Statistica Sinica 8(1998), 489-504

ITERATED BOOTSTRAP PREDICTION INTERVALS

Majid Mojirsheibani

Carleton University

Abstract: In this article we study the effects of bootstrap iteration as a method to improve the coverage accuracy of prediction intervals for 'estimator-type' statistics. Both the mechanics and the asymptotic validity of the method are discussed.

Key words and phrases: Bootstrap, calibration, nonparametric, prediction interval.

1. Introduction

Calibration, as a method to improve the coverage accuracy of confidence intervals, was first discussed by Loh (1987) and Hall (1986). When applied to a bootstrap interval, calibration is called iterated bootstrap. A general framework for bootstrap iteration and calibration is discussed by Hall and Martin (1988). Coverage accuracy of iterated bootstrap confidence intervals are also discussed by Martin (1990) and Hall (1992), sections 1.4 and 3.11. Also, see Loh (1991), who has introduced a new calibration procedure, as well as Beran (1987, 1990). In practice, calibration is usually applied to the *percentile* confidence intervals, since the resulting intervals are superior to both Efron's BC_a (for bias-corrected accelerated) and the bootstrap-t intervals. Here superior means three things: good small sample coverage properties, transformation-invariance, and range-preserving, i.e., the endpoints of the interval do not fall outside of the accepted values of the parameter. Our aim in this paper, however, is to study the effects of bootstrap calibration as a method to improve the coverage accuracy of prediction intervals for an estimator-type statistic, $\hat{\theta}_m$ say, which may be viewed as an estimator of some scalar parameter θ; here m is the future sample size. It is important to mention that in practice we have in mind a "future" statistic such as the sample mean or the sample geometric mean or perhaps a monotone transformation of a statistic for which a nonparametric prediction interval is required.

In a recent paper, Mojirsheibani and Tibshirani (1996) proposed a BC_a-type (for bias-corrected accelerated) bootstrap procedure for setting approximate prediction intervals which can be applied to a large class of statistics. In the case of a confidence interval, a brief review of the BC_a procedure is as follows. Let $\hat{\theta}$ be an estimator of θ, the parameter of interest. Let $F(\cdot)$ be the cdf of $\hat{\theta}$ and

$\hat{F}(\cdot)$ be the cdf of $\hat{\theta}^*$, the bootstrap version of $\hat{\theta}$. Then the α-endpoint of a BC$_a$ confidence interval for θ is

$$\theta_{BC_a}[\alpha] = \hat{F}^{-1}\Big[\Phi\Big(z_0 + \{z_0 + z^{(\alpha)}\}\{1 - a(z_0 + z^{(\alpha)})\}^{-1}\Big)\Big],$$

where z_0 is a bias correction factor and a is called the acceleration constant. Expressions for z_0 and a can be found in Efron (1987). Here $z^{(\alpha)} = \Phi^{-1}(\alpha)$. Note that the definitions of $\hat{F}(\cdot)$ are different in the parametric and nonparametric cases.

Mojirsheibani and Tibshirani, in the cited paper, extend the BC$_a$ procedure to the case of prediction intervals. Specifically, let $\hat{\theta}_m$ and $\hat{\theta}_n$ be the estimators of a scalar parameter θ, where n and m are the past and the future sample sizes respectively. (Here again $\hat{\theta}_m$ is an estimator-type statistic, for which a nonparametric prediction interval is desired.) Let $H(\cdot)$ and $F(\cdot)$ be the cdf's of $\hat{\theta}_m$ and $\hat{\theta}_n$, and let $\hat{H}(\cdot)$ and $\hat{F}(\cdot)$ be the cdf's of $\hat{\theta}^*_m$ and $\hat{\theta}^*_n$, the bootstrap versions of $\hat{\theta}_m$ and $\hat{\theta}_n$. Then a central 100(1-2α) percent BC$_a$-type prediction interval for $\hat{\theta}_m$ is

$$\Big(\hat{\theta}_{BC_a}[\alpha], \hat{\theta}_{BC_a}[1-\alpha]\Big), \text{ where } \hat{\theta}_{BC_a}[\alpha] = \hat{H}^{-1}\Big[\Phi\Big(\frac{u^{(\alpha)} - 1}{b} + z_1\Big)\Big]. \qquad (1)$$

Here $u^{(\alpha)}$ is the α quantile of $U = Z_1/Z_2$, where $Z_1 \sim N(1 - bz_1, b^2)$ and $Z_2 \sim N(1 - az_0, a^2)$ are two independent normal random variables. The quantities z_1 and b are given by $z_1 = \Phi^{-1}(\hat{H}(\hat{\theta}_n))$ and $b = a(n/m)^{1/2}$, and a is as before.

We say that a prediction interval is second-order correct if each endpoint of the interval, say $\hat{\theta}[\alpha]$, differs from its theoretical counterpart by $O_p([\min(m,n)]^{-3/2})$, and it is second-order accurate if $P(\hat{\theta}_m \le \hat{\theta}[\alpha]) = \alpha + O([\min(m,n)]^{-1})$. The interval given by (1) is second-order correct and second-order accurate. Furthermore, this interval is transformation-invariant (with respect to monotone transformations) as well as range-preserving, i.e., $\hat{\theta}_{BC_a}[\alpha]$ will not take values outside the support of the distribution of $\hat{\theta}_m$. When the constant b is zero in (1), we obtain a BC-type prediction interval with the α-endpoint

$$\hat{\theta}_{BC}[\alpha] = \hat{H}^{-1}\Big[\Phi\Big(z^{(\alpha)}(1+r)^{1/2} + z_0 r^{1/2}\Big)\Big], \text{ where } r = m/n \text{ and } z^{(\alpha)} = \Phi^{-1}(\alpha).$$

When both a and z_0 are zero, we have a *percentile* prediction interval whose α-endpoint is

$$\hat{\theta}_{perc}[\alpha] = \hat{H}^{-1}\Big[\Phi\Big(z^{(\alpha)}(1+r)^{1/2}\Big)\Big]. \qquad (2)$$

Unfortunately, $\hat{\theta}_{BC}[\alpha]$ and $\hat{\theta}_{perc}[\alpha]$ do not have the same second-order properties as the interval given by (1).

