

# Statistical Methods for Data Science

DATA7202

Semester 2, 2020

Lab 6

## Objectives

On completion of this laboratory session you should be able to understand and implement the bootstrap method and perform discrete event simulation.

1. Consider the nerve data (nerve.csv) from Cox, D. and Lewis, P. (1966). The Statistical Analysis of Series of Events. Chapman & Hall, where the authors reported 799 waiting times between successive pulses along a nerve fiber.

- (a) Construct and plot the empirical cdf.
- (b) Provide an estimate the fraction of waiting times between 0.4 and 0.6 seconds.
- (c) The *skewness* is a measure of asymmetry of data (or probability distribution). The skewness is given by:

$$\mathbb{E} \left[ \left( \frac{x - \mu}{\sigma} \right)^3 \right]$$

- i. Provide (and calculate) a plug-in estimate of the skewness for the nerve data.
  - ii. Use the bootstrap method to (1000 samples), to provide the 95% confidence interval for the skewness.
2. Implement the Tandem Queue example for  $1/\lambda = 3, \mu_1 = 1, \mu_2 = 2$ . Plot a figure of the occupancy of both queues as a function of time. Obtain 95% confidence interval for the following.
    - Total time in the system.
    - Average waiting time in the first queue.
    - Average waiting time in the second queue.

Repeat the analysis for the following parameters;  $1/\lambda = 2.5$ ,  $1/\lambda = 2$ , and  $1/\lambda = 1$ .