Statistics for Intelligent Systems
18.03.2021

Exercise 1. Create a new file named Noise.txt. Use the function choose.files() to load a vector of all file paths from your working directory to a variable Names. Use file.exists() on Names to confirm that all of the files are in the working directory. Remove Noise.txt using the proper command. Check the result of using file.exists() on Names.

Exercise 2. The file Books.csv contains the data of the relation between numbers of books read and the final spelling scores for third graders. Copy the file Books.csv to Books2.csv and append Books.csv to the new file. Load Books2.csv to a variable B. Check the content of B. Correct this data frame, so it has numerical values (apart of the variables names).

Exercise 3. Load the file Books.csv to a variable B. Write B to the file B3.csv using the wirte.csv() command. Try loading B3.csv using read.csv2(). What is the result?

Exercise 4. The file Life.csv contains data of US life expectancy sorted by decades. Load the file to a variable x1 and write it to a file L1.csv using write.csv2() without adding the parameter row.names=FALSE. Load the file L1.csv to x1, and repeat the writing process. Now, load Life.csv to a variable x2 and write it to a file L2.csv using write.csv2() and add the parameter row.names=FALSE. Load the file L2.csv to x2, and repeat the writing process. Compare the data in L1.csv and L2.csv.

Exercise 5. Append L1.csv to L2.csv. Try to load the result using read.csv2() without changing the fill parameter. Now load the file with the parameter fill = TRUE. Check the result.

Exercise 6. Check the names of the random number generators set on your session. Display the current seed. Change your generators to "Super-Duper" and "Box-Muller" for uniform and normal distribution respectively. Display the seed of the new generator.

Exercise 7. Use set.seed(5) to create a new random seed. Create a vector of 7 random numbers take from the normal distribution with mean 2 and standard deviation 1, then create a second vector with the same parameters. Use set.seed(5) and create another vector with the same parameters. Now set.seed(6) and create another vector with the same parameters. Compare the results.

Exercise 8. Use the proper function with prefix p to find the probability of a random value being less then -1, 0.5, 2, 5 for the following distributions: uniform distribution with min=0, max=3; uniform distribution with min=-3 and max =6; normal distribution with mean = 0 and standard deviation = 4; chi-square distribution with 5 degrees of freedom.

Exercise 9. Use the proper prefixes p and d to find the probability of having exactly 4 successes in a binomial test of size 10 and probability of success 0.6, having exactly 1 success, having at most 4 successes, and having at least 9 successes.

Exercise 10. Compare the values of every tenth quartile, using a vector of values from 0.1 to 0.9 and the prefix q for the following distributions: normal with mean 3 and standard deviation 2; uniform in the range from -3 to 3; t-Student for 6 degrees of freedom.