One can also form the normal theory prediction interval $\hat{\theta}_m$. This is obtained by inverting the asymptotically pivotal statistic $T = (\hat{\theta}_m - \hat{\theta}_n)\{(m^{-1} +$

$n^{-1})^{1/2}\hat{\sigma}_n\}^{-1}$, where $\hat{\sigma}_n^2$ is a consistent estimate of the asymptotic variance of $(m^{-1}+n^{-1})^{-1/2}(\hat{\theta}_m - \hat{\theta}_n)$. The α-endpoint of the resulting interval is

$$\hat{\theta}_{norm}[\alpha] = \hat{\theta}_n + z^{(\alpha)}\hat{\sigma}_n(m^{-1}+n^{-1})^{1/2}. \tag{3}$$

Yet, as another candidate, the bootstrap-t method can be used to construct second-order correct (and accurate) prediction intervals for $\hat{\theta}_m$. This works by inverting the bootstrap statistic

$$T^* = (\hat{\theta}_m^* - \hat{\theta}_n^*)\{(m^{-1}+n^{-1})^{1/2}\hat{\sigma}_n^*\}^{-1}, \tag{4}$$

where $\hat{\theta}_m^*$, $\hat{\theta}_n^*$, and $\hat{\sigma}_n^*$ are the bootstrap versions of $\hat{\theta}_m$, $\hat{\theta}_n$, and $\hat{\sigma}_n$. The α-endpoint of the bootstrap-t interval is then given by $\hat{\theta}_{boot.t}[\alpha] = \hat{\theta}_n + \hat{t}^{(\alpha)}\hat{\sigma}_n(m^{-1}+n^{-1})^{1/2}$, where $\hat{t}^{(\alpha)}$ is the α-quantile of the distribution of T^*.

The case of a single future observation requires special attention. Let $X_1,\dots,X_n,Y$ be independently and identically distributed random variables, and let $\bar{X}_n$ and S_n^2 be the sample mean and variance. The X's here represent the past sample and Y is a single future observation. Set $T = (Y-\bar{X}_n)\{[1+n^{-1}]^{1/2}S_n\}^{-1}$ and let $t^{(\alpha)}$ be the α quantile of T; then a $100\cdot(1-\alpha)$ percent, one-sided, prediction interval for Y is $(\bar{X}_n + S_n[1+n^{-1}]^{1/2}\cdot t^{(\alpha)},\infty)$. Since $t^{(\alpha)}$ is typically unknown, one may use the bootstrap-t method to estimate $t^{(\alpha)}$: let $X_1^*,\dots,X_n^*$, and Y^* be random samples of sizes n and 1 drawn with replacement from $X_1,\dots,X_n$. Define T^* by $T^* = (Y^*-\bar{X}_n^*)\{[1+n^{-1}]^{1/2}S_n^*\}^{-1}$, where $\bar{X}_n^*$ and S_n^* are the bootstrap versions of $\bar{X}_n$ and S_n, and let $\hat{t}^{(\alpha)}$ be the α quantile of T^*. Then one has the bootstrap-t prediction interval $(\bar{X}_n + S_n[1+n^{-1}]^{1/2}\cdot\hat{t}^{(\alpha)},\infty)$. Bai et al. (1990) have established a coverage error rate for this interval:

$$P\{Y\in(\bar{X}_n + S_n[1+n^{-1}]^{1/2}\cdot\hat{t}^{(\alpha)},\infty)\} = 1-\alpha+O(n^{-3/4+\gamma}),\quad \text{for all } \gamma>0.$$

Observe that the above results are very different from those of the bootstrap confidence intervals, where the statistic T is given by $T_c = n^{1/2}(\bar{X}_n - E(X))S_n^{-1}$. It is important to note that in the prediction context, the constant $E(X)$ is replaced by the random variable Y, and that while the distribution of $n^{1/2}(\bar{X}-\mu)S_n^{-1}$ tends to normality as n increases, the distribution of $(Y-\bar{X}_n)\{[1+n^{-1}]^{1/2}S_n\}^{-1}$ does not have to be even close to normal no matter how large n is.

One can also form a prediction interval for Y based on the sample (bootstrap) quantile $F_n^{-1}(\alpha) = X_{([n\alpha])}$. Here F_n is the empirical distribution function and $[\cdot]$ denotes the greatest integer function. If F is continuous then $P(Y\geq X_{([n\alpha])}) = 1-(n+1)^{-1}[n\alpha] = 1-\alpha+O(n^{-1})$. Furthermore, if F has a continuous density, f, then $F_n^{-1}(\alpha) = F^{-1}(\alpha)+O_p(n^{-1/2})$. Note that when the future sample size m is 1 the α-endpoint of the *percentile* prediction interval (2) becomes $F_n^{-1}[\Phi\{z^{(\alpha)}(1+n^{-1})^{1/2}\}]$. This interval, however, has the same coverage probability as $(X_{([n\alpha])}, \infty)$. To see this put $\beta = \Phi\{z^{(\alpha)}(1+n^{-1})^{1/2}\}$ and observe that

$$\begin{aligned}
&P\Big(Y \geq F_n^{-1}[\Phi\{(z^{(\alpha)}(1+1/n)^{1/2}\}]\Big)\\
&= P(Y \geq X_{([n\beta])})\\
&= 1-\beta+O(n^{-1}) \text{ by the above argument}\\
&= 1-\Phi\{z^{(\alpha)}(1+n^{-1}/2-n^{-2}/8+\cdots)\}+O(n^{-1})\\
&= 1-\alpha+O(n^{-1}).
\end{aligned}$$

Beran (1990) proposes a calibration method to improve the coverage probability of a prediction interval. His method performs well in parametric situations but fails in nonparametric cases such as the one above based on the sample quantile $F_n^{-1}(\alpha) = X_{([n\alpha])}$ (see Beran's *Example* 4 on page 721 of the cited paper). Another relevant result is that of Stine (1985) who deals with nonparametric bootstrap prediction intervals for a single future observation in a linear regression set-up.

In the rest of this article we will focus on bootstrap calibration of prediction intervals, when both the past and future sample sizes are allowed to increase.

2. Bootstrap Calibration of Prediction Intervals

Suppose that $\hat{\theta}[\alpha]$ is the α-endpoint of a prediction interval for the statistic $\hat{\theta}_m$, where m is the future sample size. If $P(\hat{\theta}_m \leq \hat{\theta}[\alpha]) \neq \alpha$, then, perhaps, there is a $\lambda = \lambda_\alpha$ such that $P(\hat{\theta}_m \leq \hat{\theta}[\lambda]) = \alpha$. In this case $\hat{\theta}[\lambda]$ is the α-endpoint of a calibrated prediction interval for $\hat{\theta}_m$. In practice, λ is unknown and the bootstrap can be used to estimate it.

