Bootstrap methods for time series

Part 2: Bootstrap Confidence Interval

Bartosz Majewski

AGH University of Krakow,
Faculty of Applied Mathematics

bartmaje@agh.edu.pl

Syllabus

1. Introduction to Bootstrap
2. Bootstrap Confidence Interval
3. Block Bootstrap Methods
4. Bootstrap for Data with Periodic Structure

Bootstrap Confidence Interval

2.1. Bootstrap-t interval

2.2. Percentile confidence intervals

In this lecture, we assume that all bootstrap samples have the same size as an original sample. That is, $m = n$.

2.1. Bootstrap-t interval

Idea

Recall that to construct confidence intervals for the mean μ in a normal population with unknown variance, we use the following fact

$$Z = \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \sim t_{n-1}.$$

The obtained confidence interval is given

$$\left[\bar{X}_n - t_{n-1}(1-\alpha/2) \cdot S_n/\sqrt{n}, \bar{X}_n - t_{n-1}(\alpha/2) \cdot S_n/\sqrt{n}\right]$$

Idea

Let

- X_n – i.i.d. data

Idea

Let

- X_n – i.i.d. data
- θ – parameter of interest

Idea

Let

- X_n – i.i.d. data
- θ – parameter of interest
- $\widehat{\theta}_n$ - estimator of θ

Idea

Let

- X_n – i.i.d. data
- θ – parameter of interest
- $\widehat{\theta}_n$ - estimator of θ
- $\sigma_n^2 = \mathrm{Var}(\widehat{\theta}_n)$ – variance of $\widehat{\theta}_n$

Idea

Let

- X_n – i.i.d. data
- θ – parameter of interest
- $\widehat{\theta}_n$ - estimator of θ
- $\sigma_n^2 = \mathrm{Var}(\widehat{\theta}_n)$ – variance of $\widehat{\theta}_n$
- $\widehat{\sigma}_n^2$ – estimator of a variance estimator σ_n^2

Idea

Let

- X_n – i.i.d. data
- θ – parameter of interest
- $\widehat{\theta}_n$ - estimator of θ
- $\sigma_n^2 = \mathrm{Var}(\widehat{\theta}_n)$ – variance of $\widehat{\theta}_n$
- $\widehat{\sigma}_n^2$ – estimator of a variance estimator σ_n^2

In the bootstrap-*t* approach, we estimate the distribution of

$$Z = \frac{\widehat{\theta}_n - \theta}{\widehat{\sigma}_n}.$$

Specifically estimate the quantiles of Z to contract the confidence interval in the same way as in the previous slide.

Algorithm

1. Generate B bootstrap samples $\boldsymbol{X}_b^*$, $b = 1, 2, \ldots, B$.

Algorithm

1. Generate B bootstrap samples $\boldsymbol{X}_b^*$, $b = 1, 2, \ldots, B$.
2. For $b = 1, 2, \ldots, B$, compute the corresponding value of the statistics in each sample $\widehat{\theta}_b^*$, their mean

$$\bar{\theta}^* = \frac{1}{B} \sum_{b=1}^{B} \widehat{\theta}_b^*.$$

   Additionally, for each bootstrap sample, compute the standard error estimate $\widehat{\sigma}_n^{*(b)}$.

1. Generate B bootstrap samples $\boldsymbol{X}_b^*$, $b = 1, 2, \ldots, B$.
2. For $b = 1, 2, \ldots, B$, compute the corresponding value of the statistics in each sample $\widehat{\theta}_b^*$, their mean
$$\bar{\theta}^* = \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_b^*.$$
Additionally, for each bootstrap sample, compute the standard error estimate $\widehat{\sigma}_n^{*(b)}$.
3. For $b = 1, \ldots, B$ let
$$Z_b^* = \frac{\widehat{\theta}^{*(b)} - \widehat{\theta}_n}{\widehat{\sigma}^{*(b)}}$$
and let $Z_{(1)}^*, \ldots, Z_{(B)}^*$ be the order statistics. Calculate the $\alpha/2$ and $1 - \alpha/2$ sample quantiles of the bootstrap estimators
$$\widehat{z}^{*(\alpha/2)} = Z^*_{(\lfloor (B+1)\alpha/2 \rfloor)}, \quad \text{and} \quad \widehat{z}^{*(1-\alpha/2)} = Z^*_{(\lfloor (B+1)(1-\alpha/2) \rfloor)}.$$

Algorithm

1. Generate B bootstrap samples $\boldsymbol{X}_b^*$, $b = 1, 2, \ldots, B$.
2. For $b = 1, 2, \ldots, B$, compute the corresponding value of the statistics in each sample $\widehat{\theta}_b^*$, their mean
$$\bar{\theta}^* = \frac{1}{B}\sum_{b=1}^{B}\widehat{\theta}_b^*.$$
Additionally, for each bootstrap sample, compute the standard error estimate $\widehat{\sigma}_n^{*(b)}$.
3. For $b = 1, \ldots, B$ let
$$Z_b^* = \frac{\widehat{\theta}^{*(b)} - \widehat{\theta}_n}{\widehat{\sigma}^{*(b)}}$$
and let $Z_{(1)}^*, \ldots, Z_{(B)}^*$ be the order statistics. Calculate the $\alpha/2$ and $1 - \alpha/2$ sample quantiles of the bootstrap estimators
$$\widehat{z}^{*(\alpha/2)} = Z^*_{(\lfloor (B+1)\alpha/2 \rfloor)}, \quad \text{and} \quad \widehat{z}^{*(1-\alpha/2)} = Z^*_{(\lfloor (B+1)(1-\alpha/2) \rfloor)}.$$
4. Construct the bootstrap-t equal-tailed confidence interval of the form
$$\left[\widehat{\theta}_n - \widehat{z}^{*(1-\alpha/2)} \cdot \widehat{\sigma}_n, \widehat{\theta}_n - \widehat{z}^{*(\alpha/2)} \cdot \widehat{\sigma}_n\right]$$

2.2. Percentile confidence intervals

Idea

Let $\widehat{\theta}_n$ be the usual plug-in estimate of a parameter θ and $\widehat{\sigma}_n$ be its estimated standard error.

