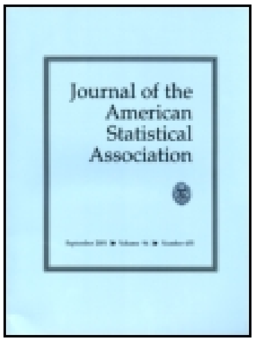

Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

Refining Bootstrap Simultaneous Confidence Sets

Rudolf Beran

To cite this article: Rudolf Beran (1990) Refining Bootstrap Simultaneous Confidence Sets, Journal of the American Statistical Association, 85:410, 417-426

To link to this article: https://doi.org/10.1080/01621459.1990.10476215

Refining Bootstrap Simultaneous Confidence Sets

RUDOLF BERAN*

Simultaneous confidence sets for a collection of parametric functions may be constructed in several different ways. These ways include: (a) the exact pivotal method that underlies Tukey's (1953) and Scheffé's (1953) simultaneous confidence intervals for linear parametric functions in the normal linear model; (b) the method of asymptotic pivots, which is an approximate extension of the pivotal method; (c) the method of bootstrapped roots developed in Beran (1988). These three methods share several features. Each method simultaneously asserts a collection of confidence sets, one confidence set for every parametric function of interest. Each method obtains the uth constituent confidence set by referring a *root* to a critical value; the uth root is a function of the sample and of the uth parametric function. Each method has the same aim: to control the overall level of the simultaneous confidence set and to keep equal the marginal levels of the individual confidence statements that make up the simultaneous confidence set. The three methods differ in how they construct the necessary critical values. The pivotal method requires that the roots be identically distributed and that the supremum over all the roots be a pivot—a function of the sample and of the unknown parameter whose distribution is completely known. The method of asymptotic pivots relaxes these assumptions slightly by requiring them to hold only asymptotically. Nevertheless, the practical scope of both methods is very narrow. In contrast, the method of bootstrapped roots, or B method, is widely applicable. Under general conditions, it generates simultaneous confidence sets that, asymptotically, are balanced and have correct overall coverage probability. Moreover, the B method has a direct Monte Carlo approximation, which makes it easy to use. Under supplementary conditions, the pivotal method turns out to be a special case of the B method. At finite sample sizes, a B method simultaneous confidence set may suffer error in its overall coverage probability and lack of balance among the marginal coverage probabilities of its constituent confidence sets. Of what orders are these errors? What can be done to reduce them? We answer both questions in this article, by introducing and studying the B^2 method—a double iteration of the B method. Under regularity conditions, the B^2 method reduces the asymptotic order of imbalance in the B method simultaneous confidence set; and at the same time, it reduces the asymptotic order of error in overall coverage probability. The B^2 method has a direct Monte Carlo approximation that makes its use straightforward, though computer-intensive.

KEY WORDS: Balance; Coverage probability; Double bootstrap; Higher-order asymptotics.

1. INTRODUCTION

In constructing an ordinary confidence set, the first goal is controlling level. When constructing a simultaneous confidence set, the initial goals are two-fold: (a) controlling the overall level; (b) controlling the relative levels of the individual confidence statements that make up the simultaneous confidence set. Methods for accomplishing both goals are the topic of this article.

Let x_n be a sample of size n whose distribution $P_{\theta,n}$ belongs to a specified parametric family of distributions. The unknown parameter θ, which can be finite or infinite-dimensional, is restricted to a parameter space Θ. Let T be a parametric function, defined on Θ, which has components labeled by an index set U. In other words, $T(\theta) = \{T_u(\theta) : u \in U\}$, where the parametric function T_u is the uth component of T. The index set U can be finite or infinite, as will be illustrated in Example 1.

Suppose that, for each value of u, $C_{n,u}$ is a confidence set for the uth component $T_u(\theta)$. By simultaneously asserting the confidence sets $\{C_{n,u} : u \in U\}$, we obtain a simultaneous confidence set C_n for the family of parametric functions $T(\theta) = \{T_u(\theta)\}$. The problem is to devise the component confidence sets $\{C_{n,u}\}$ in such a way that (a) the level of C_n as a confidence set for $T(\theta)$ is $1 - \alpha$ and (b) the level of $C_{n,u}$ as a confidence set for $T_u(\theta)$ does not vary with u. Property (b), which we term *balance*, ensures that the simultaneous confidence set treats each component $T_u(\theta)$ fairly.

We will review three related approaches to this problem and will propose a fourth method, which offers better control of overall level and of balance at moderate to large sample sizes. The first approach is the exact pivotal method that underlies Tukey's (1953) and Scheffé's (1953) simultaneous confidence sets in the normal linear model. The second method is the obvious asymptotic extension of the pivotal method. Though they are important, these two approaches turn out to have limited scope, because they require more structure than exists in many situations where we seek simultaneous confidence sets.

The third approach is the method of bootstrapped roots, or B method, which is described in Sections 1.3 and 2.1. The B method is widely applicable and generates simultaneous confidence sets that, asymptotically, are balanced and have correct overall coverage probability (Beran 1988). The fourth technique is the method of doubly bootstrapped roots, or B^2 method, which is introduced in Sections 1.4 and 2.2. The B^2 method is basically a double iteration of the B method. Under regularity conditions discussed in Section 3, the B^2 method reduces the asymptotic order of imbalance in the B-method simultaneous confidence set and, at the same time, it reduces the asymptotic order of error in overall coverage probability. In many circumstances, the exact pivotal method turns out to be a special case of both the B method and the B^2 method. More generally, the B^2 method requires a double bootstrap Monte Carlo algorithm that uses roughly the square of the computer time needed by the B method. A SUN 3/140 workstation carried out the numerical example of the B^2 method reported in Section 2.4.

* Rudolf Beran is Professor, Department of Statistics, University of California, Berkeley, CA 94720. This research was supported by National Science Foundation Grant DMS-87-01426.

Journal of the American Statistical Association
June 1990, Vol. 85, No. 410, Theory and Methods

1.1 The Pivotal Method

Let $R_{n,u} = R_{n,u}(x_n, T_u(\theta))$ be a confidence set root for $T_u(\theta)$—a function of the sample and of the parametric function $T_u(\theta)$. Suppose that the roots $\{R_{n,u} : u \in U\}$ are identically distributed and that their supremum is a continuous pivot; that is, the distribution of $\sup_u R_{n,u}$ is continuous and does not depend upon the unknown θ. Let T denote the set of all possible values of $T(\theta) = \{T_u(\theta)\}$ as θ varies over the parameter space. Every point t in this range T can be written in component form $t = \{t_u\}$, where t_u lies in the range T_u of $T_u(\theta)$.

Consider a confidence set for $T_u(\theta)$ of the form

$$C_{n,u} = \{t_u \in T_u : R_{n,u}(x_n, t_u) \leq c_u\}. \tag{1.1}$$

Simultaneously asserting the confidence sets $\{C_{n,u}\}$ generates the following simultaneous confidence set for $T(\theta)$:

$$C_n = \{t \in T : R_{n,u}(x_n, t_u) \leq c_u \text{ for every } u \in U\}. \tag{1.2}$$

Suppose that the critical value c_u is the largest $(1 - \alpha)$th quantile of the distribution of $\sup_u R_{n,u}$. Then C_n has exact level $1 - \alpha$, because $\sup_u R_{n,u}$ is a continuous pivot; and C_n is balanced, because the roots $\{R_{n,u}\}$ are identically distributed.

