Bootstrap Method

K Singh and M Xie, Rutgers University, Piscataway, NJ, USA

Glossary

Bias – Deviation between the true population value and the mean of a sample estimate (average over all possible samples).

Central limit theorem – It is the mathematical phenomenon which states that, as sample size tends to infinity, the sample mean follows normal curve (bell-shaped curve, also known as Gaussian curve) in limit.

Confidence interval – A sample data-based range such that a given population parameter lies within the range with a specified high probability (called confidence level).

Probability distribution – It describes a chance pattern according to which a random variable takes values.

Random variable – A statistical terminology used to describe a measurement or count which depends on chance.

Regression model – A statistical model used to describe the relationship of a variable of interest (called response variable) and a set of variables (called covariates) which are supposed to influence the response variable.

Sample estimator – A mathematical function of sample data used for approximating a population parameter (characteristic).

Standard error – The deviation between a true parameter value and its sample estimate is squared and averaged over all possible samples. Its square root is called standard error.

Time series – A sequence of measurements, taken at successive time points or intervals.

Introduction and the Idea

Efron (1979) introduced the bootstrap method. It spread like brush fire in statistical sciences within a couple of decades. Now if one conducts a Google search for the above title, an astounding 1.86 million records will be mentioned; scanning through even a fraction of these records is a daunting task. We attempt first to explain the idea behind the method and the purpose of it at a rather rudimentary level. The primary task of a statistician is to summarize a sample-based study and generalize the finding to the parent population in a scientific manner. A technical term for a sample summary number is (sample) statistic. Some basic sample statistics are sample mean, sample median, sample standard deviation, etc. Of course, a summary statistic like the sample mean will fluctuate from sample to sample and a statistician would like to know the magnitude of these fluctuations around the corresponding population parameter in an overall sense. This is then used in assessing margin of errors. The entire picture of all possible values of a sample statistics presented in the form of a probability distribution is called a sampling distribution. There is a plenty of theoretical knowledge of sampling distributions, which can be found in any text books of mathematical statistics. A general intuitive method applicable to just about any kind of sample statistic that keeps the user away from the technical tedium has got its own special appeal. Bootstrap is such a method.

To understand bootstrap, suppose it were possible to draw repeated samples (of the same size) from the population of interest, a large number of times. Then, one would get a fairly good idea about the sampling distribution of a particular statistic from the collection of its values arising from these repeated samples. But, that does not make sense as it would be too expensive and defeat the purpose of a sample study. The purpose of a sample study is to gather information cheaply in a timely fashion. The idea behind bootstrap is to use the data of a sample study at hand as a surrogate population, for the purpose of approximating the sampling distribution of a statistic; that is, to resample (with replacement) from the sample data at hand and create a large number of phantom samples known as bootstrap samples. The sample summary is then computed on each of the bootstrap samples (usually a few thousand). A histogram of the set of these computed values is referred to as the bootstrap distribution of the statistic.

In bootstrap's most elementary application, one produces a large number of copies of a sample statistic, computed from these phantom bootstrap samples. Then, a small percentage, say $100(\alpha/2)\%$ (usually $\alpha = 0.05$, is trimmed off from the lower as well as from the upper end of these numbers. The range of remaining $100(1-\alpha)\%$ values is declared as the confidence limits of the corresponding unknown population summary number of interest, with level of confidence $100(1-\alpha)\%$. The above method is referred to as bootstrap percentile method. We shall return to it later in the article.

