Licentiate thesis

On variance estimation and a goodness-of-fit test using the bootstrap method

Saeid Amiri

Institutionen för energi och teknik
Department of Energy and Technology

Licentiate thesis/Report 007
ISSN 1654-9406
Uppsala 2009

On variance estimation and a goodness-of-fit test using the bootstrap method

Saeid Amiri

Department of Energy and Technology, SLU,
Uppsala

Licentiate Thesis
Swedish University of Agricultural Sciences
Uppsala 2009

1654-9406

ISSN, 1654-9406
ISBN, 978-91-86197-20-9
© 2009 Saeid Amiri, Uppsala
Tryck: SLU Service/Repro, Uppsala 2009

Abstract

Saeid Amiri. *On variance estimation and a goodness-of-fit test using the bootstrap method*. Licentiate Thesis.

ISSN, 1654-9406, ISBN, 978-91-86197-20-9.

This thesis deals with the study of variance estimation using the bootstrap method, including the problem of choosing between nonparametric and parametric bootstrap methods. Paper I compares the two approaches, determines which method is preferable and analyses the accuracy of the approximations. The underlying concept of parametric bootstrap is based on the assumption of correct choice of parametric distribution. Paper II therefore considers goodness-of-fit tests and presents a new test based on the bootstrap method.

Key words: Nonparametric Bootstrap, Parametric bootstrap, Goodness-of-fit test, Variance.

Author's address: Box 7032, 750 07 Uppsala, Sweden.

Articles appended to the thesis

The thesis is based on the following articles.

I A comparison of bootstrap methods for variance estimation
Report no 2009:02 Centre of Biostochastics. Swedish University of Agriculture Sciences

II On a generalization of the Jarque-Bera test using the bootstrap method
Report no 2009:03 Centre of Biostochastics. Swedish University of Agriculture Sciences

Contents

1 Introduction

The aim of this thesis is to study estimation of variance by using bootstrap methods. Chapter I reviews and explains variance, kurtosis and the bootstrap approach used in the Papers I and II. Chapter 2 expends the obtained and presented result in Paper I and II. It includes bootstrapping of variance, the confidence interval of variance and goodness-of-fit testing by the bootstrap method.

Paper I discusses the bootstrap approach. It compares nonparametric and parametric bootstrap estimation of variance and shows that bootstrap estimations of variance either by the parametric or nonparametric method are equal but the bootstrap standard error depends on the sample kurtosis. Paper I also discusses conditions where the nonparametric bootstrap is better regardless of whether the distribution of the parametric bootstrap and the real distribution belong to the same distribution family.

Paper II considers goodness-of-fit tests by studying the underlying distribution of the sample data and supports paper I because the parametric bootstrap is based on correct choice of the sampling distribution.

1.1 Variance

Let $X_1, \cdots, X_n \overset{iid}{\sim} F$. The sample variance is then given as:

$$S_1^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2. \tag{1}$$

Its expected value and variance can be found, see Kendall and Stuart (1958),

$$E(S_1^2) = \frac{n-1}{n}\sigma^2, \tag{2}$$

$$\begin{aligned} V(S_1^2) &= \frac{(n-1)^2}{n^3}\mu_4 - \frac{(n-1)(n-3)}{n^3}\mu_2^2 \\ &= \big(\frac{n-1}{n^3}\big)(\sigma^2)^2\big((n-1)K - (n-3)\big). \end{aligned} \tag{3}$$

Where μ_i is the ith centrad moments, K is the kurtosis and σ^2 is the population variance:

$$K = \frac{E(X-\mu)^4}{(E(X-\mu)^2)^2}. \tag{4}$$

It is clear that σ^2 and K have direct effects on $V(S_1^2)$. The same discussions hold for $S_2^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$. The quantities S_1^2 and S_2^2 are the MLE and UMVUE estimators of variance for the normal distribution. Another estimator is given by Searle and Intarapanich (1990):

$$S_W^2 = \frac{n}{n(n+1) + (K-3)(n-1)}\sum_{i=1}^{n}(X_i - \overline{X})^2, \tag{5}$$

which has the minimum square error (MSE). In the case of the normal distribution, it reduces to $S_W^2 = \frac{1}{n+1}\sum_{i=1}^{n}(X_i - \overline{X})^2$.

1.2 Kurtosis

The concept of kurtosis was introduced by Pearson (1905), as a shape parameter that is a qualitative property of the distribution, as well as its skewness. Kurtosis is usually of interest only when dealing with approximately symmetric distributions. It is defined in many statistical books as e.g. a measure of tail heaviness in comparison with the normal distribution or a qualitative property of a distribution that measures peakness and which can be used to judge the deviation of a distribution from normality. The sample kurtosis is:

$$\widehat{K} = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^4/n}{\left(\sum_{i=1}^{n}(x_i - \overline{x})^2/n\right)^2}. \tag{6}$$

Based on the value of kurtosis, the distribution is defined as platykurtic, or flat-topped when $K < 3$, which is more concentrated about the mean and flatter than the normal, mesokurtic when $K = 0$ and leptokurtic when $K > 3$ and it is peaked around the mean and also has more probability in the tails of the distribution than the normal. Skewed distributions are always leptokurtic, see Everitt and Howell (2005).

Nevertheless this classification is not always correct. Chissom (1970) showed that there are many cases when $K < 3$ but the distribution is not flattened. It can be defined as a measure of unimodality versus bimodality, see Darlington (1970) and Chissom (1970) and hence there is no universal agreement about it. It can be seen that the kurtosis is defined as the expectation of the fourth power of the standardized variable which, simplifies to

$$K = Var(Z^2) + 1, \tag{7}$$

where $Z = (X - \mu)/\sigma$. This is used by Darlington (1970) to define kurtosis as bimodality, but Hildebrand (1971) gives two distributions to demonstrate that this does not always hold. Furthermore Moors (1986) explains the easy interpretation of kurtosis as a measure of dispersion around the two values $\mu \pm \sigma$. He notes that the existence of two possibilities cause the confusion about the interpretation of kurtosis, it is an inverse for concentration in these points. High kurtosis can arise when 1) concentration of probability mass is near μ (corresponding to a peak unimodal distribution) and 2) concentration of probability mass is in the tail of the distribution.

Formula (7) shows that the kurtosis is more strongly affected by the tail behavior of distribution than the center of the distribution. High kurtosis has the potential to have outliers in one or both tails of the distribution, see Everitt and Howell (2005). Nevertheless, many distributions have their own value of kurtosis which can be used to study them or to find a confidence interval for the mean, see Guttman (1948). Alternative methods are discussed in Ruppert (1987).

On the bounds of kurtosis, Johnson and Jr (1978) showed that:

$$\widehat{\gamma}^2 \leq \widehat{K} \leq n$$

where $\widehat{\gamma}$ is the sample skewness.

