SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF MATHEMATICS

UNIT – I – DISCRIPTIVE STATISTICS – SMTA1102

MEASURES OF CENTRAL TENDENCY : CONCEPTS AND FORMULAE

Mean

The mean (also know as average), is obtained by dividing the sum of observed values by the number of observations, *n*.

That is Mean $= \dfrac{Sum\ of\ all\ observations}{Number\ of\ Observations}$

Median

The median is the middle value of a set of data containing an odd number of values, or the average of the two middle values of a set of data with an even number of values.

Arrange all of the values from lowest to highest. If there are an odd number of entries, the median is the middle value. If there are an even number of entries, the median is the mean of the two middle entries.

Mode

The mode is the most frequently occurring value in the data set. In a data set where each value occurs exactly once, there is no mode.

INDIVIDUAL SERIES	**DISCRETE SERIES**	**CONTINUOUS SERIES**
ARITHMETIC MEAN: Direct Method $\bar{X} = \frac{\sum X}{N}$ Short-cut Method $\bar{X} = A + \frac{\sum d}{N}$ Step-Deviation Method $\bar{X} = A + \frac{\sum d}{N} \times i$	Direct Method $\bar{X} = \frac{\sum f\ X}{N}$ Short-cut Method $\bar{X} = A + \frac{\sum f\ d}{N}$ Step-Deviation Method $\bar{X} = A + \frac{\sum f\ d}{N} \times i$	Direct Method $\bar{X} = \frac{\sum f\ X}{N}$ Short-cut Method $\bar{X} = A + \frac{\sum f\ d}{N}$ Step-Deviation Method $\bar{X} = A + \frac{\sum f\ d}{N} \times i$
MEDIAN: Size of $\left(\frac{N+1}{2}\right)^{th}$ term	Size of $\left(\frac{N+1}{2}\right)^{th}$ term	Size of $\left(\frac{N}{2}\right)^{th}$ term Median = $L + \frac{N/2 - c.f.}{N} \times i$
MODE: Either by inspection or the value that occurs largest number of times	Grouping Method determines that value around which most of the frequencies are concentrated.	Mode = $L + \frac{f_1 - f_0}{2f_1 - (f_0 + f_2)} \times i$

EMPIRICAL RELATION: Mode = 3 Mean – 2 Median		

EXAMPLES

1. Find the mode, median and mean for this set of numbers: 3, 6, 9, 14, 3

Solution

First arrange the numbers from least to greatest: 3, 3, 6, 9, 14

Mode (number seen most often) = 3

3, 3, 6, 9, 14

Median (number exactly in the middle) = 6

3, 3, 6, 9, 14

Mean (add up all the numbers then divide by the amount of numbers) = 7

3 + 3 + 6 + 9 + 14 = 35 35 / 5 = 7

2. Find the mode, median and mean for this set of numbers: 1, 8, 23, 7, 2, 5

Solution

First arrange the numbers from least to greatest: 1, 2, 5, 7, 8, 23

Mode = no mode

Median = 6 (5 + 7 = 12 ; 12 / 2 = 6)

Range = 22 (23 – 1 = 22)

Mean = 7 4/6 (1 + 2 + 5 + 7 + 8 + 23 = 46 ; 46 / 6 = 7 4/6 or 7 2/3. This answer may also be stated in decimals.)

3. From the following data compute Arithmetic Mean

Marks	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50	50 – 60
No. of students	5	10	25	30	20	10

Solution:

Marks	Midvalue X	No. of students f	f x
0 – 10	5	5	25
10 – 20	15	10	150
20 – 30	25	25	625
30 – 40	35	30	1050
40 – 50	45	20	900
50 – 60	55	10	550
		N=100	3300

Arithmetic Mean $\overline{X} = \frac{\sum f\ X}{N} = \frac{3300}{100} = 33$

4. Calculate Arithmetic Mean from the following data

Marks	0 - 10	10 - 30	30 - 60	60 - 100
No. of students	5	12	25	8

Solution:

The class intervals are unequal but still to simplify calculations we can take 5 as common factor.

Marks	Midvalue x	No. of students f	d (x - 45) / 5	f d
0 – 10	5	5	- 8	- 40
10 – 30	15	12	- 5	- 60
30 – 60	25	25	0	0
60 - 100	35	8	7	56
		N= 50		- 44

Arithmetic Mean $\overline{X} = A + \frac{\sum f\ d}{N} \times i$

A = 45, $\sum f\ d$ = - 44, N = 50, I = 5

$$\overline{X} = 45 - \frac{44}{50} \times 5 = 45 - 4.4 = 40.6$$

5. Find the missing frequency from the following data

Marks	0 – 10	10 – 20	20 - 30	30 - 40	40 -50	50 – 60
No. of Students	5	15	20	-	20	10

The arithmetic mean is 34 marks.

Solution:

Let the missing frequency be denoted by X

Marks	Midvalue x	f`	f x
0 – 10	5	5	25
10 – 20	15	15	225
20 – 30	25	20	500
30 – 40	35	X	35X
40 -50	45	20	900
50 – 60	55	10	550
		N = 70 + X	2200+35X

$$\overline{X} = \frac{\Sigma\ f\ x}{N} \qquad 34 = \frac{2200 + 35X}{70 + X}$$

34 (70 + X) = 2200 + 35X

$$2380 + 34X = 2200 + 35X$$

$$35X - 34X = 2380 - 2200$$

$$X = 180$$

6. Calculate the Median and Mode from the following data

Central size	15	25	35	45	55	65	75	85
Frequencies	5	9	13	21	20	15	8	3

Solution:

Since we are given central values first we determine the lower and upper limits of the classes. The class interval is 10 and hence the first class would be 10 – 20.

