J. R. Statist. Soc. B (1999)
61, *Part* 2, *pp.* 353–366

The effect of Monte Carlo approximation on coverage error of double-bootstrap confidence intervals

Stephen M. S. Lee

University of Hong Kong, Hong Kong

and G. Alastair Young

University of Cambridge, UK

[Received February 1998. Revised July 1998]

Summary. A double-bootstrap confidence interval must usually be approximated by a Monte Carlo simulation, consisting of two nested levels of bootstrap sampling. We provide an analysis of the coverage accuracy of the interval which takes account of both the inherent bootstrap *and* Monte Carlo errors. The analysis shows that, by a suitable choice of the number of resamples drawn at the inner level of bootstrap sampling, we can reduce the order of coverage error. We consider also the effects of performing a finite Monte Carlo simulation on the mean length and variability of length of two-sided intervals. An adaptive procedure is presented for the choice of the number of inner level resamples. The effectiveness of the procedure is illustrated through a small simulation study.

Keywords: Bootstrap; Coverage error; Double bootstrap; Monte Carlo approximation; Percentile method; Resample; Sampling error; Simulation

1. Introduction

The double bootstrap (Hall, 1986; Beran, 1987) provides a satisfactory solution to the problem of reducing the coverage error of nonparametric bootstrap confidence intervals. As it is usually applied in this context, the double bootstrap accomplishes a calibration of the bootstrap confidence interval by making an additive adjustment to the nominal coverage of the interval. The theoretical effect of the calibration has been studied by Martin (1990), who showed that the magnitude of the coverage error is reduced by the adjustment. In practice, the double-bootstrap confidence interval must be approximated by a Monte Carlo simulation, which consists of drawing C second-level bootstrap samples from each of a series of B first-level bootstrap samples drawn from the available sample data. Booth and Hall (1994) considered the question of the optimal choice of B and C, in an analysis which considered the distance between the bootstrap interval limit constructed using an infinite number of bootstrap simulations and its Monte Carlo version based on a finite simulation. This analysis is *not* directly related to coverage accuracy. In the current paper, we provide an analysis of the coverage accuracy of the double-bootstrap confidence interval which takes account of both sampling error, imposed by the sample data, *and* Monte Carlo error, due to the finiteness of the simulation.

Address for correspondence: G. Alastair Young, Statistical Laboratory, University of Cambridge, 16 Mill Lane, Cambridge, CB2 1SB, UK.
E-mail: g.a.young@statslab.cam.ac.uk

 1369–7412/99/61353

For simplicity of presentation, we restrict attention to the situation considered by Booth and Hall (1994), where the double bootstrap is used, as described above, to make an additive adjustment to the nominal coverage of the confidence interval, and to the case of the percentile method confidence interval. An analysis similar to that carried out by Booth and Hall (1994) has been performed by Booth and Presnell (1998) for the situation where the double bootstrap is used to make an additive adjustment directly to the end points of the confidence interval. The theoretical, infinite simulation, effects of calibration of different kinds of bootstrap confidence interval were considered by Martin (1990). The analysis presented here is easily adapted to these other forms of interval.

In Section 2 we provide a detailed description of the usual Monte Carlo algorithm for approximation to the double-bootstrap percentile method interval, for both one-sided and two-sided cases. We provide, in equations (2.5) and (2.6), an asymptotic analysis of the coverage properties of the double-bootstrap interval based on a finite Monte Carlo simulation. This analysis shows that B and C must be of larger order in the sample size n, of order n^4 and n^2 respectively in the two-sided case and of order n^2 and n respectively in the one-sided case, to ensure that the coverage error of the Monte Carlo interval remains of the same order as that of the theoretical, infinite simulation, double-bootstrap interval. The analysis shows further that, by a suitable choice of C, we may reduce the order of the coverage error, in effect using the Monte Carlo error to eliminate the sampling error. We also consider the effects on mean length and variability of length of two-sided intervals of performing a finite simulation.

In Section 3 we provide a practical interpretation and illustration of our asymptotic results. We provide guidelines on the choice of B and C for practical application. Our recommendation amounts to the use of large B and to an adaptive choice of C to minimize the coverage error for the chosen B. The asymptotic forms of double-bootstrap confidence interval described by Lee and Young (1995) are suggested as a means of an empirical choice of C.

A simulation study is presented in Section 4 for the problem of constructing two-sided confidence intervals for a population variance. It is seen that our empirical procedure for the choice of Monte Carlo simulation size is effective in producing desirable coverage accuracy in a computationally efficient manner. It also strikes an effective balance between the, generally competing, demands of coverage accuracy and stability.

2. The standard Monte Carlo approach

Let $\mathcal{X} = (X_1, \ldots, X_n)$ be a random sample of size n drawn from an underlying distribution function F. We wish to construct a confidence interval of coverage α ($\frac{1}{2} < \alpha < 1$), based on $\mathcal{X}$, for a scalar parameter θ, expressible as a smooth function of the mean μ of F: $\theta = g(\mu)$. The confidence interval is to be constructed from the sample estimate $\hat{\theta} = g(\bar{X})$, where

$$\bar{X} = n^{-1} \sum_{i=1}^{n} X_i.$$

Let $\sigma^2 \equiv h(\mu)$ be the asymptotic variance of $n^{1/2}(\hat{\theta} - \theta)$ and let $\hat{\sigma}^2 = h(\bar{X})$ be its sample version.

Let $\mathcal{X}^*$ be a generic first-level bootstrap sample drawn randomly, with replacement, from $\mathcal{X}$ and similarly let $\mathcal{X}^{**}$ denote a generic second-level bootstrap sample drawn from $\mathcal{X}^*$. Let $\hat{\theta}^*$ and $\hat{\theta}^{**}$ be the versions of the statistic $\hat{\theta}$ based on $\mathcal{X}^*$ and $\mathcal{X}^{**}$ respectively.

Define $\hat{y}_\beta$ by $P(\hat{\theta}^* \leqslant \hat{y}_\beta | \mathcal{X}) = \beta$ and $\hat{y}^*_\beta$ by $P(\hat{\theta}^{**} \leqslant \hat{y}^*_\beta | \mathcal{X}, \mathcal{X}^*) = \beta$. The theoretical, infinite simulation, one-sided percentile method bootstrap confidence interval of nominal coverage α for θ is

$$P_{1,\alpha} = (-\infty, \hat{y}_\alpha).$$

The corresponding two-sided interval is

$$P_{2,\alpha} = (\hat{y}_{(1-\alpha)/2}, \hat{y}_{(1+\alpha)/2}).$$

Then we define the theoretical, infinite simulation, one-sided double-bootstrap percentile method confidence interval for θ to be

$$I_{1,\alpha} = (-\infty, \hat{y}_{\alpha+\hat{\xi}}). \tag{2.1}$$

where $\hat{\xi}$ satisfies

$$P(\hat{\theta} \leqslant \hat{y}^*_{\alpha+\hat{\xi}} | \mathcal{X}) = \alpha.$$

Similarly, the two-sided double-bootstrap percentile method interval for θ is defined to be

$$I_{2,\alpha} = (\hat{y}_{(1-\alpha-\hat{\eta})/2}, \hat{y}_{(1+\alpha+\hat{\eta})/2}), \tag{2.2}$$

where

$$P(\hat{y}^*_{(1-\alpha-\hat{\eta})/2} \leqslant \hat{\theta} \leqslant \hat{y}^*_{(1+\alpha+\hat{\eta})/2} | \mathcal{X}) = \alpha.$$