If $X \sim N(\mu, \sigma^2)$, it can be shown that $K = 3$ but in reality it is rare that its estimator becomes this value, whereas $\widehat{K} > 3$ or $\widehat{K} < 3$ are more likely to occur. Based on (3), this does not affect on the biasedness of variance but it influences the accuracy of variance. It can be seen that if $\widehat{K} < 3$ holds, $\widehat{V}(S^2)$ decreases and vice versa. In addition, Cramér (1945) shows that $E(\widehat{K}) = \frac{-6}{n+1} + 3$ which means that if the distribution is normal then $\widehat{K} < 3$ is more likely to be observed.

Finding $E(\widehat{K})$ for other distributions is rather difficult while it can easily be studied by simulations. Cramér (1945, pp. 356) proves that the estimation of kurtosis converges with $O(n^{-1})$. An and Ahmed (2008) discuss a different version of $\widehat{K}$ which reduces the biasedness.

In the case of the normal distribution it can be shown that the square of the sample variance and sample kurtosis are independent but for other distributions it is rather difficult to derive explicit results. By simulations can show that the correlation of $(\widehat{K}, S^4)$ is positive for the t and chi-square distribution but negative for the uniform distribution.

1.3 Principles of the bootstrap method

The last three decades have brought a vast new body of statistics in the form of nonparametric approaches to modeling uncertainty, in which it is not the individual parameters of the probability distribution but rather the entire distribution is sought, based on the empirical data available. The concept was first introduced in the seminal paper of Efron (1979) as the bootstrap method. Similar ideas have since been suggested in different contexts as an attempt to give a new perspective to an old and established statistical procedure known as jackknifing. Unlike Jackknifing, which is mostly concerned with calculating standard errors of the statistics of interest, Efron's Bootstrap method has the more ambitious goal of estimating not only the standard error but also the distribution of the statistics.

The idea behind the bootstrap method is not far from the traditional statistical methods and provides a complement to these. To discuss the parameter $\theta(F)$ let us look at the mean.

$$\theta = \theta(F) = \mu = \int x dF(x) = E_F(x). \tag{8}$$

The same functional of the sample distribution function $\widehat{F}$ can be used:

$$\widehat{\theta}_n = \theta(F_n) = \bar{X} = \int x dF_n(x) = E_{F_n}(x), \tag{9}$$

where $F_n(x)$ is the empirical distribution function. $\theta(F_n)$ needs some measures such as $\lambda_n(F)$ of its performance, which can be the bias of $\widehat{\theta}_n$ or the variance of $\sqrt{n}\widehat{\theta}_n$, see Lehmann (1999). The bootstrap method involves using a functional $h(F)$ by means of $h(F_n)$ and correspondingly estimating $\lambda_n(F)$ by the plug-in estimator $\lambda_n(F_n)$ via resampling. $\widehat{\theta}_n = \theta(X_1, \cdots, X_n)$ expresses $\widehat{\theta}_n$ as a function of F_n but directly as a function of sample $\mathcal{X} = (X_1, \cdots, X_n)$. The dependency of $\widehat{\theta}_n$ results from the fact that $\mathcal{X}$ is a function of F and any random sample of $\mathcal{X}$ is also a sample of F. Therefore to replace F by F_n in the distribution governing $\widehat{\theta}_n$, one should replace it by

$$\theta_n^* = \theta(X_1^*, \cdots, X_n^*), \tag{10}$$

where $(X_1^*, \cdots, X_n^*)$ is a sample from F_n, which is not the actual dataset $\mathcal{X}$ but rather a randomized or resampled version of $\mathcal{X}$. In other words, $(X_1^*, \cdots, X_n^*)$ is the set which consists of members of the original dataset, some of which appear zero times, some once, twice, and so on. This sample is conceptual sample from F_n which assign probability $\frac{1}{n}$ to each of the observed value $(x_1, \cdots, x_n)$. The most important property of the bootstrap method relies on the conditional independence of the given original sample.

From this brief discussion of bootstrap, it is obvious that the nonparametric bootstrap technique frees the analysis from the most typical assumptions, making it more attractive to researches in applied fields. In nonparametric statistics, the concept of bootstrap is known by the somewhat broader term of resampling. The bootstrap method can arguably be an instrument in understanding the structure of the random variable and in error estimation of existing models.

The aim of this work is to study variance. The followings are the steps to bootstrapping variance:
1. Suppose $\mathcal{X} = (X_1, \cdots, X_n)$ is an i.i.d. random sample of the distribution F. Assume $Var(X) = \sigma^2$.
2. We are interested in $\theta(F)$ and consider a plug-in estimation: $\widehat{\theta} = \theta(F_n)$ or $\widehat{\theta} = \widehat{\theta}(X_1, \cdots, X_n)$, where F_n is the empirical distribution function, i.e. $F_n(x) = \frac{1}{n}\sum_{j=1}^{n} I(X_j \leq x)$.
3. Generate bootstrap samples, which can be done by two different approaches:
(i) The nonparametric bootstrap method: $X_{ij}^* \overset{iid}{\sim} F_n, i = 1, \ldots, B,\ j = 1, \ldots, n$. Note that if $Z \sim F_n$ then $EZ = \overline{X}$ and $V(Z) = S_X^2$, where S_X^2 is the second moment estimator:

$$S_X^2 = \frac{1}{n}\sum_{j=1}^{n} X_j^2 - (\overline{X})^2.$$

The kurtosis of F_n is:

$$\widehat{K} = K_{F_n} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^4/n}{\Big(\sum_{i=1}^{n}(X_i - \overline{X})^2/n\Big)^2}. \tag{11}$$

(ii) The parametric bootstrap method: $X_{ij}^{\#} \overset{iid}{\sim} G_{\widehat{\lambda}}, i = 1, \ldots, B,\ j = 1, \ldots, n$ where $G_{\widehat{\lambda}} = G(.|\mathcal{X})$ is an element of a class $\{G_\lambda, \lambda \in \Lambda\}$ of distributions. The parameter λ is estimated by statistical methods such that the expectation of $G_{\widehat{\lambda}}$ equals $\overline{X}$ and the variance of $G_{\widehat{\lambda}}$ is S_X^2. It should be mentioned that the symbols $*$ and $\#$ are used for the nonparametric and parametric bootstrap, respectively, in the remainder of this work. The kurtosis of $K_{G(.|\mathcal{X})}$ is denoted by

$$K_{G(.|\mathcal{X})} = \frac{E_{\mathcal{X}}(X - \overline{x})^4}{(E_{\mathcal{X}}(X - \overline{x})^2)^2}$$

4. Calculate the bootstrap replications of the estimator:

$$\widehat{\theta}(X_i^{\times}) = \widehat{\theta}(X_{i1}^{\times}, \ldots, X_{in}^{\times})\ i = 1, \ldots, B,$$

The symbol $\times$ is used when either the nonparametric or parametric procedure holds.
5. Handle the bootstrap replications as i.i.d. random samples of $\widehat{\theta}$ and consider the sample mean and the sample variance. These are:

$$\widehat{\theta}^{\times} = \frac{1}{B}\sum_{i=1}^{B}\widehat{\theta}(X_i^{\times}), \tag{12}$$

$$\widehat{V}^{\times} = \frac{1}{B}\sum_{i=1}^{B}\left(\widehat{\theta}(X_i^{\times}) - \widehat{\theta}^{\times}\right)^2 = \frac{\sum_{i=1}^{B}\left(\widehat{\theta}(X_i^{\times})\right)^2}{B} - (\widehat{\theta}^{\times})^2. \tag{13}$$

The main question examined in this thesis is which method is more appropriate to use. A secondly question when F and $G(.|\mathcal{X})$ come from the same distribution family is whether it is possible that the nonparametric method is better. Paper I presents a comparison of the nonparametric and parametric bootstrap methods for estimating variance.

1.4 Goodness-of-fit testing

The science of statistics is based on the use of statistical distributions. The efficiency of a statistical procedure relies on how accurately the underlying distribution can be identified. Most of the statistical textbooks explain this topic. There are many different methods for evaluating the statistical distribution of data which are called goodness-of-fit tests. With the limited information contained in a sample, it is impossible to identify the underlying distribution exactly. Some methods are frequently used, such as the Q-Q plot, Shapiro-Wilk test and Jarque-Bera test. The existing software is dominated by these tests, although there are a variety of alternative tests such as as the P-P plot, chi-square, Kolmogorov-Smirnov and Anderson-Durling goodness-of-fit tests that have their own features. It is true that for a small sample size up to a moderate sample size, they are not successful, but they are fairly useful in distinguishing normal from extremely non-normal distributions with respect to symmetry and skewness, see Tiku and Akkaya (2004).

Q-Q plot is a graph that measures the conformity between the empirical distribution and the given theoretical distribution. Let $X_1, \cdots, X_n \overset{iid}{\sim} F$ with the ordered sample $X_{(1)}, \cdots, X_{(n)}$. Associated with each point is the $\frac{i}{n+1}$-quantile q_i of the given distribution. Plot $X_{(i)}$ against q_i. If it is close to a straight line then the sample qualifies as being from a plausible distribution. However, this is a subjective statistic and a formal test is preferable.

The Shapiro-Wilk test statistics is follow:

$$SW = \frac{\big(\sum_{i=1}^{n} a_i X_{(i)}\big)^2}{\sum_{i=1}^{n} \big(X_i - \overline{X}\big)^2}, \tag{14}$$

where $x_{(i)}$ is the order statistics and a_i is a constant generated from the means, variances and covariances of the order statistics of a sample of size n from the normal distribution. The Shapiro-Wilk test actually compares an estimate of the standard deviation using a linear combination of the order statistics with the usual estimate. Thode (2002) recommends it for everyday practice.

The Jarque-Bera statistics is as follow:

$$JB = \frac{n}{6}\left(\widehat{\gamma}^2 + (\widehat{K} - 3)^2/4\right), \tag{15}$$

where $\widehat{\gamma}$ and $\widehat{K}$ are the sample skewness and kurtosis. JB has an asymptotic chi-square distribution with two degrees of freedom. It is obvious that the first one is based on the order statistic but the second uses the main parameters which are constant values.

Goodness-of-fit test is covered by D'Agostino and Stephens (1986), Rayner and Best (1989) and Thode (2002).

2 Contributions

This chapter discusses the relative contributions of Paper I and II. More precisely, it can be divided into two parts; the study of bootstrap variance estimation and goodness-of-fit testing. Since the parametric bootstrap relies on the choice of distribution, study of the underlying distribution of the sample data is important.

2.1 Paper I

Paper I comprises a comparison of the two bootstrap approaches for variance estimation. As mentioned above, the bootstrap method can basically be applied in two different ways, the nonparametric and parametric method. It can be shown that there is no difference in performance between them for the mean, whereas in the case of variance, the behavior of the nonparametric and parametric bootstrap method is completely different. The object of Paper I is to explore them in detail because of the importance of variance estimation.

Although the nonparametric and parametric methods are simultaneously considered by some authors, the results of the simulations are often given without explicit discussions of their differing performance. For example, Efron and Tibshirani (1993) discuss the nonparametric and parametric bootstrap confidence intervals of variance by using an example, Ostaszewski and Rempala (2000) explains how to use the bootstrap method in the actuary sciences and Lee (1994) explains how to use it in tuning parameters to find more accurate estimations.

Since the bootstrap method is based on the sample, the expectation is that it can be used to study it. As Hall (1992a) says, the bootstrap method may be expressed as an expectation conditional on sample or equality as an integral with respect to the sample distribution function. This allows us to do a direct comparison of the nonparametric and parametric bootstrap methods. It should be mentioned that two kind of expectations are discussed, conditional and unconditional. The conditional expectation clarifies the result of the bootstrap approach whereas the unconditional expectation is the combination of the bootstrap and the frequenist approaches. As the aim of bootstrap method is to approach the parameter of interest, hence it is chosen as the criterion for that. The bootstrap methods are studied using the bias and MSE. The following theorem is one of the main results of Paper I, which describes steps to study the biasedness.

Theorem 1: Let $\mathcal{X} = (X_1, \cdots, X_n) \overset{iid}{\sim} F$ with $EX^4 < \infty$. Then for the explained bootstrap methods,

$$E(S^{2*}|\mathcal{X}) \quad = \quad E(S^{2\#}|\mathcal{X}). \tag{16}$$

$$K_{F_n} < K_{G(.|\mathcal{X})} \quad \Longleftrightarrow \quad E(V^*|\mathcal{X}) < E(V^{\#}|\mathcal{X}), \tag{17}$$

where K_{F_n} and $K_{G(.|\mathcal{X})}$ are the sample kurtosis and the kurtosis corresponding to the parametric distribution $G_{\widehat{\lambda}}$.

The theorem implies that the unconditional expectation of the bootstrap estimator of the parametric and nonparametric methods for variance estimation are equal

$$E(S^{2*}) = E(S^{2\#}) \quad = \quad \left(\frac{n-1}{n}\right)E(S_X^2) = \left(\frac{n-1}{n}\right)^2\sigma^2.$$

$$\frac{Bn^3}{(B-1)(n-1)(n^2-2n+3)}E(V^*) \quad = \quad \frac{Bn^3}{(B-1)(n^2-1)n}E(V^{\#}) = V(S_X^2).$$

It is obvious that $E(S^{2\times}) < E(S_X^2) < \sigma^2$. Relation (17) indicates that by using the sample kurtosis one can study the relative performance of the parametric and nonparametric bootstrap methods.

Theorem 2 in Paper I gives the expectation of $V^\times$. This theorem states that $E(V^*)$ depends on K, whereas $E(V^{\#})$ depends on K and $K_{G(.|\mathcal{X})}$. It should be noted that if $K_{G(.|\mathcal{X})}$ depends on the observations, for example the lognormal distribution, then it is impossible to present a closed form in general. Hence in this case the study of the performance of the parametric bootstrap is rather difficult but for the nonparametric bootstrap, it holds all the time. In the case of the normal distribution, Corollary 1 in Paper I states that:

$$E(V^*) \quad < \quad E(V^{\#}) < V(S_X^2), \tag{18}$$

$$\frac{Bn^3}{(B-1)(n-1)(n^2-2n+3)}E(V^*) \quad = \quad \frac{Bn^3}{(B-1)(n^2-1)n}E(V^{\#}) = V(S_X^2). \tag{19}$$

If the underlying distributions of F and $G(.|\mathcal{X})$ belong to the normal distribution family, it is expected that the standard error of the parametric bootstrap of variance will be close to F in comparison with the nonparametric bootstrap. Furthermore, by using the corrections given in (19), it is possible to find the unbiased estimation of parametric and nonparametric bootstrap of variance.

It is interesting that when $K_{F_n} > 3$, then $E(V^*|\mathcal{X}) > E(V^{\#}|\mathcal{X})$ and also $V(S_X^2)$ is larger than the expectation bootstrap estimation, (19). Therefore V^* is more likely to be closed to $V(S_X^2)$ than $V^{\#}$. Table 3 in Paper I explains this by simulations.

The most important result is that for the distribution with the kurtosis between 1.4 an 2, the nonparametric bootstrap has less bias than the parametric bootstrap, regardless of whether F and $G(.|\mathcal{X})$ have the same distribution. Example 2 clarifies this result.

In paper I, Lemma 1 and Lemma 2 discuss the conditional and unconditional MSE of the bootstrap variance. Lemma 3 gives the conditional MSE of $V^{\times}$ and Theorem 4 discusses the unconditional MSE of $V^{\times}$.

2.2 Bootstrap CI of variance

This section discusses the bootstrap confidence interval (CI) estimation of variance, parts of which are briefly discussed in Example 4 in Paper I.

The confidence interval (CI) estimation of variance is explained by Cojbasic and Tomovic (2007). It is based on two other papers that discuss the CI estimation of mean, Zhou (2000) and Hall (1992b). These three papers focus on the Edgeworth expansion and attempt to prove the efficiency of the Edgeworth expansion using the bootstrap method. The results of Cojbasic and Tomovic (2007) are not far from the results presented by Zhou (2000) and Hall (1992b) for the mean.

The aim of this thesis is to explain the effect of kurtosis on the bootstrap CI estimation which was ignored in previous publications. This work includes a comparison of the nonparametric and parametric bootstrap of variance, on which kurtosis has the principle effect. The parametric and nonparametric bootstrap CI are illustrated in Efron and Tibshirani (1993) by an example, but the effect of kurtosis is ignored. Their suggestion is BC_α but it is shown that for the data with a normal distribution this might be not appropriate.

As Efron and Tibshirani (1993) note, one of the principal goals of bootstrap theory is to produce good confidence interval automatically, which means that the bootstrap should be close to the exact confidence interval. It exists for the mean and variance but it is rather difficult to find for most statistics. One of the objects of this thesis is to perform a comparison of the nonparametric and parametric bootstrap approaches to determine which can achieve this aim. The following provides the standard definition of CI, then discusses the different methods of performing bootstrap CI of variance and makes comparisons of them.

2.2.1 Confidence interval

Let $(\underline{\theta}(X_1, \cdots, X_n), \overline{\theta}(X_1, \cdots, X_n))$ be a region which has a guaranteed probability of containing the unknown parameter θ, i.e. for which

$$P_\theta\left[(\underline{\theta}(X_1, \cdots, X_n) \leq \theta \leq \overline{\theta}(X_1, \cdots, X_n))\right] > 1 - \alpha \tag{20}$$

for some preassigned α. The coverage of $(\overline{\theta}, \underline{\theta})$ is the probability that the random interval covers the true parameter. It is obvious that the interval is random, not the parameter, because it is based on observations, and therefore the probability refers to X not θ. Since high coverage probability requires unnecessarily wide intervals, the following definition is more useful:

$$\underset{\theta}{inf}\, P_\theta\big[(\underline{\theta}(X_1, \cdots, X_n) \leq \theta \leq \overline{\theta}(X_1, \cdots, X_n))\big] = 1 - \alpha. \tag{21}$$

As Lehmann (1999) says, this holds when n is large enough and can be replaced by the weaker requirement:

$$\underset{\theta}{inf}\, P_\theta\big[(\underline{\theta}(X_1, \cdots, X_n) \leq \theta \leq \overline{\theta}(X_1, \cdots, X_n))\big] \longrightarrow 1 - \alpha \;\; as \;\; n \longrightarrow \infty \tag{22}$$

A still weaker condition is:

$$P_\theta\big[(\underline{\theta}(X_1, \cdots, X_n) \leq \theta \leq \overline{\theta}(X_1, \cdots, X_n))\big] \longrightarrow 1 - \alpha \;\; as \;\; n \longrightarrow \infty \tag{23}$$

There are different ways to construct CI, see Lehmann (1999), Casella and Berger (2002), Shao (1999) and Polansky (2008).

Two properties of any given CI are required: consistency which expresses that the CI contains θ as the sample size becomes larger, and accuracy which points out how much the CI of a given method is close to the exact CI, if it exists.

The bootstrap CI is based on the simulations, since the accuracy is more important than the consistency. However, most methods of bootstrap CI are constructed in such a way that the consistency is maintained but the accuracy is not. The methods can overestimate or underestimate.

2.2.2 Methods

Many different ways of constructing the bootstrap CI are discussed and are described in detail in Efron and Tibshirani (1993) and DiCiccio and Efron (1996) among many authors.

Method I. The most commonly applied method uses the central limit theorem (CLT). The quantities $(X_i - \overline{X})^2$ are not iid, hence the direct use of CLT is not possible. To use CLT, Theorem (8.16) in Lehman (1999) can be used to find an asymptotic distribution. The approximated confidence interval, by using $t_{\alpha/2,n-1}$ can be found as below.

$$S^2 \pm t_{\alpha/2,n-1} S^2 \sqrt{\frac{(\widehat{K} - 1)}{n}}, \tag{24}$$

where $\widehat{K}$ is the sample kurtosis.

Method II. Let $X_i \sim N(\mu, \sigma^2)$. It is easy to find the CI for σ^2

$$\Big(\frac{nS^2}{\chi^2_{\alpha/2,n-1}}, \frac{nS^2}{\chi^2_{1-\alpha/2,n-1}}\Big). \tag{25}$$

In the following, some bootstrap methods are given for which the complete discussion can be found in Hall (1992a), Efron and Tibshirani (1993) and Davison and

Hinkley (1996). As previously mentioned, the bootstrap method can be either nonparametric or parametric. The following includes a general discussions that holds for both of these. In all discussions "*" and "#" refer to the nonparametric and parametric method, respectively. When a statement holds for both of them, the symbol "$\times$" is used.

Method III. This method is referred to as the standard method, see Efron and Tibshirani (1993)

$$S^2 \pm t_{\alpha/2}\widehat{se}_B^{\times}, \tag{26}$$

where $\widehat{se}_B^{\times}$ is the bootstrap estimate of standard error.

Method IV. The CI of variance based on the bootstrap methods can be found using the following formula:

$$S^2 \pm t_{\alpha/2}^{\times} S^2 \sqrt{\frac{(\widehat{K}-1)}{n}}, \tag{27}$$

where $t_{\alpha/2}^{\times}$ is the $\alpha/2$ percentile of $t^{\times}$,

$$t^{\times} = \frac{S^{2\times} - S^2}{\sqrt{V(S^2)^{\times}}}, \tag{28}$$

where $S^{2\times}$ and $V(S^2)^{\times}$ are estimated by the bootstrap method. It is called the t-bootstrap interval. It is obvious that bootstrap resampling is used in estimation of the variance and the standard error. It should be noted that it is not a pure bootstrap method, since the CI includes $\sqrt{V(S^2)}$. The aim of bootstrap is to automatically find the standard error of the parameter of interest.

Method V. Using Method II, the CI is calculated asymptotically as:

$$\left(\frac{nS^2}{\chi_{\alpha/2}^{2\times}}, \frac{nS^2}{\chi_{1-\alpha/2}^{2\times}}\right) \tag{29}$$

where $\chi_{\alpha/2}^{2\times}$ is the percentile of the following quantity,

$$\chi^{2\times} = \frac{nS^{2\times}}{S^2}. \tag{30}$$

Method VI. This method is called the percentile CI,

$$[\widehat{\theta}_{\%low}, \widehat{\theta}_{\%up}] = [\widehat{G}^{-1}(\alpha/2), \widehat{G}^{-1}(1-\alpha/2)], \tag{31}$$

where $\widehat{G}^{-1}(\alpha/2) = S^{2\times}(\alpha/2)$, i.e. it is the percentile of the bootstrap resampling of variance.

Method VII. This method is called bias-corrected and accelerated, BC_α, by Efron and Tibshirani (1993), and is a substantial improvement on the percentile method in both theory and practice. It is based on the percentile of the bootstrap distribution. It is defined as:

$$[\widehat{\theta}_{\%low}, \widehat{\theta}_{\%up}] = [\widehat{G}^{-1}(\alpha/2), \widehat{G}^{-1}(1-\alpha/2)], \tag{32}$$

where

$$\alpha_1 = \Phi\Big(\widehat{z}_0 + \frac{\widehat{z}_0 + z_{\alpha/2}}{1 - \widehat{a}(\widehat{z}_0 + z_{\alpha/2})}\Big),$$
$$\alpha_2 = \Phi\Big(\widehat{z}_0 + \frac{\widehat{z}_0 + z_{1-\alpha/2}}{1 - \widehat{a}(\widehat{z}_0 + z_{1-\alpha/2})}\Big)$$

where $z_{\alpha/2}$ is the 100αth percentile point of a standard normal distribution. $\widehat{a}$ and $\widehat{z}_0$ change the percentiles used for the BC_α endpoints. These changes correct certain deficiencies of the standard and percentiles methods. The value of $\widehat{z}_0$, bias-correction, is obtained directly from the proportion of bootstrap replications less than the original estimate θ:

$$\widehat{z}_0 = \Phi^{-1}\Big(\frac{\#\{\widehat{\theta}^{\times}(b) < \widehat{\theta}\}}{B}\Big), \tag{33}$$

Φ^{-1} indicating the inverse function of a standard normal CDF. $\widehat{z}_0$ measures the median bias of $\widehat{\theta}^{\times}$, that is the discrepancy between the median $\widehat{\theta}^{\times}$ and $\widehat{\theta}$ in normal units.

The quantity of $\widehat{a}$ is called acceleration because it refers to the rate of change of the standard error of $\widehat{\theta}$ with respect to the true parameter θ. The easiest way to calculate $\widehat{a}$ is to use the jackknife value of $\widehat{\theta}$. Let X_i be the original sample with the ith value X_i deleted, $\widehat{\theta}_{(i)} = \widehat{\theta}(X_{(i)})$ and $\widehat{\theta}_{(.)} = \sum_{i=1}^{n} \widehat{\theta}_{(i)}/n$, then a simple acceleration is

$$\widehat{a} = \frac{\sum_{i=1}^{n}(\widehat{\theta}_{(.)} - \widehat{\theta}_{(i)})^3}{6\big(\sum_{i=1}^{n}(\widehat{\theta}_{(.)} - \widehat{\theta}_{(i)})^2\big)^{3/2}}. \tag{34}$$

Efron and Tibshirani (1993) explains this and also suggest it for the CI of variance.

There are other methods that can be used directly based on the bootstrap e.g. the Edgeworth expansion or the Edgeworth Bootstrap. It should be noted that the CI based on the Edgeworth expansion uses the effect of skewness to adjust t which is not considered here.

2.2.3 Results

In this section, numerical simulation of the variance bootstrap method are presented in order to clarify the results relating to the methods listed in the previous section. The data set is given in Table 1. The variable x is the spatial perception of 26 neurologically impaired children, which Efron and Tibshirani (1993) used to study the variance. The data set y is the same as x except that two observations are changed to increase the estimated kurtosis from 2.592 to 3.411 in order to reveal how much the CI bootstrap is affected by it.

Table 1: Data set used in the simulations

Variable	Data	S^2	$\widehat{K}$	Skewness
x	48 36 20 29 42 42 20 42 22 41 45 14 6 0 33 28 34 4 32 24 47 41 24 26 30 41	171.53	2.592	-0.638
y	48 36 20 29 42 42 20 42 22 41 45 14 30 0 33 28 34 24 32 24 47 41 24 26 30 41	123.07	3.411	-0.667

Efron and Tibshirani (1993, Table 14.2) analyzed the nonparametric and parametric bootstrap CI of $\theta = V(x)$. In comparisons, it is obvious that the nonparametric method has smaller length except for t-bootstrap. This does not happen randomly but has not been explained by them. This can be proved by Theorem 1 in Paper I, which states that the variance of variance, the square root of which is the standard error of variance, depends on the sample kurtosis. If the sample kurtosis is less than 3, then it is expected that the nonparametric estimate of that variance of variance will be less than the parametric estimation, or in other words that the spread of S^{2*} will be less than of $S^{2\#}$. Since the kurtosis of x is less than 3 (Table 1), this result is expected.

Table 2: Confidence interval at 95% for x

Method	Lower limit	Upper limit	Length	Shape
I	99.018	244.049	145.031	1
II	118.448	305.233	186.784	2.518
III non	100.064	243.003	142.938	1
III par	91.483	251.584	160.101	1
IV non	110.249	283.828	173.578	1.832
IV par	115.847	309.760	193.912	2.482
V non	124.379	306.281	181.902	2.857
V par	120.598	311.475	190.876	2.747
VI non	99.927	233.364	133.437	0.863
VI par	96.051	248.405	152.353	1.018
VII non	119.520	258.307	138.786	1.668
VII par	113.565	289.907	176.342	2.042

To clarify this result, Method I-VII are repeated for x and y, which have $\widehat{K} < 3$ and $\widehat{K} > 3$, respectively. Table 2 and Table 3 show the length and shape are (Upper limit-Lower limit) and (Upper limit-S^2)/(S^2-Lower limit), which are measure of wideness and asymmetry of interval. The first two lines of both tables are the standard method for constructing CI of variance, which is based on t and χ^2. It is obvious that Method I has smaller length than Method II because the former is based on the symmetrical distribution but in reality the distribution of variance is asymmetrical. Method II is known as the exact method, as a criterion which can be used to study the different methods.

Table 3: Confidence interval at 95% for y

Method	Lower limit	Upper limit	Length	Shape
I	59.050	187.094	128.043	1
II	84.984	218.999	134.014	2.518
III non	61.886	184.258	122.372	1
III par	66.084	180.060	113.976	1
IV non	75.743	295.122	219.379	3.635
IV par	77.494	236.850	159.356	2.496
V non	83.498	225.435	141.936	2.586
V par	85.460	219.482	134.022	2.563
VI non	65.223	183.094	117.870	1.037
VI par	69.262	175.156	103.660	1.009
VII non	79.792	227.236	147.443	2.406
VII par	84.531	217.723	133.192	2.455

The aim here is the comparison of nonparametric and parametric bootstrap. In Table 2, the nonparametric bootstrap CI is shorter than the parametric bootstrap CI, but in Table 3 the opposite occurs because Method III uses the square root of $V(S^2)$ which depends on the kurtosis, this method is directly affected by kurtosis. Method IV uses bootstrap resamples in t. Although Method V and VI do not use $V(S^2)^{\times}$ directly, they are based on the 5th and 95th percentiles and of course the spread is affected directly by kurtosis.

Table 4 includes a comparison of the different CI bootstrap methods in 1000 simulations. The entries in this table show how many times the parametric CI is shorter than the nonparametric CI. The results thus show how much the kurtosis affects on the spread of variance estimation of bootstrap resampling. In other words, if $\widehat{K} < 3$ then CI based on the nonparametric bootstrap will be conservative and vice versa.

Table 4: Comparison of CI at 95%.

	Data	
Method	x	y
III	0.001	0.998
IV	0.3	1
V	0.163	0.97
VI	0.029	0.912
VII	0	0.82

The entries show how many times the parametric CI is shorter than the nonparametric CI.

Let us now look at the correctness of the methods, i.e. the closeness of the given methods to the exact method. It is expected that the methods based on t will not be close to the exact method. Conversely Methods V and VI are based on the exact method and are thus expected to be closed to it. Basically, it can be seen that the parametric bootstrap is close to the exact method, which according to Corollary 1 in Paper I is as expected. The following discusses how much the nonparametric and parametric CI are affected by the kurtosis.

Table 5 shows the Euclidean distance of CI of the given methods to the exact method. The entries in this table are the result of Monte Carlo simulations of the bootstrap with B=2000 resamples. It can be seen that Method V is more acceptable in comparison with other methods, as it expresses the standard error of the parametric bootstrap is close to the exact method. Method III does not improve anything as expected. With Method IV, The parametric and nonparametric method for the first data sets are closed against the second data sets, as is obvious because when $\widehat{K} < 3$, t^* increases, since the CI based on t is shorter than the exact method, the difference between CI and the exact method decreases. However, when $\widehat{K} > 3$, t^* becomes smaller and therefore the difference between CI and the exact method increases. As can be seen, for the second data set they are different. Method VI can be expected to have good results but not much is gained. In the case of method VII, which is suggested by many authors, there is much difference between the non-parametric and parametric method for the first data set. This is because although the $E(S^{2*}) = E(S^{2\#})$ but $S^{2\#}$ is more skewed than S^{2*} regardless of the kurtosis. This is obvious because $S^{2\#}$ comes from the limited range and therefore $\widehat{z}_0$ for the parametric bootstrap is higher than that for the nonparametric bootstrap. When $\widehat{K} < 3$ then the spread of S^{2*} is shorter than $S^{2\#}$ and also $\widehat{z}_0^*$ is smaller than $\widehat{z}_0^{\#}$. This makes the CI shorter against $\widehat{K} > 3$, which makes the spread of S^{2*} wider than that of $S^{2\#}$. The fact that $\widehat{z}_0$ of the parametric bootstrap is higher than that of the nonparametric makes the nonparametric bootstrap close to the parametric bootstrap, as can be seen in Table 5. Since $\widehat{K} < 3$ is more likely to happen because $E(\widehat{K}) = 3 - \frac{6}{n+1}$ hence it can be seen that the BC_α of variance, which is often done by the nonparametric approach, is not appropriate.

Table 5: Comparison of the correctness of the CI at 95%

data	method	III	IV	V	VI	VII
x	nonparametric	240.12	16.82	9.11	72.51	44.66
	parametric	251.41	6.08	4.29	60.69	11.82
y	nonparametric	187.87	73.84	10.47	39.741	9.830
	parametric	180.43	24.38	3.02	43.81	5.501

2.3 Paper II

The second part of this work is devoted to goodness-of-fit tests. Paper II discusses a bootstrap version of the Jarque-Bera test, which can be generalized to other distributions other than the normal. Its generality and its statistical power makes it desirable. Its statistics is:

$$\mathcal{D} = \begin{bmatrix} \overline{X} - \mu & S^2 - \sigma^2 & \widehat{\gamma} - \gamma & \widehat{K} - K \end{bmatrix} W \begin{bmatrix} \overline{X} - \mu \\ S^2 - \sigma^2 \\ \widehat{\gamma} - \gamma \\ \widehat{K} - K \end{bmatrix}. \qquad (35)$$

Let $W = I$ because the reciprocal of variance of skewness and kurtosis for other distributions are not known. Since the aim of the bootstrap method is to estimate the standard error, it can be used for the parameters concerned. Therefore it can be used to estimate W, i.e. the diagonal matrix, the elements of which are the reciprocal of variance of the parameters concerned. It is referred as $\mathcal{DW}$.

Paper II also presents a new method for goodness-of-fit tests. Let $X_1, \cdots, X_n \overset{iid}{\sim} F$. The aim is to study the hypothesis of normality, for some $\sigma > 0$ and μ:

$$H_0 : F(x) \overset{d}{=} \phi(\frac{x-\mu}{\sigma}), \tag{36}$$

where ϕ is the cumulative distribution function of the standardized normal distribution. The suggested method is explained as follows:

Consider U_1, U_2, U_3, U_4 and U_5 where $U_1 < U_2 < U_3 < U_4 < U_5$ or vice versa. The aim is to find new parameters that can be used to study the goodness-of-fit test. The reasoning is based on fact that if U_i can be used to find μ_1 then U_i^j can be used to find μ_j, where $\mu_j = E(X-\mu)^j$. The new parameters can be constructed using the following equations,

$$T_1U_1 + T_2U_2 + T_3U_3 + T_4U_4 + T_5U_5 = \mu_1, \tag{37}$$
$$T_1U_1^2 + T_2U_2^2 + T_3U_3^2 + T_4U_4^2 + T_5U_5^2 = \mu_2, \tag{38}$$
$$T_1U_1^3 + T_2U_2^3 + T_3U_3^3 + T_4U_4^3 + T_5U_5^3 = \mu_3, \tag{39}$$
$$T_1U_1^4 + T_2U_2^4 + T_3U_3^4 + T_4U_4^4 + T_5U_5^4 = \mu_4, \tag{40}$$
$$T_1 + T_2 + T_3 + T_4 + T_5 = C, \tag{41}$$

Assume that C is given, which is the summation of the coefficients and actually helps to control the coefficients that plays the role as the penalty. There is a solution for T_i because $T = U^{-1}\theta$:

$$U = \begin{pmatrix} U_1 & U_2 & U_3 & U_4 & U_5 \\ U_1^2 & U_2^2 & U_3^2 & U_4^2 & U_5^2 \\ U_1^3 & U_2^3 & U_3^3 & U_4^3 & U_5^3 \\ U_1^4 & U_2^4 & U_3^4 & U_4^4 & U_5^4 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \tag{42}$$

$$\theta = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ C \end{pmatrix}. \tag{43}$$

U is a vandermonde matrix, which is nonsingular. Hence there exists a unique values for T_i.

If the distribution is standard normal, then the estimation of moments must be close to $\mu_1 = 0$, $\mu_2 = 1$, $\mu_3 = 0$ and $\mu_4 = 3$, respectively. Therefore by substituting these values in equations (37)-(41), the solution is the values that are expected if the underlying distribution of observations is standard normal. Hence they are referred to as the theoretical values, $T_i, i = 1, ..., 5$. Based on the sample, the equations are as follows,

$$O_1u_1 + O_2u_2 + O_3u_3 + O_4u_4 + O_5u_5 = \widehat{\mu}_1, \tag{44}$$

$$O_1u_1^2 + O_2u_2^2 + O_3u_3^2 + O_4u_4^2 + O_5u_5^2 = \widehat{\mu}_2, \quad (45)$$
$$O_1u_1^3 + O_2u_2^3 + O_3u_3^3 + O_4u_4^3 + O_5u_5^3 = \widehat{\mu}_3, \quad (46)$$
$$O_1u_1^4 + O_2u_2^4 + O_3u_3^4 + O_4u_4^4 + O_5u_5^4 = \widehat{\mu}_4, \quad (47)$$
$$O_1 + O_2 + O_3 + O_4 + O_5 = C, \quad (48)$$

Where $\widehat{\mu}_i$ is the estimation of μ_i. If the chosen distribution is appropriate, then it is expected that O_i is close to T_i where O is:

$$O = U^{-1}\widehat{\theta} \quad (49)$$

Any suggested criterion should include the comparison of T_i and O_i, which are referred to as the theoretical and observed value. Here the discrepancy measure is the squared distance of the theoretical and observed value:

$$\mathcal{G} = \|T - O\| = (\widehat{\theta} - \theta)'(U^{-1})'U^{-1}(\widehat{\theta} - \theta). \quad (50)$$

It is obvious that $(U^{-1})'U^{-1}$ plays the role of weight in the JB test. The main question is concerned with the distribution of the suggested criterion for the bootstrap method used.

Two different bootstrap methods are used in the study of the given statistic; the parametric and semiparametric method. The parametric bootstrap is based on the assumed distribution, which can be used for resampling, while the semiparametric bootstrap directly combines the parametric and nonparametric bootstrap in the resampling setting. It gives appropriate p_i to observations to participate in the resampling, as discussed in Paper II.

3 Future work

This thesis explores variance and the bootstrap methods. Future work will focus on the variance in Bayesian bootstrap imputation which is discussed by Rubin (1987) and Little and Rubin (1987). Imputation is a common technique for handling incomplete observations by filing in the missing values with plausible values. The Bayesian bootstrap imputation is based on using resamples from the original observed sample. Kim (2002) studied the biasedness of the multiple- imputation variance when the object is the estimation of mean. Our aim is to study the Bayesian bootstrap multiple imputation when the object is the variance.

As it mentioned in Section 2.2.3 on bootstrap CI of variance, the accuracy of nonparametric BC is weak when the kurtosis is less than 3. It is obvious that $\hat{z}_0$ is somewhat biased and studies are needed on how to improve it.

The statistical tests of variance are of interest in a number of research areas. The comparison of variances test is discussed and reviewed by Boos and Brownie (2004). The bootstrap test of variance requires in-depth study. Our aim is to study the pure nonparametric version which is referred as exponentially titled is discussed in detail by Marazzi (2002) along with its semiparametric bootstrap which is discussed in Paper II.

References

An, L., Ahmed, S. E. (2008). Improving the performance of kurtosis estimators. Computational Statistics & Data Analysis. 52, 2669-2681.

Athreya, K. B., Lahiri, S. N. (2006). Measure theory and probability theory. Springer, New York.

Boos, D., Cavell, B. (2004). Comparing variances and other measures of dispersion. 9, 571-578.

Casella, G., Berger, R. L. (2002). Statistical inference, Second Edition. Duxbury, Pacific Grove, California.

Chiodi, M. (1995). Generation of pseudo random variates from a normal distribution of order P. Pubblicato su Statistica Applicata. 7, 4, 401-416.

Chissom, B. S. (1970). Interpretation of the kurtosis statistic. The American Statistician, 24, 4, 19-22.

Cojbasic, V., Tomovic, A. (2007). Nonparametric confidence intervals for population variance of one sample and the difference of variances of two samples. Computational Statistics & Data Analysis, 2007, 51, 12, 5562-5578.

Cramér. H. (1945). Mathematical Methods of Statistics. Almqvist & Wiksells, Uppsala.

D'Agostino, R. B. and Stephens, M. A. (1986). Goodness-of-fit techniques. Marcel Dekker, New York.

Darlington, R.B. (1970). Is kurtosis really peakedness? The American Statistician. 24, 19-22.

Davison, A. C., Hinkley, D. V. (1997). Bootstrap methods and their application. Cambridge University Press. Cambridge.

DiCiccio, T. J., Efron, B. (1996). Bootstrap confidence intervals. Statistical Science, 11,3, 189-228.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife Bootstrap Methods: The Annals of Statistics. 7, 1, 1-26.

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

Efron, B., Tibshirani, R. (1993). Introduction to the Bootstrap. Chapman & Hall, New York.

Everitt, B., Howell, D. (2005). Encyclopedia of Statistics in Behavioral Science. John Wiley & Sons, New Jersey.

Good, P. (2005). Permutation, Parametric and Bootstrap Tests of Hypotheses. Springer, New York.

Guttman, L. (1948). An Inequality for kurtosis. The Annals of Mathematical Statistics, 19, 2, 277-278.

Hall, P. (1992a). The Bootstrap and Edgeworth Expansion. Springer, New York.

Hall, P. (1992b). On the removal of skewness by transformation. J. R. Statist. Soc. B. 54, 1, 221-228.

Hall, P., Härdle, W., Kleinow, T., Schmidt, P. (2000). Semiparametric Bootstrap Approach to Hypothesis Tests and Confidence Intervals for the Hurst Coefficient. Statistical Inference for Stochastic Processes. 3, 3, 263-276(14)

Hildebrand, D. K. (1971). Kurtosis measures bimodality? The American Statistician, 25, 1, pp. 42-43.

Janssen, A., Pauls,T. (2003). How do bootstrap and permutation tests work? Annals of Statistics. 31, 768-806.

Joanes, D. N., Gill, C. A. (1998). Comparing measures of sample skewness and kurtosis. The Statistician. 47, 1, 183-189.

Kendall, M, G., Stuart. A. (1958). The advanced theory of statistics. Vol 1. Charles Griffin & Company Limited London.

Lee, S. S. (1994). Optimal choice between parametric and non-parametric bootstrap estimates. Math. Proc. comb. Phil. soc, 115, 335.

Lehmann, E. L. (1999). Elements of Large Sample Theory. Springer, New York.

Little, R. J. A., Rubin, D.B. (1987) Statistical Analysis with Missing Data. J. Wiley & Sons, New York.

Mackinnon, J. G. (2002). Bootstrap inference in econometrics. Canadian Journal of Economics, 35. 615-645.

Marazzi, A. (2002). Bootstrap tests for robust means of asymmetric distribution with unequal shapes. Computational Statistics & Data Analysis, 2002, 39, 503-528.

Moors, J. A. (1988). A Quantile alternative for kurtosis. The Statistician, 37, 1 , 25-32.

Pearson, K. (1905). Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson. A rejoinder, Biometrika 4, 169-212.

Polansky, A. (2008). Observed confidence levels. Theory and application. Chapman & Hall/CRC. New York.

Ostaszewski, K., Rempala, G. A. (2000). Parametric and Nonparametric bootstrap in Actuarial practice. www.actuarialfoundation.org/research_edu.

Rayner, J. C. W., Best, D. J. (1989). Smooth tests of Goodness of fit. Oxford University Press. New York.

Rubin, D.B. (1987). Multiple imputation for nonresponse in surveys. J. Wiley & Sons, New York.

Ruppert, D. (1987). What is kurtosis?: An Influence Function Approach. The American Statistician, 41, 1, 1-5.

Searle, D. T., Intarapanich, P. (1990). A note on an estimator for variance that utilizes the kurtosis. The American Statistician. 44, 4, 295-196.

Shao, J. (1999). Mathematical Statistics. Springer, New York.

Shao, J., Tu, D. (1996). The Jackknife and Bootstrap. Springer-Verlag, New York.

Tiku, M. L., Akkaya, A. D. (2004). Robust estimation and hypothesis testing. New Age International, New Delhi.

Hode, Jr., H.C. (2002). Testing for Normality, Marcel Dekker, New York.

Zhou, X. H., (2000). Nonparametric confidence interval S^2 for the one- and two samples problems. Biostatistics (2005), 6, 2, 187-200.

Acknowledgements

Did you hear about the politician who promised that, if he was elected, he'd make certain that everybody would get an above average income? And nobody laughed... (Aaron Levenstein). The boarder between statistics (the statistician who gives the number as a statistics) and lie (liar tells lies) is thinner than a hair. This thesis is dedicated to those who have helped make me a statistician and not liar.

Most especially I would like to thank my supervisors Prof. Dietrich von Rosen and Assoc. Prof. Silvelyn Zwanzig who provided generous support and consistent encourage throughout this work.

Also I warmly thank my friends and colleagues in the Dept. of Energy and Technology for their hospitality and friendly atmosphere especially Prof. Hans liljenström, Prof. Per-anders Hansson and Dr. Tomas Thierfelder. And to Berit, Majsan and Sven for their help. Among many friends I like to thank Razaw, Zhanna, Ida, Carina, Johannes, Ali and Mehdi whom I had a wonderful time.

Finally I thank SLU University for its financial support in carrying out this work.

On variance estimation and a goodness-of-fit test using the bootstrap method

Saeid Amiri

Abstract

This thesis deals with the study of variance estimation using the bootstrap method, including the problem of choosing between nonparametric and parametric bootstrap methods. Paper I compares the two approaches, determines which method is preferable and analyses the accuracy of the approximations. The underlying concept of parametric bootstrap is based on the assumption of correct choice of parametric distribution. Paper II therefore considers goodness-of-fit tests and presents a new test based on the bootstrap method.

Key words: Nonparametric Bootstrap, Parametric bootstrap, Goodness-of-fit test, Variance.

Distribution:
Swedish University of Agricultural Sciences
Department of Energy and Technology,
750 07 Uppsala, Sweden

Uppsala
ISSN, 1654-9406
ISBN, 978-91-86197-20-9