RANDOM VARIABLE AND PROBABILITY DISTRIBUTIONS

LECTURE-6

1 Concept of Random Variable:-

Definition 1.1. A Random variable is a function that associates a real number with each element in the sample space.

Example 1.1.
Two balls are drawn in succession without replacement from an urn containing 4 red balls and 3 black balls. The possible outcomes and the values y of the random variable Y , where Y is the number of red balls, are

Sample Space	y
RR	2
RB	1
BR	1
BB	0

Definition 1.2. The random variable for which 0 and 1 are chosen to describe the two possible values is called a **Bernoulli random variable**.

Example 1.2.
Consider the simple condition in which components are arriving from the production line and they are stipulated to be defective or not defective. Define the random variable X by

$$X = \begin{cases} 1, & \text{if the component is defective,} \\ 0, & \text{if the component is not defective.} \end{cases}$$

Definition 1.3. If a sample space contains a finite number of possibilities or an un-ending sequence with as many elements as there are whole numbers, it is called a **discrete sample space**, (i.e. The set of possible outcomes is countable).

Definition 1.4. If a sample space contains an infinite number of possibilities equal to the number of points on a line segment, it is called a **continuous sample space**,(i.e. The set of possible outcomes is uncountable).

2 Discrete Probability Distributions:-

Definition 2.1. The set of ordered pairs $(x, f(x))$ is a **probability function, probability mass function, or probability distribution of the discrete random variable** X if, for each possible outcome x,
1. $f(x) \geq 0$,
2. $\sum_x f(x) = 1$,
3. $P(X = x) = f(x)$.

Definition 2.2. The cumulative distribution function $F(x)$ of a discrete random variable X with probability distribution $f(x)$ is
$F(x) = P(X \leq x) = \sum_{t \leq x} f(t)$, for $-\infty < x < \infty$.The probability distribution function(pdf) can be obtained by directly the cumulating distribution function(cdf) provided 'X' is a continuous random variable.

3 Continuous Probability Distributions

Definition 3.1.

The function $f(x)$ is a probability density function (pdf) for the continuous random variable X, defined over the set of real numbers, if
1. $f(x) \geq 0$, for all $x \in \mathbb{R}$,
2. $\int_{-\infty}^{\infty} f(x)dx = 1$,
3. $P(a < X < b) = \int_a^b f(x)dx$.

Definition 3.2. The **cumulative distribution function** $F(x)$ of a continuous random variable X with density function $f(x)$ is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)dt, \ for - \infty < x < \infty,$$

where $P(a < X < b) = F(b) - F(a)$ and $f(x) = \frac{dF(x)}{dx}$.

LECTURE-7

Problem 3.3. 7).The total number of hours, measured in units of 100 hours, that a family runs a vacuum cleaner over a period of one year is a continuous random variable X that has the density function

$$f(x) = \begin{cases} x, & 0 < x < 1, \\ 2 - x, & 1 \leq x < 2, \\ 0, & \text{elsewhere} . \end{cases}$$

Find the probability that over a period of one year, a family runs their vacuum cleaner
(a) less than 120 hours,
(b) between 50 and 100 hours.
sols:-a)

$$\begin{aligned} P(X < 1.2) &= \int_{-\infty}^{1.2} f(x)dx \\ &= \int_{-\infty}^{0} f(x)dx + \int_{0}^{1} f(x)dx + \int_{1}^{1.2} f(x)dx \\ &= 0 + \int_{0}^{1} xdx + \int_{1}^{1.2} (2-x)dx \\ &= \frac{1}{2} + 2(1.2-1) - \frac{1}{2}((1.2)^2 - 1) \\ &= \frac{1}{2} + 0.4 - 0.22 \\ &= 0.68. \end{aligned}$$

b)

$$\begin{aligned}
P(0.5 < X < 1) &= \int_{0.5}^{1} x dx \\
&= \frac{x^2}{2}\Big]_{0.5}^{1} \\
&= \frac{1}{2}(1 - 0.25) \\
&= 0.375.
\end{aligned}$$

Problem 3.4. 10). Find a formula for the probability distribution of the random variable X representing the outcome when a single die is rolled once.
solution:-

$$P(X = x) = \begin{cases} \frac{1}{6}, & X = 1, 2, 3, 4, 5, 6, \\ 0, & elesewhere. \end{cases}$$

Problem 3.5. 11) A shipment of 7 television sets contains 2 defective sets. A hotel makes a random purchase of 3 of the sets. If x is the number of defective sets purchased by the hotel, find the probability distribution of X.
solution:-

$$f(x) = \frac{{}^2c_x \ \ {}^5c_{3-x}}{{}^7c_3}, \ x = 0, 1, 2.$$

When $x = 0, \ f(x) = \frac{2}{7}, \ x = 1, \ then f(x) = \frac{4}{7}$ similarly $x = 2, \ f(x) = \frac{1}{7}$.

Problem 3.6. 12). An investment firm offers its customers municipal bonds that mature after varying numbers of years, given that the cumulative distribution function of T, the number of years to maturity for a randomly selected bond, is

$$F(t) = \begin{cases} 0, & t < 1, \\ \frac{1}{4}, & 1 \leq t < 3, \\ \frac{1}{2}, & 3 \leq t < 5, \\ \frac{3}{4}, & 5 \leq t < 7, \\ 1, & t \geq 7. \end{cases}$$

find (a) $P(T = 5)$
(b) $P(T > 3)$
(c) $P(1.4 < T < 6)$

(d) $P(T \leq 5 \mid T \geq 2)$

solutions:- a) $P(T=5) = F(5) - F(4) = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$

b) $P(T > 3) = 1 - P(T \leq 3) = 1 - F(3) = 1 - \frac{1}{2} = \frac{1}{2}$

c)$P(1.4 < T < 6) = F(6) - F(1.4) = \frac{3}{4} - \frac{1}{4} = \frac{1}{2}$

d)$P(T \leq 5 \mid T \geq 2) = \frac{P(2 \leq T \leq 5)}{P(T \geq 2)} = \frac{F(5)-F(2)}{1-F(2)} = \frac{\frac{3}{4}-\frac{1}{4}}{1-\frac{1}{4}} = \frac{1}{2} * \frac{4}{3} = \frac{2}{3}$.

Problem 3.7. 14). The waiting time, in hours, between successive speeders spotted by a radar unit is a continuous random variable with cumulative distribution function

$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-8x}, & x \geq 0. \end{cases}$$

Find the probability of waiting less than 12 minutes between successive speeders

(a) using the cumulative distribution function of X;

(b) using the probability density function of X.

solutions:- a)$P(X < 0.2) = F(0.2) = 1 - e^{-8*0.2} = 1 - e^{-1.6} = 0.7981$.

b)$f(x) = F'(x) = 8 * e^{-8x}$

$P(X < 0.2) = 8 * \displaystyle\int_0^{0.2} e^{-8x} dx = -e^{-8x}\Big]_0^{0.2} = 0.7981.$

Problem 3.8. 21). Consider the density function

$$f(x) = \begin{cases} k\sqrt{x}, & 0 < x < 1, \\ 0, & \text{elsewhere} . \end{cases}$$

(a) Evaluate k.

(b) Find $F(x)$ and use it to evaluate $P(0.3 < X < 0.6)$

solutions:- a)$\int_{-\infty}^{\infty} f(x)dx = 1$

$\Rightarrow \int_0^1 k\sqrt{x}dx = 1$

$\Rightarrow k\frac{x^{\frac{3}{2}}}{\frac{3}{2}}\Big]_0^1 = 1$

$\Rightarrow \frac{2k}{3} = 1$

$\Rightarrow k = \frac{3}{2}$

b) $F(x) = \int_{-\infty}^{x} f(t)dt$

$= \int_{-\infty}^{0} 0dt + \int_{0}^{x} \frac{3}{2}\sqrt{t}dt = x^{\frac{3}{2}}$, **where**

$$F(x) = \begin{cases} 0, \ x < 0, \\ x^{\frac{3}{2}}, \ 0 \le x < 1, \\ 1, \ x \ge 1. \end{cases}$$

$P(0.3 < X < 0.6) = F(0.6) - F(0.3) = (0.6)^{\frac{3}{2}} - (0.3)^{\frac{3}{2}} = 0.3004.$

Problem 3.9. 29).3.29 An important factor in solid missile fuel is the particle size distribution. Significant problems occur if the particle sizes are too large. From production data in the past, it has been determined that the particle size (in micrometers) distribution is characterized by

$$f(x) = \begin{cases} 3x^{-4}, & x > 1, \\ 0, & \text{elsewhere} . \end{cases}$$

(a) Verify that this is a valid density function.
(b) Evaluate $F(x)$.
(c) What is the probability that a random particle from the manufactured fuel exceeds4 micrometers?

solution:- a) $f(x) \ge 0$ and $\int_{-\infty}^{\infty} f(x)dx$
$= \int_{1}^{\infty} 3x^{-4}dx = \frac{3x^{-3}}{-3}\Big]_{1}^{\infty}$
$= -(\infty^{-3} - 1)$
$= 1$
Hence this is a valid valid density function for $x \ge 1$.
b)$F(x) = \int_{-\infty}^{x} f(t)dt$
$= \int_{-\infty}^{1} 0dt + \int_{1}^{x} 3t^{-4}dt$
$= 0 + \frac{3t^{-3}}{-3}\Big]_{1}^{x}$
$= 1 - x^{-3}$, **where**

$$F(x) = \begin{cases} 0, \ x < 1 \\ 1 - x^{-3}, \ x \ge 1. \end{cases}$$

c) $P(X > 4) = 1 - F(4) = 1 - (1 - 4^{-3}) = 4^{-3} = 0.0156.$

Problem 3.10. 30). Measurements of scientific systems are always subject to variation, some more than others. There are many structures for measurement error, and statisticians spend a great deal of time modeling these errors. Suppose the measurement error X of a certain physical quantity is decided by the density function

$$f(x) = \begin{cases} k\left(3 - x^2\right), & -1 \leq x \leq 1, \\ 0, & \text{elsewhere} . \end{cases}$$

(a) Determine k that renders $f(x)$ a valid density function.
(b) Find the probability that a random error in measurement is less than 1/2.
(c) For this particular measurement, it is undesirable if the magnitude of the error (i.e., $|x|$) exceeds 0.8. What is the probability that this occurs?

solutions:- a)$f(x) \geq 0$
$\int_{-\infty}^{\infty} f(x)dx = 1$
$\Rightarrow k\int_{-1}^{1}(3 - x^2)dx = 1$
$\Rightarrow 2k\int_{0}^{1}(3 - x^2)dx = 1$
$\Rightarrow 2k\left[3x - \frac{x^3}{3}\right]_0^1 = 1$
$\Rightarrow \frac{16}{3}k = 1$
$\Rightarrow k = \frac{3}{16}.$
b) for$-1 \leq x \leq 1$
$F(x) = \int_{-1}^{x} \frac{3}{16}(3 - t^2)dt = \frac{3}{16}\left[3t - \frac{t^3}{3}\right]_{-1}^{x}$
$= \frac{3}{16}\left[3(x+1) - \frac{1}{3}(x^3 + 1)\right]$
$= \frac{3}{16}\left(3x + 3 - \frac{x^3}{3} - \frac{1}{3}\right)$
$= \frac{9}{16}x - \frac{x^3}{16} + \frac{1}{2}$
$P(X < \frac{1}{2}) = F(\frac{1}{2}) = \frac{1}{2} + \frac{9}{16}\frac{1}{2} - \frac{1}{16}\frac{1}{8}$
$= 0.773.$
c)$P(\mid X \mid > 0.8) = P(X < -0.8) + P(X > 0.8)$
$= F(-0.8) + 1 - F(0.8)$
$= \frac{1}{2} + \frac{9}{16} * (-0.8) - \frac{1}{16}(-0.8)^3 + 1 - \frac{1}{2} + \frac{9}{16} * 0.8 - \frac{1}{16}(0.8)^3$
$= 0.164.$

Problem 3.11. 35).Suppose it is known from large amounts of historical data that X, the number of cars that arrive at a specific intersection during a 20 -second time period, is characterized by the following discrete probability function:

$$f(x) = e^{-6}\frac{6^x}{x!}, \text{ for } x = 0, 1, 2, \ldots$$

(a) Find the probability that in a specific 20 -second time period, more than 8 cars arrive at the intersection.

(b) Find the probability that only 2 cars arrive.

solutions:- a)$P(X > 8) = 1 - P(X \leq 8)$

$= 1 - \sum_{x=0}^{8} e^{-6} \frac{6^x}{x!}$

$= 1 - \left(e^{-6} \frac{6^0}{0!} + e^{-6} \frac{e^1}{1!} + \cdots + e^{-6} \frac{6^8}{8!}\right)$

$= 0.1528.$

b) $P(X = 2) = f(2) = e^{-6} \frac{6^2}{2!} = 0.0446.$

LECTURE-8

4 Joint Probability Distributions

Definition 4.1. The function $f(x, y)$ is a joint probability distribution or probability mass function of the discrete random variable X and Y if
1. $f(x, y) \geq 0$ for all (x, y),
2. $\sum_x \sum_y f(x, y) = 1$,
3. $P(X = x, Y = y) = f(x, y)$.
For any region A in the xy-plane, $P[(X, Y) \in A] = \sum \sum_A f(x, y)$.

Definition 4.2. The function $f(x, y)$ is a joint density function of the continuous random variables X and Y if
1. $f(x, y) \geq 0$, for all (x, y),
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$,
3. $P[(X, Y) \in A] = \int\int_A f(x, y) dx dy$. For any region A in the xy plane.

Definition 4.3. The marginal distributions of X alone and of Y alone are

$$g(x) = \sum_y f(x, y) \quad \text{and} \quad h(y) = \sum_x f(x, y)$$

for the discrete case, and

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad \text{and} \quad h(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

for the continuous case.

Definition 4.4. Let X and Y be two random variables, discrete or continuous. The conditional distribution of the random variable Y given that $X = x$ is

$$f(y \mid x) = \frac{f(x, y)}{g(x)}, \text{ provided } g(x) > 0$$

Similarly, the conditional distribution of X given that $Y = y$ is

$$f(x \mid y) = \frac{f(x, y)}{h(y)}, \text{ provided } h(y) > 0.$$

Definition 4.5. Let X and Y be two random variables, discrete or continuous, with joint probability distribution $f(x, y)$ and marginal distributions $g(x)$ and $h(y)$, respectively. The random variables X and Y are said to be **statistically independent** if and only if

$$f(x, y) = g(x)h(y)$$

for all (x, y) within their range.

Problem 4.6. 38. If the joint probability distribution is given by

$$f(x, y) = \frac{x + y}{30}, \qquad \text{for } x = 0, 1, 2, 3 \ \text{ and } y = 0, 1, 2 \ \text{ then}$$

find
(a) $P(X \leq 2, Y = 1)$.
(b) $P(X > 2, Y \leq 1)$.
(c) $P(X > Y)$.
(d) $P(X + Y = 4)$.
Solution:- a) $P(X \leq 2, Y = 1) = f(0, 1) + f(1, 1) + f(2, 1)$
$= \frac{1}{30} + \frac{2}{30} + \frac{3}{30} = \frac{1}{5}$.
b)$P(X > 2, Y \leq 1) = f(3, 0) + f(3, 1)$
$= \frac{3}{30} + \frac{4}{30} = \frac{7}{30}$.
c) $P(X > Y) = f(1, 0) + f(2, 0) + f(2, 1) + f(3, 0) + f(3, 1) + f(3, 2)$
$= \frac{1}{30} + \frac{2}{30} + \frac{3}{30} + \frac{3}{30} + \frac{4}{30} + \frac{5}{30}$
$= \frac{9}{15}$
d)$P(X + Y = 4) = f(2, 2) + f(3, 1) = \frac{4}{30} + \frac{4}{30} = \frac{4}{15}$.

Problem 4.7. 42. Let X and Y denote the lengths of life, in years, of two components in an electronic system. If the joint density function of these variables is

$$f(x, y) = \begin{cases} e^{-(x+y)}, & x > 0, y > 0, \\ 0, & \text{elsewhere .} \end{cases}$$

find $P(0 < X < 1 \mid Y = 2)$
Solution:- $f(x \mid y) = \frac{f(x,y)}{h(y)}$, provided $h(y) > 0$
$Now \;\; h(y) = \int_{-\infty}^{\infty} f(x, y)dx$
$= \int_0^{\infty} e^{-(x+y)}dx$
$= e^{-y} \int_0^{\infty} e^{-x}dx = e^{-y}\big[-e^{-x}\big]_0^{\infty}$

$= -e^{-y}(e^{-\infty} - e^0)$
$= e^{-y} > 0,\ y > 0.$
$f(x \mid y) = \frac{e^{-(x+y)}}{e^{-y}} = e^{-x}$
$P(0 < X < 1 \mid Y = 2) = \int_0^1 e^{-x}dx\ , x > 0$
$= -e^{-x}\Big]_0^1$
$= -(e^{-1} - 1)$
$= 1 - \frac{1}{e}.$

LECTURE-9

Problem 4.8. 49). Let X denote the number of times a certain numerical control machine will malfunction: $1, 2$, or 3 times on any given day. Let Y denote the number of times a technician is called on an emergency call. Their joint probability distribution is given as

$f(x,y)$		$x=1$	$x=2$	$x=3$
y	1	0.05	0.05	0.10
	3	0.05	0.10	0.35
	5	0.00	0.20	0.10

(a) Evaluate the marginal distribution of X.
(b) Evaluate the marginal distribution of Y.
(c) Find $P(Y = 3 \mid X = 2)$
Solution:- a).
We have $g(x) = \sum_y f(x,y)$,
where $g(1) = \sum_y f(1,y) = f(1,1) + f(1,3) + f(1,5)$
$= 0.05 + 0.05 + 0.00 = 0.1$
$g(2) = \sum_y f(2,y) = f(2,1) + f(2,3) + f(2,5)$
$= 0.05 + 0.10 + 0.20 = 0.35$
$g(3) = \sum_y f(3,y) = f(3,1) + f(3,3) + f(3,5)$
$= 0.10 + 0.35 + 0.10 = 0.55$
Here these are the marginal distributions of X.
b)
We have $h(y) = \sum_y f(x,y)$
$h(1) = f(1,1) + f(2,1) + f(3,1) = 0.05 + 0.05 + 0.10 = 0.20,$

$h(3) = f(1,3) + f(2,3) + f(3,3) = 0.05 + 0.10 + 0.35 = 0.50,$
$h(5) = f(1,5) + f(2,5) + f(3,5) = 0.00 + 0.20 + 0.10 = 0.30,$
Here these are the marginnal distribution of Y
c)we have $f(y \mid x) = \frac{f(x,y)}{g(x)}$, provided $g(x) > 0$
Hence $P(Y = 3 \mid X = 2) = \frac{f(2,3)}{g(2)} = \frac{0.10}{0.35} = \frac{10}{35} = \frac{2}{7.}$

Problem 4.9. 56). The joint density function of the random variables X and Y is

$$f(x,y) = \begin{cases} 6x, & 0 < x < 1, 0 < y < 1 - x, \\ 0, & \text{elsewhere} \end{cases}$$

(a) Show that X and Y are not independent.
(b) Find $P(X > 0.3 \mid Y = 0.5)$.
Solutions:- a) we have the X and Y are independent random variable if $f(x,y) = g(x)h(y)$ otherwise not independent

Now $g(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^{1-x} 6xdy = 6x\big[y\big]_0^{1-x} = \begin{cases} 6x(1-x), \ 0 < x < 1, \\ 0, \ elsewhere. \end{cases}$

$h(y) = \int_{-\infty}^{\infty} f(x,y)dx = \int_0^{1-y} 6xdx = 6\frac{x^2}{2}\Big]_0^{1-y} = \begin{cases} 3(1-y)^2, \ 0 < y < 1, \\ 0, \ elsewhere. \end{cases}$

as $g(x)h(y) \neq f(x,y)$
Hence X and Y are not independent.
b)we have $f(x \mid y) = \frac{f(x,y)}{h(y)}, \ provided \ h(y) > 0,$
$P(X > 0.3 \mid Y = 0.5) = \int_{0.3}^{0.5} xdx = 0.64$