A THREE-STAGE PROCEDURE BASED ON BOOTSTRAP CRITICAL POINTS

M. Aerts

Limburgs Universitair Centrum,
Diepenbeek, Belgium.

I. Gijbels

Catholic University Louvain,
Louvain-la-Neuve, Belgium.

Key words and Phrases: asymptotic consistency; asymptotic efficiency; bootstrap critical points; fixed-width confidence intervals; population mean; simulations; studentized statistic; three-stage procedure.

ABSTRACT

The construction of fixed-width confidence intervals for an unknown population parameter is often based on the normal approximation of an appropriate statistic. In this paper we follow a different approach using bootstrap approximations, which are known to have a high degree of accuracy. It is expected that such an approximation leads to better results, at least in the second order behaviour. In order to avoid computationally very costly procedures, inherent to a fully sequential method, we implement the bootstrap ideas in a three-stage procedure. Simulation results illustrate the performance of the proposed procedure for small and moderate sample sizes. To support the method we establish some basic asymptotic properties.

1 INTRODUCTION

Let $X_1, X_2, ...$ be a sequence of i.i.d. random variables from a population with distribution function (d.f.) F. Consider $\theta = \theta(F)$, the unknown (scalar) parameter of interest and let $\hat{\theta}_n = \hat{\theta}(X_1, ..., X_n)$ be an appropriate estimator for θ. Assume that $E(\hat{\theta}_n^2) < \infty$ and that for some (unknown) constant $\sigma > 0$, as $n \to \infty$,

$$\frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\sigma} \xrightarrow{d} N(0,1) \tag{1}$$

Our objective is to construct a fixed-width ($2d$) confidence interval with coverage probability $1-\alpha$, where $d > 0$ and $0 < \alpha < 1$. We focus our attention to intervals of the form

$$I_{n,d} = [\hat{\theta}_n - d, \hat{\theta}_n + d]. \tag{2}$$

Essentially, the problem of constructing such an interval is translated into a problem of determining the sample size. Indeed, for a sample size n which is too small, the interval $I_{n,d}$ will not reach the coverage probability $1 - \alpha$. There are basically two approaches to obtain a fixed-width confidence interval with prescribed coverage probability, and we now briefly describe them.

First of all there is the fully sequential procedure which was initially introduced by Anscombe (1952) and Chow and Robbins (1965) for θ equal to the mean of a random variable. In the above more general setting this procedure can be formulated as follows. Suppose that we want to determine the sample size n such that the coverage probability $P\{\theta \in I_{n,d}\}$ attains the given value $1 - \alpha$. Using the asymptotic normality result in (1), we find that

$$P_n \equiv P\{\theta \in I_{n,d}\} \approx \Phi(\sqrt{n}d/\sigma) - \Phi(-\sqrt{n}d/\sigma) = 2\Phi(\sqrt{n}d/\sigma) - 1, \tag{3}$$

and an approximate solution to the problem is provided by taking the smallest integer $n \geq 1$ such that

$$2\Phi(\sqrt{n}d/\sigma) - 1 \geq 1 - \alpha,$$

and this is equivalent with choosing the smallest $n \geq 1$ such that

$$n \geq (\frac{z\sigma}{d})^2,$$

where $z = \Phi^{-1}(1-\alpha/2)$, the $(1-\alpha/2)$-th quantile of the standard normal d.f. $\Phi(.)$. Since in general σ is not known, it is replaced by a consistent estimator, denoted by $\hat{\sigma}_n$. This results into the random stopping time introduced by Chow and Robbins (1965) (for the mean), i.e.

$$N_{CR}(d) = \inf\{n \geq n_0 : n \geq (\frac{z\hat{\sigma}_n}{d})^2\}, \tag{4}$$

where n_0 is some initial sample size.

The above fully sequential procedure involves a lot of computational effort. Stein (1945) introduced a sampling procedure which requires only two sampling operations. However, it turned out that this two-stage procedure is less efficient than the fully sequential one. Adding one extra stage to Stein's method results into a method which combines simplicity with the efficiency of the fully sequential procedure. The idea of considering three sampling operations instead of two was introduced by Hall(1981) for the normal mean.

A general three-stage procedure can be defined as follows (see e.g. Vik and Mukhopadhyay(1988)):

stage 1: take an initial pilot-sample $X_1, ..., X_{n_0}$ with the deterministic sample size

$$n_0(d) = \max\{[(\frac{z}{d})^{2\eta}] + 1, 2\}$$

where $0 < \eta < 1$ is a parameter determining the order of magnitude of this initial sample size. Here, and in the following definitions of the various sample sizes, $[x]$ denotes the greatest integer part of x;

stage 2: take additional $N_1 - n_0$ observations to get an intermediate sample $X_1, ..., X_{N_1}$ with

$$N_1(d) = \max\{[c(\frac{z\hat{\sigma}_{n_0}}{d})^2] + 1, n_0\}$$

where $0 < c < 1$ is a parameter controlling the difference $N_1 - n_0$ at this second stage;

stage 3: sample $N_2 - N_1$ more observations to obtain the final sample $X_1, ..., X_{N_2}$ with

$$N_2(d) = \max\{[(\frac{z\hat{\sigma}_{N_1}}{d})^2] + 1, N_1\};$$

finally: construct the $2d$-width confidence interval for θ :

$$[\hat{\theta}_{N_2} - d, \hat{\theta}_{N_2} + d].$$

The asymptotic properties of this three-stage procedure have been extensively studied for the population mean. For a thorough study of the triple sampling procedure we refer to Hall (1981), Hamdy(1988), Vik and Mukhopadhyay(1988), Lohr (1990) and Honda (1992).

Both, the fully sequential and the three-stage procedure, rely on the normal approximation given in (1). Hence, the accuracy of those procedures heavily depends on how good this normal approximation is. A different approach, using bootstrap ideas instead of the normal approximation, is introduced by Swanepoel, van Wijk and Venter (1984). We briefly describe their procedure. With P_n as in (3), consider

$$\inf\{n \geq n_0 : P_n \geq 1 - \alpha\},$$

and note that P_n depends on the d.f. F and hence is unknown. Therefore one estimates P_n at stage n by some estimate $\hat{P}_n$, say, and this leads to the stopping time

$$\inf\{n \geq n_0 : \hat{P}_n \geq 1 - \alpha\}.$$

If the normal approximation, given in (1), is used one ends up with estimating P_n by $2\Phi(\sqrt{n}d/\hat{\sigma}_n) - 1$ (see also (3)), and it is easily seen that the resulting stopping time reduces to the stopping time $N_{CR}(d)$ defined in (4).

Swanepoel, van Wijk and Venter (1984) suggest to use the bootstrap-estimator for P_n, denoted by $\hat{P}_n^*$, and defined as follows. Given $X_1, X_2, ..., X_n$, let $X_1^*, X_2^*, ..., X_n^*$ be n conditionally independent observations with common d.f.

$$F_n(.) = \frac{1}{n} \sum_{i=1}^{n} I\{X_i \leq .\}.$$

The estimator $\hat{\theta}_n$ evaluated in the bootstrap sample $X_1^*, ..., X_n^*$ will be denoted by $\hat{\theta}_n^*$. Further, P_n^* denotes the conditional probability under F_n.
Then

$$\hat{P}_n^* \equiv P_n^*\{\hat{\theta}_n^* - d \leq \hat{\theta}_n \leq \hat{\theta}_n^* + d\} = P_n^*\{-d \leq \hat{\theta}_n^* - \hat{\theta}_n \leq d\},$$

which leads to the bootstrap stopping time

$$N_{SWV}(d) = \inf\{n \geq n_0 : \hat{P}_n^* \geq 1 - \alpha\}, \tag{5}$$

studied in Swanepoel, van Wijk and Venter (1984) for the case $\hat{\theta}_n = \bar{X}_n$, the sample mean, and the case that $\hat{\theta}_n$ is the sample median.

The implementation of a bootstrap approximation in a fully sequential procedure is, from a computational point of view, not very attractive. Indeed, the calculation (simulation) of the bootstrap probability $\hat{P}_n^*$ increases the computation time seriously, and this extra computational effort has to be done in every sequential step! In this paper, we will propose a method which combines bootstrap ideas with the three-stage procedure. Hence, we try to implement the advantage of using the bootstrap approximation instead of the normal approximation, but at the same time we wish to avoid the lengthy computations which are inherent to the fully sequential method. It is well-known that, for a wide class of statistics, the bootstrap approximation has a high degree of accuracy. Therefore, it is expected that the proposed procedure will perform at least equally well as the classical three-stage method based on the normal approximation. Furthermore, the bootstrap approximation essentially corrects for the skewness of the sampling distribution. Hence, an appropriate bootstrap-based procedure should outperform the classical method in certain cases.

The paper is organized as follows. In Section 2 we introduce and motivate, in the above general parameter setting, our procedure for constructing $2d$-width confidence intervals with (approximate) coverage probability $1-\alpha$. First order asymptotic properties, the asymptotic consistency and efficiency of the procedure, are established in Section 3 for the case that $\theta = E(X)$ and $\hat{\theta}_n = \bar{X}_n$. The performance of the method for small and moderate sample sizes is investigated in Section 4, which contains the results of some Monte-Carlo simulations.

2 TRIPLE SAMPLING BOOTSTRAP STOPPING TIMES TO CONSTRUCT FIXED-WIDTH CONFIDENCE INTERVALS

We will now introduce our method. Denote by

$$\mathcal{L}_n(x) = P\{\mid \frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\sigma} \mid \leq x\} \tag{6}$$

the d.f. of the "symmetrized" and normalized estimator $\hat{\theta}_n$. Our preference for working with this "symmetrized" version can be translated into a preference for considering symmetric two-sided bootstrap-type confidence intervals instead of equal-tailed bootstrap-type confidence intervals. Indeed, Hall (1988a) showed that symmetric intervals have better coverage accuracy than equal-tailed intervals. Moreover, they are often shorter than equal-tailed intervals, even in highly asymmetric circumstances. For an interesting and detailed discussion on these issues we refer to Hall (1988a). Now, the optimal choice of sample size n_{opt}, assuring that the interval $I_{n,d}$, defined in (2), reaches the coverage probability $1-\alpha$, satisfies the equation:

$$\mathcal{L}_{n_{opt}}(\frac{\sqrt{n_{opt}}d}{\sigma}) = 1-\alpha$$

which leads to the implicit solution-equation:

$$n_{opt} = (\frac{\xi_{n_{opt}}(\alpha)\sigma}{d})^2 \tag{7}$$

where

$$\xi_n(\alpha) = \inf\{x : \mathcal{L}_n(x) \geq 1-\alpha\} \tag{8}$$

i.e. the $(1-\alpha)$-th quantile of the d.f. $\mathcal{L}_n$.

Equation (7) contains two unknown quantities : the unknown variance-parameter σ and the critical point (c.p.) $\xi_{n_{opt}}(\alpha)$. The fully sequential procedure and the three-stage procedure, as described in Section 1, suggest to substitute the variance-parameter by its estimator $\hat{\sigma}_{n'}$, based on a smaller available sample $X_1, ..., X_{n'}$ of size $n' < n_{opt}$, and to replace the unknown c.p. by its asymptotic value z, leading to

$$\hat{n}_{opt} = (\frac{z\hat{\sigma}_{n'}}{d})^2. \tag{9}$$

Instead of using the normal approximation (1), we suggest to use bootstrap c.p.'s $\xi^*_{n'}(\alpha)$ in order to estimate the unknown quantity $\xi_{n_{opt}}(\alpha)$.

Conditionally given $X_1, X_2, ..., X_{n'}$, let $X^*_1, X^*_2, ..., X^*_{n'}$ be n' independent r.v.'s with common d.f. $F_{n'}$, i.e. the bootstrap sample drawn with replacement from $F_{n'}$. Let $P^*_{n'}$ denote the conditional probability measure corresponding to $F_{n'}$. The estimator $\hat{\sigma}_{n'}$ evaluated in the bootstrap sample $X^*_1, ..., X^*_{n'}$ will be denoted by $\hat{\sigma}^*_{n'}$. One can basically think of three solution strategies to handle equation (7), depending on the particular bootstrap distribution one is starting from:

- Using a bootstrap approximation for the "Centered but non-standardized" estimator:

$$\sup_t \mid P^*_{n'}\{\sqrt{n'}(\hat{\theta}^*_{n'} - \hat{\theta}_{n'}) \leq t\} - P\{\sqrt{n'}(\hat{\theta}_{n'} - \theta) \leq t\} \mid \xrightarrow{\text{a.s. [P]}} 0 \quad , \quad n' \to \infty$$

then the bootstrap equivalent of equation (7) is:

$$n^*_{opt} = (\frac{\xi^*_{\mathrm{C},n'}(\alpha)}{d})^2 \tag{10}$$

where $\xi^*_{\mathrm{C},n'}(\alpha)$ is the $(1-\alpha)$-th quantile of the (conditional) d.f. of $|\sqrt{n'}(\hat{\theta}^*_{n'} - \hat{\theta}_{n'})|$ (based on the smaller available sample $X_1, X_2, ..., X_{n'}$). One hopes that the bootstrap estimator $\xi^*_{\mathrm{C},n'}(\alpha)$ will catch the unknown quantity $\xi_{n_{opt}}(\alpha)\sigma$ at least as good as the estimator $z\hat{\sigma}_{n'}$ does. Note that this approach does not require an estimator for the parameter σ, since this estimation part is included in the estimator $\xi^*_{\mathrm{C},n'}(\alpha)$.

- Using a bootstrap approximation for the "stAndardized" estimator:

$$\sup_t \mid P^*_{n'}\{\frac{\sqrt{n'}(\hat{\theta}^*_{n'} - \hat{\theta}_{n'})}{\hat{\sigma}_{n'}} \leq t\} - P\{\frac{\sqrt{n'}(\hat{\theta}_{n'} - \theta)}{\sigma} \leq t\} \mid \xrightarrow{\text{a.s. [P]}} 0 \quad , \quad n' \to \infty$$

then the bootstrap equivalent of equation (7) is:

$$n^*_{opt} = (\frac{\xi^*_{\mathrm{A},n'}(\alpha)\hat{\sigma}_{n'}}{d})^2 \tag{11}$$

where $\xi^*_{\mathrm{A},n'}(\alpha)$ is the $(1-\alpha)$-th quantile of the (conditional) d.f. of $|\frac{\sqrt{n'}(\hat{\theta}^*_{n'} - \hat{\theta}_{n'})}{\hat{\sigma}_{n'}}|$. This procedure will make use of the estimator $\hat{\sigma}_{n'}$ for σ. But it is easy to see that, a.s. $[P]$

$$\xi^*_{\mathrm{A},n'}(\alpha)\hat{\sigma}_{n'} = \xi^*_{\mathrm{C},n'}(\alpha),$$

and therefore this second proposal results in exactly the same solution as the previous one.

