Homework 2

Zhaoxia Yu

2025-04-21

Due on Monday, April 28 2025. R Code should be included as appendices.

- Problem 1: Choose a $3-by-3$ covariance matrix with non-zero covariances (the off-diagonal elements should not be 0). Also a choose a sample size n (e.g., n=100, 500, 1000, etc). Using the covariance matrix you chose, simulate 1,000 data sets from a trivariate normal distribution.

  0. Hints:

  - Hint 1: the R library MASS provides a function to generate a random sample from a multivariate normal distribution.
  - Hint 2: Make sure that the covariance matrix you choose is positive definite. You can compute the eigenvalues by the "eigen" function in R and and check whether all the eigenvalues are positive.

  1. Try to make sense of the covariance matrix by examining the pairwise scatter plots using the data you simulate.
  2. During the simulation, you will generate 1,000 Wishart distributed random matrices. Calculate the trace for each of them. Explain what distribution the traces should follow and examine their histogram.

- Problem 2: Find a good data example to conduct a two-sample Hotelling's T^2 test. Do not use the data example discussed in this course. Please (1) include visualizations as exploratory methods and (2) make conclusion in the context of the data example.