On the Performance of Confidence Intervals for Quantiles

Yijun Zuo

Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, USA
Email: zuo@msu.edu

Abstract Woodruff confidence interval for quantiles is a classical procedure and prevailing in practices and regarded as optimal one for many practitioners. This manuscript examines the performance of bootstrap based confidence interval and the classical Woodruff one for quantiles. It is found that the bootstrap procedure can outperform the Woodruff one in terms of coverage probability(accuracy) and the length of the intervals(efficiency). The validity of these theoretical findings for large sample is further confirmed in finite sample simulation studies.

Keywords: Quantile, bootstrap, Bahadur representation, confidence interval, coverage probability, length of confidence interval.

1 Introduction

Let $F$ be a given probability distribution function. The $p$th quantile of $F$, often denoted by $F^{-1}(p)$ or simply $\xi_p$ (for the fixed $F$), is defined as

$$F^{-1}(p) := \inf \{ x, F(x) \geq p \}, \quad \text{for any } p \in (0, 1).$$

It is well-known that the quantiles $\xi_p$, $p \in (0, 1)$, characterize the distribution $F$. Statistical inference about $\xi_p$ is therefore a central problem in practice.

Let $F_n$ be the empirical distribution assigning mass $1/n$ to each of i.i.d. $X_i$, $i = 1, \cdots, n$, from $F$. Then $\hat{\xi}_p := F_n^{-1}(p)$, the sample $p$th quantile, is a natural estimator of the population counterpart $\xi_p$. Indeed, when the derivative of $F$ at $\xi_p$, $F'(\xi_p)$, exists and is positive, $\hat{\xi}_p$ is consistent for $\xi_p$ and (see [14])

$$\sqrt{n} (\hat{\xi}_p - \xi_p) \xrightarrow{d} N(0, p(1-p)/(F'(\xi_p))^2).$$

The sample quantiles thus can be employed to infer the population quantiles for large sample. For example, an asymptotic $(1 - 2\alpha)$ confidence interval (CI) for $\xi_p$ is:

$$I_{\alpha,n} = [\hat{\xi}_p - z_{1-\alpha} \left( \frac{p(1-p)}{(F'(\xi_p))^2n} \right)^{1/2}, \hat{\xi}_p + z_{1-\alpha} \left( \frac{p(1-p)}{(F'(\xi_p))^2n} \right)^{1/2}],$$

where $z_\alpha$ is the $\alpha$th quantile of the standard normal distribution function $\Phi(x)$ and $0 < \alpha < 1/2$.

Unfortunately, $F'(\xi_p)$ (the density of $F$ at $\xi_p$) has to be estimated first before this inference procedure becomes practically useful and relevant. A most ingenious way to fulfill the task is to employ Bahadur’s [1] representation of the sample quantile. For $F$ twice differentiable at $\xi_p$ with $F'(\xi_p) = f(\xi_p) > 0$,

$$\hat{\xi}_p - \xi_p \approx (p - F_n(\xi_p))/f(\xi_p) + R_n,$$

where $R_n = O(n^{-3/4} \log n)$ almost surely (a.s.) as $n \to \infty$. This and other results in Bahadur [1] imply that for any integers $1 \leq k_1, k_2 \leq n$ such that

$$\frac{k_1}{n} = p - \frac{z_{1-\alpha}(p(1-p))^{1/2}}{n^{1/2}} + o(n^{-1/2}), \quad \frac{k_2}{n} = p + \frac{z_{1-\alpha}(p(1-p))^{1/2}}{n^{1/2}} + o(n^{-1/2})$$

as $n \to \infty$, the following is an asymptotic $(1 - 2\alpha)$ CI for $\xi_p$:

$$I_{\alpha,n} = [X_{(k_1)}, X_{(k_2)}],$$

Copyright © 2016 Isaac Scientific Publishing
where $X_{(1)} \leq \cdots \leq X_{(n)}$ are order statistics. Woodruff [15] first proposed this procedure empirically. The interval thus is also called Woodruff interval in the literature. Nevertheless, it is Bahadur’s representation that provides a theoretical justification of this asymptotic confidence interval procedure.

Aside from sample quantiles, Bootstrap quantiles can provide excellent approximations to sample quantiles in the inference of population quantiles; see, e.g., Bickel and Freedman [2] and Singh [13] in general and Shao and Chen [12] for the special survey situation.

The bootstrap idea leads to not only a general mechanism of generating quantiles but also a general variance and distribution estimation method. The latter offers, in turn, a bootstrap based confidence interval procedure for quantiles (see Section 2). With this bootstrap interval procedure, a natural question raised is: Is it as good (accurate or efficient) as classical Woodruff procedure? Evaluating the performance of the two confidence procedures is the major objective of this manuscript.

It is not the purpose of this manuscript (and is impossible) to review/cite all the important references on the massive topic of bootstrapping related methods. For more complete reviews/surveys, please see the specific review/survey articles on the topic, e.g., DiCiccio and Efron [3] and the references cited therein.

