The Annals of Statistics
1987, Vol. 15, No. 4, 1713–1718

ON A PARTIAL CORRECTION BY THE BOOTSTRAP[1]

By Regina Y. Liu and Kesar Singh

Rutgers University

The phenomenon of partial $n^{-1/2}$-term correction by the bootstrap in the estimation of sampling distributions of nonstandardized statistics is explained and studied in this note.

1. Partial and total correction of $n^{-1/2}$ terms. Let $X_1, \dots, X_n$ be independent and identically distributed (i.i.d.) random variables with a common distribution F. Let $T_n \equiv T(X_1, \dots, X_n)$ be an estimator of T_F, a parameter of interest which depends on F. The sampling distribution of T_n is needed in order to make inferences about T_F. In this paper, some aspects of the bootstrap approximation of the sampling distribution of T_n are studied.

When the one-term Edgeworth expansion of a sampling distribution has the form $\Phi(x) + n^{-1/2}p(x, F)\phi(x) + o(n^{-1/2})$, Φ is the c.d.f. for the standard normal distribution and $p(x, F)$ is a polynomial in x and some functions of the moments of F, the bootstrap approximation is typically correct up to $o(n^{-1/2})$. This phenomenon is well documented in the literature [see Bickel and Freedman (1981), Singh (1981), Babu and Singh (1983), Babu and Singh (1984) and Abramovitch and Singh (1985)]. We shall refer to this phenomenon as the *total $n^{-1/2}$-term correction by the bootstrap*. This is the type of expansions one has for the standardized statistics, namely, the normalized case [$\sqrt{n}(T_n - T_F)$ divided by the true standard deviation] and the studentized case [$\sqrt{n}(T_n - T_F)$ divided by an estimated standard deviation]. Consider now the following nonstandardized case where the leading term in Edgeworth expansion involves the underlying population:

$$H_n(x) \equiv P\big(\sqrt{n}(T_n - T_F) \le x\big)$$
$$= \Phi(x/S_F) + n^{-1/2}p(x, F)\phi(x/S_F) + o(n^{-1/2}).$$

Typically, $\operatorname{Var} T_n = S_F^2/n + O(n^{-2})$. The corresponding bootstrap Edgeworth expansion is

$$H_{n,\mathrm{B}}(x) = P^*\big(\sqrt{n}(T_n^* - T_n) \le x\big)$$
$$= \Phi\big(x/S_{F_n}\big) + n^{-1/2}p(x, F_n)\phi\big(x/S_{F_n}\big) + o(n^{-1/2}) \quad \text{a.s.},$$

where P^* stands for the bootstrap probability, F_n stands for the empirical distribution based on the original sample and T_n^* stands for the functional T

Received July 1986; revised March 1987.
[1]Research supported by NSF grant DMS-85-02945.
AMS 1980 *subject classifications*. Primary 62G10; secondary 62G15.
Key words and phrases. Bootstrap, total correction, partial correction.

based on the bootstrap sample. Generally, $|p(x, F) - p(x, F_n)| \to 0$ in probability; thus,

$$P\big(\sqrt{n}\,(T_n - T_F) \le x\big) - P^*\big(\sqrt{n}\,(T_n^* - T_n) \le x\big)$$
$$= \Phi(x/S_F) - \Phi\big(x/S_{F_n}\big) + o(n^{-1/2}) \quad \text{a.s.}$$

A natural estimator for $H_n(x)$ based on the CLT is $\Phi(x/S_{F_n})$. Thus, if $\sqrt{n}\,(S_{F_n}^2 - S_F^2) \to_{\mathscr{L}} N(0, a^2)$ for some a, then

$$\sqrt{n}\,(H_n(x) - H_{n,\mathrm{B}}(x)) \to_{\mathscr{L}} N(0, b^2)$$

and

$$\sqrt{n}\,\big(H_n(x) - \Phi(x/S_{F_n})\big) \to_{\mathscr{L}} N\big(p(x, F)\phi(x/S_F), b^2\big),$$

