An Annotated Bibliography for Bootstrap Resampling

For a first pass, I suggest that you concentrate on the articles marked with an asterisk. These articles are available from the ICPSR library. Other papers present theory or applications that go beyond the lecture material. Some of these are nonetheless interesting and potentially useful in special problems. (Abbreviations: *JASA* =*J. of the American Statistical Association*, *AS* =*Annals of Statistics*, and *SS* =*Statistical Science*.)

Beran, R. (1984). Jackknife approximations to bootstrap estimates. *AS,* 12, 101-18.

Explores the connection between jackknife and bootstrap methods. Technical, like most in *AS*.

——— (1987). Prepivoting to reduce level error of confidence sets. *Biometrika*, 74, 457-68.

Describes how to enhance the coverage of bootstrap confidence intervals by a clever simulation device. The idea is to improve the bootstrap by using the bootstrap approach itself.

——— (1990). Refining bootstrap simultaneous confidence sets. *JASA*, 85, 417-426.

Further study of his iterated approach to improving the coverage of bootstrap intervals.

Beran, R. and M. S. Srivastava (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. *AS*, 13, 95-115.

If you are interested in applications to structural equation models and LISREL, this is the foundation. This article also illustrates a problem in which the naive bootstrap algorithm fails.

Bickel, P. and D. Freedman (1981). Some asymptotic theory for the bootstrap. *AS*, 9 , 1196-1217.

A very elegant, technical introduction to the mathematics of the bootstrap. Useful, but more as a work of art. Recent papers weaken the assumptions of this paper and obtain more general results.

*Bollen, K. A. and R. A. Stine (1990). Direct and indirect effects: classical and bootstrap estimates of variability. *Sociological Methodology*, C. Clogg, Ed, American Sociological Association, Washington, DC.

Indirect effects are sums of products of regression coefficients in recursive models. The standard delta method for estimating variability ignores evident skewness and can be very misleading.

——— (1992). Bootstrapping goodness of fit measures in structural equation models. In *Testing Structural Equation Models*, K. A. Bollen and J. S. Long, Eds., Sage, Newbury Park.

Naive resampling fails when computing p-values, be it in simple models or structural equations.

Booth, J.G., and P. Hall (1994). Monte Carlo approximation and the iterated bootstrap. *Biometrika*, 81, 331--340.

If the bootstrap can discover problems with other estimators, why not use it to check itself – the “double bootstrap.” A computer salesperson’s dream! Well, maybe not if you are careful.

Booth, J. G. and S. Sarkar (1998). Monte Carlo approximation of bootstrap variances. *Amer Stat* 52, 354.

Efron claims you need about 200 bootstrap replications for standard error estimates. This paper argues for many more – close to 800 or more. It all depends on whether you take the marginal (Efron) or conditional approach.

Breidt, F. J., R.A.. Davis, W. T. Dunsmuir (1995). Improved bootstrap prediction intervals for autoregressions. *J. Time Series Analysis* 16, 177--200.

An updated look at using the bootstrap to compute reliable prediction intervals.

Breiman, L. (1992). The little bootstrap and other methods for dimensionality selection in regression: X-fixed prediction error.

The bootstrap can help diagnose the predictive ability of a model, if you're very careful.

Cleveland, W. S. (1985). *The Elements of Graphing Data*. Wadsworth, Monterey.

This book shows a number of schemes for presenting accurate, informative plots of data. Less elegant than Tufte, but often more useful.

*Davison, A. C. and D. V. Hinkley (1997). *Bootstrap Methods and their Application*. Cambridge Univ. Press.

A comprehensive survey of applications (with theory) of the bootstrap, including the double bootstrap, smoothing, models selection, and time series (with a library of S+ functions). Paperback!

*Diaconis, P. and B. Efron (1983). Computer intensive methods in statistics. *Scientific American*, 248:5, 116-30.

A gentle introduction to the bootstrap. It makes me wonder what biologists think about biology articles that appear in *Scientific American*.

