Improving Confidence Intervals Obtained by Sectioning in Monte Carlo Studies

Alun Pope
Department of Statistics
University of Newcastle
NSW 2308 Australia
Email: pope@maths.newcastle.edu.au

In computer simulation studies, sectioning (also called batching) has been used to estimate variances, and hence to derive confidence intervals. In this paper we give some theory for improving the sectioning technique for constructing confidence intervals. The improvements are based on Edgeworth expansions and are analogous to techniques for improving the bootstrap by pivoting and iteration. An example illustrates the method.

1 INTRODUCTION

1.1 Background

Typically, computer simulation experiments are used to construct simple random samples from a distribution, in the following way: a stream of independent and identically distributed (*iid*) random numbers is generated and passed as input to a subprogram which computes a deterministic transformation of the stream, producing as output a stream of independent random numbers which also have the same distribution, different from that of the inputs. The important point for us here is that we assume the outputs are an *iid* random sample. We assume that the goal of the simulation is to measure a scalar characteristic of the distribution of the outputs.

When computer simulation is used to obtain a point estimate of a parameter in a model, it is important also to obtain estimates of the sampling variability of the estimate. It is often the case that the process being simulated is too complex for analytical or asymptotic calculations of sampling variances to be attempted, so alternatives must be found. Lewis and Orav (1989, Chap. 9) describe *sectioning*: compute estimates based on independent subsamples of the data and use the variability of these estimates to estimate the variability of the estimate computed from the entire sample.

If measures of variability such as confidence intervals are to be constructed, then this approach is based on an assumption that the empirical distribution of subsample estimates is sufficiently close to normal. (See Lewis and Orav, 1989, p259, for example.) Often, this assumption is demonstrably false (Lewis and Orav, 1989, Figure 9.3.3, p266) so there are limitations on its use in this form. In section 2, we give a description of sectioning in more detail, and outline some improvements.

1.2 Related techniques

The ideas of "typical values" and random subsampling are also present in the work of Hartigan and Carlstein (e.g. Hartigan, 1967 and Carlstein, 1986). In their work, the subsamples are selected from a given sample rather than from simulation. However, apart from the difference that subsamples are constructed from samples in the Hartigan and Carlstein work, whereas in computer simulation it is more likely to be the other way round, the relationships between samples and subsamples are very similar in the two circumstances.

It is clear that there are also similarities between subsampling and resampling techniques such as the bootstrap. The main motivation for what follows is this comparison. We have asked what effect the improvements such as pivoting and iteration, so useful for the bootstrap, can have on the performance of the subsampling techniques. The partial answers are the basis of this paper.

2 SECTIONING

2.1 The sectioning procedure

We give a terse description of the sectioning procedure. More details may be found in Lewis and Orav (1989, Chap. 9), where applications are also described.

Let $S_1, \ldots, S_B$ be independent simple random samples of size b from the same distribution, and suppose we wish to estimate the scalar characteristic θ of the distribution. Let $\hat{\theta}_k$ denote an estimator of θ based on a sample of size k. Let θ_0 be the true value of θ and let $n = Bb$ be the sample size. We shall refer to $S = S_1 \cup \ldots \cup S_B$ as the sample and to the S_i as the subsamples, sections or blocks.

Let $\hat{\theta}_b^{(i)}$ denote an estimate of θ based on the section S_i. The subscript b indicates the sample size. The sectioning heuristic is that the distribution of the $\hat{\theta}_b^{(i)}$ (as i varies) is an approximation to the sampling distribution of $\hat{\theta}_b$, and hence provides information about the sampling distribution of $\hat{\theta}_n$.

Note that the sample sizes for $\hat{\theta}_b$ and $\hat{\theta}_n$ are different, so some adjustment has to be made for this. In the examples in Lewis and Orav (1989), this adjustment is not made explicitly but instead indirectly through a normality assumption and estimation of the standard deviation. The examples given in that book also show that the method can fail to provide good confidence intervals if the normality approximation is not good.