- Using a bootstrap approximation for the "stUdentized" estimator:

$$\sup_t \mid P^*_{n'}\{\frac{\sqrt{n'}(\hat{\theta}^*_{n'} - \hat{\theta}_{n'})}{\hat{\sigma}^*_{n'}} \leq t\} - P\{\frac{\sqrt{n'}(\hat{\theta}_{n'} - \theta)}{\hat{\sigma}_{n'}} \leq t\} \mid \xrightarrow{\text{a.s. [P]}} 0 \quad , \quad n' \to \infty$$

leading to the following bootstrap equivalent of equation (7):

$$n^*_{opt} = (\frac{\xi^*_{\mathrm{U},n'}(\alpha)\hat{\sigma}_{n'}}{d})^2 \tag{12}$$

with $\xi^*_{\mathrm{U},n'}(\alpha)$ the $(1-\alpha)$-th quantile of the (conditional) d.f. of $|\frac{\sqrt{n'}(\hat{\theta}^*_{n'}-\hat{\theta}_{n'})}{\hat{\sigma}^*_{n'}}|$. This setting also estimates separately both unknown parameters but leads to a different sample size choice.

We will refer to the above methods as the "standardized method" (case one or two, using the c.p. $\xi^*_{\mathrm{A},n'}(\alpha)$) or the "studentized method" (case three, using the c.p. $\xi^*_{\mathrm{U},n'}(\alpha)$). If no distinction is necessary, we omit the indices A and U.

We are now ready to define our triple sampling stopping times based on bootstrap c.p.'s $\xi^*_{n'}(\alpha)$. The three stages of the proposed procedure are defined as follows:

stage 1: take an initial pilot-sample $X_1, ..., X_{n_0}$, with

$$n_0(d) = \max\{[(\frac{z}{d})^{2\eta}] + 1, 2\} \tag{13}$$

where $0 < \eta < 1$;

stage 2: take additional $N_1 - n_0$ observations to get an intermediate sample $X_1, ..., X_{N_1}$ with

$$N_1(d) = \max\{[c(\frac{\xi^*_{n_0}(\alpha)\hat{\sigma}_{n_0}}{d})^2] + 1, n_0\} \tag{14}$$

where $0 < c < 1$;

stage 3: sample $N_2 - N_1$ more observations to obtain the final sample $X_1, ..., X_{N_2}$ with

$$N_2(d) = \max\{[(\frac{\xi^*_{N_1}(\alpha)\hat{\sigma}_{N_1}}{d})^2] + 1, N_1\}; \tag{15}$$

finally: construct the $2d$-width confidence interval for θ :

$$[\hat{\theta}_{N_2} - d, \hat{\theta}_{N_2} + d]. \tag{16}$$

3 ASYMPTOTIC CONSISTENCY AND EFFICIENCY

In this section, we examine the first order behaviour (consistency and efficiency) of the standardized bootstrap procedure (based on $\xi^*_{\mathrm{A},n}(\alpha)$) for the case that $\theta = E(X)$, $\hat{\theta}_n = \bar{X}_n$ the sample mean and $\hat{\sigma}^2_n = S^2_n = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ the sample variance. One would expect no difference in first order behaviour between the normal method (9), the standardized bootstrap method (11) and the studentized bootstrap method (12). For the higher order behaviour however, several existing results on bootstrap critical points have shown that the studentized bootstrap distribution outperformes the other ones (see e.g. Hall(1988b) and Bhattacharya and Qumsiyeh (1989)). Rates of convergence are not addressed here. Of interest is the comparison in performance of the above mentioned three methods for small and moderate sample sizes. In order to investigate that comparison we carried out simulation studies for the population mean. The results of these studies are presented in Section 4.

Theorem

Assume $E(X^2) < \infty$ and $\sigma^2 > 0$. Then the three-stage procedure, defined by (13) - (16) has the following properties:

$$\lim_{d \to 0} N_2(d) \overset{\text{a.s. [P]}}{=} \infty, \tag{17}$$

$$\lim_{d \to 0} \frac{N_2(d)}{\tilde{n}(d)} \overset{\text{a.s. [P]}}{=} 1 \tag{18}$$

where

$$\tilde{n}(d) = [(\frac{z\sigma}{d})^2] + 1 \tag{19}$$

and

$$\lim_{d \to 0} P\{\theta \in I_{N_2(d)}\} = 1 - \alpha. \tag{20}$$

Under the additional assumption that $E \mid X \mid^r < \infty$ for some $r > 2$, we have that

$$\lim_{d\to 0}\frac{E[N_2(d)]}{\tilde{n}(d)}=1. \tag{21}$$

Proof

Denote by

$$\mathcal{L}_n^*(x)=P_n^*\{\mid \frac{\sqrt{n}(\bar{X}_n^*-\bar{X}_n)}{S_n}\mid\leq x\}$$

the d.f. of the standardized bootstrap statistic. Again, P_n^* refers to the conditional probability measure corresponding to the d.f. F_n. As before $\xi_n^*(\alpha)$ is the $(1-\alpha)$-th quantile of the d.f. $\mathcal{L}_n^*$, i.e.: $\xi_n^*(\alpha)=\inf\{x:\mathcal{L}_n^*(x)\geq 1-\alpha\}$. We first prove that

$$\lim_{d\to 0}N_1(d)\overset{\text{a.s. [P]}}{=}\infty \tag{22}$$

which implies, see the definition of $N_2(d)$, the similar result (17) for $N_2(d)$. According to Theorem 2.1 in Bickel and Freedman(1981), we have that,

$$\lim_{n\to\infty}\mathcal{L}_n^*(x)\overset{\text{a.s. [P]}}{=}\begin{cases}2\Phi(x)-1, & \text{if } x\geq 0;\\ 0, & \text{otherwise.}\end{cases}$$

and, since $\Phi(.)$ is strictly increasing, that $\forall t\in(0,1)$,

$$\lim_{n\to\infty}\xi_n^*(t)\overset{\text{a.s.[P]}}{=}\Phi^{-1}(1-\frac{t}{2})$$

(see e.g. Lemma 1.5.6, Serfling(1980)). Note that $n_0(d)$ increases to ∞ as d tends to 0 and thus

$$\lim_{d\to 0}\xi_{n_0(d)}^*(\alpha)\overset{\text{a.s. [P]}}{=}z.$$

This together with the fact that $S_{n_0(d)}^2$ also converges a.s. to σ^2 proves (22).

In order to prove (18) we start from the following inequality, which follows from definition (15) of $N_2(d)$:

$$\begin{aligned}(\frac{\xi_{N_1}^*(\alpha)S_{N_1}}{d})^2 \quad < \quad & N_2(d)\leq(\frac{\xi_{N_1}^*(\alpha)S_{N_1}}{d})^2+1 \\ & +N_1(d)I[(\xi_{N_1}^*(\alpha)S_{N_1})^2\leq c(\xi_{n_0}^*(\alpha)S_{n_0})^2+n_0d^2+d^2]\end{aligned} \tag{23}$$

The strong consistency of $\xi_n^*(\alpha)S_n$ as an estimator for $z\sigma$ and (22) imply that

$$\lim_{n\to\infty} \xi^*_{N_1(d)}(\alpha)S_{N_1(d)} \overset{\text{a.s.[P]}}{=} z\sigma$$

(see e.g. Proposition 1 in Csörgő(1973)).
Hence, for $d > 0$ small enough,

$$I[(\xi^*_{N_1}(\alpha)S_{N_1})^2 \leq c(\xi^*_{n_0}(\alpha)S_{n_0})^2 + n_0 d^2 + d^2] \overset{\text{a.s.[P]}}{=} 0,$$

using the fact that $0 < \eta < 1$ and $0 < c < 1$.
Inserting these results in inequality (24) completes the proof of (18).

The asymptotic consistency stated in (20) guarantees that the proposed procedure converges to the right coverage probability. Justification of this statement follows immediately from (18), Slutsky's Theorem and Theorem 3 of Anscombe(1952).

We now prove the uniform integrability of the set $\{N_2(d)d^2\}_{d>0}$, under the additional assumption that $E \mid X \mid^r < \infty$ for some $r > 2$. As established in Bickel and Yahav(1968), it is sufficient to show that, for some $d_0 > 0$,

$$\sum_{l=1}^{\infty} \sup_{0<d<d_0} P(N_2(d)d^2 > l) < \infty.$$

Fix $0 < d < d_0$. For d_0 small enough, $([(\frac{z}{d})^{2\eta}]+1)d^2 \leq l$ and $2d^2 \leq l$, such that

$$\begin{aligned} P(N_2(d)d^2 > l) &\leq P(([\frac{(\xi^*_{N_1}(\alpha)S_{N_1})^2}{d^2}] + 1)d^2 > l) \\ &\quad + P(([\frac{c(\xi^*_{n_0}(\alpha)S_{n_0})^2}{d^2}] + 1)d^2 > l) \\ &\equiv P_{1,l}(d) + P_{2,l}(d). \end{aligned}$$

We first deal with the second probability.

<u>Probability $P_{2,l}(d)$</u>

Applying Markov's inequality and noting that, for d_0 small enough, $d^2 < l/2$, we can bound the second probability as follows ($\text{Var}^*_{n_0}$ refers to the conditional variance w.r.t. F_{n_0}):

$$\begin{aligned}
P_{2,l}(d) &\leq P\{\xi^*_{n_0}(\alpha)S_{n_0} > \sqrt{\frac{l-d^2}{c}}\} \\
&\leq P\{\mathcal{L}^*_{n_0}(\sqrt{\frac{l}{2c}}\frac{1}{S_{n_0}}) < 1-\alpha\} \\
&= P\{P^*_{n_0}\{\mid \sqrt{n_0}(\bar{X}^*_{n_0} - \bar{X}_{n_0}) \mid \leq \sqrt{\frac{l}{2c}}\} < 1-\alpha\} \\
&= P\{P^*_{n_0}\{\mid \sqrt{n_0}(\bar{X}^*_{n_0} - \bar{X}_{n_0}) \mid > \sqrt{\frac{l}{2c}}\} > \alpha\} \\
&\leq P\{\frac{2c}{l}n_0\mathrm{Var}^*_{n_0}(\bar{X}^*_{n_0}) > \alpha\} \\
&= P\{\frac{2c}{l}\mathrm{Var}^*_{n_0}(X^*_1) > \alpha\} \\
&\leq P\{S^2_{n_0(d)} > \frac{\alpha l}{2c}\} \\
&\leq (\frac{2c}{\alpha l})^q(\max_{m\geq n_0(d_0)} E[S^{2q}_m]) \\
&\leq Cl^{-q}E(X^{2q})
\end{aligned}$$

where $C > 0$ is some positive constant. Choosing $q = \frac{r}{2} > 1$ yields a convergent series $\sum_{l=1}^{\infty}\sup_{0<d<d_0} P_{2,l}(d)$.

Probability $P_{1,l}(d)$

Using similar arguments as before, the first probability can be bounded by:

$$P_{1,l}(d) \leq P\{S^2_{N_1(d)} > \frac{\alpha l}{2}\} \leq P\{\sup_{m\geq n_0(d)} \frac{1}{m-1}\sum_{i=1}^{m}(X_i - \bar{X}_m)^2 > \frac{\alpha l}{2}\}$$

since $P\{N_1(d) \geq n_0(d)\} = 1$. Further, remark that $\frac{1}{m-1}\sum_{i=1}^{m}(X_i - \bar{X}_m)^2$ is a U-statistic and hence a reverse martingale (see e.g. Lemma 5.1.5.B of Serfling(1980)). Using Kolmogorov's inequality for reverse martingales (Chow(1960)), we obtain

$$P_{1,l}(d) \leq (\frac{\alpha l}{2})^{-q}E[S^{2q}_{n_0(d)}]$$

leading to, as in probability $P_{2,l}(d)$, a convergent series.

The uniform integrability of the set $\{N_2(d)d^2\}_{d>0}$, together with (18) yields (21).

4 SIMULATION RESULTS

The asymptotic properties established in the previous section show that the proposed procedure, described in (13) - (16), has the same first order behaviour as the three-stage normal method (based on (9)). In this section we further investigate the procedure via Monte-Carlo simulations. The simulation results provide information about the performance of the normal method, the standardized method and the studentized method.

First of all we mention that exact computation of the involved bootstrap distributions, such as the conditional d.f. of $|\frac{\sqrt{n'}(\hat{\theta}^*_{n'}-\hat{\theta}_{n'})}{\hat{\sigma}^*_{n'}}|$, is usually difficult. Therefore, we have to rely on Monte-Carlo simulations. For ease of presentation we explain this computational part for the studentized method only. Given a sample $X_1, \cdots, X_{n'}$ from the d.f. F we proceed as follows. Take B independent random samples $X^*_1, \cdots, X^*_{n'}$ of size n' from $F_{n'}$, the empirical d.f. of $X_1, \cdots, X_{n'}$. For each of those bootstrap samples, calculate the appropriate bootstrap statistic, i.e. $|\frac{\sqrt{n'}(\hat{\theta}^*_{n'}-\hat{\theta}_{n'})}{\hat{\sigma}^*_{n'}}|$. The sampling distribution of the resulting B values of this statistic is taken as an approximation to the actual bootstrap distribution. The $(1-\alpha)$-th quantile of this sampling distribution provides an appropriate approximation to the actual bootstrap c.p. $\xi^*_{\mathrm{U},n'}(\alpha)$. The accuracy of those approximations can be improved by increasing B. However, increasing the number of bootstrap replicates is limited somehow by the computational cost. In our simulation studies we chose B equal to 2000, which is reasonable when calculating bootstrap confidence intervals (see e.g. Swanepoel (1990)).

As mentioned in Section 1, the parameters η and c appearing in (13) - (14) allow to control the sample sizes in the first two stages of the procedure. Our simulation experiences show that for the proposed procedure the choice of η has to be made with some care, whereas the choice of c is less important. This highlights the fact that the choice of the initial pilot-sample size n_0

has to be controled in the following sense. If the initial sample size n_0 is too small the approximated bootstrap critical point $\xi^*_{n_0}(\alpha)$ may differ too much from the optimal value $\xi^*_{n_{opt}}(\alpha)$. There is a certain region of small sample sizes n_0 in which the approximated bootstrap c.p.'s $\xi^*_{n_0}(\alpha)$ are subject to a lot of variation (with extreme large values). This instability and over-estimation is carried over (in a quadratic way) to the intermediate sample size N_1 and the final sample size N_2. It is clear that this region should be avoided and that the choice of the initial sample size n_0 is more delicate when using the bootstrap c.p.'s. In applying the normal method one stays on the safe-side and always uses the c.p. from the limiting normal distribution. In its turn, this safe-play leads to under-estimated final sample sizes N_2 in certain cases, as will be seen later. It turned out that the choice $\eta = 0.75$ is an appropriate one, reflecting the cautionness of an initial sample which is too small. However such a choice of η may not be sufficient in preventing for a too small initial sample. This may occur for example in situations where we sample from a skewed d.f. (such as the Gamma distribution). To protect further against this instability of the bootstrap c.p., we therefore truncated these c.p.'s at a certain truncation value, taken to be 3 in our simulation studies. This truncation only affected the estimation procedures for the Gamma distributions and for $d = 0.2$ ($n_0 = 17$) and $d = 0.3$ ($n_0 = 30$). In all our simulations we took $c = 0.5$, which seems to be an appropriate choice (see also Hall (1981)). Simulations with other choices of c lead to comparable results, and therefore are not presented here.

