Journal of Alexandria University for Administrative Sciences© Vol. 61, No. 5, September 2024

A Robust Nonparametric Yeo- Johnson- Transformation- Based Confidence Interval for Quantiles of Skewed Distributions[1]

Dr. Labiba Hassab Elnaby Alatar
Assistant Professor at Department of Statistics
Faculty of Business, Alexandria University
labiba.elatar@alexu.edu.eg

Dr. Fatma Gaber Abdelaty
Lecturer at Department of Statistics
Faculty of Business, Alexandria University
fatma.gaber@alexu.edu.eg

Dr. Mohammad Ibrahim Soliman Gaafar
Lecturer at Department of Statistics
Faculty of Business, Alexandria University
Mohammad.solayman@alexu.edu.eg
aswanybakkar@yahoo.com

ABSTRACT

The main goal of this paper is to introduce a new robust nonparametric confidence interval for population quantiles. To achieve this goal, a robustified version of an exact equal-tailed two-sided confidence interval for normal quantiles is first introduced. The proposed confidence interval uses the Yeo-Johnson family of power transformations to bring the data into approximate normality or at least symmetry. Calculating the robustified confidence interval using the transformed data then transforming back the lower and upper limits of the confidence interval, the new proposed robust nonparametric confidence interval for population quantile is obtained. Through a simulation study, the proposed confidence interval is evaluated and compared with some competitor existing confidence intervals. The criteria used to evaluate and compare the performance of confidence intervals are: the coverage probability (CP), the mean length of confidence intervals (ML), and the root mean squared deviations of confidence interval's midpoints from the true population quantiles (RMSmdp) of the confidence intervals from the true population quantile. There are no sample size restrictions on the new proposed confidence interval. Simulation results show a significant outperformance of the proposed confidence interval compared to all other competitors under investigation.

Keywords: *Robust Estimators, Nonparametric Confidence Intervals, Central and Intermediate Quantiles, Yeo-Johnson Family of Power Transformation, The Bi-weight Location and Scale Estimators, Siddiqui-Bloch-Gastwirth Estimator, Sectioning, Batching, Empirical Likelihood, Kernel Quantile Estimators, Fixed-smoothing Asymptotics.*

[1] Received in 29/5/2024, accepted in 4/7/2024.

1. INTRODUCTION

Quantiles have numerous applications in diverse areas such as: Finance, Business, Insurance, Biology, Metrology, Economics, Reliability, Quality control, Medicine, Hydrology, Engineering, Demography, etc. For example, quantiles are used to measure value-at-risk in finance and to model birth weights in demography (Kaplan, 2015).

Assume that $X_1, X_2, \cdots, X_n$ is a random sample of size (n) drawn from a continuous population with distribution function $F(\cdot)$ and density function $f(\cdot)$. Assume also that $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ is the corresponding order statistics. The pth population quantile K_p, is defined as $K_p = inf\{x: F(x) \geq p\} \quad \forall \quad 0 < p < 1$. The p^{th} sample quantile is $X_{(s)}$, where $s = s(n) = [np] + 1$ and $[np]$ equals the largest integer not exceeding np. Under mild conditions, $\sqrt{n}\left(X_{(s)} - K_p\right)$ is asymptotically normally distributed with zero mean and a variance of $\left(p(1-p)/f^2(K_p)\right)$ (Serfling, 1980). This asymptotic variance gets large in the tails indicating difficulties when estimating extreme quantiles. The Siddiqui-Bloch-Gastwirth consistent estimator of $1/f\left(K_p\right)$ is $S_{m,n} = (n/2m)\left(X_{(s+m)} - X_{(s-m)}\right)$ as $n \to \infty$, $m = m(n) \to \infty$ and $m/n \to 0$ (Siddiqui, 1960 and Bloch & Gastwirth, 1968). Hall and Sheather (1988) proved that the studentized sample quantile $\ddot{T} = \sqrt{n}\left(X_{(s)} - K_p\right)/\left(S_{m,n}\sqrt{p(1-p)}\right)$ has an asymptotic standard normal distribution. Maesono and Penev (2011) obtained the Edgeworth expansion of a *standardized* version of the kernel quantile estimator. Maesono and Penev (2013) obtained the Edgeworth expansion of a *studentized* version of the kernel quantile estimator which can be inverted to get more accurate interval estimation of quantiles. Kaplan (2015) proposed an approach based on fixed-smoothing asymptotics and a corresponding Edgeworth expansion. According to his proposed approach, only $n \to \infty$ while fixing m at its finite-sample value. He showed that the Edgeworth expansion corresponds to the fixed-m distribution contains terms capturing the variance of the quantile spacing and hence it is more accurate.

Although the statistical literature is rich in various point estimators of population quantiles, constructing confidence intervals is a more difficult problem. Beran (1987) used iterative bootstrap to reduce coverage error. Ho and Lee (2005b) used a smoothed version of Beran's (1987) procedure. The results of their simulation study showed a worse performance for their procedure compared to Chen & Hall's (1993) smoothed empirical likelihood. Nagaraja and Nagaraja (2020) surveyed numerous distribution-free methods for constructing approximate confidence intervals for central and intermediate quantiles. They divided the techniques through which most existing quantiles' confidence intervals were constructed into four main approaches: pivotal quantity, resampling, interpolation, and empirical likelihood. Examples of confidence intervals of population quantiles that are based on pivotal quantities are: Hall & Sheather (1988), Martin (1990), Goh (2004), Peng & Yang (2009), Kaplan (2015), and Nagaraja & Nagaraja (2020). Examples of confidence intervals of population quantiles that are based on resampling methods are: (a) Efron (1979, 1987), Beran (1987), and Ho & Lee (2005b) used the bootstrap approach. (b) Shao & Wu (1989), Martin (1990), and Peng & Yang (2009) used the jackknife approach. (c) Schmeiser (1982) used batching. (d) Nakayama (2014) used sectioning. Examples of confidence intervals of population quantiles that are based on interpolated order statistics are: Hettmansperger & Sheather (1986), Nyblom (1992), Beran & Hall (1993), Hutson (1999), Ho & Lee (2005a), and Goldman & Kaplan (2017). Recently, Frey and Zhang (2017) showed that in the family of continuous distributions, the minimum probability guaranteed by confidence intervals formed by interpolated order statistics do not expand the set of available guaranteed coverage probabilities. Examples of confidence intervals of population quantiles that are based on empirical likelihood method are: Owen (1988), Chen & Hall (1993), Adimari (1998), and Zhou & Jing (2003).

This paper is organized as follows. Section 2 introduces the Yeo-Johnson family of power transformations proposed by Yeo & Johnson (2000). This section gives the form of the transformation, its inverse function, and all functions required to get the maximum likelihood estimate of the transformation's parameter. Section 3 is devoted to six competitor existing confidence intervals for population quantiles. Section 3 reviews also a Gaussian-based quantile confidence interval and introduces a robustified version to employed in the new proposed confidence interval. The proposed confidence interval is described in section 4. The design of the simulation study used to evaluate and compare the empirical performance of the proposed confidence interval is described and the main results are reported in section 5. Section 6 gives concise conclusions of the study.