The Theoretical Support

Let us develop some mathematical notations for convenience. Suppose a population parameter θ is the target of a study; say, for example, θ is the household median income of a chosen community. A random sample of size n yields the data $(X_1, X_2, \ldots, X_n)$. Suppose the corresponding sample statistic computed from this data set is $\hat{\theta}$ (sample median in the case of the example). For most sample statistics, the sampling distribution of $\hat{\theta}$ for large n ($n \geq 30$ is generally accepted as large sample size) is bell shaped with center θ and standard deviation $(a/\sqrt{n})$, where the positive number a depends on the population and the type of statistic $\hat{\theta}$. This phenomenon is the celebrated central limit theorem (CLT). Often, there are serious technical complexities in approximating the required standard deviation from the data. Such is the case when $\hat{\theta}$ is sample median or sample correlation. Then bootstrap offers a bypass. Let $\hat{\theta}_B$ stand for a random quantity which represents the same statistic computed on a bootstrap sample drawn out of $(X_1, X_2, \ldots, X_n)$. What can we say about the sampling distribution of $\hat{\theta}_B$ (w.r.t. all possible bootstrap samples), while the original sample $(X_1, X_2, \ldots, X_n)$ is held fixed? The first two articles dealing with the theory of bootstrap – Bickel and Freedman (1981) and Singh (1981) – provided large sample answers for most of the commonly used statistics. In limit, as $(n \to \infty)$, the sampling distribution of $\hat{\theta}_B$ is also bell shaped with $\hat{\theta}$ as the center and the same standard deviation $(a/\sqrt{n})$. Thus, bootstrap distribution of $\hat{\theta}_B - \hat{\theta}$ approximates (fairly well) the sampling distribution of $\hat{\theta} - \theta$. Note that, as we go from one bootstrap sample to another, only $\hat{\theta}_B$ in the expression $\hat{\theta}_B - \hat{\theta}$ changes as $\hat{\theta}$ is computed on the original data $(X_1, X_2, \ldots, X_n)$. This is the bootstrap CLT. For a proof of bootstrap CLT for the mean, reader is referred to Singh (1981).

Furthermore, it has been found that if the limiting sampling distribution of a statistical function does not involve population unknowns, bootstrap distribution offers a better approximation to the sampling distribution than the CLT. Such is the case when the statistical function is of the form $(\hat{\theta}_B - \hat{\theta})/\mathrm{SE}$ where SE stands for true or sample estimate of the standard error of $\hat{\theta}$, in which case the limiting sampling distribution is usually standard normal. This phenomenon is referred to as the second-order correction by bootstrap. A caution is warranted in designing bootstrap, for second-order correction. For illustration, let $\theta = \mu$, the population mean, and $\hat{\theta} = \bar{X}$, the sample mean; σ = population standard deviation; s = sample standard deviation computed from original data; and s_B is the sample standard deviation computed on a bootstrap sample. Then, the sampling distribution of $(\bar{X} - \mu)/\mathrm{SE}$, with $\mathrm{SE} = \sigma/\sqrt{n}$, will be approximated by the bootstrap distribution of $(\bar{X}_B - \bar{X})/\mathrm{SE}$, with $\bar{X}_B$ = bootstrap sample mean and $\mathrm{SE} = s/\sqrt{n}$. Similarly, the sampling distribution of $(\bar{X} - \mu)/\mathrm{SE}$, with $\mathrm{SE} = s/\sqrt{n}$, will be approximated by the bootstrap distribution of $\bar{X}_B - \bar{X})/\mathrm{SE}_B$, with $\mathrm{SE} = s_B/\sqrt{n}$. The earliest results on second-order correction were reported in Singh (1981) and Babu and Singh (1983). In the subsequent years, a flood of large sample results on bootstrap with substantially higher depth, followed. A name among the researchers in this area that stands out is Peter Hall of Australian National University.

Primary Applications of Bootstrap

Approximating Standard Error of a Sample Estimate

Let us suppose, information is sought about a population parameter θ. Suppose $\hat{\theta}$ is a sample estimator of θ based on a random sample of size n, that is, $\hat{\theta}$ is a function of the data $(X_1, X_2, \ldots, X_n)$. In order to estimate standard error of $\hat{\theta}$, as the sample varies over the class of all possible samples, one has the following simple bootstrap approach.

Compute$(\theta_1^*, \theta_2^*, \ldots, \theta_N^*)$, using the same computing formula as the one used for $\hat{\theta}$, but now base it on N different bootstrap samples (each of size n). A crude recommendation for the size N could be $N = n^2$ (in our judgment), unless n^2 is too large. In that case, it could be reduced to an acceptable size, say $n\log_e n$. One defines

$$\mathrm{SE}_B(\hat{\theta}) = [(1/N)\sum_{i=1}^{N}(\theta_i^* - \hat{\theta})^2]^{1/2}$$

following the philosophy of bootstrap: replace the population by the empirical population.

An older resampling technique used for this purpose is Jackknife, though bootstrap is more widely applicable. The famous example where Jackknife fails while bootstrap is still useful is that of $\hat{\theta}$ = the sample median.

