The Annals of Statistics
1995, Vol. 23, No. 4, 1301–1330

ASYMPTOTIC ITERATED BOOTSTRAP CONFIDENCE INTERVALS

BY STEPHEN M. S. LEE AND G. ALASTAIR YOUNG

University of Hong Kong and University of Cambridge

An iterated bootstrap confidence interval requires an additive correction to be made to the nominal coverage level of an uncorrected interval. Such correction is usually performed using a computationally intensive Monte Carlo simulation involving two nested levels of bootstrap sampling. Asymptotic expansions of the required correction and the iterated interval endpoints are used to provide two new computationally efficient methods for constructing an approximation to the iterated bootstrap confidence interval. The first asymptotic interval replaces the need for a second level of bootstrap sampling with a series of preliminary analytic calculations, which are readily automated, and from which an approximation to the coverage correction is easily obtained. The second interval directly approximates the endpoints of the iterated interval and yields, for the first time, the possibility of constructing an approximation to an iterated bootstrap confidence interval which does not require any resampling. The theoretical properties of the two intervals are considered. The computation required for their construction is detailed and has been coded in a fully automatic user-friendly Fortran program which may be obtained by anonymous ftp. A simulation study which illustrates their effectiveness on three examples is presented.

1. Introduction. An iterated bootstrap confidence interval may be derived from the percentile method described by Efron [(1982), Section 10.4]. Beran (1987) proposed a prepivoting method to reduce coverage error, which is equivalent to making an additive correction to the nominal coverage level in the percentile method. Hall and Martin (1988) discuss this coverage correction approach within a general framework of bootstrap iteration. Our notion of an iterated bootstrap confidence interval refers to this approach. A precise definition is given in Section 2.

When constructing an iterated bootstrap confidence interval, the need to calculate a correction to the nominal coverage, or "calibrating coefficient," is usually met by a double bootstrap procedure. In principle, we can iterate the procedure to reduce coverage error by a factor of $O(n^{-1})$ successively. Hall and Martin (1988) provide a unified account of the concept of iterated bootstrap. In practice we seldom go beyond the second level of bootstrap sampling, due to the increasing computational intensity rendered by successive levels of bootstrap sampling. In this paper we consider only an iterated

Received June 1993; revised December 1994.
AMS 1991 *subject classifications*. Primary 62G09, 62G15; secondary 62G20.
Key words and phrases. Asymptotic, bootstrap, coverage error, Edgeworth expansion, iterated bootstrap, percentile method, resampling.

bootstrap confidence interval based on double bootstrapping, that is, with one level of nominal coverage correction.

The computational expense of the usual Monte Carlo construction, as considered by Booth and Hall (1994), has prompted a number of authors to develop computationally efficient procedures for approximation to the iterated bootstrap confidence interval. DiCiccio, Martin and Young (1992a, b) consider the use of analytic approximation methods to replace the inner of the two levels of resampling required by the double bootstrap procedure. An alternative approach described by Booth and Do (1993) uses the variance reduction method of balanced importance sampling at the first level of the iterated bootstrap to achieve computational savings over the standard Monte Carlo construction. Lee and Young (1993) describe an approach which performs the inner of the two levels of bootstrap sampling in a sequential manner. This approach, while still demanding, like the method of Booth and Do (1993), two levels of resampling, has the advantage of only requiring simple uniform resampling. In the current paper the goal is to develop approximations to the iterated bootstrap interval which eliminate the need for nested levels of resampling and which at the same time require only simple arithmetic computation.

Hall [(1992), Section 3.11.3] gives an account of the coverage properties of the iterated bootstrap confidence interval. Asymptotic expansions are also developed for the iterated interval endpoints, as well as for the calibrating coefficient, under his "smooth function model." Hall uses these expansions to explain theoretically why bootstrap iteration works. The complicated structure of the expansions might be felt to hold out little hope for their being of practical use. In this paper, however, we demonstrate how we may obtain explicit expressions, in terms of population moments, for the leading terms in the expansions of both the calibrating coefficient and interval endpoints. These asymptotic expressions are then used as the basis of two new computationally efficient and practically feasible procedures for constructing an approximation to the iterated bootstrap confidence interval.

The first asymptotic interval replaces the need for a second level of bootstrap sampling with a series of preliminary analytic calculations from which an approximation to the calibrating coefficient is easily obtained. The second interval directly approximates the endpoints of the iterated interval. This represents, for the first time, an approximation to the iterated bootstrap confidence interval which does not require any resampling.

The analytic computations required for construction of the asymptotic intervals are automated to provide practical procedures which are computationally extremely efficient when compared to the usual iterated bootstrap confidence interval construction which uses two nested levels of data resampling. The only input required from the user is a formula for the parameter of interest in terms of population moments, the maximum order of such moments and the dimension of the underlying distribution. In particular, techniques of exact numerical derivative evaluation are introduced to replace the need for any symbolic computation. The resulting procedures thus consist

only of simple numerical steps which can be efficiently implemented using a Fortran program. The computer time required is reduced by a factor of the order of thousands compared to a full-blown iterated bootstrap confidence interval based on 1000 outer level and 1000 inner level resamples.

Section 2 gives a formal definition of the iterated bootstrap confidence interval and presents briefly the forms of the asymptotic expansions to be used in our procedure of generating asymptotic iterated confidence intervals. Section 3 develops the theory about coverage properties of these asymptotic intervals. There they are also compared with the theoretical iterated bootstrap interval. A qualitative description of the practical procedures is presented in Section 4, with special attention paid to description of how they may be implemented by use of a general package consisting of two user-friendly Fortran programs. A simulation study has been carried out on the mean, the variance and the correlation coefficient examples, and their results are reported in Section 5, followed by a general discussion in Section 6. All the analytic computations necessary for construction of the asymptotic intervals are detailed in the Appendix.

2. Problem specification. Let $\mathscr{X} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n)$ be a random sample drawn from a d-variate distribution function F. We wish to construct a two-sided, α-level iterated bootstrap confidence interval, based on $\mathscr{X}$, for a parameter θ. Assume that $\theta = g(\boldsymbol{\mu})$ is a smooth function of the mean $\boldsymbol{\mu}$ of F. The confidence interval is to be constructed from its sample estimate $\hat{\theta} = g(\overline{\mathbf{X}})$, where $\overline{\mathbf{X}} = n^{-1}\sum_{i=1}^{n}\mathbf{X}_i$. Define y_β to be the β-th quantile of the distribution of $\hat{\theta}$, so that

$$\mathbb{P}\big(\hat{\theta} \le y_\beta\big) = \beta.$$

Let $\mathscr{X}^* = (\mathbf{X}_1^*, \ldots, \mathbf{X}_n^*)$ denote a generic bootstrap resample drawn randomly from $\mathscr{X}$ with replacement, and similarly let $\mathscr{X}^{**}$ denote a generic resample from $\mathscr{X}^*$. The conditional probability and expectation based on such uniform resampling from $\mathscr{X}$ and $\mathscr{X}^*$ are denoted by $\mathbb{P}(\cdot|\mathscr{X})$, $\mathbb{E}[\cdot|\mathscr{X}]$, $\mathbb{P}(\cdot|\mathscr{X}^*, \mathscr{X})$ and so on. Let $\hat{\theta}^*$ and $\hat{\theta}^{**}$ be the versions of the statistic $\hat{\theta}$ based on $\mathscr{X}^*$ and $\mathscr{X}^{**}$, respectively. Define the βth quantile of the distribution of $\hat{\theta}^*$ conditional on $\mathscr{X}$ to be

$$\hat{y}_\beta = \sup\big\{u\colon \mathbb{P}\big(\hat{\theta}^* \le u|\mathscr{X}\big) \le \beta\big\},$$

for $0 \le \beta \le 1$. We will assume from now on that $\hat{y}_\beta$ is the exact solution to the equation

$$\mathbb{P}\big(\hat{\theta}^* \le \hat{y}_\beta|\mathscr{X}\big) = \beta,$$

since the error due to discreteness of the sampling distribution of $\hat{\theta}^*$ is negligible compared to other error terms to be met in this section [see Hall (1986)]. Put $\xi = \frac{1}{2}(1 + \alpha)$. Let t be the solution to the equation

$$\mathbb{P}\big(\hat{y}_{1-\xi-t/2} \le \theta \le \hat{y}_{\xi+t/2}\big) = \alpha. \tag{1}$$

Similarly, the bootstrap version of t is given by $\hat{t}$ which satisfies

$$\mathbb{P}\left(\hat{y}^*_{1-\xi-\hat{t}/2} \le \hat{\theta} \le \hat{y}^*_{\xi+\hat{t}/2} | \mathscr{X}\right) = \alpha, \tag{2}$$

where

$$\mathbb{P}\left(\hat{\theta}^{**} \le \hat{y}^*_\beta | \mathscr{X}^*, \mathscr{X}\right) = \beta,$$

for $0 \le \beta \le 1$. The two-sided iterated bootstrap confidence interval of nominal coverage level α is then given by

$$I_0 = \left[\hat{y}_{1-\xi-\hat{t}/2}, \hat{y}_{\xi+\hat{t}/2}\right].$$

The sample-based quantity $\hat{t}$ may be termed the calibrating coefficient of the nominal coverage level α, using the notion introduced by Loh (1987). In fact, the interval I_0 resembles a two-sided percentile method interval with its nominal coverage level α calibrated to $\alpha + \hat{t}$.

Note that $\hat{t}$ can alternatively be obtained from the equation

$$\mathbb{P}(|2U^* - 1| \le \alpha + \hat{t} | \mathscr{X}) = \alpha, \tag{3}$$

where

$$U^* = \mathbb{P}\left(\hat{\theta}^{**} \le \hat{\theta} | \mathscr{X}^*, \mathscr{X}\right).$$

In practice, the sample-based interval I_0 may be approximated by double bootstrap resampling. An inner level of resampling is required to approximate U^* from the conditional distribution of $\hat{\theta}^{**}$, while an outer level of resampling is used to approximate the distribution of U^* itself. Such a double bootstrap scheme is usually very time-consuming and computationally intensive.

Asymptotic formulae of $\hat{t}$ and $\hat{y}_\beta$ are, however, available under some mild conditions on g and F. Suppose for the time being that both t and y_β admit Edgeworth expansions as follows:

$$t = 2n^{-1}\pi_1(z_\xi)\phi(z_\xi) + 2n^{-2}\pi_2(z_\xi)\phi(z_\xi) + \cdots \tag{4}$$

and

$$y_\beta = \theta + n^{-1/2}\sigma\left\{z_\beta + n^{-1/2}p_{11}(z_\beta) + n^{-1}p_{21}(z_\beta) + \cdots\right\}, \tag{5}$$

for $0 < \beta < 1$, where the π_j's are odd polynomials, the p_{j1}'s are polynomials of degree at most $j + 1$ and are odd for even j and even for odd j; σ^2 is the asymptotic variance of $n^{1/2}(\hat{\theta} - \theta)$; ϕ is the standard normal density function; and $z_\beta = \Phi^{-1}(\beta)$ is the usual βth standard normal quantile. A detailed derivation of expansions (4) and (5) is given in Hall [(1992), Chapter 3]. Full explicit expressions for t and y_β are detailed in the Appendix. Since σ and the coefficients of the π_j's and p_{j1}'s depend only on moments of F up to a certain order, we can easily write down the corresponding expansions for their bootstrap versions $\hat{t}$ and $\hat{y}_\beta$:

$$\hat{t} = 2n^{-1}\hat{\pi}_1(z_\xi)\phi(z_\xi) + 2n^{-2}\hat{\pi}_2(z_\xi)\phi(z_\xi) + \cdots \tag{6}$$

and

$$\hat{y}_\beta = \hat{\theta} + n^{-1/2}\hat{\sigma}\left\{z_\beta + n^{-1/2}\hat{p}_{11}(z_\beta) + n^{-1}\hat{p}_{21}(z_\beta) + \cdots\right\}, \tag{7}$$

where $\hat{\pi}_j$, $\hat{p}_{j1}$ and $\hat{\sigma}$ are obtained by substituting sample moments for population moments in the definitions of π_j, p_{j1} and σ, respectively. Define

$$\tilde{t} = 2n^{-1}\hat{\pi}_1(z_\xi)\phi(z_\xi) \tag{8}$$

and

$$\tilde{y}_\beta = \hat{\theta} + n^{-1/2}\hat{\sigma}\left\{z_\beta + n^{-1/2}\hat{p}_{11}(z_\beta) + n^{-1}\hat{p}_{21}(z_\beta)\right\}. \tag{9}$$

Both expressions (8) and (9) can be evaluated exactly from sample moments, without simulation. By this means we arrive at two possible sample-based asymptotic confidence intervals, namely,

$$I_1 = \left[\hat{y}_{1-\xi-\tilde{t}/2}, \hat{y}_{\xi+\tilde{t}/2}\right] \quad \text{and} \quad I_2 = \left[\tilde{y}_{1-\xi-\tilde{t}/2}, \tilde{y}_{\xi+\tilde{t}/2}\right].$$

Note that the interval I_1 still involves the sample quantity $\hat{y}_\beta$, which has to be approximated by one level of bootstrap resampling. However, the inner level of resampling is avoided by use of $\tilde{t}$, which is computed directly without resampling. For the interval I_2, no resampling is required at all.