The main steps for calibrating prediction intervals may be summarized as follows. Let $\hat{\theta}[\alpha]$ be the α-endpoint of a prediction interval for $\hat{\theta}_m$. Generate B bootstrap samples of size n (the past sample size), $\mathbf{X}_1^*, \ldots, \mathbf{X}_B^*$, and B bootstrap samples of size m (the future sample size), $\mathbf{Y}_1^*, \ldots, \mathbf{Y}_B^*$. For $b = 1, \ldots, B$, let $\hat{\theta}_b^*[\lambda]$ be the bootstrap version of $\hat{\theta}[\alpha]$, computed from $\mathbf{X}_b^*$ for a grid of values of λ. Similarly, let $\hat{\theta}_{m,b}^*$ be the bootstrap version of $\hat{\theta}_m$, computed from $\mathbf{Y}_b^*$. Then the bootstrap estimate $\hat{\lambda}$ of λ is the solution of the equation

$$p(\lambda) = \frac{\#\{\hat{\theta}_{m,b}^* \leq \hat{\theta}_b^*[\lambda]\}}{B} = \alpha. \tag{5}$$

In practice, one usually needs an extra B_1 bootstrap samples to form $\hat{\theta}_b^*[\lambda]$, thus requiring a total of $B \cdot B_1 + B = B(B_1+1)$.

Example 1. Consider the construction of a 90 percent, nonparametric, prediction interval for a future sample mean $\bar{Y}_m$ based on the past sample mean

$\bar{X}_n$. Here $m = 15, n = 20$, and the actual data $x_1, \ldots, x_{20}$ were generated from an Exp(1) distribution. The results are given in Table 1. For the calibrated percentile interval, we generated 700 bootstrap past samples of size $n = 20$: $\mathbf{X}_1^*, \ldots, \mathbf{X}_{700}^*$ and 700 bootstrap future samples of size $m = 15$: $\mathbf{Y}_1^*, \ldots, \mathbf{Y}_{700}^*$. For each $\mathbf{X}_b^*$, an additional 1000 bootstrap samples, $\mathbf{Y}_1^{**}, \ldots, \mathbf{Y}_{1000}^{**}$ (drawn from $\mathbf{X}_b^*$), were used to form the percentile endpoints $\bar{Y}_b^*[\lambda_\alpha]$ and $\bar{Y}_b^*[\lambda_{1-\alpha}]$. Here we allowed λ_α and $\lambda_{1-\alpha}$ to vary over a grid of 50 equally spaced values in [0.03, 0.08] and [0.93, 0.98] respectively. Note that the total number of bootstrap samples used is (700)(1000)+700, where the extra 700 bootstrap samples are used to form the $\bar{Y}_{m,b}^*$, $\; b = 1, \ldots, 700$. The estimate of λ_α that solves the equation $p(\lambda_\alpha) = \#\{\bar{Y}_{m,b}^* \leq \bar{Y}_b^*[\lambda_\alpha]\}/700 = 0.05$, was found to be $\hat{\lambda}_\alpha \doteq 0.062$. Similarly, the value $\hat{\lambda}_{1-\alpha} \doteq 0.977$ solves the equation $p(\lambda_{1-\alpha}) = \#\{\bar{Y}_{m,b}^* \leq \bar{Y}_b^*[\lambda_{1-\alpha}]\}/700 = 0.95$. Figure 1 shows the plots of λ vs. $p(\lambda)$. The results in Table 1 show that calibration has brought the percentile interval much closer to the BC$_a$-type and the bootstrap-t intervals. Note that the left-endpoint of the calibrated percentile prediction interval is very close to the BC$_a$ left-endpoint, while its right-endpoint is closer to that of the bootstrap-t interval. We repeated this entire procedure for a total of 1000 times (i.e., 1000 different samples), at a computational cost of

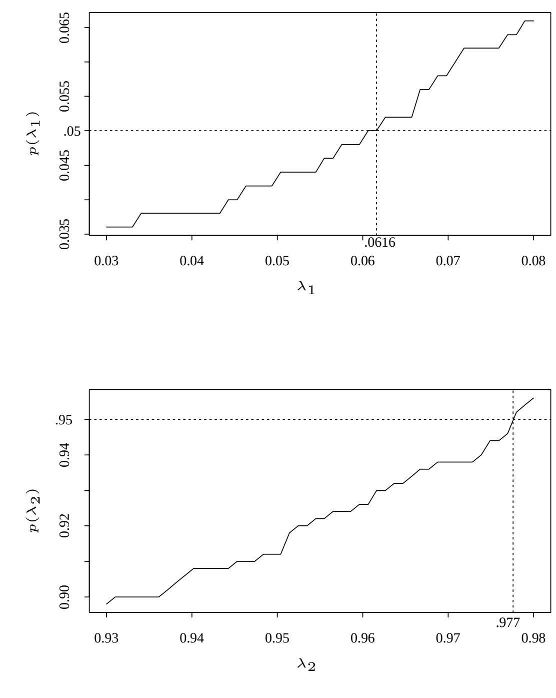

Figure 1. Plots of λ_1 and λ_2 against $p(\lambda_1)$ and $p(\lambda_2)$.

(1000)(700)(1000 + 1) bootstrap resamples. The results of these 1000 Monte Carlo runs are summarized in Table 2. Columns 2 and 3 give the average endpoints; standard deviations are given in brackets. Column 4 gives the average length, and columns 5 and 6 show the number of times (out of 1000) that the interval did not capture its corresponding future sample mean. The last column gives the overall noncoverage.

Table 1. 90 percent nonparametric P.I.'s for $\bar{Y}_m,\ m = 15$

Method	Left-end	Right-end
Bootstrap-t	0.539	1.907
BC_a-type	0.593	1.887
Calibrated percentile	0.591	2.074
Percentile	0.571	1.825

The same analysis was also carried out for the case where $m = 3$, i.e., prediction intervals for the sample mean based on a future sample of size, as small as, $m = 3$. The results appear in Table 3. Both Tables 2 and 3 show that the calibrated prediction intervals exhibit good coverage properties, this is true even for $m = 3$.

Table 2. 90 percent nonparametric P.I.'s for $\bar{Y}_m, m = 15$.

Method	Left endpoint	Right endpoint	Average length	Left noncoverage	Right noncoverage	Overall noncoverage
Bootstrap-t	0.528 (0.135)	1.749 (0.485)	1.221	37	61	98
BC_a-type	0.568 (0.139)	1.649 (0.431)	1.081	44	68	112
Calibrated percentile	0.555 (0.135)	1.768 (0.461)	1.213	43	60	103
Percentile	0.545 (0.138)	1.567 (0.378)	1.022	42	84	126

Table 3. 90 percent nonparametric P.I.'s for $\bar{Y}_m,\ m = 3$.

