Statistics and Its Interface Volume 11 (2018) 709–719

Jackknife empirical likelihood for the skewness and kurtosis

Yichuan Zhao, Anna Moss, Hanfang Yang*, and Yan Zhang

Coefficients of skewness and kurtosis provide convenient measures for describing the shape of a distribution based on a sample of independent observations. In this paper, we propose jackknife empirical likelihood (JEL) confidence intervals for the skewness and kurtosis coefficients, proving that the limiting distribution of the JEL ratio is a standard chi-squared distribution, and conduct an extensive simulation study comparing JEL with bootstrap methods. Compared with bootstrap methods, the JEL-based confidence intervals perform well in simulations with data from normal, t, γ, log-normal, and uniform distributions. We also illustrate the application of our proposed JEL methods using data from the behavioral risk factor surveillance system, an annual US health survey.

Keywords and phrases: Empirical likelihood, Jackknife pseudo-values, Kurtosis and skewness, Nonparametric confidence intervals, U-statistics.

1. INTRODUCTION

Measures of skewness and kurtosis are useful tools for examining characteristics of a given distribution and testing normality (Wilcox (1990) and D'Agostino, Belanger, & D'Agostino (1990)), which is often a basic assumption in applying statistical methods. Inference methods have mainly focused on testing hypotheses of normality, and few studies have investigated confidence interval (CI) construction, which may be due to the non-linearity and non-normality of estimators of the skewness and kurtosis. While bootstrapping is possible, at least one informal study by Wright & Herrington (2011) indicated that bootstrap methods for the skewness and kurtosis had very poor coverage. To address this gap in the literature, we develop another promising CI method, the jackknife empirical likelihood (JEL), for the skewness and kurtosis.

The JEL introduced by Jing, Yuan, & Zhou (2009) is a nonparametric method for constructing confidence intervals (CIs), which may be well-suited to the skewness and kurtosis. The JEL combines the bias reduction of the jackknife with the data-driven shape of the empirical likelihood (EL) by applying the EL method for the mean of jackknife pseudo-values. Jing, Yuan, & Zhou (2009) originally developed the JEL method for U-statistics to address computational difficulties when applying the EL method directly. Subsequently, the JEL has been successfully applied to other nonlinear statistics faced with the same computational challenges of the EL method such as Gong, Peng, & Qi (2010), Yang & Zhao (2013, 2015) and Zhao, Meng & Yang (2015). Therefore, we apply the JEL method to CI construction for the skewness and kurtosis. We then conduct a comprehensive simulation study evaluating its performance along with bootstrap methods.

The paper is organized as follows. In Section 2, we present the skewness and kurtosis and derive the original JEL, the adjusted JEL (AJEL) and the extended JEL (EJEL) methods. In Section 3, we conduct simulation studies to compare coverage probability and average length of the three JEL methods and three bootstrap methods using data from normal, gamma, t, log-normal, and uniform distributions. In Section 4, we illustrate the proposed method by applying the JEL and bootstrap methods to data from the behavioral risk factor surveillance system (BRFSS), a national health survey conducted annually in the United States. In Section 5, we present a discussion and conclusion. Proof of the theorems are put in the Appendix.

2. INFERENCE PROCEDURE

2.1 Skewness and kurtosis

Skewness and kurtosis of a random variable X, respectively denoted as g_3 and g_4, are the third and fourth standardized moments defined as

$$g_3 = \frac{\mu_3}{\sigma^3} = \frac{E\left[X-\mu\right]^3}{\left(E\left[X-\mu\right]^2\right)^{3/2}}$$

and

$$g_4 = \frac{\mu_4}{\sigma^4} = \frac{E\left[X-\mu\right]^4}{\left(E\left[X-\mu\right]^2\right)^{2}}$$

where μ_3 and μ_4 are the third and fourth central moments, E is the expectation operator, μ is the mean and σ is the standard deviation of X.

From a sample of n i.i.d. observations, $x_1, \ldots, x_n$, g_3 and g_4 are commonly estimated from $\hat{g}_3$ and $\hat{g}_4$ proposed by

*Corresponding author.

Cramer (1946), using the corresponding sample moments m_3 and m_4 as well as m_2 as follows

$$\hat{g}_3 = \frac{m_3}{m_2^{3/2}} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3/n}{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2/n\right)^{3/2}}$$

and

$$\hat{g}_4 = \frac{m_4}{m_2^2} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4/n}{\left(\sum_{i=1}^{n}(x_i - \bar{x})^2/n\right)^2}$$

where $\bar{x}$ and m_2 are the sample mean and sample standard deviation, respectively.

Under the assumption of normality, $\hat{g}_3$ and $\hat{g}_4$ have variances equal to

$$Var\,(\hat{g}_3) = \frac{6n(n-1)}{(n-2)(n+1)(n+3)}$$

and

$$Var\,(\hat{g}_4) = \frac{24n(n-1)^2}{(n-3)(n-2)(n+5)(n+3)}.$$

which may be used to construct approximately $(1-\alpha)$-level confidence intervals (CIs) as

$$\mathcal{R}_{1-\alpha} = \left\{\tilde{g}_r : \frac{(\hat{g}_r - \tilde{g}_r)^2}{Var(\hat{g}_r)} \le \chi_1^{2,1-\alpha}\right\};\ r = 3, 4,$$

for the skewness ($r = 3$) and kurtosis ($r = 4$), respectively, and $\chi_1^{2,1-\alpha}$ is the $1-\alpha$ quantile from the χ_1^2 distribution. However, in practice, the normality assumption is often implausible for highly nonlinear estimators like $\hat{g}_3$ and $\hat{g}_4$, and the resulting NA confidence intervals (CIs) can have poor coverage even for large sample sizes. Therefore, we develop CIs based on nonparametric methods, which may improve coverage probabilities and have superior performance compared with CIs based on NA methods.

2.2 Jackknife empirical likelihood

The jackknife empirical likelihood (Jing, Yuan, & Zhou (2009)) combines the jackknife proposed by Quenouille (1956) and the empirical likelihood (EL) method of Owen (1990). Like the jackknife, the statistics of interest is the sample mean of the jackknife pseudo-values based on a consistent estimator of a parameter (e.g., skewness or kurtosis). The EL method for the mean is applied to the jackknife pseudo-values to construct confidence intervals.

The JEL may be applied to skewness and kurtosis as follows. Recall that $x_1, \ldots, x_n$ are a sample of n independent observations from an unknown distribution of parameter g_r with consistent estimator $\hat{g}_r$ ($r = 3, 4$). Define the jackknife pseudo-values $V_{r,1}, \ldots, V_{r,n}$ as

$$V_{r,i} = n\,\hat{g}_r - (n-1)\hat{g}_{r,i}\,,\ i = 1, \ldots, n,$$

where $\hat{g}_{r,i}$ is computed on the sample of $n-1$ observations with the i^{th} observation removed, which is similar to $\hat{g}_r$. The jackknife estimator of g_r is

$$\hat{g}_{r,J} = \frac{1}{n}\sum_{i=1}^{n} V_{r,i},$$

which is the average of the jackknife pseudo-values.

Following Jing, Yuan, & Zhou (2009), we define the jackknife empirical likelihood ratio $R(g_r)$ at a true value of g_r as

$$R(g_r) = \max\left\{\prod_{i=1}^{n} nw_i : \sum_{i=1}^{n} w_i V_{r,i} = g_r, \sum_{i=1}^{n} w_i = 1\right\}$$

where $w_i \ge 0$ for $i = 1, \ldots, n$. Using Lagrange multipliers, we obtain

$$w_i = \frac{1}{n}\frac{1}{1 + \lambda(V_{r,i} - g_r)}$$

with λ satisfying

$$\frac{1}{n}\sum_{i=1}^{n}\frac{V_{r,i} - g_r}{1 + \lambda(V_{r,i} - g_r)} = 0.$$

Plugging w_i's into the log-likelihood ratio equation, we obtain

$$\log R(g_r) = -\sum_{i=1}^{n}\log\left\{1 + \lambda(V_{r,i} - g_r)\right\}.$$

For the true value g_r, we have the following result

Theorem 1. Under regularity conditions, $\mu_6 = E|x-\mu|^6 < \infty$, $\mu_8 = E|X-\mu|^8 < \infty$ and $\mu_2 = E|X-\mu|^2 > 0$,

$$-2\log R(g_r) \xrightarrow{d} \chi_1^2,$$

where χ_1^2 is a chi-square random variable with one degree of freedom.

