M Tech(Data Science & Engineering)
Introduction to Statistical Methods [ISM]

BITS Pilani

Pilani|Dubai|Goa|Hyderabad

innovate achieve lead

BITS Pilani

Session No 4

Random Variables – Discrete & Continuous (Multi variates)

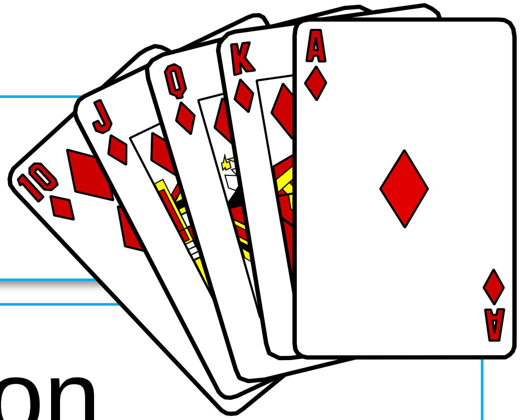
(Session 4: 27th /28th Nov 2021)

Session No 4 Course Handout

Contact Session	List of Topic Title	Reference
CS - 4	Random Variables – Discrete & Continuous (Multi variates)	T1:Chapter 3 & 4
HW	Problems on Joint RVs	T1:Chapter 3 & 4
Lab		

Probability Distribution ➡ ???

Here's what you will learn in the entire Session:



1. Discrete Joint Probability Distribution
2. Continuous Joint Probability Distribution
3. Joint Probability Distribution Function
4. Conditional Probability Distribution Function

Joint Probability Distribution

Probability Distribution → ?????

Example 1: Two balls are selected at random from a bag containing three green, two blue and four red balls.

If X and Y are respectively the numbers of green and blue balls included among the two balls drawn from the bag, find the probabilities associated with all possible pairs of value of X and Y.

Solution: Here the possible pairs are (0, 0), (0, 1), (1, 0), (1, 1), (0, 2), (2, 0).

Probability Distribution → ?????

To obtain the probability associated with (1, 0), for example, we see that we are dealing with the event of getting one of the three green balls, no blue ball and hence, one of the red ball is the number of ways in which we get this event = ${}^{3}c_{1} \times {}^{2}c_{0} \times {}^{4}c_{1} = 12$. Also, the total number of ways in which two ball are drawn out of nine = ${}^{9}c_{2} = 36$

Probability Distribution → ????

The above probabilities can also be obtained by

$$P(x, y) = \frac{{}^{3}c_{x}\,{}^{2}c_{y}\,{}^{4}c_{2-x-y}}{{}^{9}c_{2}}, \text{ for } x = y = 0, 1, 2 \text{ and } 0 \le x + y \le 2$$

x	y		
	0	1	2
0	1/6	2/9	1/36
1	1/3	1/6	0
2	1/12	0	0

To obtain the probability associated with (1, 0), we have $\boxed{{}^{3}c_{1} \times {}^{2}c_{0} \times {}^{4}c_{1} = 12.}$

Total number of ways in which two ball are drawn out of nine = $\boxed{{}^{9}c_{2} = 36.}$

As these probabilities are equally likely, the probability of the event associated with (1, 0) is 12/36 =1/3.

Probability Distribution ➡ ?????

$$P(x,y)=\frac{{}^{3}c_{x}\,{}^{2}c_{y}\,{}^{4}c_{2-x-y}}{{}^{9}c_{2}},\ \text{for}\ \ x=y=0,1,2\ \text{and}\ 0\le x+y\le 2$$

The distribution of both (X, Y) which occur jointly is called Joint Discrete Probability Distribution.

P(x, y) is called joint probability mass function (jpmf)

If X and Y are two discrete random variables, the probability distribution of their simultaneous occurrences can be represented by a function with values p(x, y) for each pair of values (x, y) within the range of X and Y. This probability distribution is called the joint probability mass function given by p(x, y) = P(X=x, Y=y).

Probability Distribution ⇒ Joint Probability Distribution

The function p(x, y) gives the probability that the outcomes x and y occur simultaneously.

