Randomized Algorithms (RA-MIRI): Assignment #2

1 Statement

In this programming assignment you will have to write programs to select the median of an array of n elements using:

1. Hoare's quickselect (see chapter 5 "Random Variables and Expectation (II)", pages 24–29, of the course slides).
2. Randomized Selection algorithm of Floy-Rivest (as given in the course slides, chapter 7 "Concentration around the Mean", pages 29–41).

Your programs will generate many arrays of size n containing each a random permutation and apply the selection algorithms to find the median of each array. In order to extract valid conclusions you should work with large arrays, with sizes ranging from 5000 to 50000 (or more). It is hence of the utmost importance that your programs are efficient, as you'll be working with very large inputs and repeat many times the experiments to draw statistically meaningful conclusions.

1.1 Quickselect

In the case of quickselect, your program should keep track of the total number of comparions between array elements and for each value of n obtain an average number of comparisons and a sample standard deviation of that number by applying the algorithm on a reasonably large number of random arrays of the given size. Study and plot the average value (and the deviation in a boxplot) as n grows and verify that the experimental results match (or not) the theoretical prediction (about $3.386n$ comparisons to find the median of an array of size n).

1.2 Randomized Selection

Measure the number of comparisons made to select the median and obtain an average number of comparisons—taking into account only those arrays in the sample where the algorithm produces a correct output. Record also the proportion of arrays in the sample for which the algorithm fails.

Compare the average number of comparisons made by Randomized Selection to that of quickselect. Verify empirically that the probability of failure decreases

as $n^{-1/4}$ at least—a loglog plot and quite large values of n will be useful at this point.

2 Instructions to deliver your work

Submit your work using the FIB-Racó. The deadline for submission is November 2nd, 2021 at 12:00. It must consist of a zip or tar file containing all your source code, auxiliary files and your report in PDF format. Include a README file that briefly describes the contents of the zip/tar file and gives instructions on how to produce the executable program(s) used and how to reproduce the experiments. The PDF file with your report must be called `YourLastName_YourFirstName-2.pdf`,

N.B. I encourage you to use LaTeX to prepare your report. For the plots you can use any of the multiple packages that LaTeX has (in particular, the bundle TikZ+PGF) or use independent software such as matplotlib and then include the images/PDF plots thus generated into your document.