Briefly, the method of Lewis and Orav is to compute the (sample) variance of the $\hat{\theta}_b^{(i)}$ and use this as an estimate of $\mathrm{var}(\hat{\theta}_b)$. Assuming that $\mathrm{var}(\hat{\theta}_n)$ is of precise order n^{-1}, they then inflate the estimate of $\mathrm{var}(\hat{\theta}_b)$ via

$$var(\hat{\theta}_n) = \left(\frac{b}{n}\right) var(\hat{\theta}_b).$$

Confidence intervals are then based on an asymptotic normality argument. A similar technique is discussed in Kleijnen (1988) where it is called batching. See also Schmeiser (1990).

2.2 Percentile confidence intervals from sectioning

If one takes the sectioning heuristic seriously, then, instead of the normal approximation, one might consider a confidence interval based on the observed distribution of the $\hat{\theta}_b^{(i)}$. It is necessary in this case also to adjust for the fact that the sample sizes b and n are different. In order to see how to do this, we first consider the Edgeworth expansion of the distribution of the estimator.

For any k, let σ^2 be the asymptotic variance (assumed finite) of $U_k = k^{\frac{1}{2}}(\hat{\theta}_k - \theta_0)$ and let F_k denote the distribution of U_k:

$$F_k(x) = \mathbb{P}\{U_k \le x\} \quad .$$

Then under sufficient regularity conditions (see Remark below), the Edgeworth expansion is valid:

$$F_k(x) = \Phi\left(\frac{x}{\sigma}\right) + k^{-\frac{1}{2}} p\left(\frac{x}{\sigma}\right) \phi\left(\frac{x}{\sigma}\right) + O(k^{-\frac{1}{2}}) \quad (1)$$

where Φ, ϕ denote respectively the *cdf*, *pdf* of a standard normal variate, and p is a polynomial independent of k.

Now we find easily that

$$\begin{aligned} & F_n(x) - \left(\frac{b}{n}\right)^{\frac{1}{2}} F_b(x) \\ = \; & \Phi\left(\frac{x}{\sigma}\right) - \left(\frac{b}{n}\right)^{\frac{1}{2}} \Phi\left(\frac{x}{\sigma}\right) + O\left(b^{-\frac{1}{2}} n^{-\frac{1}{2}}\right) \quad ,(2) \end{aligned}$$

where the term which is $O(n^{-1})$ has been dropped because $b \le n$, which makes $n^{-1} = O\left(b^{-\frac{1}{2}} n^{-\frac{1}{2}}\right)$.

Suppose we have available an estimate $\hat{\sigma}_n$ of σ which is accurate to precise order $O_p\left(n^{-\frac{1}{2}}\right)$:

$$\hat{\sigma}_n = \sigma + O_p\left(n^{-\frac{1}{2}}\right) \quad (3)$$

Estimate F_b in the natural way: by putting $\hat{F}_b(x) =$ proportion of section estimates $\le x$.

Since there are n/b sections, and the number of section estimates $\le x$ is a binomial variable, we have

$$\hat{F}_b(x) = F_b(x) + O_p\left\{\left(\frac{b}{n}\right)^{\frac{1}{2}}\right\}. \quad (4)$$

(Again the order statement is precise.)

Now put

$$\hat{F}_n(x) = \left(\frac{b}{n}\right)^{\frac{1}{2}} \hat{F}_b(x) + \left\{1 - \left(\frac{b}{n}\right)^{\frac{1}{2}}\right\} \Phi\left(\frac{x}{\hat{\sigma}_n}\right). \quad (5)$$

There are two sources of error in this as an estimator of $F_n(x)$: the error in estimating $F_b(x)$ by $\hat{F}_b(x)$ and the error in $\Phi(x/\hat{\sigma}_n)$ due to estimating σ. The former is obtained from (4); the latter is obtained by noting that it follows from (3) that

$$\Phi\left(\frac{x}{\hat{\sigma}_n}\right) = \Phi\left(\frac{x}{\sigma}\right) + O_p\left(n^{-\frac{1}{2}}\right). \qquad (6)$$