Class	Midvalue X	f	d (x – 55) / 10	f d	c.f.
10 - 20	15	5	- 4	- 20	5
20 - 30	25	9	- 3	- 27	14
30 - 40	35	13	- 2	- 26	27
40 -50	45	21	- 1	- 21	48
50 - 60	55	20	0	0	68
60 - 70	65	15	1	15	83
70 - 80	75	8	2	16	91
80 - 90	85	3	3	9	94
				Σfd = - 54	

Calculation of Median:

Med = size of $\frac{N}{2}$ th term = $\frac{94}{2}$ = 47

Median lies in the class 40 – 50

Median = $L + \frac{N/2 - c.f.}{f} \times i$

$$M = 40 + \frac{47 - 27}{21} \times 10 = 40 + 9.524 = 49.524$$

7. Calculate the median and mode of the data given below. Using then find arithmetic mean

Marks	0 – 10	10 – 20	20 - 30	30 - 40	40 -50	50 – 60
No. of Students	8	23	45	65	75	80

Solution:

Marks	f`	c.f.
0 – 10	8	8
10 – 20	15	23
20 – 30	22	45

30 – 40	20	65
40 – 50	10	75
50 – 60	5	80
	N = 80	

Calculation of Median: Med = size of $\frac{N}{2}$ th term = $\frac{80}{2}$ = 40 th item

Median lies in the class 20 – 30

$$\text{Median} = L + \frac{N/2 - c.f.}{f} \times i$$

$$M = 20 + \frac{40 - 23}{22} \times 10 = 20 + 7.73 = 27.73$$

Mode lies in the class is 20 – 30

$$\text{Mode} = L + \frac{f_1 - f_0}{2f_1 - (f_0 + f_2)} \times i = 20 + \left(\frac{22 - 15}{44 - (15 + 20)}\right) \times 10 = 27.78$$

MEASURES OF DISPERSION: CONCEPTS AND FORMULAE

Standard deviation

Standard deviation measures the variation or dispersion exists from the mean. A low standard deviation indicates that the data points tend to be very close to the mean, whereas high standard deviation indicates that the data points are spread over a large range of values.

$$\text{Standard Deviation} = \sqrt{\frac{\text{Sum of (value of entry - mean of data set)}^2}{\text{Number of Entries}}}$$

INDIVIDUAL OBERSERVATIONS	**DISCRETE& CONTINUOUS SERIES**
QUARTILE DEVIATION: Q.D. = $\frac{Q_3 - Q_1}{2}$ Coefficient	Quartile Deviation: Q.D. = $\frac{Q_3 - Q_1}{2}$ Coefficient of Q.D. =

of Q.D. = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$	$\frac{Q_3 - Q_1}{Q_3 + Q_1}$
STANDARD DEVIATION: Actual Mean Method: $\sigma = \sqrt{\frac{\sum f(X-\overline{X})^2}{N}}$ Assumed Mean Method: $\sigma = \sqrt{\frac{\sum d^2}{N} - \left(\frac{\sum d}{N}\right)^2}$ Step Deviation Method $\sigma = \sqrt{\frac{\sum d^2}{N} - \left(\frac{\sum d}{N}\right)^2} \times i$ $C.V. = \frac{\sigma}{\overline{X}} \times 100$	Actual Mean Method: $\sigma = \sqrt{\frac{\sum f(X-\overline{X})^2}{N}}$ Assumed Mean Method: $\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}$ Step Deviation Method $\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} \times i$ $C.V. = \frac{\sigma}{\overline{X}} \times 100$

EXAMPLES:

1. Find the Mean and standard deviation from the following distribution

Mid value	12.0	12.5	13.0	13.5	14	14.5	15	15.5	16
No. of Students	2	16	36	60	76	37	18	3	2

Solution:

Midvalue x	No. of Students f	d (x – 14) / 0.5	f d	f d^2
12.0	2	-4	- 8	32
12.5	16	-3	- 48	144
13	36	-2	- 72	144
13.5	60	-1	- 60	60
14	76	0	0	0

14.5	37	1	37	37
15	18	2	36	72
15.5	3	3	9	27
16.0	2	4	8	32
	N= 250		Σfd = - 98	$\Sigma fd^2 = 548$

Mean $\bar{X} = A + \frac{\sum f\,d}{N} \times i = 14 - \frac{98}{250} \times 0.5 = 13.8$

Standard deviation $\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} \times i$

$$\sigma = \sqrt{\frac{548}{250} - \left(\frac{-98}{250}\right)^2} \times .05 \;=\; 0.715$$

2. Find the Standard deviation and Coefficient of Variation from the following data

Marks	No. of students
Up to 10	12
Up to 20	30
Up to 30	65
Up to 40	107
Up to 50	157
Up to 60	202
Up to 70	222
Up to 80	230

Solution:

Class	Midvalue X	No. of Students f	d (x – 35) / 10	f d	f d^2
0 – 10	5	12	-3	- 36	108
10 - 20	15	18	-2	- 36	72
20 - 30	25	35	-1	- 35	35
30 - 40	35	42	0	0	0
40 -50	45	50	1	50	50
50 - 60	55	45	2	90	180
60 - 70	65	20	3	60	180
70 - 80	75	8	4	32	128
		N= 230		Σfd =125	$\Sigma fd^2 = 753$

Mean $\bar{X} = A + \frac{\sum f\,d}{N} \times i = 35 + \frac{125}{230} \times 10 = 40.43$

Standard deviation $\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} \times i$

$$\sigma = \sqrt{\frac{753}{230} - \left(\frac{125}{230}\right)^2} \times 10 \;=\; 17.26$$

$$C.V. = \frac{\sigma}{\overline{X}} \times 100 \quad = \quad \frac{17.26}{40.43} \times 100 = 42.69$$

3. The scores of two batsmen A and B in ten innings during a certain season are:

A	32	28	47	63	71	39	10	60	96	14
B	19	31	48	53	67	90	10	62	40	80

Find which of the two batsmen more consistent in scoring

Solution:

X	$X - \overline{X}$	$\left(X - \overline{X}\right)^2$	Y	$Y - \overline{Y}$	$\left(Y - \overline{Y}\right)^2$
32	-14	196	19	- 31	961
28	-18	324	31	- 19	361
47	1	1	48	- 2	4
63	17	289	53	3	9
71	25	625	67	17	289
39	-7	49	90	40	1600
10	- 36	1296	10	- 40	1600
60	14	196	62	12	144
96	50	2500	40	-10	100
14	- 32	1024	80	30	900
$\sum X = 460$		$\sum\left(X - \overline{X}\right)^2 = 6500$	$\sum Y = 500$		$\sum\left(Y - \overline{Y}\right)^2 = 5968$

Batsman A:

Mean $\overline{X} = \frac{\sum X}{N} = \frac{460}{10} = 46$

$$\sigma = \sqrt{\frac{\sum(X - \overline{X})^2}{N}} = \sqrt{\frac{6500}{10}} = 25.495$$

$$C.V. = \frac{\sigma}{\overline{X}} \times 100 \quad = \quad \frac{25.495}{46} \times 100 = \quad 55.42$$

Batsman B:

Mean $\overline{Y} = \frac{\sum Y}{N} = \frac{500}{10} = 50$

$$\sigma = \sqrt{\frac{\sum\left(Y - \overline{Y}\right)^2}{N}} = \sqrt{\frac{5968}{10}} = \quad 24.43$$

$$C.V. = \frac{\sigma}{\overline{Y}} \times 100 \quad = \quad \frac{24.43}{50} \times 100 = \quad 48.86$$

Since Coefficient of Variation is less in the case of Batsman B, we conclude that the Batsman B is more consistent.

4. Calculate the Quartile deviation and the coefficient of quartile deviation from the following data

Marks	No. of students
Below 20	8
Below 40	20
Below 60	50
Below 80	70
Below 100	80

Solution:

Marks	f`	c.f.
0 - 20	8	8
20 - 40	12	20
40 - 60	30	50
60 - 80	20	70
80 -100	10	80
	N= 80	

Q_1 is the size of N / 4th item.
Q_1 lies in the class 20 – 40

$$Q_1 = L + \frac{N/4 - c.f.}{f} \times i = 20 + \frac{20 - 8}{12} \times 20 = 40$$

Q_3 is the size of 3N / 4th item.
Q_3 lies in the class 60 – 80

$$Q_3 = L + \frac{3N/4 - c.f.}{f} \times i = 60 + \frac{60 - 50}{20} \times 20 = 70$$

$$Q.D. = \frac{Q_3 - Q_1}{2} = \frac{70 - 40}{2} = 15$$

$$Coefficient \ \ of \ \ QD = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{30}{110} = 0.273$$

5. Calculate the Inter-Quartile range and the coefficient of quartile deviation from the following data

Marks	No. of students
Above 0	150
Above 10	140
Above 20	100
Above 30	80
Above 40	80

Above 50	70
Above 60	30
Above 70	14
Above 80	0

Solution:

Marks	f`	c.f.
0 - 10	10	10
10 - 20	40	50
20 - 30	20	70
30 - 40	0	70
40 -50	10	80
50-60	40	120
60-70	16	136
70-80	14	150
	N = 150	

Q_1 is the size of N / 4th item. Q_1 lies in the class 10 – 20

$$Q_1 = L + \frac{N/4 - c.f.}{f} \times i = 10 + \frac{37.5 - 10}{40} \times 10 = 16.875$$

Q_3 is the size of 3N / 4th item. Q_3 lies in the class 50 – 60

$$Q_3 = L + \frac{3N/4 - c.f.}{f} \times i = 50 + \frac{112.5 - 80}{40} \times 10 = 58.25$$

$$Inter\ Quartile\ Range = Q_3 - Q_1 = 41.375$$

$$Coefficient\ of\ QD = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{41.375}{75} = 0.55$$

<u>MOMENTS</u>: FORMULAE

Moments about mean $\mu_1 = \frac{\sum(X - \overline{X})}{N} = 0$ $\mu_2 = \frac{\sum(X - \overline{X})^2}{N}$	$\mu_3 = \frac{\sum(X - \overline{X})^3}{N}$ $\mu_4 = \frac{\sum(X - \overline{X})^4}{N}$
In a Frequency distributiom $\mu_1 = \frac{\sum f\,(X - \overline{X})}{N} = 0$	$\mu_3 = \frac{\sum f\,(X - \overline{X})^3}{N}$

$\mu_2 = \frac{\sum f\,(X - \overline{X})^2}{N}$	$\mu_4 = \frac{\sum f\,(X - \overline{X})^4}{N}$
Moments about arbitrary origin $\mu_1^{'} = \frac{\sum (X - A)}{N}$ $\mu_2^{'} = \frac{\sum (X - A)^2}{N}$	$\mu_3^{'} = \frac{\sum (X - A)^3}{N}$ $\mu_4^{'} = \frac{\sum (X - A)^4}{N}$
In a frequency distribution $\mu_1^{'} = \frac{\sum f\,(X - A)}{N}$ $\mu_2^{'} = \frac{\sum f\,(X - A)^2}{N}$	$\mu_3^{'} = \frac{\sum f\,(X - A)^3}{N}$ $\mu_4^{'} = \frac{\sum f\,(X - A)^4}{N}$
Moments about mean $\mu_2 = \mu_2^{'} - \mu_1^{'2}$, $\mu_3 = \mu_3^{'} - 3\mu_1^{'}\mu_2^{'} + 2\mu_1^{'3}$	$\mu_4 = \mu_4^{'} - 4\mu_1^{'}\mu_3^{'} + 6\mu_2^{'}\left(\mu_1^{'}\right)^2 - 3\left(\mu_1^{'}\right)^4$