Method	Left endpoint	Right endpoint	Average length	Left noncoverage	Right noncoverage	Overall noncoverage
Bootstrap-t	0.212 (0.114)	2.279 (0.762)	2.067	32	60	92
BC_a-type	0.284 (0.108)	2.105 (0.594)	1.821	49	64	113
Calibrated percentile	0.255 (0.104)	2.320 (0.764)	2.065	43	58	101
Percentile	0.271 (0.106)	2.039 (0.566)	1.768	45	83	128

Example 2. Let $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ be independently and identically distributed positive random variables. We are interested in constructing nonparametric prediction intervals for the geometric mean

$$G_m = \Big(\prod_{i=1}^{m} Y_i\Big)^{1/m},$$

based on the observable (past) geometric mean $G_n = (\prod_{i=1}^{n} X_i)^{1/n}$. As in the previous example, the sample sizes are taken to be $n = 20$ and $m = 15, 3$. The actual data were generated from a Pareto distribution with density $f(x) = x^{-2}, x > 1$.

Table 4. 90 percent nonparametric P.I.'s for $G_{15} = (\prod_{i=1}^{15} Y_i)^{1/15}$.

Method	Left endpoint	Right endpoint	Average length	Left noncoverage	Right noncoverage	Overall noncoverage
Bootstrap-t	1.728 (0.248)	5.916 (3.058)	4.188	44	64	108
BC_a-type	1.778 (0.261)	5.735 (2.873)	3.957	52	67	119
Calibrated percentile	1.753 (0.245)	6.499 (4.200)	4.746	47	58	105
Percentile	1.738 (0.253)	5.242 (2.348)	3.504	45	83	128

Table 5. 90 percent nonparametric P.I.'s for $G_3 = (\prod_{i=1}^{3} Y_i)^{1/3}$.

Method	Left endpoint	Right endpoint	Average length	Left noncoverage	Right noncoverage	Overall noncoverage
Bootstrap-t	1.218 (0.622)	11.091 (7.843)	9.873	39	58	97
BC_a-type	1.332 (0.144)	10.220 (6.842)	8.888	48	62	110
Calibrated percentile	1.295 (0.136)	14.831 (16.224)	13.536	46	53	99
Percentile	1.315 (0.139)	9.483 (5.930)	8.168	48	70	118

The results are given in Tables 4 and 5. For the bootstrap-t intervals, we have used the *infinitesimal jackknife* estimate of the variance

$$\widehat{\text{VAR}}_{inf.jack}(G_n) = \sum_{i=1}^{n} U_i^2/n^2,$$

where U_i is the ith empirical influence function of G_n. It is not hard to show that

$$U_i = \Big(\prod_{i=1}^{n} X_i\Big)^{1/n} \cdot [\log X_i - n^{-1}\sum_{j=1}^{n} \log X_j].$$

Note that the bootstrap-t procedure replaces U_i by U_i^*, where U_i^* is computed from the bootstrap sample $X_1^*, \ldots, X_n^*$. As in the previous example, the results are based on 1000 Monte Carlo runs. Here again the calibrated interval is a strong competitor for both the BC_a and the bootstrap-t intervals. Note that the usual *percentile* intervals perform poorly.

3. The Effect of Calibration on Prediction Intervals

3.1. Introduction

So far, we have been concerned with the mechanics of calibration as a tool to reduce the coverage error of prediction intervals. Next, we will look at the large sample effects of calibration. Our approach is based on the following two-sample "smooth-function" model, which is suitable for the prediction problem. Let $\mathbf{X}_i, 1 \le i \le n$ and $\mathbf{Y}_j, 1 \le j \le m$ be independently and identically distributed d-vectors and put $\boldsymbol{\mu} = E(\mathbf{X}) = E(\mathbf{Y})$. For known real-valued smooth functions g and h set $\theta = g(\boldsymbol{\mu})$, $\sigma^2 = h(\boldsymbol{\mu})$, $\hat\theta_n = g(\bar{\mathbf{X}}_n)$, $\hat\theta_m = g(\bar{\mathbf{Y}}_m)$, and $\hat\sigma_n^2 = h(\bar{\mathbf{X}}_n)$, where σ^2 is the asymptotic variance of $(m^{-1}+n^{-1})^{-1/2}\{g(\bar{\mathbf{Y}}_m) - g(\bar{\mathbf{X}}_n)\}$. Let the map $S : \Re^{2d} \to \Re$ be defined by $S(\mathbf{x}', \mathbf{y}')' = [g(\mathbf{y}) - g(\mathbf{x})]/h^{1/2}(\mathbf{x})$, where $\mathbf{x}$ and $\mathbf{y}$ are d-vectors and d depends on the statistic of interest. Define T by

$$T = (m^{-1}+n^{-1})^{-1/2}\hat\sigma_n^{-1}\{g(\bar{\mathbf{Y}}_m) - g(\bar{\mathbf{X}}_n)\} = n^{1/2}(1+r^{-1})^{-1/2} S(\bar{\mathbf{X}}_n', \bar{\mathbf{Y}}_m')',$$

where $r = m/n$. Let $\min(m,n) = n$, allowing $n \to \infty$; the one-term Edgeworth expansion of the distribution of T is given by

$$P(T \le x) = \Phi(x) + n^{-1/2} q(x;r)\phi(x) + O(n^{-1}). \tag{6}$$

Here $q(x;r)$ is an even function of x, and r is allowed to act like a parameter itself. See Mojirsheibani and Tibshirani (1996) for the derivation and the exact form of $q(x;r)$. The expansion given by (6) exists under sufficient moment conditions and Cramér's continuity condition (see, for example, expression 2.34 of Hall (1992)). As a simple but important example consider the case where $g(\bar{\mathbf{X}}_n) = \bar X_n$ and $g(\bar{\mathbf{Y}}_m) = \bar Y_m$. Then if (a) $E(X^{10}) < \infty$ and (b) the characteristic function of (X, X^2) satisfies Cramér's condition, then it is not hard to show that $P(T \le x) = \Phi(x) + n^{-1/2} q_1(x;r)\phi(x) + n^{-1} q_2(x;r)\phi(x) + O(n^{-3/2})$, where

$$q_1(x;r) = -\frac{1}{6}\kappa_3\sigma^{-3}(r^{-1}+1)^{-1/2}[3 + (2+r^{-1})(x^2-1)]$$
$$q_2(x;r) = -x\{[(r+1)^{-1}/2 + (x^2-3)((r^{-1}+1) - 3(r^{-1}+1)^{-1})/24]$$

$$\times\sigma^{-4}\kappa_4 + (r^{-1}+1)^{-1}[1+(x^2-3)(r^{-1}+2)/3+(x^4-10x^2 +15)(r^{-1}+2)^2/72]\sigma^{-6}\kappa_3^2 + (r^{-1}+1)^{-1}[3(r^{-1}+1)/2 +(x^2-3)/4]\}.$$

Here $\kappa_3 = E[X-E(X)]^3$ and $\kappa_4 = E[X-E(X)]^4 - 3\sigma^4$.

3.2. One-sided intervals

It is quite straightforward to show that the second-order properties of the bootstrap-t and the BC_a-type prediction intervals, given by (1), do not hold for either the *BC* or the *percentile* intervals. In fact, standard Cornish-Fisher expansions show that

$$\begin{aligned}\hat{\theta}_{BC}[\alpha] &= \hat{H}^{-1}\Big[\Phi(z^{(\alpha)}(1+r)^{1/2}+z_0 r^{1/2})\Big]\\ &= \hat{\theta}_n + n^{-1/2}\hat{\sigma}(1+r^{-1})^{1/2}\{z^{(\alpha)} + (1+r^{-1})^{-1/2}z_0\\ &\quad -n^{-1/2}r^{-1/2}(1+r)^{-1/2}\hat{s}[z^{(\alpha)}(1+r)^{1/2}] + O_p(n^{-1})\}\end{aligned} \tag{7}$$

and

$$\begin{aligned}\hat{\theta}_{perc}[\alpha] &= \hat{H}^{-1}\Big[\Phi(z^{(\alpha)}(1+r)^{1/2})\Big]\\ &= \hat{\theta}_n + n^{-1/2}\hat{\sigma}(1+r^{-1})^{1/2}\{z^{(\alpha)}\\ &\quad -n^{-1/2}r^{-1/2}(1+r)^{-1/2}\hat{s}[z^{(\alpha)}(1+r)^{1/2}] + O_p(n^{-1})\}\end{aligned} \tag{8}$$

respectively. Here $\hat{s}[\cdot]$ is the polynomial of degree two that appears in the bootstrap Cornish-Fisher expansion of the α quantile of the distribution of $T^* = n^{1/2}\hat{\sigma}_n^{-1}(\hat{\theta}_n^* - \hat{\theta}_n)$. Using the Edgeworth expansion of T, as given by (6), we can write the one-sided coverage expansion of the *percentile* and the *BC* prediction intervals as follows. Put

$$w^{(\alpha)} = z^{(\alpha)} - n^{-1/2}r^{-1/2}(1+r)^{-1/2}s[z^{(\alpha)}(1+r)^{1/2}],$$

then

$$\begin{aligned}P(\hat{\theta}_m \le \hat{\theta}_{perc}[\alpha]) &= P\{T \le z^{(\alpha)} - n^{-1/2}r^{-1/2}(1+r)^{-1/2}\hat{s}[z^{(\alpha)}(1+r)^{1/2}]\\ &\quad +O_p(n^{-1})\}\\ &= \Phi(w^{(\alpha)}) + n^{-1/2}q(w^{(\alpha)};r)\cdot\phi(w^{(\alpha)}) + O(n^{-1})\\ &= \alpha + n^{-1/2}\{q(z^{(\alpha)};r) - r^{-1/2}(1+r)^{-1/2}s[z^{(\alpha)}(1+r)^{1/2}]\}\\ &\quad \times\phi(z^{(\alpha)}) + O(n^{-1}),\end{aligned} \tag{9}$$

and

$$\begin{aligned}P(\hat{\theta}_m \le \hat{\theta}_{BC}[\alpha]) &= P\{T \le w^{(\alpha)} + n^{-1/2}(1+r^{-1})^{-1/2}s(0)\} + O(n^{-1})\\ &= \alpha + n^{-1/2}\{q(z^{(\alpha)};r) + (1+r^{-1})^{-1/2}s(0) - r^{-1/2}(1+r)^{-1/2}\\ &\quad \times s[z^{(\alpha)}(1+r)^{1/2}]\}\cdot\phi(z^{(\alpha)}) + O(n^{-1}).\end{aligned} \tag{10}$$

In what follows , we study the effects of calibration on the percentile intervals, since they are quite easy to construct and are transformation-invariant and range-preserving. Other intervals (such as the BC or the normal theory intervals) may be handled similarly.

Let $\hat{\theta}_{perc}[\alpha]$ be the α-endpoint of the one-sided percentile prediction interval for $\hat{\theta}_m$, where $0 < \alpha < 1$. Equation (9) implies that $P(\hat{\theta}_m \leq \hat{\theta}_{perc}[\alpha]) = \alpha + O(n^{-1/2})$. Suppose $\lambda = \lambda_\alpha$ is such that

$$P(\hat{\theta}_m \leq \hat{\theta}_{perc}[\lambda]) = \alpha. \tag{11}$$

Then it is not hard to show that

$$P(\hat{\theta}_m \leq \hat{\theta}_{perc}[\hat{\lambda}]) = \alpha + O(n^{-1}), \tag{12}$$

where $\hat{\lambda}$ is the solution of (5) for $B = \infty$. The argument behind (12) is as follows. Write $\lambda = \lambda_\alpha = \delta + \alpha$, where δ is an unknown constant. From (9) we obtain

$$\begin{aligned} P(\hat{\theta}_m \leq \hat{\theta}_{perc}[\lambda]) &= P(\hat{\theta}_m \leq \hat{\theta}_{perc}[\alpha + \delta]) \\ &= \alpha + \delta + n^{-1/2}\{q(z^{(\alpha+\delta)}; r) - r^{-1/2}(1+r)^{-1/2} \\ &\quad \cdot s[z^{(\alpha+\delta)}(1+r)^{1/2}]\} \cdot \phi(z^{(\alpha+\delta)}) + O(n^{-1}). \end{aligned} \tag{13}$$

If we compare the R.H.S.'s of (13) and (11) we find

$$\delta = -n^{-1/2}\{q(z^{(\alpha+\delta)}; r) - r^{-1/2}(1+r)^{-1/2} s[z^{(\alpha+\delta)}(1+r)^{1/2}]\}\phi(z^{(\alpha+\delta)}) + O(n^{-1}),$$

that is, $\delta = O(n^{-1/2})$. In fact, since

$$q(z^{(\alpha+\delta)}; r) = q(z^{(\alpha)} + \frac{\delta}{\phi(z^{(\alpha)})} + \cdots; r) = q(z^{(\alpha)}; r) + O(n^{-1/2})$$

and

$$\begin{aligned} s[z^{(\alpha+\delta)}(1+r)^{1/2}] &= s[\{z^{(\alpha)} + \frac{\delta}{\phi(z^{(\alpha)})} + \cdots\}(1+r)^{1/2}] \\ &= s[z^{(\alpha)}(1+r)^{1/2}] + O(n^{-1/2}), \end{aligned}$$

we may write δ as

$$\delta = -n^{-1/2}\{q(z^{(\alpha)}; r) - r^{-1/2}(1+r)^{-1/2} s[z^{(\alpha)}(1+r)^{1/2}]\}\phi(z^{(\alpha)}) + O(n^{-1}). \tag{14}$$

If we set $\hat{\lambda} = \hat{\lambda}_\alpha = \hat{\delta} + \alpha$, where $\hat{\delta}$ is the sample (bootstrap) version of δ, then

$$\hat{\delta} = -n^{-1/2}\{\hat{q}(z^{(\alpha)}; r) - r^{-1/2}(1+r)^{-1/2} \hat{s}[z^{(\alpha)}(1+r)^{1/2}]\}\phi(z^{(\alpha)}) + O_p(n^{-1}), \tag{15}$$

where $\hat{q}$ and $\hat{s}$ are obtained from q and s by replacing the population moments with sample moments. Using the Taylor expansion

$$z^{(\hat{\lambda})} = z^{(\alpha+\hat{\delta})} = z^{(\alpha+\delta)} + \frac{(\hat{\delta}-\delta)}{\phi(z^{(\alpha+\delta)})} + \cdots,$$

and the fact that $\hat{\delta} - \delta = O_p(n^{-1})$, we may write

$$\begin{aligned}\hat{\theta}_{perc}[\hat{\lambda}] - \hat{\theta}_{perc}[\lambda] &= n^{-1/2}(1+r^{-1})^{1/2}\hat{\sigma}\{(z^{(\hat{\lambda})} - z^{(\lambda)}) - n^{-1/2}r^{-1/2}(1+r)^{-1/2} \\ &\quad \cdot(\hat{s}[z^{(\hat{\lambda})}(1+r)^{1/2}] - \hat{s}[z^{(\lambda)}(1+r)^{1/2}]) + O_p(n^{-1})\} \\ &= O_p(n^{-3/2}). \qquad (16)\end{aligned}$$

Therefore the coverage of the calibrated, one-sided, percentile prediction interval for $\hat{\theta}_m$ is:

$$\begin{aligned}P(\hat{\theta}_m \le \hat{\theta}_{perc}[\hat{\lambda}]) &= P(\hat{\theta}_m \le \hat{\theta}_{perc}[\lambda] + O_p(n^{-3/2})) \\ &= P(T \le z^{(\lambda)} - n^{-1/2}r^{-1/2}(1+r)^{-1/2}\hat{s}[z^{(\lambda)}(1+r)^{1/2}] + O_p(n^{-1})) \\ &= \alpha + O(n^{-1}). \qquad (17)\end{aligned}$$

In addition to improving the coverage accuracy, this method of calibration produces prediction intervals which are second-order correct. To see that calibration makes $\hat{\theta}_{perc}[\hat{\lambda}]$ a second-order correct endpoint, observe that from (16) and (8) we obtain

$$\begin{aligned}\hat{\theta}_{perc}[\hat{\lambda}] &= \hat{\theta}_{perc}[\lambda] + O_p(n^{-3/2}) \\ &= \hat{\theta}_n + n^{-1/2}\hat{\sigma}(1+r^{-1})^{1/2}\{z^{(\alpha+\delta)} \\ &\quad - n^{-1/2}r^{-1/2}(1+r)^{-1/2}\hat{s}[z^{(\alpha+\delta)}(1+r)^{1/2}]\} + O_p(n^{-3/2}) \\ &= \hat{\theta}_n + n^{-1/2}\hat{\sigma}(1+r^{-1})^{1/2}\{z^{(\alpha)} + \frac{\delta}{\phi(z^{(\alpha)})} \\ &\quad - n^{-1/2}r^{-1/2}(1+r)^{-1/2}\hat{s}[z^{(\alpha)}(1+r)^{1/2}]\} + O_p(n^{-3/2}) \\ &= \hat{\theta}_{BC_a}[\alpha] + O_p(n^{-3/2}), \qquad (18)\end{aligned}$$

where (18) follows upon replacing δ by the R.H.S of (14).

Remark A. What is the use of the expansions (14) and (15)? To answer this question, put $\lambda_\alpha = \delta_\alpha + \alpha, \lambda_{1-\alpha} = \delta_{1-\alpha} + (1-\alpha)$, $\hat{\lambda}_\alpha = \hat{\delta}_\alpha + \alpha$ and $\hat{\lambda}_{1-\alpha} = \hat{\delta}_{1-\alpha} + (1-\alpha)$, where δ_α and $\delta_{1-\alpha}$ are unknown constants and $0 < \alpha < 0.5$. Then

$$\lambda_{1-\alpha} = (1-\alpha) + \delta_\alpha + O(n^{-1}) = \lambda_\alpha + (1-2\alpha) + O(n^{-1})$$

and

$$\hat{\lambda}_{1-\alpha} = \hat{\lambda}_\alpha + (1-2\alpha) + O_p(n^{-1}); \qquad (19)$$

here we have used the fact that in (14) and (15), the functions q, s, and ϕ are even. In other words, expansions (14) and (15) show that we can simply use (19) to estimate $\hat{\lambda}_{1-\alpha}$ from $\hat{\lambda}_\alpha$ (or vice versa).

Remark B. It turns out that the bootstrap iteration, as outlined at the beginning of Section 2, does not improve the coverage accuracy of a two-sided prediction interval. In other words, for $0 < \alpha < 0.5$, suppose that λ_α and $\lambda_{1-\alpha}$ are such that

$$P(\hat{\theta}_m \leq \hat{\theta}[\lambda_\alpha]) = \alpha \text{ and } P(\hat{\theta}_m \leq \hat{\theta}[\lambda_{1-\alpha}]) = 1 - \alpha. \tag{20}$$

Then the interval $(\hat{\theta}[\lambda_\alpha], \hat{\theta}[\lambda_{1-\alpha}])$ has the two-sided coverage $(1 - \alpha) - \alpha = 1-2\alpha$. On the other hand, it can be shown that (see Appendix) for the calibrated interval, $(\hat{\theta}[\hat{\lambda}_\alpha], \hat{\theta}[\hat{\lambda}_{1-\alpha}])$,

$$P(\hat{\theta}_m \leq \hat{\theta}[\hat{\lambda}_{1-\alpha}]) - P(\hat{\theta}_m \leq \hat{\theta}[\hat{\lambda}_\alpha]) = 1 - 2\alpha + O(n^{-1}). \tag{21}$$

It is possible to calibrate a two-sided prediction interval in such a way as to improve its coverage accuracy beyond the usual level. Such calibration methods, however, do not yield second-order correct endpoints and will not be pursued here.

Remark C. To avoid the high computational cost of a direct search method for finding λ_α, one may consider the alternative method proposed by Loh (1991). In its simplest form, Loh's method works as follows. Let the statistic of interest be a smooth function of means and consider the calibration of the normal theory prediction interval (3). Generate B bootstrap samples of size n (the past sample size), $\mathbf{X}_1^*, \ldots, \mathbf{X}_B^*$, and B bootstrap samples of size m (the future sample size), $\mathbf{Y}_1^*, \ldots, \mathbf{Y}_B^*$. For $b = 1, \ldots, B$, let T_b^* be as in (4), computed utilizing $\mathbf{X}_b^*$ and $\mathbf{Y}_b^*$. Put $\hat{\beta}_b = 1 + \Phi(T_b^*)$, (the '+' sign here is correct). Then in the case of a one-sided interval take $\hat{\lambda}_\alpha$ to be the α-quantile of the set $\{\hat{\beta}_b,\ b = 1, \ldots, B\}$. Using expansions similar to (13) and (17), it is quite straightforward to show that the resulting interval has a coverage error of order $O(n^{-1})$. In fact, one can also show that the endpoint of this calibrated interval is second-order correct as well. For a two-sided prediction interval, Loh's method works by setting $\hat{\beta}_b = 1+\Phi(|T_b^*|)$ and taking $\hat{\lambda}_\alpha$ to be the 2α-quantile of the set $\{\hat{\beta}_b, b = 1, \ldots, B\}$. With much more effort, one can show that the coverage error of the resulting two-sided interval is of order $O(n^{-2})$ (i.e., fourth-order accuracy), but the endpoints, in general, are not second-order correct.

To reduce the coverage error rate of confidence interval, Loh (1991) also considers the calibration of a one-term Edgeworth-corrected interval. This approach requires, among other things, the computation of the functions $\hat{q}_1^*(x; r)$,

$\ldots, \hat{q}^*_B(x;r)$, where $\hat{q}^*_b(x;r)$ is $\hat{q}(x;r)$ computed from the bth bootstrap past sample $\mathbf{X}^*_b$, and $\hat{q}(x;r)$ is the sample version of $q(x;r)$ that appears in (6). Unfortunately, in the case of a prediction interval, the function $q(x;r)$ is quite complicated and not convenient to work with.

Acknowledgement

The author would like to thank the anonymous referee for his/her helpful suggestions. This work was supported in part by a GR-5 Grant from Carleton University.

Appendix

Derivation of (21). First note that the α-endpoint of the calibrated percentile prediction interval is given by

$$\begin{aligned}\hat{\theta}_{perc}[\hat{\lambda}_\alpha] &= \hat{\theta}_{perc}[\alpha+\hat{\delta}_\alpha]\\ &= \hat{\theta}_n + n^{-1/2}(1+r^{-1})^{1/2}\hat{\sigma}\{z^{(\alpha+\hat{\delta}_\alpha)}\\ &\quad +(1+r)^{-1/2}\sum_{j=1}^{2} n^{-j/2}r^{-j/2}\hat{s}_j[z^{(\alpha+\hat{\delta}_\alpha)}(1+r)^{1/2}] + O_p(n^{-3/2})\},\end{aligned}$$

where $\hat{s}_1$ and $\hat{s}_2$ are the polynomials in the bootstrap Cornish-Fisher expansion of the α quantile of the distribution of $T^* = n^{1/2}\hat{\sigma}_n^{-1}(\hat{\theta}^*_n - \hat{\theta}_n)$. In fact $\hat{s}_1 = -\hat{s}$, where $\hat{s}$ is as before. Since

$$\hat{s}_j[z^{(\alpha+\hat{\delta}_\alpha)}(1+r)^{1/2}] = \hat{s}_j[z^{(\alpha+\delta_\alpha)}(1+r)^{1/2}] + O_p(n^{-1}) \quad j=1,2$$

and

$$z^{(\alpha+\hat{\delta}_\alpha)} = z^{(\alpha+\delta_\alpha)} + \frac{(\hat{\delta}_\alpha-\delta_\alpha)}{\phi(z^{(\alpha+\delta_\alpha)})} + O_p(n^{-2}),$$

we may rewrite $\hat{\theta}_{perc}[\hat{\lambda}_\alpha]$ as

$$\begin{aligned}\hat{\theta}_{perc}[\hat{\lambda}_\alpha] &= \hat{\theta}_n + n^{-1/2}(1+r^{-1})^{1/2}\hat{\sigma}\{z^{(\alpha+\delta_\alpha)} + (\hat{\delta}_\alpha-\delta_\alpha)\phi^{-1}(z^{(\alpha)})\\ &\quad +(1+r)^{-1/2}\sum_{j=1}^{2} n^{-j/2}r^{-j/2}\hat{s}_j[z^{(\alpha+\delta_\alpha)}(1+r)^{1/2}] + O_p(n^{-3/2})\}.\end{aligned}$$

Now define the random variables U_1 and U_2 according to

$$\begin{aligned}U_1 &= n^{-1/2}r^{-1/2}(1+r)^{-1/2}\{\hat{s}_1[z^{(\alpha+\delta_\alpha)}(1+r)^{1/2}] - s_1[z^{(\alpha+\delta_\alpha)}(1+r)^{1/2}]\},\\ U_2 &= n(\hat{\delta}_\alpha-\delta_\alpha)\phi^{-1}(z^{(\alpha)}).\end{aligned}$$

Also, put

$$w^{(\alpha+\delta_\alpha)} = z^{(\alpha+\delta_\alpha)} + (1+r)^{-1/2}\sum_{j=1}^{2} n^{-j/2} r^{-j/2} s_j[z^{(\alpha+\delta_\alpha)}(1+r)^{1/2}].$$

Then it is straightforward to show that

$$P(\hat{\theta}_m \le \hat{\theta}_{perc}[\hat{\lambda}_\alpha]) = P(T - U_1 - n^{-1}U_2 \le w^{(\alpha+\delta_\alpha)} + O_p(n^{-3/2})),$$

(note that $w^{(\alpha+\delta_\alpha)}$ is a constant.)

On the other hand, it can be shown that the distribution of $T - U_1 - n^{-1}U_2$ admits the expansion

$$P(T-U_1-n^{-1}U_2 \le x) = P(T-U_1 \le x) + n^{-1}x\phi(x)\cdot E(TU_2) + O(n^{-3/2}). \quad (22)$$

The proof of the above expansion is similar to that in the argument leading to (3.36) of Hall (1992), and involves the following two steps:

Step (i). Let $\kappa_j(X)$ denote the jth cumulant of the random variable X; then it is easy to show that

$$\begin{aligned}
\kappa_1(T - U_1 - n^{-1}U_2) &= \kappa_1(T - U_1) + O(n^{-3/2}) \\
\kappa_2(T - U_1 - n^{-1}U_2) &= \kappa_2(T - U_1) - 2n^{-1}E(TU_2) + O(n^{-2}) \\
\kappa_3(T - U_1 - n^{-1}U_2) &= \kappa_3(T - U_1) + O(n^{-3/2}) \\
\kappa_4(T - U_1 - n^{-1}U_2) &= \kappa_4(T - U_1) + O(n^{-2})
\end{aligned}$$

Step (ii). Writing an expansion of the characteristic function of $T - U_1 - n^{-1}U_2$ in terms of its first four cumulants and then rewriting it in terms of the cumulants of $T - U_1$ (from Step (i)), and finally inverting this characteristic function will result in the probability expansion (22).