For given values of α, d, η and c, and a given random sample from F, each of the three steps in the procedure (13) - (16) can be calculated using the above explained Monte-Carlo approximations. This results in a value for $N_2(d)$ and a confidence interval $[\bar{X}_{N_2} - d, \bar{X}_{N_2} + d]$. This whole procedure, the "trial", is repeated a number of times, where trial refers to a new choice of the sample from F. An estimate for $E(N_2(d))$ is obtained by averaging over all values of $N_2(d)$ resulting from the Monte-Carlo trials. The proportion of confidence intervals $[\bar{X}_{N_2} - d, \bar{X}_{N_2} + d]$ in the trials including the population mean of F gives an idea of the obtained coverage probability. The number of trials in the presented simulation results was 100.

The investigation here focusses on 95 % confidence intervals, i.e. taking $\alpha = 0.05$. Two major types of distribution functions F were considered, the gamma distribution and the standard normal distribution. We obtained random samples from those distributions via random number generators provided by Lewis, Orav and Uribe (1986). Fortran programs were used to carry through all calculations. For each of the choices of F, namely the Gamma(k,1) d.f., with k taking the values 0.5, 1 or 1.5, and the standard normal d.f., we present simulation results for values of d 0.05, 0.10, 0.20 and 0.30. Since asymptotics here means $d \to 0$ it is expected that all procedures perform better for smaller values of d. We further mention that as k in the Gamma(k,1) d.f. increases, the skewness of the d.f. becomes less pronounced. Note also that with $k = 1$ we consider the exponential d.f.

Tables I – IV summarize the simulation results. For each value of d, the performance of the different procedures can be evaluated by looking at the average of the sample sizes $N_2(d)$ and the coverage proportion of the confidence intervals $[I_{N_2(d)} - d, I_{N_2(d)} + d]$ resulting from all 100 trials (see columns 4 and 5 for the normal method, columns 6 and 7 for the standardized method and columns 8 and 9 for the studentized method). These numbers should be compared with the exact optimal sample size n_{opt} (in column 3) and the predetermined coverage probability 0.95. The optimal sample size n_{opt} can be calculated via (7) since the d.f. is known. To quantify the accuracy of the estimated mean of $N_2(d)$, the associated standard deviation is included in the tables (between brackets).

First of all, note that if the d.f. F is normal (see Table IV), then there is no advantage in using a bootstrap approximation instead of a normal approximation (as expected).

In all tables, the results for the normal method and the standardized method are very similar. For skewed distributions and for d not too small, one can observe some under-estimation of the stopping time $N_2(d)$. This reflects the fact that using the bootstrap approximation based on the centered or standardized statistic is not an optimal way for implementing bootstrap c.p. when searching for a higher degree of accuracy. The tables show some benefit in using a studentized bootstrap approximation instead of a normal approximation; this is most visible for $d \geq 0.10$ and for skewed d.f.. These simulation results, in favour of the studentized method, are in accordance with Hall's (1988a,b) findings. It is not clear yet what the exact gain is in applying the studentized method w.r.t. the normal or standardized method. Hall(1988a,b) used Edgeworth and Cornish-Fisher expansions to quantify the optimality of the studentized approach. For the present problem of constructing a confidence interval with exact preassigned width $2d$ and approximate coverage probability $1-\alpha$, similar theoretical verifications did not reveal the same higher order preference for the studentized method. Further theoretical investigations are necessary in this respect. Finally, for relatively large values of $d \geq 0.20$, the initial sample sizes n_0 are too far from the optimal size n_{opt}, leading to extremely deviating bootstrap c.p.'s and hence sample sizes which are too large. An extra truncation argument, mentioned above, controled this partly.

TABLE I: Gamma(0.5,1) distribution

d	n_0	n_{opt}	*Normal method* Aver. $N_2(d)$	*Normal method* Estim. Covprob.	*Standard. method* Aver. $N_2(d)$	*Standard. method* Estim. Covprob.	*Student. method* Aver. $N_2(d)$	*Student. method* Estim. Covprob.
0.05	246	769	734.2 (157.31)	0.92	732.4 (162.14)	0.91	767.4 (155.84)	0.92
0.10	87	192	178.5 (57.40)	0.86	177.6 (56.79)	0.90	206.2 (69.92)	0.90
0.20	30	48	45.3 (18.24)	0.96	44.6 (17.86)	0.95	65.0 (43.68)	0.95
0.30	17	21	23.2 (10.34)	0.98	23.0 (10.76)	0.98	32.0 (18.25)	0.97