The rest of the paper is organized as follows. Section 2 introduces the bootstrap CIs for quantiles. Section 3 states two performance criteria for CIs. Section 4 presents Bahadur representations of bootstrap sample quantiles and other preliminary results. These results are utilized in Section 5 where the performance of the two types of CIs is examined with respect to two criteria in terms of their asymptotic accuracy and length and their finite sample behavior. It is found that the bootstrap procedure can be more accurate than Woodruff one for most choices of $k_n$, and is as accurate as the Woodruff one with optimal $k_n$, $i = 1, 2$ (see Section 5). The validity of these findings is confirmed at finite samples by simulation studies. The proofs of main results are reserved for the Appendix.

## 2 Bootstrap Sample Quantiles and Confidence Intervals

Now let’s introduce bootstrap CIs (also see Zuo [16]). Let $X^*_n, \cdots, X^*_n$ be a random sample from the empirical distribution $F_n$. It is often called a bootstrap sample. Denote by $F^*_n$ the empirical distribution based on this sample. The $p$th quantile based on this sample, $\hat{\xi}_{pn}$, is called the bootstrap sample $p$th quantile. In addition to Woodruff CI for $\xi_p$ in (6) we now consider a bootstrap type CI.

Let the estimator $\hat{\theta}_n = T(F_n)$ of a functional $\theta = T(F)$ satisfy $\sqrt{n}(\hat{\theta}_n - \theta) \overset{d}{\rightarrow} N(0, \sigma^2_F)$ (in our case $T(F) = F^{-1}(p)$). Note that $\sigma^2_F$ is usually unknown. Let $H_n$ be the c.d.f. of $\sqrt{n}(\hat{\theta}_n - \theta)$. Then for any finite sample

$$\left[ \hat{\theta}_n - n^{-1/2}H_n^{-1}(1 - \alpha), \; \hat{\theta}_n - n^{-1/2}H_n^{-1}(\alpha) \right]$$

is a CI for $\theta$ with an approximate level $(1 - 2\alpha)$. Since $H_n$ is still unknown in general, we thus consider its bootstrap version $\hat{H}_n$ defined as

$$H_n(x) = P_n(\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n) \leq x) = P(\sqrt{n}(\hat{\theta}_n^* - \hat{\theta}_n) \leq x | X_1, \cdots, X_n)$$

with $\hat{\theta}_n^* = T(F^*_n)$. For many $T(\cdot)$, $H_n^{-1}(t) - H_n^{-1}(t) \rightarrow 0$ a.s. as $n \rightarrow \infty$ for $t \in (0, 1)$; see, e.g., Bickel and Freedam [2] and Singh [13] (also see Section 4). Thus

$$\left[ \hat{\theta}_n - n^{-1/2}H_n^{-1}(1 - \alpha), \; \hat{\theta}_n - n^{-1/2}H_n^{-1}(\alpha) \right]$$

is a CI for $\theta$ with an approximate significance level $(1 - 2\alpha)$ for large sample size $n$. It is called the hybrid bootstrap confidence interval. This bootstrap procedure in essence provides a consistent estimator for the variance of $\hat{\theta}_n$. It is an approximation to $I_{Q_n}$ (see (3)) in our quantile case. In some cases and for small $n H_n(x)$ can be calculated directly given $F_n$; see, e.g., Efron [4]. This is true for our quantile functional case. But in general and for large $n$ it is difficult to calculate. We thus consider its empirical version based on bootstrap samples $F^*_n$, $\cdots, F^*_m$,

$$\hat{H}_B(x) = \frac{1}{m} \sum_{j=1}^{m} I\left( \sqrt{n}(T(F^*_j)) - T(F_n) \leq x \right)$$

for any $x \in \mathbb{R}$.
which converges to \( H_{\star}(x) \) a.s. and uniformly in \( x \) as \( m \to \infty \), conditional on \( F_{\star} \). Thus for large \( m \) and \( n \), an approximation to (8) (or (7)) is the CI

\[
[\hat{\theta}_n - n^{-1/2} \hat{H}_{\star}^{-1}(1 - \alpha), \hat{\theta}_n - n^{-1/2} \hat{H}_{\star}^{-1}(\alpha)].
\]  

(9)

One may also simply consider a bootstrap percentile confidence interval defined as

\[
[\hat{K}_{\star}^{-1}(\alpha), \hat{K}_{\star}^{-1}(1 - \alpha)],
\]  

(10)

which is closely related to (9), where the bootstrap sample distribution of \( T(F_{\star}^n) \) is

\[
\hat{K}_{\star}(x) = \frac{1}{m} \sum_{j=1}^{m} I(T(F_{nj}^\star) \leq x), \quad \text{for any } x \in \mathbb{R}.
\]

From now on, we focus on the case \( T(F) = F^{-1}(p) \) and the bootstrap CI (10). A natural question is: How well does (10) perform compared with the classical Woodruff one. We will employ the Bahadur representation of \( \hat{\xi}_{\star}^n \) to answer this question. The performance of (9) (not (8)) will also be examined.

We remark that the accuracy of the bootstrap confidence interval (8) (not (9) or (10)) has been studied for very smooth mean functional \( T \) thoroughly by, e.g., Hall [8] and for the quantile functional by Falk and Kaufmann [5].

3 Two Performance Criteria for Confidence Intervals

Among key performance criteria (or desirable properties) for confidence intervals are validity and optimality. And “validity” is most important, followed closely by “optimality”.

Validity means that the nominal coverage probability (confidence level) of the confidence interval should hold, either exactly or to a good approximation.

Optimality means that the procedure for constructing the CI should make as much use of the information in the data-set as possible. Recall that one could throw away half of a data-set and still be able to derive a valid confidence interval. One way of assessing optimality is by the length of the interval, so that a procedure for constructing a CI is judged better than another if it leads to intervals whose lengths are typically shorter.

4 Bahadur Representations of Bootstrap Sample Quantiles and Other Preliminary Results

To assess the performance of CIs, in this section we present Bahadur representations of bootstrap sample quantiles and other needed preliminary theoretical results. First let’s list an assumption.

(A) \( F \) is twice differentiable at \( \xi_p \) with \( F'(\xi_p) = f(\xi_p) > 0 \),

Theorem 4.1 (Zuo [16]) Let \( 0 < p < 1 \). Under (A), we have

\[
\hat{\xi}_{\star}^n = \xi_{\star} + \frac{F_n(\xi_{\star}) - F_n(\xi_p)}{f(\xi_p)} + R_{1n} = \xi_{\star} + \frac{p - F_n(\xi_p)}{f(\xi_p)} + R_{2n}
\]

(11)

where, with probability 1, \( R_{im} = O(n^{-3/4} \log n), i = 1, 2, \) as \( n \to \infty \).

This representation offers more than what the Ghosh [6] type one does: it provides an absolute (not probability) upper bound for the difference between \( \hat{\xi}_{\star}^n - \xi_{\star} \) and the average of i.i.d sum \( (F_n(\xi_p) - F_n(\xi_{\star})/f(\xi_p)) \), conditional on \( F_n \). It also leads immediately to a main theorem (Theorem 5.1) in Bickel and Freedman [2] about weak convergence of \( \sqrt{n}(\hat{\xi}_{\star}^n - \xi_{\star}) \) as a process indexed in \( p \) under their assumptions. It allows one to employ the law of the iterated logarithm conditional on \( F_n \). Below we present a more general result. Before that we need a condition (note it is different from an assumption)
(C) Integer $1 \leq k_n \leq n$ satisfies $k_n = np + o(n^{1/2}(\log n)^{3/4})$, $\delta \geq 1/2$, as $n \to \infty$.

**Theorem 4.2** (Zuo [16]) Let $0 < p < 1$. Under (A) and (C) with $1/2 \leq \delta \leq 1$, we have

$$X_{k_n} = X_{(k_n)} + \frac{F_n(\xi_p) - F_n^*(\xi_p)}{f(\xi_p)} + R_{1n} = \xi_p + \frac{k_n/n - F_n^*(\xi_p)}{f(\xi_p)} + R_{2n} \quad (12)$$

where, with probability 1, $R_{in} = O(n^{-3/4}\log n)$, $i = 1, 2$, as $n \to \infty$.

**Remark 4.3** The remainders in above theorems are bounded by $Cn^{-3/4}\log n$ a.s. for large $n$ and a constant $C > 0$ and can be improved to $O(n^{-1/2}\log n^{1/2}(\log \log n)^{3/4})$. With Kiefer [9]’s approach, they can be further improved to $O(n^{-1/2}(\log \log n)^{3/4})$. Details will not be pursued here though.

**Theorem 4.4** Under (A),

$$\sup_{x \in \mathbb{R}} |H_n(x) - \Phi(\frac{x}{\sigma_p})| = O(n^{-1/2})$$

and

$$\sup_{x \in \mathbb{R}} |H_n(x) - \Phi(\frac{x}{\sigma_p})| = O(n^{-1/2}(\log n)^{1/2}(\log \log n)^{1/4})$$

a.s. as $n \to \infty$, where $\sigma_p = (p(1-p))^{1/2}/f(\xi_p)$.

**Remark 4.5** The results in the theorem are Berry-Esseen type expansions of the sample and bootstrap sample quantiles. The $O(n^{-1/2})$ bound in the first part was also given in Reiss [10], which, however, requires $f(\xi_p) > 0$ and $\sup_{x \in \mathbb{R}} |f'(x)| < \infty$. The bound for the normal approximation to the bootstrap quantile can be improved to $O(n^{-1/2}(\log \log n)^{3/4})$ (using a result in Singh [13]) if $f'$ is bounded near $\xi_p$. Note that the standard tool, the Edgeworth and Cornish-Fisher expansions, is not applicable in the quantile case since Cramér’s (sufficient) condition is not satisfied.

An immediate consequence of this theorem is that $H_n(x) - H_n(x) \rightarrow 0$ a.s. as $n \rightarrow 0$ and uniformly in $x \in \mathbb{R}$. The rate of this convergence is controlled as follows.

$$\sup_x |H_n(x) - H_n(x)| = O(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}), \text{ a.s. as } n \to \infty. \quad (13)$$

**Theorem 4.6** Under (A),

$$\hat{H}_B^{-1}(t) = \sigma_p\Phi^{-1}(t) + \kappa_{1n}$$

for any $t \in (0, 1)$, where $\kappa_{1n} = O(n^{-1/2}(\log n)^{1/2}(\log \log n)^{1/4})$ a.s. as $n \to \infty$ for $m \geq cn^{1/2}$ and some $c > 0$.

**Remark 4.7** For simplicity we consider above (and hereafter) the case that $m \geq cn^{1/2}$ for some $c > 0$. For a general $m$, we can simply replace $\kappa_{1n}$ above by

$$k_{1nm} = O(\max\{n^{-1/2}(\log n)^{1/2}(\log \log n)^{1/4}, (\log m/m)^{1/4}\}) \text{ a.s. as } \min\{m, n\} \to \infty,$$

and the related results hereafter hold true.

## 5 Performance of the Two Types of CI for Quantiles

We now take advantage of the Bahadur representations of $\hat{\xi}_{pn}^*$ and $\hat{\xi}_{pn}$ and evaluate the performance of the Woodruff and bootstrap percentile CIs for $\xi_p$. We confine our attention to two performance criteria: validity and optimality in terms of the coverage probability and length of the intervals.

---

JAS Copyright © 2016 Isaac Scientific Publishing
5.1 Large Sample Behavior of the CIs

The accuracy of the two types of CI The accuracy of an asymptotic 1 − 2α confidence set $C = C(X_1, \ldots, X_n)$ for the unknown parameter $\theta$ measures the convergence rate of $P(\theta \in C)$ to 1 − 2α. For (10) we have

Theorem 5.1 Under (A),

$$P(\hat{K}_n^{-1}(\alpha) \leq \xi_p \leq \hat{K}_n^{-1}(1-\alpha)) = 1 - 2\alpha + \kappa_{2n},$$

where $\kappa_{2n} = O\left(n^{-\frac{1}{4}} \left(\log n\right)^{\frac{1}{2}} \left(\log \log n\right)^{\frac{1}{2}}\right)$ as $n \to \infty$ for $m \geq cn^{\frac{1}{2}}$ for some $c > 0$.

Remark 5.2 The bound in Theorem 5.1 can be improved to $O\left(\left(\log \log n/n\right)^{\frac{1}{2}}\right)$ if we strengthen (A) and assume that $f > 0$ and $f'$ is continuous near $\xi_p$.

Let $[\hat{\xi}_H, \xi_H]$ be the bootstrap confidence interval given in (9). That is,

$$\hat{\xi}_H = \hat{\xi}_0 + n^{-1/2} \hat{H}_n^{-1}(1-\alpha), \quad \xi_H = \hat{\xi}_0 + n^{-1/2} \hat{H}_n^{-1}(\alpha).$$  \hspace{1cm} (14)

Theorem 5.3 Under (A), $P(\hat{\xi}_H \leq \xi_p \leq \xi_H) = 1 - 2\alpha + \kappa_{3n}$, where $\kappa_{3n} = O\left(n^{-1/4} (\log n)^{1/2} (\log \log n)^{1/4}\right)$ as $n \to \infty$ for $m \geq cn^{1/2}$ for some $c > 0$.

Remark 5.4 The bound in Theorem 5.3 can be improved to $O\left(\left(\log \log n/n\right)^{\frac{1}{2}}\right)$ if we strengthen (A) and assume that $f > 0$ and $f'$ is continuous near $\xi_p$ and employ a result in Falk and Kaufmann [5]. Under these stronger assumptions, the latter authors proved that the coverage probability of (8) is $1 - 2\alpha + O(n^{-\frac{1}{2}})$. Note that, however, it is (9) not (8) that is used in practice in general.

For the Woodruff confidence interval in (6) with $k_{1n}$ and $k_{2n}$ in (5), we have

Theorem 5.5 Under (A),

$$P(X(k_{1n}) \leq \xi_p \leq X(k_{2n})) = 1 - 2\alpha + \kappa_{4n},$$

where $\kappa_{4n} = \epsilon_n + O\left(n^{-\frac{1}{4}} (\log n)^{\frac{1}{2}} (\log \log n)^{\frac{1}{2}}\right)$ and $\epsilon_n = o(1)$, as $n \to \infty$.

Remark 5.6 The bound in the theorem can also be improved to $\epsilon_n + O(n^{-\frac{1}{2}})$ if we strengthen (A) and assume that $f > 0$ and $f'$ is continuous near $\xi_p$.

Remark 5.7 In the light of Theorems 5.1, 5.3 and 5.5, the bootstrap and Woodruff asymptotic confidence intervals have the same accuracy order except the extra $\epsilon_n$ term in Theorem 5.5 (this term is purposely not combined with other term, the same is true for equation (15) below). This term can make a big difference. For example, if $k_{in} = np - n^{\frac{1}{2}} K_i + O(n^{\frac{1}{2}+\gamma} (\log n)^{\gamma}), i = 1, 2$, where $K_2 = -K_1 = -z_1 - \alpha (p(1-p))^{\frac{1}{2}}$, $0 < \gamma < 1/4$, $\gamma > 0$, then both $k_{1n}$ and $k_{2n}$ meet the equations in (5) and

$$P(X(k_{1n}) < \xi_p < X(k_{2n})) = 1 - 2\alpha + O(n^{-1/4+\gamma} (\log n)^{\gamma}).$$

That is, the bootstrap confidence intervals can be more accurate than Woodruff one if $\epsilon \uparrow 1/4$ and $\gamma$ is very large. On the other hand, if $\epsilon_n(= o(1)) = 0$ the latter can be as good as the bootstrap intervals in terms of the (asymptotic) accuracy order.

The length of the two asymptotic CIs Besides the coverage probability, the length is another aspect reflecting the performance of a CI. It is an efficient measure of the interval. At the same asymptotic significance level, it is obviously preferred to have a shorter CI. We now calculate the length of the two
types of CI. By the Bahadur representations for $X_{(k_i_n)}$, $i = 1, 2$, the length $L_W(n)$ of the Woodruff interval (6) satisfies
\[ L_W(n) = \frac{2K_1}{n^{2f(\xi_p)}} + \eta_n + O(n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}(\log \log n)^{\frac{1}{2}}), \quad \text{a.s. as } n \to \infty. \] (15)

where $\eta_n = o(n^{-\frac{1}{2}})$ as $n \to \infty$. It is not difficult to see that the bootstrap confidence intervals (10) and (9) have the same length $L_B(n)$. In the light of Theorem 4.6 it follows that
\[ L_B(n) = \frac{2K_1}{n^{1/2f(\xi_p)}} + O(n^{-\frac{1}{2}}(\log n)^{\frac{1}{2}}(\log \log n)^{\frac{1}{2}}), \quad \text{a.s. as } n \to \infty. \] (16)

**Remark 5.8** The rates at which the length of the two confidence intervals tends to 0 as $n \to \infty$ are the same. That is, the two intervals are asymptotically equally efficient. For a fixed $n$ (i.e. in the finite sample practice), however, $L_W(n)$ can be much larger than $L_B(n)$. Indeed, if $n_{k_in} = np - n^{1/2}K_1 + O(n^{1/4+\varepsilon}(\log n)^{\gamma})$, $i = 1, 2$, as defined in Remark 5.7, then the equations in (5) are satisfied and
\[ L_W(n) = \frac{2K_1}{n^{1/2f(\xi_p)}} + O(n^{-\frac{1}{2}+\varepsilon}(\log n)^{\gamma}), \]
which can be much larger than $L_B(n)$ given in (16).

### 5.2 Finite Sample Behavior of the CIs

To assess the validity of the asymptotic results obtained and to examine the finite sample performance of the two types of CI for $\xi_p$, we now undertake Monte Carlo studies. We will focus on the two desirable properties validity and optimality: the relative coverage frequency (the empirical confidence level) and the average length of the CIs at finite samples. We call the bootstrap hybrid CI (9) $hb$ and the bootstrap percentile CI (10) $pb$. Woodruff interval (6) with the following $k_{i_n}$ (see Remarks 5.7 and 5.8), $i = 1, 2$, is called $w_0$, $w_1$ and $w_2$, respectively (where $\lfloor \cdot \rfloor$ is the floor function).

- $w_0$: $k_{1_n} = \lfloor np - n^{1/2}K_1 - n^{1/4+1/5} \rfloor$, $k_{2_n} = \lfloor np + n^{1/2}K_1 + n^{1/4+1/5} \rfloor$.
- $w_1$: $k_{1_n} = \lfloor np - n^{1/2}K_1 + n^{1/4+1/5} \rfloor$, $k_{2_n} = \lfloor np + n^{1/2}K_1 + n^{1/4+1/5} \rfloor$.
- $w_2$: $k_{1_n} = \lfloor np - n^{1/2}K_1 \rfloor$, $k_{2_n} = \lfloor np + n^{1/2}K_1 \rfloor$.

We generate 1000 samples from $N(0, 1)$, $t(3)$, and $t(1)$ with $n = 20, 50, 100$, respectively. We set $m = 500$, $\alpha = 0.025$ and consider $p = 1/4, 1/2$ and $3/4$, respectively. Tables 1 to 3 list the relative coverage frequency (rcf) and the average length of the intervals (al). Since $w_0$ has extremely large rcf’s ($\approx 100\%$) and al’s, the corresponding results are not informative and hence skipped in the tables.

Inspecting of the tables reveals immediately that $w_1$ should be discarded since it has unacceptably low rcf’s that are far below the nominal level 95%. Results of $w_0$ and $w_1$ confirm the conclusion in Remarks 5.7 and 5.8: Woodruff interval can have very bad performance at large (as well as) small samples for special $k_{i_n}$.

The bootstrap hybrid confidence interval $hb$ shares the same al’s as the bootstrap percentile interval $pb$ but has the second lowest rcf’s which are (unexpectedly) lower than the nominal level 95%. Hence the $hb$ confidence interval (9) is not very useful practically though its coverage probability is closer to 95% when $n$ gets larger.

**Table 1** Relative coverage frequency (rcf) and average length (al) of intervals

<table>
<thead>
<tr>
<th>n</th>
<th>$p = 1/4$</th>
<th>$N(0, 1)$</th>
<th>$p = 1/2$</th>
<th>$p = 3/4$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w1</td>
<td>w2</td>
<td>hb</td>
<td>pb</td>
</tr>
<tr>
<td>20</td>
<td>ref .59</td>
<td>.89</td>
<td>.80</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>al .94</td>
<td>1.5</td>
<td>1.2</td>
<td>1.2</td>
</tr>
<tr>
<td>50</td>
<td>ref .65</td>
<td>.93</td>
<td>.85</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>al .67</td>
<td>.84</td>
<td>.75</td>
<td>.75</td>
</tr>
<tr>
<td>100</td>
<td>ref .62</td>
<td>.96</td>
<td>.85</td>
<td>.95</td>
</tr>
<tr>
<td></td>
<td>al .48</td>
<td>.56</td>
<td>.53</td>
<td>.53</td>
</tr>
</tbody>
</table>
Table 2  Relative coverage frequency (rcf) and average length (al) of intervals

<table>
<thead>
<tr>
<th>n</th>
<th>p = 1/4</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w1</td>
<td>w2</td>
<td>hb</td>
<td>pb</td>
<td>w1</td>
<td>w2</td>
<td>hb</td>
<td>pb</td>
<td>w1</td>
<td>w2</td>
</tr>
<tr>
<td>20</td>
<td>rcf</td>
<td>0.59</td>
<td>0.90</td>
<td>0.84</td>
<td>0.95</td>
<td>0.75</td>
<td>0.94</td>
<td>0.82</td>
<td>0.93</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>al</td>
<td>1.1</td>
<td>3.0</td>
<td>1.9</td>
<td>1.9</td>
<td>1.7</td>
<td>1.4</td>
<td>1.2</td>
<td>1.2</td>
<td>2.7</td>
</tr>
<tr>
<td>50</td>
<td>rcf</td>
<td>0.63</td>
<td>0.94</td>
<td>0.84</td>
<td>0.95</td>
<td>0.75</td>
<td>0.92</td>
<td>0.86</td>
<td>0.94</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>al</td>
<td>0.78</td>
<td>1.1</td>
<td>0.95</td>
<td>0.95</td>
<td>0.83</td>
<td>0.73</td>
<td>0.77</td>
<td>0.77</td>
<td>2.4</td>
</tr>
<tr>
<td>100</td>
<td>rcf</td>
<td>0.63</td>
<td>0.95</td>
<td>0.87</td>
<td>0.96</td>
<td>0.71</td>
<td>0.94</td>
<td>0.89</td>
<td>0.95</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>al</td>
<td>0.56</td>
<td>0.72</td>
<td>0.67</td>
<td>0.67</td>
<td>0.54</td>
<td>0.52</td>
<td>0.53</td>
<td>0.53</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Table 3  Relative coverage frequency (rcf) and average length (al) of intervals

<table>
<thead>
<tr>
<th>n</th>
<th>p = 1/4</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>w1</td>
<td>w2</td>
<td>hb</td>
<td>pb</td>
<td>w1</td>
<td>w2</td>
<td>hb</td>
<td>pb</td>
<td>w1</td>
<td>w2</td>
</tr>
<tr>
<td>20</td>
<td>rcf</td>
<td>0.60</td>
<td>0.88</td>
<td>0.85</td>
<td>0.93</td>
<td>0.76</td>
<td>0.94</td>
<td>0.89</td>
<td>0.92</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>al</td>
<td>1.6</td>
<td>7.2</td>
<td>5.7</td>
<td>5.7</td>
<td>3.4</td>
<td>2.1</td>
<td>1.7</td>
<td>1.7</td>
<td>3.4</td>
</tr>
<tr>
<td>50</td>
<td>rcf</td>
<td>0.63</td>
<td>0.93</td>
<td>0.85</td>
<td>0.94</td>
<td>0.77</td>
<td>0.92</td>
<td>0.88</td>
<td>0.93</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>al</td>
<td>1.1</td>
<td>2.5</td>
<td>1.7</td>
<td>1.7</td>
<td>1.1</td>
<td>0.90</td>
<td>0.94</td>
<td>0.94</td>
<td>1.4</td>
</tr>
<tr>
<td>100</td>
<td>rcf</td>
<td>0.62</td>
<td>0.95</td>
<td>0.89</td>
<td>0.95</td>
<td>0.69</td>
<td>0.94</td>
<td>0.90</td>
<td>0.95</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td>al</td>
<td>0.80</td>
<td>1.3</td>
<td>1.2</td>
<td>1.2</td>
<td>0.67</td>
<td>0.63</td>
<td>0.64</td>
<td>0.64</td>
<td>2.5</td>
</tr>
</tbody>
</table>

We thus need only to focus on the Woodruff w2 and the bootstrap percentile confidence interval (10) pb. The performance of the two is roughly the same when \( n \geq 50 \). Both can reach (or are close to) the nominal level 95% when \( n \geq 50 \). On the other hand, when \( p = 1/4 \), w2 usually has a wider CI and a lower coverage probability. The latter is especially true for \( n = 20 \) while the interval of w2 can get extremely wider when the tails of the distribution get heavier. When \( p = 1/2 \), w2 has a slightly wider interval and higher coverage probability than pb for \( n = 20 \). This is reversed when \( n \geq 50 \). When \( p = 3/4 \), the two perform roughly the same.

Findings above indicate that the bootstrap percentile procedure (10) can perform as well as the Woodruff procedure with optimal choices of \( k_{1n} \) and \( k_{2n} \). It performs better when the sample size or \( p \) is small.

Simulation studies for asymmetric (e.g., \( \chi^2 \)) distributions and for very large \( m, n \) and replication number were conducted. The above conclusions remain valid.

Finally, we conclude that overall the bootstrap percentile procedure (10) is a much favorable alternative to Woodruff confidence procedure.

References


**Appendix: Proofs of main results**

**Proof of Theorem 4.4** First consider \(|x| \geq (\log n/(p(1-p)))^{1/2} := x_n. If x \geq x_n, then
\[
\Phi(x) \geq 1 - \exp(-x^2/(2p(1-p))/(\sqrt{2\pi}x) \geq 1 - (2\sqrt{2\pi}(\log n)^{1/2}n^{-1})^{-1},
\]
and for large \(n\),
\[
1 - H_n(\sigma_p x) = P(F_n(\sigma_p x + \xi_p) - F(\sigma_p x + \xi_p) < p - F(\sigma_p x + \xi_p)) < n^{-1/2},
\]
by Hoeffding’s inequality. Thus for large \(n\), \(\sup_{x \geq x_n} |H_n(\sigma_p x) - \Phi(x)| < n^{-1/2}. Likewise we can obtain the counterpart for \(x \leq -x_n\). Hence for large \(n\)
\[
\sup_{|x| \geq x_n} |H_n(\sigma_p x) - \Phi(x)| < n^{-1/2}. (17)
\]
Now consider \(|x| < x_n. Let \(\xi_n(x) = \sigma_p x n^{-1/2} + \xi_p and p_n(x) = F(\xi_n(x)). Then
\[
\sup_{|x| < x_n} |H_n(\sigma_p x) - \Phi(x)| = \sup_{|x| < x_n} |P(p_n(x) - F_n(\xi_n(x)) \leq p_n(x) - p - \Phi(x)|
\]
\[
\leq \sup_{|x| < x_n} |\Phi((p_n(x) - p)/(p_n(x)(1 - p_n(x)))^{1/2}) - \Phi(x)| + r_n(x) (18)
\]
by Berry-Esséen Theorem for all \(n\), where \(r_n(x) = 33/(4(n p_n(x)(1 - p_n(x)))^{1/2}). Expanding p_n(x) leads to
\[
p_n(x) = p + (p(1-p))^{1/2} x (1 + (p(1-p))^{1/2} x n^{1/2} f^2(\xi_p) + o(1))
\]
and
\[
(p_n(x) - p)/(p_n(x)(1 - p_n(x)))^{1/2} = x + (p(1-p))^{1/2} x^2 n^{1/2} f^2(\xi_p) + o(1) - (1 - 2p) x^2 (p(1-p)n)^{1/2}/n o(x^3)
\]
Thus
\[
\sup_{|x| < x_n} |\Phi(\frac{n^{1/2}(p_n(x) - p)}{(p_n(x)(1 - p_n(x)))^{1/2}} - \Phi(x)| \leq \sup_{|x| < x_n} |\Phi(x) x^2 O(n^{-1/2})| = O(n^{-1/2}).
\]
This, (18), and (17) lead to the first part of the desired result.

By Bahadur representations $\hat{\xi}_{mn}$ (Theorem 4.2) and Remark 4.3, we have

$$\sqrt{n}(\hat{\xi}_{mn} - \xi_{mn}) = \sqrt{n}(F_n(\xi_p) - F_n^*(\xi_p)) / f(\xi_p) + r_{2n}$$

(19)

where $r_{2n} = O(n^{-1/4} \log n)^{1/2} (\log \log n)^{1/4}$ a.s. as $n \to \infty$. Define

$$W_n^* = \sqrt{n}(F_n(\xi_p) - F_n^*(\xi_p)) / (F_n(\xi_p)(1 - F_n(\xi_p)))^{1/2}. $$

(20)

Then by Berry-Esseen theorem, we have

$$\sup_x |F_{W_n^*}(x) - \Phi(x)| = O(1/\sqrt{n}) \text{ conditional on } X_1, X_2, \ldots X_n.$$ (21)

By the representation and the result in Bahadur [1], we have

$$F_n(\xi_p) = p + f(\xi_p)(\xi_p - \hat{\xi}_{mn}) + O(n^{-3/4} \log n) = p + O((\log \log n/n)^{1/2}), \text{ a.s.}$$

(22)

as $n \to \infty$. Let $\sigma_n = (F_n(\xi_p)(1 - F_n(\xi_p)))^{1/2} / f(\xi_p)$. It is readily seen that

$$\sigma_n = \sigma_p^{-1}(1 + \eta_n), \text{ a.s.}$$

(23)

for $\eta_n = O((\log \log n)^{1/2})$ as $n \to \infty$. This, (19), (20) and (21) imply that

$$\sup_{x \in \mathbb{R}} |H_n(x) - \Phi(x)| = \sup_{x \in \mathbb{R}} |F_{W_n^*}((x - r_{2n})/\sigma_n) - \Phi(x/\sigma_p)|$$

$$\leq \sup_{x \in \mathbb{R}} |E[\Phi((x - r_{2n})/\sigma_n) + O(n^{-1/2})] - \Phi(x/\sigma_p)|$$

$$\leq \sup_{x \in \mathbb{R}} |\Phi(x/\sigma_p) - \Phi(x/\sigma_p)O(r_{2n}) + o(r_{2n}) + O(n^{-1/2}) - \Phi(x/\sigma_p)|.$$ (24)

This completes the proof. \hfill \Box

**Proof of Theorem 4.6** By the DKW inequality (see, e.g., Theorem 2.1.3 A of Serfling [11]),

$$P_x(\sup_{x \in \mathbb{R}} |\hat{H}_B(x) - H_*(x)| > (\log m/m)^{1/2}) \leq C/m^2,$$

for a constant $C$ and all $m = 1, 2, \ldots$. Since the right side is independent of $F_n$, the inequality holds for $P$ replacing $P_x$. By the Borel-Cantelli lemma for large $m$

$$\sup_{x \in \mathbb{R}} |\hat{H}_B(x) - H_*(x)| \leq (\log m/m)^{1/2}.$$

By Theorem 4.4, we conclude that

$$\sup_{x \in \mathbb{R}} |\hat{H}_B(x) - \Phi(x/\sigma_p)| = \eta_1_n.$$ (24)

with $\eta_1_n = O(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4})$, a.s. as $n \to \infty$. Thus for any $t \in (0, 1)$,

$$\hat{H}_B(\sigma_p\Phi^{-1}(t - 2\eta_1_n)) < t, \quad \hat{H}_B(\sigma_p\Phi^{-1}(t + 2\eta_1_n)) > t.$$

Thus $\sigma_p\Phi^{-1}(t - 2\eta_1_n) < \hat{H}_B^{-1}(t) \leq \sigma_p\Phi^{-1}(t + 2\eta_1_n)$. Let $\hat{H}_B^{-1}(t) - \sigma_p\Phi^{-1}(t) = \kappa_1_n$. Then the desired result follows immediately from Taylor’s expansion theorem. \hfill \Box

**Proof of Theorem 5.1** For any random variable $S_n$, we have

$$P(S_n < s) = P\left(\bigcup_{k=1}^{\infty} \left(S_n \leq s - 1/k\right)\right) = \lim_{k \to \infty} P(S_n \leq s - 1/k).$$
Let \( G_X(x-) = P(X < x) \) for any \( X \sim G_X \). By the proof of Theorem 4.4, we have
\[
\sup_{x \in \mathbb{R}} |H_n(x-) - \Phi(x/\sigma)| = O(n^{-1/2}) \quad \text{as} \quad n \to \infty.
\] (25)

This and Theorems 4.4 and 4.6 imply that
\[
P(\hat{K}_B^{-1}(\alpha) \leq \xi_p \leq \hat{K}_B^{-1}(1 - \alpha)) \leq P\left(-\hat{H}_B^{-1}(1 - \alpha) \leq n^{1/2}(\hat{\xi}_{pn} - \xi_p) \leq -\hat{H}_B^{-1}(\alpha)\right)
\]
\[= H_n(\sigma z_{1-\alpha} + \kappa_{1n1}) - H_n(\sigma z_\alpha + \kappa_{1n2} -)
\]
\[= \Phi(z_{1-\alpha} + \frac{\kappa_{1n1}}{\sigma}) - \Phi(z_\alpha + \frac{\kappa_{1n2}}{\sigma}) + O(n^{-1/2})
\]
\[= 1 - 2\alpha + \kappa_{2n},
\]
where \( \kappa_{1n1} \) and \( \kappa_{2n} \) are of order \( O(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}) \), \( i = 1, 2 \).

\[\square\]

**Proof of Theorem 5.3** The proof is similar to that of Theorem 5.1 and hence skipped.

\[\square\]

**Proof of Theorem 5.5** To prove the result in the theorem, it suffices to show that as \( n \to \infty \)
\[
P(X_{(k_{1n})} > \xi_p) = \alpha + o(1) + O(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}),
\] (26)
since a similar result can be obtained for \( P(X_{(k_{2n})} < \xi_p) \) and hence the desired result follows. Let \( K_1 = z_{1-\alpha}(p(1-p))^{1/2} \). Then \( k_{1n}/n = p - K_1/n^{1/2} + o(n^{-1/2}) \). By the Bahadur representation results for \( X_{(k_{1n})} \) and for \( \hat{\xi}_{pn} \), we have
\[
X_{(k_{1n})} - \hat{\xi}_{pn} = -K_1/(n^{1/2}f(\xi_p)) + o(n^{-1/2}) + \kappa_{5n},
\] (27)
where \( \kappa_{5n} = O(n^{-1/2}(\log n)^{1/2}(\log \log n)^{1/2}) \) a.s as \( n \to \infty \). This and Theorem 4.4 imply
\[
P(X_{(k_{1n})} > \xi_p) = P\left(\sqrt{n}(\hat{\xi}_{pn} - \xi_p) > K_1/f(\xi_p) - o(1) - n^{-1/2}\kappa_{5n}\right)
\]
\[= \Phi(-z_{1-\alpha} + o(1) + \kappa_{6n}) + O(n^{-1/2})
\]
\[= \alpha + o(1) + O(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}),
\]
where \( \kappa_{6n} = O(n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}) \). This completes the proof.

\[\square\]