STATISTICAL METHODS

MODULE 4: Random Variables and sampling:

Discrete, continuous and mixed random variables, probability mass, probability density and cumulative distribution functions, mathematical expectation, moments

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Concept of a Random Variable

A **random variable (RV)** is a rule that assign numerical value to each outcome in a sample space.

Illustration: In a random experiment of tossing 2 coins, the sample space be $S = \{HH, HT, TH, TT\}$.

If X is a RV denoting the "number of heads", then we have the followings:

$X(HH) = 2, X(HT) = 1, X(TH) = 1, X(TT) = 0,$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI
STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Concept of a Random Variable

A **random variable (RV)** is a rule that assign numerical value to each outcome in a sample space.

Illustration: In a random experiment of tossing 2 coins, the sample space be $S = \{HH, HT, TH, TT\}$.

If X is a RV denoting the "number of heads", then we have the followings:

$X(HH) = 2, X(HT) = 1, X(TH) = 1, X(TT) = 0,$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI
STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Types of Random Variable

- Random Variable
  - Discrete RV
  - Continuous RV

Discrete Random Variable

Definition: A random variable which takes finite or at most countable number of values is called discrete random variable.

Example :

- Number of head obtained when two coins are tossed
- Number of goals scored in a soccer match
- Number of cars sold by a dealership in a day
- Number of defective items in a batch

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI
STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Example-Discrete Random Variable

Example 1: Suppose we are playing the board game "Snakes and Ladders," where the moves are determined by the sum of two independent throws with a die.

The sample space S is

$$S = \{(\omega_1, \omega_2): \omega_1, \omega_2 \in \{1, 2, \ldots, 6\}\}$$
$$= \{(1,1), (1,2), \ldots, (1,6), (2,1), \ldots, (6,5), (6,6)\}.$$

However, as players of the game, we are only interested in the sum of the outcomes of the two throws

So, $X(\omega_1, \omega_2) = \omega_1 + \omega_2$ for $(\omega_1, \omega_2) \in S$.

The possible outcomes are listed in the table alongside.

We denote the event that the function X attains the value k by $\{X = k\}$ i.e.,

	ω_1					
ω_2	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$\{X = k\} = \{(\omega_1, \omega_2) \in S : X(\omega_1, \omega_2) = k\}$$

Quick exercise: List the outcomes in the event $\{X = 8\}$.

Ans: {(2,6), (3,5), (4,4), (5,3), (6,2)}

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI
STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

The function X is a example of what we call discrete random variable.

We denote the probability of the event $\{X = k\}$ by

$$P(X = k)$$

In our example, X attains only the values $k = 2, 3, \ldots, 12$ with positive probability. For example,

$$P(X = 2) = P((1,1)) = \frac{1}{36}$$

$$P(X = 3) = P(\{(1,2),(2,1)\}) = \frac{2}{36}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

The probability distribution of a discrete random variable

Definition: Let X be a random variable on S taking values $a_1, a_2, \ldots, a_n$ with probabilities $P_1, P_2, P_3, \ldots, P_n$ respectively. The set of values of X together with their corresponding probabilities is called probability mass function on X.

The **probability mass function** or **probability function** or **probability distribution** f of a discrete random variable X is the function $f: \Re \rightarrow [0,1]$, defined by

$$f(a) = P(X = a) \text{ for } -\infty < a < \infty$$

NOTE: If X is a discrete random variable that takes on the values $a_1, a_2, \ldots, a_n$ then

$$f(a_i) \geq 0,$$

$$f(a_1) + f(a_2) + \cdots = 1$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Example- Probability distribution of a discrete random variable

Example 2: A shipment of 20 similar laptop computers to a retail outlet contains 3 that are defective. If a school makes a random purchase of 2 of these computers, find the probability distribution for the number of defectives.

Solution : Let X be a random variable whose values x are the possible numbers of defective computers purchased by the school. Then x can only take the numbers 0, 1, and 2.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

The probability distribution of a discrete random variable

Now

$$f(0) = P(X = 0) = \frac{\binom{3}{0}\binom{17}{2}}{\binom{20}{2}} = \frac{68}{95}$$

$$f(1) = P(X = 1) = \frac{\binom{3}{1}\binom{17}{1}}{\binom{20}{2}} = \frac{51}{190}$$

$$f(2) = P(X = 2) = \frac{\binom{3}{2}\binom{17}{0}}{\binom{20}{2}} = \frac{3}{190}$$

Thus, the probability distribution of X is

x	0	1	2
$f(x)$	$\frac{68}{95}$	$\frac{51}{190}$	$\frac{3}{190}$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Cumulative Distribution Function (CDF) of a discrete random variable

Definition: The cumulative distribution function (CDF) for a discrete random variable *X*, in terms of *i*, is expressed as the sum of the probabilities for all values of *X* up to *i*.

If the possible values of *X* are $x_1, x_2, \dots, x_n$ then the CDF is:

$$F(i) = P(X \leq i) = \sum_{x_j \leq i} P(X = x_j)$$

Here, *F(i)* gives the probability that the random variable *X* takes a value less than or equal to *i*, where $P(X = x_j)$ is the probability that *X* equals a specific value x_j . The function accumulates these probabilities as *i* increases.

