Draft_1

April 17, 2022

0.1 Introduction

Bootstrapping was developed in the 20th century by Bradley Efron, an American statistician(Efron, 1979). This method assumes that the sample has the same relationship to the population asit has to an empirical distribution that is created by resampling with replacement from the original distribution N samples of the same size as the original sample. By creating this empirical distribution and comparing the sample statistic to it, the researcher can gauge the accuracy of the inferences on the population parameter. Over the last four decades, bootstrapping has become widely used and has been expanded to include various types of bootstrapping, such as parametric.

0.2 Idea of boostrapping (Theoretical)

Let $x = (x_1, x_2, ..., x_n)$ be an i.i.d. sample f size n from a population $F(x)$, and let $\hat{\theta} = T(X)$ be a sample statistic (such as $\bar{x}$ or any sort of estimator).Our goal is to determine the sampling distribution of $\hat{\theta}$.

The steps of the bootstrap are as follows:
1 Draw a sample x^* with replacement of size k from x. This is called a resample.
2 Calculate $\hat{\theta}^* = T(X^*)$.
3 Repeat steps 1 – 2 for m times, obtaining a series of outputs $\hat{\theta}_1^*, \hat{\theta}_2^*, ..., \hat{\theta}_m^*$

When you are done, the vector $(\hat{\theta}_1^*, \hat{\theta}_2^*, ..., \hat{\theta}_m^*)$ will be like simulated values of $\hat{\theta}$. Thus, we have obtained an approximation of the sampling distribution of $\hat{\theta}$.

0.3 Application (Pratical)

0.3.1 intro our data

0.3.2 Exploratory Data Analysis

0.3.3 boostrapping

0.3.4 goal

bias

variance

confidence interval)

0.4 pros and cons