Problem Sheet 2

Mean Vectors & Covariance Matrices

Your Name (Reg No)

Consider the data set crossSecData(You may download the data set fom <https://github.com/davidrajdgl/MultivariteAnalysis/blob/master/DataSets/crossSecData.xlsx>. The details of the variables are given below.

* Id : Patient Id
* Diabetes: Diabetes Present : 1; Obsent : 0
* Fat\_intake: The amount of fat intake
* Height : Height of the patient
* Weight : Weight of the patient
* SBP : Systollic Blood Pressure
* DBP : Diastolic blood pressure

**Important:** Dont forget to save the dataset in your working directory. Set the working directory before you proceed to answer any questions below.

**Load the data set Here:**

library(readxl)  
data= read\_excel('crossSecData.xlsx')  
head(data) #Show only first 6 rows of the data

Data Set

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Id | Diabetes | Fat\_intake | Height | Weight | SBP | DBP |
| 1 | 0 | 25 | 162 | 57 | 134 | 90 |
| 2 | 0 | 150 | 151 | 102 | 110 | 70 |
| 3 | 0 | 34 | 150 | 56 | 124 | 70 |
| 4 | 0 | 24 | 150 | 59 | 130 | 80 |
| 5 | 0 | 50 | 156 | 53 | 126 | 80 |
| 6 | 0 | 50 | 180 | 63 | 120 | 80 |

**Question**  What is a Random vector?

**Question** What is the Random vector in this data set?

meanVec = colnames(data)

The mean vector of the data is (Id, Diabetes, Fat\_intake, Height, Weight, SBP, DBP ).

**Question 1:**  Generate a new data set consisits of Fat\_intak, Height, SBP and DBP only. Save the data set as ‘newCrossSec’.

*Answer:*

**Question 2:**  What is the Random Vector in the new data set?