Example 1. Consider the classical normal linear model, in which the sample x_n has an $N(A\beta, \sigma^2 I)$ distribution, β is r-dimensional, the regression matrix A has rank r, and I is the identity matrix. The unknown parameter $\theta = (\beta, \sigma^2)$ has the usual estimate $\hat{\theta}_n = (\hat{\beta}_n, \hat{\sigma}_n^2)$ from least squares theory. Suppose that $T_u(\theta)$ is a linear combination

$$T_u(\theta) = u'\beta, \tag{1.3}$$

where u is an r-dimensional vector, and that the corresponding root is

$$R_{n,u} = |u'(\hat{\beta}_n - \beta)|\hat{\sigma}_{n,u}, \tag{1.4}$$

where $\hat{\sigma}_{n,u}^2 = u'(A'A)^{-1}u\hat{\sigma}_n^2$. The roots $\{R_{n,u}\}$ are identically distributed, each having a t distribution with $n - r$ df folded over at the origin.

It remains to specify U, the set of r-dimensional vectors over which u is allowed to range. Consider two special cases.

Case 1: U Is a Subspace of Dimension q. Then $\sup_u R_{n,u}$ is a continuous pivot, distributed as $(qW)^{1/2}$, where W has an F distribution with q and $n - r$ df (Miller 1966, chap. 2, sec. 2). The pivotal method's confidence set C_n is just Scheffé's (1953) simultaneous confidence intervals for the linear combinations $\{u'\beta : u \in U\}$.

Case 2: U Is All Pairwise Contrasts in a Balanced One-Way Layout. The parameter β is just the vector of means in this specialization of the linear model. Then $\sup_u R_{n,u}$ is a continuous pivot, distributed as $2^{-1/2}R$, where R has the studentized range distribution with parameters r and $n - r$ (Miller 1966, chap. 2, sec. 1). Now the pivotal method yields Tukey's (1953) simultaneous confidence intervals for all pairwise differences in means.

1.2 The Method of Asymptotic Pivots

An obvious approximate extension of the pivotal method is possible when only the asymptotic distributions of the roots $\{R_{n,u}\}$ are identical and when the asymptotic distribution of $\sup_u R_{n,u}$ is continuous and does not depend on θ. In (1.2), take the critical value c_u to be the largest $(1 - \alpha)$th quantile of the asymptotic distribution of $\sup_u R_{n,u}$. Then, asymptotically, the simultaneous confidence set C_n is balanced and has overall coverage probability $1 - \alpha$. Both statements are true pointwise in θ. Convergence of overall and component *levels* may or may not occur; see Section 4 for further discussion of this point.

Example 2. Consider the following nonparametric generalization of the normal linear model in Example 1. The sample x_n is distributed as $A\beta + e_n$, where e_n is an n vector of iid errors whose common distribution P has mean 0 and finite variance. Here the unknown parameter θ is the pair (β, P). Let the roots $\{R_{n,u}\}$ and the index set U be as in Case 1 of Example 1.

In this nonparametric setting, Scheffé's simultaneous confidence intervals for the linear combinations $\{u'\beta\}$ no longer have overall level $1 - \alpha$; nor are they balanced. Nevertheless, under certain assumptions on the regression matrix A (Huber 1973), the overall coverage probability of Scheffé's confidence intervals C_n converges to $1 - \alpha$, pointwise in θ; and the probability that the component interval $C_{n,u}$ contains the linear combination $u'\beta$ converges pointwise in θ to a value independent of u. Both assertions about limits remain true if the F distribution with q and $n - r$ df is replaced, in calculating the critical values, by q^{-1} times the chi-squared distribution with q df.

1.3 The *B* Method

In general, both the exact and asymptotic distributions of the roots $\{R_{n,u}\}$ depend on u and on θ; and $\sup_u R_{n,u}$ is not a continuous pivot, even asymptotically. The two methods described in Sections 1.1 and 1.2 then fail to produce a simultaneous confidence set that is asymptotically balanced and of correct overall coverage probability.

Example 3. Suppose that x_n is a sample of n iid r vectors whose common distribution P has finite mean vector and finite covariance matrix. The unknown parameter θ in this case is the distribution P. Suppose that the eigenvalues of the covariance matrix are distinct and strictly positive. Define $T_u(\theta)$ to be the uth eigenvector of the covariance matrix, corresponding to the uth largest eigenvalue and expressed as a vector of unit length whose direction is determined by a preset rule. The r ordered orthonormal eigenvectors then define $T(\theta)$. The range T of $T(\theta)$ is the set of all orthonormal systems in r space.

To construct simultaneous confidence cones for the eigenvectors $\{T_u(\theta) : 1 \leq u \leq r\}$, consider the roots

$$R_{n,u} = n(1 - |d_{n,u}'T_u(\theta)|), \tag{1.5}$$

where $d_{n,u}$ is the uth sample eigenvector, standardized to unit length. The asymptotic distribution of $R_{n,u}$ is complicated and depends on both u and the unknown distribution θ (Davis 1977). The asymptotic distribution of $\sup_u R_{n,u}$ depends on θ and is analytically intractable. Neither the pivotal method nor the method of asymptotic pivots can handle this example.

The method of bootstrapped roots, or B method, overcomes distributional difficulties such as those in Example 3. Let $H_{n,u}(\cdot, \theta)$ and $H_n(\cdot, \theta)$ be the left-continuous cumulative distribution functions of $R_{n,u}$ and of $\sup_u H_{n,u}(R_{n,u}, \theta)$, respectively. Suppose that $\hat{\theta}_n$ is a good consistent estimate of θ. Let $\hat{H}_{n,u} = H_{n,u}(\cdot, \hat{\theta}_n)$ and $\hat{H}_n = H_n(\cdot, \hat{\theta}_n)$ be the natural plug-in estimates of the respective cdf's. In Efron's (1979) terminology, $\hat{H}_{n,u}$ and $\hat{H}_n$ are bootstrap estimates.

Let $\hat{H}_{n,u}^{-1}(t)$ and $\hat{H}_n^{-1}(t)$ denote the largest tth quantiles of $\hat{H}_{n,u}$ and $\hat{H}_n$, respectively. The B method sets the critical values of C_n in (1.2) to be

$$c_u = \hat{H}_{n,u}^{-1}[\hat{H}_n^{-1}(1 - \alpha)]. \tag{1.6}$$

An intuitive rationale for this choice is presented in Section 2.1. Each critical value c_u is obtained in two steps: (a) Find the largest $(1 - \alpha)$th quantile of $\hat{H}_n$ and call it b, and (b) find the largest bth quantile of $\hat{H}_{n,u}$. This is c_u. Monte Carlo approximations for $\hat{H}_{n,u}$ and $\hat{H}_n$ are discussed in Section 2.4.

Under general conditions, the B method generates a simultaneous confidence set that, asymptotically, is balanced and has correct overall coverage probability. Example 3 satisfies these conditions (Beran 1988). The B method is, in fact, a logical extension of the exact pivotal method described earlier. If the roots $\{R_{n,u}\}$ are identically distributed pivots and $\sup_u R_{n,u}$ is a continuous pivot, then the B method is equivalent to the pivotal method. In particular, applying the B method to Example 1 yields the Scheffé and Tukey simultaneous confidence intervals in the normal linear model.