As noted, in practice the theoretical intervals $I_{1,\alpha}$ and $I_{2,\alpha}$ must be approximated by a Monte Carlo simulation. We draw a collection of B independent bootstrap samples $\mathcal{X}^*_1$, $\mathcal{X}^*_2, \ldots, \mathcal{X}^*_B$ from $\mathcal{X}$. From each $\mathcal{X}^*_b$ ($b = 1, 2, \ldots, B$) we then draw C second-level bootstrap samples $\mathcal{X}^{**}_{b1}, \mathcal{X}^{**}_{b2}, \ldots, \mathcal{X}^{**}_{bC}$. Denote by $\hat{\theta}^*_b$ and $\hat{\theta}^{**}_{bc}$ the realizations of $\hat{\theta}^*$ and $\hat{\theta}^{**}$ based on $\mathcal{X}^*_b$ and $\mathcal{X}^{**}_{bc}$ respectively, $b = 1, 2, \ldots, B$ and $c = 1, 2, \ldots, C$.

Let $[\cdot]$ denote the integer part function and define

$$k = [(B+1)\alpha],$$

$$U^* = P(\hat{\theta}^{**} \leqslant \hat{\theta} | \mathcal{X}, \mathcal{X}^*), \qquad U^*_b = P(\hat{\theta}^{**} \leqslant \hat{\theta} | \mathcal{X}, \mathcal{X}^*_b), \qquad U^*_{(1)} \leqslant \ldots \leqslant U^*_{(B)},$$

$$\hat{U}^*_b = \frac{1}{C} \sum_{c=1}^{C} \mathbf{1}\{\hat{\theta}^{**}_{bc} \leqslant \hat{\theta}\}, \qquad \hat{U}^*_{(1)} \leqslant \hat{U}^*_{(2)} \leqslant \ldots \leqslant \hat{U}^*_{(B)},$$

$$l = [(B+1)\hat{U}^*_{(k)}], \qquad V^*_b = |2U^*_b - 1|, \qquad V^*_{(1)} \leqslant \ldots \leqslant V^*_{(B)},$$

$$\hat{V}^*_b = |2\hat{U}^*_b - 1|, \qquad \hat{V}^*_{(1)} \leqslant \hat{V}^*_{(2)} \leqslant \ldots \leqslant \hat{V}^*_{(B)},$$

$$m' = [\tfrac{1}{2}(B+1)(1 + \hat{V}^*_{(k)})], \qquad m'' = [\tfrac{1}{2}(B+1)(1 - \hat{V}^*_{(k)})].$$

Here $\mathbf{1}$ denotes the indicator function, and subscripts in parentheses denote ordered values. Then the Monte Carlo approximation to $I_{1,\alpha}$ is

$$\tilde{I}_{1,\alpha} = (-\infty, \hat{\theta}^*_{(l)}), \tag{2.3}$$

where $\hat{\theta}^*_{(1)} \leqslant \hat{\theta}^*_{(2)} \leqslant \ldots \leqslant \hat{\theta}^*_{(B)}$ are the ordered values of $\hat{\theta}^*_1, \hat{\theta}^*_2, \ldots, \hat{\theta}^*_B$.

The Monte Carlo approximation to $I_{2,\alpha}$ is

$$\tilde{I}_{2,\alpha} = (\hat{\theta}^*_{(m'')}, \hat{\theta}^*_{(m')}). \tag{2.4}$$

The corresponding Monte Carlo approximation to $P_{1,\alpha}$ is $\tilde{P}_{1,\alpha} = (-\infty, \hat{\theta}^*_{(k)})$ and that to $P_{2,\alpha}$ is $\tilde{P}_{2,\alpha} = (\hat{\theta}^*_{(k'')}, \hat{\theta}^*_{(k')})$, where

$$k' = [(B+1) \times \tfrac{1}{2}(1+\alpha)],$$

$$k'' = [(B+1) \times \tfrac{1}{2}(1-\alpha)].$$

In Appendix A we prove the following results concerning the coverages of the intervals $\tilde{I}_{1,\alpha}$ and $\tilde{I}_{2,\alpha}$:

$$P(\theta \in \tilde{I}_{1,\alpha}) = P(\theta \in I_{1,\alpha}) + C^{-1}(\alpha - \tfrac{1}{2}) + o(B^{-1/2} + C^{-1}), \tag{2.5}$$

$$P(\theta \in \tilde{I}_{2,\alpha}) = P(\theta \in I_{2,\alpha}) + C^{-1}\alpha + o(B^{-1/2} + C^{-1}). \tag{2.6}$$

Martin (1990) showed that the coverage error of the theoretical one-sided interval $I_{1,\alpha}$, $P(\theta \in I_{1,\alpha}) - \alpha$, is of order $O(n^{-1})$, compared with the coverage error $O(n^{-1/2})$ of $P_{1,\alpha}$. It is therefore seen from equation (2.5) that C must be of order n, or larger, and B of order n^2, or larger, to ensure that the use of a finite simulation does not result in the approximate interval $\tilde{I}_{1,\alpha}$ having a coverage error of lower order in n than that of $I_{1,\alpha}$. Taking C of order n and B of order n^2 gives a coverage error of $\tilde{I}_{1,\alpha}$ of order $O(n^{-1})$. The coverage error of the theoretical two-sided interval $I_{2,\alpha}$ is of order $O(n^{-2})$: see Martin (1990). Formula (2.6) shows that in this two-sided case taking C to be of order n^2 and B to be of order n^4 ensures that the coverage error of $\tilde{I}_{2,\alpha}$ remains of order $O(n^{-2})$.

Typically, as in the example considered in Section 4 later, the theoretical intervals $I_{1,\alpha}$ and $I_{2,\alpha}$ undercover, so the coverage errors $P(\theta \in I_{1,\alpha}) - \alpha$ and $P(\theta \in I_{2,\alpha}) - \alpha$ are negative. In these circumstances, it is clear from equations (2.5) and (2.6) that the coverage errors of $\tilde{I}_{1,\alpha}$ and $\tilde{I}_{2,\alpha}$ can be reduced to orders $o(n^{-1})$ and $o(n^{-2})$ respectively by a suitable choice of C. In the one-sided case, take

$$C = (\tfrac{1}{2} - \alpha) \big/ \{P(\theta \in I_{1,\alpha}) - \alpha\} \tag{2.7}$$

and in the two-sided case take

$$C = -\alpha / \{P(\theta \in I_{2,\alpha}) - \alpha\}. \tag{2.8}$$

We discuss a practical estimation of these optimal values of C in Section 3.

A more detailed analysis than that given in Appendix A shows that a similar reduction in the order of the coverage error of the intervals $\tilde{I}_{1,\alpha}$ and $\tilde{I}_{2,\alpha}$ can be achieved also in the case where the infinite simulation interval has coverage which exceeds α. In the one-sided case the Monte Carlo algorithm may be modified by replacing $\hat{U}_b^*$ by $(C\hat{U}_b^* + a)/(C+b)$ for a and b with $(b-1)\alpha - a > -\tfrac{1}{2}$ and choosing

$$C = \{\tfrac{1}{2} + (b-1)\alpha - a\}/\{P(\theta \in I_{1,\alpha}) - \alpha\}.$$

In the two-sided case, replace $\hat{V}_b^*$ by $(C\hat{V}_b^* + a)/(C+b)$ for a and b with $(b-1)\alpha - a > 0$, and choose

$$C = \{(b-1)\alpha - a\}/\{P(\theta \in I_{2,\alpha}) - \alpha\}.$$

The analysis given in Appendix A examines the effect on the mean interval length and the variance of the interval length of the use of a finite Monte Carlo simulation. Define $L = \hat{y}_{(1+\alpha+\hat{\eta})/2} - \hat{y}_{(1-\alpha-\hat{\eta})/2}$ to be the length of the, infinite simulation, interval $I_{2,\alpha}$ and $\tilde{L} = \hat{\theta}^*_{(m')} - \hat{\theta}^*_{(m'')}$ to be the length of its Monte Carlo approximation $\tilde{I}_{2,\alpha}$. Appendix A shows that

$$E(\tilde{L}) = E(L) + n^{-1/2}C^{-1}\sigma\,\phi(z_{(1+\alpha)/2})^{-1}\alpha + o\{n^{-1/2}(C^{-1} + B^{-1/2})\}, \tag{2.9}$$

and

$$\mathrm{var}(\tilde{L}) = \mathrm{var}(L) + n^{-1}B^{-1}\sigma^2\,\phi(z_{(1+\alpha)/2})^{-2} \times 4\alpha(1-\alpha) + o\{n^{-3/2}(C^{-1} + B^{-1/2}) + n^{-1}(C^{-2} + B^{-1})\}, \tag{2.10}$$

where $z_\beta = \Phi^{-1}(\beta)$. We may readily show that $E(L)$ is of order $O(n^{-1/2})$ and $\mathrm{var}(L)$ is of order $O(n^{-2})$. We have argued above that, to ensure that the coverage of $\tilde{I}_{2,\alpha}$ remains of the same order, $O(n^{-2})$, as that of $I_{2,\alpha}$, we should take B of order n^4 and C of order n^2. We see from equations (2.9) and (2.10) that such a choice of B and C has an asymptotically negligible effect on the mean length and the variance of length: $E(\tilde{L}) - E(L) = O(n^{-5/2})$ and $\mathrm{var}(\tilde{L}) - \mathrm{var}(L) = o(n^{-7/2})$.

A somewhat crude interpretation of the expansions (2.9) and (2.10) is that $E(\tilde{L})$ is mainly affected by C, whereas $\mathrm{var}(\tilde{L})$ is affected by both B and C, for the sort of B and C advocated on coverage considerations.

3. An illustration and recommendation

Figs 1 and 2 illustrate the effects of varying B and C on the coverage, mean length and variance of length characteristics of the interval $\tilde{I}_{2,\alpha}$ when $\alpha = 0.9$ and the parent population F is folded normal, $|N(0, 1)|$ for two situations:

(a) where the parameter of interest θ is the population mean and $n = 15$ (Fig. 1);
(b) where θ is the population variance and $n = 35$ (Fig. 2).

The coverage, mean length and variance of length figures plotted are all based on a series of 1600 replications of the double-bootstrap procedure. The plots shown are typical of those obtained for a range of estimation problems.

In both situations shown, and others that we have studied, the coverage probability of the double-bootstrap interval increases with decreasing C, in accordance with our expectations from the asymptotic analysis of Section 2. The coverage probability depends to a lesser extent on B. For values of C other than those very small, the coverage error decreases with increasing B, providing evidence that the use of large B is required. The price paid for improving the coverage by reducing C is that the average length of the interval typically increases with decreasing C, again as we expect from the asymptotic analysis. The picture as regards the variance of the length of the interval is less clear, though the variance may actually decrease as C is reduced. We note also that, typically, the average length increases with increasing B, though this effect is only pronounced for very small values of C. We might expect from equation (2.10) that the variance of the length of the interval would decrease with increasing B. It is clear from Figs 1 and 2 that the practical effect of increasing B on the variance of length is problem specific and dependent on the choice of C. We have evidence that increasing B decreases the variance when C is sufficiently large, as has been expected from equation (2.10).

We summarize our theoretical and empirical findings in a specific recommendation on the choice of B and C. In doing so, recall that our primary objective in the paper has been to analyse the effect of B and C on the coverage error, with a view to recommending a strategy for the optimal choice of B and C specifically in those terms. The coverage error of the double-bootstrap interval is generally smallest, and more stable to varying C, for large B. We therefore interpret our findings in terms of a recommendation for the choice of large B, in line

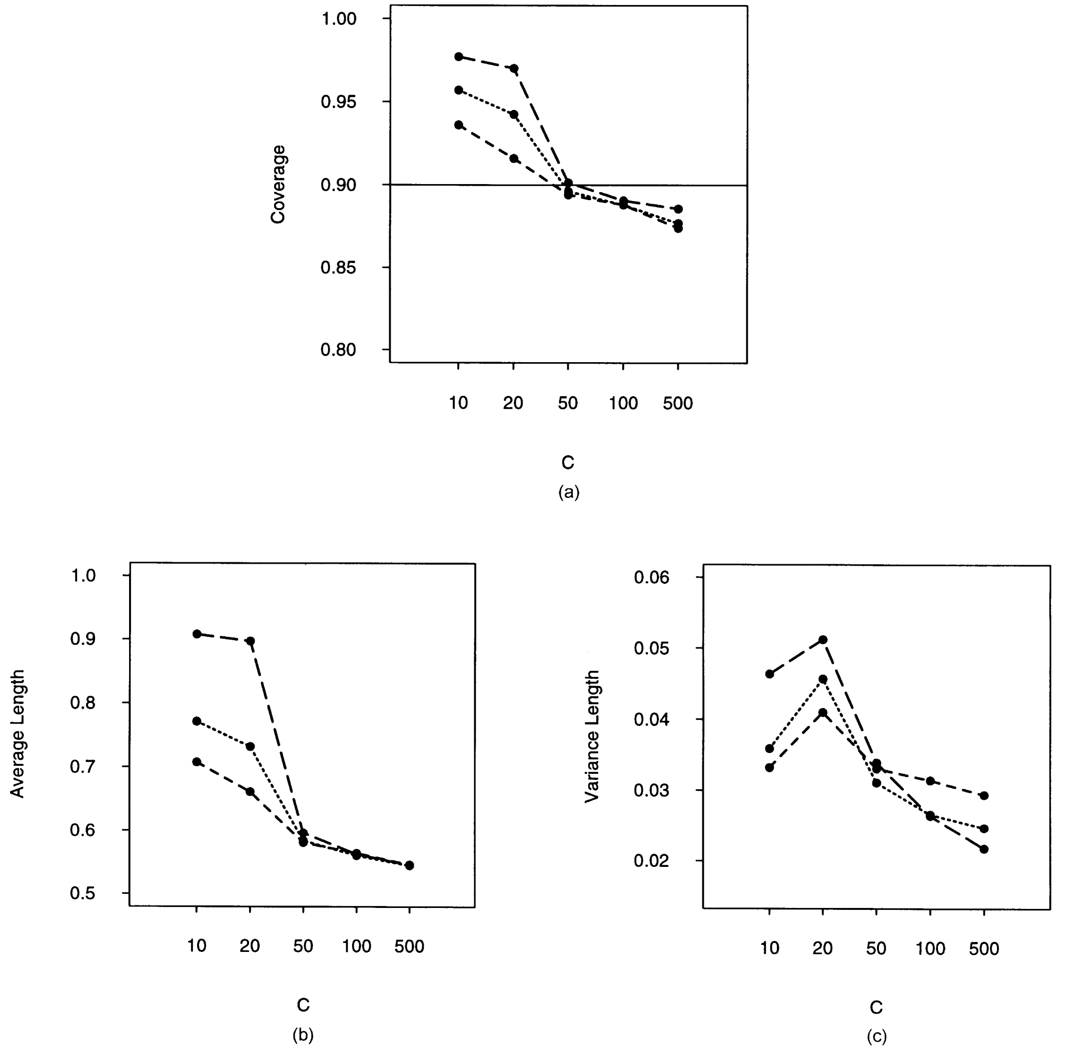

Fig. 1. Effects of varying B and C on (a) the coverage, (b) the mean length and (c) the variance of the length for $n = 15$, mean: -------, $B = 100$; ··········, $B = 200$; — — —, $B = 1000$

with standard practice and previous recommendations, such as those provided by Booth and Hall (1994) and Booth and Presnell (1998). Computational resource considerations would lead to the use of, say, $B = 1000$. The novel aspect to our analysis is the discovery that, for a given value of B, the coverage error may be reduced by an appropriate choice of C, typically much smaller than values advocated in previous work, without necessarily producing any significant increase in either the interval length or the variability in interval length.

The optimal choice of C, asymptotically, unfortunately depends, as shown in equations (2.7) and (2.8), on the coverage error of the theoretical, infinite simulation, interval. An estimation of this coverage error is therefore necessary as part of a practical strategy for the choice of C.

Lee and Young (1995) presented methods for direct approximation, without a Monte

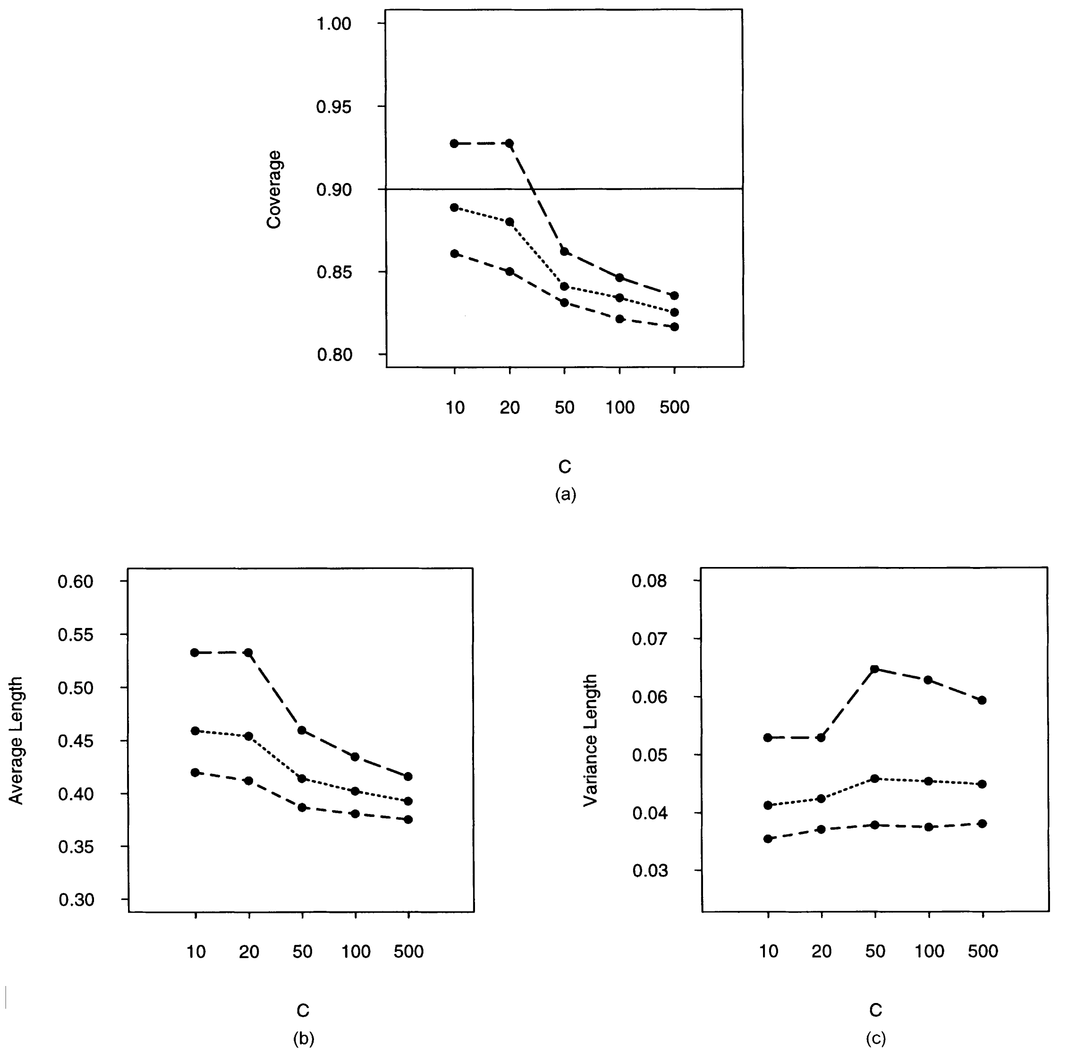

Fig. 2. Effects of varying B and C on (a) the coverage, (b) the mean length and (c) the variance of the length for $n = 35$, variance: -------, $B = 100$; ·········, $B = 200$; — — —, $B = 1000$

Carlo simulation, to the double-bootstrap intervals $I_{1,\alpha}$ and $I_{2,\alpha}$. Their methods amount to direct substitution of quantities computed from the data into truncated asymptotic expansions for the end points of the intervals. These approximate forms of double-bootstrap interval may be used in the context of the current paper to obtain estimates of the coverage errors $P\{\theta \in I_{1,\alpha}\} - \alpha$ and $P\{\theta \in I_{2,\alpha}\} - \alpha$. These estimates may then be used to choose C empirically via equations (2.7) and (2.8).

Such an adaptive choice of C requires a preliminary Monte Carlo simulation, which will consist of the drawing of, say, D bootstrap samples from $\mathcal{X}$. To estimate, say, $P\{\theta \in I_{2,\alpha}\}$ the analytic interval of Lee and Young (1995) is constructed for each of these D samples: the coverage is estimated by the proportion of the intervals which contain $\hat{\theta}$, the sample estimate constructed from $\mathcal{X}$. We advocate that moderate D, say of the order of a few hundreds, is adequate for an estimation of the coverage error. Since the intervals of Lee and Young (1995) require simple arithmetic calculations, the preliminary Monte Carlo simulation is negligible in computational cost compared with the nested levels of bootstrap sampling required in the construction of the double-bootstrap interval once B and C have been set. Of course, an adaptive choice of C in this way will generally lead to the use of small C, and therefore significant computational gains over conventional choices of C, which would generally set C large, say 500 or 1000, without regard to the possible benefits, in terms of a reduction in coverage error, of the use of smaller C.