DiCicco, T. and R. Tibshirani (1987). Bootstrap confidence intervals and bootstrap approximations. *JASA*, 82, 163-70.

Explains accelerated bootstrap confidence intervals and the reasons that they can be effective.

*Efron, B. (1979). Computers and the theory of statistics: thinking the unthinkable. *Siam Review*, 21, 460-80.

The underlying concepts of the bootstrap appear in this early article in which Efron anticipates the effects of cheap computing on statistics.

——— (1979). Bootstrap methods: another look at the jackknife. *AS*, 7, 1-26.

The first, largely read introduction to the bootstrap. The bootstrap had been around, but this was the big introduction.

——— (1982). *The Jackknife, the Bootstrap, and Other Resampling Plans*. CBMS 38, SIAM-NSF.

This monograph details the bootstrap and its relationship to 'more familiar' methods like the jackknife and cross validation.

——— (1986). How biased is the apparent error rate of a prediction rule. *JASA*, 81, 461-469.

If you use discriminant analysis or logistic regression, look at this paper to appreciate how optimistic the claimed classification rates can be. The bootstrap is related to cross-validation, and this paper elaborates the tie.

——— (1987). Better bootstrap confidence intervals (with discussion). *JASA*, 82, 171-200.

Efron's description of accelerated BS intervals. Not easy reading, but the introduction is OK.

——— (1988). Bootstrap confidence intervals: good or bad? (with disc). *Psychology Bulletin,* 104, 293-296.

Part of an interesting discussion of the merits of bootstrap confidence intervals.

——— (1990). More efficient bootstrap computations. *JASA*, 85, 79-89.

As the bootstrap has gained acceptance, more research has been devoted to speeding the necessary computations. Here is Efron's look at how to make things happen faster. The material depends upon the ANOVA decomposition described in Efron's manuscript.

——— (1994). Missing data, imputation, and the bootstrap (with discussion). *JASA*, 89, 463-479.

Everything from the title, plus a nice overview of bootstrap confidence intervals.

*Efron, B. and G. Gong (1983). A leisurely look at the bootstrap, the jackknife, and cross-validation. *American Statistician*, 37, 36-48.

A survey article, with a nice application to model selection (picking variables) in logistic regression.

*Efron, B. and R. Tibshirani (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *SS*, 1, 54-75.

A look at the bootstrap in time series models, survival analysis, and confidence intervals.

*——— (1991). Statistical analysis in the computer age. *Science*, 253, 390-395.

A short, clear introduction to bootstrapping, written for a results-oriented group of scientists.

*——— (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York.

A comprehensive review of bootstrap resampling, circa 1993. While occasionally technical, the authors motivate the mathematics with numerous illustrative examples. The bibliography is quite complete, but you might prefer the more recent book of Davison and Hinkley.

Faraway, J.J. (1992). On the cost of data analysis. *J. Comp. & Graphical Statis*, 213-229.

This paper assesses the effects on SE's of "all those little things" that we do when we analyze data.

*Finifter, B. M. (1972). The generation of confidence: evaluating research findings by random subsample replication. *Sociological Methodology*, H. L. Costinue, Ed., Jossey-Bass, San Francisco.

Believe it or not, Finifter anticipates the bootstrap from a data analysis/survey analysis perspective. He even coins the word *bootstrap*.

Fisher, N. I. and P. Hall (1990). On bootstrap hypothesis testing. *Australian Journal of Statistics*, 32, 177-190.

Fisher and Hall show examples of the bootstrap as a means to better p-values, mostly in ANOVA, but easily extensible to regression. Proof that drinking large quantities of beer is not harmful.

Fox, J. (1991). *Regression Diagnostics*. Sage, Newbury Park.

An inexpensive, concise look at regression diagnostics, with examples using interesting data. The methods include leverage, influence, transformation, and the various types of residual plots.

Fox, J. and J. S. Long (1990). *Modern Methods of Data Analysis*. Sage, Newbury Park.

A nice collection of articles (after all, I wrote the one on bootstrapping) including bootstrapping, regression, nonparametric regression, plotting, and robust regression.

Freedman, D. A. (1981). Bootstrapping regression models. *AS*, 9, 1218-28.

The initial definitive technical exposition of bootstrapping in regression models. Includes an example of problems with the bootstrap in models without an intercept.

——— (1984). On bootstrapping two-stage least squares estimates in stationary linear models. *AS,* 12, 827-42.

So you really want to use a lagged dependent variable in that regression equation...

——— (1987). As others see us: a case study in path analysis. *J. of Educational Statistics*, 12, 101-128.

Freedman takes on modern social science with this one. Read it if you dare. The discussion is very interesting – how would you defend yourself in this case?

*Freedman, D. A. and S. C. Peters (1982). Bootstrapping a regression equation: some empirical results. *JASA*, 79, 97-106.

An application of bootstrap methods to simultaneous equations with heteroscedasticity and repeated measures. Classical estimates of variability are far too small, and even the bootstrap comes up short.

Hall, P. (1986). On the number of bootstrap simulations required to construct a confidence interval. *AS*, 14, 1453-62.

Who says that the bootstrap requires a lot of simulation. Hall shows that you may only need 10 or 20 replications. After that, you may be impressing your friends, but just gilding the lily.

——— (1992). *The Bootstrap and Edgeworth Expansion.* Springer, New York.

Presents and unified, theoretical perspective on the bootstrap from the point of view of Edgeworth expansions, a variation of Taylor series from calculus. Nice collection of results on the coverage of CI's.

Hall, P., J. L. Horowitz, B-Y. Jing (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* 82, 561--574.

A recent update on the use of blocking to handle dependent data.

Hall, P. and M. A. Martin (1988). On bootstrap resampling and iteration. *Biometrika* 75, 661–671.

Peter Hall's version of the double bootstrap. Has a very appealing analysis and clever presentation.

Hampel, F. R., E. M. Ronchetti, P. J. Rousseuw and W. A. Stahel (1986). *Robust Statistics*. Wiley, New York.

Still not using robust methods after all of these years? Take a look here for useful motivation.

Hastie, T. J. and R. J. Tibshirani (1990). *Generalized Additive Models*. Chapman and Hall, New York.

A good introduction to smoothing methods in regression, with quite a few examples. It is helpful to know something about generalized linear models (GLIM, not GLS) before reading, but not necessary. GLIM includes logistic regression as a special case.

Hjorth, J. S. U. (1994). *Computer Intensive Statistical Methods*. Chapman and Hall, London.

Lots of Fortran code for examples, and the underlying message of picking the right model.

Hoaglin, D. C. and D. S. Moore (1992). *Perspectives on Contemporary Statistics*. Mathematical Association of America, Washington, D. C.

A collection of mostly readable articles on the changing state of statistics, with lots of emphasis on the role of computing, graphics, and models.

Lambert, Z., A. R. Wildt, and R. M. Durant (1990). Assessing sampling variation relative to number of factors criteria. *Educational and Psychological Measurement*, 50, 33-48.

Zarrel was a participant in the first bootstrap course at ICPSR.

*Leger, C. and D.N. Politis (1992). Bootstrap technology and applications. *Technometrics*, 34, 378-398.

Technical in places, but it gives you an overview of some recent research objectives.

LePage, R. and L. Billiard (1992). *Exploring the Limits of Bootstrap*. Wiley, New York.

An edited volume of papers on the bootstrap, including a very provocative article (at least within this problem domain) by Efron discussing how the bootstrap as led to a new look at old problems.

*Mooney, C. Z. and R. D. Duval (1993). *Bootstrapping: A Nonparametric Approach to Statistical Inference*. Sage, Newbury Park, CA.

This monograph moves slowly, but its leisurely pace is relaxing after looking at some references. Most of the discussion concerns confidence intervals in straightforward applications.