1.4 The B^2 Method

At finite sample sizes, a B-method simultaneous confidence set C_n may suffer error in its overall coverage probability and lack of balance among the marginal coverage probabilities of its constituent confidence sets $\{C_{n,u}\}$. The B^2 method, or method of doubly bootstrapped roots, has the important property that it reduces both kinds of error, at least for moderate to large sample sizes. It is defined as follows. Let

$$S_{n,u} = \hat{H}_n[\hat{H}_{n,u}(R_{n,u})]. \tag{1.7}$$

Apply the B method to the transformed roots $\{S_{n,u}\}$ rather than to the original roots $\{R_{n,u}\}$. The simultaneous confidence set so obtained is called the B^2-method confidence set for $T(\theta)$.

More explicitly, let $K_{n,u}(\cdot, \theta)$ and $K_n(\cdot, \theta)$ be the left-continuous cdf's of $S_{n,u}$ and $\sup_u K_{n,u}(S_{n,u}, \theta)$, respectively. Let $\hat{K}_{n,u} = K_{n,u}(\cdot, \hat{\theta}_n)$ and $\hat{K}_n = K_n(\cdot, \hat{\theta}_n)$ denote the corresponding bootstrap estimates of these cdf's, and let $\hat{K}_{n,u}^{-1}$ and $\hat{K}_n^{-1}$ denote the quantile functions of these estimates, defined as in Section 1.3. The B^2 method sets the critical values in (1.2) to be

$$c_n = \hat{H}_{n,u}^{-1}[\hat{H}_n^{-1}\{\hat{K}_{n,u}^{-1}[\hat{K}_n^{-1}(1 - \alpha)]\}]. \tag{1.8}$$

The evaluation of (1.8) is more complicated but otherwise similar to that of (1.6). An intuitive argument for the critical values (1.8) is given in Section 2.2. Monte Carlo approximations for $\hat{K}_{n,u}$, $\hat{K}_n$, $\hat{H}_{n,u}$, and $\hat{H}_n$ are described in Section 2.4.

The B^2 method has several noteworthy properties. It is widely applicable. Under asymptotic expansion assumptions discussed in Section 3, the B^2 method reduces the asymptotic order of imbalance in the B-method simultaneous confidence set; at the same time it reduces the asymptotic order of error in overall coverage probability. Moreover, like the B method, the B^2 method is equivalent to the pivotal method in the circumstances already described at the end of Section 1.3. In the special case in which the index set U has only one element, the B^2 method reduces to the double bootstrap confidence set studied in Beran (1987). In this sense, the B^2 method generalizes the prepivoting transformation discussed in that paper.

Example 4. Suppose that the sample x_n, the parametric functions $\{T_u(\theta)\}$, and the estimate $\hat{\theta}_n = (\hat{\beta}_n, \hat{\sigma}_n^2)$ are as in Example 1, but that the roots are not studentized:

$$R_{n,u} = |u'(\hat{\beta}_n - \beta)|. \tag{1.9}$$

Consider the B method and B^2 method in the case in which U is a subspace of dimension q.

Let Ψ denote the cdf of $|Z|$, where Z has a standard normal distribution. Evidently,

$$H_{n,u}(\cdot, \theta) = \Psi(\cdot/\sigma_u), \tag{1.10}$$

with $\sigma_n^2 = u'(A'A)^{-1}u\sigma^2$. Let χ_q be the cdf of $V^{1/2}$, where V has a chi-squared distribution with q df. Then, by the Cauchy–Schwarz inequality,

$$H_n(\cdot, \theta) = \chi_q \cdot \Psi^{-1}(\cdot). \tag{1.11}$$

In this setting, the B method thus yields an asymptotic approximation to Scheffé's (1953) simultaneous confidence intervals for the linear combinations $\{u'\beta\}$, in which q times the $(1 - \alpha)$th quantile of the F distribution with q and $n - r$ df is replaced by the $(1 - \alpha)$th quantile of the chi-squared distribution with q df. This simultaneous confidence set is not exact but is asymptotically correct, both in level and in balance.

On the other hand, the induced roots $\{S_{n,u}\}$, defined in (1.7), are given here by

$$S_{n,u} = \chi_q[|u'(\hat{\beta}_n - \beta)| / \hat{\sigma}_{n,u}], \tag{1.12}$$

with $\hat{\sigma}_{n,u}$ as in Example 1. Consequently, the B^2 method in this setting yields Scheffé's exact simultaneous confidence intervals for the $\{u'\beta\}$; the argument is similar to that for Case 1 of Example 1. This simultaneous confidence set is balanced and has correct asymptotic level at every sample size.

2. BASIC THEORY

Discussed in this section are three related topics: intuitive derivations for the B-method and B^2-method simultaneous confidence sets; estimates for the imbalance and overall coverage probability of the B-method simultaneous confidence set; and relevant Monte Carlo approximations. The concepts and notation from Section 1 are retained.

2.1 Motivating the B Method

Consider a simultaneous confidence set C_n of the form (1.2), where the critical values $\{c_u\}$ are constants to be determined. The choice of the $\{c_u\}$ is subject to the requirements

$$P_{\theta,n}[C_n \ni T(\theta)] = P_{\theta,n}[R_{n,u} \leq c_u \text{ for every } u \in U] = 1 - \alpha \quad (2.1)$$

and, for every u in U,

$$P_{\theta,n}[C_{n,u} \ni T_u(\theta)] = P_{\theta,n}[R_{n,u} \leq c_u] = \beta, \quad (2.2)$$

where β is unspecified but does not depend on u. Condition (2.1) sets the overall coverage probability, and Condition (2.2) ensures balance. If θ is known, (2.1) and (2.2) determine the critical values $\{c_u\}$. Indeed, when $H_{n,u}$ and H_n are continuous, (2.2) has the solution $c_u = H_{n,u}^{-1}(\beta, \theta)$. Then (2.1) implies that $\beta = H_n^{-1}\ (1 - \alpha, \theta)$. Consequently,

$$c_u = H_{n,u}^{-1}[H_n^{-1}\ (1 - \alpha, \theta), \theta] \quad (2.3)$$

solves the system of Equations (2.1) and (2.2).

Since θ is actually unknown, it is reasonable to substitute the estimate $\hat{\theta}_n$ for θ in the critical values (2.3). The end result of this argument is the B-method simultaneous confidence set for $T(\theta)$,

$$C_{n,B} = \{t \in T : R_{n,u}(x_n, t_u) \leq \hat{H}_{n,u}^{-1}[\hat{H}_n^{-1}(1 - \alpha)] \text{ for every } u \in U\}, \quad (2.4)$$

which was previously introduced in Section 1.3. Let $S_{n,u} = \hat{H}_n[\hat{H}_{n,u}(R_{n,u})]$ as in Section 1.4. Another way of writing $C_{n,B}$ is

$$C_{n,B} = \{t \in T : S_{n,u}(x_n, t_u) \leq 1 - \alpha \text{ for every } u \in U\}. \quad (2.5)$$

The basic asymptotic theory for $C_{n,B}$ rests on representation (2.5) and two convergences that occur under general conditions: (a) The asymptotic distribution of $\hat{H}_{n,u}(R_{n,u})$ is uniform on (0, 1) for every value of u, and (b) the asymptotic distribution of $\sup_u S_{n,u}$ is also uniform on (0, 1). Property (b) ensures that the asymptotic coverage probability of $C_{n,B}$ for $T(\theta)$ is $1 - \alpha$. Property (a) ensures the asymptotic balance of $C_{n,B}$. For more details and for proofs, see Beran (1988).