Using Theorem 1, we can construct JEL confidence intervals as

$$\mathcal{R}_{JEL,1-\alpha} = \left\{\tilde{g}_r : -2\log R(\tilde{g}_r) \le \chi_1^{2,1-\alpha}\right\},$$

where $\chi_1^{2,1-\alpha}$ is the $1-\alpha$ quantile from the χ_1^2 distribution.

2.3 Adjusted jackknife empirical likelihood

The adjusted empirical likelihood of Chen, Variyath, & Abraham (2008) expands the domain of the EL method by adding a single artificial data point to the set of sample values. The data point is constructed so that the true parameter value will be located in the interior of the convex hull of the augmented data set, and thus eliminating the so-called "empty set" problem. As a benefit, this adjustment can improve coverage probabilities of the EL method particularly in small samples. For pseudo-values $V_{r,1}, \ldots, V_{r,n}$, we perform the adjusted JEL (AJEL) by defining a function $g_{r,i}^{ad} = g^{ad}(V_{r,i}, g_r) = V_{r,i} - g_r$, for $i = 1, \ldots, n$, where g_r is the skewness (r=3) or kurtosis (r=4) parameter. Then an artificial point $g_{r,n+1}^{ad}$ is constructed as $-a_n \bar{g}_{r_n}$ with $a_n = \max\{1, \log(n)/2\}$, and $\bar{g}_{r_n} = \sum_{i=1}^n g_{r,i}^{ad}/n = \hat{g}_{r,J} - g_r$.

We define the adjusted jackknife empirical likelihood ratio, $W(g_r)$, at the true value g_r as

$$W(g_r) = \max\left\{\prod_{i=1}^{n+1}(n+1)w_i : \sum_{i=1}^{n+1} w_i g_{r,i}^{ad} = 0, \sum_{i=1}^{n+1} w_i = 1\right\}$$

where $w_i \geq 0$ for $i = 1, \ldots, n+1$. Using Lagrange multipliers, we obtain

$$w_i = \frac{1}{(n+1)} \frac{1}{1 + \lambda g_{r,i}^{ad}}$$

with λ satisfying

$$\frac{1}{(n+1)} \sum_{i=1}^{n+1} \frac{g_{r,i}^{ad}}{1 + \lambda g_{r,i}^{ad}} = 0.$$

We plug the w_i's into the adjusted jackknife empirical log-likelihood ratio to obtain $\log W(g_r)$ as

$$\log W(g_r) = -\sum_{i=1}^{n+1} \log\left\{1 + \lambda g_{r,i}^{ad}\right\}.$$

For the true value g_r, we establish the following Wilks' theorem,

Theorem 2. Under regularity conditions, $\mu_6 < \infty$, $\mu_8 < \infty$ and $\mu_2 > 0$,

$$-2\log W(g_r) \xrightarrow{d} \chi_1^2.$$

Using Theorem 2, we can construct JEL confidence intervals as

$$\mathcal{R}_{AJEL,1-\alpha} = \left\{\tilde{g}_r : -2\log W(\tilde{g}_r) \leq \chi_1^{2,1-\alpha}\right\};\ r = 3, 4.$$

2.4 Extended jackknife empirical likelihood

We perform the extended JEL (EJEL) using extended empirical likelihood method of Tsao & Wu (2013). This method also eliminates the empty set problem and achieves the domain expansion by mapping the EL domain Θ_n onto the full parameter space, which is $\mathbb{R}$ for both skewness and kurtosis. The extended EL method can also improve the coverage of EL-based confidence intervals. By applying it to the JEL, we define a bijective mapping $h_n : \Theta_n \to \mathbb{R}$ as

$$h_n(g_r) = \hat{g}_r + \gamma(l(g_r), n)(g_r - \hat{g}_r);\ g_r \in \Theta_n$$

where $\gamma(l(g_r), n) = 1 + l/(2n)$ with $l = l(g_r) = -2\log R(g_r)$. In this case, g_r is the parameter value (i.e., the skewness or kurtosis) in the EL domain Θ_n. For parameter $g_r \in \mathbb{R}$, we perform the JEL with $h_n^{-1} \in \Theta_n$ like Tsao and Wu (2014) to obtain the extended jackknife empirical log-likelihood ratio $l^*(g_r)$ as

$$l^*(g_r) = -2\log R(h_n^{-1}).$$

For the true value $g_r \in \mathbb{R}$, we have the following theorem,

Theorem 3. Under regularity conditions, $\mu_6 < \infty$, $\mu_8 < \infty$ and $\mu_2 > 0$,

$$l^*(g_r) \xrightarrow{d} \chi_1^2.$$

Using Theorem 3, we can construct EJEL confidence intervals as follows,

$$\mathcal{R}_{EJEL,1-\alpha} = \left\{\tilde{g}_r : l^*(\tilde{g}_r) \leq \chi_1^{2,1-\alpha}\right\};\ r = 3, 4.$$

3. NUMERICAL STUDIES

We conducted a simulation study to evaluate the performance of the three JEL CI methods: the original JEL, AJEL, and EJEL, for the skewness and kurtosis. We compared performance of the JEL methods to three bootstrap CI methods: original, percentile, and bias-corrected accelerated (BCa) bootstrap. We generated samples from normal, $N(\mu = 0,\ \sigma = 1)$, gamma ($\alpha = 4,\ \beta = 1$), t_{10}, log-normal ($\mu = 0,\ \sigma = 0.25$), and uniform (0, 1) distributions of samples sizes n = 30, 60, 120 and 240. For each setting we calculated the coverage probability and average interval length for confidence levels of 1-α = 0.90, 0.95 and 0.99 from 5000 simulations. For bootstrap methods, CIs were calculated from 400 bootstrap samples taken with replacement. Tables 1 - 10 display the results for coverage probability and average length of CIs for the skewness (Tables 1-5) and the kurtosis (Tables 6-10) for all five distributions.

For the skewness, all methods performed as expected with increasing coverage and decreasing length with the increasing sample size. For the normal data, JEL and bootstrap methods had similar coverage reaching nominal levels for the largest sample size of $n = 240$. For t distribution, only the BC_a bootstrap achieved nominal level coverage rates and only for the largest sample size $n = 240$. For asymmetrical distributions, gamma and log-normal, JEL performed about the same as bootstrap, but all methods demonstrated under-coverage for all sample sizes and confidence levels. For the uniform distribution, both JEL and bootstrapping achieved nominal coverage levels for all sample sizes. With the exception of the uniform distribution, bootstrap methods tended to produce slightly shorter intervals than the JEL methods. With uniformly distributed data, the original JEL method produced shortest intervals.

For the kurtosis, both JEL and bootstrapping consistently under-covered true values for all distributions except the uniform data. Under the uniform sampling, both JEL and bootstrap methods achieved nominal coverage for all confidence levels and sample sizes. With other distributions, under-coverage was generally worse with bootstrapping and particularly severe for t-distributed data. Also notable for

Table 1. Coverage probability (average interval length) for the skewness with $N(0,1)$

n	JEL	AJEL	EJEL	Bootstrap	Percentile	BCa
$\alpha = 0.90$						
30	0.82 (1.41)	0.84 (1.51)	0.83 (1.48)	0.90 (1.15)	0.90 (1.14)	0.84 (1.15)
60	0.85 (1.05)	0.87 (1.08)	0.86 (1.07)	0.89 (0.86)	0.90 (0.86)	0.87 (0.86)
120	0.85 (0.76)	0.86 (0.78)	0.86 (0.77)	0.88 (0.66)	0.88 (0.65)	0.87 (0.66)
240	0.88 (0.54)	0.88 (0.55)	0.90 (0.54)	0.89 (0.48)	0.88 (0.48)	0.88 (0.49)
$\alpha = 0.95$						
30	0.89 (1.68)	0.91 (1.79)	0.91 (1.78)	0.95 (1.35)	0.95 (1.35)	0.91 (1.37)
60	0.91 (1.27)	0.92 (1.32)	0.91 (1.31)	0.94 (1.03)	0.94 (1.02)	0.92 (1.03)
120	0.92 (0.91)	0.92 (0.93)	0.92 (0.93)	0.93 (0.77)	0.93 (0.76)	0.93 (0.77)
240	0.93 (0.66)	0.94 (0.67)	0.94 (0.66)	0.94 (0.57)	0.94 (0.57)	0.94 (0.58)
$\alpha = 0.99$						
30	0.97 (2.35)	0.98 (2.53)	0.99 (2.61)	0.99 (1.80)	0.99 (1.87)	0.98 (1.87)
60	0.98 (1.77)	0.98 (1.84)	0.98 (1.87)	0.99 (1.35)	0.99 (1.37)	0.98 (1.38)
120	0.98 (1.30)	0.98 (1.33)	0.98 (1.34)	0.99 (1.03)	0.99 (1.04)	0.99 (1.04)
240	0.98 (0.92)	0.98 (0.93)	0.98 (0.93)	0.99 (0.76)	0.99 (0.77)	0.99 (0.77)

