J. R. Statist. Soc. B (2000)
62, *Part* 3, *pp.* 479–491

Importance of interpolation when constructing double-bootstrap confidence intervals

Peter Hall,

Australian National University, Canberra, Australia

Stephen M.-S. Lee

University of Hong Kong, People's Republic of China

and G. Alastair Young

University of Cambridge, UK

[Received August 1998. Revised October 1999]

Summary. We show that, in the context of double-bootstrap confidence intervals, linear interpolation at the second level of the double bootstrap can reduce the simulation error component of coverage error by an order of magnitude. Intervals that are indistinguishable in terms of coverage error with theoretical, infinite simulation, double-bootstrap confidence intervals may be obtained at substantially less computational expense than by using the standard Monte Carlo approximation method. The intervals retain the simplicity of uniform bootstrap sampling and require no special analysis or computational techniques. Interpolation at the first level of the double bootstrap is shown to have a relatively minor effect on the simulation error.

Keywords: Confidence interval; Coverage error; Edgeworth expansion; Iterated bootstrap; Monte Carlo simulation; Resample; Simulation

1. Introduction

The double bootstrap provides a satisfactory theoretical solution to the problem of constructing nonparametric confidence intervals of low coverage error. The basic idea is to use the bootstrap itself to estimate the coverage error of a bootstrap confidence interval, and by that means to recalibrate the nominal coverage of the interval, to produce intervals of lower coverage error: see Hall (1986) and Beran (1987). The recalibration is best performed on simple forms of confidence interval, such as the percentile method interval, as introduced by Efron (1979). In these circumstances the double-bootstrap interval retains the stability properties of the uncalibrated interval but achieves a significant reduction in the coverage error: see, for example, Hall (1992), section 3.11. In practice, the drawback to the use of the double bootstrap is its computational expense, the construction of the double-bootstrap interval generally requiring a Monte Carlo simulation involving two nested levels of data resampling.

In a recent investigation of the coverage accuracy of double-bootstrap confidence intervals, Lee and Young (1999) gave formulae for the main effects of employing finite numbers of bootstrap simulations at both levels of the Monte Carlo algorithm. Denoting those numbers

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

 1369–7412/00/62479

by B and C, at the first and second levels respectively, Lee and Young's argument showed that the respective terms in formulae for the coverage error are of sizes $o(B^{-1/2})$ and $O(C^{-1})$.

A novel aspect of Lee and Young's work is the discovery that, by first using a pilot version of the double-bootstrap algorithm and then choosing C adaptively, depending on the results of the pilot study, to balance the main simulation and sampling components of coverage error, intervals of high coverage accuracy may be obtained from rather smaller Monte Carlo simulations than those previously advocated. This adaptive method, based on the sophisticated, though computationally simple, analytic methods described by Lee and Young (1995), is a corner-stone of the recommendations of Lee and Young (1999) for practical implementations of the double bootstrap.

In the present paper we show that a more elementary approach, involving only linear interpolation at the second level of the double bootstrap, reduces the effect of finite C on coverage error from $O(C^{-1})$ to $O(C^{-2}+C^{-1}n^{-1/2})$ for one-sided confidence intervals, and to $O(C^{-2}+C^{-1}n^{-1})$ for the two-sided case. (Here, n denotes the sample size.) We provide a theoretical argument verifying these improvements, and we corroborate the theory by numerical work. The benefits of interpolating at the first level of the double bootstrap are less marked; again, this is evident from both theory and simulation.

The current paper takes a somewhat different approach to the implementation of the double bootstrap from that of Lee and Young (1999) and should be viewed as complementary to that paper. Lee and Young (1999) were concerned with the optimal choice, in terms of a reduction of coverage error, of simulation size in the conventional Monte Carlo approach. Here we show that the interpolation device provides an alternative Monte Carlo approach which enables the construction of approximations to theoretical, infinite simulation, double-bootstrap intervals from a smaller simulation than required when using the conventional Monte Carlo algorithm. By contrast with other methods which have been proposed, including that of Lee and Young (1999), the method presented here has strong advantages and represents a simpler solution to the question of how to implement the double bootstrap. No sophisticated analytic method or preliminary bootstrap analysis is required, and the Monte Carlo simulation remains based on simple uniform resampling. Use of the interpolation device yields a coverage accuracy that is indistinguishable from that of infinite simulation intervals, using small values of C. We can put forward a very simple recommendation that $C=50$ is generally perfectly adequate, provided that the interpolation is applied. Intervals constructed according to this prescription are, in terms of computational speed, highly competitive with the procedures described by Lee and Young (1999), in many circumstances involving significant computational savings, without any loss in coverage accuracy.

It may be shown that the simulation component of coverage error that remains after the adaptive approach of Lee and Young (1999) to the choice of C also includes contributions of sizes C^{-2} and $C^{-1}n^{-1/2}$, the latter reducing to $C^{-1}n^{-1}$ in the two-sided case. Therefore, in order-of-magnitude terms our method provides as much reduction of simulation error as that of Lee and Young (1999), with significantly less computational expense. Note that the method of Lee and Young typically requires C of the order of n in the one-sided case and n^2 in the two-sided case. A version of Lee and Young's approach, requiring C of smaller order $n^{1/2}$ in the one-sided case and of order n in the two-sided case, might be employed to reduce further the contribution of C to the simulation error associated with the interpolation method. However, the pilot method would now require an estimation of early terms in an Edgeworth expansion and so would be less attractive.

Techniques which reduce the computational expense of the Monte Carlo simulation required by the double bootstrap have been considered by many researchers. The use of standard Monte Carlo variance reduction methods has been considered by, for example, Hinkley and Shi (1989). The use of analytic approximations based on saddlepoint methods to replace Monte Carlo simulation in the construction of double-bootstrap confidence intervals has been considered by DiCiccio *et al.* (1992a, b). Lee and Young (1996) suggested techniques for reducing computational cost based on sequential analysis of the Monte Carlo simulation. Lee and Young (1995) used truncated expansions to bootstrap quantities to provide direct, Monte Carlo simulation free, approximations to double-bootstrap confidence intervals. Ventura (1997) discussed the use of the Monte Carlo technique of recycling in nonparametric bootstrap applications.

Section 2 describes interpolation methods and outlines their properties. A derivation of the theoretical results of Section 2 is sketched in Appendix A.

2. Methodology

2.1. Double-bootstrap confidence intervals

Let $\mathcal{X} = \{X_1, \ldots, X_n\}$ denote the data, and write $\hat{\theta}$ for an estimator, computed from $\mathcal{X}$, of a parameter θ. Let $\mathcal{X}_1^*, \ldots, \mathcal{X}_B^*$ be independent resamples drawn by sampling randomly, with replacement from $\mathcal{X}$; and, conditional on $\mathcal{X}_b^*$, let $\mathcal{X}_{b1}^{**}, \ldots, \mathcal{X}_{bC}^{**}$ be independent resamples drawn in the same manner from $\mathcal{X}_b^*$. Let $\hat{\theta}_b^*$ and $\hat{\theta}_{bc}^{**}$ denote the values of $\hat{\theta}$ computed from $\mathcal{X}_b^*$ and $\mathcal{X}_{bc}^{**}$ respectively, rather than $\mathcal{X}$. Rank the sequences $\hat{\theta}_1^*, \ldots, \hat{\theta}_B^*$ and $\hat{\theta}_{b1}^{**}, \ldots, \hat{\theta}_{bC}^{**}$, obtaining $\hat{\theta}_{(1)}^* \leqslant \ldots \leqslant \hat{\theta}_{(B)}^*$ and $\hat{\theta}_{b(1)}^{**} \leqslant \ldots \leqslant \hat{\theta}_{b(C)}^{**}$ respectively.