Idea

Let $\widehat{\theta}_n$ be the usual plug-in estimate of a parameter θ and $\widehat{\sigma}_n$ be its estimated standard error.

Consider the standard normal confidence interval

$$[\theta_{low}, \theta_{up}] = [\widehat{\theta}_n - z^{(1-\alpha/2)} \cdot \widehat{\sigma}_n, \widehat{\theta}_n - z^{(\alpha/2)} \cdot \widehat{\sigma}_n].$$

Idea

Let $\widehat{\theta}_n$ be the usual plug-in estimate of a parameter θ and $\widehat{\sigma}_n$ be its estimated standard error.

Consider the standard normal confidence interval

$$[\theta_{low}, \theta_{up}] = [\widehat{\theta}_n - z^{(1-\alpha/2)} \cdot \widehat{\sigma}_n, \widehat{\theta}_n - z^{(\alpha/2)} \cdot \widehat{\sigma}_n].$$

The endpoints of this interval can be described in a way that is particularly convenient for bootstrap calculations.

Let $\widehat{\theta}_n^*$ indicate a random variable from the distribution $N(\widehat{\theta}_n, \widehat{\sigma}_n^2)$,

Idea

Let $\widehat{\theta}_n$ be the usual plug-in estimate of a parameter θ and $\widehat{\sigma}_n$ be its estimated standard error.

Consider the standard normal confidence interval

$$[\theta_{low}, \theta_{up}] = [\widehat{\theta}_n - z^{(1-\alpha/2)} \cdot \widehat{\sigma}_n, \widehat{\theta}_n - z^{(\alpha/2)} \cdot \widehat{\sigma}_n].$$

The endpoints of this interval can be described in a way that is particularly convenient for bootstrap calculations.

Let $\widehat{\theta}_n^*$ indicate a random variable from the distribution $N(\widehat{\theta}_n, \widehat{\sigma}_n^2)$,

Then

- $\widehat{\theta}_{low} = \widehat{\theta}_n^*(\alpha/2) = 100 \cdot (\alpha/2)^{th}$ percentile of $\widehat{\theta}_n^*$'s distribution
- $\widehat{\theta}_{up} = \widehat{\theta}_n^*(1-\alpha/2) = 100 \cdot (1-\alpha/2)^{th}$ percentile of $\widehat{\theta}_n^*$'s distribution

Idea

Assume that we have a dataset $\boldsymbol{X}_n^*$ generated using the bootstrap method and computed bootstrap replications $\widehat{\theta}_n^*$.

Let the function $\widehat{G}_n$ be the distribution function of $\widehat{\theta}^*$.

Then, the $1-\alpha$ confidence interval is defined as

$$[\widehat{\theta}_{\%,low}, \widehat{\theta}_{\%,up}] = [\widehat{G}_n^{-1}(\alpha/2), \widehat{G}_n^{-1}(1-\alpha/2)]$$

Notice that $\widehat{G}_n^{-1}(\alpha) = \widehat{\theta}_n^*(\alpha)$ is the $100 \cdot \alpha$ percentile of the bootstrap distribution. Therefore, the percentile confidence interval can be written as:

$$[\widehat{\theta}_{\%,low}, \widehat{\theta}_{\%,up}] = [\widehat{\theta}_n^*(\alpha), \widehat{\theta}_n^*(1-\alpha)]$$

The above situation refers to the case where we have an infinite number of replications.

In practice, we use a finite number B.

Algorithm

1. Generate B bootstrap samples X_b^*, $b = 1, 2, \ldots, B$.
2. For $b = 1, 2, \ldots, B$, compute the corresponding value of the statistics in each sample $\widehat{\theta}_b^*$.
3. Let $\widehat{\theta}^*_{(1)}, \ldots, \widehat{\theta}^*_{(B)}$ be the order statistics. Calculate the $\alpha/2$ and $1 - \alpha/2$ sample quantiles of the bootstrap statistics

$$\widehat{\theta}^{*(\alpha/2)} = \widehat{\theta}^*_{(\lfloor (B+1)\alpha/2 \rfloor)}, \quad \text{and} \quad \widehat{\theta}^{*(1-\alpha/2)} = \widehat{\theta}^*_{(\lfloor (B+1)(1-\alpha/2) \rfloor)}.$$

4. The percentile confidence interval is given by

$$[\widehat{\theta}^{*(\alpha/2)}, \widehat{\theta}^{*(1-\alpha/2)}]$$

Literature

- Efron, B., & Tibshirani, R.J. (1994). *An Introduction to the Bootstrap*. Chapman and Hall/CRC.

- Simar, L. (2008). *An Invitation to the Bootstrap: Panacea for Statistical Inference?* Institut de Statistique, Université Catholique de Louvain Louvain-la-Neuve, Belgium. [link]

Course materials

Go to my website (https://home.agh.edu.pl/~ bartmaje/) to the section Teaching.

You can use the QR code below:

Thank you!