The next section investigates the asymptotic properties of all three intervals I_0, I_1 and I_2.

3. Theory. First we establish some notation. Define, for any vector $\mathbf{x} \in \mathbb{R}^m$,

$$\mathbf{x}^{(i)} = \text{the } i\text{th component of } \mathbf{x} \quad \text{and} \quad \|\mathbf{x}\|^2 = \sum_{i=1}^{m} (\mathbf{x}^{(i)})^2.$$

We may sometimes for convenience write $\mathbf{x}^{(i)}$ as x_i, provided no ambiguity is introduced in the context.

Throughout this section we shall make free use of Edgeworth expansions up to some order as deemed necessary in the context. Conditions for their validity are described in full detail in Hall [(1992), Chapter 5]. Briefly, sufficient conditions for the results derived in this section to hold can be summarised as follows:

1. g has eight bounded continuous derivatives in a neighbourhood around $\boldsymbol{\mu}$;
2. for some sufficiently large $\nu > 0$, $\mathbb{E}[\|\mathbf{X}\|^\nu] < \infty$;
3. F satisfies Cramér's condition, that is,

$$\limsup_{\|\mathbf{t}\| \to \infty} \chi_F(\mathbf{t}) < 1,$$

where $\chi_F(t) = \mathbb{E}[\exp(i\mathbf{t}^T\mathbf{X})]$ is the characteristic function of F.

In particular, Cramér's condition holds if F has a nondegenerate absolutely continuous component. In order to facilitate development of theoretical results about the asymptotic intervals we will assume conditions 1–3 from now on.

For convenience, append to the random vector $\mathbf{X}$ products of the form

$$\mathbf{X}^{(i_1)}\mathbf{X}^{(i_2)}, \mathbf{X}^{(i_1)}\mathbf{X}^{(i_2)}\mathbf{X}^{(i_3)}, \ldots, \mathbf{X}^{(i_1)} \cdots \mathbf{X}^{(i_p)},$$

up to a suitable order p, which must be at least the maximum order of moments of F we come across in this section. For reasons explained in the Appendix, it suffices to take $p = 6$.

Suppose now that $\mathbf{X}$ has this extended form, with dimension $d^\dagger$, say, and that $\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}] \in \mathbb{R}^{d^\dagger}$. Noting that the original $\mathbf{X}$ and $\boldsymbol{\mu}$ are exactly the first d components of their respective extended versions, we can redefine the function $g(\mathbf{x})$ on the more general domain $\mathbb{R}^{d^\dagger}$ in an obvious way so that g depends on the first d components of $\mathbf{x}$ only. Define also

$$\mu_{i_1 i_2 \cdots i_r} = \mathbb{E}\left[(\mathbf{X} - \boldsymbol{\mu})^{(i_1)}(\mathbf{X} - \boldsymbol{\mu})^{(i_2)} \cdots (\mathbf{X} - \boldsymbol{\mu})^{(i_r)}\right],$$

where $i_j = 1, 2, \ldots, d^\dagger$.

The asymptotic variance of $n^{1/2}(\hat{\theta} - \theta)$ may then be obtained as

$$\sigma^2 = \sum_{i,j=1}^{d} \left.\frac{\partial g}{\partial x_i}\right|_{\mathbf{x}=\boldsymbol{\mu}} \left.\frac{\partial g}{\partial x_j}\right|_{\mathbf{x}=\boldsymbol{\mu}} \mu_{ij}.$$

It is easy to see that σ^2 is a function of $\boldsymbol{\mu}$ if the dimension $d^\dagger$ of $\boldsymbol{\mu}$ is large enough. We can therefore define functions h and B on $\mathbb{R}^{d^\dagger}$ such that

$$h(\boldsymbol{\mu}) = \sigma \quad \text{and} \quad B(\mathbf{x}) = \frac{g(\mathbf{x}) - g(\boldsymbol{\mu})}{h(\mathbf{x})}.$$

Then $n^{1/2}B(\tilde{\mathbf{X}}) = n^{1/2}(\hat{\theta} - \theta)/\hat{\sigma}$ represents a standardized version of $\hat{\theta}$, where $\hat{\sigma}$ is the sample version of σ.

Define also

$$b_{i_1 i_2 \cdots i_r} = \left.\frac{\partial^r B(\mathbf{x})}{\partial x_{i_1} \cdots \partial x_{i_r}}\right|_{\mathbf{x}=\boldsymbol{\mu}},$$

for $i_j = 1, 2, \ldots, d^\dagger$. It is easily seen that

$$\sum_{i,j=1}^{d^\dagger} b_i b_j \mu_{ij} = 1. \tag{10}$$

We now consider the coverage error of I_0. It may be shown using the general arguments of Hall and Martin (1988) that the coverage error of I_0 is of order $O(n^{-2})$. Martin (1990) provides an explicit expression for this leading $O(n^{-2})$ term, defined in terms of coefficients of an expansion of the joint density of

$$\left[n^{1/2}(\hat{\theta} - \theta)/\hat{\sigma},\ n^{1/2}(\hat{\pi}_1(z) - \pi_1(z))\right].$$

We present here a different approach to obtaining an explicit leading term for the coverage error of I_0, which depends only on derivatives of the smooth function g. In fact, this approach may be used for any form of interval with endpoints expressible in terms of a known Edgeworth expansion, as will be appreciated later in this section. A detailed comparison can then be made between I_0 and its asymptotic competitors, I_1 and I_2, through explicit evaluation of the leading terms of their respective coverage errors.

The coverage error of a nominal α-level confidence interval I for θ is defined simply as $\mathbb{P}(\theta \in I) - \alpha$. Set $T_n = n^{1/2}(\hat{\theta} - \theta)/\hat{\sigma}$. Using (7) and noting that

$$\hat{t} - t = 2n^{-1}\big(\hat{\pi}_1(z_\xi) - \pi_1(z_\xi)\big)\phi(z_\xi) + O_p(n^{-5/2}),$$

we may write

$$\mathbb{P}\big(\theta \le \hat{y}_{\beta+\hat{t}/2}\big) = \mathbb{P}\big(S_n \ge -z_{\beta+t/2} - n^{-1/2}\hat{p}_{11}(z_{\beta+t/2}) - \cdots\big) + O(n^{-5/2}), \tag{11}$$

where $S_n = T_n + \delta_n$ and

$$\delta_n = \phi(z_\beta)^{-1}\phi(z_\xi)\big\{n^{-1}\big(\hat{\pi}_1(z_\xi) - \pi_1(z_\xi)\big) + n^{-3/2}p'_{11}(z_\beta)\big(\hat{\pi}_1(z_\xi) - \pi_1(z_\xi)\big)\big\}.$$

The order term $O(n^{-5/2})$ is deduced via the delta method [see Hall (1992), Section 2.7]. We first detail a few lemmas before we establish an Edgeworth expansion of the distribution of S_n.

LEMMA 1. *For any statistic $\Delta_n = O_p(n^{-3/2})$ under distribution F, define*

$$W_n = T_n + \Delta_n.$$

Let κ_j and κ'_j be the jth cumulants of T_n and W_n, respectively. Then the difference $(\kappa'_j - \kappa_j)$ is given, up to order $O_p(n^{-5/2})$, by the coefficient of $(it)^{j-1}/j!$ in the power series expansion in (it) of the function

$$Q_n(it) = \frac{\mathbb{E}[\Delta_n \exp(itT_n)]}{\mathbb{E}[\exp(itT_n)]}.$$

PROOF. Note that κ'_j and κ_j are given by coefficients of $(it)^j/j!$ in the power series of $\log \mathbb{E}[\exp(itW_n)]$ and $\log \mathbb{E}[\exp(itT_n)]$, respectively. Consider

$$\log \mathbb{E}[e^{itW_n}] = \log\left\{\sum_{j\ge 0}\frac{(it)^j}{j!}\sum_{k=0}^{j}\binom{j}{k}\mathbb{E}\big[T_n^{j-k}\Delta_n^k\big]\right\} = \log \mathbb{E}[e^{itT_n}] + itQ_n(it) + O(n^{-3}).$$

The last equality holds because $Q_n(it)$ is of order $O(n^{-3/2})$. Thus κ'_j and κ_j differ by the coefficient of $(it)^j/j!$ in $itQ_n(it)$ plus an error of order $O(n^{-3})$. □

COROLLARY 1. *Define $m_j = \mathbb{E}[T_n^j]$ for $j = 1, 2, \ldots$. Then we have*

$$\kappa'_j - \kappa_j = \mathbb{E}\big[f_j(T_n)\Delta_n\big] + O(n^{-5/2}),$$

where the $f_j(T_n)$'s are polynomials in T_n. In particular, we have the following:

(i) $f_1(T_n) = 1$;

(ii) $f_2(T_n) = 2(T_n - m_1)$;

(iii) $f_3(T_n) = 3T_n^2 - 6T_n m_1 - 3m_2$;

(iv) $f_4(T_n) = 4T_n^3 - 12T_n^2 m_1 - 12T_n m_2 + 24m_1 m_2 - 4m_3$;

(v) $f_5(T_n) = 5T_n^4 - 20T_n^3 m_1 - 30T_n^2 m_2 + 120T_2 m_1 m_2 + 30m_2^2 - 20T_n m_3 - 5m_4$;

(vi) $f_6(T_n) = 6T_n^5 - 30T_n^4 m_1 - 60T_n^3 m_2 + 360T_n^2 m_1 m_2 + 180T_n m_2^2 - 540m_1 m_2^2 - 60T_n^2 m_3 + 120m_2 m_3 - 30T_n m_4 + 60m_1 m_4 - 6m_5$.

PROOF. The proof is a tedious exercise of expanding $Q_n(it)$ in a power series in (it) and using the fact that m_j has order $O(n^{-1/2})$ for odd j and $O(1)$ for even j. □

COROLLARY 2. *The following hold for the polynomials f_j*:

(i) $\mathbb{E}[f_1(T_n)] = 1$;
(ii) $\mathbb{E}[f_j(T_n)] = O(n^{-1})$, *for* $j = 2, 3, \ldots$.