Given $\alpha \in [k/(C+1), (k+1)/(C+1)]$, construct a second-level bootstrap estimator of the αth quantile of the distribution of $\hat{\theta}$, by interpolating between adjacent values of $\hat{\theta}_{b(j)}^{**}$:

$$\hat{\xi}_b^*(\alpha) = \{k + 1 - (C+1)\alpha\}\hat{\theta}_{b(k)}^{**} + \{(C+1)\alpha - k\}\hat{\theta}_{b(k+1)}^{**}.$$

This construction is crucial. The construction which does not interpolate would use $\hat{\theta}_{b(k)}^{**}$ or $\hat{\theta}_{b(k+1)}^{**}$ instead of $\hat{\xi}_b^*(\alpha)$, and our theoretical results are not valid unless the latter is employed. It might be noted that the methods considered by Lee and Young (1999) correspond to the use of $\hat{\theta}_{b(k)}^{**}$ instead of $\hat{\xi}_b^*(\alpha)$, but in the numerical study described in Section 2.3 we compare the interpolation method with two constructions which involve no interpolation, these using $\hat{\theta}_{bk}^{**}$ and $\hat{\theta}_{b(k+1)}^{**}$ respectively, it not being immediately clear which of these two constructions is most appropriate. Intuitively, which is best depends on the value of α, yet the whole idea of the double bootstrap, as formalized below, is to vary α in a neighbourhood of the nominal desired coverage.

Given $\alpha \in [j/(B+1), (j+1)/(B+1)]$, put

$$\hat{\xi}(\alpha) = \{j + 1 - (B+1)\alpha\}\hat{\theta}_{(j)}^* + \{(B+1)\alpha - j\}\hat{\theta}_{(j+1)}^*, \tag{2.1}$$

and let

$$\hat{\beta}(\alpha) = B^{-1}\sum_{b=1}^{B} I\{\hat{\theta} \leqslant \hat{\xi}_b^*(\alpha)\},$$

representing an approximation to the probability that $\theta \leqslant \hat{\xi}(\alpha)$.

Alternatively, we may define $\hat{\beta}$ by interpolation, as follows. Put $\hat{\alpha}_b^* = \hat{\xi}_b^{*-1}(\hat{\theta})$, and let $\hat{\alpha}_{(1)}^* \leqslant \ldots \leqslant \hat{\alpha}_{(B)}^*$ denote the ranked values of $\hat{\alpha}_b^*$. If $\alpha \in [\hat{\alpha}_{(j)}^*, \hat{\alpha}_{(j+1)}^*]$ then

$$\hat{\beta}_1(\alpha) = \frac{(\hat{\alpha}_{(j+1)}^* - \alpha)\,\hat{\beta}(\hat{\alpha}_{(j)}^*) + (\alpha - \hat{\alpha}_{(j)}^*)\,\hat{\beta}(\hat{\alpha}_{(j+1)}^*)}{\hat{\alpha}_{(j+1)}^* - \hat{\alpha}_{(j)}^*}$$

is the interpolated form of $\hat{\beta}(\alpha)$. Interpolation at this level of the double bootstrap does not reduce the order of the contribution from B in our expansion of the coverage error, however. Likewise, using interpolation to construct $\hat{\xi}(\alpha)$ is not essential; our theoretical results continue to hold if $\hat{\xi}(\alpha)$ is replaced by either $\hat{\theta}^*_{(j)}$ or $\hat{\theta}^*_{(j+1)}$.

Given $0 < \gamma < 1$, define

$$\hat{\alpha}_\gamma = \hat{\beta}^{-1}(\gamma) = \inf\{\alpha\colon \hat{\beta}(\alpha) \geqslant \gamma \text{ and } 0 < \alpha < 1\}.$$

Put

$$\bar{\beta}(\alpha) = \lim_{B,C\to\infty} \{\hat{\beta}(\alpha)\},$$
$$\bar{\xi}(\alpha) = \lim_{B\to\infty} \{\hat{\xi}(\alpha)\}$$

and $\bar{\alpha}_\gamma = \bar{\beta}^{-1}(\gamma)$. Then, $\hat{\xi}(\hat{\alpha}_\gamma)$ is our double-bootstrap approximation to the upper γ-level quantile of the distribution of $\hat{\theta}$, and $\bar{\xi}(\bar{\alpha}_\gamma)$ is the version of $\hat{\xi}(\hat{\alpha}_\gamma)$ that we would obtain if we were to do an infinite number of Monte Carlo simulations at both levels of the double bootstrap.

We shall show in the next section that the coverage error that arises through employing $\hat{\xi}(\hat{\alpha}_\gamma)$ instead of $\bar{\xi}(\bar{\alpha}_\gamma)$ to construct a one-sided confidence interval for θ is of the order $B^{-3/4} + C^{-2} + C^{-1}n^{-1/2}$. Moreover, owing to parity properties of terms in Edgeworth expansions, the contribution of size $C^{-1}n^{-1/2}$ vanishes when two one-sided intervals are intersected to form an equal-tailed, two-sided interval. In this case the effect on the coverage error of using $\hat{\xi}(\hat{\alpha}_\gamma)$ instead of $\bar{\xi}(\bar{\alpha}_\gamma)$ reduces to $O(B^{-3/4} + C^{-2} + C^{-1}n^{-1})$. The same effect is noted for the two-sided interval derived from a two-sided version of $\hat{\beta}(\alpha)$, based on a single calibration.

2.2. *Theoretical properties*

First we describe the assumptions that we make of the quantities introduced in Section 2.1. We suppose that the true value of the parameter $\theta = \theta(\mu)$ is defined in terms of the 'smooth function model' (e.g. Hall (1992), page 52 and following feature), where $\mu = E(X)$ is a mean and the data in $\mathcal{X}$ may be interpreted as independent and identically distributed observations of X. (In this context, the ith datum might not be the ith recorded observation, but a function of that quantity, chosen in a way that is convenient for placing the problem into the 'smooth function model' framework.) Then, $\hat{\theta} = \theta(\bar{X})$, where $\bar{X}$ is the mean of the data in $\mathcal{X}$. Assuming sufficiently many moments of X (representing a generic datum) and sufficient smoothness of $\theta(\cdot)$, the asymptotic variance σ^2 of $n^{1/2}\hat{\theta}$ may also be expressed through the smooth function model, so that $\sigma^2 = \sigma^2(\mu)$. (This generally involves adjoining further components to μ, and hence to X.) Let $\hat{\sigma}^2$ and $\hat{\sigma}^{*2}_b$ denote the bootstrap estimators of σ^2 from $\mathcal{X}$ and $\mathcal{X}^*_b$ respectively. Then, if

the distribution of X satisfies Cramér's condition, $\theta(\cdot)$ has sufficiently many derivatives, X has sufficiently many moments, and $\sigma^2(\mu) > 0$, (2.2)

the Edgeworth expansions

$$P\{n^{1/2}(\hat{\theta} - \theta)/\sigma \leqslant x\} = \Phi(x) + n^{-1/2}\,\psi_1(x) + O(n^{-1}),$$
$$P\{n^{1/2}(\hat{\theta} - \theta)/\hat{\sigma} \leqslant x\} = \Phi(x) + n^{-1/2}\,\psi_2(x) + O(n^{-1})$$

are valid, where $\psi_j(x) = (a_j + b_j x^2)\,\phi(x)$ for constants a_j and b_j, and Φ and ϕ are respectively the standard normal distribution and density functions. Define $a = a_1 + a_2$ and $b = b_1 + b_2$.