Table 2. Coverage probability (average interval length) for the skewness with t_{10}

n	JEL	AJEL	EJEL	Bootstrap	Percentile	BCa
$\alpha = 0.90$						
30	0.69 (2.08)	0.73 (2.22)	0.72 (2.17)	0.86 (1.46)	0.87 (1.43)	0.78 (1.42)
60	0.72 (1.84)	0.74 (1.91)	0.73 (1.88)	0.84 (1.28)	0.85 (1.24)	0.83 (1.25)
120	0.76 (1.53)	0.77 (1.57)	0.76 (1.55)	0.84 (1.11)	0.86 (1.09)	0.86 (1.10)
240	0.79 (1.25)	0.80 (1.27)	0.80 (1.26)	0.87 (0.96)	0.86 (0.94)	0.88 (0.95)
$\alpha = 0.95$						
30	0.78 (2.55)	0.82 (2.72)	0.81 (2.71)	0.93 (1.74)	0.93 (1.70)	0.87 (1.69)
60	0.80 (2.27)	0.81 (2.35)	0.81 (2.34)	0.91 (1.52)	0.92 (1.46)	0.90 (1.46)
120	0.83 (1.90)	0.84 (1.94)	0.84 (1.93)	0.91 (1.32)	0.92 (1.27)	0.93 (1.30)
240	0.86 (1.55)	0.86 (1.57)	0.86 (1.56)	0.93 (1.14)	0.93 (1.11)	0.94 (1.11)
$\alpha = 0.99$						
30	0.90 (3.54)	0.93 (3.81)	0.94 (3.94)	0.98 (2.28)	0.99 (2.26)	0.96 (2.25)
60	0.90 (3.21)	0.91 (3.33)	0.92 (3.38)	0.97 (2.00)	0.98 (1.88)	0.98 (1.87)
120	0.92 (2.70)	0.93 (2.76)	0.93 (2.78)	0.97 (1.74)	0.98 (1.63)	0.98 (1.61)
240	0.94 (2.21)	0.94 (2.23)	0.94 (2.24)	0.98 (1.50)	0.98 (1.42)	0.99 (1.40)

Table 3. Coverage probability (average interval length) for the skewness with $gamma\ (4,1)$

n	JEL	AJEL	EJEL	Bootstrap	Percentile	BCa
$\alpha = 0.90$						
30	0.70 (1.60)	0.73 (1.71)	0.72 (1.67)	0.69 (1.26)	0.69 (1.24)	0.71 (1.22)
60	0.74 (1.32)	0.75 (1.37)	0.75 (1.35)	0.72 (1.01)	0.72 (0.99)	0.74 (0.99)
120	0.77 (0.99)	0.78 (1.01)	0.77 (1.00)	0.73 (0.79)	0.73 (0.78)	0.75 (0.78)
240	0.80 (0.77)	0.81 (0.78)	0.81 (0.78)	0.78 (0.64)	0.78 (0.63)	0.80 (0.64)
$\alpha = 0.95$						
30	0.76 (1.98)	0.79 (2.12)	0.79 (2.11)	0.76 (1.52)	0.76 (1.49)	0.79 (1.50)
60	0.81 (1.57)	0.83 (1.63)	0.82 (1.62)	0.77 (1.19)	0.77 (1.16)	0.79 (1.16)
120	0.84 (1.25)	0.85 (1.27)	0.85 (1.27)	0.80 (0.96)	0.81 (0.93)	0.82 (0.94)
240	0.87 (0.94)	0.88 (0.95)	0.88 (0.94)	0.83 (0.76)	0.83 (0.74)	0.84 (0.75)
$\alpha = 0.99$						
30	0.87 (2.65)	0.89 (2.86)	0.90 (2.95)	0.87 (1.98)	0.89 (2.01)	0.89 (2.00)
60	0.91 (2.23)	0.92 (2.32)	0.93 (2.36)	0.89 (1.59)	0.89 (1.56)	0.90 (1.54)
120	0.93 (1.69)	0.93 (1.73)	0.93 (1.74)	0.89 (1.24)	0.89 (1.21)	0.90 (1.20)
240	0.95 (1.31)	0.95 (1.33)	0.95 (1.33)	0.91 (1.00)	0.91 (0.98)	0.91 (0.97)

Table 4. Coverage probability (average interval length) for the skewness with $lognormal\,(0, 0.25)$

n	JEL	AJEL	EJEL	Bootstrap	Percentile	BCa
$\alpha = 0.90$						
30	0.72 (1.59)	0.75 (1.69)	0.74 (1.66)	0.74 (1.23)	0.74 (1.23)	0.73 (1.22)
60	0.75 (1.26)	0.77 (1.30)	0.76 (1.29)	0.74 (0.98)	0.74 (0.97)	0.75 (0.96)
120	0.79 (1.00)	0.80 (1.02)	0.80 (1.01)	0.78 (0.80)	0.78 (0.78)	0.79 (0.79)
240	0.82 (0.75)	0.82 (0.76)	0.82 (0.75)	0.80 (0.63)	0.80 (0.62)	0.82 (0.63)
$\alpha = 0.95$						
30	0.80 (1.93)	0.83 (2.07)	0.82 (2.06)	0.82 (1.49)	0.82 (1.48)	0.82 (1.47)
60	0.83 (1.54)	0.84 (1.59)	0.84 (1.59)	0.81 (1.17)	0.81 (1.15)	0.82 (1.14)
120	0.86 (1.22)	0.87 (1.25)	0.87 (1.24)	0.84 (0.95)	0.84 (0.93)	0.84 (0.93)
240	0.88 (0.92)	0.88 (0.93)	0.88 (0.93)	0.86 (0.75)	0.86 (0.74)	0.86 (0.74)
$\alpha = 0.99$						
30	0.90 (2.65)	0.92 (2.85)	0.93 (2.94)	0.91 (1.95)	0.91 (2.00)	0.91 (1.99)
60	0.92 (2.13)	0.93 (2.21)	0.93 (2.25)	0.91 (1.53)	0.90 (1.51)	0.91 (1.51)
120	0.95 (1.70)	0.95 (1.74)	0.95 (1.75)	0.92 (1.24)	0.92 (1.21)	0.92 (1.20)
240	0.95 (1.28)	0.95 (1.29)	0.95 (1.29)	0.93 (0.98)	0.93 (0.96)	0.93 (0.96)

Table 5. Coverage probability (average interval length) for the skewness with $U(0, 1)$

n	JEL	AJEL	EJEL	Bootstrap	Percentile	BCa
$\alpha = 0.90$						
30	0.92 (0.88)	0.94 (0.94)	0.93 (0.92)	0.92 (0.92)	0.91 (0.91)	0.89 (0.92)
60	0.91 (0.61)	0.92 (0.63)	0.92 (0.63)	0.91 (0.63)	0.91 (0.62)	0.90 (0.63)
120	0.90 (0.43)	0.91 (0.44)	0.91 (0.44)	0.91 (0.43)	0.90 (0.44)	0.90 (0.44)
240	0.90 (0.30)	0.91 (0.31)	0.94 (0.31)	0.90 (0.31)	0.90 (0.31)	0.90 (0.31)
$\alpha = 0.95$						
30	0.96 (1.04)	0.98 (1.11)	0.98 (1.11)	0.97 (1.09)	0.96 (1.11)	0.94 (1.11)
60	0.96 (0.73)	0.96 (0.75)	0.96 (0.75)	0.96 (0.75)	0.95 (0.75)	0.94 (0.76)
120	0.95 (0.51)	0.96 (0.52)	0.96 (0.52)	0.95 (0.52)	0.95 (0.53)	0.94 (0.53)
240	0.95 (0.36)	0.96 (0.37)	0.97 (0.37)	0.95 (0.37)	0.96 (0.37)	0.95 (0.37)
$\alpha = 0.99$						
30	1.00 (1.35)	1.00 (1.46)	1.00 (1.50)	0.99 (1.43)	0.99 (1.53)	0.99 (1.54)
60	0.99 (0.95)	1.00 (0.99)	1.00 (1.00)	0.99 (0.98)	0.99 (1.02)	0.98 (1.03)
120	0.99 (0.67)	0.99 (0.69)	0.99 (0.69)	0.99 (0.68)	0.99 (0.71)	0.99 (0.71)
240	0.99 (0.48)	0.99 (0.48)	0.99 (0.48)	0.99 (0.48)	0.99 (0.50)	0.99 (0.50)

Table 6. Coverage probability (average interval length) for the kurtosis with $N(0, 1)$

n	JEL	AJEL	EJEL	Bootstrap	Percentile	BCa
$\alpha = 0.90$						
30	0.72 (2.40)	0.75 (2.56)	0.74 (2.51)	0.78 (2.02)	0.82 (1.93)	0.81 (2.34)
60	0.78 (1.79)	0.79 (1.86)	0.78 (1.83)	0.78 (1.48)	0.80 (1.45)	0.82 (1.66)
120	0.81 (1.38)	0.82 (1.41)	0.81 (1.39)	0.80 (1.16)	0.81 (1.14)	0.83 (1.28)
240	0.84 (1.00)	0.85 (1.02)	0.84 (1.01)	0.83 (0.88)	0.83 (0.87)	0.86 (0.97)
$\alpha = 0.95$						
30	0.78 (2.82)	0.80 (3.01)	0.80 (3.00)	0.83 (2.36)	0.89 (2.31)	0.88 (2.75)
60	0.84 (2.18)	0.85 (2.26)	0.85 (2.25)	0.83 (1.77)	0.87 (1.73)	0.88 (2.02)
120	0.87 (1.61)	0.87 (1.64)	0.87 (1.63)	0.84 (1.34)	0.87 (1.31)	0.89 (1.50)
240	0.90 (1.21)	0.91 (1.22)	0.90 (1.22)	0.88 (1.04)	0.89 (1.02)	0.91 (1.17)
$\alpha = 0.99$						
30	0.86 (3.95)	0.89 (4.26)	0.89 (4.39)	0.92 (3.16)	0.98 (3.32)	0.97 (3.62)
60	0.91 (2.99)	0.92 (3.11)	0.92 (3.15)	0.91 (2.32)	0.96 (2.35)	0.96 (2.57)
120	0.94 (2.31)	0.94 (2.36)	0.94 (2.37)	0.93 (1.81)	0.95 (1.80)	0.96 (1.98)
240	0.96 (1.68)	0.96 (1.70)	0.96 (1.71)	0.94 (1.38)	0.96 (1.37)	0.97 (1.51)

Table 7. Coverage probability (average interval length) for the kurtosis with t_{10}

n	JEL	AJEL	EJEL	Bootstrap	Percentile	BCa
$\alpha = 0.90$						
30	0.41 (4.19)	0.42 (4.46)	0.42 (4.38)	0.32 (3.00)	0.35 (2.86)	0.34 (3.64)
60	0.47 (4.36)	0.48 (4.52)	0.47 (4.46)	0.36 (2.89)	0.37 (2.74)	0.26 (3.33)
120	0.52 (4.18)	0.52 (4.27)	0.52 (4.23)	0.41 (2.80)	0.43 (2.63)	0.23 (3.11)
240	0.56 (3.94)	0.57 (3.98)	0.57 (3.96)	0.48 (2.79)	0.49 (2.61)	0.14 (3.08)
$\alpha = 0.95$						
30	0.47 (5.12)	0.48 (5.46)	0.48 (5.44)	0.38 (3.58)	0.44 (3.45)	0.49 (4.29)
60	0.53 (5.35)	0.54 (5.56)	0.55 (5.52)	0.41 (3.46)	0.44 (3.22)	0.43 (3.98)
120	0.59 (5.14)	0.59 (5.25)	0.59 (5.22)	0.47 (3.33)	0.48 (3.04)	0.38 (3.71)
240	0.65 (4.84)	0.65 (4.90)	0.65 (4.88)	0.53 (3.30)	0.54 (3.00)	0.30 (3.65)
$\alpha = 0.99$						
30	0.56 (7.06)	0.58 (7.60)	0.59 (7.84)	0.48 (4.71)	0.64 (4.70)	0.70 (5.17)
60	0.63 (7.48)	0.65 (7.78)	0.65 (7.90)	0.50 (4.53)	0.58 (4.19)	0.63 (4.65)
120	0.70 (7.23)	0.71 (7.39)	0.71 (7.43)	0.55 (4.38)	0.58 (3.84)	0.61 (4.26)
240	0.77 (6.82)	0.77 (6.90)	0.77 (6.91)	0.61 (4.34)	0.64 (3.72)	0.58 (4.13)

Table 8. Coverage probability (average interval length) for the kurtosis with $gamma\,(4, 1)$

n	JEL	AJEL	EJEL	Bootstrap	Percentile	BCa
$\alpha = 0.90$						
30	0.46 (4.58)	0.47 (4.88)	0.47 (4.79)	0.45 (3.30)	0.49 (3.10)	0.57 (3.89)
60	0.57 (4.58)	0.59 (4.75)	0.58 (4.68)	0.53 (3.17)	0.56 (3.01)	0.61 (3.70)
120	0.63 (3.84)	0.63 (3.92)	0.63 (3.88)	0.57 (2.77)	0.58 (2.66)	0.64 (3.13)
240	0.70 (3.33)	0.70 (3.37)	0.70 (3.35)	0.64 (2.53)	0.65 (2.45)	0.70 (2.87)
$\alpha = 0.95$						
30	0.52 (5.78)	0.54 (6.18)	0.54 (6.15)	0.51 (4.00)	0.58 (3.79)	0.66 (4.66)
60	0.64 (5.41)	0.65 (5.62)	0.65 (5.59)	0.57 (3.68)	0.62 (3.46)	0.69 (4.28)
120	0.71 (4.93)	0.72 (5.03)	0.71 (5.01)	0.63 (3.40)	0.67 (3.17)	0.74 (3.84)
240	0.77 (4.02)	0.77 (4.07)	0.77 (4.06)	0.69 (2.98)	0.71 (2.81)	0.77 (3.43)
$\alpha = 0.99$						
30	0.59 (7.71)	0.61 (8.30)	0.62 (8.56)	0.62 (5.17)	0.76 (5.15)	0.79 (5.56)
60	0.71 (7.85)	0.73 (8.16)	0.73 (8.29)	0.67 (4.16)	0.77 (4.71)	0.80 (5.20)
120	0.80 (6.62)	0.80 (6.75)	0.81 (6.80)	0.71 (4.96)	0.78 (4.02)	0.82 (4.44)
240	0.86 (5.75)	0.86 (5.82)	0.86 (5.83)	0.78 (4.34)	0.81 (3.64)	0.86 (4.06)

Table 9. Coverage probability (average interval length) for the kurtosis with $lognormal\,(0, 0.25)$

n	JEL	AJEL	EJEL	Bootstrap	Percentile	BCa
$\alpha = 0.90$						
30	0.49 (3.91)	0.51 (4.17)	0.50 (4.09)	0.48 (2.88)	0.51 (2.73)	0.60 (3.43)
60	0.56 (3.66)	0.57 (3.80)	0.57 (3.75)	0.51 (2.60)	0.54 (2.48)	0.61 (3.02)
120	0.64 (3.38)	0.65 (3.45)	0.64 (3.42)	0.58 (2.43)	0.60 (2.32)	0.67 (2.73)
240	0.69 (2.83)	0.69 (2.87)	0.69 (2.85)	0.64 (2.16)	0.65 (2.09)	0.71 (2.46)
$\alpha = 0.95$						
30	0.54 (4.77)	0.57 (5.10)	0.56 (5.08)	0.54 (3.44)	0.62 (3.31)	0.69 (4.06)
60	0.63 (4.49)	0.64 (4.66)	0.64 (4.63)	0.57 (3.10)	0.62 (2.94)	0.69 (3.61)
120	0.71 (4.16)	0.72 (4.25)	0.72 (4.25)	0.64 (2.89)	0.67 (2.71)	0.74 (3.29)
240	0.76 (3.48)	0.77 (3.52)	0.77 (3.51)	0.69 (2.57)	0.72 (2.44)	0.78 (2.95)
$\alpha = 0.99$						
30	0.63 (6.57)	0.66 (7.07)	0.67 (7.29)	0.65 (4.52)	0.80 (4.56)	0.82 (4.95)
60	0.72 (6.25)	0.73 (6.50)	0.73 (6.60)	0.66 (4.07)	0.76 (3.89)	0.81 (4.32)
120	0.80 (5.83)	0.80 (5.96)	0.81 (5.99)	0.72 (3.80)	0.78 (3.47)	0.83 (3.87)
240	0.86 (4.89)	0.87 (4.95)	0.87 (4.96)	0.77 (3.38)	0.81 (3.09)	0.86 (3.45)

Table 10. Coverage probability (average interval length) for the kurtosis with $U(0,1)$

n	JEL	AJEL	EJEL	Bootstrap	Percentile	BCa
$\alpha = 0.90$						
30	0.91 (0.98)	0.91 (1.04)	0.91 (1.02)	0.99 (1.14)	0.96 (1.07)	0.97 (0.98)
60	0.92 (0.59)	0.93 (0.61)	0.92 (0.60)	0.96 (0.65)	0.93 (0.63)	0.96 (0.60)
120	0.91 (0.38)	0.92 (0.39)	0.91 (0.39)	0.94 (0.40)	0.92 (0.40)	0.96 (0.39)
240	0.91 (0.26)	0.91 (0.26)	0.94 (0.26)	0.92 (0.26)	0.91 (0.26)	0.96 (0.26)
$\alpha = 0.95$						
30	0.94 (1.16)	0.95 (1.24)	0.95 (1.24)	1.00 (1.35)	0.99 (1.34)	0.99 (1.20)
60	0.95 (0.73)	0.96 (0.73)	0.96 (0.73)	0.99 (0.77)	0.97 (0.77)	0.99 (0.72)
120	0.95 (0.46)	0.96 (0.47)	0.96 (0.46)	0.97 (0.48)	0.96 (0.48)	0.98 (0.47)
240	0.96 (0.31)	0.96 (0.31)	0.96 (0.31)	0.96 (0.32)	0.96 (0.32)	0.98 (0.31)
$\alpha = 0.99$						
30	0.98 (1.52)	0.98 (1.65)	0.98 (1.69)	1.00 (1.78)	1.00 (1.95)	1.00 (1.71)
60	0.98 (0.92)	0.99 (0.96)	0.99 (0.97)	1.00 (1.01)	1.00 (1.07)	1.00 (1.00)
120	0.99 (0.60)	0.99 (0.61)	0.99 (0.62)	1.00 (0.63)	0.99 (0.66)	1.00 (0.64)
240	0.99 (0.40)	0.99 (0.41)	0.99 (0.41)	0.99 (0.41)	0.99 (0.43)	1.00 (0.42)

Table 11. Interval estimates of coefficients of skewness and kurtosis for body weight distribution

	Skewness ($\hat{g}_3 = 0.72$)		Kurtosis ($\hat{g}_4 = 2.84$)	
Method	95% CI	Length	95% CI	Length
JEL	0.43, 1.02	0.59	2.13, 3.57	1.44
AJEL	0.43, 1.03	0.60	2.11, 3.58	1.47
EJEL	0.42, 1.02	0.60	2.11, 3.58	1.47
Bootstrap	0.43, 0.99	0.56	2.16, 3.51	1.35
Percetile Bootstrap	0.45, 1.00	0.55	2.26, 3.65	1.39
BCa Bootstrap	0.49, 1.03	0.54	2.32, 3.70	1.38

t-distributed data was the BC_a bootstrap with decreasing the coverage (with increasing sample size) for all confidence levels. As with the skewness, bootstrap produced shorter intervals compared with the JEL methods except for uniform data, where the original JEL produced shortest intervals.

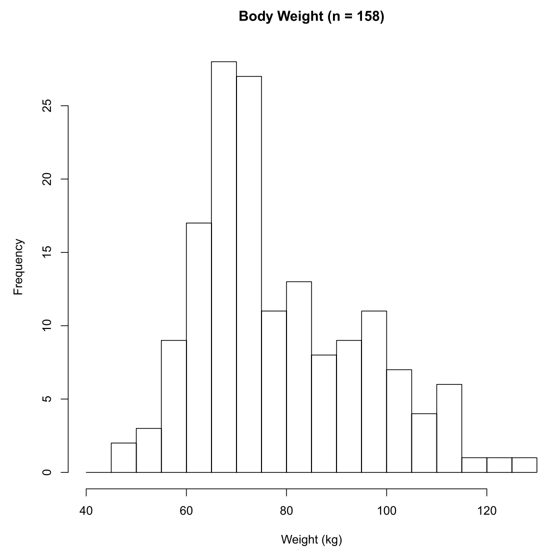

Figure 1. Histogram of self-reported body weight.

4. DATA ANALYSIS

The BRFSS is an annual US survey collecting data on health risks and chronic conditions from adults in all US states and territories. The BRFSS is a national dataset, which provides state-level health and risk factor data for all US states and territories. Both US national and state health agencies use BRFSS data to monitor health trends in the US population. For example, the BRFSS has been used to estimate the growing obesity epidemic in the US over the past few decades. Consequently, an abundance of health studies have analyzed the body weight, which is typically assumed to have a normal distribution in statistical models and tests. For this data analysis, we characterized the distribution of body weight in kilograms (kg) among women, aged 18 to 25, from Georgia, US ($n = 160$). A histogram of the distribution of the variable for self-reported weight in kilograms (kg) indicates a slightly skewed and highly peaked distribution (Figure 1). Point estimates for the skewness and kurtosis were $\hat{g}_3 = 0.72$ and $\hat{g}_4 = 2.84$, respectively. Results in Table 5 show that JEL and bootstrap methods produce similar interval estimates for skewness and kurtosis coefficients; the JEL methods produce slightly wider intervals compared with bootstrap methods. The CIs also confirm observations from the histogram. For the skewness, all CIs exclude the value of zero indicating that body weight among young women in Georgia may be more skewed than a normal distribution. For the kurtosis, all CIs include the value of 3 indicating that the central peakedness of the distribution of body weight is consistent with that of a normal distribution. The CIs provide evidence that further models and analysis assuming normality or symmetry in the body weight distribution may not be appropriate in this population. But beyond the informal hypothesis test of normality, the CIs also provide meaningful and robust interval estimates for these population parameters of body weight among young women in Georgia, US.

5. CONCLUSION

Coefficients of skewness and kurtosis provide information about moment characteristics of a probability distribution; however, accurate estimation of these quantities can be difficult due to highly nonlinear behavior of the existing estimators, and very few studies have evaluated nonparametric methods for interval estimation. To address these drawbacks, we developed JEL methods, proving the limiting chi-square distribution for the JEL ratio, and compared performance of JEL-based methods with standard bootstrap methods through extensive simulations. Our simulation studies showed that the JEL methods can attain superior coverage probabilities compared with bootstrap methods particularly for small sample sizes. However, in some cases the bootstrap had better coverage than the JEL, and coverage probabilities for both JEL and bootstrap methods were substantially lower than nominal levels for small sample size.

For the kurtosis, all methods suffer from under-coverage for distributions with long tails including the normal distribution. The poor performance is likely due to the difficulty sampling low probability observations in the tails. Therefore, small to moderate sample sizes may not include a sufficient number of tail points for $\hat{g}_4$ to obtain an adequate estimate of kurtosis and thus compromise JEL and bootstrap methods, which are centered around these estimates. Optimization of point estimators through weighting or other strategies would improve both JEL and bootstrap methods.

