

1. Consider the law school data provided on the class website. Write a matlab program to generate $n = 15$ sample pairs (LSAT and GPA) from the population of $N = 82$. Having obtained this sample write a matlab program to compute the following:

- (a) Estimate the means (μ_x and μ_y) and correlation coefficient from the sampled data.
- (b) Estimate the standard errors of each of these estimators for $B = 25, 50, 100, 200, 500, 1000$, and 2000 . Either tabulate them or plot curves.
- (c) Plot a histogram of the bootstrap replicates of (estimated) correlation coefficients for the case $B = 2000$.

To compare, generate $n = 15$ random samples from the original population and compute the estimated correlation coefficient. Generate 2000 such values by repeated sampling of the population to form a second histogram. Compare the two histograms.

2. Consider an artificial data set consisting of the 10 numbers:

1 2 3.5 4 7 7.3 8.6 12.4 13.8 18.1 .

Let $\hat{\theta}$ be the 25% trimmed mean, computed by deleting the smallest two numbers and the largest two numbers, and then taking the average of the remaining six numbers.

- (a) Calculate \hat{se}_B for $B = 25, 100, 200, 500, 1000, 2000$.
 - (b) Repeat part (a) using ten different random number seeds and assess the variability in the estimates. How large should we take B to provide reasonable accuracy?
3. Generate n values from a standard normal random variable. Let $\theta = \text{mean}(X)$, and the estimator be the median of the observed values. Use bootstrap method to estimate the bias of this estimator for $n = 10, 20$ and 100 .