JEPPIAAR INSTITUTE OF TECHNOLOGY

"Self-Belief | Self Discipline | Self Respect"

DEPARTMENT

OF

COMPUTER SCIENCE AND ENGINEERING

LECTURE NOTES-MA8402

PROBABILITY AND QUEUING THEORY

(Regulation 2017)

Unit II

TWO DIMENSIONAL RANDOM VARIABLES

- Introduction
- Joint distribution
- Marginal and Conditional Distribution
- Covariance
- Correlation Coefficient
- Linear Regression
- Transformation of random variables

Introduction

In the previous chapter we studied various aspects of the theory of a single R.V. In this chapter we extend our theory to include two R.V's one for each coordinator axis X and Y of the XY Plane.

DEFINITION : Let S be the sample space. Let X = X(S) & Y = Y(S) be two functions each assigning a real number to each outcome s ∈ S. hen (X, Y) is a two dimensional random variable.

1 Types of random variables

1. Discrete R.V.'s
2. Continuous R.V.'s

Discrete R.V.'s (Two Dimensional Discrete R.V.'s)

If the possible values of (X, Y) are finite, then (X, Y) is called a two dimensional discrete R.V. and it can be represented by (xi, y), i = 1,2,….,m.

In the study of two dimensional discrete R.V.'s we have the following 5 important terms.

- Joint Probability Function (JPF) (or) Joint Probability Mass Function.
- Joint Probability Distribution.
- Marginal Probability Function of X.
- Marginal Probability Function of Y.
- Conditional Probability Function.

1.1 Joint Probability Function of discrete R.V.'s X and Y

The function $P(X = x_i, Y = y_j) = P(x_i, y_j)$ is called the joint probability function for discrete random variable X and Y is denote by pij.

Note

1. $P(X = x_i, Y = y_j) = P[(X = x_i) \cap (Y = y_j)] = p_{ij}$
2. It should satisfies the following conditions

(i) $p_{ij} \geq \forall\ i, j$ (ii) $\Sigma_j \Sigma_i\, p_{ij} = 1$

1.2 Marginal Probability Function of X

If the joint probability distribution of two random variables X and Y is given then the marginal probability function of X is given by

$P_x(x_i) = p_i$ (marginal probability function of Y)

Conditional Probabilities

The conditional probabilities function of X given $Y = y_j$ is given by

$$P[X = x_i / Y = y_j] = \frac{P[X = x_i / Y = y_j]}{P[Y = y_j]} = \frac{p_{ij}}{p_{.j}}$$

The set $\{x_i, p_{ij} / p_{.j}\}$, i = 1, 2, 3, ...is called the conditional probability distribut given $Y = y_j$.

The conditional probability function of Y given $X = x_i$ is given by

$$P[Y = y_i / X = x_j] = \frac{P[Y = y_i / X = x_j]}{P[X = x_j]} = \frac{p_{ij}}{p_{i.}}$$

The set $\{y_i, p_{ij} / p_{i.}\}$, j = 1, 2, 3, ...is called the conditional probability distribut given $X = x_i$.

SOLVED PROBLEMS ON MARGINAL DISTRIBUTION

Example:2.1.1

From the following joint distribution of X and Y find the marginal distributions.

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

Solution

X \ Y	0	2	$P_Y(y) = p(Y=y)$
0	3/28 P(0,0)	3/28 P(2,0)	$15/28 = P_y(0)$
1	3/14 P(0, 1)	3/14 P(1,1)	$6/14 = P_y(1)$
2	1/28 P(0,2)	0 P(2,2)	$1/28 = P_y(2)$
$P_X(X) = P(X=x)$	10/28 = 5/14 $P_X(0)$	3/28 $P_X(2)$	1

The marginal distribution of X

$$P_X(0) = P(X = 0) = p(0,0) + p(0,1) + p(0,2) = 5/14$$
$$P_X(1) = P(X = 1) = p(1,0) + p(1,1) + p(1,2) = 15/28$$
$$P_X(2) = P(X = 2) = p(2,0) + p(2,1) + p(2,2) = 3/28$$

Marginal probability function of X is

$$P_X(x) = \begin{cases} \frac{5}{14}, & x = 0 \\ \frac{15}{28}, & x = 1 \\ \frac{3}{28}, & x = 2 \end{cases}$$

The marginal distribution of Y

$$P_Y(0) = P(Y = 0) = p(0,0) + p(1,0) + p(2,0) = 15/28$$
$$P_Y(1) = P(Y = 1) = p(0,1) + p(2,1) + p(1,1) = 3/7$$

$P_Y(2) = P(Y = 2) = p(0,2) + p(1,2) + p(2,2) = 1/28$

Marginal probability function of Y is

$$P_Y(y) = \begin{cases} \frac{15}{28}, & y = 0 \\ \frac{3}{7}, & y = 1 \\ \frac{1}{28}, & y = 2 \end{cases}$$

3 CONTINUOUS RANDOM VARIABLES

- Two dimensional continuous R.V.'s

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

- Joint probability density function :

(i) $f_{XY}(x,y) \geq 0$; (ii) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dydx = 1$

- Joint probability distribution function
  $F(x,y) = P[X \leq x, Y \leq y]$
  $$= \int_{-\infty}^{x}\left\{\int_{-\infty}^{y} f(x,y)dx\right\}dy$$
- Marginal probability density function
  $f(x) = f_X(x) = \int_{-\infty}^{\infty} f_{x,y}(x,y)dy$ (Marginal pdf of X)
  $f(y) = f_Y(x) = \int_{-\infty}^{\infty} f_{x,y}(x,y)dy$ (Marginal pdf of Y)
- Conditional probability density function
  (i) $P(Y = y / X = x) = f(y/x) = \dfrac{f(x,y)}{f(x)}, f(x) > 0$
  (ii) $P(X = x / Y = y) = f(x/y) = \dfrac{f(x,y)}{f(y)}, f(y) > 0$

Example :2.3.1

Show that the function $f(x,y)=\begin{cases}\frac{2}{5}(2x+3y), & 0<x<1, \\ 0 & \text{otherwise}\end{cases}$

is a joint density function of X and Y.

Solution

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

(i) $f(x,y)\geq 0$ (ii) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)=1$, then f(x,y) is a jdf

Given $f(x,y)=\begin{cases}\frac{2}{5}(2x+3y), & 0<x<1, \quad 0<y<1 \\ 0 & \text{otherwise}\end{cases}$

(i) $f(x,y)\geq 0$ in the given interval $0\leq (x,y)\leq 1$

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

$$=\frac{2}{5}\int_0^1\int_0^1\left[2\frac{x^2}{2}+3xy\right]_0^1 dy$$

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

$$=\frac{2}{5}\left(\frac{5}{2}\right)=1$$

Since f(x,y) satisfies the two conditions it is a j.d.f.

Example :2.3.2

The j.d.f of the random variables X and Y is given

$$f(x,y)=\begin{cases} 8xy, & 0<x<1, \quad 0<y<x \\ 0, & \text{otherwise} \end{cases}$$

Find (i) $f_X(x)$ (ii) $f_Y(y)$ (iii) $f(y/x)$

Solution

We know that

(i) The marginal pdf of 'X' is

$$f_X(x) = f(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^x 8xy\,dy = 4x^3$$

$$f(x)=4x^3, \; 0<x<1$$

(ii) The marginal pdf of 'Y' is

$$f_Y(y) = f(y) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^1 8xy\,dy = 4y$$

$$f(y)=4y, \; 0<y<\alpha$$

(iii) We know that

$$f(y/x) = \frac{f(x,y)}{f(x)}$$

$$= \frac{8xy}{4x^3} = \frac{2y}{x^2}, \; 0<y<x, \; 0<x<1$$

Result

Marginal pdf g	**Marginal pdf y**	**F(y/x)**
$4x^3$, 0<x<1	4y, 0<y<x	$\frac{2y}{x^2}, 0<y<x, 0<x<1$

4 REGRESSION

* Line of regression

The line of regression of X on Y is given by

* Line of regression

The line of regression of X on Y is given by

$$x-\bar{x}=r.\frac{\sigma y}{\sigma x}(y-\bar{y})$$

The line of regression of Y on X is given by

$$y-\bar{y}=r.\frac{\sigma y}{\sigma x}(x-\bar{x})$$

* Angle between two lines of Regression.

$$\tan\theta=\frac{1-r^2}{r}\left(\frac{\sigma_y\sigma_x}{\sigma_{x^2}+\sigma_{y^2}}\right)$$

* Regression coefficient

Regression coefficients of Y on X

$$r.\frac{\sigma y}{\sigma x}=b_{YX}$$

Regression coefficient of X and Y

$$r.\frac{\sigma x}{\sigma y}=b_{XY}$$

$\therefore$ Correlation coefficient $r=\pm\sqrt{b_{XY}\times b_{YX}}$

Example:2.4.1

1. From the following data, find

(i) The two regression equation

(ii) The coefficient of correlation between the marks in Economic and Statistics.

(iii) The most likely marks in statistics when marks in Economic are 30.

Marks in Economics	25	28	35	32	31	36	29	38	34	32
Marks in Statistics	40	46	49	41	36	32	31	30	33	39

Solution

X	**Y**	$X-\overline{X}=X-32$	$X-\overline{Y}=Y-38$	$\left(X-X\right)^{2}$	$\left(Y-\overline{Y}\right)^{2}$	$\left(X-\overline{X}\right)^{2}$
25	43	-7	5	49	25	-35
28	46	-4	8	16	64	-32
35	4	3	11	9	121	33
32	41	0	3	0	9	0
31	36	-1	-2	1	4	2
36	32	4	-6	16	36	-24
29	31	-3	-7	09	49	+21
38	30	6	-8	36	64	-48
34	33	2	-5	4	25	-48
32	39	0	1	0	1	100
320	380	0	0	140	398	-93

Here $\overline{X} = \frac{\sum X}{n} = \frac{320}{10} = 32$ and $\overline{Y} = \frac{\sum Y}{n} = \frac{380}{10} = 38$

Coefficient of regression of Y on X is

$$b_{YX} = \frac{\sum(X-\overline{X})(Y-\overline{Y})}{\sum(X-\overline{X})^2} = \frac{-93}{140} = -0.6643$$

Coefficient of regression of X on Y is

$$b_{XY} = \frac{\sum(X-\overline{X})(Y-\overline{Y})}{\sum(Y-\overline{Y})^2} = \frac{-93}{398} = -0.2337$$

Equation of the line of regression of X and Y is

$$X - \overline{X} = b_{XY}(Y - \overline{Y})$$
$$X - 32 = -0.2337\,(y - 38)$$
$$X = -0.2337\,y + 0.2337 \times 38 + 32$$
$$X = -0.2337\,y + 40.8806$$

Equation of the line of regression of Y on X is

$$Y - \overline{Y} = b_{YX}(X - \overline{X})$$
$$Y - 38 = -0.6643\,(x - 32)$$
$$Y = -0.6643\,x + 38 + 0.6643 \times 32$$
$$= -0.6642\,x + 59.2576$$

Coefficient of Correlation

Coefficient of Correlation

$$r^2 = b_{YX} \times b_{XY}$$
$$= -0.6643 \times (-0.2337)$$
$$r = 0.1552$$
$$r = \pm\sqrt{0.1552}$$
$$r = \pm\sqrt{0.394}$$

Now we have to find the most likely mark, in statistics (Y) when marks in economics

$$y = -0.6643\,x + 59.2575$$

Put x = 30, we get

$$y = -0.6643 \times 30 + 59.2536$$
$$= 39.3286$$
$$y \simeq 39$$

5 COVARIANCE

Def : If X and Y are random variables, then Covariance between X and Y is defined as

Cov (X, Y) = E(XY) – E(X) . E(Y)

Cov (X, Y) = 0 [If X & Y are independent]

6 CORRELATION

Types of Correlation

- Positive Correlation

  (If two variables deviate in same direction)

- Negative Correlation

  (If two variables constantly deviate in opposite direction)

7 KARL-PEARSON'S COEFFICIENT OF CORRELATION

Correlation coefficient between two random variables X and Y usually denoted by r(X, Y) is a numerical measure of linear relationship between them and is defined as

$$r(X,Y) = \frac{Cov(X,Y)}{\sigma_X . \sigma_Y},$$

$$\text{Where } Cov(X, Y) = \frac{1}{n}\sum XY - \overline{X}\,\overline{Y}$$

$$\sigma_X = \frac{\sum X}{n}; \quad \sigma_Y = \frac{\sum Y}{n}$$

* Limits of correlation coefficient

$-1 \leq r \leq 1.$

X & Y independent, $\therefore$ r(X, Y) = 0.

Note : Types of correlation based on 'r'.

Values of 'r'	Correlation is said to be
r = 1	perfect and positive
0<r<1	positive
-1<r<0	negative
r = 0	Uncorrelated

SOLVED PROBLEMS ON CORRELATION

Example :2.6.1

Calculated the correlation coefficient for the following heights of fathers X and their sons Y

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

Solution

X	**Y**	**U = X – 68**	**V = Y – 68**	**UV**	U^2	V^2
65	67	-3	-1	3	9	1
66	68	-2	0	0	4	0
67	65	-1	-3	3	1	9
67	68	-1	0	0	1	0
68	72	0	4	0	0	16
69	72	1	4	4	1	16
70	69	2	1	2	4	1
72	71	4	3	12	16	9

$\sum U = 0$ $\quad \sum V = 0$ $\quad \sum UV = 24$ $\sum U^2 = 36$ $\sum V^2 = 52$

Now

Now

$$\overline{U} = \frac{\sum U}{n} = \frac{0}{8} = 0$$

$$\overline{V} = \frac{\sum V}{n} = \frac{8}{8} = 1$$

Cov (X, Y) = Cov (U, V)

$$\Rightarrow \frac{\sum UV}{n} - \overline{U}\,\overline{V} = \frac{24}{8} - 0 = 3 \qquad (1)$$

$$\sigma_U = \sqrt{\frac{\sum U^2}{n} - \overline{U}^2} = \sqrt{\frac{36}{8} - 0} = 2.121 \qquad (2)$$

$$\sigma_V = \sqrt{\frac{\sum V^2}{n} - \overline{V}^2} = \sqrt{\frac{52}{8} - 1} = 2.345 \qquad (3)$$

$$\therefore r(X, Y) = r(U, V) = \frac{Cov(U, V)}{\sigma_U . \sigma_V} = \frac{3}{2.121 \times 2.345}$$

$$= 0.6031 \qquad \text{(by 1, 2, 3)}$$

Example :2.6.2

Let X be a random variable with p.d.f. $f(x) = \frac{1}{2}, -1 \le x \le 1$ $Y = x^2$, find the correlation coefficient between X and Y.

Solution

$$E(X) = \int_{-\infty}^{\infty} x.f(x)\,dx$$

$$= \int_{-1}^{1} x.\frac{1}{2}\,dx \quad = \frac{1}{2}\left(\frac{x^2}{2}\right)_{-1}^{1} \quad = \frac{1}{2}\left(\frac{1}{2} - \frac{1}{2}\right) = 0$$

$$E(X) = 0$$

$$E(Y) = \int_{-\infty}^{\infty} x^2.f(x)\,dx$$

$$= \int_{-1}^{1} x^2.\frac{1}{2}\,dx \quad = \frac{1}{2}\left(\frac{x^3}{3}\right)_{-1}^{1} \quad = \frac{1}{2}\left(\frac{1}{3} + \frac{1}{3}\right) = \frac{1}{2}.\frac{2}{3} = \frac{1}{3}$$

$$E(XY) = E(XX^2)$$

$$= E(X^3) \qquad = \int_{-\infty}^{\infty} x^3.f(x)\,dx = \left(\frac{x^4}{4}\right)_{-1}^{1} = 0$$

$$E(XY) = 0$$

$$\therefore r(X, Y) = \rho(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = 0$$

$$\rho = 0.$$

Note : E(X) and E(XY) are equal to zero, noted not find $\sigma_x \& \sigma_y$.

8 TRANSFORMS OF TWO DIMENSIONAL RANDOM VARIABLE

Formula:

Formula:

$$f_U(u) = \int_{-\infty}^{\infty} f_{u,v}(u,v)\,dv$$

$$\& \qquad f_V(u) = \int_{-\infty}^{\infty} f_{u,v}(u,v)\,du$$

$$f_{UV}(u,V) = f_{XY}(x,y)\left|\frac{\partial(x,y)}{\partial(u,v)}\right|$$

Example : 1

If the joint pdf of (X, Y) is given by $f_{xy}(x, y) = x+y$, $0 \le x, y \le 1$, find the pdf of

Solution

Given $f_{xy}(x, y) = x + y$

Given $U = XY$

Let $V = Y$

$$x = \frac{u}{v} \ \& \ y = V$$

$$\frac{\partial x}{\partial u} = \frac{1}{V} . \frac{\partial x}{\partial v} = \frac{-u}{V^2}; \frac{\partial y}{\partial u} = 0; \frac{\partial y}{\partial v} = 1 \tag{1}$$

$$\therefore J = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \begin{vmatrix} \frac{\partial y}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{V} & \frac{-u}{V^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{v} - 1 = \frac{1}{v}$$

$$\Rightarrow |J| = \frac{1}{V} \tag{2}$$

The joint p.d.f. (u, v) is given by

$$f_{uv}(u,v) = f_{xy}(x,y)\,|J|$$

$$= (x+y)\frac{1}{|v|}$$

$$= \frac{1}{V}\left(\frac{u}{v} + u\right) \tag{3}$$

The range of V :

Since $0 \le y \le 1$, we have $0 \le V \le 1$ $\quad (\therefore V = y)$

The range of u :

Given $\quad 0 \le x \le 1$

$\Rightarrow \quad 0 \le \frac{u}{v} \le$

$\Rightarrow \qquad 0 \le \frac{u}{v} \le$

$\Rightarrow \qquad 0 \le u \le v$

Hence the p.d.f. of (u, v) is given by

$$f_{uv}(u,v) = \frac{1}{v}\left(\frac{u}{v}+v\right),\ 0 \le u \le v,\ 0 \le v \le 1$$

Now

$$f_U(u) = \int_{-\infty}^{\infty} f_{u,v}(u,v)\,dv$$

$$= \int_u^1 f_{u,v}(u,v)\,dv$$

$$= \int_u^1 \left(\frac{u}{v^2}+1\right) dv$$

$$= \left[v + u.\frac{v^{-1}}{-1}\right]_u^1$$

$\therefore f_u(u) = 2(1-u),\ 0 < u < 1$

p.d.f of (u, v) p.d.f of u = XY

$$f_{uv}(u,v) = \frac{1}{v}\left(\frac{u}{v}+v\right)$$

$0 \le u \le v,\ 0 \le v \le 1$

$f_u(u) = 2(1-u),\ 0 < u < 1$

TUTORIAL QUESTIONS

1. The jpdf of r.v X and Y is given by f(x,y)=3(x+y),0<x<1,0<y<1,x+y<1 and 0 otherwise. Find the marginal pdf of X and Y and ii) Cov(X,Y).
2. Obtain the correlation coefficient for the following data:

X: 68	64	75	50	64	80	75	40	55	64
Y: 62	58	68	45	81	60	48	48	50	70

3. The two lines of regression are 8X-10Y+66=0, 4X-18Y-214=0. The variance of x is 9 find i) The mean value of x and y. ii) Correlation coefficient between x and y.

4. If X1,X2,…Xn are Poisson variates with parameter $\lambda=2$, use the central limit theorem to find P(120≤Sn≤160) where Sn=X1+X2+…Xn and n=75.

5. If the joint probability density function of a two dimensional random variable (X,Y) is

given by f(x, y) = x^2 + , 0<x<1,0<y<2= 0, elsewhere Find (i) P(X>1/2)(ii) P(Y<X) and (iii)

P(Y<1/2/ X<1/2).

6. Two random variables X and Y have joint density Find Cov (X,Y).

7. If the equations of the two lines of regression of y on x and x on y are respectively 7x-16y+9=0; 5y-4x-3=0, calculate the coefficient of correlation.

WORKEDOUT EXAMPLES

Example 1

The j.d.f of the random variables X and Y is given

f (x, y) = 8xy, 0 <x <1, 0 y< x

otherwise 0,

Find (i) fX(x) (ii) fY(y) (iii) f(y/x)

Solution

We know that

(i) The marginal pdf of 'X' is

$$f_X(x) = f(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^x 8xy\,dy = 4x^3$$

$$f(x)=4x^3,\ 0 < x < 1$$

(ii) The marginal pdf of 'Y' is

$$f_Y(y) = f(y) = \int_{-\infty}^{\infty} f(x,y)dy = \int_0^1 8xy\,dy = 4y$$

$$f(y)=4y,\ 0 < y < \alpha$$

(iii) We know that

$$f(y/x) = \frac{f(x,y)}{f(x)}$$

$$= \frac{8xy}{4x^3} = \frac{2y}{x^2},\ 0 < y < x,\ 0 < x < 1$$

Example 2

Let X be a random variable with p.d.f. $f(x) = \frac{1}{2}, -1 \leq x$ ... $Y = x^2$, find the correlation coefficient between X and Y.

Solution

$$E(X) = \int_{-\infty}^{\infty} x.f(x)\,dx$$

$$= \int_{-1}^{1} x.\frac{1}{2}\,dx \quad = \frac{1}{2}\left(\frac{x^2}{2}\right)_{-1}^{1} \quad = \frac{1}{2}\left(\frac{1}{2} - \frac{1}{2}\right) = 0$$

$$E(X) = 0$$

$$E(Y) = \int_{-\infty}^{\infty} x^2.f(x)\,dx$$

$$= \int_{-1}^{1} x^2.\frac{1}{2}\,dx \quad = \frac{1}{2}\left(\frac{x^3}{3}\right)_{-1}^{1} \quad = \frac{1}{2}\left(\frac{1}{3} + \frac{1}{3}\right) = \frac{1}{2}.\frac{2}{3} = \frac{1}{3}$$

$$E(XY) = E(XX^2)$$

$$= E(X^3) \quad = \int_{-\infty}^{\infty} x^3.f(x)\,dx = \left(\frac{x^4}{4}\right)_{-1}^{1} = 0$$

$$E(XY) = 0$$

$$\therefore r(X, Y) = \rho(X, Y) = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} = 0$$

$$\rho = 0.$$

Note : E(X) and E(XY) are equal to zero, noted not find σ_x&σ_y.

Result

Marginal pdf g	Marginal pdf y	F(y/x)
$4x^3$, 0<x<1	4y, 0<y<x	$\frac{2y}{x^2}, 0 < y < x, 0 < x <$