MA515 - Foundations of Data Science

Exercise 0

August 6, 2025

---

1. Suppose marks obtained by 27 students in MA515 class in ascending order are as follows: 23, 43, 54, 56, 61, 62, 66, 68, 69, 69, 70, 71, 72, 77, 78, 79, 85, 87, 88, 89, 93, 95, 96, 98, 99, 99, 121.

   (i) Find the 30th, 60th and 90th percentile of the data.

   (ii) Also find out the quartiles $Q1, Q2$ and $Q3$.

   (iii) Identify the outliers in the data based on the $1.5IQR$ method. Also, if we use 1.2IQR method, identify the outliers.

2. Based on histogram plot, Q-Q plot, and Kolmogorov-Smirnov test identify if the above data is normally distributed. Use Python for this purpose.

3. Using python get the Box-Whisker plot of the given data. Identify the outliers.

4. Let (X,Y) is bivariate Gaussian such that $\mathbb{E}X = 1, \mathbb{E}Y = 5$, $\text{Var}(X) = 4$, $\text{Var}(Y) = 9$ and $\rho(X,Y) = 0.6$. Sketch the contour plots of the joint density function. Further calculate $\mathbb{E}(XY)$ and $\text{Var}(X+Y)$.

5. In the above bivariate Gaussian, write the expression for $\mathbb{P}(X+Y > 1)$.