Let (C) denote the intersection of condition (2.2) with the assumption that, for some $\lambda > 0$,

$$\left.\begin{aligned} B(n)/\log(n) &\to \infty, \\ C(n)/\log(n) &\to \infty, \\ B(n) + C(n) &= O(n^{\lambda}). \end{aligned}\right\} \tag{2.3}$$

Let $z_\gamma = \Phi^{-1}(\gamma)$ denote the γ-level critical point of the standard normal distribution, and let $\omega_j(\gamma) = \omega_j(\gamma;\, B,\, C,\, n)$ denote quantities satisfying $0 \leqslant \omega_j(\gamma) \leqslant 1$, $j = 1, 2, 3$.

Theorem 1. Under conditions (C),

$$\begin{aligned} P\{\theta \leqslant \hat{\xi}(\hat{\alpha}_\gamma)\} = P\{\theta \leqslant \bar{\xi}(\bar{\alpha}_\gamma)\} - \tfrac{1}{4} C^{-2}\, \omega_1(\gamma) z_\gamma\, \phi(z_\gamma)^{-1} \\ + \tfrac{1}{2} C^{-1} n^{-1/2} \gamma(1-\gamma)(2b - a - 3bz_\gamma^2)\, \phi(z_\gamma)^{-1} + O(B^{-3/4} + C^{-3} + C^{-1} n^{-1}) \end{aligned} \tag{2.4}$$

uniformly in $\epsilon < \gamma < 1 - \epsilon$ for any $0 < \epsilon < \frac{1}{2}$.

The left-hand side of equation (2.4) equals the coverage of the confidence interval $\hat{I}_{1,\gamma} = (-\infty, \hat{\xi}(\hat{\alpha}_\gamma)]$, and the first term on the right-hand side represents the coverage of $\bar{I}_{1,\gamma} = (-\infty, \bar{\xi}(\bar{\alpha}_\gamma)]$.

As a corollary we obtain, for $\frac{1}{2} < \gamma < 1$, the following coverage expansion for the two-sided confidence intervals $\hat{I}_{2,\gamma} = \hat{I}_{1,(1+\gamma)/2} \setminus \hat{I}_{1,(1-\gamma)/2}$ and $\bar{I}_{2,\gamma} = \bar{I}_{1,(1+\gamma)/2} \setminus \bar{I}_{1,(1-\gamma)/2}$:

$$P(\theta \in \hat{I}_{2,\gamma}) = P(\theta \in \bar{I}_{2,\gamma}) - \tfrac{1}{2} C^{-2}\, \omega_2(\gamma) z_{(1+\gamma)/2}\, \phi(z_{(1+\gamma)/2})^{-1} + O(B^{-3/4} + C^{-3} + C^{-1} n^{-1}). \tag{2.5}$$

A two-sided interval may alternatively be defined via a two-sided version of $\hat{\beta}(\alpha)$,

$$\hat{\beta}_2(\alpha) = B^{-1} \sum_{b=1}^{B} I\{\hat{\xi}_b^*(1-\alpha) \leqslant \hat{\theta} \leqslant \hat{\xi}_b^*(\alpha)\}.$$

Putting $\hat{\alpha}_{2,\gamma} = \hat{\beta}_2^{-1}(\gamma)$, we obtain a γ-level two-sided interval $\hat{J}_{2,\gamma} = [\hat{\xi}(1 - \hat{\alpha}_{2,\gamma}),\ \hat{\xi}(\hat{\alpha}_{2,\gamma})]$, which is an approximation to $\bar{J}_{2,\gamma} = \lim_{B,C\to\infty}(\hat{J}_{2,\gamma})$ and is based on a single calibration of the nominal coverage level. Using arguments similar to those establishing equation (2.4), we have

$$P(\theta \in \hat{J}_{2,\gamma}) = P(\theta \in \bar{J}_{2,\gamma}) - \tfrac{1}{2} C^{-2}\, \omega_3(\gamma) z_{(1+\gamma)/2}\, \phi(z_{(1+\gamma)/2})^{-1} + O(B^{-3/4} + C^{-3} + C^{-1} n^{-1}), \tag{2.6}$$

which closely resembles equation (2.5).

The claims made about the coverage error at the end of Section 2.1 follow from equations (2.4)–(2.6). As in Lee and Young (1999), it may be proved that the mean length of $\hat{I}_{2,\gamma}$ (or $\hat{J}_{2,\gamma}$) minus the mean length of $\bar{I}_{2,\gamma}$ (or $\bar{J}_{2,\gamma}$) is asymptotic to a constant multiple of $C^{-1} n^{-1/2}$, and that the difference of length variances is asymptotic to a constant multiple of $B^{-1} n^{-1}$, up to terms of the same orders as those in Lee and Young's formulae (2.9) and (2.10) respectively.

Our proposals in this paper are based on linear interpolation between bootstrap percentiles. More sophisticated methods, based on interpolation between adjacent percentiles using higher order polynomials, are possible. If we employ cubic interpolation at the first bootstrap level then it may be proved that the $O(B^{-3/4})$ term is actually of size B^{-1}, and it seems likely that this is also true for lower orders of interpolation. (Indeed, a longer argument than that in Appendix A will show that the $O(B^{-3/4})$ term is actually $o(B^{-3/4})$.) That the B^{-1}-term persists despite high orders of interpolation may be seen from the fact that, as part of a Taylor expansion argument connected with the delta method, non-degenerate terms of size B^{-1} appear in Edgeworth expansions. This suggests that there is relatively little advantage in interpolating at the first level of the double bootstrap.

2.3. Numerical properties

Numerical illustrations of our results are provided by Figs 1–4. A series of 1600 replications of the double-bootstrap procedure, with and without the use of interpolation, was carried out for the construction of both one-sided and two-sided confidence intervals of nominal coverage $\gamma = 0.9$ for the population variance. Figs 1 and 2 refer to one-sided intervals based on samples of size $n = 50$ from the folded normal distribution $|N(0, 1)|$, whereas Figs 3 and 4 refer to two-sided intervals based on samples of size $n = 20$ from an exponential distribution of mean 1. The observed coverages over the 1600 replications of the intervals as functions of B and C are represented in Figs 1 and 2 for the one-sided interval $\hat{I}_{1,\gamma}$ and in Figs 3 and 4 for the two-sided interval $\hat{J}_{2,\gamma}$.

Our primary methodological point in the paper has been to argue that the use of interpolation at the inner level of the double-bootstrap construction has a marked effect on the rate at which the simulation component of the coverage error converges to 0 as C increases. Fig. 1 therefore compares coverages for intervals $\hat{I}_{1,\gamma}$, constructed using the interpolation device as described in Section 2.1, with those which do not use interpolation. Two no-interpolation intervals are considered. In the notation of Section 2.1, these employ $\hat{\theta}^{**}_{b(k)}$ ('neither k') and $\hat{\theta}^{**}_{b(k+1)}$ ('neither $k+1$') respectively, instead of $\hat{\xi}^{*}_{b}(\alpha)$ ('inner'). All intervals are constructed defining $\hat{\xi}(\alpha)$ to be $\hat{\theta}^{*}_{(j)}$, so that no interpolation is employed at the outer level

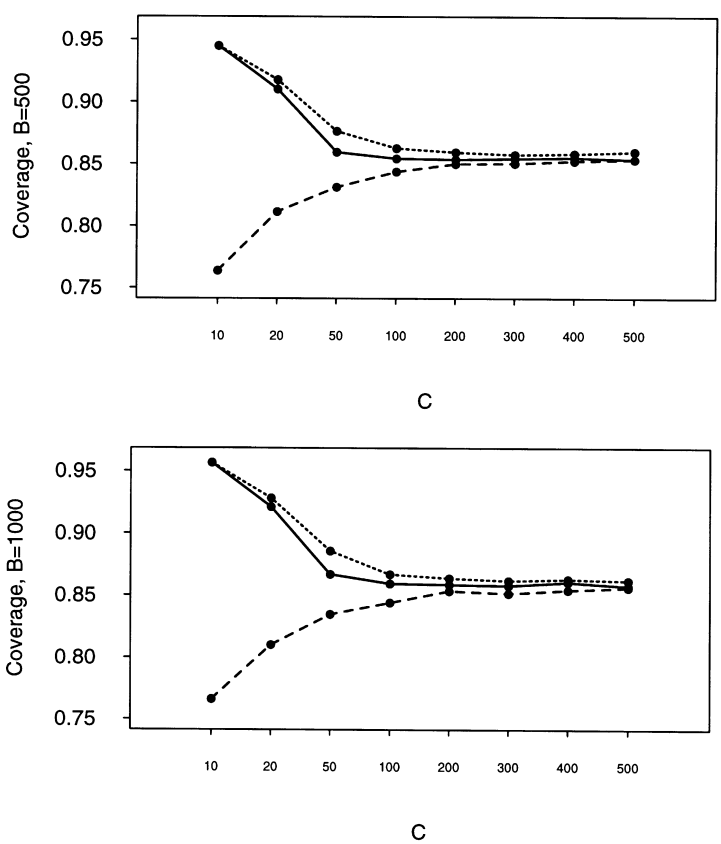

Fig. 1. Effect of inner interpolation: one-sided intervals (———, inner; — — —, neither k; ·········, neither $k+1$)

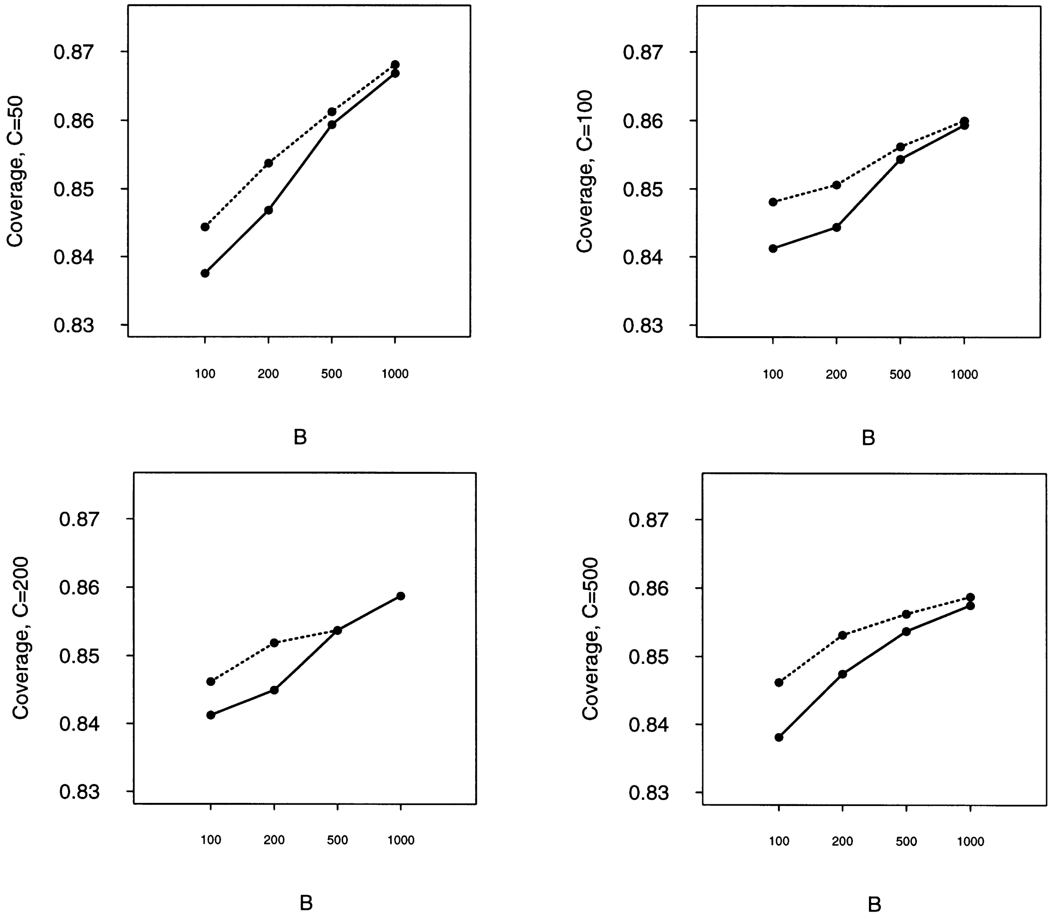

Fig. 2. Effect of outer interpolation: one-sided intervals (———, inner; ··········, both)

of the nested simulation. Fig. 1 shows, for $B = 500$ and $B = 1000$, how the observed coverage changes with the chosen value of C.

It is very clear from the Fig. 1, and the corresponding Fig. 3, which represents coverages for the two-sided interval $\hat{J}_{2,\gamma}$ for samples from the exponential distribution, that interpolation at the inner level of the Monte Carlo simulation is effective at reducing the simulation error. Interpolation ensures that the coverage of the double-bootstrap interval converges to its limiting, C infinite, value quicker than using the standard Monte Carlo approach, which does not interpolate. We note, in particular, the different qualitative behaviour in the two examples of the no-interpolation method which employs $\hat{\theta}^{**}_{b(k+1)}$. We have remarked previously that the question of which percentile, $\hat{\theta}^{**}_{b(k)}$ or $\hat{\theta}^{**}_{b(k+1)}$, is the appropriate replacement for $\hat{\xi}^{*}_{b}(\alpha)$ in the situation where no interpolation is used is subtle and may have no direct answer. The interpolation provides a safeguard against undesirably slow convergence of coverage properties with increasing C and can reduce very significantly the substantial errors that can arise from the use of a particular percentile.

These illustrative figures, together with other examples which we have studied but which are not reported here, lead us to suggest that small values of C, say 50 or at most 100, can be

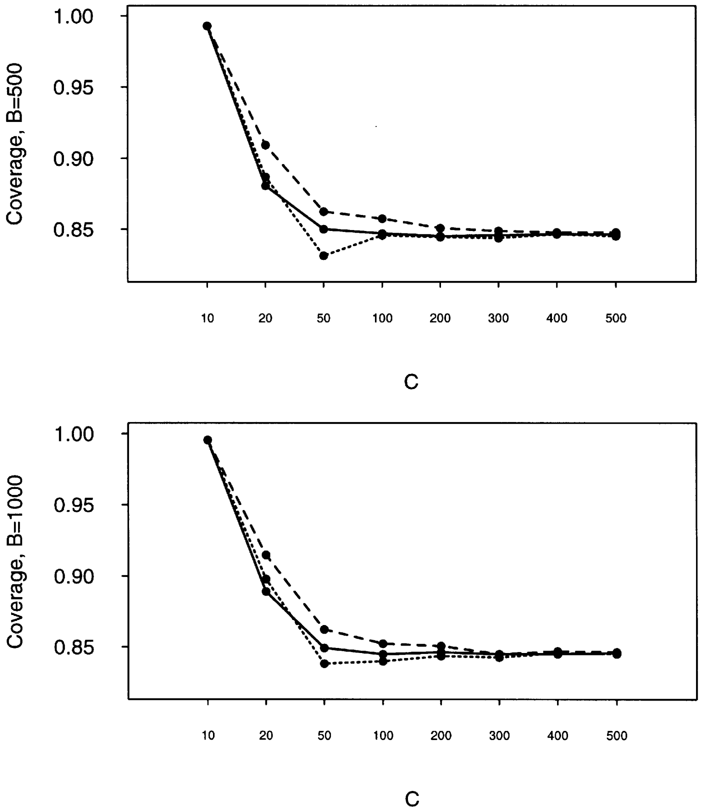

Fig. 3. Effect of inner interpolation: two-sided intervals (———, inner; — — —, neither *k*; ··········, neither $k+1$)

recommended for use in the Monte Carlo algorithm, provided that the interpolation device is used. A value of $C = 50$ is generally quite adequate to imitate the limiting case. Without interpolation, much larger C is required.