Bias Correction by Bootstrap

The mean of sampling distribution of $\hat{\theta}$ often differs from θ, usually by an amount $= c/n$ for large n. In statistical language, one writes

$$Bias(\hat{\theta}) = E(\hat{\theta}) - \theta \approx O(1/n)$$

A bootstrap based approximation to this bias is

$$\frac{1}{N}\sum_{i=1}^{N}\theta_i^* - \hat{\theta} = Bias_B\ (\hat{\theta})\ (\text{say})$$

where θ_*^i are bootstrap copies of $\hat{\theta}$, as defined in the earlier subsection. Clearly, this construction is also based on the standard bootstrap thinking: replace the population by the empirical population of the sample. The bootstrap bias-corrected estimator is $\hat{\theta}_c = \hat{\theta} - Bias_B\ (\hat{\theta})$. It needs to be pointed out that the older resampling technique called Jackknife is more popular with statisticians for the purpose of bias estimation.

Bootstrap Confidence Intervals

Confidence intervals for a given population parameter θ are sample-based range $[\hat{\theta}_1, \hat{\theta}_2]$ given out for the unknown number θ. The range possesses the property that θ would lie within its bounds with a high (specified) probability. The latter is referred to as confidence level. Of course, this probability is with respect to all possible samples, each sample giving rise to a confidence interval which thus depends on the chance mechanism involved in drawing the samples. The two mostly used levels of confidence are 95% and 99%. We limit ourselves to the level 95% for our discussion here. Traditional confidence intervals rely on the knowledge of sampling distribution of $\hat{\theta}$, exact or asymptotic as $n \to \infty$. Here are some standard brands of confidence intervals constructed using bootstrap.

Bootstrap percentile method

This method was mentioned in the introduction itself, because of its popularity which is primarily due to its simplicity and natural appeal. Suppose one settles for 1000 bootstrap replications of $\hat{\theta}$, denoted by $(\theta_1^*, \theta_2^*, \ldots, \theta_{1000}^*)$. After ranking from bottom to top, let us denote these bootstrap values as $(\theta_{(1)}^*, \theta_{(2)}^*, \ldots, \theta_{(1000)}^*)$. Then the bootstrap percentile confidence interval at 95% level of confidence would be $[\theta_{(25)}^*, \theta_{(975)}^*]$. Turning to the theoretical aspects of this method, it should be pointed out that the method requires the symmetry of the sampling distribution of $\hat{\theta}$ around θ. The reason is that the method approximates the sampling distribution of $\hat{\theta} - \theta$ by the bootstrap distribution of $\hat{\theta} - \hat{\theta}_B$, which is contrary to the bootstrap thinking that the sampling distribution of $\hat{\theta} - \theta$ could be approximated by the bootstrap distribution of $\hat{\theta}_B - \hat{\theta}$. Interested readers are referred to Hall (1988).

Centered bootstrap percentile method

Suppose the sampling distribution of $\hat{\theta} - \theta$ is approximated by the bootstrap distribution of $\hat{\theta}_B - \hat{\theta}$, which is what the bootstrap prescribes. Denote 100s-th percentile of $\hat{\theta}_B$ (in bootstrap replications) by B_s. Then, the statement that $\hat{\theta} - \theta$ lies within the range $B_{0.025} - \hat{\theta}$, $B_{0.975} - \hat{\theta}$ would carry a probability $\approx$0.95. But, this statement easily translates to the statement that θ lies within the range $(2\hat{\theta} - B_{0.975}, 2\hat{\theta} - B_{0.025})$. The latter range is what is known as centered bootstrap percentile confidence interval (at coverage level 95%). In terms of 1000 bootstrap replications $B_{0.025} = \theta^*_{(25)}$ and $B_{0.975} = \theta^*_{(975)}$.

Bootstrap-t methods

As it was mentioned in section 'The theoretical support', bootstrapping a statistical function of the form $T = (\hat{\theta} - \theta)/\mathrm{SE}$, where SE is a sample estimate of the standard error of $\hat{\theta}$, brings extra accuracy. This additional accuracy is due to so-called one-term Edgeworth correction by the bootstrap. The reader could find essential details in Hall (1992b). The basic example of T is the standard t-statistics (from which the name bootstrap-t is derived): $t = (\bar{X} - \mu)/s/\sqrt{n}$, which is a special case with $\theta = \mu$ (the population mean), $\hat{\theta} = \bar{X}$ (the sample mean), and s standing for the sample standard deviation. The bootstrap counterpart of such a function T is $T_B = (\hat{\theta}_B - \theta)/\mathrm{SE}_B$ where SE_B is exactly like SE but computed on a bootstrap sample. Denote the 100s-th bootstrap percentile of T_B by b_s and consider the statement: T lies within $[b_{0.025}, b_{0.975}]$. After the substitution $T = (\hat{\theta} - \theta)/\mathrm{SE}$, the above statement translates to 'θ lies within $(\hat{\theta} - \mathrm{SE}\ b_{0.975}, \hat{\theta} - \mathrm{SE}\ b_{0.025})$'. This range for θ is called bootstrap-t-based confidence interval for θ at coverage level 95%. Such an interval is known to achieve higher accuracy than the earlier method, which is referred to as second-order accuracy in technical literature.

We conclude the section with a remark that B. Efron proposed correction to the rudimentary percentile method to bring in extra accuracy. These corrections are known as Efron's 'bias-correction' and 'accelerated bias-correction'. The details could be found in Efron and Tibshirani (1993). The bootstrap-t automatically takes care of such corrections, although the bootstrapper needs to look for a formula for SE which is avoided in the percentile method.

Some Real Data Example

Example 1 (Skewed Univariate Data). In the first example, the data are taken from Hollander and Wolfe (1999: 63), which represent the effect of illumination (difference between counts with and without illumination) on the rate of beak-clapping among chick-embryos. The boxplot suggests lack of normality of the population. We have carried out bootstrap analysis on the median and the mean. A noteworthy finding is the lack of symmetry of bootstrap-t histogram, which differs from limiting normal curve. The 95% level confidence intervals coming from our analysis for both mean and the median (centered bootstrap percentile method) cover the range [10, 30], roughly speaking. This range represents overall difference (increase) in the beak-clapping counts per minute due to illumination.

Example 2 (Bivariate Data). In this example, the data are from Collins *et al.* (1999), which assess body fat in collegiate football players (Devore, 2003: 553). We study correlation between the BOD and HW measurements; see the data at the end of this section. Here, BOD is BOD POD, a whole body air-displacement plethysmograph, and HW refers to hydrostatic weighing. The sample size is modest, but reasonable for bootstrap methods. As bivariate data consist of n pairs of data, say (X_i, Y_i), for $i = 1, \ldots, n$, one draws a pair of data randomly at a time in the bootstrap resampling. For instance, the first draw could be (X_7, Y_7) followed by (X_3, Y_3), etc. The box plots of original data in **Figure 1(a)** suggest lack of normality of the underlying populations.

