# Bootstrapping

MICHAEL SCHARKOW
*Zeppelin University, Germany*

## Basics of the bootstrap method

The bootstrap method was introduced by Efron (1979) as a generalization of the older jackknife approach which is also based on repeated resampling of the original data. The basic idea of bootstrapping is simple: If, in the frequentist tradition, inferences are based on repeatedly running the same experiment or sampling from the same population, then a researcher can mimic this process by repeatedly sampling (with replacement) from the data and analyzing these *bootstrap samples* in lieu of independent samples from the population. Put differently, if bootstrap samples relate to the actual sample data in the same way that the sample data relate to the unknown population, it is possible to estimate the actual sampling distribution from the bootstrap sample distribution.

The basic (nonparametric) bootstrap consists of the following steps:

1 From a given sample of size *n*, draw a sample with *n* elements, which requires sampling with replacement. This is called a bootstrap sample.
2 From the bootstrap sample, compute one or multiple statistics of interest, such as the median, mean, correlation, or regression coefficient.
3 Repeat steps 1 and 2 multiple times and save the statistics of interest for every replication. Typically a few hundred or thousand bootstrap samples are generated, depending on the complexity of the computation required in step 2.
4 Summarize the distribution of the statistics of interest across all bootstrap samples, often in terms of confidence intervals using percentiles from the bootstrap distribution.

In the parametric bootstrap variant, the first (resampling) step differs from the basic bootstrap approach: Instead of sampling cases with replacement from the data, the bootstrap samples are generated from the empirical distribution function of the data, for example, a normal distribution with the sample mean and variance.

The bootstrap method has the advantage that it is conceptually and technically very simple, although the procedure can be computationally expensive. However, with modern computer hardware, many statistics can be bootstrapped a thousand times or more in mere seconds or minutes. Since a bootstrap distribution can be computed for any sample statistic and any derived quantities of interest, bootstrapping can

*The International Encyclopedia of Communication Research Methods*. Jörg Matthes (General Editor), Christine S. Davis and Robert F. Potter (Associate Editors).

DOI: 10.1002/9781118901731.iecrm0017

always be used as an alternative to asymptotic approaches. Moreover, bootstrapping is often used in cases where no closed-form solution is feasible and/or the sampling distribution of the statistic of interest is unknown, for example the median. Finally, even when sample distributions are skewed or truncated, using bootstrapping with a nonparametric confidence interval guarantees that confidence limits are within the range of the empirical data.

Although bootstrapping requires researchers to make relatively few assumptions, there are some issues to consider when using the method for statistical inference (see Mooney & Duval, 1993; Davison & Hinkley, 1997): (i) Just like asymptotic inference, bootstrapping assumes that the sample data are representative of the population data. If the available data are not a representative sample of the population, resampling from these cases will not lead to accurate estimations of the true population values. (ii) Even if the sample is representative, bootstrapping can result in an underestimation of rare events or extreme values, leading to confidence intervals that are too narrow. (iii) In very small samples ($n < 20$), simple nonparametric bootstrapping can lead to bias and additional correction mechanisms are required. (iv) In contrast to asymptotic procedures, there are no guarantees for coverage, Type I and II errors, and so on, when using bootstrap inference. Therefore, Monte Carlo simulations are often used in order to judge the performance of different approaches for a given statistical model and/or sampling distribution. Since bootstrapping depends on the quality of the sample data, it is often advisable to check for influential cases as well as distributional assumptions when using parametric approaches.

## Bootstrap confidence intervals

Although it is possible to compute bootstrap *p*-values by simply counting the occurrence of a specific result, for example, a statistic being larger or smaller than a defined value, across all bootstrap samples, confidence intervals are frequently reported when using the bootstrap method. In general, one can distinguish parametric and nonparametric approaches when constructing the confidence interval. While the former only uses mean and variance estimates from the bootstrap sampling distribution to construct conventional CI, the latter uses the data from the bootstrap replications directly to establish confidence limits. A large number of different methods of constructing these bootstrap confidence intervals have been proposed in the literature, among them the parametric standard and Studentized CI, as well as the nonparametric percentile, bias-corrected (BC) and the bias-corrected and accelerated (BCa) CI.

*Standard bootstrap CI*: A standard confidence interval is obtained by using the statistic of interest and its standard deviation across all bootstrap samples with a standard-normal distribution and computing the confidence limit as usual. A variant of this, the Studentized bootstrap CI, uses Student's *t*- rather than a *z* distribution. This approach assumes that the bootstrap distribution is approximately normal.

*Percentile bootstrap CI*: The limits of the confidence interval are simply looked up from the bootstrap distribution using the required percentile, for example, the 2.5th and 97.5th percentile for a 95% confidence interval. This approach assumes that the bootstrap distribution is symmetrical and that the bootstrap distribution mean is equal to the sample mean.