The cumulative distribution function (CDF) of random variable X is defined as

$$F_X(x) = P(X \leq x), \text{ for all } x \in \mathbb{R}.$$

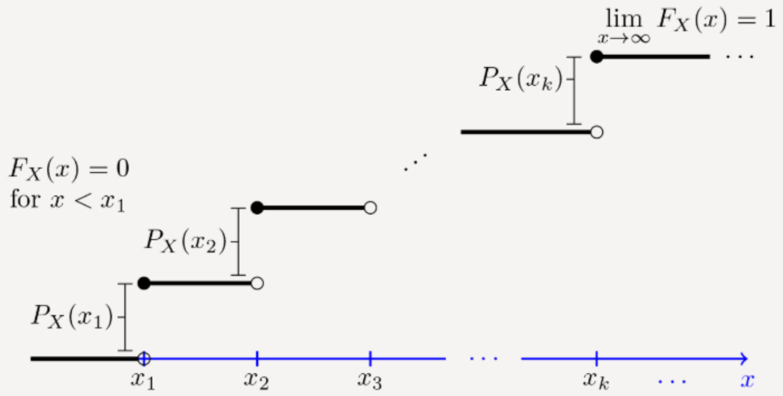
$$F_X(x) = F_X(x_k), \text{ for } x_k \leq x < x_{k+1}.$$

The CDF jumps at each x_k. In particular, we can write

$$F_X(x_k) - F_X(x_k - \epsilon) = P_X(x_k), \text{ For } \epsilon > 0 \text{ small enough.}$$

Thus, the CDF is always a non-decreasing function, i.e., if $y \geq x$ then $F_X(y) \geq F_X(x)$. Finally, the CDF approaches 1 as x becomes large. We can write

$$\lim_{x \to \infty} F_X(x) = 1.$$



Example – CDF of discrete random variable

Example 3: Find the cumulative distribution function of the random variable X. Hence verify that $f(2) = 3/8$. Find F(0), F(1), F(2), F(3), and F(4)

x	0	1	2	3	4
$f(x)$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

Solution. From the definition cumulative distribution function $F(x)$ we have,

$$F(0) = f(0) = \frac{1}{16}$$

$$F(1) = f(0) + f(1) = \frac{5}{16}$$

$$F(2) = f(0) + f(1) + f(2) = \frac{11}{16}$$

$$F(3) = f(0) + f(1) + f(2) + f(3) = \frac{15}{16}$$

$$F(4) = f(0) + f(1) + f(2) + f(3) + f(4) = 1$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Example – CDF of discrete random variable

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

- Hence,

$$F(x) \begin{cases} 0, & \text{for } x < 0, \\ \frac{1}{16}, & \text{for } 0 \leq x < 1, \\ \frac{5}{16}, & \text{for } 1 \leq x < 2, \\ \frac{11}{16}, & \text{for } 2 \leq x < 3, \\ \frac{15}{16}, & \text{for } 3 \leq x < 4, \\ 1, & \text{for } x \geq 4. \end{cases}$$

Now $f(2) = F(2) - F(1) = \frac{11}{16} - \frac{5}{16} = \frac{3}{8}$

Example – CDF of discrete random variable

Example 4: Suppose a random variable X has a probability mass function given by P(1) = 1/2 , P(2) = 1/3, and P(3) = 1/6. Determine the cumulative distribution function.

X	1	2	3
P(X)	1/2	1/3	1/6

Solution: Here,

The CDF, $F(x) = P(X \leq x)$, is the cumulative probability up to each value of x.

1. For $x < 1$:

$$F(x) = 0 \quad \text{(since there is no probability mass below 1)}$$

2. For $x = 1$:

$$F(1) = P(X \leq 1) = P(X = 1) = \frac{1}{2}$$

3. For $x = 2$:

$$F(2) = P(X \leq 2) = P(X = 1) + P(X = 2) = \frac{1}{2} + \frac{1}{3} = \frac{3}{6} + \frac{2}{6} = \frac{5}{6}$$

4. For $x = 3$:

$$F(3) = P(X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{5}{6} + \frac{1}{6} = 1$$

$$F(x) = \begin{cases} 0, & \text{for } x < 1 \\ \frac{1}{2}, & \text{for } 1 \leq x < 2 \\ \frac{5}{6}, & \text{for } 2 \leq x < 3 \\ 1, & \text{for } x \geq 3 \end{cases}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI
STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Example – CDF of discrete random variable

Example 5:

A random variable X has the following probability function.

Values of X	0	1	2	3	4	5	6	7	8
$P(X = x)$	a	3a	5a	7a	9a	11a	13a	15a	17a

(i) **Find the value of a** (ii) **Find** $P(X < 3)$, $P(0 < X < 3)$, $P(X \geq 3)$

(iii) Find the distribution function of X.

Solution:

i) We know that , $\sum_i p_i = 1$

$$a + 3a + 5a + 7a + 9a + 11a + 13a + 15a + 17a = 1$$

$$81a = 1 \quad \Rightarrow a = \frac{1}{81}$$

ii) $P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2) = a + 3a + 5a$

$$= 9a = \frac{1}{9}$$

$$P(0 < X < 3) = P(X = 1) + P(X = 2) = 3a + 5a = 8a = \frac{8}{81}$$

$$P(X \geq 3) = 1 - P(X < 3) = 1 - \frac{1}{9} = \frac{8}{9}$$

iii) Distribution function F(x) of X

Values of X	0	1	2	3	4	5	6	7	8
$F(x) = P(X \leq x)$	a $\frac{1}{81}$	4a $\frac{4}{81}$	9a $\frac{9}{81}$	16a $\frac{16}{81}$	25a $\frac{25}{81}$	36a $\frac{36}{81}$	49a $\frac{49}{81}$	64a $\frac{64}{81}$	81a 1

