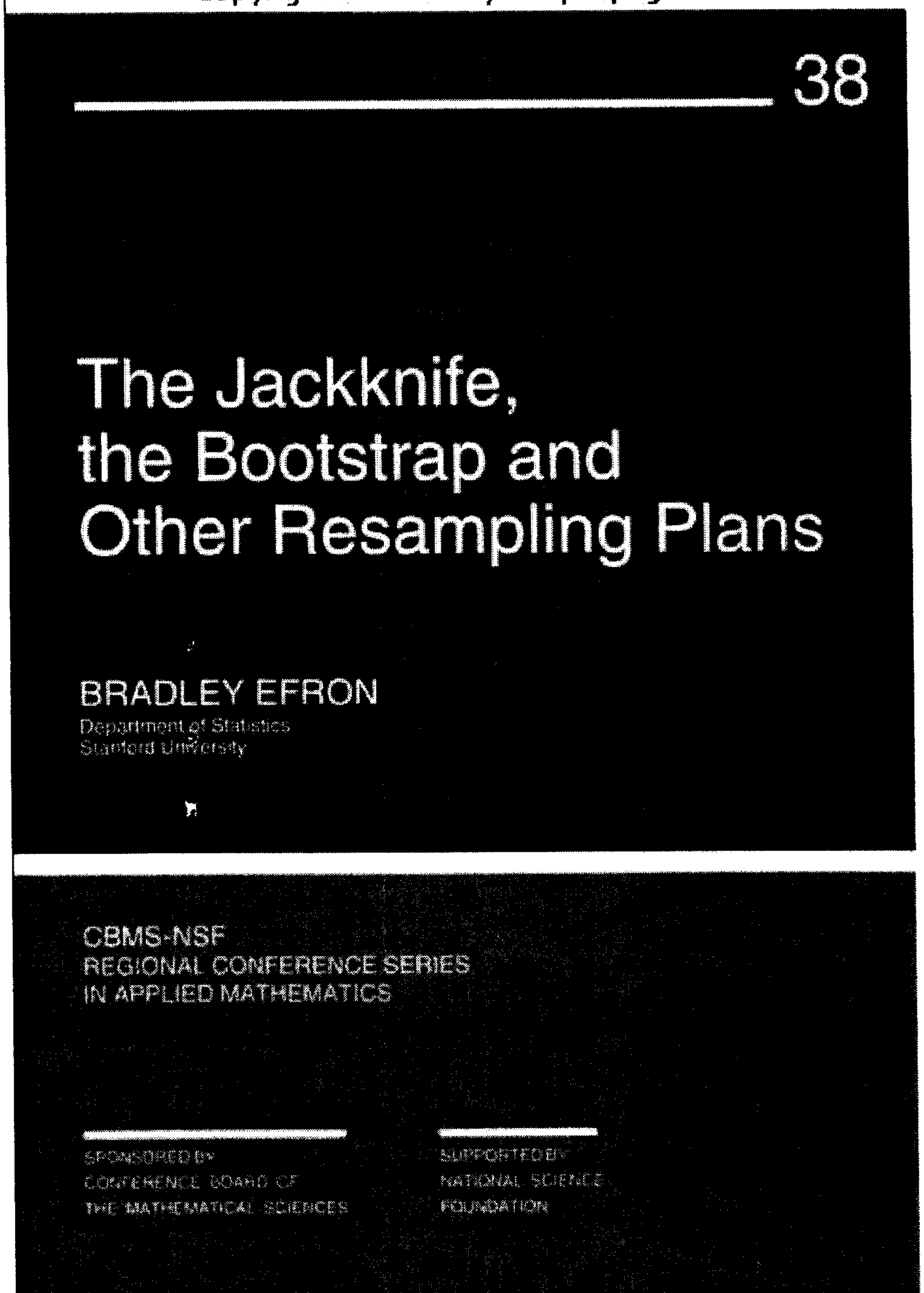
7

73771

Copyrighted material; sample page 1 of 7



Previous page   (+) Zoom in   Next page   See all 7 s

Back to book details   Look inside the book   Conti

Look for similar books by subject:

Browse for books in:

Contents

Previous page + Zoom in Next page See all 7

Back to book details Look inside the book Co

Look for similar books by subject:

CONTENTS

CHAPTER 1

Introduction

Our goal is to understand a collection of ideas concerning the nonparametric estimation of bias, variance and more general measures of error. Historically the subject begins with the Quenouille–Tukey jackknife, which is where we will begin also. In fact it would be more logical to begin with the bootstrap, which most clearly exposes the simple idea underlying all of these methods. (And in fact underlies many common parametric methods as well, such as Fisher's information theory for assigning a standard error to a maximum likelihood estimate.) *Good* simple ideas, of which the jackknife is a prime example, are our most precious intellectual commodity, so there is no need to apologize for the easy mathematical level. The statistical ideas run deep, sometimes over our head at the current level of understanding. Chapter 10, on nonparametric confidence intervals, is particularly speculative in nature.