Comparing (5) with (6), we see that

$$\hat{F}_n(x) = F_n(x) + O_p\left(\frac{b}{n}\right) + O_p\left(n^{-\frac{1}{2}}\right). \qquad (7)$$

Let b be of precise order n^{β}. From (6), we see that, provided $0 \leq \beta \leq \frac{1}{2}$, the error is $O_p\left(n^{-\frac{1}{2}}\right)$, while if $\beta > \frac{1}{2}$, the error is $O_p\left(n^{\beta-1}\right)$.

We see that this method approximates to the same order of accuracy as the normal approximation, provided $0 \leq \beta \leq \frac{1}{2}$. If we use quantiles of $\hat{F}_n$ to construct confidence intervals then they will have the same order of coverage accuracy as the normal approximation, and the endpoints of the intervals will also have the same order of accuracy as those constructed via the normal approximation. We do not pursue this further because we are now in a position to see from our derivation of (6) how to improve our estimation procedure. The key to this improvement is to consider the empirical distribution of a pivot instead of U_k. We turn to this in the next subsection.

Remark 1 *We assume that all Edgeworth expansions that we write down make sense: in particular, they exist as asymptotic series, and have uniform error terms. For definiteness, the reader may safely assume that the estimator $\hat{\theta}_n$ is a "smooth function of means". Details of this idea are in (Hall, 1992, Chapter 2) but it is worth pointing out here that means, variances, covariances and correlations are all examples of such statistics. An important example not included in this formulation is the studentized quantile. (Compare Hall, 1992, Appendix IV.)*

2.3 Using a pivot

It is well known that for the bootstrap one way to achieve better accuracy in the estimation of confidence intervals by percentile methods is to apply these methods to a pivot instead of the generally non-pivotal statistic U_k in the last subsection. See Hall (1992) for a detailed exposition. We apply these ideas to the case of sectional estimates.

We define $T_k = U_k/\hat{\sigma}_k$. Then T_k is asymptotically a pivot (i.e. has an asymptotic distribution independent of the values of the parameters). We set

$$G_k(x) = \mathbb{P}\{T_k \leq x\}.$$

Then, under regularity conditions, we have

$$G_k(x) = \Phi(x) + k^{-\frac{1}{2}}p(x)\phi(x) + O(k^{-1}) \qquad (8)$$

where Φ, ϕ are as before, p is a polynomial independent of k. Now (2) is replaced by

$$G_n(x) - \left(\frac{b}{n}\right)^{\frac{1}{2}} G_b(x) \qquad (9)$$

$$= \left\{1 - \left(\frac{b}{n}\right)^{\frac{1}{2}}\right\}\Phi(x) + O\left(b^{-\frac{1}{2}}n^{-\frac{1}{2}}\right). \qquad (10)$$

We estimate G_b by $\hat{G}_b$: $\hat{G}_b(x) =$ proportion of sections in which $T_b \leq x$. As before,

$$\hat{G}_b(x) = G_b(x) + O_p\left\{\left(\frac{b}{n}\right)^{\frac{1}{2}}\right\} \qquad (11)$$

and so if we set

$$\hat{G}_n(x) = \left(\frac{b}{n}\right)^{\frac{1}{2}} \hat{G}_b(x) + \Phi(x)\left\{1 - \left(\frac{b}{n}\right)^{\frac{1}{2}}\right\}, \qquad (12)$$

then

$$\hat{G}_n(x) = G_n(x) + O_p\left(\frac{b}{n}\right). \qquad (13)$$

(The ignored order term in (8) is of smaller order.)