SKEWNESS AND KURTOSIS

Karl Pearson's Skewness = Mean - Mode
Bowley's Skewness = $Q_3 + Q_1$- 2 Med
Karl Pearson's coefficient of Skewness $= \frac{\textbf{Mean} - \textbf{Mode}}{\sigma}$
Bowley's coefficient of Skewness $= \frac{Q_3 + Q_1 - 2\,\mathrm{Med}}{Q_3 - Q_1}$
$\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \beta_2 = \frac{\mu_4}{\mu_2^2}$
$\gamma_1 = \frac{\mu_3}{\mu_3^{\frac{3}{2}}}, \gamma_2 = \beta_2 - 3$

1. Calculate the coefficient of skewness by Karl Pearson's method and the values of β_1 and β_2 from the following data

Profits (in lakhs)	No. of companies
10 - 20	18
20 - 30	20
30 - 40	30
40 – 50	22
50 – 60	10

Solution :

Class	Midvalue x	No. of Students F	d (x – 35) / 10	f d	f d^2	f d^3	f d^4
10 - 20	15	18	-2	- 36	72	-144	288
20 - 30	25	20	-1	- 20	35	- 20	20
30 - 40	35	30	0	0	0	0	0
40 -50	45	22	1	22	50	22	22
50 - 60	55	10	2	20	180	80	160
		N= 100		$\Sigma fd = -14$	$\Sigma fd^2 = 154$	$\Sigma fd^3 = -62$	$\Sigma fd^4 = 490$

$$\overline{X} = A + \frac{\sum f\,d}{N} \times i = 35 - \frac{14}{100} \times 10 = 33.6$$

Modal class 30 -40

$$\text{Mode} = L + \frac{f_1 - f_0}{2f_1 - (f_0 + f_2)} \times i = 30 + \frac{30 - 20}{60 - (20 + 22)} = 35.56$$

$$\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2} \times i$$

$$= \sqrt{\frac{154}{100} - \left(\frac{-14}{100}\right)^2} \times 10 = 12.33$$

$$\text{Karl Pearson's coefficient of Skewness} = \frac{\text{Mean} - \text{Mode}}{\sigma}$$

$$= \frac{33.6 - 35.56}{12.33} = -\,0.159$$

$$\mu_1' = \frac{\sum fd}{N} = -0.14$$

$$\mu_2' = \frac{\sum f\,d^2}{N} = 1.54$$

$$\mu_3' = \frac{\sum f\,d^3}{N} = -0.62$$

$$\mu'_4 = \frac{\sum f\, d^4}{N} = 4.9$$

$$\mu_2 = \mu'_2 - \mu'^2_1 = 1.5204\,,$$

$$\mu_3 = \mu'_3 - 3\mu'_1\mu'_2 + 2\mu'^3_1 = 0.0213$$

$$\mu_4 = \mu'_4 - 4\mu'_1\mu'_3 + 6\mu'_2\left(\mu'_1\right)^2 - 3\left(\mu'_1\right)^4 = 4.735$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{0.00045}{3.51458} = 0.000128$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{4.735}{2.312} = 2.048$$

2. By measuring the quartiles find a measure of skewness for the following distribution

Annual Sales	No. of firms
Less than 20	30
Less than 30	225
Less than 40	465
Less than 50	580
Less than 60	634
Less than 70	644
Less than 80	650
Less than 90	665
Less than 100	680

Solution:

Sales	f`	c.f.
10 - 20	30	30
20 - 30	195	225
30 - 40	240	465
40 - 50	115	580
50 - 60	54	634
60 - 70	10	644
70 - 80	6	650
80 - 90	15	665
90 - 100	15	680
	N = 680	

Q_1 lies in the class 20-30

$$Q_1 = L + \frac{N/4 - c.f.}{f} \times i = 20 + \frac{170 - 30}{195} \times 10 = 27.18$$

. Q_3 lies in the class 40 – 50

$$Q_3 = L + \frac{3N/4 - c.f.}{f} \times i = 40 + \frac{510 - 465}{115} \times 10 = 43.9$$

$$\text{Inter Quartile Range} = Q_3 - Q_1 = 41.375$$

$$\text{Coefficient of QD} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{41.375}{75} = 0.55$$

Median class 30 - 40

$$\text{Median} = L + \frac{N/2 - c.f.}{f} \times i$$

$$= 30 + \frac{340 - 225}{240} \times 10 = 34.79$$

$$\text{Bowley's coefficient of Skewness} = \frac{Q_3 + Q_1 - 2\,\text{Med}}{Q_3 - Q_1}$$

$$= \frac{43.9 + 27.18 - 2(34.79)}{43.9 - 27.18} = 0.09$$

3. Calculate the first four moments about the mean from the following data and also calculate the values of β_1 and β_2

Marks	No. of students
0 - 10	5
10 - 20	12
20 - 30	18
30 - 40	40
40 - 50	15
50 - 60	7
60 - 70	3

Solution :

