IC-252 Lab Assignment 6

Q1) Let a continuous random variable X follows a Gaussian probability density function (PDF) f_X (x) with mean μ and variance σ^2, then

(a) Write a code to plot the PDF f_X (x) of the random variable X with mean μ = 0 and variance σ^2 = 1 (Use the PDF expression to plot).

(b) Write a code to generate data X, N = 100, 000 sample points, following a Normal PDF with mean μ = 0 and variance σ^2 = 1(Can use the numpy function **np.random.randn**)

(c) Redo the part 1b for decreasing value of the bin width of the histogram and state your observations. Can the two axes be representative of something? Hence, explain the significance of the obtained histograms in relation to the f_X (x) of part 1a.

Q2) Suppose that X is a continuous random variable whose probability distribution function is given by :

$$f(x) = C(4x - 2x^2)e^{-x} \quad ; 0 < x < 2$$
$$= 0 \quad ; otherwise$$

(a) Calculate the constant C using the fact that

$$\int_{-\infty}^{\infty} dx\, f(x) = 1$$

You might want to use **scipy.integrate** method for performing integration.

(b) Using **Matplotlib**, $f(x)$ plot as a function of x.

(c) Calculate the probability P(1 < X < 2).

Q3) The normal distribution is the most common type of distribution used in statistical analysis. In many fields which deal with large data you will notice that the data can be approximated to follow Normal distribution. Here, you are given a **marks.csv** file which contains the marks and grades of the students from a certain course at IIT Mandi. You are expected to do the following with the given data :

(a) Find the mean and standard deviation of the given marks and check if the marks distribution follows the 3-sigma rule.

(b) Assign the grades to the students based on the following rules :-

- Score >Average +1.5 SD -gives O
- Average +1.5 SD>Score >Average + SD -gives A
- Average + SD>Score >Average +0.5 SD -gives B
- Average + 0.5 SD>Score >Average -gives C
- Average>Score >Average-0.5 SD -gives D
- Average - 0.5 SD>Score >Average - SD - gives E
- Average - SD >Score >Average -1.5 SD -gives F

Check whether the originally assigned grades match the grades you assigned using this scheme.

(c) Check whether the given data follows Normal distribution or not. Plot a histogram of the marks data. Generate random numbers following normal distribution with mean and standard deviation that you calculated in part (a). Now plot these random numbers by superimposing the curve over the histogram and see if the distribution fits our data or not.

(d) Standardize the random variable, marks, by applying the Z-transformation, and calculate the mean and standard deviation of new variable.

(e) Repeat a-d for IC252 marks.

Q4) A Galton Board is a device used for statistical experiments and demonstrations. It consists of a vertical board with a series of evenly spaced pegs or nails arranged in a triangular pattern. At the top of the board, a large number of small balls or beads are dropped, and they bounce off the pegs as they descend, randomly falling into slots or compartments at the bottom of the board.
Let the number of balls be N and the number of levels be L (i.e number of times the ball will hit pegs). Simulate a simple galton board in python under the following assumptions.
i) The balls do not interact with each other.

ii) The probability of a ball going to either side of a peg on impact is 0.5 each.

Now using the simulated galton board experiment, do the following tasks.

a) Plot the distribution of the balls in the slots for L = [2, 6, 20, 100]
b) Find the Mean and Variance of the distributions.
c) Difference in odd and even L.