Exercises for Lecture 9

Statistical Computing with R, 2022-23

Exercise 1

Use the code below to create two vectors:

```
set.seed(9)
x1 = rbinom(n = 20000, size = 5, prob = 0.3)
x2 = rgamma(n = 30000, shape = 2, rate = 1)
```

1. Use the `mean( )` function to compute the mean of `x1` and of `x2`. Can you guess what the functions `rbinom` and `rgamma` in the chunk above do?
2. Write a function that computes the mean of a vector x by: 1) computing the sum of x through a `for` loop and (2) by dividing the sum thus obtained by the length of x
3. Use the `benchmark` function to compare the execution time of `mean( )` to that of the function created at point (2). Do it both using `x1` and `x2`. Set the number of replicates equal to 500. Which of the two solutions is faster, and how much faster is it?

Exercise 2

Use the code below to create a "large" matrix:

```
set.seed(9)
n = 2000; p = 500
m1 = matrix( rnorm(n*p, mean = 4.7, sd = 0.5), ncol = p )
```

In this exercise, we will compare the performance of different ways to compute the mean of each column in `m1`:

a. using the `apply( )` function;
b. using the `colMeans( )` function;
c. using a `for` loop where the vector of outputs is preallocated;
d. using a `for` loop where the vector of outputs is not preallocated, but it is instead augmented at each iteration.

1. Use the `benchmark` function to compare the performance of the 4 alternatives mentioned above. Use at least 100 replications;
2. Which solution is the fastest? And the slowest?

Exercise 3

An apartment building consists of 127 flats. Its parking lot has the capacity to host 44 cars. If the probability that the inhabitants of a randomly picked flat own a car is 38%, what is the probability that there won't be enough parking spaces for all the cars?

Exercise 4

Let $X \sim Poi(\lambda = 6)$ and $Y \sim Gamma(\alpha = 3, \beta = 2)$. Use `R` to compute the following quantities:

1. $P(X = 7)$
2. $P(Y = 3)$
3. $P(2 < X < 5)$
4. $P(1 < Y < 3)$
5. $F_X(5)$
6. $F_X(3) + F_Y(10)$

Exercise 5

Let $X \sim N(\mu = 3, \sigma = 1.4)$ and $Y \sim Beta(\alpha = 2, \beta = 2)$. Define

$$Z = \frac{X}{Y}.$$

1. Draw 10000 random realization from $X \sim f_X(x)$, and 10000 random realizations from $Y \sim f_Y(y)$.
2. Compute $z = \frac{x}{y}$ for each of the 10000 random realizations from X and Y.
3. Use the random realizations created at point (2) to estimate $E(Z)$ and $Var(Z)$.