Some material has been deliberately omitted for these notes. This includes most of the detailed work on the jackknife, especially the asymptotic theory. Miller (1974a) gives an excellent review of the subject.

What are the jackknife and the bootstrap? It is easiest to give a quick answer in terms of a problem where neither is necessary, that of estimating the standard deviation of a sample average. The data set consists of an independent and identically distributed (i.i.d.) sample of size n from an unknown probability distribution F on the real line.

$$X_1, X_2, \cdots, X_n \overset{\text{ind}}{\sim} F. \tag{1.1}$$

Having observed values $X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n$, we compute the sample average $\bar{x} = \sum_{i=1}^{n} x_i/n$, for use as an estimate of the expectation of F.

An interesting fact, and a crucial one for statistical applications, is that the data set provides more than the estimate $\bar{x}$. It also gives an estimate of the accuracy of $\bar{x}$, namely,

$$\hat{\sigma} = \left[\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{1/2}; \tag{1.2}$$

$\hat{\sigma}$ is the estimated standard deviation of $\bar{X}$, the root mean square error of estimation.

The trouble with formula (1.2) is that it doesn't, in any obvious way, extend to estimators other than $\bar{X}$, for example the sample median. The jackknife and

1

◀ Previous page + Zoom in Next page ▶ **See all 7**

Back to book details **Look inside the book** Co

Look for similar books by subject:

the bootstrap are two ways of making this extension. Let

$$\bar{x}_{(i)} = \frac{n\bar{x} - x_i}{n-1} = \frac{1}{n-1} \sum_{j \neq i} x_j, \tag{1.3}$$

the sample average of the data set deleting the ith point. Also, let $\bar{x}_{(\cdot)} = \sum_{i=1}^{n} \bar{x}_{(i)}/n$, the average of these deleted averages. Actually $\bar{x}_{(\cdot)} = \bar{x}$, but we need the dot notation below. The jackknife estimate of standard deviation is

$$\hat{\sigma}_{\text{JACK}} = \left(\frac{n-1}{n} \sum_{i=1}^{n} (\bar{x}_{(i)} - \bar{x}_{(\cdot)})^2 \right)^{1/2}. \tag{1.4}$$

The reader can verify that this is the same as (1.3). The advantage of (1.4) is that it can be generalized to any estimator $\hat{\theta} = \hat{\theta}(X_1, X_2, \cdots, X_n)$. The only change is to substitute $\hat{\theta}_{(i)} = \hat{\theta}(X_1, X_2, \cdots, X_{i-1}, X_{i+1}, \cdots, X_n)$ for $\bar{x}_{(i)}$ and $\hat{\theta}_{(\cdot)}$ for $\bar{x}_{(\cdot)}$.

The bootstrap generalizes (1.3) in an apparently different way. Let $\hat{F}$ be the empirical probability distribution of the data, putting probability mass $1/n$ on each x_i, and let $X_1^*, X_2^*, \cdots, X_n^*$ be an i.i.d. sample from $\hat{F}$.

$$X_1^*, X_2^*, \cdots, X_n^* \overset{\text{iid}}{\sim} \hat{F}. \tag{1.5}$$

In other words, the X_i^* are a random sample drawn *with replacement* from the observed values $x_1, x_2, \cdots, x_n$. Then $\bar{X}^* = \sum X_i^*/n$ has variance

$$\text{Var}_* \bar{X}^* = \frac{1}{n^2} \sum_{i=1}^{n} (x_i - \bar{x})^2, \tag{1.6}$$

