

# Homework 2

Simulation and Performance Evaluation – University of Trento

**DEADLINE: April 27, 2023**

---

You can solve the following assignments using any programming language. In doing so, try to implement the formulas explained in class. You can use built-in functions in your programming language to draw random variates, compute means and standard deviations, as well as for any utility purposes (e.g., for managing data, sorting, plotting, printing messages etc).

## Exercise 1

Load the data from the CSV file `data_ex1.csv`.

1. Compute the coefficient of variation (CoV) for the data, Jain's fairness index and the Lorenz curve gap
2. Discuss the source of the difference between the Lorenz curve gap and Jain's fairness index.
3. Implement the bootstrap algorithm seen in class
4. Use bootstrap to compute 95% and 99% confidence intervals for:
  - (a) the mean of the data;
  - (b) the standard deviation of the data,
  - (c) the Lorenz curve gap;
  - (d) Jain's fairness index;
5. Compare the obtained confidence interval for the mean with the one obtained via the asymptotic formulas and discuss the results.

## Exercise 2

Consider the following “weird” probability density function:

$$f(x) = \frac{1}{A} e^{-|x|/4} (\sin(|x|) + 1), \quad -\frac{5\pi}{2} \leq x \leq \frac{5\pi}{2},$$

where  $A = 8.69336125$  is a normalization factor such that  $\int_{-5\pi/2}^{5\pi/2} f(t) dt = 1$ , so  $f(x)$  is in fact a PDF.  
(Question: do you really need to know the value of  $A$ ?)

1. Employ rejection sampling to draw a large number of samples from the above PDF.
2. Plot the resulting empirical PDF (e.g., through a histogram) and compare it against the theoretical PDF. Make sure you draw a sufficiently large number of samples, so that the histogram convincingly fits the theoretical PDF.
3. Take 1000 draws from the above distribution. Apply the bootstrap procedure of Exercise 1 to compute a 99% confidence interval for the mean and standard deviation of the dataset you drew.

## Exercise 3

Load the iid data samples from the CSV file `data_ex3.csv` and provide the value of all quartiles, and of the 10th and 90th percentiles of the data, along with a 95% confidence interval for all of them.