PROOF. For any statistic Y and nonrandom constant C, the second- and higher-order cumulants of Y and $Y + C$ coincide. The result follows by letting $\Delta_n = n^{-3/2}C$ in Corollary 1. □

LEMMA 2. *Let $\omega(\mathbf{x})$ be a smooth function on $\mathbb{R}^{d^\dagger}$. Suppose ω has three bounded continuous derivatives in a neighbourhood of $\boldsymbol{\mu}$. Set*

$$\Gamma_n = n^{1/2}\big(\omega(\bar{\mathbf{X}}) - \omega(\boldsymbol{\mu})\big) \quad \text{and} \quad \omega_{i_1 i_2 \cdots i_r} = \left.\frac{\partial^r \omega(\mathbf{x})}{\partial x_{i_1} \cdots \partial x_{i_r}}\right|_{\mathbf{x}=\boldsymbol{\mu}},$$

for $i_j = 1, 2, \ldots, d^\dagger$. Then

(i) $$\mathbb{E}[f_1(T_n)\Gamma_n] = n^{-1/2}\tfrac{1}{2}\sum_{i,j=1}^{d^\dagger} \omega_{ij}\mu_{ij} + O(n^{-3/2});$$

(ii) $$\mathbb{E}[f_2(T_n)\Gamma_n] = 2\sum_{i,j=1}^{d^\dagger} b_i \omega_j \mu_{ij} + O(n^{-1});$$

$$\text{(iii)}\quad \mathbb{E}[f_3(T_n)\Gamma_n] = n^{-1/2}3\Bigg\{\sum_{i,j,k=1}^{d^\dagger} b_i b_j \omega_k \mu_{ijk} + \sum_{i,j,k,l=1}^{d^\dagger} (2b_i b_{jk}\omega_l \mu_{ij}\mu_{kl} + b_i b_j \omega_{kl}\mu_{ik}\mu_{jl})\Bigg\} + O(n^{-3/2});$$

$$\text{(iv)}\quad \mathbb{E}[f_j(T_n)\Gamma_n] = O(n^{-1}) \quad \text{for } j = 4,5,6.$$

PROOF. Again, the results are obtained by expanding $f_j(T_n)$ and Γ_n in Taylor series with terms $(\bar{\mathbf{X}} - \boldsymbol{\mu})^{(i_1)}(\bar{\mathbf{X}} - \boldsymbol{\mu})^{(i_2)} \cdots (\bar{\mathbf{X}} - \boldsymbol{\mu})^{(i_r)}$, using Corollary 1 and identity (10), and taking expectations. □

Now recall that $\pi_1(z)$ is a polynomial in z with coefficients depending on $\boldsymbol{\mu}$ only. Therefore, $\pi_1(z)$ can be generalized to a function $\pi_1(z;\mathbf{x})$ on $\mathbb{R} \times \mathbb{R}^{d^\dagger}$, with $\pi_1(z;\boldsymbol{\mu}) = \pi_1(z)$. Define

$$\eta_{i_1 i_2 \cdots i_r}(z) = \frac{\partial^r \pi_1(z;\mathbf{x})}{\partial x_{i_1} \cdots \partial x_{i_r}}\bigg|_{\mathbf{x}=\boldsymbol{\mu}}$$

for $i_j = 1,2,\ldots,d^\dagger$. For convenience, we write $\eta_{i_1 \cdots i_r} = \eta_{i_1 \cdots i_r}(z_\xi)$. Letting $\Delta_n = \delta_n$ in Corollary 1, and setting

$$\omega(\mathbf{x}) = n^{-3/2}\pi_1(z_\xi;\mathbf{x})\phi(z_\beta)^{-1}\phi(z_\xi)$$

and

$$\omega(\mathbf{x}) = n^{-2}p'_{11}(z_\beta)\pi_1(z_\xi;\mathbf{x})\phi(z_\beta)^{-1}\phi(z_\xi)$$

successively in Lemma 2, we deduce

$$\begin{aligned}
\kappa'_1 - \kappa_1 &= n^{-2}\tfrac{1}{2}\tau_1\phi(z_\beta)^{-1}\phi(z_\xi) + O(n^{-5/2}),\\
\kappa'_2 - \kappa_2 &= 2\big(n^{-3/2} + n^{-2}p'_{11}(z_\beta)\big)\tau_2\phi(z_\beta)^{-1}\phi(z_\xi) + O(n^{-5/2}),\\
\kappa'_3 - \kappa_3 &= n^{-2}3\tau_3\phi(z_\beta)^{-1}\phi(z_\xi) + O(n^{-5/2}),\\
\kappa'_j - \kappa_j &= O(n^{-5/2}) \quad \text{for } j = 4,5,6,
\end{aligned} \tag{12}$$

where κ_j and κ'_j denote the jth cumulants of T_n and S_n, respectively, and

$$\begin{aligned}
\tau_1 &= \sum_{i,j=1}^{d^\dagger} \eta_{ij}\mu_{ij},\\
\tau_2 &= \sum_{i,j=1}^{d^\dagger} b_i \eta_j \mu_{ij} \text{ and}\\
\tau_3 &= \sum_{i,j,k=1}^{d^\dagger} b_i b_j \eta_k \mu_{ijk} + \sum_{i,j,k,l=1}^{d^\dagger} (2b_i b_{jk}\eta_l \mu_{ij}\mu_{kl} + b_i b_j \eta_{kl}\mu_{ik}\mu_{jl}).
\end{aligned} \tag{13}$$

Note that $\mathbb{P}(T_n \le x)$ admits an Edgeworth expansion

$$\mathbb{P}(T_n \le x) = \Phi(x) + n^{-1/2}q_1(x)\phi(x) + n^{-1}q_2(x)\phi(x) + n^{-3/2}q_3(x)\phi(x) + n^{-2}q_4(x)\phi(x) + O(n^{-5/2}), \tag{14}$$

where each q_j is a polynomial with coefficients depending on cumulants of F, and q_j is odd for even j and even for odd j. In particular, we can show that

$$\begin{aligned}
q_1(x) &= -\left\{k_{12} + \frac{k_{31}(x^2-1)}{6}\right\},\\
q_2(x) &= -x\left\{\frac{(k_{12}^2 + k_{22})}{2} + \frac{(4k_{12}k_{31} + k_{41})(x^2-3)}{24} + \frac{k_{31}^2(x^4 - 10x^2 + 15)}{72}\right\},\\
q_3(x) &= -\left\{k_{13} + \frac{(3k_{12}k_{22} + k_{32})(x^2-1)}{6} + \frac{k_{22}k_{31}(x^4 - 6x^2 + 3)}{12}\right\}\\
&\quad + r_3(x; k_{12}, k_{31}, k_{41}, k_{51}),\\
q_4(x) &= -x\left\{\frac{(2k_{12}k_{13} + k_{23})}{2} + \frac{(6k_{12}^2k_{22} + 3k_{22}^2 + 4k_{13}k_{31} + 4k_{12}k_{32})(x^2-3)}{24}\right.\\
&\quad + \frac{(12k_{12}k_{22}k_{31} + 4k_{31}k_{32} + 3k_{22}k_{41})(x^4 - 10x^2 + 15)}{144}\\
&\quad \left. + \frac{k_{22}k_{31}^2(x^6 - 21x^4 + 105x^2 - 105)}{144}\right\}\\
&\quad + r_4(x; k_{12}, k_{31}, k_{41}, k_{42}, k_{51}, k_{61}),
\end{aligned} \tag{15}$$

where the k_{ij} are constants depending on F such that

$$\kappa_j = n^{-(j-2)/2}\left(k_{j1} + n^{-1}k_{j2} + n^{-2}k_{j3} + \cdots\right) \tag{16}$$

for $j = 1, 2, 3, \ldots$, with $k_{11} = 0$ and $k_{21} = 1$. The remainders $r_i(x; k_{12}, \ldots)$ are polynomials with coefficients depending only on those k_{ij} included in their brackets. In fact, the coefficients of each r_i are unaffected by changes (12) up to an error of order $O(n^{-5/2})$ and so need not be considered in the distribution of S_n on the right-hand side of (11). Hall [(1992), Theorem 2.1] verifies (16) by an elaborate proof. Knowing expressions (15) and (16), we can

therefore incorporate the cumulant differences as given by (12) into expansion (14) and obtain a corresponding Edgeworth expansion for S_n:

$$\begin{aligned}\mathbb{P}(S_n \le x) &= \mathbb{P}(T_n \le x) - n^{-3/2}\tau_2\phi(z_\beta)^{-1}\phi(z_\xi)x\phi(x) \\ &\quad - n^{-2}\left\{\tfrac{1}{2}\tau_1 + \left(k_{12}\tau_2 + \tfrac{1}{2}\tau_3\right)(x^2-1)\right. \\ &\qquad \left. + \tfrac{1}{6}k_{31}\tau_2(x^4 - 6x^2 + 3) + p'_{11}(z_\beta)\tau_2 x\right\} \\ &\quad \times \phi(z_\beta)^{-1}\phi(z_\xi)\phi(x) + O(n^{-5/2}).\end{aligned} \tag{17}$$

Now set

$$\delta'_n = n^{-1/2}\hat{p}_{11}(z_{\beta+t/2}) + n^{-1}\hat{p}_{21}(z_{\beta+t/2}) + \cdots,$$
$$S'_n = S_n + \delta'_n \quad \text{and} \quad T'_n = T_n + \delta'_n.$$

Using similar arguments as in Lemma 1, we have

$$\log \mathbb{E}[\exp(itS'_n)] = \log \mathbb{E}[\exp(itT'_n)] + itQ_n + O(n^{-5/2}), \tag{18}$$

where $Q_n = \mathbb{E}[\delta_n \exp(itT_n)]/\mathbb{E}[\exp(itT_n)]$. Thus, by similar arguments as those leading to (12), we know that cumulants of S'_n differ from those of T'_n by the same differences as given by (12), up to errors of order $O(n^{-5/2})$. Therefore the distributions of S'_n and T'_n are related in a similar way as that between S_n and T_n as given by (17), with k_{12} and k_{31} replaced by the corresponding k'_{12} and k'_{31} for T'_n. Since

$$\mathbb{E}[T'_n] = \kappa_1 + n^{-1/2}p_{11}(z_\beta) + O(n^{-1}),$$

we can adjust k_{12} to be

$$k'_{12} = k_{12} + p_{11}(z_\beta) + O(n^{-1/2}). \tag{19}$$

Similarly, by considering the third cumulant of T'_n, we have

$$k'_{31} = k_{31} + O(n^{-1/2}). \tag{20}$$

Substituting (19) and (20) for k_{12} and k_{31} in (17), we obtain

$$\begin{aligned}\mathbb{P}(S'_n \le x) &= \mathbb{P}(T'_n \le x) - n^{-3/2}\tau_2\phi(z_\beta)^{-1}\phi(z_\xi)x\phi(x) \\ &\quad - n^{-2}\left\{\tfrac{1}{2}\tau_1 + \left(k'_{12}\tau_2 + p_{11}(z_\beta)\tau_2 + \tfrac{1}{2}\tau_3\right)(x^2-1)\right. \\ &\qquad \left. + \tfrac{1}{6}k_{31}\tau_2(x^4 - 6x^2 + 3) + p'_{11}(z_\beta)\tau_2 x\right\} \\ &\quad \times \phi(z_\beta)^{-1}\phi(z_\xi)\phi(x) + O(n^{-5/2}).\end{aligned} \tag{21}$$

Setting $x = -z_{\beta+t/2}$ and using (11), we deduce

$$\begin{aligned}\mathbb{P}\left(\theta \le \hat{y}_{\beta+\hat{t}/2}\right) &= \mathbb{P}\left(T'_n \ge -z_{\beta+t/2}\right) - n^{-3/2}z_\beta\tau_2\phi(z_\xi) \\ &\quad + n^{-2}\lambda(\beta)\phi(z_\xi) + O(n^{-5/2}),\end{aligned} \tag{22}$$

where

$$\begin{aligned}\lambda(\beta) &= \tfrac{1}{2}\tau_1 + \big(k_{12}\tau_2 + p_{11}(z_\beta)\tau_2 + \tfrac{1}{2}\tau_3\big)\big(z_\beta^2 - 1\big)\\ &\quad + \tfrac{1}{6}k_{31}\tau_2\big(z_\beta^4 - 6z_\beta^2 + 3\big)\\ &\quad - z_\beta p'_{11}(z_\beta)\tau_2.\end{aligned} \tag{23}$$

Changing the signs of t and $\hat{t}$ throughout, we obtain

$$\begin{aligned}\mathbb{P}\big(\theta \le \hat{y}_{\beta - \hat{t}/2}\big) &= \mathbb{P}\big(T_n' \ge -z_{\beta - t/2}\big) + n^{-3/2}z_\beta\tau_2\phi(z_\xi)\\ &\quad - n^{-2}\lambda(\beta)\phi(z_\xi) + O(n^{-5/2}).\end{aligned} \tag{24}$$

Finally, setting $\beta = \xi$ in (22) and $\beta = 1 - \xi$ in (24) and then subtracting, we have

$$\begin{aligned}\mathbb{P}(\theta \in I_0) &= \mathbb{P}\big(\theta \le \hat{y}_{\xi + \hat{t}/2}\big) - \mathbb{P}\big(\theta \le \hat{y}_{1 - \xi - \hat{t}/2}\big)\\ &= \mathbb{P}\big(\theta \in \big[\hat{y}_{1-\xi-t/2}, \hat{y}_{\xi+t/2}\big]\big)\\ &\quad + n^{-2}2\lambda(\xi)\phi(z_\xi) + O(n^{-5/2}),\end{aligned} \tag{25}$$

using the facts that $z_{1-\xi} = -z_\xi$ and $\lambda(\xi) = \lambda(1 - \xi)$. Noting that $[\hat{y}_{1-\xi-t/2}, \hat{y}_{\xi+t/2}]$ is the correct α-level interval, we have proved the following proposition.

PROPOSITION 1. *The coverage error of I_0 is given by*

$$\mathbb{P}(\theta \in I_0) - \alpha = n^{-2}2\lambda(\xi)\phi(z_\xi) + O(n^{-5/2}),$$

where $\xi = (1 + \alpha)/2$ and $\lambda(\xi)$ is defined as in (23).

By modifying the definition of δ_n slightly, we may obtain the coverage errors of I_1 and I_2 as given by the following propositions.

PROPOSITION 2. *The coverage error of I_1 is given by*

$$\mathbb{P}(\theta \in I_1) - \alpha = n^{-2}2\big\{\lambda(\xi) - \pi_2(z_\xi)\big\}\phi(z_\xi) + O(n^{-5/2}),$$

where $\xi = (1 + \alpha)/2$, and $\lambda(\xi)$ and $\pi_2(z_\xi)$ are defined as in (23) *and* (4), *respectively.*

PROOF. Take

$$\begin{aligned}\delta_n = \phi(z_\beta)^{-1}\phi(z_\xi)\big\{&n^{-1}\big(\hat{\pi}_1(z_\xi) - \pi_1(z_\xi)\big)\\ &+ n^{-3/2}p'_{11}(z_\beta)\big(\hat{\pi}_1(z_\xi) - \pi_1(z_\xi)\big) - n^{-2}\pi_2(z_\xi)\big\}.\end{aligned}$$

Clearly, from Corollary 2, we have here the same situation as in the case of I_0, except that now

$$\kappa_1' - \kappa_1 = n^{-2}\big(\tfrac{1}{2}\tau_1 - \pi_2(z_\xi)\big)\phi(z_\beta)^{-1}\phi(z_\xi) + O(n^{-5/2})$$

in (12). The result follows immediately. □

Define

$$\zeta(z) = \sum_{i,j=1}^{d^\dagger} b_i \mu_{ij} \frac{\partial p_{31}(z;\mathbf{x})}{\partial x_j}\bigg|_{\mathbf{x}=\boldsymbol{\mu}} \tag{26}$$

and

$$\begin{aligned}\lambda_0(\beta) = &\left\{k_{12} + p_{11}(z_\beta) + \tfrac{1}{6}k_{31}\left(z_\beta^2 - 3\right)\right\} z_\beta p_{31}(z_\beta) \\ &+ z_\beta \zeta(z_\beta) - p_{41}(z_\beta),\end{aligned} \tag{27}$$

where $p_{31}(z;\mathbf{x})$ is obtained by generalising $p_{31}(z)$ as in the case of $\pi_1(z;\mathbf{x})$, and $p_{41}(z_\beta)$ is obtained from (5).

PROPOSITION 3. *The coverage error of I_2 is given by*

$$\mathbb{P}(\theta \in I_2) - \alpha = n^{-2} 2\{\lambda(\xi) - \pi_2(z_\xi) + \lambda_0(\xi)\}\phi(z_\xi) + O(n^{-5/2}),$$

where $\xi = \frac{1}{2}(1+\alpha)$, and the functions $\lambda(\xi)$, $\pi_2(z_\xi)$ and $\lambda_0(\xi)$ are given by formulae (23), (4) *and* (27), *respectively.*

PROOF. Take

$$\begin{aligned}\delta_n = \phi(z_\beta)^{-1}\phi(z_\xi)\{&n^{-1}(\hat{\pi}_1(z_\xi) - \pi_1(z_\xi)) \\ &+ n^{-3/2} p'_{11}(z_\beta)(\hat{\pi}_1(z_\xi) - \pi_1(z_\xi)) - n^{-2}\pi_2(z_\xi)\} \\ &- n^{-3/2}\hat{p}_{31}(z_\beta) - n^{-2}p_{41}(z_\beta).\end{aligned}$$

The changes in (12) then become

$$\begin{aligned}\kappa'_1 - \kappa_1 = &-n^{-3/2}p_{31}(z_\beta) + n^{-2}\left(\tfrac{1}{2}\tau_1 - \pi_2(z_\xi)\right)\phi(z_\beta)^{-1}\phi(z_\xi) \\ &- n^{-2}p_{41}(z_\beta) + O(n^{-5/2}),\end{aligned}$$

and

$$\begin{aligned}\kappa'_2 - \kappa_2 = &2\left(n^{-3/2} + n^{-2}p'_{11}(z_\beta)\right)\tau_2\phi(z_\beta)^{-1}\phi(z_\xi) \\ &- 2n^{-2}\zeta(z_\beta) + O(n^{-5/2}).\end{aligned}$$

It follows from (22) that

$$\begin{aligned}&\mathbb{P}\left(\theta \le \tilde{y}_{\beta+\tilde{t}/2}\right) \\ &\quad = \mathbb{P}\left(T'_n \ge -z_{\beta+t/2}\right) - n^{-3/2}z_\beta\tau_2\phi(z_\xi) - n^{-3/2}p_{31}(z_\beta)\phi(z_\beta) \\ &\qquad + n^{-2}\{\lambda(\beta) - \pi_2(z_\xi)\}\phi(z_\xi) + n^{-2}\lambda_0(\beta)\phi(z_\beta) + O(n^{-5/2}).\end{aligned}$$

Similarly, we may obtain an expression for $\mathbb{P}(\theta \le \tilde{y}_{\beta-\tilde{t}/2})$. The result follows by setting $\beta = \xi$ in the first expression and $1-\xi$ in the second and then subtracting, using the facts that $\lambda_0(1-\xi) = -\lambda_0(\xi)$, p_{31} is an even function and p_{41} is odd. □

All three intervals I_0, I_1 and I_2 enjoy a coverage error of $O(n^{-2})$. Among them, I_2 is the most efficient to evaluate, since it requires no resampling but

simply substitution of sample moments into formulae (8) and (9). The next section and the Appendix describe how these formulae can be calculated in practice through an automatic procedure requiring no preliminary analytic calculation from the user.

It should be noted that the last term in (9) is essential for a coverage error of $O(n^{-2})$, as it corrects for kurtosis. Scrapping this term from I_2 results in an extra coverage error of $-2n^{-1}p_{21}(z_\xi)\phi(z_\xi)$, as can be deduced using arguments similar to those given earlier.

A potential drawback in the use of I_2 is the possibility of its giving an empty interval when $\tilde{y}_{1-\xi-\tilde{t}/2} \geq \tilde{y}_{\xi+\tilde{t}/2}$. This problem is most significant if the theoretical iterated bootstrap interval I_0 has a poor coverage in the first place. Therefore, the interval I_2 may have potential use for diagnostic purposes in validating the bootstrap method itself.

On the other hand, the interval I_1 requires, in general, a single outer level of resampling and is approximated by

$$\left[\hat{\theta}^*_{([B(1-\xi-\tilde{t}/2)]+1)}, \hat{\theta}^*_{([B(\xi+\tilde{t}/2)]+1)}\right],$$

where $[\cdot]$ denotes the integer-part function, possibly with some adjustment to ensure well-ordering of the end points. The details of such adjustment are given at the end of the Appendix. The need for one level of resampling makes I_1 less efficient to calculate than I_2. Also, the Monte Carlo construction introduces a simulation error of size $O(n^{-1/2}B^{-1/2})$ to its endpoints [see Booth and Hall (1994)]. This is inherited by the coverage error as a term of $O(B^{-1/2})$, according to a delta method calculation. The analytic approximation method suggested by DiCiccio, Martin and Young (1992a, b), which estimates t by analytic means, provides an alternative to the interval I_1.

The leading coverage error terms of all three of the intervals considered in this section consist mainly of leading terms from $\mathbb{E}[T_n^j(\hat{t}-t)]$ for $j=0,1,2$, as well as those from the mean and skewness of T_n. The half-asymptotic, half-resampled interval I_1 brings in an extra error term $-n^{-2}2\pi_2(z_\xi)\phi(z_\xi)$, owing to approximation of $\hat{t}$ by the leading term in its asymptotic expansion. The coverage error of the completely asymptotic interval I_2 includes other terms involving higher-order cumulants of T_n. There is no straight clear message from these coverage error expressions about the relative performance of these intervals. Since $\delta_n = O(n^{-3/2})$ in all three cases, we have

$$\hat{y}_{\xi+\hat{t}/2} - \hat{y}_{\xi+t/2} = \hat{y}_{\xi+\tilde{t}/2} - \hat{y}_{\xi+t/2} = \tilde{y}_{\xi+\tilde{t}/2} - \hat{y}_{\xi+t/2} = O(n^{-2}),$$

which implies that all the intervals are third-order correct compared with the true nominal α-level interval $[\hat{y}_{1-\xi-t/2}, \hat{y}_{\xi+t/2}]$, using terminology given in Hall [(1992), Section 3.4].

To conclude this section, we remark that the $O(n^{-2})$ coverage errors of the asymptotic intervals, as given by Propositions 2 and 3, can be reduced further by explicit Edgeworth correction. Hall (1983) gives an account of this approach. However, complexity of the functions λ and λ_0 makes Edgeworth correction a formidable task. Moreover, the gain from such correction is very

much asymptotic and dubious for small samples. Also, we know from Hall [(1992), Section 3.8] that a two-sided normal approximation interval corrected for skewness also has coverage error of order $O(n^{-2})$. It is interesting to compare such an Edgeworth corrected interval with our asymptotic iterated intervals. Heuristically, the Edgeworth corrected interval tries to correct an already poor interval, such as one obtained from normal approximation, whereas our asymptotic iterated intervals approximate the correct α-level interval $[\hat{y}_{1-\xi-t/2}, \hat{y}_{\xi+t/2}]$ and should therefore result in more accurate coverages. We shall return to this point in the simulation study of Section 5.

4. Practical evaluation of asymptotic iterated intervals. In this section we give a user-oriented qualitative description of an automatic procedure, which we have programmed in Fortran and which takes care of all the lengthy algebra and calculations necessary for obtaining the endpoints of I_1 and I_2 for a general parameter. Special attention is paid to a built-in automatic derivative evaluation routine for computing the exact numerical values of any required derivatives, and by use of which the need for execution of symbolic operations is avoided. No separate software such as *Mathematica* is needed. To set the program running, the user needs only to enter a few trivial parameter values and the formula of the function $g(\boldsymbol{\mu})$ in a comprehensible coded form. The detailed calculations involved in the automatic procedure are described in the Appendix. The Fortran programs, which also simulate coverage probabilities under any predetermined underlying distributions, are available by anonymous ftp. Details of how to obtain the programs are given at the end of this section.

