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.
EFRON, B. and TIBSHIRANI, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy (with discussion). *Statist. Sci.* **1** 54–77.

DEPARTMENT OF STATISTICS
SEQUOIA HALL
STANFORD UNIVERSITY
STANFORD, CALIFORNIA 94305-4065

WEI-YIN LOH[1]

University of Wisconsin-Madison

It was a pleasure to read Professor Hall's paper, which so effectively analyzes the relative performance of Efron's many recipes for bootstrap interval construction, under the assumption that the parameter is a function of vector means. In this setting, Hall shows that the percentile-t and accelerated bias-corrected methods tie for first place, the main reason being that one consults "Studentized" tables, and the other looks up "ordinary" tables, after employing an analytical correction to adjust the critical points.

I share Hall's prejudice that computer-intensive methods such as the bootstrap should not have to appeal to tedious analytic corrections and therefore agree with his preference of the percentile-t over the accelerated bias-corrected method in the present situation. Because all the bias-corrected methods look up tables "backwards," the percentile-t may also be preferred in those nonlinear and nonsmooth problems where the asymptotic distributions are asymmetric, if we know how to Studentize. One such problem is discussed in Loh (1984).

On the other hand, I believe that the idea of looking up standard tables using adjusted levels has intrinsic merit on its own, and I will present a simple way of doing this which does not look up tables backwards and does not involve difficult analytic manipulations. It turns out that, under the "smooth" model of Hall, this method yields one-sided intervals that are second-order equivalent to the STUD and ABC methods and two-sided intervals that possess coverage errors which are an order of magnitude *smaller* than those of all the methods examined in the paper. Furthermore, it requires no more bootstrap sampling than the rest.

The method I propose has its origin in the "calibrated" method introduced in Loh (1987), and the basic idea is as follows. Starting with any reasonable interval procedure, bootstrap its coverage probability $\pi(\alpha)$. (This is a distinct departure from the other bootstrap recipes because the latter all call for bootstrapping the distribution of a statistic.) After the bootstrap estimate $\hat{\pi}(\alpha)$ is obtained, a corrected α^* is computed that is then used in place of α in the original formula.

[1]Work begun while visiting the T. J. Watson Research Center for a year as an IBM Junior Faculty Research Fellow. Research also partially supported by NSF Grant DMS-85-02303.

To illustrate, I will use the normal-theory confidence bound $\hat{\theta}_{\text{Norm}}(\alpha) \equiv \hat{\theta} + n^{-1/2}\hat{\sigma} z_\alpha$ as the initial procedure. The method may therefore be literally described as "looking up the wrong tables with corrected confidence levels." I will use Hall's notation throughout.

One-sided intervals. An expansion for the coverage probability $\pi_{\text{Norm}}(\alpha)$ of this confidence bound is given in Section 4.5 of the paper and its bootstrap estimate has the corresponding expansion

$$\hat{\pi}_{\text{Norm}}(\alpha) = \alpha - n^{-1/2}\hat{q}_1(z_\alpha)\phi(z_\alpha) + n^{-1}\hat{q}_2(z_\alpha)\phi(z_\alpha) + O_p(n^{-3/2}).$$

Let $\psi(\cdot)$ be any strictly increasing unbounded function on the unit interval with continuous third derivative and define

$$\begin{aligned}\delta_1 &\equiv \psi\{\hat{\pi}_{\text{Norm}}(\alpha)\} - \psi\{\alpha\} \\ &= -\{n^{-1/2}\hat{q}_1(z_\alpha) - n^{-1}\hat{q}_2(z_\alpha)\}\phi(z_\alpha)\psi'(\alpha) \\ &\quad + \tfrac{1}{2}n^{-1}\hat{q}_1^2(z_\alpha)\phi^2(z_\alpha)\psi''(\alpha) + O_p(n^{-3/2})\end{aligned}$$

to be the excess between estimated and true coverage on the ψ-scale.

Define the adjusted nominal level α^* to be the solution of the equation $\psi(\alpha^*) = \psi(\alpha) - \delta_1$. Then

$$\begin{aligned}\alpha^* &= \alpha + \phi(z_\alpha)\big[n^{-1/2}\hat{q}_1(z_\alpha) - n^{-1}\{\hat{q}_2(z_\alpha) + \hat{q}_1^2(z_\alpha)\phi(z_\alpha)\psi''(\alpha)/\psi'(\alpha)\}\big] \\ &\quad + O_p(n^{-3/2})\end{aligned}$$

and

$$\begin{aligned}z_{\alpha^*} &= z_\alpha + n^{-1/2}\hat{q}_1(z_\alpha) - n^{-1}\{\hat{q}_2(z_\alpha) + \hat{q}_1^2(z_\alpha)\phi(z_\alpha)\psi''(\alpha)/\psi'(\alpha) - \tfrac{1}{2}z_\alpha\hat{q}_1^2(z_\alpha)\} \\ &\quad + O_p(n^{-3/2}).\end{aligned}$$

This yields the adjusted confidence bound $\hat{\theta}_{\text{ADJ1}(\psi)}(\alpha) \equiv \hat{\theta}_{\text{Norm}}(\alpha^*)$, which has the expansion

$$\begin{aligned}&\hat{\theta} + n^{-1/2}\hat{\sigma}\big[z_\alpha + n^{-1/2}\hat{q}_1(z_\alpha) - n^{-1}\hat{q}_2(z_\alpha) \\ &\qquad + n^{-1}\{\tfrac{1}{2}z_\alpha - \phi(z_\alpha)\psi''(\alpha)/\psi'(\alpha)\}\hat{q}_1^2(z_\alpha)\big] + O_p(n^{-2}).\end{aligned}$$

Therefore, $\hat{\theta}_{\text{ADJ1}(\psi)}(\alpha)$ is second-order equivalent to $\hat{\theta}_{\text{STUD}}(\alpha)$ and $\hat{\theta}_{\text{ABC}}(\alpha)$, and the interval $(-\infty, \hat{\theta}_{\text{ADJ1}(\psi)}(\alpha)]$ has coverage

$$\begin{aligned}\pi_{\text{ADJ1}(\psi)}(\alpha) &= \alpha - n^{-1}\big[q_1(z_\alpha)\{q_1'(z_\alpha) + (\phi(z_\alpha)\psi''(\alpha)/\psi'(\alpha) - z_\alpha)q_1(z_\alpha)\} \\ &\qquad + u_q z_\alpha\big]\phi(z_\alpha) + O(n^{-3/2}).\end{aligned} \tag{1}$$

It follows that the one-sided normal-theory interval with α adjusted to α^* has coverage error of the same order as that of the ABC and STUD methods.

Note that if $\psi(\alpha) = \Phi^{-1}(\alpha) = z_\alpha$, then $\phi(z_\alpha)\psi''(\alpha)/\psi'(\alpha) - z_\alpha \equiv 0$ and (1) reduces to

$$\pi_{\text{ADJ1}(\Phi^{-1})}(\alpha) = \alpha - n^{-1}\{q_1(z_\alpha)q_1'(z_\alpha) + u_q z_\alpha\}\phi(z_\alpha) + O(n^{-3/2}).$$

Note also that except for terms of order n^{-1} and higher, the value of z_{α^*} is equal to the one-term Edgeworth-corrected critical point studied in Hall (1983, 1986).

EXAMPLE 1. NONPARAMETRIC MEAN. For the nonparametric estimation of the mean of a distribution, expansion (1) simplifies to

$$\pi_{\mathrm{ADJ1}(\psi)}(\alpha) = \alpha - n^{-1}\big[\gamma^2\{2z_\alpha/3 + (\phi(z_\alpha)\psi''(\alpha)/\psi'(\alpha) - z_\alpha)(2z_\alpha^2+1)/6\} + z_\alpha(\kappa - 3\gamma^2/2)\big](2z_\alpha^2+1)\phi(z_\alpha)/6 + O(n^{-3/2}).$$

If we choose ψ to be the function

$$\psi(\alpha) = \arctan\{\sqrt{2}\,\Phi^{-1}(\alpha)\}, \tag{2}$$

which is the solution of the equation

$$q_1'(z_\alpha) + \{\phi(z_\alpha)\psi''(\alpha)/\psi'(\alpha) - z_\alpha\}q_1(z_\alpha) = 0,$$

then $\pi_{\mathrm{ADJ1}(\arctan\{\sqrt{2}\Phi^{-1}\})}(\alpha) = \pi_{\mathrm{STUD}}(\alpha) + O(n^{-3/2})$ and $\hat\theta_{\mathrm{ADJ1}(\arctan\{\sqrt{2}\Phi^{-1}\})}(\alpha) = \hat\theta_{\mathrm{STUD}}(\alpha) + O_p(n^{-2})$.

EXAMPLE 2. EXPONENTIAL MEAN. In the parametric estimation of the exponential mean,

$$\pi_{\mathrm{ADJ1}(\psi)}(\alpha) = \alpha - (9n)^{-1}\big[(2z_\alpha^2+1)\{\phi(z_\alpha)\psi''(\alpha)/\psi'(\alpha) - z_\alpha\} + 4z_\alpha\big] \times (2z_\alpha^2+1)\phi(z_\alpha) + O(n^{-3/2}).$$

The order n^{-1} term vanishes if ψ is chosen to be the function in (2). Thus the coverage error of $\hat\theta_{\mathrm{ADJ1}(\arctan\{\sqrt{2}\Phi^{-1}\})}$ is an order of magnitude smaller than that of $\hat\theta_{\mathrm{ABC}}$.

Two-sided intervals. The equal-tailed two-sided interval

$$\big[\hat\theta_{\mathrm{ADJ1}(\psi)}(\alpha), \hat\theta_{\mathrm{ADJ1}(\psi)}(1-\alpha)\big]$$

has coverage error of order n^{-1}, which is the same order as that of the equal-tailed two-sided STUD, HYB, BACK, BC and ABC intervals.

Another way to obtain a calibrated two-sided interval is to adjust directly the value of α in the two-sided normal-theory interval $[\hat\theta_{\mathrm{Norm}}(\alpha), \hat\theta_{\mathrm{Norm}}(1-\alpha)]$. The formulas in Section 4.5 of the paper imply that

$$\hat\pi_{\mathrm{Norm}}(1-\alpha) - \hat\pi_{\mathrm{Norm}}(\alpha) = 1 - 2\alpha + 2n^{-1}\hat q_2(z_{1-\alpha})\phi(z_{1-\alpha}) + O_p(n^{-2}).$$

Now let ψ be any strictly increasing unbounded function on (0, 1) with continuous second derivative and define

$$\delta_2 \equiv \psi\{\hat\pi_{\mathrm{Norm}}(1-\alpha) - \hat\pi_{\mathrm{Norm}}(\alpha)\} - \psi\{1-2\alpha\} = 2n^{-1}\hat q_2(z_{1-\alpha})\phi(z_{1-\alpha})\psi'(1-2\alpha) + O_p(n^{-2}).$$

Let the adjusted level be $1 - 2\alpha^{**}$, where $\psi(1-2\alpha^{**}) \equiv \psi(1-2\alpha) - \delta_2$. Then

$$\alpha^{**} = \alpha + n^{-1}\hat q_2(z_{1-\alpha})\phi(z_{1-\alpha}) + O_p(n^{-2}).$$

TABLE 1
Length and coverage of two-sided 95% intervals in the nonparametric estimation of a mean.

Method	$s(z_{1-\alpha})$ (length)	$t(z_{1-\alpha})$ (coverage error)
ADJ1(arctan$\{\sqrt{2}\,\Phi^{-1}\}$)	$-0.14\kappa + 1.96\gamma^2 + 3.35$	$-2.84\kappa + 4.25\gamma^2$
ADJ1(Φ^{-1})	$-0.14\kappa + 0.064\gamma^2 + 3.35$	$-2.84\kappa + 2.36\gamma^2$
ADJ2(ψ)	$-0.14\kappa + 2.12\gamma^2 + 3.35$	0

Because $\hat{q}_2$ is an odd polynomial and $z_\alpha = -z_{1-\alpha}$, we have for *all* ψ,

$$z_{\alpha^{**}} = z_\alpha - n^{-1}\hat{q}_2(z_\alpha) + O_p(n^{-2}).$$

Therefore, if we define $\hat{\theta}_{\mathrm{ADJ2}(\psi)}(\alpha) \equiv \hat{\theta}_{\mathrm{Norm}}(\alpha^{**})$,

$$\hat{\theta}_{\mathrm{ADJ2}(\psi)}(\alpha) = \hat{\theta} + n^{-1/2}\hat{\sigma}\left[z_\alpha - n^{-1}\hat{q}_2(z_\alpha)\right] + O_p(n^{-2})$$

and the interval

$$\left[\hat{\theta}_{\mathrm{ADJ2}(\psi)}(\alpha), \hat{\theta}_{\mathrm{ADJ2}(\psi)}(1-\alpha)\right] \qquad (3)$$

has coverage probability $1 - 2\alpha + O(n^{-2})$, which incurs an error of a *smaller* order than that of any of the intervals considered in the paper. The length of (3) exceeds that of the corresponding normal-theory interval by the amount $-2n^{-3/2}\hat{\sigma}q_2(z_{1-\alpha}) + O_p(n^{-2})$.

REMARK. Two-term Edgeworth-corrected statistics like those in Hall (1983) or Abramovitch and Singh (1985) can also achieve the same order of improvement in the coverage probability of two-sided intervals as $\hat{\theta}_{\mathrm{ADJ2}(\psi)}$. The expansion for $z_{\alpha^{**}}$ however implies that a *one-term* correction suffices, if that term is chosen appropriately.

EXAMPLE 1 (continued). Table 1 supplements Hall's Table 4.1 and shows the differences in length and coverage errors of the proposed intervals for the nonparametric estimation of the mean when $\alpha = 0.025$. Because $q_1(x)$ vanishes when $\gamma = 0$, ADJ2 and the "ideal" interval Stud are equal in length to order n^{-2} when F is symmetric.

EXAMPLE 2 (continued). Table 2 supplements Hall's Table 4.2 for the parametric estimation of an exponential mean. It turns out that when ψ is defined by

TABLE 2
Length and coverage of two-sided 95% intervals in the parametric estimation of an exponential mean.

Method	$s(z_{1-\alpha})$ (length)	$t(z_{1-\alpha})$ (coverage error)
ADJ1(arctan$\{\sqrt{2}\,\Phi^{-1}\}$)	3.64	0
ADJ2(ψ)	4.29	0

(2), $\hat{\theta}_{\mathrm{ADJ1}(\psi)}(\alpha)$ is *third-order* equivalent to both $\hat{\theta}_{\mathrm{STUD}}(\alpha)$ and $\hat{\theta}_{\mathrm{ABC}}(\alpha)$ in the sense that their expansions match up to and including the term of order $n^{-3/2}$.

REFERENCES

ABRAMOVITCH, L. and SINGH, K. (1985). Edgeworth corrected pivotal statistics and the bootstrap. *Ann. Statist.* **13** 116–132.

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

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

LOH, W.-Y. (1984). Estimating an endpoint of a distribution with resampling methods. *Ann. Statist.* **12** 1543–1550.

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

DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN
1210 WEST DAYTON STREET
MADISON, WISCONSIN 53706

J. ROBINSON

University of Sydney

My major comment concerns the relative importance of the approximation of the critical points and of coverage error. It appears to me that much greater emphasis should be placed on the accuracy of the approximation of the bootstrap critical points to the theoretical points. The theoretical critical points based on θ should have been chosen as the best ones, in the nontechnical sense that they are thought to be better than any others available and the interval based on these has exact coverage α. Then, because we are not, in fact, able to find these critical points, we need an approximation; this can be provided by the bootstrap. Then we need to examine first the closeness of the approximating confidence interval to the theoretical one which we would use if we could. Finally, the coverage error for the approximation can be examined.

The point is more strongly made with reference to bootstrap simulations, which are not directly referred to here, the assumption being in this work that the number of simulations B, say, is very large. There is a discussion of them in Hall (1986), where it is shown that even, for small B, say 19 for a 95% one-side interval, the coverage error is of the same order as if we simulated an infinite number of times. However, the accuracy of the approximation to the simulated critical point in the Studentized case is of order $n^{-1/2}B^{-1/2}$ compared to an accuracy for the infinitely resampled bootstrap of order $n^{-3/2}$. So B must be at least of size n^2 to make these approximations comparable. The reason for the accuracy of the coverage is that an averaging over all possible bootstrap simulations of size B has taken place in its calculation. Thus the particular approximation based on B simulations, compared to the real bootstrap approximation, may be gravely in error, although an average of these errors, taken over an inappropriate set, is small. I believe that in this case it is apparent that the accuracy