The p(x, y) is called joint probability mass function if it satisfies the following conditions:

(i) p(x, y) ≥ 0, for each pair of values (x, y) within its domain.

(ii) $$\sum_x \sum_y p(x, y) = 1$$

where the summation extends over all possible pair of values (x, y) within its domain.

Joint Probability Distribution Function

If X and Y are random variables, the function given by

$$\mathbf{F(x,y) = P(X \le x, Y \le y) = \sum_{s \le x} \sum_{t \le y} p(x,y)}$$

Probability Distribution → Joint Probability Distribution

If X and Y are discrete random variables, the joint probability distribution function is given by

$$\mathbf{F(x,y) = P(X \leq x, Y \leq y) = \sum_{s \leq x} \sum_{t \leq y} p(x,y)}$$

X	Y		
	0	1	2
0	1/6	2/9	1/36
1	1/3	1/6	0
2	1/12	0	0

Find P(X ≤ 1, Y ≤1)

P(X=0, Y=0) + P(X=0, Y=1) +

P(X=1, Y=0) + P(X=1, Y=1)

P(X ≤ 1, Y ≤1) = 16/18

Discrete Marginal Distributions

If X and Y are random variables, then the marginal distribution of X=x is given by

$$g(\mathrm{x}) = \sum_{y} p(x, y)$$

If X and Y are random variables, then the marginal distribution of Y=y is given by

$$h(\mathrm{y}) = \sum_{x} p(x, y)$$

Probability Distribution ⇨ Marginal probability distribution

Find the marginal probability mass function of X and Y

X	Y		
	0	1	2
0	1/6	2/9	1/36
1	1/3	1/6	0
2	1/12	0	0

X	Y			g(x)
	0	1	2	
0	1/6	2/9	1/36	15/36
1	1/3	1/6	0	3/6
2	1/12	0	0	1/12
h(y)	7/12	7/18	1/36	1

$$\sum_{x=0}^{2} g(x) = \sum_{y=0}^{2} h(y) = 1$$

Summarizing the results:

The marginal distribution of X=x is

x	0	1	2
g(x)	7/12	7/18	1/36

The marginal distribution of Y=y is

h	0	1	2
h(y)	15/36	3/6	1/12

Example 2: If the joint probability mass function of two discrete random variables is given below, obtain the marginal distribution of X and Y. Also find P(X=1/Y=0) & P(Y=1/X=1)

X	Y		
	0	1	2
0	3/28	3/14	1/28
1	9/28	3/14	0
2	3/28	0	0

Marginal distributions of g(X=x) and h(Y = y)

X	Y			g(x)
	0	1	3	
0	3/28	3/14	1/28	5/14
1	9/28	3/14	0	15/28
2	3/28	0	0	3/28
h(y)	15/28	3/7	1/28	1

Summarizing the results:

The marginal distribution of X=x is

x	0	1	2
g(x)	15/28	3/7	1/28

The marginal distribution of Y=y is

h	0	1	2
h(y)	5/14	15/28	3/28

Conditional Distributions

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ P(B)>0}$$

If X and Y are random variables, then the conditional distribution of X=x given Y=y is

$$P(X = x/Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x,y)}{h(y)}, \mathrm{h(y)} > 0$$

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

$$P(Y = y/X = x) = \frac{P(X = x, Y = y)}{P(X = x)} = \frac{p(x,y)}{g(x)}, \mathrm{g(x)} > 0$$

Probability Distribution ⇨ Conditional probability distribution

Find $P(X=2 \mid Y=0)$ and $P(Y=1 \mid X=1)$

y	X		
	0	1	2
0	1/6	1/3	1/12
1	2/9	1/6	0
2	1/36	0	0

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \ P(B)>0$$

Find P(X=2 | Y=0) and P(Y=1 | X=1)