Example – CDF of discrete random variable

Example 6: A random variable X has the following probability of X

X	0	1	2	3	4	5	6	7
P(X)	0	k	2k	2k	3k	k^2	$2k^2$	$7k^2 + k$

i. Find value of k.

ii. Evaluate $P(X < 6)$, $P(X \geq 6)$ & $P(0 < X < 5)$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

To find the value of k for the random variable X with the given probability conditions, we can start by simplifying the equation:

$$P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4)+P(X=5)+P(X=6)+P(X=7)=1$$

Substituting the probabilities provided:

$$0+k+2k+2k+3k+k^2+2k^2+(7k^2+k)=1$$

Combine like terms:

$$0+(k+2k+2k+3k+k)+(k^2+2k^2+7k^2)=1$$

This simplifies to:

$$9k+10k^2=1$$

Step 1: Rearranging the Equation

Rearranging gives us a quadratic equation:

$$10k^2+9k-1=0$$

Step 2: Solve the Quadratic Equation

We can use the quadratic formula:

$$k=\frac{-b\pm\sqrt{b^2-4ac}}{2a}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

This gives us two possible solutions:

1. $k = \frac{2}{20} = \frac{1}{10} = 0.1$
2. $k = \frac{-20}{20} = -1$ (not valid since probabilities cannot be negative)

So, the value of k is:

$$k = 0.1$$

Step 3: Evaluate the Probabilities

Now, we can find the required probabilities using $k = 0.1$:

1. **Calculate Individual Probabilities:**
   - $P(X = 0) = 0$
   - $P(X = 1) = k = 0.1$
   - $P(X = 2) = 2k = 0.2$
   - $P(X = 3) = 2k = 0.2$
   - $P(X = 4) = 3k = 0.3$
   - $P(X = 5) = k^2 = 0.01$
   - $P(X = 6) = 2k^2 = 0.02$
   - $P(X = 7) = 7k^2 + k = 0.07 + 0.1 = 0.17$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

2. **Calculate** $P(X < 6)$:

$$P(X < 6) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$$

$$= 0 + 0.1 + 0.2 + 0.2 + 0.3 + 0.01 = 0.81$$

3. **Calculate** $P(X \geq 6)$:

$$P(X \geq 6) = P(X = 6) + P(X = 7) = 0.02 + 0.17 = 0.19$$

4. **Calculate** $P(0 < X < 5)$:

$$P(0 < X < 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

$$= 0.1 + 0.2 + 0.2 + 0.3 = 0.8$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI
STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Some practice questions

Ques 1) The probability mass function of a random variable X is defined as $P(X=0) = 3C^2$, $P(X=1) = 4C - 10C^2$, and $P(X=2) = 5C - 1$, where $C > 0$ and $P(X = r) = 0$, if $r \neq 0,1,2$,

(i) Solve to find the value of C, Ans: C= 2/7

(ii) Solve to find $P\left(0 < X < 2/X > 0\right)$. Ans: 16/37

Ques 2) If the probability mass function of a random variable X is given by $P(X=r) = kr^3$, r=1, 2, 3, 4, then

(i) Solve to find the value of k, and Ans) k = 1/100

(ii) Solve to find $P\left(\frac{1}{2} < X < \frac{5}{2}/X > 1\right)$. Ans) k=1/11

Ques 3) A discrete random variable X has the following probability distribution:

X	1	2	3	4	5	6	7	8	9
P (X)	b	2b	4b	6b	8b	10b	12b	14b	16b

(i) Solve to find the value of b,

(ii) Solve to find $P\left(X \leq 4\right)$.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Some practice questions

Ques 1) The probability mass function of a random variable X is defined as P(X=0) =3C^2, P(X=1) =4C-10C^2, and P(X=2) =5C-1, where C > 0 and P(X = r) =0, if $r \neq 0,1,2$,

(i) Solve to find the value of C, Ans: C= 2/7

(ii) Solve to find P $(0 < X < 2/X > 0)$. Ans: 16/37

Step 1: Find the value of C

We are given the probability mass function (PMF) of a random variable X, with the following probabilities:

- $P(X=0)=3C^2$
- $P(X=1)=4C-10C^2$
- $P(X=2)=5C-1$
- $P(X=r)=0$ for all other values of r (i.e., $r \neq 0,1,2$).

Since these are probabilities, the sum of all probabilities must equal 1:

$$P(X=0)+P(X=1)+P(X=2)=1$$

Substituting the given values:

$$3C^2+(4C-10C^2)+(5C-1)=1$$

Simplify the equation:

$$3C^2+4C-10C^2+5C-1=1$$

Combine like terms:

$$-7C^2+9C-1=1$$

Simplify:

$$C=\frac{-9\pm\sqrt{81-56}}{-14}$$

$$C=\frac{-9\pm\sqrt{25}}{-14}$$

$$C=\frac{-9\pm 5}{-14}$$

Now, calculate the two possible values of C:

1. $C=\frac{-9+5}{-14}=\frac{-4}{-14}=\frac{2}{7}$
2. $C=\frac{-9-5}{-14}=\frac{-14}{-14}=1$

Since the problem states that $C>0$, we discard $C=1$ and choose $C=\frac{2}{7}$.

Thus, the value of C is:

$$C=\frac{2}{7}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI
STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Some practice questions

Step 2: Find $P(0 < X < 2 \mid X > 0)$

This is the conditional probability $P(0 < X < 2 \mid X > 0)$, which simplifies to $P(X = 1 \mid X > 0)$ because $X = 1$ is the only value satisfying $0 < X < 2$.

The conditional probability formula is:

$$P(X = 1 \mid X > 0) = \frac{P(X = 1)}{P(X > 0)}$$

Step 2.1: Find $P(X = 1)$

Substitute $C = \frac{2}{7}$ into the expression for $P(X = 1)$:

$$P(X = 1) = 4C - 10C^2$$

$$P(X = 1) = 4 \times \frac{2}{7} - 10 \times \left(\frac{2}{7}\right)^2$$

$$P(X = 1) = \frac{8}{7} - 10 \times \frac{4}{49}$$

$$P(X = 1) = \frac{8}{7} - \frac{40}{49}$$

Convert to a common denominator:

$$P(X = 1) = \frac{56}{49} - \frac{40}{49} = \frac{16}{49}$$

Step 2.2: Find $P(X > 0)$

We need to find $P(X > 0) = P(X = 1) + P(X = 2)$.

Substitute $C = \frac{2}{7}$ into the expression for $P(X = 2)$:

$$P(X = 2) = 5C - 1$$

$$P(X = 2) = 5 \times \frac{2}{7} - 1$$

$$P(X = 2) = \frac{10}{7} - 1 = \frac{10}{7} - \frac{7}{7} = \frac{3}{7}$$

Now, add $P(X = 1)$ and $P(X = 2)$:

$$P(X > 0) = P(X = 1) + P(X = 2) = \frac{16}{49} + \frac{3}{7}$$

Convert to a common denominator:

$$P(X > 0) = \frac{16}{49} + \frac{21}{49} = \frac{37}{49}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI
STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Step 2.3: Compute the Conditional Probability

Finally, substitute into the conditional probability formula:

$$P(X = 1 \mid X > 0) = \frac{P(X = 1)}{P(X > 0)} = \frac{\frac{16}{49}}{\frac{37}{49}} = \frac{16}{37}$$

Final Answer:

$$P(0 < X < 2 \mid X > 0) = \frac{16}{37}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Since X is a discrete random variable, the sum of all probabilities must be equal to 1:

$$\sum_{r=1}^{4} P(X = r) = 1$$

$$k(1^3) + k(2^3) + k(3^3) + k(4^3) = 1$$

$$k(1 + 8 + 27 + 64) = 1$$

$$k(100) = 1$$

$$k = \frac{1}{100}$$

Step 2: Solve for $P(1/2 < X < 5/2 \mid X > 1)$

This means calculating the conditional probability:

$$P(1/2 < X < 5/2 \mid X > 1)$$

Step 2.1: Find $P(1/2 < X < 5/2)$

The values of X that satisfy $1/2 < X < 5/2$ are $X = 1$ and $X = 2$, so:

$$P(1/2 < X < 5/2) = P(X = 1) + P(X = 2)$$

$$= k(1^3) + k(2^3)$$

$$= \frac{1}{100}(1 + 8) = \frac{9}{100}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Step 2.2: Find $P(X > 1)$

The values of X that satisfy $X > 1$ are $X = 2, 3, 4$, so:

$$P(X > 1) = P(X = 2) + P(X = 3) + P(X = 4)$$

$$= k(2^3) + k(3^3) + k(4^3)$$

$$= \frac{1}{100}(8 + 27 + 64) = \frac{99}{100}$$

Step 2.3: Compute the Conditional Probability

$$P(1/2 < X < 5/2 \mid X > 1) = \frac{P(1/2 < X < 5/2)}{P(X > 1)}$$

$$= \frac{\frac{9}{100}}{\frac{99}{100}} = \frac{9}{99} = \frac{1}{11}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Continuous Radom Variable

Definition: A continuous random variable is a type of random variable that can take any value within a given range or interval.

Unlike discrete random variables, which can only take specific, isolated values (like the outcome of rolling a die), continuous random variables can assume an infinite number of possible values.

A random variable X is continuous if for some function $f : S \rightarrow \Re$ and for any numbers a and b with $a \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx$$

The function f has to satisfy the following two conditions:

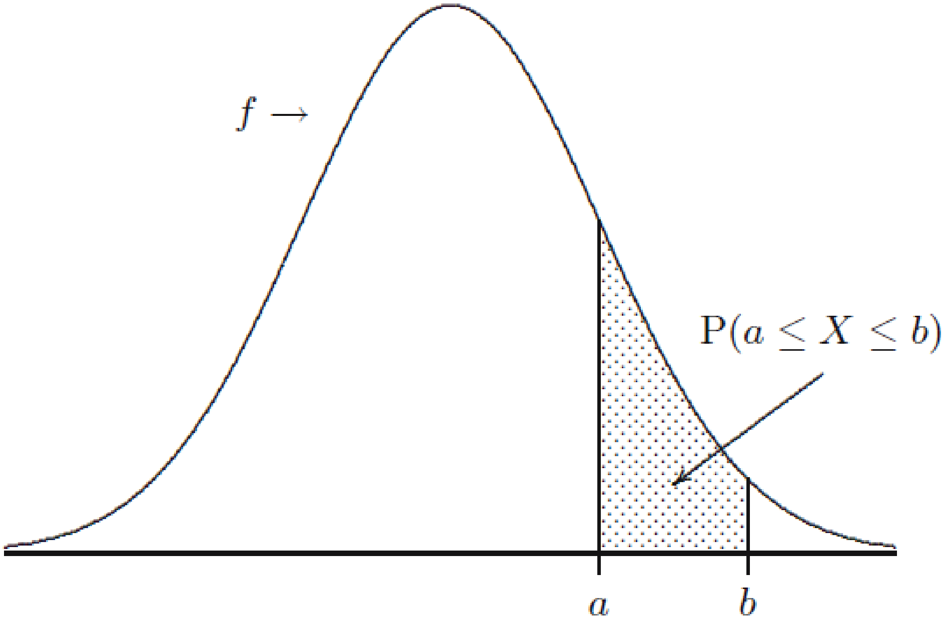
1. $f(x) \geq 0$ for all x and
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$

We call f the **probability density function (pdf)** (or **probability density**) of X.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Continuous Radom Variable



Area under a probability density function f on the interval $[a, b]$.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI
STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Example –Continuous Radom Variable

Example 7: Suppose that the error in the reaction temperature, in °C, for a controlled laboratory experiment is a continuous random variable X having the probability density function

$$f(x) = \begin{cases} \frac{x^2}{3}, & \text{for } -1 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

i. Verify that $f(x)$ is a density function.

ii. Find $P(0 < X \leq 1)$.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Solution: We use the definition for density function,

i) Obviously, $f(x) \geq 0$. To verify the next condition,

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-1}^{2} \frac{x^2}{3} dx = \frac{x^3}{9}\bigg]_{-1}^{2} = \frac{8}{9} + \frac{1}{9} = 1$$

ii) To find $P(0 < X \leq 1)$, we use the following formula.

$$P(0 < X \leq 1) = \int_{0}^{1} \frac{x^2}{3} dx = \frac{x^3}{9}\bigg]_{0}^{1} = \frac{1}{9}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Example – Continuous Radom Variable

Example 8: Show that $f(x) = \begin{cases} x, & 0 < x \le 1 \\ 2 - x, & 1 \le x \le 2 \\ 0, & x > 2 \end{cases}$ is a probability density function.

1. **Non-negativity**: $f(x) \ge 0$ for all x.
2. **Normalization**: The total area under the curve (integral of $f(x)$ over its entire range) equals 1.

Step 1: Non-negativity

- For $0 < x \le 1$:

$$f(x) = x \ge 0$$

- For $1 < x \le 2$:

$$f(x) = 2 - x \ge 0 \quad (\text{since } 2 - x \text{ is non-negative for } 1 < x \le 2)$$

- For $x > 2$:

$$f(x) = 0 \ge 0$$

Thus, $f(x)$ is non-negative for all x.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI
STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Step 2: Normalization

We need to compute the integral of $f(x)$ over the entire range:

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 f(x)\,dx + \int_1^2 f(x)\,dx$$

Calculating the integrals:

1. **For** $0 < x \leq 1$:

$$\int_0^1 f(x)\,dx = \int_0^1 x\,dx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1^2}{2} - \frac{0^2}{2} = \frac{1}{2}$$

2. **For** $1 < x \leq 2$:

$$\int_1^2 f(x)\,dx = \int_1^2 (2-x)\,dx$$

$$= \left[2x - \frac{x^2}{2}\right]_1^2 = \left(2(2) - \frac{2^2}{2}\right) - \left(2(1) - \frac{1^2}{2}\right)$$

$$= (4-2) - \left(2 - \frac{1}{2}\right) = 2 - (2 - 0.5) = 2 - 1.5 = 0.5$$

Now, adding the two results:

$$\int_{-\infty}^{\infty} f(x)\,dx = \frac{1}{2} + \frac{1}{2} = 1$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

The cumulative distribution function of a continuous random variable

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

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

$$\text{for } -\infty < x < \infty$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Example 9: Let X be a continuous random variable with P.D.F given by

$$f(x) = \begin{cases} ax, & 0 \le x < 1 \\ a, & 1 \le x \le 2 \\ -ax + 3a, & 2 \le x \le 3 \\ 0, & 3 < x \end{cases}$$

1. Determine the constant a
2. Compute P(X ≤ 1.5)

Step 1: Determine the Constant a

To ensure that $f(x)$ is a valid probability density function, it must satisfy the condition:

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

Since $f(x) = 0$ for $x > 3$, we only need to integrate over the intervals where $f(x)$ is defined:

$$\int_0^1 ax\,dx + \int_1^2 a\,dx + \int_2^3 (-ax + 3a)\,dx = 1$$

Calculating each integral:

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

1. **First Integral**:

$$\int_0^1 ax\,dx = \left[\frac{a}{2}x^2\right]_0^1 = \frac{a}{2}$$

2. **Second Integral**:

$$\int_1^2 a\,dx = a\,[x]_1^2 = a(2-1) = a$$

3. **Third Integral**:

$$\int_2^3 (-ax+3a)\,dx = \int_2^3 -ax\,dx + \int_2^3 3a\,dx$$

- For the first part:

$$\int_2^3 -ax\,dx = \left[-\frac{a}{2}x^2\right]_2^3 = -\frac{a}{2}(9-4) = -\frac{5a}{2}$$

- For the second part:

$$\int_2^3 3a\,dx = 3a\,[x]_2^3 = 3a(3-2) = 3a$$

Combining these results:

$$\int_2^3 (-ax+3a)\,dx = -\frac{5a}{2} + 3a = \frac{1a}{2}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Summing Up the Integrals