Our secondary methodological point in the paper has been to argue that interpolation at the outer level of the double bootstrap has a relatively minor effect on the simulation error, in that it has little effect on the way that the simulation component of coverage error depends on B. Fig. 2 compares, for each of four values of C, the coverages, as functions of B, the number of bootstrap samples drawn at the outer level of the simulation, of the intervals which use interpolation at the inner level, but not at the outer level, as considered above, with those which employ interpolation at both the inner and the outer levels. The latter intervals, denoted 'both' in Figs 2 and 4, are constructed defining $\hat{\xi}(\alpha)$ as in equation (2.1), rather than simply as $\hat{\theta}^*_{(j)}$. It is clear from Fig. 2 and the corresponding two-sided case represented in Fig. 4, that interpolation at the outer level has little, if any, significant effect on the speed at which the coverage figures approach their limiting, B infinite, values. We might reasonably judge that interpolation at the outer level of the simulation makes little difference to the coverage characteristics of the double-bootstrap intervals. Note also that the figures, especially Fig. 4, provide further confirmation of the view expressed by Lee and Young (1999) that B needs to be large, say of the order 500 or 1000, to ensure satisfactory coverage properties. The use of smaller B risks a substantial coverage error compared with the infinite simulation case.

The interpolation method that we have described provides the basis for a Monte Carlo

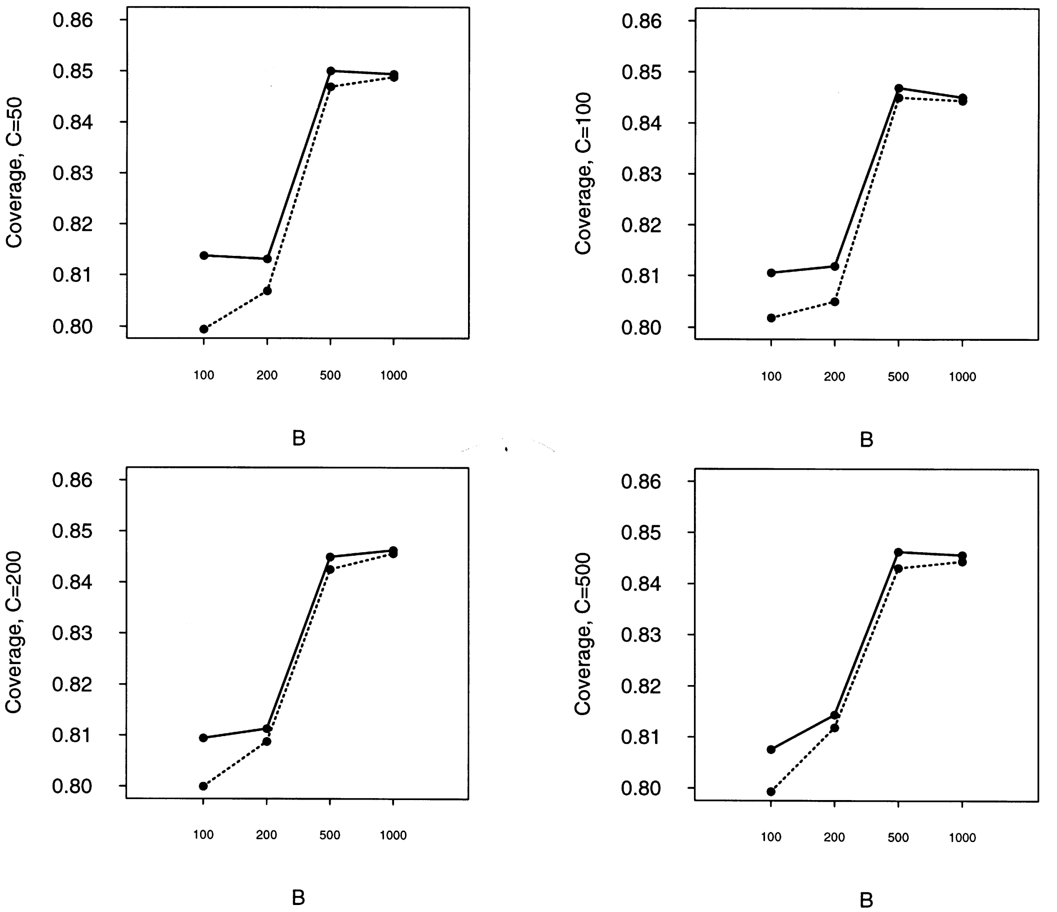

Fig. 4. Effect of outer interpolation: two-sided intervals (———, inner; ··········, both)

construction of the double-bootstrap interval which provides, quite reliably, coverage accuracy that is indistinguishable from that obtained by the theoretical, infinite simulation, double bootstrap, using small values of C. We can justify the use of $C = 50$, with the interpolation, as a general rule for the choice of simulation size across the range of problems for which the double bootstrap is indicated. Being based on such small values of C, the construction is, in terms of computational speed, highly competitive, even compared with the methods described by Lee and Young (1999). For the situation illustrated in Fig. 4, for example, the Lee and Young (1999) methodology is of the order of a tenth of the speed of the interpolation-based construction, and even for substantially larger sample sizes n the methods of the present paper remain faster, yet coverage accuracies are very similar.

Acknowledgement

The research of Stephen Lee was supported by a grant from the Hong Kong Research Grants Council.

Appendix A: Proof of theorem 1

Our derivation involves the repeated application of the 'delta method' for Edgeworth expansion, and on each occasion the arguments that are needed to justify the application are standard (although complex to apply). Therefore, they are not given. Put $\hat{\pi}(\alpha) = E\{\hat{\beta}(\alpha)|\mathcal{X}\} = P\{\hat{\theta} \leqslant \hat{\xi}_1^*(\alpha)|\mathcal{X}\}$ and

$$\Delta(\alpha) = B^{-1} \sum_{b=1}^{B} [I\{\hat{\theta} \leqslant \hat{\xi}_b^*(\alpha)\} - \hat{\pi}(\alpha)].$$