2. THE YEO-JOHNSON FAMILY OF POWER TRANSFORMATIONS

Under location-scale models, the main goal of a transformation is to achieve some degree of normality, or at least symmetry, required for the validity of applying a certain statistical technique. This section presents the Yeo-Johnson family of power transformations introduced by Yeo & Johnson (2000) which is valid to handle both positive and negative data. This transformation is defined as:

$$T_{YJ}(X;\lambda) = \begin{cases} \left(1-(1-X)^{2-\lambda}\right)/(2-\lambda) & if \quad X \le 0 \ \ and\ \lambda \ne 2 \\ -ln(1-X) & if \quad X \le 0 \ \ and\ \lambda = 2 \\ \left((1+X)^{\lambda}-1\right)/\lambda & if \quad X > 0 \ \ and\ \lambda \ne 0 \\ ln(1+X) & if \quad X > 0 \ \ and\ \lambda = 0 \end{cases} \tag{2.1}$$

This transformation is a special case of the generalized modulus power transformation introduced by Halawa (1989).

Van Zwet (1964) proved that a non-decreasing concave (convex) transformation of a random variable extends the lower (upper) part of the support and squeezes the upper (lower) part. Accordingly, $T_{YJ}(X;\lambda)$ decreases the skewness to the right when $\lambda \in (0,1)$ while it decreases the skewness to the left when $\lambda \in (1,\infty)$.

The inverse transformation of $T_{YJ}(X;\lambda)$ *is given as:*

$$X = T_{YJ}^{-1}(T_{YJ}) = \begin{cases} \left(1-\left(1-(2-\lambda)T_{YJ}(X;\lambda)\right)^{(1/(2-\lambda))}\right) & if \quad T_{YJ}(X;\lambda) \le 0 \ \ and\ \lambda \ne 2 \\ 1-e^{-T_{YJ}(X;\lambda)} & if \quad T_{YJ}(X;\lambda) \le 0 \ \ and\ \lambda = 2 \\ \left(\left(1+\lambda T_{YJ}(X;\lambda)\right)^{(1/\lambda)}-1\right) & if \quad T_{YJ}(X;\lambda) > 0 \ \ and\ \lambda \ne 0 \\ e^{T_{YJ}(X;\lambda)}-1 & if \quad T_{YJ}(X;\lambda) > 0 \ \ and\ \lambda = 0 \end{cases} \tag{2.2}$$

Assume that there exists unknown value of the transformation parameter (λ) such that the model

$$T_{YJ}\left(X-\widehat{M};\lambda\right) = \mu_{YJ} + \sigma_{YJ}\varepsilon \tag{2.3}$$

holds with ε having a standard normal distribution, where ($\widehat{M}$) refers to the sample median.

The Jacobian of the transformation $T_{YJ}\left(X-\widehat{M};\lambda\right)$ is given as:

$$T_{YJ}\left(X-\widehat{M};\lambda\right) = \left(\left|X-\widehat{M}\right|+1\right)^{(\lambda-1)sgn(X-\widehat{M})}$$

and hence, under model (2.3), the log-likelihood function used to estimate the parameters vector $\theta_{YJ} = \left(\mu_{YJ}, \sigma_{YJ}, \lambda\right)^t$ is given as:

$$l_{YJ}\left(\theta_{YJ}; x_1, \dots, x_n\right) = \frac{-n}{2} In(2\pi) - nIn\left(\sigma_{YJ}\right) - \frac{1}{2\sigma_{YJ}^2} \sum_{i=1}^{n} \left[T_{YJ}\left(X - \widehat{M}; \lambda\right) - \mu_{YJ}\right]^2 + (\lambda - 1) \sum_{i=1}^{n} sgn\left(X - \widehat{M}\right) ln\left(\left|X - \widehat{M}\right| + 1\right) \quad (2.4)$$

The partial derivatives of (2.4) with respect to μ_{YJ}, σ_{YJ}, and λ are respectively,

$$\frac{\partial l_{YJ}}{\partial \mu_{YJ}} = \frac{1}{\sigma_{YJ}^2} \sum_{i=1}^{n} \left[T_{YJ}\left(X - \widehat{M}; \lambda\right) - \mu_{YJ}\right], \quad (2.5)$$

$$\frac{\partial l_{YJ}}{\partial \sigma_{YJ}} = \frac{-n}{\sigma_{YJ}} + \frac{1}{\sigma_{YJ}^3} \sum_{i=1}^{n} \left[T_{YJ}\left(X - \widehat{M}; \lambda\right) - \mu_{YJ}\right]^2, \quad (2.6)$$

$$\frac{\partial l_{YJ}}{\partial \lambda} = \frac{-1}{\sigma_{YJ}^2} \sum_{i=1}^{n} \left[T_{YJ}\left(X - \widehat{M}; \lambda\right) - \mu_{YJ}\right] \times \frac{\partial T_{YJ}(X - \widehat{M}; \lambda)}{\partial \lambda} + \sum_{i=1}^{n} sgn\left(X - \widehat{M}\right) ln\left(\left|X - \widehat{M}\right| + 1\right) \quad (2.7)$$

where,

$$\frac{\partial T_{YJ}}{\partial \lambda} = \begin{cases} \left[T_{YJ}\left(X - \widehat{M}; \lambda\right) + \left(1 + \widehat{M} - x_i\right)^{(2-\lambda)} ln\left(1 + \widehat{M} - x_i\right)\right] / (2 - \lambda) & if \quad x_i \le \widehat{M} \ and\ \lambda \ne 2 \\ 0.5\left[ln\left(1 + \widehat{M} - x_i\right)\right]^2 & if \quad x_i \le \widehat{M} \ and\ \lambda = 2 \\ \left[\left(1 - \widehat{M} + x_i\right)^{\lambda} ln\left(1 - \widehat{M} + x_i\right) - T_{YJ}\left(X - \widehat{M}; \lambda\right)\right] / \lambda & if \quad x_i > \widehat{M} \ and\ \lambda \ne 0 \\ 0.5\left[ln\left(1 - \widehat{M} + x_i\right)\right]^2 & if \quad x_i > \widehat{M} \ and\ \lambda = 0 \end{cases}, \quad (2.8)$$

For fixed (λ), setting (2.5) and (2.6) equal to zero, we get the estimates,

$$\hat{\mu}_{YJ} = \frac{1}{n} \sum_{i=1}^{n} T_{YJ}\left(X - \widehat{M}; \lambda\right), \quad (2.9)$$

and $\hat{\sigma}_{YJ}^2 = \frac{1}{n} \sum_{i=1}^{n} \left[T_{YJ}\left(X - \widehat{M}; \lambda\right) - \hat{\mu}_{YJ}\right]^2$ (2.10)

Setting (2.7) equal to zero and substituting the estimators from (2.9) and (2.10) for μ_{YJ} and σ_{YJ} respectively, an estimate $\left(\hat{\lambda}\ \right)$ of (λ) can be obtained using an algorithm for solving nonlinear systems of equations. The normal likelihood vector of estimators $\left(\hat{\theta}_{YJ}\right)$ is determined by iteratively updating (2.9) and (2.10) by $\left(\hat{\lambda}\ \right)$ and updating (2.7) by $\hat{\mu}_{YJ}$ and $\hat{\sigma}_{YJ}$ until a certain convergence criterion is reached.

3. COMPETITOR QUANTILE CONFIDENCE INTERVALS

In this section, an exact equal-tailed confidence interval based on the uniformly minimum variance unbiased estimator (UMVUE) of normal quantiles is reviewed and a robustified version is introduced. Also, a review of six of the best existing quantile confidence intervals are considered.