The important difference between (6) and (12) is that the $O_p(n^{-\frac{1}{2}})$ term is missing from (12). If we take b to be of precise order n^{β} then provided $0 \leq \beta \leq \frac{1}{2}$, the estimate (12) is better than that in (19). Indeed by taking β small we can obtain accuracy arbitrarily close ($\beta = 0$ gives same accuracy) to the estimate that is obtained via the percentile-t bootstrap, which has error $O_p(n^{-1})$. In the case of the bootstrap, expressions like (12) lead to higher order accuracy of confidence intervals. (See Hall, 1992, Chap. 3.) From (12) it is now easy to construct confidence intervals and to compute their coverage accuracies. We illustrate this in the next subsection for a particular case.

2.4 The percentile-t method

For $0 \leq \gamma \leq 1$, let $v_{\gamma}^{(n)}, v_{\gamma}^{(b)}$ denote the γ-quantiles of G_n, G_b, respectively. In this subsection we show how these two quantities are related, and how this makes it possible to see that the percentile-t confidence intervals have coverage properties better than those obtained by the normal approximation. The discussion is based on the account in Chapter 3 of Hall (1992), in which the theory of bootstrap confidence intervals is developed. Our example shows how the bootstrap theory is easily adapted to our setting.

Because of its simplicity, we consider first the one-sided level-α confidence interval:

$$I_{\alpha} = (-\infty, \hat{\theta}_n - n^{-\frac{1}{2}}\hat{\sigma}_n v_{1-\alpha}^{(n)}). \qquad (14)$$

Since

$$1 - \alpha = G_n(v_{1-\alpha}^{(n)})$$

$$= \mathbb{P}\left\{T_n \le v^{(n)}_{1-\alpha}\right\}$$
$$= \mathbb{P}\left\{n^{\frac{1}{2}}(\hat\theta_n - \theta_0)/\hat\sigma_n \le v^{(n)}_{1-\alpha}\right\}$$
$$= \mathbb{P}\left\{\theta_0 \ge \hat\theta_n - n^{-\frac{1}{2}}\hat\sigma_n v^{(n)}_{1-\alpha}\right\},$$

we see I_α has coverage probability α as required.

The difficulty with I_α is of course that we do not know the value of $v^{(n)}_{1-\alpha}$. We consider now the question of estimating it.

Cornish-Fisher expansions, inverting the Edgeworth expansions (7), lead to

$$v^{(n)}_{1-\alpha} = z_{1-\alpha} - n^{-\frac{1}{2}}p(z_{1-\alpha})\phi(z_{1-\alpha}) + O(n^{-1}). \quad (15)$$

and

$$v^{(b)}_{1-\alpha} = z_{1-\alpha} - b^{-\frac{1}{2}}p(z_{1-\alpha})\phi(z_{1-\alpha}) + O(b^{-1}). \quad (16)$$

Here p is the same polynomial as appears in (7). (Compare Hall 1992, p 88.) From these equations we obtain

$$v^{(n)}_{1-\alpha} = \left(\frac{b}{n}\right)^{\frac{1}{2}} v^{(b)}_{1-\alpha} + \left\{1 - \left(\frac{b}{n}\right)^{\frac{1}{2}}\right\} z_{1-\alpha} \quad (17)$$
$$+O\left(b^{-\frac{1}{2}}n^{-\frac{1}{2}}\right), \quad (18)$$

which gives us the relationship between the theoretical quantiles. We have available a natural estimator of $v^{(b)}_{1-\alpha}$:

$$\hat v^{(b)}_{1-\alpha} = (1-\alpha)\text{-quantile of } \hat G_b.$$

Since

$$\hat v^{(b)}_{1-\alpha} = v^{(b)}_{1-\alpha} + O_p\left(\left(\frac{b}{n}\right)^{\frac{1}{2}}\right),$$

if we put

$$\hat v^{(n)}_{1-\alpha} = \left(\frac{b}{n}\right)^{\frac{1}{2}} \hat v^{(b)}_{1-\alpha} + \left\{1 - \left(\frac{b}{n}\right)^{\frac{1}{2}}\right\} z_{1-\alpha} \quad, \quad (19)$$

we obtain

$$\hat v^{(n)}_{1-\alpha} = v^{(n)}_{1-\alpha} + O_p\left(\frac{b}{n}\right). \quad (20)$$

We consider now how using this estimate affects the coverage probability of the corresponding interval

$$\hat I_\alpha = (-\infty, \hat\theta_n - n^{-\frac{1}{2}}\hat\sigma_n \hat v^{(n)}_{1-\alpha}). \quad (21)$$

We see immediately that the upper end-point of the interval $\hat I_\alpha$ differs from that of I_α by a term which is $O_p(bn^{-\frac{3}{2}})$, because of (20), and so we can write $\hat I_\alpha$ in the form

$$(-\infty, \hat\theta_n - n^{-\frac{1}{2}}\hat\sigma_n(v^{(n)}_{1-\alpha} + \delta)) \quad ,$$

where $\delta = O_p(bn^{-1})$, from which it follows by a Taylor series argument that the coverage probability of $\hat I_\alpha$ is

$$\mathbb{P}\left\{\hat\theta \in \hat I_\alpha\right\} = \alpha + O_p(bn^{-1}). \quad (22)$$

To see that $\hat I_\alpha$ constructed in this way is an improvement over the confidence interval constructed by a normal approximation, let b be of precise order n^β and suppose $0 \le \beta \le \frac{1}{2}$. (This is the more favourable case. See remarks following (6).) Then the normal approximation can be seen to have coverage errors of order $O_p(n^{-\frac{1}{2}})$, by a Taylor series argument similar to that leading to (22); on the other hand, in this case the coverage error of $\hat I_\alpha$ is $O_p(n^{\beta-1})$, by (22), which can be made arbitrarily close to $O_p(n^{-1})$, the bootstrap interval error, by choosing $\beta \ge 0$ as small as we please.

Similar remarks apply to the various other types of confidence interval described in Hall (1992, Chap. 3). These results can be summarised by saying that for a percentile-t bootstrap confidence interval whose coverage error rate is $O_p(n^{-j/2})$, $j > 1$, the corresponding error rate will be $O_p(bn^{-j/2})$ in the sectioning analogue. It must be remembered in interpreting these facts that these results are asymptotic in nature. Finite sample behaviour of percentile-t sectioning estimates is discussed in a simulation study which will be reported elsewhere.

3 ITERATED SECTIONING

3.1 Introduction

Despite the theoretical appeal of methods based on pivoting, the scope for such methods in complicated situations is rather limited, because in such cases it can be hard to find an appropriate pivot. (This will typically be the case if the asymptotic variance of the estimator can not be reliably estimated in closed form.) Since complicated situations are what one expects to find in simulation studies, it is clear that alternative methods must be sought.

In subsection 2.3, we described the construction of percentile confidence intervals from sectioning, and observed that such intervals do not have coverage error asymptotically better than the normal approximation intervals. This state of affairs is analogous with that obtaining for the bootstrap. A way to improve the coverage properties of bootstrap percentile confidence intervals is to calibrate them. We consider in the next subsections how this idea can be applied to the sectioning percentile intervals.

3.2 Calibration of confidence intervals

There are several ways of calibrating bootstrap confidence intervals: bootstrap iteration (see Hall, 1992, Section 3.11, for example), pre-pivoting (Beran, 1988) and those discussed in Loh (1987).

The analogies between the bootstrap and sectioning have been repeatedly mentioned; what would the analogue of bootstrap iteration be? In bootstrap iteration, each bootstrap sample is itself re-sampled, so what we would expect to do with sectioning is to divide each section into subsections. Here we see an advantage of the bootstrap over sectioning: the sample sizes in each section are smaller than the original sample size, and similarly the subsections are of smaller size than the sections. It is clear that estimates based on small subsections will be very variable. With the bootstrap however all the sample sizes can be kept the same. Because of this problem of small sample sizes in the subsection, we cannot recommend this analogue of the iterated bootstrap.