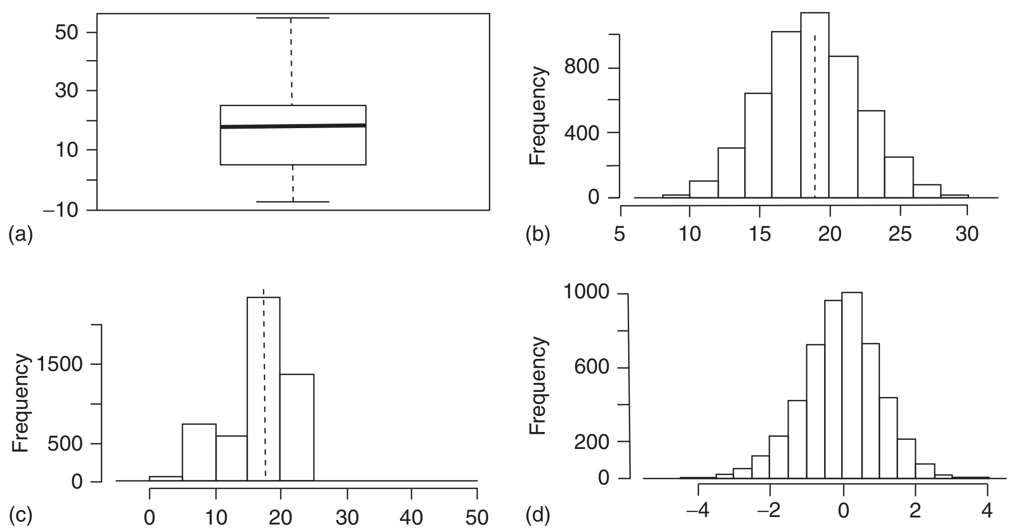

Figure 1 Boxplot of the measurement is presented in (a). Bootstrap distributions of the sample mean, sample median and t^* statistic are plotted in (b)–(d), respectively. The dotted lines in (b) and (c) correspond respectively to the sample mean and sample median. Based the bootstrap distributions, the 95% confidence interval for the population median by the bootstrap percentile method is (4.7000, 24.7000), by the centered boostrap percentile is (10.5000, 30.5000). The 95% confidence interval for the population mean by the percentile bootstrap method is (10.0960, 28.1200) and by the centered bootstrap method is (9.4880, 27.51200). The bootstrap-*t* 95% CI for the population mean is (12.9413, 30.8147). Note that the bootstrap-*t* on the mean shows skewed histogram of the *t*-distribution.

The histogram for correlations computed on bootstrap bivariate data is plotted in Figure 2(c) which is asymmetric (skewed to the left). For this reason, the centered bootstrap percentile confidence interval appears more appropriate. According to our bootstrap analysis, the two measurements have at least a correlation of 0.78 in the population.

Data for Example 1:

-8.5-4.6-1.8-0.8 1.9 3.9 4.7 7.1 7.5 8.5 14.8 16.7 17.6 19.7 20.6 21.9 23.8 24.7 24.7 25.0 40.7 46.9 48.3 52.8 54.0

Data for Example 2:

BOD

2.5 4.0 4.1 6.2 7.1 7.0 8.3 9.2 9.3 12.0 12.2 12.6 14.2 14.4 15.1 15.2 16.3 17.1 17.9 17.9

HW

8.0 6.2 9.2 6.4 8.6 12.2 7.2 12.0 14.9 12.1 15.3 14.8 14.3 16.3 17.9 19.5 17.5 14.3 18.3 16.2

Engineering: A Fitting Bootstrap

A sizable amount of journal literature on the topic is directed toward proposal and study of bootstrap schemes which will produce decent results in various statistical situations. The setup that has been the basis of forgoing discussion is basic and there are many types of departures from it. How to bootstrap in case of two-stage sampling or a stratified sampling? Natural schemes are not hard to think of. Bootstrapping in the case of data with regression models has attracted a lot of attention. There are two schemes which stand out: in one of which the covariate(s) and the response variable are resampled together (called paired bootstrap), and the other one bootstraps the residuals (=response − fitted model value) and then reconstructs the bootstrap regression data by plugging in the estimated regression parameters (called residual bootstrap). Paired bootstrap remains valid – in the sense of correct outcome in the limit as $n \to \infty$, even if the error variances in the model are unequal; a property which the residual bootstrap lacks. The shortcoming is compensated by the fact that the latter scheme brings additional accuracy in the estimation of standard error. This is the classic tug of war between efficiency and robustness in statistics (see Liu and Singh, 1992a).

A lot harder to bootstrap are the time series data. Needless to say, time series analysis is of critical importance in several disciplines, especially in econometrics. The sources of difficulty are twofold: (1) Time series data possess serial dependence, that is, X_{T+1} has dependence on X_T, X_{T-1}, etc; (2) the statistical population changes with time, and that is known as nonstationarity. It was noted very early on (see Singh (1981) for *m*-dependent data) that the classical bootstrap cannot handle dependent data. A fair amount of research has been dedicated to modifying the bootstrap so that it could automatically bring in the dependence structure of the original sampling into bootstrap samples. The scheme of moving–block bootstrap has become quite well known (invented in Kunch (1989) and Liu and Singh (1992). Potitis and Romano are well-known authors on the topic, whose contributions have led to significant advancements on the topic of resampling, in general. In a moving block

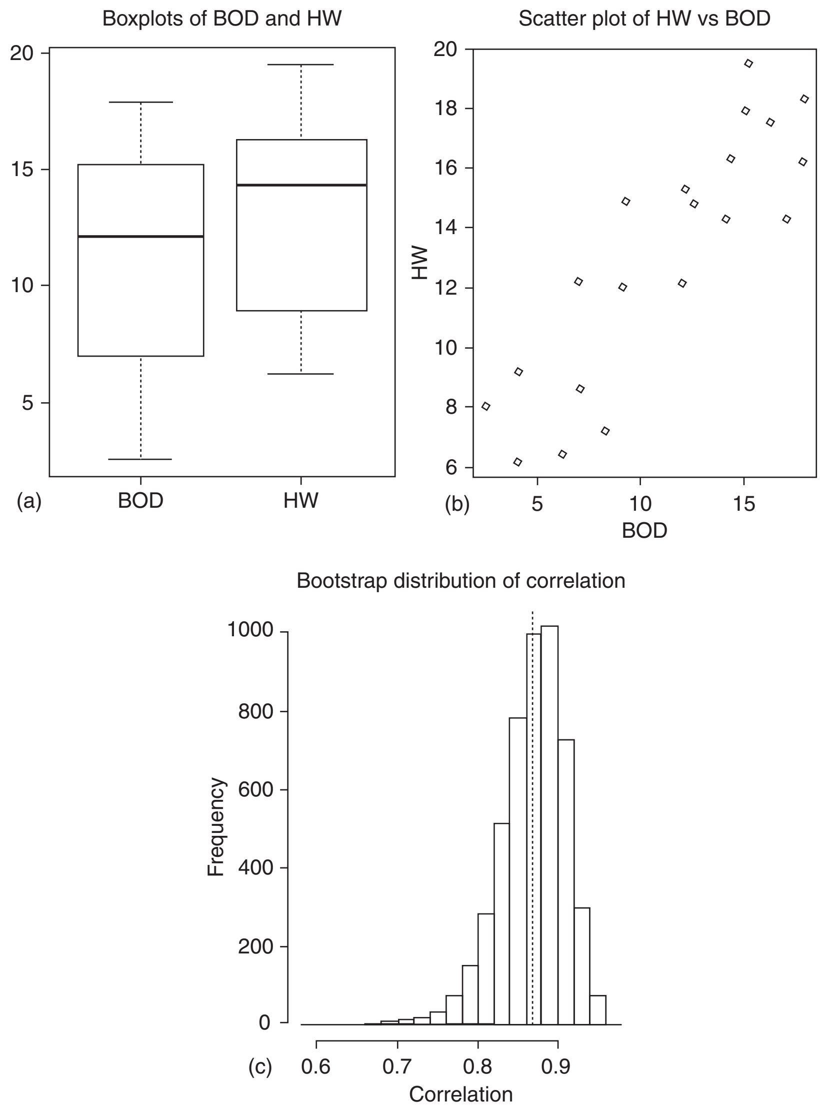

Figure 2 Boxplots of BOD and HW in (a) suggest somewhat non-normal data. Scatter plot in (b) indicates they are highly correlated. Bootstrap inference on the correlation between BOD and HW is presented in (c), which shows the bootstrap distribution (in histogram) of correlation. In particular, sample correlation of BOD and HW is 0.8679 which corresponds to the dotted vertical line in (c). The SE of the correlation is 0.0412 with an estimated bias of 0.0003. The 95% confidence interval of the correlation by the bootstrap percentile method is (0.7222, 0.9490) and the 95% confidence interval by the centered bootstrap percentile method is (0.7868, 1.0136).

bootstrap scheme, one draws a block of data at a time, instead of one of the X_i's at a time, in order to preserve the underlying serial dependence structure that is present in the sample. There is plenty of ongoing research in the area of bootstrap methodology on econometric data.

The Great *m* out of *n* Bootstrap with ($m/n \to 0$)

There are various types of conditions under which the straightforward bootstrap becomes inconsistent, meaning that the bootstrap estimate of sampling distribution and the true sampling distribution do not approach to the same limit, as the sample size n tends to ∞. That means, for large samples, one is bound to end up with an inaccurate statistical inference. The examples include, just to name a few, bootstrapping sample minimum or sample maximum which estimate endpoint of a population distribution (Bickel and Freedman, 1981), the case of sample mean when the population variance is ∞ (Athreya, 1986) and bootstrapping sample eigenvalues when population eigenvalues have multiplicity (Eaton and Tyler, 1991), the case of sample median when the population density is discontinuous at the population median (Huang *et al.*, 1996). Luckily, a general remedy exists and that is to keep the bootstrap sample size m much lower than the original size. Mathematically speaking, one requires $m \to \infty$ and $m/n \to 0$, as $n \to \infty$. In theory it fixes the problem; however for users, it is somewhat troublesome. How to choose m? An obvious suggestion would be settle for a fraction of n, say 20% or so. It should be pointed out that in good situations, where the regular bootstrap is fine, such an m is not advisable as it will result in loss of efficiency (see Bickel (2003), for a recent survey on the topic).

