Bootstrap: A Statistical Method

Kesar Singh and Minge Xie

Rutgers University

Abstract

This paper attempts to introduce readers with the concept and methodology of bootstrap in Statistics, which is placed under a larger umbrella of resampling. Major portion of the discussions should be accessible to any one who has had a couple of college level applied statistics courses. Towards the end, we attempt to provide glimpses of the vast literature published on the topic, which should be helpful to someone aspiring to go into the depth of the methodology. A section is dedicated to illustrate real data examples. A technical appendix is included, which contains a short proof of "bootstrap Central Limit Theorem" for the means. It should inspire a mathematical minded reader to study further. We think the selected set of references cover the greater part of the developments on this subject matter.

1. Introduction and the Idea

B. Efron introduced a statistical method, which is called Bootstrap, published in 1979 (Efron 1979). 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, for instance, 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; i.e. 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 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 given out as the confidence limits of the corresponding unknown population summary number of interest, with level of confidence for this range to include the unknown equal to $100(1-\alpha)\%$ approximately. The above method is referred to as bootstrap percentile method. We shall return to it later in the article.

2. The Theoretical Support

Let us develop some mathematical notations for convenience. Suppose population parameter θ is some type of population summary number which is a target of the sample study; say for example, the household median income of a chosen community. A random sample of size n yields the data $(X_1, X_2, ..., X_n)$. Suppose, the corresponding sample statistic computed from this data set is $\hat{\theta}$ (sample median in the case of the example). For most types of sample statistics (with plentiful exceptions), 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 $(X_1, X_2, ..., X_n)$. What can we say about the sampling distribution of $\hat{\theta}_B$ (w.r.t. all possible bootstrap samples), while the original sample $(X_1, X_2, ..., X_n)$ 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 "usual" set of 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 $(X_1, X_2, ..., X_n)$. This is the bootstrap Central Limit Theorem. In the technical appendix at the end of this article (for mathematical minded readers) we include a proof of bootstrap CLT for the mean from Singh (1981), which possesses exceptional elegance.

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 $(\hat{\theta}_B - \hat{\theta})/SE$ 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 $(\bar{X} - \mu)/SE$, with $SE = \sigma/\sqrt{n}$, will be approximated by the bootstrap distribution of $(\bar{X}_B - \bar{X})/\widehat{SE}$, with $\bar{X}_B$ = bootstrap sample mean and $\widehat{SE} = s/\sqrt{n}$. Similarly, the sampling distribution of $(\bar{X} - \mu)/\widehat{SE}$, with $\widehat{SE} = s/\sqrt{n}$, will be approximated by the bootstrap distribution of $(\bar{X}_B - \bar{X})/SE_B$**,** with $SE_B = s_B/\sqrt{n}$. 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.

3. Primary Applications of Bootstrap

3.1 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, i.e. $\hat{\theta}$ is a function of the data $(X_1, X_2, ..., X_n)$. 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 $(\theta_1^*, \theta_2^*, ..., \theta_N^*)$, 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). Here N is the number of bootstrap replications. 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 $SE_B(\hat{\theta}) = [(1/N)\sum_{i=1}^{N}(\theta_i^* - \hat{\theta})^2]^{1/2}$ 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.

3.2 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

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

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 $\hat{\theta}_c = \hat{\theta} - \widehat{Bias}_B(\hat{\theta})$. It needs to be pointed out that the older resampling technique called Jackknife is more popular with statisticians, for the purpose of bias estimation.

3.3 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 standard two 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 singled out to be 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 $(\theta_1^*, \theta_2^*, ..., \theta_{1000}^*)$. After ranking from bottom to top, let us denote these bootstrap values as $(\theta_{(1)}^*, \theta_{(2)}^*, ..., \theta_{(1000)}^*)$. Then the bootstrap percentile confidence interval at 95% level of confidence would be $[\theta_{(50)}^*, \theta_{(950)}^*]$. 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. As a matter of fact, the method approximates the sampling distribution of $\hat{\theta} - \theta$ by the bootstrap distribution of $\hat{\theta} - \hat{\theta}_B$, and NOT $\hat{\theta}_B - \hat{\theta}$ as the bootstrap thinking would dictate. Interested readers may check out 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 $B_{.025} - \hat{\theta}$, $B_{.975} - \hat{\theta}$ would carry a probability $\approx .95$. But, this statement easily translates to the statement that θ lies within the range $(2\hat{\theta} - B_{.975}, 2\hat{\theta} - B_{.025})$. The latter range is what is known as centered bootstrap percentile confidence interval (at coverage level 95%). In terms of 1000 bootstrap replications $B_{.025} = \theta_{(25)}^*$ and $B_{.975} = \theta_{(975)}^*$. If $\hat{\theta}_B - \hat{\theta}$ is replaced by $\hat{\theta} - \hat{\theta}_B$ in this construction, the outcome is the earlier bootstrap percentile confidence interval.

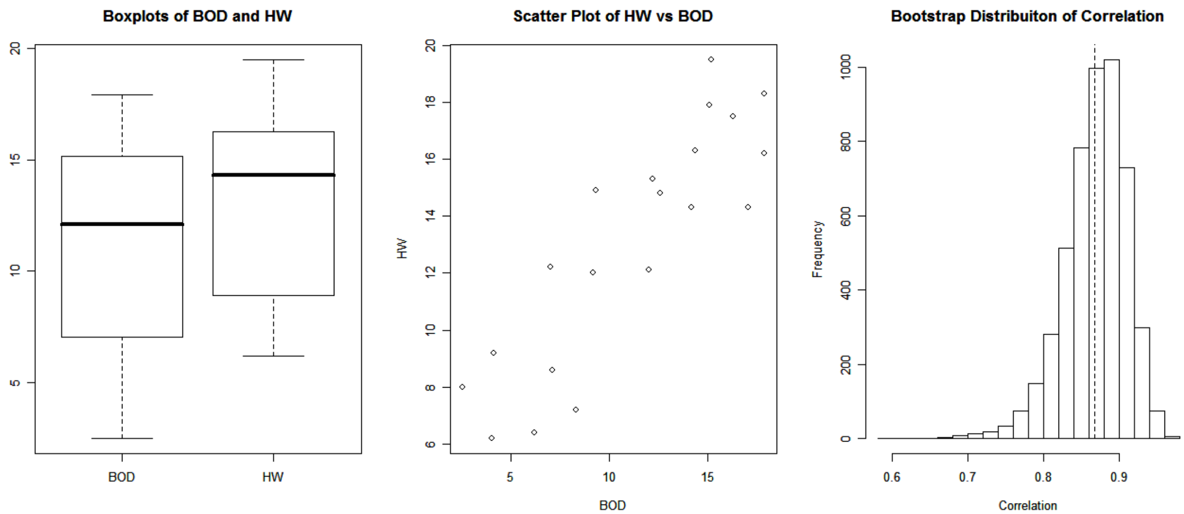
Bootstrap-t Methods:

As it was mentioned in section 2, bootstrapping a statistical function of the form $T = (\hat{\theta} - \theta)/SE$ 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 (1992). 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 $T_B = (\hat{\theta}_B - \hat{\theta}) / SE_B$ 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 $[b_{.025}, b_{.975}]$. After the substitution $T = (\hat{\theta} - \theta) / SE$, the above statement translates to 'θ lies within $(\hat{\theta} - SE\ b_{.975},\ \hat{\theta} - SE\ b_{.025})$'. This range for θ is called bootstrap-t based confidence interval for θ at coverage level 95%. Such an interval is known to have "second order accuracy" in coverage error at both ends of the interval.

We end 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 Tibishirani (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.

4. Some Real Data Example

Example 1. (Bivariate Data) In our first example, the data are from the article "Evaluating BOD POD(R) for assessing body fat in collegiate football players" published in Medicine and Science in Sports and Exercise (1999), page 1350-56 (as presented in PROBABILITY AND STATISTICS for Engineering and Science by J. L. Devore p 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. The box plots in Figure 1 (a) suggest lack of normality. The bootstrap histogram in Figure 1 (c) 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.



(a) Box plots (b) Scatter plot (c) Histogram

Figure 1. 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 0.8678753, which corresponds to the dotted vertical line in (c). The SE of the correlation is 0.0411610281 with an estimated bias of 0.0002804083. The 95% confidence interval of the correlation by the bootstrap percentile method is (0.7221611596, 0.9489691691) and the 95% confidence interval by the centered bootstrap percentile method is (0.7867814879, 1.0135894974).

Example2. (Skewed Univariate Data) In the second example, the data are taken from an article by R.W. Oppenhein (Ani. Beh. 1968 vol. 16, p. 276-280). The data represent the effect of illumination on the rate of beak-clapping among chick-embryos (presented in NONPARAMETRIC STATISTICAL METHODS by M. Hollander and D.A. Wolfe, p. 42); See the end of the section. The boxplot suggests lack of normality of the population. We have carried out bootstrap analysis on the median and on 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 arising from our analysis for both mean and the median (centered bootstrap percentile method) cover the range [10, 30], roughly speaking. This range represents overall rise in the mean number of beak-clapping per minute due to illumination.

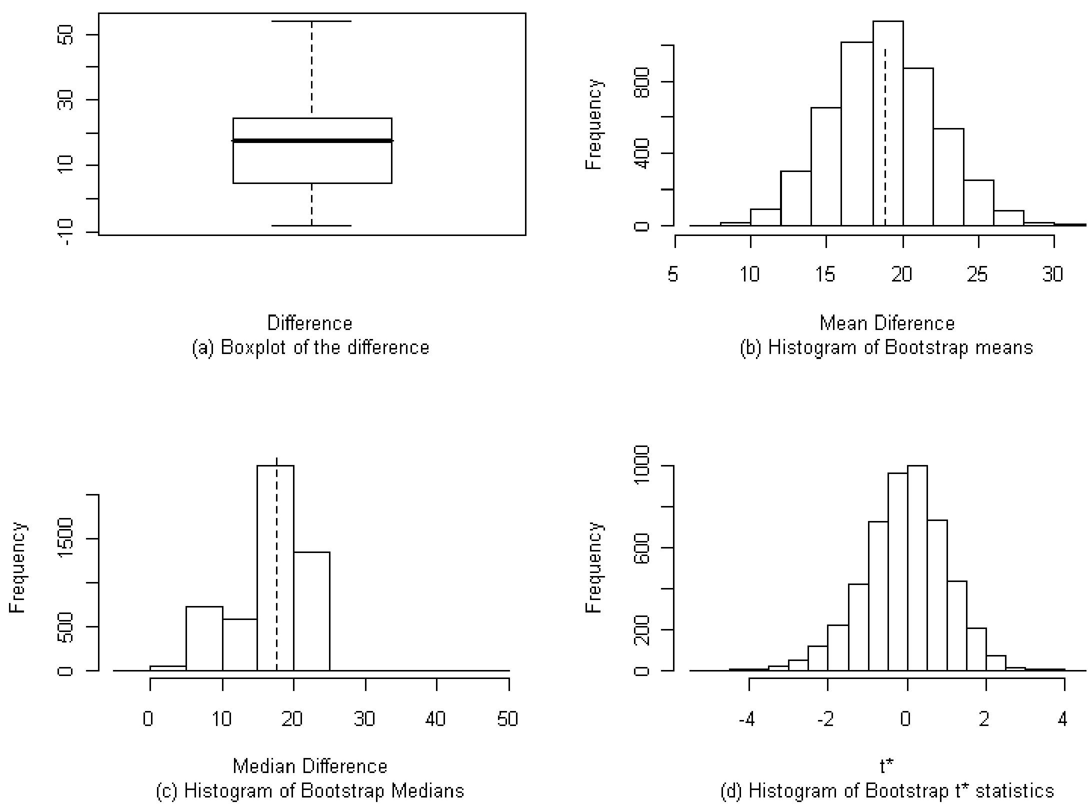


Figure 2. Boxplot of the measurement is presented in (a). Bootstrap distributions of the sample mean, sample median and t* statistic are plotted in (b)-(d), respectively. The dotted lines in (b) and (c) correspond respectively to the sample mean and sample median. Based the bootstrap distributions, the 95% confidence interval for the population median by the bootstrap percentile method is (4.700000, 24.700000), by the centered boostrap percentile is (10.500000, 30.500000). The 95% confidence interval for the population mean by the percentile bootstrap method is (10.0960000, 28.1200000), by the centered bootstrap method is (9.4880000, 27.5120000). The Bootstrap-t 95% CI for the population mean is (12.94131, 30.81469). Note that the bootstrap t on the mean show skewed histogram of the t-distribution.

Data for Example 1:

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

Data for Example 2:

-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

5. Engineering A Fitting Bootstrap

A sizable amount of journal literature on the topic is directed towards proposal and study of bootstrap schemes which will produce decent results in various statistical situations. The set up that has been the basis of forgoing discussion is basic and there are many types of departures from it. To find an example, one has to look no further than just the case of "sampling without replacement" used in drawing the original sample. If one carries out sampling without replacement (n draws) on the original sample $(X_1, X_2, ..., X_n)$, one would end up with just another permutation of this sample! It would be a nice exercise for the reader to come up with a remedy; see Bickel and Freedman (1984) for a reference. How to bootstrap in case of two stage sampling or a stratified sampling? Natural schemes are not hard to think of. Bootstrapping in 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 (1992)).

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 two-fold: (I) Time series data possess serial dependence i.e. X_{T+1} has dependence on X_T, X_{T-1} etc; (II)The statistical population changes with time, and that is known as non stationarity. It was noted very early on (see Singh (1981) for m-dependent data) that the classical bootstrap can not 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 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.

6. The great m out n bootstrap with ($m/n \to 0$)

Imagine a treatment which will rise to the occasion regardless of lies at the root of illness! A statement of this sort could be made about the 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. 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 $m/n \to 0$, 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 m is not advisable as it will result is loss of efficiency. Let us try to understand by an example how the m out of n bootstrap provides the remedy. Suppose, the statistical function of interest is $f(\bar{X}) - f(\mu)$ for a smooth function f, where $\bar{X}$ and μ are the sample mean and the population mean, respectively. Using the Taylor's expansion (δ–method) the above expression is approximated by $f'(\mu)(\bar{X} - \mu)$, with $f'(\cdot)$ denoting the first derivative. Now suppose, it has been stipulated that for the underlying population $f'(\mu) = 0$, in which case one moves on to the next term in the Taylor's expansion, which is $f''(\mu)(\bar{X} - \mu)^2/2$. Now, if $f''(\mu) \neq 0$, the limiting distribution of $n\{f(\bar{X}) - f(\mu)\}$ is the distribution of a constant times a χ^2 random variable with df 1. Consider now, the corresponding regular bootstrap statistic $f(\bar{X}_B) - f(\bar{X})$, where $\bar{X}_B$ stands for bootstrap mean. In the first term of the expansion, one encounters $f'(\bar{X})$ which is close to zero ≈ constant times $1/\sqrt{n}$, but not exactly zero. This is precisely the source of above mentioned inconsistency. This leading term after we multiply it by n, does not vanish in limit. Turning to the modified bootstrap with sample size m, with ($m/n \to 0$), the lead term (multiplied with m)

$$= \sqrt{m} f'(\bar{X})\sqrt{m}(\bar{X}_B - \bar{X})$$

which goes to 0 in limit and the next term in the expansion yields the right limit. Heuristically speaking, with a much smaller bootstrap sample size, the empirical population "feels" much closer to the true population and that does the trick! The example is from Babu (1984), in essence. See Bickel (2003), for a recent survey on the topic.

We close this section with some well known examples of bootstrap failures where "m out of n" is a remedy: Bootstrapping sample minimum or sample maximum which estimate end-point of a population distribution (Bickel and Freedman (1981)); the case of sample mean when the population variance is ∞ (Athreya (1981)); 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)).

Technical Appendix

We include here a proof of bootstrap Central Limit Theorem (CLT), taken from Singh (1981), which brings together two celebrated theorems from mathematical statistics and probability with impressive exactness. The two theorems we are pointing to are Berry-Esseen Bound and Marcinkiewicz-Zigmund Strong Law of Large Numbers (MZ-SLLN). The classical CLT asserts: If $(X_1, X_2, ...)$ are i.i.d. random variables with $\sigma^2 = Var(X_1) < \infty$, then as $n \to \infty$,

$$\sup_x | P\left(\sqrt{n}(\bar{X} - \mu)/\sigma \le x\right) - \Phi(x) | \to 0$$

with $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$, $\mu = E(X_1)$ and $\Phi(\bullet)$ denoting the cdf of standard normal distribution. Now, the corresponding theorem for bootstrap asserts:

Theorem (Bootstrap CLT). Let $(X_1, X_2, ...)$ be i.i.d random variables with $\sigma^2 = Var(X_1) < \infty$. Let $\bar{X}_B$ denote the mean of n i.i.d draws from $(X_1, X_2, ..., X_n)$. Then as $n \to \infty$,

$$\sup_x |P_B\left(\sqrt{n}(\bar{X}_B - \bar{X})/s_n \le x\right) - \Phi(x)| \to 0,$$

with probability 1, w.r.t $(X_1, X_2, ...)$. Here $\bar{X}$ and s_n stand for the mean and SD of $(X_1, X_2, ..., X_n)$ and P_B denotes bootstrap probability (probability associated with random drawing from $(X_1, X_2, ..., X_n)$).

Proof: According to the Berry-Esseen bound, the left hand side of the above expression is bounded by $(K/\sqrt{n})(1/n)\sum_{i=1}^{n}|X_i - \bar{X}|^3$**,** where K is an universal constant. As is well known, $s_n^2 \to \sigma^2$, with probability 1 if $\sigma^2 < \infty$. Now, let us note that $|X_i - \bar{X}|^3 \le 4(|X_i|^3 + |\bar{X}|^3)$ and $n^{-1/2}|\bar{X}|^3 \to 0$, with probability 1. Thus, it suffices to conclude that $(1/n^{3/2})\sum_{i=1}^{n}|X_i|^3 \to 0$, with probability 1. A part of MZ-SLLN states: If $(W_1, W_2, ...)$ are i.i.d random variables with $E|W_1|^r \le \infty$ with $0 < r < 1$, then $n^{-1/r}\sum_{i=1}^{n} W_i \to 0$, with probability 1. Let us apply this with $W_i = |X_i|^3$ and $r = 2/3$, since one has the condition $E(|W_i|^{2/3}) < \infty$, as $|W_i|^{2/3} = X_i^2$. Now, MZ-SLLN exactly fits to our situation. The bootstrap CLT for other statistics can be established via linear approximation of the statistics when available.

References:

Athreya, K.B. (1986). Bootstrap of the mean in the infinite variance case. *Ann. Stat*. 14, 724-731.

Azzalini, A. and Hall, P. (2000). Reducing variability using bootstrap methods with quantitative constraints. *Biometrika*, 87, 895-906.

Babu, G.J. (1984). Bootstrapping statistics with linear combination of Chi-square as weak limit. *Sankhya* A. 46, 85-93.

Babu, G.J. and Singh, K. (1983). Inference on means using the bootstrap. *Ann. Stat*. 11, 999-1003.

Beran, R. (1984). Prepivoting to reduce level errors of confidence sets. *Biometrika*. 74, 151-173.

Beran, R. (1990) Refining bootstrap simultaneous confidence sets. *Jour. Amer. Stat. Assoc*. 85, 417-428.

Bickel, P.J. (2003). Unorthodox bootstraps (invited papers). *J. of Korean Stat. Soc*. 32, 213-224.

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

Bickel, P.J. and Freedman, D (1984). Asymptotic normality and the bootstrap in stratified sampling. *Ann. Stat*. 12, 470-482.

Boos, D.D. and Brownie, C. (1989). Bootstrap methods for testing homogeneity of variances. Technometics. 31, 69-82.

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. *Ann. Stat*. 16, 1709-1722.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, 26, 123-140.

Buhlmann, P. (1994). Bootstrap empirical process for stationary sequences. *Ann. Stat*. 22, 995-1012.

Buhlmann, P. (2002). Sieve bootstrap with variable length – Markov chains for stationary categorical Time series (with discussions) *Jour. Amer. Stat. Assoc.* 97, 443-455.

Burr, D. (1994). A comparison of certain bootstrap confidence intervals in Cox model. *Jour. Amer. Stat. Assoc*. 89, 1290-1302.

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

DiCiccio, T.J. and Romano, J.P. (1988). A review of bootstrap confidence intervals (with discussions). *J. R. Stat. Soc*. B. 50, 538-554.

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. *Ann. Stat.* **19**, 260–271.

Efron, B. (1979). Bootstrap methods: Another look at jackknife. *Ann. Stat*. 7, 1-26.

Efron, B. (1987). Better bootstrap confidence intervals (with discussions). *Jour. Amer. Stat. Assoc*. 82, 171-200.

Efron, B. (1992). Jackknife-after-bootstrap standard errors and influences functions (with discussions). *J.R. Stat. Soc*. B. 54, 83-127.

Efron, B. (1994). Missing data, imputation and the bootstrap (with discussions). *Jour. Amer. Stat. Assoc*. 89, 463-479.

Efron, B. and Tibshirani, R.J. (1993). *AN INTRODUCTION TO THE BOOTSTRAP*, Chapman and Hall New York.

Freedman, D.A. (1981) Bootstrapping Regression models. *Ann. Stat*. 9, 1281- 1228.

Hall, P. (1989). On efficient bootstrap simulation. *Boimetrika*. 76, 613-617.

Hall, P. (1992). Bootstrap confidence intervals in nonparametric regression. *Ann. Stat*. 20, 695-711.

Hall, P. (1992). THE *BOOTSTRAP AND EDGEWORTH EXPANSION*. Springer Verlag, N.Y.

Hall. P. (1988). Theoretical comparison of bootstrap confidence intervals (with discussions). *Ann. Stat*., 16, 927-953.

Hinkley, D.V. (1988). Bootstrap methods (with discussions*). J. Roy. Stat. Soc*. B, 50, 321-337.

Hwang, J.S., Sen, P.K. and Shao, J. (1996). Bootstrapping a sample quantile when the density has a jump. *Stat. Sinica*. 6, 1996.

Kunch, H.R. (1989). The jackknife and bootstrap for general stationary observations. *Ann. Stat*. 17, 1217-1241.

Lahiri, S.N. (1993). Bootstrapping the studentized sample mean of Lattice variables. *J. Mult. Analy*. 45, 247-256.

Lahiri, S.N. (1993). On the moving block bootstrap under long range dependence. *Stat. Prob. Letters*. 18, 405-413.

Liu, R.Y. and Singh, K. (1992). Efficiency and Robustness in re sampling. *Ann. Stat*. 20, 370-384.

Liu, R.Y. and Singh, K. (1992). Moving block jackknife and bootstrap capture weak dependence. *EXPLORING THE LIMITS OF BOOTSTRAP*, R. Lepage and L. Billard edited. Wiley, N.Y.

Lunneborg, EE. (2000). *DATA ANALYSIS BY RESAMPLING: CONCEPTS AND APLLICATIONS*. Duxbury Press.

Mamman, E. (1992). *WHEN DOES BOOTSTRAP WORK. ASYMPTOTOC RESULTS AND SIMULATIONS*. Springer Verlag, N.Y.

Politis, D.N. and Romano, J.P. (1994). The stationary bootstrap. *Jour. Amer. Stat. Assoc*. 89, 1303 – 1313.

Rubin, D.B. (1981). The Bayesian bootstrap. *Ann. Stat*. 9, 130-134.

Shao, J. and Tu, D. (1995). *THE JACKKNIFE AND BOOTSTRAP*, Springer, Verlag, N.Y.

Singh, K. (1981). On Asymptotic accuracy of Efron's bootstrap. *Ann. Stat*. 9, 1187-1195.

Singh, K (1998). Breakdown theory for bootstrap quantiles. *Ann. Stat*. 26, 1719-1732.

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

Taylor, C.C. (1989). Bootstrap choice of smoothing parameter in kernel density estimation. *Biometrika*. 76, 705-712.

Tibshirani, R.J. (1988). Variance stabilization and the bootstrap. *Biometrika*. 75, 433-444.

Wu, C.F.J. (1986). Jackknife, bootstrap and other resampling procedures (with discussions). *Ann. Stat*. 14, 1261-1350.

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