3.1. The Binomial Confidence Interval

For integers r_1 and r_2 satisfying $0 \le r_1 < r_2 \le n$, the most well-known classical nonparametric $100(1-\alpha)\%$ confidence interval for K_p is given as:

$$\left(X_{(r_1)},\ X_{(1+r_2)}\right) \tag{3.1}$$

where r_1 and r_2 are defined as: $r_1 = \omega(\alpha/2), r_2 = \omega\big(1-(\alpha/2)\big)$, and $\omega(\tau) = The\ (\tau)^{th}\ quantile\ of\ Bin(n,p)$.

Due to the discreteness of the binomial distribution, the coverage probability (or the exact attained confidence level) of this interval cannot be generally rendered equal to a predetermined confidence level. The exact attained confidence level of this interval is greater than or equal to the nominal confidence level $(1-\alpha)$. The error of the coverage probability of this interval is of order $O(n^{-0.5})$.

When the quantile level (p) is very close to 0 or 1, r_1 and r_2 may take on the values 0 and (n) respectively. To overcome the case where $r_1 = 0$, the lower limit of the confidence interval can be replaced by the first order statistics $X_{(1)}$ or any proper monotone and unbounded functions of one or more smallest extreme order statistics. Another plausible strategy to overcome the case where $r_1 = 0$, is to construct a lower-tail confidence interval $\left(-\infty,\ X_{(1+r_3)}\right)$, where $r_3 = \omega(1-\alpha)$. Similarly, to overcome the case where $r_2 = n+1$, the upper limit of the confidence interval can be replaced by the last order statistics $X_{(n)}$ or any proper monotone and unbounded functions of one or more large extreme order statistics. Another plausible strategy to overcome the case where $r_2 = n+1$, is to construct an upper-tail confidence interval $\left(X_{(r_4)}\ ,\ \infty\right)$, where $r_4 = \omega(\alpha)$.

3.2. Nyblom's Confidence Interval

Through a convex combination of adjacent order statistics, Nyblom (1992) extended the work of Hettmansperger & Sheather (1986) and proposed an equal-tailed two-sided confidence interval for the p^{th} population quantile K_p. This interval can be rewritten as follows:

$$\left(\widehat{K}^{Nyb}\big(p,(\alpha/2)\big),\ \widehat{K}^{Nyb}\big(p,(1-\alpha/2)\big)\right) \tag{3.2}$$

Where,

$\widehat{K}^{Nyb}(p,\tau) = \left(1-\lambda^{Nyb}(\tau)\right)X_{(\omega(\tau))} + \lambda^{Nyb}(\tau)X_{(1+\omega(\tau))}$,

$\omega(\tau) = The\ (\tau)^{th}\ quantile\ of\ Bin(n,p)$,

$$\lambda^{Nyb}(\tau) = \left\{1+\frac{(1-p)\omega(\tau)\left[\mathrm{B}_{\omega(\tau)}-\tau\right]}{p\left(n-\omega(\tau)\right)\left[\tau-\mathrm{B}_{\omega(\tau)-1}\right]}\right\}^{-1},$$

$\mathrm{B}_j = pr(\mathrm{B} \leq \mathrm{j})$.

The number of sample observations less than K_p, denoted by B, is binomially distributed $Bin(n,p)$.

3.3 Beran and Hall's Confidence Interval

Beran and Hall (1993) proposed a confidence interval for the p^{th} population quantile, K_p, based on a simple linear interpolation between two adjacent classical nonparametric confidence intervals of the form described in (3.1). Their interval can be rewritten as:

$$\left(\widehat{K}^{BH}\left(p,(\alpha/2)\right),\ \widehat{K}^{BH}\left(p,(1-\alpha/2)\right)\right), \quad (3.3)$$

Where,

$\widehat{K}^{BH}(p,\tau) = \left(1-\lambda^{BH}(\tau)\right)X_{(\omega(\tau))} + \lambda^{BH}(\tau)X_{(1+\omega(\tau))}$,

$\omega(\tau) = The\ (\tau)^{th}\ quantile\ of\ Bin(n,p)$,

$$\lambda^{BH}(\tau) = \frac{\tau-\mathrm{B}_{\omega(\tau)-1}}{\mathrm{B}_{\omega(\tau)}-\mathrm{B}_{\omega(\tau)-1}},$$

The order of the coverage probability error of this interpolated confidence interval reduces to $O(n^{-1})$ *for all smooth distributions.*

3.4. Hutson's Confidence Interval

Based on fractional order statistics defined by Stigler (1977), Hutson (1999) used a different interpolation weight to propose a new confidence interval for population quantile. This confidence interval can be rewritten as follows:

$$\left(\widehat{K}^{Hut}\left(p,(\alpha/2)\right),\ \widehat{K}^{Hut}\left(p,(1-\alpha/2)\right)\right), \quad (3.4)$$

Where,

$\widehat{K}^{Hut}(p,\tau) = \left(1-\lambda^{Hut}(\tau)\right)X_{(\lfloor(n+1)u(\tau)\rfloor)} + \lambda^{Hut}(\tau)X_{(1+\lfloor(n+1)u(\tau)\rfloor)}$,

$\lambda^{Hut}(\tau) = (n+1)u(\tau) - \lfloor(n+1)u(\tau)\rfloor$,

and $u(\tau)$ can be determined by solving the following equation:

$$\mathrm{I}_p\left((n+1)u(\tau), (n+1)\left(1-u(\tau)\right)\right) = 1-\tau. \quad (3.5)$$

The regularized incomplete beta function $\mathrm{I}_v(a,b)$ is defined as:

$$I_v(a,b) = \left[\int_0^v t^{a-1}(1-t)^{b-1}.dt\right] \Big/ \beta eta(a,b)$$

It should be noted that this confidence interval can*not* be obtained for extreme quantiles where the quantile level (p) is very close to 0 or 1.

3.5. Hoo and Lee's Confidence Interval

To reduce the coverage probability error of the confidence interval introduced by Beran & Hall (1993), Ho and Lee (2005a) improved the asymptotic results by calibrating the nominal coverage probability. Their interval can be rewritten as:

$$\left(\widehat{K}^{HL}(p,(\alpha/2)),\ \widehat{K}^{HL}(p,(1-\alpha/2))\right), \quad (3.6)$$

Where,

$\widehat{K}^{HL}(p,\tau) = \widehat{K}^{BH}(p,\tilde{\tau}_{HL})$,

$$\tilde{\tau}_{HL} = \tau - \frac{1}{n}\left\{\frac{\lambda^{BH}(\tau)[1-\lambda^{BH}(\tau)] \times Z_\tau \times \phi(Z_\tau)}{p(1-p)}\right\},$$

Z_τ and $\phi(\cdot)$ are the p[th] quantile and probability density function of the standard normal distribution respectively.

The order of the coverage probability error of this calibrated interpolated confidence interval reduces to $O(n^{-1.5})$ *for all smooth distributions.*

It worth mentioning that when the quantile level (p) is very close to 0 or 1, $\omega(\tau)$ or $\omega(\tilde{\tau})$ may take on the values 0 and (n) respectively. The same strategies described for the binomial confidence intervals to overcome this problem can be used for Nyblom's (1992), Beran & Hall's (1993) , and Ho & Lee's (2005) confidence intervals.