X	Y			g(x)
	0	1	2	
0	1/6	2/9	1/36	15/36
1	1/3	1/6	0	3/6
2	1/12	0	0	1/12
h(y)	7/12	7/18	1/36	1

$$P(X=2 \mid Y=0)=\frac{P(X=2,Y=0)}{P(Y=0)}=\frac{\frac{1}{12}}{\frac{7}{12}}=\frac{1}{7}$$

$$P(Y=1 \mid X=1)=\frac{P(X=1,Y=1)}{P(X=1)}=\frac{\frac{1}{6}}{\frac{3}{6}}=\frac{1}{3}$$

$$P(Y=2 \mid X=0)=\frac{P(X=2,Y=0)}{P(Y=0)}=\frac{\frac{1}{36}}{\frac{15}{36}}=\frac{1}{15}$$

Example 2: **If the joint probability mass function of two discrete random variables is given below, obtain the marginal distribution of X and Y. Also find P(X=1/Y=0) & P(Y=1/X=1)**

X	Y		
	0	1	2
0	3/28	3/14	1/28
1	9/28	3/14	0
2	3/28	0	0

$$P(X=1 \mid Y=0) = \frac{P(X=1, Y=0)}{P(X=1)} = \frac{\frac{9}{28}}{\frac{15}{28}} = \frac{3}{5}$$

$$P(Y=2 \mid X=0) = \frac{P(X=0, Y=2)}{P(Y=0)} = \frac{\frac{1}{28}}{\frac{6}{28}} = \frac{1}{6}$$

Example 3: If the joint probability mass function of two discrete random variables is given below:

X	Y		
	0	1	2
-1	1/15	3/15	2/15
0	2/15	2/15	1/15
1	1/15	1/15	2/15

$P(X = -1 \mid Y = 2) = 2/5$

$P(X = 0 \mid Y = 2) = 1/5$

$P(X = 1 \mid Y = 2) = 2/5$

Find the conditional distribution of X given Y = 2

Example

The probability distribution of random variables X and Y is given by

Independence of random variables

If X and Y are two random variables, with joint probability distributions p(x, y) and the marginal distributions g(x) and h(y) respectively, then they are called as statistically independent if and only if

$$p(x, y) = g(x)h(y) \text{ for all } (x, y) \text{ within their range.}$$

Note: X and Y may be discrete or continuous.

Let X and Y be continuous random variables. A joint Probability density function $f(x, y)$ for these two variables is a function satisfying $f(x, y) \geq 0$

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

Also,

$$\mathbf{P(a \leq X \leq b, c \leq Y \leq d) = \int_{a}^{b}\int_{c}^{d} f(x, y)\, dy\, dx}$$

Marginal densities of continuous random variables

Let f(x, y) be the joint density function of the continuous random variables X and Y, then the marginal densities of X and Y will be obtained by integrating over the range of Y for X and over the range of X for Y, i.e.,

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

The joint probability distribution function of two continuous random variables (X, Y) is given f(x, y) = P(a ≤ x ≤ b, c ≤ y ≤ d)

The joint cumulative probability distribution function of two continuous random variables (X, Y) is given

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(x, y)\,dx\,dy$$

Two continuous random variables X and Y are said to be **independent** if for every pair of x and y values

$$f(x, y) = f_X(x) \cdot f_Y(y)$$

If it is not satisfied, then X and Y are said to be dependent.

Conditional Distributions

Let X and Y be two continuous rv's with joint pdf $f(x, y)$ and marginal probability distribution of X is $f_X(x)$.

Then for any X value of x for which $f_X(x) > 0$, the conditional probability density function of Y for given X is $f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$

Distribution Functions

Let (X, Y) be a two dimensional random variable then their joint distribution function is given by:

In case of continuous rv's

- $\mathbf{F_{XY}(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y) dxdy}$

In case of discrete rv's

- $\mathbf{F_{XY}(x, y) = P(X \leq x, Y \leq y) = \sum_{X \leq x} \sum_{Y \leq y} p(x, y)}$

Marginal density functions

In case of continuous rv's

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

In case of discrete rv's

- $$F_X(x) = \sum_{X \le x, \forall y} P(X \le x, Y = y)$$
- $$F_Y(y) = \sum_{Y \le y, \forall x} P(X = x, Y \le y)$$

Joint probability mass functions of n random variables

Let $X_1, X_2, . \quad . \quad ., X_n$, be n discrete random variables which are assuming the values $x_1, x_2, . \quad . \quad ., x_n$. The joint probability mass function of n discrete random variables is given by