However, we do have another proposal to offer. The idea of using the bootstrap to calibrate confidence intervals can be applied to many confidence interval constructions. It is natural to apply it in our context.

Specifically, let us suppose we wish to construct a one-sided level-α confidence interval by calibrating a percentile sectioning interval as in subsection 2.3 above. Use the notation of section 2.

An ideal level-β confidence interval for θ would be $(-\infty, y_\beta^{(n)})$, where

$$y_\beta^{(n)} = \hat{\theta}_n + n^{-\frac{1}{2}} \sigma u_\beta^{(n)}. \tag{23}$$

and $u_\beta^{(n)} = \beta$-quantile of F_n.

Unfortunately this will not do, as F_n has to be estimated. The percentile method of subsection 2.3 replaces $u_\beta^{(n)}$ by an estimate based on the quantiles of $\hat{F}_b$, the sectioning approximation of F_b. That is, we construct the interval

$$\hat{I}_\beta = (-\infty, \hat{y}_\beta^{(n)}) \tag{24}$$

where

$$\hat{y}_\beta^{(n)} = \hat{\theta}_n + n^{-\frac{1}{2}} \sigma \hat{u}_\beta^{(n)}, \tag{25}$$

and by anology with (15), (16) we obtain

$$\hat{u}_\beta^{(n)} = \left(\frac{b}{n}\right)^{\frac{1}{2}} \hat{u}_\beta^{(b)} + \left\{1 - \left(\frac{b}{n}\right)^{\frac{1}{2}}\right\} z_\beta, \tag{26}$$

where $\hat{u}_\beta^{(b)}$ is the β-quantile of $\hat{F}_b$.

Since $\hat{y}_\alpha^{(n)}$ is an estimate, the true coverage of the interval $\hat{I}_\alpha$ will only be approximately α. This idea of adjusting α, say to $\alpha + \xi_n$, by estimating the coverage by bootstrap, while new in this context, may be found in Hall (1986), Beran (1987), and Loh (1987). For a detailed discussion, see Hall (1992, Section 3.11).

We outline briefly how the algorithm works, before explaining the theory behind it.

Let S_{ij}^*, $j = 1, ..., B$ be a random sample of size b drawn with replacement from the section S_i and let θ_{ij}^* denote the corresponding estimate of θ.

For $0 \leq \beta \leq 1$, let $u_i^*(\beta)$ be the β-quantile of the set $\{\theta_{ij}^* : j = 1, ..., \beta\}$. Let $p^*(\beta)$ be the proportion of occasions i (out of n/b) on which $\hat{\theta}_n < u_i^*(\beta)$. Thus $p^*(\beta)$ estimates the coverage probability of an interval constructed using the percentile method with β as the nominal coverage level. The calibration idea is to use, instead of α, the $\alpha + \xi$ for which

$$p^*(\alpha + \xi) = \alpha,$$

at least, as nearly as this equation can be solved, allowing for the discreteness of the sampling distributions involved.

If we examine the Edgeworth expansions for F_n, F_b, and invert them as we did to obtain (15), (16), we obtain for any β

$$u_\beta^{(b)} = z_\beta - b^{-\frac{1}{2}} p(z_\beta)\phi(z_\beta) + O(b^{-1}), \tag{27}$$

with an analogous equation for $u_\beta{}^{(n)}$.