*Bias-corrected bootstrap (and accelerated) CI*: The BC and BCa confidence intervals extend the percentile CI by correcting the percentiles for the confidence limits when the bootstrap mean deviates from the sample mean, and accounting for the skew in the bootstrap distribution by estimating an acceleration constant. Both BC and BCa have been shown to be more robust when the bootstrap distribution deviates strongly from normality, for example, in very small samples.

The choice of confidence interval depends on given data and statistic(s) of interest as well as the requirements of the statistical analysis, that is, whether one values Type I or Type II errors, coverage, or the width of the interval (Mooney & Duval, 1993). For example, in mediation analysis (see later), the bias-corrected bootstrap CI has most power to detect an indirect effect, but an inflated Type I error rate, whereas the percentile bootstrap has better coverage and Type I error rates. In general, the percentile bootstrap CI is the easiest to implement and understand, and performs well in most circumstances. If the statistic of interest likely has a skewed sampling distribution, bias corrected confidence intervals are preferred.

## Applications in communication research

The most frequent application of bootstrapping in communication research is the estimation of confidence intervals for indirect effects in mediation models. Popularized by Preacher and Hayes (2004) who provided macros for SPSS and SAS, the use of bootstrap CI when judging the statistical significance of an indirect effect has become the de facto standard for mediation analysis. Since the bootstrap method is very versatile and simple to implement, it is widely applicable to all kinds of conditional process models (Hayes, 2013), which includes serial and parallel mediation as well as moderated mediation.

A second frequent application of bootstrapping lies in the area of structural equation modeling (SEM), an approach that is increasingly used in communication research and other social sciences (Holbert & Stephenson, 2002). Since SEM assumes multivariate normality of the data, which is almost always not the case, bootstrapping provides a robust alternative to asymptotic inference for both parameter estimates such as factor loadings or path coefficients, as well as goodness-of-fit measures (Bollen & Stine, 1992; Finney & DiStefano, 2006). In addition, the partial least squares (PLS) approach to fitting path models also relies on bootstrapping as the primary means of statistical inference.

A third application of the bootstrap method has been suggested by Hayes and Krippendorff (2007) in the context of intercoder reliability testing, in particular for Krippendorff's alpha coefficient, for which a SPSS macro was provided. In the same way, one can use bootstrapping in order to compute confidence intervals for any other agreement measure, provided that the reliability test is based on a representative sample of the coding data. However, since little is known about the sampling distributions of common intercoder reliability statistics, and reliability tests are often conducted with very small samples, more research on appropriate confidence intervals is needed.

## Availability in common software packages

Most statistical software packages used in communication research provide bootstrapping methods for common models and test statistics. SPSS introduced the Bootstrapping module in version 20, which adds the option to compute bootstrap standard errors and confidence intervals for the most frequently used types of analysis. Moreover, Hayes' (2013) *Process* macro for SPSS provides bootstrap CI for mediation and moderation analyses by default. Likewise, Stata (since version 13) and R (via the included *boot* package) provide generic functions for bootstrapping any statistic of interest. Bootstrapping of parameter estimates and/or the Bollen-Stine chi-squared goodness-of-fit statistic is available in all commonly used software packages for structural equation modeling such as LISREL, Mplus, EQS, AMOS, or the *lavaan* package for R.

SEE ALSO: Conditional Process Modeling (Mediation Analysis, Moderated Mediation Analysis, Moderation Analysis, and Mediated Moderation Analysis); Statistical Significance (Testing); Statistics, Inferential; Structural Equation Modeling

### References

Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research, 21*(2), 205–229. doi:10.1177/0049124192021002004

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application. Cambridge series in statistical and probabilistic mathematics*. Cambridge, UK: Cambridge University Press.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics, 7*(1), 1–26. doi:10.1214/aos/1176344552

Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In *Structural equation modeling: A second course* (pp. 269–314). Charlotte, NC: Information Age Publishing.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*(1), 77–89. doi:10.1080/19312450709336664

Holbert, R. L., & Stephenson, M. T. (2002). Structural equation modeling in the communication sciences, 1995–2000. *Human Communication Research, 28*(4), 531–551. doi: 10.1111/j.1468-2958.2002.tb00822.x

Mooney, C. Z., & Duval, R. (1993). *Bootstrapping: A non-parametric approach to statistical inference*. Newbury Park, CA: SAGE.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers, 36*(4), 717–731. doi:10.3758/BF03206553

## Further reading

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application. Cambridge series in statistical and probabilistic mathematics*. Cambridge, UK: Cambridge University Press.
Mooney, C. Z., & Duval, R. (1993). *Bootstrapping: A non-parametric approach to statistical inference*. Newbury Park, CA: SAGE.

**Michael Scharkow** (PhD, University of the Arts Berlin) is professor of communication at Zeppelin University, Germany. His research interests include empirical research methods, online communication, and media use.