4. Simulation study and remarks

Tables 1 and 2 summarize the results of a simulation study involving the construction of nominal 90% coverage confidence intervals for the population variance, for four underlying populations and three sample sizes $n = 20, 35, 100$. For each combination of distribution and sample size, a series of 1600 random samples were drawn and from each of these various confidence intervals constructed, coverages being estimated by the proportion of the 1600 intervals containing the true population variance. The intervals considered were the percentile interval $\tilde{P}_{2,0.9}$ based on $B = 1000$ outer level bootstrap samples, iterated intervals $\tilde{I}^1_{2,0.9}$ and $\tilde{I}^2_{2,0.9}$ based on $B = 1000$ and respectively $C = 500$ and $C = 100$ inner level bootstrap samples, and the interval $\tilde{I}^A_{2,0.9}$ based on $B = 1000$ and an adaptive choice of C. For $\tilde{I}^A_{2,0.9}$ a preliminary simulation of $D = 500$ bootstrap samples were drawn from each of the 1600 parent samples to estimate the coverage error $P\{\theta \in I_{2,0.9}\} - 0.9$. Letting $\hat{I}^d_{2,0.9}$ denote the approximation to

Table 1. Variance example — estimated coverage probabilities

Interval	*Coverage, n = 20*	*Coverage, n = 35*	*Coverage, n = 100*
Normal data N(0, 1) (no skewness, no kurtosis)			
$\tilde{P}_{2,0.9}$	0.727	0.793	0.857
$\tilde{I}^1_{2,0.9}$	0.849	0.863	0.888
$\tilde{I}^2_{2,0.9}$	0.866	0.868	0.896
$\tilde{I}^A_{2,0.9}$	0.851 (56.54)	0.869 (80.34)	0.889 (102.33)
Folded normal data |N(0, 1)| (high skewness, low kurtosis)			
$\tilde{P}_{2,0.9}$	0.686	0.753	0.843
$\tilde{I}^1_{2,0.9}$	0.814	0.835	0.880
$\tilde{I}^2_{2,0.9}$	0.825	0.846	0.887
$\tilde{I}^A_{2,0.9}$	0.809 (36.36)	0.846 (51.23)	0.892 (83.41)
Double-exponential data $\frac{1}{2}\exp(-\|x\|)$ (no skewness, high kurtosis)			
$\tilde{P}_{2,0.9}$	0.698	0.776	0.834
$\tilde{I}^1_{2,0.9}$	0.827	0.854	0.878
$\tilde{I}^2_{2,0.9}$	0.838	0.865	0.885
$\tilde{I}^A_{2,0.9}$	0.832 (43.17)	0.867 (54.57)	0.884 (79.72)
Log-normal data exp{N(0, 1)} (high skewness, high kurtosis)			
$\tilde{P}_{2,0.9}$	0.416	0.504	0.608
$\tilde{I}^1_{2,0.9}$	0.544	0.631	0.722
$\tilde{I}^2_{2,0.9}$	0.546	0.641	0.733
$\tilde{I}^A_{2,0.9}$	0.544 (23.14)	0.636 (28.56)	0.754 (38.46)

Table 2. Variance example—simulated mean interval length and variance of interval length†

Interval	*n = 20*		*n = 35*		*n = 100*	
Normal data N(0, 1) (no skewness, no kurtosis)						
$\tilde{P}_{2,0.9}$	0.848	0.111	0.701	4.648×10^{-2}	0.447	6.933×10^{-3}
$\tilde{I}^{1}_{2,0.9}$	1.274	0.362	0.900	0.129	0.489	1.126×10^{-2}
$\tilde{I}^{2}_{2,0.9}$	1.344	0.400	0.945	0.164	0.501	1.204×10^{-2}
$\tilde{I}^{A}_{2,0.9}$	1.326	0.435	0.974	0.203	0.510	2.323×10^{-2}
Folded normal data |N(0, 1)| (high skewness, low kurtosis)						
$\tilde{P}_{2,0.9}$	0.342	3.628×10^{-2}	0.289	1.668×10^{-2}	0.189	2.937×10^{-3}
$\tilde{I}^{1}_{2,0.9}$	0.530	0.107	0.416	5.931×10^{-2}	0.224	7.946×10^{-3}
$\tilde{I}^{2}_{2,0.9}$	0.545	0.107	0.435	6.285×10^{-2}	0.232	9.718×10^{-3}
$\tilde{I}^{A}_{2,0.9}$	0.537	0.111	0.442	6.545×10^{-2}	0.253	1.546×10^{-2}
Double-exponential data $\frac{1}{2}\exp(-\|x\|)$ (no skewness, high kurtosis)						
$\tilde{P}_{2,0.9}$	2.311	2.510	2.011	1.270	1.361	0.259
$\tilde{I}^{1}_{2,0.9}$	3.735	7.628	2.983	4.283	1.705	0.868
$\tilde{I}^{2}_{2,0.9}$	3.873	7.696	3.118	4.566	1.777	1.022
$\tilde{I}^{A}_{2,0.9}$	3.799	7.726	3.156	4.678	1.748	0.754
Log-normal data $\exp\{N(0, 1)\}$ (high skewness, high kurtosis)						
$\tilde{P}_{2,0.9}$	9.411	9408.940	8.777	3659.902	6.879	573.833
$\tilde{I}^{1}_{2,0.9}$	13.864	15130.088	14.633	12019.490	11.404	1577.404
$\tilde{I}^{2}_{2,0.9}$	13.908	15129.914	14.731	12018.776	11.584	1576.358
$\tilde{I}^{A}_{2,0.9}$	13.838	15129.936	14.639	12019.287	11.732	1574.956

†The first figure shown in each cell is the mean length; the second the variance of length.

the double-bootstrap interval for the dth such bootstrap sample ($d = 1, \ldots, D$) constructed as described by Lee and Young (1995), the coverage error of $I_{2,0.9}$ is estimated by

$$\hat{\pi} = D^{-1} \sum_{d=1}^{D} \mathbf{1}\{\hat{\theta} \in \hat{I}^{d}_{2,0.9}\} - 0.9.$$

Then the interval $\tilde{I}^{A}_{2,0.9}$ is based on $C = -0.9/\hat{\pi}$, if $\hat{\pi} < 0$, and $C = 0.45/\hat{\pi}$, if $\hat{\pi} > 0$. In the latter case, as described in Section 2, the Monte Carlo algorithm is modified, once C has been computed, by modifying $\hat{V}^{*}_{b}$ to $CV^{*}_{b}/(C + 1.5)$.

The coverages of the intervals are given in Table 1, together with figures showing the average number of inner level bootstrap samples chosen by the adaptive method. The mean interval length and the variance of interval length over the 1600 replications is shown for each of the interval types in Table 2. Note that the simulation size used in the study, 1600 random samples, ensures that the standard error of each simulated coverage is of the order of 0.01. A detailed comparison between the various intervals is possible, however, since each is constructed from the same set of parent samples.

The coverage figures in Table 1 indicate that the adaptive procedure for the choice of C works well, producing greater coverage accuracy than the interval $\tilde{I}^{1}_{2,0.9}$, yet typically drawing far fewer second-level bootstrap samples, and coverage accuracy comparable with that of the interval $\tilde{I}^{2}_{2,0.9}$, but with the automatic detection of the appropriate number of second-level samples. Although there might be a slight tendency for the average length and, in particular, the variability of length to increase over a fixed choice of C, any such increase is insignificant compared with increases attached to the use of the double-bootstrap interval rather than the uncalibrated percentile interval.

Martin (1990) showed that calibration of the percentile interval $P_{2,\alpha}$ to produce the double-bootstrap interval $I_{2,\alpha}$ is accompanied by a change in mean length which is proportional to the coverage error of the original interval. In examples such as that being considered here, the coverage error of the original interval is substantial, and the double-bootstrap interval is of considerably greater mean length.

In other examples which we have studied, such as the mean, the coverage error of the percentile interval is typically smaller than in the variance example considered here. In such a case, the optimal C is generally larger than the values that are optimal when the coverage error is large. In particular, the fixed choice of $C = 100$, which Table 1 suggests is reasonable in the variance example, may no longer be adequate. However, our adaptive procedure for the choice of C is effective in identifying the need for larger C in such cases, though the average number of second-level resamples C drawn remains generally lower than the large values that are conventionally used. The larger the coverage error, the smaller C can be to obtain a satisfactory result. In complex cases, where the coverage error might be substantial, very low values of C may be reasonable. However, we see that in the variance example reducing C to very low levels may lead to serious overcoverage. What sensible lower limit might be set on C depends on the problem in question, and we recommend as a general strategy the adaptive procedure for the choice of C, combined with the avoidance of very low values of C, say less than 20.

The computational cost of the use of the asymptotic interval constructions of Lee and Young (1995) in an adaptive choice of C is negligible in comparison with the computational cost of the Monte Carlo simulation once B and C have been set. The question arises whether there is any advantage to the use of the Monte Carlo construction, rather than direct use of the asymptotic intervals. A comparison of the coverage figures in Table 1 with the corresponding figures for the asymptotic intervals themselves, as given by Lee and Young (1995), shows that the Monte Carlo intervals implemented using the asymptotic intervals to choose C have a significantly lower coverage error than the asymptotic intervals have.

The motivation for the use of the double-bootstrap confidence interval is as a means of reducing the coverage error of the uncalibrated bootstrap interval. In this paper we have provided an analysis of the consequences for coverage error of approximating the double-bootstrap interval by a finite Monte Carlo simulation. The analysis shows that the number of bootstrap samples drawn must be of higher order in the sample size n than is the case for the raw uncalibrated interval. Adapting the analysis of Hall (1986) to the case of the percentile method interval, we may show that taking B to be of order $n^{1/2}$ in the one-sided case, or of order n in the two-sided case, guarantees that the Monte Carlo approximation to the percentile interval has a coverage error of the same order as that of the infinite simulation interval. Nevertheless, we have demonstrated that in practice a realistic goal is to reduce the coverage error by control of the number of second-level bootstrap samples C.

In the same way that the infinite simulation double-bootstrap interval may be viewed as an adjustment to the uncalibrated bootstrap interval, the Monte Carlo construction of the interval may be viewed as providing an adjustment to the infinite simulation interval. By a suitable choice of C, the adjustment can be directed to improve the coverage accuracy. In general, a reduction in the value of C provides the appropriate calibration of coverage error. In effect, we may use the Monte Carlo approximation as an opportunity to eliminate the coverage error of the theoretical double-bootstrap interval, without causing significant damage in terms of interval length or stability. A by-product is improved computational efficiency compared with the conventional practice of setting *both* B and C large, especially in problems where the coverage error of the theoretical double-bootstrap interval remains large.

Appendix A

A.1. Coverage error of one-sided interval

Define $\hat{S}_n(\beta)$ by

$$P\{U^* \leqslant \beta + n^{-1/2}\,\hat{S}_n(\beta)|\mathcal{X}\} = \beta$$

and $\hat{R}_n(\beta)$ by

$$P(U^* \leqslant \beta|\mathcal{X}) = \beta + n^{-1/2}\,\hat{R}_n(\beta).$$

Then $\hat{\xi} = n^{-1/2}\,\hat{S}_n(\alpha)$ by definition.

Consider

$$\mathbb{E}[\hat{U}^*_{(k)}|\mathcal{X}] = C^{-1} \sum_{i=0}^{C-1} \sum_{j=0}^{k-1} \binom{B}{j} \hat{p}_i^{\,j}(1-\hat{p}_i)^{B-j}. \tag{A.1}$$

Here

$$\begin{aligned}\hat{p}_i &= P(C\hat{U}^*_1 \leqslant i|\mathcal{X}) \\ &= \mathbb{E}\left[\sum_{j=0}^{i}\binom{C}{j} U^{*m}(1-U^*)^{C-m}|\mathcal{X}\right] \\ &= \frac{i+1}{C+1} + n^{-1/2}\int_0^1 \hat{R}_n\{\tau(v)\}\,\mathrm{d}v,\end{aligned}$$

where

$$\tau(v) = \beta_i - C^{-1/2}\{\beta_i(1-\beta_i)\}^{1/2} z_v + \tfrac{1}{6}C^{-1}(1-2\beta_i)(1+2z_v^2) + o(C^{-1}),$$

and

$$\beta_i = C^{-1}(i+\tfrac{1}{2}).$$

The formula for $\tau(v)$ follows from the definition of $\tau(v, v)$ in Hall (1986).

Expanding the integral $\int_0^1 \hat{R}_n\{\tau(v)\}\,\mathrm{d}v$, we have

$$\hat{p}_i = \beta_i + n^{-1/2}\,\hat{R}_n(\beta_i) + \tfrac{1}{2}C^{-1}(1-2\beta_i) + o(C^{-1}). \tag{A.2}$$

Substituting equation (A.2) in equation (A.1), using the trapezoidal rule and Taylor expansion,

$$\mathbb{E}[\hat{U}^*_{(k)}|\mathcal{X}] = \int_0^1 [g\{u + n^{-1/2}\,\hat{R}_n(u)\} + \tfrac{1}{2}C^{-1}(1-2u)\,g'\{u + n^{-1/2}\,\hat{R}_n(u)\}]\,\mathrm{d}u + o_p(C^{-1}), \tag{A.3}$$

where

$$g(\beta) := \sum_{j=0}^{k-1}\binom{B}{j}\beta^j(1-\beta)^{B-j}.$$

Noting that $g(0)=1$, $g(1)=0$ and $\hat{R}_n(0) = \hat{R}_n(1) = 0$, we can show that

$$\int_0^1 (1-2u)\,g'\{u + n^{-1/2}\,\hat{R}_n(u)\}\,\mathrm{d}u = 2\int_0^1 g\{u + n^{-1/2}\,\hat{R}_n(u)\}\,\mathrm{d}u - 1, \tag{A.4}$$

and

$$\int_0^1 g\{u+n^{-1/2}\,\hat{R}_n(u)\}\,\mathrm{d}u=\frac{k}{B+1}+n^{-1/2}\,\hat{S}_n\left(\frac{k-\frac{1}{2}}{B}\right)+O_p(n^{-1/2}B^{-1}). \tag{A.5}$$

It follows from equations (A.3)–(A.5) that

$$\mathbb{E}[\hat{U}^*_{(k)}|\mathcal{X}]=\frac{k}{B+1}+C^{-1}\left(\frac{k}{B+1}-\frac{1}{2}\right)+n^{-1/2}\,\hat{S}_n\left(\frac{k}{B+1}\right)+O_p(n^{-1/2}B^{-1})+o_p(C^{-1}). \tag{A.6}$$

According to Hall (1992), $\hat{y}_\beta=\hat{\theta}+n^{-1/2}\,\hat{\sigma}\{z_\beta+n^{-1/2}\,\hat{p}_{11}(z_\beta)+\ldots\}$. Thus, expanding about $\alpha+\hat{\xi}$, using equation (A.2) in Booth and Hall (1994), and substituting equation (A.6),

$$\hat{y}_{\hat{U}^*_{(k)}}-\hat{y}_{\alpha+\hat{\xi}}=n^{-1/2}\,\hat{\sigma}\,\phi(z_\alpha)^{-1}\{U^*_{(k)}-\mathbb{E}[U^*_{(k)}|\mathcal{X}]+C^{-1}(\alpha-\tfrac{1}{2})+o_p(B^{-1/2}+C^{-1})\}. \tag{A.7}$$

Using Bahadur's representation and Chebyshev's inequality,

$$\hat{\theta}^*_{(l)}-\hat{y}_{\hat{U}^*_{(k)}}=\hat{\theta}^*_{(k)}-\hat{y}_\alpha+o_p(n^{-1/2}B^{-1/2}). \tag{A.8}$$

Define $S^*\equiv n^{1/2}(\hat{\theta}^*-\hat{\theta})/\hat{\sigma}$ and $\hat{u}_\beta$ by $P(S^*\leqslant\hat{u}_\beta|\mathcal{X})=\beta$. By repeated use of Edgeworth and Taylor expansions,

$$\hat{U}^*_{(k)}-\mathbb{E}[U^*_{(k)}|\mathcal{X}]=-(S^*_{(B-k+1)}-\hat{u}_{1-\alpha})\,\phi(z_\alpha)+o_p(B^{-1/2}), \tag{A.9}$$

where $S^*_{(1)}\leqslant\ldots\leqslant S^*_{(B)}$, and S^*_b is the realization of S^* on the bth bootstrap resample.

Combining equations (A.7)–(A.9), we then obtain

$$P(\theta\leqslant\hat{\theta}^*_{(l)}|\mathcal{X})=P\{S^*_{(k)}-\hat{u}_\alpha-(S^*_{(B-k+1)}-\hat{u}_{1-\alpha})\geqslant\hat{\omega}|\mathcal{X}\}, \tag{A.10}$$

where

$$\hat{\omega}=-\frac{n^{1/2}(\hat{\theta}-\theta)}{\hat{\sigma}}-\hat{u}_{\alpha+\hat{\xi}}-C^{-1}(\alpha-\tfrac{1}{2})\,\phi(z_\alpha)^{-1}+o_p(B^{-1/2}+C^{-1}).$$

By expansion of the joint distribution of order statistics,

$$P[B^{1/2}\{S^*_{(k)}-\hat{u}_\alpha-(S^*_{(B-k+1)}-\hat{u}_{1-\alpha})\}\leqslant x|\mathcal{X}]=\Phi(x/\hat{\nu})+B^{-1/2}\,H(x/\hat{\nu})+o_p(B^{-1/2}), \tag{A.11}$$

where ν^2 is the asymptotic variance of $B^{1/2}\{S^*_{(k)}-\hat{u}_\alpha-(S^*_{(B-k+1)}-\hat{u}_{1-\alpha})\}$ and H satisfies $H(u)\to 0$ as $u\to\pm\infty$.

Using equation (A.11) and taking the expectation of equation (A.10),

$$\begin{aligned}
P(\theta\leqslant\hat{\theta}^*_{(l)})&=P(\hat{\omega}\leqslant 0)+o(B^{-1/2})\\
&=P\left\{\frac{n^{1/2}(\hat{\theta}-\theta)}{\hat{\sigma}}+\hat{u}_{\alpha+\hat{\xi}}\geqslant\ -C^{-1}(\alpha-\tfrac{1}{2})\,\phi(z_\alpha)^{-1}\right\}+o(B^{-1/2}+C^{-1})\\
&=P\left\{\frac{n^{1/2}(\hat{\theta}-\theta)}{\hat{\sigma}}+\hat{u}_{\alpha+\hat{\xi}}\geqslant 0\right\}+C^{-1}(\alpha-\tfrac{1}{2})+o(B^{-1/2}+C^{-1})\\
&=P(\theta\leqslant\hat{y}_{\alpha+\hat{\xi}})+C^{-1}(\alpha-\tfrac{1}{2})+o(B^{-1/2}+C^{-1}),
\end{aligned} \tag{A.12}$$

as required.

A.2. Coverage error of two-sided interval

Consider

$$\mathbb{E}[\hat{V}^*_{(k)}|\mathcal{X}]=1-\frac{2}{C}\sum_{i=0}^{[C/2]-1}\sum_{j=0}^{B-k}\binom{B}{j}\hat{q}_i^{\,j}(1-\hat{q}_i)^{B-j},$$

where

$$\hat{q}_i \equiv P\left\{\frac{C}{2}(1-\hat{V}_1^*) \leqslant i|\mathcal{X}\right\}$$

$$= 2\left(\frac{i+1}{C+1}\right) + n^{-1/2}\left\{\hat{R}_n\left(\frac{i+\frac{1}{2}}{C}\right) - \hat{R}_n\left(\frac{C-i-\frac{1}{2}}{C}\right)\right\} + o(C^{-1}).$$

It then follows from the trapezoidal rule, Taylor expansion and term-by-term integration that

$$\mathbb{E}[\hat{V}_{(k)}^*|\mathcal{X}] = \alpha - n^{-1/2}\hat{A}_n(1-\alpha) + C^{-1}\alpha + o_p(C^{-1}+B^{-1/2}), \tag{A.13}$$

where $\hat{A}_n(\beta)$ satisfies

$$\hat{A}_n(\beta) + \hat{R}_n\left\{\frac{\beta+n^{-1/2}\hat{A}_n(\beta)}{2}\right\} - \hat{R}_n\left\{1-\frac{\beta+n^{-1/2}\hat{A}_n(\beta)}{2}\right\} = 0. \tag{A.14}$$

Putting $\beta = 1-\alpha$ in equation (A.14), we have

$$P[\tfrac{1}{2}\{1-\alpha+n^{-1/2}\hat{A}_n(1-\alpha)\} \leqslant U^* \leqslant \tfrac{1}{2}\{1+\alpha-n^{-1/2}\hat{A}_n(1-\alpha)\}|\mathcal{X}]$$

$$= \alpha - n^{-1/2}\hat{A}_n(1-\alpha) + n^{-1/2}\left[\hat{R}_n\left\{\frac{1+\alpha-n^{-1/2}\hat{A}_n(1-\alpha)}{2}\right\} - \hat{R}_n\left\{\frac{1-\alpha+n^{-1/2}\hat{A}_n(1-\alpha)}{2}\right\}\right]$$

$$= \alpha.$$

Therefore $\alpha - n^{-1/2}\hat{A}_n(1-\alpha) = \alpha + \hat{\eta}$, or $\hat{\eta} = -n^{-1/2}\hat{A}_n(1-\alpha)$.

