Bootstrap and Resampling Methods: Assignment 1.
Due date: 15th of April

Exercise 1

The goal of this exercise is to apply the theory for the bootstrap that we have presented in class.

Let X be a real-valued random variable ($X \in \mathbb{R}$) with density f_X. The Fisher-Pearson moment coefficient of skewness of f_X is defined as

$$sk := \mathbb{E}\left\{\left(\frac{X-\mu}{\sigma}\right)^3\right\},$$

where μ and σ are the expectation and standard deviation of X.
If f_X is symmetric about μ then $sk = 0$.
Given an observed iid sample $\{X_i\}_{i=1}^n$, we want to test if f_X is symmetric about μ, so we want to test the hypotheses

$$\mathcal{H}_0 : sk = 0 \text{ Versus } \mathcal{H}_1 : sk \neq 0\,.$$

(We are testing an implication of symmetry). We assume that X has finite 6th order moments.

1. Build a statistic, say S_n, that is asymptotically normal under $\mathcal{H}_0$, but not necessarily asymptotically pivotal (i.e. show that such a statistic converges in distribution to $\mathcal{N}(0, \Psi)$, where $\Psi > 0$). You do not need to specify the expression of Ψ.

2. Imagine that we are too "lazy" to compute Ψ, so we want to implement a bootstrap test. Propose a bootstrap test based on the *pairwise resampling scheme*.

3. Prove that the bootstrapped statistic estimates consistently the distribution of S_n under $\mathcal{H}_0$ (Hint: here you should just use the theorems we have presented in class)

Exercise 2

The goal of this exercise is to see the limitations of the theory we have presented in class and to understand how we can overcome these limits. To this end, we study the behavior of the bootstrap for the square of a sample mean. This case is a simplified version of more complex cases encountered in statistics.

We assume to observe an iid sample $\{X_i\}_{i=1}^n$ and we want to test

$$\mathcal{H}_0 : \mu = 3 \text{ Versus } \mathcal{H}_1 : \mu \neq 3\,,$$

where $\mu = \mathbb{E}X$. To this end, we use the statistic

$$S_{n,1} := \sqrt{n}\left[\left(\overline{X}\right)^2 - \mu_0^2\right] \quad \text{with } \mu_0 = 3$$

and $\overline{X}$ denoting the sample mean of X.

1. Show that $S_{n,1}$ is asymptotically normal under $\mathcal{H}_0$; propose a bootstrap version of $S_{n,1}$, say $S^*_{n,1}$; show that the bootstrap distribution of $S^*_{n,1}$ estimates consistently the distribution of $S_{n,1}$ under $\mathcal{H}_0$.

From now henceforth, we modify the null hypothesis and we want to test

$$\mathcal{H}_0 : \mu = 0 \text{ Versus } \mathcal{H}_1 : \mu \neq 0$$

by using the statistic

$$S_{n,2} = \sqrt{n}\left[\left(\overline{X}\right)^2 - \mu_0^2\right] \text{ with } \mu_0 = 0\,.$$

2. What is the asymptotic distribution of $S_{n,2}$ under $\mathcal{H}_0$ $(\mu = 0)$ and what happens to its bootstrap counterpart $S^*_{n,2}$? Can they be used for testing $\mu = 0$?

Consider now the statistic

$$S_{n,3} := \left[\sqrt{n}(\overline{X} - \mu_0)\right]^2 \text{ with } \mu_0 = 0\,.$$

3. Obtain the asymptotic distribution of $S_{n,3}$ under $\mathcal{H}_0\,(\mu = 0)$
4. Propose a bootstrap counterpart $S^*_{n,3}$ of $S_{n,3}$. Show that the bootstrap distribution of $S^*_{n,3}$ estimates consistently the distribution of $S_{n,3}$ under $\mathcal{H}_0 : \mu = 0$

Exercise 3

Imagine we observe an iid sample $\{X_i\}_{i=1}^n$ and we want to test the null hypothesis about the population mean $\mathbb{E}X$

$$\mathcal{H}_0 \,:\, \mathbb{E}X = 0$$

We can use the statistic

$$S_n := \sqrt{n}\,\overline{X}\,.$$

To bootstrap the above statistic, we can use a ***weighted bootstrap*** scheme. This is similar to the *wild bootstrap* for regressions but has wider applications, as it does not need a regression structure. The **weighted bootstrap** version of S_n is

$$S^*_n := \sqrt{n}\left[\,\overline{vX} - \overline{X}\,\right]\,,$$

where $\overline{vX} := (1/n)\sum_{i=1}^n v_i X_i$ and $\{v_i\}_{i=1}^n$ is a sequence of iid bootstrap weights independent from the sample data and drawn from a distribution set up by the researcher. Now, put yourself at the place of such a researcher who must decide such a distribution.

1. What is the mean and variance that the weights $\{v_i\}_{i=1}^n$ must have for the weighted bootstrap to be consistent? In other words, you should decide the mean and variance that v_i should have so that the *weighted bootstrap distribution of* S^*_n estimates consistently the *distribution of* S_n *under* $\mathcal{H}_0$. Please, motivate your answer.

   (Hint: Use Proposition 4.2.1 at page 56 of Chapter 4, i.e. the *the Conditional Multiplier Central Limit Theorem.*)