Problem Set 2

Probability and Statistics 2022
Indian Institute of Management, Udaipur

3^{rd} August, 2022

This is a Student's Activity Task containing a few multiple type questions and a number of descriptive type problems. Please feel free to answer as much as you can. The activity is not a graded component of the course. Its objective is to encourage students learning while solving problems.

Part I: Short answer type questions.

1. **Statement:** The function $F(x) = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}(x)$ for all $x \in \mathbb{R}$ is a valid CDF function.

   Ans: TURE FALSE

2. Let $\mathbb{I}_B(\cdot)$ denotes the indicator function. Is it true that $\mathbb{I}_{(0,4)}(1) = 0$?

   Ans: TURE FALSE

3. **Statement:** If X, Y are independent and identically distributed jointly continuous random variables then $P(X > Y) = \frac{1}{2}$.

   Ans: TURE FALSE

4. Let X follows Exponential distribution with mean 2. What is $E(X^2)$?

5. LIf X_1, X_2 be independent random variables such that $X_1 \sim \text{Gamma}(10, 2)$ and $X_1 \sim \text{Gamma}(15, 2)$ then what is the distribution of the random variable $Y := X_1 + X_2$.

6. What are the values of $\Gamma(5)$ and $\Gamma(5.5)$?

7. Let X_1 and X_2 statistically independent random variables such that $X_1 \sim \chi^2_{5df}$ and $X_2 \sim \chi^2_{3df}$. Then what is the distribution of $X_1 + X_2$?

8. What is the value of the integral $\int_0^1 x^{10}(1-x)^{15}dx$?

9. The achievement scores for a college entrance examination are normally distributed with mean 75 and standard deviation 10. What fraction of the scores lies between 80 and 90?

10. **Statement:** If $Z \sim N(0,1)$ and $V \sim \chi^2_{\nu\text{df}}$ and Z,V are **statistically independent** then the random variable $Y := \frac{Z}{\sqrt{V/\nu}} \sim t_{\nu\text{ df}}$.

Ans: TURE FALSE

11. If $V_1 \sim \chi^2_{5\text{ df}}$ and $V_2 \sim \chi^2_{20\text{df}}$ and V_1, V_2 are **statistically independent** then what is the distribution of the random variable $Y := \frac{4V_1}{V_2}$?

12. If $V_1 \sim \chi_{5\text{ df}}$ and $V_2 \sim \chi^2_{10\text{df}}$ and V_1, V_2 are **statistically independent** then the random variable $Y := \frac{2V_1}{V_2}$. Let $X \sim \text{Normal}(\mu = 10, \sigma^2 = 4)$. X is independent to both V_1, V_2. What is the distribution of $U = \frac{\sqrt{5}(X-10)}{\sqrt{2V_2}}$?

13. Let X,Y,Z are three independent and identically distributed random variables with mean 0, variance σ^2. Define $U = X+Y$, $V = X-Y$ and $W = Z+1$. Represent the

$$\text{Var}\left(3U + 4V + W + 2\right)$$

in terms of σ^2.

14. **Statement:** Let X,Y be two random variables that are statistically independent to each other. Then covariance between X and Y is 0.

Ans: TURE FALSE

15. Let the joint probability density function of the random vector (X,Y) is given as

$$f(x,y) := 8xy \text{ when } 0 \le y \le x \le 1$$

and 0 elsewhere. Are the random variables statistically independent?

16. Let X be a random variable with the moment generating function $M_X(t) := e^{4t^2}$ for $t \in \mathbb{R}$. What is the distribution of the random variable X?

17. Let (X,Y) be a random vector with the joint probability density function

$$f_{X,Y}(x,y) = \frac{1}{\sqrt{\pi}} e^{-x} e^{-y^2}$$

for $x > 0$ and $y \in \mathbb{R}$. Are the random variables statistically independent?

18. Let X be a random variable such that $\mathrm{E}(X) = 5$ and $\mathrm{Var}(X) = 5$. What is the value of $E(X^2)$?

19. Let the moment generating function of a random variable is $M_X(t) := e^{4(e^t-1)}$. What is $E(X)$?

20. Let X be a random variable with support $(0,\infty)$. Assume that $E(X)$ and $E(\log(X))$ are finite. Is the following statement true?

$$E(X \log(X)) \leq E(X) \log(E(X))$$

21. Let X be a random variable with support $(0,\infty)$. Assume that $E(e^X)$ is finite. Is it true that

$$E(e^X) \geq e^{E(X)}?$$

Part II: Descriptive Problems.

1. Let $X \sim$ Uniform(-1, 1). Derive the probability density function of the random variable $Y = X^2$.

2. A multiple-choice examination has 15 questions, each with five possible answers, only one of which is correct. Suppose that one of the students who takes the examination answers each of the questions with an independent random guess. What is the probability that he answers at least ten questions correctly?

3. A food manufacturer uses an extruder (a machine that produces bite-size cookies and snack food) that yields revenue for the firm at a rate of ₹5000 per hour when in operation. However, the extruder breaks down an average of two times every day it operates. If Y denotes the number of breakdowns per day, the daily revenue generated by the machine is $R = 40000 - 50Y^2$. Find the expected daily revenue for the extruder.

4. A geological study indicates that an exploratory oil well drilled in a particular region should strike oil with probability .2. Find the probability that the third oil strike comes on the fifth well drilled.

5. A geological study indicates that an exploratory oil well drilled in a particular region should strike oil with probability .2. Find the probability that the third oil strike comes on the fifth well drilled.

6. A company that manufactures and bottles apple juice uses a machine that automatically fills 16-ounce bottles. There is some variation, however, in the amounts of liquid dispensed into the bottles that are filled. The amount dispensed has been observed to be approximately normally distributed with mean 16 ounces and standard deviation 1 ounce. to determine the proportion of bottles that will have more than 17 ounces dispensed into them.

   Ans:

7. The median of the distribution of a continuous random variable Y is the value m_Y such that $P(Y \leq m_Y) = 0.5$. What is the median of the Exponential distribution with mean 10.

   Ans:

8. The magnitude of earthquakes recorded in a region of North America can be modeled as having an exponential distribution with mean 2.4, as measured on the Richter scale. Find the probability that an earthquake striking this region will
   (a) exceed 3.0 on the Richter scale.
   (b) fall between 2.0 and 3.0 on the Richter scale.

   Ans:

9. The length of time Y necessary to complete a key operation in the construction of houses has an exponential distribution with mean 10 hours. The formula $U = 100 + 40Y + 3Y^2$ relates the cost U of completing this operation to the square of the time to completion. Find the mean and variance of U.

Ans:

10. A company that manufactures and bottles apple juice uses a machine that automatically fills 16-ounce bottles. There is some variation, however, in the amounts of liquid dispensed into the bottles that are filled. The amount dispensed has been observed to be approximately normally distributed with mean 16 ounces and standard deviation 1 ounce. to determine the proportion of bottles that will have more than 17 ounces dispensed into them.

Ans:

11. The velocities of gas particles, V, can be modeled by the Maxwell distribution, whose probability density function is given by

$$f_V(v) := 4\pi \left(\frac{m}{2\pi KT}\right)^{\frac{3}{2}} v^2 e^{-v^2 \frac{m}{2KT}} \mathbb{I}_{\mathbb{R}_+}(v)$$

where m is the mass of the particle, K is Boltzmann's constant, and T is the absolute temperature.

(a) Find the mean velocity of these particles.

(b) The kinetic energy of a particle is given by $\frac{1}{2}mV^2$. Find the mean kinetic energy for a particle.

Ans:

12. A supplier of kerosene has a 150-gallon tank that is filled at the beginning of each week. His weekly demand shows a relative frequency behavior that increases steadily up to 100 gallons and then levels off between 100 and 150 gallons. If Y denotes weekly demand in hundreds of gallons, the relative frequency of demand can be modeled by

$$f(y) := \begin{cases} y & \text{if } 0 \leq y \leq 1 \\ 1 & \text{if } 1 \leq y \leq 1.5 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $F(y)$, the CDF of the random variable Y.

(b) Find the probability that the station will pump between 80 and 120 gallons in a particular week.

(c) Given that the station pumped more than 100 gallons in a particular week, find the probability that the station pumped more than 120 gallons during the week.

Ans:

13. we considered the random variables Y_1 (the proportional amount of gasoline stocked at the beginning of a week) and Y_2 (the proportional amount of gasoline sold during the week). The joint density function of Y_1 and Y_2 is given by

$$f_{Y_1,Y_2}(y_1,y_2) := \begin{cases} 3y_1 & \text{if } 0 \leq y_2 \leq y_1 \leq 1. \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the probability density function for $U = Y_1 - Y_2$, the proportional amount of gasoline remaining at the end of the week.

(b) Use the density function of U to find $E(U)$.

Ans:

14. Consider the joint probability mass function

$$p_{X,Y} = \frac{1}{2^{x-1}3^y}$$

for $x = 1,2,\ldots$ and $y = 1,2,3,\ldots$.

(a) What is the support of the distribution $\mathscr{S}_{X,Y}$.

(b) Find the marginal distribution of X.

Ans:

15. Let X and Y have joint density function

$$f_{X,Y}(x,y) = \begin{cases} e^{-x-y} & \text{if } x > 0, y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(a) What is P(1 < X < 2.5)? and P(1 < Y < 2.5)?

(b) What is P(X>2Y)?

Ans:

16. Suppose X and Y have joint density function

$$f_{X,Y}(x,y) = \begin{cases} 6x^2y & \text{if } 0 \leq x \leq y, x+y \leq 2 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Show that the marginal distribution of X is Beta($\alpha = 3, \beta = 2$).

(b) Derive the conditional density of Y given X=x, for some $x \in (0,1)$.

(c) Find $E(Y \mid X = 0.6)$.

Ans:

17. Is it true that: If $\mathrm{Cor}(X,Y) = 0$ then the random variables X and Y are independent. If not true then provide appropriate counter example.

Ans:

18. Consider the joint pdf of the random vectors (X,Y)

$$f(x,y) := 2 \text{ for } 0 \leq y \leq x \leq 1$$

Prove that the random variables X,Y are not statistically independent?

Ans:

19. Let X be a continuous random variable with the probability density function

$$f_X(x) := \frac{x}{50}\mathbb{I}_{[0,10]}(x).$$

(a) Evaluate the Cumulative Distribution Function of the Random variable.
(b) Find the mean of the distribution.
(c) Find the median of the distribution. (Note that CDF in this case will be strictly increasing, hence inverse function of the cdf exists.)

20. Let (X,Y) be random vector such that the marginal density of Y is

$$f_Y(y) := 2e^{-2y}$$

for $y > 0$, and zero elsewhere. Additionally, the density for the conditional distribution of X given $Y = y$ is

$$f_{X|Y}(x \mid y) := \sqrt{y}e^{-\pi x^2 y}$$

for $x \in \mathbb{R}$.
(a) What is the joint density of the random vector (X,Y)?
(b) What is the marginal density of the random variable X?

Ans:

21. Let (X,Y) be random vector with the corresponding joint density

$$f_{X,Y}(x,y) := e^{-x-y} \text{ for } x > 0, y > 0.$$

Consider a one-to-one bivariate transformation

$$U = X+Y, V = Y.$$

(a) Derive the joint density of the random vector (U,V)?
(b) What is the marginal density of U?

Ans:

22. Let the total number of items a specific Amazon seller sells in a week follows a Poisson distribution with mean $\lambda = 100$. From experience, the seller knows that proportion of items that are returned is 5%.

(a) What is the Expected number of items that are returned during a week?

(b) Let us assume that the seller makes a 25% profit on each items that are sold but not returned. However, if a product is returned then the seller needs to issue a refund of entire amount to the specific customer. Additionally the seller needs to pay for the cost of return shipping of amount D per items returned. If the price at which the seller sells each item is S, then what is expected net profit for the seller over a period of a week?

Ans:

23. If X, Y are independent and identically distributed continuous random variables then prove that $P(X > Y) = \frac{1}{2}$.

24. Let X be a continuous random variable with the corresponding pdf $f_X(x) := e^{-x}$ for $x > 0$, and 0 otherwise. Therefore, the support of the random variable $\mathscr{S}_X := \mathbb{R}_+$. Consider the transformation $Y = [X]$ where the notation $[x]$ refers to the largest integer less than or equal to x. For example, $[2] = 2$ and $[2.9] = 2$.

(a) What is the support of the random variable Y?

(b) Derive a formula for the pdf/pmf for Y.

Ans: