The Annals of Statistics
2018, Vol. 46, No. 1, 438–456
https://doi.org/10.1214/17-AOS1557

ON THE ASYMPTOTIC THEORY OF NEW BOOTSTRAP CONFIDENCE BOUNDS

BY CHARL PRETORIUS AND JAN W. H. SWANEPOEL[1]

North-West University

We propose a new method, based on sample splitting, for constructing bootstrap confidence bounds for a parameter appearing in the regular smooth function model. It has been demonstrated in the literature, for example, by Hall [*Ann. Statist.* **16** (1988) 927–985; *The Bootstrap and Edgeworth Expansion* (1992) Springer], that the well-known *percentile-t* method for constructing bootstrap confidence bounds typically incurs a coverage error of order $O(n^{-1})$, with n being the sample size. Our version of the percentile-t bound reduces this coverage error to order $O(n^{-3/2})$ and in some cases to $O(n^{-2})$. Furthermore, whereas the standard *percentile* bounds typically incur coverage error of $O(n^{-1/2})$, the new bounds have reduced error of $O(n^{-1})$. In the case where the parameter of interest is the population mean, we derive for each confidence bound the exact coefficient of the leading term in an asymptotic expansion of the coverage error, although similar results may be obtained for other parameters such as the variance, the correlation coefficient, and the ratio of two means. We show that equal-tailed confidence intervals with coverage error at most $O(n^{-2})$ may be obtained from the newly proposed bounds, as opposed to the typical error $O(n^{-1})$ of the standard intervals. It is also shown that the good properties of the new percentile-t method carry over to regression problems. Results of independent interest are derived, such as a generalisation of a delta method by Cramér [*Mathematical Methods of Statistics* (1946) Princeton Univ. Press] and Hurt [*Apl. Mat.* **21** (1976) 444–456], and an expression for a polynomial appearing in an Edgeworth expansion of the distribution of a Studentised statistic for the slope parameter in a regression model. A small simulation study illustrates the behavior of the confidence bounds for small to moderate sample sizes.

1. Introduction. Since its introduction by Efron [6] in the 1970s, the bootstrap method has provided an ever-increasing number of automated methods tailored for inference, including methods that may be used to construct confidence bounds or intervals for an unknown population parameter. Standard methods include the well-known backwards percentile bound (denoted in this paper by $\hat{I}_B$), a hybrid percentile bound ($\hat{I}_H$) and the percentile-t bound ($\hat{J}$), as well as refinements such as the bias-corrected and the accelerated bias-corrected bounds (see [7]). A very informative theoretical review is given in [9], in which the author

Received May 2016; revised January 2017.

[1]Supported in part by the National Research Foundation of South Africa.

MSC2010 subject classifications. Primary 62G09, 62G20; secondary 62G15.

Key words and phrases. Confidence bounds, sample splitting, coverage error, smooth function model, Edgeworth polynomials, Cornish–Fisher expansion, regression.

demonstrates that using these standard methods to construct *one-sided confidence bounds* typically results in coverage errors of order $O(n^{-1/2})$, except in the case of the percentile-t and accelerated bias-corrected bounds, which incur errors of $O(n^{-1})$.

In [4], Chang and Lee show that it is possible to reduce the coverage error of the standard *percentile* bounds by employing the m/n bootstrap, which were studied by [2, 13], among others. Their method for constructing percentile bounds reduces the coverage error to $O(n^{-1})$. Although in a different way, our new method for constructing bounds also relies on the successes of the m/n bootstrap. We show that our new percentile bounds offer reduced coverage error of $O(n^{-1})$ as well. However, our method may be used to obtain new *percentile-t* bounds with reduced coverage error of size $O(n^{-3/2})$ and in some cases $O(n^{-2})$. These improvements are achieved by the new bounds without computationally intensive bootstrap iteration or parametric assumptions required for most higher-order likelihood or saddlepoint methods.

In the arguments of Hall [9], the order of coverage error of confidence bounds is primarily determined by a *random* distance, for example, $\hat{\theta}_n - \theta = O_p(n^{-1/2})$, where $\hat{\theta}_n$ is some estimator for the parameter θ. The rationale behind our idea rests upon the construction of a confidence bound in such a way that the order of coverage error is essentially determined by a *constant* distance, which is typically of the form $\mathbb{E}(\hat{\theta}_n - \theta) = O(n^{-1})$. This may be accomplished by *splitting* the sample into two independent sets. The method of construction relies partly on the fact that, if Y and Z are two independent random variables in $\mathbb{R}$ and we let $\Psi(z) := \mathbb{P}(Y \geq z)$, $z \in \mathbb{R}$, we may write

$$\mathbb{P}(Y \geq Z) = \mathbb{E}\big(\Psi(Z)\big). \tag{1.1}$$

The remainder of the paper is organised as follows. In Section 2, we briefly discuss the standard bootstrap methods. The construction of the new confidence bounds is presented in Section 3. Section 4 contains a discussion on the asymptotic coverage probabilities of the new hybrid and backwards *percentile* bounds. Section 5 presents a similar discussion on the asymptotic behavior of the new hybrid and backwards *percentile-t* bounds. As an illustrative example, we provide in Section 6 details of the asymptotics of the proposed confidence bounds when the parameter of interest is the mean of a univariate population. As shown in Section 7, the new results may be extended to the linear regression setup, where the slope parameter is of interest. Section 8 contains a brief discussion on how the results for bounds may be used to obtain similar asymptotic results for equal-tailed confidence intervals. Section 9 provides a small simulation study, illustrating the behavior of the confidence bounds for small to moderate samples.

2. The standard methods. To fully appreciate the construction of the new confidence bounds, it is worth stating the standard bounds in terms of bootstrap quantiles. Consider a random sample $\mathcal{X}_n = \{\mathbf{X}_1, \dots, \mathbf{X}_n\}$ from an unknown p-dimensional distribution depending on a scalar parameter θ. The aim is to construct a $(1-\alpha)$-level upper confidence bound for θ, based on some appropriate point estimator $\hat{\theta}_n$ for θ. Denote by $\mathcal{X}_n^* = \{\mathbf{X}_1^*, \dots, \mathbf{X}_n^*\}$ a random sample of size n taken *with replacement* from $\mathcal{X}_n$ and let $\hat{\theta}_n^*$ be the same function of $\mathcal{X}_n^*$ as $\hat{\theta}_n$ is of $\mathcal{X}_n$. In what follows σ^2 denotes the asymptotic variance of $n^{1/2}\hat{\theta}_n$, for which an estimator $\hat{\sigma}_n^2$ exists. Let $\hat{\sigma}_n^*$ be the bootstrap version of $\hat{\sigma}_n$.

In terms of this notation, the two standard *percentile* $(1-\alpha)$-level bootstrap confidence bounds for θ may then be written as

$$\hat{I}_H(\alpha) := \big(-\infty, \hat{\theta}_n - n^{-1/2}\hat{\sigma}_n\hat{\xi}_{n,\alpha}\big],$$
$$\hat{I}_B(\alpha) := \big(-\infty, \hat{\theta}_n + n^{-1/2}\hat{\sigma}_n\hat{\xi}_{n,1-\alpha}\big],$$

where $\hat{\xi}_{n,\alpha}$ is the α-level quantile of the bootstrap distribution of the standardised $\hat{\theta}_n^*$, i.e., $\mathbb{P}^*(n^{1/2}(\hat{\theta}_n^* - \hat{\theta}_n)/\hat{\sigma}_n \le \hat{\xi}_{n,\alpha}) = \alpha$, where $\mathbb{P}^*$ refers to the conditional probability law of $\mathcal{X}_n^*$ given $\mathcal{X}_n$. The subscripts H and B allude to the terms *hybrid* and *backwards* often used to refer to these two types of bounds (cf. [9]). Typically,

$$\mathbb{P}\big(\theta \in \hat{I}_B(\alpha)\big) = 1 - \alpha + O\big(n^{-1/2}\big) = \mathbb{P}\big(\theta \in \hat{I}_H(\alpha)\big).$$

The so-called *percentile-t* bound, favored by [9], may be expressed as

$$\hat{J}(\alpha) := \big(-\infty, \hat{\theta}_n - n^{-1/2}\hat{\sigma}_n\hat{\eta}_{n,\alpha}\big],$$

where $\hat{\eta}_{n,\alpha}$ is the α-level quantile of the bootstrap distribution of the Studentised $\hat{\theta}_n^*$, that is, $\mathbb{P}^*(n^{1/2}(\hat{\theta}_n^* - \hat{\theta}_n)/\hat{\sigma}_n^* \le \hat{\eta}_{n,\alpha}) = \alpha$. Typically,

$$\mathbb{P}\big(\theta \in \hat{J}(\alpha)\big) = 1 - \alpha + O\big(n^{-1}\big).$$

REMARK 2.1. Although only *upper* confidence bounds are studied in this paper, the results immediately hold also for *lower* confidence bounds by noting that if, for example, $\hat{J}(\alpha)$ is an upper $(1-\alpha)$-level confidence bound for θ, then

$$\mathbb{R} \setminus \hat{J}(1-\alpha) = \big(\hat{\theta}_n - n^{-1/2}\hat{\sigma}_n\hat{\eta}_{n,1-\alpha}, \infty\big)$$

is a lower $(1-\alpha)$-level confidence bound for θ.

