R.M.K
GROUP OF ENGINEERING INSTITUTIONS

R.M.K
GROUP OF
INSTITUTIONS

R.M.K
GROUP OF
INSTITUTIONS

Please read this disclaimer before proceeding:

22MA401

PROBABILITY AND STATISTICS

DEPARTMENT	**COMPUTER SCIENCE AND ENGINEERING**
BATCH/YEAR	**2022-2026/ II**
CREATED BY	**Mrs. A. K. ARULMOZHI, Dr. S . VIDHYA**
DATE	**10.01.2024**

Table of content

Course Objectives

S. No.	**Course Objectives**
1	To Provide the necessary basic concepts of random variables and to introduce some standard distributions.
2	To introduce the basic concepts of two dimensional random variables
3	To test the hypothesis for small and large samples.
4	To introduce the concepts of Analysis of Variances.
5	To understand the concept of statistical quality control.

PREREQUISITES

S.No.	TOPICS	COURSE NAME WITH CODE
1	Differentiation & Integration	Higher Secondary level
2	Basic Probabilities	
3	Knowledge in set theory	

Syllabus

22MA401 PROBABILITY AND STATISTICS (Theory Course with Laboratory Component) — **L T P C 3 2 0 4**

UNIT I ONE DIMENSIONAL RANDOM VARIABLES 15

Basic probability definitions- Independent events- Conditional probability (revisit) - Random variable - Discrete and continuous random variables – Moments – Moment generating functions – Binomial, Poisson, Geometric, Uniform, Exponential and Normal distributions.

Experiments using R Programming:

1. Finding conditional probability.
2. Finding mean, variance and standard deviation.

UNIT II TWO DIMENSIONAL RANDOM VARIABLES 15

Joint distributions – Marginal and conditional distributions – Covariance – Correlation and linear regression – Transformation of random variables.

Experiments using R Programming:

1. Finding marginal density functions for discrete random variables
2. Calculating correlation and regression

UNIT III TESTING OF HYPOTHESIS 15

Sampling distributions - Estimation of parameters - Statistical hypothesis - Large sample tests based on Normal distribution for single mean and difference of means -Tests based on t and F distributions for mean and variance – Chisquare - Contingency table (test for independence) - Goodness of fit.

Experiments using R Programming:

1. Testing of hypothesis for given data using Z – test.
2. Testing of hypothesis for given data using t – test.

UNIT IV DESIGN OF EXPERIMENTS 15

One way and Two way classifications - Completely randomized design – Randomized block design – Latin square design.

Experiments using R Programming:

1. Perform one- way ANOVA test for the given data.
2. Perform two-way ANOVA test for the given data.

UNIT V STATISTICAL QUALITY CONTROL 15

Control charts for measurements (X and R charts) – Control charts for attributes (p, c and np charts) – Tolerance limits.

Experiments using R Programming:

1. Interpret the results for $\overline{X}$-Chart for variable data
2. Interpret the results for R-Chart for variable data

TOTAL: 75 PERIODS

Course Outcomes

CO's	Course Outcomes	Highest Cognitive Level
CO1	Calculate the statistical measures of standard distributions.	K2
CO2	Compute the correlation and regression for two dimensional random variables.	K2
CO3	Apply the concept of testing the hypothesis.	K2
CO4	Implement the concept of analysis of variance for various experimental designs.	K3
CO5	Demonstrate the control charts for variables and attributes.	K2

CCO-PO/CO-PSO Mapping

CO's	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12
CO1	3	2	-	-	-	-	-	-	-	-	-	-
CO2	3	2	-	-	-	-	-	-	-	-	-	-
CO3	3	2	-	-	-	-	-	-	-	-	-	-
CO4	3	2	-	-	-	-	-	-	-	-	-	-
CO5	3	2	-	-	-	-	-	-	-	-	-	-

CO's	PSO1	PSO2	PSO3
CO1	-	-	-
CO2	-	-	-
CO3	-	-	-
CO4	-	-	-
CO5	-	-	-

Lecture Plan

S.No	Topics to be covered	No. of periods	Proposed Date	Actual Date	CO	Taxonomy Level	Mode of Delivery
1	Joint distributions for Discrete random variable	1	25.1.2024		CO2	K3	PPT,Black Board & Chalk
2	Problems	1	27.1.2024		CO2	K3	PPT,Black Board & Chalk
3	Joint distributions for Continuous random variable	1	29.1.2024		CO2	K3	PPT,Black Board & Chalk
4	Problems	1	30.1.2024		CO2	K3	PPT,Black Board & Chalk
5	Covariance	1	31.1.2024		CO2	K3	PPT,Black Board & Chalk
6	Correlation	1	01.2.2024		CO2	K3	PPT,Black Board & Chalk
7	Problems	1	02.2.2024		CO2	K3	PPT,Black Board & Chalk
8	linear regression	1	03.2.2024		CO2	K3	PPT,Black Board & Chalk
9	Problems	1	06.2.2024		CO2	K3	PPT,Black Board & Chalk
10	Transformation of random variables	1	07.2.2024		CO2	K3	PPT,Black Board & Chalk
11.	Problems	1	07.2.2024		CO2	K3	PPT,Black Board & Chalk
12	Central limit theorem	1	08.2.2024		CO2	K3	PPT,Black Board & Chalk
13	Problems	1	08.2.2024		CO2	K3	PPT PPT,Black Board & Chalk
14	Lab	1	09.2.2024		CO2	K3	PPT,Black Board & Chalk
15	Lab	1	09.2.2024		CO2	K3	PPT,Black Board & Chalk

ACTIVITY BASED LEARNING

Activity based learning enhances students' critical thinking and collaborative skills. Experiential learning being the core, various activities such as quiz competitions, group discussion, etc. are conducted for all the five units to enhance the learning abilities of students. The students are the center of the activities, where student's opinions are valued, questions are encouraged, and discussions are done. These activities empower the students to explore and learn by themselves.

S.No.	TOPICS	Activity	Link/Question
1	Joint probability density	Activity	Suppose a certain local bank had three deposit or withdrawal counters. Two clients arrive at the counters at different times when the counters are serving no other customers. Each client chooses a counter at random, independently of the other. Calculate the joint probability function of X and Y. If X denotes the number of clients to select counter one and Y to represent the number of clients to select counter two.
2	Correlation & Regression	Assignment	https://drive.google.com/file/d/13wtB6E9yDiu0ykjT8Pn_D3kY_wO6Qgb2/view?usp=sharing
3	Transformation of random variables	Quiz and Flash card	

UNIT II

TWO - DIMENSIONAL RANDOM VARIABLES

UNIT II - TWO DIMENSIONAL RANDOM VARIABLES

2.1 Introduction:

In the previous chapter we studied various aspects of the theory of a one dimensional random variable. Now we extend our theory into two random variables one for each coordinate axis X and Y of the XY plane.

A two-dimensional random variable is mapping of the points in the sample space to ordered pairs (x, y). Usually, when dealing with a pair of random variables, the sample space naturally partitions itself so that it can be viewed as a combination of two simpler sample spaces. For example, if the experiment was to observe the density and weight of soil. The range of soil density could fall in some set, which we will call sample space S1, while the range of the weight would fall within sample space S2. The overall sample space of the experiment could be viewed as S = (S1 X S2). The outcome of the experiment, the pair of random variables (X, Y) is merely a mapping of the outcome S to a pair of numerical values (x(S), y(S)), where x(S) to be the density and y(S) is the weight of soil

2.2 Two Dimensional Random Variables:

Definition

Let S be the sample space. Let $X = X(s)$ and $Y = Y(s)$ be two functions each assigning a real number to each outcome $s \in S$. Then (X, Y) is a two-dimensional random variable.

Types of Random Variables:

1. Discrete random variables

2. Continuous random variables

Two-Dimensional Discrete Random Variable:

If the possible values of(X,Y) are finite, then (X,Y) is called a two-dimensional discrete random variable and it can be represented by$(x_i, y_j), i = 1,2, \dots n; j = 1,2, \dots m$.

Joint Probability Mass Function of the Discrete Random Variable:

For two discrete random variables $X$and Y, we write the probability that Xwill take the value $x_i$and Y will take the value $y_j$as$P(X = x_i, Y = y_j)$. Consequently, $P(X = x_i, Y = y_j)$ is the probability of intersection of the events$X = x_i$and$Y = y_j$.

$$(i.e)., P(X = x_i, Y = y_j) = P[(X = x_i) \cap (Y = y_j)]$$

The function$P(X = x_i, Y = y_j) = p(x_i, y_j)$ is called the joint probability function or joint probability mass function for discrete random variables $X$and $Y$and is denoted by p_{ij} and it satisfies the following conditions,

$i) p_{ij} \geq 0 \forall i, j ii) \sum_j \sum_i p_{ij} = 1$

Note: Joint Probability Distribution

The set of triples $\{x_i, y_j, p_{ij}\}$ is called Joint Probability Distribution.

Y X	Y1 Y2 Y3 …
X1	P11 p12……
X2	P21 p22 …
X3	
……	

Example: If the joint p.d.f of (X, Y) is given by $p(x, y) = K(2x + 3y), x = 0,1,2; y = 1,2,3$. Find 'K'.

Solution: Given

$$p(x, y) = K(2x + 3y), x = 0,1,2; y = 1,2,3.$$

$$Now, p(0,1) = K(0 + 3) = 3K$$

$$p(0,2) = K(0 + 6) = 6K$$

$$p(0,3) = K(0 + 9) = 9K$$

$$p(1,1) = K(2 + 3) = 5K$$

$$p(1,2) = K(2 + 6) = 8Kp(1,3) = K(2 + 9) = 11K$$

$$p(2,1) = K(4 + 3) = 7K$$

$$p(2,2) = K(4 + 6) = 10K$$

$$p(2,3) = K(4 + 9) = 13K$$

We know that, Total probability = 1

$$\left.\begin{array}{l} 3K + 6K + 9K + 5K + 8K \\ +11K + 7K + 10K + 13K \end{array}\right\} = 1$$

$$72K = 1$$

$$\therefore K = \frac{1}{72}$$

Two-Dimensional Continuous Random Variable:

If (X, Y) can take all the values in a region R in the XY plane, then (X, Y) is called a two dimensional continuous random variable.

Joint Probability Density Function of a Continuous Random Variable (P.D.F)

The joint probability density function of a two-dimensional continuous random variables (X, Y) is denoted by $f(x, y)$ and it satisfies the following conditions

$$i) f(x, y) \geq 0$$

$$\text{ii)} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dxdy = 1$$

Example: Show that the function$f(x,y) = 4xye^{-(x^2+y^2)}, x > 0, y > 0$is a joint p.d.f of $X$and Y.

Solution:

We know that if$f(x,y)$ satisfies the conditions

$$i) f_{XY}(x,y) \geq 0 ii) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)dxdy = 1 \text{ ,then } f(x,y) \text{is a joint density function.}$$

Given$f(x,y) = 4xye^{-(x^2+y^2)}, x > 0, y > 0$

$$i) f(x,y) \geq 0, x, y > 0$$

$$ii) \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = \int_0^{\infty}\int_0^{\infty} 4xye^{-(x^2+y^2)}dxdy$$

$$= 4\int_0^{\infty}\int_0^{\infty} xye^{-(x^2+y^2)}dxdy$$

$$= 4\left(\int_0^{\infty} xe^{-x^2}dx\right)\left(\int_0^{\infty} ye^{-y^2}dy\right)$$

$$= 4\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\because \int_0^{\infty} e^{-x^2}dx = \frac{1}{2}\right)$$

$$= 1$$

Since $f(x,y)$ satisfies the two conditions, it is a joint density function.

Joint Cumulative Distribution Function

If (X,Y) is a two-dimensional discrete random variable such that $F(x,y) = P(X \leq x,\ Y \leq y)$, is called the joint cumulative distribution Function of (X,Y)

1) For discrete Random variable, $F(x,y) = \sum_{y_j \leq y}\sum_{x_i \leq x} p_{ij}$
2) For continuous random variables, $F(x,y) = \int_{-\infty}^{y}\int_{-\infty}^{x} \boldsymbol{f(x,y)dxdy}$

2.3 MARGINAL PROBABILITY FUNCTION

Marginal Probability Function of X:

If (X, Y) is a two-dimensional discrete random variable, then the marginal probability function of X is defined as $P(X = x_i) = p_{i1} + p_{i2} + p_{i3} + \cdots = p_{i*}$

Marginal Probability Function of Y:

If (X, Y) is a two-dimensional discrete random variable then the marginal probability function of Y is defined as $P\left(Y = y_j\right) = p_{1j} + p_{2j} + p_{3j} + \cdots = p_{*j}$

Note:

1. Collection of pair $\{x_{i\,,}\, P(X = x_i)\}$ is called the marginal probability distribution of X.
2. The collection of pair $\{\, y_{j\,,}\;\; P\left(Y = y_j\right)\}$ is called the marginal probability distribution of Y.

