KEN4258: Computational Statistics

Homework Assignment 2 (graded)

Christof Seiler

Posted on March 14 at 18:00 / Due on March 22 at 20:00

Logistics

The final submission in the form of a report should be one single pdf that you can upload to Canvas. One submission per group. The group can have a maximum of 5 members. Keep the report short. The main text should not exceed three pages. You can add an appendix to mention details that you find important. Otherwise, the structure of the report is up to you. You might want to add a link to your GitHub repository with your code (in R or Python).

Randomization Test [100 points]

1) Reproduce Figure 1 from (Candès et al. 2018). [20 points]
   Use this version of the paper: https://arxiv.org/abs/1610.02351.
2) What is the problem that Figure 1 tries to illustrate? [20 points]
3) Propose a solution to address the problem. You can implement conditional randomization test (Candès et al. 2018), another existing method, or come up with your own idea. [20 points]
4) Go back to Figure 1 and show that your solution fixes the problem. [20 points]
5) Find a real dataset and apply your method. [20 points]

References

Candès, Emmanuel, Yingying Fan, Lucas Janson, and Jinchi Lv. 2018. "Panning for Gold: 'Model-X' Knockoffs for High Dimensional Controlled Variable Selection." *Journal of the Royal Statistical Society Series B* 80 (3): 551–77.