Ann Inst Stat Math (2017) 69:1059–1073
DOI 10.1007/s10463-016-0576-7

Smoothed jackknife empirical likelihood for the difference of two quantiles

Hanfang Yang[1] · **Yichuan Zhao**[2]

Received: 17 July 2015 / Revised: 8 May 2016 / Published online: 27 July 2016

Abstract In this paper, we propose a smoothed estimating equation for the difference of quantiles with two samples. Using the jackknife pseudo-sample technique for the estimating equation, we propose the jackknife empirical likelihood (JEL) ratio and establish the Wilk's theorem. Due to avoiding estimating link variables, the simulation studies demonstrate that JEL method has computational efficiency compared with traditional normal approximation method. We carry out a simulation study in terms of coverage probability and average length of the proposed confidence intervals. A real data set is used to illustrate the JEL procedure.

Keywords Difference of quantiles · Jackknife · Kernel smoothing · Two samples

1 Introduction

The quantile is an attractive statistical measure because of its versatility with robustness against the extreme value. It is defined by $F^{-1}(p) = \inf\{x \in \mathbb{R} : F(x) \geq p\}$, where $0 < p < 1$. Motivated by the excellent property of quantiles, the researchers from different fields proposed many useful methods as their fundamental tools, such as Value-at-Risk (VaR) in risk management, quantile regression in econometrics, etc. Csörgo (1987) established the theoretical foundation of a quantile estimation. To overcome the analytical difficulty in the discreteness of quantile function, Sheather and Marron (1990) introduced the kernel method to estimate an empirical quantile.

✉ Yichuan Zhao
yichuan@gsu.edu

1 Center for Applied Statistics and School of Statistics, Renmin University of China, Beijing 100872, China

2 Department of Mathematics and Statistics, Georgia State University, Atlanta 30303, GA, USA

Inheriting those favorable properties of the quantile, the difference of two quantiles is used as a supplementary measure for comparing two distributions besides the Q–Q plot. Kosorok (1999) developed two-sample quantile non-parametric tests for a variety of empirical distribution functions for censored data and repeated measures data. However, drawbacks of the asymptotic normal approximation method cannot be ignored, such as the difficulty in estimating variance and generating symmetric confidence interval.

The empirical likelihood (EL) is a popular nonparametric methodology. Owen (1988, 1990, 2001) seminally built up the framework of EL as a new philosophy of statistics. Chen and Hall (1993) studied the quantile estimation using the EL method. Qin and Lawless (1994) proposed EL for the general estimating equation incorporating side information. As an application, Baysal and Staum (2008) developed the EL inference for the value-at-risk and expected shortfall. However, using the traditional EL method for the difference of two quantiles, people need to solve multiple nonlinear estimating equations. In order to reduce heavy computational intensity of existing methods (see Zhou and Jing 2003), we construct the jackknife pseudo-samples and derive the EL based on those pseudo-samples, i.e., the jackknife EL proposed by Jing et al. (2009). It is a more feasible method, appreciating its computational efficiency in the small sample. Wang et al. (2013) develop the JEL method for the high dimensional two-mean problem, while the JEL method for a discrete estimation problem, such as the difference of two-sample quantiles $\theta(p) = F_1^{-1}(p) - F_2^{-1}(p)$, needs to be investigated further.

In order to apply the JEL method for the difference of two-sample quantiles $\theta(p), 0 < p < 1$, we need to propose a smoothed estimator rather than apply the classical discrete estimator directly (see Gong et al. 2010; Yang and Zhao 2013). We propose a smoothed nonparametric estimating equation for $\theta(p)$ to solve discrete problems. Using the smooth jackknife empirical likelihood method, confidence intervals for $\theta(p)$ are constructed. The Wilk's theorem is proved under suitable conditions. Moreover, we obtain the asymptotical results of $\theta(p)$, and demonstrate the computational accuracy using JEL in the small sample.

The rest of the paper is organized as follows. In Sect. 2, we show that the jackknife empirical log-likelihood ratio for the difference of two quantiles with two samples is an asymptotically Chi-squared distribution. To illustrate the performance of the proposed method, the coverage probability and average length of confidence intervals are reported in Sect. 3. In Sect. 4, we use a real data example to illustrate the proposed JEL method. Furthermore, we give a discussion in Sect. 5. The proofs are provided in the Appendix.

2 New procedure

Let X_1 and X_2 be independent random variables with distribution functions $F_1(x)$ and $F_2(x)$, respectively. The difference of quantiles at p can be written as $\theta(p) = F_1^{-1}(p) - F_2^{-1}(p)$, where $0 < p < 1$ and F_j^{-1} denotes quantile function of $F_j(x)$, $j = 1, 2$. Define $D\{\theta(p), p\} = F_1\{\theta(p) + F_2^{-1}(p)\}$. $D^{'}\{\theta(p), p\}$ is the first derivative of $D\{\theta(p), p\}$ with respect to p.

Let $X_{1,i}$, $i = 1, \ldots, m$ and $X_{2,i}$, $i = 1, \ldots, n$ be two samples independently observed from the distribution functions $F_1(x)$ and $F_2(x)$, respectively. The empirical estimators of distribution functions $F_1(x)$ and $F_2(x)$ are defined by $F_{m,1}(x) = 1/m \sum_{i=1}^{m} I(X_{1,i} \leq x)$, $F_{n,2}(x) = 1/n \sum_{i=1}^{n} I(X_{2,i} \leq x)$, respectively. Let $K(p)$ be the smooth distribution function which satisfies $K(p) = \int_{u \leq p} w(u)\mathrm{d}u$, where $w(u)$ is a symmetric density function. We propose a new smooth estimating equation for the difference of two quantiles

$$\Pi_{m,n}(p, \theta(p)) = \frac{1}{m}\sum_{j=1}^{m} K\left\{\frac{p - F_{n,2}(X_{1,j} - \theta(p))}{h}\right\} - p,$$

where $h = h(m) > 0$ is a bandwidth. In order to develop the JEL method, we need to overcome the difficulty for analyzing the discrete empirical quantile estimator because the jackknife variance estimator for a quantile estimator is not consistent (see Miller 1974). Thus, we employ implicit estimating equation motivated by $D\{\theta(p), p\}$ and kernel smoothing technique to avoid discrete non-parametric quantile estimators. We assume the following conditions.

C.1. $F_j(x)$ and their first derivatives $f_j(x)$, $j = 1, 2$ are continuous and bounded. Assume $f_1\{F_1^{-1}(p)\} > 0$, $f_2\{F_2^{-1}(p)\} > 0$ for a given $p \in (0, 1)$;
C.2. $D\{\theta(p), p\}$ and its first derivative $D^{'}\{\theta(p), p\}$ are bounded and continuous for the p in condition C.1.
C.3. $m/n \to r$, where $r > 0$; $h = h(m) \to 0$, $mh^2/\log m \to \infty$, $mh^4 \to 0$ as $m \to \infty$;
C.4. $w(u)$ is a symmetric density function with support $[-1, 1]$ and $w^{'}(u)$ is bounded, continuous for $u \in [-1, 1]$.

Condition C.1 is a regular condition in order to derive theorems. Because $D\{\theta(p), p\}$ and its derivative exist in the following theorems, we need its properties in condition C.2. Conditions C.3–C.4 are common conditions for both two-sample problems and kernel methods. We need asymptotic properties of this estimation equation $\Pi_{m,n}\{p, \theta(p)\}$ to obtain the consistency (Theorem 1) and the asymptotic normality (Theorem 2).

Theorem 1 *Assume conditions C.1–C.4 hold. We have*

$$\Pi_{m,n}(p, \theta(p)) \xrightarrow{\mathcal{P}} 0,$$

where $\theta(p)$ is a true value of the difference of two quantiles at p.

Theorem 2 *Under the conditions C.1–C.4, one has*

$$\sqrt{m+n}\,\Pi_{m,n}\{p, \theta(p)\} \xrightarrow{\mathfrak{D}} N\{0, \sigma^2(p)\},$$

where

$$\sigma^2(p) = \frac{1+r}{r}(1-p)p + (1+r)(1-p)pD^{'}\{\theta(p), p\}^2.$$

After we establish the asymptotic normality theorem for the kernel estimating equation of the difference of quantiles, we need the empirical estimator of the variance $\sigma^2(p)$ as Sheather and Marron (1990) and Csörgo (1987) did. However, from Theorem 2, due to the complicated variance $\sigma^2(p)$, normal approximation methods could not be widely applied in statistical inference regardless of using the smoothed estimator or an equation of the difference of two quantiles. The JEL method is interesting for the small sample problem because the JEL confidence interval is automatically adapted by the data set in an asymmetrical manner (Jing et al. 2009; Gong et al. 2010). First, we propose a procedure to generate jackknife pseudo-sample. Denote

$$\Pi_{m,n,i}\{p,\theta(p)\} = \begin{cases} \frac{1}{m-1} \sum\limits_{1\le j\le m, j\ne i} K\left(\dfrac{p - F_{n,2}\{X_{1,j} - \theta(p)\}}{h}\right) - p, & \text{if } \ 1 \le i \le m \\ \frac{1}{m} \sum\limits_{j=1}^{m} K\left(\dfrac{p - F_{n,2,m-i}\{X_{1,j} - \theta(p)\}}{h}\right) - p, & m+1 \le i \le m+n, \end{cases}$$

where

$$F_{n,2,m-i}(y) = \frac{1}{n-1} \sum_{1\le j\le n, j\ne i} I(X_{2,j} \le y), \quad i = m+1, \ldots, m+n.$$

The jackknife pseudo-sample is defined as

$$\hat{V}_i\{p,\theta(p)\} = (m+n)\Pi_{m,n}\{p,\theta(p)\} - (m+n-1)\Pi_{m,n,i}\{p,\theta(p)\}, i = 1, \ldots, m+n.$$