Moulton, L.H. and S. L. Zegar (1989). Analyzing repeated measures on generalized linear models via the bootstrap. *Biometrics*, 45, 381-394.

This use of the bootstrap "automatically" captures the standard approximations used for estimating the variance of coefficient estimates in longitudinal models.

Newton, M. A., C. J. Geyer (1994). Bootstrap recycling: a Monte Carlo alternative to the nested bootstrap. *JASA*, 89, 905--912.

Much of the recent work focuses on faster calculations, particularly when using the bootstrap to improve the bootstrap. This paper suggests a method for dodging some double bootstrap calculations.

Politis, D. N., J. P. Romano (1994). The stationary bootstrap. *JASA*, 89 , 1303--1313.

A general purpose approach to handing dependent data in a less model-dependent way. In the family of blocked bootstrap methods.

Politis, D. N., J. P. Romano, and M. Wolf (1999). *Subsampling*. Springer, New York.

Bootstrap resampling samples with replacement from the data. Subsampling estimates the variation of a statistic without using samples drawn with replacement, instead computing the statistic on many subsamples of the data. These methods avoid some of the problems that the bootstrap encounters with dependence.

Rao, J. N. K. and C. F. J. Wu (1988). Resampling inference with complex survey data. *JASA*, 83, 231-241.

Most work on the bootstrap assumes a simple random sample. Here you get a look at what happens when the data have a more complex structure, such as multistage designs with stratification.

Schenker, N. (1985). Qualms about bootstrap confidence intervals. *JASA*, 80, 360-1.

Nat lit a fire with this article, and hastened the arrival of the accelerated bootstrap interval.

Shao, J. and D. Tu (1995). *The Jackknife and Bootstrap*. Springer, New York.

A bit too theoretical for most, but a huge collection (about 600) references to the literature.

Stine, R. A. (1985). Bootstrap prediction intervals for regression. *JASA*, 80, 1026-1031

What can I say? It was a thesis chapter.

*——— (1989). An introduction to bootstrap methods: examples and ideas. S*ociological Methods in Research*, 16, 243-291.

My favorite introductory overview!

Stine, R. A. and J. Fox (1997). *Statistical Computing Environments for Social Research*. Sage, Thousand Oaks.

The edited collection describes seven alternative computing environments for doing your data analysis (APL, Gauss, Lisp-Stat, Mathematica, SAS, Stata, and S+), with introductions to AXIS, R-code, and Vista (freely available systems derived from LispStat).

Tierney, L. (1990). *Lisp-Stat*. Wiley, New York.

This book describes the FREE software which is the basis for the *AXIS* software used in the presentations. Not as comprehensive as S, SPSS or SAS, but Lisp-Stat is the nicest, most extensible package of the lot. Lisp expands your mind (see comparisons in Stine and Fox).

Tukey, J. W. T. (1958). Bias and confidence in not-quite large samples. *Annals of Mathematical Statistics,* 614.

The most famous abstract in statistics. It presented the jackknife in one paragraph.

——— (1962). The future of data analysis. *Annals of Math. Statistics*, 33, 1-67.

Your summer reading is not complete without a look a this one.

——— (1986). Sunset salvo. *American Statistician*, 40, 72-76.

Classic Tukey, with some great quotes about fancy answers to wrong questions.

Vinod, H. D. and B. Raj (1988).Economic issues in Bell System divestiture: a bootstrap application. *Applied Statistics*, 37, 251-261.

Draws heuristic inferences from the stem-and-leaf plot of the bootstrap distribution.

Young, G. A. (1994). Bootstrap: More than a stab in the dark (with discussion). *Statistical Science*, 9, 382-415.

This article looks at some limitations, and what progress toward applications has been done.

Weinberg, S., J. D. Carroll, and H. S. Cohen (1984). Confidence regions for Indscal using the jackknife and bootstrap techniques. *Psychometrika*, 49, 475-91.

An early application in the social sciences. Limits its use of the bootstrap to standard errors.