2.2 Motivating the B^2 Method

In the form (2.5), the B-method simultaneous confidence set refers $\sup_u S_{n,u}$ to the $(1 - \alpha)$th quantile of the uniform distribution on (0, 1). To improve upon this choice of critical values for $C_{n,B}$, consider more generally a simultaneous confidence set for $T(\theta)$ of the form

$$\{t \in T : S_n(x_n, t_u) \leq d_u \text{ for every } u \in U\}, \quad (2.6)$$

where the $\{d_u\}$ are to be determined. Repeating the argument for the B method in Section 2.1 yields the critical values

$$d_u = \hat{K}_{n,u}^{-1}[\hat{K}_n^{-1}(1 - \alpha)], \quad (2.7)$$

where $\hat{K}_{n,u}$ and $\hat{K}_n$ are the bootstrap distributions defined in Section 1.4.

The simultaneous confidence set for $T(\theta)$ that this reasoning generates can be written in two other equivalent forms:

$$\begin{aligned} C_{n,B^2} &= \{t \in T : \hat{K}_n[\hat{K}_{n,u}\{S_{n,u}(x_n, t_u)\}] \leq 1 - \alpha \text{ for every } u \in U\} \\ &= \{t \in T : R_{n,u}(x_n, t_u) \leq \hat{H}_{n,u}^{-1}[\hat{H}_n^{-1}\{\hat{K}_{n,u}^{-1}[\hat{K}_n^{-1}(1 - \alpha)]\}] \text{ for every } u \in U\}. \end{aligned} \quad (2.8)$$

The second expression in (2.8) coincides with the B^2-method simultaneous confidence set defined in Section 1.4.

Since the B^2 method is a double iteration of the B method, the confidence set C_{n,B^2} inherits the basic asymptotic properties of $C_{n,B}$—asymptotic balance and asymptotic coverage probability $1 - \alpha$—under general conditions. That C_{n,B^2} is actually asymptotically better than $C_{n,B}$ can be seen in two ways: through studying special cases, such as Example 4 of Section 1.4 and Example 5 of Section 2.4, and through the analysis in Section 3, which is based on asymptotic expansions.

2.3 Estimating Coverage Probability and Imbalance

As in Section 1.1, let C_n be any simultaneous confidence set for $T(\theta)$ constructed from component confidence sets $\{C_{n,u}\}$. The overall coverage probability of C_n and the uth component coverage probability of C_n are

$$CP(\theta) = P_{\theta,n}[C_n \ni T(\theta)]$$
$$CP_u(\theta) = P_{\theta,n}[C_{n,u} \ni T_u(\theta)], \quad (2.9)$$

respectively. Define the *imbalance* of C_n to be

$$IM(\theta) = |\sup_u CP_u(\theta) - \inf_u CP_u(\theta)|. \quad (2.10)$$

The confidence set C_n is balanced iff the imbalance $IM(\theta)$ vanishes.

Good natural estimates for the coverage probabilities and for the imbalances are $CP(\hat{\theta}_n)$, $CP_u(\hat{\theta}_n)$, and $IM(\hat{\theta}_n)$, respectively. For the B-method simultaneous confidence set $C_{n,B}$, these bootstrap estimates can be rewritten usefully. Define the cdf's $K_{n,u}(\cdot, \theta)$ and $\hat{K}_{n,u}$ as in Section 1.4. Let $K_n'(\cdot, \theta)$ be the right-continuous cdf of $\sup_u S_{n,u}$, and let $\hat{K}_n' = K_n'(\cdot, \hat{\theta}_n)$ be its bootstrap estimate. Evidently, the coverage probability of $C_{n,B}$ is

$$CP_B(\theta) = K_n'(1 - \alpha, \theta), \quad (2.11)$$

and the uth component coverage probability and the imbalance of $C_{n,B}$ are

$$CP_{B,u}(\theta) = K_{n,u}[(1-\alpha)+, \theta]$$
$$IM_B(\theta) = |\sup_u CP_{B,u}(\theta) - \inf_u CP_{B,u}(\theta)|. \quad (2.12)$$

Replacing $K_{n,u}$ and K_n' in (2.11) and (2.12) with $\hat{K}_{n,u}$ and $\hat{K}_n'$, respectively, yields the bootstrap estimates $CP_B(\hat{\theta}_n)$, $CP_{B,u}(\hat{\theta}_n)$, and $IM_B(\hat{\theta}_n)$. Section 3.3 analyzes the rates of convergence of these estimates. The Monte Carlo algorithm that approximates the B^2-method simultaneous confidence set also approximates $CP_B(\hat{\theta}_n)$, $CP_{B,u}(\hat{\theta}_n)$, and $IM_B(\hat{\theta}_n)$. Section 2.4 presents the details.

Example 5. Consider the normal r-sample model, in which x_n is the union of r independent subsamples of sizes $\{n_u : 1 \le u \le r\}$, respectively, and n denotes the overall sample size. The uth subsample consists of n_u iid observations from the $N(\mu_u, \sigma_u^2)$ distribution. The unknown parameter θ is $(\mu_1, \ldots, \mu_r, \sigma_1^2, \ldots, \sigma_r^2)$. Let $(\hat{\mu}_{n,u}, \hat{\sigma}_{n,u}^2)$ denote the usual estimates of (μ_u, σ_u^2). Suppose that the parametric functions of interest are the means

$$T_u(\theta) = \mu_u, \qquad 1 \le u \le r, \quad (2.13)$$

and that the corresponding roots are

$$R_{n,u} = n_u^{1/2}|\hat{\mu}_{n,u} - \mu_u|. \quad (2.14)$$

For Ψ as in Example 4,

$$H_{n,u}(x, \theta) = \Psi(x/\sigma_u)$$
$$H_n(x, \theta) = x^r \quad \text{for } 0 \le x \le 1. \quad (2.15)$$

Consequently, the B-method simultaneous confidence set for $\mu = (\mu_1, \ldots, \mu_r)$ is

$$C_{n,B} = \{\mu : n_u^{1/2}|\hat{\mu}_{n,u} - \mu_u| \le \hat{\sigma}_{n,u}\Psi^{-1}[(1-\alpha)^{1/r}] \quad \text{for } 1 \le u \le r\}. \quad (2.16)$$

Let Ψ_ν denote the cdf of $|W|$ when W has a t distribution with ν df. The marginal level of the uth component confidence interval in (2.16) is $\Psi_{n_u-1} \cdot \Psi^{-1}[(1-\alpha)^{1/r}]$. The overall level of $C_{n,B}$ is the product of the r marginal levels. In this example, $CP_B(\theta)$ and $IM_B(\theta)$ do not depend on θ; consequently, their natural estimates are exact.

From (2.14) and (2.15), it follows that the transformed roots $\{S_{n,u}\}$ for the B^2 method are

$$S_{n,u} = \{\Psi[n_u^{1/2}\,|\hat{\mu}_{n,u} - \mu_u|\,/\hat{\sigma}_{n,u}]\}^r. \quad (2.17)$$

Thus

$$C_{n,B^2} = \{\mu : n_u^{1/2}\,|\hat{\mu}_{n,u} - \mu_u| \le \hat{\sigma}_{n,u}\Psi_{n_u-1}^{-1}[(1-\alpha)^{1/r}] \quad \text{for } 1 \le u \le r\}. \quad (2.18)$$

The overall level of C_{n,B^2} is $1-\alpha$; the marginal levels are each $(1-\alpha)^{1/r}$. The B^2-method simultaneous confidence set is exact and balanced in this example. Table 1, on the other hand, records the imbalance and overall level of $C_{n,B}$ in the two-sample case ($r = 2$), when sample sizes are not large. In this situation, the improvement achieved by the B^2 method is worthwhile.

Table 1. Overall Level, Imbalance, and Marginal Levels of $C_{n,B}$ and C_{n,B^2} in the Two-Sample Normal Model of Example 5

Confidence set	Subsample sizes (n_1, n_2)	Overall level	Imbalance	Marginal levels for μ_1	μ_2
$C_{n,B}$	(5, 5)	.830	0	.911	.911
	(5, 10)	.864	.037	.911	.948
	(5, 20)	.877	.051	.911	.962
	(10, 10)	.898	0	.948	.948
	(10, 20)	.912	.014	.948	.962
	(20, 20)	.926	0	.962	.962
C_{n_1,B^2}	(n_1, n_2)	.950	0	.975	.975

NOTE: The nominal overall level is .95.

2.4 Monte Carlo Approximations

Apart from simple special cases, closed-form expressions are not available for the bootstrap cdf's $\hat{H}_{n,u}$, $\hat{H}_n$, $\hat{K}_{n,u}$, and $\hat{K}_n$ that are required by the B method and the B^2 method. Analytical approximations, based on asymptotic expansions, are occasionally available for some or all of these distribution functions. In general, Monte Carlo methods provide useful approximations to these bootstrap cdf's and to the estimates of overall coverage probability and of imbalance that were introduced in Section 2.3.

Underlying the Monte Carlo algorithm are the following representations. Given the sample x_n, let x_n^* be a bootstrap sample of size n drawn from the fitted model $P_{\hat{\theta}_n,n}$. Let $\theta_n^* = \hat{\theta}_n(x_n^*)$ denote the estimate of θ recalculated from the bootstrap sample x_n^*. Given x_n and x_n^*, let x_n^{**} be a bootstrap sample drawn from the re-fitted model $P_{\theta_n^*,n}$. Write

$$R_{n,u}^* = R_{n,u}(x_n^*, \hat{\theta}_n)$$
$$R_{n,u}^{**} = R_{n,u}(x_n^{**}, \theta_n^*)$$
$$S_{n,u}^{**} = S_{n,u}(x_n^*, \hat{\theta}_n). \quad (2.19)$$

Then

$$\hat{H}_{n,u}(x) = P_{\hat{\theta}_n,n}[R_{n,u}^* < x \mid x_n]$$
$$\hat{H}_n(x) = P_{\hat{\theta}_n,n}[\sup_u \hat{H}_{n,u}(R_{n,u}^*) < x \mid x_n]. \quad (2.20)$$

Define

$$H_{n,u}^*(x) = H_{n,u}(x, \theta_n^*) = P_{\theta^*_n,n}[R_{n,u}^{**} < x \mid x_n^*, x_n]$$
$$H_n^*(x) = H_n(x, \theta_n^*)$$
$$= P_{\theta^*_n,n}[\sup_u H_{n,u}^{**}(R_{n,u}^{**}) < x \mid x_n^*, x_n]. \quad (2.21)$$

Since $S_{n,u} = \hat{H}_n \cdot \hat{H}_{n,u}(R_{n,u})$, it follows from (2.21) that

$$S_{n,u}^* = H_n^* \cdot H_{n,u}^*(R_{n,u}^*)$$
$$= P_{\theta^*_n,n}[\sup_u H_{n,u}^*(R_{n,u}^{**}) < H_{n,u}^*(R_{n,u}^*) \mid x_n^*, x_n]. \quad (2.22)$$

Then

$$\hat{K}_{n,u}(x) = P_{\hat{\theta}_n,n}[S_{n,u}^* < x \mid x_n]$$
$$\hat{K}_n(x) = P_{\hat{\theta}_n,n}[\sup_u \hat{K}_{n,u}(S_{n,u}^*) < x \mid x_n]$$
$$\hat{K}_n'(x) = P_{\hat{\theta}_n,n}[\sup_u S_{n,u}^* \le x \mid x_n]. \quad (2.23)$$

These representations for $\hat{H}_{n,u}$, $\hat{H}_n$, $\hat{K}_{n,u}$, and $\hat{K}_n$ show that the B method is a single bootstrap procedure whereas the B^2 method is a double bootstrap procedure.

The representations also motivate the following Monte Carlo algorithm for approximating the cdf's $\hat{H}_{n,u}$, $\hat{H}_n$, $\hat{K}_{n,u}$, and $\hat{K}_n$ required by the B method and the B^2 method and for approximating the additional cdf $\hat{K}_n'$ required by the coverage probability estimate $CP_B(\hat{\theta}_n)$ of Section 2.3. The algorithm has five steps.

1. Let $y_1^*, \ldots, y_M^*$ be M bootstrap samples of size n, each drawn from the fitted model $P_{\hat{\theta}_n,n}$. These samples are conditionally independent, given the original sample x_n. Set $R_{u,j}^* = R_{n,u}(y_j^*, \hat{\theta}_n)$, the uth root recalculated from the jth bootstrap sample and the estimate $\hat{\theta}_n$. The left-continuous empirical cdf of the values $\{R_{u,j}^* : 1 \le j \le M\}$ approximates $\hat{H}_{n,u}$ for sufficiently large M.
2. Let
$$V_j = M^{-1}\sup_u[\operatorname{rank}(R_{u,j}^*) - 1], \tag{2.24}$$
the rank being calculated over the subscript j. The left-continuous empirical cdf of the values $\{V_j : 1 \le j \le M\}$ approximates $\hat{H}_n$ for sufficiently large M.
3. Let $\theta_j^* = \hat{\theta}_n(y_j^*)$ denote the estimate of θ recomputed from the jth bootstrap sample. For every j, let $y_{j,1}^{**}, \ldots, y_{j,N}^{**}$ be N further bootstrap samples of size n, each drawn from the refitted model $P_{\theta_j^*,n}$. These samples are conditionally independent, given the original sample x_n and the first-round bootstrap samples $\{y_j^* : 1 \le j \le M\}$. Set $R_{u,j,k}^{**} = R_{n,u}(y_{j,k}^{**}, \theta_j^*)$, the uth root recalculated from the (j, k)th second-round bootstrap sample and the estimate θ_j^*. For every u and j, let $Z_{u,j}$ be the fraction of the values $\{R_{u,j,k}^{**} : 1 \le k \le N\}$ that are less than $R_{u,j}^*$. Let
$$V_{j,k} = N^{-1}\sup_u[\operatorname{rank}(R_{u,j,k}^{**}) - 1], \tag{2.25}$$
the rank being calculated over the subscript k. For every u and j, let $S_{u,j}^*$ be the fraction of the values $\{V_{j,k} : 1 \le k \le N\}$ that are less than $Z_{u,j}$. The left-continuous empirical cdf of the values $\{S_{u,j}^* : 1 \le j \le M\}$ approximates $\hat{K}_{n,u}$ for sufficiently large M and N.
4. Repeat step 2 with the $\{S_{u,j}^* : 1 \le j \le M\}$ in place of the $\{R_{u,j}^* : 1 \le j \le M\}$. This yields an approximation to $\hat{K}_n$ for sufficiently large M and N.
5. For every j, let
$$S_j^* = \sup_u S_{u,j}^*. \tag{2.26}$$
The right-continuous cdf of the values $\{S_j^* : 1 \le j \le M\}$ approximates $\hat{K}_n'$ for sufficiently large M and N.

