CORRECTNESS AND ACCURACY OF BOOTSTRAP CONFIDENCE INTERVALS

Linda Bohdan, Kubanová Jana
Ústav matematiky, FES, Univerzita Pardubice

Abstract
This paper describes some theoretical backgrounds of confidence interval construction. The order of accuracy that reach the bounds of the confidence interval determined by a standard normal method, bootstrap t-method and BC_a method is analyzed in a detail way.

Key words: correctness and accuracy of the first and second order

1 Accuracy and correctness

Let's assume that $X_1, \ldots, X_n$ is a random sample from distribution with distribution function *F,* and θ is some parameter of this distribution. Our objective is to create the confidence interval for this parameter. The generally used technique is to determine simultaneously lower and upper limit of this interval. Now we will consider in successive steps apart the lower and apart the upper limit. For the lower limit $\hat{\theta}[\alpha]$ is valid:

$$P(\theta \le \hat{\theta}[\alpha]) \approx \alpha$$

for any α .

We say that the confidence interval limit $\hat{\theta}[\alpha]$ has the *first-order accuracy*, if

$$P(\theta \le \hat{\theta}[\alpha]) = \alpha + O(n^{-1/2})$$

and the *second-order accuracy*, if

$$P(\theta \le \hat{\theta}[\alpha]) = \alpha + O(n^{-1}) .$$

In common use standard normal and Student's t intervals are of the first-order accuracy, but they aren't of the second order accuracy (only in the case of normal distribution of random variable *X*). Some bootstrap methods make possible to determinate intervals of the second order accuracy, without dependence on distribution of random variable *X.*

Correctness is another from the basic terms. Correctness refers, to what degree the potential confidence interval limit responds to exact limit. Let's $\hat{\theta}_{\text{exact}}[\alpha]$ is the exact confidence interval limit, which matches relation $P(\theta \le \hat{\theta}_{\text{exact}}[\alpha]) = \alpha$. The confidence interval limit $\hat{\theta}[\alpha]$ has the *first-order correctness*, if

$$\hat{\theta}[\alpha] = \hat{\theta}_{exact}[\alpha] + O_p(n^{-1}) \quad (1)$$

and the *second-order correctness*, if

$$\hat{\theta}[\alpha] = \hat{\theta}_{exact}[\alpha] + O_p\left(n^{-3/2}\right). \quad (2)$$

Equivalently, the confidence interval limit has the *first-order correctness*, if

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

and the *second-order correctness*, if

$$\hat{\theta}[\alpha] = \hat{\theta}_{exact}[\alpha] + O_p(n^{-1}).\hat{\sigma} \tag{4}$$

where $\hat{\sigma}$ is standard error estimate $\hat{\theta}$. Because of $\hat{\sigma}$ has usually order $n^{-1/2}$, the relations (1) and (2) are equivalent to (3) and (4).
It can be shown, that correctness of given order implies accuracy of this order. In situation, when it is possible to identify the exact confidence interval limits, then the limits obtained for standard normal and Student's *t*- intervals have only the first-order correctness, while some bootstrap methods enable obtaining of the second order correctness limits.

2 Approximate pivots

Model $\theta = g(\mu)$, where *g* is any smooth function and μ is mean of distribution *F,* is usually used at studying of bootstrap confidence intervals. If we estimate mean μ with average $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, then the estimation of the parameter θ is statistics $\hat{\theta} = g(\overline{X})$.

Let's further consider the variation of estimation $\hat{\theta}$ in the form $\hat{\theta} = \frac{\tau^2}{n}$, whereas we again assume, that τ^2 is any smooth function of μ, it is $\tau^2 = h(\mu)$. The sample estimation of τ^2 is $\hat{\tau}^2 = h(\overline{X})$.
The following problem is, under what conditions is it possible to use the central limit theorem for estimation of $\hat{\theta}$. We will consider following four variables:

$$P = \sqrt{n}(\hat{\theta} - \theta); \qquad Q = \sqrt{n}(\hat{\theta} - \theta)/\tau;$$
$$\hat{P} = \sqrt{n}(\hat{\theta}^* - \hat{\theta}); \qquad Q = \sqrt{n}(\hat{\theta}^* - \hat{\theta})/\hat{\tau}^*$$

