximization is achieved at $p_i = n^{-1} \{1 + \lambda'(t_i - \theta)\}^{-1}$ and the log empirical likelihood ratio $l^j(\alpha) = -2 \log L^j(\theta)$ is given by

$$l^j(\theta) = \sum_{i=1}^n \log \{1 + \lambda'(t_i - \theta)\}, \tag{2.2}$$

Where λ satisfies

$$\frac{1}{n} \sum_{i=1}^n \frac{t_i - \theta}{1 + \lambda'(t_i - \theta)} = 0.$$

We say that Wilks' theorem holds if $l'(\theta_0)$ converges in distribution to a chi-square distribution

With d degrees of freedom, where θ_0 is the true value of the parameter θ . Wilks' theorem can be

used to construct $1 - \alpha$ confidence regions for θ

$$I_{1-\alpha} = \left\{ \theta : l'(\theta) \leq \chi_d^2(\alpha) \right\},$$

Where $\chi_d^2(\alpha)$ is the α level critical value for a chi-square distribution with d degrees of freedom, or

we can reject the null hypothesis $H_0 : \theta = \theta_0$ if $l'(\theta_0) > \chi_d^2(\alpha)$.

Two-sample JEL

Suppose that X_1, \dots, X_{n_1} and Y_1, \dots, Y_{n_2} are two independent samples with distribution function F and G , respectively. Write $X = (X_1, \dots, X_{n_1})$, $Y = (Y_1, \dots, Y_{n_2})$, and $n = n_1 + n_2$. Assume that the parameter θ we are interested in can be estimated by $T_{n_1, n_2}(X, Y)$, and (2.1) holds for this estimate. The two-sample jackknife pseudo-sample is defined as

$$t_i = nT_{n_1, n_2}(X, Y) - (n-1)T_{n_1-1, n_2}(X_{(-i)}, Y), \quad i = 1, \dots, n_1$$

and

$$t_{n_1+j} = nT_{n_1, n_2}(X, Y) - (n-1)T_{n_1, n_2-1}(X, Y_{(-j)}), \quad j = 1, \dots, n_2.$$

Then, based on the jackknife sample $\{t_1, \dots, t_n\}$, we define the log empirical likelihood ratio $l'(\alpha)$ as in (2.2). When the parameters of interest are functionals of one distribution function F or two distribution functions F and G , the estimates of the parameters can naturally defined as functionals of the empirical distributions for F and G .

Smoothed JEL

If the functionals of the empirical distributions are used to estimate the parameters of interest and the asymptotic covariance matrices in (2.1) depend on local properties such as the density functions of the underlying distributions, the sample covariance matrices based on resulting jackknife samples are usually not consistent estimates of the asymptotic covariances in the central limit theorem (2.1), and in this case, Wilks' theorem will fail to hold. One should consider smoothed JEL methods. The so-called smoothed JEL uses the functionals of the smoothed empirical distributions for F and G to generate jackknife samples, which can overcome the problem of variance inconsistency.

Applications

In addition to the work by Jing et al. [1] for one- and two-sample U-statistics, there are a few applications of the standard JEL methods in statistics. Wang et al. [7] propose JEL based test for equality of two high dimensional means, Zhang et al. [8]

investigate the population mean with ranked set samples. JEL based confidence intervals for the mean absolute deviation and difference of two Gini indices are studied by Zhao et al. [9] and Wang & Zhao [10], respectively. The JEL methods have been applied in many problems in insurance and actuarial sciences. For instance, JEL-based confidence intervals for copulas are proposed by Peng et al. [11], and Wang et al. [12]. Wilks' theorem for JEL methods for Spearman's rho and a class of risk measures are proved by Wang & Peng [13] and Peng, et al. [14], respectively. For the tail copulas and difference of two quantiles, Wilks' theorem are shown to be valid for smoothed JEL methods by Peng & Qi [15] and Yang & Zhao [16,17].

In diagnostic medicine, the accuracy of a diagnostic test in discriminating diseased patients from non-diseased ones is measured by the receiver operating characteristic (ROC) curve when the response of a test is continuous. Let F and G be the distribution functions of the diseased and non-diseased populations, respectively. Then the ROC curve can be written as $\theta = \theta(t) = 1 - F(G^{-1}(1-t))$, $0 < t < 1$, which is a functional of two populations. Gong et al. [18] propose smoothed JEL method and construct the confidence intervals for $\theta(t)$, and Yang and Zhao [19] extend the method under the setup of missing data. Adimari & Chiogna [20] and An & Zhao [21] employ JEL methods for confidence intervals for partial areas under ROC curves and the difference of two volumes under ROC surfaces, respectively. JEL based confidence regions for quantities in sensitivity and specificity for continuous-scale diagnostic tests are investigated in Wang & Qin [22]. To construct confidence regions when many nuisance parameters are present, a profile empirical likelihood method has to be employed, which is computationally costly in general. Li et al. [23] and Peng [24] propose JEL method for estimating equations to avoid heavy computational burden. Further, Zhang et al. [25] propose jackknife-blockwise empirical likelihood methods for data under dependence.

JEL methods have also been applied in regression models. JEL based confidence intervals for the regression parameters in accelerated failure time model with censored observations and in linear transformation models under right censoring are studied by Bouadoumou et al. [26] and Yang et al. [27], respectively. JEL based confidence intervals for the error variances in linear model and in partially linear varying-coefficient errors-in-variables models are investigated by Lin et al. [28] and Liu and Liang [29]. The JEL based intervals of mean with regression imputation is considered by Zhong & Chen [30]. In summary, a common feature for these applications is that the JEL methods maintain the advantages of empirical likelihood methods over normal approximation methods and perform very well for small samples [31]. Computationally JEL methods are easy and straight forward even for complicated statistical problems.

