

Assignment 1

CS-UH-2218: Algorithmic Foundations of Data Science

Assignments are to be submitted in groups of two or three. Upload the solutions on NYU classes as one PDF file for theoretical assignments and separate source code files for each programming assignment. Submit only one copy per group. Clearly mention the participant names on each file you submit.

Mandatory Reading: Chapters 1 and 2 of the textbook [MMDS].

Optional Reading: Section 12.4 of the textbook [FoDS] (<https://www.cs.cornell.edu/jeh/book.pdf>) provides an excellent short review of basic probability theory.

Problem 1 (10 points).

Learn to use `pyplot` from https://matplotlib.org/users/pyplot_tutorial.html (or any other resource you prefer). Plot the function $f(x) = 1 - (1 - x^r)^b$ for $(r, b) \in \{(1, 1), (2, 2), (5, 5), (1, 5), (5, 1), (5, 20)\}$ and $x \in [0, 1]$. Plot the six functions obtained for the different values of r and b as subplots of the same figure.

Experiment with different values of r and b to determine a pair of (not too large) positive integers r and b so that the function $f(x)$ has the maximum slope around $x = 0.4$.

Problem 2 (10 points).

Solve Exercise 1.2.1 of [MMDS].

Problem 3 (10 points).

Solve Exercise 2.3.1 of [MMDS].