$\hat{\theta}^*$ and $\hat{\tau}^*$ are bootstrap analogies of $\hat{\theta}$ and $\hat{\tau}$ calculated of bootstrap sample. If we have known *F*, the exact confidence interval limits could be based on distribution of *U* or *V*. Let's *C*(*x*) and *D*(*x*) are distribution functions of variables *U* and *V* when sampling from *F* and let's $x^{(\alpha)} = C^{-1}(\alpha)$ and $y^{(\alpha)} = D^{-1}(\alpha)$ are α-quantiles of *C*(*x*) and *D*(*x*). Then the exact confidence interval limits based on pivot variable *U* are $C(x) = P\{n^{1/2}(\hat{\theta} - \theta) \le x\} = P\{\theta \ge \hat{\theta} - n^{-1/2}x\}$

and we get $\hat{\theta}_{ns}[\alpha] = \hat{\theta} - n^{-1/2}x^{(1-\alpha)}$ and similarly for *V* we get $\hat{\theta}_{st}[\alpha] = \hat{\theta} - n^{-1/2}\hat{\tau}y^{(1-\alpha)}$.

The first limit is „non-studentized" based on *U*, while the second one is „bootstrap-*t*" or studentized limit based on *V*.
Standard normal confidence interval limit is $\hat{\theta}_{st}$ with normal quantile $z^{(1-\alpha)}$ substituting $y^{(1-\alpha)}$, while common *t* interval use α- quantile of Student's–*t* distribution with *n-1* degrees of freedom.

The distribution function *F* is generally unknown. Bootstrap use estimates $\hat{C}$, $\hat{D}$ of distribution $\hat{U}$ and $\hat{V}$ from estimated distribution $\hat{F}$ to estimate *C* and *D*. If $\hat{x}^{(\alpha)} = \hat{C}^{-1}(\alpha)$

and $\hat{y}^{(\alpha)} = \hat{D}^{-1}(\alpha)$, the estimated confidence intervals limits are: $\hat{\theta}_{NS}[\alpha] = \hat{\theta} - n^{-1/2}\hat{x}^{(1-\alpha)}$ and $\hat{\theta}_{ST}[\alpha] = \hat{\theta} - n^{-1/2}\hat{\tau}\hat{y}^{(1-\alpha)}$.

Following relations are valid for these confidence intervals limits:

$$\hat{\theta}_{NS} = \hat{\theta}_{ns} + O_p(n^{-1}); \qquad \hat{C}(x) = C(x) + O_p(n^{-1/2}),$$
$$\hat{\theta}_{ST} = \hat{\theta}_{st} + O_p(n^{-3/2}); \qquad \hat{D}(x) = D(x) + O_p(n^{-1}).$$

In the other words the confidence interval limit $\hat{\theta}_{ST}$ based on V is of the second-order accuracy and if we consider $\hat{\theta}_{st}$ to be the "proper" interval, so it is as well of the second-order accuracy. It differs from $\hat{\theta}_{st}$ in the term of order $O(n^{-3/2})$. The confidence interval limit $\hat{\theta}_{NS}$ based on U is only of the first-order accuracy.
The bootstrap confidence interval has to be based on a studentized variable in order to have better properties then intervals based on normal approximation.

It is possible to show the reason why studentization is so important. The first four cumulants of U are

$$\begin{aligned} E(U) &= \frac{g_1(\theta)}{\sqrt{n}} + O(n^{-1}), \\ var(U) &= g_2(\theta) + O(n^{-1}), \\ SK(U) &= \frac{g_3(\theta)}{\sqrt{n}} + O(n^{-3/2}), \\ EK(U) &= O(n^{-1}), \end{aligned} \qquad (5)$$

when the common conditions are fulfilled and *SK* and *EK* are symbols for skewness and kurtosis.