Class	Midvalue x	No. of Students F	d (x – 35) / 10	f d	f d^2	f d^3	f d^4
0 - 10	5	5	-3	-15	45	-135	405
10 - 20	15	12	-2	-24	48	- 96	192
20 - 30	25	18	-1	-18	18	- 18	18
30 - 40	35	40	0	0	0	0	0
40 - 50	45	15	1	15	15	15	15
50 - 60	55	7	2	14	28	56	112
60 - 70	65	3	3	9	27	81	243
		N= 100		$\Sigma f d = 19$	$\Sigma f d^2 = 181$	$\Sigma f d^3 = 97$	$\Sigma f d^4 = 985$

$$\mu_1' = \frac{\sum fd}{N} \times i = -1.9$$

$$\mu_2' = \frac{\sum f\,d^2}{N} \times i^2 = 181$$

$$\mu_3' = \frac{\sum f\,d^3}{N} \times i^3 = -970$$

$$\mu_4' = \frac{\sum f\,d^4}{N} \times i^4 = 98500$$

$$\mu_2 = \mu_2' - \mu_1'^{2} = 177.39,$$

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^{3} = 47.982$$

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_2'\left(\mu_1'\right)^2 - 3\left(\mu_1'\right)^4 = 95009.364$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{2302.27}{5581968.75} = 0.0004$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{95009.364}{31467.212} = 3.02$$

4. The first four moments of a distribution of a distribution about x = 2 are -2, 12, -20 and 100. Calculate the moment about mean. Also calculate β_2 and find whether the distribution is leptokurtic or platykurtic.

Solution:

μ_1'= -2, μ_2'= 12, μ_3'= - 20, μ_4'= 100

$$\mu_2 = \mu_2' - \mu_1'^{2} = 8,$$

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^{3} = 36$$

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_2'\left(\mu_1'\right)^2 - 3\left(\mu_1'\right)^4 = 20$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} == 0.3125$$

Since β_2 is less than 3 the distribution is platykurtic.

<u>EXERCISE PROBLEMS:</u>

1. Define measure of central tendency.
2. State Karl Pearson's coefficient of skewness.
3. Find the mode of: 4, 8, 3, 8, 8, 9, 1, 8, 3.
4. Define Kurtosis and write the measures of kurtosis.

5. Compute the median for the following frequency distribution

Class: 0-9 10-19 20-29 30-39 40-49 50-59 60-69 70-79 80-89	
F : 32 65 100 184 288 167 98 46 20	

6. For a group of 200 candidates, the mean and standard deviation of scores were found to be 40 and 15 respectively. Later on it was discovered that the scores were found to be 43 and 35 were missed as 34 and 53 respectively. Find the correlated mean and standard deviation corresponding to the corrected figures.

7. Calculate the Arithmetic M can of the following frequency distribution:

X	0-10	10-20	20-30	30-40	40-50	50-60
F	12	18	27	20	17	6

8. In ten cricket matches two batsmen A and B scored as follows

A	12	115	6	73	7	19	119	36	84	29
B	47	12	16	42	4	51	37	48	13	0

Who is better scorer and who is more consistent

9. Calculate the coefficient of skewness and kurtosis on the moments for the following distribution

x	4.5	14.5	24.5	34.5	44.5	54.5	64.5	74.5	84.5	94.5
f	1	5	12	22	17	9	4	3	1	1

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF MATHEMATICS

UNIT – II – CORRELATION AND EGRESSION – SMTA1102

Karl Pearson Coefficient of Correlation

As a measure of intensity or degree of linear relationship between two variables, Karl Pearson developed a formula called Correlation coefficient (also called as product moment correlation coefficient).

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

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

$$r(X,Y) = \frac{\dfrac{\sum_{i=1}^{n} x_i y_i}{n} - (\bar{x})(\bar{y})}{\sqrt{\dfrac{\sum_{i=1}^{n} x_i^2}{n} - (\bar{x})^2}\ \sqrt{\dfrac{\sum_{i=1}^{n} y_i^2}{n} - (\bar{y})^2}}$$

$$r(X,Y) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 . \sum_{i=1}^{n}(y_i - \bar{y})^2\right]^{\frac{1}{2}}}$$

The correlation coefficient is a dimensionless number; it has no units of measurement. The maximum value r can achieve is 1, and its minimum value is –1. Therefore, for any given set of observations, $-1 \leq r \leq 1$.

EXAMPLES:

1. Calculate the correlation coefficient between X and Y from the following data:

$$\sum_{i=1}^{15}(X_i - \bar{X})^2 = 136 \quad \sum_{i=1}^{15}(Y_i - \bar{Y})^2 = 138 \qquad \sum_{i=1}^{15}(X_i - \bar{X})(Y_i - \bar{Y}) = 122$$

Solution:

We have $r(X,Y) = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2\right]^{\frac{1}{2}}} = \dfrac{122}{\sqrt{136}\sqrt{138}}$ r(X,Y) = 0.89

Example 2. Some health researchers have reported an inverse relationship between central nervous system malformations and the hardness of the related water supplies. Suppose the data were collected on a sample of 9 geographic areas with the following results:

C.N.S. malformation rate (per 1000 births)	9	8	5	1	4	2	3	6	7
Water hardness(ppm)	120	130	90	150	160	100	140	80	200

Calculate the Correlation Coefficient between the C.N.S. malformation rate and Water hardness.

