Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 1:

Calculate the mean and median of the following data

Marks <	10	20	30	40	50	60	70	80	90
No. of students	5	15	98	242	367	405	425	438	439

Solution:

To find the mean let's consider

Assumed mean (a) = 45

Class width (c) = 10

Class Interval	Class matrix (x_i)	Frequency f_i	Cumulative frequency	$u_i = \frac{x_i - a}{c}$	f_iu_i
0-10	5	5	5	-4	-20
10-20	15	10	15	-3	-30
20-30	25	83	98	-2	-166
30-40	35	144	242	-1	-144
40-50	45	125	367	0	0
50-60	55	38	405	1	38
60-70	65	20	425	2	40
70-80	75	13	438	3	39
80-90	85	1	439	4	4
		N=$\sum f_i = 439$			$\sum f_iu_i = -239$

To calculate mean

$$\bar{x} = a + c\,\frac{\sum f_i u_i}{N} \qquad where\ N = \sum f_i \qquad and\ u_i = \frac{x_i - a}{c}$$

Hence

$$\bar{x} = a + c\,\frac{\sum f_i u_i}{N}$$

$$\bar{x} = 45 + 10 * \frac{-239}{439}$$

$$\bar{x} = 39.56$$

To calculate median

$$N = \sum f_i = 439$$

$$\therefore \quad \frac{N}{2} = 219.5$$

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

Median class (l_1 –l_2) = (30-40)

Frequency of median class (f) = 144

Cumulative frequency up to but not including the median class (F) =98

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

$$Median = 30 + \frac{(40 - 30)}{144}(219.5 - 98)$$

$$Median = 38.43$$

∴ For given data
Mean = 39.56
Median = 38.43

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 2:

Calculate the median of the following data

Marks	0-20	20-40	40-60	60-80	80-100
No. of students	5	8	15	16	6

Solution:

We prepare following table

Class Interval	Frequency f_i	Cumulative frequency
0-20	5	5
20-40	8	13
40-60	15	28
60-80	16	44
80-100	6	50
	$N=\sum f_i = 50$	

To calculate median

$$N = \sum f_i = 50$$

$$\therefore \quad \frac{N}{2} = 25$$

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

Median class ($l_2 - l_1$) = (40-60)

Frequency of median class (f) = 15

Cumulative frequency up to but not including the median class (F) =13

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

$$Median = 40 + \frac{60 - 40}{15}(25 - 13)$$

$$Median = 56$$

$\therefore$ For given data

Median = 56

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3:

Calculate the median of the following data

Marks	10-12	12-14	14-16	16-18	18-20	20-22	22-24
No. of students	11	17	20	22	10	10	10

Solution:

We prepare following table

Class Interval	Frequency f_i	Cumulative frequency
10-12	11	11
12-14	17	28
14-16	20	48
16-18	22	70
18-20	10	80
20-22	10	90
22-24	10	100
	$N=\sum f_i = 100$	

To calculate median

$$N = \sum f_i = 100$$

$$\therefore \quad \frac{N}{2} = \frac{100}{2} = 50$$

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

Median class ($l_2 - l_1$) = (16-18)

Frequency of median class (f) = 22

Cumulative frequency up to but not including the median class (F) = 48

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

$$Median = 16 + \frac{(18-16)}{22}(50 - 48)$$

$$Median = 16.1818$$

∴ For given data Median = 16.1818

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4:
The following table shows the percentage of ash content in 280 wagons tests of a certain kind of coal. Find the mode of the distribution.

% of ash content	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-11	11-12
Frequency	1	7	28	78	84	45	28	7	2

Solution:
We prepare following table

Class Interval	Frequency f_i	
3-4	1	
4-5	7	
5-6	28	
6-7	78	
7-8	84	Modal class
8-9	45	
9-10	28	
10-11	7	
11-12	2	

$$Mode = l_1 + \left(\frac{d_1}{d_1 + d_2}\right)(l_2 - l_1)$$

Modal class (l_2 –l_1) =7-8

Frequency of modal class (f) = 84

Frequency of pre-modal class (f_1) = 78

Frequency of post-modal class (f_2) = 45

Difference between frequency of modal class and of previous class
d_1 = (f- f_1) = 84-78 = 6

Difference between frequency of modal class and of following class
d_2 = (f- f_2) = 84 -45 = 39

$$Mode = l_1 + \left(\frac{d_1}{d_1 + d_2}\right)(l_2 - l_1)$$

$$Mode = 7 + \left(\frac{6}{6 + 39}\right)(8 - 7)$$

$$Mode = 7.1333$$

$\therefore$ For given data

Mode = 7.1333

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5:

The following frequency distribution of marks of students in an examination, calculate the value of Karl Pearson's coefficient of skewness.

Marks less than	10	20	30	40	50
No. of students	5	12	32	44	50

Solution:

We prepare following table

Assumed mean (a) = 25

Class width (c) = 10

Class Interval	Class marks (x_i)	Freq^n f_i	Cumu. Freq^n	$u_i = \frac{x_i - a}{c}$	$f_i u_i$	$f_i u_i^2$
0-10	5	5	5	-2	-10	20
10-20	15	7	12	-1	-7	7
20-30	25	20	32	0	0	0
30-40	35	12	44	1	12	12
40-50	45	6	50	2	12	24
		N= $\sum f_i = 50$			$\sum f_i u_i = 7$	$\sum f_i u_i^2 = 63$

To calculate mean

$$\bar{x} = a + c\,\frac{\sum f_i u_i}{N} \qquad where\ N = \sum f_i \qquad and\ u_i = \frac{x_i - a}{c}$$

Hence

$$\bar{x} = a + c\,\frac{\sum f_i u_i}{N}$$

$$\bar{x} = 25 + 10 * \frac{7}{50}$$

$$\bar{x} = 26.4$$

Standard deviation

$$\sigma_x = c\sqrt{\frac{\sum f_i {u_i}^2}{N} - \left(\frac{\sum f_i u_i}{N}\right)^2} \qquad \text{where } N = \Sigma f_i$$

$$\sigma_x = 10\sqrt{\frac{63}{50} - \left(\frac{7}{50}\right)^2}$$

$$\sigma_x = 11.1373$$

To find mode

$$Mode = l_1 + \left(\frac{d_1}{d_1 + d_2}\right)(l_2 - l_1)$$

Modal class (l_1 –l_2) = (20-30)

Frequency of modal class (f) = 20

Frequency of pre-modal class (f_1) = 7

Frequency of post-modal class (f_2) = 12

Difference between frequency of modal class and of previous class
d_1 = (f- f_1) = 20 -7 = 13

Difference between frequency of modal class and of following class
d_2 = (f- f_2) = 20 – 12 = 8

$$Mode = 20 + \left(\frac{13}{13+8}\right)(30-20)$$

$$Mode = 20 + \left(\frac{13}{21}\right)(10)$$

$$Mode = 26.19$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = \frac{Mean - Mode}{Standard\ Deviation}$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = \frac{26.4 - 26.19}{11.1373}$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = 0.0188$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 6:
A prospective buyer tested the bursting pressure of the sample of polythene bags received from a manufacturer. The test gives the following results

Bursting pressure	5-10	10-15	15-20	20-25	25-30	30-35
No. of bags	2	10	30	50	6	2

Find Karl Pearson's coefficient of skewness for bursting pressure
Solution:
We prepare following table
Assumed mean (a) = 17.5
Class width (c) = 5

Class Interval	Class pressure (x_i)	Freq^n f_i	Cumu. Freq^n	$u_i = \frac{x_i - a}{c}$	$f_i u_i$	$f_i u_i^2$
5-10	7.5	2	2	-2	-4	8
10-15	12.5	10	12	-1	-10	10
15-20	17.5	30	42	0	0	0
20-25	22.5	50	92	1	50	50
25-30	27.5	6	98	2	12	24
30-35	32.5	2	100	3	6	18
		N= $\sum f_i = 100$			$\sum f_i u_i = 54$	$\sum f_i u_i^2 = 110$

To calculate mean

$$\bar{x} = a + c\,\frac{\sum f_i u_i}{N} \qquad where\ N = \sum f_i \qquad and\ u_i = \frac{x_i - a}{c}$$

Hence

$$\bar{x} = 17.5 + 5\,\frac{54}{100}$$

$$\bar{x} = 20.2$$

Standard deviation

$$\sigma_x = c\sqrt{\frac{\sum f_i {u_i}^2}{N} - \left(\frac{\sum f_i u_i}{N}\right)^2} \qquad \text{where } N = \Sigma f_i$$

$$\sigma_x = 5\sqrt{\frac{110}{100} - \left(\frac{54}{100}\right)^2}$$

$$\sigma_x = 4.4956$$

To find mode

$$Mode = l_1 + \left(\frac{d_1}{d_1 + d_2}\right)(l_2 - l_1)$$

Modal class (l_1 –l_2) = (20-25)

Frequency of modal class (f) = 50

Frequency of pre-modal class (f_1) =30

Frequency of post-modal class (f_2) =6

Difference between frequency of modal class and of previous class
d_1 = ($f - f_1$) = 50-30 = 20

Difference between frequency of modal class and of following class
d_2 = ($f - f_2$) =50 – 6 = 44

$$Mode = 20 + \left(\frac{20}{20 + 44}\right)(25 - 20)$$

$$Mode = 20 + \left(\frac{20}{64}\right)(5)$$

$$Mode = 21.5625$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = \frac{Mean - Mode}{Standard\ Deviation}$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = \frac{20.2 - 21.5625}{4.4956}$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = -0.3031$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 7:

From the following data on age of employee, calculate the Karl Pearson's coefficient of skewness

Age (years)	20-25	25-30	30-35	35-40	40-45	45-50	50-55
No. of employees	8	12	20	25	15	12	8

Solution:

We prepare following table

Assumed mean (a) = 37.5

Class width (c) = 5

Class Interval	Class age (x_i)	Freq^n f_i	Cumu. Freq^n	$u_i = \frac{x_i - a}{c}$	f_iu_i	$f_iu_i^2$
20-25	22.5	8	8	-3	-24	72
25-30	27.5	12	20	-2	-24	48
30-35	32.5	20	40	-1	-20	20
35-40	37.5	25	65	0	0	0
40-45	42.5	15	80	1	15	15
45-50	47.5	12	92	2	24	48
50-55	52.5	8	100	3	24	72
		N=$\sum f_i = 100$			$\sum f_iu_i =$ -5	$\sum f_iu_i^2 = 275$

To calculate mean

$$\bar{x} = a + c\,\frac{\sum f_i u_i}{N} \qquad where\ N = \sum f_i \qquad and\ u_i = \frac{x_i - a}{c}$$

Hence

$$\bar{x} = a + c\,\frac{\sum f_i u_i}{N}$$

$$\bar{x} = 37.5 + 5\,\frac{-5}{100}$$

$$\bar{x} = 37.25$$

Standard deviation

$$\sigma_x = c\sqrt{\frac{\sum f_i {u_i}^2}{N} - \left(\frac{\sum f_i u_i}{N}\right)^2} \qquad \text{where } N = \Sigma f_i$$

$$\sigma_x = 5\sqrt{\frac{275}{100} - \left(\frac{-5}{100}\right)^2}$$

$$\sigma_x = 8.2878$$

To find mode

$$Mode = l_1 + \left(\frac{d_1}{d_1 + d_2}\right)(l_2 - l_1)$$

Modal class (l_1 –l_2) = (35-40)

Frequency of modal class (f) = 25

Frequency of pre-modal class (f_1) =20

Frequency of post-modal class (f_2) =15

Difference between frequency of modal class and of previous class
d_1 = ($f - f_1$) = 25-20 =5

Difference between frequency of modal class and of following class
d_2 = ($f - f_2$) = 25-15 = 10

$$Mode = 35 + \left(\frac{5}{5+10}\right)(40-35)$$

$$Mode = 36.6667$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = \frac{Mean - Mode}{Standard\ Deviation}$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = \frac{37.25 - 36.6667}{8.2878}$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = 0.0704$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 8:
For a moderately skewed frequency distribution of retail prices for men's shoes it is found that the mean price is Rs. 20 and median price is Rs. 17. If the coefficient of variation is 20%, find the Pearson's coefficient of skewness.

Solution:
Given

Mean =20
Median = 17
Coefficient of variation = 20%

$$Coefficient\ of\ variation = \frac{Standard\ Deviation}{Mean}\ X\ 100\ \%$$

$$20 = \frac{Standard\ Deviation}{20}\ X\ 100$$

Standard Deviation = 4

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = \frac{3(Mean - Median)}{Standard\ Deviation}$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = \frac{3(20 - 17)}{4}$$

$$Karl\ Pearson's\ \ coefficient\ of\ skewness = 2.25$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 9:
From the following frequency distribution of marks of students in the examination calculate the Bowley's coefficient of skewness

Marks less than	10	20	30	40	50
No. of students	5	12	32	44	50

Solution:
We prepare following table

Class Interval	Freq^n f_i	Cumu. Freq^n
0-10	5	5
10-20	7	12
20-30	20	32
30-40	12	44
40-50	6	50
	N=$\sum f_i = 50$	

To calculate Q1, Q2, Q3

$$N = \sum f_i = 50$$

For first quartile (Q1)

$$\therefore \quad \frac{N}{4} = 12.5$$

$$Q1 = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{4} - F\right)$$

Q1 class(l_1 –l_2) =(20-30)

Frequency of Q1 class (f) = 20

Cumulative frequency up to but not including the Q1 class (F) = 12

$$Q1 = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{4} - F\right)$$

$$Q1 = 20 + \frac{30 - 20}{20}(12.5 - 12)$$

$$Q1 = 20.25$$

For second quartile Q2 (median)

To calculate median

$$N = \sum f_i = 50$$

$$\therefore \quad \frac{N}{2} = 25$$

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

Median/Q2 class(l_1 –l_2) = (20-30)

Frequency of median class (f) = 20

Cumulative frequency up to but not including the median class (F) =12

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

$$Median = 20 + \frac{30 - 20}{20}(25 - 12)$$

$$Median = 26.5$$

$$i.e.\ Q2 = 26.5$$

For third quartile (Q3)

$$\therefore \quad \frac{3N}{4} = 37.5$$

$$Q3 = l_l + \frac{l_2 - l_1}{f}\left(\frac{3N}{4} - F\right)$$

Q3 class(l_1 –l_2) = 30-40

Frequency of Q3 class (f) = 12

Cumulative frequency up to but not including the Q3 class (F) =32

$$Q3 = l_l + \frac{l_2 - l_1}{f}\left(\frac{3N}{4} - F\right)$$

$$Q3 = 30 + \frac{40 - 30}{12}(37.5 - 32)$$

$$Q3 = 34.5833$$

$$Bowley's \;\; coefficient \; of \; skewness = \frac{(Q_3 + Q_1 - 2 \, X \, Median)}{(Q_3 - Q_1)}$$

$$Bowley's \;\; coefficient \; of \; skewness = \frac{(34.5833 + 20.25 - 2 \, X \, 26.5)}{(34.5833 - 20.25)}$$

$$Bowley's \;\; coefficient \; of \; skewness = 0.1279$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 10:

The following data gives the number of car accidents in the city during a random time period. Calculate Bowley's coefficient of skewness for the following distribution

Class	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Frequency	7	9	16	22	14	12	3

Solution:

We prepare following table

Class Interval	Freq^n f_i	Cumu. Freq^n
5-10	7	7
10-15	9	16
15-20	16	32
20-25	22	54
25-30	14	68
30-35	12	80
35-40	3	83
	N= $\sum f_i = 83$	

To calculate Q1, Q2, Q3

$$N = \sum f_i = 83$$

For first quartile (Q1)

$$\therefore \quad \frac{N}{4} = 20.75$$

$$Q1 = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{4} - F\right)$$

Q1 class(l_1 –l_2) =15-20

Frequency of Q1 class (f) = 16

Cumulative frequency up to but not including the Q1 class (F) =16

$$Q1 = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{4} - F\right)$$

$$Q1 = 15 + \frac{20 - 15}{16}(20.75 - 16)$$

$$Q1 = 16.4843$$

For second quartile Q2 (median)

To calculate median

$$N = \sum f_i = 83$$

$$\therefore \quad \frac{N}{2} = 41.5$$

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

Median/Q2 class(l_1 –l_2) = 20-25

Frequency of median class (f) = 22

Cumulative frequency up to but not including the median class (F) =32

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

$$Median = 20 + \frac{25 - 20}{22}(41.5 - 32)$$

$$Median = 22.1590$$

$$i.e.\ Q2 = 22.1590$$

For third quartile (Q3)

$$\therefore \quad \frac{3N}{4} = 62.25$$

$$Q3 = l_l + \frac{l_2 - l_1}{f}\left(\frac{3N}{4} - F\right)$$

Q3 class(l_1 –l_2) = 25-30

Frequency of Q3 class (f) = 14

Cumulative frequency up to but not including the Q3 class (F) =54

$$Q3 = l_l + \frac{l_2 - l_1}{f}\left(\frac{3N}{4} - F\right)$$

$$Q3 = 25 + \frac{30 - 25}{14}(62.25 - 54)$$

$$Q3 = 27.9464$$

$$Bowley's \ \ coefficient \ of \ skewness = \frac{(Q_3 + Q_1 - 2 \, X \, Median)}{(Q_3 - Q_1)}$$

$$Bowley's \ \ coefficient \ of \ skewness$$
$$= \frac{(27.9464 + 16.4843 - 2 \, X \, 22.1590)}{(27.9464 - 16.4843)}$$

$$Bowley's \ \ coefficient \ of \ skewness = 0.0098$$

Quartile devitation = (Q3 –Q1)/(Q3 +Q1)

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 11:

Calculate Bowley's coefficient of skewness for the following distribution

X	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50
f	2	5	7	13	21	16	8	3

Solution:

We prepare following table

Class Interval	Freq^n f_i	Cumu. Freq^n
10-15	2	2
15-20	5	7
20-25	7	14
25-30	13	27
30-35	21	48
35-40	16	64
40-45	8	72
45-50	3	75
	N= $\sum f_i = 75$	

To calculate Q1, Q2, Q3

$$N = \sum f_i = 75$$

For first quartile (Q1)

$$\therefore \quad \frac{N}{4} = 18.75$$

$$Q1 = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{4} - F\right)$$

Q1 class (l_1 –l_2) = 25-30

Frequency of Q1 class (f) = 13

Cumulative frequency up to but not including the Q1 class (F) =14

$$Q1 = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{4} - F\right)$$

$$Q1 = 25 + \frac{30 - 25}{13}(18.75 - 14)$$

$$Q1 = 26.8269$$

For second quartile Q2 (median)

To calculate median

$$N = \sum f_i = 75$$

$$\therefore \quad \frac{N}{2} = 37.5$$

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

Median/Q2 class (l_1 –l_2) = 30-35

Frequency of median class (f) = 21

Cumulative frequency up to but not including the median class (F) =27

$$Median = l_l + \frac{l_2 - l_1}{f}\left(\frac{N}{2} - F\right)$$

$$Median = 30 + \frac{35 - 30}{21}(37.5 - 27)$$

$$Median = 32.5$$

$$i.e.\ Q2 = 32.5$$

For third quartile (Q3)

$$\therefore \quad \frac{3N}{4} = 56.25$$

$$Q3 = l_l + \frac{l_2 - l_1}{f}\left(\frac{3N}{4} - F\right)$$

Q3 class (l_1 –l_2) = 35-40

Frequency of Q3 class (f) = 16

Cumulative frequency up to but not including the Q3 class (F) =48

$$Q3 = l_l + \frac{l_2 - l_1}{f}\left(\frac{3N}{4} - F\right)$$

$$Q3 = 35 + \frac{40 - 35}{16}(56.25 - 48)$$

$$Q3 = 37.5781$$

$$Bowley's \;\; coefficient\ of\ skewness = \frac{(Q_3 + Q_1 - 2\,X\,Median)}{(Q_3 - Q_1)}$$

$$Bowley's \;\; coefficient\ of\ skewness = \frac{(37.5781 + 26.8269 - 2\,X\,32.5)}{37.5781 - 26.8269}$$

$$Bowley's \;\; coefficient\ of\ skewness = -0.0553$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.1:

Obtain the equation of regression line for the following values of x and y

x	1	2	3	4	5
y	2	5	3	8	7

Find the value of corresponding regression coefficient.

If x = 2.45, find the value of y.

Solution:

We prepare following table

Here n = 5

x	y	x^2	y^2	xy
1	2	1	4	2
2	5	4	25	10
3	3	9	9	9
4	8	16	64	32
5	7	25	49	35
$\sum x = 15$	$\sum y = 25$	$\sum x^2 = 55$	$\sum y^2 = 151$	$\sum xy = 88$

$$\bar{x} = \frac{\sum x_i}{n} = 3$$

$$\bar{y} = \frac{\sum y_i}{n} = 5$$

$$b_{yx} = \frac{\frac{\sum x_i y_i}{n} - \frac{\sum x_i}{n}\frac{\sum y_i}{n}}{\frac{\sum {x_i}^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}$$

$$b_{yx} = \frac{\frac{88}{5} - \frac{15}{5}\frac{25}{5}}{\frac{55}{5} - \left(\frac{15}{5}\right)^2}$$

$$b_{yx} = \frac{17.6 - 15}{11 - 9}$$

$$b_{yx} = 1.3$$

Regression of y on x is

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

$$(y - 5) = 1.3(x - 3)$$

y-5 = 1.3 x -3.9

y=1.3 x + 1.1

$y = 1.3\,x + \ \ 1.1$

$when\ x = 2.45$

Y = 4.285

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 2.2:
Obtain the equation of regression line for the following values of x and y

x	1	3	4	6	8	9	11	14
y	1	2	4	4	5	7	8	9

Find the value of corresponding regression coefficient.

Solution:
We prepare following table
Here n = 8

x	y	x^2	y^2	xy
1	1	1	1	1
3	2	9	4	6
4	4	16	16	16
6	4	36	16	24
8	5	64	25	40
9	7	81	49	63
11	8	121	64	88
14	9	196	81	126
$\sum x = 56$	$\sum y = 40$	$\sum x^2 = 524$	$\sum y^2 = 256$	$\sum xy = 364$

$$\bar{x} = \frac{\sum x_i}{n} = 7$$

$$\bar{y} = \frac{\sum y_i}{n} = 5$$

$$b_{yx} = \frac{\frac{\sum x_i y_i}{n} - \frac{\sum x_i}{n}\frac{\sum y_i}{n}}{\frac{\sum {x_i}^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}$$

$$b_{yx} = \frac{\frac{364}{8} - \frac{56}{8}\frac{40}{8}}{\frac{524}{8} - \left(\frac{56}{8}\right)^2}$$

$$b_{yx} = 0.6364$$

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

$$b_{xy} = \frac{\frac{364}{8} - \frac{56}{8}\frac{40}{8}}{\frac{256}{8} - \left(\frac{40}{8}\right)^2}$$

$$b_{xy} = 1.5$$

Regression of y on x is

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

$$(y - 5) = 0.6364(x - 7)$$

$$y = 0.6364x + 0.5452$$

Regression of x on y is

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

$$(x - 7) = 1.5(y - 5)$$

$$x = 1.5y - 0.5$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.3:

The table given below is of production (in thousand tons) of a sugar factory

Year	1969	1970	1971	1972	1973	1974	1975
Production	77	88	94	85	91	98	90

also find Karl pearson's coefficient of correlation.

Solution:

We prepare following table

Here n = 7

year	X= year-1969	y	x^2	y^2	xy
1969	0	77	0	5929	0
1970	1	88	1	7744	88
1971	2	94	4	8836	188
1972	3	85	9	7225	255
1973	4	91	16	8281	364
1974	5	98	25	9604	490
1975	6	90	36	8100	540
	$\sum x = 21$	$\sum y = 623$	$\sum x^2 = 91$	$\sum y^2 = 55719$	$\sum xy = 1925$

$$\bar{x} = \frac{\sum x_i}{n} = 3$$

$$\bar{y} = \frac{\sum y_i}{n} = 89$$

$$b_{yx} = \frac{\frac{\sum x_i y_i}{n} - \frac{\sum x_i}{n}\frac{\sum y_i}{n}}{\frac{\sum {x_i}^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}$$

$b_{yx} = 2$

Regression of y on x is

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

$$(y - 89) = 2(x - 3)$$

$$y = 2x + 83$$

Karl pearson's coefficient of correlation is given by

$$\gamma(x, y) = \frac{\frac{\sum xy}{n} - \frac{\sum x}{n}\frac{\sum y}{n}}{\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}\sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}}$$

$$\gamma\ (x, y) = \frac{\frac{1925}{7} - \frac{21}{7}\,\frac{623}{7}}{\sqrt{\frac{91}{7} - \left(\frac{21}{7}\right)^2}\ \sqrt{\frac{55719}{7} - \left(\frac{623}{7}\right)^2}}$$

$$\gamma\ (x, y) = \frac{275 - 267}{\sqrt{13 - 9}\,\sqrt{7959.86 - 7921}}$$

$$\gamma\ (x, y) = 0.6417$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.4:

Following data represents assets of a multinational company in crores of rupees for the year 1995 to 2000.

Year	1995	1996	1997	1998	1999	2000
Asset	83	92	71	90	110	115

Find the regression of asset on year. Estimate the asset for the year 2002. Also find Karl pearson's coefficient of correlation.

Solution:

We prepare following table

Here n = 6

year	X= year-1995	y	x^2	y^2	xy
1995	0	83	0	6889	0
1996	1	92	1	8464	92
1997	2	71	4	5041	142
1998	3	90	9	8100	270
1999	4	110	16	12100	440
2000	5	115	25	13225	575
	$\sum x = 15$	$\sum y = 561$	$\sum x^2 = 55$	$\sum y^2 = 53819$	$\sum xy = 1519$

$$\bar{x} = \frac{\sum x_i}{n} = 2.5$$

$$\bar{y} = \frac{\sum y_i}{n} = 93.5$$

$$b_{yx} = \frac{\frac{\sum x_i y_i}{n} - \frac{\sum x_i}{n} \frac{\sum y_i}{n}}{\frac{\sum {x_i}^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}$$

$$b_{yx} = \frac{\frac{1519}{6} - \frac{15}{6} \frac{561}{6}}{\frac{55}{6} - \left(\frac{15}{6}\right)^2}$$

$$b_{yx} = 6.6571$$

Regression of y on x is

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

$$(y - 93.5) = 6.6571(x - 2.5)$$

$$y = 6.6571\, x + 76.8571$$

To estimate asset in the year 2002

X= 2002-1995

X=7

Hence y-93.5 = byx (7-2.5)

Y=123.4574

Karl pearson's coefficient of correlation is given by

$$\gamma\,(x,y) = \frac{\frac{\sum xy}{n} - \frac{\sum x}{n}\frac{\sum y}{n}}{\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}\sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}}$$

$$\gamma\,(x,y) = \frac{\frac{1519}{6} - \frac{15}{6}\frac{561}{6}}{\sqrt{\frac{55}{6} - \left(\frac{15}{6}\right)^2}\sqrt{\frac{53819}{6} - \left(\frac{561}{6}\right)^2}}$$

$$\gamma\,(x,y) = \frac{253.1667 - 233.75}{\sqrt{9.1667 - 6.25}\sqrt{8969.8333 - 8742.25}}$$

$\gamma\,(x,y)$ =0.7536

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 2.5:
Find Karl Pearson's coefficient of correlation for the following data.

X	1	2	3
Y	6	5	10

Solution:
We prepare following table
Here n = 3

X	y	x^2	y^2	xy
1	6	1	36	6
2	5	4	25	10
3	10	9	100	30
$\sum x = 6$	$\sum y = 21$	$\sum x^2 = 14$	$\sum y^2 = 161$	$\sum xy = 46$

Karl Pearson's coefficient of correlation is given by

$$\gamma\,(x,y) = \frac{\frac{\sum xy}{n} - \frac{\sum x}{n}\frac{\sum y}{n}}{\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}\sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}}$$

$$\gamma\ (x, y) = \frac{\frac{46}{3} - \frac{6}{3}\ \frac{21}{3}}{\sqrt{\frac{14}{3} - \left(\frac{6}{3}\right)^2}\ \sqrt{\frac{161}{3} - \left(\frac{21}{3}\right)^2}}$$

$$\gamma\ (x, y) = 0.7559$$

◈◈◈◈◈◈◈

MUMBAI EDUCATIONAL TRUST
MET Institute of Computer Science

विद्याधनं सर्वधन प्रधानम् ।

THE MET LEAGUE OF COLLEGES MET AS SHARP AS YOU CAN GET

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.6:

The following data represents 10 students marks in statistics (x) and probability (y). Find Karl pearson's coefficient of correlation for the following data.

X	56	55	58	58	57	56	60	54	59	57
Y	68	67	67	70	65	68	70	66	68	66

Solution:

We prepare following table

Here n = 10

X	y	x^2	y^2	xy
56	68	3136	4624	3808
55	67	3025	4489	3685
58	67	3364	4489	3886
58	70	3364	4900	4060
57	65	3249	4225	3705
56	68	3136	4624	3808
60	70	3600	4900	4200
54	66	2916	4356	3564
59	68	3481	4624	4012
57	66	3249	4356	3762
$\sum x = 570$	$\sum y = 675$	$\sum x^2 = 32520$	$\sum y^2 = 45587$	$\sum xy = 38490$

Karl Pearson's coefficient of correlation is given by

$$\gamma\,(x,y) = \frac{\frac{\sum xy}{n} - \frac{\sum x}{n}\,\frac{\sum y}{n}}{\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}\,\sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}}$$

$$\gamma\,(x,y) = \frac{\frac{38490}{10} - \frac{570}{10}\,\frac{675}{10}}{\sqrt{\frac{32520}{10} - \left(\frac{570}{10}\right)^2}\,\sqrt{\frac{45587}{10} - \left(\frac{675}{10}\right)^2}}$$

$$\gamma\,(x,y) = 0.5533$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.7:

The marks obtained by 9 students in OS and DS are given below

Marks in OS	35	47	23	6	17	10	43	9	28
Marks in DS	30	46	33	4	23	8	48	12	31

Compute the ranks in two subjects and the coefficient of correlation of ranks.

Solution:

We prepare following table

Here n = 9

Marks in OS	Marks in DS	Rank in OS (r_1)	Rank in DS (r_2)	d = $\lvert r_1 - r_2 \rvert$	d^2
35	30	3	5	2	4
47	46	1	2	1	1
23	33	5	3	2	4
6	4	9	9	0	0
17	23	6	6	0	0
10	8	7	8	1	1
43	48	2	1	1	1
9	12	8	7	1	1
28	31	4	4	0	0
					$\sum d^2 = 12$

Spearman's rank correlation coefficient is given by

$$R = 1 - \frac{6 \sum_{i=1}^{n} {d_i}^2}{n\,(n^2 - 1)}$$

$$R = 1 - \frac{6\ X\ 12}{9\ (81 - 1)}$$

$$R = 1 - \frac{72}{720}$$

$$R = \frac{9}{10}$$

$$R = 0.9$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.8:

Find Spearman's rank correlation for the following data

Marks in CPP	64	50	44	42	56	65	59
Marks in SEPM	80	60	37	51	30	75	44

Solution:

We prepare following table

Here n = 7

Marks in CPP	Marks in SEPM	Rank in CPP (r_1)	Rank in SEPM (r_2)	d = $\lvert r_1 - r_2 \rvert$	d^2
64	80	2	1	1	1
50	60	5	3	2	4
44	37	6	6	0	0
42	51	7	4	3	9
56	30	4	7	3	9
65	75	1	2	1	1
59	44	3	5	2	4
					$\sum d^2 = 28$

Spearman's rank correlation coefficient is given by

$$R = 1 - \frac{6 \sum_{i=1}^{n} {d_i}^2}{n\,(n^2 - 1)}$$

$$R = 1 - \frac{6\ X\ 28}{7\,(49 - 1)}$$

$$R = 1 - \frac{168}{336}$$

$$R = \frac{168}{336}$$

$$R = 0.5$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 2.9:

In a sample of 12 fathers and their eldest sons gave the following data about their heights in inches.

Father(x)	65	63	67	64	68	62	70	66	68	67	69	71
Son(y)	68	66	68	65	69	66	68	65	71	67	68	70

Calculate coefficient of rank correlation between x and y.

Solution:

We prepare following table, Here n = 12

Father (x)	Son (y)	Rank in x (r_1)	Rank in y (r_2)	d = $\lvert r_1 - r_2 \rvert$	d^2
65	68	9	5.5	3.5	12.25
63	66	11	9.5	1.5	2.25
67	68	6.5	5.5	1	1
64	65	10	11.5	1.5	2.25
68	69	4.5	3	1.5	2.25
62	66	12	9.5	2.5	6.25
70	68	2	5.5	3.5	12.25
66	65	8	11.5	3.5	12.25
68	71	4.5	1	3.5	12.25
67	67	6.5	8	1.5	2.25
69	68	3	5.5	2.5	6.25
71	70	1	2	1	1
					$\sum d^2 = 72.5$

(i) Rank 4.5 repeated m = 2 times

Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 0.5$

(ii) Rank 6.5 is repeated m=2 times

Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 0.5$

(iii) Rank 5.5 is repeated m=4 times

Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 5$

(iv) Rank 9.5 is repeated m=2 times

Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 0.5$

(v) Rank 11.5 is repeated m=2 times

Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 0.5$

Corrected $\sum d^2 = 72.5 + 0.5 + 0.5 + 5 + 0.5 + 0.5$

$$\sum d^2 = 79.5$$

Spearman's rank correlation coefficient is given by

$$R = 1 - \frac{6\sum_{i=1}^{n} {d_i}^2}{n\,(n^2-1)}$$

$$R = 1 - \frac{6\ X\ 79.5}{12\ (144 - 1)}$$

$$R = 1 - \frac{477}{1716}$$

$$R = 0.7220$$

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.10:

Find Spearman's rank correlation for the following data

Student	A	B	C	D	E	F	G	H
Marks in Test1	52	34	47	65	43	34	54	65
Marks in Test2	65	59	65	68	82	60	57	58

Solution:

We prepare following table, Here n = 8

Marks in Test1	Marks in Test2	Rank in Test1 (r_1)	Rank in Test2 (r_2)	d = $\|r_1 - r_2\|$	d^2
52	65	4	3.5	0.5	0.25
34	59	7.5	6	1.5	2.25
47	65	5	3.5	1.5	2.25
65	68	1.5	2	0.5	0.25
43	82	6	1	5	25
34	60	7.5	5	2.5	6.25
54	57	3	8	5	25
65	58	1.5	7	5.5	30.25
					$\sum d^2$ =91.5

(i) Rank 1.5 repeated m = 2 times

Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 0.5$

(ii) Rank 7.5 is repeated m=2 times

Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 0.5$

(iii) Rank 3.5 is repeated m=2 times

Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 0.5$

Corrected $\sum d^2 = 91.5 + 0.5 + 0.5 + 0.5$

$$\sum d^2 = 93$$

Spearman's rank correlation coefficient is given by

$$R = 1 - \frac{6 \sum_{i=1}^{n} {d_i}^2}{n\,(n^2 - 1)}$$

$$R = 1 - \frac{6\, X\, 93}{8\,(64 - 1)}$$

$$R = 1 - \frac{558}{504}$$

$$R = -0.1071$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.11:

Calculate the coefficient of rank correlation from the following data:

X	48	33	40	9	16	16	65	24	16	57
Y	13	13	24	6	15	4	20	9	6	19

Solution:

We prepare following table, Here n = 10

X	Y	Rank in X (r_1)	Rank in Y (r_2)	d = $\lvert r_1 - r_2 \rvert$	d^2
48	13	3	5.5	2.5	6.25
33	13	5	5.5	0.5	0.25
40	24	4	1	3	9
9	6	10	8.5	1.5	2.25
16	15	8	4	4	16
16	4	8	10	2	4
65	20	1	2	1	1
24	9	6	7	1	1
16	6	8	8.5	0.5	0.25
57	19	2	3	1	1
					$\sum d^2 = 41$

(i) Rank 5.5 repeated m = 2 times
Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 0.5$

(ii) Rank 8.5 is repeated m=2 times
Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 0.5$

(iii) Rank 8 is repeated m=3 times
Hence the correlation coefficient $\frac{m(m^2-1)}{12} = 2$

Corrected $\sum d^2 = 41 + 0.5 + 0.5 + 2$

$$\sum d^2 = 44$$

Spearman's rank correlation coefficient is given by

$$R = 1 - \frac{6 \sum_{i=1}^{n} {d_i}^2}{n\,(n^2-1)}$$

$$R = 1 - \frac{6\ X\ 44}{10\,(100-1)}$$

$$R = 1 - \frac{264}{990}$$

$$R = 0.7333$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.12:

The ranking of ten students in two subjects A and B as follows:

Subject A	3	5	8	4	7	10	2	1	6	9
Subject B	6	4	9	8	1	2	3	10	5	7

Find Spearman's Rank correlation coefficient.

Solution:

We prepare following table, Here n = 10

Rank in subject A (r_1)	Rank in subject B (r_2)	d = $\lvert r_1 - r_2 \rvert$	d^2
3	6	3	9
5	4	1	1
8	9	1	1
4	8	4	16
7	1	6	36
10	2	8	64
2	3	1	1
1	10	9	81
6	5	1	1
9	7	2	4
			$\sum d^2 = 214$

Spearman's rank correlation coefficient is given by

$$R = 1 - \frac{6 \sum_{i=1}^{n} {d_i}^2}{n\,(n^2 - 1)}$$

$$R = 1 - \frac{6\, X\, 214}{10\,(100 - 1)}$$

$$R = 1 - \frac{1284}{990}$$

$$R = -0.29697$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.13:

Compute the quadratic regression equation of following data. Check its best fitness.

x	-3	-2	-1	0	1	2	3
y	7.5	3	0.5	1	3	6	14

Solution:

We prepare following table, Here N=7

x	y	x^2	x^3	x^4	xy	x^2y
-3	7.5	9	-27	81	-22.5	67.5
-2	3	4	-8	16	-6	12
-1	0.5	1	-1	1	-0.5	0.5
0	1	0	0	0	0	0
1	3	1	1	1	3	3
2	6	4	8	16	12	24
3	14	9	27	81	42	126
$\sum x = 0$	$\sum y = 35$	$\sum x^2 = 28$	$\sum x^3 = 0$	$\sum x^4 = 196$	$\sum xy = 28$	$\sum x^2y = 233$

$$\begin{bmatrix} N & \sum x_i & \sum {x_i}^2 \\ \sum x_i & \sum {x_i}^2 & \sum {x_i}^3 \\ \sum {x_i}^2 & \sum {x_i}^3 & \sum {x_i}^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum {x_i}^2 y_i \end{bmatrix}$$

The equations are

$$a_0 N \quad + \; a_1 \sum x_i \quad + \; a_2 \sum {x_i}^2 \quad = \sum y_i$$

$$a_0 \sum x_i + \; a_1 \sum {x_i}^2 \; + \; a_2 \sum {x_i}^3 \quad = \sum x_i y_i$$

$$a_0 \sum {x_i}^2 + \; a_1 \sum {x_i}^3 \; + \; a_2 \sum {x_i}^4 \quad = \sum {x_i}^2 y_i$$

Substituting values the equations are

$7a_0$ + 0 + $28a_2$ = 35 -----eq1

0 + $28a_1$ + 0 = 28 ------eq2

$28a_0$ + 0 + $196a_2$ = 233 -------eq3

Solving the simultaneous equations eq1 and eq3

$7a_0$ + $28a_2$ = 35

$28a_0$ + $196a_2$ = 233

Hence a_2 = 1.107142 and a_0 = 0.571429

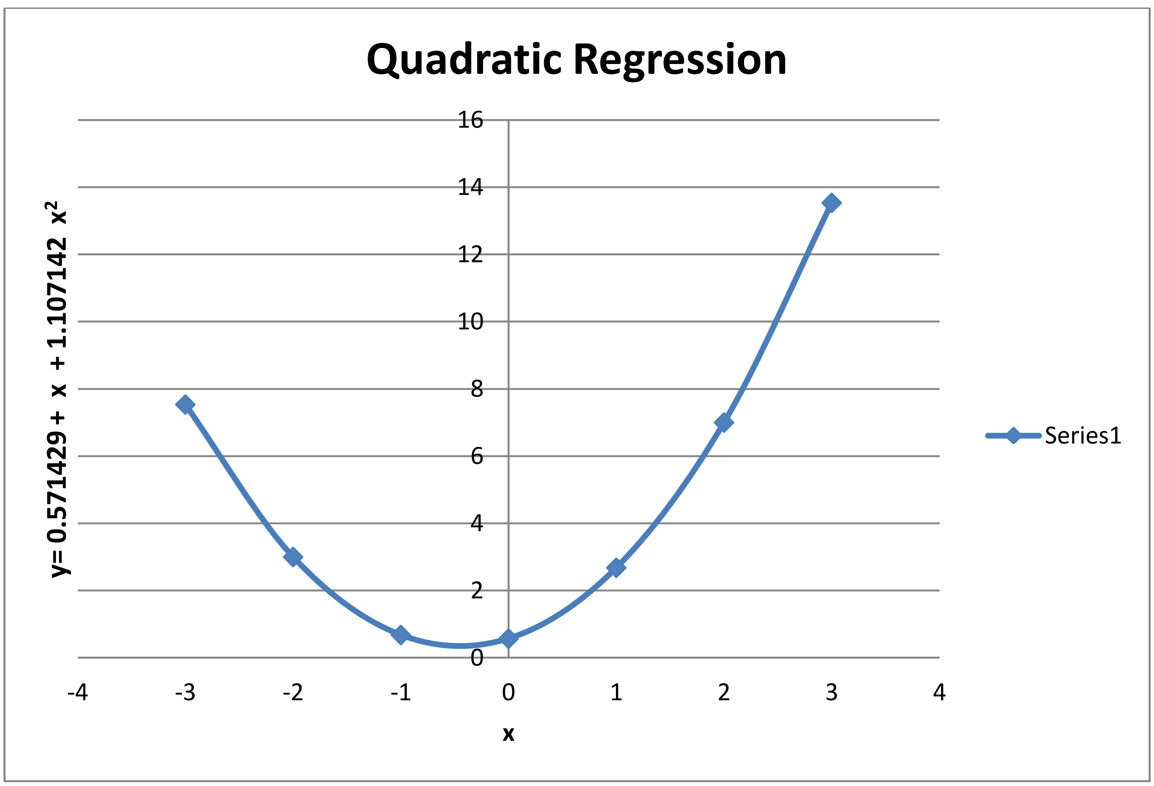
We get $a_1 = 1$ using equation 2

Hence the quadratic equation is

$$y = 0.571429 + x + 1.107142\ x^2$$

$$y = 1.107142\ x^2 + x + 0.571429$$

The following plot is just for reference. If asked then only draw it.



◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 2.14:

Fit a least-squares parabola of the form $Y = a_0 + a_1x + a_2x^2$ to the set of data in Table

X	1.2	1.8	3.1	4.9	5.7	7.1	8.6	9.8
y	4.5	5.9	7.0	7.8	7.2	6.8	4.5	2.7

Solution:

We prepare following table, Here N=8

x	y	x^2	x^3	x^4	xy	x^2y
1.2	4.5	1.44	1.728	2.0736	5.4	6.48
1.8	5.9	3.24	5.832	10.498	10.62	19.12
3.1	7	9.61	29.791	92.352	21.7	67.27
4.9	7.8	24.01	117.649	576.48	38.22	187.3
5.7	7.2	32.49	185.193	1055.6	41.04	233.9
7.1	6.8	50.41	357.911	2541.2	48.28	342.8
8.6	4.5	73.96	636.056	5470.1	38.7	332.8
9.8	2.7	96.04	941.192	9223.7	26.46	259.3
$\sum x =$ 42.2	$\sum y =$ 46.4	$\sum x^2 =$ 291.2	$\sum x^3 =$ 2275.35	$\sum x^4 =$ 18971.92	$\sum xy =$ 230.42	$\sum x^2y =$ 1449

$$\begin{bmatrix} N & \sum x_i & \sum {x_i}^2 \\ \sum x_i & \sum {x_i}^2 & \sum {x_i}^3 \\ \sum {x_i}^2 & \sum {x_i}^3 & \sum {x_i}^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \sum {x_i}^2 y_i \end{bmatrix}$$

The equations are

$$a_0 N \quad + \; a_1 \sum x_i \quad + \; a_2 \sum {x_i}^2 \quad = \sum y_i$$

$$a_0 \sum x_i + \; a_1 \sum {x_i}^2 \; + \; a_2 \sum {x_i}^3 \quad = \sum x_i y_i$$

$$a_0 \sum {x_i}^2 + \; a_1 \sum {x_i}^3 \; + \; a_2 \sum {x_i}^4 \quad = \sum {x_i}^2 y_i$$

Since N = 8, the normal equations becomes

$$8a_0 + \quad 42.2a_1 + \quad 291.20a_2 = 46.4$$

$$42.2a_0 + \; 291.20a_1 + \; 2275.35a_2 = 230.42$$

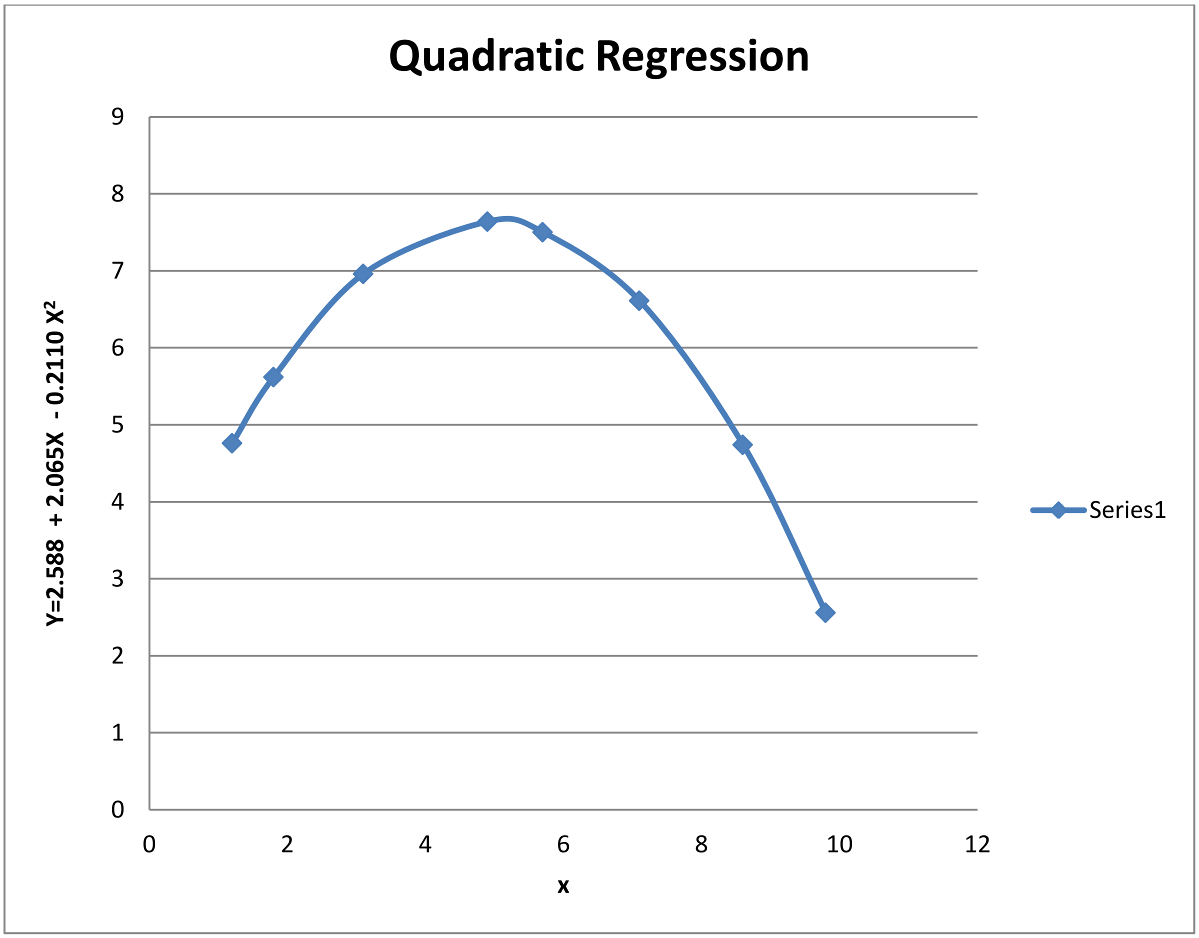
$$291.20a_0 + 2275.35a_1 + 18971.92a_2 = 1449.00$$

https://onlinemschool.com/math/assistance/equation/gaus/

Solving the equations a_0 =2.5887, a_1 = 2.0644 and a_2 = -0.2110; hence the required least-squares parabola has the equation

$Y=2.5887 + 2.0644X - 0.2110 X^2$

The following plot is just for reference. If asked then only draw it.



◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.15:

The revenue generated by a company in $billions is provided in following table, Find a quadratic least square fit for the data and estimate 2006 revenue generated

year	$billion
2000	236
2001	214
2002	207
2003	250
2004	300
2005	375

Solution:

We prepare following table, Here N=6

X=year-2000

year	x	y	x^2	x^3	x^4	xy	x^2y
2000	0	236	0	0	0	0	0
2001	1	214	1	1	1	214	214
2002	2	207	4	8	16	414	828
2003	3	250	9	27	81	750	2250
2004	4	300	16	64	256	1200	4800
2005	5	375	25	125	625	1875	9375
	$\sum x =$ 15	$\sum y =$ 1582	$\sum x^2 =$ 55	$\sum x^3 =$ 225	$\sum x^4 =$ 979	$\sum xy =$ 4453	$\sum x^2y =$ 17467

The equations are

$$a_0 N + a_1 \sum x_i + a_2 \sum x_i^2 = \sum y_i$$

$$a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 = \sum x_i y_i$$

$$a_0 \sum x_i^2 + a_1 \sum x_i^3 + a_2 \sum x_i^4 = \sum x_i^2 y_i$$

Hence we will get

$$6a_0 + 15a_1 + 55a_2 = 1582$$

$$15a_0 + 55a_1 + 225a_2 = 4453$$

$$55a_0 + 225a_1 + 979a_2 = 17467$$

Solving the equations $a_0 = 234.9643$, $a_1 = -35.2036$ and $a_2 = 12.7321$; hence the required least-squares parabola has the equation

$$\mathbf{Y = 234.9643 - 35.2036X + 12.7321\,X^2}$$

The expenses of 2006 will be substituting x= 6

Y=482.0983

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.16:

Find a cubic least square fit for the following data.

xi	yi
-3	7.5
-2	3
-1	0.5
0	1
1	3
2	6
3	14

Solution:

We prepare following table, Here N=7

x	y	x^2	x^3	x^4	x^5	x^6	xy	x^2y	X^3y
-3	7.5	9	-27	81	-243	729	-22.5	67.5	-202.5
-2	3	4	-8	16	-32	64	-6	12	-24
-1	0.5	1	-1	1	-1	1	-0.5	0.5	-0.5
0	1	0	0	0	0	0	0	0	0
1	3	1	1	1	1	1	3	3	3
2	6	4	8	16	32	64	12	24	48
3	14	9	27	81	243	729	42	126	378
$\sum x = 0$	$\sum y = 35$	$\sum x^2 = 28$	$\sum x^3 = 0$	$\sum x^4 = 196$	$\sum x^5 = 0$	$\sum x^6 = 1588$	$\sum xy = 28$	$\sum x^2y = 233$	$\sum x^3y = 202$

$$\begin{bmatrix} N & \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \sum x_i^5 \\ \sum x_i^3 & \sum x_i^4 & \sum x_i^5 & \sum x_i^6 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i x_i^2 \\ \sum y_i x_i^3 \end{bmatrix}$$

The equations are (here N is number of data points given)

$$a_0 N + a_1 \sum x_i + a_2 \sum x_i^2 + a_3 \sum x_i^3 = \sum y_i$$

$$a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 + a_3 \sum x_i^4 = \sum x_i y_i$$

$$a_0 \sum x_i^2 + a_1 \sum x_i^3 + a_2 \sum x_i^4 + a_3 \sum x_i^5 = \sum x_i^2 y_i$$

$$a_0 \sum x_i^3 + a_1 \sum x_i^4 + a_2 \sum x_i^5 + a_3 \sum x_i^6 = \sum x_i^3 y_i$$

Solving the simultaneous equation obtain the values of coefficients a_0 , a_1, a_2 and a_3

$$\begin{bmatrix} 7 & 0 & 28 & 0 \\ 0 & 28 & 0 & 196 \\ 28 & 0 & 196 & 0 \\ 0 & 196 & 0 & 1588 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 35 \\ 28 \\ 233 \\ 202 \end{bmatrix}$$

$$7a_0 + 0 + 28a_2 + 0 = 35$$

$$0 + 28a_1 + 0 + 196\, a_3 = 28$$

$$28a_0 + 0 + 196a_2 + 0 = 233$$

$$0 + 196a_1 + 0 + 1588\, a_3 = 202$$

$$7a_0 \quad + \ 28a_2 \ = \ 35 \qquad \text{---eq1}$$

$$28a_1 \ + \ 196\,a_3 \ = \ 28 \qquad \text{---eq2}$$

$$28a_0 + \ 196a_2 \ = \ 233 \qquad \text{---eq3}$$

$$196a_1 \ + \ 1588\,a_3 \ = \ 202 \quad \text{--eq4}$$

Solve **eq1 and eq3**

$$7a_0 \quad + \ 28a_2 \ = \ 35$$

$$28a_0 + \ 196a_2 \ = \ 233$$

Solve **eq2 and eq4**

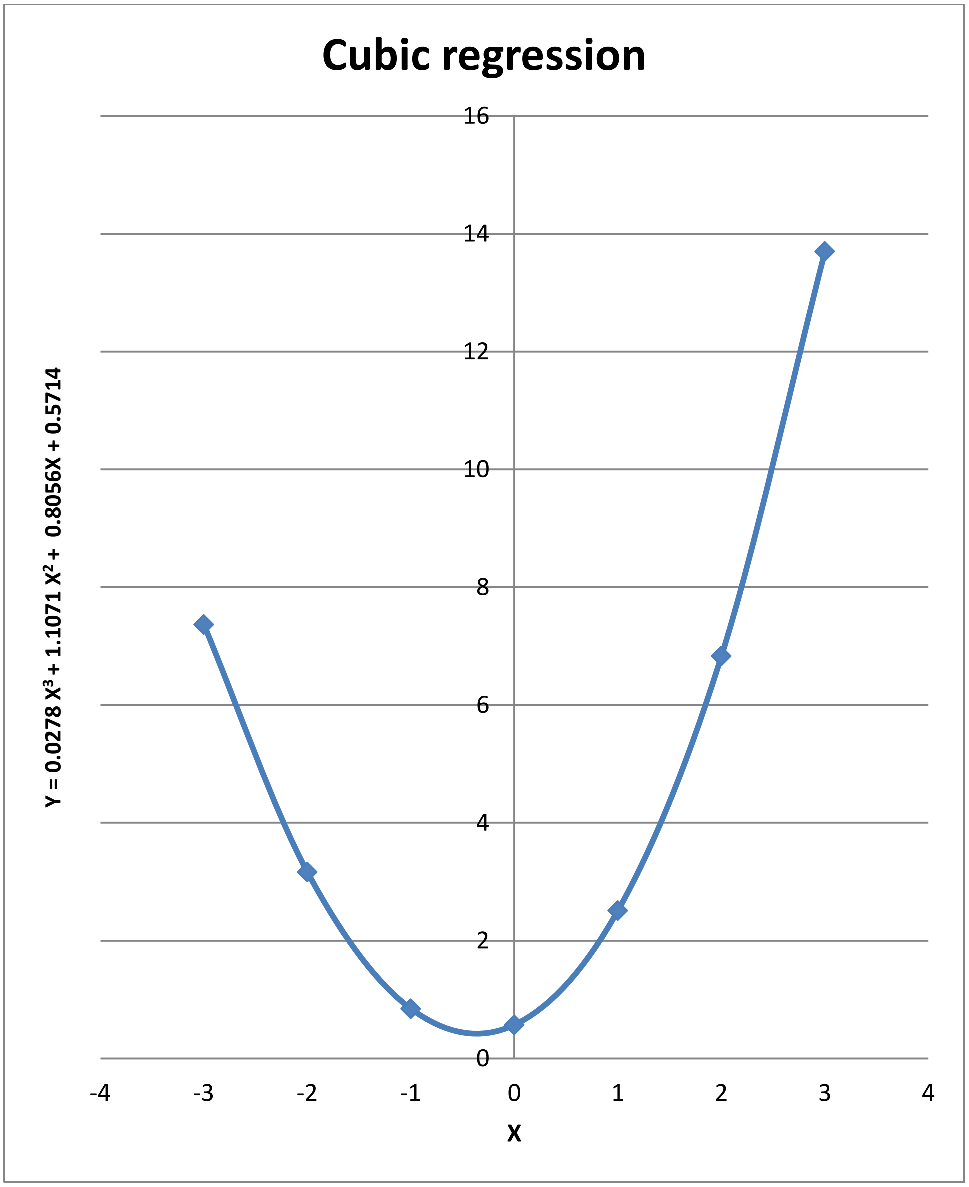
$$28a_1 \ + \ 196\,a_3 \ = \ 28$$

$$196a_1 \ + \ 1588\,a_3 \ = \ 202$$

Solving the equations a_0 =0.5714, a_1 = 0.8056, a_2 =1.1071 and a_3 = 0.0278; hence the required least-squares cubic curve is

$Y = 0.5714 + 0.8056X + 1.1071\,X^2 + 0.0278\,X^3$

$Y = 0.0278\,X^3 + 1.1071\,X^2 + 0.8056X + 0.5714$


Cubic regression
Y = 0.0278 X³ + 1.1071 X² + 0.8056X + 0.5714
16
14
12
10
8
6
4
2
0
-4
-3
-2
-1
0
1
2
3
4
X

◈◈◈◈◈◈◈

MUMBAI EDUCATIONAL TRUST
MET Institute of Computer Science

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 2.17:

Find the cubic regression for the following data

X	-3	-2	-1	0	1	2	3
y	-27	-8	-1	0	1	8	27

Solution:

We prepare following table, Here N=7

x	y	x^2	x^3	x^4	x^5	x^6	xy	x^2y	X^3y
-3	-27	9	-27	81	-243	729	81	-243	729
-2	-8	4	-8	16	-32	64	16	-32	64
-1	-1	1	-1	1	-1	1	1	-1	1
0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1
2	8	4	8	16	32	64	16	32	64
3	27	9	27	81	243	729	81	243	729
$\sum x = 0$	$\sum y = 0$	$\sum x^2 = 28$	$\sum x^3 = 0$	$\sum x^4 = 196$	$\sum x^5 = 0$	$\sum x^6 = 1588$	$\sum xy = 196$	$\sum x^2y = 0$	$\sum x^3y = 1588$

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \sum x_i^5 \\ \sum x_i^3 & \sum x_i^4 & \sum x_i^5 & \sum x_i^6 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i x_i^2 \\ \sum y_i x_i^3 \end{bmatrix}$$

The equations are (here N is number of data points given)

$$a_0 N + a_1 \sum x_i + a_2 \sum x_i^2 + a_3 \sum x_i^3 = \sum y_i$$

$$a_0 \sum x_i + a_1 \sum x_i^2 + a_2 \sum x_i^3 + a_3 \sum x_i^4 = \sum x_i y_i$$

$$a_0 \sum x_i^2 + a_1 \sum x_i^3 + a_2 \sum x_i^4 + a_3 \sum x_i^5 = \sum x_i^2 y_i$$

$$a_0 \sum x_i^3 + a_1 \sum x_i^4 + a_2 \sum x_i^5 + a_3 \sum x_i^6 = \sum x_i^3 y_i$$

Solving the simultaneous equation obtain the values of coefficients a_0 , a_1, a_2 and a_3

$$\begin{bmatrix} 7 & 0 & 28 & 0 \\ 0 & 28 & 0 & 196 \\ 28 & 0 & 196 & 0 \\ 0 & 196 & 0 & 1588 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 196 \\ 0 \\ 1588 \end{bmatrix}$$

$7\,a_0 + 28\,a_2 = 0$ ----------eq1

$28\,a_1 + 196\,a_3 = 196$ --------eq2

$28\,a_0 + 196\,a_2 = 0$ ---------eq3

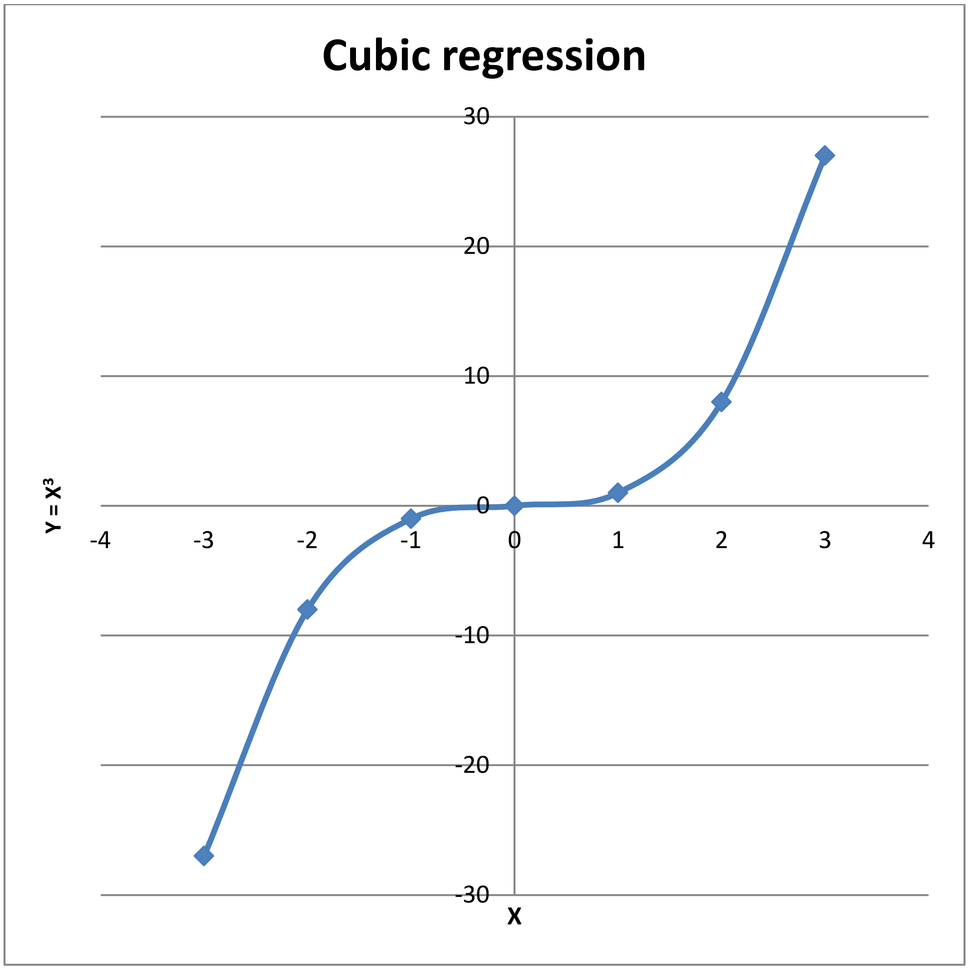
$196\,a_1 + 1588\,a_3 = 1588$ ----eq4

Solving **eq1 and eq3** , solving **eq2 and eq4** we get

a_0 =0, a_1 = 0 , a_2 =0 and a_3 = 1; hence the required least-squares cubic curve is

$$Y = 1\,X^3 + 0\,X^2 + 0\,X + 0$$

i.e. $y = x^3$



◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 2.18:

Find the cubic regression for the following data

X	-1	0	2	3
y	0	-12	0	0

Solution:

We prepare following table, Here N=4

x	y	x^2	x^3	x^4	x^5	x^6	xy	x^2y	X^3y
-1	0	1	-1	1	-1	1	0	0	0
0	-12	0	0	0	0	0	0	0	0
2	0	4	8	16	32	64	0	0	0
3	0	9	27	81	243	729	0	0	0
$\sum x =$ 4	$\sum y =$ -12	$\sum x^2 =$ 14	$\sum x^3 =$ 34	$\sum x^4 =$ 98	$\sum x^5 =$ 274	$\sum x^6 =$ 794	$\sum xy =$ 0	$\sum x^2 y$ 0	$\sum x^3 y =$ 0

$$\begin{bmatrix} n & \sum x_i & \sum x_i^2 & \sum x_i^3 \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \sum x_i^4 \\ \sum x_i^2 & \sum x_i^3 & \sum x_i^4 & \sum x_i^5 \\ \sum x_i^3 & \sum x_i^4 & \sum x_i^5 & \sum x_i^6 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i x_i^2 \\ \sum y_i x_i^3 \end{bmatrix}$$

The equations are (here N is number of data points given)

$$a_0 N \; + a_1 \sum x_i \; + a_2 \sum {x_i}^2 + \; a_3 \sum {x_i}^3 \; = \sum y_i$$

$$a_0 \sum x_i + a_1 \sum {x_i}^2 + a_2 \sum {x_i}^3 + \; a_3 \sum {x_i}^4 \; = \sum x_i y_i$$

$$a_0 \sum {x_i}^2 + a_1 \sum {x_i}^3 + a_2 \sum {x_i}^4 + a_3 \sum {x_i}^5 \; = \sum {x_i}^2 y_i$$

$$a_0 \sum {x_i}^3 + a_1 \sum {x_i}^4 + a_2 \sum {x_i}^5 + a_3 \sum {x_i}^6 \; = \sum {x_i}^3 y_i$$

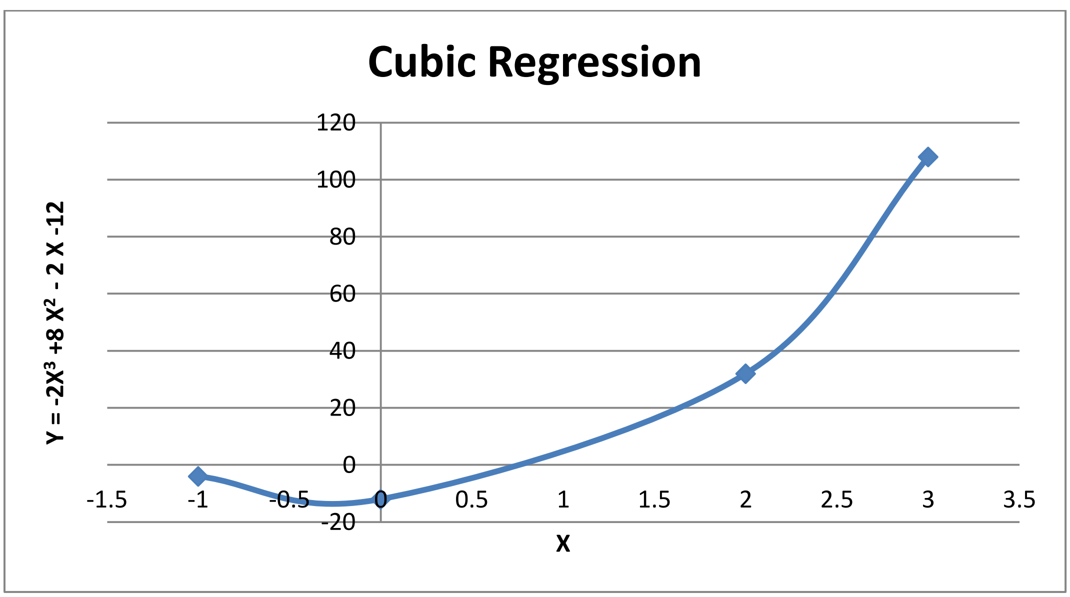
Solving the simultaneous equation obtain the values of coefficients a_0 , a_1, a_2 and a_3

$$\begin{bmatrix} 4 & 4 & 14 & 34 \\ 4 & 14 & 34 & 98 \\ 14 & 34 & 98 & 274 \\ 34 & 98 & 274 & 794 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -12 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving the equations
$a_0 = -12$, $a_1 = -2$, $a_2 = 8$, $a_3 = -2$

Hence the required least-squares cubic curve is

$$\mathbf{Y = -2X^3 + 8\,X^2 - 2\,X - 12}$$



MUMBAI EDUCATIONAL TRUST

MET Institute of Computer Science

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.1:

A bag contains 2 red, 5 white and 8 blue balls. Two balls are drawn at random from it. What is the probability that one is white and other is blue?

Solution:

Sample space = 2 + 5 + 8 = 15 balls

$$\therefore n(S) = {}^{15}C_2$$

$${}^{n}C_r = \frac{n!}{(n-r)!r!}$$

$$n(S) = \frac{15!}{13!\ 2!}$$

$$n(S) = \frac{15\ x\ 14\ x\ 13!}{13!\ 2!}$$

$$n(S) = 105$$

Event A = 1 ball is white and 1 ball is blue

$$\therefore\ n(A) = {}^{5}C_1 \text{ X } {}^{8}C_1$$

$$\therefore\ n(A) = \frac{5!}{4!\,1!}\ X\ \frac{8!}{7!\,1!}$$

$$\therefore\ n(A) = \frac{5\ x\ 4!}{4!\,1!}\ X\ \frac{8\ x\ 7!}{7!\,1!}$$

$$\therefore\ n(A) = 40$$

$$\therefore\ P(A) = \frac{n(A)}{n(S)}$$

$$\therefore\ P(A) = \frac{40}{105}$$

$$\therefore P(A) = \frac{8}{21}$$

$$\therefore P(A) = 0.38095$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.2:

Box A contains 5 red marbles and 3 blue marbles and the box B contains 3 red and 2 blue marbles. A marble is drawn at random from each box. Find the probability that

(i) Both marbles are red

(ii) One is red and one is blue

Solution:

Sample space (S) = (one marble is drawn from box A containing 8 marbles) and (one marble is drawn from box B containing 5 marbles)

$$\therefore n(S) = {}^{8}C_{1} \text{ X } {}^{5}C_{1}$$

$${}^{n}C_{r} = \frac{n!}{(n-r)!r!}$$

$$n(S) = \frac{8!}{7!\ 1!}\ X\ \frac{5!}{4!\ 1!}$$

$$n(S) = \frac{8\ x\ 7!}{7!\ 1!}\ X\ \frac{5\ x\ 4!}{4!\ 1!}$$

$$n(S) = 40$$

(i) Event A = both marbles are red

$$\therefore\ n(A) = {}^{5}C_{1}\ X\ {}^{3}C_{1}$$

$$\therefore\ n(A) = \frac{5!}{4!\,1!}\ X\ \frac{3!}{2!\,1!}$$

$$\therefore\ n(A) = \frac{5\ x\ 4!}{4!\,1!}\ X\ \frac{3\ x\ 2!}{2!\,1!}$$

$$\therefore\ n(A) = 15$$

$$\therefore\ P(A) = \frac{n(A)}{n(S)}$$

$$\therefore\ P(A) = \frac{15}{40}$$

$$\therefore P(A) = \frac{3}{8}$$

$$\therefore P(A) = 0.375$$

(ii) Event B = (1 red from box A and 1 blue from box B) or (1 red from box B and 1 blue from box A)

$$\therefore\ n(B) = {}^{5}C_{1}\ X\ {}^{2}C_{1}\ +\ {}^{3}C_{1}\ X\ {}^{3}C_{1}$$

$$\therefore\ n(B) = \frac{5!}{4!\,1!}\ X\ \frac{2!}{1!\,1!}\ +\ \frac{3!}{2!\,1!}\ X\ \frac{3!}{2!\,1!}$$

$$\therefore\ n(B) = \frac{5\ x\ 4!}{4!\,1!}\ X\ \frac{2\ x\ 1!}{1!\,1!}\ +\ \frac{3\ x\ 2!}{2!\,1!}\ X\ \frac{3\ x\ 2!}{2!\,1!}$$

$$\therefore\ n(B) = 10\ +\ 9$$

$$\therefore\ n(B) = 19$$

$$\therefore\ P(B) = \frac{n(B)}{n(S)}$$

$$\therefore\ P(B) = \frac{19}{40}$$

$$\therefore\ P(B) = 0.475$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.3:

A box contains 2 white socks and 2 blue socks. Two are drawn at random. Find the probability, p, that they are a match.

Solution:

Sample space (S) = 2 socks are drawn from 4

$$\therefore n(S) = {}^{4}C_{2}$$

$$^{n}C_{r} = \frac{n!}{(n-r)!r!}$$

$$n(S) = \frac{4!}{2!\ 2!}$$

$$n(S) = \frac{4\ x\ 3\ x\ 2!}{2!\ 2!}$$

$$n(S) = 6$$

Event A = both socks are white or both are blue

$$\therefore\ n(A) = {}^{2}C_{2} + {}^{2}C_{2}$$

$$\therefore\ n(A) = \frac{2!}{0!\,2!} + \frac{2!}{0!\,2!}$$

$$\therefore\ n(A) = 1 + 1$$

$$\therefore\ n(A) = 2$$

$$\therefore\ P(A) = \frac{n(A)}{n(S)}$$

$$\therefore\ P(A) = \frac{2}{6}$$

$$\therefore P(A) = \frac{1}{3}$$

$$\therefore P(A) = 0.3333$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.4:

A box contains 15 chips where 5 are defective. If the random samples of 3 chips are drawn, what is the probability that exactly two are defective?

Solution:

Sample space (S) = 3 chips are drawn randomly from a box containing 15 chips

$$\therefore n(S) = {}^{15}C_3$$

$$ {}^{n}C_r = \frac{n!}{(n-r)!r!}$$

$$n(S) = \frac{15!}{12!\ 3!}$$

$$n(S) = \frac{15\ x\ 14\ x\ 13\ x\ 12!}{12!\ 3\ x\ 2\ x1}$$

$$n(S) = 455$$

Event A = 3 chips are drawn and 2 are defective

$$\therefore\ n(A) = {}^{5}C_2\ X\ {}^{10}C_1$$

$$\therefore\ n(A) = \frac{5!}{3!\,2!} X \frac{10!}{9!\,1!}$$

$$\therefore\ n(A) = \frac{5\ x\ 4\ x\ 3!}{3!\,2!} X \frac{10\ x\ 9!}{9!\,1!}$$

$$\therefore\ n(A) = 10\ x\ 10$$

$$\therefore\ n(A) = 100$$

$$\therefore\ P(A) = \frac{n(A)}{n(S)}$$

$$\therefore\ P(A) = \frac{100}{455}$$

$$\therefore P(A) = \frac{20}{91}$$

$$\therefore P(A) = 0.21978$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.5:

Five horses are in race. X person picks two of the horses at random and bets on them. Find the probability that X picked the winner?

Solution:

Sample space (S) = picking 2 horses from a set of 5

$$\therefore n(S) = {}^{5}C_{2}$$

$$ {}^{n}C_{r} = \frac{n!}{(n-r)!r!} $$

$$n(S) = \frac{5!}{3!\ 2!}$$

$$n(S) = \frac{5\ x\ 4\ x\ 3!}{3!\ \ 2\ x 1}$$

$$n(S) = 10$$

Event A = 1 horse is winner and other is not

$$\therefore\ n(A) = {}^{1}C_{1}\ \text{X}\ {}^{4}C_{1}$$

$$\therefore\ n(A) = \frac{1!}{0!\,1!} X \frac{4!}{3!\,1!}$$

$$\therefore\ n(A) = 1X4$$

$$\therefore\ n(A) = 4$$

$$\therefore\ P(A) = \frac{n(A)}{n(S)}$$

$$\therefore\ P(A) = \frac{4}{10}$$

$$\therefore P(A) = 0.4$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.6:

Two cards are drawn from a pack of 52 playing cards, find the probability that both the cards are kings.

Solution:

Sample space (S) = 2 cards are drawn from a pack of 52 cards

$$\therefore n(S) = {}^{52}C_2$$

$${}^{n}C_{r} = \frac{n!}{(n-r)!r!}$$

$$n(S) = \frac{52!}{50!\ 2!}$$

$$n(S) = \frac{52\ x\ 51\ x\ 50!}{50!\ \ 2\ x1}$$

$$n(S) = 1326$$

Event A = Both cards are king

$$\therefore\ n(A) = {}^{4}C_2$$

$$\therefore\ n(A) = \frac{4!}{2!\,2!}$$

$$\therefore\ n(A) = 6$$

$$\therefore\ P(A) = \frac{n(A)}{n(S)}$$

$$\therefore\ P(A) = \frac{6}{1326}$$

$$\therefore\ P(A) = \frac{1}{221}$$

$$\therefore P(A) = 0.0045$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.7:

4 cards are drawn from a pack of 52 playing cards. Find the probability that

(i) All are spade cards
(ii) There is 1 card of each suit
(iii) 2 spade and 2 diamonds.

Solution:

Sample space (S) = 4 cards are drawn from a pack of 52 cards

$$\therefore n(S) = {}^{52}C_4$$

$${}^{n}C_r = \frac{n!}{(n-r)!r!}$$

$$n(S) = \frac{52!}{48!\ 4!}$$

$$n(S) = \frac{52\ x\ 51\ x\ 50\ x\ 49\ x\ 48!}{48!\ \ 4\ x\ 3\ x\ 2\ x1}$$

$$n(S) = 270725$$

(i) Event A = all 4 cards are spade cards

$$\therefore\ n(A) = {}^{13}C_4$$

$$\therefore\ n(A) = \frac{13!}{9!\,4!}$$

$$\therefore\ n(A) = 715$$

$$\therefore\ P(A) = \frac{n(A)}{n(S)}$$

$$\therefore\ P(A) = \frac{715}{270725}$$

$$\therefore P(A) = 0.002641$$

(ii) Event B= there is 1 card of each suit

$$\therefore\ n(B) = {}^{13}C_1 \text{ X } {}^{13}C_1 \text{ X } {}^{13}C_1 \text{ X } {}^{13}C_1$$

$$\therefore\ P(B) = \frac{n(B)}{n(S)}$$

$$\therefore\ P(B) = \frac{13\ X\ 13\ X\ 13\ X\ 13}{270725}$$

$$\therefore\ P(B) = 0.105498$$

(iii) Event C = 2 spade and 2 diamonds

$$\therefore\ n(C) = {}^{13}C_2 \text{ X } {}^{13}C_2$$

$$\therefore\ n(C) = \frac{13\,!}{11\,!\ 2!}\ X\ \frac{13\,!}{11\,!\ 2!}$$

$$\therefore\ n(C) = \frac{13\ x\ 12\ x\ 11\,!}{11\,!\ 2!}\ X\ \frac{13\ x\ 12\ x\ 11\,!}{11\,!\ 2!}$$

$$\therefore\ n(C) = 78\ X\ 78$$

$$\therefore\ P(C) = \frac{n(C)}{n(S)}$$

$$\therefore\ P(C) = \frac{78\ x\ 78}{270725}$$

$$\therefore\ P(C) = 0.022472$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.8:

A box contains 75 good ICs and 25 defective. If 12 ICs are selected at random, find the probability that at least 1 chip is defective?

Solution:

Probability that at least 1 chip is defective = 1 – probability that no chips are defective

$$\therefore n(S) = {}^{100}C_{12}$$

$$\boxed{{}^{n}C_{r} = \frac{n!}{(n-r)!r!}}$$

$$n(S) = \frac{100!}{80!\ 12!}$$

Event A = no chips are defective

$$\therefore\ n(A) = {}^{75}C_{12}$$

$$\therefore\ n(A) = \frac{75!}{63!\ \ 12!}$$

$$\therefore\ P(at\ least\ 1\ chip\ is\ defective) = 1 - P(A)$$

$$\therefore\ P(A) = \frac{n(A)}{n(S)}$$

$$\therefore\ P(A) = {}^{75}C_{12}\ /\ {}^{100}C_{12}$$

$$P(at\ least\ 1\ chip\ is\ defective) = 1 - [\ {}^{75}C_{12}\ /\ {}^{100}C_{12}\ ]$$

$$\therefore P(at\ least\ 1\ chip\ is\ defective) = 0.9751$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 3.9:

A bag contains 8 red and 5 white balls. Two successive drawings of 3 balls are made such that

Case I : balls are replaced before the second trial.

Case II: balls are not replaced before the second trial.

Find the probability that the first drawing will give 3 white balls and 2nd drawing will give 3 red balls for case I and case II

Solution:

Let A = three white balls are drawn

B= three red balls are drawn

Case I : balls are replaced before second trial

$$P(A \cap B) = P(A).\ P(B)$$

$$P(A) = {}^{5}C_{3} / {}^{13}C_{3}$$

$${}^{n}C_{r} = \frac{n!}{(n-r)!r!}$$

$$P(A) = 10/286$$

$$P(B) = {}^{8}C_{3} / {}^{13}C_{3}$$

$$P(B) = 56/286$$

$$P(A \cap B) = P(A).\ P(B)$$

$$P(A \cap B) = 10/286\ \ X\ \ 56/286$$
$$P(A \cap B) = 0.0068$$

Case II : balls are not replaced before the second trial

$$P(A \cap B) = P(A).\ P(B|A)$$

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

$$P(A) = {}^{5}C_{3} / {}^{13}C_{3}$$

$$P(A) = 10/286$$

$$P(B|A) = {}^{8}C_{3} / {}^{10}C_{3}$$

$$P(B|A) = 56 / 120$$

$$P(A \cap B) = P(A).\ P(B|A)$$
$$P(A \cap B) = 10/286\ \ X\ \ 56 / 120$$
$$P(A \cap B) = 0.016$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 3.10:

Among the digits 3, 4, 5, 6, 7 first digit is chosen and then second digit is chosen from the remaining four. Find the probability that an odd digit will be selected

(i) As first digit
(ii) As second digit
(iii) As both digits

Solution:

Sample space (S) = two digit number is selected where first digit from 5 digit and second from remaining 4 digits.

$$\therefore n(S)= {}^{5}C_{1} \text{ X } {}^{4}C_{1}$$

$$\therefore n(S)= 5 \text{ X } 4$$

$$\therefore n(S)=20$$

(i) Event A = an odd digit will be selected as first digit and second digit is selected from remaining digits

$$n(A) = {}^{3}C_{1} \text{ X } {}^{4}C_{1}$$

$$\therefore n(A)= 3 \text{ X } 4$$

$$\therefore n(A)= 12$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{12}{20}$$

$$P(A) = 0.6$$

(ii) Event B = an odd digit will be selected as second digit

= an even digit is selected as first digit and odd digit is selected as second digit OR

an odd digit is selected as first digit and odd digit is selected as second digit

$$n(B) = {}^{2}C_{1} \text{ X } {}^{3}C_{1} + {}^{3}C_{1} \text{ X } {}^{2}C_{1}$$

$$n(B) = (2\ x\ 3) + (3\ x\ 2)$$

$$n(B) = 12$$

$$P(B) = \frac{n(B)}{n(S)}$$

$P(B) = \frac{12}{20}$ i.e. $P(B) = 0.6$

(iii) Event C= odd digits are selected as both digits

$$n(C) = {}^{3}C_{1} \times {}^{2}C_{1}$$

$$n(C) = 3 \times 2$$

$$n(C) = 6$$

$$P(C) = \frac{n(C)}{n(S)}$$

$$P(C) = \frac{6}{20} \qquad \text{i.e.} \quad P(C) = 0.3$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.11:

Find the probability of constructing a two digit even number using the digits 2,3,4,5,6 if

(i) You can use the same digit again

(ii) You cannot use a digit more than once

Solution:

Sample space (S) = two digit number with 5 digits with repetition allowed

$$\therefore n(S) = {}^{5}C_{1} \text{ X } {}^{5}C_{1}$$

$$\therefore n(S) = 5 \text{ X } 5$$

$$\therefore n(S) = 25$$

(i) Event A = two digit even number

at units place = {2,4,6 } = 3 ways

at tens place = {2,3,4,5,6} = 5 ways

$$n(A) = {}^{5}C_{1} \text{ X } {}^{3}C_{1}$$

$$\therefore n(A) = 5 \text{ X } 3$$

$$\therefore n(A) = 15$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{15}{25}$$

$$P(A) = 0.6$$

(ii) Event B = repetition of digits is not allowed and required number is even number

$n(S) = {}^{5}C_1 \times {}^{4}C_1 = 20$

$n(B) = {}^{4}C_1 \times {}^{3}C_1$

$$n(B) = 4 X 3$$

$$n(B) = 12$$

$$P(B) = \frac{n(B)}{n(S)}$$

$P(B) = \frac{12}{20}$ i.e. $P(B) = 0.6$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.12:

Find the probability of constructing a two digit even number using the digits 1,2,3,4,5,6,7,8,9 if

(i) Repetition of digits is allowed

(ii) Repetition of digits is not allowed

Solution:

Sample space (S) = two digit number with 9 digits with repetition allowed

$$\therefore n(S) = {}^{9}C_{1} \text{ X } {}^{9}C_{1}$$

$$\therefore n(S) = 9 \text{X } 9$$

$$\therefore n(S) = 81$$

(i) Event A = repetition of digits is allowed

at units place = {2,4,6 ,8} = 4 ways

at tens place = {1,2,3,4,5,6,7,8,9} = 9 ways

$$\therefore n(A) = 9 \text{ X } 4$$

$$\therefore n(A) = 36$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = 36/81$$

$$P(A) = 0.4444$$

(ii) Event B = repetition of digits is not allowed and required number is even number

$n(S) = {}^9C_1 \ X \ {}^8C_1 = 72$

$n(B) = {}^8C_1 \ X \ {}^4C_1$

$$n(B) = 8\ X\ 4$$

$$n(B) = 32$$

$$P(B) = \frac{n(B)}{n(S)}$$

$P(B) = \frac{32}{72}$ i.e. $P(B) = 0.4444$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.13:

Find the probability of constructing a two digit even number using the digits 3,4,5,6, and 7

Assume first that you may use same digit again, then repeat the question, assuming that you may not use a digit more than once

Solution:

Sample space (S) = two digit number with 5 digits with repetition allowed

$$\therefore n(S)= {}^{5}C_{1} \text{ X } {}^{5}C_{1}$$

$$\therefore n(S)= 5\text{X } 5$$

$$\therefore n(S)=25$$

(i) Event A = two digit even number , repetition of digits is allowed

at units place = {4,6} = 2 ways

at tens place = {3,4,5,6,7} = 5 ways

$$\therefore n(A)= 5 \text{ X } 2$$

$$\therefore n(A)= 10$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = 10/25$$

$$P(A) = 0.4$$

Event B = two digit even number, repetition of digits is not allowed and required number is even number

$n(S) = {}^{5}C_{1} \times {}^{4}C_{1} = 20$

$n(B) = {}^{4}C_{1} \times {}^{2}C_{1}$

$$n(B) = 4 X 2$$

$$n(B) = 8$$

$$P(B) = \frac{n(B)}{n(S)}$$

$P(B) = \frac{8}{20}$ i.e. $P(B) = 0.4$

◈◈◈◈◈◈◈

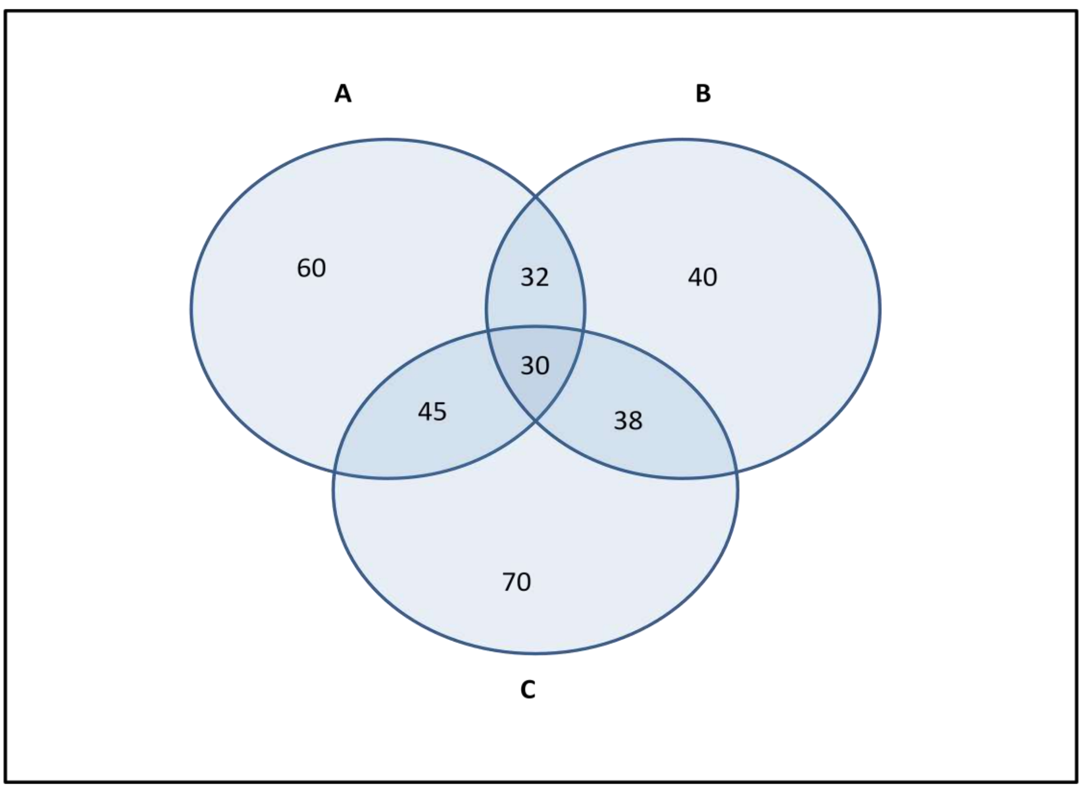
Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 3.14:

A sample survey was taken to check which newspaper (A,B, C) people read. In a sample of 100 people the following results are obtained, 60 read A, 40 read B, 70 read C, 45 read A and C, 32 read A and B, 38 read B and C, 30 read A, B and C. If a person is selected at random, find the probability that

(a) He reads only A newspaper.
(b) He reads at least two newspapers
(c) He reads at most 1 newspaper
(d) He doesn't read any paper

Solution:



$$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(B \cap C) - n(A \cap C) + n(A \cap B \cap C)$$

$$n(A \cup B \cup C) = 60 + 40 + 70 - 32 - 38 - 45 + 30$$

$$n(A \cup B \cup C) = 85$$

(i) Event A = he reads only A newspaper

$\therefore$ n(A)= 60 – n(A $\cap$B) – n(A $\cap$C) + n(A$\cap$B$\cap$C)

$\therefore$ n(A)= 60-32-45+30

n(A) = 13

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = 13/100$$

$$P(A) = 0.13$$

(ii) Event B = he reads at least two newspapers (means he may read 2 or more newspapers) n(B) = (45-30) + (32-30) + (38-30) + 30

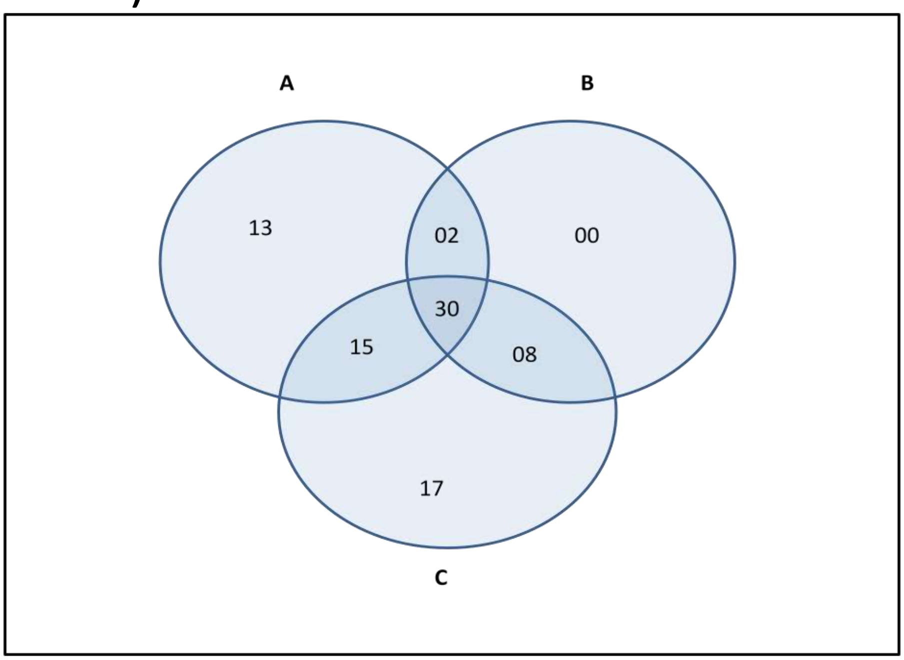
$$n(B) = 55$$

$$P(B) = \frac{n(B)}{n(S)}$$

$P(B) = \frac{55}{100}$ i.e. $P(B) = 0.55$

(iii) Event C = he reads at most 1 newspaper (means 1 or less than 1 newspaper)

n(C) = (60-45-32+30) + (40-32 -38+ 30) + (70-45-38+30) + 15



n(C)= 13+0+17+15

n(C)=45

$$P(C) = \frac{n(C)}{n(S)}$$

$P(C) = \frac{45}{100}$ i.e. $P(C) = 0.45$

(iv) Event D = he doesn't read any newspaper

$$P(D) = \frac{n(D)}{n(S)}$$

$$P(D) = \frac{15}{100}$$

$$P(D) = 0.15$$

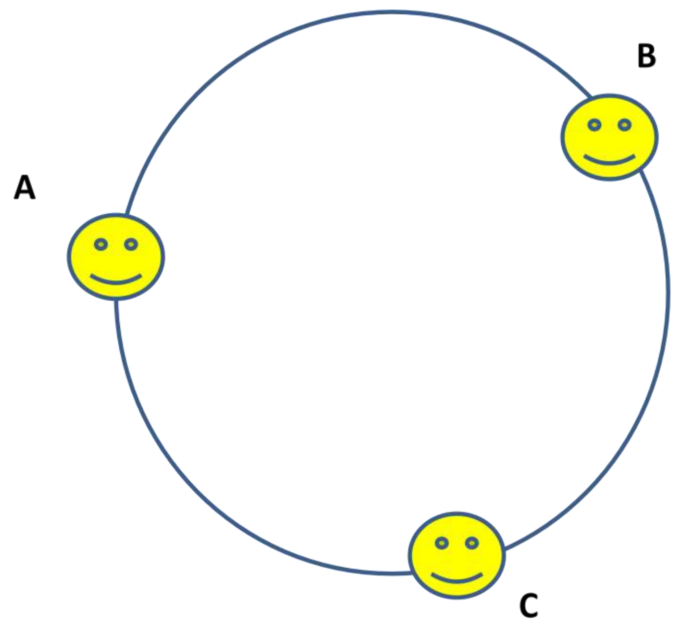
Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.15:

'n' persons are seated on 'n' chairs on a round table. Find the probability that two specific persons are sitting next to each other.

Solution:



For a round table n persons and n chairs arrangements

Sample space S= total arrangements of sitting

$\therefore n(S) = (n-1)!$

Event A = two specific persons are sitting next to each other in arranegment

$\therefore$ n(A)= (n-2)! X 2!

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{(n-2)!\,2!}{(n-1)!}$$

$$P(A) = \frac{(n-2)!\,2!}{(n-1)(n-2)!}$$

$$P(A) = \frac{2}{(n-1)}$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 3.16:
What is the probability that 4 S's come consecutively in the arrangement of the letters in the word 'MISSISSIPPI'

Solution:
Sample space S= total arrangements of letters in word 'MISSISSIPI'

S= arrangement of 4 letters S, 4 letters I, 2 letters P, 1 letter M

$$n(S) = \frac{11!}{4!\,4!\,2!\,1!}$$

Event A = 4 S's come consecutively
= arrangement of (4S) in 1 group , 4 letters I, 2 letters P, 1 letter M

$$n(A) = \frac{8!}{4!\,2!\,1!\,1!}$$

$$P(A) = \frac{n(A)}{n(S)}$$

$P(A) = \frac{4}{165}$ $\therefore P(A) = 0.02424$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.17:

What is the probability that 4 A's come consecutively in the arrangement of the letters in the word 'MAHARASHTRA'

Solution:

Sample space S= total arrangements of letters in word 'MAHARASHTRA'

S= arrangement of

M- 1

A-4

H-2

R-2

S-1

T-1

$$n(S) = \frac{11!}{4!\,2!\,2!\,1!\,1!\,1!}$$

Event A = 4 A's come consecutively

= arrangement of (4A) in 1 group , 2 letters H, 2 letters R, 1 letter M, 1 letter S, 1 letter T

$$n(A) = \frac{8!}{2!\,2!\,1!\,1!\,1!\,1!}$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{4}{165} \qquad \therefore P(A) = 0.02424$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 3.18:
What is the probability that all vowels come consecutively in the arrangement of the letters in the word 'PROBABILITY'

Solution:
Sample space S= total arrangements of letters in word 'PROBABILITY'
S= arrangement of
P- 1
R-1
O-1
B-2
A-1
L-1
I-2
T-1
Y-1

$$n(S) = \frac{11!}{2!\,2!}$$

Event A = all vowels (O,A, I, I) come consecutively
P, R, [O, A, I, I], B , B, L, T, Y

$$n(A) = \frac{8!}{2!\,1!\,1!\,1!\,1!\,1!\,1!}\ X\ \frac{4!}{2!\,1!\,1!}$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{4}{165} \qquad \therefore P(A) = 0.02424$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.19:

What is the probability that 4 A's come consecutively in the arrangement of the letters in the word 'MAHANAGAR'

Solution:

Sample space S= total arrangements of letters in word 'MAHANAGAR'

S= arrangement of

M- 1

A-4

H-1

N-1

G-1

R-1

$$n(S) = \frac{9!}{4!\,1!\,1!\,1!\,1!\,1!}$$

Event A = 4A's come consecutively

M, [A, A, A, A,] N, H, G, R

$$n(A) = \frac{6!}{1!}$$

$$P(A) = \frac{n(A)}{n(S)}$$

$P(A) = \frac{1}{21}$ $\therefore P(A) = 0.0476$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.20:

What is the probability that all vowels come consecutively in the arrangement of the letters in the word 'COMMERCE'

Solution:

Sample space S= total arrangements of letters in word 'COMMERCE'

S= arrangement of

C- 2

O-1

M-2

E-2

R-1

$$n(S) = \frac{8!}{2!\,2!\,2!\,1!\,1!}$$

Event A = all vowels come consecutively

C, [O, E,E,] M, M, R,C

$$n(A) = \frac{6!}{2!\,2!\,1!\,1!}\ X\,\frac{3!}{2!\,1!}$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{\frac{6!\,3!}{2!\,2!\,2!}}{\frac{8!}{2!\,2!\,2!}}$$

$$P(A) = \frac{6!\,3!}{8x7x6!}$$

$$P(A) = \frac{3}{28}$$

$$\therefore P(A) = 0.1071$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 3.21:
Letters of the word 'FAILURE' are arranged at random. Find the probability that the consonants may occupy only odd position.

Solution:
Sample space S= total arrangements of letters in word 'FAILURE'
S= arrangement of
F- 1
A-1
I-1
L-1
U-1
R-1
E-1

$$n(S) = \frac{7!}{1!\,1!\,1!\,1!\,1!\,1!\,1!}$$

$$n(S) = 7!$$

Event A = all consonants may occupy only odd positions (there are odd positions 1^{st}, 3^{rd}, 5^{th}, 7^{th} and 3 consonants viz. F, L, R)

$$n(A) = {}^{4}P_{3} \text{ X } 4!$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = [{}^{4}P_{3} \times 4!] / 7!$$

$$P(A) = \frac{4}{35}$$

$$\therefore P(A) = 0.1143$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.22:

What is the probability that all vowels come consecutively in the arrangement of the letters in the word 'MATHEMATICS'

Solution:

Sample space S= total arrangements of letters in word 'MATHEMATICS'

S= arrangement of M= 2 , A= 2 , T= 2 ,H=1 , I =1 , E=1 ,C= 1 , S=1

$$n(S) = \frac{11!}{2!\,2!\,2!\,1!\,1!\,1!\,1!\,1!}$$

Event A = all vowels (A,E,I) come consecutively

M, T, M,H, T, C, S, [A, A, E, I]

$$n(A) = \frac{8!}{2!\,2!\,1!\,1!\,1!\,1!}\ X\ \frac{4!}{2!\,1!\,1!}$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{\frac{8!\,4!}{2!\,2!\,2!}}{\frac{11!}{2!\,2!\,2!}}$$

$$P(A) = \frac{8!\,4\ x\ 3\ x\ 2\ x\ 1}{11\ x\ 10\ x\ 9\ x\ 8!}$$

$P(A) = \frac{4}{165}$ $\therefore P(A) = 0.02424$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 3.23:
What is the probability that all vowels come consecutively in the arrangement of the letters in the word 'AHMEDNAGAR'

Solution:
Sample space S= total arrangements of letters in word 'AHMEDNAGAR'
S= arrangement of A= 3 , H= 1 , M= 1 ,E=1 , D =1 , N=1 ,G= 1 R=1

$$n(S) = \frac{10!}{3!}$$

Event A = all vowels (A,A,A,E) come consecutively
H,M,D,N,G,R ,[A, A, A, E]

$$n(A) = 7!\ X\ \frac{4!}{3!\,1!}$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{\frac{7!\,4!}{3!}}{\frac{10!}{3!}}$$

$$P(A) = \frac{7!\,4\ x\ 3\ x\ 2\ x\ 1}{10\ x\ 9\ x\ 8\ x\ 7!}$$

$P(A) = \frac{1}{30}$ $\therefore P(A) = 0.033$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 3.24:
Ten balls are distributed at random among 4 boxes. What is the probability that first box contains 4 balls.

Solution:
Sample space S= 10 balls are distributed at random among 4 boxes
Every ball have 4 choices. hence 10 balls have

$$n(S) = 4^{10}$$

Event A = first box will contain 4 balls
For the first box 4 balls can be selected from 10 balls and from remaining boxes remaining balls can be distributed in 3^6 ways

$$n(A) = {}^{10}C_4 \text{ X } 3^6$$

$$P(A) = \frac{n(A)}{n(S)}$$

P(A) =$({}^{10}C_4 \text{ X } 3^6) / 4^{10}$

$$\therefore P(A) = 0.146$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 3.25:

Find the probability that randomly chosen 3-letter sequence will not have any repeated letters.

Solution:

Sample space S= 3-letters are randomly chosen among 26 alphabets/letters

$$n(S) = 26 \ x \ 26 \ x \ 26$$

Event A = 3 letter sequence is chosen without any repetition letters.

$$n(A) = 26 \ x \ 25 \ x \ 24$$

$$P(A) = \frac{n(A)}{n(S)}$$

$$P(A) = \frac{26 \ x \ 25 \ x \ 24}{26 \ x \ 26 \ x \ 26}$$

$$\therefore P(A) = 0.8876$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 3.26:
In a party of five persons, compute the probability that at least two have the same birthday (month/day), assume a 365 day a year.

Solution:
Sample space S= A day is a birthday of 5 persons

$$n(S) = 365^5$$

$$n(S) = 365\ x\ 365\ x\ 365\ x\ 365\ x\ 365$$

Event A = at least two have same birthday.

$$\bar{A} = None\ have\ same\ birthday$$

$$n(\bar{A}) = 365\ x\ 364\ x\ 363\ x\ 362\ x\ 361$$

$$P(\bar{A}) = \frac{n(\bar{A})}{n(S)}$$

$$P(A) = \frac{365\ x\ 364\ x\ 363\ x\ 362\ x\ 361}{365\ x\ 365\ x\ 365\ x\ 365\ x\ 365}$$

$$\therefore P(\bar{A}) = 0.9728$$

$$P(A) = 1 - P(\bar{A})$$

$$P(A) = 1 - 0.9728$$

$$P(A) = 0.0272$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.1:

Consider an experiment "three coins are tossed".
Let the random variable X= 'number of heads'

a) Find the values of X
b) Find the probability of X
c) Find the probability mass function
d) Find the cumulative distribution function

Solution:

The sample space is

S = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}

Event	HHH	HHT	HTH	HTT	THH	THT	TTH	TTT
X	3	2	2	1	2	1	1	0

X	P(X)	pmf P(X)	cdf F(X)
0	1/8	1/8	1/8
1	3/8	3/8	4/8
2	3/8	3/8	7/8
3	1/8	1/8	8/8 i.e. 1

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.2:

A box of 6 ICs contains 2 defective. A computer center makes a random purchase of 3 of the ICs. If X is the number of defective chips purchased by the computer center, find the probability distribution of X.

Solution:

There are only 2 defective ICs in a box

Therefore X can take values 0,1,2

$$P(x=r) = P(\text{choosing exactly r defective chips})$$
$$= P(\text{choosing r defective and (3-r) good chips})$$
$$= ({}^{2}C_{r} \,.\, {}^{4}C_{3-r})/{}^{6}C_{3} \qquad r=0,1,2$$

The probability distribution is represented in the following table

X=r	P(X=r)
0	1/5
1	3/5
2	1/5
Total=	1

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.3:

Two dice are rolled. Let X denote the random variable which counts the total number of points on the upturned faces. Construct a table giving the non-zero values of probability mass function?

Solution:

X=r	P(X=r)
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.4:

A random variable X takes the values 1, 2, 3, and 4 such that $4P(X=1)=2P(X=2)=3P(X=3)=P(X=4)$

Find the probability and cumulative distribution function of X

Solution:

Let $4P(X=1)=2P(X=2)=3P(X=3)=P(X=4)=k$

Therefore

$P(X=1) = k/4$

$P(X=2) = k/2$

$P(X=3) = k/3$

$P(X=4) = k$

Since $\Sigma P(x_i) = 1$

We get $k/4 + k/2 + k/3 + k = 1$

$25k/12 = 1$

$k=12/25$

x	PMF P(X)	CDF F(X)
1	3/25	3/25
2	6/25	9/25
3	4/25	13/25
4	12/25	25/25 = 1

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.5:

The probability distribution P(X) and cumulative distribution F(X) are given in the following table

X	X=1	X=2	X=3	X=4	X=5
Pmf	1/15	2/15	3/15	4/15	5/15
Cdf	1/15	3/15	6/15	10/15	15/15 =1

$$\left.\begin{aligned} P(X) &= \frac{x}{15} \qquad x = 1,2,3,4,5 \\ P(X) &= 0 \qquad otherwise \end{aligned}\right\}$$

Find (i) $P\left\{\frac{1}{2} < x < \frac{5}{2} \middle| x > 1\right\}$

(ii) $P\{\, x = 1 \; or \; 2\}$

Solution:

(i)

$$P\left\{\frac{1}{2} < x < \frac{5}{2} \middle| x > 1\right\} = \frac{P\{\frac{1}{2} < x < \frac{5}{2} \quad \cap \quad x > 1\,\}}{P(x > 1)}$$

$$P\left\{\frac{1}{2} < x < \frac{5}{2} \middle| x > 1\right\} = \frac{P\{\, x = 2 \,\}}{P(x > 1)}$$

$$P\left\{\frac{1}{2} < x < \frac{5}{2} \middle| x > 1\right\} = \frac{\frac{2}{15}}{\sum_{x=2}^{5} \frac{x}{15}}$$

$$P\left\{\frac{1}{2} < x < \frac{5}{2} \middle| x > 1\right\} = \frac{\frac{2}{15}}{\frac{14}{15}}$$

$$P\left\{\frac{1}{2} < x < \frac{5}{2} \middle| x > 1\right\} = \frac{1}{7}$$

(ii) $P\{\, x = 1 \; or \; 2\} = P(x = 1) + P(x = 2)$

$$P\{\, x = 1 \; or \; 2\} = \frac{1}{15} + \frac{2}{15} = \frac{3}{15}$$

$$P\{\, x = 1 \; or \; 2\} = \frac{1}{5}$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.6:

A random variable X has the following probability distribution function

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

(i) Find k

(ii) Evaluate P(x < 6), P(x >= 6), and P(0 < x < 5)

(iii) If $P(x \leq C) > ½$, find minimum value of C, and

(iv) Determine the distribution function CDF of x

Solution:

X	0	1	2	3	4	5	6	7
P(X)	0	k	2k	2k	3k	k^2	$2k^2$	$7k^2 + k$
F(X)	0	k	3k	5k	8k	$8k+ k^2$	$8k+ 3k^2$	$9k+ 10k^2$

(i) Since $\sum_{i=0}^{7} P(x) = 1$ we have

$k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$

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

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

$10k(k+1) - 1(k+1) = 0$

$(10k - 1)(k+1) = 0$

K = -1 or k = 1/10
But k is probability so it cannot be negative hence k = 1/10

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

$P(X < 6) = 0 + 1/10 + 2/10 + 2/10 + 3/10 + 1/100$
$P(x < 6) = 81/100$

$\therefore P(X \geq 6) = 1 - P(X < 6)$
$\therefore P(X \geq 6) = 1 - 81/100$

$\therefore P(X \geq 6) = 19/100$

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

$P(0 < x < 5) = 8k$

$P(0 < x < 5) = 8/10$

$P(0 < x < 5) = 4/5$

(iii) To find minimum value of C such that
$P(x \leq C) > ½$

Consider $P(x \leq 0) = P(X=0) = 0$

$P(x \leq 1) = P(X=0) + P(X=1)$
$= 0 + k$

$P(X \leq 1) = k = 1/10$

$P(x \leq 2) = P(X=0) + P(X=1) + P(X=2)$
$P(x \leq 2) = 0 + k + 2k$
$P(x \leq 2) = 3k = 3/10$

$P(x \leq 3) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$
$P(x \leq 3) = 0 + k + 2k + 2k$
$P(x \leq 3) = 5k = 5/10$

$P(x \leq 4) = P(X=0) + P(X=1) + P(X=2) + P(X=3) + P(X=4)$
$P(x \leq 4) = 0 + k + 2k + 2k + 3k$
$P(x \leq 4) = 8k = 8/10$
Hence minimum value of C such that $P(x \leq C) > ½$ is equal to 4.

(iv) The cumulative distribution F(X) is given below

X	$F_x(X) = P(X \leq x)$
0	0
1	$k = 1/10$
2	$3k = 3/10$
3	$5k = 5/10$

4	$8k = 8/10$
5	$8k + k^2 = 81/100$
6	$8k + 3k^2 = 83/100$
7	$9k + 10k^2 = 1$

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.7:

The probability mass function of a random variable X is zero except at points x=0, 1, 2. At these points it has the values $P(0) = 3C^2$, $P(1) = 4C - 10C^2$ and $P(2)= 5C - 1$, for some $C > 0$

(i) Determine the value of C
(ii) Compute the following probabilities $P[x < 2]$ and $P[1 < x \leq 2]$
(iii) Find the largest x such that $F(x) < ½$
(iv) Find the smallest x such that $F(x) \geq 1/3$

Solution:

(i) Since $\sum P(X) = 1$ we have

$P(0) + P(1) + P(2)= 1$
$3C^2 + 4C - 10C^2 + 5C - 1 = 1$
$-7C^2 + 9C - 2 = 0$
$7C^2 - 9C + 2 = 0$
$7C^2 - 7C - 2C + 2 = 0$
$7C(C-1) - 2(C-1) = 0$
$(C-1)(7C-2) = 0$
$C= 1$ or $C= 2/7$

But for C=1 gives P(0) = 3 which is impossible hence C= 2/7

(ii) $P(X < 2) = P(X=0) + P(X=1)$
$P(X < 2) = 3C^2 + 4C - 10C^2$

$P(X < 2) = -7C^2 + 4C$

Substituting C= 2/7 we get

$P(X < 2) = 4/7$

$P(1 < x \le 2) = P(x=2)$

$P(1 < x \le 2) = 5C - 1 = 3/7$

(iii) To find largest X s.t. F(X) < ½

Note that P(0) =12/49

P(1) = 16/49

P(2) = 3/7

$$F(X) = \sum_{x=0}^{2} P(X)$$

$$F(0) = \frac{12}{49} < \frac{1}{2} \quad \therefore x = 0$$

(iv) To find the smallest x s.t. $F(X) \ge 1/3$

F(1) = 12/49 + 16/49

$F(1) = 28/49 \ge 1/3$

$\therefore$ x= 1

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.8:

A continuous random variable has pdf

$$f(x) = \begin{cases} k(2-x) & 0 \le x < 2 \\ kx(x-2) & 2 \le x < 3 \\ 0 & otherwise \end{cases}$$

Find k and median of the distribution

Solution:

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

$$\int_0^2 k(2-x)dx + \int_2^3 kx(x-2)dx = 1$$

$$\int_0^2 (k2 - kx)dx + \int_2^3 (kx^2 - 2kx)\, dx = 1$$

$$\left[k2x - k\frac{x^2}{2}\right]_0^2 + \left[k\frac{x^3}{3} - 2k\,\frac{x^2}{2}\right]_2^3 = 1$$

$$[4k - 2k] + \left[9k - 9k - \frac{8}{3}k + 4k\right] = 1$$

$$2k + 4k - \frac{8}{3}k = 1$$

$$6k - \frac{8}{3}k = 1$$

$$\frac{10}{3}k = 1$$

$$k = \frac{3}{10}$$

Now F(0) = 0 and F(3) = 1

$$F(2) = \int_0^2 k(2-x)dx = \int_0^2 \frac{3}{10}(2-x)dx$$

$$F(2) = \frac{3}{10}\left[2x - \frac{x^2}{2}\right]_0^2$$

$$F(2) = \frac{3}{10}\,[4-2]$$

$$F(2) = \frac{6}{10}$$

$$F(2) = \frac{3}{5} \quad > \frac{1}{2}$$

$$\therefore Median\ lies\ between\ x = 0\ and\ x = 2$$
$$Let\ median\ be\ m$$

$$\therefore \int_0^m f(x)dx = \frac{1}{2}$$

$$\therefore \int_0^m k(2-x)dx = \frac{1}{2}$$

$$\therefore \int_0^m \frac{3}{10}(2-x)dx = \frac{1}{2}$$

$$\frac{3}{10}\left[2x - \frac{x^2}{2}\right]_0^m = \frac{1}{2}$$

$$\therefore \frac{3}{10}\left[2m - \frac{m^2}{2}\right] = \frac{1}{2}$$

$$\frac{6m}{10} - \frac{3m^2}{20} = \frac{1}{2}$$

$$\frac{12m - 3m^2}{20} = \frac{1}{2}$$

$$3m^2 - 12m + 10 = 0$$

a=3, b=-12, c=10

$$roots = \frac{-b \mp \sqrt{b^2 - 4ac}}{2a}$$

$$\therefore m = \frac{12 \mp \sqrt{144 - 4x3x10}}{6}$$

$$\therefore m = \frac{12 \mp \sqrt{144 - 120}}{6}$$

$$\therefore m = \frac{12 \mp \sqrt{24}}{6}$$

$$\therefore m = \frac{12 \mp 2\sqrt{6}}{6}$$

$$\therefore m = \frac{2(6 \mp \sqrt{6})}{6}$$

$$\therefore m = \frac{6 \mp \sqrt{6}}{3}$$

$$\therefore Median = \frac{6 - \sqrt{6}}{3} \quad (median\ lies\ between\ 0\ and\ 2)$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.9:

Suppose that the error in the reaction temperature in ^{0}C, for a controlled experiment is a continuous random variable X having the probability function.

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

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

(ii) Find P(0 < x ≤ 1)

Solution:

(i)

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-1}^{2} \frac{x^2}{3}\, dx$$

$$\int_{-\infty}^{\infty} f(x)dx = \left[\frac{x^3}{3*3}\right]_{-1}^{2}$$

$$\int_{-\infty}^{\infty} f(x)dx = \left[\frac{x^3}{9}\right]_{-1}^{2}$$

$$\int_{-\infty}^{\infty} f(x)dx = \left[\frac{8}{9} - \frac{-1}{9}\right]$$

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

(ii)

$$P(0 < x \le 1) = \int_{0}^{1} \frac{x^2}{3} dx$$

$$P(0 < x \le 1) = \left[\frac{x^3}{3 * 3}\right]_0^1$$

$$P(0 < x \le 1) = \left[\frac{x^3}{9}\right]_0^1$$

$$P(0 < x \le 1) = \frac{1}{9}$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.10:

The life in hours (x) of a certain electronic component is a continuous random variable with pdf

$$f(x) = \begin{cases} \frac{150}{x^2} & x \geq 150 \\ 0 & otherwise \end{cases}$$

Determine the form of the distribution function of F(x). Also answer the following questions

(i) What is the probability that a component would last for (a) at most 250 hours (b) at least 200 hours (c) more than 200 hours but less than 500 hours.

(ii) If a certain device involves five such components in parallel, find the probability that the device would be functioning at the end of 500 hours assuming that five components work independently.

Solution:

(i)

$$F(x) = \int_{150}^{x} f(x)dx = \int_{150}^{x} \frac{150}{x^2}\, dx$$

$$F(x) = \int_{150}^{x} 150\ (x)^{-2}\, dx$$

$$F(x) = \left[150\ \frac{x^{-1}}{-1}\right]_{150}^{x}$$

$$F(x) = \left[-\frac{150}{x}\right]_{150}^{x}$$

$$F(x) = -\frac{150}{x} + \frac{150}{150}$$

$$F(x) = 1 - \frac{150}{x}$$

$$\therefore\ F(x) = P(X \le x) = 1 - \frac{150}{x} \qquad x \ge 150$$

a) $P(X \le 250) = 1 - \frac{150}{250} = \frac{2}{5}$

b) $P(X \ge 200) = 1 - P(X \le 200) = 1 - \left(1 - \frac{150}{200}\right)$

$$P(X \ge 200) = 1 - \left(1 - \frac{3}{4}\right) = \frac{3}{4}$$

c) $P(200 \le X \le 500) = F(500) - F(200)$

$$P(200 \le X \le 500) = \left(1 - \frac{150}{500}\right) - \left(1 - \frac{150}{200}\right)$$

$$P(200 \le X \le 500) = \left(1 - \frac{3}{10}\right) - \left(1 - \frac{3}{4}\right)$$

$$P(200 \le X \le 500) = 1 - \frac{3}{10} - 1 + \frac{3}{4}$$

$$P(200 \le X \le 500) = \frac{3}{4} - \frac{3}{10}$$

$$P(200 \le X \le 500) = \frac{15 - 6}{20}$$

$$P(200 \le X \le 500) = \frac{9}{20}$$

(ii) $P(device\ is\ functioning) = 1 - P(device\ is\ not\ functioning)$

$$P(device\ is\ functioning) = 1 - P(all\ 5\ components\ are\ not\ functioning\ at\ the\ end\ of\ 500\ hours)$$

$$P(device\ is\ functioning) = 1 - [P(X \le 500)]^5$$

$$P(device\ is\ functioning) = 1 - [F(500)]^5$$

$$P(device\ is\ functioning) = 1 - (\frac{7}{10})^5$$

$$P(device\ is\ functioning) = 1 - 0.1681$$

$$P(device\ is\ functioning) = 0.8319$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.11:

Suppose that life in hours of a certain type of radio tube is a continuous random variable X with p.d.f given by

$$f(x) = \begin{cases} \frac{100}{x^2} & x \geq 100 \\ 0 & otherwise \end{cases}$$

(i) What is the probability that all of the original three tubes in a given set will have to be replaced in the first 150 hours of operations?

(ii) What is the probability that a tube will last less than 200 hours if it is known that the tube is still functioning after 150 hours of service?

Solution:

First find the CDF , F(X)

(i)

$$F(x) = \int_{100}^{x} f(x)dx = \int_{100}^{x} \frac{100}{x^2}\,dx$$

$$F(x) = \left[\frac{100}{-x}\right]_{100}^{x}$$

$$F(x) = 1 - \frac{100}{x}$$

$$\therefore\ F(x) = P(X \le x) = 1 - \frac{100}{x} \qquad x \ge 100$$

a) $[P(X \le 150)]^3 = [1 - \frac{100}{150}]^3 = [1 - \frac{2}{3}]^3 = \frac{1}{27}$

b) $P[X < 200 \mid X > 150] = \frac{P[150 < X < 200]}{P[X > 150]}$

$$P[X < 200 \mid X > 150] = \frac{F(X = 200) - F(X = 150)}{1 - F(X = 150)}$$

$$P[X < 200 \mid X > 150] = \frac{\left(1 - \frac{100}{200}\right) - \left(1 - \frac{100}{150}\right)}{1 - \left(1 - \frac{100}{150}\right)}$$

$$P[X < 200 \mid X > 150] = \frac{\frac{1}{2} - \frac{1}{3}}{1 - \frac{1}{3}}$$

$$P[X < 200 \mid X > 150] = \frac{\frac{1}{6}}{\frac{2}{3}}$$

$$P[X < 200 \mid X > 150] = \frac{1}{4}$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.12:

The diameter of an electric cable say X, is assumed to be a continuous random variable with pdf

f(x) = 6x(1-x) $0 \le x \le 1$

(i) Check that above is p.d.f

(ii) Determine a number b s.t. P(x<b) = P(x >b)

Solution:

(i) For $0 \le x \le 1$ $f(x) \ge 0$

$$F(x) = \int_{-\infty}^{\infty} f(x)dx \quad i.e. \quad \int_0^1 f(x)\ dx = 1$$

Hence

$$\int_0^1 6x(1-x)dx = \int_0^1 (6x - 6x^2)dx$$

$$\int_0^1 6x(1-x)dx = \left[\frac{6x^2}{2} - \frac{6x^3}{3}\right]_0^1$$

$$\int_0^1 6x(1-x)dx = [3x^2 - 2x^3]_0^1$$

$$\int_0^1 6x(1-x)dx \quad = 3-2$$

$$\int_0^1 6x(1-x)dx \quad = 1$$

$\therefore$ f(x) is p.d.f. of random variable X.

(ii) $P(X < b) = P(X > b)$

0 to 1 interval is split at b

$$\int_0^b f(x)dx \quad = \int_b^1 f(x)dx$$

$$\int_0^b 6x(1-x)dx \quad = \int_b^1 6x(1-x)dx$$

$$[3x^2 - 2x^3]_0^b = [3x^2 - 2x^3]_b^1$$

$$3b^2 - 2b^3 = 3 - 2 - 3b^2 + 2b^3$$

$$3b^2 - 2b^3 - 1 + 3b^2 - 2b^3 = 0$$

$$6b^2 - 4b^3 - 1 = 0$$

$$-4b^3 + 6b^2 - 1 = 0$$

$$4b^3 - 6b^2 + 1 = 0$$

$$2b(2b^2 - 3) + 1 = 0$$

$$2b(2b^2 - 3b + 1) + (1 - 2b) = 0$$

$$2b[(2b^2 - 2b - b + 1)] + (1 - 2b) = 0$$

$$2b[2b(b - 1) - 1(b - 1)] + (1 - 2b) = 0$$

$$2b[(b - 1)(2b - 1)] + (1 - 2b) = 0$$

$$2b(b - 1)(2b - 1) - (2b - 1) = 0$$

$$(2b - 1)[2b(b - 1) - 1] = 0$$

$$(2b - 1)[2b^2 - 2b - 1] = 0$$

$$(2b - 1)[2b^2 - 2b - 1] = 0$$

$\therefore$ 2b -1 = 0

Hence b = ½   real number between (0, 1)

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.13:

A continuous random variable x has p.d.f.

$f(x) = 3x^2 \quad 0 \leq x \leq 1$

Find a and b such that

(i) $P(x \leq a) = P(x > a)$

(ii) $P(x > b) = 0.05$

Solution:

(i) Since $P(x \leq a) = P(x > a)$ each must be ½ as total probability is 1.

$$P[x \leq \ a\ ] = \frac{1}{2}$$

$$\int_0^1 f(x)\ dx = \frac{1}{2}$$

Hence

$$\int_0^1 3x^2\ dx \quad = \left[\frac{3\ x^3}{3}\right]_0^a$$

$$\int_0^1 3x^2\ dx \quad = \left[\frac{3\ x^3}{3}\right]_0^a$$

$$\int_0^1 3x^2\,dx \quad = [x^3]_0^a$$

$$\int_0^1 3x^2\,dx \quad = \quad a^3$$

$$\therefore \quad a^3 \quad = \frac{1}{2}$$

$$\therefore \quad a = \frac{1}{2}^{1/3}$$

$$\therefore \quad a = 0.7937$$

(ii) $\qquad P[x > b\,] = \; 0.05$

$$\int_b^1 f(x)\; dx = 0.05$$

$$\therefore \; [x^3]_b^1 \; = 0.05$$

$$\therefore \; 1 - b^3 \; = 0.05$$

$$\therefore \; b^3 \; = 0.95$$

$$b = 0.9830$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.14:

A continuous random variable x has p.d.f.

$$f(x) = \begin{cases} kx & 0 \le x \le 1 \\ k & 1 \le x \le 2 \\ -kx + 3k & 2 \le x \le 3 \\ 0 & otherwise \end{cases}$$

(i) Determine the constant k

(ii) Compute $P(X \le 1.5)$

(iii) Determine cumulative density function, F(x)

Solution:

(i) Constant 'k' is determined from consideration that total probability is unity.

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

$$\therefore \int_0^1 f(x)\ dx + \int_1^2 f(x)\ dx + \int_2^3 f(x)\ dx = 1$$

$$\therefore \int_0^1 kx\ dx + \int_1^2 k\ dx + \int_2^3 (-kx + 3k)\ dx = 1$$

$$\left[k\frac{x^2}{2}\right]_0^1 + [kx]_1^2 + \left[-k\frac{x^2}{2} + 3kx\right]_2^3 = 1$$

$$\frac{k}{2} + (2k - k) + \left[-\frac{9k}{2} + 9k + 2k - 6k\right] = 1$$

$$\frac{k}{2} + k - \frac{9k}{2} + 5k = 1$$

$$\frac{k}{2} + \frac{2k}{2} - \frac{9k}{2} + \frac{10k}{2} = 1$$

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

$$\therefore k = \frac{1}{2}$$

(ii) $P[x \leq 1.5] = \int_0^{1.5} f(x)\, dx$

$$P[x \leq 1.5] = \int_0^{1} f(x)\, dx + \int_1^{1.5} f(x)\, dx$$

$$P[x \leq 1.5] = \int_0^{1} kx\, dx + \int_1^{1.5} k\, dx$$

$$P[x \leq 1.5\,] = \left[k\,\frac{x^2}{2}\right]_0^1 + [kx]_1^{1.5}$$

$$P[x \leq 1.5\,] = \frac{k}{2} + 1.5k - k$$

$$P[x \leq 1.5\,] = \frac{k}{2} + \frac{3}{2}k - k$$

$$P[x \leq 1.5\,] = \frac{k}{2} + \frac{3k}{2} - \frac{2k}{2}$$

$$P[x \leq 1.5\,] = k$$

$$P[x \leq 1.5\,] = \frac{1}{2} \qquad since\ k = \frac{1}{2}$$

(iii) Determine CDF
Substituting k

$$f(x) = \begin{cases} \frac{x}{2} & 0 \leq x \leq 1 \\ \frac{1}{2} & 1 \leq x \leq 2 \\ -\frac{x}{2} + \frac{3}{2} & 2 \leq x \leq 3 \\ 0 & otherwise \end{cases}$$

$$\therefore \int_0^1 f(x)\ dx + \int_1^2 f(x)\ dx + \int_2^3 f(x)\ dx = 1$$

$$\therefore \int_0^1 \frac{x}{2}\ dx + \int_1^2 \frac{1}{2}\ dx + \int_2^3 (-\frac{x}{2} + \frac{3}{2})\ dx = 1$$

$$\left[\frac{x^2}{4}\right]_0^1 + \left[\frac{x}{2}\right]_1^2 + \left[-\frac{x^2}{4} + \frac{3x}{2}\right]_2^3 = 1$$

$$\frac{1}{4} + 1 - \frac{2}{4} - \frac{9}{4} + \frac{18}{4} + 1 - 3 = 1$$

i.e. 1 = 1

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.15:

A continuous random variable x has p.d.f.

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

(i) Determine the constant a

(ii) Compute P(X < 1.5)

(iii) Find P(1.5 < X < 2.5)

Solution:

(i) Constant 'a' is determined from consideration that total probability is unity.

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

$$\therefore \int_0^1 f(x)\ dx + \int_1^2 f(x)\ dx + \int_2^3 f(x)\ dx = 1$$

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

$$\left[a\frac{x^2}{2}\right]_0^1 + [ax]_1^2 + \left[-a\frac{x^2}{2} + 3ax\right]_2^3 = 1$$

$$\frac{a}{2} + (2a - a) + \left[-\frac{9a}{2} + 9a + 2a - 6a\right] = 1$$

$$\frac{a}{2} + a - \frac{9a}{2} + 5a = 1$$

$$\frac{a}{2} + \frac{2a}{2} - \frac{9a}{2} + \frac{10a}{2} = 1$$

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

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

(ii) $P[x < 1.5] = \int_{-\infty}^{1.5} f(x)\, dx$

$$P[x < 1.5] = \int_0^1 f(x)\, dx + \int_1^{1.5} f(x)\, dx$$

$$P[x < 1.5] = \int_0^1 ax\, dx + \int_1^{1.5} a\, dx$$

$$P[x < 1.5\,] = \left[a\,\frac{x^2}{2}\right]_0^1 + \,[ax]_1^{1.5}$$

$$P[x < 1.5\,] = \frac{a}{2} + \,1.5a - a$$

$$P[x < \,1.5\,] = \frac{a}{2} + \frac{3}{2}a - a$$

$$P[\,x < 1.5\,] = \frac{a}{2} + \frac{3a}{2} - \frac{2a}{2}$$

$$P[x < 1.5\,] = a$$

$$P[x < \,1.5\,] = \frac{1}{2} \qquad\qquad since\ a = \frac{1}{2}$$

(iii) Determine P(1.5 < x <2.5)

$$P[1.5 < x < 2.5] = \int_{1.5}^{2.5} f(x)\ dx$$

$$\therefore\ P[1.5 < x < 2.5] = \int_{1.5}^{2} f(x)\ dx\ + \int_{2}^{2.5} f(x)\ dx$$

$$\therefore\ P[1.5 < x < 2.5] = [ax]_{1.5}^{2} + \left[-\frac{ax^2}{2} + 3ax\right]_{2}^{2.5}$$

$$\therefore\ P[1.5 < x < 2.5] = [2a - 1.5a] + [-\frac{6.25a}{2} + 7.5a + 2a - 6a]$$

$$\therefore\ P[1.5 < x < 2.5] = 0.5a - \frac{6.25a}{2} + 3.5a$$

$$\therefore\ P[1.5 < x < 2.5] = 4a - \frac{6.25a}{2}$$

$$\therefore\ P[1.5 < x < 2.5] = 0.4375$$

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.16:

Find the value of k so that

$$f(x) = \begin{cases} kx^2(1-x^3) & 0 < x < 1 \\ 0 & otherwise \end{cases}$$

Is a proper density function of a continuous variable.

Solution:

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

$$\therefore \int_0^1 f(x)\ dx = 1$$

$$\therefore \int_0^1 kx^2(1-x^3)\ dx = 1$$

$$\therefore k\int_0^1 x^2(1-x^3)\ dx = 1$$

$$\therefore k\int_0^1 (x^2-x^5)\ dx = 1$$

$$\therefore k\left[\frac{x^3}{3}-\frac{x^6}{6}\right]_0^1 = 1$$

$$\therefore\ k\left[\frac{1}{3}-\frac{1}{6}\right]=\ 1$$

$$\therefore\ k\frac{1}{6}=\ 1$$

$$\therefore\ k=6$$

◈◈◈◈◈◈◈

MUMBAI EDUCATIONAL TRUST

MET Institute of Computer Science

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.17:

For the following what is bivariate probability distribution of X and Y

Find (i) $P(X \leq 1, y=2)$

(ii) $P(X \leq 1)$

(iii) $P(Y=3)$

(iv) $P(Y \leq 3)$

(v) $P(X < 3, Y \leq 4)$

X \ Y	1	2	3	4	5	6
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

Solution:

X \ 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
$P_y(Y)$	3/32	3/32	11/64	13/64	6/32	16/64	1

(i) $P(X \leq 1, y=2) = P(X=0, y=2) + P(X=1, y=2)$

$P(X \leq 1, y=2) = 0 + 1/16$

$P(X \leq 1, y=2) = 1/16$

(ii) $P(X \leq 1) = P(X=0) + P(X=1)$
$= 8/32 + 10/16$
$P(X \leq 1) = 7/8$

(iii) $P(Y=3) = 11/64$

(iv) $P(Y \leq 3) = P(Y=1) + P(Y=2) + P(Y=3)$
$= 3/32 + 2/32 + 11/64$
$P(Y \leq 3) = 23/64$

(v) $P(X < 3, Y \leq 4) = P(X=0, Y \leq 4) + P(X=1, Y \leq 4) + P(X=2, Y \leq 4)$

$$P(X < 3,\ Y \leq 4) = \left(\frac{1}{32} + \frac{2}{32}\right) + \left(\frac{1}{16} + \frac{1}{16} + \frac{1}{8} + \frac{1}{8}\right) + \left(\frac{1}{32} + \frac{1}{32} + \frac{1}{64} + \frac{1}{64}\right)$$

$$P(X < 3, Y \leq 4) = \left(\frac{3}{32}\right) + \left(\frac{2}{16} + \frac{2}{8}\right) + \left(\frac{2}{32} + \frac{2}{64}\right)$$

$$P(X < 3, Y \leq 4) = \frac{6 + 8 + 16 + 4 + 2}{64}$$

$$P(X < 3, Y \leq 4) = \frac{36}{64}$$

$$P(X < 3, Y \leq 4) = \frac{9}{16}$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.18:

Given the following what is bivariate probability distribution of X and Y obtain

(i) Marginal distribution of X and Y

(ii) The conditional distribution of X given Y = 2

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

Solution:

Y \ X	-1	0	1	$P(Y) = \sum_x P(x,y)$
0	1/15	2/15	1/15	4/15
1	3/15	2/15	1/15	6/15
2	2/15	1/15	2/15	5/15
$P(X) = \sum_y P(x,y)$	6/15	5/15	4/15	1

(i) Marginal distribution of X

P(X=-1) = 6/15 = 2/5

P(X=0) = 5/15 = 1/3

P(X=1) = 4/15

Marginal distribution of Y
P(Y=0) = 4/15
P(Y=1) = 6/15 = 2/5
P(Y=2) = 5/15 = 1/3

(ii) Conditional distribution of X GIVEN Y=2

$$P(X = x \cap Y = 2) = P(Y = 2) * P(X = x | Y = 2)$$

$$P(X = x | Y = 2) = \frac{P(X = x \cap Y = 2)}{P(Y = 2)}$$

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

$$P(X = -1 | Y = 2) = \frac{P(X = -1 \cap Y = 2)}{P(Y = 2)}$$

$$P(X = -1 | Y = 2) = \frac{\frac{2}{15}}{\frac{5}{15}}$$

$$P(X = -1 | Y = 2) = \frac{2}{5}$$

$$P(X = 0\ |Y = 2) = \frac{P(X = 0\ \cap Y = 2)}{P(Y = 2)}$$

$$P(X = 0\ |Y = 2) = \frac{\frac{1}{15}}{\frac{5}{15}}$$

$$P(X = 0\ |Y = 2) = \frac{1}{5}$$

$$P(X = 1\ |Y = 2) = \frac{P(X = 1\ \cap Y = 2)}{P(Y = 2)}$$

$$P(X = 1\ |Y = 2) = \frac{\frac{2}{15}}{\frac{5}{15}}$$

$$P(X = 1\ |Y = 2) = \frac{2}{5}$$

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.19:
The joint probability distribution of two random variable X and Y is given by P(X=0,Y=1) = 1/3 , P(X=1, Y=-1) = 1/3 , P(X=1,Y=1) = 1/3
Find
(i) The marginal distribution of X and Y
(ii) The conditional distribution of X given Y = 1

Solution:
(i) Following table is joint probability distribution of X and Y and also marginal density function of X and Y.

Y \ X	0	1	$P(Y) = \sum_{x} P(x,y)$
-1	0	1/3	1/3
1	1/3	1/3	2/3
$P(X) = \sum_{y} P(x,y)$	1/3	2/3	1

(i) Conditional distribution of X given Y=1 is given by

$$P(X|Y) = P(X|Y = 1) = \frac{P(X,Y)}{P(Y = 1)}$$

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

$$P(X = 0 \mid Y = 1) = \frac{\frac{1}{3}}{\frac{2}{3}}$$

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

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

$$P(X = 1 \mid Y = 1) = \frac{\frac{1}{3}}{\frac{2}{3}}$$

$$P(X = 1 \mid Y = 1) = \frac{1}{2}$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.20:

Two discrete random variable X and Y have joint pmf given by the following table

X \ Y	1	2	3
1	2/16	2/16	1/16
2	3/16	2/16	1/16
3	2/16	1/16	2/16

(i) Compute the probability of each of the following events
   a. $X \leq 1\frac{1}{2}$
   b. X is odd
   c. XY is even

(ii) Marginal distribution of X and Y

(iii) The conditional distribution of X given Y = 2

Solution:

X \ Y	1	2	3	P(X)
1	2/16	2/16	1/16	5/16
2	3/16	2/16	1/16	6/16
3	2/16	1/16	2/16	5/16
P(Y)	7/16	5/16	4/16	1

(i) Compute the probability of each of the following events

a. P(X ≤ 1½) = P(X=1) = 2/16 + 2/16 + 1/16 = 5/16
b. P(X is odd) = P(X=1) + P(X=3) = 5/16 + 5/16 = 10/16
c. P(XY is even) = P(X=1, Y=2) + P(X=2, Y=1) + P(X=2, Y=2) + P(X=3, Y=2) + P(X=2, Y=3)
P(XY is even) = 2/16 + 3/16 + 2/16 + 1/16 + 1/16 = 9/16

(ii) Marginal distributions are given by

X \ Y	1	2	3	P(X)
1	2/16	2/16	1/16	5/16
2	3/16	2/16	1/16	6/16
3	2/16	1/16	2/16	5/16
P(Y)	7/16	5/16	4/16	1

(iii) Conditional distribution of X given Y=2
We have P(X=x ∩ Y=2) = P(Y=2) * P(X=x | Y=2)

$$P(X = x \,|Y = 2) = \frac{P(X = x \ \cap Y = 2)}{P(Y = 2)}$$

$$P(X = 1 \,|Y = 2) = \frac{P(X = 1 \ \cap Y = 2)}{P(Y = 2)}$$

$$P(X = 1 \,|Y = 2) = \frac{\frac{2}{16}}{\frac{5}{16}}$$

$$P(X = 1 \mid Y = 2) = \frac{2}{5}$$

$$P(X = 2 \mid Y = 2) = \frac{P(X = 2 \cap Y = 2)}{P(Y = 2)}$$

$$P(X = 2 \mid Y = 2) = \frac{\frac{2}{16}}{\frac{5}{16}}$$

$$P(X = 2 \mid Y = 2) = \frac{2}{5}$$

$$P(X = 3 \mid Y = 2) = \frac{P(X = 3 \cap Y = 2)}{P(Y = 2)}$$

$$P(X = 3 \mid Y = 2) = \frac{\frac{1}{16}}{\frac{5}{15}}$$

$$P(X = 3 \mid Y = 2) = \frac{1}{5}$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.21:

X and Y are random variables having joint density function

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

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

Solution:

The joint probability function

$f(x,y) = \frac{1}{27}\,(2x+y) \qquad x = 0,1,2 \;\; ; \;\; y = 0,1,2$

Gives the following table of joint probability distribution of X and Y

X \ Y	0	1	2	$P(x) = f_x(X)$
0	0	1/27	2/27	3/27
1	2/27	3/27	4/27	9/27
2	4/27	5/27	6/27	15/27
P(Y)	6/27	9/27	12/27	1

The marginal distribution of X is given by

$$f_x(x) = \sum_y f(x,y)$$

Is tabulated as last column in the above table.

The Conditional distribution of Y for X=x is given by

$$f_{Y|X}(Y = y \,|X = x) = \frac{f(x,y)}{f_x(x)}$$

Is given by following table

X \ Y	0	1	2
0	0	1/3	2/3
1	2/9	3/9	4/9
2	4/15	5/15	6/15

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.22:

A candy company distributes boxes of chocolates with a mixture of cream, toffees and nuts coated in both light and dark chocolate. For a randomly selected box, Let X and Y, respectively be the proportion of light and dark chocolate that are creams and suppose that the joint density function is

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

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

(ii) Find $P[(x,y) \in A]$ $where\ A = \{(x,y)|0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$

Solution:

(i)

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dx\,dy = \int_{0}^{1}\int_{0}^{1} \frac{2}{5}(2x+3y)\,dx\,dy$$

$$\int_{0}^{1}\int_{0}^{1} f(x,y)dx\,dy = \int_{0}^{1}\left[\frac{2}{5}\left(2xy+\frac{3y^2}{2}\right)\right]_0^1 dx$$

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

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

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

$$\int_0^1\int_0^1 f(x,y)dx\,dy = \frac{2}{5}\left[\frac{5}{2}\right]$$

$$\int_0^1\int_0^1 f(x,y)dx\,dy = 1$$

(ii) $P[(x,y) \in A] = \int_{x=0}^{\frac{1}{2}}\int_{y=\frac{1}{4}}^{\frac{1}{2}} \frac{2}{5}(2x+3y)dx\,dy$

$$since\ A = \{(x,y)|0 < x < \frac{1}{2}, \frac{1}{4} < y < \frac{1}{2}\}$$

$$P[(x,y) \in A] = \frac{2}{5}\int_{x=0}^{\frac{1}{2}}\left[2xy+\frac{3y^2}{2}\right]_{\frac{1}{4}}^{\frac{1}{2}} dx$$

$$P[(x,y)\ \in A\ ] = \frac{2}{5}\int_{x=0}^{\frac{1}{2}}\left[x+\frac{3}{8}-\frac{x}{2}-\frac{3}{32}\right]dx$$

$$P[(x,y)\ \in A\ ] = \frac{2}{5}\int_{x=0}^{\frac{1}{2}}\left[\frac{x}{2}+\frac{9}{32}\right]dx$$

$$P[(x,y)\ \in A\ ] = \frac{2}{5}\left[\frac{x^2}{4}+\frac{9x}{32}\right]_0^{\frac{1}{2}}$$

$$P[(x,y)\ \in A\ ] = \frac{2}{5}\left[\frac{1}{16}+\frac{9}{64}\right]$$

$$P[(x,y)\ \in A\ ] = \frac{2}{5}\left[\frac{13}{64}\right]$$

$$P[(x,y)\ \in A\ ] = \frac{13}{160}$$

$$P[(x,y)\ \in A\ ] = 0.08125$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.23:
The joint PDF of (x, y) is given by

$$f(x, y) = \quad k \qquad 0 \le x \le y \le 2$$

Find k and also marginal and conditional density functions

Solution:
Given joint pdf

$$f(x, y) = \quad k \qquad 0 \le x \le y \le 2$$

To find k let's use the property
Total probability = 1

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

$$\int_{y=0}^{2} \int_{x=0}^{y} k\, dx\, dy \quad = \; 1$$

$$\int_{y=0}^{2} [kx]_0^y \; dy \; = 1$$

$$\int_{y=0}^{2} ky \; dy \; = 1$$

$$\left[\frac{ky^2}{2}\right]_0^2 \; = 1$$

$$2k \; = 1$$

$$\therefore \; k \; = \frac{1}{2}$$

The marginal density of x is given by

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

$$f(x) = \int_{y=x}^{2} k \; dy$$

$$f(x) = k \; [y]_x^2$$

$$f(x) = 2k - kx$$

$$f(x) = k(2 - x)$$

$$f(x) = \frac{1}{2}(2 - x) \qquad since\ k = \frac{1}{2}$$

$$f(x) = 1 - \frac{x}{2}$$

$$f(x) = \frac{2 - x}{2}$$

Marginal density of y is given by

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

$$f(y) = \int_{x=0}^{y} k\ dx$$

$$f(y) = k\ [x]_0^y$$

$$f(y) = ky$$

$$\therefore\ f(y) = \frac{y}{2} \qquad since\ k = \frac{1}{2}$$

The conditional density function Y given X

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

$$f(y \mid x) = \frac{\frac{1}{2}}{\frac{2-x}{2}}$$

$$f(y \mid x) = \frac{1}{2-x}$$

The conditional density function X given Y

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

$$f(x \mid y) = \frac{\frac{1}{2}}{\frac{y}{2}}$$

$$f(x \mid y) = \frac{1}{y}$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.24:

The joint PDF of a two dimensional random variable (x, y) is given by

$$f(x,y) = \begin{cases} 2 & 0 < x < 1, 0 < y < x \\ 0 & otherwise \end{cases}$$

(i) Find the marginal and conditional density functions of X and Y
(ii) Find the conditional density function of Y|X and X|Y
(iii) Check for independence of X and Y

Solution:

Evidently f(x,y) ≥ 0 and

$$\int_0^1 \int_0^x f(x,y)dx\,dy = \int_0^1 \int_0^x 2\,dx\,dy$$

$$\int_0^1 \int_0^x f(x,y)dx\,dy = 2\int_0^1 x\,\,dx$$

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

$$\int_0^1 \int_0^x f(x,y)dx\,dy = 1$$

(i) The marginal pdf of X and Y are

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

$$f_x(x) = \int_0^x 2\,dy$$

$$f_x(x) = [2y]_0^x$$

$$f_x(x) = 2x \qquad 0 < x < 1$$

$$f_x(x) = 0 \quad otherwise$$

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

$$f_y(y) = \int_y^1 2\,dx$$

$$f_y(y) = [2x]_y^1$$

$$f_y(y) = 2 - 2y \qquad 0 < y < x$$

$$f_y(y) = 2\,(1-y) \qquad 0 < y < x$$

$$f_y(y) = 0 \quad otherwise$$

(ii) The conditional density function of Y given X

$$f_{y|x}(y \mid x) = \frac{f_{xy}(x, y)}{f_x(x)}$$

$$f_{y|x}(y \mid x) = \frac{2}{2x}$$

$$f_{y|x}(y \mid x) = \frac{1}{x} \qquad 0 < x < 1$$

$$f(y \mid x) = \frac{1}{2-x}$$

The conditional density function X given Y

$$f_{x|y}(x \mid y) = \frac{f_{xy}(x, y)}{f_y(y)}$$

$$f_{x|y}(x \mid y) = \frac{2}{2(1-y)}$$

$$f_{x|y}(x \mid y) = \frac{1}{1-y} \qquad 0 < y < x$$

(iii) Since

$$f_x(x)f_y(y) = (2x)\,2(1-y) \quad \neq f_{xy}(x,y)$$

$$x \text{ and } y \text{ are not independent}$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.25:

The joint PDF of a two dimensional random variable (x, y) is given by

$$f_{xy}(x,y) = \begin{cases} \dfrac{x^3y^3}{16} & 0 \le x \le 2, \quad 0 \le y \le 2 \\ 0 & otherwise \end{cases}$$

Find the marginal densities of X and Y. Also find the cumulative distribution functions of X and Y.

Solution:

$$f_x(x) = \int_0^2 f_{xy}(x,y)dy$$

$$f_x(x) = \int_0^2 \frac{x^3y^3}{16}\,dy$$

$$f_x(x) = \left[\frac{x^3y^4}{64}\right]_0^2$$

$$f_x(x) = \frac{x^3}{4} \qquad 0 \le x \le 2$$

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

$$f_y(y) = \int_0^2 \frac{x^3y^3}{16}\, dx$$

$$f_y(y) = \left[\frac{x^4y^3}{64}\right]_0^2$$

$$f_y(y) = \frac{y^3}{4} \qquad\qquad 0 \le y \le 2$$

As given

$$f_{xy}(x,y) = \begin{cases} \dfrac{x^3y^3}{16} & 0 \le x \le 2, \qquad 0 \le y \le 2 \\ 0 & otherwise \end{cases}$$

$$F_x(X) = \int_0^x f_x(x)dx$$

$$F_x(X) = \int_0^x \frac{x^3}{4} dx$$

$$F_x(X) = \left[\frac{x^4}{16}\right]_0^x$$

$$F_x(X) = \frac{x^4}{16} \qquad 0 \le x \le 2$$

$$F_x(X) = \begin{cases} 0 & x < 0 \\ \dfrac{x^4}{16} & 0 \le x \le 2 \\ 1 & x > 2 \end{cases}$$

$$F_y(Y) = \int_0^y f_y(y)dy$$

$$F_y(Y) = \int_0^y \frac{y^3}{4}dy$$

$$F_y(Y) = \left[\frac{y^4}{16}\right]_0^y$$

$$F_y(Y) = \frac{y^4}{16} \qquad 0 \le y \le 2$$

$$F_y(Y) = \begin{cases} 0 & y < 0 \\ \dfrac{y^4}{16} & 0 \le y \le 2 \\ 1 & y > 2 \end{cases}$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.26:

The joint PDF of a two dimensional random variable (x, y) is given by

$$f(x,y) = \begin{cases} \frac{8}{9}xy & 1 \le x \le y \le 2 \\ 0 & otherwise \end{cases}$$

(i) Find the marginal densities of X and Y.

(ii) Find the conditional density function of Y given X=x, and the conditional density function of X given Y=y

Solution:

(i)

$$f_x(x) = \int_x^2 f(x,y)dy$$

$$f_x(x) = \int_x^2 \frac{8}{9}xy\, dy$$

$$f_x(x) = \left[\frac{8xy^2}{18}\right]_x^2$$

$$f_x(x) = \frac{4}{9}[xy^2]_x^2 \qquad 1 \leq x \leq 2$$

$$f_x(x) = \frac{4}{9}[4x - \; x^3\,]$$

$$f_x(x) = \frac{4}{9}x[4 - \; x^2\,]$$

$$f_y(y) = \int_1^y f(x, y)dx$$

$$f_y(y) = \int_1^y \frac{8}{9}xy \; dx$$

$$f_y(y) = \left[\frac{8x^2y}{18}\right]_1^y$$

$$f_y(y) = \frac{4}{9}\,[x^2y]_1^y$$

$$f_y(y) = \frac{4}{9}\,[y^3 \; - y]$$

$$f_y(y) = \frac{4}{9}y\,[y^2 \; - 1]$$

(ii)

$$f_{x|y}(x \mid y) = \frac{f_{xy}(x, y)}{f_y(y)} \qquad 1 \leq x \leq y$$

$$f_{x|y}(x \mid y) = \frac{\frac{8}{9}xy}{\frac{4}{9}y\,(y^2 - 1)}$$

$$f_{x|y}(x \mid y) = \frac{2x}{(y^2 - 1)}$$

$$f_{y|x}(y \mid x) = \frac{f_{xy}(x, y)}{f_x(x)}$$

$$f_{y|x}(y \mid x) = \frac{\frac{8}{9}xy}{\frac{4}{9}x(4 - x^2)}$$

$$f_{y|x}(y \mid x) = \frac{2y}{(4 - x^2)}$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.27:

Consider an experiment “two coins are tossed”

Let the random variable X=’number of heads’

(i) Find the values of X.

(ii) Find the probability of X.

(iii) Find the probability mass function.

(iv) Find the cumulative density function.

Solution:

The sample space is S = { HH, HT , TH, TT }

Sample space (S)	HH	HT	TH	TT
Number of heads	2	1	1	0

X	P(X)	pmf P(X)	CDF F(X)
0	1/4	1/4	1/4
1	2/4	2/4	3/4
2	1/4	1/4	4/4 = 1

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.28:

Given the following bivariate probability distribution obtain

(i) Marginal distribution of X and Y

(ii) The conditional distributions of X given Y=1

(iii) P(X +Y) <4

Y \ X	1	2	3
1	0.1	0.1	0.2
2	0.2	0.3	0.1

Solution:

(i) Marginal distributions are given below

Y \ X	1	2	3	$P(Y) = \sum_x P(x,y)$
1	0.1	0.1	0.2	0.4
2	0.2	0.3	0.1	0.6
$P(X) = \sum_y P(x,y)$	0.3	0.4	0.3	1

(ii) The conditional distribution of X given Y=1

We have P(X=x $\cap$ Y=1) = P(Y=1) P(X=x | Y=1)

$$P(X = 1 \mid Y = 1) = \frac{P(X = 1 \ \cap Y = 1)}{P(Y = 1)}$$

$$P(X = 1\,|Y = 1) = \frac{0.1}{0.4}$$

$$P(X = 1\,|Y = 1) = 0.25$$

$$P(X = 2\,|Y = 1) = \frac{P(X = 2\ \cap Y = 1)}{P(Y = 1)}$$

$$P(X = 2\,|Y = 1) = \frac{0.1}{0.4}$$

$$P(X = 2\,|Y = 1) = 0.25$$

$$P(X = 3\,|Y = 1) = \frac{P(X = 3\ \cap Y = 1)}{P(Y = 1)}$$

$$P(X = 3\,|Y = 1) = \frac{0.2}{0.4}$$

$$P(X = 3\,|Y = 1) = 0.5$$

(iii) $P(\,X + Y\,) < 4$

$$P(\,X + Y\,) = P(X = 1, Y = 1) + P(X = 1, Y = 2) + P(X = 2, Y = 1)$$

$$P(\,X + Y\,) = 0.1 + 0.2 + 0.1$$
$$P(\,X + Y\,) = 0.4$$
$$P(\,X + Y\,) < 4$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.29:

If the random variable X takes the values 0!, 1!, 2!, ---- with probability law $P(X = x!) = \frac{e^{-x}}{x!}$ $\quad x = 0,1,2$ then

$$E(X) = \sum_{x=0}^{\infty} x!\, P(X = x!)$$

Solution:

$$E(X) = \sum_{x=0}^{\infty} x!\, P(X = x!)$$

$$E(X) = \sum_{x=0}^{\infty} x!\, \frac{e^{-x}}{x!}$$

$$E(X) = \sum_{x=0}^{\infty} e^{-x}$$

$$E(X) = e^{-x} \sum_{x=0}^{\infty} 1 \qquad \therefore E(X) does\ not\ exist$$

$$E(X) = \infty \qquad since \sum_{x=0}^{\infty} 1 = \infty$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.30:

Consider a random variable with pdf

$$f(x) = \frac{1}{\pi} \frac{1}{1+x^2} \qquad -\infty < x < \infty$$

Solution:

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

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

Let us consider $1+x^2 = u$

$\therefore du = 2x\, dx$

$$\frac{du}{2} = x\, dx$$

$$since \quad \frac{\frac{du}{2}}{u} \quad i.e. \frac{1}{2} \left[\frac{1}{u}\, du\right]$$

$$E(X) = \frac{1}{\pi} \frac{1}{2} \int_{-\infty}^{\infty} \frac{1}{u}\, du$$

$$E(X) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{u}\, du$$

$$E(X) = \frac{1}{2\pi} [\log u\,]_{-\infty}^{\infty}$$

$$E(X) = \frac{1}{2\pi} [\log\ (1 + x^2)\,]_{-\infty}^{\infty}$$

$$E(X) = \infty$$

E(X) does not exist.

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.31:

Let X be a continuous random variable with an exponential density function given by

$$f(x) = \lambda e^{-\lambda x} \qquad -\lambda > 0$$

Then $E(X) = ?$

Solution:

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

$$E(X) = \int_{0}^{\infty} x\, \lambda\, e^{-\lambda x}\, dx$$

Let $\lambda x = t$ $\quad \therefore \lambda dx = dt$

$$E(X) = \frac{1}{\lambda} \int_0^\infty t\, e^{-t}\, dt \qquad \int u\, dv = uv - \int v\, du$$

$$E(X) = \frac{1}{\lambda} [-te^{-t} + e^{-t}]_0^\infty \qquad let\ us\ consider\ u = t$$

$$E(X) = \frac{1}{\lambda} [e^{-t} - te^{-t}]_0^\infty \qquad dv = e^{-t}\, dt \qquad v = -e^{-t}$$

$$E(X) = \frac{1}{\lambda}$$

❖❖❖❖❖❖❖

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Addition theorem of Expectation

Example 4.32:

If X and Y are random variables, then E(X+Y) = E(X) + E(Y) provided all the expectations exist.

Solution:

Let X and Y be continuous random variable with joint pdf $f_{xy}(x,y)$ and marginal pdf's $f_x(x)$ and $f_y(y)$ respectively, then by definition

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

And

$$E(Y) = \int_{-\infty}^{\infty} y\, f_y(y)\, dy$$

$$\therefore\ E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\ f_{xy}(x,y)\, dx\, dy$$

$$E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\ f_{xy}(x,y)\, dx\, dy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\ f_{xy}(x,y)\, dx\, dy$$

$$E(X+Y) = \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f_{xy}(x,y)\, dy \right] dx \;+\; \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f_{xy}(x,y)\, dx \right] dy$$

$$E(X+Y) = \int_{-\infty}^{\infty} x f_x(x)\, dx \;+\; \int_{-\infty}^{\infty} y f_y(y)\, dy$$

$$\therefore \quad E(X+Y) = E(X) + E(Y)$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Multiplication theorem of Expectation

Example 4.33:

If X and Y are random variables, then E(XY) = E(X).E(Y) provided all the expectations exists.

Solution:

Let X and Y be continuous random variable with joint pdf $f_{xy}(x,y)$ and marginal pdf's $f_x(x)$ and $f_y(y)$ respectively, then by definition

$$\therefore\ E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\ f_{xy}(x,y)\,dx\,dy$$

$$\therefore E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_x(x) f_y(y) dx\,dy \quad since \quad x\ and\ y\ are\ independent$$

$$\therefore E(XY) = \int_{-\infty}^{\infty} x\, f_x(x)dx \quad \int_{-\infty}^{\infty} y\, f_y(y)dy$$

$\therefore E(XY) = E(X).E(Y)$ provided X and Y are independent.

$$\therefore \quad E(XY) = E(X).E(Y)$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.34:

If X is a random variable 'a' is a constant then,

(i) E[a g(x)] = a E[g(x)]

(ii) E[g(x) + a] = E[g(x)] + a

Where g(x) is a function of X, is a random variable and all the expectations exists.

Solution:

i)

$$E[a\ g(X)] = \int_{-\infty}^{\infty} a\ g(x) f(x)\ dx$$

$$E[a\ g(X)] = a \int_{-\infty}^{\infty} g(x) f(x)\ dx$$

$$E[a\ g(X)] = a\ E[\ g(x)]$$

ii)

$$E[\ g(x) + a] = \int_{-\infty}^{\infty} [g(x) + a]\ f(x)\ dx$$

$$E[\,g(x) + \,a] = \int_{-\infty}^{\infty} g(x)f(x)\,dx \,+\, a\int_{-\infty}^{\infty} f(x)dx$$

$$E[\,g(x) + \,a] = E[g(x)] + a \qquad since \int_{-\infty}^{\infty} f(x)dx \,= 1$$

$$E[\,g(x) + \,a] = E[g(x)] + a$$

◈◈◈◈◈◈◈

Corollary 1

If g(x) = X then E[aX] = a E(X) and E[X+a] = E[X] + a

Corollary 2

If g(x) = 1 then E[a] = a

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.35:

Theorem: If X is a discrete random variable with pmf P(X), then prove that

(i) E(aX+b) = aE(X) + b and
V(aX+b) = a^2 V(X) where a and b are constants.

(ii) E(aX-b) = aE(X) - b and
V(aX-b) = a^2 V(X) where a and b are constants.

Theorem: If X is a continuous random variable with pmf f(X), then prove that

(i) E(aX+b) = aE(X) + b and
V(aX+b) = a^2 V(X) where a and b are constants.

(ii) E(aX-b) = aE(X) - b and
V(aX-b) = a^2 V(X) where a and b are constants.

Proof:

By definition

$$E[ax+b] = \int_{-\infty}^{\infty} (ax+b) f(x)\, dx$$

$$E[ax+b] = a \int_{-\infty}^{\infty} x f(x)\, dx \ + b \int_{-\infty}^{\infty} f(x) dx$$

$$E[ax + b] = aE(x) + b$$

◈◈◈◈◈◈◈

Corollary

If b = 0 then E[aX] = a E(X)

Corollary

If a = 1 , b = $-\bar{X} = -E(X)$ then $E(X - \bar{X}) = 0$

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.36:

Theorem: If X is a random variable then
$V(aX+b) = a^2 V(X)$ where a and b are constants.

Proof:

Let $Y = aX + b$

$\therefore E(Y) = E(aX + b)$

$E(Y) = a E(X) + b$

$\therefore Y - E(Y) = (aX + b) - [aE(X) + b]$

$\therefore Y - E(Y) = aX + b - aE(X) - b$

$\therefore Y - E(Y) = aX - aE(X)$

$\therefore Y - E(Y) = a [X - E(X)]$

Squaring both sides

$\therefore [Y - E(Y)]^2 = a^2 [X - E(X)]^2$

Take expectation s of both sides

$\therefore$ $E[\,Y - E(Y)\,]^2 = a^2\, E[\,X - E(X)\,]^2$

$V(Y) = a^2\, V(X)$

$V(aX + b) = a^2\, V(X)$

Case (i) If $b = 0$ then $V(aX) = a^2 V(X)$
i.e. variance is not independent of scale

Case (ii) If $a = 0$ then $V(b) = 0$
i.e. variance of constant = 0

Case (iii) If $a = 1$ then $V(X+b) = V(X)$
i.e. variance is independent of change of origin

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.37:

If the mean of the following frequency distribution is 4.876 find k.

X	3.2	5.8	7.9	4.5
Frequency	k	k+2	k-3	k+6

Proof:

The mean i.e. expected value is given by

Mean = E(X) = $\sum$ x P(x)

4.876 = 3.2 k + 5.8 (k+2) + 7.9 (k-3) +4.5 (k+6)

4.876 = 21.4 k + 14.9

$\therefore$ k =-10.024/21.4

$\therefore$ k = -0.4684

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.38:

Suppose a random variable X, takes values -3, -1, 2 & 5 with probabilities (2k-3)/10 , (k-2)/10 , (k-1)/10, (k+1)/10

(i) Determine the distribution of X

(ii) Find the expected value E(X) of X

Proof:

Since $\sum$ x P(x) = 1

$$\frac{2k-3}{10}+\frac{k-2}{10}+\frac{k-1}{10}+\frac{k+1}{10}=1$$

$$\frac{2k-3+k-2+k-1+k+1}{10}=1$$

$$\frac{5k-5}{10}=1$$

$$5k-5 = 10$$
$$5(k-1) = 10$$
$$k-1 = 2$$
$$\therefore\ k = 3$$

(i) Distribution of X is given by

X	F(X=x)= P(X $\leq$ x)
-3	3/10
-1	4/10

2	6/10
5	1

$$P(X \leq -3) = \frac{2k-3}{10} = \frac{3}{10} \quad since\ k = 3$$

$$P(X \leq -1) = P(X = -3) + P(X = -1)$$

$$P(X \leq -1) = \frac{2k-3}{10} + \frac{k-2}{10}$$

$$P(X \leq -1) = \frac{2k-3+k-2}{10}$$

$$P(X \leq -1) = \frac{3k-5}{10}$$

$$P(X \leq -1) = \frac{4}{10}$$

$$P(X \leq 2) = P(X = -3) + P(X = -1) + P(X = 2)$$

$$P(X \leq 2) = \frac{2k-3}{10} + \frac{k-2}{10} + \frac{k-1}{10}$$

$$P(X \leq 2) = \frac{4k-6}{10}$$

$$P(\, X \leq\; 2) = \frac{6}{10}$$

$$P(\, X \leq\; 5) = P(X = -3) + \; P(X = \; -1) + \; P(X = 2) + \; P(X = 5)$$

$$P(\, X \leq\; 5) = \frac{2k-3}{10} + \frac{k-2}{10} + \frac{k-1}{10} + \frac{k+1}{10}$$

$$P(\, X \leq\; 5) = \frac{5k-5}{10}$$

$$P(\, X \leq\; 5) = 1$$

(ii) The expected value is given by

E(X) = $\sum$ x P(x)

$$E(X) = (-3)\; \frac{2k-3}{10} + \; (-1)\frac{k-2}{10} + \; 2\; \frac{k-1}{10} + \; 5\; \frac{k+1}{10}$$

$$E(X) = (-3)\; \frac{3}{10} + \; (-1)\frac{1}{10} + \; 2\; \frac{2}{10} + \; 5\; \frac{4}{10}$$

$$E(X) = \; -\frac{9}{10} \; -\frac{1}{10} + \frac{4}{10} + \frac{20}{10}$$

$$E(X) = \frac{14}{10}$$

$$E(X) = \; 1.4$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.40:
A random variable X is defined as the sum of faces when a pair of dice is thrown find the probability distribution of X and

(i) Expectation value of X
(ii) Probability that the sum is less than 4
(iii) Probability 6 < x < 10

Proof:
Random variable X = the sum of faces when a pair of dice is thrown
The probability distribution of X is

X	2	3	4	5	6	7	8	9	10	11	12
P(X)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

(i) $E(X) = \sum x\, P(x)$

$$E(X) = \frac{2}{36} + \frac{6}{36} + \frac{12}{36} + \frac{20}{36} + \frac{30}{36} + \frac{42}{36} + \frac{40}{36} + \frac{36}{36} + \frac{30}{36} + \frac{22}{36} + \frac{12}{36}$$

$$E(X) = \frac{252}{36}$$

$$E(X) = 7$$

(ii) $P(X < 4) = P(X=2) + P(X=3)$

$$P(X < 4) = \frac{1}{36} + \frac{2}{36}$$

$$P(X < 4) = \frac{3}{36}$$

$$P(X < 4) = 0.0833$$

(iii) P(6 < X < 10) = P(X=7) + P(X=8) + P(X=9)

$$P(6 < X < 10) = \frac{6}{36} + \frac{5}{36} + \frac{4}{36}$$

$$P(6 < X < 10) = \frac{15}{36}$$

$$P(6 < X < 10) = 0.4166$$

◈◈◈◈◈◈◈

MUMBAI EDUCATIONAL TRUST
MET Institute of Computer Science

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.41:

The number of hardware failure system in a week of operation has the following probability mass function

No. of failures	0	1	2	3	4	5	6
Probability	0.18	0.28	0.25	0.18	0.06	0.04	0.01

Find the expectation and variance of the number of failure

Proof:

E(X) = $\sum$ x P(x)

$$E(X) = (0 * 0.18) + (1 * 0.28) + (2 * 0.25) + (3 * 0.18) + (4 * 0.06) + (5 * 0.04) + (6 * 0.01)$$

$$E(X) = 1.82 \quad failures\ per\ week$$

$$E(X) = 1.82$$

$$E(X^2) = (0 * 0.18) + (1 * 0.28) + (4 * 0.25) + (9 * 0.18) + (16 * 0.06) + (25 * 0.04) + (36 * 0.01)$$

$$E(X^2) = 5.22$$

Variance

$$V(X) = E(X^2) - [E(X)]^2$$

$$V(X) = 5.22 - (1.82)^2$$

$$V(X) = 5.22 - 3.3124$$

$$V(X) = 1.9076$$

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.42:

Let X be random variable with following probability distribution

X	-3	6	9
P(X=x)	1/6	1/2	1/3

Find E(X) and $E(X^2)$ and using the laws of expectation, evaluate $E(2X+1)^2$

Proof:

$E(X) = \sum x\, P(x)$

$$E(X) = (-3) * \frac{1}{6} + 6 * \frac{1}{2} + 9 * \frac{1}{3}$$

$$E(X) = -\frac{1}{2} + 3 + 3$$

$$E(X) = 6 - \frac{1}{2}$$

$$E(X) = \frac{11}{2}$$

$$E(X) = 5.5$$

$$E(X^2) = 9 * \frac{1}{6} + 36 * \frac{1}{2} + 81 * \frac{1}{3}$$

$$E(X^2) = \frac{3}{2} + 18 + 27$$

$$E(X^2) = 45 + 1.5$$

$$E(X^2) = 46.5$$

$$E(2X + 1)^2 = E(4X^2 + 4X + 1)$$

$$E(2X + 1)^2 = 4E(X^2) + 4\,E(X) + E(1)$$

$$E(2X + 1)^2 = 4 * 46.5 + 4 * 5.5 + 1$$

$$E(2X + 1)^2 = 186 + 22 + 1$$

$$E(2X + 1)^2 = 209$$

❖❖❖❖❖❖❖

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.43:

Let X be a discrete random variable with the following pdf

X	0	1	2	3
P(X=x)	1/3	1/2	1/24	1/8

Find E(Y) where Y = $(X-1)^2$

Proof:

E(X) = $\sum$ x P(x)

$$E(X) = 0 * \frac{1}{3} + 1 * \frac{1}{2} + 2 * \frac{1}{24} + 3 * \frac{1}{8}$$

$$E(X) = \frac{1}{2} + \frac{1}{12} + \frac{3}{8}$$

$$E(X) = \frac{12 + 2 + 9}{24}$$

$$E(X) = \frac{23}{24}$$

$$E(X^2) = 0 * \frac{1}{3} + 1 * \frac{1}{2} + 4 * \frac{1}{24} + 9 * \frac{1}{8}$$

$$E(X^2) = \frac{1}{2} + \frac{4}{24} + \frac{9}{8}$$

$$E(X^2) = \frac{12 + 4 + 27}{24}$$

$$E(X^2) = \frac{43}{24}$$

$$E(X-1)^2 = E(\,X^2 - 2X + 1)$$

$$E(X-1)^2 = E(X^2) - 2\,E(X) + E(1)$$

$$E(X-1)^2 = \frac{43}{24} - \; 2 * \frac{23}{24} \; + 1$$

$$E(X-1)^2 = \frac{43 - 46 + 24}{24}$$

$$E(X-1)^2 = \frac{21}{24}$$

$$E(X-1)^2 = \frac{7}{8}$$

◈◈◈◈◈◈◈

MUMBAI EDUCATIONAL TRUST
MET Institute of Computer Science

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.44:

Let X be a random variable with probability distribution as follows

X	0	1	2	3
f(X)	1/3	1/2	0	1/6

Find E(Y) where Y = $(X-1)^2$

Proof:

E(X) = $\sum$ x P(x)

$$E(X) = 0 * \frac{1}{3} + 1 * \frac{1}{2} + 2 * 0 + 3 * \frac{1}{6}$$

$$E(X) = 0 + \frac{1}{2} + 0 + \frac{3}{6}$$

$$E(X) = \frac{6}{6}$$

$$E(X) = 1$$

$$E(X^2) = 0 * \frac{1}{3} + 1 * \frac{1}{2} + 4 * 0 + 9 * \frac{1}{8}$$

$$E(X^2) = \frac{1}{2} + \frac{3}{2}$$

$$E(X^2) = 2$$

$$E(X-1)^2 = E(\,X^2 - 2X + 1)$$

$$E(X-1)^2 = E(X^2) - 2\,E(X) + E(1)$$

$$E(X-1)^2 = 2 - \ 2 * 1 \ + 1$$

$$E(X-1)^2 = 1$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.45:

Find the expectation of a number on a die when thrown. If two unbiased dice are thrown, find the expectation value of the sum of number of points on them. Also find the variance of the sum.

Proof:

Consider one dice is thrown

Let X be the number of points on it

E(X) = $\sum$ x P(x)

$$E(X) = 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6}$$

$$E(X) = (1 + 2 + 3 + 4 + 5 + 6) * \frac{1}{6}$$

$$E(X) = \frac{21}{6}$$

$$E(X) = \frac{7}{2}$$

Let X_i = number obtained on i^{th} dice (i=1,2)

$\therefore$ S = X1 + X2

E(S) = E(X1) + E(X2)

E(S) = 7/2 + 7/2

E(S) = 7

$$E(X^2) = \sum x^2 P(x)$$

$$E(X^2) = (1+4+9+16+25+36) * \frac{1}{6}$$

$$E(X^2) = \frac{91}{6}$$

Variance

$V(X) = E(X^2) - [E(X)]^2$

$$V(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2$$

$$V(X) = \frac{91}{6} - \frac{49}{4}$$

$$V(X) = \frac{35}{12}$$

V(S) = V(X1 + X2)

V(S) = V(X1) + V(X2)

$$V(S) = \frac{35}{12} + \frac{35}{12}$$

$$V(S) = \frac{70}{12}$$

$$V(S) = \frac{35}{6}$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.46:

A coin is tossed four times. Calculate the expectation value and variance of the number of heads obtained.

Proof:

Srno	Outcome	Value of X
1	HHHH	4
2	HHHT	3
3	HHTH	3
4	HHTT	2
5	HTHH	3
6	HTHT	2
7	HTTH	2
8	HTTT	1
9	THHH	3
10	THHT	2
11	THTH	2
12	THTT	1
13	TTHH	2
14	TTHT	1
15	TTTH	1
16	TTTT	0

Hence the probability table obtained is as follows

X	0	1	2	3	4
P(X)	1/16	4/16	6/16	4/16	1/16

$$E(X) = \sum x\, P(x)$$

$$E(X) = 0 * \frac{1}{16} + 1 * \frac{4}{16} + 2 * \frac{6}{16} + 3 * \frac{4}{16} + 4 * \frac{1}{16}$$

$$E(X) = \frac{4}{16} + \frac{12}{16} + \frac{12}{16} + \frac{4}{16}$$

$$E(X) = \frac{32}{16}$$

$$E(X) = 2$$

$$E(X)^2 = \sum x^2\, P(x)$$

$$E(X^2) = 0 * \frac{1}{16} + 1 * \frac{4}{16} + 4 * \frac{6}{16} + 9 * \frac{4}{16} + 16 * \frac{1}{16}$$

$$E(X^2) = \frac{4}{16} + \frac{24}{16} + \frac{36}{16} + \frac{16}{16}$$

$$E(X^2) = \frac{80}{16}$$
$$E(X^2) = 5$$

Variance
V(X) = $E(X^2) - [E(X)]^2$

$$V(X) = 5 - 2^2$$

$$V(X) = 5 - 4$$

$$V(X) = 1$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.47:

A coin is tossed until a head appears. What is the expectation of the number of tosses required?

Proof:

Let X be the number of tosses required to get the first head

$$E(X) = \sum_{X=0}^{\infty} x\,P(x)$$

Event	X	P(X)
1	H	½
2	TH	½ * ½ = ¼
3	THH	½ * ½ * ½ = 1/8
.	.	.
.	.	.
.	.	.

$$E(X) = 1 * \frac{1}{2} + 2 * \frac{1}{4} + 3 * \frac{1}{8} + ----$$

Let

$$S = 1 * \frac{1}{2} + 2 * \frac{1}{4} + 3 * \frac{1}{8} + ----$$

$$S = \frac{1}{2} + \frac{2}{4} + \frac{3}{8} + ---- \qquad eq\ 1$$

Divide both sides by 2

$$\frac{1}{2}S = 1*\frac{1}{4} + 2*\frac{1}{8} + 3*\frac{1}{16} + ----$$

$$\frac{S}{2} = \frac{1}{4} + \frac{2}{8} + \frac{3}{16} + ---- \qquad eq\ 2$$

Eq 1 – Eq 2

$$\left(S - \frac{S}{2}\right) = \left(\frac{1}{2} + \frac{2}{4} + \frac{3}{8} + ---\right) - \left(\frac{1}{4} + \frac{2}{8} + \frac{3}{16} + ---\right)$$

$$\left(S - \frac{S}{2}\right) = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + ----$$

$$\frac{S}{2} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + ----$$

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

$$\therefore S = 2$$

$$\therefore E(X) = 2$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.48:

Let variable X have the distribution

P(X=0) = P(X=2) = P

P(X=1) = 1 – 2P for $0 \leq P \leq ½$

For what P, is the var(X) a maximum?

Proof:

Here random variable X takes values 0,1,2 with respective probabilities P, 1-2P , P for $0 \leq P \leq ½$

$$E(X) = \sum_{X=0}^{2} x\,P(x)$$

$$E(X) = 0 * P + 1 * (1 - 2P) + 2 * P$$

$$E(X) = (1 - 2P) + 2P$$

$$E(X) = 1$$

$$E(X^2) = 0 * P + 1 * (1 - 2P) + 4 * P$$

$$E(X^2) = (1 - 2P) + 4P$$

$$E(X^2) = 1 + 2P$$

$$Var\ (X) = E(X^2)\ -\ [E(X)]^2$$

$$Var\ (X) = 1 + 2P\ -\ [1]^2$$

$$Var\ (X) = 1 + 2P - 1$$

$$Var\ (X) = 2P$$

$$[Var\ (X)]_{max} = 2 * \frac{1}{2} = 1$$

$$\therefore Var(X)\ is\ maximum\ at\ P = \frac{1}{2}$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.49:

Let X be a random variable for which E(X) = 10 and V(X) = 25. Find the values of a and b such that Y=aX – b has expectation zero and variance 1.

Proof:

Given E(X) = 10

V(X) = 25

but

$$Var\,(X) = E(X^2) \;-\; [E(X)]^2$$

Now Y = aX – b and E(Y) = 0

$$\therefore\; E(aX\; - b) = 0$$

$$aE(X) - \; b = 0$$

$$a * 10 - \; b = 0$$

$$10a \;= b$$

$$b = 10a \qquad - - - - eq\; 1$$

$$and \qquad\qquad V(Y) = 1$$

$$\therefore\ V(aX - b) = 1$$

$$a^2 V(X) = 1$$

$$a^2 * 25 = 1$$

$$a^2 = \frac{1}{25}$$

$$a = \pm \frac{1}{5} \qquad --- eq\ 2$$

Therefore using equation 1 and 2

$$b = 10a$$

$$b = 10 * \frac{1}{5} \quad or \quad b = 10 * \frac{1}{-5}$$

$$b = 2 \quad or \quad b = -2$$

$$b = \pm 2$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.50:

Let X be a random variable for which E(X) = 24 and V(X) = 16. Find the values of a and b such that Y=aX – b has expectation zero and variance 2.

Proof:

Given E(X) = 24

V(X) = 16

but

$$Var\ (X) = E(X^2)\ -\ [E(X)]^2$$

Now Y = aX – b and E(Y) = 0

$$\therefore\ E(aX\ - b) = 0$$

$$aE(X) -\ b = 0$$

$$a * 24 -\ b = 0$$

$$24a\ = b$$

$$b = 24a \qquad ----eq\ 1$$

$$and \qquad V(Y) = 2$$

$$\therefore \; V(aX - b) = 2$$

$$a^2 V(X) = 2$$

$$a^2 * 16 = 2$$

$$a^2 = \frac{1}{8}$$

$$a = \pm \frac{1}{2\sqrt{2}} \qquad ---eq\ 2$$

Therefore using equation 1 and 2

$$b = 24a$$

$$b = 24 * \frac{1}{2\sqrt{2}} \quad or \quad b = 24 * \frac{1}{-2\sqrt{2}}$$

$$b = \frac{12}{\sqrt{2}} \quad or \quad b = -\frac{12}{\sqrt{2}}$$

$$b = \pm 6\sqrt{2}$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.51:
Let X be a random variable for which E(X) = 24 and V(X) = 2. Find the values of a and b such that Y=aX – b has expectation 20 and variance 8.

Proof:
Given E(X) = 24
V(X) = 2
but
$$Var\ (X) = E(X^2)\ -\ [E(X)]^2$$

Now Y = aX – b and E(Y) =20

$$\therefore\ E(aX\ - b) = 20$$

$$aE(X) - \ b = 20$$

$$a * 24 - \ b = 20$$

$$24a - b = 20$$

$$b = 24a - 20 \qquad - - - -eq\ 1$$

$$and \qquad V(Y) = 8$$

$$\therefore\ V(aX-b)=8$$

$$a^2V(X)=8$$

$$a^2 * 2 = 8$$

$$a^2 = 4$$

$$a = \pm 2 \qquad ---eq\ 2$$

Therefore using equation 1 and 2

$$b = 24a - 20$$

$$b = 24*2-20 \quad or \quad b = 24*(-2)-20$$

$$b = 28 \quad or \quad b = -68$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.52:

Let X_1 and X_2 be two stochastic random variable having variance k and 2 respectively. If variance of $Y=3X_1 - X_2$ is 25 find k.

Proof:

Given $V(X_1) = k$
$V(X_2) = 2$
$V(Y) = 25$
$Y = 3X_1 - X_2$

$$\therefore V(3X_1 - X_2) = 25$$

$$\therefore 3^2\, V(X_1) + (-1)^2\, V(X_2) = 25$$

$$\therefore 9\, V(X_1) + V(X_2) = 25$$

$$\therefore 9k + 2 = 25$$

$$9k = 23$$

$$\therefore k = \frac{23}{9}$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.53:

The probability distribution of a bivariate random variable (X,Y) is given below.

Y \ X	1	2	3	Total
1	0.1	0.1	0.2	0.4
2	0.2	0.3	0.1	0.6
Total	0.3	0.4	0.3	1

Find E(X+Y) and E(XY)

Proof:

$$E(X+Y) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} (x+y)P(x,y)$$

$$E(X+Y) = (1+1)*0.1 + (1+2)*0.2 + (2+1)*0.1 + (2+2)*0.3 + (3+1)*0.2 + (3+2)*0.1$$

$$E(X+Y) = 0.2 + 0.6 + 0.3 + 1.2 + 0.8 + 0.5$$

$$E(X+Y) = 3.6$$

$$E(XY) = \sum_{x=-\infty}^{\infty} \sum_{y=-\infty}^{\infty} (xy)P(x,y)$$

$$E(XY) = (1*1)*0.1 + (1*2)*0.2 + (2*1)*0.1 + (2*2)*0.3 + (3*1)*0.2 + (3*2)*0.1$$

$$E(XY) = 0.1 + 0.4 + 0.2 + 1.2 + 0.6 + 0.6$$

$$E(XY) = 3.1$$

◈◈◈◈◈◈◈

MUMBAI EDUCATIONAL TRUST
MET Institute of Computer Science

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.54:

Consider discrete random variables X and Y with the joint pmf as follows

X \ Y	-1	0	3
-2	1/16	1/16	1/16
-1	1/8	1/16	1/8
1	1/8	1/16	1/8
2	1/16	1/16	1/16

Are X and Y independent? Are they uncorrelated?

Proof:

Marginal pmf's are calculated as follows

X \ Y	-1	0	3	P_X
-2	1/16	1/16	1/16	3/16
-1	1/8	1/16	1/8	5/16
1	1/8	1/16	1/8	5/16
2	1/16	1/16	1/16	3/16
P_Y	6/16	4/16	6/16	1

Marginal pmf

X	-2	-1	1	2
P_X	3/16	5/16	5/16	3/16

And

Y	-1	0	1
P_Y	6/16	4/16	6/16

Computing the expectations

$$E(X) = \sum xP(X)$$

$$E(X) = (-2) * \frac{3}{16} + (-1) * \frac{5}{16} + 1 * \frac{5}{16} + 2 * \frac{3}{16}$$

$$E(X) = -\frac{6}{16} - \frac{5}{16} + \frac{5}{16} + \frac{6}{16}$$

$$E(X) = 0$$

Similarly

$$E(Y) = \sum yP(Y)$$

$$E(Y) = (-1) * \frac{6}{16} + 0 * \frac{4}{16} + 1 * \frac{6}{16}$$

$$E(Y) = -\frac{6}{16} + \frac{6}{16}$$

$$E(Y) = 0$$

$$E(XY) = \sum xyP(X,Y)$$

$$E(XY) = \frac{2}{16} + 0 + \left(-\frac{2}{16}\right) + \frac{1}{8} + 0 + \left(-\frac{1}{8}\right) - \frac{1}{8} + 0 + \frac{1}{8} + \left(-\frac{2}{16}\right) + 0 + \frac{2}{16}$$

$$E(XY) = 0$$

This implies that E(XY) = E(X).E(Y)
Hence X and Y are un-correlated.

However P(-2,-1) = 1/16

$P_x(-2)\ P_y(-1) = 3/16 * 3/8$

$P(-2, -1) \neq P_x(-2)\ P_y(-1)$

$\therefore$ X and Y are not independent.

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.55:

Two discrete random variables X and Y with the joint pmf as follows

X \ Y	1	2	3
1	2/16	2/16	1/16
2	3/16	2/16	1/16
3	2/16	1/16	2/16

Are X and Y independent? Are they uncorrelated?

Proof:

Marginal pmf's are calculated as follows

X \ Y	1	2	3	P_X
1	2/16	2/16	1/16	5/16
2	3/16	2/16	1/16	6/16
3	2/16	1/16	2/16	5/16
P_Y	7/16	5/16	4/16	1

Marginal pmf

X	1	2	3
P_X	5/16	6/16	5/16

And

Y	1	2	3
P_Y	7/16	5/16	4/16

Computing the expectations

$$E(X) = \sum xP(X)$$

$$E(X) = 1 * \frac{5}{16} + 2 * \frac{6}{16} + 3 * \frac{5}{16}$$

$$E(X) = \frac{5}{16} + \frac{12}{16} + \frac{15}{16}$$

$$E(X) = \frac{32}{16}$$

$$E(X) = 2$$

Similarly

$$E(Y) = \sum yP(Y)$$

$$E(Y) = 1 * \frac{7}{16} + 2 * \frac{5}{16} + 3 * \frac{4}{16}$$

$$E(Y) = \frac{7}{16} + \frac{10}{16} + \frac{12}{16}$$

$$E(Y) = \frac{29}{16}$$

$$E(XY) = \sum xyP(X,Y)$$

$$E(XY) = 1\left(\frac{2}{16}\right) + 2\left(\frac{3}{16}\right) + 3\left(\frac{2}{16}\right) + 2\left(\frac{2}{16}\right) + 4\left(\frac{2}{16}\right) + 6\left(\frac{1}{16}\right) + 3\left(\frac{1}{16}\right) + 6\left(\frac{1}{16}\right) + 9\left(\frac{2}{16}\right)$$

$$E(XY) = \frac{2}{16} + \frac{6}{16} + \frac{6}{16} + \frac{4}{16} + \frac{8}{16} + \frac{6}{16} + \frac{3}{16} + \frac{6}{16} + \frac{18}{16}$$

$$E(XY) = \frac{59}{16}$$

$$\therefore E(XY) \neq E(X)E(Y)$$

Hence X and Y are un-correlated.
However P(1,1) = 2/16

$P_x(1)\, P_y(1) = 5/16 * 7/16$

$P(1,1) \neq P_x(1)\, P_y(1)$

$\therefore$ X and Y are not independent.

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 4.56:
Two discrete random variables X and Y with the joint pmf as follows

X \ Y	0	1	2
-1	1/27	2/27	3/27
0	2/27	3/27	4/27
1	3/27	4/27	5/27

Are X and Y independent? Are they uncorrelated?

Proof:
Marginal pmf's are calculated as follows

X \ Y	0	1	2	P_X
-1	1/27	2/27	3/27	6/27
0	2/27	3/27	4/27	9/27
1	3/27	4/27	5/27	12/27
P_Y	6/27	9/27	12/27	1

Marginal pmf

X	-1	0	1
P_X	6/27	9/27	12/27

And

Y	0	1	2
P_Y	6/27	9/27	12/27

Computing the expectations

$$E(X) = \sum xP(X)$$

$$E(X) = -1 * \frac{6}{27} + 0 * \frac{9}{27} + 1 * \frac{12}{27}$$

$$E(X) = -\frac{6}{27} + \frac{12}{27}$$

$$E(X) = \frac{6}{27}$$

Similarly

$$E(Y) = \sum yP(Y)$$

$$E(Y) = 0 * \frac{6}{27} + 1 * \frac{9}{27} + 2 * \frac{12}{27}$$

$$E(Y) = \frac{9}{27} + \frac{24}{27}$$

$$E(Y) = \frac{33}{27}$$

$$E(XY) = \sum xyP(X,Y)$$

$$E(XY) = -1\left(\frac{2}{27}\right) + 1\left(\frac{4}{27}\right) + (-2)\left(\frac{3}{27}\right) + 2\left(\frac{5}{27}\right)$$

$$E(XY) = -\frac{2}{27} + \frac{4}{27} - \frac{6}{27} + \frac{10}{27}$$

$$E(XY) = \frac{6}{27}$$

$$E(XY) = E(X)E(Y) = \frac{6}{27} * \frac{33}{27}$$

$$\therefore E(XY) \neq E(X)E(Y)$$

Hence X and Y are un-correlated.
However P(-1,0) = 1/27

$P_x(-1)\, P_y(0) = 6/27 * 6/27$

$P(-1,0) \neq P_x(-1)\, P_y(0)$

$\therefore$X and Y are not independent.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.57:

Monthly demand for transistors is known to have the following probability distribution

Demand	1	2	3	4	5	6
Probability	0.10	0.15	0.20	0.25	0.18	0.12

(i) Determine the expected demand of transistor
(ii) Obtain the variance
(iii) Suppose that the cost (C) of producing n transistors is given by the rule C

Proof:

Computing the expectations

$$E(X) = \sum xP(X)$$

$$E(X) = 1 * 0.10 + 2 * 0.15 + 3 * 0.20 + 4 * 0.25 + 5 * 0.18 + 6 * 0.12$$

$$E(X) = 0.10 + 0.30 + 0.60 + 1.00 + 0.90 + 0.72$$

$$E(X) = 3.62$$

$$E(X^2) = \sum x^2 P(X)$$

$$E(X^2) = 1 * 0.10 + 4 * 0.15 + 9 * 0.20 + 16 * 0.25 + 25 * 0.18 + 36 * 0.12$$

$$E(X^2) = 15.32$$

$$Var\ (X) = E(X^2)\ \ -\ [E(X)]^2$$

$$Var\ (X) = 15.32\ \ -\ [3.62]^2$$

$$Var\ (X) = 2.2156$$

Demand	1	2	3	4	5	6
Cost	10500	11000	11500	12000	12500	13000
Probability	0.10	0.15	0.20	0.25	0.18	0.12

$$E(C) = \sum C\ P(X)$$

$$E(C) = 10500 * 0.10 + 11000 * 0.15 + 11500 * 0.20 + 12000 * 0.25 + 12500 * 0.18 + 13000 * 0.12$$

$$E(C) = 11810$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 4.58:

Two random variables X and Y have a joint pdf as

$$f(x,y) = \begin{matrix} 2-x-y & 0 \le x \le 1\ , 0 \le y \le 1 \\ 0 & otherwise \end{matrix}$$

Find

(i) Marginal densities of X and Y

(ii) Conditional densities of X and Y

(iii) Variances of X and Y

(iv) Co-variance between X and Y

Proof:

(i) $f(x) = \int_{Y=0}^{1} f(x,y)dy = \int_0^1 (2-x-y)dy$

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

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

$$f(x) = \begin{cases} \left[\frac{3}{2} - x\right] & 0 \le x \le 1 \\ 0 & otherwise \end{cases}$$

$$f(y) = \int_{x=0}^{1} f(x,y)dx = \int_{0}^{1} (2 - x - y)dx$$

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

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

$$f(y) = \begin{cases} \left[\frac{3}{2} - y\right] & 0 \le y \le 1 \\ 0 & otherwise \end{cases}$$

(ii) Conditional densities of X and Y

$$f(x|y) = \frac{f(x,y)}{f(y)} = \frac{2 - x - y}{\frac{3}{2} - y} = \frac{2(2 - x - y)}{3 - 2y}$$

$$f(y|x) = \frac{f(x,y)}{f(x)} = \frac{2 - x - y}{\frac{3}{2} - x} = \frac{2(2 - x - y)}{3 - 2x}$$

(iii) Variances of X and Y

$$E(X^2) = \int_{0}^{1} x^2 f(x)dx$$

$$E(X^2) = \int_0^1 x^2\left(\frac{3}{2} - x\right)dx$$

$$E(X^2) = \int_0^1 \frac{3}{2}x^2 - x^3\,dx$$

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

$$E(X^2) = \frac{1}{2} - \frac{1}{4}$$

$$E(X^2) = \frac{1}{4}$$

$$E(X) = \int_0^1 xf(x)dx$$

$$E(X) = \int_0^1 x\left(\frac{3}{2} - x\right)dx$$

$$E(X) = \int_0^1 \frac{3}{2}x - x^2\,dx$$

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

$$E(X) = \frac{3}{4} - \frac{1}{3}$$

$$E(X) = \frac{5}{12}$$

$$Var\ (X) = E(X^2)\ -\ [E(X)]^2$$

$$Var\ (X) = \frac{1}{4}\ -\ \left[\frac{5}{12}\right]^2$$

$$Var\ (X) = \frac{1}{4}\ - \frac{25}{144}$$

$$Var\ (X) = \frac{11}{144}$$

Similarly

$Var\ (Y) = \frac{11}{144}$ $\qquad\qquad$ $E(Y) = \frac{5}{12}$

(iv) Covariance between X and Y

$$E(XY) = \int_{x=0}^{1}\int_{y=0}^{1} xyf(x,y)\,dx\,dy$$

$$E(XY) = \int_{x=0}^{1}\int_{y=0}^{1} xy(2-x-y)\,dx\,dy$$

$$E(XY) = \int_{x=0}^{1}\int_{y=0}^{1} (2xy - x^2y - xy^2)\,dx\,dy$$

$$E(XY) = \int_{x=0}^{1}\left[\frac{2xy^2}{2} - \frac{x^2y^2}{2} - \frac{xy^3}{3}\right]_0^1 dx$$

$$E(XY) = \int_{x=0}^{1}\left[x - \frac{x^2}{2} - \frac{x}{3}\right] dx$$

$$E(XY) = \left[\frac{x^2}{2} - \frac{x^3}{6} - \frac{x^2}{6}\right]_0^1$$

$$E(XY) = \left[\frac{1}{2} - \frac{1}{6} - \frac{1}{6}\right]$$

$$E(XY) = \frac{1}{6}$$

$$COV(X,Y) = E(XY) - E(X).E(Y)$$

$$COV(X,Y) = \frac{1}{6} - \frac{5}{12}.\frac{5}{12}$$

$$COV(X,Y) = \frac{1}{6} - \frac{25}{144}$$

$$COV(X,Y) = -\frac{1}{144}$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 5.1:
A die is rolled 3 times. What is the probability of
(a) No fives turning up?
(b) 1 five?
(c) 3 fives?

Solution:
This is a binomial distribution because there are only 2 possible outcomes (we get a 5 or we don't know, n=3 for each part.)
Let X= number of fives appearing

(a) Here, X=0
$P(X = 0) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$
$P(X = 0) = {}^{3}C_{0}\ (1/6)^{0}\ (5/6)^{3} = 125/216 = 0.5787$

(b) Here , X=1
$P(X = 1) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$
$P(X = 1) = {}^{3}C_{1}\ (1/6)^{1}\ (5/6)^{2} = 75/216 = 0.3472$

(c) Here, X=3
$P(X = 3) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$
$P(X = 3) = {}^{3}C_{3}\ (1/6)^{3}\ (5/6)^{0} = 1/216 = 0.0046$

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 5.2:

Hospital records show that of patients suffering from a certain disease, 75% die of it. What is the probability that of 6 randomly selected patients, 4 will recover?

Solution:

This is a binomial distribution because there are only 2 possible outcomes (The patient die or does not die)
Let X= number who recover
Here n=6, and x=4

Let p=25% (success, i.e. they live)
q=75% (failure, i.e. they die)

The probability that 4 will recover

$P(X = 4) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$
$P(X = 4) = {}^{6}C_{4}\ (0.25)^{4}\ (0.75)^{2}$
$P(X = 4) = 15 * 0.0039 * 0.5625$
$P(X = 4) = 0.0329$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.3:

In the old days, there was a probability of 0.8 of success in any attempt to make a telephone call.
(This often depended on the importance of the person making the call, or the operator's curiosity!)
Calculate the probability of having 7 successes in 10 attempts.

Solution:

The probability of success p =0.8 and q=0.2
n=10
X= success in getting through

The probability of 7 successes in 10 attempts

$P(X = 7) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$
$P(X = 7) = {}^{10}C_{7}\, (0.8)^{7}\, (0.2)^{3}$
$P(X = 7) = 120 * 0.2098 * 0.008$
$P(X = 7) = 0.2014$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.4:

A manufacturer of metal pistons finds that on the average, 12% of his pistons are rejected because they are either oversize or undersize. What is the probability that a batch of 10 pistons will contain

(a) no more than 2 rejects?

(b) at least 2 rejects?

Solution:

Let X= number of rejected pistons

In this case, "success" means rejection!

Here n=10, p=0.12 , q=0.88

(a) No rejects

$P(X = 0) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$

$P(X = 0) = {}^{10}C_{0}\ (0.12)^{0}\ (0.88)^{10}$

$P(X = 0) = 1 * 1 * 0.2785$

$P(X = 0) = 0.2785$

(b) One reject

$P(X = 1) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$

$P(X = 1) = {}^{10}C_{1}\ (0.12)^{1}\ (0.88)^{9}$

$P(X = 1) = 10 * 0.12 * 0.3165$

$P(X = 1) = 0.3798$

(c) Two rejects

$P(X = 2) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$

$P(X = 2) = {}^{10}C_{2}\ (0.12)^{2}\ (0.88)^{8}$

$P(X = 2) = 45 * 0.0144 * 0.3596$

$P(X = 2) = 0.2330$

So the probability of getting no more than 2 rejects is

Probability $= P(X \leq 2)$

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

$= 0.2785 + 0.3798 + 0.2330$

$P(X \leq 2) = 0.8913$

Probability of at least 2 rejects = $1 - P(X \leq 1)$

$= 1 - [\ P(X=0) + P(X=1)]$

$= 1 - (0.2785 + 0.3798)$

$= 1 - 0.6583$

Probability of at least 2 rejects = 0.3417

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.5:

A fair coin is tossed 7 times.

Find the probabilities of obtaining various numbers of heads.

Solution:

A fair coin is tossed 7 times.

p= probability of appearing head = ½

q= 1 – p = 1 – ½ = ½

Let X= number of heads

(a) No head

$P(X = 0) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$

$P(X = 0) = {}^{7}C_{0}\, (1/2)^{0}\, (1/2)^{7}$

$P(X = 0) = 1 * 1 * 1/128$

$P(X = 0) = \frac{1}{128}$

(b) One head

$P(X = 1) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$

$P(X = 1) = {}^{7}C_{1}\, (1/2)^{1}\, (1/2)^{6}$

$P(X = 1) = 7 * 1/128$

$P(X = 1) = \frac{7}{128}$

(c) Two heads

$P(X=2) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$

$P(X=2) = {}^{7}C_{2}\, (1/2)^{2}\, (1/2)^{5}$

$P(X=2) = 21 * 1/128$

$P(X=2) = \frac{21}{128}$

(d) Three heads

$P(X=3) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$

$P(X=3) = {}^{7}C_{3}\, (1/2)^{3}\, (1/2)^{4}$

$P(X=3) = 35 * 1/128$

$P(X=3) = \frac{35}{128}$

(e) Four heads

$P(X=4) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$

$P(X=4) = {}^{7}C_{4}\, (1/2)^{4}\, (1/2)^{3}$

$P(X=4) = 35 * 1/128$

$P(X=4) = \frac{35}{128}$

(f) Five heads

$P(X=5) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$

$P(X=5) = {}^{7}C_{5}\, (1/2)^{5}\, (1/2)^{2}$

$P(X=5) = 21 * 1/128$

$P(X=5) = \frac{21}{128}$

(g) Six heads

$P(X = 6) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$

$P(X = 6) = {}^{7}C_{6}\ (1/2)^{6}\ (1/2)^{1}$

$P(X = 6) = 7 * 1/128$

$P(X = 6) = \frac{7}{128}$

(h) Seven heads

$P(X = 7) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$

$P(X = 7) = {}^{7}C_{7}\ (1/2)^{7}\ (1/2)^{0}$

$P(X = 7) = 1 * 1/128$

$P(X = 7) = \frac{1}{128}$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.6:

A quiz consists of 10 multiple-choice questions. Each question has 5 possible answers, only one of which is correct. Pratik plans to guess the answer to each question. Find the probability that Pratik gets

(a) One answer correct
(b) All 10 answers correct

Solution:

n=10,
p=0.2
q=1-p =1 -0.2 = 0.8
Let X= number of correct answers

(a) One correct answer

$P(X = 1) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$
$P(X = 1) = {}^{10}C_{1}\ (0.2)^{1}\ (0.8)^{9}$
$P(X = 1) = 10*0.2*0.1342$
$P(X = 1) = 0.2648$

(b) All correct answers

$P(X = 10) = {}^{n}C_{x}\ p^{x}\ q^{n-x}$
$P(X = 10) = {}^{10}C_{10}\ (0.2)^{10}\ (0.8)^{0}$
$P(X = 10) = 1*0.0000001*1$
$P(X = 10) = 0.0000001$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.7:

The Probability that a baby is born a boy is 0.51 . A mid-wife delivers 10 babies. Find

(i) The Probability that exactly 4 are male

(ii) The Probability that at least 8 are male

Solution:

n=10,

p=0.51

q=1-p =1 -0.51 = 0.49

Let X= number of male babies

(i) Exactly 4 male babies

$P(X = 4) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$

$P(X = 4) = {}^{10}C_{4}\, (0.51)^{4}\, (0.49)^{6}$

$P(X = 4) = 210 * 0.0677 * 0.0138$

$P(X = 4) = 0.1966$

(ii) At least 8 male babies

$P(X \geq 8) = P(X = 8) + P(X = 9) + P(X = 10)$

$P(X \geq 8) = {}^{10}C_{8}(0.51)^{8}\, (0.49)^{2} + {}^{10}C_{9}(0.51)^{9}\, (0.49)^{1} + {}^{10}C_{10}\, (0.51)^{10}\, (0.49)^{0}$

$P(X \geq 8) = 0.0495 + 0.0114 + 0.0012$

$P(X \geq 8) = 0.0621$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 5.8:

In a college 20% of students are girls. In a random sample of 5 students, find the probability that there are at most 2 girls?

Solution:

n=5,
p=0.2
q=1-p =1 -0.2 = 0.8
Let X= number of girls

At most 2 girls = P(X <= 2)

$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$
$P(X \leq 2) = {}^5C_0\,(0.2)^0\,(0.8)^5 + {}^5C_1\,(0.2)^1\,(0.8)^4 + {}^5C_2\,(0.2)^2\,(0.8)^3$
$P(X \leq 2) = 0.3277 + 0.4096 + 0.2048$
$P(X \leq 2) = 0.9421$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.9:

Determine the binomial distribution for which mean is 4 and variance is 3.

Solution:

Let X ~ B(n,p)
given mean=np = 4
And variance npq = 3
Consider

$$\frac{npq}{np} = \frac{3}{4}$$

$$q = \frac{3}{4}$$

$$p = 1 - q = 1 - \frac{3}{4} = \frac{1}{4}$$

As np=4
n* ¼ = 4
We get n=16
The binomial distribution has parameters n=16 and p=1/4 . that is the random variable X~B(16,1/4)

$$P(X = x) = {}^{16}C_x\,(1/4)^x\,(3/4)^{16-x} \qquad x=0,1,2,----,16$$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 5.10:
On a particular river, overflow floods occur once every 100 years on average. Calculate the probability of k = 0, 1, 2, 3, 4, 5, or 6 overflow floods in a 100-year interval, assuming the Poisson model is appropriate.

Solution:
Because the average event rate is one overflow flood per 100 years, $\lambda = 1$

$$P(k\ overflow\ floods\ in\ 100\ years) = \frac{e^{-\lambda}\lambda^{k}}{k!} = \frac{e^{-1}1^{k}}{k!}$$

$$P(k = 0\ overflow\ floods\ in\ 100\ years) = \frac{e^{-1}1^{0}}{0!} = \frac{e^{-1}}{1} = 0.3679$$

$$P(k = 1\ overflow\ floods\ in\ 100\ years) = \frac{e^{-1}1^{1}}{1!} = \frac{e^{-1}}{1} = 0.3679$$

$$P(k = 2\ overflow\ floods\ in\ 100\ years) = \frac{e^{-1}1^{2}}{2!} = \frac{e^{-1}}{2} = 0.1839$$

The table below gives the probability for 0 to 6 overflow floods in a 100-year period.

k	P(k overflow floods in 100 years)
0	0.3679
1	0.3679
2	0.1839
3	0.0613
4	0.0153
5	0.0031
6	0.0005

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.11:

James and colleagues report that the average number of goals in a World Cup soccer match is approximately 2.5 and the Poisson model is appropriate. Calculate the probability of k = 0, 1, 2, 3, 4, 5, 6, 7 goals in a world cup soccer match.

Solution:

Because the average event rate is 2.5 goals per match, λ = 2.5.

$$P(k\ goals\ in\ a\ match) = \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-2.5}2.5^k}{k!}$$

$$P(k = 0\ goals\ in\ a\ match) = \frac{e^{-2.5}2.5^0}{0!} = \frac{e^{-2.5}}{1} = 0.0821$$

$$P(k = 1\ goals\ in\ a\ match) = \frac{e^{-2.5}2.5^1}{1!} = \frac{e^{-2.5} * \ 2.5}{1} = 0.2052$$

$$P(k = 2\ goals\ in\ a\ match) = \frac{e^{-2.5}2.5^2}{2!} = \frac{e^{-25} * 6.25}{2}$$
$$= 0.2565$$

The table below gives the probability for 0 to 7 goals in a match

k	P(k goals in a match)
0	0.0821
1	0.2052
2	0.2565
3	0.2138
4	0.1336
5	0.0668
6	0.0278
7	0.0099

NOTE: $P(X = n) = \frac{\lambda^n}{n!} P(X = 0)$

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 5.12:
The average number of homes sold by the Acme Realty company is 2 homes per day. What is the probability that exactly 3 homes will be sold tomorrow?

Solution:
Because the average event rate is 2 homes are sold per day, λ = 2

$$P(x,\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-2}2^x}{x!}$$

$$P(x=3,\lambda=2) = \frac{e^{-2}2^3}{3!} = \frac{e^{-2} * 8}{6} = 0.1804$$

thus, the probability of selling 3 homes tomorrow is 0.1804

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.13:

The mean and variance of Poisson distribution are given as E(X)=λ and V(X)=λ ?

Solution:

We know by definition

$$\text{Mean} = E(X) = \sum_{x=0}^{\infty} xP(X = x)$$

$$E(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$E(X) = e^{-\lambda} \sum_{x=0}^{\infty} x \frac{\lambda^{x}}{x!}$$

$$E(X) = e^{-\lambda}[\ 0 + \textstyle\sum_{x=1}^{\infty} x \frac{\lambda^{x}}{x!}] \qquad \lambda^{x} = \lambda\lambda^{x-1}$$

$$E(X) = e^{-\lambda}\lambda[\ \textstyle\sum_{x=1}^{\infty} x \frac{\lambda^{x-1}}{x(x-1)!}]$$

$$E(X) = e^{-\lambda}\lambda[\ \textstyle\sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}]$$

$$E(X) = e^{-\lambda}\lambda\left[\ 1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \ \ldots.\right]$$

$$E(X) = e^{-\lambda}\lambda\ e^{\lambda}$$

$$E(X) = \lambda$$

$$E(X^2)\ = \sum_{x=0}^{\infty} x^2 P(X = x)$$

$$E(X^2) = \ \textstyle\sum_{x=0}^{\infty} x^2 P(X = x)$$

$x^2 = x^2 - x + x\ = x(x-1) + x$

$$E(X^2)\ = \sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!}$$

$$E(X^2) = \sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^x}{x!} + \ \lambda$$

Consider

$$\sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^{x}}{x!} = \sum_{x=2}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^{x-2}\lambda^{2}}{x(x-1)(x-2)!}$$

$$\sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^{x}}{x!} = e^{-\lambda}\lambda^{2}\sum_{x=2}^{\infty} x(x-1)\frac{\lambda^{x-2}}{x(x-1)(x-2)!}$$

$$\sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^{x}}{x!} = e^{-\lambda}\lambda^{2}\sum_{x=2}^{\infty}\frac{\lambda^{x-2}}{(x-2)!}$$

$$\sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^{x}}{x!} = e^{-\lambda}\lambda^{2}[\,1+\frac{\lambda}{1!}+\frac{\lambda^{2}}{2!}+ ---]$$

$$\sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^{x}}{x!} = e^{-\lambda}\lambda^{2}e^{\lambda}$$

$$\sum_{x=0}^{\infty} x(x-1)\frac{e^{-\lambda}\lambda^{x}}{x!} = \lambda^{2}$$

$$\therefore E(X^{2}) = \lambda^{2} + \lambda$$

$$varience = E(X^{2}) - [E(X)]^{2}$$

$$varience = \lambda^{2} + \lambda - [\lambda]^{2}$$

$$varience = \lambda$$

Hence proved

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.14:

In a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005, and accidents are independent of each other.

(i) What is the probability that in any given period of 400 days, there will be an accident on one day?

(ii) What is the probability that there are at most three days with an accident?

Solution:

Let the probability of an accident on any given day be p=0.005

Given number of days n= 400

Since n is large and p is small, we can approximate this to a Poisson distribution with mean λ = np = 400 X 0.005 = 2

Now if X is a random variable that number of accidents, then X follows a Poisson distribution with mean λ =2

$$P(x,\lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}\ ,\ \mathrm{x=0,1,2,----}$$

(i) Probability that there is an accident on a given day

$$P(X = 1,2) = \frac{e^{-2}2^{1}}{1!} = 0.271$$

(ii) Probability that there are at most 3 days with an accident is given day

$$P(X \leq 3,2) = \sum_{x=0}^{3} \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$P(X \leq 3,2) = \sum_{x=0}^{3} \frac{e^{-2}2^{x}}{x!}$$

$$P(X \leq 3,2) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

$$P(X \leq 3,2) = \frac{e^{-2}2^{0}}{0!} + \frac{e^{-2}2^{1}}{1!} + \frac{e^{-2}2^{2}}{2!} + \frac{e^{-2}2^{3}}{3!}$$

$$P(X \leq 3,2) = 0.857$$

❖❖❖❖❖❖❖

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.15:

If X is Poisson variate with parameter λ such that P(X=2) = 9P(X=4)+90P(X=6), then find variance of X.

Solution:

Given that

P(X=2) = 9P(X=4) + 90P(X=6)

$$\frac{e^{-\lambda}\lambda^2}{2!} = 9\ \frac{e^{-\lambda}\lambda^4}{4!} + 90\ \frac{e^{-\lambda}\lambda^6}{6!}$$

$$\frac{e^{-\lambda}\lambda^2}{2!} = 9\ \frac{e^{-\lambda}\lambda^4}{4!} + 90\ \frac{e^{-\lambda}\lambda^6}{6!}$$

$$\frac{e^{-\lambda}\lambda^2}{2!} = 9\lambda^4\ [\frac{e^{-\lambda}}{4!} + 10\ \frac{e^{-\lambda}\lambda^2}{6!}]$$

$$\frac{e^{-\lambda}}{2} = 9\lambda^2\ [\frac{e^{-\lambda}}{24} + 10\ \frac{e^{-\lambda}\lambda^2}{720}]$$

$$\frac{e^{-\lambda}}{2} = 9\lambda^2\ [\frac{e^{-\lambda}}{24} + \frac{e^{-\lambda}\lambda^2}{72}]$$

$$1 = 9\lambda^2\ [\frac{1}{12} + \frac{\lambda^2}{36}]$$

$$1 = 9\lambda^2 \left[\frac{3}{36} + \frac{\lambda^2}{36}\right]$$

$$1 = \lambda^2 \left[\frac{3 + \lambda^2}{4}\right]$$

$$4 = 3\lambda^2 + \lambda^4$$

$$\lambda^4 + 3\lambda^2 - 4 = 0$$

$$1\lambda^4 + 0\lambda^3 + 3\lambda^2 + 0\lambda^1 - 4\lambda^0 = 0$$

Factorization method

1	1	0	3	0	-4
		1	1	4	4
	1	1	4	4	$\boxed{0}$

-1	1	0	3	0	-4
		-1	1	-4	4
	1	-1	4	-4	$\boxed{0}$

Hence factors are $(\lambda - 1) = 0$ and $(\lambda + 1) = 0$
Then $\lambda = -1$ or 1
But we know that $\lambda > 0$ hence $\lambda = 1$

OR

$$\lambda^4 + 3\lambda^2 - 4 = 0$$
$$\lambda^4 + 4\lambda^2 - \lambda^2 - 4 = 0$$

$$\lambda^2(\lambda^2 + 4) - 1(\lambda^2 + 4) = 0$$

$$(\lambda^2 - 1)(\lambda^2 + 4) = 0$$

$$(\lambda^2 - 1) = 0$$

Then λ = -1 or 1

$$(\lambda^2 + 4) = 0$$
$$\lambda^2 = -4$$

λ is imaginary number, hence not valid.

We know that $\lambda > 0$ hence $\lambda = 1$

Since in Poisson distribution mean and variance = λ
Variance = 1

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.16:

It is know that 5% of the books bound at a bindery have defective bindings. Find the probability that 2 out of 100 books bound by this bindery will have defective bindings.

Solution:

Let the probability of defective bindings p=5% = 0.05
Number of books bound n= 100
Since n is large and p is small , we can approximate this to a Poisson distribution as mean λ = np = 100*0.05 = 5

If X is a random variable that number of books are defective then X follows a Poisson distribution with mean λ=5

$$P(x, \lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$P(X = x, 5) = \frac{e^{-5}5^x}{x!}$$

$$P(X = 2,5) = \frac{e^{-5}5^2}{2!}$$

$$P(X = 2, 5) = 0.0842$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.17:

Calculate the probability density function of normal distribution using the following data. x = 3, μ = 4 and σ = 2.

Solution:

Given, variable, x = 3

Mean (μ) = 4 and

Standard deviation (σ) = 2

By the formula of the probability density of normal distribution, we can write;

$$f(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

$$f(3,4,2) = \frac{1}{2\sqrt{2\pi}}\, e^{\frac{-(3-4)^2}{2*2^2}}$$

$$f(3,4,2) = \frac{1}{2\sqrt{2\pi}}\, e^{\frac{-1}{8}}$$

Hence, f(3,4,2) = 0.1760.

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 5.18:

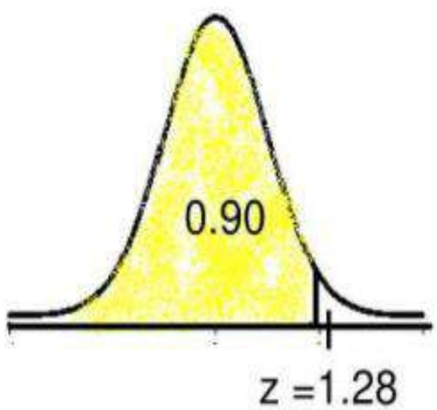
Suppose scores X on a test follow a normal distribution with mean 430 and standard deviation 100.
Find 90^{th} percentile of the scores, that is find score X such that $P(X \le x) = 0.9$

Solution:

Since we start with normal but not standard normal distribution we have to standardize at some point.
$\mu = 430$, $\sigma = 100$

Standard normal distribution $Z = (x-\mu)/\sigma$

$$0.9 = P(X \le x) = P\left(\underbrace{\frac{X-430}{100}}_{Z} \le \frac{x-430}{100}\right) = P(Z < \underbrace{\frac{x-430}{100}}_{z}).$$



get equation: $\frac{x-430}{100} = 1.28$

$x - 430 = 128$

$x = 558$

90% of students scored 558 or less.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.19:

X is a normally distributed variable with mean $\mu = 30$ and standard deviation $\sigma = 4$. Find

a) $P(x < 40)$

b) $P(x > 21)$

c) $P(30 < x < 35)$

Solution:

Given

mean $\mu = 30$ and standard deviation $\sigma = 4$.

Standard normal curve.

$Z=(x-\mu)/\sigma$

a) For $x = 40$, the z-value $z = (40 - 30) / 4 = 2.5$

Hence $P(x < 40) = P(z < 2.5)$

$P(x<40)$ = [area to the left of 2.5] = 0.9938

$P(x < 40) =0.9938$

b) For $x = 21$, $z = (21 - 30) / 4 = -2.25$

Hence $P(x > 21) = P(z > -2.25)$

$P(x > 21)$ = [total area] - [area to the left of -2.25]

$P(x>21) = 1 - 0.0122$

$P(x > 21) =0.9879$

c) For $x = 30$, $z = (30 - 30) / 4 = 0$ and for x = 35, z = (35 - 30) / 4 = 1.25

Hence P(30 < x < 35) = P(0 < z < 1.25)
P(30 < x < 35) = [area to the left of z = 1.25] - [area to the left of 0]
P(30 < x < 35) = 0.8944 - 0.5
P(30 < x < 35) =0.3944

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.20:

A radar unit is used to measure speeds of cars on a motorway. The speeds are normally distributed with a mean of 90 km/hr and a standard deviation of 10 km/hr. What is the probability that a car picked at random is travelling at more than 100 km/hr?

Solution:

Given

Let x be the random variable that represents the speed of cars.

x has $\mu = 90$ and $\sigma = 10$.

We have to find the probability that x is higher than 100 or $P(x > 100)$

For $x = 100$, $z = (100 - 90) / 10 = 1$

$P(x > 100) = P(z > 1)$

$P(x > 100)$ = [total area] - [area to the left of $z = 1$]

$P(x > 100) = 1 - 0.8413$

$P(x > 100) = 0.1587$

The probability that a car selected at a random has a speed greater than 100 km/hr is equal to 0.1587

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.21:

For a certain type of computers, the length of time between charges of the battery is normally distributed with a mean of 50 hours and a standard deviation of 15 hours. John owns one of these computers and wants to know the probability that the length of time will be between 50 and 70 hours.

Solution:

Let x be the random variable that represents the length of time.
It has a mean of 50 and a standard deviation of 15.
We have to find the probability that x is between 50 and 70 or
$P(50 < x < 70)$

For $x = 50$, $z = (50 - 50) / 15 = 0$

For $x = 70$, $z = (70 - 50) / 15 = 1.33$ (rounded to 2 decimal places)

$P(50 < x < 70) = P(0 < z < 1.33)$
$P(50 < x < 70) =$ [area to the left of $z = 1.33$] - [area to the left of $z = 0$]
$P(50 < x < 70) = 0.9082 - 0.5$
$P(50 < x < 70) = 0.4082$

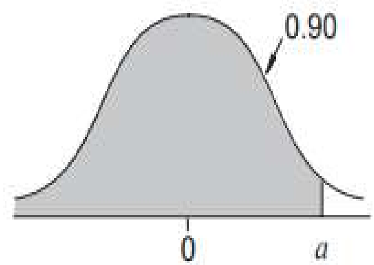
The probability that John's computer has a length of time between 50 and 70 hours is equal to 0.4082.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 5.22

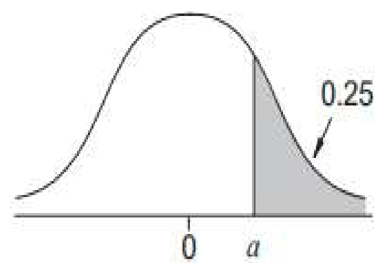


If $Z \sim N(0,1)$, find a such that

(a) $P(Z < a) = 0.90$

(b) $P(Z > a) = 0.25$

Solution



(a) Here $\Phi(a) = 0.90$, and from the tables

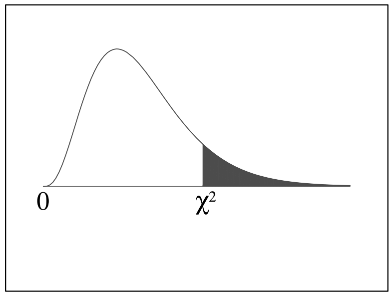
$$a \approx 1.28$$

(b) Here $\Phi(a) = 1 - 0.25 = 0.75$ and from the tables

$$a \approx 0.67$$

◈◈◈◈◈◈◈

Chi-Square Distribution Table



The shaded area is equal to α for $\chi^2 = \chi^2_\alpha$.

df	$\chi^2_{.995}$	$\chi^2_{.990}$	$\chi^2_{.975}$	$\chi^2_{.950}$	$\chi^2_{.900}$	$\chi^2_{.100}$	$\chi^2_{.050}$	$\chi^2_{.025}$	$\chi^2_{.010}$	$\chi^2_{.005}$
1	0.000	0.000	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1.610	9.236	11.070	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688	29.819
14	4.075	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4.601	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578	32.801
16	5.142	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000	34.267
17	5.697	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191	38.582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566	39.997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	46.459	74.397	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	55.329	85.527	90.531	95.023	100.425	104.215
80	51.172	53.540	57.153	60.391	64.278	96.578	101.879	106.629	112.329	116.321
90	59.196	61.754	65.647	69.126	73.291	107.565	113.145	118.136	124.116	128.299
100	67.328	70.065	74.222	77.929	82.358	118.498	124.342	129.561	135.807	140.169

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 6.1:
The specified diameter of a cylindrical part of a machine is 3cm. A sample of 900 such parts shows an average diameter of 2.99 cm with standard deviation of 0.01cm. Does the product differ the specification?
Given: 1% level of significance $Z_\alpha = 2.58$

Solution:

Population mean (μ) = 3 cm

Sample size (n) = 900

Sample mean ($\bar{x}$) = 2.99 cm

Sample standard deviation (S) = 0.01 cm

$H_0 : \mu = 3$ cm

$H_1 : \mu \neq 3$ cm

$$z = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

$$z = \frac{2.99 - 3}{\frac{0.01}{\sqrt{900}}}$$

$$z = -30$$

Since $|Z| > Z_{\alpha}$ hence reject null hypothesis

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 6.2:
Random sample of 100 students gave mean weight of 58 kg with standard deviation of 4 kg. Test the hypothesis that the mean weight in the population is 60 kg. Use 1% level of significance.
Given: 1% level of significance $Z_\alpha = 2.58$

Solution:

Population mean (μ) = 60 kg
Sample size (n) = 100
Sample mean ($\bar{x}$) = 58 kg
Sample standard deviation (S) = 4 kg

$H_0 : \mu = 60$ kg

$H_1 : \mu \neq 60$ kg

$$z = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

$$z = \frac{58 - 60}{\frac{4}{\sqrt{100}}}$$

$$z = -5$$

Since calculated value of Z statistic is more than 2.58, it is significant at 1% level of significance.

Therefore, H_0 is rejected at all levels of significance which implies that mean weight of population is not 60 kg.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.3:

A random sample of 100 students gave a mean weight of 64 kg with a standard deviation of 16 kg. Test the hypothesis that the mean weight in the population is 60 kg.

Given: 5% level of significance $Z_\alpha = 1.96$

Solution:

Population mean (μ) = 60 kg

Sample size (n) = 100

Sample mean ($\bar{x}$) = 64 kg

Sample standard deviation (S) = 16 kg

H_0: μ = 60 kg. , i.e. the mean weight in the population is 60 kg.

H_1: $\mu \neq 60$ kg

$$z = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

$$z = \frac{64 - 60}{\frac{16}{\sqrt{100}}}$$

$$z = 2.5$$

Since calculated value of Z statistic is more than 1.96, it is significant at 5% level of significance.

Therefore, H_0 is rejected at all levels of significance which implies that mean weight of population is not 60 kg.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.4:

A sample of 50 cows in a herd has average lactation yield 1290 litres. Test whether the sample has been drawn from the population having herd average lactation yield of 1350 litres with a standard deviation of 65 litres.

Solution:

Population mean (μ) = 1350 litres

Sample size (n) = 50

Sample mean ($\bar{x}$) = 1290 litres

Sample standard deviation (S) = 65 litres

H_0: $\mu = 1350$

H_1: $\mu \neq 1350$

$$z = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

$$z = \frac{1290 - 1350}{\frac{65}{\sqrt{50}}}$$

$$z = \frac{-60}{7.0711}$$

$$z = -8.48$$

$|Z| = 8.48$

Since calculated value of Z statistic is more than 3, it is significant at all levels of significance.

Therefore, H_0 is rejected at all levels of significance which implies that the sample has not been drawn from the population having mean lactation milk yield as 1350 litres or there is a significant difference between sample mean and population mean.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.5:

The heights of college students in Chennai are normally distributed with standard deviation 6 cm and sample of 100 students had their mean height 158 cm. test the Hypothesis that the mean height of college students in Chennai is 160 cm at 1% level of significance.

Solution:

Population mean (μ) = 160 cm

Sample size (n) = 100

Sample mean ($\bar{x}$) = 158 cm

Sample standard deviation (S) = 6 cm

H_0: μ = 160 i.e., there is no difference between sample mean and hypothetical population mean.

H_1: $\mu \neq 160$

$$z = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

$$z = \frac{158 - 160}{\frac{6}{\sqrt{100}}}$$

$$z = \frac{-2}{0.6}$$

$$z = -3.33$$

$|Z| = 3.33$

at 1% level of significance the value for $Z_\alpha = 2.58$

since $|Z| > Z_\alpha$ Therefore, H_0 is rejected.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.6:

A manufacturer of chocolates claims that the mean weight of a certain box of chocolates is 368 grams. The standard deviation of the box's weight is known to be σ = 10 grams. If a sample of 49 boxes has sample mean $\bar{x}$ = 364 grams, test the hypothesis that the mean weight of the boxes is less than 368 grams. Use α = 0.05 level of significance.

Solution:

Population mean (μ) = 368 grams

Sample size (n) = 49

Sample mean ($\bar{x}$) = 364 grams

Sample standard deviation (σ) = 10 grams

H_0: μ = 368

H_1: μ < 368

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$z = \frac{364 - 368}{\frac{10}{\sqrt{49}}}$$

$$z = -2.8$$

We find the rejection region. Here we use significance level $\alpha = 0.05$, therefore the rejection region is when $z < -1.645$.

Since $|Z| > Z_\alpha$ reject H_0

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 6.7:

A clinical trial with 400 subjects was conducted to test whether the average weight loss using a new diet pill is significant. You know that the standard deviation of the test subject's weight loss was 5 lbs. Give an example of sample mean weight loss which would result in rejecting the null hypothesis at the 5% level (1.64)but not rejecting it at the 1% level(2.34).
Hint: We only care if the subjects lost weight and not gained it, think about what tails are appropriate and start by writing the appropriate hypotheses.

Solution:

Population mean (μ) = 0
Sample size (n) = 400
Sample mean ($\bar{x}$) = ?
Sample standard deviation (S) = 5

H_0: $\mu_{diff} = 0$

H_1: $\mu_{diff} > 0$

In order to reject the hypothesis at the 5% significance level we would need a T-score of 1.65 or higher. Using the following we can calculate the minimum required $\bar{x}$ to reject H_0 at 5% significance level.

$$z = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

$$1.65 = \frac{\bar{x} - 0}{\frac{5}{\sqrt{400}}}$$

$$1.65 * \frac{5}{20} = \bar{x}$$

$$\bar{x} = 0.4125$$

Next, we need to figure out the required T-score for significance at 1% level. Using the following we can calculate the minimum required $\bar{x}$ to reject H0 at 5% significance level.

$$z = \frac{\bar{x} - \mu}{\frac{S}{\sqrt{n}}}$$

$$2.34 = \frac{\bar{x} - 0}{\frac{5}{\sqrt{400}}}$$

$$2.34 * \frac{5}{20} = \bar{x}$$

$$\bar{x} = 0.585$$

So average weight loss between 0.4125 pounds and 0.585 pounds will result in rejecting H_0 at 5% significance level but not at 1% significance level.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.8:

A coin is tossed 130 times. Heads occurred 72 times. Is that coin unbiased? [The value of Z_α at 5% level of significance is 1.96]

Solution:

Hypothesis, the coin is unbiased i.e. P=0.5

H_0: P=0.5 coin is unbiased

H_1: P≠0.5 coin is biased

Given n=130

$$P' = \frac{72}{130} = 0.5538$$

$$\therefore \quad z = \frac{P' - P}{\sqrt{\frac{PQ}{n}}}$$

$$\therefore \quad z = \frac{0.5538 - 0.5}{\sqrt{\frac{0.5 * 0.5}{130}}}$$

$$\therefore \quad z = \frac{0.0538}{0.04385}$$

$$\therefore \quad z = 1.2269$$

Since $|Z| \leq Z_{critical}$, the hypothesis that coin is unbiased is accepted.

i.e. 1.2269 < 1.96. accept null hypothesis

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.9:

In a certain manufacturing process it's known from previous experience that the defectives are 25%. In a lot of 500, 165 are observed to be defective. Is it necessary to revise the hypothesis? The value of Z_α at 5% level of significance is 1.96

Solution:

Hypothesis, the defectives are 25% i.e. P=0.25

H_0: P=0.25

H_1: P≠0.25

Given n=500

$$P' = \frac{165}{500} = 0.33$$

$$z = \frac{P' - P}{\sqrt{\frac{PQ}{n}}}$$

$$\therefore \quad z = \frac{0.33 - 0.25}{\sqrt{\frac{0.25 * 0.75}{500}}}$$

$$\therefore \quad z = \frac{0.08}{\sqrt{0.000375}}$$

$$\therefore \quad z = 4.132$$

Since $|Z| > Z_{critical}$, the null hypothesis is rejected.

i.e. 4.132 > 1.96.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.10:

A manufacturer claims that 20% of the public preferred her product. A sample of 100 persons is taken to check her claim. It is found that 8 of these 100 persons preferred her product.

a. Find the p-value of the test (use a two-tailed test).

b. Using the 0.05 level of significance test her claim

Solution:

We test the following hypothesis:

$H_0 : p = 0.20$

$H_a : p \neq 0.20$

We compute the test statistic z:

Given n=100

$$P' = \frac{8}{100} = 0.08$$

$$z = \frac{P' - P}{\sqrt{\frac{PQ}{n}}}$$

$$\therefore \quad z = \frac{0.08 - 0.20}{\sqrt{\frac{0.20 * 0.80}{100}}}$$

$$\therefore \quad z = \frac{-0.12}{\sqrt{0.0016}}$$

$$\therefore \quad z = \frac{-0.12}{0.04}$$

$$\therefore \quad z = -3$$

Therefore the p−value is:

p − value = 2*P(p' < 0.08) = 2*P(Z < −3.0) = 2*(0.0013) = 0.0026.

We reject H_0 because p−value= 0.0026 < 0.05.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.11:

Principal at school claims that students in his school are above average intelligence and a random sample of 30 students IQ scores have a mean score of 112.5 and mean population IQ is 100 with a standard deviation of 15. Is there sufficient evidence to support the principal claim?

Solution:

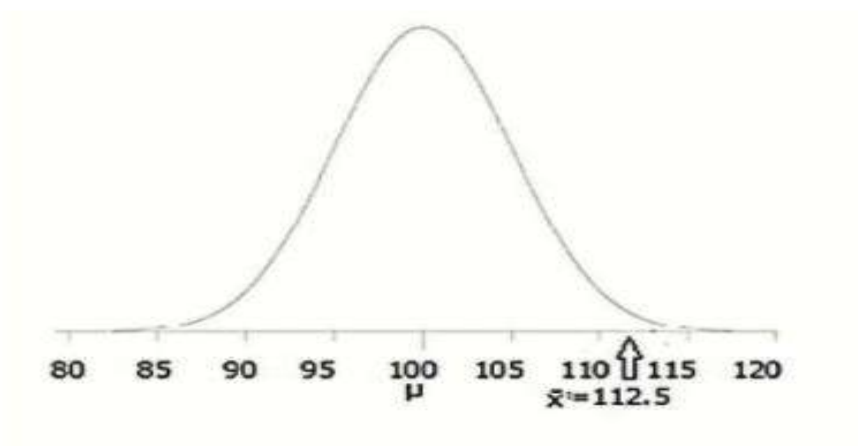
Given

Sample size (n) =30

Sample mean ($\bar{x}$) = 112.5

Population mean (μ) = 100

Population standard deviation (σ) = 15



We test the following hypothesis:

$H_0 : \mu = 100$

$H_a : \mu > 100$

We compute the test statistic z:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$z = \frac{112.5 - 100}{\frac{15}{\sqrt{30}}}$$

$$\therefore \quad z = \frac{12.5}{2.7386}$$

$$\therefore \quad z = 4.5643$$

Lets consider α level 0.05 hence Z_α =1.64

|Z| > 1.64 hence reject null hypothesis

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.12:

Suppose an investor looking to analyze the average daily return of the stock of one the company is greater than 1% or not? So investors picked up a random sample of 50 and return is calculated and has a mean of 0.02 and investors considered standard deviation of mean is 0.025.

Solution:

Given

Sample size (n) =50

Sample mean ($\bar{x}$) = 0.02

Population mean (μ) = 1% = 0.01

Population standard deviation (σ) = 0.025

We test the following hypothesis:

H_0 : μ = 0.01 , the average daily return of the stock is 1%.

H_a : μ > 0.01 , the average daily return of the stock is more than 1%.

We compute the test statistic z:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$z = \frac{0.02 - 0.01}{\frac{0.025}{\sqrt{50}}}$$

$$\therefore \quad z = \frac{0.01}{0.0035}$$

$$\therefore \quad z = 2.8284$$

Lets consider α level 0.05 hence Z_α =1.64

|Z| > 1.64 hence reject null hypothesis

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.13:

An insurance company is currently reviewing its current policy rates when originally settings the rate they believe that the average claim amount will be a maximum of Rs 180000. The company is concern about that true mean actually higher than this. The company randomly selects 40 sample claims and calculate sample mean of Rs 195000 assuming a standard deviation of Claim is Rs 50000 and set alpha as 0.05. Perform z test to see insurance company should be concerned or not.

Solution:

Given

Sample size (n) =40

Sample mean ($\bar{x}$) = 195000

Population mean (μ) = 180000

Population standard deviation (σ) = 50000

We test the following hypothesis:

$H_0 : \mu = 180000$

$H_a : \mu > 180000$, company should be concerned

We compute the test statistic z:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$z = \frac{195000 - 180000}{\frac{50000}{\sqrt{40}}}$$

$$\therefore \quad z = \frac{15000}{7905.69}$$

$$\therefore \quad z = 1.8973$$

Lets consider α level 0.05 hence Z_{α} =1.64

|Z| > 1.64 hence reject null hypothesis

Hence the insurance company should be concerned about their current policies.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.14:

It is claimed that the mean mileage of a certain type of vehicle is 35 miles per gallon of gasoline with population standard deviation σ = 5 miles. What can be concluded using α = 0.01 about the claim if a random sample of 49 such vehicles has sample mean x = 36 miles?

Solution:

Given

Sample size (n) =49

Sample mean ($\bar{x}$) = 36

Population mean (μ) = 35

Population standard deviation (σ) = 5

We test the following hypothesis:

$H_0 : \mu = 35$

$H_a : \mu \neq 35$

We compute the test statistic z:

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$z = \frac{36 - 35}{\frac{5}{\sqrt{49}}}$$

$$\therefore \quad z = \frac{7}{5}$$

$$\therefore \quad z = 1.4$$

Lets consider α level 0.05 hence Z_{α} =1.64

|Z| < 1.64 hence accept the null hypothesis

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 6.15:
A sample of 100 electric fuses produced by a manufacturer A showed mean life time of 1290 days and standard deviation 9 days. A sample of 85 electric fuses produced by a manufacturer B showed life time of 1295 days and standard deviation 15 days.
Is there any difference between mean life time of the two brands at 5% level of significance?
[value of Z_α at 5% level of significance is 1.96]

Solution:
Given

$$n_1 = 100 \qquad \overline{x_1} = 1290 \qquad S_1 = 9$$

$$n_2 = 85 \qquad \overline{x_2} = 1295 \qquad S_2 = 15$$

We test the following hypothesis:
H_0 : There is no difference between the average life time of two brands of electric fuses i.e. $m_1 = m_2$

We compute the test statistic z:

$$Z = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}}$$

$$Z = \frac{1290 - 1295}{\sqrt{\frac{9^2}{100} + \frac{15^2}{85}}}$$

$$\therefore \quad z = \frac{-5}{\sqrt{0.81 + 2.6470}}$$

$$\therefore \quad z = -\frac{5}{1.8593}$$

$$Z = -2.69$$

Since |Z| > Z_α

i.e. |Z| > 1.96 hence reject the null hypothesis at 5% level of significance.

Hence the difference between the average life time of two brands of electric fuses is significant

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.16:

The lengths in cm of 10 nails produced by a certain machine are as :
5.10, 4.98, 5.03, 4.99, 5.00, 5.07, 5.04, 5.03, 4.91, 4.97
Can it be concluded that average length of a nail produce by the machine is 5 cm? [the value of t_α at 5% level of significance for 9 degrees of freedom is 1.833]

Solution:

Population mean μ = 5 cm
Sample size n = 10

X	X^2
5.10	26.0100
4.98	24.8004
5.03	25.3009
4.99	24.9001
5.00	25.0000
5.07	25.7049
5.04	25.4016
5.03	25.3009
4.91	24.1081
4.97	24.7009
$\sum x = 50.12$	$\sum x^2 = 251.2278$

$$Sample\ Mean\ \bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{50.12}{10}$$

$$\bar{x} = 5.012\ cm$$

$$Sample\ standard\ deviation\ (s) = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$S = \sqrt{\frac{251.2278}{10} - (5.012)^2}$$

$$S = 0.05134$$

We test the following hypothesis:

H_0 : Average length of nail produced = 5 cm

H_a: Average length of nail produced ≠ 5 cm

We compute the test statistic t (as n < 30)

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}}$$

$$t = \frac{5.012 - 5}{\frac{0.05134}{\sqrt{10-1}}}$$

$$t = \frac{0.012}{0.01711}$$

$$t = 0.7013$$

Level of significance $\alpha = 0.05$

The critical value at 5% level of significance for 9 degrees of freedom is $t_\alpha = 1.833$

If $|t| \le t_\alpha$ then null hypothesis is accepted otherwise rejected.

Since calculated value of $|t| = 0.7013 < 1.833$

$\therefore$ Accept the null hypothesis

$\therefore$ the average length of nail is 5 cm

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.17:

A certain injection administrated to 12 patients resulted in the following changes of blood pressure 5,2,8,-1,3,0,6,-2, 1, 5, 0, 4 Can it be concluded that the injection will be in general accompanied by an increase in blood pressure ?
[the value of t_α at 5% level of significance for 11 degrees of freedom is 2.201]

Solution:

Population mean $\mu = 0$
Sample size n = 12

X	X^2
5	25
2	4
8	64
-1	1
3	9
0	0
6	36
-2	4
1	1
5	25
0	0
4	16
$\sum x = 31$	$\sum x^2 = 185$

$$Sample\ Mean\ \ \bar{x} = \frac{\sum x}{n}$$

$$\bar{x} = \frac{31}{12}$$

$$\bar{x} = 2.58$$

$$Sample\ standard\ deviation\ (s) = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$S = \sqrt{\frac{185}{12} - (2.58)^2}$$

$$S = 2.96$$

We test the following hypothesis:

H_0 : Change in blood pressure = 0

H_a: Change in blood pressure $\neq$ 0

We compute the test statistic t (as n < 30)

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}}$$

$$t = \frac{2.58 - 0}{\frac{2.96}{\sqrt{12-1}}}$$

$$t = 2.89$$

Level of significance α = 0.05

The critical value at 5% level of significance for 11 degrees of freedom is t_α = 2.201

If $|t| \leq t_\alpha$ then null hypothesis is accepted otherwise rejected.

Since calculated value of $|t|$ =2.89 > 2.201

$\therefore$ Reject the null hypothesis

$\therefore$ the change is blood pressure $\neq 0$

Since
$\bar{x}$ *is positive , the blood pressure increases after injection*

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 6.18:
A random sample of 1000 farms in a year gives an average yield of wheat of 2500 kg per hector with a standard deviation 200 kg. A random sample of 1000 farms in the following year gives an average yield of wheat 2700 kg per hector with standard deviation 250 kg. Can it be inferred that there is a significant increase in the mean yield?
(Given the value of Z_α at 1% level of significance is 2.58)

Solution:
Given

$$n_1 = 1000 \qquad \overline{x_1} = 2500 \qquad S_1 = 200$$

$$n_2 = 1000 \qquad \overline{x_2} = 2700 \qquad S_2 = 250$$

We test the following hypothesis:
H_0 : There is no difference between the average yields in the two years i.e. $m_1 = m_2$

We compute the test statistic z:

$$Z = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{{s_1}^2}{n_1} + \frac{{s_2}^2}{n_2}}}$$

$$Z = \frac{2500 - 2700}{\sqrt{\frac{200^2}{1000} + \frac{250^2}{1000}}}$$

$$\therefore \quad z = \frac{-200}{\sqrt{102.5}}$$

$$Z = -19.75$$

Since |Z| > Z_α

i.e. |Z| > 2.58 hence reject the null hypothesis at 1% level of significance.

Hence there is a significant increase in the mean yield.

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.19:

A random sample of 400 members is found to have mean of 4.45 cm. Can it be reasonably regarded as a sample from a large population where mean is 5 cm and whose variance is 4? Use 5% level of significance, which is 1.96.

Solution:

Sample size (n) = 400
Sample mean ($\bar{x}$) = 4.45
Population mean (μ) = 5
Population variance $\sigma^2 = 4$
$\therefore \sigma = 2$

H_0: $\mu = 5$

H_1: $\mu \neq 5$

Numbers of samples are large and population standard deviation is given, therefore apply z test

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

$$z = \frac{4.45 - 5}{\frac{2}{\sqrt{400}}}$$

$$z = -\frac{0.55}{0.1}$$

$$z = -5.5$$

|Z| > 1.96   i.e. 5.5 > 1.96

Since calculated value of Z statistic is more than 1.96, it is significant at 5% level of significance.

Therefore, Null hypothesis is rejected.

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 6.20:

The following table gives the number of accidents in a city during 10 days of time. Find whether the accidents are uniformly distributed over that period.

Day	1	2	3	4	5	6	7	8	9	10
No. of accidents	8	8	10	9	12	8	10	14	10	11

(given for 9 degrees of freedom at 5% level of significance, the table value of χ^2 is 16.9)

Solution:

Computation for Chi-square test E=N/n =100/10 = 10

Day	O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
1	8	10	-2	4	0.4
2	8	10	-2	4	0.4
3	10	10	0	0	0
4	9	10	-1	1	0.1
5	12	10	2	4	0.4
6	8	10	-2	4	0.4
7	10	10	0	0	0
8	14	10	4	16	1.6
9	10	10	0	0	0
10	11	10	1	1	0.1
	$\sum = 100$				$\sum = 3.4$

Calculate

$$\chi_0{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0{}^2 = 3.4$$

$$Critical\ \chi^2{}_{0.05,\ 9} = 16.9$$

$$Since\ the\ calculate\ value\ Critical\ \chi_0{}^2 \leq \chi^2{}_{0.05,\ 9}$$

$$the\ hypothesis\ that\ accidents\ are\ uniformly\ distributed\ over\ time\ period$$

◈◈◈◈◈◈◈

MUMBAI EDUCATIONAL TRUST

MET Institute of Computer Science

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.21:

The following table gives the number of accidents in a city during a week. Find whether the accidents are uniformly distributed over a week.

Day	Sun	Mon	Tue	Wed	Thu	Fri	Sat
No. of accidents	13	15	9	11	12	10	14

(given for 6 degrees of freedom at 5% level of significance, the table value of χ^2 is 12.59)

Solution:

Computation for Chi-square test ($E_i = N/n = 84/7 = 12$)

Day	O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$
Sun	13	12	1	1
Mon	15	12	-3	9
Tue	9	12	-3	9
Wed	11	12	-1	1
Thu	12	12	0	0
Fri	10	12	-2	4
Sat	14	12	2	4
	$N=\sum O_i = 84$			$\sum = 28$

$$E = \frac{\sum O_i}{n}$$

$$E = \frac{84}{7}$$

$$E = 12$$

$$\chi_0{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0{}^2 = \frac{28}{12}$$

$$\chi_0{}^2 = 2.33$$

$$Critical\ \chi^2{}_{0.05,\ 6} = 12.59$$

$$Since\ the\ calculate\ value\ Critical\ \chi_0{}^2 \leq \chi^2{}_{0.05,6}$$

$$the\ hypothesis\ that\ accidents\ are\ uniformly\ distributed$$

$$over\ a\ week\ is\ accepted$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.22:

The following table gives the number of accidents in a city during a week. Find whether the accidents are uniformly distributed over a week.

Day	Sun	Mon	Tue	Wed	Thu	Fri	Sat
No. of accidents	10	8	11	9	12	10	10

(given for 6 degrees of freedom at 5% level of significance, the table value of χ^2 is 12.59)

Solution:

Computation for Chi-square test $E_i = N/n = 70/7=10$

Day	O_i	E_i	$\lvert O_i - E_i \rvert$	$(O_i - E_i)^2$
Sun	10	10	0	0
Mon	8	10	-2	4
Tue	11	10	1	1
Wed	9	10	-1	1
Thu	12	10	2	4
Fri	10	10	0	0
Sat	10	10	0	0
	$\sum = 70$			$\sum = 10$

$$E = \frac{\sum O_i}{n}$$

$$E = \frac{70}{7}$$

$$E = 10$$

$$\chi_0{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0{}^2 = \frac{10}{10}$$

$$\chi_0{}^2 = 1$$

$$Critical\ \chi^2{}_{0.05,\ 6} = 12.59$$

$$Since\ the\ calculate\ value\ Critical\ \chi_0{}^2 \leq \chi^2{}_{0.05,6}$$

$$the\ hypothesis\ that\ accidents\ are\ uniformly\ distributed$$

$$over\ a\ week\ is\ accepted$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.23:

The following table gives the number of car accidents in a city during a week. Find whether the accidents are uniformly distributed over a week.

Day	Sun	Mon	Tue	Wed	Thu	Fri	Sat
No. of accidents	10	17	11	13	17	14	16

(given for 6 degrees of freedom at 5% level of significance, the table value of χ^2 is 12.59)

Solution:

Computation for Chi-square test

Day	O_i	E_i	$\lvert O_i - E_i \rvert$	$(O_i - E_i)^2$
Sun	10	14	4	16
Mon	17	14	3	9
Tue	11	14	3	9
Wed	13	14	1	1
Thu	17	14	3	9
Fri	14	14	0	0
Sat	16	14	2	4
	$\sum = 98$			$\sum = 48$

$$E = \frac{\sum O_i}{n}$$

$$E = \frac{98}{7}$$

$$E = 14$$

$$\chi_0{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0{}^2 = \frac{48}{14}$$

$$\chi_0{}^2 = 3.4285$$

$$Critical\ \ \chi^2{}_{0.05,\ 6} = 12.59$$

$$Since\ the\ calculate\ value\ Critical\ \ \chi_0{}^2 \le \chi^2{}_{0.05,6}$$

$$the\ hypothesis\ that\ accidents\ are\ uniformly\ distributed$$

$$over\ a\ week\ is\ accepted$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.24:

The following table gives the number of aircraft accidents that occurred during various days of the week. Find whether the accidents are uniformly distributed over a week.

Day	Mon	Tue	Wed	Thu	Fri	Sat
No. of accidents	15	19	13	12	16	15

Test if the accidents are uniformly distributed over the week. (given for 5 degrees of freedom at 5% level of significance, the table value of χ^2 is 11.07)

Solution:

Computation for Chi-square test

Day	O_i	E_i	$\lvert O_i - E_i \rvert$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
Mon	15	15	0	0	0
Tue	19	15	4	16	16/15
Wed	13	15	2	4	4/15
Thu	12	15	3	9	9/15
Fri	16	15	1	1	1/15
Sat	15	15	0	0	0
	$\sum = 90$			$\sum = 30$	$\sum = 2$

$$E = \frac{\sum O_i}{n}$$

$$E = \frac{90}{6}$$

$$E = 15$$

$$\chi_0{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0{}^2 = 2$$

$$Critical\ \ \chi^2{}_{0.05,\ 6} = 11.07$$

$$Since\ the\ calculate\ value\ Critical\ \ \chi_0{}^2 < \chi^2_{0.05,6}$$

$\therefore$ Null hypothesis is accepted.

$$the\ hypothesis\ that\ accidents\ are\ uniformly\ distributed$$

$$over\ a\ week\ is\ accepted$$

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.25:

The following figure shows the distribution of digits in numbers chosen at random from telephone directory.

Digit	0	1	2	3	4	5	6	7	8	9
Frequency	1026	1107	997	966	1075	933	1107	972	964	853

Test whether the digits may be taken to occur equally frequently in the directory. (given for 9 degrees of freedom at 5% level of significance, the table value of χ^2 is 16.92)

Solution:

Computation for Chi-square test $E_i = N/n = 10000/10 = 1000$

Digit	O_i	E_i	$\lvert O_i - E_i \rvert$	$(O_i - E_i)^2$
0	1026	1000	26	676
1	1107	1000	107	11449
2	997	1000	3	9
3	966	1000	34	1156
4	1075	1000	75	5625
5	933	1000	67	4489
6	1107	1000	107	11449
7	972	1000	28	784
8	964	1000	36	1296
9	853	1000	147	21609
	$\sum = 10000$			$\sum = 58542$

$$E = \frac{\sum O_i}{n}$$

$$E = \frac{10000}{10}$$

$$E = 1000$$

Null hypothesis (H_0): Given data is uniformly distributed
Alternate hypothesis (H_a): Given data is not uniformly distributed

$$\chi_0{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0{}^2 = \frac{58542}{1000}$$

$$\chi_0{}^2 = 58.542$$

$$Critical\ \ \chi^2{}_{0.05,9} = 16.92$$

$$Since\ the\ calculate\ value\ Critical\ \ \chi_0{}^2 > \chi^2{}_{0.05,9}$$

$\therefore$ Null hypothesis is rejected.

$$Hence\ the\ digits\ are\ not\ uniformly\ distributed.$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.26:

The observed and expected frequencies in tossing a die 120 times are given below. Test the hypothesis that the die is fair
(Given level of significance =0.01 , 5 degrees of freedom is 15.086)

No. observed	1	2	3	4	5	6
Frequency	17	14	20	17	17	15

At 0.01 level of significance determine whether the die is true (or uniform)

Solution:

Null hypothesis (H_0): The die is uniform
Alternate hypothesis (H_1): the die is not uniform
Total number of trials (N)=100
Number of categories (n)=6

Number	O_i	E_i	$\lvert O_i - E_i \rvert$	$(O_i - E_i)^2$
1	17	16.67	0.33	0.1089
2	14	16.67	2.67	7.1289
3	20	16.67	3.33	11.0889
4	17	16.67	0.33	0.1089
5	17	16.67	0.33	0.1089
6	15	16.67	1.67	2.7889
	$\sum = 100$			$\sum = 21.3334$

$$E = \frac{\sum O_i}{n}$$

$$E = \frac{100}{6}$$

$$E = 16.67$$

$$\chi_0{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0{}^2 = \frac{21.3334}{16.67}$$

$$\chi_0{}^2 = 1.2797$$

$$Critical\ \chi^2{}_{0.05,\ 5} = 15.086$$

$$Since\ the\ calculate\ value\ Critical\ \chi_0{}^2 \leq \chi^2{}_{0.05,5}$$

$\therefore$ Null hypothesis is accepted.

$$Hence\ the\ die\ is\ true\ or\ uniform.$$

◈◈◈◈◈◈◈

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.27:

Calculate the chi-square value for the following data:

	Male (M)	**Female (F)**
Full Stop (FS)	6(observed) 6.24 (expected)	6 (observed) 5.76 (expected)
Rolling Stop (RS)	16 (observed) 16.12 (expected)	15 (observed) 14.88 (expected)
No Stop (NS)	4 (observed) 3.64 (expected)	3 (observed) 3.36 (expected)

Solution:

data	Observed= O_i	Expected = E_i	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
M/FS	6	6.24	-0.24	0.0576	0.009231
F/FS	6	5.76	0.24	0.0576	0.01
M/RS	16	16.12	-0.12	0.0144	0.000893
F/RS	15	14.88	0.12	0.0144	0.000968
M/NS	4	3.64	0.36	0.1296	0.035604
F/NS	3	3.26	-0.26	0.0676	0.020736
	$\sum =$ 50				$\sum =$ 0.077432

Chi-square test formula

$$\chi_0{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0{}^2 = 0.077432$$

Semester: I

MCA11: Mathematical Foundation for Computer Science 1

Example 6.28:

Suppose we wish to determine if there is a relationship between a voter's opinion concerning a new tax reform bill and their level of income. Conduct a chi-square test to determine association. (5% level of significance at 2 degrees of freedom is given, the table value of χ^2 is 5.99)

Income Level

Tax Bill	Low	Medium	High	
For	213 (209.9)	203 (187.2)	182 (200.9)	598
Against	138 (141.1)	110 (125.8)	154 (135.1)	402
	351	313	336	1000

Solution:

data	Observed= O_i	Expected = E_i	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
F/L	213	209.9	3.1	9.61	0.045784
F/M	203	187.2	15.8	249.64	1.333547
F/H	182	200.9	-18.9	357.21	1.778049
A/L	138	141.1	-3.1	9.61	0.068108
A/M	110	125.8	-15.8	249.64	1.98442
A/H	154	135.1	18.9	357.21	2.644041
	$\sum$= 1000				$\sum$= 7.853948

Null hypothesis (H_0) : voter's opinion concerning a new tax reform bill and their level of income are dependent.

Alternative hypothesis(H_a) : voter's opinion concerning a new tax reform bill and their level of income are independent.

Chi-square test formula

$$\chi_0{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0{}^2 = 7.853948$$

$$Critical\ \ \chi^2{}_{0.05,2} = 5.99$$

$$Since\ the\ calculate\ value\ Critical\ \ \chi_0{}^2 > \chi^2_{0.05,2}$$

$\therefore$ Null hypothesis is rejected.

Hence we conclude voter's opinion concerning a new tax reform bill and their level of income are independent.

◈◈◈◈◈◈◈

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Example 6.29:

Chi-Square is one way to show the relationship between two categorical variables. Which pet will you prefer?

	Cat	Dog
Men	207	282
Women	231	242

(1 degree of freedom at 5% level of significance is 3.84)

Solution:

	Cat	Dog	Row Total
Men	207	282	489
Women	231	242	473
Col. total	438	524	**N=962**

$$\text{Expected value} = \frac{\text{row total} * \text{column total}}{\text{N}}$$

data	Observed= O_i	Expected = E_i	$O_i - E_i$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
M/C	207	222.64	-15.64	244.6096	1.098678
M/D	282	266.36	15.64	244.6096	0.918342
W/C	231	215.36	15.64	244.6096	1.135817
W/D	242	257.64	-15.64	244.6096	0.949424
	$\sum$= 962				$\sum$= 4.102261

Null hypothesis (H_0): gender and pet are dependent
Alternate hypothesis (H_a): gender and pet are independent

Chi-square test formula

$$\chi_0{}^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0{}^2 = 4.102261$$

$$Critical\ \ \chi^2{}_{0.05,1} = 3.84$$

$$Since\ the\ calculate\ value\ Critical\ \ \chi_0{}^2 > \chi^2{}_{0.05,1}$$

$\therefore$ Null hypothesis is rejected.

Hence we conclude Gender and pet are independent.

◈◈◈◈◈◈◈

MUMBAI EDUCATIONAL TRUST
MET Institute of Computer Science

Semester: I
MCA11: Mathematical Foundation for Computer Science 1

Q.1 The average height of 16 students is 170cm with a standard deviation of 10cm. Test at $\alpha =$ 5% whether the average height of the population is 172 cm. (The value of t_α at 5% level of significance for 15 degrees of freedom is 2.131)

Q.2 A soap manufacturing company was distributing a particular brand of soap through a large number of retail shops. Before a heavy advertising campaign the mean sales per week per shop was 140 dozens. After the campaign a sample of 26 shops was taken and the mean sale was found to be 147 dozens with standard deviation of 16. Can you consider the advertising is effective?
(The value of t_α at 5% level of significance for 25 degrees of freedom is 2.06)

Q.3 The mean weekly sales of soap bars in independent departmental stores were 146.3 bars per store. After an advertising campaign the mean weekly sales in 22 stores for a typical week increased to 153.7 and showed a standard deviation of 17.2. Was the advertising campaign successful?
(The value of t_α at 1% level of significance for 21 degrees of freedom is 1.721)

Q.4 A random sample of size 16 from a normal population showed a mean of 103.75 cm and sum of squares of deviations from the mean 843.75 cm^2. Can we say that the population has mean of 108.75 cm?
(The value of t_α at 5% level of significance for 15 degrees of freedom is 1.753)

Q.5 The machine is designed to produce insulation washers for electric devices of average thickness of 0.020 cm. A random sample of 10 washers was found to have an average thickness of 0.024 cm with standard deviation of 0.002 cm. Test the significance of the deviation. (The value of t_{α} for 9 degrees of freedom at 5% level of significance is 2.262)

Q.6 The average marks scored by 32 boys are 72 with a standard deviation of 8 while for 36 girls is 70 with a standard deviation of 6. Did the boys perform better than the girls?
(Z value for right tailed test and 1% level of significance is 2.33)

Q.7 The random sample of 900 items is found to have mean of 63.3 cm. Can it be regarded as a sample from a large population whose mean is 66.2 cm and standard deviation of 5.6 cm? Use 5% level of significance.

Q.8 In the population, the average IQ is 100 with a standard deviation of 15. A team of scientists want to test a new medication to see if it has either a positive or negative effect on intelligence or not effect at all. A sample of 30 participants who have taken the medication has a mean of 140. Did the medication affect intelligence?

Q.9 A professor wants to know if her introductory statistics class has a good grasp of basic math. Six students are chosen at random form the call and given a math proficiency test. The professor wants the class to be able to score above 70 on the test. The six students get the following scores: 62, 92, 75, 68, 83, 95.
Can the professor have 90% confidence that the mean score for the class on the test would be above 70.

Q.10 What will be the z value when the given parameters are sample mean = 600, population mean = 585, the standard deviation is 100 and the sample size is 150?

Q.11 Jane has just begun her new job as on the sales force of a very competitive company. In a sample of 16 sales calls it was found that she closed the contract for an average value of 108 dollars with a standard deviation of 12 dollars. Test at 5% significance that the population mean is at least 100 dollars against the alternative that it is less than 100 dollars. Company policy requires that new members of the sales force must exceed an average of $100 per contract during the trial employment period. Can we conclude that Jane has met this requirement at the significance level of 95%?

Q.12 The following table gives the number of bike accidents that occurred during various days of the week. Find whether the accidents are uniformly distributed over a week.

Day	Mon	Tue	Wed	Thu	Fri	Sat
No. of accidents	15	17	10	15	16	14

Test if the bike accidents are uniformly distributed over the week. (Given 5 degrees of freedom at 5% level of significance, χ^2 is 11.07)

Q.13 Which car will you prefer?

	Audi	BMW	Range Rover
Men	55	34	51
Women	67	50	43

(2 degree of freedom at 5% level of significance, χ^2 is 5.99)

◈◈◈◈◈◈◈

Table entry for p and C is the critical value t^* with probability p lying to its right and probability C lying between $-t^*$ and t^*.

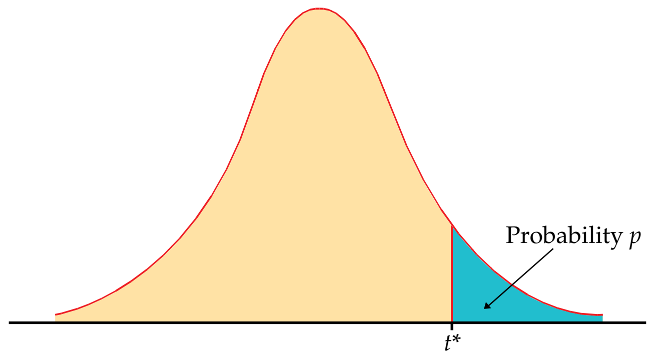


TABLE D

t distribution critical values

	Upper-tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	0.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	0.765	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	0.741	0.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	0.727	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	0.718	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	0.711	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	0.706	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	0.703	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	0.700	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	0.697	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	0.695	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	0.694	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	0.692	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	0.691	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	0.690	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	0.689	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	0.688	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	0.688	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	0.687	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	0.686	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	0.686	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	0.685	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	0.685	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	0.684	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	0.684	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	0.684	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	0.683	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	0.683	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	0.683	0.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	0.681	0.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	0.679	0.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	0.679	0.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	0.678	0.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	0.677	0.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	0.675	0.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
z^*	0.674	0.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

Z Score Table- chart value corresponds to area below z score.

z	0.09	0.08	0.07	0.06	0.05	0.04	0.03	0.02	0.01	0.00
–3.4	0.0002	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
–3.3	0.0003	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005	0.0005
–3.2	0.0005	0.0005	0.0005	0.0006	0.0006	0.0006	0.0006	0.0006	0.0007	0.0007
–3.1	0.0007	0.0007	0.0008	0.0008	0.0008	0.0008	0.0009	0.0009	0.0009	0.0010
–3.0	0.0010	0.0010	0.0011	0.0011	0.0011	0.0012	0.0012	0.0013	0.0013	0.0013
–2.9	0.0014	0.0014	0.0015	0.0015	0.0016	0.0016	0.0017	0.0018	0.0018	0.0019
–2.8	0.0019	0.0020	0.0021	0.0021	0.0022	0.0023	0.0023	0.0024	0.0025	0.0026
–2.7	0.0026	0.0027	0.0028	0.0029	0.0030	0.0031	0.0032	0.0033	0.0034	0.0035
–2.6	0.0036	0.0037	0.0038	0.0039	0.0040	0.0041	0.0043	0.0044	0.0045	0.0047
–2.5	0.0048	0.0049	0.0051	0.0052	0.0054	0.0055	0.0057	0.0059	0.0060	0.0062
–2.4	0.0064	0.0066	0.0068	0.0069	0.0071	0.0073	0.0075	0.0078	0.0080	0.0082
–2.3	0.0084	0.0087	0.0089	0.0091	0.0094	0.0096	0.0099	0.0102	0.0104	0.0107
–2.2	0.0110	0.0113	0.0116	0.0119	0.0122	0.0125	0.0129	0.0132	0.0136	0.0139
–2.1	0.0143	0.0146	0.0150	0.0154	0.0158	0.0162	0.0166	0.0170	0.0174	0.0179
–2.0	0.0183	0.0188	0.0192	0.0197	0.0202	0.0207	0.0212	0.0217	0.0222	0.0228
–1.9	0.0233	0.0239	0.0244	0.0250	0.0256	0.0262	0.0268	0.0274	0.0281	0.0287
–1.8	0.0294	0.0301	0.0307	0.0314	0.0322	0.0329	0.0336	0.0344	0.0351	0.0359
–1.7	0.0367	0.0375	0.0384	0.0392	0.0401	0.0409	0.0418	0.0427	0.0436	0.0446
–1.6	0.0455	0.0465	0.0475	0.0485	0.0495	0.0505	0.0516	0.0526	0.0537	0.0548
–1.5	0.0559	0.0571	0.0582	0.0594	0.0606	0.0618	0.0630	0.0643	0.0655	0.0668
–1.4	0.0681	0.0694	0.0708	0.0721	0.0735	0.0749	0.0764	0.0778	0.0793	0.0808
–1.3	0.0823	0.0838	0.0853	0.0869	0.0885	0.0901	0.0918	0.0934	0.0951	0.0968
–1.2	0.0985	0.1003	0.1020	0.1038	0.1056	0.1075	0.1093	0.1112	0.1131	0.1151
–1.1	0.1170	0.1190	0.1210	0.1230	0.1251	0.1271	0.1292	0.1314	0.1335	0.1357
–1.0	0.1379	0.1401	0.1423	0.1446	0.1469	0.1492	0.1515	0.1539	0.1562	0.1587
–0.9	0.1611	0.1635	0.1660	0.1685	0.1711	0.1736	0.1762	0.1788	0.1814	0.1841
–0.8	0.1867	0.1894	0.1922	0.1949	0.1977	0.2005	0.2033	0.2061	0.2090	0.2119
–0.7	0.2148	0.2177	0.2206	0.2236	0.2266	0.2296	0.2327	0.2358	0.2389	0.2420
–0.6	0.2451	0.2483	0.2514	0.2546	0.2578	0.2611	0.2643	0.2676	0.2709	0.2743
–0.5	0.2776	0.2810	0.2843	0.2877	0.2912	0.2946	0.2981	0.3015	0.3050	0.3085
–0.4	0.3121	0.3156	0.3192	0.3228	0.3264	0.3300	0.3336	0.3372	0.3409	0.3446
–0.3	0.3483	0.3520	0.3557	0.3594	0.3632	0.3669	0.3707	0.3745	0.3783	0.3821
–0.2	0.3859	0.3897	0.3936	0.3974	0.4013	0.4052	0.4090	0.4129	0.4168	0.4207
–0.1	0.4247	0.4286	0.4325	0.4364	0.4404	0.4443	0.4483	0.4522	0.4562	0.4602
–0.0	0.4641	0.4681	0.4721	0.4761	0.4801	0.4840	0.4880	0.4920	0.4960	0.5000

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998