Now combine all the integrals:

$$\frac{a}{2} + a + \frac{a}{2} = 1$$

This simplifies to:

$$\frac{a}{2} + \frac{2a}{2} + \frac{a}{2} = 1 \Rightarrow 2a = 1 \Rightarrow a = \frac{1}{2}$$

Step 2: Compute $P(X \leq 1.5)$

To compute $P(X \leq 1.5)$, we split it into two intervals: from 0 to 1 and from 1 to 1.5.

$$P(X \leq 1.5) = \int_0^1 \frac{1}{2}x\,dx + \int_1^{1.5} \frac{1}{2}\,dx$$

1. $\int_0^1 \frac{1}{2}x\,dx$

$$\int_0^1 \frac{1}{2}x\,dx = \frac{1}{2}\left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

2. $\int_1^{1.5} \frac{1}{2}\,dx$

$$\int_1^{1.5} \frac{1}{2}\,dx = \frac{1}{2} \times (1.5 - 1) = \frac{1}{2} \times 0.5 = \frac{1}{4}$$

Total probability:

$$P(X \leq 1.5) = \frac{1}{4} + \frac{1}{4} = \frac{2}{4} = \frac{1}{2}$$

Final Answer:

- The constant $a = \frac{1}{2}$.
- $P(X \leq 1.5) = \frac{1}{2}$.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Some practice questions

Ques) A continuous random variable has the pdf f (x) = kx^4, -1 < x < 0.

i. Solve to find the value of k, and
ii. Also solve to find $P(X > -\frac{1}{2} / X < -\frac{1}{4})$.

Ques) A continuous random variable has the pdf $f(x) = kx^2(1-x), \text{ for } 0 \le x \le 1.$

i. Solve to find the value of k. k=12
ii. Solve to find P(0.2 < X < 0.8 / X>0.5). 0.599.

Ques) The pdf of a continuous random variable X is f(X) = k (2−x), for 0 ≤ x ≤ 2

i. Solve to find the value of k. k=1/2
ii. Solve to find P(X>1.5 / X<1.8)P(X>1.5 / X<1.8). 0.05303.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Some practice questions

Ques) A continuous random variable has the pdf f (X) = kx^4, -1 < x < 0.

i. Solve to find the value of k, and

ii. Also solve to find $P(X > -\frac{1}{2} / X < -\frac{1}{4})$.

Step 1: Find the value of k

The pdf $f(X) = kx^4$ is defined for the interval $-1 < x < 0$. For this function to be a valid probability density function, it must satisfy the normalization condition:

$$\int_{-1}^{0} f(x)\, dx = 1$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

$$\int x^4\,dx = \frac{x^5}{5}$$

Now, evaluating the definite integral:

$$\int_{-1}^{0} x^4\,dx = \left[\frac{x^5}{5}\right]_{-1}^{0} = \left(\frac{0^5}{5}\right) - \left(\frac{(-1)^5}{5}\right) = 0 - \left(-\frac{1}{5}\right) = \frac{1}{5}$$

So, we have:

$$\int_{-1}^{0} kx^4\,dx = k \cdot \frac{1}{5} = 1$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

$$k = 5$$

Step 2: Find $P(X > -\frac{1}{2} \mid X < -\frac{1}{4})$

To find this conditional probability, we can use the formula:

$$P(X > a \mid X < b) = \frac{P(X > a \cap X < b)}{P(X < b)}$$

1. **Calculate** $P(X < -\frac{1}{4})$

$$P(X < -\frac{1}{4}) = \int_{-1}^{-\frac{1}{4}} f(x)\,dx = \int_{-1}^{-\frac{1}{4}} 5x^4\,dx$$

Calculating the integral:

$$= 5\left[\frac{x^5}{5}\right]_{-1}^{-\frac{1}{4}} = [x^5]_{-1}^{-\frac{1}{4}} = \left(-\frac{1}{4}\right)^5 - (-1)^5$$

Calculating:

$$= -\frac{1}{1024} + 1 = 1 - \frac{1}{1024} = \frac{1024-1}{1024} = \frac{1023}{1024}$$

2. **Calculate** $P(X > -\frac{1}{2} \cap X < -\frac{1}{4})$

$$P(X > -\frac{1}{2} \cap X < -\frac{1}{4}) = \int_{-\frac{1}{2}}^{-\frac{1}{4}} f(x)\,dx = \int_{-\frac{1}{2}}^{-\frac{1}{4}} 5x^4\,dx$$

Calculating the integral:

$$= 5\left[\frac{x^5}{5}\right]_{-\frac{1}{2}}^{-\frac{1}{4}} = [x^5]_{-\frac{1}{2}}^{-\frac{1}{4}} = \left(-\frac{1}{4}\right)^5 - \left(-\frac{1}{2}\right)^5$$

Calculating:

$$= -\frac{1}{1024} - \left(-\frac{1}{32}\right) = -\frac{1}{1024} + \frac{1}{32}$$

Converting $\frac{1}{32}$ to a denominator of 1024:

$$= -\frac{1}{1024} + \frac{32}{1024} = \frac{31}{1024}$$

3. **Calculate the conditional probability**:

Now, substituting back into the conditional probability formula:

$$P(X > -\frac{1}{2} \mid X < -\frac{1}{4}) = \frac{P(X > -\frac{1}{2} \cap X < -\frac{1}{4})}{P(X < -\frac{1}{4})} = \frac{\frac{31}{1024}}{\frac{1023}{1024}} = \frac{31}{1023}$$

Final Answers:

- The value of k is **5**.
- The conditional probability $P(X > -\frac{1}{2} \mid X < -\frac{1}{4})$ is $\frac{31}{1023}$.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Some practice questions

Ques) A continuous random variable has the pdf f (X) = kx^4, -1 < x < 0.

i. Solve to find the value of k, and Ans) k=5

ii. Also solve to find $P(X > -\frac{1}{2} / X < -\frac{1}{4})$. Ans) 31/1023

Ques) A continuous random variable has the pdf f(X) $= kx^2(1-x)$, for $0 \leq x \leq 1$.

i. Solve to find the value of *k*.

ii. Solve to find P(0.2 < X < 0.8 / X>0.5)

Ques) The pdf of a continuous random variable X is f(X) = k (2−x), for 0 ≤ x ≤ 2

i. Solve to find the value of k.

ii. Solve to find P(X>1.5 / X<1.8)P(X>1.5 / X<1.8).

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Joint Probability Distribution

Definition: A joint probability distribution describes the probability of two or more random variables occurring simultaneously. For two random variables X and Y, the joint probability distribution specifies the probability that X takes a specific value and Y takes a specific value at the same time.

1. Joint Probability Mass Function (for Discrete Random Variables)

For **discrete** random variables, the **joint probability mass function** (PMF) $P(X = x, Y = y)$ gives the probability that X is equal to x and Y is equal to y. It satisfies the following conditions:

- $P(X = x, Y = y) \geq 0$ for all values of x and y,
- The sum of all joint probabilities is 1:

$$\sum_x \sum_y P(X = x, Y = y) = 1.$$

2. Joint Probability Density Function (for Continuous Random Variables)

For **continuous** random variables, the **joint probability density function (pdf)** $f(x, y)$ represents the probability density at a specific pair (x, y). The probability that X and Y fall within certain intervals is found by integrating the joint pdf over those intervals.

- The joint pdf must satisfy:
  - $f(x, y) \geq 0$,
  - The total probability over the entire space is 1:

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

Example (Joint pdf):

Let X and Y have the joint pdf:

$$f(x, y) = 6(1 - x - y), \quad 0 < x < 1, 0 < y < 1.$$

This means the probability of X and Y both lying in a region is the integral of this function over that region.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Example – Joint Probability Calculation

Example 9: Let X and Y be continuous random variables with P.D.F given by

$$f(x,y) = \begin{cases} \frac{2}{5}(2x+3y), & 0 \le x \le 1, 0 \le y \le 1, \\ 0\,, & \text{elsewhere} \end{cases}$$

i. Verify for the joint probability density function total probability sum is 1.

ii. Find P{(x , y) | 0 < x < $\frac{1}{2}$, $\frac{1}{4}$ < y < $\frac{1}{2}$}.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Example – Joint Probability Calculation

Step 1: Verify that the total probability is 1

We need to compute the integral of the pdf over the region where it is defined:

$$\int_0^1 \int_0^1 f(x,y)\,dy\,dx = \int_0^1 \int_0^1 \frac{2}{5}(2x+3y)\,dy\,dx$$

Integrate with respect to y:

$$\int_0^1 (2x+3y)\,dy = \left[2xy + \frac{3y^2}{2}\right]_0^1 = 2x(1) + \frac{3(1)^2}{2} - 0 = 2x + \frac{3}{2}$$

Now, substituting this result back into the outer integral:

$$\int_0^1 \frac{2}{5}(2x + \frac{3}{2})\,dx = \frac{2}{5}\left[x^2 + \frac{3}{2}x\right]_0^1$$

Calculating this:

$$= \frac{2}{5}\left[1^2 + \frac{3}{2}(1) - 0\right] = \frac{2}{5}\left(1 + \frac{3}{2}\right) = \frac{2}{5} \cdot \frac{5}{2} = 1$$

Thus, the total probability integrates to 1, verifying that $f(x,y)$ is a valid joint pdf.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Step 2: Find $P(0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2})$

We now compute the probability for the specified ranges:

$$P\left(0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\right) = \int_0^{1/2} \int_{1/4}^{1/2} \frac{2}{5}(2x + 3y)\, dy\, dx$$

Integrate with respect to y:

$$\int_{1/4}^{1/2} (2x + 3y)\, dy = \left[2xy + \frac{3y^2}{2}\right]_{1/4}^{1/2}$$

Calculating this:

$$= \left[2x \cdot \frac{1}{2} + \frac{3}{2} \cdot \left(\frac{1}{2}\right)^2\right] - \left[2x \cdot \frac{1}{4} + \frac{3}{2} \cdot \left(\frac{1}{4}\right)^2\right]$$

$$= \left[x + \frac{3}{8}\right] - \left[\frac{x}{2} + \frac{3}{32}\right]$$

$$= x + \frac{3}{8} - \frac{x}{2} - \frac{3}{32} = \frac{x}{2} + \frac{12}{32} - \frac{3}{32} = \frac{x}{2} + \frac{9}{32}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Now, substitute this back into the integral with respect to x:

$$P = \int_0^{1/2} \frac{2}{5} \left(\frac{x}{2} + \frac{9}{32} \right) dx = \frac{2}{5} \left[\frac{x^2}{4} + \frac{9}{32} x \right]_0^{1/2}$$

Calculating this:

$$= \frac{2}{5} \left[\frac{(1/2)^2}{4} + \frac{9}{32} \cdot \frac{1}{2} \right] = \frac{2}{5} \left[\frac{1/16}{4} + \frac{9}{64} \right] = \frac{2}{5} \left[\frac{1}{64} + \frac{9}{64} \right]$$

$$= \frac{2}{5} \cdot \frac{10}{64} = \frac{2}{5} \cdot \frac{5}{32} = \frac{2}{32} = \frac{1}{16}$$

Final Result:

- The total probability confirms that the joint pdf is valid.
- The calculated probability $P(0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2})$ is $\frac{1}{16}$.

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Example – Joint Probability Calculation

Example 9: The joint pdf of a two-dimensional random variable (X, Y) is given by

$$f(x, y) = \begin{cases} k(6 - x - y), 0 < x < 2, 2 < y < 4 \\ 0, \text{elsewhere} \end{cases}$$

i. Solve to find P (X < 1, Y < 3) and
ii. Solve to find P (X < 1/Y < 3)

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Step 1: Find k

We are given the joint pdf:

$$f(x, y) = k(6 - x - y), \quad 0 < x < 2, \quad 2 < y < 4$$

To find k, we use the total probability condition:

$$\int_2^4 \int_0^2 k(6 - x - y)\, dx\, dy = 1$$

Inner Integral:

$$\int_0^2 (6 - x - y) dx$$

Computing:

$$\left[6x - \frac{x^2}{2} - xy\right]_0^2$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

$$= \left(6(2) - \frac{2^2}{2} - 2y\right) - (0)$$

$$= (12 - 2 - 2y) = (10 - 2y)$$

Outer Integral:

$$\int_2^4 k(10 - 2y)dy$$

Computing:

$$k\left[10y - y^2\right]_2^4$$

$$= k\left((10(4) - 4^2) - (10(2) - 2^2)\right)$$

$$= k\left((40 - 16) - (20 - 4)\right)$$

$$= k(24 - 16) = k(8)$$

Setting this equal to 1:

$$8k = 1 \Rightarrow k = \frac{1}{8}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Step 2: Compute $P(X < 1, Y < 3)$

$$P(X < 1, Y < 3) = \int_2^3 \int_0^1 f(x,y)dxdy$$

Substituting $f(x,y) = \frac{1}{8}(6 - x - y)$:

$$\int_2^3 \int_0^1 \frac{1}{8}(6 - x - y)dxdy$$

Computing the inner integral:

$$\frac{1}{8}\int_0^1 (6 - x - y)dx$$

$$= \frac{1}{8}\left[6x - \frac{x^2}{2} - xy\right]_0^1$$

$$= \frac{1}{8}\left((6(1) - \frac{1^2}{2} - 1y) - (0)\right)$$

$$= \frac{1}{8}\left(6 - \frac{1}{2} - y\right)$$

$$= \frac{1}{8}\left(\frac{12}{2} - \frac{1}{2} - y\right)$$

$$= \frac{1}{8}\left(\frac{11}{2} - y\right)$$

Now, integrating over y:

$$\frac{1}{8}\int_2^3 \left(\frac{11}{2} - y\right)dy$$

$$= \frac{1}{8}\left[\frac{11}{2}y - \frac{y^2}{2}\right]_2^3$$

$$= \frac{1}{8}\left(\left(\frac{11}{2}(3) - \frac{3^2}{2}\right) - \left(\frac{11}{2}(2) - \frac{2^2}{2}\right)\right)$$

$$= \frac{1}{8}\left(\left(\frac{33}{2} - \frac{9}{2}\right) - \left(\frac{22}{2} - \frac{4}{2}\right)\right)$$

$$= \frac{1}{8}\left(\frac{24}{2} - \frac{18}{2}\right)$$

$$= \frac{1}{8}\left(\frac{6}{2}\right) = \frac{3}{8}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

Step 3: Compute $P(X < 1 \mid Y < 3)$

$$P(X < 1 \mid Y < 3) = \frac{P(X < 1, Y < 3}{P(Y < 3)}$$

We already found $P(X < 1, Y < 3) = \frac{3}{8}$.

Now, compute $P(Y < 3)$:

$$P(Y < 3) = \int_2^3 P(X \text{ over its range})d$$

$$= \int_2^3 \int_0^2 f(x, y)dxdy$$

$$= \int_2^3 \int_0^2 \frac{1}{8}(6 - x - y)dxdy$$

Computing the inner integral:

$$\frac{1}{8}\int_0^2 (6 - x - y)dx$$

$$= \frac{1}{8}\left[6x - \frac{x^2}{2} - xy\right]_0^2$$

$$= \frac{1}{8}\left((6(2) - \frac{2^2}{2} - 2y) - (0)\right)$$

$$= \frac{1}{8}(12 - 2 - 2y)$$

$$= \frac{1}{8}(10 - 2y)$$

Now, integrating over y:

$$\frac{1}{8}\int_2^3 (10 - 2y)dy$$

$$= \frac{1}{8}\left[10y - y^2\right]_2^3$$

$$= \frac{1}{8}\left((10(3) - 3^2) - (10(2) - 2^2)\right)$$

$$= \frac{1}{8}\left((30 - 9) - (20 - 4)\right)$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

$$= \frac{1}{8}(21 - 16) = \frac{5}{8}$$

Thus,

$$P(X < 1 \mid Y < 3) = \frac{\frac{3}{8}}{\frac{5}{8}} = \frac{3}{5}$$

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI

STATISTICAL METHOD NOTES BY: PROF. POOJA BHAKUNI