References

- Jing BY, Yuan J, Zhou W (2009) Jackknife empirical likelihood. Journal of the American Statistical Association 104(487): 1224-1232.

2. Thomas DR, Grunkemeier GL (1975) Confidence interval estimation of survival probabilities for censored data. *Journal of the American Statistical Association* 70: 865-871.
3. Owen AB (1988) Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* 75(2): 237-249.
4. Owen AB (1990) Empirical likelihood ratio confidence regions. *Annals of Statistics* 18: 90-120.
5. Shao J, Tu D (1995) *The Jackknife and Bootstrap*, Springer, UK.
6. Tukey JW (1958) Bias and confidence in not-quite large samples. *The Annals of Statistics* 29: 614-623.
7. Wang R, Peng L and Qi Y (2013) Jackknife empirical likelihood test for equality of two high dimensional means. *Statistica Sinica* 23: 667-690.
8. Zhang Z, Liu T, Zhang B (2016) Jackknife empirical likelihood inferences for the population mean with ranked set samples. *Statistics & Probability Letters* 108: 16-22.
9. Zhao Y, Meng X, Yang H (2015) Jackknife empirical likelihood inference for the mean absolute deviation. *Computational Statistics & Data Analysis* 91: 92-101.
10. Wang D, Zhao Y (2016) Jackknife empirical likelihood for comparing two Gini indices. *Canadian Journal of Statistics* DOI: 10.1002/cjs.11275.
11. Peng L, Qi Y, Van Keilegom I (2012) Jackknife empirical likelihood method for copulas. *Test* 21: 74-92.
12. Wang R, Peng L, Yang J (2013) Jackknife empirical likelihood for parametric copulas. *Scandinavian Actuarial Journal* 1: 325-339.
13. Wang R, Peng L (2011) Jackknife empirical likelihood intervals for Spearman's rho. *North American Actuarial Journal* 15: 475-486.
14. Peng L, Qi Y, Wang R, Yang J (2012) Jackknife empirical likelihood methods for some risk measures and related quantities. *Insurance: Mathematics and Economics* 51: 142-150.
15. Peng L, Qi Y (2010) Smoothed jackknife empirical likelihood method for tail copulas. *Test* 19: 514-536.
16. Yang H, Zhao Y (2017) Smoothed jackknife empirical likelihood for the difference of two quantiles. *Ann Inst Statist Math* 69: 1059-1073.
17. Yang H, Zhao Y (2018) Smoothed jackknife empirical likelihood for the one-sample difference of quantiles. *Computational Statistics & Data Analysis* 120: 58-69.
18. Gong Y, Peng L, Qi Y (2010) Smoothed jackknife empirical likelihood method for ROC curve. *Journal of Multivariate Analysis* 101: 520-531.
19. Yang H, Zhao Y (2015) Smoothed jackknife empirical likelihood inference for ROC curves with missing data. *Journal of Multivariate Analysis* 140: 123-138.
20. Adimari G, Chiogna M (2012) Jackknife empirical likelihood based confidence intervals for partial areas under ROC curves. *Statistica Sinica* 22: 1457-1477.
21. An Y, Zhao Y (2017) Jackknife empirical likelihood for the difference of two volumes under ROC surfaces. *Annals of the Institute of Statistical Mathematics* Pp. 1-8.
22. Wang B, Qin G (2016) Jackknife empirical likelihood confidence regions for the evaluation of continuous-scale diagnostic tests with verification bias. *Stat Methods Med Res* 25(5): 2120-2137.
23. Li M, Peng L, Qi Y (2011) Reduce computation in profile empirical likelihood method. *The Canadian Journal of Statistics* 38: 370-384.
24. Peng L (2012) Approximate jackknife empirical likelihood method for estimating equations. *Canadian Journal of Statistics* 40: 110-123.
25. Zhang R, Peng L, Qi Y (2012) Jackknife-blockwise empirical likelihood methods under dependence. *Journal of Multivariate Analysis* 104: 56-72.
26. Bouadoumou M, Zhao Y, Lu Y (2015) Jackknife empirical likelihood for the accelerated failure time model with censored data. *Communications in Statistics-Simulation and Computation* 44(7): 1818-1832.
27. Yang H, Liu S, Zhao Y (2016) Jackknife empirical likelihood for linear transformation models with right censoring. *Annals of the Institute of Statistical Mathematics* 68(5): 1095-1109.
28. Lin HL, Li Z, Wang D, Zhao Y (2017) Jackknife empirical likelihood for the error variance in linear models. *Journal of Nonparametric Statistics* 29(2): 151-166.
29. Liu AA, Liang HY (2017) Jackknife empirical likelihood of error variance in partially linear varying-coefficient errors-in-variables models. *Statistical Papers* 58: 95-122.
30. Zhong PS, Chen S (2014) Jackknife empirical likelihood inference with regression imputation and survey data. *Journal of Multivariate Analysis* 129: 193-205.
31. Hall P, La Scala B (1990) Methodology and algorithms of empirical likelihood. *International Statistical Review* 58: 109-127.



This work is licensed under Creative Commons Attribution 4.0 License
DOI: [10.19080/BBOJ.2018.07.555708](https://doi.org/10.19080/BBOJ.2018.07.555708)

Your next submission with Juniper Publishers will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats
(Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission

<https://juniperpublishers.com/online-submission.php>