Future research could theoretically focus on improving the performance of nonparametric inference methods for skewness and kurtosis parameters. For example, the kernel method may be incorporated into the JEL framework. Moreover, it is possible to estimate the skewness and kurtosis by JEL methods under the multivariate p-dimensional distribution or even the high-dimensional setting. The extended empirical likelihood method similar to Tsao and Wu (2014) could be applied to increase the accuracy in the second order. In terms of application, our approach could be extended to a general JEL theorem of other statistics constructed from the skewness and kurtosis, such as Jarque-Bera test statistics, which is to test whether sample skewness and kurtosis fit the normal distribution assumption.

APPENDIX A. APPENDIX: PROOFS OF THEOREMS

Let $x_1, \ldots, x_n$ be a sample of independent and identically-distributed observations. Define $\bar{x} = 1/n\sum_{j=1}^n x_j$ and $\hat{\sigma}^2 = 1/n\sum_{j=1}^n (x_j - \bar{x})^2$ and $\hat{\mu}_r = 1/n\sum_{i=1}^n (x_i - \bar{x})^r$, $r = 3, 4$. Their deleting one empirical estimators are represented as follows,

$$\begin{aligned}\bar{x}_i &= \frac{1}{n-1}\sum_{j=1 j\neq i}^{n} x_j,\ i = 1, \ldots, n,\\ \hat{\sigma}_i^2 &= \frac{1}{n-1}\sum_{j=1 j\neq i}^{n} (x_j - \bar{x}_i)^2,\ i = 1, \ldots, n,\\ \hat{\mu}_{r,i} &= \frac{1}{n-1}\sum_{j=1 j\neq i}^{n} (x_j - \bar{x}_i)^r,\ i = 1, \ldots, n;\ r = 3, 4.\end{aligned}$$

We consider the function $f_r(z, y) = zy^{-r/2}$ for $z, y \in \mathbb{R}$ and $r = 3, 4$. Skewness and kurtosis denoted in Section 2.2 can be summarized as $g_r = f(\mu_r, \sigma^2)$, where $\mu_r = E[(X - \mu)^r]$, $\sigma^2 = E[(X - \mu)^2]$, and $\mu = E[X]$. The empirical estimator based on the full sample denoted in Section 2.2 can be represented as $\hat{g}_r = f_r(\hat{\mu}_r, \hat{\sigma}^2)$. The jackknife pseudo-sample based on the empirical estimators of skewness and kurtosis is $V_{r,i} = nf_r(\hat{\mu}_r, \hat{\sigma}^2) - (n-1)f_r(\hat{\mu}_{r,i}, \hat{\sigma}_i^2)$ for $i = 1, \ldots, n$, where $f_r(\hat{\mu}_{r,i}, \hat{\sigma}_i^2) = \hat{g}_{r,i}$, $\hat{\mu}_{r,i}$ and $\hat{\sigma}_i^2$ are estimated by original data except the ith observation. The corresponding jackknife estimator defined in Section 2.2 is $\hat{g}_{r,J} = 1/n\sum_{i=1}^n V_{r,i}$. The jackknife sample variance is

$$\hat{\sigma}^2_{g_r,J} = \frac{1}{n-1}\sum_{i=1}^{n}(V_{r,i} - \hat{g}_{r,J})^2$$

and the jackknife variance with given true value is

$$\hat{\sigma}^2_{g_r,T} = \frac{1}{n}\sum_{i=1}^{n}(V_{r,i} - g_r)^2.$$

Lemma 1. Under regularity conditions,

$$\frac{\sqrt{n}(\hat{g}_{r,J} - g_r)}{\hat{\sigma}_{g_r,T}} \xrightarrow{d} N(0, 1),$$

where $r = 3, 4$.

Proof of Lemma 1. From Theorem 1 of Heffernan (1997) and Lee (1990) the r^{th} central moments may be represented as U-statistics. Abassi (2009) found U-statistics representations of $\hat{\sigma}^2, \hat{\mu}_3$ and $\hat{\mu}_4$ which we denote as $\hat{U}_2$, $\hat{U}_3$ and $\hat{U}_4$ as follows.

$$\begin{aligned}&\hat{U}_2 = \frac{n}{n-1}\hat{\sigma}^2,\\ &\hat{U}_3 = \frac{n^2}{(n-1)(n-2)}\hat{\mu}_3,\\ &\hat{U}_4 = \frac{n^3}{(n-1)(n-2)(n-3)}\hat{\mu}_4 + \frac{-2n^2+3n}{(n-1)(n-2)(n-3)}\overline{x^4}\\ &\quad + \frac{8n^2-12n}{(n-1)(n-2)(n-3)}\overline{x^3}\overline{x} + \frac{-6n^2+9n}{(n-1)(n-2)(n-3)}\overline{x^2}^2.\end{aligned}$$

Hence, we have

$$\begin{aligned}\hat{\sigma}^2 &= \hat{U}_2 + \hat{\sigma}^2 O_p(n^{-1}),\\ \hat{\mu}_3 &= \hat{U}_3 + \hat{\mu}_3 O_p(n^{-1}),\\ \hat{\mu}_4 &= \hat{U}_4 + \left(\hat{\mu}_4 + \overline{x^4} + \overline{x^3}\,\overline{x} + \overline{x^2}^2\right) O_p(n^{-1}). \quad (1)\end{aligned}$$

From the properties of $f_r(\cdot)$, Taylor series expansion, and Equation (1), we have

$$\begin{aligned}\hat{g}_3 &= f_3(\hat{\mu}_3, \hat{\sigma}^2) = f_3(\hat{U}_3, \hat{U}_2) + (\hat{\sigma}^2 + \hat{\mu}_3) O_p(n^{-1}),\\ \hat{g}_4 &= f_4(\hat{\mu}_4, \hat{\sigma}^2)\\ &= f_4(\hat{U}_4, \hat{U}_2) + \left(\hat{\mu}_4 + \overline{x^4} + \overline{x^3}\,\overline{x} + \overline{x^2}^2\right) O_p(n^{-1}). \quad (2)\end{aligned}$$

Denote $\hat{U}_{r,i}$ and $\hat{U}_{2,i}$ which is obtained by the deleting-i data set. From Equation (2), the jackknife estimator $\hat{g}_{r,J}$ is as follows,

$$\begin{aligned}&\hat{g}_{r,J}\\ =\ & \hat{g}_r + \frac{n-1}{n}\sum_{i=1}^{n}(\hat{g}_r - \hat{g}_{r,i})\\ =\ & f_r(\hat{U}_r, \hat{U}_2) + \frac{n-1}{n}\sum_{i=1}^{n}(\hat{g}_r - \hat{g}_{r,i}) + O_p(n^{-1})\\ =\ & f_r(\hat{U}_r, \hat{U}_2) + \frac{n-1}{n}\sum_{i=1}^{n}\left(f_r(\hat{U}_r, \hat{U}_2) - f_r(\hat{U}_{r,i}, \hat{U}_{2,i})\right)\\ +\ & \frac{n-1}{n}\sum_{i=1}^{n}\left(\hat{g}_r - \hat{g}_{r,i} - f_r(\hat{U}_r, \hat{U}_2) + f_r(\hat{U}_{r,i}, \hat{U}_{2,i})\right)\\ +\ & O_p(n^{-1})\\ =\ & f_r(\hat{U}_r, \hat{U}_2) + \frac{n-1}{n}\sum_{i=1}^{n}\left(f_r(\hat{U}_r, \hat{U}_2) - f_r(\hat{U}_{r,i}, \hat{U}_{2,i})\right)\\ (3) +\ & O_p(n^{-1}).\end{aligned}$$

Denote kernels of $\hat{U}_2$, $\hat{U}_3$, and $\hat{U}_4$ as $\hat{\phi}_2$, $\hat{\phi}_3$, and $\hat{\phi}_4$ with expectations η_2, η_3, and η_4. We know that

(4)
$$g_r = f_r(\mu_r, \sigma^2) = f_r(E(\hat{\mu}_r) + O_p(n^{-1}), E(\hat{\sigma}^2) + O_p(n^{-1})).$$

$\hat{U}_2$, $\hat{U}_3$, and $\hat{U}_4$ are unbiased estimates of η_2, η_3, and η_4. From Equation (1) we have that

$$\begin{aligned}g_r &= f_r\left(E(\hat{U}_r) + O_p(n^{-1}), E(\hat{U}_2) + O_p(n^{-1})\right)\\ &= f_r\left(\eta_r + O_p(n^{-1}), \eta_2 + O_p(n^{-1})\right)\\ &= f_r(\eta_r, \eta_2) + O_p(n^{-1}). \quad (5)\end{aligned}$$

Based on (3) and (5), one has that

$$\sqrt{n}(\hat{g}_{r,J} - g_r) = \sqrt{n}\left(f_J - f_r(\eta_r, \eta_2)\right) + o_p(1), \quad (6)$$

where $f_J = f_r(\hat{U}_r, \hat{U}_2) + (n-1)/n\sum_{i=1}^{n}\left(f_r(\hat{U}_r, \hat{U}_2) - f_r(\hat{U}_{r,i}, \hat{U}_{2,i})\right)$. The skewness and kurtosis are real-valued functions f_r of two arguments U_r and U_2. f_r has continuous first partial derivatives and bounded second partial derivatives in the neighborhood of (η_3, η_2) or (η_4, η_2) for $r = 3$ or 4, respectively. By Theorem 8 from Arvesen (1969), we obtain

$$\sqrt{n}\left(f_J - f_r(\eta_r, \eta_2)\right) \xrightarrow{d} N(0, \sigma^2\left(f_r(\eta_r, \eta_2)\right);\ r = 3, 4, \quad (7)$$

where $\sigma^2\left(f_r(\eta_r, \eta_2)\right)$ is defined as in Theorem 8 of Arvesen (1969) as

$$\sigma^2\left(f_r(\eta_r, \eta_2)\right) = \dot{f}_r \Omega \dot{f}_r^T$$

with

$$\Omega = \begin{bmatrix}\mathbf{cov}(\hat{\phi}_r, \hat{\phi}_r) & \mathbf{cov}(\hat{\phi}_r, \hat{\phi}_2)\\ \mathbf{cov}(\hat{\phi}_r, \hat{\phi}_2) & \mathbf{cov}(\hat{\phi}_2, \hat{\phi}_2)\end{bmatrix}$$

and

$$\dot{f}_r = \begin{bmatrix} r\frac{\partial f_r(\eta_r, \eta_2)}{\partial \eta_r} & 2\frac{\partial f_r(\eta_r, \eta_2)}{\partial \eta_2}\end{bmatrix}.$$

From Equation (6), one has

$$\sqrt{n}\left(\hat{g}_{r,J} - g_r\right) \xrightarrow{d} N(0, \sigma^2\left(f_r(\eta_r, \eta_2)\right),\ r = 3, 4. \quad (8)$$

In addition,

$$\begin{aligned}V_{r,i} &= nf_r(\hat{\mu}_r, \hat{\sigma}^2) - (n-1)f_r(\hat{\mu}_{r,i}, \hat{\sigma}_i^2)\\ &= nf_r(\hat{U}_r, \hat{U}_2) - (n-1)f_r(\hat{U}_{r,i}, \hat{U}_{2,i})\\ &+ n(\hat{\mu}_r - \hat{U}_r)\frac{\partial f_r}{\partial \eta_r} + n(\hat{\sigma}^2 - \hat{U}_2)\frac{\partial f_r}{\partial \eta_2}\\ &- (n-1)(\hat{\mu}_{r,i} - \hat{U}_{r,i})\frac{\partial f_r}{\partial \eta_r}\\ &- (n-1)(\hat{\sigma}_i^2 - \hat{U}_{2,i})\frac{\partial f_r}{\partial \eta_2} + o_p(1)\\ &= nf_r(\hat{U}_r, \hat{U}_2) - (n-1)f_r(\hat{U}_{r,i}, \hat{U}_{2,i})\\ &+ \left(n(\hat{\mu}_r - \hat{U}_r) - (n-1)(\hat{\mu}_{r,i} - \hat{U}_{r,i})\right)\frac{\partial f_r}{\partial \eta_r}\\ &+ \left(n(\sigma^2 - \hat{U}_2) - (n-1)(\hat{\sigma}_i^2 - \hat{U}_{2,i}\right)\frac{\partial f_r}{\partial \eta_2} + o_p(1).\end{aligned}$$

By Equation (1) and similar results in jackknife samples

$$V_{r,i} = nf_r(\hat{U}_r, \hat{U}_2) - (n-1)f_r(\hat{U}_{r,i}, \hat{U}_{2,i}) + o_p(1). \quad (9)$$

Incorporating Equations (3) and (9), the jackknife variance is as follows,

$$\begin{aligned}\hat{\sigma}^2_{g_r,J} &= \frac{1}{n-1}\sum_{i=1}^{n}(V_{r,i} - \hat{g}_{r,J})^2\\ &= \frac{1}{n-1}\sum_{i=1}^{n}\left(nf_r(\hat{U}_r, \hat{U}_2) - (n-1)f_r(\hat{U}_{r,i}, \hat{U}_{2,i}) - f_J\right)^2\\ &+ o_p(1). \quad (10)\end{aligned}$$

The first term of Equation (10) is the jackknife variance estimator of a function of several U-statistics. By the convergence of functions of U-statistics in Theorem 9 of Arvesen (1969), we have that

$$\hat{\sigma}^2_{g_r,J} \xrightarrow{p} \sigma^2(f_r(\eta_r,\eta_2)),\ r=3,4. \tag{11}$$

Then, one has that

$$\begin{aligned}
\hat{\sigma}^2_{g_r,J} &= \frac{1}{n-1}\sum_{i=1}^{n}(V_{r,i}-\hat{g}_{r,J})^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}(V_{r,i}-g_r-\hat{g}_{r,J}+g_r)^2 \\
&= \frac{1}{n-1}\sum_{i=1}^{n}(V_{r,i}-g_r)^2 \\
&+ \frac{1}{n-1}\sum_{i=1}^{n}(\hat{g}_{r,J}-g_r)(\hat{g}_{r,J}-g_r-2V_{r,i}+2g_r) \\
&= \frac{1}{n-1}\sum_{i=1}^{n}(V_{r,i}-g_r)^2+O_p(\hat{g}_{r,J}-g_r).
\end{aligned}$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}(V_{r,i}-g_r)^2+o_p(1). \tag{12}$$

From Equations (8) and (12) and Slutsky's Theorem, Lemma 1 holds. □

Proof of Theorem 1. Based on Lemma 1, we finish the proof of Theorem 1 by the standard arguments of Owen (1990). □

Proof of Theorem 2. Following Chen, Variyath & Abraham (2008) and Lin, Li, Wang & Zhao (2017), we obtain Theorem 2 from Theorem 1. First define $Z_n = \max_{1\le i\le n}|V_{r,i}-g_r|$. Denote $\bar{g}_{r,n} = 1/n\sum_{i=1}^{n} g^{ad}_{r,i}$. We have

$$\begin{aligned}
0 &= \frac{1}{n+1}\sum_{i=1}^{n+1}\frac{g^{ad}_{r,i}}{1+\lambda g^{ad}_{r,i}} \\
&= \frac{1}{n+1}\sum_{i=1}^{n+1} g^{ad}_{r,i} + \frac{\lambda}{n+1}\sum_{i=1}^{n+1}\frac{(g^{ad}_{r,i})^2}{1+\lambda g^{ad}_{r,i}} \\
&\le \bar{g}_{r,n}\left(1-\frac{a_n}{n}\right) - \frac{\lambda}{n(1+\lambda Z_n)}\sum_{i=1}^{n}(g^{ad}_{r,i})^2 \\
&= \bar{g}_{r,n}\left(1-\frac{a_n}{n}\right) - \frac{\lambda}{n(1+\lambda Z_n)}\sum_{i=1}^{n}(V_{r,i}-g_r)^2 \\
&= \bar{g}_{r,n} - \frac{\lambda\hat{\sigma}^2_{g_r,T}}{1+\lambda Z_n} + O_p(n^{-\frac{3}{2}}a_n).
\end{aligned}$$

Then, $\lambda = O_p(n^{-1/2})$ and $a_n = o_p(n)$ and $\lambda = (\hat{\sigma}^2_{g_r,T})^{-1}\bar{g}_{r,n} + o_p(n^{-1/2})$. The adjusted JEL ratio $-2\log W(g_r)$ is as follows,

$$\begin{aligned}
-2\log W(g_r) &= 2\sum_{i=1}^{n+1}\log(1+\lambda g^{ad}_{r,i}) \\
&= 2\sum_{i=1}^{n+1}\left(\lambda g^{ad}_{r,i} - \frac{\lambda^2(g^{ad}_{r,i})^2}{2}\right) + o_p(1) \\
&= n(\hat{\sigma}^2_{g_r,T})^{-1}\bar{g}^2_{r,n} + o_p(1) \\
&= \left(\frac{1}{n}\sum_{i=1}^{n}(V_{r,i}-g_r)^2\right)^{-1} n\bar{g}^2_{r,n} + o_p(1) \\
&\xrightarrow{d} \chi^2_1.
\end{aligned}$$

□

Proof of Theorem 3. Following Tsao and Wu (2013, 2014) and Lin, Li, Wang & Zhao (2017), we prove Theorem 3 for the parameter skewness g_3 or kurtosis g_4. Define the generalized inverse function h_n^{-1} like Tsao and Wu (2013, 2014). We have the extended EL ratio $l^*(g_r)$ as $l^*(g_r) = l(h_n^{-1})$. For any $\tilde{g}_r \in \{\tilde{g}_r : \|\tilde{g}_r - g_r\| \le n^{-1/2}\}$, we apply the Taylor expansion to obtain

$$\begin{aligned}
l(\tilde{g}_r) &= l(\tilde{g}_r + g_r - g_r) \\
&= l(g_r) + \frac{\mathrm{d}l(g_r)}{\mathrm{d}g_r}(\tilde{g}_r - g_r) + O_p(1).
\end{aligned}$$

It implies that $l(\tilde{g}_r) = O_p(1)$. We denote $h_n^{-1}(g_r) = \hat{g}_r$. Following Tsao and Wu (2014), one has $\hat{g}_r - g_r = O_p(n^{-3/2})$. By the Taylor expansion,

$$\begin{aligned}
l^*(g_r) &= l(\hat{g}_r) \\
&= l(\hat{g}_r + g_r - g_r) \\
&= l(g_r) + \frac{\mathrm{d}l(g_r)}{\mathrm{d}g_r}(\hat{g}_r - g_r) + O_p(n^{-1}).
\end{aligned}$$

Thus, the $l^*(g_r)$ has the same limiting distribution as the $l(g_r)$. Hence, we prove Theorem 3. □

ACKNOWLEDGEMENTS

Yichuan Zhao is supported by the National Science Foundation(NSF) Grant (DMS-1613176) and the National Security Agency (NSA) Grant (H98230-12-1-0209). Hanfang Yang is supported by National Natural Science Foundation of China (No. 11501567). The authors would like to thank the editor-in-chief, the associate editor and two referees for their insightful suggestions and useful comments, which improved the current version of the paper substantially.

Received 6 February 2017

REFERENCES

[1] Abbasi, N. (2009). Asymptotic distribution of coefficients of skewness and kurtosis. *Journal of Mathematics and Statistics*, 5(4), 365–368.

[2] Adimari, G., & Chiogna, M. (2012). Jackknife empirical likelihood based confidence intervals for partial areas under the ROC curves. *Statistica Sinica*, 22, 1457–1477. MR3027095

[3] Ankarali, H., Canan-Yazici, A., & Ankarali, S. (2009). A bootstrap confidence interval for skewness and kurtosis and properties of t-test in small samples from normal distribution. *Medical Journal of Trakya Universitesi Tip Fakultesi Dergisi*, 26(4), 297–305.

[4] Arvesen, J. (1969). Asymptotic distribution of coefficients of skewness and kurtosis. *Annals of Mathematical Statistics*, 40(6), 2076–2100.

[5] Balanda, K. P., & MacGillivray, H. (1998). Kurtosis: a critical review. *The American Statistician*, 42(2), 111–119.

[6] Carpenter, J., & Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? a practical guide for medical statisticians. *Statistics in Medicine*, 19(9), 1141–1164.

[7] Chen, J., Variyath, A., & Abraham, B. (2008). Adjusted empirical likelihood and its properties. *Journal of Computational and Graphical Statistics*, 17, 426–443. MR2439967

[8] Cramer, H. (1946). *Mathematical Methods of Statistics*, Princeton University Press, Princeton. MR0016588

[9] D'Agostino, R. B., Belanger, A., & D'Agostino R. (1990). A suggestion for using powerful and informative tests of normality. *The American Statistician*, 44(4), 316–321.

[10] Efron, B. & Tibrishirani, R. (1993). *An Introduction to the Bootstrap*, 2nd ed., Chapman & Hall/CRC Press, Boco Raton. MR1270903

[11] Gong, Y., Peng, L., & Qi, Y. (2010). Smoothed jackknife empirical likelihood method for ROC curve. *Journal of Multivariate Analysis*, 101, 1520–1531.

[12] Heffernan, P. (1997). Unbiased estimation of central moments using U-statistics. *Journal of the Royal Statistical Society. Series B*, 59(4), 861–863.

[13] Jing, B., Li, Z., Qin, J., Zhou, W. (2012). Jackknife empirical likelihood method for case-control studies with gene-environment independence of controls. *Statistics and Its Interface*, 5, 293–302.

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

[15] Lin, H.-L., Li, Z., Wang, D. & Zhao, Y. (2017). Jackknife empirical likelihood for the error variance in linear models. *Journal of Nonparametric Statistics*, 29, 151–166. MR3635008

[16] Owen, A. (1990). Empirical likelihood ratio confidence regions. *Annals of Statistics*, 18(1), 90–120.

[17] Owen, A. (1988). Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, 75(2), 237–249.

[18] Owen, A. (2001). *Empirical Likelihood*, Chapman & Hall/CRC Press, Boco Raton.

[19] Quenouille, M. (1956). Notes on bias in estimation. *Biometrika*, 43(3), 353–360.

[20] Tsao, M. (2013). Extending the empirical likelihood by domain expansion. *The Canadian Journal of Statistics*, 41(2), 257–274.

[21] Tsao, M., & Wu, F., (2013). Empirical likelihood on the full parameter space. *The Annals of Statistics*, 41(4), 2176–2196.

[22] Tsao, M., & Wu, F., (2014). Extended empirical likelihood for general estimating equations. *Biometrika*, 1, 1306–1493.

[23] Tukey, J., (1957). Bias and confidence in not-quite large samples. *Annals of Mathematical Statistics*, 29(2), 614.

[24] Wang, H., & Zhao, Y. (2009). A comparison of some confidence intervals for the mean quality-adjusted lifetime with censored data. *Computational Statistics and Data Analysis*, 53(7), 2733–2739.

[25] Wilcox, R. (1990). Comparing the means of two independent groups. *Biometrical Journal*, 32(7), 771–780.

[26] Wright, D. & Herrington, D. (2011). Problematic standard errors and confidence intervals for skewness and kurtosis. *Behavioral Research*, 43, 8–17.

[27] Yang, H., & Zhao, Y. (2013). Smoothed jackknife empirical likelihood inference for the difference of ROC curves. *Journal of Multivariate Analysis*, 115, 270–284.

[28] Yang, H., & Zhao, Y. (2015). Smoothed jackknife empirical likelihood inference for ROC curves with missing data. *Journal of Multivariate Analysis*, 140, 123–138. MR3372558

[29] Zhao, Y., Meng, X. & Yang, H. (2015). Jackknife empirical likelihood inference for the mean absolute deviation. *Computational Statistics and Data Analysis*, 91, 92–101.

[30] Zhong, P., & Chen, S. (2014). Jackknife empirical likelihood inference with regression imputation and survey data. *Journal of Multivariate Analysis*, 129, 193–205.

Yichuan Zhao
Georgia State University
USA

Anna Moss
Georgia State University
USA

Hanfang Yang
Renmin University of China
China
E-mail address: hyang@ruc.edu.cn

Yan Zhang
LexisNexis Risk Solutions
USA