Marginal Density Function of X:

1. If (X, Y) is a two-dimensional continuous random variable, then the marginal density function of X is defined by $f_x(X) = \int_{-\infty}^{\infty} f(x, y)\, dy$

1. If (X, Y) is a two-dimensional continuous random variable, then the marginal density function of Y is defined by $f_y(Y) = \int_{-\infty}^{\infty} f(x, y)\, dx$

Note:

1. $P[a \le X \le b] = \int_a^b f_x(X) dx$
2. $P[c \le Y \le d] = \int_a^b f_y(Y) dy$
3. $P[a \le X \le b, c \le Y \le d\,] = \int_c^d \int_a^b \boldsymbol{f(x, y) dxdy}$

Conditional Probability

- If (X, Y) is a two-dimensional discrete random variable, then the conditional probability of X given Y is $P\left(X = x_i / Y = y_j\right) = \frac{P\left(X=x_i, Y=y_j\right)}{P\left(Y=y_j\right)}$
- If (X, Y) is a two-dimensional discrete random variable, then the conditional probability of Y given X is $\left(Y = y_j / X = x_i\right) = \frac{P\left(X=x_i, Y=y_j\right)}{P(X=x_i)}$

For continuous random variable

- If (X, Y) is a two-dimensional continuous random variable, then the conditional probability of X given Y is $f(X/Y) = \frac{f(x,y)}{f_y(Y)}$
- If (X, Y) is a two dimensional continuous random variable then the conditional probability of Y given X is $f(Y/X) = \frac{f(x,y)}{f_x(X)}$

Note:

- If (X, Y) are two dimensional independent discrete random variables, then

$$P\left(X = x_i,\ Y = y_j\right) = P(X = x_i)\ P\left(Y = y_j\right)$$

- If (X, Y) are two dimensional independent continuous random variables, then $f(x, y) = f_x(X) f_y(Y)$

Problems

1. For the following bivariate probability distribution of (x, y), find

x/y		Y					
		1	2	3	4	5	6
X	0	0	0	1/32	2/32	2/32	3/32
	1	1/16	1/16	1/8	1/8	1/8	1/8
	2	1/32	1/32	1/64	1/64	0	2/64

(I) Find all the marginal probability distributions and (ii) conditional probability distributions of X given Y. Also find (iii) $P(X \leq 1)$, (iv) $P(Y \leq 3)$, (v) $P(X \leq 1/\, Y \leq 3)$, (vi) $P(X + Y \leq\ 4\,)$.

Solution:

x/y		Y						
		1	2	3	4	5	6	P(X)
X	0	0	0	1/32	2/32	2/32	3/32	8/32
	1	1/16	1/16	1/8	1/8	1/8	1/8	10/16
	2	1/32	1/32	1/64	1/64	0	2/64	8/64
	P(Y)	3/32	3/32	11/64	13/64	6/32	16/64	

i) Marginal probability distribution of X.

X	$P(X)$
0	8/32
1	10/16
2	8/64

II) Marginal probability distribution of Y

Y	$P(Y)$
1	3/32
2	3/32
3	11/64
4	13/64
5	6/32
6	16/64

Conditional probability distribution of X given Y

$$P(X=0/Y=1)=\frac{P(X=0,\ Y=1)}{P(Y=1)}=\frac{p_{01}}{P(Y=1)}=0$$

$$P(X=1/Y=1)=\frac{P(X=1,\ Y=1)}{P(Y=1)}=\frac{p_{11}}{P(Y=1)}=\frac{{}^{1}/_{16}}{{}^{3}/_{32}}={}^{2}/_{3}$$

$$P(X=2/Y=1)=\frac{P(X=2,\ Y=1)}{P(Y=1)}=\frac{p_{21}}{P(Y=1)}=\frac{{}^{1}/_{32}}{{}^{3}/_{32}}={}^{1}/_{3}$$

$$P(X=0/Y=2)=\frac{P(X=0,\ Y=2)}{P(Y=2)}=\frac{p_{02}}{P(Y=2)}=0$$

$$P(X=1/Y=2)=\frac{P(X=1,\ Y=2)}{P(Y=2)}=\frac{p_{12}}{P(Y=2)}=\frac{1/_{16}}{3/_{32}}={}^{2}/_{3}$$

$$P(X=2/Y=2)=\frac{P(X=2,\ Y=2)}{P(Y=2)}=\frac{p_{22}}{P(Y=2)}=\frac{1/_{32}}{3/_{32}}={}^{1}/_{3}$$

$$P(X=0/Y=3)=\frac{P(X=0,\ Y=3)}{P(Y=3)}=\frac{p_{03}}{P(Y=3)}=\frac{1/_{32}}{11/_{64}}={}^{2}/_{11}$$

$$P(X=1/Y=3)=\frac{P(X=1,\ Y=3)}{P(Y=3)}=\frac{p_{13}}{P(Y=3)}=\frac{1/_{8}}{11/_{64}}={}^{8}/_{11}$$

$$P(X=2/Y=3)=\frac{P(X=2,\ Y=3)}{P(Y=3)}=\frac{p_{23}}{P(Y=3)}=\frac{1/_{64}}{11/_{64}}={}^{1}/_{11}$$

$$P(X=0/Y=4)=\frac{P(X=0,\ Y=4)}{P(Y=4)}=\frac{p_{04}}{P(Y=4)}=\frac{2/_{32}}{13/_{64}}={}^{4}/_{13}$$

$$P(X=1/Y=4)=\frac{P(X=1,\ Y=4)}{P(Y=4)}=\frac{p_{14}}{P(Y=4)}=\frac{1/_{8}}{13/_{64}}={}^{8}/_{13}$$

$$P(X=2/Y=4)=\frac{P(X=2,\ Y=4)}{P(Y=4)}=\frac{p_{24}}{P(Y=4)}=\frac{1/_{64}}{13/_{64}}={}^{1}/_{13}$$

$$P(X=0/Y=5)=\frac{P(X=0,\ Y=5)}{P(Y=5)}=\frac{p_{05}}{P(Y=5)}=\frac{2/32}{6/32}=1/3$$

$$P(X=1/Y=5)=\frac{P(X=1,\ Y=5)}{P(Y=5)}=\frac{p_{15}}{P(Y=5)}=\frac{1/8}{6/32}=2/3$$

$$P(X=2/Y=5)=\frac{P(X=2,\ Y=5)}{P(Y=5)}=\frac{p_{25}}{P(Y=5)}=\frac{0}{6/32}=0$$

$$P(X=0/Y=6)=\frac{P(X=0,\ Y=6)}{P(Y=6)}=\frac{p_{06}}{P(Y=6)}=\frac{3/32}{16/64}=3/8$$

$$P(X=1/Y=6)=\frac{P(X=1,\ Y=6)}{P(Y=6)}=\frac{p_{16}}{P(Y=6)}=\frac{1/8}{16/64}=1/2$$

$$P(X=2/Y=6)=\frac{P(X=2,\ Y=6)}{P(Y=6)}=\frac{p_{26}}{P(Y=6)}=\frac{2/64}{16/64}=1/8$$

(iii) $\mathrm{P}(\mathrm{X}\le 1)=\mathrm{P}(\mathrm{X}=0)+\mathrm{P}(\mathrm{X}=1)$

$$={}^{8}/_{32}+{}^{10}/_{16}={}^{28}/_{32}={}^{7}/_{8}$$

(iv) $\mathrm{P}(\mathrm{Y}\le 3)=\mathrm{P}(\mathrm{Y}=1)+\mathrm{P}(\mathrm{Y}=2)++\mathrm{P}(\mathrm{Y}=3)$

$$={}^{3}/_{32}+{}^{3}/_{32}+{}^{11}/_{64}={}^{23}/_{64}$$

(v) $\mathrm{P}(\mathrm{X} \leq 1/\mathrm{Y} \leq 3) = \dfrac{\mathrm{P}(\mathrm{X} \leq 1, \mathrm{Y} \leq 3)}{P(\mathrm{Y} \leq 3)}$

$\mathrm{P}(\mathrm{X} \leq 1, \mathrm{Y} \leq 3) = P_{01} + P_{02} + P_{03} + P_{03} + P_{11} + P_{12} + P_{13}$

$$= 0 + 0 + {}^{1}/_{32} + {}^{1}/_{16} + {}^{1}/_{16} + {}^{1}/_{8} = {}^{9}/_{32}$$

$$\mathrm{P}(\mathrm{X} \leq 1/\mathrm{Y} \leq 3) = \frac{\mathrm{P}(\mathrm{X} \leq 1, \mathrm{Y} \leq 3)}{P(\mathrm{Y} \leq 3)} = \frac{{}^{9}/_{32}}{{}^{23}/_{64}} = {}^{18}/_{23}$$

vi) $\mathrm{P}(\mathrm{X} + \mathrm{Y} \leq 4)$.

Probability distribution of $\mathrm{X} + \mathrm{Y}$

X+Y	1	2	3	4	5	6	7	8
P(X+Y)	0	1/16	4/32	7/32	13/64	15/64	1/8	2/64

$\mathrm{P}(\mathrm{X} + \mathrm{Y} \leq 4) = 0 + {}^{1}/_{16} + {}^{4}/_{32} + {}^{7}/_{32} = {}^{13}/_{32}$

1. The joint probability mass function of (X, Y) is given by $p(x, y) = k(2x + 3y), x = 0,1,2, y = 1,2,3$. i) Find all the marginal probability distributions. Also find ii) the probability distribution of $X + Y$, iii) $P(X + Y > 3)$, iv) $P(X \leq 1, Y \leq 2)$, v) $P(X \leq 1/Y \leq 2)$

Solution:

The joint probability mass function of (X, Y) is given by $p(x, y) = k(2x + 3y), x = 0,1,2,\ y = 1,2,3$

Probability distribution is

Y \ X	1	2	3
0	3k	6k	9k
1	5k	8k	11k
2	7k	10k	13k

$$\Rightarrow \sum\sum P_{ij} = 1$$

$$\Rightarrow 72k = 1$$

$$\Rightarrow k = {}^{1}/_{72}$$

Marginal probability distributions

Marginal probability distributions of X

X	$P(X)$
0	18K=18/72
1	24K=24/72
2	30K=30/72

Marginal probability distributions of Y

Y	$P(Y)$
1	15K=15/72
2	24K=24/72
3	33K=33/72

ii) Probability distribution of $X + Y$

$X+Y$	1	2	3	4	5
$P(X+Y)$	3/72	11/72	24/72	21/72	13/72

iii) $P(X+Y>3)$ =34/72

iv) $\mathrm{P(X \le 1, Y \le 2)} = P_{01} + P_{02} + P_{11} + P_{12}$

$$= {}^{3}/_{72} + {}^{6}/_{72} + {}^{5}/_{72} + {}^{8}/_{72} = {}^{22}/_{72} = {}^{11}/_{36}$$

$$\mathrm{v)P(X \le 1/Y \le 2)} = \frac{\mathrm{P(X \le 1, Y \le 2)}}{\mathrm{P(Y \le 2)}} = \frac{{}^{11}/_{36}}{{}^{39}/_{72}} = \frac{22}{39}$$

1. The joint probability mass function of a bivariate discrete random variable (X,Y) is given by $P(X,Y) = K(2x+y)$, $x = 0,1.2;\ y = 1,2.3.$
   Find the marginal probability mass function of X and Y.
   Also find probability distribution of $(X+Y)$.

Solution:

The joint probability mass function of (X,Y) is given by

$P(X,Y) = K(2x+y)$, $x = 0,1.2;\ y = 1,2.3.$

Probability distribution is

Y / X	1	2	3
0	k	2k	3k
1	3k	4k	5k
2	5k	6k	7k

$$\Rightarrow \sum\sum P_{ij} = 1$$

$$\Rightarrow 36k = 1$$

$$\Rightarrow k = {}^{1}/_{36}$$

i) Marginal probability distributions

Marginal probability distributions of X

X	$P(X)$
0	6K=6/36
1	12K=12/36
2	18K=18/36

Marginal probability distributions of Y

Y	$P(Y)$
1	9K=9/36
2	12K=12/36
3	15K=15/36

Probability distribution of X + Y

$X+Y$	1	2	3	4	5
$P(X+Y)$	1/36	5/36	12/36	11/36	7/36

1. The two-dimensional random variable (X, Y) has the joint density function $f(x, y) = \frac{x+y}{21}$, $X = 1,2,3;\ Y = 1,2.$

   (i) Find all the marginal and conditional distributions of X and Y (ii) probability distribution of $(X + Y)$, (iii) $P(X + Y \leq 2)$

Solution:

The joint probability mass function of (X,Y) is given by$f(x,y)=\frac{x+y}{21}$, $X=1,2,3;\ Y=1,2.$

Probability distribution is

X \ Y	1	2
1	2/21	3/21
2	3/21	4/21
3	4/21	5/21

i) Marginal probability distributions

Marginal probability distributions of X

X	$P(X)$
1	5/21
2	7/21
3	9/21

Marginal probability distributions of Y

Y	$P(Y)$
1	9/21
2	12/21

Conditional probability distribution of X given Y

Conditional probability distribution of X given Y=1

Conditional probability distribution of Y given X=2

$$P(X = 1/Y = 1) = \frac{P(X = 1,\ Y = 1)}{P(Y = 1)} = \frac{p_{11}}{P(Y = 1)} = \frac{2/_{21}}{9/_{21}} = 2/_{9}$$

$$P(X = 2/Y = 1) = \frac{P(X = 2,\ Y = 1)}{P(Y = 1)} = \frac{p_{21}}{P(Y = 1)} = \frac{3/_{21}}{9/_{21}} = 3/_{9}$$

$$P(X = 3/Y = 1) = \frac{P(X = 3,\ Y = 1)}{P(Y = 1)} = \frac{p_{31}}{P(Y = 1)} = \frac{4/_{21}}{9/_{21}} = 4/_{9}$$

Conditional probability distribution of X given Y=2

$$P(X = 1/Y = 2) = \frac{P(X = 1,\ Y = 2)}{P(Y = 2)} = \frac{p_{12}}{P(Y = 2)} = \frac{3/_{21}}{12/_{21}} = 3/_{12}$$

$$P(X = 2/Y = 2) = \frac{P(X = 2,\ Y = 2)}{P(Y = 2)} = \frac{p_{22}}{P(Y = 2)} = \frac{4/_{21}}{12/_{21}} = 4/_{12}$$

$$P(X = 3/Y = 2) = \frac{P(X = 3,\ Y = 2)}{P(Y = 2)} = \frac{p_{32}}{P(Y = 2)} = \frac{5/_{21}}{12/_{21}} = 5/_{12}$$

Conditional probability distribution of Y given X

Conditional probability distribution of Y given X=1

$$P(Y = 1/X = 1) = \frac{P(X = 1,\ Y = 1)}{P(X = 1)} = \frac{p_{11}}{P(X = 1)} = \frac{2/_{21}}{5/_{21}} = 2/_{5}$$

$$P(Y = 2/X = 1) = \frac{P(X = 1,\ Y = 2)}{P(X = 1)} = \frac{p_{12}}{P(X = 1)} = \frac{3/_{21}}{5/_{21}} = 3/_{5}$$

$$P(Y=1/X=2)=\frac{P(X=2,\ Y=1)}{P(X=2)}=\frac{p_{21}}{P(X=2)}=\frac{3/_{21}}{7/_{21}}=3/_{7}$$

$$P(Y=2/X=2)=\frac{P(X=2,\ Y=2)}{P(X=2)}=\frac{p_{22}}{P(X=2)}=\frac{4/_{21}}{7/_{21}}=4/_{7}$$

