STA 3180 Statistical Modelling: Bootstrapping

Bootstrapping for STA 3180 Statistical Modelling

Definition

Bootstrapping is a statistical technique used to estimate the sampling distribution of an estimator by resampling with replacement from the original sample. It is a non-parametric approach to statistical inference that can be used to assess the accuracy of a given estimator.

Key Concepts

1. Resampling with replacement: This is a key concept in bootstrapping. It means that when a sample is drawn, it is replaced back into the population before the next sample is drawn. This allows for the same data points to be sampled multiple times, which is necessary for bootstrapping.

2. Sampling Distribution: This is the distribution of the values of a statistic (such as the mean or median) calculated from all possible samples of the same size from a population. The sampling distribution of an estimator is used to assess the accuracy of the estimator.

3. Estimator: An estimator is a statistic used to estimate a population parameter. Examples of estimators include the sample mean and sample variance.

Coding Example

```
Start of Code
# Bootstrapping example in R
# Generate a sample of 10 observations from a normal distribution
set.seed(123)
sample <- rnorm(10)
# Create a function to calculate the sample mean
mean_func <- function(x){
    mean(x)
}
# Use the bootstrap function to generate 1000 bootstrap samples
boot_samples <- boot(sample, mean_func, R=1000)
# Calculate the 95% confidence interval for the mean
boot.ci(boot_samples, type="bca")
End of Code
```

Practice Multiple Choice Questions

Q1. What is the purpose of bootstrapping?

A. To estimate the sampling distribution of an estimator

B. To estimate the population parameter

C. To estimate the standard deviation of a sample

D. To estimate the mean of a sample

Answer: A. To estimate the sampling distribution of an estimator