This procedure is distribution-free in the sense that it does not depend upon the shape of the parent distribution(s) from which the observations were drawn. It does, of course, depend upon the (necessarily symmetric) permutation distribution generated from the sample data.

However, the value being estimated is a population value and, whatever procedure is used, random sampling is desirable. With matched pairs, if the two members of each pair were selected in the same way and given equal probabilities of the two possible assignments to treatments, it may seem plausible to treat the sample of differences as more representative of the (hypothetical) population of differences than the two individual samples are of their parent populations.

Permutation intervals are obtainable for parameters other than the mean. All that is necessary is that they can be estimated by sample statistics having unique values with each permutation of the data. Thus intervals can be constructed for variances and medians, but not for modes. The search for the confidence limit again renders the method inefficient in more than one dimension. If there is a single parameter of interest, as before we might choose to condition on the estimates of the nuisance parameters.

## REFERENCES

GARTHWAITE, P. H. and BUCKLAND, S. T. (1988). Generating Monte Carlo confidence intervals by the Robbins–Monro process. Unpublished.

ROBBINS, H. and MONRO, S. (1951). A stochastic approximation method. *Ann. Math. Statist.* **22** 400–407.

STEPHEN T. BUCKLAND
SASS, ABERDEEN UNIT
MACAULAY LAND USE RESEARCH INSTITUTE
CRAIGIEBUCKLER
ABERDEEN AB9 2QJ
UNITED KINGDOM

PAUL H. GARTHWAITE AND HOWARD G. LOVELL
DEPARTMENT OF STATISTICS
UNIVERSITY OF ABERDEEN
EDWARD WRIGHT BUILDING
DUNBAR STREET
OLD ABERDEEN AB9 2TY
UNITED KINGDOM

THOMAS J. DICICCIO AND JOSEPH P. ROMANO

*Stanford University*

We congratulate Hall for a most stimulating paper. Hall has presented bootstrappers with a useful framework within which to compare resampling methods.

Before getting to the main topic of our discussion, we would like to raise two issues involving smoothing and uniformity. Outside of the obvious problem that a stable estimate of variance may be difficult to obtain, one may question the extent to which the results contribute to a complete theory of confidence intervals. It is known, for example, that outside of the "smooth function" model, the percentile-$t$ method may be inconsistent. For instance, bootstrap confidence intervals for functionals of a density (based on percentile-$t$ or other proposed methods) will generally be inconsistent unless resampling is performed from an appropriately smooth distribution; see Romano (1986). In other problems, even if the percentile-$t$ is consistent, smoothing may conceivably lower coverage error. In summary, the issue of smoothing before resampling deserves investigation, even though the benefit of smoothing falls outside of the smooth function model. Also, see Silverman (1987). A second point that may deserve further attention is the uniformity of coverage errors of various bootstrap confidence intervals over appropriately broad classes of distributions. Indeed, if the level of a confidence interval $\hat{I}$ is defined to be the infimum over all $F$ of the probability under $F$ that $\hat{I}$ contains $\theta(F)$, then uniform results in $F$ of coverage errors are required in comparing the level of various procedures. Of course, lack of uniformity in bootstrap procedures is sometimes unavoidable in nonparametric problems; see Bahadur and Savage (1956).

Within the smooth function model, Hall provides strong evidence against the use of certain resampling techniques. Emerging as the two most promising techniques are the percentile-$t$ and Efron's acccelerated bias-corrected ($\mathrm{BC}_a$) method. Hall prefers the percentile-$t$ over the $\mathrm{BC}_a$ due to the following two criticisms of the $\mathrm{BC}_a$. First, the $\mathrm{BC}_a$ requires calculation of an analytical adjustment, namely the acceleration constant, which automatic bootstrap methods should ideally avoid. Second, for any given sample, the $\mathrm{BC}_a$ intervals shrink to a point as coverage increases. We wish to describe an alternative procedure which grew out of an attempt to reconstruct the $\mathrm{BC}_a$ limits without having to explicitly calculate $a$ and does not suffer from these same criticisms.

**1. Scalar parameter models.** Suppose the model is parametrized by a scalar parameter $\theta$ and the interest lies in constructing an upper $1-\alpha$ confidence limit for $\theta$ based on an estimator $\hat{\theta}$. Let $G_\theta(\cdot)$ be the distribution function of $\hat{\theta}$ under $\theta$. To motivate the method, suppose that $G_{\hat{\theta}}(\theta)$ is a pivot; that is, its distribution is independent of $\theta$ with $1-\alpha$ quantile denoted $u_{1-\alpha}$. For simplicity, assume that $G_\theta(t)$ is strictly increasing in $t$ and strictly decreasing in $\theta$. Then,

$$P_\theta\{G_{\hat{\theta}}(\theta) \le u_{1-\alpha}\} = P_\theta\{\theta \le G_{\hat{\theta}}^{-1}(u_{1-\alpha})\} = 1-\alpha.$$

Hence, $G_{\hat{\theta}}^{-1}(u_{1-\alpha})$ is an upper $1-\alpha$ confidence limit for $\theta$. All that is needed is a method to compute $u_{1-\alpha}$. To do this, fix any $\theta_0$ and simulate the distribution of $G_{\hat{\theta}}(\theta_0)$ under $\theta_0$ and compute its $1-\alpha$ quantile. Let $h(\hat{\theta}) = G_{\hat{\theta}}(\theta_0)$ and let $\theta_0' = G_{\theta_0}^{-1}(\alpha)$. Then the upper $1-\alpha$ quantile of $h(\hat{\theta})$ under $\theta_0$ is $h(\theta_0') = G_{\theta_0'}(\theta_0)$, since $h$ is monotone decreasing. Hence, the upper $1-\alpha$ limit for $\theta$ is $\theta_1$, where $\theta_1$ is defined by

$$\theta_1 = G_{\hat{\theta}}^{-1}\{G_{\theta_0'}(\theta_0)\}. \tag{1}$$

The assumption that $G_{\hat{\theta}}(\theta)$ is pivotal holds, for example, when for some monotone transformation $g$ and constants $a$ and $z_0$, the distribution of

$$\frac{g(\hat{\theta}) - g(\theta)}{1 + ag(\theta)} + z_0 \tag{2}$$

is independent of $\theta$. Efron actually assumes the distribution of (2) is standard Gaussian to motivate the $\mathrm{BC}_a$. Several equivalent necessary and sufficient conditions for the exactness of the upper limit given by (1) are given in DiCiccio and Romano (1987). For example, (1) is exact in the exponential mean example considered by Hall.

Typically, $\theta_1$ given by (1) is an approximate limit. In practice, a reasonable choice for the initial value $\theta_0$ is the percentile limit or the BC limit. Note that if the exact upper limit $\theta_U$ is used as $\theta_0$, where $\theta_U$ satisfies $G_{\theta_U}(\hat{\theta}) = \alpha$, then $\theta_1 = \theta_U$ is obtained. Generally, $\theta_1$ is second-order correct in the sense of Efron (1987) and the coverage error is of order $n^{-1}$. No analytical adjustments are required in (1), but the distribution of $\hat{\theta}$ must be obtained at three values: $\hat{\theta}$, $\theta_0$ and $\theta_0'$. By choosing $\theta_0 = \hat{\theta}$, this number can be reduced to two. The procedure lends itself to iteration and it can be shown that, under certain conditions involving the existence of Edgeworth expansions, each round of iteration brings the approximate limit closer to $\theta_U$ and brings the coverage probability closer to $1 - \alpha$ by an order of $n^{-1/2}$.

**2. Multiparameter families.** The procedure given by (1) can be extended to set approximate confidence limits for a scalar parameter in multiparameter families. Consider a family of distributions indexed by the vector parameter $\lambda = (\lambda^1, \dots, \lambda^p)$ and suppose that the scalar parameter $\theta = f(\lambda)$ is of interest. Let $\hat{\lambda}$ be an estimator of $\lambda$ and let $\hat{\theta} = f(\hat{\lambda})$ be the corresponding estimator of $\theta$. The distribution function of $\hat{\theta}$ is given by $G_\lambda(s) = P_\lambda(\hat{\theta} \le s)$ and the bootstrap distribution for $\hat{\theta}$ is $G_{\hat{\lambda}}$.

First, suppose the model has been parametrized so that $\lambda^1 = \theta$ and $\lambda^2, \dots, \lambda^p$ are orthogonal to $\theta$, i.e., suppose that

$$\operatorname{cov}\left(n^{1/2}\hat{\theta}, n^{1/2}\hat{\lambda}^i\right) = O(n^{-1}), \qquad i = 2, \dots, p.$$

Let $\nu = (\lambda^2, \dots, \lambda^p)$. Then (1) can be adapted to find an approximate upper $1 - \alpha$ confidence limit for $\theta$ by the following procedure. Start with an initial value $\theta_0$ of $\theta$, perhaps the percentile limit $G_{\hat{\lambda}}^{-1}(1 - \alpha)$, set $\theta_0' = G_{(\theta_0, \hat{\nu})}^{-1}(\alpha)$ and finally take

$$\theta_1 = G_{(\hat{\theta}, \hat{\nu})}^{-1}\left\{G_{(\theta_0', \hat{\nu})}(\theta_0)\right\}.$$

In many situations, it is difficult to transform a given model to achieve an orthogonal parametrization. For such cases, the least favorable family construction used by Efron (1987) allows the procedure to be implemented in terms of the original parameters. Suppose that

$$\operatorname{cov}\left(n^{1/2}\hat{\lambda}^i, n^{1/2}\hat{\lambda}^j\right) = \kappa^{i,j} + O(n^{-1}), \qquad i, j = 1, \dots, p,$$

and let $f_i = \partial f / \partial \lambda^i$ and $\mu^i = \Sigma \kappa^{i,j} f_j$. Consider the line in the parameter space, parametrized by $t$, given by $\lambda(t) = \hat{\lambda} + t\hat{\mu}$, where $\hat{\mu} = (\hat{\mu}^1, \dots, \hat{\mu}^p)$ and $\mu^i = \hat{\mu}^i(\hat{\lambda})$. Then $\hat{\lambda} = \lambda(0)$ and the bootstrap distribution for $\hat{\theta}$ is $G_{\lambda(0)}$. To implement the procedure, commence with an initial value $\theta_0$ for $\theta$, choose $t_0$ to satisfy $\theta_0 =$

$f\{\lambda(t_0)\}$ and let $\theta_0' = G^{-1}_{\lambda(t_0)}(\alpha)$. The approximate limit for $\theta$ is

$$\theta_1 = G_{\hat{\lambda}}^{-1}\left\{G_{\lambda(t_0')}(\theta_0)\right\},$$

where $t_0'$ satisfies $\theta_0' = f\{\lambda(t_0')\}$. It can be shown that the error in the coverage level of $\theta_1$ is $O(n^{-1})$; in fact, $\theta_1$ differs from Hall's theoretical limit $\theta_{\text{Stud}}$ by order $n^{-3/2}$ in probability. However, because of the presence of nuisance parameters, the order of this error cannot be reduced by further iteration.

**3. The nonparametric case.** Consider the problem of constructing a confidence limit for a functional $\theta = \theta(F)$ of the unknown distribution $F$ in a nonparametric setting. The normal estimate of $\theta(F)$ is $\theta(\hat{F}_n)$, where $\hat{F}_n$ is the empirical distribution based on a sample $X_1, \dots, X_n$ from $F$. As in Efron (1987), we reduce the nonparametric problem to the scalar parameter case by considering an appropriate one-dimensional least favorable subfamily of distributions, constructed as follows. Let

$$U_i = \lim_{\Delta \to 0} \frac{\theta\left((1-\Delta)\hat{F}_n + \Delta\delta_i\right) - \theta\left(\hat{F}_n\right)}{\Delta}, \qquad i = 1, 2, \dots, n,$$

where $\delta_i$ is the distribution assigning mass 1 to $X_i$. Let $F_t$ denote the distribution assigning mass

$$p_i(t) = \exp(tU_i) \Big/ \sum_{j=1}^{n} \exp(tU_j)$$

to $X_i$. This family of distributions, indexed by $t$, is the least favorable family; see Efron (1987). Now, set $\theta(t) = \theta(F_t)$. Also, let $G_t(s)$ be the distribution function of $\hat{\theta}$ under $F_t$ (conditional on the data). To construct the nonparametric version of (1), let $\theta_0$ be an initial choice for an upper $1 - \alpha$ confidence limit. Let $t_0$ be the value of $t$ satisfying $\theta(t) = \theta_0$, let $\theta_0' = G_{t_0}^{-1}(\alpha)$ and let $t_0'$ be the value of $t$ satisfying $\theta(t) = \theta_0'$. The proposed limit is then

$$\theta_1 = G_0^{-1}\{G_{t_0'}(\theta_0)\}. \tag{3}$$

As in the multiparameter case, commencing with a reasonable choice of $\theta_0$ [so that $\theta_0$ is within $\theta(F)$ by order $n^{-1/2}$ in probability] results in a second-order accurate limit $\theta_1$; that is, $\theta_1$ differs from $\theta_{\text{Stud}}$ by order $n^{-3/2}$ in probability. This result holds under Hall's smooth function model and will be proved in a forthcoming technical report. As in the parametric case, one may iterate the procedure by starting with $\theta_1$ to yield a new limit $\theta_2$ and so on. Unfortunately, this does not reduce the error of the procedure outside of the scalar parameter case.

In the case of nonparametric confidence intervals for the mean, $\theta_i$ approaches the limit $\theta_{\text{TILT}}$ obtained by Efron's (1981) nonparametric tilting method as the number of iterations increases. Specifically, if $t_U$ is the value of $t$ satisfying $\hat{\theta} = G_t^{-1}(\alpha)$, then $\theta_{\text{TILT}} = \theta(t_U)$. We have obtained expansions for $\theta_1$ and $\theta_{\text{TILT}}$ corresponding to Hall's analysis in Sections 4.3 and 4.5. The expansions for $\theta_1$ to third order depend on the choice of $\theta_0$. Assuming we have started with a "good" choice for $\theta_0$ [or that we have iterated (3) once more], the third-order properties of $\theta_1$ and $\theta_{\mathrm{TILT}}$ are the same. In particular, the upper $\alpha$ nonparametric tilting limit satisfies

$$\hat{\theta}_{\mathrm{TILT}}(\alpha) = \hat{\theta} + n^{-1/2}\hat{\sigma}\big[z_\alpha + n^{-1/2}\tfrac{1}{6}\hat{\gamma}(2z_\alpha^2 + 1) + n^{-1}z_\alpha\big\{\tfrac{1}{72}\gamma^2(-7z_\alpha^2 + 4) + \tfrac{1}{24}\kappa(3z_\alpha + 1)\big\}\big] + O_p(n^{-2}).$$

Furthermore, the coverage probability associated with $\hat{\theta}_{\mathrm{TILT}}$ satisfies

$$\pi_{\mathrm{TILT}}(\alpha) = \alpha - n^{-1}\phi(z_\alpha)z_\alpha(z_\alpha^2 + 3)\tfrac{1}{8}(\kappa - \gamma^2 + 2) + O(n^{-3/2})$$

and, corresponding to Table 1,

$$t(z_{1-\alpha}) = -1.68\kappa + 1.68\gamma^2 - 3.35.$$

## REFERENCES


BAHADUR, R. and SAVAGE, L.J. (1956). The nonexistence of certain statistical procedures in nonparametric problems. *Ann. Math. Statist.* **27** 1115–1122.

DICICCIO, T. and ROMANO, J. (1987). Accurate bootstrap confidence limits. Technical Report 281, Dept. Statistics, Stanford Univ.

EFRON, B. (1981). Nonparametric standard errors and confidence intervals (with discussion). *Canad. J. Statist.* **9** 139–172.

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

ROMANO, J. (1986). On bootstrapping the joint distribution of the location and size of the mode. Ph.D. dissertation, Dept. Statistics, Univ. California, Berkeley.

SILVERMAN, B. (1987). The bootstrap: To smooth or not to smooth? *Biometrika* **74** 469–479.



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



## B. EFRON

*Stanford University*


This is the latest in Peter Hall's series of impressive bootstrap papers. One effect of these papers [and those by other authors, in particular Abramovitch and Singh (1985)] has been to renew interest in bootstrap-$t$ confidence intervals. My original enthusiasm for bootstrap-$t$ intervals, as naively expressed in Remark F of Efron (1979) and slightly less naively in Section 10.10 of Efron (1982), foundered on a list of their substantial drawbacks: noninvariance under transformations, occasional numerical instability and, worst of all, the need to compute auxiliary estimates of standard deviation $\hat{\sigma}$ and $\hat{\sigma}^*$. The good properties demonstrated in this paper and others make it worthwhile to pursue the practical details of applying the bootstrap-$t$ method on a routine basis.

Figure 1 concerns "looking up tables backwards." It is natural to assume that if an error distribution is long-tailed to the right, then the corresponding