No good theoretical guidelines yet exist for the choice of M and N in this algorithm. A practical strategy is to increase M and N until the desired critical values for the B-method and B^2-method simultaneous confidence sets stabilize. Preliminary trials of the algorithm suggest taking both M and N equal to at least 2,000 for 90% simultaneous confidence intervals.

Example 6. Mardia, Kent, and Bibby (1979) reported test scores for 88 college students, each of whom took two closed-book and three open-book tests. Were there differences among the open-book tests? We can address this question by constructing simultaneous confidence intervals for the pairwise differences of the three open-book test score means, on the supposition that the test scores for each student are a random vector drawn from a common unknown distribution, and by also constructing simultaneous confidence intervals for the pairwise ratios of the three open-book test score standard deviations.

To better bring out the effects of the B^2 method as a refinement of the B method, we artificially base the analysis on only the first 22 of the 88 test score vectors. For this subsample, the sample mean vector and sample covariance matrix of the open-book test scores are

$$\hat{\mu} = \begin{pmatrix} 62.55 \\ 61.09 \\ 59.64 \end{pmatrix}, \quad \hat{\Sigma} = \begin{pmatrix} 40.74 & 21.38 & 61.64 \\ 21.38 & 34.66 & 42.23 \\ 61.64 & 42.23 & 194.34 \end{pmatrix}. \tag{2.27}$$

The pairwise differences of sample means are thus

$$\hat{\mu}_1 - \hat{\mu}_2 = 1.46, \quad \hat{\mu}_1 - \hat{\mu}_3 = 2.91, \quad \hat{\mu}_2 - \hat{\mu}_3 = 1.45, \tag{2.28}$$

and the pairwise ratios of sample standard deviations are

$$\hat{\sigma}_1/\hat{\sigma}_2 = 1.08, \quad \hat{\sigma}_1/\hat{\sigma}_3 = .458, \quad \hat{\sigma}_2/\hat{\sigma}_3 = .422. \tag{2.29}$$

Let the $\{\mu_i\}$ and $\{\sigma_i\}$ denote the population means and standard deviations. As roots for the desired simultaneous confidence sets, we use

$$R_{n,i,j} = |(\hat{\mu}_i - \hat{\mu}_j) - (\mu_i - \mu_j)| \tag{2.30}$$

in the case of the differences in means and

$$R_{n,i,j} = \max(r_{i,j}, r_{i,j}^{-1}), \tag{2.31}$$

where

$$r_{i,j} = (\hat{\sigma}_i/\hat{\sigma}_j)/(\sigma_i/\sigma_j), \tag{2.32}$$

in the case of the ratios of standard deviations. In both cases, the index $u = (i, j)$ ranges over the set of values $U = \{(1, 2), (1, 3), (2, 3)\}$. Table 2 records the critical values generated for these two sets of roots (2.30) and (2.31), at nominal level .90 by the B method and the B^2 method. Table 3 reports the *estimated* overall coverage probability, marginal coverage probabilities, and imbalance of the B-method simultaneous confidence sets, using the bootstrap estimates of Section 2.3.

The estimates in Table 3 indicate that the B-method simultaneous confidence sets for differences of means and

Table 2. Critical Values Defining $C_{n,B}$ and C_{n,B^2} for the Two Sets of Roots in Example 6

Sample size n	Confidence set	Critical values for differences of means (1, 2)	(1, 3)	(2, 3)	Critical values for ratios of standard deviations (1, 2)	(1, 3)	(2, 3)
22	$C_{n,B}$	2.27	4.27	4.77	1.66	1.44	1.40
	C_{n,B^2}	2.45	4.86	4.91	1.64	1.55	1.42

NOTE: The nominal overall level is .90 in each case. The Monte Carlo approximation is based on $2{,}000^2$ double bootstrap samples.

Table 3. Estimated Overall Coverage Probability, Imbalance, and Marginal Coverage Probabilities of $C_{n,B}$ for the Two Sets of Roots in Example 6

Sample size n	Confidence set for	Overall coverage probability	Imbalance	Marginal coverage probabilities for (1, 2)	(1, 3)	(2, 3)
22	Differences of means	.851	.009	.937	.930	.939
	Ratios of standard deviations	.828	.027	.923	.913	.940

NOTE: The nominal overall level is .90 in each case. The Monte Carlo approximation is based on $2{,}000^2$ double bootstrap samples.

for ratios of standard deviations are nearly balanced, despite the small sample size; however, the overall coverage probability in both cases is well below the intended .90. The B^2 method accordingly adjusts most critical values upwards, as seen in Table 2. The last digits in Table 2 are not numerically stable, in the sense that increasing the numbers M and N of bootstrap samples beyond $M = N = 2{,}000$ might cause these last digits to change. $M = N = 2{,}000$ was the practical limit, however, on the SUN 3/140 workstation used to perform the calculation for this example.

For the subsample under discussion, the B^2 method at nominal level .90 yields the simultaneous confidence intervals

$$(.66, 1.77), \quad (.30, .71), \quad (.30, .60), \tag{2.33}$$

for the pairwise ratios of standard deviations, taken in the order of (2.29). The corresponding B-method intervals are

$$(.65, 1.79), \quad (.32, .66), \quad (.30, .64). \tag{2.34}$$

Scores from the third open-book test are thus significantly more variable than scores from the other two open-book tests. On the other hand, neither method finds interesting pairwise differences among the three mean scores.

3. ASYMPTOTIC ANALYSIS

The analysis in this section compares the overall and marginal coverage probabilities of the B-method and B^2-method simultaneous confidence sets. Rates of covergence are determined for the estimated coverage probabilities and estimated imbalance of the B-method simultaneous confidence set. The treatment rests on formal asymptotic expansions.

3.1 Coverage Probabilities of the B Method

Suppose that the estimate $\hat{\theta}_n$ is $n^{1/2}$-consistent and that for some nonnegative integers r and s, the following asymptotic expansions hold, uniformly in x and u and locally uniformly in θ:

$$H_{n,u}(x, \theta) = H_{n,u}(x) + n^{-r/2}h_u(x, \theta) + O(n^{-(r+1)/2})$$
$$H_n(x, \theta) = H_n(x) + n^{-s/2}h(x, \theta) + O(n^{-(s+1)/2}). \tag{3.1}$$

The first term in each expansion does not depend on θ and is of order $O(1)$ in n. The nonnegative integers r and s thus indicate the order of dependence on θ. Suppose, as is typically true, that the functions on the right side of (3.1) are differentiable in x and θ. Let $t = \min(r, s)$. We will argue here that the overall coverage probability and imbalance of $C_{n,B}$ normally satisfy

$$CP_B(\theta) = 1 - \alpha + O(n^{-(t+1)/2})$$
$$IM_B(\theta) = O(n^{-(t+1)/2}) \tag{3.2}$$

pointwise in θ.

This is the leading case, which arises in constructing simultaneous one-sided confidence intervals. Other cases arise in which some of the higher-order terms in Expansions (3.1) may vanish. The analysis of these other possibilities parallels the treatment given here for the leading case.