3.6. Goldman and Kaplan's Confidence Interval

Going on the same lines of Ho & Lee (2005), Goldman and Kaplan (2017) applied a calibration method to improve the asymptotic performance of the confidence interval introduce by Hutson (1999). Their interval can be rewritten as:

$$\left(\hat{K}^{GK}\left(p,(\alpha/2)\right),\ \hat{K}^{GK}\left(p,(1-\alpha/2)\right)\right), \tag{3.7}$$

Where,

$$\hat{K}^{GK}(p,\tau) = \hat{K}^{Hut}(p,\tilde{\tau}_{GK}),$$

$$\tilde{\tau}_{GK} = \tau - \frac{1}{n}\left\{\frac{\lambda^{Hut}(\tau)[1-\lambda^{Hut}(\tau)] \times Z_\tau \times \phi(Z_\tau)}{p(1-p)}\right\}.$$

3.7. An Exact Equal-tailed T-Confidence Interval for Normal Quantiles

Based on the uniformly minimum variance unbiased estimator (UMVUE) of normal quantile, Chakraborti and Li (2007) introduced an exact equal-tailed $100(1-\alpha)\%$ confidence interval of normal quantiles given as:

$$\left(\hat{K}^{CL}\left(p,(\alpha/2)\right),\ \hat{K}^{CL}\left(p,(1-\alpha/2)\right)\right), \tag{3.8}$$

where,

$$\hat{K}^{CL}(p,\tau) = \bar{X} + C_n Z_p S - \Delta_{(1-\tau)} S\sqrt{\left(\left[1+nZ_p^2(C_n^2-1)\right]/n\right)}\ ,$$

$$C_n = \left(\sqrt{0.5(n-1)}\Gamma\left(0.5(n-1)\right)/\Gamma(0.5n)\right),$$

$\Delta_{(\tau)}$ is determined by solving the following equation in y_2 :

$$F_{\tilde{T}}\left(\left[y_2 \times \sqrt{\left(\left[1+nZ_p^2(C_n^2-1)\right]/n\right)}\right] - \left[C_n \times Z_p \times \sqrt{n}\right]\right) = 1-\tau.$$

$F_{\tilde{T}}(\cdot)$ refers to the distribution function of the random variable $\tilde{T}$ which follows a noncentral *t*-distribution with $(n-1)$ degrees of freedom and noncentrality parameter $\left(-\sqrt{n}Z_p\right)$.

3.8. A Robutstified Version of the Exact Equal-tailed T-Confidence Interval for Normal Quantiles

Replacing $\bar{X}$ and S in the confidence interval of Chakraborti & Li (2007) by corresponding robust bi-weight location and scale estimators, a robustified equal-tailed $100(1-\alpha)\%$ confidence interval for the p^{th} population quantile $K_p$can be written as:

$$\left(\widehat{K}^{RobCL}(p,(\alpha/2)),\ \ \widehat{K}^{RobCL}(p,(1-\alpha/2))\right), \tag{3.9}$$

where,

$$\widehat{K}^{RobCL}(p,\tau)$$
$$= T_{bi}(X,3.44) + C_n Z_p S_{bi}(X,3.44) - \Delta_{(1-\tau)} S\sqrt{([1+nZ_p^2(C_n^2-1)]/n)}\ ,$$

$T_{bi}(X,3.44)$ is the bi-weight location estimator with tuning constant 3.44 and $S_{bi}(X,3.44)$ is the bi-weight estimator of dispersion with the same tuning constant 3.44.

$$C_n = \left(\sqrt{0.5(n-1)}\Gamma(0.5(n-1))/\Gamma(0.5n)\right),$$

$\Delta_{(\tau)}$ is determined by solving the following equation in y_2 :

$$F_{\tilde{T}}\left(\left[y_2 \times \sqrt{([1+nZ_p^2(C_n^2-1)]/n)}\right] - \left[C_n \times Z_p \times \sqrt{n}\right]\right) = 1-\tau.$$

$F_{\tilde{T}}(\cdot)$ refers to the distribution function of the random variable $\tilde{T}$ which follows a noncentral *t*-distribution with $(n-1)$ degrees of freedom and noncentrality parameter $\left(-\sqrt{n}Z_p\right)$.

As will be shown in the next section, this robust confidence interval will be used in constructing the new proposed confidence interval.

4. The Proposed Quantile Confidence Interval

In this section, a new proposed robust $100(1-\alpha)\%$ confidence interval for the p^{th} population quantile K_p, $\left(\widehat{K}^{LSF}(\alpha/2),\ \ \widehat{K}^{LSF}(p,(1-\alpha/2))\right)$, is described. Whether the distribution is positively or negatively skewed slightly affects how to get the interval. The proposed robust $100(1-\alpha)\%$ confidence interval can be obtained through the following detailed four steps in a random sample of size (n):

[1] Apply the transformation $T_{YJ}(X-\widehat{M};\lambda)$ to the original sample observations and use equations (2.7) up to (2.10) to get an estimate $(\hat{\lambda})$ of the transformation parameter (λ).

[2] Find the transformed observations, Y_1, defined as $Y_1 = T_{YJ}(X-\widehat{M};\lambda)$.

[3] Calculate the confidence interval $\left(\widehat{K}^{RobCL}(p,(\alpha/2)),\ \widehat{K}^{RobCL}(p,(1-\alpha/2))\right)$ based on Y_1.

[4] Calculate the new proposed robust $100(1-\alpha)\%$ confidence interval for the pth population quantile K_p through the following equation:

$$\left(\widehat{K}^{LSF}(p,(\alpha/2)),\ \widehat{K}^{LSF}\left(\mathrm{p},\left(1-(\alpha/2)\right)\right)\right) = \begin{pmatrix} \widehat{M}+T_{YJ}^{-1}\left(\widehat{K}^{RobCL}(p,(\alpha/2));\hat{\lambda}\right), \\ \widehat{M}+T_{YJ}^{-1}\left(\widehat{K}^{RobCL}\left(\mathrm{p},\left(1-(\alpha/2)\right)\right);\hat{\lambda}\right) \end{pmatrix} \tag{3.10}$$

5. SIMULATION STUDY

The simulation study was executed using R version 4.3.3. For each experimental situation described below, 10000 pseudo random samples, of sizes n=20, n=50, and n=100, were generated with initial seed 9831815. These samples are then used to calculate 95% confidence intervals for population quantiles using all competitor and proposed confidence intervals described in Sections 3 and 4 at five different levels, $p = 0.10, 0.25, 0.50, 0.75, 0.90$. The empirical performance of confidence intervals is evaluated and compared based on three criteria, the coverage probability, the average length or width, and the square root of the mean squared deviations of the midpoints of a confidence interval from the true p^{th} population quantile K_p. Four degrees of skewness (2, 4, 5, and 6.18) are covered, the standard exponential distribution, the generalized lambda distribution (GLD), a member of Johnson's unbounded system of distributions, and the standard lognormal distribution respectively.

