DS313: Statistical Foundations of Data Science
Assignment No: 5

Course Instructor: Siddhartha Sarma *Date:* 11 Apr 2025

Submission deadline: 23 Apr 2025, 10 AM (Bring hardcopies to the classroom)

Related topics:

- **Continuous random variables and their distributions**
- **Functions of a random variable**

Notations:

- $F_X(.)$, $f_X(.)$ denote cumulative distribution function(CDF) and probability density function (pdf) of a random variable X, respectively.
- $\mathbf{E}[X], \mathrm{Var}(X)$ denote the expectation and variance of random variable X, respectively.
- $\mathrm{Exp}(\lambda)$ denotes exponential random variable with parameter λ, i.e, $f_X(x) = \lambda e^{-\lambda x}, x \geq 0$.
- $\mathrm{Unif}(a, b)$ denotes continuous uniform distribution with a and b as left and right limits, respectively.
- $N(\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho_{XY}), \sigma_X, \sigma_Y > 0, |\rho_{XY}| < 1$, represents the *joint Gaussian distribution* of random variables X and Y, and the corresponding probability density function is

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho_{XY}^2}} \exp\left(-\frac{1}{2(1-\rho_{XY}^2)}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} - \frac{2\rho_{XY}(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\right)\right), \quad -\infty < x, y < \infty.$$

Instructions & Information:

- For the questions prefixed with [PA], you need to solve them on paper and write code to simulate the underlying experiment as well. You are free to choose the programming language for writing your code. As discussed in the second tutorial session, you should plot the normalised frequencies of the events under consideration as bar plots and compare them to their corresponding probabilities calculated analytically. Please refer to the course moodle page for examples.
- *The programming assignments should be submitted through moodle only. You also need to indicate in moodle quiz corresponding to the Assignment#5 the questions you have submitted in hard copy.*

- You must include a table similar to the one below on the front page of your submission to indicate the page numbers of the answer to the assignment questions.

Question no.	1	2	3	4	5	6	7	8	9	10
Page no.										

---

1. Let X_1 and X_2 be independent random variables, with X_i being exponentially distributed with parameter λ_i. Find the pdf of (a) $Z = \min\{X_1, X_2\}$, (b) $V = X_1 - X_2$, (c) $W = X_1/X_2$.

2. Let the random variables X and Y be jointly uniformly distributed over the region shown in Fig. 1.

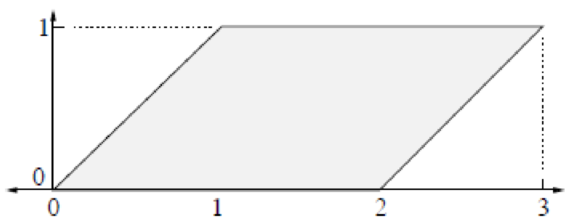

Figure 1: X and Y are uniformly distributed over the shaded region. check Q.2

   (a) Determine the value of $f_{XY}(x, y)$ on the region shown.
   (b) Find $f_X(x)$, the marginal pdf of X.
   (c) Find the mean and variance of X.
   (d) Find the conditional pdf of Y given that $X = x$, for $0 \leq x \leq 1$.
   (e) Find the conditional pdf of Y given that $X = x$, for $1 \leq x \leq 2$.
   (f) Find and sketch $\mathbf{E}[Y|X = x]$ as a function of x. Specify the range of x for which this conditional expectation is well defined for.

3. [PA] If $Y = X^2$, then express $f_Y(y|X \geq \alpha)$ in terms of CDF, PDF of X. Cross-check your answer with the help of simulation by choosing at least 5 different values for α and list the values corresponding to your analytical answer and simulation in tabular form. You may assume $X \sim \text{Exp}(1)$ or $X \sim \text{Unif}(0, 1)$.

4. (a) The function $g(x)$ is monotone increasing and $Y = g(X)$. Show that

$$F_{XY}(x, y) = \begin{cases} F_X(x), & y > g(x) \\ F_Y(y), & y < g(x) \end{cases}$$

   (b) Find $F_{XY}(x, y)$ if $g(x)$ is monotone decreasing.

5. Show that if X, Y are two random variables with densities $f_X(x), f_Y(y)$, respectively, then

$$\mathbf{E}[\ln(f_X(X))] \geq \mathbf{E}[\ln(f_Y(X))]$$

6. Let X and Y are jointly Gaussian with parameters $N(\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho_{XY}), \sigma_X, \sigma_Y > 0$. Find a necessary and sufficient condition for $X + Y$ and $X - Y$ to be independent.

7. X and Y are jointly Gaussian with parameters $N(\mu_X, \mu_Y, \sigma_X, \sigma_Y, \rho_{XY}), \sigma_X, \sigma_Y > 0$. Find (a) $\mathbf{E}[Y|X = x]$ and (b) $\mathbf{E}[X^2|Y = y]$.

8. Suppose U, V has joint pdf

$$f_{UV}(u, v) = \begin{cases} 9u^2v^2, & 0 \leq u \leq 1, 0 \leq v \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

   Let $X = 3U$ and $Y = UV$. (a) Find the joint pdf of X and Y and specify where the joint pdf is zero. (b) Using the joint pdf of X and Y, find the conditional pdf, $f_{Y|X}(y|x)$, of Y given X. (Be sure to indicate which values of x the conditional pdf is well defined for, and for each such x specify the conditional pdf for all real values of y.)

   [PA] Plot $f_{UV}(u, v)$ and $f_{XY}(x, y)$ using the appropriate 3D plotting utilities (functions or methods) in Python or MATLAB.

9. [PA] If $U \sim \text{Unif}(0, 2\pi)$ and $Z \sim \text{Exp}(1)$ are independent, then X, Y defined below are independent standard normal random variables.

$$X = \sqrt{2Z} \cos U, \quad Y = \sqrt{2Z} \sin U$$

10. [PA] Suppose X, Y, Z are i.i.d., with the distribution being $\text{Unif}(0, 1)$. What is the probability that all the roots of the equation $X\zeta^2 + Y\zeta + Z = 0$ are real? What is the probability that at least one root is greater than 1 given that the roots are real?

Some questions were taken from the following books

- Papoulis & Pillai, "Probability, Random Variables and Stochastic Processes", McGraw Hill, 4th Ed
- Sheldon Ross, "A first course in Probability", 8th Ed
- Lecture notes of Prof. Hajek, University of Illinois Urbana-Champaign