We first set up some notation. Suppose G is an r-variate distribution and $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_r)^T$ is a random vector from G. The parameter of interest $\theta = g(\boldsymbol{\mu})$ is assumed to depend only on moments of G up to order s. For any positive integer m, we can, by means of a recursive algorithm, list all the distinct r-sequences

$$(i_1, i_2, \ldots, i_r)$$

such that $\sum_{j=1}^{r} i_j = m$, with the i_j being nonnegative integers. For instance, if $m = 2$ and $r = 3$, then we have a list

$$\{(2,0,0), (1,1,0), (1,0,1), (0,2,0), (0,1,1), (0,0,2)\}.$$

Define $d(r, m)$ to be the number of such r-sequences corresponding to order m, and define $\rho_i(r, m)$ to be the ith sequence in this list, for $i = 1, 2, \ldots, d(r, m)$. Define also $\mathbf{Y}^{(i_1, i_2, \ldots, i_r)}$ to be the product $Y_1^{i_1} Y_2^{i_2} \cdots Y_r^{i_r}$. We can then extend the random vector $\mathbf{Y}$ to some $\mathbf{X}$, say, so that

$$\mathbf{X} = \left(\mathbf{Y}^{\rho_1(r,1)}, \mathbf{Y}^{\rho_2(r,1)}, \ldots, \mathbf{Y}^{\rho_{d(r,1)}(r,1)}, \mathbf{Y}^{\rho_1(r,2)}, \ldots, \right.$$
$$\left. \mathbf{Y}^{\rho_{d(r,2)}(r,2)}, \ldots, \mathbf{Y}^{\rho_1(r,6s)}, \ldots, \mathbf{Y}^{\rho_{d(r,6s)}(r,6s)}\right)^T,$$

that is, $\mathbf{X}$ is a sequence of products $Y_1^{i_1} \cdots Y_r^{i_r}$ up to order $6s$. Let F be the distribution of $\mathbf{X}$ under G. Define

$$d = \sum_{m=1}^{s} d(r,m), \qquad d' = \sum_{m=1}^{2s} d(r,m), \qquad d'' = \sum_{m=1}^{3s} d(r,m)$$

and

$$d''' = \sum_{m=1}^{6s} d(r,m).$$

Also, set $\boldsymbol{\mu} = \mathbb{E}[\mathbf{X}]$ and extend the domain of g to $\mathbb{R}^{d'''}$ so that $g(\boldsymbol{\mu}) = \theta$. In fact, g depends on $\boldsymbol{\mu}$ only through its first d elements.

The preliminary Fortran program `prelim.f` computes automatically the values of d, d', d'', and d''', which are stored, respectively, in the variables `ND0`, `ND1`, `ND2` and `ND3`. The extended vector $\mathbf{X}$ is also printed in terms of the defining $\mathbf{Y}$ products. This is particularly useful for later reference when we feed into the main program the formula of $g(\boldsymbol{\mu})$ in terms of the $\mathbf{X}$ components. The only input required by `prelim.f` from the user are

(28) `NVAR` $= r$, `NORD` $= s$ and `NORD6` $= 6s$,

defined via the `PARAMETER` command line.

The generality of this setup lies in the fact that the parameter of interest $g(\boldsymbol{\mu})$ in a smooth function model usually has a closed-form algebraic expression in terms of moments of G up to a certain order. The program `prelim.f` orders the moments systematically to form the vector $\mathbf{X}$ and so eliminates any ambiguity in the definition of $g(\boldsymbol{\mu})$ via moments. The trivial input `NVAR`, `NORD` and `NORD6` also makes `prelim.f` extremely straightforward to implement.

Next follows the execution of the main program `asymp.f`. Its basic input includes the variables mentioned above, namely,

(29) `NVAR, NORD, NORD6, ND0, ND1, ND2, ND3,`

as well as the formula of $g(\boldsymbol{\mu})$, the parameter of interest, which is entered through the command

(30) `CALL EVAL(`g`(IX(1), IX(2), ..., IX(NDO)), VECTOR).`

Here the function g`(IX(1), ..., IX(NDO))` stands for a Fortran coded form of $g(\boldsymbol{\mu})$, where `IX` is a Fortran integer array to be identified with the first d components of $\mathbf{X}$. To specify $g(\cdot)$, we first break up its definition into a sequence of basic algebraic operations such as addition, subtraction, multiplication, division, logarithm and so on. To each basic operation corresponds a Fortran user-defined function: for example, the addition operation goes with the function `IADD`$(\cdot,\cdot)$, multiplication goes with `IMUL`$(\cdot,\cdot)$, logarithm goes with `ILOG`$(\cdot)$ and so on. These user-defined functions constitute a kind of routine library which enables us to define most common parameters with algebraic formulae. A fairly comprehensive library of algebraic operations is already attached to the program `asymp.f`. See the comments in the program

for more details about this library. Here we illustrate its use with the variance case in which $g(x_1, x_2) = x_2 - x_1^2$ so that $g(\mathbb{E}Y, \mathbb{E}Y^2)$ is the variance of the random variable Y. The corresponding Fortran instruction (30) is

```
CALL EVAL(ISUB(IX(2),IMUL(IX(1),IX(1))), VECTOR).
```

The subroutine `EVAL` evaluates the exact numerical values of $g(\mathbf{x})$ and its partial derivatives up to the third order, all evaluated at the sample moments, and stores the output in a Fortran real array `VECTOR`. The details of this automatic derivative evaluation procedure are discussed in Kagiwada, Kalaba, Rasakhoo and Spingarn [(1986), Chapters 1 and 2] and follow the approach described by Wexler (1988). We have generalised the procedure to deal with any user-input dimension of the domain of g and any partial derivatives up to the third order. The automatic differentiation done by `EVAL` makes any preliminary symbolic calculation unnecessary. The output, which is stored in `VECTOR`, is used in subsequent straightforward algebraic calculations in `asymp.f` to produce the endpoints of the asymptotic iterated intervals. The Appendix details step by step the calculations carried out by the program.

The program `asymp.f` is applicable to a variety of smooth function models and no problem-specific recoding is necessary. The programs `prelim.f` and `asymp.f` may be obtained by anonymous ftp from `ftp.statslab.cam.ac.uk`. Give `anonymous` as your user name and your full e-mail address as your password. Then type `cd pub/CI`: the two programs are fetched by typing `mget prelim.f asymp.f`.

5. Simulation study. A simulation study was conducted to investigate coverage properties of the asymptotic intervals I_1 and I_2. Three examples of smooth function models were studied: mean, variance and correlation coefficient. In all examples, I_1 and I_2 were compared with the percentile method confidence interval I_P. Because of the extremely intensive computation involved in its construction, the "full-blown" standard iterated confidence interval I_F was only included for comparison in the mean example and for small sample sizes in the variance example. A variety of other asymptotic confidence intervals were also studied alongside I_1 and I_2 in the three examples. These include the Edgeworth-inversion-based confidence interval I_W of Withers (1983), the two-sided Edgeworth-corrected confidence interval I_{EC}, the approximate bootstrap confidence (ABC) interval I_{ABC} and the asymptotic "short" confidence interval I_{SH}. See Hall [(1992), Section 3.8] and DiCiccio and Efron (1992) for details of I_{EC} and I_{ABC}, respectively. The interval I_{SH} was taken to be the asymptotic version of the "short" interval proposed by Hall (1988), namely,

$$I_{SH} = \Big[\hat{\theta} - n^{-1/2}\hat{\sigma}\big(z_\xi + n^{-1/2}\hat{a}_1 + n^{-1}\hat{a}_2\big),$$
$$\hat{\theta} - n^{-1/2}\hat{\sigma}\big(-z_\xi + n^{-1/2}\hat{a}_1 - n^{-1}\hat{a}_2\big)\Big],$$

where the endpoints are obtained by asymptotically expanding the theoretical "short" interval endpoints up to order $O(n^{-3/2})$. Finally, the accelerated bias-corrected confidence interval I_{BC_a}, which requires one level of bootstrap resampling, was also included together with I_{ABC} for interest. See Efron (1987) for details of its construction.

It should be remarked that all the confidence intervals thus far mentioned were constructed in such a way that they have coverage error of order $O(n^{-2})$, except for I_P and I_{SH}, which have coverage error of order $O(n^{-1})$.

Perhaps we should emphasize here that it is our main interest to demonstrate in the simulation study the practical feasibility and advantages of approximating the standard iterated bootstrap confidence interval by asymptotic expansions, while maintaining its desirable accuracy. We do not attempt in this paper to provide a thorough comparison involving all the existing competing confidence intervals of the asymptotic kind. However, our simulation results are encouraging in that they show the asymptotic intervals I_1 and I_2 to outperform many existing asymptotic forms of confidence interval.

We now describe in detail the setting of our simulation study. In all examples the coverage probabilities of the various confidence intervals were approximated from 1600 random samples, so that each figure has a standard error of approximately 0.01. Intervals I_1, I_P and I_{BC_a} were constructed using $B = 1000$ bootstrap resamples. The full-blown iterated interval I_F was constructed using $C = 1000$ inner level bootstrap resamples. The asymptotic intervals I_2, I_W, I_{EC}, I_{ABC} and I_{SH} generally require no resampling. However, in cases of erratic asymptotic endpoints where, for example, the lower limit exceeds the upper limit, certain kinds of bootstrap intervals were substituted for those erratic asymptotic intervals. To be specific, we replaced any erratic I_2 by I_1, any erratic I_W by I_P, and any erratic I_{ABC} by I_{BC_a}. For any erratic I_{EC} or I_{SH}, $B = 1000$ bootstrap resamples were drawn and the interval rectified to $[\hat{\theta}^*_{(1)}, \hat{\theta}^*_{(B)}]$. This adjustment is largely pragmatic, as the theoretical bootstrap interval is likely to undercover in this situation and the widest interval possible is given by the extreme pair of bootstrap quantities $\hat{\theta}^*_{(1)}$ and $\hat{\theta}^*_{(B)}$. Throughout the whole study, the nominal level α was taken to be 0.9.

EXAMPLE 1 (Mean). In the mean example, the parameter of interest θ and its estimate $\hat{\theta}$ were taken to be the population and sample means, respectively. For input to the automatic procedure described in Section 4 we have NVAR = NORD = 1. Our asymptotic iterated intervals I_1 and I_2 were compared with I_P, I_F, I_W, I_{ABC} and I_{BC_a}. Four different underlying distributions, normal, folded normal, negative exponential and lognormal, were studied using sample sizes $n = 15$ and 30. The results are tabulated and shown in Table 1. In the cases of I_2, I_W and I_{ABC}, the proportions of adjustments by bootstrap resampling required out of the 1600 simulations are given in parentheses following the corresponding estimated coverage probabilities.

It is found that the full-blown standard iterated interval I_F gives the most accurate coverage probabilities in all cases. The asymptotic intervals I_1 and

TABLE 1

Example 1—*estimated coverage probabilities for mean, based on* 1,600 *random samples of sizes* $n = 15$ *and* 30 *drawn from each of four different distributions. Intervals* I_1, I_P *and* I_{BC_a} *are based on* 1,000 *outer level bootstrap resamples. For asymptotic intervals* I_2, I_W *and* I_{ABC}, *bootstrap intervals* I_1, I_P *and* I_{BC_a} *are substituted respectively if their endpoints turn out to be faulty. The frequencies of such adjustments are given in parentheses following the corresponding estimated coverages. The full-blown iterated interval* I_F *is based on* 1,000 *outer and* 1,000 *inner level bootstrap resamples*

Interval:	I_1	I_2		I_P	I_F	I_W		I_{ABC}		I_{BC_a}
					Normal data $N(0, 1)$					
$n = 15$	0.891	0.894	(0.000)	0.860	0.897	0.875	(0.000)	0.862	(0.000)	0.859
$n = 30$	0.901	0.903	(0.000)	0.892	0.902	0.899	(0.000)	0.893	(0.000)	0.889
					Folded normal $\|N(0, 1)\|$					
$n = 15$	0.876	0.873	(0.000)	0.839	0.883	0.866	(0.000)	0.861	(0.000)	0.855
$n = 30$	0.885	0.881	(0.000)	0.869	0.888	0.879	(0.000)	0.875	(0.000)	0.882
					Negative exponential data exp(1)					
$n = 15$	0.862	0.867	(0.004)	0.819	0.874	0.834	(0.000)	0.826	(0.000)	0.824
$n = 30$	0.896	0.898	(0.000)	0.876	0.901	0.881	(0.000)	0.875	(0.000)	0.876
					Log normal data exp $(N(0, 1))$					
$n = 15$	0.813	0.811	(0.048)	0.765	0.829	0.780	(0.000)	0.788	(0.000)	0.783
$n = 30$	0.847	0.846	(0.016)	0.815	0.853	0.827	(0.000)	0.819	(0.000)	0.826

I_2 are excellent approximations to I_F. They are more accurate than I_P, I_{BC_a} and the other asymptotic intervals by a considerable margin. Also, their coverages are consistently smaller than those of I_F, although by a very small magnitude. No adjustments by bootstrap resampling due to erratic endpoints were recorded, with the exception of a few cases related to I_2.

EXAMPLE 2 (Variance). We consider next the variance example studied by Schenker (1985) and DiCiccio, Martin and Young (1992a). The parameter of interest θ is the population variance, and the estimate $\hat{\theta}$ is the (biased) sample variance. Here we set NVAR = 1 and NORD = 2 in our automatic procedure. Intervals chosen for study in parallel with I_1 and I_2 included I_P, I_F, I_W, I_{EC} and I_{SH}. Four different underlying distributions with various degrees of skewness and kurtosis were used: the standard normal $N(0, 1)$, with no skewness and no kurtosis; the folded normal $|N(0, 1)|$, with high skewness and low kurtosis; the double exponential of unit rate with no skewness and high kurtosis; and, finally, the lognormal, $\exp(N(0, 1))$, which has high skewness and high kurtosis. The variances are, respectively, 1, $1 - 2/\pi$, 2 and $e(e - 1)$. Three different sample sizes were taken: $n = 20$, 35, and 100, respectively. The full-blown iterated interval I_F was not constructed for $n = 100$ due to its immense computational demands in this case.

The simulation results are reported in Table 2. Again, for those asymptotic intervals which might need bootstrap adjustments, the proportions of such adjustments are given in parentheses. We observe that the full-blown inter-

TABLE 2

Example 2—estimated coverage probabilities for variance, based on 1600 *random samples of sizes* n = 20, 35 *and* 100 *drawn from each of four different distributions: intervals* I_1 *and* I_P *are based on* 1000 *outer level bootstrap resamples; for the other asymptotic intervals, outer level resampling is carried out only if the asymptotic endpoints turn out to be faulty; the frequency of such failures is given in parentheses following the estimated coverage; the full-blown iterated interval* I_F *is based on* 1000 *outer and* 1000 *inner level bootstrap resamples*

Interval	Coverage, n = 20	Coverage, n = 35	Coverage, n = 100
	Normal data $N(0, 1)$ (no skew, no kurtosis)		
I_1	0.833	0.854	0.883
I_2	0.832 (0.161)	0.853 (0.014)	0.884 (0.000)
I_P	0.727	0.793	0.857
I_F	0.848	0.859	—
I_W	0.804 (0.000)	0.843 (0.000)	0.883 (0.000)
I_{EC}	0.802 (0.000)	0.841 (0.000)	0.883 (0.000)
I_{SH}	0.821 (0.000)	0.853 (0.000)	0.887 (0.000)
	Folded normal data $\lvert N(0, 1)\rvert$ (high skew, low kurtosis)		
I_1	0.803	0.821	0.874
I_2	0.800 (0.285)	0.819 (0.101)	0.880 (0.003)
I_P	0.686	0.753	0.843
I_F	0.815	0.834	—
I_W	0.759 (0.000)	0.800 (0.000)	0.868 (0.000)
I_{EC}	0.779 (0.000)	0.810 (0.000)	0.874 (0.000)
I_{SH}	0.799 (0.000)	0.824 (0.000)	0.881 (0.000)
	Double exponential data $\frac{1}{2}\exp(-\lvert x\rvert)$ (no skew, high kurtosis)		
I_1	0.811	0.846	0.869
I_2	0.809 (0.304)	0.848 (0.118)	0.872 (0.013)
I_P	0.698	0.776	0.834
I_F	0.826	0.854	—
I_W	0.746 (0.000)	0.806 (0.000)	0.855 (0.000)
I_{EC}	0.771 (0.000)	0.832 (0.000)	0.868 (0.000)
I_{SH}	0.787 (0.000)	0.844 (0.000)	0.879 (0.000)
	Lognormal data $\exp\{N(0, 1)\}$ (high skew, high kurtosis)		
I_1	0.526	0.602	0.696
I_2	0.526 (0.533)	0.602 (0.393)	0.696 (0.216)
I_P	0.416	0.504	0.608
I_F	0.544	0.630	—
I_W	0.430 (0.000)	0.523 (0.000)	0.662 (0.000)
I_{EC}	0.485 (0.000)	0.559 (0.000)	0.659 (0.000)
I_{SH}	0.510 (0.000)	0.576 (0.000)	0.681 (0.000)

val I_F gives the most accurate coverage and is consistently undercovering. Our asymptotic intervals I_1 and I_2 offer close approximations to I_F. They have virtually the same coverage and generally outperform the other asymptotic intervals I_W, I_{EC} and I_{SH}, especially when n is small or the underlying distribution is of the lognormal type with high skewness and high kurtosis. The percentile method interval I_P is noticeably poor in coverage. It should be

noted that the coverage of I_{SH} is quite similar to that of I_1 or I_2 for large sample sizes, despite its having a coverage error of greater order asymptotically. This observation, of course, demonstrates the inadequacy of using asymptotic theory alone in an assessment of different confidence interval construction procedures. With the lognormal distribution, all seven confidence intervals display very poor coverage properties. However, it is in this case where the most pronounced difference between performances of the intervals can be observed: here the asymptotic iterated intervals I_1 and I_2 clearly approximate the full-blown I_F best.

Another point should be made about the proportions of adjustments by bootstrap resampling required by the intervals I_2, I_W, I_{EC} and I_{SH}. Asymptotic endpoint failures are reported only in the case of I_2 where the failure (or adjustment) proportions reflect remarkably the general coverage error of the bootstrap confidence intervals. A poor coverage generally results in a higher chance of getting a faulty pair of endpoints for I_2 and having to adjust by resampling, as pointed out in our discussion at the end of Section 3.

To appreciate the huge impact on coverage error made by different underlying distributions, we computed the theoretical leading terms of the expansions of the calibrating coefficient t and coverage error, for the theoretical iterated bootstrap confidence interval I_0. These are, respectively, $2\pi_1(z_\xi)\phi(z_\xi)n^{-1}$ for t, as given by (4), and $2\lambda(\xi)\phi(z_\xi)n^{-2}$ for the coverage error, as given by Proposition 1. The results are listed in Table 3. We can easily understand why the lognormal distribution yields relatively huge coverage error and why the asymptotics are hardly revealing in this case.

EXAMPLE 3 (Correlation coefficient). We repeated the simulation study conducted by DiCiccio, Martin and Young (1992b) concerning the construction of confidence intervals for a correlation coefficient. Our parameter of interest θ and its corresponding estimate $\hat{\theta}$ were taken to be the population and sample correlation coefficients, respectively. The input parameters for the automatic procedure were therefore NVAR = NORD = 2. Intervals I_1, I_2, I_P, I_{EC} and I_{SH} were constructed in this example. The much more computationally intensive full-blown interval I_F was not examined here. Nevertheless,

TABLE 3

Example 2—theoretical leading terms in asymptotic expansions of calibrating coefficient and coverage error corresponding to the standard iterated bootstrap confidence interval I_0 under distributions studied in the variance example

True distribution	Calibrating coefficient t	Coverage error $\mathbb{P}(\theta \in I_0) - \alpha$
Standard normal, $N(0, 1)$	$3.109\, n^{-1}$	$-1.499 \times 10^2\, n^{-2}$
Folded normal, $\lvert N(0, 1)\rvert$	$6.498\, n^{-1}$	$-1.370 \times 10^3\, n^{-2}$
Double exponential, $\exp(-\lvert x\rvert)/2$	$1.206 \times 10\, n^{-1}$	$-1.240 \times 10^4\, n^{-2}$
Lognormal, $\exp(N(0, 1))$	$1.411 \times 10^6\, n^{-1}$	$-2.488 \times 10^{20}\, n^{-2}$

our conclusions are not compromised very much by such omission, as we observe coverage results for I_1 and I_2 which, compared to the other examples, are relatively accurate, and the full-blown figures might therefore reasonably be expected to be close to those of the asymptotic intervals in this example. Six different underlying distributions were combined with four sample sizes $n = 15$, 20, 30 and 50 in the complete simulation study.

Let W_1, W_2 and W_3 be independent standard normal variates. The underlying distributions were taken to be, respectively, as follows:

1. $Y = |W_1|$ and $Z = |W_2|$ so that (Y, Z) are independent folded normal variates;
2. $Y = |W_1| + |W_3|$ and $Z = |W_2| + |W_3|$ so that (Y, Z) are correlated folded normal variates with $\rho = 0.5$;
3. $Y = W_1$ and $Z = W_2$ so that (Y, Z) are independent normal variates;
4. $Y = (W_1 + \sqrt{3}\,W_2)/2$ and $Z = (-W_1 + \sqrt{3}\,W_2)/2$ so that (Y, Z) are correlated normal variates with $\rho = 0.5$;
5. $Y = \exp(W_1)$ and $Z = \exp(W_2)$ so that (Y, Z) are independent lognormal variates;
6. $Y = \exp\{(W_1 + W_3)/\sqrt{2}\}$ and $Z = \exp\{(W_2 + W_3)/\sqrt{2}\}$ so that (Y, Z) are correlated lognormal variates with $\rho = (e^{3/2} - e)/(e^2 - e)$.

Table 4 reports the complete simulation results, with proportions of adjustments by bootstrap resampling given in parentheses. It is clear that our asymptotic iterated intervals I_1 and I_2 outperform the other three methods by a considerable margin. The coverage of I_1 slightly dominates that of I_2 in all but a few rare cases. Also, our asymptotic iterated intervals result in coverage rather stable against the various sample sizes considered in the study. The percentile interval I_P performs moderately well, quite unlike the poor coverage revealed in the variance example. Its dependence on sample size is only token. The Edgeworth corrected interval I_{EC} and asymptotic short interval I_{SH} turn out to be the poorest in terms of coverage error, especially for lognormal data or for small to moderate sample sizes such as $n = 15$ or 20. Their coverage depends more strongly on sample size and thus enables them to regain some advantage over I_P in a few cases when $n = 30$ and 50. Again, the asymptotically more accurate interval I_{EC} is, contrary to expectation, always outperformed by the interval I_{SH}, which is asymptotically much inferior.

As in the variance example, the failure proportions of intervals I_2, I_{EC} and I_{SH} have an apparent diagnostic quality concerning coverage accuracy.

6. Comments. The simulation study reported above suggests that, in terms of coverage accuracy, the asymptotic iterated intervals I_1 and I_2 approximate the standard iterated bootstrap confidence interval excellently and are more effective than the percentile method and many other existing asymptotic methods.

