

MATH205, Fall 2022
Midterm (Simulations), Section 051L
Wednesday, Oct 26, 2:30 pm

Instructions

Your solution needs to be submitted on Canvas (Assignment “Midterm–Simulations”), including:

- An R script of the work
- Screenshots of the results and the figures

The plots should have clear titles and all axes labeled.

Problem 1 (20pts)

Consider a continuous distribution with the following probability density function

$$f(x) = \begin{cases} \frac{3}{8}x^2, & \text{for } x \in [0, 2] \\ 0 & \text{otherwise} \end{cases}$$

- Simulate a dataset of $n = 5000$ random draws from this distribution.
- Compute the mean, the median and the standard deviation of the dataset
- Produce a histogram of the dataset.

Problem 2 (20pts)

The famous Anderson’s *iris* data set gives the measurements in centimeters of the variables sepal length and width and petal length and width, respectively, for 50 flowers from each of 3 species of iris. The species are *Iris setosa*, *versicolor*, and *virginica*.

- Create a histogram to describe the distribution of the *sepal width*
- Make a scatter plot to visualize the relationship between two variables: *petal length* and *petal width*
- The function

boxplot(y ~ grp)

allows us to create a comparative boxplot that describe a (continuous) variable y across different subgroups according to the (discrete) grouping variables grp .

Use this function to produce a comparative boxplot that represent the relationship between *sepal length* and *species*.