From the jackknife pseudo-sample, the empirical log-likelihood ratio at $\theta(p)$ is defined as

$$L\{p,\theta(p)\} = \frac{\sup\{\prod_{i=1}^{m+n} p_i : \sum_{i=1}^{m+n} p_i = 1, \sum_{i=1}^{m+n} p_i \hat{V}_i\{p,\theta(p)\} = 0, p_i > 0, i = 1, \ldots, m+n\}}{\sup\{\prod_{i=1}^{m+n} p_i, \sum_{i=1}^{m+n} p_i = 1, p_i > 0, i = 1, \ldots, m+n\}}.$$

Following the standard Lagrange multiplier method, we have

$$l(p,\theta(p)) = -2\log L\{p,\theta(p)\} = 2\sum_{i=1}^{m+n} \log\{1 + \lambda \hat{V}_i(p)\},$$

where Lagrange multiplier λ satisfies the equation

$$\sum_{i=1}^{m+n} \frac{\hat{V}_i\{p,\theta(p)\}}{1 + \lambda \hat{V}_i\{p,\theta(p)\}} = 0. \tag{1}$$

Assuming the conditions C.1–C.4, we establish the Wilks' theorem for $\theta(p)$ as follows.

Theorem 3 *Suppose the conditions C.1–C.4 hold. We have*

$$l\{p, \theta(p)\} \xrightarrow{\mathcal{D}} \chi_1^2,$$

where $\theta(p)$ is the true value of the difference of quantiles at the fixed $p \in (0, 1)$.

Thus, using smoothed JEL method, the asymptotic $100(1-\alpha)$ % confidence interval for the difference of two quantiles is as follows

$$I(p) = \left\{\tilde{\theta}(p) : l\{p, \tilde{\theta}(p)\} \le \chi_1^2(\alpha)\right\},$$

where $\chi_1^2(\alpha)$ is the upper α-quantile of χ_1^2.

3 Numerical studies

We carry out a comprehensive simulation study to illustrate our method. The coverage probabilities and average lengths are studied under various scenarios by the different values of p and the distribution functions F_1 and F_2. For scenario A, $F_1(x)$ is selected as a normal distribution function with mean 0.2 and standard deviation 0.5, and $F_2(x)$ is a $N(0, 0.5^2)$. For scenario B, X_1 is simulated from an exponential distribution function with parameter 1, and $F_2(x)$ is a $N(1, 0.5^2)$. For scenario C, $F_1(x)$ and $F_2(x)$ are two exponential distribution functions with parameter 1. In our simulation studies, we select $p = 0.4, 0.6$, and sample sizes m and n are chosen as (50, 70), (80, 60), (70, 100), (100, 100), (120, 100) and (150, 150). The data sets are simulated with 1000 repetitions. The bandwidth is determined as $h = c * m^{-1/3}$, where c is selected (see Chen et al. 2009) to minimize the mean-squared error. We use the Epanechnikov kernel function in the simulation study,

$$w(u) = \begin{cases} \frac{3}{4}(1-u^2) & \text{if } \ |u| \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

$\hat{\theta}(p)$ is the empirical estimator of difference of two quantile from two samples. Recall that f_1 and f_2 are two density functions defined before. From properties of the difference of two quantiles from two independent samples, we have

$$\sqrt{m+n}\{\hat{\theta}(p) - \theta(p)\} \xrightarrow{\mathcal{D}} N(0, \sigma^2(\theta(p))),$$

where

$$\sigma^2(\theta(p)) = \frac{(1+r)(1-p)p}{r f_1^2\{F_1^{-1}(p)\}} + \frac{(1+r)(1-p)p}{f_2^2\{F_2^{-1}(p)\}},$$

and f_1 and f_2 are estimated by the common kernel method. Based on this result, we construct normal approximation (NA) confidence interval for two-sample difference of

Table 1 Coverage probability of 95 % confidence interval for the difference of quantiles with two samples at p

p	m	n	JEL (A)	NA (A)	JEL (B)	NA (B)	JEL (C)	NA (C)
0.4	50	70	0.953	0.985	0.942	0.970	0.942	0.977
0.4	80	60	0.925	0.950	0.937	0.956	0.941	0.953
0.4	70	100	0.944	0.980	0.943	0.976	0.953	0.981
0.4	100	100	0.933	0.961	0.953	0.965	0.942	0.957
0.4	120	100	0.950	0.960	0.949	0.956	0.944	0.952
0.4	150	150	0.933	0.955	0.944	0.955	0.953	0.960
0.6	50	70	0.932	0.980	0.960	0.952	0.954	0.965
0.6	80	60	0.929	0.944	0.953	0.941	0.942	0.925
0.6	70	100	0.953	0.982	0.950	0.952	0.944	0.960
0.6	100	100	0.946	0.968	0.949	0.947	0.940	0.938
0.6	120	100	0.940	0.954	0.952	0.943	0.939	0.934
0.6	150	150	0.952	0.961	0.954	0.946	0.938	0.935