From (3.1)

$$H_n[H_{n,u}(x, \theta), \theta] = H_n \cdot H_{n,u}(x) + n^{-t/2}i_u(x, \theta) + O(n^{-(t+1)/2}) \tag{3.3}$$

uniformly in x and u and locally uniformly in θ. Hence

$$S_{n,u} = \hat{H}_n \cdot \hat{H}_{n,u}(R_{n,u}) = H_n\,[H_{n,u}(R_{n,u}, \theta), \theta] + O_p(n^{-(t+1)/2}). \tag{3.4}$$

Since the jumps in $H_{n,u}(\cdot, \theta)$ are at most of order $O(n^{-(r+1)/2})$,

$$P_{\theta,n}[H_{n,u}(R_{n,u}, \theta) \le x] = U(x) + O(n^{-(r+1)/2}), \tag{3.5}$$

where U is the uniform (0, 1) cdf. Thus

$$P_{\theta,n}[H_n\{H_{n,u}(R_{n,u}, \theta), \theta\} \le x] = H_n^{-1}(x, \theta) + O(n^{-(r+1)/2}), \tag{3.6}$$

Consequently, in regular cases $K_{n,u}$, the left-continuous cdf of $S_{n,u}$, has an expansion of the form

$$K_{n,u}(x, \theta) = H_n^{-1}(x, \theta) + n^{-(t+1)/2}k_u(x, \theta) + O(n^{-(t+2)/2}) \tag{3.7}$$

uniformly in x and u.

Consider next K_n', the right-continuous cdf of $\sup_u S_{n,u}$. In regular cases, K_n' differs by a term of order $O(n^{-(t+1)/2})$ from the right-continuous cdf of $H_n[\sup_u H_{n,u}(R_{n,u}, \theta), \theta]$ because of (3.4). Hence

$$K_n'(x, \theta) = U(x) + n^{-(t+1)/2}v(x, \theta) + O(n^{-(t+1)/2}) \tag{3.8}$$

uniformly in x.

On the other hand, it follows from (3.7) and the definition of K_n in Section 1.4 that K_n differs by a term of order $O(n^{-(t+1)/2})$ from the left-continuous cdf of $H_n^{-1}(\sup_u S_{n,u}, \theta)$. In regular cases, because of (3.8),

$$K_n(x, \theta) = H_n(x, \theta) + n^{-(t+1)/2}k(x, \theta) + O(n^{-(t+2)/2}) \tag{3.9}$$

uniformly in x.

From (3.8), the overall coverage probability of $C_{n,B}$ for $T(\theta)$ satisfies the first line of (3.2). From (3.7), the marginal coverage probability of the uth component confidence set in $C_{N,B}$ is

$$CP_{B,u}(\theta) = H_n^{-1}(1 - \alpha, \theta) + n^{-(t+1)/2}k_u((1 - \alpha)+, \theta) + O(n^{-(t+2)/2}) \quad (3.10)$$

uniformly in u. Hence the imbalance $IM_B(\theta)$ satisfies the second relation in (3.2).

3.2 Coverage Probabilities of the B^2 Method

By continuing the argument begun in the previous section, we will show that the overall coverage probability and imbalance of C_{n,B^2} normally satisfy

$$CP_{B^2}(\theta) = 1 - \alpha + O(n^{-(t+2)/2})$$
$$IM_{B^2}(\theta) = O(n^{-(t+2)/2}) \quad (3.11)$$

pointwise in θ. Comparing (3.11) with (3.2) establishes the principal conclusion of this article: Under regularity conditions, the B^2 method reduces the asymptotic order of imbalance in the B-method simultaneous confidence set, and at the same time it reduces the asymptotic order of error in overall coverage probability.

Define $T_{n,u} = \hat{K}_n[\hat{K}_{n,u}(S_{n,u})]$. In view of (2.8), the B^2-method simultaneous confidence set can be written as

$$C_{n,B^2} = \{t \in T : \sup_u T_{n,u}(x_n, t_u) \le 1 - \alpha\}. \quad (3.12)$$

From (3.7) and (3.9),

$$K_n[K_{n,u}(x, \theta), \theta] = U(x) + n^{-(t+1)/2}j_u(x, \theta) + O(n^{-(t+2)/2}) \quad (3.13)$$

uniformly in x and u and locally uniformly in θ. Hence

$$T_{n,u} = K_n[K_{n,u}(S_{n,u}, \theta), \theta] + O_p(n^{-(t+2)/2}). \quad (3.14)$$

Moreover, by analogy with (3.6),

$$P_{\theta,n}[K_n\{K_{n,u}(S_{n,u}, \theta), \theta\} \le x] = K_n^{-1}(x, \theta) + O(n^{-(t+2)/2}), \quad (3.15)$$

the discontinuities in $K_{n,u}$ being at most of order $O(n^{-(t+2)/2})$, assuming k_u is continuous in x. Consequently, in regular cases $M_{n,u}$, the left-continuous cdf of $T_{n,u}$ has an expansion of the form

$$M_{n,u}(x, \theta) = K_n^{-1}(x, \theta) + n^{-(t+2)/2}m_u(x, \theta) + O(n^{-(t+3)/2}) \quad (3.16)$$

uniformly in x and u.

Consider next M_n', the right-continuous cdf of $\sup_u T_{n,u}$. In regular cases, M_n' differs by a term of order $O(n^{-(t+2)/2})$ from the right-continuous cdf of $K_n[\sup_u K_{n,u}(S_{n,u}, \theta), \theta]$ because of (3.14). Hence

$$M_n'(x, \theta) = U(x) + n^{-(t+2)/2}w(x, \theta) + O(n^{-(t+3)/2}) \quad (3.17)$$

uniformly in x.

On the other hand, it follows from (3.16) that M_n, the left-continuous cdf of $\sup_u M_{n,u}(T_{n,u}, \theta)$, differs by a term of order $O_p(n^{-(t+2)/2})$ from left-continuous cdf of $K_n^{-1}(\sup_u T_{n,u}, \theta)$. In regular cases, because of (3.17),

$$M_n(x, \theta) = K_n(x, \theta) + n^{-(t+2)/2}m(x, \theta) + O(n^{-(t+3)/2}) \quad (3.18)$$

uniformly in x.

The assertions in (3.11) about the overall coverage probability and imbalance of C_{n,B^2} follow from (3.17) and (3.16), respectively.

3.3 Convergence of Coverage Probability Estimates

How good are the bootstrap estimates $CP_B(\hat{\theta}_n)$ and $IM_B(\hat{\theta}_n)$ of the overall coverage probability $CP_B(\theta)$ and imbalance $IM_B(\theta)$ of the B-method simultaneous confidence set? It follows from (2.11) and (3.8) that

$$CP_B(\hat{\theta}_n) = CP_B(\theta) + O_p(n^{-(t+2)/2}). \quad (3.19)$$

The imbalance $IM_B(\theta)$ converges to 0 asymptotically. From (2.12) and (3.7),

$$IM_B(\hat{\theta}_n) = IM_B(\theta) + O_p(n^{-(t+2)/2}). \quad (3.20)$$

By contrast, consider the simpler asymptotic estimates of overall coverage probability and of imbalance: $\hat{CP}_B = 1 - \alpha$ and $\hat{IM}_B = 0$. It is immediate from (3.2) that

$$\hat{CP}_B = CP_B(\theta) + O(n^{-(t+1)/2}) \quad (3.21)$$

and

$$\hat{IM}_B - IM_B(\theta) + O(n^{-(t+1)/2}). \quad (3.22)$$

Thus the bootstrap estimates dominate the simple asymptotic estimates of $CP_B(\theta)$ and $IM_B(\theta)$ for all sufficiently large n.

Convergence of $CP_{B,u}(\hat{\theta}_n)$ to the uth component coverage probability $CP_{B,u}(\theta)$ is a more complicated matter. Because of (3.10) and (3.1),

$$CP_{B,u}(\hat{\theta}_n) = CP_{B,u}(\theta) + O_p(n^{-(s+1)/2}) + O_p(n^{-(t+2)/2}). \quad (3.23)$$

The rate of convergence in (3.23) can be slower than that in (3.19).

4. COMPLEMENTS

This final section treats several further questions concerning the B^2 method: the distinction between level and coverage probability; computational strategies for B^2-method critical values when the index set U is infinite; and the asymptotic properties of higher-order iterates of the B method.

4.1 Level and Imbalance

By the analysis of Section 3.2, confidence set C_{n,B^2} has the property that its overall coverage probability converges *pointwise* to $1 - \alpha$,

$$CP_{B^2}(\theta) = 1 - \alpha + O(n^{-(t+2)/2}) \quad \text{for every } \theta, \quad (4.1)$$

and that its imbalance converges *pointwise* to 0,

$$IM_{B^2}(\theta) = O(n^{-(t+2)/2}) \quad \text{for every } \theta. \qquad (4.2)$$

Here t is a nonnegative integer, defined in Section 3.1, which depends on the parametric model and the roots $\{R_{n,u}\}$. These two convergences are not very strong, because the remainder terms on the right side of (4.1) and (4.2) could remain large for some value of θ no matter how great the sample size n might be.

More satisfactory would be uniform rates of convergence for overall coverage probability and imbalance:

$$\sup_\theta |CP_{B^2}(\theta) - (1-\alpha)| = O(n^{-(t+2)/2})$$
$$\sup_\theta IM_{B^2}(\theta) = O(n^{-(t+2)/2}). \qquad (4.3)$$

The asymptotic expansions in Sections 3.1 and 3.2 typically converge uniformly in θ, as long as θ is restricted to a compact set. Consequently, (4.3) holds if the parameter space Θ lies within a compact set. In particular, the first line of (4.3) implies a rate of convergence for the overall level of C_{n,B^2}:

$$\inf_\theta CP_{B^2}(\theta) = 1 - \alpha + O(n^{-(t+2)/2}). \qquad (4.4)$$

Unfortunately, when θ is infinite dimensional, the parameter space Θ usually does not lie within a compact and (4.3) breaks down.

The notion of controlling the overall level and the marginal levels of a simultaneous confidence set thus seems too strong. More accessible is the goal of controlling *average* coverage probabilities. We conjecture that, under regularity conditions,

$$\int CP_{B^2}(\theta)\pi(d\theta) = 1 - \alpha + O(n^{-(t+2)/2}), \qquad (4.5)$$

where π is any member of some large class of prior distributions supported on Θ. The average coverage probability (4.5) makes sense if we think of nature selecting the true θ from a prior distribution π. For an interesting technical discussion of average coverage probabilities, see Woodroofe (1986).

4.2 Critical Values for Infinite U

The Monte Carlo algorithm for B^2-method critical values (see Sec. 2.4) requires calculation of four suprema over U. Doing this exactly may not be possible when the number of elements in U is infinite. A natural approximation strategy is to replace supremum over U by maximum over a well-chosen finite subset of U, which may be deterministic or random.

Standard numerical algorithms for maximization search over a deterministic subset of U. Alternatively, the search subset might consist of a random sample drawn from a distribution on U that has full support (i.e., the distribution gives positive probability to every nonempty open subset of U). More complex random search schemes also exist. When the dimension of U is not low, a random search usually needs fewer points to approximate a supremum adequately than does a deterministic search. For further analysis of this property, in connection with the B method, see Beran (1988).

4.3 The B^j Method

Two iterations of the B method gave the B^2 method. The process may be continued: j iterations of the B method define the B^j method. We illustrate this by describing the B^3 method explicitly. As in Section 3.2, let

$$T_{n,u} = \hat{K}_n[\hat{K}_{n,u}(S_{n,u})], \qquad (4.6)$$

where $S_{n,u}$ is defined by (1.7). The transformed roots $\{T_{n,u}\}$ are the outcome of the B^2 method. Applying the B method to the roots $\{T_{n,u}\}$ yields the B^3-method simultaneous confidence set.

In other words, let $L_{n,u}(\cdot, \theta)$ and $L_n(\cdot, \theta)$ be the left-continuous cdf's of $T_{n,u}$ and $\sup_u L_{n,u}(T_{n,u}, \theta)$, respectively. Let $\hat{L}_{n,u} = L_{n,u}(\cdot, \hat{\theta}_n)$ and $\hat{L}_n = L_n(\cdot, \hat{\theta}_n)$ denote the corresponding bootstrap estimates of these cdf's. The B^3 method sets the critical values in the simultaneous confidence set (1.2) to be

$$c_u = \hat{H}_{n,u}^{-1}[\hat{H}_n^{-1}\{\hat{K}_{n,u}^{-1}[\hat{K}_n^{-1}\{\hat{L}_{n,u}^{-1}[\hat{L}_n^{-1}(1-\alpha)]\}]\}]. \qquad (4.7)$$

This expression may be compared with (1.8) for the B^2 method and with (1.6) for the B method.

Under regularity conditions, the overall coverage probability and imbalance of C_{n,B^j} satisfy

$$CP_{B^j}(\theta) = 1 - \alpha + O(n^{-(t+j)/2})$$
$$IM_{B^j}(\theta) = O(n^{-(t+j)/2}) \qquad (4.8)$$

pointwise in θ. The argument is an iteration of that in Section 3.2. Several features of this result are noteworthy.

1. For all sufficiently large n, the error in coverage probability and the imbalance of $C_{n,B^{j+1}}$ will be smaller than their counterparts for C_{n,B^j}. For fixed sample size n, however—the situation usually encountered in practice—there is no assurance that the error in coverage probability and imbalance of C_{B^j} will decrease as j increases.
2. Constructing C_{B^j}, whether exactly or by some combination of Monte Carlo and analytical approximations, is rarely feasible, at present, for $j \geq 3$.
3. The B^j method falls into a very general class of iterated bootstrap methods recently identified by Hall and Martin (1988). A common feature of these bootstrap methods is that they seek to solve a system of equations analogous to the system (2.1) and (2.2).

[*Received February 1988. Revised July 1989.*]

REFERENCES

Beran, R. (1987), "Prepivoting to Reduce Level Error of Confidence Sets," *Biometrika*, 74, 457–468.

——— (1988), "Balanced Simultaneous Confidence Sets," *Journal of the American Statistical Association*, 83, 679–697.

Davis, A. W. (1977), "Asymptotic Theory for Principal Component Analysis: Nonnormal Case," *The Australian Journal of Statistics*, 19, 206–212.

Efron, B. (1979), "Bootstrap Methods: Another Look at the Jackknife," *The Annals of Mathematical Statistics*, 7, 1–26.

Hall, P., and Martin, M. A. (1988), "On Bootstrap Resampling and Iteration," *Biometrika*, 75, 661–671.

Huber, P. J. (1973), "Robust Regression: Asymptotics, Conjectures, and Monte Carlo," *The Annals of Statistics*, 1, 799–821.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979), *Multivariate Analysis*, New York: Academic Press.

Miller, R. (1966), *Simultaneous Statistical Inference*, New York: McGraw-Hill.

Scheffé, H. (1953), "A Method for Judging All Contrasts in the Analysis of Variance," *Biometrika*, 40, 87–104.

Tukey, J. W. (1953), "The Problem of Multiple Comparisons," unpublished preprint.

Woodroofe, M. (1986), "Very Weak Expansions for Sequential Confidence Levels," *The Annals of Statistics*, 14, 1049–1067.