$P(x_1, x_2, . \quad . \quad ., x_n) = P(X_1{=}x_1, X_2{=}x_2, . \quad . \quad ., X_n{=}x_n)$, which satisfies the properties

- (i) $P(x_1, x_2, . \quad . \quad ., x_n) \geq 0$
- (ii) $\sum\sum. \; . \; . \sum P(x_1, x_2, . \quad . \quad ., xn) = 1$

Joint probability density functions of n random variables

Let $X_1, X_2, . \quad . \quad ., X_n$, be n continuous random variables which are assuming the values $x_1, x_2, . \quad . \quad ., x_n$. The joint probability density function of n discrete random variables is given by

$$P(a_1 \le x_1 \le b_1, a_1 \le x_2 \le b_1, . \quad . \quad ., a_1 \le x_n \le b_1) =$$

$$\int_{a_1}^{b_1} \int_{a2}^{b_2} . \quad . \quad ., \int_{a_n}^{b_n} f(x_1, x_2, . \quad . \quad ., xn) dx_1 xd_2 . \quad . \quad . dx_n$$

Joint probability density functions of n random variables

which satisfies the properties

(i) $f(x_1, x_2, . \quad . \quad ., x_n) \geq 0$

(ii) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}. \quad . \quad ., \int_{-\infty}^{\infty} f(x_1, x_2, . \quad . \quad ., xn)dx_1xd_2. \quad . \quad .dx_n = 1$

Joint probability distribution functions of n random variables

Let F $(x_1, x_2, . \quad . \quad ., x_n)$ denote the distribution function of n continuous random variables, then

$$\int_{-\infty}^{x_1}\int_{-\infty}^{x_2} . \quad . \quad ., \int_{-\infty}^{x_n} f(x_1, x_2, . \quad . \quad ., xn)\mathrm{dx}_1\mathrm{xd}_2. \quad . \quad .\mathrm{dx}_n$$

Expected value

Let X and Y be jointly distributed rv's with pmf $p\,(x, y)$ or pdf $f\,(x, y)$ according to whether the variables are discrete or continuous. Then the expected value of a function $h\,(X, Y)$, where $h\,(X, Y)$ is a random variable is denoted by $E\,[h(X, Y)]$ or $\mu_{h(x,y)}$ and it given by

$$E[h(X,Y)] = \begin{cases} \sum_{x}\sum_{y} h(x,y)\cdot p(x,y), if\ X, Y\ are\ discrete \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x,y)\cdot f(x,y)dxdy, if\ X\ and\ Y\ are\ continuous \end{cases}$$

Covariance

- Cov(X, Y)= E(XY) - E(X). E(Y) where
- $E(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$
- $E(Y) = \int_{-\infty}^{\infty} y\, f_Y(y)\, dy$
- $E(XY) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f(x, y)\, dx\mathrm{dy}$
- $V(X) = E(X^2) - [E(X)]^2$
- $V(Y) = E(Y^2) - [E(Y)]^2$

Example

Example :

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

(i) Find P[X≤2/3, Y≤1/4]?

innovate achieve lead

Example

Solution

Exercise 1

$$f(x) = \begin{cases} kx(2x+3y), & 0 < x < 1; 0 < y < 1 \\ 0, & elsewhere \end{cases}$$

(a) Find the value of k

(b) Find the marginal probability density function of X and Y

(c) E(X), E(Y), V(X) and V(Y)

Exercise 2

A large insurance agency services a number of customers who have purchased both a homeowner's policy and an automobile policy from the agency. For each type of policy, a deductible amount must be specified. For an automobile policy, the choices are $100 and $250, whereas for a homeowner's policy, the choices are 0, $100, and $200. Suppose an individual with both types of policy is selected at random from the agency's files. Let X the deductible amount on the auto policy and Y the deductible amount on the homeowner's policy

Suppose the joint probability mass function is given by

X	Y		
	0	100	200
100	0.2	0.1	0.2
250	0.05	0.15	0.3

Find

(i) Marginal probabilities of X and Y.

