DS313: Statistical Foundations of Data Science
Assignment No: 4

Course Instructor: Siddhartha Sarma *Date:* 28 Mar 2024
Submission deadline: 11 Apr 2024, 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.
- The characteristic function and the Moment generating function of a random variable X is defined as $\Phi_X(u) = \mathbf{E}[e^{juX}]$ and $M_X(s) = \mathbf{E}[e^{sX}]$, 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.
- $N(\mu, \sigma)$ denotes Gaussian distribution with mean μ and standard deviation $\sigma > 0$ and $Q(x) = (1/\sqrt{2\pi}) \int_x^\infty e^{-u^2/2} du$

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#4 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. A system consisting of one original unit plus a spare can function for a random amount of time X. If the density of X is given (in units of months) by

$$f(x) = \begin{cases} Cxe^{-x/2}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

what is the probability that the system functions for at least 5 months?

2. [PA] Let $X = \begin{cases} U - 0.5, & U \geq 0.5 \\ 0, & U < 0.5 \end{cases}$, where $U \sim \text{Unif}(0, 1)$

(a) Find and carefully sketch the CDF $F_X(.)$. In particular, what is $F_X(0)$?

(b) Find the characteristic function $\Phi_X(u)$.

3. [PA] Let X be exponentially distributed with mean $1/\lambda$. Find and carefully sketch the distribution functions for the random variables $Y = \exp(X)$ and $Z = \min\{X, 3\}$.

4. [PA] If $X \sim \text{Unif}(-2\pi, 2\pi)$, then find $f_Y(y)$ if (a) $Y = X^3$, (b) $Y = X^4$, and (c) $Y = 2\sin(3X + 40°)$.

5. [PA] Let $F_X(x)$ be the CDF of a random variable X. Find the distribution (CDF) of random variable $Y = F_X(X)$. Crosscheck you answer for the following: $X \sim \text{Exp}(\lambda)$, $X \sim N(\mu, \sigma)$.

6. Find the mean and variance of random variables with the following characteristic functions: (a) $\Phi(u) = \exp(-5u^2 + 2ju)$, (b) $\Phi(u) = (\exp(ju) - 1)/ju$, and (c) $\Phi(u) = \exp(\lambda(\exp(ju) - 1))$.

7. Show that for a continuous random variable X (*recall Q.1 of Assignment#3*)

$$\mathbf{E}[X] = \int_0^\infty (1 - F_X(x))dx - \int_{-\infty}^0 F_X(x)dx$$

8. Express each of the given probabilities in terms of the standard Gaussian complementary CDF $Q(.)$ (a) $\Pr(X \geq 16)$, (b) $\Pr(X^2 \geq 16)$, where $X \sim N(10, 3)$.

9. If X has mean μ and standard deviation σ, the ratio $\rho := |\mu|/\sigma$ is called the measurement signal-to-noise ratio of X. The idea is that X can be expressed as $X = \mu + (X - \mu)$, with μ representing the signal and $X - \mu$ the noise. If we define $D := |(X - \mu)/\mu|$ as the relative deviation of X from its signal (or mean) μ, show that, for $\alpha > 0$,

$$\Pr(D \leq \alpha) \geq 1 - \frac{1}{\rho^2\alpha^2}$$

10. Show that if m is the median of X, then

$$\mathbf{E}[|X-a|] = \mathbf{E}[|X-m|] + 2\int_a^m (x-a) f_X(x)dx$$

for any a. Find c such that $\mathbf{E}[|X-c|]$ is minimum.

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