Solution:

Let us denote the C.N.S. malformation rate by x and water hardness by y. The mean of the x series $\bar{x}$ = 5 and the mean of the y series $\bar{y}$ =130, hence we can use the formula (2.1)

Calculation of correlation coefficient

x	**y**	**(x – $\bar{x}$)= x – 5**	**(y – $\bar{y}$) = y – 130**	**(x – $\bar{x}$)^2**	**(y – $\bar{y}$)^2**	**(x – $\bar{x}$) (y – $\bar{y}$)**
9	120	4	– 10	16	100	– 40
8	130	3	0	9	0	0
5	90	0	– 40	0	1600	0
1	150	– 4	20	16	400	– 80
4	160	– 1	30	1	900	– 30
2	100	– 3	– 30	9	900	90
3	140	– 2	10	4	100	– 20
6	80	1	– 50	1	2500	– 50
7	200	2	70	4	4900	140
				Σ(x – $\bar{x}$)^2= 60	**Σ(y – $\bar{y}$)^2= 11400**	**Σ(x – $\bar{x}$) (y – $\bar{y}$)= 10**

$$r(X,Y)=\frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 \cdot \sum_{i=1}^{n}(y_i-\bar{y})^2\right]^{\frac{1}{2}}} = \frac{10}{[60x11400]^{\frac{1}{2}}}$$

r(X,Y) = 0.012

Therefore, the correlation coefficient between the C.N.S. malformation rate and water hardness is 0.012.

Example 3: Find the product moment correlation for the following data

X	57	62	60	57	65	60	58	62	56
Y	71	70	66	70	69	67	69	63	70

Solution:

X	**Y**	**XY**	**X^2**	**Y^2**
57	71	4047	3249	5041
62	70	4340	3844	4900
60	66	3960	3600	4356
57	70	3990	3249	4900
65	69	4485	4225	4761
60	67	4020	3600	4489
58	69	4002	3364	4761
62	63	3906	3844	3969
56	70	3920	3136	4900
537	**615**	**36670**	**32111**	**42077**

Thus we have, n = 9, $\sum X = 537$, $\sum Y = 615$, $\sum XY = 36670$, $\sum X^2 = 32111$, $\sum Y^2 = 42077$

$$r(X,Y)=\frac{\frac{\sum_{i=1}^{n} x_i y_i}{n}-(\bar{x})(\bar{y})}{\sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}-(\bar{x})^2}\ \sqrt{\frac{\sum_{i=1}^{n} y_i^2}{n}-(\bar{y})^2}} = -0.414$$

Example 4: A computer operator while calculating the coefficient of correlation between two variables X and Y for 25 pairs of observations obtained the following constants: $\sum X = 125$, $\sum Y = 100$, $\sum XY = 508$, $\sum X^2 = 650$, $\sum Y^2 = 460$. However it was later discovered at the time of checking that he had copied two pairs as (6,14) and (8,6) while the correct pairs were (8,12) and (6,8). Obtain the correct correlation coefficient.

Solution:

The formula involved with the given data is, $r(X,Y)=\dfrac{\dfrac{\sum_{i=1}^{n}x_iy_i}{n}-(\bar{x})(\bar{y})}{\sqrt{\dfrac{\sum_{i=1}^{n}x_i^2}{n}-(\bar{x})^2}\sqrt{\dfrac{\sum_{i=1}^{n}y_i^2}{n}-(\bar{y})^2}}$

The Corrected $\sum X$ = Incorrect $\sum X$ – (6+8) + (8+6) = 125

Corrected $\sum Y$ = Incorrect $\sum Y$ – (14+6) + (12+8) = 100

Corrected $\sum X^2$ = Incorrect $\sum X^2$ – $(6^2+8^2) + (8^2+6^2)$ = 650

Corrected $\sum Y^2$ = Incorrect $\sum Y^2$ – $(14^2+6^2) + (12^2+8^2)$ = 436

Corrected $\sum XY$ = Incorrect $\sum XY$ – (84+48) + (96+48) = 520

Now the correct value of correlation coefficient is, $r(X,Y)=\dfrac{\dfrac{520}{25}-(5\times 4)}{\sqrt{\dfrac{650}{25}-5^2}\sqrt{\dfrac{436}{25}-4^2}}$ = 0.67

Partial Correlation Coefficient:

Partial correlation coefficient provides a measure of the relationship between the dependent variable and other variables, with the effect of the most of the variables eliminated.

Let $r_{12.3}$ be the coefficient of partial correlation between X_1 and X_2 keeping X_3 constant, then

$$r_{12.3}=\frac{r_{12}-r_{13}r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}$$

Similarly,

$$r_{13.2}=\frac{r_{13}-r_{12}r_{23}}{\sqrt{(1-r_{12}^2)(1-r_{23}^2)}}$$

where $r_{13.2}$ is the coefficient of partial correlation between X_1 and X_3 keeping X_2 constant.

$$r_{23.1}=\frac{r_{23}-r_{12}r_{13}}{\sqrt{(1-r_{12}^2)(1-r_{13}^2)}}$$

where $r_{23.1}$ is the coefficient of partial correlation between X_2 and X_3 keeping X_1 constant.

Problems:

1. If $r_{12} = 0.8$, $r_{13} = 0.4$ and $r_{23} = 0.56$, find the value of $r_{12.3}$, $r_{13.2}$ and $r_{23.1}$

Solution:

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{\left(1 - r_{13}^2\right)\left(1 - r_{23}^2\right)}}$$

Substituting the given values,

$$r_{12.3} = \frac{0.8 - 0.4 \times 0.56}{\sqrt{1 - (0.4)^2}\sqrt{1 - (0.56)^2}}$$

$= 0.7586$

$$r_{13.2} = \frac{r_{13} - r_{12}r_{23}}{\sqrt{\left(1 - r_{12}^2\right)\left(1 - r_{23}^2\right)}}$$

$$= \frac{0.4 - (0.8)(0.56)}{\sqrt{1 - (0.8)^2}\sqrt{1 - (0.56)^2}}$$

$= -0.0966$

$$r_{23.1} = \frac{r_{23} - r_{12}r_{13}}{\sqrt{\left(1 - r_{12}^2\right)\left(1 - r_{13}^2\right)}}$$

$$= \frac{0.56 - (0.8)(0.4)}{\sqrt{1 - (0.8)^2}\sqrt{1 - (0.4)^2}}$$

$= 0.4364$

2. The correlation between a general intelligence test and school achievement in a group of children from 6 to 15 years old is 0.80. The correlation between the general intelligence test and age in the same group is 0.70 and the correlation between school achievement and age is 0.60. What is the correlation between general intelligence and school achievement in children of the same age?

Solution:

Let X_1 denote general intelligence test.

X_2 denote school achievement.

X_3 denote age.

We are given $r_{12} = 0.8$, $r_{13} = 0.7$ and $r_{23} = 0.6$

We have to find $r_{12.3}$

$$r_{12.3} = \frac{r_{12} - r_{13}r_{23}}{\sqrt{(1-r_{13}^2)(1-r_{23}^2)}}$$

$$= \frac{0.8-(0.7)(0.6)}{\sqrt{1-(0.7)^2}\sqrt{1-(0.6)^2}}$$

$$= 0.6651$$

Multiple Correlation

The coefficient of multiple linear correlation is represented by R, and it is common to add subscripts designating the variable involved. Thus $R_{1.23}$ would represent the coefficient of multiple linear correlation between X_1, on the one hand, and X_2 and X_3 on the other hand. The subscript of the dependent variable is always to the left of the point.

The coefficient of multiplication correlation can be expressed in terms of r_{12}, r_{13} and r_{23} as follows:

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1-r_{23}^2}}$$

$$R_{2.13} = \sqrt{\frac{r_{12}^2 + r_{23}^2 - 2r_{12}r_{13}r_{23}}{1-r_{13}^2}}$$

$$R_{3.12} = \sqrt{\frac{r_{13}^2 + r_{23}^2 - 2r_{12}r_{13}r_{23}}{1-r_{12}^2}}$$

The coefficient of multiple correlation lies between 0 and 1.

1. If $r_{12} = 0.09$, $r_{13} = 0.75$ and $r_{23} = 0.7$, find $R_{1.23}$

Solution:

We have to calculate the multiple correlation coefficient treating first variable as dependent and second and third variables as independent, that is we have to find $R_{1.23}$

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1-r_{23}^2}}$$

Substituting the given values,

$$R_{1.23} = \sqrt{\frac{0.9^2 + 0.75^2 - 2 \times 0.9 \times 0.75 \times 0.7}{1 - 0.7^2}}$$

$= 0.9156$

Spearman's Rank Correlation Coefficient

If X and Y are qualitative variables then Karl Pearson's coefficient of correlation will be meaningless. In this case, we use Spearman's rank correlation coefficient which is defined as follows:

$$\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^{\,2}}{n\left(n^2 - 1\right)}$$ where d is the difference in ranks.

Problems:

1. The ranks of same 16 students in Mathematics and Physics are as follows. The numbers within brackets denote the ranks of the students in Mathematics and Physics. (1,1), (2,10), (3,3), (4,4), (5,5), (6,7), (7,2), (8,6), (9,8), (10,11) (11. 15), (12,9), (13,14), (14,12), (15,16), (16,13). Calculate the rank correlation coefficient for the proficiencies of this group in Mathematics and Physics.

Solution:

Ranks in Maths(X)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Total
Ranks in Physics(Y)	1	10	3	4	5	7	2	6	8	11	15	9	14	12	16	13	
d = X – Y	0	–8	0	0	0	–1	5	2	1	–1	–4	3	–1	2	–1	3	0
d^2	0	64	0	0	0	1	25	4	1	1	16	9	1	4	1	9	136

Spearman's Rank Correlation Coefficient is given by, $\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^{\,2}}{n\left(n^2 - 1\right)} = 1 - \frac{6 \times 136}{16\left(16^2 - 1\right)} =$ 0.8

2. The coefficient of rank correlation between the marks in Statistics and Mathematics obtained by a certain group of students is 2/3 and the sum of the squares of the differences in ranks is 55. Find the number of students in the group.

Solution: Spearman's rank correlation coefficient is given by $\rho = 1 - \frac{6\sum_{i=1}^{n} d_i^{\,2}}{n\left(n^2 - 1\right)}$

Here ρ = 2/3, $\sum d^2 = 55$, N = ? Therefore $\frac{2}{3} = 1 - \frac{6 \times 55}{n(n^2 - 1)}$ Solving the above equation we get n = 10.

Repeated Ranks:

If any two or more individuals are equal in the series then Spearman's formula for calculating the rank correlation coefficients breaks down. In this case, common ranks are given to the repeated ranks. This common rank is the average of the ranks which these items would have assumed if they are slightly different from each other and the next item will get the rank next the ranks already assumed. As a result of this, following adjustment is made in the formula: add the factor $\frac{m(m^2 - 1)}{12}$ to $\sum d^2$ where m is the number of items an item is repeated. This correction factor is to be added for each repeated value.