3. Construction of the new confidence bounds. We first introduce some notation in the regular smooth function model framework of [1]. For $k = 1, \dots, n$, set $\mathbf{W}_k = (f_1(\mathbf{X}_k), \dots, f_d(\mathbf{X}_k))$, where $f_1, \dots, f_d$ are real-valued Borel measurable functions on $\mathbb{R}^p$. Define $\boldsymbol{\nu} = \mathbb{E}(\mathbf{W}_1)$. Assume that the parameter of interest is of the form $\theta = g_s(\boldsymbol{\nu})$, where $g_s : \mathbb{R}^d \to \mathbb{R}$ is a known smooth, Borel measurable function.

Our new method involves splitting the sample in two disjoint sets, say $\mathcal{W}_\ell = \{\mathbf{W}_1, \dots, \mathbf{W}_\ell\}$ and $\mathcal{W}_r = \{\mathbf{W}_{\ell+1}, \dots, \mathbf{W}_n\}$, for some integer $2 \le \ell \le n-2$, with

$r := n - \ell$. Let $\bar{\mathbf{W}}_\ell = \ell^{-1} \sum_{k=1}^{\ell} \mathbf{W}_k$ and $\bar{\mathbf{W}}_r = r^{-1} \sum_{k=\ell+1}^{n} \mathbf{W}_k$. Let $\hat{\theta}_\ell := g_s(\bar{\mathbf{W}}_\ell)$ be an estimator for θ, which we assume has an asymptotic variance of the form $\ell^{-1}\beta^2 = \ell^{-1}h_s^2(\boldsymbol{\nu})$, for some known smooth, Borel measurable function $h_s : \mathbb{R}^d \to \mathbb{R}$. Two possible estimators for β are $\hat{\beta}_\ell := h_s(\bar{\mathbf{W}}_\ell)$ and $\hat{\beta}_r := h_s(\bar{\mathbf{W}}_r)$.

Throughout, assume that $\mathbf{W}_1$ satisfies Cramér's continuity condition, that is,

$$\limsup_{\|\mathbf{t}\|\to\infty} \left|\chi(\mathbf{t})\right| < 1, \tag{3.1}$$

where $\chi(\mathbf{t})$ denotes the characteristic function of $\mathbf{W}_1$. Then, if g and h are sufficiently smooth and $\mathbf{W}_1$ has sufficiently many bounded moments, [1] showed rigorously that the statistics $S_\ell := \ell^{1/2}(\hat{\theta}_\ell - \theta)/\beta$ and $T_\ell := \ell^{1/2}(\hat{\theta}_\ell - \theta)/\hat{\beta}_\ell$ admit the Edgeworth expansions

$$\mathbb{P}(S_\ell \le x) = \Phi(x) + \ell^{-1/2} p_1(x)\phi(x) + \ell^{-1} p_2(x)\phi(x) + \cdots, \tag{3.2}$$

$$\mathbb{P}(T_\ell \le x) = \Phi(x) + \ell^{-1/2} q_1(x)\phi(x) + \ell^{-1} q_2(x)\phi(x) + \cdots, \tag{3.3}$$

uniformly in $x \in \mathbb{R}$, where the p_j and q_j are polynomials of degree $3j - 1$, odd/even for even/odd j, with coefficients depending on moments of $\mathbf{W}_1$ up to order $j + 2$.

It was shown by [4] that valid expansions analogous to (3.2) and (3.3) can be obtained for statistics obtained via the m/n bootstrap. Let $\mathcal{W}_{m,r}^* = \{\mathbf{W}_1^*, \ldots, \mathbf{W}_m^*\}$ denote a *resample* of size m drawn randomly *with replacement* from $\mathcal{W}_r$. Throughout we will assume that $m = O(r)$ and $m \to \infty$ as $r \to \infty$. We do not require the more restrictive assumption $m = o(r)$, as is usually done in the m/r bootstrap literature when considering *nonregular* cases. This means that when we apply the m/r bootstrap we can indeed also take resamples of sizes m larger than r. In fact, several papers have appeared in the literature in which the resample size is chosen larger than the original sample (see, e.g., [3]). In the simulation study in Section 9, we have also considered choices of m larger than r. Now define m/r bootstrap estimators for θ and β as

$$\hat{\theta}_{m,r}^* = g_s(\bar{\mathbf{W}}_{m,r}^*) \quad \text{and} \quad \hat{\beta}_{m,r}^* = h_s(\bar{\mathbf{W}}_{m,r}^*),$$

respectively, where $\bar{\mathbf{W}}_{m,r}^* = m^{-1} \sum_{k=1}^{m} \mathbf{W}_k^*$. Standardised and Studentised versions of the estimator $\hat{\theta}_{m,r}^*$ are

$$S_{m,r}^* := \frac{m^{1/2}(\hat{\theta}_{m,r}^* - \hat{\theta}_r)}{\hat{\beta}_r} \quad \text{and} \quad T_{m,r}^* := \frac{m^{1/2}(\hat{\theta}_{m,r}^* - \hat{\theta}_r)}{\hat{\beta}_{m,r}^*}.$$

Under conditions stated by [4], we may obtain Edgeworth expansions (as power series in $m^{-1/2}$) for $\mathbb{P}^*(S_{m,r}^* \le x)$ and $\mathbb{P}^*(T_{m,r}^* \le x)$ analogous to (3.2) and (3.3), which depend on polynomials $\hat{p}_{j,r}$ and $\hat{q}_{j,r}$ obtained by substituting population moments appearing in p_j and q_j for sample moments calculated from the subsample $\mathcal{W}_r$. Moreover, if we denote the α-level quantiles of the bootstrap distribution of $S^*_{m,r}$ and $T^*_{m,r}$ by $\hat{\xi}_{m,r,\alpha}$ and $\hat{\eta}_{m,r,\alpha}$ respectively, one may obtain the Cornish–Fisher expansions

$$\hat{\xi}_{m,r,\alpha} = z_\alpha + m^{-1/2}\hat{p}^{cf}_{1,r}(z_\alpha) + m^{-1}\hat{p}^{cf}_{2,r}(z_\alpha) + O_p(m^{-3/2}),$$
$$\hat{\eta}_{m,r,\alpha} = z_\alpha + m^{-1/2}\hat{q}^{cf}_{1,r}(z_\alpha) + m^{-1}\hat{q}^{cf}_{2,r}(z_\alpha) + m^{-3/2}\hat{q}^{cf}_{3,r}(z_\alpha) + O_p(m^{-2}),$$

where $z_\alpha = \Phi^{-1}(\alpha)$ denotes the α-level quantile of the standard normal distribution, and $\hat{p}^{cf}_{j,r}$ and $\hat{q}^{cf}_{j,r}$ are polynomials completely determined by the Edgeworth polynomials $\hat{p}_{j,r}$ and $\hat{q}_{j,r}$ (see Lemma 1 in the supplementary material [12]). These expansions hold uniformly in $\varepsilon \le \alpha \le 1 - \varepsilon$ for any $\varepsilon \in (0, \frac{1}{2})$.

We are now ready to propose our new *percentile* $(1-\alpha)$-level upper confidence bounds for θ. Define a hybrid version by

$$\hat{I}^N_H(m,\alpha) := \bigl(-\infty, \hat{\theta}_\ell - \ell^{-1/2}\hat{\beta}_r\tilde{\xi}_{m,r,\alpha}\bigr]$$

and a backwards version by

$$\hat{I}^N_B(m,\alpha) := \bigl(-\infty, \hat{\theta}_\ell + \ell^{-1/2}\hat{\beta}_r\tilde{\xi}_{m,r,1-\alpha}\bigr],$$

where

$$\tilde{\xi}_{m,r,\alpha} := z_\alpha + m^{-1/2}\hat{p}^{cf}_{1,r}(z_\alpha) + m^{-1}\hat{p}^{cf}_{2,r}(z_\alpha). \tag{3.4}$$

Analogously, we define a hybrid and a backwards version of the *percentile-t* type bounds by

$$\hat{J}^N_H(m,\alpha) := \bigl(-\infty, \hat{\theta}_\ell - \ell^{-1/2}\hat{\beta}_\ell\tilde{\eta}_{m,r,\alpha}\bigr]$$

and

$$\hat{J}^N_B(m,\alpha) := \bigl(-\infty, \hat{\theta}_\ell + \ell^{-1/2}\hat{\beta}_\ell\tilde{\eta}_{m,r,1-\alpha}\bigr],$$

where

$$\tilde{\eta}_{m,r,\alpha} := z_\alpha + m^{-1/2}\hat{q}^{cf}_{1,r}(z_\alpha) + m^{-1}\hat{q}^{cf}_{2,r}(z_\alpha) + m^{-3/2}\hat{q}^{cf}_{3,r}(z_\alpha). \tag{3.5}$$

In the following section, we investigate the asymptotic properties of these newly proposed bounds.

4. Asymptotic properties of the percentile bounds. In the following two subsections we derive, under some regularity assumptions, the asymptotic coverage probabilities of the hybrid and backwards percentile bounds. Among others, it is shown that $\hat{I}^N_H$ has coverage error of $O(n^{-1})$, compared to the coverage error of $O(n^{-1/2})$ of the standard bootstrap bound $\hat{I}_H$. As far as the backwards bound is concerned, we show that $\hat{I}^N_B$ has coverage error of $O(n^{-1/2})$, but in some cases also has coverage error of $O(n^{-1})$.

4.1. *Hybrid bound coverage probability.* The next theorem presents an asymptotic expansion for the coverage probability of $\hat{I}_H^N$.

THEOREM 4.1. *Suppose that* $\mathbf{W}_1$ *satisfies* (3.1) *and has sufficiently many finite moments such that* (A1)–(A7) *stated in the supplement hold. Also, assume that* g_s *and* h_s *are continuously differentiable up to a sufficiently high order in an open neighborhood of* $\boldsymbol{\nu}$. *Then, if* $m = \ell = O(r)$ *and* $\ell \to \infty$ *as* $n \to \infty$, *we have that*

$$\mathbb{P}\big(\theta \in \hat{I}_H^N(\ell, \alpha)\big) = 1 - \alpha + \frac{C_\theta(z_\alpha)}{r} + O\big(\ell^{-3/2}\big), \tag{4.1}$$

where $C_\theta(z_\alpha)$ *is the coefficient of* r^{-1} *in a power series expansion of*

$$-z_\alpha \phi(z_\alpha)\beta^{-1}\mathbb{E}(\hat{\beta}_r - \beta) + \frac{1}{2} z_\alpha^3 \phi(z_\alpha)\beta^{-2}\mathbb{E}\{(\hat{\beta}_r - \beta)^2\}.$$

Moreover, if we choose $\ell = \lfloor \gamma n^\psi \rfloor$ *for some* $\gamma > 0$ *and* $\frac{2}{3} < \psi < 1$, *then*

$$\mathbb{P}\big(\theta \in \hat{I}_H^N(\ell, \alpha)\big) = \begin{cases} 1 - \alpha + \dfrac{C_\theta(z_\alpha)}{n} + O\big(n^{-(2-\psi)} + n^{-3\psi/2}\big) & \text{if } C_\theta(z_\alpha) \neq 0, \\ 1 - \alpha + O\big(n^{-3\psi/2}\big) & \text{if } C_\theta(z_\alpha) = 0. \end{cases} \tag{4.2}$$

In the case where $\psi = 1$ *and* $0 < \gamma < 1$,

$$\mathbb{P}\big(\theta \in \hat{I}_H^N(\ell, \alpha)\big) = 1 - \alpha + \frac{C_\theta(z_\alpha)}{(1-\gamma)n} + O\big(n^{-3/2}\big).$$

4.2. *Backwards bound coverage probability.* The next theorem presents an asymptotic expansion for the coverage probability of $\hat{I}_B^N$.

THEOREM 4.2. *Under the assumptions of Theorem* 4.1, *it follows that*

$$\mathbb{P}\big(\theta \in \hat{I}_B^N(\ell, \alpha)\big) = 1 - \alpha + \frac{K_1(z_\alpha)}{\ell^{1/2}} + \frac{K_2(z_\alpha)}{\ell} + \frac{C_\theta(z_\alpha)}{r} + O\big(\ell^{-3/2}\big),$$

where

$$K_1(z_\alpha) = -2p_1(z_\alpha)\phi(z_\alpha), \qquad K_2(z_\alpha) = p_1(z_\alpha)K_1'(z_\alpha).$$

Further, if we choose $\ell = \lfloor \gamma n \rfloor$ *for some* $0 < \gamma < 1$, *then*

$$\mathbb{P}\big(\theta \in \hat{I}_B^N(\ell, \alpha)\big) = 1 - \alpha + \frac{K_1(z_\alpha)}{(\gamma n)^{1/2}} + O\big(n^{-1}\big). \tag{4.3}$$

In the case where $K_1(z_\alpha) = K_2(z_\alpha) = 0$, *all the results of Theorem* 4.1 *hold for* $\hat{I}_B^N$.

REMARK 4.1. In Section 6, we apply the results of this section to the case where the parameter of interest is the mean of a univariate population. We also derive exact expressions for the constants $C_\theta(z_\alpha)$, $K_1(z_\alpha)$ and $K_2(z_\alpha)$. A case where $K_1(z_\alpha) = K_2(z_\alpha) = 0$ is given in Example 6.2.

We now move on to derive corresponding results for the percentile-t bounds $\hat{J}_H^N$ and $\hat{J}_B^N$. It will be seen that $\hat{J}_H^N$ has asymptotic behavior that is superior to that of the percentile bounds.

5. Asymptotic properties of the percentile-*t* bounds.

In this section, we derive asymptotic expressions for the coverage probabilities of the hybrid and backwards percentile-t type bounds. We demonstrate that, typically, the newly proposed hybrid bound $\hat{J}_H^N$ leads to a coverage error of $O(n^{-3/2})$ and in some cases even to $O(n^{-2})$. This is an improvement over the standard percentile-t bootstrap bound $\hat{J}$, which has coverage error $O(n^{-1})$.

5.1. *Hybrid bound coverage probability.* The next theorem presents an asymptotic expansion for the coverage probability of $\hat{J}_H^N$.

THEOREM 5.1. *Suppose that* $\mathbf{W}_1$ *satisfies* (3.1) *and has sufficiently many finite moments such that* (B1)–(B7) *stated in the supplement hold. Also, assume that* g_s *and* h_s *have sufficiently many continuous derivatives in an open neighborhood of* $\boldsymbol{\nu}$. *Then, if* $m = \ell = O(r)$ *and* $\ell \to \infty$ *as* $n \to \infty$, *we have that*

$$\mathbb{P}(\theta \in \hat{J}_H^N(\ell, \alpha)) = 1 - \alpha + \frac{D_\theta(z_\alpha)}{\ell^{1/2} r} + O(\ell^{-2}), \tag{5.1}$$

where $D_\theta(z_\alpha)$ *is the coefficient of* r^{-1} *in a power series expansion of*

$$\phi(z_\alpha)\mathbb{E}\{\hat{q}_{1,r}(z_\alpha) - q_1(z_\alpha)\}. \tag{5.2}$$

Moreover, if we choose $\ell = \lfloor \gamma n^\psi \rfloor$ *for some* $\gamma > 0$ *and* $\frac{2}{3} < \psi < 1$, *then*

$$\begin{aligned} &\mathbb{P}(\theta \in \hat{J}_H^N(\ell, \alpha)) \\ &\quad = \begin{cases} 1 - \alpha + \dfrac{D_\theta(z_\alpha)}{\gamma^{1/2} n^{(2+\psi)/2}} + O(n^{-(4-\psi)/2} + n^{-2\psi}) & \text{if } D_\theta(z_\alpha) \neq 0, \\ 1 - \alpha + O(n^{-2\psi}) & \text{if } D_\theta(z_\alpha) = 0. \end{cases} \end{aligned} \tag{5.3}$$

In the case where $\psi = 1$ *and* $0 < \gamma < 1$,

$$\mathbb{P}(\theta \in \hat{J}_H^N(\ell, \alpha)) = 1 - \alpha + \frac{D_\theta(z_\alpha)}{\gamma^{1/2}(1-\gamma) n^{3/2}} + O(n^{-2}).$$

REMARK 5.1. As will be shown in Example 6.3, it might occur naturally that $D_\theta(z_\alpha) = 0$. In such cases, the order of coverage error is reduced to $O(n^{-2\psi})$, for $\frac{2}{3} < \psi \leq 1$.

5.2. *Backwards bound coverage probability.* The next theorem presents an asymptotic expansion for the coverage probability of $\hat{J}_B^N$.

THEOREM 5.2. *Under the assumptions of Theorem 5.1, it follows that*

$$\mathbb{P}(\theta \in \hat{J}_B^N(\ell,\alpha)) = 1-\alpha+\frac{K_3(z_\alpha)}{\ell^{1/2}}+\frac{K_4(z_\alpha)}{\ell}+\frac{K_5(z_\alpha)}{\ell^{3/2}}-\frac{D_\theta(z_\alpha)}{\ell^{1/2}r}+O(\ell^{-2}),$$

where $K_3(z_\alpha) = -2q_1(z_\alpha)\phi(z_\alpha)$, $K_4(z_\alpha) = q_1(z_\alpha)K_3'(z_\alpha)$, *and*

$$K_5(z_\alpha) = \frac{1}{2}q_1^2(z_\alpha)K_3''(z_\alpha) + q_2^{cf}(z_\alpha)K_3'(z_\alpha) - 2q_3(z_\alpha)\phi(z_\alpha).$$

Furthermore, if we choose $\ell = \lfloor \gamma n \rfloor$ *for some* $0 < \gamma < 1$, *then*

$$\mathbb{P}(\theta \in \hat{J}_B^N(\ell,\alpha)) = 1-\alpha+\frac{K_3(z_\alpha)}{(\gamma n)^{1/2}}+O(n^{-1}). \tag{5.4}$$

In the case where $K_3(z_\alpha) = K_4(z_\alpha) = K_5(z_\alpha) = 0$, *all the results of Theorem 5.1 hold for* $\hat{J}_B^N$.

6. Some illustrative examples. In this section, we provide a detailed discussion for the case where the parameter θ is the *mean* of a univariate population. The results derived in Sections 4 and 5 hold in general for any parameter θ which can be expressed in the regular smooth function model framework of [1], including, for example, the variance, the correlation coefficient, and the ratio of two means.

To be able to derive rigorously exact asymptotic expressions for the expectations in Theorems 4.1, 4.2, 5.1 and 5.2, and the assumptions (A1)–(A7) and (B1)–(B7), calls for a special form of the so-called “delta method”. One convenient result (see [11]) states formal conditions under which the expectation of a Taylor approximation of a bounded function g of statistics accurately approximates the expectation of the function itself up to an arbitrary order. The theorem we prove below extends the result derived by [11] in that it allows the restriction of boundedness of g to be relaxed. Furthermore, the theorem is also a generalization of a result by [5].

THEOREM 6.1. *For any positive integer* s, *let* $g:\mathbb{R}^q \to \mathbb{R}$ *be a function having bounded* $(s+1)$*-order partial derivatives in an open neighborhood of some point* $\boldsymbol{\nu} \in \mathbb{R}^q$. *Suppose* $\mathbf{V}$ *is a* q*-vector of real-valued statistics* (*determined by a sample of size* n) *such that* $|g(\mathbf{V})| \le Cn^{\delta/2}$ *a.s. for* $n \ge n_0$, *with* $n_0 \ge 1$, $C > 0$ *and* $\delta \ge 0$ *some finite constants. If* $\mathbf{V}$ *has finite moments up to order* $2k = (2s) \vee (\delta+s+1)$ *and* $\mathbb{E}(V_i-\nu_i)^{2k} = O(n^{-k})$, $i=1,\dots,q$, *then*

$$\mathbb{E}\{g(\mathbf{V})\} = g(\boldsymbol{\nu}) + \sum_{1\le|\alpha|\le s}\frac{1}{\alpha!}\mathbb{E}\{(\mathbf{V}-\boldsymbol{\nu})^\alpha\}\partial^\alpha g(\boldsymbol{\nu}) + O(n^{-(s+1)/2}),$$

where $|\alpha| = \alpha_1+\cdots+\alpha_q$, $\alpha! = \alpha_1!\cdots\alpha_q!$, $(\mathbf{V}-\boldsymbol{\nu})^\alpha = \prod_{i=1}^q (V_i-\nu_i)^{\alpha_i}$, *and*

$$\partial^\alpha g(\boldsymbol{\nu}) = \frac{\partial^{|\alpha|}}{\partial \nu_1^{\alpha_1}\cdots\partial\nu_q^{\alpha_q}} g(\nu_1,\dots,\nu_q),$$

for any multi-index $\alpha = (\alpha_1,\dots,\alpha_q) \in \mathbb{N}_0^q$.

As a consequence of this theorem, we have the following useful result, which will be required in the examples that follow. To the best of our knowledge, the coefficients of the terms of order n^{-1} do not appear in the existing literature.

COROLLARY 6.1. *Let $X_1, \dots, X_n$ denote i.i.d. random variables such that $\mathbb{E}(|X_1|^k) < \infty$ for some sufficiently large k. Define $\mu = \mathbb{E}(X_1)$, $\sigma^2 = \mathrm{Var}(X_1) > 0$ and denote by κ_j the jth cumulant of $(X_1 - \mu)/\sigma$. Consider the following estimators for κ_3, κ_4 and κ_5:*

$$\hat{\kappa}_{3,n} = \frac{m_3}{m_2^{3/2}}, \qquad \hat{\kappa}_{4,n} = \frac{m_4}{m_2^2} - 3 \quad \textit{and} \quad \hat{\kappa}_{5,n} = \frac{m_5}{m_2^{5/2}} - 10\hat{\kappa}_{3,n},$$

respectively, where $m_j = n^{-1}\sum_{i=1}^n (X_i - \bar{X}_n)^j$ and $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$. It then follows that

$$\mathbb{E}(\hat{\kappa}_{3,n} - \kappa_3) = -\frac{1}{8n}\{12\kappa_5 - 15\kappa_4\kappa_3 + 54\kappa_3\} + O(n^{-2}), \quad \textit{and} \tag{6.1}$$

$$\mathbb{E}(\hat{\kappa}_{4,n} - \kappa_4) = -\frac{1}{n}\{2\kappa_6 - 3\kappa_4^2 + 15\kappa_4 + 12\kappa_3^2 + 6\} + O(n^{-2}).$$

Furthermore, $\mathbb{E}\{(\hat{\kappa}_{3,n} - \kappa_3)^4\} = O(n^{-2})$, $\mathbb{E}\{(\hat{\kappa}_{4,n} - \kappa_4)^2\} = O(n^{-1})$, and $\mathbb{E}\{\hat{\kappa}_{5,n} - \kappa_5\} = O(n^{-1})$.

EXAMPLE 6.1 (Hybrid percentile bound). Let $X_1, \dots, X_n$ denote a random sample from an unknown univariate distribution with mean μ and variance $0 < \sigma^2 < \infty$. We would like to construct the confidence bound $\hat{I}_H^N$ for the population mean μ, which may be expressed in the smooth function model setting as follows. In the notation of Section 3, set $\mathbf{W}_k = (X_k, X_k^2)$, $k = 1, \dots, n$. Then $\boldsymbol{\nu} = \mathbb{E}(\mathbf{W}_1) = (\mu, \mu^2 + \sigma^2)$,

$$\bar{\mathbf{W}}_\ell = \left(\ell^{-1}\sum_{k=1}^{\ell} X_k, \ell^{-1}\sum_{k=1}^{\ell} X_k^2\right), \qquad \bar{\mathbf{W}}_r = \left(r^{-1}\sum_{k=\ell+1}^{n} X_k, r^{-1}\sum_{k=\ell+1}^{n} X_k^2\right).$$

Let $g_s(x_1, x_2) = x_1$ and $h_s^2(x_1, x_2) = x_2 - x_1^2$ so that $\theta = g_s(\boldsymbol{\nu}) = \mu$ and $\beta^2 = h_s^2(\boldsymbol{\nu}) = \sigma^2$. The appropriate estimators for θ and β are then given by $\hat{\theta}_\ell = \ell^{-1}\sum_{k=1}^{\ell} X_k$, $\hat{\theta}_r = r^{-1}\sum_{k=\ell+1}^{n} X_k$, $\hat{\beta}_\ell^2 = \ell^{-1}\sum_{k=1}^{\ell}(X_k - \hat{\theta}_\ell)^2$ and $\hat{\beta}_r^2 = r^{-1}\sum_{k=\ell+1}^{n}(X_k - \hat{\theta}_r)^2$.

For the case of the mean it has been shown in the literature (see, e.g., [10]) that the polynomials p_1 and p_2 in (3.2) are given by

$$p_1(x) = -\frac{1}{6}\kappa_3(x^2 - 1),$$

$$p_2(x) = -x\left\{\frac{1}{24}\kappa_4(x^2 - 3) + \frac{1}{72}\kappa_3^2(x^4 - 10x^2 + 15)\right\},$$

where κ_3 and κ_4 denote the third and fourth cumulants of $(X_1 - \mu)/\sigma$, respectively. Sample versions $\hat{p}_{1,r}$ and $\hat{p}_{2,r}$ of these polynomials may be obtained by substituting κ_3 and κ_4 for their respective estimators based on the subsample $\mathcal{X}_r$. Explicitly,

$$\hat{\kappa}_{3,r} = \frac{r^{-1}\sum_{k=\ell+1}^{n}(X_k - \hat{\theta}_r)^3}{\hat{\beta}_r^3}, \qquad \hat{\kappa}_{4,r} = \frac{r^{-1}\sum_{k=\ell+1}^{n}(X_k - \hat{\theta}_r)^4}{\hat{\beta}_r^4} - 3. \tag{6.2}$$

If it is assumed that X_1 has sufficiently many finite moments, it follows by Corollary 6.1 and Lemma 3 in the supplementary material [12] that assumptions (A1)–(A7) are satisfied. The results of Theorem 4.1 therefore hold for the case of the mean, and it follows immediately that the coefficient $C_\theta(z_\alpha)$ is given by

$$C_\theta(z_\alpha) = \frac{1}{8}\{\kappa_4 + 6 + z_\alpha^2(\kappa_4 + 2)\}z_\alpha\phi(z_\alpha).$$

EXAMPLE 6.2 (Backwards percentile bound). Applying Theorem 4.2 in the setting of Example 6.1, it follows readily that

$$K_1(z_\alpha) = \frac{1}{3}\kappa_3(z_\alpha^2 - 1)\phi(z_\alpha) \quad \text{and} \quad K_2(z_\alpha) = \frac{1}{18}\kappa_3^2 z_\alpha(z_\alpha^2 - 1)(z_\alpha^2 - 3)\phi(z_\alpha).$$

The coefficient $C_\theta(z_\alpha)$ is given in Example 6.1. Notice that if, for example, the sample originated from a symmetric distribution, then $\kappa_3 = 0$ and $K_1(z_\alpha) = K_2(z_\alpha) = 0$ so that the two confidence bounds $\hat{I}_H^N(\ell, \alpha)$ and $\hat{I}_B^N(\ell, \alpha)$ have the same order of coverage error.

EXAMPLE 6.3 (Hybrid percentile-t bound). Suppose $X_1, \dots, X_n$ are i.i.d. random variables from an unknown univariate distribution with mean μ and variance $0 < \sigma^2 < \infty$. We again consider the case where the parameter of interest is $\theta = \mu$. Denoting by κ_j the jth cumulant of $(X_1 - \mu)/\sigma$, it is well known (see [10]) that the polynomials q_1 and q_2 in (3.3) are given by

$$q_1(x) = \frac{1}{6}\kappa_3(2x^2 + 1),$$
$$q_2(x) = x\left\{\frac{1}{12}\kappa_4(x^2 - 3) - \frac{1}{18}\kappa_3^2(x^4 + 2x^2 - 3) - \frac{1}{4}(x^2 + 3)\right\}.$$

More recently, the Edgeworth polynomial q_3 has been derived by [8], which is reproduced here in a form more convenient for our purposes:

$$\begin{aligned} q_3(x) = &-\frac{1}{40}\kappa_5(2x^4 + 8x^2 + 1) - \frac{1}{144}\kappa_4\kappa_3(4x^6 - 30x^4 - 90x^2 - 15) \\ &+ \frac{1}{1296}\kappa_3^3(8x^8 + 28x^6 - 210x^4 - 525x^2 - 105) \\ &+ \frac{1}{24}\kappa_3(2x^6 - 3x^4 - 6x^2). \end{aligned}$$

Sample versions $\hat{q}_{1,r}$, $\hat{q}_{2,r}$ and $\hat{q}_{3,r}$ of these polynomials may be obtained by substituting the population cumulants for their respective estimators based on the subsample $\mathcal{X}_r$. $\hat{\kappa}_{3,r}$ and $\hat{\kappa}_{4,r}$ are given in (6.2), and (see [5], page 187)

$$\hat{\kappa}_{5,r} = \frac{r^{-1}\sum_{k=\ell+1}^{n}(X_k - \hat{\theta}_r)^5}{\hat{\beta}_r^5} - 10\hat{\kappa}_{3,r}.$$

By making use of the results of Corollary 6.1, it is a trivial task to show that assumptions (B1)–(B7) in the supplementary material [12] are satisfied. For this example, the coefficient $D_\theta(z_\alpha)$ in Theorem 5.1 is given by

$$D_\theta(z_\alpha) = -\frac{1}{48}\{12\kappa_5 - 15\kappa_4\kappa_3 + 54\kappa_3\}(2z_\alpha^2 + 1)\phi(z_\alpha).$$

Note that $D_\theta(z_\alpha) = 0$ if X_1 has a symmetric distribution. In this case, the order of coverage error of $\hat{J}_H^N(\ell, \alpha)$ will be significantly reduced to $O(\ell^{-2})$, which becomes $O(n^{-2})$ if $\ell = \lfloor \gamma n \rfloor$, $0 < \gamma < 1$. See Remark 5.1.

EXAMPLE 6.4 (Backwards percentile-t bound). As a final example, we apply Theorem 5.2 in the setting of Example 6.3 under the supposition that X_1 has a symmetric distribution. In this case $\kappa_3 = \kappa_5 = 0$, whence $q_1(x) = q_3(x) = 0$, $\forall x \in \mathbb{R}$. Consequently, $K_3(z_\alpha) = K_4(z_\alpha) = K_5(z_\alpha) = D_\theta(z_\alpha) = 0$ so that the coverage error of $\hat{J}_B^N(\ell, \alpha)$ reduces to $O(\ell^{-2})$. See Remark 5.1.

In the next section, we demonstrate that the results of the newly proposed confidence bounds may be extended to the linear regression setup.

7. Linear regression. It has been shown in the literature (see [10]) that the good properties of both the *standard* percentile and percentile-t bootstrap methods carry over to regression problems. For example, confidence bounds for the slope parameter constructed using the traditional methods $\hat{I}_H$ and $\hat{J}$ have reduced coverage errors of $O(n^{-1})$ and $O(n^{-3/2})$, respectively. In this section, we investigate only the performance of our new hybrid percentile-t bound (the two percentile and the backwards percentile-t bounds can be treated similarly) in the linear regression setup. We show that the coverage error of this bound is typically $O(n^{-2})$. To facilitate exposition, we consider only simple linear regression, but the results may be extended to multiple linear regression.

Suppose we observe pairs $\mathcal{X}_n = \{(x_1, Y_1), \ldots, (x_n, Y_n)\}$ generated by the simple linear regression model

$$Y_i = c + (x_i - \bar{x}_n)d + \varepsilon_i,$$

where c and d are unknown, nonrandom constants, $\bar{x}_n = n^{-1}\sum_{i=1}^{n} x_i$, and $\{\varepsilon_1, \ldots, \varepsilon_n\}$ is a sequence of i.i.d. random variables from an unknown distribution with zero mean and constant variance $0 < \sigma^2 < \infty$. Throughout, we assume that the x_i are fixed.

The least-squares estimator for d is $\hat{d}_n = (n\sigma_{x,n}^2)^{-1}\sum_{k=1}^n (x_k - \bar{x}_n)Y_k$, where $\sigma_{x,n}^2 = n^{-1}\sum_{k=1}^n (x_k - \bar{x}_n)^2 > 0$. Furthermore, the estimator for σ^2 is the mean squared residuals, viz. $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{k=1}^n (Y_k - \bar{Y}_n - (x_k - \bar{x}_n)\hat{d}_n)^2$, with $\bar{Y}_n = n^{-1}\sum_{k=1}^n Y_k$. Also, define

$$\begin{aligned}
&\gamma_{x,n} = \frac{1}{n\sigma_{x,n}^3}\sum_{k=1}^n (x_k - \bar{x}_n)^3, \qquad \kappa_{x,n} = \frac{1}{n\sigma_{x,n}^4}\sum_{k=1}^n (x_k - \bar{x}_n)^4 - 3, \\
&\tau_{x,n} = \frac{1}{n\sigma_{x,n}^5}\sum_{k=1}^n (x_k - \bar{x}_n)^5 - 10\gamma_{x,n}.
\end{aligned} \tag{7.1}$$

In [10] it is shown that, if $\limsup_n \max_{1\le i\le n} |x_i - \bar{x}_n| < \infty$, one may obtain the Edgeworth expansion

$$\begin{aligned}
\mathbb{P}\biggl(\frac{n^{1/2}(\hat{d}_n - d)\sigma_{x,n}}{\hat{\sigma}_n} \le x\biggr) &= \Phi(x) + n^{-1/2}q_{1,n}(x)\phi(x) + n^{-1}q_{2,n}(x)\phi(x) \\
&\quad + n^{-3/2}q_{3,n}(x)\phi(x) + \cdots,
\end{aligned} \tag{7.2}$$

uniformly in $x \in \mathbb{R}$, where the $q_{j,n}$ are the appropriate polynomials with coefficients depending on moments of (x_i, Y_i). In particular,

$$\begin{aligned}
q_{1,n}(u) &= -\frac{1}{6}\kappa_3'\gamma_{x,n}He_2(u), \\
q_{2,n}(u) &= -\frac{1}{24}\kappa_4'\kappa_{x,n}He_3(u) - \frac{1}{72}(\kappa_3')^2\gamma_{x,n}^2 He_5(u) - \frac{1}{4}(u^2+5)u,
\end{aligned} \tag{7.3}$$

with κ_j' denoting the jth cumulant of ε_1/σ and $He_j(u)$ the jth Hermite polynomial. We shall also require the third Edgeworth polynomial $q_{3,n}$, which apparently does not appear in the existing literature. It may be shown by laborious algebra (see Lemma 4 in the supplementary material [12]) that

$$\begin{aligned}
q_{3,n}(u) = &-\frac{1}{120}\kappa_5'\{\tau_{x,n}He_4(u) - 30\gamma_{x,n}He_2(u)\} \\
&- \frac{1}{144}\kappa_4'\kappa_3'\{\kappa_{x,n}\gamma_{x,n}He_6(u) + 45\gamma_{x,n}He_2(u)\} \\
&- \frac{1}{1296}(\kappa_3')^3\gamma_{x,n}^3 He_8(u) - \frac{1}{24}\kappa_3'\gamma_{x,n}(u^2-1)u^4.
\end{aligned} \tag{7.4}$$

We may now construct our new hybrid percentile-t confidence bound for d. As before, split the original sample in two disjoint subsets

$$\mathcal{X}_\ell = \{(x_1, Y_1), \dots, (x_\ell, Y_\ell)\} \quad \text{and} \quad \mathcal{X}_r = \{(x_{\ell+1}, Y_{\ell+1}), \dots, (x_n, Y_n)\},$$

for some integer $2 \le \ell \le n-2$. Writing $\sigma_{x,\ell}^2 = \ell^{-1}\sum_{k=1}^\ell (x_k - \bar{x}_\ell)^2$, with $\bar{x}_\ell = \ell^{-1}\sum_{k=1}^\ell x_k$, the least squares estimators (based solely on $\mathcal{X}_\ell$) for d and c are given by $\hat{d}_\ell = (\ell\sigma_{x,\ell}^2)^{-1}\sum_{k=1}^{\ell}(x_k - \bar{x}_\ell)Y_k$, and $\hat{c}_\ell = \bar{Y}_\ell$, where $\bar{Y}_\ell = \ell^{-1}\sum_{k=1}^{\ell} Y_k$. Let $\gamma_{x,r}$, $\kappa_{x,r}$ and $\tau_{x,r}$ be the same functions of $\mathcal{X}_r$ as $\gamma_{x,n}$, $\kappa_{x,n}$ and $\tau_{x,n}$ are of $\mathcal{X}_n$. Also, define $\gamma_{x,\ell}$, $\kappa_{x,\ell}$, $\tau_{x,\ell}$ as functions of $\mathcal{X}_\ell$.

Since the variance of $\hat{d}_\ell$ is $\sigma^2/(\ell\sigma_{x,\ell}^2)$, the new $(1-\alpha)$-level percentile-t confidence bound for d (corresponding to $\hat{J}_H^N$) is given by

$$\hat{K}_H^N(m,\alpha) := \left(-\infty, \hat{d}_\ell - \ell^{-1/2}\sigma_{x,\ell}^{-1}\hat{\sigma}_\ell\tilde{\eta}_{m,r,\alpha}\right],$$

where $\hat{\sigma}_\ell^2 = \ell^{-1}\sum_{k=1}^{\ell}\hat{\varepsilon}_k^2 := \ell^{-1}\sum_{k=1}^{\ell}(Y_k - \bar{Y}_\ell - (x_k - \bar{x}_\ell)\hat{d}_\ell)^2$, and

$$\tilde{\eta}_{m,r,\alpha} := z_\alpha + m^{-1/2}\hat{q}_{1,r}^{cf}(z_\alpha) + m^{-1}\hat{q}_{2,r}^{cf}(z_\alpha) + m^{-3/2}\hat{q}_{3,r}^{cf}(z_\alpha).$$

The Cornish–Fisher polynomials $\hat{q}_{j,r}^{cf}$ appearing in this expression are completely determined by the Edgeworth polynomials $\hat{q}_{j,r}$ through the relations given in Lemma 1 in the supplementary material [12], where $\hat{q}_{j,r}$ are given by

$$\begin{aligned}
\hat{q}_{1,r}(u) &= -\frac{1}{6}\hat{\kappa}_{3,r}'\gamma_{x,r}He_2(u),\\
\hat{q}_{2,r}(u) &= -\frac{1}{24}\hat{\kappa}_{4,r}'\kappa_{x,r}He_3(u) - \frac{1}{72}\left(\hat{\kappa}_{3,r}'\right)^2\gamma_{x,r}^2He_5(u) - \frac{1}{4}\left(u^2+5\right)u,\\
\hat{q}_{3,r}(u) &= -\frac{1}{120}\hat{\kappa}_{5,r}'\left\{\tau_{x,r}He_4(u) - 30\gamma_{x,r}He_2(u)\right\}\\
&\quad - \frac{1}{144}\hat{\kappa}_{4,r}'\hat{\kappa}_{3,r}'\left\{\kappa_{x,r}\gamma_{x,r}He_6(u) + 45\gamma_{x,r}He_2(u)\right\}\\
&\quad - \frac{1}{1296}\left(\hat{\kappa}_{3,r}'\right)^3\gamma_{x,r}^3He_8(u) - \frac{1}{24}\hat{\kappa}_{3,r}'\gamma_{x,r}\left(u^2-1\right)u^4,
\end{aligned}$$

with $m_{j,r}' = r^{-1}\sum_{k=\ell+1}^{n}\hat{\varepsilon}_k^j$, $\hat{\kappa}_{3,r}' = m_{3,r}'(m_{2,r}')^{-3/2}$, $\hat{\kappa}_{4,r}' = m_{4,r}'(m_{2,r}')^{-2} - 3$, and $\hat{\kappa}_{5,r}' = m_{5,r}'(m_{2,r}')^{-5/2} - 10\hat{\kappa}_{3,r}'$.

THEOREM 7.1. *Suppose that ε_1 has sufficiently many finite moments and satisfies Cramér's condition. Assume* $\limsup_{n\to\infty}\max_{1\le i\le n}|x_i - \bar{x}_n| < \infty$, $\gamma_{x,r} - \gamma_{x,\ell} = O(n^{-(1+\delta)})$ *for some* $\delta > 0$, $\kappa_{x,r} - \kappa_{x,\ell} = O(n^{-1})$, *and* $\tau_{x,r} - \tau_{x,\ell} = O(n^{-1})$. *Then, if* $m = \ell = O(r)$ *and* $\ell \to \infty$ *as* $n \to \infty$, *we have that*

$$\mathbb{P}\left(d \in \hat{K}_H^N(\ell,\alpha)\right) = 1 - \alpha + \frac{E_d(z_\alpha)}{\ell^{1/2}r} + O\left(\ell^{-2} + \ell^{-1/2}n^{-(1+\delta)}\right), \tag{7.5}$$

with $E_d(z_\alpha) = \frac{1}{48}\gamma_{x,r}(12\kappa_5' - 15\kappa_4'\kappa_3' + 66\kappa_3')(z_\alpha^2 - 1)\phi(z_\alpha)$, *where* κ_j' *denotes the* j*th cumulant of* ε_1/σ. *Moreover, if we choose* $\ell = \lfloor\gamma n^\psi\rfloor$ *for some* $\gamma > 0$

and $\frac{2}{3} < \psi < 1$, *then*

$$\begin{aligned}
&\mathbb{P}\big(d \in \hat{K}_H^N(\ell,\alpha)\big)\\
&\quad = \begin{cases} 1-\alpha + \dfrac{E_d(z_\alpha)}{\gamma^{1/2} n^{(2+\psi)/2}} + O\big(n^{-\min\{2-\psi/2, 2\psi, 1+\delta+\psi/2\}}\big) & \text{if } E_d(z_\alpha) \neq 0,\\ 1-\alpha + O\big(n^{-2\psi} + n^{-(1+\delta+\psi/2)}\big) & \text{if } E_d(z_\alpha) = 0. \end{cases}
\end{aligned}$$

In the case where $\psi = 1$ *and* $0 < \gamma < 1$,

$$\mathbb{P}\big(d \in \hat{K}_H^N(\ell,\alpha)\big) = 1-\alpha + \frac{E_d(z_\alpha)}{\gamma^{1/2}(1-\gamma)n^{3/2}} + O\big(n^{-2} + n^{-(3/2+\delta)}\big),$$

which becomes $\mathbb{P}(d \in \hat{K}_H^N(\ell,\alpha)) = 1-\alpha + O(n^{-2} + n^{-(3/2+\delta)})$ *if* ε_1 *has a symmetric distribution around zero.*

REMARK 7.1. If the design points are regularly spaced, say $x_i = u\frac{i}{n} + v$, $i = 1,\dots,n$, for some constants u and v, then the assumptions on the x_i in Theorem 7.1 can easily be verified. In fact, since in this case $\gamma_{x,r} = \gamma_{x,\ell} = 0$, we can take $\delta = \infty$. Consequently, $E_d(z_\alpha) = 0$ so that the coverage error reduces to $O(n^{-2})$, even if the errors have an asymmetric distribution.

8. Equal-tailed confidence intervals. The one-sided upper and lower confidence bounds may be used to construct equal-tailed confidence intervals. For example, in the notation of Section 2 the standard bootstrap percentile-t $(1-2\alpha)$-level confidence interval for θ is given by

$$\hat{J}(\alpha) \setminus \hat{J}(1-\alpha) = \big(\hat{\theta}_n - n^{-1/2}\hat{\sigma}_n\hat{\eta}_{n,1-\alpha}, \hat{\theta}_n - n^{-1/2}\hat{\sigma}_n\hat{\eta}_{n,\alpha}\big].$$

The order of coverage error of this interval is typically $O(n^{-1})$, except in the case where $\kappa_3 = \kappa_4 = 0$, which reduces the error to $O(n^{-2})$ (see [9], page 949). Moreover, [9] shows that equal-tailed confidence intervals constructed from $\hat{I}_H$ and $\hat{I}_B$, as well as intervals constructed from the bias-corrected and accelerated bias-corrected bounds, also incur coverage errors of order $O(n^{-1})$.

We now show that equal-tailed confidence intervals with a reduced coverage error of $O(n^{-2})$ may be obtained using the newly proposed hybrid percentile-t bound $\hat{J}_H^N$, without the assumption that $\kappa_3 = \kappa_4 = 0$. We have from Theorem 5.1 that

$$\mathbb{P}\big(\theta \in \hat{J}_H^N(\ell,\alpha) \setminus \hat{J}_H^N(\ell,1-\alpha)\big) = 1-2\alpha + \frac{D_\theta(z_\alpha) - D_\theta(z_{1-\alpha})}{\ell^{1/2}r} + O\big(\ell^{-2}\big).$$

Recalling that ϕ, q_1 and $\hat{q}_{1,r}$ are even functions, it follows immediately from (5.2) that $D_\theta(z_\alpha) = D_\theta(z_{1-\alpha})$, so that

$$\mathbb{P}\big(\theta \in \hat{J}_H^N(\ell,\alpha) \setminus \hat{J}_H^N(\ell,1-\alpha)\big) = 1-2\alpha + O\big(\ell^{-2}\big).$$

If we now choose $\ell = \lfloor \gamma n^{\psi} \rfloor$ for some $\gamma > 0$ and $\frac{2}{3} < \psi \leq 1$, then

$$\mathbb{P}\big(\theta \in \hat{J}_H^N(\ell, \alpha) \setminus \hat{J}_H^N(\ell, 1-\alpha)\big) = 1 - 2\alpha + O\big(n^{-2\psi}\big).$$

A similar argument may be used to show that equal-tailed confidence intervals with coverage error of order $O(n^{-1})$ can be constructed from the other newly proposed types of bounds $\hat{I}_H^N$, $\hat{I}_B^N$ and $\hat{J}_B^N$. In contrast to one-sided confidence *bounds* constructed by means of the backwards method, additional assumptions (such as symmetry) are not needed to achieve this order of coverage error (see Example 6.2).

Similar confidence intervals can be constructed for the slope parameter in the linear regression model of Section 7. Coverage errors of $O(n^{-2})$ and even smaller (in the case of symmetric errors) can be obtained.

9. Simulation study. A modest simulation study was carried out to compare the standard upper bounds $\hat{I}_H$, $\hat{I}_B$, $\hat{J}$ and the upper bound proposed by Chung and Lee [4], which we denote by C-L, with the newly developed upper bounds $\hat{I}_H^N$, $\hat{I}_B^N$, $\hat{J}_H^N$ and $\hat{J}_B^N$, where the parameter of interest is the population mean. Monte Carlo estimates were calculated for the non-coverage probability (NC) and expected size of the upper bound (EUB) resulting from each method. We considered the performance of the different bounds for samples of sizes $n = 50, 100, 200$ drawn from the uniform(0, 1), standard Laplace, χ_3^2 and $F_{5,8}$ distributions. The new bounds were evaluated for $\alpha = 5\%$ and different choices of ℓ such that the assumption $\ell = O(r)$ required by the theorems is satisfied. Each entry in Tables 1–5 is based on 100,000 independent Monte Carlo trials, each comprising 10,000 bootstrap samples. Standard errors were found to be negligibly small and are not reported. All calculations were done in R.

Recall that for distributions with $\kappa_3 = 0$ the standard *percentile* bounds $\hat{I}_H$ and $\hat{I}_B$ have coverage errors of order $O(n^{-1})$ (see [9]), which is of the same order as the coverage errors produced by the newly proposed percentile bounds $\hat{I}_H^N$ and $\hat{I}_B^N$. Therefore, for the two symmetric distributions we report in Tables 1 and 2 results only for the *percentile-t* type bounds $\hat{J}$ and $\hat{J}_H^N$, which have coverage errors of order $O(n^{-1})$ and $O(n^{-2})$, respectively. We omit the results for $\hat{J}_B^N$, since its behavior is almost identical to that of $\hat{J}_H^N$ (see Example 6.4). We do not consider distributions with $\kappa_4 = 0$ (e.g., the normal distribution), since in this case the various

TABLE 1
Results of the existing percentile-t method $\hat{J}$ for two symmetric distributions

	n = 50		**n = 100**		**n = 200**	
Distribution	**NC**	**EUB**	**NC**	**EUB**	**NC**	**EUB**
Uniform	0.045	0.568	0.048	0.548	0.049	0.534
Laplace	0.059	0.334	0.056	0.234	0.054	0.165

TABLE 2
Results of the new hybrid percentile-t method $\hat{J}_H^N$ for two symmetric distributions

	$n = 50$			$n = 100$			$n = 200$		
Distribution	ℓ	**NC**	**EUB**	ℓ	**NC**	**EUB**	ℓ	**NC**	**EUB**
Uniform	25	0.050	0.598	50	0.050	0.568	100	0.050	0.548
	30	0.050	0.589	60	0.050	0.562	120	0.050	0.544
	35	0.051	0.582	70	0.050	0.557	140	0.050	0.540
	40	0.050	0.577	80	0.050	0.554	160	0.050	0.538
Laplace	30	0.050	0.436	60	0.050	0.304	120	0.050	0.213
	35	0.051	0.402	70	0.050	0.281	140	0.050	0.197
	40	0.050	0.375	80	0.050	0.263	160	0.050	0.185
	45	0.050	0.352	90	0.050	0.248	180	0.050	0.174

confidence bounds have almost identical performance in terms of coverage error. For the uniform and Laplace distributions $\kappa_4 = -1.2$ and $\kappa_4 = 3$, respectively.

Comparing Tables 1 and 2 it is evident that, for both the uniform and Laplace distributions, the new bound $\hat{J}_H^N$ significantly outperforms the standard percentile-t bound $\hat{J}$ in terms of coverage error for all sample sizes considered. This striking performance is visible even for a relatively small sample. Although the upper bound $\hat{J}_H^N$ is slightly larger than $\hat{J}$ in each case (as expected), a suitable choice of ℓ greatly diminishes this difference. Note that a larger choice of ℓ corresponds to a smaller upper bound, which agrees with the definition of $\hat{J}_H^N$.

The results for the skewed distributions presented in Tables 3–5 show that for most choices of ℓ the newly proposed *percentile* bounds $\hat{I}_H^N$ and $\hat{I}_B^N$ significantly outperform the standard percentile bounds $\hat{I}_H$ and $\hat{I}_B$ in terms of coverage error.

TABLE 3
Results of the existing methods for two skewed distributions

		$n = 50$		$n = 100$		$n = 200$	
Distribution	**Type**	**NC**	**EUB**	**NC**	**EUB**	**NC**	**EUB**
χ_3^2	$\hat{I}_H$	0.092	3.537	0.077	3.388	0.068	3.278
	$\hat{I}_B$	0.080	3.576	0.068	3.390	0.062	3.289
	C-L	0.064	3.641	0.057	3.436	0.053	3.304
	$\hat{J}$	0.056	3.674	0.052	3.453	0.051	3.309
$F_{5,8}$	$\hat{I}_H$	0.135	1.612	0.112	1.539	0.096	1.484
	$\hat{I}_B$	0.115	1.650	0.097	1.562	0.084	1.498
	C-L	0.090	1.732	0.079	1.599	0.070	1.517
	$\hat{J}$	0.080	1.772	0.070	1.627	0.064	1.531

TABLE 4
Results of the new methods for the χ_3^2 distribution

	$n = 50$			$n = 100$			$n = 200$		
Type	ℓ	**NC**	**EUB**	ℓ	**NC**	**EUB**	ℓ	**NC**	**EUB**
$\hat{I}_H^N$	20	0.065	3.820	40	0.058	3.599	80	0.054	3.433
	25	0.068	3.734	50	0.059	3.536	100	0.055	3.388
	30	0.074	3.667	60	0.062	3.489	120	0.055	3.354
	35	0.081	3.610	70	0.065	3.451	140	0.058	3.327
$\hat{I}_B^N$	20	0.050	3.909	40	0.046	3.649	80	0.045	3.459
	25	0.055	3.802	50	0.049	3.575	100	0.047	3.408
	30	0.061	3.720	60	0.052	3.520	120	0.049	3.371
	35	0.070	3.651	70	0.057	3.476	140	0.051	3.341
$\hat{J}_H^N$	20	0.059	4.174	40	0.053	3.770	80	0.050	3.514
	25	0.059	4.003	50	0.053	3.669	100	0.051	3.451
	30	0.060	3.883	60	0.054	3.597	120	0.051	3.407
	35	0.062	3.791	70	0.054	3.543	140	0.052	3.372

Furthermore, it is clear that the bound C-L, which also has coverage error $O(n^{-1})$, performs slightly better than $\hat{I}_H^N$, but slightly worse than $\hat{I}_B^N$. The performance of the new *percentile-t* bound $\hat{J}_H^N$ is comparable to that of the standard percentile-t bound $\hat{J}$. We omit the results for $\hat{J}_B^N$, as its coverage error $O(n^{-1/2})$ compares

TABLE 5
Results of the new methods for the $F_{5,8}$ distribution

	$n = 50$			$n = 100$			$n = 200$		
Type	ℓ	**NC**	**EUB**	ℓ	**NC**	**EUB**	ℓ	**NC**	**EUB**
$\hat{I}_H^N$	20	0.088	1.748	40	0.073	1.644	80	0.065	1.563
	25	0.096	1.706	50	0.078	1.612	100	0.068	1.540
	30	0.105	1.671	60	0.085	1.588	120	0.072	1.522
	35	0.118	1.640	70	0.093	1.566	140	0.077	1.507
$\hat{I}_B^N$	20	0.063	1.828	40	0.052	1.695	80	0.048	1.594
	25	0.075	1.764	50	0.060	1.650	100	0.053	1.563
	30	0.087	1.715	60	0.069	1.617	120	0.059	1.540
	35	0.103	1.672	70	0.080	1.589	140	0.066	1.521
$\hat{J}_H^N$	20	0.074	2.070	40	0.062	1.830	80	0.057	1.666
	25	0.078	1.937	50	0.066	1.748	100	0.060	1.616
	30	0.083	1.848	60	0.069	1.693	120	0.062	1.582
	35	0.088	1.781	70	0.073	1.651	140	0.064	1.556

poorly to the error $O(n^{-3/2})$ attained by $\hat{J}_H^N$ (see Theorem 5.2). Again, the size of the upper bound can be decreased with an appropriate choice of ℓ. Notice that, in agreement with theory, the coverage errors of all considered bounds converge to the nominal coverage error α as the sample size n is increased.

Interestingly, the simulation study shows that the coverage of the backwards percentile bound $\hat{I}_B^N$ seems to be better than that of the hybrid percentile bound $\hat{I}_H^N$ for the skewed distributions χ_3^2 and $F_{5,8}$. However, this does not contradict the results derived in Theorems 4.1 and 4.2. The main reason behind this observation appears to be the magnitude of the constants $K_1(z_\alpha)$ and $C_\theta(z_\alpha)$ appearing in the theorems relative to the sample sizes chosen in this study. Similarly, in the case of the χ_3^2 distribution, the slight underperformance of the proposed percentile-t bound $\hat{J}_H^N$ when compared to the standard percentile-t bound $\hat{J}$ can be ascribed to the fact that the constant $D_\theta(z_\alpha)$ in Theorem 5.1 is relatively large, but its effect on coverage diminishes quickly as the sample size increases. A more detailed discussion on these two observations is given in Section 2 of the supplementary material [12].

Overall, it is clear that the improvement in coverage accuracy comes at the cost of a larger upper bound. However, by making a suitable choice of ℓ when splitting the sample one may achieve a significantly improved coverage probability with only a slight increase in the magnitude of the upper bound. Ideally, a data-based choice of ℓ is needed which, however, will require deeper analysis and we leave a detailed study for future research.

Acknowledegments. The authors would like to thank the Associate Editor and three referees for their insightful and constructive comments which led to significant improvement of the paper.

SUPPLEMENTARY MATERIAL

Supplement to "On the asymptotic theory of new bootstrap confidence bounds" (DOI: 10.1214/17-AOS1557SUPP; .pdf). In the online supplement [12], we supply proofs for all theorems found in the main text.

REFERENCES

[1] BHATTACHARYA, R. N. and GHOSH, J. K. (1978). On the validity of the formal Edgeworth expansion. *Ann. Statist.* **6** 434–451. MR0471142

[2] BICKEL, P. J., GÖTZE, F. and VAN ZWET, W. R. (1997). Resampling fewer than n observations: Gains, losses, and remedies for losses. *Statist. Sinica* **7** 1–31. MR1441142

[3] CHANG, C. C. and POLITIS, D. N. (2011). Bootstrap with larger resample size for root-n consistent density estimation with time series data. *Statist. Probab. Lett.* **81** 652–661. MR2783862

[4] CHUNG, K.-H. and LEE, S. M. S. (2001). Optimal bootstrap sample size in construction of percentile confidence bounds. *Scand. J. Stat.* **28** 225–239. MR1844358

[5] CRAMÉR, H. (1946). *Mathematical Methods of Statistics*. *Princeton Mathematical Series* **9**. Princeton Univ. Press, Princeton, NJ. MR0016588

[6] EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann*. *Statist*. **7** 1–26. MR0515681

[7] EFRON, B. and TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap*. *Monographs on Statistics and Applied Probability* **57**. Chapman & Hall, New York. MR1270903

[8] FINNER, H. and DICKHAUS, T. (2010). Edgeworth expansions and rates of convergence for normalized sums: Chung's 1946 method revisited. *Statist*. *Probab. Lett*. **80** 1875–1880. MR2734254

[9] HALL, P. (1988). Theoretical comparison of bootstrap confidence intervals. *Ann*. *Statist*. **16** 927–985. MR0959185

[10] HALL, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer, New York. MR1145237

[11] HURT, J. (1976). Asymptotic expansions of functions of statistics. *Apl*. *Mat*. **21** 444–456. MR0418309

[12] PRETORIUS, C. and SWANEPOEL, J. W. H. (2018). Supplement to "On the asymptotic theory of new bootstrap confidence bounds". DOI:10.1214/17-AOS1557SUPP.

[13] SWANEPOEL, J. W. H. (1986). A note on proving that the (modified) bootstrap works. *Comm*. *Statist*. *Theory Methods* **15** 3193–3203. MR0860478

DEPARTMENT OF STATISTICS
NORTH-WEST UNIVERSITY
POTCHEFSTROOM
SOUTH AFRICA
E-MAIL: cpretorius@gmail.com
jan.swanepoel@nwu.ac.za