Applied Statistics - Exercise 9

1. Basic bootstrapping (Theory)

We generate a single bootstrap dataset $x_1^*, \ldots, x_n^*$ from the empirical distribution function of

$$1, 4, 6, 7, 8, 11, 15, 19$$

a. What is the probability that the bootstrap sample mean is equal to 19?
b. What is the probability that the minimum of the bootstrap dataset is 1?
c. What is the probability that in the bootstrap sample exactly two elements are ≤ 6 and all the other are ≥ 15?

2. Bright stars (R)

Consider the `brightness` dataset from the `UsingR` package, which collects the brightness of 966 stars. Using empirical bootstrap, estimate the probability

$$Pr[|\bar{X}_n - \mu| > 0.1]$$

where μ is the *true* mean of the distribution. *Hint*: as we did in class, you will need to approximate this probability by replacing the sample mean with the bootstrapped mean, and μ with the sample mean.

3. Parametric bootstrap (R)

The dataset `arctic.oscillations` (in package UsingR) contains a time series from January to June 2002 of sea-level pressure measurement at the arctic, relative to some base line. Use parametric bootstrap to judge whether it is safe to assume that the measurements are samples from normal distribution or not. *Hint*: use parametric bootstrap in combination with the Kolmogorov-Smirnov distance, as we did in class.

4. Unbiased estimators (Theory)

Consider a random sample $X_1, \ldots, X_n$ from a uniform distribution in the interval $[-\theta, \theta]$, where θ is an unknwon parameter. You are interested in estimating the values of θ.

a. Show that
$$\hat{\Theta} = \frac{2}{n}(|X_1| + |X_2| + \cdots + |X_n|)$$
is an unbiased estimator for θ. *Hint*: you may need to use the *change of variable* formula (cfr. Chapter 7 of the book).

b. Consider instead the problem of estimating θ^2. Show that
$$T = \frac{3}{n}(X_1^2 + X_2^2 + \cdots + X_n^2)$$
is an unbiased estimator for θ^2

c. Is $\sqrt{T}$ an unbiased estimator for θ? If not, discuss whether it has positive or negative bias.