Bootstrap

Ivan Jacob Agaloos Pesigan

August 28, 2023

References

Andrews: Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space **Andrews-2000**

Donald W. K. Andrews. "Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space". In: *Econometrica* 68.2 (Mar. 2000), pp. 399–405. DOI: 10.1111/1468-0262.00114.

Barnard et al.: Nonparametric standard errors and confidence intervals: Discussion **Barnard-Collins-Farewell-etal-1981**

George A. Barnard et al. "Nonparametric standard errors and confidence intervals: Discussion". In: *The Canadian Journal of Statistics / La Revue Canadienne de Statistique* 9.2 (1981), pp. 158–170. DOI: 10.2307/3314609.

Beran: The impact of the bootstrap on statistical algorithms and theory **Beran-2003**

Rudolf Beran. "The impact of the bootstrap on statistical algorithms and theory". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994972.

Abstract: Bootstrap ideas yield remarkably effective algorithms for realizing certain programs in statistics. These include the construction of (possibly simultaneous) confidences sets and tests in classical models for which exact or asymptotic distribution theory is intractable. Success of the bootstrap, in the sense of doing what is expected under a probability model for data, is not universal.

Modifications to Efron's definition of the bootstrap are needed to make the idea work for modern procedures that are not classically regular.

Boos: Introduction to the bootstrap world **Boos-2003**

Dennis D. Boos. "Introduction to the bootstrap world". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994971.

Abstract: The bootstrap has made a fundamental impact on how we carry out statistical inference in problems without analytic solutions. This fact is illustrated with examples and comments that emphasize the parametric bootstrap and hypothesis testing.

Efron: Bootstrap methods: Another look at the jackknife **Efron-1979a**

Bradley Efron. "Bootstrap methods: Another look at the jackknife". In: *The Annals of Statistics* 7.1 (Jan. 1979). DOI: 10.1214/aos/1176344552.

Abstract: We discuss the following problem: given a random sample $\boldsymbol{X} = (X_1, X_2, \ldots, X_n)$ from an unknown probability distribution F, estimate the sampling distribution of some prespecified random variable $R(\boldsymbol{X}, F)$, on the basis of the observed data $\boldsymbol{x}$. (Standard jackknife theory gives an approximate mean and variance in the case $R(\boldsymbol{X}, F) = \theta\left(\hat{F}\right) - \theta(F)$, θ some parameter of interest.) A general method, called the "bootstrap" is introduced, and shown to work satisfactorily on a variety of estimation problems. The jackknife is shown to be a linear approximation method for the bootstrap. The exposition proceeds by a series of examples: variance of the sample median, error rates in a linear discriminant analysis, ratio estimation, estimating regression parameters, etc.

Efron: Computers and the theory of statistics: Thinking the unthinkable **Efron-1979b**

Bradley Efron. "Computers and the theory of statistics: Thinking the unthinkable". In: *SIAM Review* 21.4 (Oct. 1979), pp. 460–480. DOI: 10.1137/1021092.

Abstract: This is a survey article concerning recent advances in certain areas of statistical theory, written for a mathematical audience with no background in statistics. The topics are chosen to illustrate a special point: how the advent of the high-speed computer has affected the development of statistical theory. The topics discussed include nonparametric methods, the jackknife, the bootstrap, cross-validation, error-rate estimation in discriminant analysis, robust estimation, the influence function, censored data, the EM algorithm, and Cox's likelihood function. The exposition is mainly by example, with only a little offered in the way of theoretical development.

Efron: Nonparametric standard errors and confidence intervals **Efron-1981a**

Bradley Efron. "Nonparametric standard errors and confidence intervals". In: *Canadian Journal of Statistics / La Revue Canadienne de Statistique* 9.2 (1981), pp. 139–158. DOI: 10.2307/3314608.

Abstract: We investigate several nonparametric methods; the bootstrap, the jackknife, the delta method, and other related techniques. The first and simplest goal is the assignment of nonparametric standard errors to a real-valued statistic. More ambitiously, we consider setting nonparametric confidence intervals for a real-valued parameter. Building on the well understood case of confidence intervals for the median, some hopeful evidence is presented that such a theory may be possible.

Efron: Nonparametric standard errors and confidence intervals: Rejoinder **Efron-1981b**

Bradley Efron. "Nonparametric standard errors and confidence intervals: Rejoinder". In: *The Canadian Journal of Statistics / La Revue Canadienne de Statistique* 9.2 (1981), pp. 170–172. DOI: 10.2307/3314610.

Efron: Better bootstrap confidence intervals **Efron-1987**

Bradley Efron. "Better bootstrap confidence intervals". In: *Journal of the American Statistical Association* 82.397 (Mar. 1987), pp. 171–185. DOI: 10.1080/01621459.1987.10478410.

Abstract: We consider the problem of setting approximate confidence intervals for a single parameter θ in a multiparameter family. The standard approximate intervals based on maximum likelihood theory, $\hat{\theta} \pm \hat{\sigma} z^{(\alpha)}$, can be quite misleading. In practice, tricks based on transformations, bias corrections, and so forth, are often used to improve their accuracy. The bootstrap confidence intervals discussed in this article automatically incorporate such tricks without requiring the statistician to think them through for each new application, at the price of a considerable increase in computational effort. The new intervals incorporate an improvement over previously suggested methods, which results in second-order correctness in a wide variety of problems. In addition to parametric families, bootstrap intervals are also developed for nonparametric situations.

Efron: Bootstrap confidence intervals: Good or bad? **Efron-1988**

Bradley Efron. "Bootstrap confidence intervals: Good or bad?" In: *Psychological Bulletin* 104.2 (1988), pp. 293–296. DOI: 10.1037/0033-2909.104.2.293.

Abstract: The bootstrap is a nonparametric technique for estimating standard errors and approximate confidence intervals. Rasmussen has used a simulation experiment to suggest that bootstrap confidence intervals perform very poorly in the estimation of a correlation coefficient. Part of Rasmussen's simulation is repeated. A careful look at the results shows the bootstrap intervals performing quite well. Some remarks are made concerning the virtues and defects of bootstrap intervals in general.

Efron: Bayesian inference and the parametric bootstrap **Efron-2012**

Bradley Efron. "Bayesian inference and the parametric bootstrap". In: *The Annals of Applied Statistics* 6.4 (Dec. 2012). DOI: 10.1214/12-aoas571.

Abstract: The parametric bootstrap can be used for the efficient computation of Bayes posterior distributions. Importance sampling formulas take on an easy form relating to the deviance in exponential families and are particularly simple starting from Jeffreys invariant prior. Because of the i.i.d. nature of bootstrap sampling, familiar formulas describe the computational accuracy of the Bayes estimates. Besides computational methods, the theory provides a connection between Bayesian and frequentist analysis. Efficient algorithms for the frequentist accuracy of Bayesian inferences are developed and demonstrated in a model selection example.

Ernst et al.: Utilizing a quantile function approach to obtain exact bootstrap solutions Ernst-Hutson-2003

Michael D. Ernst and Alan D. Hutson. "Utilizing a quantile function approach to obtain exact bootstrap solutions". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994978.

Abstract: The popularity of the bootstrap is due in part to its wide applicability and the ease of implementing resampling procedures on modern computers. But careful reading of Efron (1979) will show that at its heart, the bootstrap is a "plug-in" procedure that involves calculating a functional $\theta\left(\hat{F}\right)$ from an estimate of the c.d.f. F. Resampling becomes invaluable when, as is often the case, $\theta\left(\hat{F}\right)$ cannot be calculated explicitly. We discuss some situations where working with the sample quantile function, $\hat{Q}$, rather than $\hat{F}$, can lead to explicit (exact) solutions to $\theta\left(\hat{F}\right)$.

Hall: A short prehistory of the bootstrap Hall-2003

Peter Hall. "A short prehistory of the bootstrap". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994970.

Abstract: The contemporary development of bootstrap methods, from the time of Efron's early articles to the present day, is well documented and widely appreciated. Likewise, the relationship of bootstrap techniques to certain early work on permutation testing, the jackknife and cross-validation is well understood. Less known, however, are the connections of the bootstrap to research on survey sampling for spatial data in the first half of the last century or to work from the 1940s to the 1970s on subsampling and resampling. In a selective way, some of these early linkages will be explored, giving emphasis to developments with which the statistics community tends to be less familiar.

Particular attention will be paid to the work of P. C. Mahalanobis, whose development in the 1930s and 1940s of moving-block sampling methods for spatial data has a range of interesting features, and to contributions of other scientists who, during the next 40 years, developed half-sampling, subsampling and resampling methods.

Hesterberg: What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum **Hesterberg-2014**

Tim C. Hesterberg. *What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum*. 2014. arXiv: 1411.5279 [stat.OT]. URL: https://arxiv.org/abs/1411.5279.

Abstract: I have three goals in this article:

1. To show the enormous potential of bootstrapping and permutation tests to help students understand statistical concepts including sampling distributions, standard errors, bias, confidence intervals, null distributions, and P-values.
2. To dig deeper, understand why these methods work and when they don't, things to watch out for, and how to deal with these issues when teaching.
3. To change statistical practice—by comparing these methods to common t tests and intervals, we see how inaccurate the latter are; we confirm this with asymptotics. $n \geq 30$ isn't enough—think $n \geq 5000$.

Resampling provides diagnostics, and more accurate alternatives. Sadly, the common bootstrap percentile interval badly under-covers in small samples; there are better alternatives. The tone is informal, with a few stories and jokes.

Hesterberg: What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum **Hesterberg-2015**

Tim C. Hesterberg. "What teachers should know about the bootstrap: Resampling in the under-

graduate statistics curriculum". In: *The American Statistician* 69.4 (Oct. 2015), pp. 371–386. DOI: 10.1080/00031305.2015.1089789.

Abstract: Bootstrapping has enormous potential in statistics education and practice, but there are subtle issues and ways to go wrong. For example, the common combination of nonparametric bootstrapping and bootstrap percentile confidence intervals is less accurate than using t-intervals for small samples, though more accurate for larger samples. My goals in this article are to provide a deeper understanding of bootstrap methods–how they work, when they work or not, and which methods work better-and to highlight pedagogical issues. Supplementary materials for this article are available online.

Holmes: Bootstrapping phylogenetic trees: Theory and methods **Holmes-2003a**

Susan Holmes. "Bootstrapping phylogenetic trees: Theory and methods". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994979.

Abstract: This is a survey of the use of the bootstrap in the area of systematic and evolutionary biology. I present the current usage by biologists of the bootstrap as a tool both for making inferences and for evaluating robustness, and propose a framework for thinking about these problems in terms of mathematical statistics.

Holmes: Bradley Efron: A conversation with good friends **Holmes-2003b**

Susan Holmes. "Bradley Efron: A conversation with good friends". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994981.

Abstract: Bradley Efron is Professor of Statistics and Biostatistics at Stanford University. He works on a combination of theoretical and applied topics, including empirical Bayes, survival analysis, exponential families, bootstrap and jackknife methods and confidence intervals. Most of his applied work has originated in biomedical consulting projects at the Stanford Medical School, mixed

in with a few papers concerning astronomy and physics. Even his theoretical papers usually begin with specific applied problems. All three of the interviewers here have been close scientific collaborators. Brad was born in St. Paul, Minnestora, May 1938, to Esther and Miles Efron, Jewish-Russian immigrants. A Merit Scholarship, in the program's inaugural year, brought him to Caltech, graduating in Mathematics in 1960. He arrived at Stanford that Fall, eventually gaining his Ph.D., under the direction of Rupert Miller and Herb Solomon, in the Statistics Department, whose faculty also included Charles Stein, Herman Chernoff, Manny Parzen, Lincoln Moses and Ingram Olkin. Brad has lived at Stanford since 1960, with sabbaticals at Harvard, Imperial College and Berkeley. He has held several administrative positions in the university: Chair of Statistics, Associate Dean of Science, Chairman of the University Advisory Board and Chair of the Faculty Senate. He is currently Chair of the Undergraduate Program in Applied Mathematics. Honors include doctorates from Chicago, Madrid and Oslo, a MacArthur Prize Fellowship, membership in the National Academy of Sciences and the American Academy of Arts and Sciences, fellowship in the IMS and ASA, the Wilks Medal, Parzen Prize, the newly inaugurated Rao Prize and the outstanding statistician award from the Chicago ASA chapter. He has been the Rietz, Wald, and Fisher lecturers and holds the Max H. Stein endowed chair as Professor of Humanities and Sciences at Stanford. Professional service includes Theory and Methods Editor of JASA and President of the IMS. Currently he is President-Elect of the American Statistical Association, becoming President in 2004.

Horowitz: The bootstrap in econometrics — Horowitz-2003

Joel L. Horowitz. "The bootstrap in econometrics". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994976.

Abstract: This paper presents examples of problems in estimation and hypothesis testing that demonstrate the use and performance of the bootstrap in econometric settings. The examples are illustrated with two empirical applications. The paper concludes with a discussion of topics on which further research is needed.

Lahiri: On the impact of bootstrap in survey sampling and small-area estimation **Lahiri-2003**

Partha Lahiri. "On the impact of bootstrap in survey sampling and small-area estimation". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994975.

Abstract: Development of valid bootstrap procedures has been a challenging problem for survey samplers for the last two decades. This is due to the fact that in surveys we constantly face various complex issues such as complex correlation structure induced by the survey design, weighting, imputation, small-area estimation, among others. In this paper, we critically review various bootstrap methods developed to deal with these challenging issues. We discuss two applications where the bootstrap has been found to be effective.

Lele: Impact of bootstrap on the estimating functions **Lele-2003**

Subhash R. Lele. "Impact of bootstrap on the estimating functions". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994973.

Abstract: Estimating functions form an attractive statistical methodology because of their dependence on only a few features of the underlying probabilistic structure. They also put a premium on developing methods that obtain model-robust confidence intervals. Bootstrap and jackknife ideas can be fruitfully used toward this purpose. Another important area in which bootstrap has proved its use is in the context of detecting the problem of multiple roots and searching for the consistent root of an estimating function. In this article, I review, compare and contrast various approaches for bootstrapping estimating functions.

Politis: The impact of bootstrap methods on time series analysis **Politis-2003**

Dimitris N. Politis. "The impact of bootstrap methods on time series analysis". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994977.

Abstract: Sparked by Efron's seminal paper, the decade of the 1980s was a period of active research on bootstrap methods for independent data–mainly i.i.d. or regression set-ups. By contrast, in the 1990s much research was directed towards resampling dependent data, for example, time series and random fields. Consequently, the availability of valid nonparametric inference procedures based on resampling and/or subsampling has freed practitioners from the necessity of resorting to simplifying assumptions such as normality or linearity that may be misleading.

Rasmussen: Estimating correlation coefficients: Bootstrap and parametric approaches Rasmussen-1987

Jeffrey L. Rasmussen. "Estimating correlation coefficients: Bootstrap and parametric approaches". In: *Psychological Bulletin* 101.1 (1987), pp. 136–139. DOI: 10.1037/0033-2909.101.1.136.

Abstract: The bootstrap, a computer-intensive approach to statistical data analysis, has been recommended as an alternative to parametric approaches. Advocates claim it is superior because it is not burdened by potentially unwarranted normal theory assumptions and because it retains information about the form of the original sample. Empirical support for its superiority, however, is quite limited. The present article compares the bootstrap and parametric approaches to estimating confidence intervals and Type I error rates of the correlation coefficient. The parametric approach is superior to the bootstrap under both assumption violation and nonviolation. The bootstrap results in overly restricted confidence intervals and overly liberal Type I error rates.

Rousselet et al.: The percentile bootstrap: A primer with step-by-step instructions in R Rousselet-Pernet-Wilcox-2021

Guillaume A. Rousselet, Cyril R. Pernet, and Rand R. Wilcox. "The percentile bootstrap: A primer with step-by-step instructions in R". In: *Advances in Methods and Practices in Psychological Science* 4.1 (Jan. 2021), pp. 1–10. DOI: 10.1177/2515245920911881.

Abstract: The percentile bootstrap is the Swiss Army knife of statistics: It is a nonparametric method based on data-driven simulations. It can be applied to many statistical problems, as a substitute to standard parametric approaches, or in situations for which parametric methods do not exist. In this Tutorial, we cover `R` code to implement the percentile bootstrap to make inferences about central tendency (e.g., means and trimmed means) and spread in a one-sample example and in an example comparing two independent groups. For each example, we explain how to derive a bootstrap distribution and how to get a confidence interval and a p value from that distribution. We also demonstrate how to run a simulation to assess the behavior of the bootstrap. For some purposes, such as making inferences about the mean, the bootstrap performs poorly. But for other purposes, it is the only known method that works well over a broad range of situations. More broadly, combining the percentile bootstrap with robust estimators (i.e., estimators that are not overly sensitive to outliers) can help users gain a deeper understanding of their data than they would using conventional methods.

Schenker: Better bootstrap confidence intervals: Comment **Schenker-1987**

Nathaniel Schenker. “Better bootstrap confidence intervals: Comment”. In: *Journal of the American Statistical Association* 82.397 (Mar. 1987), p. 192. DOI: 10.2307/2289150.

Shao: Impact of the bootstrap on sample surveys **Shao-2003**

Jun Shao. “Impact of the bootstrap on sample surveys”. In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994974.

Abstract: This article discusses the impact of the bootstrap on sample surveys and introduces some of the main developments of the bootstrap methodology for sample surveys in the last twenty five years.

P. S. Soltis et al.: Applying the Bootstrap in Phylogeny Reconstruction

Soltis-Soltis-2003

Pamela S. Soltis and Douglas E. Soltis. "Applying the Bootstrap in Phylogeny Reconstruction". In: *Statistical Science* 18.2 (May 2003). DOI: 10.1214/ss/1063994980.

Abstract: With the increasing emphasis in biology on reconstruction of phylogenetic trees, questions have arisen as to how confident one should be in a given phylogenetic tree and how support for phylogenetic trees should be measured. Felsenstein suggested that bootstrapping be applied across characters of a taxon-by-character data matrix to produce replicate "bootstrap data sets," each of which is then analyzed phylogenetically, with a consensus tree constructed to summarize the results of all replicates. The proportion of trees/replicates in which a grouping is recovered is presented as a measure of support for that group. Bootstrapping has become a common feature of phylogenetic analysis. However, the interpretation of bootstrap values remains open to discussion, and phylogeneticists have used these values in multiple ways. The usefulness of phylogenetic bootstrapping is potentially limited by a number of features, such as the size of the data matrix and the underlying assumptions of the phylogeny reconstruction program. Recent studies have explored the application of bootstrapping to large data sets and the relative performance of bootstrapping and jackknifing.