where $b^2 = \phi^2(x/S_F)x^2a^2/S_F^6$. Thus, we see that the bootstrap estimator is superior to $\Phi(x/S_{F_n})$ in terms of the asymptotic mean squared error (m.s.e.), even though the rate of convergence for both estimators is $n^{-1/2}$. The amount of the reduction in the m.s.e. $[= p^2(x, F)\phi^2(x/S_F)]$ due to the bootstrap generally depends on the bias in T_n as an estimator of T_F and the skewness in the sampling distribution of T_n. For example, if T_n is the sample mean $\overline{X}$, then $a^2 = \mu_4 - \sigma^4$ and $p(x, F) = -(\mu_3/6\sigma^3)(x^2 - 1)$, where μ_i is the ith central moment and σ^2 is the variance of F. This fact was first observed in Beran [(1982), Remark (e), page 218]. It appears that his remark did not receive the attention it deserved. For instance, Hartigan (1986) states that, in the nonstudentized cases, the bootstrap estimator has no advantage over the estimator based on the normal approximation, since the convergence rates are the same. We shall refer to this correction of $n^{-1/2}p(x, F)\phi(x/S_F)$ term as *the partial $n^{-1/2}$-term correction by the bootstrap*.

The total $n^{-1/2}$-term correction improves the accuracy of the estimated p-value as well as the achieved α in the following sense. In constructing a two-sided confidence interval, the difference between the actual lower (also the upper) tail probability and $\alpha/2$ is reduced to $o(n^{-1/2})$ by the bootstrap compared to $O(n)^{-1/2}$ from the normal approximation. However, the total tail probability from the normal approximation is $\alpha + o(n^{-1/2})$ due to the cancellation of the $O(n)^{-1/2}$ term in the Edgeworth expansion. The extra accuracy of the bootstrap, therefore, is in terms of the following: (i) distributing the error probability more evenly in the two tails of the two-sided confidence intervals and (ii) bringing the estimated tail probability closer to the target in the case of one-sided confidence intervals (which are needed for hypothesis testing). These results on α in the case of total correction by the bootstrap are in Hall (1986) and Babu and Bose (1986). The claim for the p-value can be established easily. Our main aim in this note is to see if the preceding partial correction by the bootstrap brings any improvement in the estimation of p-values in hypothesis testing problems and in the coverage probabilities of bootstrap confidence intervals. We shall see in the following sections that there is an affirmative answer for the case of p-values and a negative answer for α.

We conclude this section with two remarks on the partial correction:

REMARK 1. From a one-term Edgeworth expansion for a sampling distribution, one immediately obtains a one-term expansion for a quantile of the sampling distribution. It clearly follows that the partial correction improves the estimate of a quantile of the sampling distribution in the sense of reducing the asymptotic m.s.e.

REMARK 2. For studentized statistics, there exists a phenomenon of partial n^{-1}-term correction by the bootstrap. It follows along the lines of Beran's remark that for studentized statistics the bootstrap is even superior to a one-term Edgeworth expansion (after replacing the population moments with the sample moments) in reducing the asymptotic m.s.e.

2. Partial correction and *p*-values. If $E(T_n) = T_F + o(n^{-1})$, the one-term Edgeworth expansion for $H_n(x)$ has the form

$$H_n(x) = \Phi\left(\frac{x}{S_F}\right) - \frac{\kappa_3}{6S_F^3\sqrt{n}}\left(\frac{x^2}{S_F^2} - 1\right)\phi\left(\frac{x}{S_F}\right) + o(n^{-1/2}), \tag{1}$$

where $\kappa_3/\sqrt{n}$ is the leading term in the cumulant of $\sqrt{n}(T_n - T_F)$. For testing an hypothesis with H_0: $T_F = T_0$ versus H_1: $T_F > T_0$, one can use the two estimates of the sampling distribution H_n, $\Phi(x/S_{F_n})$ and $H_{n,\mathrm{B}}$, to obtain approximate p-values, respectively,

$$\hat{p}_{\mathrm{CLT},n} = 1 - \Phi\big(\sqrt{n}(T_n - T_0)/S_{F_n}\big)$$

and

$$\hat{p}_{\mathrm{B},n} = 1 - H_{n,\mathrm{B}}\big(\sqrt{n}(T_n - T_0)\big).$$

The true p-value p_n based on $H_n(x)$ is $H_n(\sqrt{n}(T_n - T_0))$. (Hereafter, for simplicity, the subscript n in $\hat{p}_{\mathrm{CLT},n}$, $\hat{p}_{\mathrm{B},n}$ and p_n is suppressed.) We present a theorem now which establishes the fact that $\hat{p}_{\mathrm{B}}$ is closer to p than $\hat{p}_{\mathrm{CLT}}$ is to p in an asymptotic sense.

THEOREM. *Let* (1) *hold. Under* H_0, *if* $(\sqrt{n}(T_n - T_0), \sqrt{n}(S_{F_n}^2 - S_F^2))$ *has a bivariate normal weak limit with means* 0 *and if* κ_{3,F_n} *converges to* $\kappa_{3,F}$ *in probability, then* $\sqrt{n}(\hat{p}_{\mathrm{B}} - p)$ *and* $\sqrt{n}(\hat{p}_{\mathrm{CLT}} - p)$ *have weak limits. If* ξ *and* η *denote the two weak limits, then* $E(\xi^2) = A^2$ *and* $E(\eta^2) = A^2 + B^2$ *for some numbers A and B.*

PROOF. Let (Z, W) denote the bivariate normal weak limit of $(\sqrt{n}(T_n - T_0)/S_F, \sqrt{n}(S_{F_n}^2 - S_F^2))$. Both Z and W have mean 0; Z has variance 1. Using (1) and some Taylor expansions, one can easily check that

$$\sqrt{n}(\hat{p}_{\mathrm{B}} - p) \to_{\mathscr{L}} -\frac{ZW}{2S_F^2}\phi(Z)$$

and

$$\sqrt{n}(\hat{p}_{\mathrm{CLT}} - p) \to_{\mathscr{L}} -\frac{ZW}{2S_F^2}\phi(Z) - \frac{\kappa_{3,F}}{6S_F^3}(Z^2 - 1)\phi(Z).$$

The result clearly follows if we show that $E[ZW\phi(Z)][(Z^2-1)\phi(Z)]=0$. Since $E(W|Z)=c_0Z$ for some constant c_0, it suffices to show that

$$E(Z^4-Z^2)\phi^2(Z)=0 \quad \text{i.e.,} \quad \int(z^4-z^2)\phi^3(z)\,dz=0.$$

This is clearly so in view of the fact that $\phi^3(Z)$ is proportional to the normal density with mean 0 and variance $\frac{1}{3}$. □

We must admit that the preceding proof does not seem to provide much insight on the validity of such a result. The form (1) for the expansion seems crucial. It is not clear at the moment if there is any extension of the result if the bias of T_n is $O(n^{-1})$, in which case, of course, the expansion is different from (1). We conclude this section with the comment that, under moderate conditions, the expansion (1) holds in particular for the class of U-statistics [see Callaert, Janssen and Vevaverbeke (1980) and Bickel, Götze and van Zwet (1986)]. The expansion is also expected to hold for all (n^{-1}-term) bias corrected statistics, in particular jackknife statistics.

3. Partial correction and α. A natural question about partial correction is whether the confidence intervals based on $H_{n,\mathrm{B}}$ are better in terms of coverage probabilities than the ones based on $\Phi(x/S_{F_n})$. It is known that in the case of total correction, each tail error probability is $\alpha/2+o(n^{-1/2})$ under the bootstrap, whereas if the normal approximation is used for forming the confidence intervals, the tail probabilities are $\alpha/2+O(n^{-1/2})$ [see Hall (1986) and Babu and Bose (1986)]. Thus, the total correction implies a more even distribution of the error probability in the two tails. In order to see if such a result holds even for the partial correction, we expand the error probabilities in the cases of partial correction and normal approximation when $T_n=\overline{X}$ and $T_F=\mu$. The confidence interval based on the normal approximation is

$$\left(\overline{X}-z_{1-\alpha/2}S_n/\sqrt{n},\ \overline{X}+z_{1-\alpha/2}S_n/\sqrt{n}\right),$$

where S_n is the sample standard deviation and $z_\alpha=\Phi^{-1}(\alpha)$. The confidence interval based on $H_{n,\mathrm{B}}$ is

$$\left(\overline{X}-H_{n,\mathrm{B}}^{-1}(1-\alpha/2)/\sqrt{n},\ \overline{X}-H_{n,\mathrm{B}}^{-1}(\alpha/2)/\sqrt{n}\right).$$

It turns out that

$$(2)\quad P\left(\mu<\overline{X}-\frac{z_{\alpha/2}S_n}{\sqrt{n}}\right)=\alpha/2-\frac{\mu_3}{6\sigma^3\sqrt{n}}\left(2z_{\alpha/2}^2+1\right)\phi(z_{\alpha/2})+o(n^{-1/2})$$

and

$$(3)\quad \begin{aligned}&P\left(\mu<\overline{X}-H_{n,\mathrm{B}}^{-1}(1-\alpha/2)/\sqrt{n}\right)\\&\quad=\frac{\alpha}{2}-\frac{\mu_3}{2\sigma^3\sqrt{n}}z_{\alpha/2}^2\phi(z_{\alpha/2})+o(n^{-1/2}),\end{aligned}$$

under moderate conditions on the population. The same expansion holds for the upper one-sided probabilities (with different signs before the $n^{-1/2}$ term). Thus, if $z_{1-\alpha/2} > 1$,

$$|P(\mu < \overline{X} - H_{n,\mathrm{B}}^{-1}(1-\alpha/2)/\sqrt{n}) - \alpha/2|$$

$$= \left| P\left(\mu < \overline{X} - \frac{z_{\alpha/2} S_n}{\sqrt{n}}\right) - \frac{\alpha}{2} \right| + (z_{\alpha/2}^2 - 1)\frac{|\mu_3|}{6\sigma^3\sqrt{n}}\phi(z_{\alpha/2}) + o(n^{-1/2}).$$

Hence, it follows that the confidence interval based on the normal approximation in fact distributes the error probability more evenly if $z_{1-\alpha/2} > 1$, which typically is the case. The partial correction, therefore, does not seem to lead to better confidence intervals in terms of each tail error probability. On the other hand, the total coverage probabilities for both the confidence intervals are $(1-\alpha) + o(n^{-1/2})$. Incidentally, a more even distribution of α in the two sides also means that the achieved level of significance would be closer to that intended if the one-sided confidence interval is to be used for a one-sided hypothesis testing problem.

We conclude this note by sketching the proofs for the expansions (2) and (3). The expansion (2) is an easy consequence of the one-term Edgeworth expansion for the studentized mean,

$$(4) \qquad P\left(\sqrt{n}\,\frac{\overline{X}-\mu}{S_n} \le x\right) = \Phi(x) + \frac{\mu_3}{6\sigma^3\sqrt{n}}(2x^2+1)\phi(x) + o(n^{-1/2}).$$

One set of sufficient conditions for (4) to hold is (a) F is continuous and (b) F has a finite sixth moment. To establish (3), one requires

$$(5) \quad P^*(\sqrt{n}(\overline{Y} - \overline{X}) \le x) = \Phi\left(\frac{x}{S_n}\right) - \frac{\mu_3}{6\sigma^3\sqrt{n}}\left(\frac{x^2}{S_n^2} - 1\right)\phi\left(\frac{x}{S_n}\right) + \gamma_n(x) \quad \text{a.s.},$$

where

$$(6) \qquad P\left(\sup_x |\gamma_n(x)| > \varepsilon n^{-1/2}\right) = o(n^{-1/2}) \quad \text{for any } \varepsilon > 0.$$

$\overline{Y}$ stands for the bootstrap sample mean based on $Y_1, \dots, Y_n$, an i.i.d. sample from the empirical distribution of X_i's. Using (5), one obtains a one-term expansion for $H_{n,\mathrm{B}}^{-1}(1-\alpha/2)$ with a probability bound on the remainder. After substituting this expansion, one needs to use (4) to deduce the required result (3). The result (5) with (6) can be deduced following the standard proof given for the one-term Edgeworth expansion for the sample mean. One set of conditions sufficient for (5) is F is nonlattice and $E|X|^{9/2} < \infty$.

Some results of this section were also mentioned in Singh (1986).

REFERENCES

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

BABU, G. J. and BOSE, A. (1986). Accuracy of the bootstrap approximation. Unpublished.

BABU, G. J. and SINGH, K. (1983). Inference on means using the bootstrap. *Ann. Statist.* **11** 999–1003.

BABU, G. J. and SINGH, K. (1984). On one term Edgeworth correction by Efron's bootstrap. *Sankhyā Ser. A* **46** 219–232.

BERAN, R. J. (1982). Estimated sampling distributions: The bootstrap and competitors. *Ann. Statist.* **10** 212–215.

BICKEL, P. J. and FREEDMAN, D. A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217.

BICKEL P. J., GÖTZE, F. and VAN ZWET, W. R. (1986). The Edgeworth expansion for *U*-statistics of degree two. *Ann. Statist.* **14** 1463–1484.

CALLAERT, H., JANSSEN, P. and VEVAVERBEKE, N. (1980). An Edgeworth expansion for *U*-statistics. *Ann. Statist.* **8** 299–312.

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

HARTIGAN, J. (1986). Comment on "Bootstrap methods for standard errors, confidence intervals and other measures of statistical accuracy" by B. Efron and R. Tibshirani. *Statist. Sci.* **1** 75–77.

SINGH, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187–1195.

SINGH, K. (1986). Discussion of "Jackknife, bootstrap and other resampling methods in regression analysis" by C. F. J. Wu. *Ann. Statist.* **14** 1328–1330.

DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
BUSCH CAMPUS
NEW BRUNSWICK, NEW JERSEY 08903