Table 2 Average length of 95 % confidence interval for the difference of quantiles with two samples at p

p	m	n	JEL (A)	NA (A)	JEL (B)	NA (B)	JEL (C)	NA (C)
0.4	50	70	0.5088	0.5478	0.6003	0.6263	0.6025	0.6872
0.4	80	60	0.4631	0.4291	0.5279	0.4831	0.5494	0.5358
0.4	70	100	0.4436	0.4577	0.5166	0.5156	0.5242	0.5686
0.4	100	100	0.4040	0.3786	0.4691	0.4272	0.4761	0.4647
0.4	120	100	0.3922	0.3464	0.4437	0.3853	0.4583	0.4216
0.4	150	150	0.3609	0.3080	0.3983	0.3403	0.4066	0.3705
0.6	50	70	0.5108	0.5420	0.8291	0.7314	0.8708	0.8913
0.6	80	60	0.3922	0.4267	0.6954	0.5861	0.7953	0.7136
0.6	70	100	0.4427	0.4535	0.7073	0.6273	0.7496	0.7652
0.6	100	100	0.4060	0.3782	0.6160	0.5294	0.6842	0.6463
0.6	120	100	0.3917	0.3443	0.5809	0.4848	0.6543	0.5895
0.6	150	150	0.3605	0.3062	0.5134	0.4314	0.5790	0.5325

quantiles. The coverage probabilities of JEL and NA in Table 1 is close to the nominal level 95 %. However, using the normal approximation method, coverage probabilities perform over-coverage under some small sample cases.

Furthermore, we apply the bisection method to obtain the upper bound and lower bound of the difference of two quantiles. The average lengths of confidence intervals are displayed in Table 2. We find that the confidence intervals with larger sample size have shorter average length. From our results, we find that average lengths with the JEL method are slightly narrower than that with the NA method at small sample settings and wider than that with the NA method for large sample sizes in most cases.

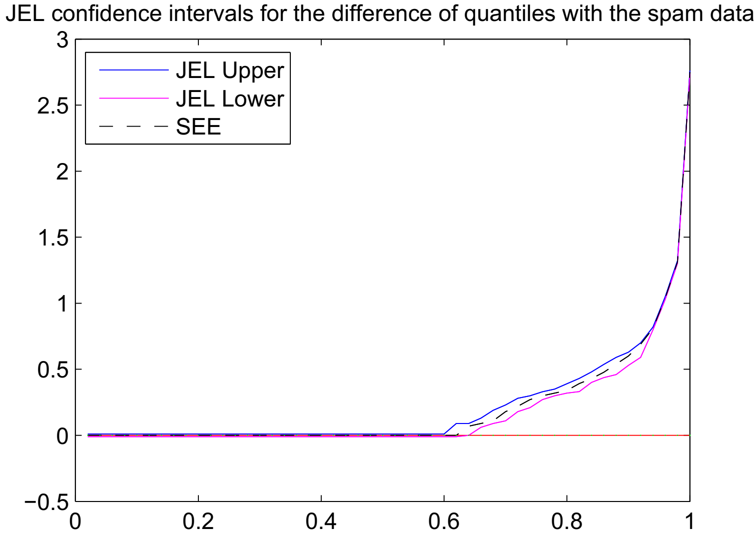

Fig. 1 95 % point-wise JEL confidence intervals for the difference of quantiles from the 24th attribute spam data, where *JEL Upper* indicates the upper bound of JEL confidence intervals, *JEL Lower* indicates the lower bound of JEL confidence intervals and SEE means the smoothed empirical estimator

4 Data analysis

In this section, we investigate a real data set to illustrate the proposed method. The data set is from the web site of the Center for Machine Learning and Intelligent Systems at UC, Irvine. It contains 4601 observations with one indicator variable for spam e-mails, which are the advertisements for products or web sites, make money fast schemes and pornography. Most of those attributes are measured by percentages of certain words appearing in the e-mail. In this paper, we focus on the 24th attribute with zero value of around 60 % observations, which recorded the appearance frequency of word "money" in every email. We split the "money" attribute into two groups by the spam indicator variable. The Epanechnikov kernel function is used for the data analysis as well. Figure 1 shows the point-wise confidence intervals for the difference of quantiles between the spam group and the non-spam group. Confidence intervals for the difference of two quantiles are above the x-axis from the 60 % quantile to the maximum significantly. Thus, the 24th attribute can distinguish the spam observation from the non-spam observation.

5 Discussion

Motivated by the challenge and importance of the difference of quantiles, we develop smoothed JEL methods. The main contribution of this paper is to develop an implicit smoothed estimating equation for the difference of two quantiles and establish the Wilks' theorem for the JEL method. The advantage of the proposed JEL method is that it reduces the number of variables in the optimization and makes the computation less intensive tremendously. Our simulation studies show this new procedure achieves the accuracy in terms of coverage probability and average length in most cases. The proposed JEL inference procedures can be applied to other problems like the low income proportion, value-at-risk and expected shortfall as well (cf. Yang and Zhao 2015).

Acknowledgements This research was supported by National Natural Science Foundation of China (No. 11501567). Yichuan Zhao acknowledges the support from the NSF Grant (DMS-1406163) and NSA Grant (H98230-12-1-0209). The authors would like to thank an associate editor, the two referees for the helpful comments, which has improved the quality of the paper significantly.

Appendix: Proofs of Theorems

Proof of Theorem 1 We can decompose $\Pi_{m,n}\{p,\theta(p)\}$ as

$$\Pi_{m,n}\{p,\theta(p)\} = \frac{1}{m}\sum_{j=1}^{m} K\left\{\frac{p - F_{n,2}\{X_j - \theta(p)\}}{h}\right\} - p - \Pi_m\{p,\theta(p)\} + \Pi_m\{p,\theta(p)\}, \tag{2}$$

where

$$\Pi_m\{p,\theta(p)\} = \frac{1}{m}\sum_{j=1}^{m} K\left\{\frac{p - F_2\{X_j - \theta(p)\}}{h}\right\} - p. \tag{3}$$

The Eq. (3) is simplified as follows

$$\begin{aligned}
\Pi_m\{p,\theta(p)\} &= \frac{1}{m}\sum_{j=1}^{m} K\left\{\frac{p - F_2\{X_j - \theta(p)\}}{h}\right\} - p \\
&= \int_{-\infty}^{\infty} K\left\{\frac{p - F_2\{x - \theta(p)\}}{h}\right\} \mathrm{d}F_{m,1}(x) - p \\
&= K\left\{\frac{p - F_2\{x - \theta(p)\}}{h}\right\} F_{m,1}(x)|_{-\infty}^{\infty} \\
&\quad - \int_{-\infty}^{\infty} F_{m,1}(x)\mathrm{d}K\left\{\frac{p - F_2\{x - \theta(p)\}}{h}\right\} - p \\
&= \frac{1}{h}\int_{-\infty}^{\infty} F_{m,1}(x) w\left\{\frac{p - F_2\{x - \theta(p)\}}{h}\right\} \mathrm{d}F_2\{x - \theta(p)\} - p \\
&= \int_{-1}^{1} F_{m,1}\{F_2^{-1}(p + uh) + \theta(p)\} w\,(u)\,\mathrm{d}u - p \\
&= \int_{-1}^{1} [F_{m,1}\{F_2^{-1}(p + uh) + \theta(p)\} - F_1\{F_2^{-1}(p + uh) + \theta(p)\} \\
&\quad + F_1\{F_2^{-1}(p + uh) + \theta(p)\} - F_1\{F_2^{-1}(p) + \theta(p)\}] w\,(u)\,\mathrm{d}u \\
&= o_p(1).
\end{aligned} \tag{4}$$

The above equation is obtained by the Glivenko–Cantelli Theorem of F_1 and the bounded derivative of $D\{\theta(p), p\} = F_1\{F_2^{-1}(p) + \theta(p)\}$. By Eqs. (10) and (11) in Gong et al. (2010), we can extend their result in our case, i.e., $\Pi_m\{p, \theta(p)\} - \Pi_{m,n}\{p, \theta(p)\} = o_p(1)$. Thus, considering Eqs. (3) and (4), we finish the proof about

$$\Pi_{m,n}\{p, \theta(p)\} \xrightarrow{\mathcal{P}} 0. \tag{5}$$

□

Proof of Theorem 2: One has that

$$\begin{aligned}\sqrt{m+n}\Pi_{m,n}\{p, \theta(p)\} &= \frac{\sqrt{m+n}}{\sqrt{m}}\sqrt{m}[\Pi_m\{p, \theta(p)\}] \\ &\quad + \frac{\sqrt{m+n}}{\sqrt{n}}\sqrt{n}[\Pi_{m,n}\{p, \theta(p)\} - \Pi_m\{p, \theta(p)\}]. \end{aligned}\tag{6}$$

For the first term of (6), one has

$$\begin{aligned}&\sqrt{m}[\Pi_m\{p, \theta(p)\}] \\ &= \int_{-1}^{1} \sqrt{m}[F_{m,1}\{F_2^{-1}(p+uh) + \theta(p)\} - F_1\{F_2^{-1}(p+uh) + \theta(p)\}]w\,(u)\,\mathrm{d}u \\ &\quad + \sqrt{m}\int_{-1}^{1} F_1\{F_2^{-1}(p+uh) + \theta(p)\} - F_1\{F_2^{-1}(p) + \theta(p)\}]w\,(u)\,\mathrm{d}u \\ &= \int_{-1}^{1} W_{F_1}\{F_2^{-1}(p+uh) + \theta(p)\}w\,(u)\,\mathrm{d}u + \sqrt{m}\int_{-1}^{1} D^{'}\{\theta(p), p\}uhw\,(u)\,\mathrm{d}u \\ &\quad + O_p(\sqrt{m}h^2) = I + II + O_p(\sqrt{m}h^2), \end{aligned}\tag{7}$$

where $W_{F_1}(t) = \sqrt{m}\{F_{m,1}(t) - F_1(t)\}$. Because of the symmetric property of kernel function, the second term of (7) is equal to zero.

By arguments in pp. 266 and 269 in van der Vaart (2000), the weak convergence of $F_{m,1}(x)$ and $F_{n,2}(x)$ is true.

$$\sqrt{m}\{F_{m,1}(x) - F_1(x)\} \Longrightarrow B_1(F_1(x)), \quad \sqrt{n}\{F_{n,2}(x) - F_2(x)\} \Longrightarrow B_2(F_2(x)),$$

where $B_1(\cdot)$ and $B_2(\cdot)$ are two independent Brownian bridge on $[0, 1]$. Thus, $B_1(F_1(x))$ and $B_2(F_2(x))$ are independent. Due to the Donsker theorem and similar proofs for equation 9 in Gong et al. (2010), $I \xrightarrow{\mathfrak{D}} B_1(F_1\{F_2^{-1}(p) + \theta(p)\})$. Using $F_1^{-1}(p) = F_2^{-1}(p) + \theta(p)$, it is clear that

$$\sqrt{m}[\Pi_m\{p, \theta(p)\}] \xrightarrow{\mathfrak{D}} B_1(p). \tag{8}$$

For the second term of Eq. (6), under condition C.1, we adopt the procedure similar to that in Gong et al. (2010),

$$\begin{aligned}
&\sqrt{n}[\Pi_{m,n}\{p,\theta(p)\} - \Pi_m\{p,\theta(p)\}] \\
&\quad = -\int_{-\infty}^{\infty} W_{F_2}(x) w\left\{\frac{p - F_2\{x-\theta(p)\}}{h}\right\} \mathrm{d}F_1(x) + O_p(n^{-1/2}h^{-1}) \\
&\quad = \int_{-1}^{1} W_{F_2}\{F_2^{-1}(p)\} w(u) D^{'}\{p,\theta(p)\}\mathrm{d}u + O_p(n^{-1/2}h^{-1}) \\
&\quad \xrightarrow{\mathfrak{D}} B_2(F_2\{F_2^{-1}(p)\}) D^{'}\{p,\theta(p)\}, \\
&\quad = B_2(p) D^{'}\{p,\theta(p)\},
\end{aligned} \tag{9}$$

where $W_{F_2}(t) = \sqrt{n}\{F_{n,2}(t) - F_2(t)\}$. Combining (7), (8), (9) and the independence of $B_1(F_1(x))$ and $B_2(F_2(x))$, one has that

$$\sqrt{m+n}\,\Pi_{m,n}\{p,\theta(p)\} \xrightarrow{\mathfrak{D}} N(0,\sigma^2(p)). \tag{10}$$

□

Before we prove Theorem 3, we need to obtain the asymptotic normality of jackknife estimator and the consistency of jackknife variance estimator. Those asymptotic properties are given in Lemmas 1 and 2.

Lemma 1 *Suppose conditions C.1–C.4 hold. We have*

$$\sqrt{m+n}\left\{\frac{1}{m+n}\sum_{i=1}^{m+n} \hat{V}_i\{p,\theta(p)\}\right\} \xrightarrow{\mathcal{D}} N\{0,\sigma^2(p)\},$$

where $\sigma^2(p)$ is defined in Theorem 2.

Proof of Lemma 1: First, we introduce some properties of $F_{n,2,-i}$ as follows:

$$F_{n,2,-i}(X_j) - F_{n,2}(X_j) = \frac{1}{n-1}\{F_{n,2}(X_j) - I(Y_i \le X_j)\} = O_p\left(\frac{1}{n-1}\right), i = 1,\ldots,n \tag{11}$$

and

$$\sum_{i=1}^{n}\{F_{n,2,-i}(X_j) - F_{n,2}(X_j)\} = 0,$$

because

$$
\begin{aligned}
&F_{n,2,-i}(X_j) - F_{n,2}(X_j) \\
&\quad = \frac{1}{n-1} \sum_{k=1,k\neq i}^{n} I(Y_k \le X_j) - \frac{1}{n}\sum_{k=1}^{n} I(Y_k \le X_j) \\
&\quad = \frac{1}{n-1} \left\{ \sum_{k=1,k\neq i}^{n} I(Y_k \le X_j) - \sum_{k=1}^{n} I(Y_k \le X_j) \right\} + \left(\frac{1}{n-1} - \frac{1}{n} \right) \sum_{k=1}^{n} I(Y_k \le X_j) \\
&\quad = \frac{1}{n-1} \{F_{n,2}(X_j) - I(Y_i \le X_j)\}.
\end{aligned} \tag{12}
$$

For the pseudo-sample, based on equation (16) in Gong et al. (2010) and Eqs. (11) and (12), one has that

$$
\begin{aligned}
&\left\{ \frac{1}{m+n} \sum_{i=1}^{m+n} \hat{V}_i\{p, \theta(p)\} \right\} \\
&\quad = \frac{1}{m+n} \sum_{i=1}^{m} \left[\frac{m+n}{m} \sum_{j=1}^{m} K\left\{ \frac{p - F_{n,2}\{X_j - \theta(p)\}}{h} \right\} \right. \\
&\qquad - \frac{m+n-1}{(m+n)(m-1)} \sum_{i=1}^{m} \left\{ \sum_{j=1, j\neq i}^{m} K\left\{ \frac{p - F_{n,2}\{X_j - \theta(p)\}}{h} \right\} \right\} - p \\
&\qquad + \sum_{i=m+1}^{m+n} \left\{ \frac{1}{m} \sum_{j=1}^{m} K\left\{ \frac{p - F_{n,2}\{X_j - \theta(p)\}}{h} \right\} \right. \\
&\qquad \left. - \frac{m+n-1}{(m+n)m} \sum_{j=1}^{m} K\left\{ \frac{p - F_{n,2,m-i}\{X_j - \theta(p)\}}{h} \right\} \right\} \\
&\quad = \frac{1}{m+n} \left[\sum_{j=1}^{m} K\left\{ \frac{p - F_{n,2}\{X_j - \theta(p)\}}{h} \right\} + \frac{n}{m} \sum_{j=1}^{m} K\left\{ \frac{p - F_{n,2}\{X_j - \theta(p)\}}{h} \right\} \right] \\
&\qquad + \frac{m+n-1}{(m+n)m} \sum_{i=m+1}^{m+n} \sum_{j=1}^{m} \left[K\left\{ \frac{p - F_{n,2}\{X_j - \theta(p)\}}{h} \right\} \right. \\
&\qquad \left. - K\left\{ \frac{p - F_{n,2,m-i}\{X_j - \theta(p)\}}{h} \right\} \right] - p \\
&\quad = \frac{1}{m} \sum_{j=1}^{m} K\left\{ \frac{p - F_{n,2}\{X_j - \theta(p)\}}{h} \right\} - p + O_p\left\{ \frac{mn}{(m+n)(n-1)^2 h} \right\}.
\end{aligned} \tag{13}
$$

Using (10) and (13), it is clear that

$$\sqrt{m+n}\left\{\frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i\{p,\theta(p)\}\right\} = \sqrt{m+n}\,\Pi_{m,n}\{p,\theta(p)\} + o_p(1)$$
$$\xrightarrow{\mathfrak{D}} N(0,\sigma^2(p)). \tag{14}$$

□

In order to obtain the Wilks' theorem for the JEL procedure, we need to check the consistency of jackknife pseudo-sample variance in addition to Eq. (14). Define the pseudo-sample variance

$$v_{m,n}^2\{p,\theta(p)\} = \frac{1}{m+n}\sum_{i=1}^{m+n}\left\{\hat{V}_i\{p,\theta(p)\} - \frac{1}{m+n}\sum_{i=1}^{m+n}\hat{V}_i\{p,\theta(p)\}\right\}^2.$$

Lemma 2 *Under the conditions C.1–C.4, one has that*

$$v_{m,n}^2\{p,\theta(p)\} \xrightarrow{\mathcal{P}} \sigma^2(p).$$

Proof of Lemma 2: For $1 \le i \le m$,

$$\begin{aligned}
&\hat{V}_i\{p,\theta(p)\}\\
&\quad= \frac{m+n}{m}\sum_{j=1}^{m}K\left\{\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right\} - \frac{m+n-1}{m-1}\sum_{j=1,j\neq i}^{m}K\left\{\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right\} - p\\
&\quad= \frac{m+n-1}{m-1}K\left\{\frac{p-F_{n,2}(X_i-\theta(p))}{h}\right\} - \frac{n}{m(m-1)}\sum_{j=1}^{m}K\left\{\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right\} - p
\end{aligned}$$

and

$$\begin{aligned}
&\hat{V}_i^2\{p,\theta(p)\}\\
&= \left[\frac{m+n-1}{m-1}K\left\{\frac{p-F_{n,2}(X_i-\theta(p))}{h}\right\}\right]^2 + \left[\frac{n}{m(m-1)}\sum_{j=1}^{m}K\left\{\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right\}\right]^2\\
&\quad - \frac{2(m+n-1)n}{m(m-1)^2}K\left\{\frac{p-F_{n,2}(X_i-\theta(p))}{h}\right\}\sum_{j=1}^{m}K\left\{\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right\} + p^2\\
&\quad - 2p\left[\frac{m+n-1}{m-1}K\left\{\frac{p-F_{n,2}(X_i-\theta(p))}{h}\right\} - \frac{n}{m(m-1)}\sum_{j=1}^{m}K\left\{\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right\}\right].
\end{aligned}$$

By the similar argument in Gong et al. (2010), one has

$$\frac{1}{m+n}\sum_{j=1}^{m}\hat{V}_i^2\{p,\theta(p)\}\xrightarrow{\mathcal{P}}\frac{r+1}{r}p(1-p). \tag{15}$$

For $m+1\le i\le m+n$, one has that

$$\begin{aligned}&\hat{V}_i\{p,\theta(p)\}\\&\quad=\frac{m+n-1}{m}\sum_{j=1}^{m}\left[K\left\{\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right\}-K\left\{\frac{p-F_{n,2,m-i}\{X_j-\theta(p)\}}{h}\right\}\right]\\&\qquad+\frac{1}{m}\sum_{j=1}^{m}K\left\{\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right\}-p,\end{aligned}$$

and

$$\begin{aligned}\hat{V}_i^2\{p,\theta(p)\}&=\left\{\frac{m+n-1}{m}\right\}^2\left[\sum_{j=1}^{m}K\left\{\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right\}\right.\\&\quad\left.-K\left\{\frac{p-F_{n,2,m-i}\{X_j-\theta(p)\}}{h}\right\}\right]^2+o_p(1)=\left\{\frac{m+n-1}{m}\right\}^2\\&\quad\times\left[\sum_{j=1}^{m}w\left\{\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right\}\frac{F_{n,2}\{X_j-\theta(p)\}-F_{n,2,m-i}\{X_j-\theta(p)\}}{h}\right]^2\\&\quad+o_p(1).\end{aligned}$$

Under condition C.1, we follow the argument which is similar to Gong et al. (2010),

$$\begin{aligned}&\frac{1}{m+n}\sum_{j=m+1}^{m+n}\hat{V}_i^2\{p,\theta(p)\}=\frac{m+n}{nh^2}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{l=1}^{m}\{F_{n,2}(X_j\\&\quad-\theta(p))F_{n,2}(X_l-\theta(p))-F_{n,2}\{X_j-\theta(p)\}I(Y_i\le X_l-\theta(p))\\&\quad-F_{n,2}(X_l-\theta(p))I(Y_i\le X_j-\theta(p))+I(Y_i\le X_j-\theta(p))I(Y_i\le X_l-\theta(p))\}\\&\quad w\left(\frac{p-F_{n,2}\{X_j-\theta(p)\}}{h}\right)w\left(\frac{p-F_{n,2}(X_l-\theta(p))}{h}\right)+o_p(1)\\&\quad=\frac{m+n}{nh^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{F_{n,2}(x_1\wedge x_2-\theta(p))-F_{n,2}(x_1-\theta(p))F_{n,2}(x_2-\theta(p))\}\\&\qquad w\left(\frac{p-F_{n,2}(x_1-\theta(p))}{h}\right)w\left(\frac{p-F_{n,2}(x_2-\theta(p))}{h}\right)\mathrm{d}F_{m,1}(x_1)\mathrm{d}F_{m,1}(x_2)+o_p(1)\\&\quad=\frac{m+n}{nh^2}\int_{-1}^{1}\int_{-1}^{1}\{F_2\{F_2^{-1}(p-u_1h)\wedge F_2^{-1}(p-u_2h)\}\end{aligned}$$

$$
\begin{aligned}
&\quad - F_2\{F_2^{-1}(p-u_1h)\}F_2\{F_2^{-1}(p-u_2h)\}\}\\
&\quad w\,(u_1)\,w\,(u_2)\,\mathrm{d}F_1\{F_2^{-1}(p-u_1h)+\theta(p)\}\mathrm{d}F_1\{F_2^{-1}(p-u_2h)+\theta(p)\}+o_p(1)\\
&=\frac{m+n}{n}\int_{-1}^{1}\int_{-1}^{1}p(1-p)\{D^{'}\{p,\theta(p)\}\}^2w(u_1)w(u_2)\mathrm{d}u_1\mathrm{d}u_2\\
&=\frac{m+n}{n}p(1-p)\{D^{'}\{p,\theta(p)\}\}^2+o_p(1). \qquad (16)
\end{aligned}
$$

Thus, based on Eqs. (15) and (16),

$$\frac{1}{m+n}\sum_{j=1}^{m+n}\hat{V}_i^2\{p,\theta(p)\}\xrightarrow{\mathcal{P}}\sigma^2(p).$$

From Eqs. (5) and (13),

$$v_{m,n}^2\{p,\theta(p)\}\overset{\mathcal{P}}{\to}\sigma^2(p).$$

□

Proof of Theorem 3: Combining Lemmas 1 and 2, we can easily prove Theorem 3 by the standard arguments in Owen (1990). The details of the proof are omitted. □

References

Baysal, E., Staum, J. (2008). Empirical likelihood for value-at-risk and expected shortfall. *Journal of Risk*, *11*, 3–32.

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

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

Csörgo, M. (1987). *Quantile processes with statistical applications*. Philadelphia, PA: SIAM Publications.

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

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

Kosorok, M. (1999). Two-sample quantile tests under general conditions. *Biometrika*, *86*, 909–921.

Miller, R. G. (1974). The jackknife: a review. *Biometrika*, *61*, 1–15.

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

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

Owen, A. (2001). *Empirical likelihood*. New York: Chapman & Hall.

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

Reiss, R. D. (1989). *Approximate distributions of order statistics*. New York: Springer.

Sheather, S., Marron, J. (1990). Kernel quantile estimators. *Journal of the American Statistical Association*, *85*, 410–416.

van der Vaart, A. W., (2000). *Asymptotic Statistics*. Cambridge, UK: Cambridge University Press.

Wang, R., Peng, L., Qi, Y. (2013). Jackknife empirical likelihood test for equality of two high-dimensional means. *Statistica Sinica*, *23*(2), 667–690.

Yang, H., Zhao, Y. (2013). Smoothed jackknife empirical likelihood inference for the difference of ROC curves. *Journal of Multivariate Analysis*, *140*, 270–284.

Yang, H., Zhao, Y. (2015). Smoothed jackknife empirical likelihood inference for ROC curves with missing data. *Journal of Multivariate Analysis*, *115*, 123–138.

Zhou, W., Jing, B.-Y. (2003). Adjusted empirical likelihood method for quantiles. *Annals Institute of Statistical Mathematics*, *55*, 689–703.