In this notation,

$$\hat{\beta}(\alpha) = \hat{\pi}(\alpha) + \Delta(\alpha). \tag{A.1}$$

Assume for the time being that $\hat{\pi}$ is three times differentiable, with bounded derivatives. Given $0 < \gamma < 1$, put $\hat{c} = \hat{\pi}^{-1}(\gamma)$ and $\hat{d} = 1/\hat{\pi}'(\hat{c}) = 1 + O_p(C^{-1} + n^{-1/2})$, and define $\hat{\eta}$ by

$$\hat{\beta}^{-1}(\gamma) = \hat{c} - \hat{d}\,\Delta(\hat{c}) + \hat{\eta}.$$

Since $\Delta(\hat{c}) = O_p(B^{-1/2})$, and the jumps in $\hat{\beta}$ are $O_p(B^{-1})$ in size, then by equation (A.1)

$$\begin{aligned}\gamma + O_p(B^{-1}) &= \hat{\beta}\{\hat{\beta}^{-1}(\gamma)\} \\ &= \hat{\pi}(\hat{c}) + \{-\hat{d}\,\Delta(\hat{c}) + \hat{\eta}\}\,\hat{\pi}'(\hat{c}) + \Delta\{\hat{c} - \hat{d}\,\Delta(\hat{c}) + \hat{\eta}\} + O_p(B^{-1} + \hat{\eta}^2) \\ &= \gamma + \Delta\{\hat{c} - \hat{d}\,\Delta(\hat{c}) + \hat{\eta}\} - \Delta(\hat{c}) + \hat{\eta}\hat{d}^{-1} + O_p(B^{-1} + \hat{\eta}^2).\end{aligned}$$

Therefore,

$$\hat{\eta} + O_p(\hat{\eta}^2) = -\hat{d}[\Delta\{\hat{c} - \hat{d}\,\Delta(\hat{c}) + \hat{\eta}\} - \Delta(\hat{c})] + O_p(B^{-1}). \tag{A.2}$$

In reality, $\hat{\pi}$ has jump discontinuities, but they are so close together that the function is, in effect, infinitely differentiable. To treat this problem formally we smooth the double-bootstrap distribution at both levels, as follows. To each datum X_{bi}^* in $\mathcal{X}_b^* = \{X_{b1}^*, \ldots, X_{bn}^*\}$, add an independent copy of $n^{-\lambda}N$, where the random vector N is of the same length as the data vectors in $\mathcal{X}$ but with a standard normal distribution, and $\lambda > 0$. Resampling from this smoothed version of $\mathcal{X}_b^*$, add to each datum X_{bci}^{**} in $\mathcal{X}_{bc}^{**} = \{X_{bc1}^{**}, \ldots, X_{bcn}^{**}\}$ another independent copy of $n^{-\lambda}N$. Since the original data vectors satisfied Cramér's condition (see expression (2.2)) then, no matter how large the value of λ, the resulting smoothed version of $\hat{\pi}$ is infinitely differentiable; and by choosing λ sufficiently large we may render asymptotically negligible the effect of smoothing.

This argument may be made rigorous by following lines in Hall (1986). The third assumption at expression (2.3) is necessary to ensure that, through taking λ sufficiently large, we may be sure that the error committed by smoothing is of smaller order than any given power of $B^{-1} + C^{-1} + n^{-1}$. The first two assumptions at expression (2.3) guarantee that the atoms of both first- and second-level bootstrap distributions are of the order $n^{-\lambda}$ for all $\lambda > 0$. Hence, in view of the last part of expression (2.3), the atoms are of orders $B^{-\lambda} + C^{-\lambda}$ for all $\lambda > 0$. Arguing in this manner we may verify equation (A.2).

It may be proved that $\Delta(\hat{c}) = O_p(B^{-1/2})$ and

$$\Delta(\hat{c} + B^{-1/2}t) - \Delta(\hat{c}) = O_p(B^{-3/4}),$$

uniformly in t in any given compact interval. Hence, by equation (A.2) and since $\hat{d} = 1 + o_p(1)$, $\hat{\eta} = O_p(B^{-3/4})$. Therefore,

$$\hat{\beta}^{-1}(\gamma) = \hat{c} - \hat{d}\,\Delta(\hat{c}) + O_p(B^{-3/4}). \tag{A.3}$$

As a prelude to describing properties of $\hat{\pi}$ we first address an analogous but simpler problem, as follows. Let $V_1, \ldots, V_C$ be independent random variables and $\mathcal{F} = \mathcal{F}(C)$ be a σ-field, with the property that, conditional on $\mathcal{F}$, $V_1, \ldots, V_C$ are independent and identically distributed with distribution function $\mathcal{F}_{\mathcal{F}}$, say. Given $\alpha \in [k/(C+1), (k+1)/(C+1)]$, let $t = k + 1 - (C+1)\alpha$ and $\xi(\alpha) = tV_{(k)} + (1-t)V_{(k+1)}$. Let w be a real number; put $\pi(\alpha) = P\{w \leqslant \xi(\alpha)\}$; conditionally on $\mathcal{F}$, let $U_{(k)}$ denote the kth-largest order statistic of a random sample of size C from the uniform distribution on the interval [0, 1]; and let $G_{\mathcal{F}} = F_{\mathcal{F}}^{-1}$, $W = F_{\mathcal{F}}(w)$ and $H(s) = P(W \leqslant U_{(k+1)}s)$. Assume that, for $j = 0, \ldots, 4$,

$$(\partial/\partial\alpha)^j\, G_{\mathcal{F}}(\alpha) = (\partial/\partial\alpha)^j \{Q_n + qn^{-1/2}\,\Phi^{-1}(\alpha)\} + O_p(n^{-1}), \tag{A.4}$$

where Q_n is an $\mathcal{F}$-measurable random variable and q is a non-zero constant. We can think of Q_n as the expectation of V_i conditional on $\mathcal{F}$ and q as a scale factor such that $n^{1/2}(V_i - Q_n)/q$ has an asymptotic standard normal distribution conditional on $\mathcal{F}$. Note that, by equation (A.4),

$$G''_{\mathcal{F}}(\alpha)\, G'_{\mathcal{F}}(\alpha)^{-1} = z_\alpha \phi(z_\alpha)^{-1} + O_p(n^{-1/2}). \tag{A.5}$$

We claim that

$$\pi(\alpha) = H(1 - tk^{-1}) + k^{-2}[\{t + (t - t^2)\alpha z_\alpha\, \phi(z_\alpha)^{-1}\}\, H'(1) + \tfrac{1}{2} t^2\, H''(1)] + O(C^{-3} + C^{-2}n^{-1/2}) \tag{A.6}$$

as $C \to \infty$.

To establish equation (A.6), note that without loss of generality we may choose the $U_{(k)}$ so that $V_{(k)} = G_{\mathcal{F}}(U_{(k)})$; and observe that $U_{(k)} = \exp(-\Sigma_{k \leqslant j \leqslant C}\, Z_j/j)$, where $Z_1, \ldots, Z_n$ are independent standard exponential random variables. In this notation,

$$(U_{(k+1)} - U_{(k)})/U_{(k+1)} = 1 - \exp(-Z_k/k).$$

(Note particularly that the right-hand side is stochastically independent of $U_{(k+1)}$.) Observe also that if $\rho(u) = u\, G''_{\mathcal{F}}(u)\, G'_{\mathcal{F}}(u)^{-1}$ and $\rho_1(u) = uz_u\, \phi(u)^{-1}$ then by equation (A.5) $\rho(u) = \rho_1(u) + O(n^{-1/2})$. Hence, by Taylor expansion,

$$\begin{aligned}
V_{(k+1)} - V_{(k)} &= (U_{(k+1)} - U_{(k)})\, G'_{\mathcal{F}}(U_{(k+1)}) - \tfrac{1}{2}(U_{(k+1)} - U_{(k)})^2\, G''_{\mathcal{F}}(U_{(k+1)}) + O_p(C^{-3}n^{-1/2}),\\
V_{(k+1)} - t(V_{(k+1)} - V_{(k)}) &= G_{\mathcal{F}}\{U_{(k+1)} - t(U_{(k+1)} - U_{(k)}) + \tfrac{1}{2}(t - t^2)(U_{(k+1)} - U_{(k)})^2\, \rho(U_{(k+1)})U_{(k+1)}^{-1}\}\\
&\quad + O_p(C^{-3}n^{-1/2}),\\
\pi(\alpha) &= P[W \leqslant F_{\mathcal{F}}\{V_{(k+1)} - t(V_{(k+1)} - V_{(k)})\}]\\
&= P\{W \leqslant U_{(k+1)} - t(U_{(k+1)} - U_{(k)}) + \tfrac{1}{2}(t - t^2)(U_{(k+1)} - U_{(k)})^2\, \rho(U_{(k+1)})U_{(k+1)}^{-1}\}\\
&\quad + O(C^{-3})\\
&= E(H[1 - t\{Z_k/k - \tfrac{1}{2}(Z_k/k)^2\} + \tfrac{1}{2}(t - t^2)(Z_k/k)^2\, \rho(\alpha)]) + O(C^{-3})\\
&= E(H(1) + [-t\{Z_k/k - \tfrac{1}{2}(Z_k/k)^2\} + \tfrac{1}{2}(t - t^2)(Z_k/k)^2\, \rho(\alpha)]\, H'(1)\\
&\quad + \tfrac{1}{2}(tZ_k/k)^2\, H''(1)) + O(C^{-3})\\
&= H(1) - tk^{-1}\, H'(1) + k^{-2}[\{t + (t - t^2)\, \rho_1(\alpha)\}\, H'(1) + t^2\, H''(1)]\\
&\quad + O(C^{-3} + C^{-2}n^{-1/2})\\
&= H(1 - tk^{-1}) + k^{-2}[\{t + (t - t^2)\, \rho_1(\alpha)\}\, H'(1) + \tfrac{1}{2} t^2\, H''(1)]\\
&\quad + O(C^{-3} + C^{-2}n^{-1/2}).
\end{aligned}$$

The latter formula implies equation (A.6).

Now consider applying equation (A.6) in the case where $V_k = \hat{\theta}^{**}_{bk}$ (so that $V_{(k)} = \hat{\theta}^{**}_{b(k)}$), $w = \hat{\theta}$ and probability is taken conditionally on $\mathcal{X}$. Then, the roles of $\pi(\alpha)$, $\mathcal{F}$, $F_{\mathcal{F}}(z)$, w, W and $H(z)$ are played by $\hat{\pi}(\alpha) = P\{\hat{\theta} \leqslant \hat{\xi}^*_b(\alpha)|\mathcal{X}\}$, the σ-field generated by $\mathcal{X}^*_b$, $\hat{F}^*_b(z) = P(\hat{\theta}^{**}_{bc} \leqslant z|\mathcal{X}^*_b)$, $\hat{\theta}$, $\hat{w}^*_b = \hat{F}^*_b(\hat{\theta})$ and $\hat{H}(s) = P(\hat{w}^*_b \leqslant U_{(k+1)}s|\mathcal{X})$ respectively, where we take $U_{(k+1)}$ to be independent of both $\mathcal{X}$ and $\mathcal{X}^*_b$. Then,

$$\begin{aligned}
\hat{F}^*_b(z) &= P\{n^{1/2}(\hat{\theta}^{**}_{bc} - \hat{\theta}^*_b)/\sigma \leqslant (z - \hat{\theta}^*_b)/\sigma|\mathcal{X}^*_b\}\\
&= \Phi\{n^{1/2}(z - \hat{\theta}^*_b)/\sigma\} + O_p(n^{-1/2}),
\end{aligned} \tag{A.7}$$

and so $(\hat{F}^*_b)^{-1}(\alpha) = \hat{\theta}^*_b + n^{-1/2}\sigma z_\alpha + O_p(n^{-1})$, which is the version of equation (A.4) in the case $j = 0$. The smoothing argument given immediately below equation (A.2), for both levels of the bootstrap algorithm, permits us to extend equation (A.4) to arbitrary $j \geqslant 0$. Hence, by equation (A.6),

$$\hat{\pi}(\alpha) = \hat{H}(1 - tk^{-1}) + k^{-2}[\{t + (t - t^2)\alpha z_\alpha\, \phi(z_\alpha)^{-1}\}\, \hat{H}'(1) + \tfrac{1}{2} t^2\, \hat{H}''(1)] + O_p(C^{-3} + C^{-2} n^{-1/2}). \tag{A.8}$$

Let $\hat{R}(u) = P(\hat{w}_b^* \leqslant u | \mathcal{X}) - u$,

$$S(u) = P(U_{(k+1)} > u) = \sum_{i=0}^{k} \binom{C}{i} u^i (1-u)^{C-i}$$

and $\hat{T}(s) = \int S(u/s)\, \mathrm{d}\hat{R}(u) = E\{\hat{R}(U_{(k+1)}s) | \mathcal{X}\}$. Then,

$$\hat{H}(s) = \int S\left(\frac{u}{s}\right) \{\mathrm{d}u + \mathrm{d}\hat{R}(u)\} = \frac{s(k+1)}{C+1} + \hat{T}(s).$$

It follows that

$$\begin{aligned}\hat{H}(1 - tk^{-1}) &= \alpha - \frac{t}{\alpha C^2} + \int \{S(u) + utk^{-1}\, S'(u)\}\, \mathrm{d}\hat{R}(u) + O_p(C^{-3} + C^{-2} n^{-1/2}) \\ &= \alpha + \hat{T}(1) - \frac{t}{\alpha C}\, \hat{T}'(1) - \frac{t}{\alpha C^2} + O_p(C^{-3} + C^{-2} n^{-1/2}),\end{aligned}$$

$\hat{H}'(1) = \alpha + O_p(C^{-1} + n^{-1/2})$ and $\hat{H}''(1) = O_p(C^{-1} + n^{-1/2})$. Therefore, by equation (A.8),

$$\hat{\pi}(\alpha) = \alpha + \hat{T}(1) - \frac{t}{\alpha C}\, \hat{T}'(1) + C^{-2}(t - t^2) z_\alpha\, \phi(z_\alpha)^{-1} + O_p(C^{-3} + C^{-2} n^{-1/2}). \tag{A.9}$$

Put $\zeta_b^* = n^{1/2}(\hat{\theta}_b^* - \hat{\theta})/\hat{\sigma}_b^*$ and observe that

$$\begin{aligned}P(\hat{\theta}_{bc}^{**} \leqslant \hat{\theta} | \mathcal{X}_b^*) &= P\{n^{1/2}(\hat{\theta}_{bc}^{**} - \hat{\theta}_b^*)/\hat{\sigma}_b^* \leqslant -\zeta_b^* | \mathcal{X}_b^*\} \\ &= \Phi(-\zeta_b^*) + n^{-1/2}\, \psi_1(-\zeta_b^*) + O_p(n^{-1}), \\ \hat{R}(u) + u &= P\{P(\hat{\theta}_{bc}^{**} \leqslant \hat{\theta} | \mathcal{X}_b^*) \leqslant u | \mathcal{X}\} \\ &= P\{\Phi(-\zeta_b^*) + n^{-1/2}\, \psi_1(-\zeta_b^*) + O_p(n^{-1}) \leqslant u | \mathcal{X}\} \\ &= P[-\zeta_b^* \leqslant \Phi^{-1}(u) - n^{-1/2}\, \psi_1\{\Phi^{-1}(u)\}\, \phi\{\Phi^{-1}(u)\}^{-1} | \mathcal{X}] + O_p(n^{-1}) \\ &= 1 - (\Phi[-\Phi^{-1}(u) + n^{-1/2}\, \psi_1\{\Phi^{-1}(u)\}\, \phi\{\Phi^{-1}(u)\}^{-1}] + n^{-1/2}\, \psi_2\{-\Phi^{-1}(u)\}) + O_p(n^{-1}) \\ &= 1 - [1 - u + n^{-1/2}\, \psi_1\{\Phi^{-1}(u)\} + n^{-1/2}\, \psi_2\{\Phi^{-1}(u)\}] + O_p(n^{-1}) \\ &= u - n^{-1/2}(\psi_1 + \psi_2)\, \Phi^{-1}(u) + O_p(n^{-1}).\end{aligned}$$

Therefore, defining $\chi(u) = (\psi_1 + \psi_2)(z_u)$, we have

$$\begin{aligned}-n^{1/2}\, \hat{T}(s) + O_p(n^{-1/2}) &= E\{\chi(U_{(k+1)}s)\} \\ &= \chi\left(\frac{k+1}{C+1} s\right) + \frac{1}{2} E\left(U_{(k+1)} - \frac{k+1}{C+1}\right)^2 s^2\, \chi''(\alpha s) + O(C^{-2}) \\ &= \chi(\alpha s) + \left(\frac{k+1}{C+1} - \alpha\right) s\, \chi'(\alpha s) + \tfrac{1}{2} C^{-1} \alpha(1-\alpha) s^2\, \chi''(\alpha s) + O(C^{-2}) \\ &= \chi(\alpha s) + C^{-1}\{st\, \chi'(\alpha s) + \tfrac{1}{2}\alpha(1-\alpha) s^2\, \chi''(\alpha s)\} + O(C^{-2}).\end{aligned}$$

Similarly, $-n^{1/2}\, \hat{T}'(1) = \alpha\, \chi'(\alpha) + O_p(C^{-1} + n^{-1/2})$, and so, by equation (A.9),

$$\begin{aligned}\hat{\pi}(\alpha) - C^{-2}(t - t^2) z_\alpha\, \phi(z_\alpha)^{-1} &= \alpha - n^{-1/2}[\chi(\alpha) + C^{-1}\{t\, \chi'(\alpha) + \tfrac{1}{2}\alpha(1-\alpha)\, \chi''(\alpha)\}] \\ &\quad + C^{-1} n^{-1/2} (t/\alpha) \alpha\, \chi'(\alpha) + O_p(C^{-3} + C^{-2} n^{-1/2} + n^{-1}) \\ &= \alpha - n^{-1/2}\, \chi(\alpha) - \tfrac{1}{2} C^{-1} n^{-1/2} \alpha(1-\alpha)\, \chi''(\alpha) + O_p(C^{-3} + n^{-1}).\end{aligned} \tag{A.10}$$

Defining $\psi = \psi_1 + \psi_2$ we have $\chi = \psi(\Phi^{-1})$, $\chi' = \psi'(\Phi^{-1})/\phi(\Phi^{-1})$ and

$$\chi'' = \psi''(\Phi^{-1})\,\phi(\Phi^{-1})^{-2} - \psi'(\Phi^{-1})\,\phi'(\Phi^{-1})\,\phi(\Phi^{-1})^{-3}.$$

Therefore, since $\psi(z) = (a + bz^2)\,\phi(z)$, then

$$\begin{aligned}\chi''(\alpha)\,\phi(z_\alpha) &= 2b - a - 3bz_\alpha^2 - z_\alpha^2(2b - a - bz_\alpha^2) - z_\alpha(2b - a - bz_\alpha^2)(-z_\alpha)\\ &= 2b - a - 3bz_\alpha^2.\end{aligned}$$

Hence, defining $\kappa(z) = \frac{1}{2}\,\alpha(1-\alpha)(2b - a - 3bz^2)$, letting $\bar{\pi} = \lim_{C\to\infty}(\hat{\pi})$ denote the version of $\hat{\pi}$ after an infinite amount of Monte Carlo simulation and noting that $C^{-2}n^{-1/2} = O(C^{-3} + C^{-1}n^{-1})$, we see from equation (A.10) that

$$\hat{\pi}(\alpha) = \bar{\pi}(\alpha) + C^{-2}t(1-t)z_\alpha\,\phi(z_\alpha)^{-1} - C^{-1}n^{-1/2}\,\kappa(z_\alpha)\,\phi(z_\alpha)^{-1} + O_p(C^{-3} + C^{-1}n^{-1}). \qquad \text{(A.11)}$$

Note that $0 \leqslant t \leqslant 1$, and so $0 \leqslant t(1-t) \leqslant \frac{1}{4}$. Therefore, by equation (A.11),

$$\hat{\pi}^{-1}(\gamma) = \bar{\pi}^{-1}(\gamma) + \{C^{-1}n^{-1/2}\,\kappa(z_\gamma) - \tfrac{1}{4}C^{-2}\,\hat{\omega}(\gamma)z_\gamma\}[\phi(z_\gamma)\bar{\pi}'\{\bar{\pi}^{-1}(\gamma)\}]^{-1} + O_p(C^{-3} + C^{-1}n^{-1}), \qquad \text{(A.12)}$$

where the random variable $\hat{\omega}(\gamma) = 4t(1-t)$ satisfies $0 \leqslant \hat{\omega}(\gamma) \leqslant 1$. Note that inversion of $\hat{\pi}$ has turned t into a random variable. Combining equations (A.3) and (A.12), and noting that $\bar{\beta} = \bar{\pi}$, we may derive an expansion of the kind given in equation (A.7) of Lee and Young (1999), where their $\hat{U}^*_{(k)}$ and $\alpha + \hat{\xi}$ are replaced respectively by $\hat{\beta}^{-1}(\gamma)$ and $\bar{\pi}^{-1}(\gamma)$. The proof can now be completed as in Lee and Young (1999). In this regard, note particularly that although the variable $\hat{d}\,\Delta(\hat{c})$ at equation (A.3) is of order $B^{-1/2}$, it has zero mean conditional on $\mathcal{X}$. This means that the contribution to the Edgeworth expansion of the main effect of B is of order at most that of the $B^{-3/4}$-term at equation (A.3), not of order $B^{-1/2}$.

References

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

DiCiccio, T. J., Martin, M. A. and Young, G. A. (1992a) Analytic approximations for iterated bootstrap confidence intervals. *Statist. Comput.*, **2**, 161–171.

——— (1992b) Fast and accurate approximate double bootstrap confidence intervals. *Biometrika*, **79**, 285–295.

Efron, B. (1979) Bootstrap methods: another look at the jackknife. *Ann. Statist.*, **7**, 1–26.

Hall, P. (1986) On the number of bootstrap simulations required to construct a confidence interval. *Ann. Statist.*, **14**, 1453–1462.

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

Hinkley, D. V. and Shi, S. (1989) Importance sampling and the iterated bootstrap. *Biometrika*, **76**, 435–446.

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

——— (1996) Sequential iterated bootstrap confidence intervals. *J. R. Statist. Soc.* B, **58**, 235–251.

——— (1999) The effect of Monte Carlo approximation on coverage error of double-bootstrap confidence intervals. *J. R. Statist. Soc.* B, **61**, 353–366.

Ventura, V. (1997) Likelihood inference by Monte Carlo methods and efficient nested bootstrapping. *DPhil Thesis*. University of Oxford, Oxford.