TABLE II: Gamma(1.0,1) distribution

d	n_0	n_{opt}	*Normal method* Aver. $N_2(d)$	*Normal method* Estim. Covprob.	*Standard. method* Aver. $N_2(d)$	*Standard. method* Estim. Covprob.	*Student. method* Aver. $N_2(d)$	*Student. method* Estim. Covprob.
0.05	246	1537	1487.2 (165.31)	0.91	1480.2 (176.96)	0.92	1512.8 (179.64)	0.91
0.10	87	385	370.0 (81.99)	0.93	373.2 (83.34)	0.95	388.7 (90.86)	0.97
0.20	30	96	87.9 (32.99)	0.89	87.1 (32.49)	0.88	104.8 (46.32)	0.95
0.30	17	42	34.5 (15.97)	0.86	34.0 (15.54)	0.85	51.6 (32.18)	0.91

TABLE III: Gamma(1.5,1) distribution

d	n_0	n_{opt}	*Normal method* Aver. $N_2(d)$	Estim. Covprob.	*Standard. method* Aver. $N_2(d)$	Estim. Covprob.	*Student. method* Aver. $N_2(d)$	Estim. Covprob.
0.05	246	2305	2296.2 (162.09)	0.94	2284.6 (182.78)	0.94	2323.1 (180.98)	0.93
0.10	87	577	549.5 (90.12)	0.91	547.3 (92.75)	0.90	573.7 (91.87)	0.91
0.20	30	144	131.7 (47.59)	0.88	130.1 (46.37)	0.89	157.3 (65.21)	0.91
0.30	17	64	49.2 (24.58)	0.85	48.5 (23.85)	0.86	72.0 (43.03)	0.92

TABLE IV: N(0,1) distribution

d	n_0	n_{opt}	*Normal method* Aver. $N_2(d)$	Estim. Covprob.	*Standard. method* Aver. $N_2(d)$	Estim. Covprob.	*Student. method* Aver. $N_2(d)$	Estim. Covprob.
0.05	246	1537	1540.5 (70.47)	0.98	1541.6 (97.24)	0.99	1540.1 (87.63)	0.99
0.10	87	385	384.2 (34.84)	0.95	384.4 (40.83)	0.95	390.5 (40.37)	0.97
0.20	30	97	90.5 (18.95)	0.91	91.2 (20.46)	0.92	97.8 (17.35)	0.94
0.30	17	43	38.0 (11.37)	0.92	37.8 (11.89)	0.92	44.1 (11.16)	0.94

ACKNOWLEDGEMENTS

This work was completed while the second author was visiting the Mathematical Sciences Research Institute, Berkeley, California, as Senior Research Assistant at the National Fund for Scientific Research (Belgium).

BIBLIOGRAPHY

Anscombe, F.J. (1952). Large-sample theory of sequential estimation. Proc. Cambridge Philos. Soc., 48, 600-607.

Bhattacharya, R. and Qumsiyeh, M. (1989). Second order and L^p-comparisons between the bootstrap and empirical Edgeworth expansion methodologies. Ann. Statist., 17, 160-169.

Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. Ann. Statist., 9, 1196-1217.

Bickel, P.J. and Yahav, J.A. (1968). Asymptotic optimal Bayes and minimax procedures in sequential estimation. Ann. Statist., 39, 442-456.

Chow, Y.S. (1960). A martingale inequality and the law of large numbers. Proc. Amer. Math. Soc., 11, 107-110.

Chow, Y.S. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. Ann. Math. Statist., 36, 457-462.

Csörgő, M. (1973). Glivenko-Cantelli type theorems for distance functions based on the modified empirical distribution function of M. Kac and for the empirical process with random sample size in general. Lecture Notes in Mathematics, 296, Springer-Verlag, New York, 149-164.

Hall, P. (1981). Asymptotic theory of triple sampling for sequential estimation of a mean. Ann. Statist., 9, 1229-1238.

Hall, P. (1988a). On symmetric bootstrap confidence intervals. J. R. Statist. Soc., Series B, 1, 35-45.

Hall, P. (1988b). Theoretical comparison of bootstrap confidence intervals. Ann. Statist., 16, 927-953.

Hamdy, H.I. (1988). Remarks on the asymptotic theory of triple stage estimation of the normal mean. Scand. J. Statist., 15, 303-310.

Honda, T. (1992). Estimation of the mean by the three stage procedure. Commun. in Statist., -Seq. Analys.-, 11, 73-89.

Lohr, S.L. (1990). Accurate multivariate estimation using triple sampling. Ann. Statist., 18, 1615-1633.

Lewis, P.A.W., Orav, E.J. and Uribe, L. (1986). Advanced simulation and statistics package. IBM professional fortran version. Wadsworth & Brooks/ Cole, Monterey, California.

Serfling, R.J. (1980). Approximation theorems of mathematical statistics. Wiley, New York.

Stein, C. (1945). A two sample test for a linear hypothesis whose power is independent of the variance. Ann. Math. Statist., 43, 243-258.

Swanepoel, J.W.H. (1990). A review of bootstrap methods. South African Statist. J., 24, 1-34.

Swanepoel, J.W.H., van Wijk, J.W.J. and Venter, J.H. (1984). Fixed-width confidence intervals based on bootstrap procedures. Commun. in Statist., -Seq. Analys.-, 2, 289-310.

Vik, G. and Mukhopadhyay, N. (1988). Triple sampling to construct fixed-width confidence intervals for estimable parameters based on U-statistics. Metron, Intern. J. Statist., XLVI, 165-174.