Conditional probability distribution of Y given X=3

$$P(Y=1/X=3)=\frac{P(X=3,\ Y=1)}{P(X=3)}=\frac{p_{31}}{P(X=3)}=\frac{4/_{21}}{9/_{21}}=4/_{9}$$

$$P(Y=2/X=3)=\frac{P(X=3,\ Y=2)}{P(X=3)}=\frac{p_{32}}{P(X=3)}=\frac{5/_{21}}{9/_{21}}=5/_{9}$$

i) Probability distribution of X + Y

$X+Y$	2	3	4	5
$P(X+Y)$	2/21	6/21	8/21	5/21

ii) $P(X+Y\le\ 2)=2/_{21}$

Problems for Practice

1. The joint probability mass function of a bivariate discrete random variable (X, Y) is given by

x y	1	2
1	0.1	0.2
2	0.3	0.4

(i) Find the conditional distribution of X given Y=1

(ii) Find the conditional distribution of Y given X= 2

Also find probability distribution of $(X + Y)$

2. The joint probability mass function of (X, Y) is given by $P(x, y) = k(x + 2y)$, $x = 0,1,2$, and $y = 0,1,2$.Find (i) all the marginal and conditional distributions of $X and Y$(ii)probability distribution of $(X + Y)$, (iii) $P(X + Y \leq 2)$

3. The two-dimensional random variable(X, Y) has the joint density function $f(x, y) = \frac{x+2y}{27}$, $X = 0,1,2$; $Y = 0,1,2$. (i) Find all the marginal and conditional distributions of $X$and$Y$. (ii)probability distribution of $(X + Y)$, (iii) $P(X + Y \leq 2)$

Problems in Continuous Random Variable

1. The joint p.d.f of (X, Y) is given by $(x, y) = \frac{1}{8}(6 - x - y), 0 \leq x \leq 2, 2 \leq y \leq 4$.

(i) Find the marginal density of X and Y

(ii) Find the conditional density of X and Y

Also find

(iii) $P\ (X < 1 \cap Y < 3)$,

(iv) $P(x + y < 3)$

(v) $P(^{X < 1}/_{Y < 3})$

Solution:

The joint probability mass function of (X, Y) is given by $f(x, y) = \frac{1}{8}(6 - x - y), 0 \leq x \leq 2, 2 \leq y \leq 4$

i) Marginal density of X

$$f_x(X) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

$$= \int_{2}^{4} \frac{1}{8}(6 - x - y)\, dy$$

$$= \frac{1}{8}\left[6y - xy - \frac{y^2}{2}\right]_2^4$$

$$= \frac{1}{8}[24 - 4x - 8 - 12 + 2x + 2]$$

$$f_x(X) = \frac{1}{8}[6 - 2x]$$

Marginal density of Y

$$f_y(Y) = \int_{-\infty}^{\infty} f(x, y)\, dx$$

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

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

$$= \frac{1}{8}[12-2-2y-0]$$

$$= \frac{1}{8}[10-2y]$$

$$f_y(Y) = \frac{1}{4}[5-y]$$

i) Conditional density of X

$$f\left({}^{X}\!/_{Y}\right) = \frac{f(x,y)}{f_y(Y)}$$

$$= \frac{\frac{1}{8}(6-x-y)}{\frac{1}{4}[5-y]} = \frac{(6-x-y)}{2[5-y]}$$

Conditional density of Y

$$f\left({}^{Y}\!/_{X}\right) = \frac{f(x,y)}{f_y(Y)}$$

$$= \frac{\frac{1}{8}(6-x-y)}{\frac{1}{8}[6-2x]} = \frac{(6-x-y)}{[6-2x]}$$

ii) $P\ (X < 1 \cap Y < 3) = P(X < 1, Y < 3)$

$$= \iint f(x,y)dxdy$$

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

$$= \int_2^3 \left[6x - \frac{x^2}{2} - yx\right]_0^1 dy$$

$$= \int_2^3 \left[6 - \frac{1}{2} - y\right] dy$$

$$= \left[6y - \frac{y}{2} - \frac{y^2}{2}\right]_2^3 = \left[18 - \frac{3}{2} - \frac{9}{2} - 12 + 1 + 2\right] = \frac{3}{8}$$

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

i) $P(x + y < 3) = \iint f(x, y)dxdy$

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

$$= \int_2^3 \left[6x - \frac{x^2}{2} - yx\right]_0^{3-y} dy$$

$$= \int_2^3 \left[6(3 - y) - \frac{(3 - y)^2}{2} - y(3 - y)\right] dy$$

$$\int_2^3 \left[18 - 6y - \frac{(9 + y^2 - 6y)}{2} - 3y + y^2\right] dy$$

$$= \left[18y - \frac{6y^2}{2} - \frac{9y}{2} - \frac{y^3}{6} + \frac{6y^2}{4} - \frac{3y^2}{2} + \frac{y^3}{3}\right]_2^3$$

$$= \frac{5}{24}$$

ii) $P\left({}^{X < 1}/_{Y < 3}\right) = \frac{P(X<1, Y<3)}{P(\,Y<3)}$

$$P(\,Y < 3) = \int_2^3 f_y(Y)\,dy$$

$$= \int_2^3 \frac{1}{4}[5-y]\,dy$$

$$= \frac{1}{4}\left[5y - \frac{y^2}{2}\right]_2^3$$

$$= \frac{1}{4}\left[15 - \frac{9}{2} - 10 + 2\right]$$

$$P(\,Y < 3) = \frac{5}{8}$$

$$P\left(X < 1/_{Y\,<\,3}\right) = \frac{\frac{3}{8}}{\frac{5}{8}} = \frac{3}{5}$$

2. The joint probability density function of the random variable (X, Y) is given by $f(x,y) = kxye^{-(x^2+y^2)}, x > 0, y > 0$. Find the value of k and also prove that X and Y are independent

Solution:

The joint probability density function of the random variable (X, Y) is given by $f(x,y) = kxye^{-(x^2+y^2)}, x > 0, y > 0.$

$$\Rightarrow \iint_R f(x,y)dxdy = 1$$

$$\Rightarrow \int_0^\infty \int_0^\infty k\,x\,y\,e^{-\left(x^2+y^2\right)}\,dxdy = 1$$

$$\Rightarrow k\int_0^\infty\int_0^\infty xye^{-(x^2+y^2)}\,dx\,dy = 1$$

$$\Rightarrow k\int_0^\infty xe^{-x^2}dx\int_0^\infty ye^{-y^2}dy = 1 \qquad (\because \int_0^\infty xe^{-x^2}dx = {}^1/_2)$$

$$\Rightarrow k({}^1/_2)({}^1/_2) = 1$$

$$\Rightarrow k = 4$$

To prove X and Y are independent, we have to prove that $f(x,y) = f_x(X)f_y(Y)$

$$f_x(X) = \int_0^\infty f(x,y)\,dy$$

$$f_x(X) = \int_0^\infty 4xe^{-x^2}ye^{-y^2}\,dy$$

$$f_x(X) = 4xe^{-x^2}\int_0^\infty ye^{-y^2}\,dy$$

$$= 4xe^{-x^2}\left(\frac{1}{2}\right) = 2xe^{-x^2}$$

$$f_x(X) = 2xe^{-x^2}$$

$$f_y(Y) = \int_{-\infty} f(x,y)\,dx$$

$$= \int_0^\infty 4xe^{-x^2}ye^{-y^2}\,dx$$

$$= 4ye^{-y^2}\int_0^\infty xe^{-x^2}\,dx$$

$$= 4ye^{-y^2}\left(\frac{1}{2}\right) = 2ye^{-y^2}$$

$$f_y(Y) = 2ye^{-y^2}$$

$$f_x(X)f_y(Y) = 2xe^{-x^2}2ye^{-y^2} = 4xye^{-(x^2+y^2)} = f(x,y)$$

$\therefore X$ and Y are independent random variables.

3. The joint probability density function of the random variable(x,y) is given by $f(x,y) = \begin{cases} 2, & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0, & elsewhere \end{cases}$. Prove that X and Y are independent.
Also find the conditional density function of X and Y

Solution:

The joint probability density function of the random variable (x,y) is given by

$$f(x,y) = \begin{cases} 2, & 0 \le x \le 1, \quad 0 \le y \le 1 \\ 0, & elsewhere \end{cases}$$

$$f_x(X) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$= \int_0^{\mathrm{x}} 2\,dy$$

$$= 2[\mathrm{y}]_0^{\mathrm{x}}$$

$$f_x(X) = 2x$$

$$f_y(Y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

$$= \int_0^{2} 2\,dx$$

$$= 2[\mathrm{x}]_{\mathrm{y}}^{1}$$

$$f_y(Y) = 2[1-y]$$

$$f_x(X)f_y(Y) = 2x * 2[1-y] = 4x[1-y] \neq f(x,y)$$

Conditional density function of X

$$f\left({}^{X}/_{Y}\right) = \frac{f(x,y)}{f_y(Y)}$$

$$f\left({}^{X}/_{Y}\right) = \frac{2}{2[1-y]} = \frac{1}{[1-y]}$$

Conditional density of Y

$$f\left({}^{Y}/_{X}\right) = \frac{f(x,y)}{f_x(X)}$$

$$f\left({}^{Y}/_{X}\right) = \frac{2}{2x} = \frac{1}{\mathrm{x}}$$

4. The joint probability density function of the random variable(x,y) is given by $f(x,y) = \begin{cases} \mathrm{cx(x-y)}, & 0 \le x \le 2, -\mathrm{x} \le y \le \mathrm{x} \\ 0, & elsewhere \end{cases}$

Find
(i) The value of c
(ii) $f_x(X), f_y(Y)$
(iii) $f\left({}^{Y}/_{X}\right)$

Solution:

The joint probability density function of the random variable (x,y) is given by

$$f(x,y) = \begin{cases} \mathrm{cx(x-y)}, & 0 \le x \le 2, -\mathrm{x} \le y \le \mathrm{x} \\ 0, & elsewhere \end{cases}$$

We know that the total probability is 1

$$\Rightarrow \iint_R f(x,y)dxdy = 1$$

$$\Rightarrow \int_0^2 \int_{-\mathrm{x}}^{\mathrm{x}} \mathrm{cx(x-y)}dxdy = 1$$

$$\Rightarrow c\int_0^2 \left[x^2y - \frac{y^2}{2}\right]_{-\mathrm{x}}^{\mathrm{x}} dy = 1$$

$$\Rightarrow c\int_0^2 \left[\left(x^3 - \frac{x^2}{2} + x^3 + \frac{x^2}{2}\right)\right] dy = 1$$

$$\Rightarrow c\int_0^2 [(2x^3)]dy = 1$$

$$\Rightarrow c\,2\left[\frac{x^4}{4}\right]_0^2 = 1$$

$$\Rightarrow c2\left(\frac{16}{4}\right) = 1$$

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

$$\therefore f(x,y) = \begin{cases} \frac{1}{8}\mathrm{x}(\mathrm{x}-\mathrm{y}), & 0 \le x \le 2, -\mathrm{x} \le y \le \mathrm{x} \\ 0, & elsewhere \end{cases}$$

ii) Marginal density of X

$$f_x(X) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

$$= \int_{-\mathrm{x}}^{\mathrm{x}} \frac{1}{8}\mathrm{x}(\mathrm{x}-\mathrm{y})\, dy$$

$$= \frac{1}{8}\left[x^2 y - \frac{y^2}{2}\right]_{-\mathrm{x}}^{\mathrm{x}}$$

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

$$f_x(X) = \frac{2x^3}{8} = \frac{x^3}{4}$$

Marginal density of Y

In region $R1$, $-2 \le y \le 0$

$$f_y(Y) = \int_{-\infty}^{\infty} f(x,y)\, dx$$

$$= \int_{-\mathrm{y}}^{2} \frac{1}{8}\mathrm{x}(\mathrm{x}-\mathrm{y})\, dX$$

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

$$= \frac{1}{8}\left[\frac{8}{3} - 2\mathrm{y} + \frac{5y^3}{6}\right]$$

$$f_y(Y) = \frac{1}{3} - \frac{\mathrm{y}}{4} + \frac{5y^3}{48}$$

In region $R2$, $0 \le y \le 2$

$$f_y(Y) = \int_{-\infty}^{\infty} f(x,y)\, dx$$

$$= \int_{\mathrm{y}}^{2} \frac{1}{8}\mathrm{x}(\mathrm{x}-\mathrm{y})\, dX$$

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

$$=\frac{1}{8}\left[\frac{8}{3}-\frac{4\mathrm{y}}{2}-\frac{y^3}{3}+\frac{y^3}{2}\right]$$

$$f_y(Y)=\frac{1}{3}-\frac{\mathrm{y}}{4}+\frac{y^3}{48}$$

$$f_y(Y)=\begin{cases}\frac{1}{3}-\frac{\mathrm{y}}{4}+\frac{5y^3}{48}, & -2\leq y\leq 0\\ \frac{1}{3}-\frac{\mathrm{y}}{4}+\frac{y^3}{48}, & 0\leq y\leq 2\end{cases}$$

iii. Conditional density of Y given X

$$f\left({}^{Y}\!/_{X}\right)=\frac{f(x,y)}{f_x(x)}$$

$$f\left({}^{Y}\!/_{X}\right)=\frac{\frac{1}{8}\mathrm{x}(\mathrm{x}-\mathrm{y})}{\left(\frac{x^3}{4}\right)}=\frac{(\mathrm{x}-\mathrm{y})}{2\mathrm{x}^2}$$

5. The joint probability density function of the random variable(x,y) is given by

$$f(x,y)=\begin{cases}\mathrm{x}y^2+\frac{x^2}{8}, & 0\leq x\leq 2, 0\leq y\leq 1\\ 0, & elsewhere\end{cases}$$

(i) Find the marginal density of X and Y

Also find,

(ii) $P(X>1)$

(iii) $P\left(Y<\frac{1}{2}\right)$

(iv) $P\left(X>1\big/_{Y<\frac{1}{2}}\right)$

(v) $P(X<Y)$

(vi) $P(X+Y\leq 1)$.

Solution:

The joint probability density function of the random variable (x,y) is given by

$$f(x,y) = \begin{cases} \mathrm{x}y^2 + \frac{x^2}{8}, & 0 \le x \le 2, 0 \le y \le 1 \\ 0, & elsewhere \end{cases}$$

i) Marginal density of X

$$f_x(X) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$= \int_0^1 \left(\mathrm{x}y^2 + \frac{x^2}{8}\right) dy$$

$$= \left[\frac{\mathrm{x}y^3}{3} - \frac{x^2 y}{8}\right]_0^1$$

$$f_x(X) = \frac{\mathrm{x}}{3} - \frac{x^2}{8}$$

Marginal density of Y

$$f_y(Y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

$$= \int_0^2 \left(\mathrm{x}y^2 + \frac{x^2}{8}\right) dX$$

$$= \left[\frac{x^2 y^2}{2} + \frac{x^3}{24}\right]_0^2$$

$$= \left[\left[\frac{4y^2}{2} + \frac{8}{24}\right]\right] = \left(2y^2 + \frac{1}{3}\right)$$

$$f_y(Y) = 2y^2 + \frac{1}{3}$$

ii) $P(X > 1) = \int_1^2 f_x(X)\,dx$

$$= \int_1^2 \left(\frac{\mathrm{x}}{3} - \frac{x^2}{8}\right) dx = \left[\frac{x^2}{6} + \frac{x^3}{24}\right]_1^2$$

$$= \left[\frac{4}{6} + \frac{8}{24} - \frac{1}{6} - \frac{1}{24}\right] = \frac{19}{24}$$

i) $$P\left(Y<\frac{1}{2}\right)=\int_0^{\frac{1}{2}} f_y(Y)\,dy$$

$$=\int_0^{\frac{1}{2}}\left(2y^2+\frac{1}{3}\right)dy$$

$$=\left[\frac{2y^3}{3}+\frac{y}{3}\right]_0^{\frac{1}{2}}=\frac{2}{24}+\frac{1}{6}=\frac{1}{4}$$

ii) $$P\left(X>1\Big/_{Y<\frac{1}{2}}\right)=\frac{P\left(X>1,Y<\frac{1}{2}\right)}{P\left(Y<\frac{1}{2}\right)}$$

$$P\left(X>1,Y<\frac{1}{2}\right)=\int_0^{\frac{1}{2}}\int_1^2\left(\mathrm{x}y^2+\frac{x^2}{8}\right)\boldsymbol{dxdy}$$

$$=\int_0^{\frac{1}{2}}\left[\frac{x^2y^2}{2}+\frac{x^3}{24}\right]_1^2 dy$$

$$=\int_0^{\frac{1}{2}}\left(2y^2+\frac{8}{24}-\frac{y^2}{2}-\frac{1}{24}\right)dy=\int_0^{\frac{1}{2}}\left(\frac{7}{24}+\frac{3y^2}{2}\right)dy$$

$$=\left[\frac{7\mathrm{y}}{24}+\frac{y^3}{2}\right]_0^{\frac{1}{2}}=\left[\frac{7}{48}+\frac{1}{16}\right]=\frac{10}{48}=\frac{5}{24}$$

$$P\left(X>1,Y<\frac{1}{2}\right)=\frac{5}{24}$$

$$\therefore P\left(X>1\Big/_{Y<\frac{1}{2}}\right)=\frac{P\left(X>1,Y<\frac{1}{2}\right)}{P\left(Y<\frac{1}{2}\right)}=\frac{\frac{5}{24}}{\frac{1}{4}}=\frac{5}{6}$$

iii) $P(X<Y)=\iint_R f(x,y)dxdy$ $\quad\therefore$ *limit of R is* $,0\le x\le y\ \&\ 0\le y\le 1$

$$=\int_0^1\int_0^{\mathrm{y}}\left(\mathrm{x}y^2+\frac{x^2}{8}\right)\boldsymbol{dxdy}$$

$$=\int_0^1\left[\frac{x^2y^2}{2}+\frac{x^3}{24}\right]_0^y dy$$

$$= \int_0^1 \left(\frac{y^4}{2} + \frac{y^3}{24}\right) dy$$

$$= \left[\frac{y^5}{10} + \frac{y^4}{96}\right]_0^1 = \frac{1}{10} + \frac{1}{96} = \frac{53}{480}$$

i) $P(X + Y \leq 1) = \iint_R f(x,y)dxdy \quad \therefore limit\ of\ R\ is\ 0 \leq x \leq 1 - y\ \&\ 0 \leq y \leq$

$$= \int_0^1 \int_0^{1-y} \left(\mathrm{x}y^2 + \frac{x^2}{8}\right) dxdy$$

$$= \int_0^1 \left[\frac{x^2y^2}{2} + \frac{x^3}{24}\right]_0^{1-y} dy$$

$$= \int_0^1 \left(\frac{y^2(1-y)^2}{2} + \frac{(1-y)^3}{24}\right) dy$$

$$= \int_0^1 \left(\frac{y^2 - 2y^3 + y^4}{2} + \frac{(1-y)^3}{24}\right) dy = \frac{13}{480}$$

$$\Rightarrow P(X + Y \leq 1) = \frac{13}{480}$$

6. The joint density function of X and Y $f(x,y) = \begin{cases} 8xy 0 < x < y < 1 \\ 0 elsewhere \end{cases}$

(i) Find the marginal density of X and Y

(ii) Find the conditional density of X and Y

(iii) Are X and Y independent.

Solution:

The joint probability density function of the random variable (x,y) is given by

$$f(x,y) = \begin{cases} 8xy 0 < x < y < 1 \\ 0 elsewhere \end{cases}$$

i) Marginal density of X

$$f_x(X) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$= \int_{x}^{1} 8xy\,dy$$

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

$$f_x(X) = 4x(1 - x^2)$$

Marginal density of Y

$$f_y(Y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

$$= \int_{0}^{y} 8xy\,dx$$

$$= 8y\left(\frac{x^2}{2}\right)_0^y = 8y\left(\frac{y^2}{2}\right) = 4y^3$$

$$f_y(Y) = 4y^3$$

i) Conditional density function

Conditional density function of X

$$f\left({}^{X}\!/_{Y}\right) = \frac{f(x,y)}{f_y(Y)}$$

$$= \frac{8xy}{4y^3}$$

$$f\left({}^{X}\!/_{Y}\right) = \frac{2x}{y^2}$$

Conditional density function of Y

$$f\left({}^{Y}\!/_{X}\right) = \frac{f(x,y)}{f_x(x)}$$

$$= \frac{8xy}{4x(1-x^2)}$$

$$f\left({}^{Y}/_{X}\right) == \frac{2x}{(1-x^2)}$$

$$f_x(X)f_y(Y) = \left(4x(1-x^2)(4y^3)\right) \neq \boldsymbol{f(x,y)}$$

$\therefore X$ and Y are not independent Random variables

Exercises

1. If the joint distribution function of a two dimensional random variable (X,Y) is given by

$$F(x,y) = \begin{Bmatrix} (1-e^{-x})(1-e^{-y}) & x > 0 \quad and \quad y > 0 \\ 0 & elsewhere \end{Bmatrix}$$

Find

i) Marginal density of X and Y

ii) $P(X < 1/Y < 3)$

iii) $P(X > Y)$

iv) $P(X + Y < 1)$

2. If the joint pdf of a two dimensional random variable (X,Y) is given by

$$f(x,y) = \begin{cases} x^2 + \frac{xy}{3}, 0 < x < 1, 0 < y < 2 \\ 0, otherwise \end{cases}$$

(i) Find the marginal density of X and Y

(ii) Find the conditional density of X and Y

(iii) Find $(a) P(X > \frac{1}{2})$ $(b) P(Y < X)$ $(c) P(Y < \frac{1}{2} / X < \frac{1}{2})$

3. The joint probability density function of the two dimensional random variable (X,Y) is given by $f(x,y) = \begin{cases} \frac{x}{4}(1+3y^2), 0 < x < 2, 0 < y < 1 \\ 0,\text{elsewhere} \end{cases}$.

Find

a) Conditional probability density functions of X given Y=y and Y given X=x.

b) $P\left[\frac{\frac{1}{4}<X<\frac{1}{2}}{Y=\frac{1}{3}}\right]$.

2.4 MEASURES OF TWO-DIMENSIONAL RANDOM VARIABLE (X,Y) OR EXPECTATION:

If (X,Y) is a two dimensional discrete random variable, then the marginal expectation of the variable *X, Y* is given by

$$E(X) = \sum_i xP(X = x_i)$$

$$E(Y) = \sum_j yP(Y = y_j)$$

$$E(X^2) = \sum_i x^2 P(X = x_i)$$

$$E(Y^2) = \sum_j y^2 P(Y = y_j)$$

$$E(XY) = \sum_j \sum_i x\,yP_{ij}$$

If *(X,Y)* is a two-dimensional continuous random variable, then the marginal expectation of the variable *X, Y* is given by

i) $E(X) = \int x f_X(x)\,dx$

ii) $E(Y) = \int y\, f_Y(y)dy$

iii) $E(X^2) = \int x^2 f_X(x)\,dx$

Variance:

i) Variance of X is given by $Var(X) = E(X^2) - (E(X))^2$

ii) Variance of Y is given by $Var(Y) = E(Y^2) - (E(Y))^2$

2.5 CORRELATION

So far, we learned about the joint probability distribution of two random variables $X \text{ and } Y$. Here, we will extend our study of the relationship between two random variables by learning how to quantify the degree to which two random variables X and Y are associated or correlated. Here, we will begin our attempt to compute the dependence between two random variables and by exploring what is called the covariance between the two random variables. Now we will start with a formal definition of the covariance.

Covariance

Let X and Y be random variables (discrete or continuous!) with means $\mu_x$and μ_y

. The covariance of X and Y, denoted $COV(X,Y)$ or σ_{XY}, is defined as:

$$COV(X,Y) = \sigma_{XY} = E[(X - \mu_x)\,(Y - \mu_y)]$$

Note: Covariance of X and Y is also calculated by

$$COV(X,Y) = \sigma_{XY} = E(XY) - E(X)E(Y)$$

Now that we know how to calculate the covariance between two random variables, X and Y, let's turn our attention to seeing how the covariance helps us calculate what is called the correlation coefficient.

Correlation Coefficient

Let (X, Y) be any two random variables (discrete or continuous!) with standard deviations $\sigma_x$and σ_y respectively. The correlation coefficient of (X, Y) denoted ρ_{XY} or (the greek letter "rho") is defined as:

$$\rho_{XY} = \frac{COV(X, Y)}{\sigma_x \sigma_y}$$

NOTE

- If $\rho_{XY} = 1$, then X and Y are perfectly, positively, linearly correlated.
- If$\rho_{XY} = -1$, then X and Y are perfectly, negatively, linearly correlated.
- If $\rho_{XY} = 0$, then X and Y are completely, un-linearly correlated. That is, X and Y may be perfectly correlated in some other manner, in a parabolic manner, perhaps, but not in a linear manner.
- If $\rho_{XY} > 0$,, then X and Y are positively, linearly correlated, but not perfectly so.
- If $\rho_{XY} < 0$,, then X and Y are negatively, linearly correlated, but not perfectly so.
- If X and Y are independent then $E(XY) = E(X)\,E(Y) \Longrightarrow \rho_{XY} = 0$

2.6 REGRESSION LINES

A regression line is a straight line that describes how a response variable y changes as an explanatory variable x changes. We often use a regression line to predict the value of y for a given value of x.

(i) Regression Line X on Y

Regression line X on Y is given by

$$(x-\bar{x}) = b_{xy}(y-\bar{y})$$

Where b_{xy} is called regression coefficient X on Y and $b_{xy} = \rho\frac{\sigma_x}{\sigma_y}\ or\ \frac{Cov(X,Y)}{{\sigma_y}^2}$

$$\bar{x} = E(X) = \frac{\sum x_i}{n},\ \bar{y} = E(Y) = \frac{\sum y_i}{n}$$

(i) Regression line X on Y

Regression line X on Y is given by

$$(x-\bar{x}) = b_{xy}(y-\bar{y})$$

Where b_{xy} is called regression coefficient X on Y and $b_{xy} = \rho\frac{\sigma_x}{\sigma_y}\ or\ \frac{Cov(X,Y)}{{\sigma_y}^2}$

$$\bar{x} = E(X) = \frac{\sum x_i}{n},\ \bar{y} = E(Y) = \frac{\sum y_i}{n}$$

(ii) Regression line Y on X

Regression line Y on X is given by

$$(y-\bar{y}) = b_{yx}(x-\bar{x})$$

Where b_{yx} is called regression coefficient Y on X and $b_{yx} = \rho\frac{\sigma_y}{\sigma_x}\ or\ \frac{Cov(X,Y)}{{\sigma_x}^2}$

$$\bar{x} = E(X) = \frac{\sum x_i}{n},\ \bar{y} = E(Y) = \frac{\sum y_i}{n}$$

Note: Direct formula

i) For Correlation Coefficient

$$r_{xy}\,(\mathrm{or})\rho_{xy} = \frac{n\sum xy - \sum x\sum y}{\sqrt{\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

ii) For regression coefficient X on Y

$$b_{xy} = \frac{n\sum xy - \sum x\sum y}{n\sum y^2 - (\sum y)^2}$$

iii) For regression coefficient Y on X

$$b_{yx} = \frac{n\sum xy - \sum x\sum y}{\sum x^2 - (\sum x)^2}$$

Problems in Correlation Coefficient

1. The joint density function of X and Y is

$f(x,y) = 3(x+y), 0 \le x \le 1, 0 \le y \le 1, x+ \le 1$, Find $COV\ (x,y)$

Solution:

Given $f(x,y) \ = \ 3(x+y)$

To find $COV(x,y) \ = \ E(xy) - E(x)E(y)$

$$E(xy) = \iint\limits_R xyf(x,y)dxdy$$

$$= \int_0^1 \int_0^{1-y} xy\ 3(x+)dxdy$$

$$= 3\int_0^1 y\int_0^{1-y}(x2 + yx)dxdy$$

$$= 3\int_0^1 y\,[\tfrac{x3}{3} + \tfrac{yx\,2}{2}]_{0}^{1-y}\ dy33$$

$$= 3\int_0^1 y\,[\frac{(1-y)^3}{3} + \frac{y(1-y)^2}{2}]dy$$

$$= 3\int_0^1 y\quad\left[\frac{(1-y)^3}{3} + \frac{y^2(1-y)^2}{2}\right]dy$$

$$= 3\int_0^1 [\frac{Y}{3}(1 - y^3 - 3y + 3y^2] + \frac{y^2(1 + y^2 - 2Y)}{2}]dy$$

$$= 3\int_0^1 \frac{1}{3}(y - y^4 - 3y^2 + 3y^3) + \frac{(y^2 + y^4 - 2y^3)}{2}]dy$$

$$= \left[\left(\frac{y^2}{2} - \frac{y^5}{5} - 3\frac{y^3}{3} + 3y^3\right) + \frac{3}{2}\left(\frac{y^3}{3} + \frac{y^5}{5} - 2\frac{y^4}{4}\right)\right]$$

$$= \left[\frac{1}{2} - \frac{1}{5} - 1 + \frac{3}{4} + \frac{-1}{2} + \frac{3}{10} - \frac{3}{4}\right] = \frac{1}{10}$$

$$E(xy) = \frac{1}{10}$$

$$= f_x\,(x) = \int_{y=0}^{y=1-x} f(x,y)\,dy$$

$$= \int_0^{1-x} 3\,(x+y)\,dy$$

$$= 3\left[xy + \frac{y^2}{2}\right]\begin{matrix}1-x\\0\end{matrix}\ = \ 3\ \left[x(1-x) + \frac{(1-x)^2}{2}\right]$$

$$3\left[x - x^2 + \frac{1+\ x^2\ - 2x}{2}\right]$$

$$= \frac{3}{2}\,[2x - 2x^2 + 1 + x^2 - 2x] = \frac{3}{2}\,[1 - x^2]$$

$$f_y(y) = \int_0^{1-x} f(x,y)dx = \int_0^{1-y} 3(x+y)dx$$

$$= 3\left[\frac{x^2}{2} + yx\right] \begin{matrix} 1-y \\ 0 \end{matrix}$$

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

$$f_y(y) = \frac{3}{2}\;[1 - y^2]$$

$$E(x) = \int_0^1 x f_x(x)\,dx = \int_0^1 x\frac{3}{2}\,[1 - x^2]\,dx = \frac{3}{2}\left[\frac{x^2}{2} - \frac{x^4}{4}\right]_0^1 = \frac{3}{4} - \frac{3}{8} = \frac{3}{8}$$

$$E(y) = \int_0^1 y f_y(y)\,dy = \int_0^1 y\frac{3}{2}\;[1 - y^2]\,dy = \frac{3}{2}\left[\frac{y^2}{2} - \frac{y^4}{4}\right]_0^1 = \frac{3}{4} - \frac{3}{8} = \frac{3}{8}$$

$$Cov(x,y) = E(xy) - E(x)E(y) = \frac{1}{10} - \left(\frac{3}{8}\right)\left(\frac{3}{8}\right) = \frac{1}{10} - \frac{9}{64} = -\frac{13}{320}$$

2. The Joint P.D.F of (x,y) is given by f(x,y)= $\begin{cases} 2 - x - y, & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0, & elsewhere \end{cases}$.

Find $r(x,y)$

Solution:

$$\boldsymbol{f_x(x) = \int f(x,y)dy} = \int_{\mathbf{0}}^{\mathbf{1}} (2 - x - y)dy = \left[2y - xy - \frac{y^2}{2}\right]_0^1$$

$$= \left(2 - x - \frac{1}{2}\right) = \frac{3}{2} - x$$

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

$$= \left(2 - y - \frac{1}{2}\right) = \frac{3}{2} - y$$

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

$$= \frac{3}{4} - \frac{1}{3} = \frac{9-4}{12} = \frac{5}{12}$$

$$E(y) = \int_0^1 y f_y(y)\, dy = \int_0^1 y\left(\frac{3}{2} - y\right) dy = \left[\frac{3}{2}\frac{y^2}{2} - \frac{y^3}{3}\right]_0^1$$

$$= \frac{3}{4} - \frac{1}{3} = \frac{9-4}{12} = \frac{5}{12}$$

$$E(x^2) = x^2 f_x(x)dx = \int_0^1 x^2\left(\frac{3}{2} - x\right) dx = \left[\frac{3}{2}\frac{x^3}{3} - \frac{x^4}{4}\right]_0^1$$

$$= \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

$$E(y^2) = y^2 f_y(y)dy = \int_0^1 y^2\left(\frac{3}{2} - y\right) dy = \left[\frac{3}{2}\frac{y^3}{3} - \frac{y^4}{4}\right]_0^1$$

$$= \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

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

$$= \int_0^1 \left(y - \frac{y}{3} - \frac{y^2}{2}\right) dy$$

$$= \left[\frac{y^2}{2} - \frac{y^2}{6} - \frac{y^3}{6}\right]_0^1 = \left[\frac{1}{2} - \frac{1}{6} - \frac{1}{6}\right] = \left[\frac{1}{2} - \frac{1}{3}\right] = \frac{1}{6}$$

$$Cov(x,y) = E(xy) - E(x)E(y) = \frac{1}{6} - \frac{5}{12}\frac{5}{12} = \frac{1}{6} - \frac{25}{144} = -\frac{1}{144}$$

$$\sigma_x^2 = E(x^2) - \left(E(x)\right)^2 = \frac{1}{4} - \left(\frac{5}{12}\right)^2 = \frac{1}{4} - \frac{25}{144} = \frac{36-25}{144} = \frac{11}{144}$$

$$\sigma_y^2 = E(y^2) - \left(E(y)\right)^2 = \frac{1}{4} - \left(\frac{5}{12}\right)^2 = \frac{1}{4} - \frac{25}{144} = \frac{36-25}{144} = \frac{11}{144}$$

$$\rho(xy) = \frac{-\frac{1}{144}}{\sqrt{\frac{11}{144}}\sqrt{\frac{11}{144}}} = -\frac{1}{11}$$

3. The joint probability density function of (x,y) is

$$f(x,y) = \begin{cases} x+y & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & Elsewhere \end{cases}$$

Compute $\rho(x,y)$

Solution:

$$fx^x = \int_0^1 f(x,y)\,dy = \int_0^1 (x+y)\,dy = \left(xy + \frac{y^2}{2}\right)_0^1 = x + \frac{1}{2}$$

$$f_y\,(y) = \int_0^1 f(x,y)\,dx = \int_0^1 (x+y)\,dx = \left(y + \frac{1}{2}\right)$$

$$E(x) = \int_0^1 \left(x+\frac{1}{2}\right) dx = \int_0^1 \left(x^2+\frac{x}{2}\right) dx = \left(\frac{x^{23}}{3}+\frac{x^2}{4}\right) = \frac{1}{3}+\frac{1}{4} = \frac{7}{12}$$

$$E(y) = \int_0^1 y\left(y+\frac{1}{2}\right) dy = \frac{7}{12}$$

$$E(x^2) = \int_0^1 x^2\left(x+\frac{1}{2}\right) dx = \left[\frac{x^4}{4}+\frac{1}{2}\,\frac{x3}{3}\right]_0^1 = \left(\frac{6+4}{24}+\frac{10}{24}\right) = \frac{5}{12}$$

$$E(y^2) = \int_0^1 y^2\, f_y\,(y)dy = \frac{5}{12}$$

$$E(xy) = \iint xy\, f(x,y)dxdy$$

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

$$= \int_0^1 y\,\left(\frac{1}{3}+\frac{y}{2}\right)dy = \left[\frac{y^2}{6}+\frac{y^3}{6}\right]_0^1$$

$$\frac{1}{6}+\frac{1}{6} = \frac{2}{6} = \frac{1}{3}$$

$$\rho(x,y) = \frac{Cov(x,y)}{\sigma_x\,\sigma_y}$$

$$Cov(x,y) = E(xy) - E(x) - E(y)$$

$$= \frac{1}{3} - \left(\frac{7}{12}\right)\left(\frac{7}{12}\right) = \frac{1}{3} = \frac{49}{144}$$

$$= \frac{48-49}{144} = \frac{-1}{144}$$

$$\sigma_x^2 = E(x^2) - \left(E(x)\right)^2 = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{5}{12} - \frac{49}{144}$$

$$= \frac{60-49}{144} = \frac{11}{144}$$

$$\sigma_y^2 = \frac{11}{144}$$

$$\rho(x,y) = \frac{\frac{-1}{144}}{\frac{11}{144}} = \frac{-1}{11}$$

4. Compute the coefficient of correlation between X and Y using the following data

X	1	3	5	7	8	10
Y	8	12	15	17	18	20

Solution:

	x	y	X^2	Y^2	xy
	1	8	1	64	8
	3	12	9	144	36
	5	15	25	225	75
	7	17	49	289	119
	8	18	64	324	144
	19	20	100	400	200
Total =	34	90	248	1446	582

$$\eta = 6$$

$$\sum x = 34$$

$$\sum y = 90$$

$$\sum x^2 = 248$$

$$\sum y^2 = 1446$$

$$\sum xy = 582$$

$$r_{xy} = \frac{n\sum xy - \sum x \sum y}{\sqrt{\sum x^2 - (\sum x)^2}\sqrt{n\sum y^2 - (\sum y)^2}}$$

$$= \frac{(6 \times 582) - (34)(90)}{\sqrt{(6 \times 248) - (34)^2}\sqrt{n(6 \times 1446) - (90)^2}}$$

$$r_{xy} = 0.9879$$

5. The regression line of x on y is $3x + y = 10$ and the regression line of y on x is $3x + 4y = 12$, find mean of x, y. Also find $r(x, y)$

Solution:

Regression line x on y $\Rightarrow 3x + y = 10$

$$3x = 10 - y$$

$$x = \frac{10}{3} - \frac{y}{3}$$

$$\therefore\ b_{xy} = \frac{-1}{3}$$

Regression line y on x $\Rightarrow 3x + 4y = 12$

$$3x = 10 - y$$

$$y = \frac{12}{4} - \frac{3x}{4}$$

$$\therefore\ b_{xy} = \frac{-3}{4}$$

$$r = \pm\sqrt{b_{xy}\ b_{yx}}$$

$$= -\sqrt{\left(\frac{-1}{3}\right)\left(\frac{-3}{4}\right)}$$

$$= -\sqrt{\frac{1}{4}} = \frac{-1}{2}$$

Since the lines are passing through the posts $(\overline{x}, \overline{y})$

$$3\overline{x} + \overline{y} = 10$$

$$3\overline{x} + 4\overline{y} = 12$$

Solving, we get

$$\overline{x} = \frac{28}{9}, \qquad \overline{y} = \frac{2}{3}$$

Exercises:

1. The joint p.d.f. of (x, y) is given by $f(x,y) = \begin{Bmatrix} \frac{3}{2}(x^2+y^2) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & Elsew\ here \end{Bmatrix}$

   Find the correlation coefficient

2. Let x and y be random variables in the p.d.f.

$$\begin{Bmatrix} \frac{x+y}{3} & 0 \le x \le 1, 0 \le y \le 2 \\ 0 & Elsewhere \end{Bmatrix}$$

Find $r(x, y)$

3. Find the coefficient of correlation between X and Y using the following data

X	5	10	15	20	25
Y	16	19	23	26	30

Ans: -0.9907

4. Find the correlation coefficient between X and Y

X	1	2	3	4	5	6	7	8	9
Y	9	8	10	12	11	13	14	16	15

Also find the reg. lines and estimate the value of y when $x = 6.2$.

5. Find the coefficient of correlation and obtain the lines of regression from data given below.

X	65	66	67	67	68	69	70	72
Y	67	68	65	68	72	72	69	71

Problems.

1. The lines of regression are $8x - 10y + 66 = 0, 40x - 18y - 214 - 0.$ The variance of x is 9. Find the covariance between $x + y$ and $x - y$

Solution:

If x and y are independent, then $E(xy) = E(x)\,E(y)$

$$Cov(x, y) = 0$$

$$Var(x) = 36, Var(y) = 16$$

Let $u = x + y, v = (x = y)$

To find $r(u, v) = \frac{Cov\ (u,v)}{\sigma_u\ \sigma_v}$

$$Cov\ (u, v) = E(uv) - \ E(u)E(v)$$

$$= E[(x + y)(x - y)] - \ E\ (x + y)(x - y)$$

$$= E(x^2 - \ y^2) - \ \left(E(x) + \ E(y)\right) E(x) - \ E(y)$$

$$= [E(x^2) - \ E(y^2)] - \ [(E(x)^2) - \ E(y^2)]$$

$$= \left[E(x^2) - \ \left(E(x)\right)^2 - \ E(y^2) - \ \left(E(y)\right)^2\right]$$

$$= Var(x) - \ Var(y) = 36 - 16 = 20$$

$$\sigma_u^2\ = E(u^2) - \ \left(E(u)\right)^2$$

$$= E(x + y)^2 - \ \left(E(x + y)\right)^2$$

$$= \ [E(x^2) + \ E(y^2) + 2E(xy)] - \left[\left(E(x)\right)^2 + \ \left(E(y)\right)^2 + 2E(x)E(y)\right]$$

$$= E(x^2) - \left(E(x)\right)^2 + \ E(y^2) - \ \left(E(y)\right)^2$$

$$\sigma_u^2 = Var(x) + \ Var(y) = 36 + 16 = 52$$

$$\sigma_v^2 = E(v^2) - \left(E(v)\right)^2$$

$$= E(x-y)^2 - \left(E(x-y)\right)^2$$

$$= [E(x^2) + E(y^2) - 2E(xy)] - \left[\left(E(x)\right)^2 + \left(E(y)\right)^2 - 2E(x)E(y)\right]$$

$$= E(x^2) - \left(E(x)\right)^2 + E(y^2) - \left(E(y)\right)^2$$

$$\sigma_v^2 = Var(x) + Var(y) = 36 + 16 = 52$$

$$r(u,v) = \frac{20}{\sqrt{52}\sqrt{52}} = \frac{20}{52} = \frac{5}{13}$$

2. If X, Y and Z are uncorrelated RV's with zero mean and standard deviation 5, 12 and 9 respectively, and if $U = X + Y$ and $V = Y + Z$, find the correlation coefficient between U and V.

Solution:

$$E(x) = E(y) = E(z) = 0,\ \sigma_x = 5,\ \sigma_z = 12,\ \sigma_z = 9$$

$$Var(x) = E(x^2) - \left(E(x)\right)^2 = (5)^2$$

$$\Rightarrow E(x^2) = 25$$

$$E(y^2) = 144, E(z^2) = 8$$

X & Y are uncorrelated, $\therefore E(xy)=\mathrm{E}(x)\mathrm{E}(y) = 0$

$$E(yz)=\mathrm{E}(y)\mathrm{E}(z) = 0$$

$$E(xz)=\mathrm{E}(x)\mathrm{E}(z) = 0$$

$$\mathrm{E}(u) = \mathrm{E}(x+y) = E(x) + E(y) = 0$$

$$\mathrm{E}(v) = 0$$

$$\mathrm{E}(uv) = E\big((x+y)\,(y+z)\big) = E(xy + xz + y^2 + yz)$$

$$= E(xy) + E(xz) + E(y^2) + E(yz) = E(y^2)$$

$$E(uv) = 144$$

$$Var(u) = Var(x+y) = Var(x) + Var(y) = 25 + 144 = 169$$

$$Var(v) = Var(y+z) = Var(y) + Var(z) = 144 + 81 = 225$$

$$\therefore Cor\,(u,v) = \frac{Cor(uv)}{\sigma_u\ \sigma_v}$$

$$\frac{144-0}{\sqrt{169}\sqrt{225}} = \frac{144}{13\ \times 15} = \frac{48}{65}$$

Exercises:

1. If x and y are independent random variables with mean 2 and 3 and variance 1 and 2. Find the mean and variance of $Z = 2x - 5y$.

2. If x, y, and z are uncorrelated random variable with same variance. Find the correlation coefficient between $(x+y)$ and $(y+z)$
3. If x and y are uncorrelated random variables with zero means and same variance, prove that $u = x\cos\propto\ + y\sin\propto$ and $V = x\sin\propto - y\cos\propto$ are also uncorrelated.

4. The random variables x, y are defined as $Y = 4x + 9$. Find $r(x,y)$

5. The regression equation for variables x and y are $7x - 3y - 18 = 0$ and $4x - y - 11 = 0$.

   i) What is the mean of x and y?

ii) Find the correlation coefficient in between x and y.

Ans: $\overline{x} = 3, \overline{y} = 1$ & r(x,y)=0.7638

Problems in Correlation Coefficient for Discrete Random Variable

1. The Joint probability mass function of (x, y) is given by $p(x, y) = k(x, y), x = 1,2,3,4\,, y = 1,2,3$. Find $r(x, y)$

Solution

Given probability mass function $p(x, y) = k(x, y),\ x = 1,2,3,4\,,\ y = 1,2,3$

Y X	1	2	3	P(x=x$_i$)	xP(x)	x^2P(x)
1	2k	3k	4k	9k	9k	9k
2	3k	4k	5k	12k	24k	48k
3	4k	5k	6k	15k	45k	135k
4	5k	6k	7k	18k	72k	288k
P(y=yi)	14k	18k	24k	$\sum\sum P_{ij} = 54k$	E(x)=150k	E(x^2)=480k
yP(y)	14k	36k	66k	E(y)=116k		
y^2P(y)	14k	72k	198k	E(y^2)=284k		

$$\sum\sum P_{ij} = 1 \Rightarrow 54k = 1 \Longrightarrow k = \frac{1}{54}$$

$$E(x) = \sum \mathrm{xP(x)} = 150k = \frac{150}{56}$$

$$E(y) = \sum \mathrm{yP(y)} = 116k = \frac{116}{56}$$

$$E(x^2) = \sum x^2 \mathrm{P(x)} = 480k = \frac{480}{56}$$

$$E(y^2) = \sum y^2 \mathrm{P(y)} = 284k = \frac{284}{56}$$

$$E(xy) = \sum\sum \mathrm{xyP(xy)}$$

$$= 2k + 6k + 12k + 6k + 16k + 30k + 12k + 54k + 20k + 48 + 84k$$

$$\Longrightarrow\ 320k = \frac{320}{54}$$

$$cov(x, y) = E(xy) - E(x)E(y) = \frac{320}{54} - \frac{150}{54}\frac{116}{54} = \frac{160}{27} - \frac{75}{27}\frac{58}{27}$$

$$= \frac{4320 - 4350}{729} = -\frac{30}{729}$$

$$\sigma_x^2 = E(x^2) - [E(x)]^2 = \frac{480}{54} - \left[\frac{150}{54}\right]^2 = \frac{240}{27} - \left[\frac{75}{27}\right]^2$$

$$= \frac{6480 - 5625}{729} = \frac{855}{729}$$

$$\sigma_y^2 = E(y^2) - [E(y)]^2 = \frac{284}{54} - \left[\frac{116}{54}\right]^2 = \frac{160}{27} - \left[\frac{58}{27}\right]^2$$

$$= \frac{3834 - 3364}{729} = \frac{470}{729}$$

$$r(x, y) = \frac{-\frac{30}{729}}{\sqrt{\frac{855}{729}\frac{470}{729}}} = -\frac{30}{\sqrt{401850}} = -0.047$$

Exercises

1. Find the joint probability density function of (x, y) is

$$f(x, y) = x + \frac{y}{21}\ , x = 1,2,3,\ \ y = 1,2$$

2. Find $r(x, y)$ for the following

X Y	5	15
10	0.2	0.4
20	0.3	0.1

2.7 TRANSFORMATION OF RANDOM VARIABLES

If (x, y) is a two-dimensional random variable with probability density function $f(x, y)$ and if $z = g(x, y)$ and $w = h(x, y)$ are two other random variable, then the joint probability density function of (z, w) is given by $f(z, w) = |J| f(x, y)$ where

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix}$$

Note:

If only one relation $z = g(x, y)$ is given, then assume the other relation as $w = y$

1. If X and Y are independent random variable with probability density function $e^{-x}, x \geq 0 \; and \; e^{-y}, y \geq 0$ respectively. Find the density function of $u = \frac{x+y}{2}$

Solution

Given $f_x(x) = e^{-x}, x \geq 0\,,\quad f_y(y) = e^{-y}, y \geq 0$

$$\Longrightarrow f(x,y) = f_x(x)f_y(y) = e^{-(x+y)} \quad (\because\ \text{X, Y are independent})$$

Let $v = y$

$$u = \frac{x+y}{2} \Rightarrow x = 2u - y \Rightarrow x = 2u - v$$

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 2 & -1 \\ 0 & 1 \end{vmatrix} = 2$$

$$f(u,v) = 2e^{-(x+y)} = 2e^{-2u}, u \geq 0, v \geq 0, \frac{v}{2} \leq u$$

To find $f_u(u)$

Range space

$v = y \geq 0$

$x = 2u - v \geq 0, 2u \geq v$

$\Longrightarrow v \leq 2u \Rightarrow \frac{v}{2} \leq u$

$u \rightarrow \frac{v}{2}\ to\ \infty$

$v \rightarrow 0\ to\ 2u$

$$f_u(u) = \int f(u,v)\,dv$$

$$= \int_0^{-2u} 2e^{-2u} dv = 2e^{-2u}[v]_0^{-2u} = 4ue^{-2u}$$

$$f_u(u) = 4ue^{-2u}$$

2. If X and Y are independent random variables with $f_x(x) = e^{-x}u(x)$ and $f_y(y) = 3e^{-3y}u(y)$. Find $f_z(z)\ if\ z = \frac{x}{y}$

Solution

Here X and Y are independent, $\therefore f(x, y) = f_x(x) f_y(y) = 3e^{-(x+3y)}, x, y \geq 0$

$$\text{Let W} = \text{Y}$$

$$Z = \frac{X}{Y} \Rightarrow \text{X} = \text{ZW}$$

$$J = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} w & z \\ 0 & 1 \end{vmatrix} = w$$

The joint p.d.f. of $(\text{z}, \text{w}) = \text{f}(\text{z}, \text{w}) = |J| f(x, y) = w3e^{-(x+3y)} = w3e^{-(zw+3w)}$

$$\text{f}(\text{z}, \text{w}) = 3we^{-w(z+3)}, z, w \geq 0$$

$$f_z(z) = \int_0^\infty f(z, w)\, dw$$

$$= \int_0^\infty 3we^{-w(z+3)} dw$$

Put $w(z+3) = t \;\; \Rightarrow \;\; (z+3)dw = dt$

$$= 3 \int_0^\infty \frac{t}{z+3} e^{-t} \frac{dt}{z+3} = \frac{3}{(z+3)^2} [(-te^{-t}) - e^{-t}]_0^\infty$$

$$= \frac{3}{(z+3)^2} [0+1] = \frac{3}{(z+3)^2}$$

$$f_z(z) = \frac{3}{(z+3)^2}, z \geq 0$$

3. If X and Y each follow an exponential distribution with parameter 1 and are independent. Find the probability density function of $U = X - Y$

Solution:

X and Y follow exponential distribution with λ=1

$\therefore f(x) = e^{-x}, x \geq 0, \quad f(y) = e^{-y}, y \geq 0$

$\therefore f(x, y) = f(x)f(y) = e^{-(x+y)}, x, y \geq 0$, (since x,y are independent)

Given $\mathrm{U = X - Y} \Rightarrow X = U + Y \Rightarrow X = U + V$

Assume $V = Y \Rightarrow Y = V$

We know that $\mathrm{f(u, v)} = |J| f(x, y)$

Where

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1$$

$$f(u, v) = 1. e^{-(x+y)} = e^{-(u+v+v)}$$

$$f(u, v) = e^{-(u+2v)}$$

To find $f_u(u)$

Range space: $x \geq 0, u + v \geq 0 \Rightarrow u \geq -v$

$y \geq 0, v \geq 0$

Case i) $u < 0$

$$f_u(u) = \int_{-u}^{\infty} e^{-(u+2v)} dv$$

$$= e^{-u} \int_{-u}^{\infty} e^{-2v} dv$$

$$= e^{-u} \left[\frac{e^{-2v}}{-2}\right]_{-u}^{\infty}$$

$$= e^{-u} \left[0 + \frac{e^{2u}}{2}\right]$$

$$f_u(u) = \frac{e^u}{2}, u < 0$$

Case ii) $u > 0$

$$f_u(u) = \int_0^{\infty} e^{-(u+2v)} dv$$

$$= e^{-u} \left[\frac{e^{-2v}}{-2}\right]_0^{\infty}$$

$$= e^{-u} \left[0 + \frac{1}{2}\right]$$

$$f_u(u) = \frac{e^{-u}}{2} \, , u > 0$$

$$\therefore f_u(u) = \begin{cases} \frac{e^{u}}{2}, & u < 0 \\ \frac{e^{-u}}{2}, & u > 0 \end{cases}$$

4. If X and Y are independent random variables with p.d.f $e^{-x}, x \geq 0 \; and \; e^{-y}, y \geq 0$ respectively, find the density function of $U = \frac{X}{X+Y}\, , V = X + Y$. Are U and V independent?

Solution:

Given $f(x) = e^{-x}, x \geq 0, \; f(y) = e^{-y}, y \geq 0$

$\therefore f(x,y) = f(x)f(y) = e^{-(x+y)}\, , x, y \geq 0,$ (since x,y are independent)

Given $U = \frac{X}{X+Y}\, , V = X + Y.$

$\Longrightarrow U = \frac{X}{V}\, , X = UV.$ and $, V = UV + Y \Longrightarrow Y = V - UV \Longrightarrow Y = V(1 - U)$

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} = v(1-u) + vu$$

$$|J| = v$$

$$f(u,v) = v.e^{-(x+y)} = ve^{-(v)}$$

To find $f_u(u), f_v(v)$

Range space: $x \geq 0, \Longrightarrow uv \geq 0 \Rightarrow u \geq 0, v \geq 0$

$$y \geq 0, \Longrightarrow v \geq 0, 1 - u \geq 0, u \leq 1$$

$$\Longrightarrow v \geq 0, 0 \leq u \leq 1$$

$$f_u(u) = \int_0^\infty ve^{-(v)} dv$$

$$= [(-ve^{-v}) - e^{-v}]_0^\infty = 1$$

$$f_u(u) = 1, \; 0 \leq u \leq 1$$

$$f_v(v) = \int_0^1 ve^{-(v)} du$$

$$= ve^{-v}[u]_0^1 = ve^{-v}, \; v \geq 0$$

Also, $f_u(u) f_v(v) = 1.ve^{-v} = ve^{-v} = f(u,v)$

$\therefore$ U and V are independent

Exercises

1. If the joint probability density function of (X,Y) is given by $f(x,y) = x + y, 0 \leq x, y \leq 1$, find the probability density function of $U = XY$
2. If X and Y each follow an exponential distribution with parameter 2 and 3 respectively and are independent, find the probability density function of $U = X + Y$

2.8 Practice Quiz

1. The joint probability mass function of two dimensional RV (X,Y) is given by P(x,y) = ***k(x+y),*** *x*=1,2,3 and *y*=1,2 where *k* is a constant. Find the value of k.

a) 21 b) 22 c) 1/21 d) 1/22

2. Find the marginal probability mass function of X , of the two dimensional discrete random variable (X,Y) whose joint probability mass function is

$$P[X=x,Y=y]=\begin{cases}\dfrac{x+2y}{18}, & x=1,2;\ y=1,2\\ 0 & ,otherwise\end{cases}.,$$

a) P(X=1) = 8/18, P(X=2) = 10/18 b) P(X=1) = 6/18, P(X=2) = 12/18 c) P(X=1) = 4/18, P(X=2) = 10/18 d) P(X=1) = 6/18, P(X=2) = 12/18

3. Find the marginal probability mass function of Y, of the two dimensional discrete random variable (X,Y) whose joint probability mass function is

$$P[X=x,Y=y]=\begin{cases}\dfrac{x+2y}{18}, & x=1,2;\ y=1,2\\ 0 & ,otherwise\end{cases}.,$$

a) P(Y=1) = 7/18, P(Y=2) = 11/18 b) P(Y=1) = 5/18, P(Y=2) = 13/18
c) P(Y=1) = 2/18, P(Y=2) = 16/18 d) P(Y=1) = 4/18, P(Y=2) = 12/18

4. The joint probability mass function of X and Y is

X / Y	0	1	2
0	0.1	0.04	0.02
1	0.08	0.20	0.06
2	0.06	0.14	0.30

Find $P(X \le 1, Y \le 1)$

a)0.32 b) 0.42 c) 0.34 d) 0.53

5. If X, Y are independent Random variables then

a) E(XY) = 0 b) E(X) =0 c) E(XY) = E(X) E(Y)
d) E(XY) = E(X)+

6. If the joint p.d.f of (X,Y) is given by $f(x,y) = \frac{1}{8}(6 - x - y), 0 \le x \le 2, 2 \le y \le 4$, then the marginal density of Y

a) $\frac{(3-x)}{4}$ b) $\frac{(5-y)}{4}$ c) 2y d) 5x

7.	A perfect positive correlation is signified by: (a) 0 (b) -1 (c) +1 (d) -1 to +1
8.	If *X* is a random variable '*a*' and '*b*' are constants then *E(aX + b)* is a. E(*a* X)+E(*b*) b. E(*a* X)+*b* c. *a* E(X)+*b* d. None of these
9.	The slope of the regression line of Y on X is also called the: (a) Correlation coefficient of X on Y (b) Correlation coefficient of Y on X (c) Regression coefficient of X on Y (d) Regression coefficient of Y on X
10.	The cumulative probability distribution shows the probability a. that a random variable is less than or equal to a particular value. b. of two or more events occurring at once. c. of all possible events occurring d. that a random variable takes on a particular value given that another event has happened.
11.	If X, Y are independent Random variables then a) E(XY) = 0 b) E(X) =0 c) E(XY) = E(X) E(Y) d) E(XY) = E(X)+E(Y)
12	If X, Y are independent Random variables then the correlation coefficient of (X,Y) a) r(X,Y) = 0 b) r(X,Y) =1 c) r(X,Y) = -1 d) none
13	. Find the value of K, if the joint density function of (x,y) is give by $f(x,y) = \left\{\begin{matrix} K(1-x)(1-y) & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & otherwise \end{matrix}\right\}$ a) 1/32 b) 1/23 c)32 d)1/64
14	If the joint density function of (x,y) is give by $f(x,y) = \left\{Kxye^{-(x^2+y^2)} \quad x > 0, y > 0\right\}$ Find the value of K a)1/4 b) 4 c)32 d)1/64
15	Given the joint p.d.f of (X,Y) as f(X,Y)=1/6, 0<X<2, 0<Y<3, determine the marginal density function of X. a)1/2 b) 4 c) 2 d)1/6
16	Consider $f(x,y) = \begin{cases} 2x & 0 < x < 1,\ 0 < y < 1 \\ 0 & otherwise \end{cases}$. Find marginal density function of X a)3x b) 2x c) 2 d) 1/2x

17. Consider $f(x,y)=\begin{cases}3y & 0<x<1,\ 1<y<2\\ 0 & otherwise\end{cases}$. Find all the marginal density function of X

a)3x b) 9 c) 2

18. The two regression equations are $4x-5y+33=0\ and\ 20x-9y=107$. find the means of x,y

a)$\bar{x}=\ 13\ \ ,\bar{y}=12$ b) $\bar{x}=13\ \ ,\bar{y}=17$ c) $\bar{x}=12\ \ ,\bar{y}=6$ d) $\bar{x}=\ 8\ \ ,\bar{y}=12$

19. If X and Y are independent random variables with variance 2and 3. Find the variance of 3 X +4Y .

a)31 b) 55 c) 66 d) 60

20. If the independent random variables X and Y have the variances 36 an respectively. Find the covariance between $(X+Y)$ and $(X-Y)$

a)30 b) 12 c) 20 d) 60

1	2	3	4	5	6	7	8	9	10
c	a	a	b	c	b	c	c	d	a
11	12	13	14	15	16	17	18	19	20
a	a	a	b	a	b	b	b	c	c

2.9 Assignment-I

Two dimensional D. R. V & C.R.V

1. The two dimensional random variable (X,Y) has the joint probability mass function $f(x,y)=\dfrac{x+2y}{27}$, $x = 0,1,2;\quad y = 0,1,2.$ (i) Find all the marginal distributions of X and Y (ii) probability distribution of (X+Y) (iii) $P(X + Y \leq\ 2)$.

2. The joint pdf of a two dimensional random variable (X,Y) is given by

   $f(x,y)=xy^2+\dfrac{x^2}{8},0\le x\le 2;0\le y\le 1.$

   Compute (1) $P[X>1]$ (2) $P\left[Y<\dfrac{1}{2}\right]$ (3) $P[X<Y]$

3. Two dimensional random variable (X,Y) have the joint probability density function $f(x,y)=\begin{cases}8xy, & 0<x<y<1\\ 0\ , & \text{elsewhere}\end{cases}$.

   i) Find the Marginal and Conditional distributions

   ii) Are X and Y independent iii) Find $P\left(X<\dfrac{1}{4}\middle/\dfrac{1}{2}<Y<\dfrac{3}{4}\right)$.

4. For the following Bivariate probability distribution of (x,y), find

x/y				Y			
		1	2	3	4	5	6
X	0	0	0	1/32	2/32	2/32	3/32
	1	1/16	1/16	1/8	1/8	1/8	1/8
	2	1/32	1/32	1/64	1/64	0	2/64

   Find all the marginal and conditional probability distributions.
   Also find (i)P(X≤1) (ii)P(Y≤3) (iii)P(X≤1/ Y≤3) (iv)P(X+Y≤ 4).

5. If the joint distribution function of a two dimensional random variable (X,Y) is given by

$$F(x,y) = \left\{\begin{matrix}(1-e^{-x})(1-e^{-y}) & x>0\quad and\quad y>0\\ 0 & elsewhere\end{matrix}\right\}$$

Find

i) marginal density of X and Y
ii) $P(X < 1/Y < 3)$ iii) $P(X > Y)$ iv) $P(X + Y < 1)$

Assignment-II

Correlation & Regression

1. If the independent random variables X and Y have variances 36 and 16 respectively. Find the correlation coefficient r_{UV} where $U = X + Y$ and $V = X - Y$.

2. A joint probability mass function of the discrete random variable X and Y is given as $P(X = x, = y) = \left\{ \frac{x+y}{32}, \; x = 1, 2; \; y = 1, 2, 3, 4. \right.$ Compute the

   a) Cov(X,Y) b) Correlation of X and Y

3. The two lines of regression are $8x - 10y + 66 = 0$ and $40x - 18y - 214 = 0$. The variance of x is 9. Find

   i) The mean values of 'x' and 'y'.

   ii) Correlation co-efficient between 'x' and 'y'

4. Given that $X = 4Y + 5$ and Y=kX+4 are regression lines of X on Y and Y on X respectively. Show that $0 \le k \le \frac{1}{4}$. If $k = \frac{1}{16}$, find the means of X and Y and the correlation coefficient r_{XY}.

5. If the joint pdf of (X,Y) is given by $f(x, y) = e^{-(x+y)}, x > 0, \; y > 0.$ Prove that X and Y are uncorrelated

2.10 Part A Questions & Answers

<table>
<tr><th>Q. No.</th><th>Questions &Answers</th><th>K Level</th><th>CO</th></tr>
<tr><td>1.</td><td>
The joint probability mass function of a bivariate discrete random variable (X,Y) is given by the following table:
<table>
<tr><td>x
y</td><td>1</td><td>2</td><td>3</td></tr>
<tr><td>1</td><td>0.1</td><td>0.1</td><td>0.2</td></tr>
<tr><td>2</td><td>0.2</td><td>0.3</td><td>0.1</td></tr>
</table>
Find the marginal probability mass function of X and Y.

Solution: The joint probability mass function of (X,Y) is
<table>
<tr><td>x
y</td><td>1</td><td>2</td><td>3</td><td>$p_y(y)$</td></tr>
<tr><td>1</td><td>0.1</td><td>0.1</td><td>0.2</td><td>0.4</td></tr>
<tr><td>2</td><td>0.2</td><td>0.3</td><td>0.1</td><td>0.6</td></tr>
<tr><td>$p_x(x)$</td><td>0.3</td><td>0.4</td><td>0.3</td><td>1</td></tr>
</table>
The marginal probability mass function of X is,
When
$x=1,\quad p_x(1)=\sum_{y_j} p_{xy}(1,y_j)=p_{xy}(1,1)+p_{xy}(1,2)=0.1+0.2=0.3$
Similarly, $P_x(2)=0.4 \quad P_x(3)=0.3$
The marginal p.d.f. of Y is given by
$P_Y(1)=0.4$, $P_Y(2)=0.6$ (follows from the table)
</td><td>K1</td><td>CO2</td></tr>
<tr><td>2.</td><td>
If the joint p.d.f. of (X, Y) is given by $f(x,y)=6e^{-2x-2y}$ $,x\geq 0, y\geq 0$ and $f(x,y)=0$ otherwise.
Find the marginal density of X and the conditional density of Y given $X=x$.
Solution: The marginal density function of X is given by
$$f_x(x)=\int_{-\infty}^{\infty} f(x,y)dy=\int_{0}^{\infty} 6e^{-2x-3y}dy$$
</td><td>K1</td><td>CO2</td></tr>
</table>

	$= 6\int_0^{\infty} e^{-2x}.e^{-3y}dy$ $6e^{-2x}\int_0^{\infty} e^{-3y}dy = 6e^{-2x}\left[\frac{e^{-3y}}{-3}\right]_0^{\infty} = 2e^{-2x}, x \ge 0$ The Conditional density of y given $X = x$ is given by $f(y/x) = \frac{f(x,y)}{f_x(x)} = \frac{6e^{-2x-3y}}{2e^{-2x}} = 3e^{-3y}, y \ge 0$		
3.	The joint p.d.f. of (X,Y) is given by $f(x,y) = e^{-(x+y)}\ x \ge 0, y \ge 0$ and $f(x,y) = 0$ otherwise. Are X and Y independent? Why? **Solution:** Two random variables X and Y are independent if and only if $f(x,y) = f_x(x).f_y(y)$. Given that $f(x,y) = e^{-(x+y)} \quad = e^{-x}.e^{-y} = f_x(x).f_y(y)$ i.e., f(x, y) being written as a product of functions of X and Y . $\therefore$ X and Y are independent	**K2**	**CO2**
4.	. If the joint p.d.f. of (X,Y) is given by $f(x,y) = 2 - x - y, 0 \le x < y \le 1$ and $f(x,y) = 0$ otherwise Find E(X) and E(Y). **Solution:** $E[X] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f(x,y)dxdy = \int_0^1\int_0^y x(2-x-y)dx\,dy$ $= \int_0^1\left(y^2 - \frac{5y^3}{6}\right)dy = \frac{1}{8}$ Similarly, $E[y] = \frac{1}{8}$.	**K1**	**CO2**
5.	The joint p.d.f. of (X,Y) is given by $f(x,y) = \begin{cases} 24xy & ,x > 0, y > 0, x + y \le 1 \\ 0 & ,\text{elsewhere} \end{cases}$ Find the Conditional mean and the variance of Y given X . **Solution:** $f_x(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^{1-x} 24\,xy\,dy = 12x(1-x)^2, 0 < x < 1$ Now $f(y/x) = \frac{f(x,y)}{f_x(x)} = \frac{2y}{(1-x)^2}, 0 < y < 1 - x.$		

	$E[Y/X=X]=\int_0^{1-x} y\,f(y/x)\,dy=\int_0^{1-x}\frac{2y^2}{(1-x)^2}dy=\frac{2}{3}(1-x)$ $E[Y^2/X=X]=\int_0^{1-x} y^2\,f(y/x)\,dy=\int_0^{1-x}\frac{2y^3}{(1-x)^2}dy=\frac{1}{2}(1-x)^2$ $Var[Y/X=0]=E[y^2/X=0]-\{E[y/X=0]\}^2$ $=\frac{1}{2}(1-x)^2-\frac{4}{9}(1-x)^2=\frac{1}{18}(1-x)^2$.		
6.	Prove that the $Cov(aX,bY)=ab\,Cov(X,Y)$. Solution: $Cov(aX,bY)=E[\{ax-E(ax)\}\{by-E(by)\}]$ $=E[\{ax-aE(x)\}\{by-bE(y)\}]$ $=E[ab\{x-E(x)\}\{y-E(y)\}]$ $=abE[\{x-E(x)\}\{y-E(y)\}]$ $=ab\,Cov(X,Y)$	K2	CO2
7.	Show that the correlation coefficient $r(x,y)$ lies between -1 +1, ie. $-1\le r(x,y)\le 1$ Solution: Let $E[X]=\overline{X}$ and $E[Y]=\overline{Y}$. We know that $E\left[\left(\frac{X-\overline{X}}{\sigma_X}\right)\pm\left(\frac{Y-\overline{Y}}{\sigma_Y}\right)\right]^2\ge 0$ ie. $E\left(\frac{X-\overline{X}}{\sigma_X}\right)^2+E\left(\frac{Y-\overline{Y}}{\sigma_Y}\right)^2\pm 2E\left[\frac{(X-\overline{X})(Y-\overline{Y})}{\sigma_X\sigma_Y}\right]\ge 0$ $E\left(\frac{X-\overline{X}}{\sigma_X}\right)^2=\frac{1}{\sigma_X{}^2}E\left[\left(X-\overline{X}\right)\right]^2=1,\ E\left(\frac{Y-\overline{Y}}{\sigma_Y}\right)^2=\frac{1}{\sigma_Y{}^2}E\left[\left(Y-\overline{Y}\right)\right]^2=1$ $\therefore 1+1\pm 2\frac{1}{\sigma_X\sigma_Y}E\left[\left(X-\overline{X}\right)\left(Y-\overline{Y}\right)\right]\ge 0.$ $\Rightarrow 1+1\pm 2r(X,Y)\ge 0,\quad \Rightarrow -1\le r(X,Y)\le 1$	K1	CO2
8.	If the independent random variables X and Y have the variances 36 and 16 respectively. Find the covariance between $(X+Y)$ and $(X-Y)$. Solution: Let $U=X+Y$ and $V=X-Y$ $E[U]=E[X]+E[Y]$, $E[V]=E[X]-E[Y]$ and $E[UV]=E[X^2-Y^2]=E[X^2]-E[Y^2]$	K1	CO2

	$Cov(U,V)=E[UV]-E[U]E[V]$ $=E[X^2]-E[Y^2]-\{(E(X))^2-(E(Y))^2\}$ $=\{E(X^2)-(E(X))^2\}-\{E(Y^2)-(E(Y))^2\}$ $=\sigma_X^2-\sigma_Y^2=36-16=20.$		
9.	X and Y are independent random variables with variance 2 a respectively . Find the variance of $3X+4Y$. **Solution:** Given that X and Y are independent random variables with varia 2 and 3. i.e., $Var(X)=2$ and $Var(Y)=3$, Consider $Var(2X+4Y)=3^2Var(X)+4^2Var(Y)\quad(\because Var(aX)=a^2Var(X))$ $=9\times 2+16\times 3=66.$	**K1**	**CO2**
10.	The tangent of the angle between the lines of regres y on x and x on y is 0.6 and $\sigma_x=\frac{1}{2}\sigma_y$; find the correlation coefficient between X and Y. **Solution:** $\tan\theta=\left(\frac{1-r^2}{r}\right)\frac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}$ $0.6=\frac{3}{5}=\frac{1-r^2}{r}\cdot\frac{2}{5}\quad\Rightarrow 2r^2+3r-2=0\quad\Rightarrow r=\frac{1}{2}.$	**K2**	**CO2**
11.	The regression line of x on y is $3x+y=10$ and the regression li y on x is $3x+4y=12$ .find the Co-efficient of Correlation between X and Y. **Solution:** The regression coefficient of X on Y is give $b_{xy}=-1/3$ and the regression coefficient of Y on X is give $b_{yx}=-3/4$. $\therefore$ The coefficient of correlation between X and given by $r_{xy}=\sqrt{b_{xy}\times b_{yx}}=\sqrt{(-1/3)(-3/4)}=1/2=0.5$	**K1**	**CO2**
12.	Define joint cumulative distribution Function. If (X,Y) is a two dimensional discrete Random Variable such that $F(x,y)=P(X\le x,\ Y\le y)$, is called the joint cumulative distribution Function of (X,Y) 1) For discrete Random variable,	**K1**	**CO2**

	$$F(x,y) = \sum_{y_j \le y} \sum_{x_i \le x} p_{ij}$$ 1) For continuous random variables, $$F(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} \boldsymbol{f(x,y)dxdy}$$		
13.	Define two dimensional random variable Let S be the sample space. Let $X = X(s)$ and $Y = Y(s)$ be two functions each assigning a real number to each outcome $s \in S$. Then (X,Y) is a two dimensional random variable	**K1**	**CO2**

2.11 Part B Questions

Q. No.	Questions	K Level	CO
1.	If the joint p.d.f of (X,Y) is given $f(x,y)=\begin{cases}\frac{9(1+x+y)}{2(1+x)^4(1+y)^4}, & x>0, y>0.\\ 0 & elsewhere\end{cases}$ (i) Find the marginal densities of X and Y (ii) Find the conditional densities of X and Y (iii) Are X and Y independent **Ans.**: (i) $f_x(x)=\frac{3(2x+3)}{4(1+x)^4}, x>0$ $f_y(y)=\frac{3(2y+3)}{4(1+y)^4}, y>0$ $f(x/y)=\frac{6(1+x+y)}{(2x+3)(1+y)^4}$ $f(y/x)=\frac{6(1+x+y)}{(2y+3)(1+x)^4}, x>0, y>0$ (iii) X and Y dependent.	K1	CO2
2.	The joint density function of X and Y $f(x,y)=\begin{cases}8xy & 0<x<y<1\\ 0 & elsewhere\end{cases}$ (i) Find $P\left[X<\frac{1}{2}\cap Y<\frac{1}{4}\right]$ (ii) $Var\left[Y/X=x\right]$ (iii) Are X and Y are correlated; if so find the coefficient of the correlation. **Ans.**: $\rho=0.490$	K1	CO2
3.	Two dice are thrown. Let X_1 be the random variable which denote the outcome of the first die and X_2 denote the outcome of the second die. Let $Y=\max(X_1,X_2)$. (i) Form the distribution table of X_1 and Y. (ii) (ii) Find the Var(Y) (iii) Find the Cov(X, Y) **Ans.**: Var(Y)=1.97145 Cov(X,Y)= -76.80556	K2	CO2
4.	. X, Y, Z are uncorrelated random variables with zero means and standard deviation 5, 12 and 9 respectively and if $U=X+Y, V=Y+Z$. Find the Correlation coefficient between U and V **Ans.**: $r_{UV}=\frac{48}{65}$	K2	CO2

<table>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>5.</td><td>A distribution with unknown mean μ has variance equal to 1.5.Use Central limit theorem to find how large a sample should be taken from the distribution in order that the probability will be at least 0.95 that the sample mean will be within 0.5 of the population.
Ans.: $n = 24$</td><td>K2</td><td>CO2</td></tr>
<tr><td>6.</td><td>A random sample of size 100 is taken from a population Whose mean is 60 and variance is 400 using central limit theorem with what probability can we assert that the mean of the sample will not differ from $\mu = 60$ by more than 4 ?
Ans.: 0.9546</td><td>K2</td><td>CO2</td></tr>
<tr><td>7.</td><td>If X and Y are independent simple Gamma variates with parameters K_1 and K_2 respectively. Show that the variables U=X+Y, $V = \frac{X}{X+Y}$ are independent and U is a simple Gamma variate with parameter $K_1 + K_2$.</td><td>K2</td><td>CO2</td></tr>
<tr><td>8.</td><td>If X and Y are independent random variables each normally distributed with zero mean and variance σ^2.
Find the density function of $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1}\left(\frac{y}{x}\right)$.
Ans.: $f_r(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, r \ge 0$, $f_\theta(\theta) = \frac{1}{2\pi}, 0 \le \theta \le 2\pi$.</td><td>K2</td><td>CO2</td></tr>
<tr><td>9.</td><td>Find the coefficient of correlation and obtain the lines of regression from the data given below:
<table>
<tr><td>X</td><td>62</td><td>64</td><td>65</td><td>69</td><td>70</td><td>71</td><td>72</td><td>74</td></tr>
<tr><td>Y</td><td>126</td><td>125</td><td>139</td><td>145</td><td>165</td><td>152</td><td>180</td><td>208</td></tr>
</table>
Ans.: $Y = 6.02X - 256.62$ $X = 0.13Y + 48.23$ & $r(X,Y) = 0.$</td><td>K1</td><td>CO2</td></tr>
<tr><td>10.</td><td>If the joint density of X_1 and X_2 is given by
$f(x_1, x_2) = \begin{cases} 6e^{-3x_1 - 2x_2} & \text{for } x_1 > 0, x_2 > 0 \\ 0 & \text{otherwise} \end{cases}$
Find the p.d.f. of $Y = X_1 + X_2$ and its mean.</td><td>K2</td><td>CO2</td></tr>
</table>

	Ans.: $f_y(y)=6\left[e^{-2y}-e^{-3y}\right]$		
11.	Two dimensional random variables X and Y have joint pdf $f(x,y)=\begin{cases}\frac{xy}{9} & 0\le x\le 4,1\le y\le 5\\ 0 & otherwise\end{cases}$ Find Cov (X,Y)	**K2**	**CO2**
12.	. Two dimensional random variables X and Y have joint pdf $f(x,y)=\begin{cases}2-x-y & 0\le x\le 1,0\le y\le 1\\ 0 & otherwise\end{cases}$ (i) Find the marginal densities of X and Y (ii) Find the conditional densities of X and Y (iii) Find Var(X) and Var(Y)	**K1**	**CO2**
13.	The random variable X and Y have a joint p.d.f $f(x,y)=\begin{cases}x^2+\frac{xy}{3} & 0\le x\le 1,0\le y\le 2\\ 0 & otherwise\end{cases}$ (i) Are X and Y independent? (ii) Find the conditional p.d.f of X given Y	**K2**	**CO2**

Supportive Online Certification Courses

- Online Course: NPTEL
  Course Name: NOC:Introduction to probability and Statistics
  Course Instructor: Prof. G. Srinivasan, IIT Madras
  Duration : 4 weeks
  https://nptel.ac.in/courses/111/106/111106112/

- Online Course: NPTEL
  Course Name: NOC: Introduction to Probability Theory and
  Stochastic Processes
  Course Instructor: Dr. S. Dharmaraja, IIT Delhi
  Duration : 12 weeks
  https://nptel.ac.in/courses/111/102/111102111/

- Online Course: NPTEL
  Course Name: NOC:Probability and Statistics
  Course Instructor: Prof. Somesh Kumar
  Department of MathematicsIIT Kharagpur
  Duration : 10 weeks
  https://nptel.ac.in/courses/111/105/111105090/

- Online Course: Coursera
  Course Name: Probability Theory: Foundation for Data Science

Course Instructor: Anne Dougherty, Senior Instructor and Teaching Professor,
University of Colorado Boulder.
Jem Corcoran, Associate Professor,
University of Colorado Boulder
Duration : 4 weeks
https://in.coursera.org/learn/probability-theory-foundation-for-data-science

Real Time Applications

- Concept of Probability - Basics- Why Learn Probability?
  https://www.youtube.com/watch?v=74zR4OQ-ByY

- Concepts of Empirical Probability - Basics - Understanding Empirical Probability
  https://www.youtube.com/watch?v=oliMtbLoDTE

- Concept of Probability - Basics - Application Question
  https://www.youtube.com/watch?v=C0dslb8h5yQ
  https://www.youtube.com/watch?v=GSQWPHd3-CM
  https://www.youtube.com/watch?v=Ru1s7K4wed4
  https://www.youtube.com/watch?v=74zR4OQ-ByY
  https://www.youtube.com/watch?v=YoM87Td5_os
  https://www.youtube.com/watch?v=f2kc88oqUOg
  https://www.youtube.com/watch?v=vNhDhkMTqcI
  https://www.youtube.com/watch?v=9t9WnQErTbk

- Applications of Empirical Probability - Real Life Applications - Application
  https://www.youtube.com/watch?time_continue=30&v=_rmswvrmZ8w&feature=emb_logo
  https://www.youtube.com/watch?v=2rR-vFT2r7M
  https://www.youtube.com/watch?v=PO5wG5_x3Oc
  https://www.youtube.com/watch?v=-6kdwzYgl_c
  https://www.youtube.com/watch?v=HM9-HUVGl2c
  https://www.youtube.com/watch?v=1uvJuBc9l_8
  https://www.youtube.com/watch?v=J7QCWreTio4
  https://www.youtube.com/watch?v=9qkII19r1Ac

- Applications of conditional probability
  https://www.youtube.com/watch?v=TamclE-b0WY
  https://www.youtube.com/watch?v=rqN-ybI2MUA

- Understanding the applications of Probability in Machine Learning
  https://www.datasciencecentral.com/profiles/blogs/understanding-the-applications-of-probability-in-machine-learning

- probability and applications videos
  https://www.bing.com/videos/search?q=probability+and+applications&qpvt=probability+and+applications&FORM=VDRE

MINI PROJECT

COLLECT THE SAMPLES FOR YOURSELVES AND FIND THE CORRELATION AND REGRESSION FOR THE QUESTIONS GIVEN BELOW

- Comparison and relation between petroleum prices with food prices
- Relationship between exercises as well as a reduction in overall medical expenses
- The impact of information and communication technology on the teaching and learning of statistics in tertiary institutions

Prescribed Text Books & Reference Books

PROBABILITY AND STATISTICS	
S. No.	**TEXT BOOKS**
1.	Miller and Freund’s Probability and Statistics for Engineers, Johnson, R.A., Miller, I and Freund J., Pearson Education, Asia, 8th Edition, 2015.
2.	Introduction to Probability and Statistics, Milton. J. S. and Arnold. J.C., Tata McGraw Hill,4 th Edition, 2017.
	REFERENCES :
1.	Devore. J.L., "Probability and Statistics for Engineering and the Sciences‖,Cengage Learning, New Delhi, 9th Edition, 2016.
2.	Ross, S.M., "Introduction to Probability and Statistics for Engineers and Scientists", 6thEdition, Elsevier, 2020.
3.	Spiegel. M.R., Schiller. J. and Srinivasan, R.A., "Schaum‵s Outline of Theory and Problems of Probability and Statistics", Tata McGraw Hill Edition, 2004.
4.	Walpole. R.E., Myers. R.H., Myers. S.L. and Ye. K., "Probability and Statistics for Engineers and Scientists", Pearson Education, Asia, 9th Edition, 2012.

Thank you