

Exercise 5

Problem 1: Frequency Analysis

The data set SCIENCEDOCTORATES.txt contains the number of doctors graduated from different fields of science. The data is from USA between the years 1960-1975.

Find the attraction repulsion matrix and interpret the results.

Problem 2: Covariance Matrix

Let x be a p -variate continuous random variable. Show that $\text{Cov}[x]$ is positive semidefinite.

Homework Assignment 5: Frequency Analysis

The data SMOKING1.txt contains a 2-dimensional frequency table, where the employees of a firm have been categorized according to their position (5 categories: SM = Senior Managers, JM = Junior Managers, SE = Senior Employees, JE = Junior Employees, SC = Secretaries). Furthermore, the smoking of the employees have 4 categories (None, Light, Medium, Heavy).

- a) Form the theoretical frequencies under independence.
- b) Calculate the attraction repulsion matrix.
- c) Is smoking more frequent among a specific group in this company?
Hint: Think about the complements of the different categories.