

Bootstrap Methods

Carsten Jentsch

Business and Social Statistics

Department of Statistics

TU Dortmund University

Summer Term 2024

The Bootstrap

Bootstrapping (Statistics) (engl. Wikipedia):

Bootstrapping is any test or metric that uses **random sampling with replacement** (e.g. mimicking the sampling process), and falls under the broader class of **resampling methods**. Bootstrapping assigns measures of accuracy (bias, variance, confidence intervals, prediction error, etc.) to sample estimates. This technique allows **estimation of the sampling distribution** of almost any statistic using random sampling methods.

A seminal idea

Pioneering work:

- Efron, B. (1979). The Bootstrap: another look at the jackknife, *The Annals of Statistics*, 7, 1-26.
- Efron, B. (1992) Bootstrap Methods: Another Look at the Jackknife. In: Kotz S., Johnson N.L. (eds) *Breakthroughs in Statistics. Springer Series in Statistics (Perspectives in Statistics)*. Springer, New York, NY.
- Efron, B. & Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman & Hall/CRC.

Huge impact (google scholar):

- The paper has about 22,000 citations
- The book has about 47,000 citations
- 780,000 sources when searching for 'Bootstrap' & 'statistics'

Recent paper:

- Andrew Gelman and Aki Vehtari (2020). What are the most important statistical ideas of the past 50 years?
 - ▶ 'Bootstrapping and simulation-based inference' is mentioned in second place (after 'counterfactual causal inference')

Course Content

General focus:

- This lecture discusses the origins of bootstrap methods pioneered by Bradley Efron in his seminal paper Efron (1979, Annals of Statistics). Besides its predecessor **Jackknife**, several **variants of bootstrap** resampling suitable for different data scenarios will be discussed. The main focus will be on the **bootstrap methodology** itself and on (asymptotic) properties of bootstrap methods. Their derivation requires **good (mathematical) knowledge of probability theory and asymptotic theory**. Students will gain insights why, how and when the bootstrap as a universal and powerful tool can be used to conduct **(valid) statistical inference**.

Course Content

Particular topics covered are:

- The Jackknife (bias-) estimator
- Efron's (original) Bootstrap
- Parametric and non-parametric bootstrap approaches
- Asymptotic bootstrap theory (first and higher order consistency, CLT)
- Bootstrap inconsistencies
- Bootstrap confidence intervals/hypothesis tests
- Bootstrap approaches for linear regression models and time series

Literature

- Davison & Hinkley (1997). Bootstrap Methods and Their Applications
- Efron & Tibshirani (1993). An Introduction to Bootstrap.
- Shao & Tu (1993). The Jackknife and Bootstrap.
- Chernick (2008). Bootstrap Methods - Guide for Practitioners and Researchers
- Politis, Romano & Wolf (1999). Subsampling.
- Horowitz (2001). The Bootstrap. Handbook of Econometrics, (Chapter 52), Vol. 5, 3159-3228.
- Lahiri (2003). Resampling of dependant data. Springer.

Admin

Lectures:

- on campus
- mixture of slides and writing on white/black board
- two lectures per week
- Tuesday 14-16, Thursday 10-12 (both in CDI 120)
- Moodle Room

Tutorials:

- M.Sc. Maxime Faymonville
- one tutorial per week
- on campus (if possible)
- weekly exercise sheet; solutions to be handed in
- corrections if number of participants allows, otherwise two graded tests during the semester
- 40% of the total points required to be eligible for the exam

Exam:

- oral exams (if number of participants allows)