(ii) $\mathrm{P}(Y \geq 100)$

(iii) Are X and Y independent ?

(iv) Conditional distribution of Y=100 for X=x.

Exercise 3

(a) If X and Y are independent with $p_x(0) = 0.5$,

$p_x(1) = 0.3$, $p_x(0) = 0.2$ and $p_y(0) = 0.6$, $p_y(1) = 0.1$,

$p_y(2) = p_y(3) = 0.05$ and $p_y(4) = 0.2$

Display the joint pmf of (X, Y).

b) Compute $P(X \leq 1 \; and \; Y \leq 1)$ from joint pmf and verify that it equals the product $P(X \leq 1) \cdot P(Y \leq 1)$.

c) What is $P(X + Y = 0)$

d) Compute $P(X + Y \leq 1)$

Exercise 4

X and Y are two rv's having the joint pmf

$$f(x, y) = \frac{1}{27}(2x + y),$$

where x and y can assume only the integer values 0,1 and 2. Find the conditional distribution of Y for X=x.

Exercise 5

A bank operates both a drive-up facility and a walk-up window. On a randomly selected day, let X the proportion of time that the drive-up facility is in use (at least one customer is being served or waiting to be served) and Y the proportion of time that the walk-up window is in use. Suppose the joint pdf of (X, Y) is given by

$$f(x,y)=\begin{cases}\frac{6}{5}(x+y^2), & 0\le x\le 1, 0\le y\le 1\\ 0, & otherwise\end{cases}$$

Verify

a) Is this legitimate pdf

b) The probability that neither facility is busy more than one-quarter of the time.

c) The marginal pdf of *X*, which gives the probability distribution of busy time for the drive-up facility without reference to the walk-up window

Exercise 6

If X and Y are two random variables having joint density function

$$f(x,y) = \begin{cases} \frac{1}{8}(6-x-y), 0 \leq x \leq 2, 2 \leq y \leq 4 \\ 0, \quad otherwise \end{cases}$$

Find (i) P(X<1 and Y<3)

(ii) P(X+Y<3)

(iii) P(X<1|Y<3)

Exercise 7

Each front tire on a particular type of vehicle is supposed to be filled to a pressure of 26 psi. Suppose the actual air pressure in each tire is a random variable—*X* for the right tire and *Y* for the left tire, with joint pdf

$$f(x,y)=\begin{cases}k(x^2+y^2), & 20\le x\le 30, 20\le y\le 30\\ 0, & otherwise\end{cases}$$

a) What is the value of K?

b) What is the probability that both tires are under filled?

c) What is the probability that the difference in air pressure between the two tires is at most 2 psi?

d) Determine the (marginal) distribution of air pressure in the right tire alone.

e) Are X and Y independent rv's?

Exercise 8

If the joint probability density of two continuous random variables X_1 and X_2 is given by

$$f\,(x_1, x_2) = \begin{cases} 6e^{-2x_1-3x_2}, & \text{for } x_1 > 0 \text{ and } x_2 > 0 \\ 0, & \text{Elsewhere} \end{cases}$$

Find the joint distribution function of two random variables and use it to find out the probability that both random variables will take on values less than 1.

Exercise 9

- Five friends have purchased tickets to a certain concert. If the tickets are for seats 1–5 in a particular row and the tickets are randomly distributed among the five, what is the expected number of seats separating any particular two of the five?
- Let X and Y denote the seat numbers of the first and second individuals, respectively. Possible (X, Y) pairs are $\{(1, 2), (1, 3), \ldots, (5, 4)\}$ and the joint pmf of (X, Y) is

Exercise 9 (contd)

$$p(x,y) = \begin{cases} \frac{1}{20}, & x = 1 \; to \; 5, y = 1 \; to \; 5, x \neq y \\ 0, & otherwise \end{cases}$$

The number of seats separating the two individuals is $h(X,Y) = |X - Y| - 1$.

Exercise 10

Two rv's X and Y have the following joint pdf

$$f(x,y) = \begin{cases} 2 - x - y, 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, \quad otherwise \end{cases}$$

Find Co-variance between X and Y

Thanks