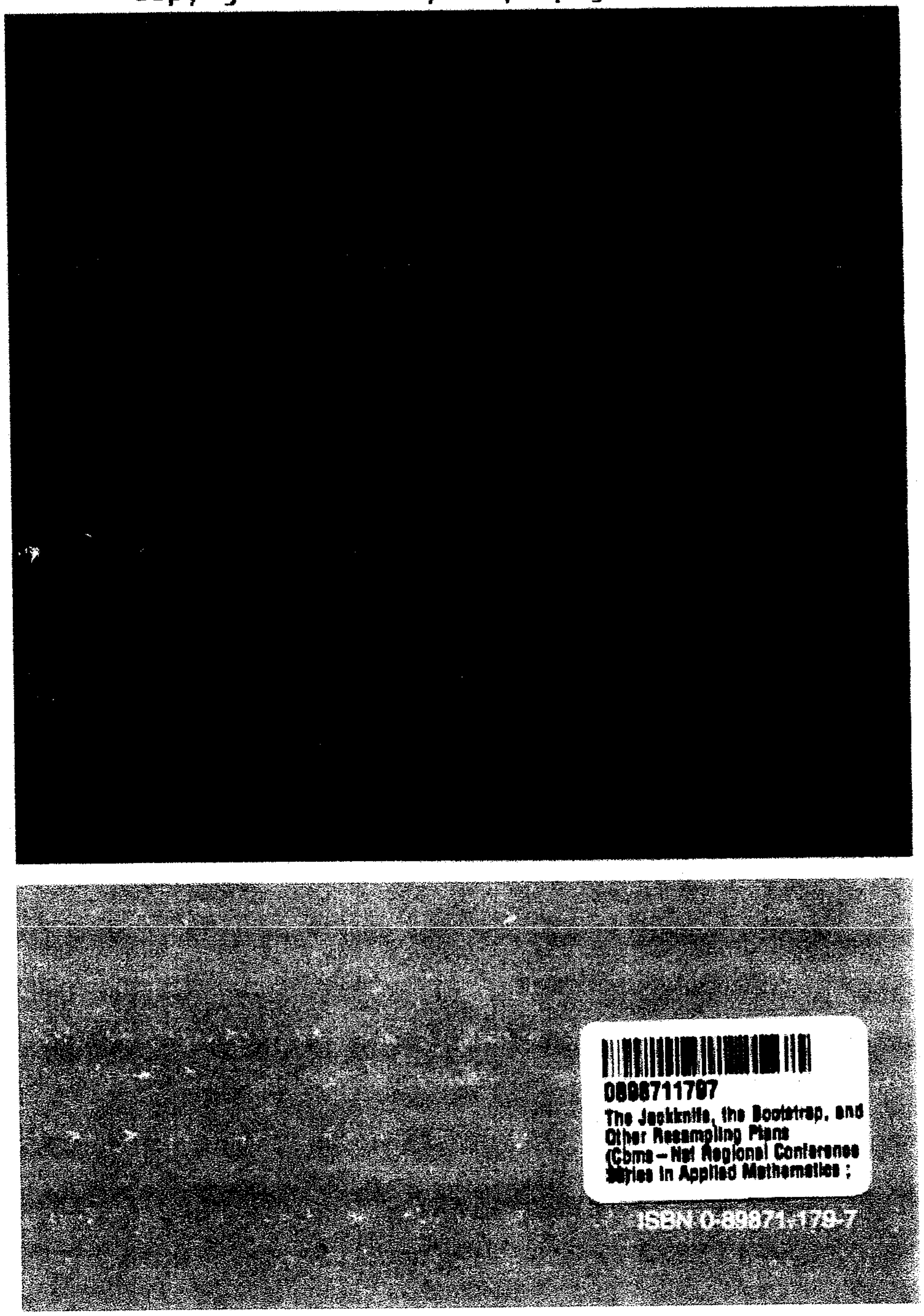
Var_* indicating variance under sampling scheme (1.5). The bootstrap estimate of standard deviation for an estimator $\hat{\theta}(X_1, X_2, \cdots, X_n)$ is

$$\hat{\sigma}_{\text{BOOT}} = [\text{Var}_* \hat{\theta}(X_1^*, X_2^*, \cdots, X_n^*)]^{1/2}. \tag{1.7}$$

Comparing (1.7) with (1.2), we see that $\hat{\sigma}_{\text{BOOT}} = [(n-1)/n]^{1/2}\hat{\sigma}$ for $\hat{\theta} = \bar{X}$. We could make $\hat{\sigma}_{\text{BOOT}}$ exactly equal to $\hat{\sigma}$, for $\hat{\theta} = \bar{X}$, simply by adjusting definition (1.7) with the factor $[n/(n-1)]^{1/2}$, but there turns out to be no advantage in doing so. A simple algorithm described in Chapter 4 allows the statistician to compute $\hat{\sigma}_{\text{BOOT}}$ no matter how complicated $\hat{\theta}$ may be.

Many other generalizations of (1.3) have been put forth. All such methods turn out to be closely related in theory, but not necessarily in their numerical consequences for a specific data set. We shall be investigating the theoretical and practical aspects of this collection of methods. Other measures of statistical error besides standard deviation are also considered: bias, prediction error and confidence intervals.

From a traditional point of view, all of the methods discussed here are prodigious computational spendthrifts. We blithely ask the reader to consider techniques which require the usual statistical calculations to be multiplied a thousand times over. None of this would have been feasible twenty-five years



Previous page   + Zoom in   Next page   See all 7 sa

Back to book details   **Look inside the book**   Conti

Look for similar books by subject:

Browse for books in: