

Directions

Each step 1pt - 6pts total.

Data frame in Janka.xlsx files contain two meaningful attributes. Please write the R program that reads file janka.csv and then computes:

1. mean, variance and standard deviation for every attribute
2. covariance between attributes (i.e. covariance matrix),
3. histograms for each attribute assuming samples are drawn from probability distribution
4. Experimental CDF for each attribute and plot it
5. quintiles for each attribute with a step 0.1
6. quantile-to-quantile plot between attributes

Plots should be saved to files named as a plot itself (e.g. 'histogram-'your last name'.jpg')

Assume that you are reading files from and writing files to R's working directory.

The program should be submitted as saved in the file Janka-Vizual-'your last name'.R

Notes:

- To read .xlsx file into R you may need to install xlsx package. Then use the following operators:
 - `require(xlsx)`
 - `data = read.xlsx2("janka.xlsx", 1)`where data is a variable that will contain imported data as data.frame type. However, if you do it literally as described above then all your attributes will default to being factors while you need numerical values. So read carefully description of read.xlsx2 and use it properly.
- Experimental CDF is computed by `ecdf` built-in function/object. You can use `plot` method on it to plot eCDFs. Use `?` function to learn about `ecdf` and its methods.
- There are 2 methods of saving plots in .jpg files:
 1. Saving from R-studio. Click on export in plot pane and save
 2. Use `jpeg` operator, e.g. `jpeg("ecdf-density.jpg")`. If you do so then after saving the plot you need to turn the device invoked for the plot off by including `dev.off()` operator.

Note: All functions are described in the programs that we had in class you just need to modify them