Bibliography

Athreya, K. B. (1986). Bootstrap of the mean in the infinite variance case. *Annals of Statistics* **14**, 724–731.

Babu, G. J. and Singh, K. (1983). Inference on means using the bootstrap. *Annals of Statistics* **11**, 999–1003.

Bickel, P. J. (2003). Unorthodox bootstraps (invited papers). *Journal of the Korean Physical Society* **32**, 213–224.

Bickel, P. J. and Freedman, D. (1981). Some asymptotic theory for the bootstrap. *Annals of Statistics* **9**, 1196–1217.

Collins, M. A., Millard-Stafford, M. L., Sparling, P. B., *et al.* (1999). Evaluating BOD POD(R) for assessing body fat in collegiate football players. *Medicine and Science in Sports and Exercise* **31**, 1350–1356.

Davison, A. C. and Hinkley, D. V. (1988). Saddle point approximations in resampling method. *Biometrika* **75**, 417–431.

Devore, L. J. (2003). *Probability and Statistics for Engineering and the Science*. Boston, MA: Duxbury Press.

Eaton, M. L. and Tyler, D. E. (1991). On Wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix. *Annals of Statistics* **19**, 260–271.

Efron, B. (1979). Bootstrap methods: Another look at jackknife. *Annals of Statistics* **7**, 1–26.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. New York: Chapman and Hall.

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals (with discussions). *Annals of Statistics* **16**, 927–953.

Hall, P. (1992a). Bootstrap confidence intervals in nonparametric regression. *Annals of Statistics* **20**, 695–711.

Hall, P. (1992b). *The Bootstrap and Edgeworth Expansion*. New York: Springer.

Hollander, M. and Wolfe, D. A. (1999). *Nonparametric Statistical Methods*, 2nd edn. New York: Wiley.

Huang, J. S., Sen, P. K., and Shao, J. (1996). Bootstrapping a sample quantile when the density has a jump. *Statistica Sinica* **6**, 299–309.

Liu, R. Y. and Singh, K. (1992a). Efficiency and robustness in re sampling. *Annals of Statistics* **20**, 370–384.

Liu, R. Y. and Singh, K. (1992b). Moving block jackknife and bootstrap capture weak dependence. In Lepage, R. and Billard, L. (eds.) *Exploring the Limits of Bootstrap*, pp 225–248. New York: Wiley.

Kunch, H. R. (1989). The jackknife and bootstrap for general stationary observations. *Annals of Statistics* **17**, 1217–1241.

Singh, K. (1981). On asymptotic accuracy of Efron's bootstrap. *Annals of Statistics* **9**, 1187–1195.

Further Reading

Beran, R. (1990). Refining bootstrap simultaneous confidence sets. *Journal of the American Statistical Association* **85**, 417–428.

Bickel, P. J. and Freedman, D. (1984). Asymptotic normality and the bootstrap in stratified sampling. *Annals of Statistics* **12**, 470–482.

Boos, D. D. and Munahan, J. F. (1986). Bootstrap methods using prior information. *Biometrika* **73**, 77–83.

Bose, A. (1988). Edgeworth correction by bootstrap in autoregressions. *Annals of Statistics* **16**, 1709–1722.

DiCiccio, T. J. and Romano, J. P. (1988). A review of bootstrap confidence intervals (with discussions). *Journal of the Royal Statistical Society: Series B* **50**, 538–554.

Efron, B. (1987). Better bootstrap confidence intervals (with discussions). *Journal of the American Statistical Association* **82**, 171–200.

Efron, B. (1994). Missing data, imputation and the bootstrap (with discussions). *Journal of the American Statistical Association* **89**, 463–479.

Freedman, D. A. (1981). Bootstrapping regression models. *Annals of Statistics* **9**, 1218–1228.

Hinkley, D. V. (1988). Bootstrap methods (with discussions). *Journal of the Royal Statistical Society: Series B* **50**, 321–337.

Rubin, D. B. (1981). The Bayesian bootstrap. *Annals of Statistics* **9**, 130–134.

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

Singh, K. (1998). Breakdown theory for bootstrap quantiles. *Annals of Statistics* **26**, 1719–1732.

Singh, K. and Xie, M. (2003). Bootlier-plot-Bootstrap based outlier detection plot. *Sankhya* **65**, 532–559.

Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling procedures (with discussions). *Annals of Statistics* **14**, 1261–1350.

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