These cumulants for V are:

$$\begin{aligned} E(V) &= \frac{g_4(\theta)}{\sqrt{n}} + O(n^{-1}), \\ var(V) &= 1 + O(n^{-1}), \\ SK(V) &= \frac{g_5(\theta)}{\sqrt{n}} + O(n^{-3/2}), \\ EK(V) &= O(n^{-1}). \end{aligned}$$

Functions $g_1(\theta), g_2(\theta), g_3(\theta), g_4(\theta)$ and $g_5(\theta)$ are dependent on θ, but not on n.
All other cumulants are $O(n^{-1})$ or smaller. Variation $var(V)$ doesn't involve any function $g_i(\theta)$. The details are described in DiCiccio and Romano (1988) and Hall(1988).

Applications of $\hat{C}$ and $\hat{D}$ to estimates C and D is equivalent to estimate $\hat{\theta}$ of parameter θ in these functions and the result is error $O_p(n^{-1/2})$, which is $g_1(\hat{\theta}) = g_1(\theta) + O_p(n^{-1/2})$, $g_2(\hat{\theta}) = g_2(\theta) + O_p(n^{-1/2})$, and so on. If these functions are substituted into equation (5), mean, standardized skewness and kurtosis of $\widehat{U}$ are only $O(n^{-1})$, far away from corresponding cumulants of U, but

$$var(\widehat{U}) = var(U) + O(n^{-1/2}).$$

This causes that confidence interval limit based on $\widehat{U}$ is only first-order accuracy. On the other hand $var(V)=1+O(n^{-1})$, $var(\hat{\theta}) = 1 + O(n^{-1})$, so that the estimate doesn't run the error $O(n^{-1/2})$. The result is confidence interval limit based on $\widehat{V}$, that is of the second-order accuracy.

3 BC_a confidence interval limits

the α-percentage BC_a confidence interval limit is given with relation

$$\widehat{\theta}_{BC_a}[\alpha] = \widehat{G}^{-1}(\Phi(\widehat{z}_0 + \frac{\widehat{z}_0 + z^{(\alpha)}}{1 - \widehat{a}(\widehat{z}_0 + z^{(\alpha)})})),$$

where $\hat{G}$ is distribution function of booststrap replications $\hat{\theta}^*$, $\hat{z}_0$ and $\hat{a}$ are bias a acceleration and Φ is distribution function of standard normal distribution. It is possible to show that BC_a interval is also the second-order accuracy:

$$P(\theta \le \widehat{\theta}_{BC_a}[\alpha]) = \alpha + O(n^{-1}).$$

The bootstrap-*t* confidence interval limits and BC_a confidence interval limits are second-order accuracy:

$$\widehat{\theta}_{\mathrm{BCa}}[\alpha] = \widehat{\theta}_{\mathrm{STUD}}[\alpha] + O(n^{-1}),$$

so that according to definition of correctness accepted in the introductory part also $\hat{\theta}_{\mathrm{BC_a}}$ is of the second-order correctness. The proof of these matters is based on Edgeworth series of $C(x)$ and $D(x)$ and it is possible to read it in Hall (1988).

As bootstrap *t*- method and BC_a method enable us to obtain confidence intervals of the second-order accuracy, the main advantage of the BC_a method is its property of transformation respecting. BC_a interval parameter $\gamma = r(\theta)$, based on $\widehat{\gamma} = r(\widehat{\theta})$ (where r is increasing function) can be obtained, when we use transformation $r(.)$ for BC_a interval limits for θ based on $\hat{\theta}$. Bootstrap *t*-method doesn't respect transformation and can work not successfully if it is used with wrong scale. This method is correct for parameter's determination. The difficulty of its application is in specification of transformation $r(.)$. This problem can be solved by stabilization of variance method. Interval that arises by the help of this technique is as well of the second order accuracy and correctness.

Principles of BC_a interval construction

BC_a interval is based on following model. We assume the existence of such a transformation $r(.)$, that $\gamma = r(\theta)$, $\widehat{\gamma} = r(\widehat{\theta})$ $\phi = m(\theta)$, $\hat{\phi} = m(\hat{\theta})$ and stands

$$\frac{\widehat{\gamma} - \gamma}{se_\gamma} \approx N(-z_0, 1) \tag{6}$$

where $se_{\gamma} = se_{\gamma 0}.[1 + a(\gamma - \gamma_0)]$. γ_0 is some basic point in the scale of γ values.

The generalization involves three components that capture bias from ideal transformation: *transformation r(.)*, *bias correction* z_0 and *acceleration a.*

The percentile method generalizes normal approximation with admitting of transformation of *r(.)* by the help of θ and $\hat{\theta}$. BC$_a$ method add further frames of z_0 and *a,* both are $O_p(n^{-1/2})$. Bias correction z_0 explains the possible bias of estimation of $\hat{\gamma}$, while acceleration constant *a* explains possible changes of standard deviation of estimation of $\hat{\gamma}$.

Model (6) is used at solution of many problems and it is of the second-order correctness, which means that the approximation error is in principle $O_p(n^{-1})$. On the other hand the error of normal approximation is generally $O_p(n^{-1/2})$. This means that confidence intervals constructed under assumption (6) will be in principle of the second order accuracy and second order correctness. All three components in (6) are needed to reduction of error to $O_p(n^{-1})$.
We assume now that model (6) is exact. Then the exact limit of 1-α percentage confidence interval for γ is

$$\gamma[\alpha] = \hat{\gamma} + se_{\hat{\gamma}} \frac{z_0 + z^{(\alpha)}}{1 - a(z_0 + z^{(\alpha)})}.$$

Let's G is distribution function of bootstrap variable $\hat{\theta}$.
If we mark the interval limit $\gamma[\alpha]$ in θ scale by the help of inverse transformation $r^{-1}(.)$, we get

$$\theta[\alpha] = G^{-1}(\Phi(z_0 + \frac{z_0 + z^{(\alpha)}}{1 - a(z_0 + z^{(\alpha)})}).$$

This is the exact limit of BC$_a$ confidence interval. The distribution function G can be estimated with bootstrap distribution function G^*.

Let's Z is a variable with standard normal distribution with distribution function Φ, the estimate of z_0 can be obtained

$$P_\theta\{\hat{\theta} \le \theta\} = P_\gamma\{\hat{\gamma} \le \gamma\} = P\{Z \le z_0\} = \Phi(z_0).$$

We get by substitution of parameter θ with its estimation $\hat{\theta}$

$$\hat{z}_0 = \Phi^{-1}(P_{\hat{\gamma}}\{\hat{\theta}^* \le \hat{\theta}\}) = \Phi^{-1}(G(\hat{\theta})).$$

The acceleration constant a has meaning given by relation (6). It measures degree of changes of standard error in standardized scale. The good approximation for a in the one-parameter model is

$$\hat{a} = \frac{1}{6} SK_\theta = \ln L(\theta, X_i)$$

where $\ln L(\theta, X_i)$ is the likelihood function for estimation of θ.

4 Literature

[1.] DICICCIO,T.J., ROMANO,J.P. (1988) A review of bootstrap confidence intervals (with discussions). *J.Royal. Statisti. Soc*. B 50, 338-370

[2.] HALL,P.(1988) Theoretical comparison of bootstrap confidence intervals (with Discussion). *Annals of Statistics* 16, 927-985.

[3.] KUBANOVÁ,J.(2002) BC_a confidence intervals obtained with Unistat. In: *Sborník 11. medzinárodný seminár Výpočtová štatistika*. Bratislava,s.51-53, ISBN 80-88946-20-4, EAN 9788088946205

[4.] KUBANOVÁ,J.,LINDA,B.(2002) Confidence intervals based on bootstrap method. In.: *7th International Scientific Conference, Proceedings Section 1: Applied Mathematics*. Košice 2002, s.35-40, ISBN 80-7099-808-3

Contact addresses

doc. RNDr. Bohdan Linda, CSc.
Ústav matematiky, FES, Univerzita Pardubice
53210 Pardubice
tel. 00420 46 603 6020, e-mail: Bohdan.Linda@upce.cz

Paed.Dr. Jana Kubanová, CSc.
Ústav matematiky, FES, Univerzita Pardubice
53210 Pardubice
tel. 00420 46 603 6046, e-mail: Jana.Kubanova@upce.cz

Recenzoval: prof. RNDr. Otakar Prachař, CSc., Ústav matematiky, FES, UPa