Inserting this result in equation (A.13), we obtain

$$\mathbb{E}[\hat{V}_{(k)}^*|\mathcal{X}] = \alpha + \hat{\eta} + C^{-1}\alpha + o_p(C^{-1}+B^{-1/2}). \tag{A.15}$$

By an argument that is very similar to that used in deriving equation (A.10),

$$P(\theta \leqslant \hat{\theta}_{(m')}^*|\mathcal{X}) = P[\phi(z_{(1+\alpha)/2})^{-1}\{-(U_{(B+1-k')}^* - \mathbb{E}[U_{(B+1-k')}^*|\mathcal{X}]) + \tfrac{1}{2}(V_{(k)}^* - \mathbb{E}[V_{(k)}^*|\mathcal{X}])\} \geqslant \hat{v}|\mathcal{X}], \tag{A.16}$$

where

$$\hat{v} = -\frac{n^{1/2}(\hat{\theta}-\theta)}{\hat{\sigma}} - \hat{u}_{(1+\alpha+\hat{\eta})/2} - \frac{1}{2}C^{-1}\alpha\,\phi(z_{(1+\alpha)/2})^{-1} + o_p(B^{-1/2}+C^{-1}),$$

and

$$k' = [\tfrac{1}{2}(B+1)(1+\alpha)],$$

as before. Using the asymptotic normality of order statistics and properties of their concomitants, and taking the expectation of equation (A.16),

$$P(\theta \leqslant \hat{\theta}_{(m')}^*) = P\left\{\frac{n^{1/2}(\hat{\theta}-\theta)}{\hat{\sigma}} + \hat{u}_{(1+\alpha+\hat{\eta})/2} \geqslant 0\right\} + \frac{1}{2}C^{-1}\alpha + o(B^{-1/2}+C^{-1})$$

$$= P(\theta \leqslant \hat{y}_{(1+\alpha+\hat{\eta})/2}) + \tfrac{1}{2}C^{-1}\alpha + o(B^{-1/2}+C^{-1}). \tag{A.17}$$

Similarly for the lower tail

$$P(\theta \leqslant \hat{\theta}_{(m'')}^*) = P(\theta \leqslant \hat{y}_{(1-\alpha-\hat{\eta})/2}) - \tfrac{1}{2}C^{-1}\alpha + o(B^{-1/2}+C^{-1}). \tag{A.18}$$

Subtracting equation (A.18) from equation (A.17), we obtain

$$P(\hat{\theta}_{(m'')}^* \leqslant \theta \leqslant \hat{\theta}_{(m')}^*) = P(\hat{y}_{(1-\alpha-\hat{\eta})/2} \leqslant \theta \leqslant \hat{y}_{(1+\alpha+\hat{\eta})/2}) + C^{-1}\alpha + o(B^{-1/2}+C^{-1}), \tag{A.19}$$

which is equation (2.6).

A.3. Length of two-sided interval

From expansions of $\hat{\theta}^*_{(m')}$ and $\hat{\theta}^*_{(m'')}$,

$$\tilde{L} \equiv \hat{\theta}^*_{(m')} - \hat{\theta}^*_{(m'')} = \hat{y}_{(1+\alpha+\hat{\eta})/2} - \hat{y}_{(1-\alpha-\hat{\eta})/2} + n^{-1/2}\hat{\sigma}\,\phi(z_{(1+\alpha)/2})^{-1}\{-(U^*_{(B+1-k')} - \mathbb{E}[U^*_{(B+1-k')}|\mathcal{X}]) + U^*_{(B+1-k'')} - \mathbb{E}[U^*_{(B+1-k'')}|\mathcal{X}] + V^*_{(k)} - \mathbb{E}[V^*_{(k)}|\mathcal{X}] + C^{-1}\alpha + o_p(C^{-1} + B^{-1/2})\}, \tag{A.20}$$

where $k' = [\frac{1}{2}(B+1)(1+\alpha)]$ and $k'' = [\frac{1}{2}(B+1)(1-\alpha)]$.

Taking the expectation of equation (A.20),

$$\mathbb{E}[\tilde{L}] = \mathbb{E}[\hat{y}_{(1+\alpha+\hat{\eta})/2} - \hat{y}_{(1-\alpha-\hat{\eta})/2}] + n^{-1/2}C^{-1}\sigma\,\phi(z_{(1+\alpha)/2})^{-1}\alpha + o\{n^{-1/2}(C^{-1} + B^{-1/2})\}. \tag{A.21}$$

Expansion (A.21) confirms equation (2.9). Using properties of concomitant statistics,

$$\mathrm{cov}(V^*_{(k)}, U^*_{(k')}|\mathcal{X}) = \tfrac{1}{2}B^{-1}\alpha(1-\alpha) + o(B^{-1})$$

and

$$\mathrm{cov}(V^*_{(k)}, U^*_{(k'')}|\mathcal{X}) = -\tfrac{1}{2}B^{-1}\alpha(1-\alpha) + o(B^{-1}). \tag{A.22}$$

Using equation (A.22) and properties of order statistics,

$$\mathrm{var}(\hat{\theta}^*_{(m')} - \hat{\theta}^*_{(m'')}|\mathcal{X}) = n^{-1}\hat{\sigma}^2\,\phi(z_{(1+\alpha)/2})^{-2}\{4B^{-1}\alpha(1-\alpha) + o_p(C^{-2} + B^{-1})\}. \tag{A.23}$$

Clearly,

$$\mathbb{E}[\hat{\theta}^*_{(m')} - \hat{\theta}^*_{(m'')}|\mathcal{X}] = \hat{y}_{(1+\alpha+\hat{\eta})/2} - \hat{y}_{(1-\alpha-\hat{\eta})/2} + n^{-1/2}\hat{\sigma}\,\phi(z_{(1+\alpha)/2})^{-1}\{C^{-1}\alpha + o_p(C^{-1} + B^{-1/2})\}. \tag{A.24}$$

It follows from equations (A.23) and (A.24) that

$$\mathrm{var}(\tilde{L}) = \mathrm{var}(\hat{y}_{(1+\alpha+\hat{\eta})/2} - \hat{y}_{(1-\alpha-\hat{\eta})/2}) + n^{-1}B^{-1}\sigma^2\,\phi(z_{(1+\alpha)/2})^{-2} \times 4\alpha(1-\alpha) + o\{n^{-3/2}(C^{-1} + B^{-1/2}) + n^{-1}(C^{-2} + B^{-1})\}.$$

References

Beran, R. (1987) Prepivoting to reduce level error of confidence sets. *Biometrika*, **74**, 457–468.

Booth, J. G. and Hall, P. (1994) Monte Carlo approximation and the iterated bootstrap. *Biometrika*, **81**, 331–340.

Booth, J. G. and Presnell, B. (1998) Allocation of Monte Carlo resources for the iterated bootstrap. *J. Comput. Graph. Statist.*, **7**, 92–112.

Hall, P. (1986) On the bootstrap and confidence intervals. *Ann. Statist.*, **14**, 1431–1452.

——— (1992) *The Bootstrap and Edgeworth Expansion*. New York: Springer.

Lee, S. M. S. and Young, G. A. (1995) Asymptotic iterated bootstrap confidence intervals. *Ann. Statist.*, **23**, 1301–1330.

Martin, M. A. (1990) On bootstrap iteration for coverage correction in confidence intervals. *J. Am. Statist. Ass.*, **85**, 1105–1118.