For all confidence intervals and for each distribution at each and every combination of the considered different levels of "n and p", the three criteria are given in tables 1 up to 4. The main results deducted from these four tables can be summarized as follows:

1. Increasing the sample size substantially reduces the coverage probability errors of all confidence intervals.
2. As the degree of skewness increases or the value of ($\boldsymbol{p}$) gets closer to the longer tail, the empirical performance of all confidence intervals gets worse.
3. As the value of ($\boldsymbol{p}$) gets closer to (0.5), the empirical performance of all confidence intervals improves.
4. The proposed new robust nonparametric confidence interval shows the best performance compared to all other competitors.

6. CONCLUSIONS

In this paper, a new nonparametric robust confidence interval for population quantiles is proposed. There are four main advantages to the new proposed confidence interval. First, no distributional assumptions are required. Second, its robustness. Third, no sample size restrictions. Finally, it can be used for estimating central as well as extreme and near extreme quantiles. Simulation results show that the proposed confidence interval outperforms all other competitors in most cases covered. Accordingly, it is recommended to use the new proposed confidence interval for estimating population quantiles.

Table 1: The coverage probability (CP), the mean length of confidence intervals (ML), and the root mean squared deviations of confidence interval's midpoints from the true population quantiles (RMSmdp) under the standard exponential distribution (Coefficient of skewness $(\gamma) = 2$).

n	p	Criterion	Confidence Intervals						
			Binomial	Nyblom	Beran_Hall	Hutson	Ho_Lee	Goldman_Kaplan	LSF
20	0.10	CP	0.8626	0.8516	0.9850	0.9674	0.8546	0.8378	0.9662
		ML	0.3018	0.2549	0.5715	0.5384	0.2630	0.2377	0.1377
		RMSmdp	0.1266	0.1054	0.1039	0.1140	0.1087	0.0993	0.0888
	0.25	CP	0.9598	0.9508	0.9548	0.9498	0.9532	0.9474	0.9523
		ML	0.5637	0.5046	0.5257	0.5006	0.5151	0.4915	0.3815
		RMSmdp	0.1577	0.1361	0.1425	0.1343	0.1390	0.1322	0.1119
	0.50	CP	0.9608	0.9538	0.9572	0.9556	0.9558	0.9516	0.9564
		ML	0.9582	0.9032	0.9310	0.9161	0.9186	0.8924	0.5587
		RMSmdp	0.2541	0.2447	0.2492	0.2467	0.2471	0.2430	0.1919
	0.75	CP	0.9602	0.9454	0.9514	0.9446	0.9480	0.9408	0.948
		ML	1.8184	1.7348	1.7686	1.7365	1.7528	1.7189	1.0055
		RMSmdp	0.5242	0.5413	0.5373	0.5474	0.5401	0.5473	0.3743
	0.90	CP	0.8668	0.8528	0.9338	0.9152	0.8562	0.8400	0.9245
		ML	2.3147	2.1730	2.5255	2.4257	2.1975	2.1209	1.8209
		RMSmdp	0.7272	0.7592	0.7997	0.8334	0.7528	0.7739	0.6726
50	0.10	CP	0.9682	0.9452	0.9576	0.9484	0.9532	0.9470	0.953
		ML	0.2001	0.1803	0.1854	0.1813	0.1830	0.1805	0.0903
		RMSmdp	0.0413	0.0483	0.0464	0.0480	0.0473	0.0479	0.04
	0.25	CP	0.9656	0.9482	0.9532	0.9476	0.9508	0.9460	0.9504
		ML	0.3555	0.3199	0.3263	0.3194	0.3231	0.3172	0.2158
		RMSmdp	0.0845	0.0797	0.0794	0.0789	0.0795	0.0795	0.0631
	0.50	CP	0.9698	0.9532	0.9588	0.9528	0.9564	0.9504	0.9558
		ML	0.6200	0.5619	0.5737	0.5614	0.5679	0.5570	0.3934
		RMSmdp	0.1454	0.1410	0.1416	0.1410	0.1413	0.1408	0.0999
	0.75	CP	0.9696	0.9538	0.9582	0.9540	0.9564	0.9526	0.9561
		ML	1.0979	1.0008	1.0269	1.0071	1.0147	0.9934	0.6133
		RMSmdp	0.2600	0.2560	0.2607	0.2597	0.2586	0.2558	0.1831
	0.90	CP	0.9696	0.9474	0.9590	0.9502	0.9552	0.9492	0.9546
		ML	2.8189	1.9079	2.1368	1.9500	2.0285	1.9325	0.9798
		RMSmdp	1.0293	0.5749	0.6613	0.5885	0.6169	0.5840	0.3556
100	0.10	CP	0.9556	0.9482	0.9506	0.9484	0.9494	0.9472	0.9495
		ML	0.1346	0.1297	0.1314	0.1299	0.1306	0.1285	0.0701
		RMSmdp	0.0324	0.0315	0.0317	0.0314	0.0316	0.0312	0.0222
	0.25	CP	0.9570	0.9450	0.9474	0.9450	0.9464	0.9436	0.9462
		ML	0.2435	0.2264	0.2287	0.2263	0.2276	0.2254	0.1567
		RMSmdp	0.0565	0.0548	0.0548	0.0547	0.0548	0.0548	0.0457
	0.50	CP	0.9664	0.9508	0.9538	0.9508	0.9526	0.9506	0.9523
		ML	0.4271	0.3953	0.3987	0.3949	0.3970	0.3938	0.2233
		RMSmdp	0.0990	0.0984	0.0984	0.0984	0.0984	0.0984	0.0711
	0.75	CP	0.9632	0.9500	0.9522	0.9502	0.9510	0.9488	0.9512
		ML	0.7365	0.6917	0.6995	0.6927	0.6958	0.6887	0.3944
		RMSmdp	0.1708	0.1720	0.1727	0.1726	0.1723	0.1719	0.1313
	0.90	CP	0.9536	0.9470	0.9496	0.9466	0.9482	0.9450	0.9481
		ML	1.2995	1.2629	1.2775	1.2675	1.2709	1.2563	0.7653
		RMSmdp	0.3403	0.3404	0.3414	0.3424	0.3413	0.3419	0.1761

Table 2:The coverage probability (CP), the mean length of confidence intervals (ML), and the root mean squared deviations of confidence interval's midpoints from the true population quantiles (RMSmdp) under the generalized lambda distribution $gld\ (-0.74575, -0.28726, -0.00989, -0.18315)\quad (\gamma = 4)$.

n	p	Criterion	Confidence Intervals						
			Binomial	Nyblo	Beran_Hall	Hutson	Ho_Lee	Goldman_Kapla	LSF
20	0.10	CP	0.8710	0.8584	0.9842	0.9632	0.8606	0.8424	0.9626
		ML	0.2792	0.2422	0.5550	0.5289	0.2486	0.2286	0.1892
		RMSmdp	0.0942	0.0814	0.1270	0.1374	0.0831	0.0783	0.0671
	0.25	CP	0.9564	0.9472	0.9526	0.9454	0.9508	0.9422	0.9565
		ML	0.4516	0.4063	0.4226	0.4034	0.4144	0.3963	0.2977
		RMSmdp	0.1190	0.1037	0.1081	0.1024	0.1057	0.1010	0.0902
	0.50	CP	0.9582	0.9512	0.9566	0.9542	0.9552	0.9496	0.9504
		ML	0.7698	0.7242	0.7472	0.7349	0.7370	0.7152	0.6933
		RMSmdp	0.2186	0.2085	0.2135	0.2108	0.2112	0.2067	0.1908
	0.75	CP	0.9630	0.9490	0.9560	0.9472	0.9530	0.9424	0.9514
		ML	1.6749	1.6045	1.6340	1.6080	1.6205	1.5919	1.2604
		RMSmdp	0.5865	0.5982	0.5969	0.6049	0.5985	0.6032	0.4686
	0.90	CP	0.8668	0.8524	0.9290	0.9102	0.8556	0.8388	0.9468
		ML	2.5427	2.4218	2.7664	2.6813	2.4427	2.3774	2.0178
		RMSmdp	1.1132	1.1475	1.1990	1.2288	1.1412	1.1615	0.9821
50	0.10	CP	0.9692	0.9444	0.9580	0.9486	0.9530	0.9476	0.9592
		ML	0.2319	0.1863	0.1979	0.1884	0.1924	0.1873	0.1442
		RMSmdp	0.0480	0.0432	0.0425	0.0430	0.0427	0.0430	0.0413
	0.25	CP	0.9668	0.9478	0.9528	0.9464	0.9498	0.9452	0.9598
		ML	0.2796	0.2515	0.2568	0.2514	0.2542	0.2494	0.1783
		RMSmdp	0.0643	0.0612	0.0610	0.0607	0.0611	0.0611	0.0533
	0.50	CP	0.9702	0.9522	0.9576	0.9520	0.9540	0.9508	0.9506
		ML	0.4789	0.4331	0.4424	0.4327	0.4378	0.4292	0.4090
		RMSmdp	0.1152	0.1110	0.1116	0.1110	0.1113	0.1108	0.0998
	0.75	CP	0.9676	0.9504	0.9562	0.9506	0.9538	0.9488	0.9562
		ML	0.9497	0.8652	0.8891	0.8717	0.8779	0.8588	0.7422
		RMSmdp	0.2508	0.2430	0.2490	0.2470	0.2463	0.2425	0.2188
	0.90	CP	0.9698	0.9494	0.9610	0.9530	0.9578	0.9516	0.9490
		ML	3.4026	2.0322	2.3761	2.0954	2.2134	2.0702	1.1532
		RMSmdp	1.5817	0.7203	0.8921	0.7475	0.8041	0.7377	0.4642
100	0.10	CP	0.9562	0.9508	0.9534	0.9518	0.9528	0.9490	0.9560
		ML	0.1328	0.1285	0.1301	0.1287	0.1293	0.1275	0.1005
		RMSmdp	0.0294	0.0292	0.0293	0.0292	0.0292	0.0292	0.0184
	0.25	CP	0.9558	0.9432	0.9458	0.9432	0.9448	0.9420	0.9520
		ML	0.1900	0.1770	0.1788	0.1770	0.1779	0.1762	0.1244
		RMSmdp	0.0444	0.0434	0.0433	0.0433	0.0434	0.0434	0.0381
	0.50	CP	0.9638	0.9480	0.9500	0.9474	0.9496	0.9470	0.9578
		ML	0.3260	0.3016	0.3041	0.3012	0.3029	0.3004	0.2993
		RMSmdp	0.0762	0.0752	0.0752	0.0752	0.0752	0.0751	0.0652
	0.75	CP	0.9644	0.9536	0.9556	0.9530	0.9544	0.9514	0.9544
		ML	0.6210	0.5837	0.5905	0.5847	0.5872	0.5811	0.5179
		RMSmdp	0.1510	0.1513	0.1522	0.1519	0.1518	0.1511	0.1389
	0.90	CP	0.9526	0.9460	0.9490	0.9462	0.9476	0.9446	0.9558
		ML	1.3270	1.2908	1.3057	1.2961	1.2991	1.2847	0.8185
		RMSmdp	0.3940	0.3924	0.3943	0.3951	0.3939	0.3940	0.3295

Table 3: The coverage probability (CP), the mean length of confidence intervals (ML), and the root mean squared deviations of confidence interval's midpoints from the true population quantiles (RMSmdp) under the Johnson's unbounded system of distributions $SU\ (0, 1,\ -0.835,\ 1)\ \ (\gamma = 5)$.

n	p	Criterion	Confidence Intervals						
			Binomial	Nyblo	Beran_Hall	Hutso	Ho_Lee	Goldman_Kaplan	LSF
20	0.10	CP	0.8706	0.8544	0.9552	0.9368	0.8594	0.8442	0.9551
		ML	1.6596	1.5493	1.8899	1.8123	1.5684	1.5088	1.2239
		RMSmdp	0.6306	0.6465	0.6891	0.7104	0.6431	0.6545	0.4276
	0.25	CP	0.9616	0.9500	0.9558	0.9488	0.9532	0.9438	0.9540
		ML	1.5572	1.4419	1.4852	1.4382	1.4641	1.4178	1.1619
		RMSmdp	0.3702	0.3715	0.3706	0.3744	0.3713	0.3745	0.3206
	0.50	CP	0.9562	0.9466	0.9500	0.9486	0.9488	0.9446	0.9528
		ML	1.7852	1.6766	1.7314	1.7020	1.7070	1.6553	1.4512
		RMSmdp	0.4857	0.4653	0.4752	0.4698	0.4707	0.4618	0.4189
	0.75	CP	0.9636	0.9510	0.9578	0.9494	0.9546	0.9446	0.9516
		ML	3.9826	3.8176	3.8877	3.8275	3.8558	3.7887	2.7386
		RMSmdp	1.4939	1.5142	1.5147	1.5313	1.5170	1.5250	1.0469
	0.90	CP	0.8684	0.8498	0.8958	0.8756	0.8558	0.8392	0.9316
		ML	6.7686	6.4933	6.8966	6.7027	6.5409	6.3921	5.5799
		RMSmdp	3.9209	3.9905	4.0195	4.0743	3.9778	4.0182	2.9709
50	0.10	CP	0.9664	0.9440	0.9582	0.9480	0.9528	0.9466	0.9508
		ML	2.1031	1.3059	1.5061	1.3427	1.4114	1.3277	0.7379
		RMSmdp	0.9389	0.4300	0.5280	0.4451	0.4772	0.4397	0.2709
	0.25	CP	0.9664	0.9452	0.9500	0.9464	0.9480	0.9428	0.9526
		ML	0.8888	0.8029	0.8225	0.8055	0.8132	0.7964	0.6934
		RMSmdp	0.1966	0.1951	0.1962	0.1962	0.1957	0.1952	0.1827
	0.50	CP	0.9702	0.9526	0.9570	0.9526	0.9540	0.9506	0.9444
		ML	1.0917	0.9861	1.0076	0.9852	0.9969	0.9771	0.8746
		RMSmdp	0.2532	0.2461	0.2471	0.2461	0.2466	0.2459	0.2296
	0.75	CP	0.9648	0.9504	0.9542	0.9490	0.9524	0.9478	0.9560
		ML	2.1856	1.9901	2.0461	2.0057	2.0200	1.9751	1.5848
		RMSmdp	0.6082	0.5849	0.6007	0.5951	0.5936	0.5833	0.5535
	0.90	CP	0.9746	0.9538	0.9666	0.9570	0.9626	0.9558	0.9546
		ML	9.6606	5.4173	6.4817	5.6129	5.9782	5.5361	2.8151
		RMSmdp	5.1284	2.1306	2.7363	2.2262	2.4258	2.1910	1.1788
100	0.10	CP	0.9566	0.9516	0.9528	0.9516	0.9520	0.9482	0.9414
		ML	0.8434	0.8187	0.8286	0.8218	0.8241	0.8142	0.5118
		RMSmdp	0.2284	0.2278	0.2287	0.2292	0.2285	0.2287	0.1936
	0.25	CP	0.9620	0.9502	0.9528	0.9498	0.9508	0.9488	0.9594
		ML	0.5958	0.5571	0.5631	0.5575	0.5602	0.5546	0.4872
		RMSmdp	0.1348	0.1345	0.1346	0.1346	0.1345	0.1345	0.1268
	0.50	CP	0.9576	0.9390	0.9424	0.9390	0.9408	0.9386	0.9490
		ML	0.7388	0.6826	0.6885	0.6818	0.6856	0.6799	0.6068
		RMSmdp	0.1730	0.1716	0.1716	0.1716	0.1716	0.1716	0.1699
	0.75	CP	0.9652	0.9532	0.9548	0.9534	0.9538	0.9522	0.9534
		ML	1.4100	1.3244	1.3402	1.3268	1.3326	1.3185	1.0640
		RMSmdp	0.3501	0.3492	0.3517	0.3508	0.3505	0.3488	0.3100
	0.90	CP	0.9620	0.9580	0.9598	0.9582	0.9590	0.9556	0.9570
		ML	3.3019	3.2137	3.2504	3.2274	3.2342	3.1996	1.8199
		RMSmdp	1.0117	1.0045	1.0107	1.0119	1.0090	1.0080	0.7299

Table 4: The coverage probability (CP), the mean length of confidence intervals (ML), and the root mean squared deviations of confidence interval's midpoints from the true population quantiles (RMSmdp) under the standard lognormal distribution LogNormal (0,1) ($\gamma = 6.1849$).

n	p	Criterion	Confidence Intervals						
			Binomial	Nyblom	Beran_Hall	Hutson	Ho_Lee	Goldman_Kaplan	LSF
20	0.10	CP	0.8706	0.8570	0.9840	0.9642	0.8608	0.8448	0.9600
		ML	0.4013	0.3464	0.6660	0.6273	0.3559	0.3262	0.2478
		RMSmdp	0.1474	0.1249	0.1231	0.1329	0.1282	0.1189	0.1036
	0.25	CP	0.9616	0.9510	0.9562	0.9496	0.9536	0.9448	0.9670
		ML	0.7206	0.6447	0.6719	0.6398	0.6583	0.6279	0.5530
		RMSmdp	0.1984	0.1710	0.1791	0.1688	0.1747	0.1662	0.1574
	0.50	CP	0.9562	0.9464	0.9500	0.9488	0.9488	0.9446	0.9566
		ML	1.3045	1.2254	1.2653	1.2439	1.2476	1.2098	0.9078
		RMSmdp	0.4041	0.3834	0.3935	0.3880	0.3889	0.3796	0.3199
	0.75	CP	0.9636	0.9512	0.9578	0.9494	0.9544	0.9448	0.9658
		ML	3.3213	3.1923	3.2480	3.2018	3.2229	3.1703	1.9622
		RMSmdp	1.2945	1.3087	1.3104	1.3233	1.3118	1.3172	0.7717
	0.90	CP	0.8684	0.8498	0.8982	0.8780	0.8558	0.8392	0.9412
		ML	5.7980	5.5698	5.9551	5.7944	5.6092	5.4859	4.6085
		RMSmdp	3.3997	3.4587	3.4930	3.5396	3.4480	3.4822	2.1817
50	0.10	CP	0.9664	0.9416	0.9554	0.9436	0.9486	0.9428	0.9466
		ML	0.3087	0.2635	0.2750	0.2656	0.2696	0.2643	0.2101
		RMSmdp	0.0570	0.0626	0.0600	0.0621	0.0611	0.0622	0.0578
	0.25	CP	0.9664	0.9450	0.9508	0.9462	0.9482	0.9426	0.9558
		ML	0.4319	0.3877	0.3959	0.3874	0.3919	0.3844	0.3464
		RMSmdp	0.1023	0.0973	0.0970	0.0965	0.0971	0.0972	0.0558
	0.50	CP	0.9702	0.9528	0.9568	0.9528	0.9540	0.9506	0.9164
		ML	0.7973	0.7202	0.7358	0.7195	0.7280	0.7136	0.5407
		RMSmdp	0.1991	0.1908	0.1920	0.1907	0.1914	0.1903	0.1732
	0.75	CP	0.9648	0.9504	0.9542	0.9490	0.9524	0.9478	0.9018
		ML	1.8141	1.6528	1.6998	1.6664	1.6779	1.6404	1.1394
		RMSmdp	0.5190	0.4973	0.5115	0.5062	0.5051	0.4958	0.3721
	0.90	CP	0.9746	0.9538	0.9666	0.9570	0.9626	0.9560	0.8390
		ML	8.3242	4.6508	5.5722	4.8202	5.1363	4.7537	2.4856
		RMSmdp	4.4452	1.8445	2.3706	1.9275	2.1010	1.8969	0.9036
100	0.10	CP	0.9566	0.9514	0.9532	0.9518	0.9520	0.9484	0.8718
		ML	0.1927	0.1861	0.1885	0.1864	0.1873	0.1845	0.1616
		RMSmdp	0.0429	0.0423	0.0425	0.0423	0.0424	0.0421	0.0333
	0.25	CP	0.9620	0.9502	0.9528	0.9500	0.9508	0.9490	0.9476
		ML	0.2964	0.2759	0.2786	0.2758	0.2773	0.2747	0.2508
		RMSmdp	0.0702	0.0683	0.0682	0.0681	0.0682	0.0683	0.0667
	0.50	CP	0.9576	0.9390	0.9422	0.9390	0.9408	0.9386	0.9080
		ML	0.5391	0.4981	0.5024	0.4975	0.5003	0.4962	0.3740
		RMSmdp	0.1307	0.1287	0.1289	0.1287	0.1288	0.1287	0.1113
	0.75	CP	0.9652	0.9532	0.9548	0.9534	0.9538	0.9522	0.9078
		ML	1.1681	1.0976	1.1108	1.0997	1.1045	1.0927	0.7751
		RMSmdp	0.2947	0.2936	0.2959	0.2949	0.2948	0.2932	0.2327
	0.90	CP	0.9620	0.9580	0.9598	0.9582	0.9590	0.9556	0.8354
		ML	2.8290	2.7537	2.7850	2.7655	2.7713	2.7417	1.6706
		RMSmdp	0.8729	0.8665	0.8719	0.8729	0.8705	0.8695	0.5710

REFERENCES

Adimari, G. (1998). An empirical likelihood statistic for quantiles. Journal of Statistical Computation and Simulation, 60(1), 85–95.

Beran, R. (1987). Pre-pivoting to reduce level error of confidence sets. Biometrika, 74(3), 457–468.

Beran, R., & Hall, P. (1993). Interpolated nonparametric prediction intervals and confidence intervals. Journal of the Royal Statistical Society Series B: Statistical Methodology, 55(3), 643-652.

Bloch, D. A. & Gastwirth, J. L. (1968). On a simple estimate of the reciprocal of the density function. The Annals of Mathematical Statistics, 39(3), 1083-1085.