It follows that the coverage probability $\pi_n(\beta)$ of $\hat{I}_\beta$ is

$$\mathbb{P}\left\{\theta_o \leq \hat{\theta}_n + n^{-\frac{1}{2}} \sigma (z_\beta - n^{-\frac{1}{2}} p(z_\beta)\phi(z_\beta) + O(n^{-1})\right\} \tag{28}$$

$$= \beta + n^{-\frac{1}{2}} p(z_\beta)\phi(z_\beta) + O(n^{-1}).$$

(Compare Hall (1992, p 102).) Thus if we wish to chose ξ_n such that $\pi_n(\alpha + \xi_n) = \alpha$, then we should take

$$\xi_n = -n^{\frac{1}{2}} p(z_b)\phi(z_\alpha) + O(n^{-1}). \tag{29}$$

Hence we can write

$$\xi_n = \left(\frac{b}{n}\right)^{\frac{1}{2}} \xi_b + O\left(n^{-\frac{1}{2}} b^{-\frac{1}{2}}\right), \tag{30}$$

where ξ_b is obtained by writing b for n in (29). The advantage of using ξ_b is that it may be estimated by the bootstrap construction. For, the bootstrap enables us to replace the Edgeworth expansion which leads to the analogue of (3.7) for b with an Edgeworth expansion which is identical except for the fact that the coefficients of the polynomials are estimated (rather than theoretical values) from a section of size b. From this it follows that we may write

$$\hat{\xi}_b = -b^{-\frac{1}{2}} \hat{p}(z_\alpha)\phi(z_\alpha) + O_p(b^{-\frac{1}{2}})$$

where $\hat{p}$ denotes the estimated polynomial p. Since the coefficients of $\hat{p}$ are within $O_p(b^{-\frac{1}{2}})$ of those of p,

$$\hat{\xi}_b = \xi_b + O_p(b^{-1}).$$

It is $\hat{\xi}_b$ that is produced by the bootstrap algorithm outlined above.

If we now compute the coverage probability of the interval $\left(-\infty, \hat{y}_{\alpha+\hat{\xi}_n}^{(n)}\right)$, where we have put $\hat{\xi}_n = (b/n)^{\frac{1}{2}} \hat{\xi}_b$, we find

$$\mathbb{P}\left\{\theta_0 \leq \hat{y}_{\alpha+\hat{\xi}_n}^{(n)}\right\}$$

$$= \mathbb{P}\left\{\theta_0 \leq \hat{\theta}_n + n^{-\frac{1}{2}} \sigma \hat{u}_{\alpha+\hat{\xi}_n}^{(n)}\right\}$$

$$= \mathbb{P}\Big\{\theta_0 \leq \hat{\theta}_n + n^{-\frac{1}{2}} \sigma \Big\{ z_{\alpha+\hat{\xi}_n} - p\left(z_{\alpha+\hat{\xi}_n}\right) \phi\left(z_{\alpha+\hat{\xi}_n}\right) +$$

$$+O_p\left(n^{-\frac{1}{2}}b^{-\frac{1}{2}}\right)\Big\}\Big\}$$

$$=\mathbb{P}\Big\{n^{\frac{1}{2}}\left(\hat{\theta}_n-\theta_o\right)/\sigma > z_{\alpha+\hat{\xi}_n}$$

$$+p\left(z_{\alpha+\hat{\xi}_n}\right)\phi\left(z_{\alpha+\hat{\xi}_n}\right)+O_p\left(n^{-\frac{1}{2}}b^{-\frac{1}{2}}\right)\Big\}$$

$$=\mathbb{P}\Big\{n^{\frac{1}{2}}\left(\hat{\theta}_n-\theta_o\right)/\sigma > z_{\alpha+\xi_n}$$

$$+p\left(z_{\alpha+\xi_n}\right)\phi\left(z_{\alpha+\xi_n}\right)+$$

$$+O_p\left(n^{-\frac{1}{2}}b^{-\frac{1}{2}}\right)\Big\},$$

since $\hat{\xi}_n=\xi_n+O_p\left(b^{-\frac{1}{2}}n^{-\frac{1}{2}}\right)$.

The last display is equal to

$$\mathbb{P}\left\{\frac{n^{\frac{1}{2}}\left(\hat{\theta}_n-\theta_0\right)}{\sigma} > z_{\alpha+\xi_n}+p\left(z_{\alpha+\xi_n}\right)\phi\left(z_{\alpha+\xi_n}\right)+\dots\right\}$$

$$+O\left(n^{-\frac{1}{2}}b^{-\frac{1}{2}}\right),$$

where ... denotes the remainder of the Cornish-Fisher expansion for $u^{(n)}_{\alpha+\xi_n}$, and hence we find

$$\mathbb{P}\left\{\theta_0\in\left(-\infty,\hat{y}^{(n)}_{\alpha+\hat{\xi}_n}\right)\right\}=\alpha+O\left(n^{-\frac{1}{2}}b^{-\frac{1}{2}}\right).$$

We see that the coverage accuracy has been improved by the calibration procedure. The same approach may be applied to two-sided intervals, symmetric intervals, percentile-t intervals, but we do not pursue that further here.

4 A SIMULATION EXPERIMENT

In Figure 1 we show the result of a simulation experiment in applying the sectioning percentile confidence interval (section 2.2) to estimation of the mean of a standard (mean 1) exponential distribution. One hundred samples of size n = 100were generated , with 20 sections of 5 used to compute a nominal 90% confidence interval. The upper display shows each of these confidence intervals, plotted from left to right in order of increasing mid-point. Also shown is the true mean as a dotted horizontal line. The estimated coverage probability is 1 (all the intervals contain 1). The lower display shows the same thing except that this time a bootstrap calibration was employed. Now the coverage probability is 0.89 (SE= 0.03). As well as being more accurate, the intervals are as often over (5 times) as under (6 times). The simple asymptotic Normal approximation performs poorly, with coverage probability 0.80 (SE= 0.04).

4.1 REFERENCES

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

Carlstein, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Ann. Statist.* **14**, 1171-1179.

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

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion.* Springer-Verlag, New York.

Hartigan, J. (1969). Using subsample values as typical values. *J. Amer. Statist. Assoc.* **64**, 1303-1317.

Kleijnen, J.P.C. (1988). *Statistical Tools for Simulation Practitioners.* Marcel Dekker, New York.

Lewis, P.A.W. and Orav, E.J. (1989). *Simulation Methodology for Statisticians, Operations Analysts, and Engineers: Vol. 1.* Wadsworth & Brooks/Cole, Pacific Grove. CA.

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

Schmeiser, B. (1990). *Simulation Experiments, in Stochastic Models*, D.P. Heyman and M.J. Sobel (eds), North-Holland, Amsterdam.

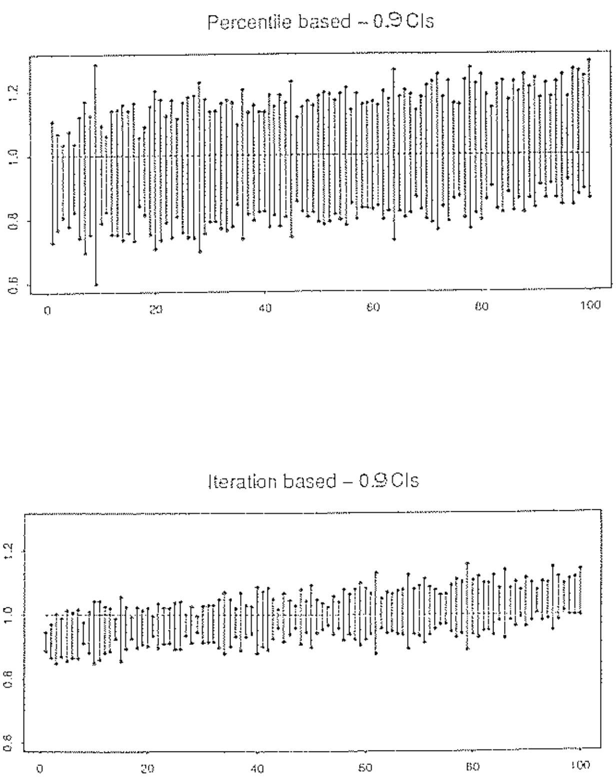

Figure 1. Results of simulation experiment.