DS201/DSL253: Statistical Programming

Assignment 05

27.02.2025

Instructions for Submission: You can submit your solution as a Jupyter Notebook/Matlab file with comments and discussions on the results obtained in each step.

1. Follow Standard Report Format: Include sections like Introduction, Data, Methodology, Results, Discussion, and Conclusion.
2. File Naming Convention: Adhere to the specified naming convention for each file you submit (e.g., RollNumber FirstName Asg1).
3. Refrain from using zip files. If necessary, submit multiple files.
4. Include comments in the code explaining the logic and any assumptions made.
5. Include References: Cite any external sources or references used in your assignment.
6. Code Quality: Ensure your code follows best practices, is well-organized, and avoid plagiarism as a plagiarism check will be conducted.
7. Be aware that late submissions are not permitted; ensure timely submission.
8. Coding can be done in any language.

1. Let X and Y have a bivariate normal distribution with parameters:

$$\mu_x = 3, \quad \mu_y = 1, \quad \sigma_x^2 = 16, \quad \sigma_y^2 = 25, \quad \rho_{xy} = \frac{3}{5}$$

where μ_x and μ_y represents mean of X and Y, σ_x^2 and σ_y^2 represents variance of X and Y and ρ_{xy} represents correlation coefficient between of X and Y.

Determine the following probabilities:

(a) $P(3 < Y < 8)$.

(b) $P(3 < Y < 8 \mid X = 7)$

(c) $P(-3 < X < 3)$

(d) $P(-3 < X < 3 \mid Y = -4)$

NOTE: Program shall be written to accommodate any given value of μ_x, μ_y, σ_x^2, σ_y^2, ρ_{xy} during the time of evaluation.

2.

(a) Write a program to generate P samples from multinomial random variable $X \in \mathbb{R}^n$, having multivariate normal distribution $N_n(\mu, \Sigma)$, where $\mu \in \mathbb{R}^n$ denote mean vector and $\Sigma \in \mathbb{R}^{n \times n}$ covariance matrix of X.

(b) Using P generated samples in part (a) get new samples using the following equation:

$$Y = (X - \mu)^T \Sigma^{-1} (X - \mu)$$

Observe the distribution of Y for different values of n and P.

(c) Compute the probability that Prob[$(x - \mu)^T \Sigma^{-1} (x - \mu) \leq c^2$] for given c.

3. The probability distributions of two different classes are known to follow a Normal distribution with the following parameters:

$$\mu_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \mu_2 = \begin{bmatrix} -2 \\ -3 \end{bmatrix}, \quad \Sigma_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 2 \end{bmatrix}, \quad \Sigma_2 = \begin{bmatrix} 2 & -0.3 \\ -0.3 & 1 \end{bmatrix}$$

Use Bayes' Theorem to perform classification for the datapoints given in the attached file "File_Datapoints.txt". Demonstrate the result using 2D diagram, illustrate classes with different colors.

NOTE: Program shall be written to accommodate any given value of $\mu_1, \mu_2, \Sigma_1, \Sigma_2$ during the time of evaluation.