Chakraborti, S. & Li, J. (2007). Confidence interval estimation of a normal percentile. The American Statistician, 61(4), 331-336.

Chen, S.X. & Hall, P. (1993). Smoothed empirical likelihood confidence intervals for quantiles. The Annals of Statistics, 21(3), 1166–1181.

David, H. A., & Nagaraja, H. N. (2004). *Order statistics*. John Wiley & Sons.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. The Annals of Statistics, 7(1), 1–26.

Efron, B. (1987). Better bootstrap confidence intervals. Journal of the American statistical Association, 82(397), 171-185.

Efron, B., & Tibshirani, R. J. (1993). An introduction to the bootstrap Chapman & Hall. New York.

Frey, J., & Zhang, Y. (2017). What do interpolated nonparametric confidence intervals for population quantiles guarantee? The American Statistician, 71(4), 305-309.

Goldman, M., & Kaplan, D. M. (2017). Fractional order statistic approximation for nonparametric conditional quantile inference. Journal of Econometrics, 196(2), 331-346.

Goh, C. (2004). Smoothing choices and distributional approximations for econometric inference. University of California, Berkeley.

Halawa, A.M. (1989). *Testing for location after transformation to normality* [doctoral thesis]. Oregon State University.

Hall, P., & Sheather, S. J. (1988). On the distribution of a studentized quantile. Journal of the Royal Statistical Society: Series B (Methodological), 50(3), 381-391.

Hettmansperger, T. P., & Sheather, S. J. (1986). Confidence intervals based on interpolated order statistics. Statistics & Probability Letters, 4(2), 75-79.

Ho, Y. H., & Lee, S. M. (2005a). Calibrated interpolated confidence intervals for population quantiles. Biometrika, 92(1), 234-241.

Ho, Y. H. S. & Lee, S.M.S. (2005b). Iterated smoothed bootstrap confidence intervals for population quantiles. The Annals of Statistics, 33(1), 437–462.

Hutson, A. D. (1999). Calculating nonparametric confidence intervals for quantiles using fractional order statistics. Journal of Applied Statistics, 26(3), 343-353.

John, J.A. & Draper, N.R. (1980). An alternative family of transformations. Applied Statistics, 29(2), 190 – 197.

Kaplan, D. M. (2015). Improved quantile inference via fixed-smoothing asymptotics and Edgeworth expansion. Journal of Econometrics, 185(1), 20-32.

Lehmann, E. L., & Casella, G. (2006). *Theory of point estimation*. Springer Science & Business Media.

Martin, M. A. (1990). On using the jackknife to estimate quantile variance. Canadian Journal of Statistics, 18(2), 149-153.

Maesono, Y. & Penev, S. (2011). Edgeworth expansion for the kernel quantile estimator. Annals of the Institute of Statistical Mathematics, 63, 617-644.

Maesono, Y., & Penev, S. (2013). Improved confidence intervals for quantiles. Annals of the Institute of Statistical Mathematics, 65, 167-189.

Nakayama, M. K. (2014). Confidence intervals for quantiles using sectioning when applying variance-reduction techniques. ACM Transactions on Modeling and Computer Simulation (TOMACS), 24(4), 1-21.

Nagaraja, C. H., & Nagaraja, H. N. (2020). Distribution-free Approximate Methods for Constructing Confidence Intervals for Quantiles. International Statistical Review, 88(1), 75-100.

Nyblom, J. (1992). Note on interpolated order statistics. Statistics & probability letters, 14(2), 129-131.

Owen, A. B. (1988). Empirical likelihood ratio confidence intervals for a single functional. Biometrika, 75(2), 237-249.

Peng, L., & Yang, J. (2009). Jackknife method for intermediate quantiles. Journal of statistical planning and inference, 139(7), 2373-2381.

Schmeiser, B. (1982). Batch size effects in the analysis of simulation output. Operations Research, 30(3), 556-568.

Serfling, R. J. (2009). Approximation theorems of mathematical statistics. John Wiley & Sons.

Shao, J., & Wu, C. J. (1989). A general theory for jackknife variance estimation. The annals of Statistics, 17(3), 1176-1197.

Siddiqui, M. M. (1960). Distribution of quantiles in samples from a bivariate population. Journal of Research of the NBS-B, Mathematics and Mathematical Physics, 64B(3), 145-150.

Stigler, S. M. (1977). Fractional order statistics, with applications. Journal of the American Statistical Association, 72(359), 544-550.

Van Zwet, W. R. (1964). Convex transformations of random variables. Amsterdam: Mathematisch Centrum.

Wilcox, R. R. (2017). *Introduction to robust estimation and hypothesis testing (4th ed.)*. Burlington, MA: Elsevier.

Yeo, I. K., & Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry. Biometrika, 87(4), 954-959.

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

مقدر فترة ثقة متين لا معلمي لمئينات التوزيعات الملتوية بالاعتماد على تحويلة يو- جونسون

د. لبيبة حسب النبي العطار   د. فاطمة جابر عبد العاطي   د. محمد إبراهيم سليمان جعفر

ملخص البحث باللغة العربية

يتم من خلال هذا البحث تقديم مقدر فترة ثقة قوي (أو متين) لا معلمي لمئينات التوزيعات الملتوية. حيث يتم أولاً تقديم نسخة متينة لأحد مقدرات فترات الثقة لمئينات المجتمعات المعتدلة. وتبدأ المرحلة الأولي بتقريب البيانات من الاعتدال، أو على الأقل من التماثل، وذلك باستخدام تحويلة يو-جونسون. يتم بعد ذلك إيجاد فترة ثقة لمئين البيانات المحوَّلة باستخدام النسخة المتينة لفترة الثقة والتي تم تقديمها في هذا البحث. يلي ذلك استخدام الدالة العكسية لتحويلة يو-جونسون وذلك للوصول إلى مقدر فترة ثقة مقترح متين لا معلمي لمئين المجتمع. وقد تم تقييم أداء مقدر فترة الثقة المقترح ومقارنته مع بعض مقدرات فترات الثقة المنافسة من خلال دراسة محاكاة شملت مجموعة من العوامل بمستويات مختلفة. وقد اعتمدت المقارنة على ثلاثة معايير هي: مستوي الثقة التجريبي، متوسط طول فترة الثقة، والجذر التربيعي لمتوسط مربعات انحر افات مركز فترة الثقة عن قيمة المئين الحقيقية. لا توجد قيود على حجم العينة عند استخدام المقدر المقترح. وقد أظهرت نتائج دراسة المحاكاة تفوقاً معنوياً لمقدر فترة الثقة المقترح على جميع المقدرات محل المقارنة.

الكلمـات الدالـة: المقدرات المتينة Robust، فترات ثقة لا معلمية، المئينات Quantiles، تحويلة يو-جونسون، مقدرات Bi-weight للموضع والمقياس، مقدر Siddiqui-Bloch-Gastwirth، التقسيم Sectioning، تكوين حزم Batching.

Suggested Citation according to the APA Style

Alatar, L. H. E., Abdelaty, F. G., Gaafar, M. I. S. (2024). A Robust Nonparametric Yeo-Johnson-Transformation-Based Confidence Interval for Quantiles of Skewed Distributions. *Journal of Alexandria University for Administrative Sciences, 61*(5), 157- 176.