Now we may use (22) to write

$$\begin{aligned}
P(\hat{\theta}_m \le \hat{\theta}_{perc}[\hat{\lambda}_\alpha]) &= P(\hat{\theta}_m \le \hat{\theta}_{perc}[\alpha + \hat{\delta}_\alpha]) \\
&= P(T - U_1 - n^{-1}U_2 \le w^{(\alpha+\delta_\alpha)}) + O(n^{-3/2}) \\
&= P(T - U_1 \le w^{(\alpha+\delta_\alpha)}) \\
&\quad + n^{-1}w^{(\alpha+\delta_\alpha)}\phi(w^{(\alpha+\delta_\alpha)})E(TU_2) + O(n^{-3/2}) \\
&= \alpha + n^{-1}z^{(\alpha)}C_{\delta_\alpha} + O(n^{-3/2}), \qquad (23)
\end{aligned}$$

where C_{δ_α} is such that

$$E\Big(n^{1/2}(1+r^{-1})^{-1/2}\hat{\sigma}^{-1}(\hat{\theta}_m - \hat{\theta}_n)\cdot n(\hat{\delta}_\alpha - \delta_\alpha)\Big) = C_{\delta_\alpha} + O(n^{-1}). \quad (24)$$

(Here, the right hand side of (24) follows by taking the expectation of the product of the Taylor expansions of the $O_p(1)$ random variables $T = n^{1/2}(1 + r^{-1})^{-1/2}\hat{\sigma}^{-1}(\hat{\theta}_m - \hat{\theta}_n)$ and $V = n(\hat{\delta}_\alpha - \delta_\alpha)$.)

The justification of (23) follows from the following facts:

1. $P(T-U_1 \le w^{(\alpha+\delta_\alpha)}) = P(\hat\theta_m \le \hat\theta_{perc}[\alpha+\delta_\alpha]) + O(n^{-3/2}) = \alpha + O(n^{-3/2})$,
2. $w^{(\alpha+\delta_\alpha)} = z^{(\alpha+\delta_\alpha)} + O(n^{-1/2}) = z^{(\alpha)} + O(n^{-1/2})$,

   and
3.

$$\begin{aligned} n^{-1}\phi(w^{(\alpha+\delta_\alpha)})\cdot E(TU_2) &= n^{-1}\phi(w^{(\alpha+\delta_\alpha)})\cdot\phi^{-1}(z^{(\alpha)})[C_{\delta_\alpha}+O(n^{-1})] \\ &= n^{-1}\phi(z^{(\alpha)}+O(n^{-1/2}))\cdot\phi^{-1}(z^{(\alpha)})[C_{\delta_\alpha}+O(n^{-1})] \\ &= n^{-1}C_{\delta_\alpha}+O(n^{-3/2}). \end{aligned}$$

Similarly, we have

$$\begin{aligned} P(\hat\theta_m \le \hat\theta_{perc}[\hat\lambda_{1-\alpha}]) &= P(\hat\theta_m \le \hat\theta_{perc}[(1-\alpha)+\hat\delta_{1-\alpha}]) \\ &= 1-\alpha+n^{-1}z^{(1-\alpha)}C_{\delta_{1-\alpha}}+O(n^{-3/2}), \end{aligned} \tag{25}$$

where $C_{\delta_{1-\alpha}}$ is such that $E(n^{1/2}(1+r^{-1})^{-1/2}\hat\sigma^{-1}(\hat\theta_m-\hat\theta_n)\cdot n(\hat\delta_{1-\alpha}-\delta_{1-\alpha})) = C_{\delta_{1-\alpha}}+O(n^{-1})$. Now the coverage of the calibrated, two-sided, percentile prediction interval for $\hat\theta_m$ is given by:

$$\begin{aligned} &P(\hat\theta_m \le \hat\theta_{perc}[(1-\alpha)+\hat\delta_{1-\alpha}]) - P(\hat\theta_m \le \hat\theta_{perc}[\alpha+\hat\delta_\alpha]) \\ &= 1-\alpha+n^{-1}z^{(1-\alpha)}C_{\delta_{1-\alpha}} - (\alpha+n^{-1}z^{(\alpha)}C_{\delta_\alpha}) + O(n^{-3/2}) \\ &= 1-2\alpha-n^{-1}z^{(\alpha)}(C_{\delta_{1-\alpha}}+C_{\delta_\alpha})+O(n^{-3/2}), \end{aligned} \tag{26}$$

where the term $C_{\delta_{1-\alpha}}+C_{\delta_\alpha}$ does not vanish in (26).

References

Bai, C., Bickel, P. and Olshen, R. (1990). Hyperaccuracy of bootstrap based prediction. In *Probability in Banach Spaces VII Proceedings of the Seventh International Conference* (Edited by E. Eberlein, J. Kuelbs and M.B. Marcus), 31-42, Birkhauser Boston, Cambridge, Massachusetts.

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

Beran, R. (1990). Calibrating prediction regions. *J. Amer. Statist. Assoc.* **85**, 715-723.

DiCiccio, T. and Tibshirani, R. (1987). Bootstrap confidence intervals and bootstrap approximations. *J. Amer. Statist. Assoc.* **82**, 163-170.

Efron, B. (1987). Better bootstrap confidence intervals. *J. Amer. Statist. Assoc.* **82**, 171-185.

Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap.* Chapman and Hall. New York.

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

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals. (With discussion.) *Ann. Statist.* **16**, 927-985.

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

Hall, P. and Martin, M. A. (1988). On bootstrap resampling and iteration. *Biometrika* **75**, 661-671.

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

Loh, W.-Y. (1991). Bootstrap calibration for confidence interval construction and selection. *Statist. Sinica* **1**, 477-491.

Martin, M. A. (1990). On bootstrap iteration for coverage correction in confidence intervals. *J. Amer. Statist. Assoc.* **85**, 1105-1118.

Mojirsheibani, M. and Tibshirani, R. (1996). Some results on bootstrap prediction intervals. *Canad. J. Statist.* **24**, 549-568.

Stine, R. (1985). Bootstrap prediction intervals for regression. *J. Amer. Statist. Assoc.* **80**, 1026-1031.

Department of Mathematics and Statistics, Herzberg Bldg, Carleton University, Ottawa, Ontario, K1S 5B6, Canada.

E-mail: majidm@math.carleton.ca

(Received December 1996; accepted June 1997)