We note also that the improvement in coverage accuracy made by I_1 and I_2 over the percentile method is similar to that which results from the analytic

TABLE 4

Example 3—estimated coverage probabilities for correlation coefficient ρ, based on 1600 random samples of sizes $n = 15, 20, 30$ and 50 drawn from each of six different distributions: intervals I_1 and I_P are based on 1000 outer level bootstrap resamples; for the asymptotic intervals, outer level resampling is used only in case the endpoints turn out to be faulty; the proportion of using outer level resampling is given in parentheses following the estimated coverage probability in this case

Interval	$n = 15$	$n = 20$	$n = 30$	$n = 50$
		Folded normal data, $\rho = 0$		
I_1	0.889	0.882	0.892	0.893
I_2	0.877 (0.0000)	0.879 (0.0000)	0.893 (0.0000)	0.893 (0.0000)
I_P	0.866	0.857	0.872	0.877
I_{EC}	0.841 (0.0006)	0.844 (0.0000)	0.866 (0.0000)	0.878 (0.0000)
I_{SH}	0.850 (0.0006)	0.855 (0.0000)	0.874 (0.0000)	0.883 (0.0000)
		Folded normal data, $\rho = 0.5$		
I_1	0.874	0.873	0.890	0.895
I_2	0.864 (0.0006)	0.868 (0.0000)	0.887 (0.0000)	0.897 (0.0000)
I_P	0.841	0.840	0.868	0.872
I_{EC}	0.810 (0.0000)	0.824 (0.0000)	0.856 (0.0000)	0.883 (0.0000)
I_{SH}	0.820 (0.0000)	0.833 (0.0000)	0.867 (0.0000)	0.888 (0.0000)
		Normal data, $\rho = 0$		
I_1	0.901	0.895	0.901	0.896
I_2	0.892 (0.0000)	0.893 (0.0000)	0.897 (0.0000)	0.893 (0.0000)
I_P	0.873	0.873	0.875	0.874
I_{EC}	0.849 (0.0000)	0.856 (0.0000)	0.878 (0.0000)	0.889 (0.0000)
I_{SH}	0.863 (0.0000)	0.871 (0.0000)	0.888 (0.0000)	0.894 (0.0000)
		Normal data, $\rho = 0.5$		
I_1	0.884	0.889	0.904	0.898
I_2	0.871 (0.0006)	0.887 (0.0000)	0.899 (0.0000)	0.894 (0.0000)
I_P	0.853	0.869	0.883	0.887
I_{EC}	0.824 (0.0000)	0.858 (0.0000)	0.876 (0.0000)	0.887 (0.0000)
I_{SH}	0.839 (0.0000)	0.865 (0.0000)	0.895 (0.0000)	0.893 (0.0000)
		Lognormal data, $\rho = 0$		
I_1	0.873	0.863	0.862	0.859
I_2	0.836 (0.0031)	0.835 (0.0019)	0.846 (0.0025)	0.849 (0.0006)
I_P	0.851	0.839	0.841	0.845
I_{EC}	0.705 (0.0013)	0.723 (0.0006)	0.750 (0.0006)	0.786 (0.0000)
I_{SH}	0.706 (0.0019)	0.728 (0.0006)	0.760 (0.0006)	0.789 (0.0000)
		Lognormal data, $\rho = 0.377541$		
I_1	0.863	0.849	0.852	0.857
I_2	0.834 (0.0081)	0.821 (0.0069)	0.843 (0.0069)	0.856 (0.0044)
I_P	0.815	0.806	0.812	0.819
I_{EC}	0.689 (0.0050)	0.718 (0.0050)	0.748 (0.0019)	0.779 (0.0019)
I_{SH}	0.697 (0.0050)	0.723 (0.0056)	0.753 (0.0025)	0.791 (0.0025)

approximations proposed by DiCiccio, Martin and Young (1992a, b). Moreover, the problem of overcovering by analytic approximations in the correlation coefficient example reported by DiCiccio, Martin and Young (1992b) does not exist for our asymptotic iterated methods.

The relative performances of I_1 and I_2 depend on the particular parameter of interest. Since it is the computationally more efficient method, the interval I_2 is very attractive when both I_1 and I_2 have similar coverage error. The need for bootstrap resampling typically renders construction of I_1 slower than that of I_2 by a factor of around 100 on an HP9000 workstation, depending on the number of bootstrap samples drawn. Nevertheless, both intervals I_1 and I_2 are far quicker to compute than the standard full-blown iterated bootstrap confidence interval, which requires two levels of bootstrap resampling. For example, in the case of constructing a confidence interval for the mean based on a sample of size $n = 15$, an HP9000 workstation took about 0.0012 second to compute I_2 using the automatic package described in Section 4, and about 155 seconds to compute I_F, the full-blown iterated confidence interval based on 1000 outer and 1000 inner level bootstrap resamples. Here the automatic package outruns the double bootstrap procedure by a factor of almost 130,000. Of course, the above comparison might have exaggerated the general computational savings offered by the automatic procedure, given that the population mean is a particularly simple parameter of interest. Nevertheless, the computational savings may still be expected in most problems to be as dramatic as a factor of thousands. Moreover, the automatic procedure has been packaged by means of a single Fortran program which takes care of all analytic calculations. The simple input requirements of the program makes construction of the asymptotic intervals practically feasible for general use.

It should also be remarked that the built-in exact derivative evaluation routine utilised in our automatic procedure for construction of asymptotic iterated percentile bootstrap confidence intervals can in fact be incorporated into any asymptotic confidence interval construction procedure, especially those based on the smooth function model. Its numerical yet exact nature makes any symbolic calculation or numerical approximations unnecessary. It thus opens up the possibility of packaging many existing confidence interval construction methods for general application, in a way that demands no analytic calculation from the user.

Finally, as observed from our simulation results, the interval I_2 is more likely to be erratic (and to require adjustment by the method described above) if the coverage of common bootstrap intervals fails to approximate the nominal level accurately. In this case, the real fault is not due to the asymptotic nature of I_2, but rather to the general failure of bootstrap intervals. Therefore, the use of bootstrap resampling to substitute for the asymptotic endpoints does little to rectify the coverage error. Nevertheless, such need for resampling adjustment may be seen as a sign of bootstrap inadequacy and is therefore of some diagnostic value. Further research in

this direction may be beneficial to our general understanding of bootstrap methods as applied to confidence interval construction.

APPENDIX

Detailed calculations for constructing I_1 and I_2. We now give a complete account of calculations coded in the program `asymp.f`. The notation follows that introduced in Section 4.

Consider first the infinite list $\mathscr{L}$ of sequences of the $\rho_i(r, j)$,

$$\mathscr{L} = \{\rho_1(r,1), \rho_2(r,1), \ldots, \rho_{d(r,1)}(r,1), \rho_1(r,2), \ldots,$$
$$\rho_{d(r,2)}(r,2), \ldots, \rho_1(r,m), \ldots, \rho_{d(r,m)}(r,m), \ldots\}.$$

Let $\rho_i(r)$ denote the ith sequence in $\mathscr{L}$. Define an operation $\otimes$ on $\mathbb{N} \times \mathbb{N}$ such that, for any $i, j \in \mathbb{N}$, we have

$$i \otimes j = k \quad \text{if } \rho_i(r) + \rho_j(r) \text{ is the } k\text{th sequence in } \mathscr{L}.$$

Therefore, we have

$$\mathbf{X}^{(i)}\mathbf{X}^{(j)} = \mathbf{X}^{(i \otimes j)},$$

provided $i \otimes j$ does not exceed the dimension of $\mathbf{X}$. To perform this operation we set up an initial list

$$\mathscr{L}_0 = \{\rho_1(r,1), \rho_2(r,1), \ldots, \rho_{d(r,1)}(r,1), \rho_1(r,2), \ldots,$$
$$\rho_{d(r,2)}(r,2), \ldots, \rho_1(r,6s), \ldots, \rho_{d(r,6s)}(r,6s)\}$$

as the basis underlying the operation. An order of $6s$ is adequate for the entire range of moments likely to be encountered.

Suppose now we observe a random sample $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ from G. Let $\mathbf{X}_i$ denote the extended vector $\mathbf{X}$ corresponding to $\mathbf{Y}_i$. Our automatic procedure operates as follows:

1. Calculate

$$\overline{\mathbf{X}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i.$$

2. Partially differentiate $g(\mathbf{x})$ up to the third order. Define

$$g_i(\mathbf{x}) = \frac{\partial g(\mathbf{x})}{\partial x_i}, \qquad g_{ij}(\mathbf{x}) = \frac{\partial^2 g(\mathbf{x})}{\partial x_i\, \partial x_j} \quad \text{and} \quad g_{ijk}(\mathbf{x}) = \frac{\partial^3 g(\mathbf{x})}{\partial x_i\, \partial x_j\, \partial x_k},$$

   for $i, j, k = 1, 2, \ldots, d'$.
3. Evaluate the sample values of $g(\overline{\mathbf{X}})$, $g_i(\overline{\mathbf{X}})$, $g_{ij}(\overline{\mathbf{X}})$ and $g_{ijk}(\overline{\mathbf{X}})$.
4. For $\mathbf{x} \in \mathbb{R}^{d^m}$, define the following functions:

$$\mu_{ij}(\mathbf{x}) = x_{i \otimes j} - x_i x_j,$$

for $i = 1, 2, \ldots, d'$ and $j = 1, 2, \ldots, d''$;

$$\mu_{ijk}(\mathbf{x}) = x_{i\otimes j\otimes k} - x_{i\otimes j}x_k - x_{j\otimes k}x_i - x_{k\otimes i}x_j + 2x_ix_jx_k,$$

for $i, j, k = 1, 2, \ldots, d'$; and

$$\begin{aligned}\mu_{ijkl}(\mathbf{x}) = {} & x_{i\otimes j\otimes k\otimes l} - x_{i\otimes j\otimes k}x_l - x_{j\otimes k\otimes l}x_i - x_{k\otimes l\otimes i}x_j - x_{l\otimes i\otimes j}x_k \\ & + x_{i\otimes j}x_kx_l + x_{i\otimes k}x_lx_j + x_{i\otimes l}x_jx_k + x_{j\otimes k}x_lx_i + x_{j\otimes l}x_ix_k \\ & + x_{k\otimes l}x_ix_j - 3x_ix_jx_kx_l,\end{aligned}$$

for $i, j, k, l = 1, 2, \ldots, d$.

5. Partially differentiate μ_{ij} and μ_{ijk} to get the following:

$$\partial_l \mu_{ij}(\mathbf{x}) = \frac{\partial \mu_{ij}(\mathbf{x})}{\partial x_l},$$

for $i, j = 1, 2, \ldots, d$ and $l = 1, 2, \ldots, d'$; and

$$\partial_l \mu_{ijk}(\mathbf{x}) = \frac{\partial \mu_{ijk}(\mathbf{x})}{\partial x_l},$$

for $i, j, k = 1, 2, \ldots, d$ and $l = 1, 2, \ldots, d''$.

6. Evaluate the sample values of all the functions defined in steps 4 and 5, that is, $\mu_{ij}(\overline{\mathbf{X}}), \ldots$ and so on.

To simplify notation, we now drop the bracketed arguments of the sample values of functions so far evaluated, by setting, for example, $g = g(\overline{\mathbf{X}})$, $g_i = g_i(\overline{\mathbf{X}}), \ldots$.

7. Evaluate

$$h = \left\{ \sum_{i,j=1}^{d} g_i g_j \mu_{ij} \right\}^{1/2}.$$

8. Evaluate

$$a_i = \frac{g_i}{h}, \qquad a_{ij} = \frac{g_{ij}}{h} \quad \text{and} \quad a_{ijk} = \frac{g_{ijk}}{h},$$

for $i, j, k = 1, 2, \ldots, d'''$.

9. Evaluate

$$h_k = (2h)^{-1} \sum_{i,j=1}^{d} \{2g_{ik}g_j\mu_{ij} + g_ig_j\,\partial_k\mu_{ij}\}$$

and

$$\begin{aligned}h_{kl} = {} & h^{-1} \sum_{i,j=1}^{d} \{g_{ikl}g_j\mu_{ij} + g_{ik}g_{jl}\mu_{ij} + g_{ik}g_j\,\partial_l\mu_{ij} + g_{il}g_j\,\partial_k\mu_{ij}\} \\ & - h^{-1}\{h_kh_l + g_kg_l\},\end{aligned}$$

for $k, l = 1, 2, \ldots, d'$.

10. Evaluate

$$b_i = h^{-1}g_i,$$

for $i = 1, 2, \ldots, d$,

$$b_{ij} = h^{-1}g_{ij} - h^{-2}(g_i h_j + g_j h_i),$$

and

$$b_{ijk} = h^{-1}g_{ijk} - h^{-2}(g_{ij}h_k + g_{ik}h_j + g_{jk}h_i + g_i h_{jk} + g_j h_{ik} + g_k h_{ij}) + 2h^{-3}(g_i h_j h_k + g_j h_k h_i + g_k h_i h_j),$$

for $i, j, k = 1, 2, \ldots, d'$.

11. Evaluate

$$l_{12} = \tfrac{1}{2}\sum_{i,j=1}^{d} a_{ij}\mu_{ij},$$

$$l_{31} = \sum_{i,j,k=1}^{d} a_i a_j a_k \mu_{ijk} + 3\sum_{i,j,k,l=1}^{d} a_i a_j a_{kl}\mu_{ik}\mu_{jl},$$

$$l_{22} = \sum_{i,j,k,l=1}^{d}\left\{\tfrac{1}{2}a_{ij}a_{kl}\mu_{ik}\mu_{jl} + a_i a_{jkl}\mu_{ij}\mu_{kl}\right\} + \sum_{i,j,k=1}^{d} a_i a_{jk}\mu_{ijk}$$

and

$$l_{41} = \sum_{i,j,k,l=1}^{d} a_i a_j a_k a_l \mu_{ijkl} + 12\sum_{i,j,k,l,m=1}^{d} a_i a_j a_k a_{lm}\mu_{ijl}\mu_{km} + \sum_{i,j,k,l,m,n=1}^{d}\{4a_i a_j a_k a_{lmn}\mu_{il}\mu_{jm}\mu_{kn} + 12a_i a_j a_{kl}a_{mn}\mu_{ik}\mu_{jm}\mu_{ln}\} - 3.$$

12. Similarly, evaluate

$$k_{12} = \tfrac{1}{2}\sum_{i,j=1}^{d'} b_{ij}\mu_{ij},$$

$$k_{31} = \sum_{i,j,k=1}^{d} b_i b_j b_k \mu_{ijk} + 3\sum_{i,j=1}^{d}\sum_{k,l=1}^{d'} b_i b_j b_{kl}\mu_{ik}\mu_{jl},$$

$$k_{22} = \tfrac{1}{2}\sum_{i,j,k,l=1}^{d'} b_{ij}b_{kl}\mu_{ik}\mu_{jl} + \sum_{i=1}^{d}\sum_{j,k,l=1}^{d'} b_i b_{jkl}\mu_{ij}\mu_{kl} + \sum_{i=1}^{d}\sum_{j,k=1}^{d'} b_i b_{jk}\mu_{ijk}$$

and

$$k_{41} = \sum_{i,j,k,l=1}^{d} b_i b_j b_k b_l \mu_{ijkl} + 12 \sum_{i,j,k=1}^{d} \sum_{l,m=1}^{d'} b_i b_j b_k b_{lm} \mu_{ijl} \mu_{km} + 4 \sum_{i,j,k=1}^{d} \sum_{l,m,n=1}^{d'} b_i b_j b_k b_{lmn} \mu_{il} \mu_{jm} \mu_{kn} + 12 \sum_{i,j=1}^{d} \sum_{k,l,m,n=1}^{d'} b_i b_j b_{kl} b_{mn} \mu_{ik} \mu_{jm} \mu_{ln} - 3.$$

13. Evaluate

$$\partial_m l_{12} = \frac{1}{2} \sum_{i,j=1}^{d} (a_{ijm} \mu_{ij} + a_{ij} \partial_m \mu_{ij}) - \frac{l_{12} h_m}{h},$$

for $m = 1, 2, \ldots, d'$, and

$$\partial_m l_{31} = \sum_{i,j,k,l=1}^{d} \{6 a_{im} a_j a_{kl} \mu_{ik} \mu_{jl} + 3 a_i a_j a_{klm} \mu_{ik} \mu_{jl} + 6 a_i a_j a_{kl} \mu_{jl} \partial_m \mu_{ik}\} + \sum_{i,j,k=1}^{d} \{3 a_{im} a_j a_k \mu_{ijk} + a_i a_j a_k \partial_m \mu_{ijk}\} - \frac{3 l_{31} h_m}{h},$$

for $m = 1, 2, \ldots, d''$.

14. Set $\xi = \frac{1}{2}(1 + \alpha)$. Evaluate $z_\xi = \Phi^{-1}(\xi)$.

15. Evaluate

$$a_\xi = \sum_{i=1}^{d} \sum_{j=1}^{d'} \mu_{ij} b_i \partial_j l_{12} + \tfrac{1}{6}(z_\xi^2 - 1) \sum_{i=1}^{d} \sum_{j=1}^{d''} \mu_{ij} b_i \partial_j l_{31}.$$

16. Evaluate

$$p_1(z_\xi) = -\left\{ l_{12} + \frac{l_{31}(z_\xi^2 - 1)}{6} \right\},$$

$$p_2(z_\xi) = -z_\xi \left\{ \frac{l_{12}^2 + l_{22}}{2} + \frac{(4 l_{12} l_{31} + l_{41})(z_\xi^2 - 3)}{24} + \frac{l_{31}^2 (z_\xi^4 - 10 z_\xi^2 + 15)}{72} \right\},$$

$$p_1'(z_\xi) = -\frac{l_{31} z_\xi}{3}$$

and

$$p_2'(z_\xi) = -\left\{ \frac{l_{12}^2 + l_{22}}{2} + \frac{(4 l_{12} l_{31} + l_{41})(3 z_\xi^2 - 3)}{24} + \frac{l_{31}^2 (5 z_\xi^4 - 30 z_\xi^2 + 15)}{72} \right\}.$$

Evaluate also $q_1(z_\xi)$, $q_2(z_\xi)$, $q_1'(z_\xi)$ and $q_2'(z_\xi)$ by replacing l_{ij} with k_{ij} in the above formulae for the p_i and p_i'.

17. Evaluate

$$\pi_1(z_\xi) = p_2(z_\xi) - q_2(z_\xi) - p_1(z_\xi)\{p_1'(z_\xi) - z_\xi p_1(z_\xi) + q_1'(z_\xi) - z_\xi q_1(z_\xi)\} + a_\xi z_\xi$$

and

$$\tilde{t} = n^{-1} 2\pi_1(z_\xi)\phi(z_\xi).$$

If $0 < \xi + \tilde{t}/2 < 1$, then evaluate

$$\tilde{y}_\beta = g + n^{-1/2}h\left\{z_\beta - n^{-1/2}p_1(z_\beta) + n^{-1}\left(\frac{p_1(z_\beta)p_1'(z_\beta) - p_2(z_\beta) - z_\beta p_1^2(z_\beta)}{2}\right)\right\}$$

for $\beta = \xi + \tilde{t}/2$ and $1 - \xi - \tilde{t}/2$, respectively, where the functions p_1, p_2 and p_1' are recalculated using formulae in the last step.

Note that the notation used in the above description follows that given in Section 3, except that now all quantities represent sample versions. Only steps 2 and 5 require anything other than arithmetic operations: these can be most conveniently handled by an exact derivative evaluation routine, as described in Section 4 above and Kagiwada, Kalaba, Rasakhoo and Spingarn [(1986), Chapters 1 and 2]. This routine evaluates derivatives exactly and numerically, thus saving a great deal of analytic effort and/or computer time spent on symbolic operations. Note that the term $\mu_{d'd'd'}(\mathbf{x})$, defined in step 4 and evaluated subsequently in step 6, depends on a moment of order $6s$, the highest throughout the whole procedure. This observation explains our choice of the maximum dimension d'''.

As pointed out in step 17, it may happen that $\xi + \tilde{t}/2 \notin (0, 1)$ so that $z_{\xi+\tilde{t}/2}$ is infinite and the formula for $\tilde{y}_{\xi+\tilde{t}/2}$ is not defined. Also, if $\tilde{y}_{1-\xi-\tilde{t}/2} \geq \tilde{y}_{\xi+\tilde{t}/2}$, the interval I_2 would be empty and meaningless. In these situations, we suggest obtaining an approximation to I_1 by means of bootstrap resampling. The calibrated coverage level $\alpha + \tilde{t}$ should also be forced to its nearest point within the interval $[0, 1]$. To be more specific, we draw B random resamples $\chi_1^*, \ldots, \chi_B^*$ from χ, calculate the version of $\hat{\theta}$, $\hat{\theta}_b^*$ say, based on each χ_b^*, order these quantities as $\hat{\theta}_{(1)}^* \leq \cdots \leq \hat{\theta}_{(B)}^*$, and then approximate I_1 by

$$\left[\hat{\theta}_{(m_1)}^*, \hat{\theta}_{(m_2)}^*\right],$$

where $m_1 = \max\{1, [(B + 1)(1 - \xi')]\}$, $m_2 = \min\{B, [(B + 1)\xi']\}$, $[\cdot]$ denotes the integer part, and $\xi' = \max\{1/2, \xi + \tilde{t}/2\}$. Such choices of m_1, m_2 and ξ' ensure that $1 \leq m_1 \leq m_2 \leq B$.

REFERENCES

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

BOOTH, J. G. and DO, K.-A. (1993). Simple and efficient methods for constructing bootstrap confidence intervals. *Computational Statistics* **8** 333–346.

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

DICICCIO, T. J. and EFRON, B. (1992). More accurate confidence intervals in exponential families. *Biometrika* **79** 231–245.

DICICCIO, T. J., MARTIN, M. A. and YOUNG, G. A. (1992a). Analytical approximations for iterated bootstrap confidence intervals. *Statistics and Computing* **2** 161–171.

DICICCIO, T. J., MARTIN, M. A. and YOUNG, G. A. (1992b). Fast and accurate approximate double bootstrap confidence intervals. *Biometrika* **79** 285–295.

EFRON, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*. SIAM, Philadelphia.

EFRON, B. (1987). Better bootstrap confidence intervals (with discussion). *J. Amer. Statist. Assoc.* **82** 171–200.

HALL, P. (1983). Inverting an Edgeworth expansion. *Ann. Statist.* **11** 569–576.

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

HALL, P. (1988). Theoretical comparison of bootstrap confidence intervals (with discussion). *Ann. Statist.* **16** 927–985.

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

HALL, P. and MARTIN, M. A. (1988). On bootstrap resampling and iteration. *Biometrika* **75** 661–671.

KAGIWADA, H., KALABA, R., RASAKHOO, N. and SPINGARN, K. (1986). *Numerical Derivatives and Nonlinear Analysis*. Plenum, New York.

LEE, S. M. S. and YOUNG, G. A. (1993). Sequential iterated bootstrap confidence intervals Research Report 93-17, Statistical Laboratory, Univ. Cambridge.

LOH, W.-Y. (1987). Calibrating confidence coefficients. *J. Amer. Statist. Assoc.* **82** 155–162.

MARTIN, M. A. (1990). On bootstrap iteration for coverage correction in confidence intervals. *J. Amer. Statist. Assoc.* **85** 1105–1118.

SCHENKER, N. (1985). Qualms about bootstrap confidence intervals. *J. Amer. Statist. Assoc.* **80** 360–361.

WEXLER, A. S. (1988). An algorithm for exact evaluation of multivariate functions and their derivatives to any order. *Comput. Statist. Data Anal.* **6** 1–6.

WITHERS, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals. *Ann. Statist.* **11** 577–587.

DEPARTMENT OF STATISTICS
UNIVERSITY OF HONG KONG
POKFULAM ROAD
HONG KONG

STATISTICAL LABORATORY
UNIVERSITY OF CAMBRIDGE
16 MILL LANE, CAMBRIDGE CB2 1SB
UNITED KINGDOM