3. Obtain the rank correlation coefficient for the following data:

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

Solution:

X	**Y**	**Rank X**	**Rank Y**	**D = X – Y**	**D^2**
68	62	4	5	– 1	1
64	58	6	7	– 1	1
75	68	2.5	3.5	– 1	1
50	45	9	10	– 1	1
64	81	6	1	5	25
80	60	1	6	– 5	25
75	68	2.5	3.5	– 1	1
40	48	10	9	1	1
55	50	8	8	0	0
64	70	6	2	4	16
					72

In X series 75 is repeated twice which are in the positions 2nd and 3rd ranks. Therefore common ranks 2.5 (which is the average of 2 and 3) is given for each 75. The corresponding correction factor is $C.F = \frac{2(2^2 - 1)}{12} = \frac{1}{2}$. Also in the X series 64 is repeated thrice which are in the position 5^{th}, 6^{th} and 7^{th} ranks. Therefore, common ranks 6(which is the average of 5, 6 and 7) is given for each 64. The corresponding correction factor is $C.F = \frac{3(3^2-1)}{12} = 2$. Similarly, in the Y series, 68 is repeated twice which are in the positions 3^{rd} and 4^{th} ranks. Therefore, common ranks(which is the average of 3 and 4) is given for each 68. The corresponding correction factor is $C.F = \frac{2(2^2-1)}{12} = \frac{1}{2}$. Rank correlation coefficient is $\rho = 1 - \frac{6\left(\Sigma d^2 + TotalCorrectionFactor\right)}{n\left(n^2-1\right)}$ $= 1 - \frac{6\left(72 + \frac{1}{2} + 2 + \frac{1}{2}\right)}{10(10-1)}$ = 0.5454.

Regression Analysis

Regression analysis helps us to estimate or predict the value of one variable from the given value of another. The known variable(or variables) is called independent variable(s). The variable we are trying to predict is the dependent variable.

Regression equations

Prediction or estimation of most likely values of one variable for specified values of the other is done by using suitable equations involving the two variables. Such equations are known as Regression Equations

Regression equation of y on x:

y – $\bar{y}$ = b_{yx} (x – $\bar{x}$) where y is the dependent variable and x is the independent variable and b_{yx} is given by

$$b_{yx} = \frac{\sum_{i=1}^{n}(x-\bar{x})(y-\bar{y})}{\sum_{i=1}^{n}(x-\bar{x})^2} \quad \text{or} \quad b_{yx} = r\frac{\sigma_y}{\sigma_x} = \frac{\frac{\sum_{i=1}^{n} xy}{n} - (\bar{x}\bar{y})}{\frac{\sum_{i=1}^{n} x^2}{n} - \left(\bar{x}\right)^2}$$

Regression equation of x on y:

x – $\bar{x}$ = b_{xy} (y – $\bar{y}$) where y is the dependent variable and x is the independent variable and b_{yx} is given by

$$b_{xy} = \frac{\sum_{i=1}^{n}(x-\bar{x})(y-\bar{y})}{\sum_{i=1}^{n}(y-\bar{y})^2} \quad \text{or} \quad b_{xy} = r\frac{\sigma_x}{\sigma_y} = \frac{\frac{\sum_{i=1}^{n} xy}{n} - (\bar{x}\bar{y})}{\frac{\sum_{i=1}^{n} y^2}{n} - \left(\bar{y}\right)^2}$$

b_{yx} and b_{xy} are called as regression coefficients of y on x and x on y respectively.

Relation between correlation and regression coefficients:

$b_{yx} = r\frac{\sigma_y}{\sigma_x}$ and $b_{xy} = r\frac{\sigma_x}{\sigma_y}$ $b_{yx} \cdot b_{xy} = r\frac{\sigma_y}{\sigma_x} \; r\frac{\sigma_x}{\sigma_y} = r^2$ Hence

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

Note: In the above expression the components inside the square root is valid only when b_{yx} and b_{xy} have the same sign. Therefore the regression coefficients will have the same sign.

Problems:

1. In trying to evaluate the effectiveness of its advertising campaign a company compiled the following information. Calculate the regression line of sales on advertising.

Year	1980	1981	1982	1983	1984	1985	1986	1987
Advertisement in 1000 rupees	12	15	15	23	24	38	42	48
Sales in lakhs of rupees	5	5.6	5.8	7.0	7.2	88	9.2	9.5

Solution : Let x be advertising amount and y be the sales amount.

Here, n = 8, $\bar{x} = \frac{217}{8} = 27.1$, $\bar{y} = \frac{58.1}{8} = 7.26$

We know that, Regression equation of y on x is given by $y - \bar{y} = b_{yx}(x - \bar{x})$ where

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = \frac{\frac{\sum_{i=1}^{n} xy}{n} - (\bar{x}\bar{y})}{\frac{\sum_{i=1}^{n} x^2}{n} - (\bar{x})^2}$$

X	Y	X^2	XY
12	5	144	60
15	5.6	225	84
15	5.8	225	87
23	7.0	529	161
24	7.2	576	172.8
38	8.8	1444	334.4
42	9.2	1764	386.4
48	9.5	2304	456
217	**58.1**	**7211**	**1741.6**

Therefore $b_{yx} = 0.125$ Substituting this value in the y on x equation, we get,

y – 7.26 = 0.125(x – 27.1) Therefore the required equation of Sales on Advertisement is y = 3.87 + 0.125 x

2. In a study of the effect of a dietary component on plasma lipid composition, the following ratios were obtained on a sample of experimental anumals

Measure of dietary component (X)	1	5	3	2	1	1	7	3
Measure of plasma lipid level (Y)	6	1	0	0	1	2	1	5

(i) obtain the two regression lines and hence predict the ratio of plasma lipid level with 4 dietary component.

(ii) find the correlation coefficient between X and Y

Solution: (i)

X	**Y**	**XY**	**X^2**	**Y^2**
1	6	6	1	36
5	1	5	25	1
3	0	0	9	0
2	0	0	4	0
1	1	1	1	1
1	2	2	1	4
7	1	7	49	1
3	5	15	9	25
23	**16**	**36**	**99**	**68**

Here n = 8 $\bar{x}$ = 2.875 $\bar{y}$ = 2 The Regression equation of y on x is given by y – $\bar{y}$ = b_{yx} (x – $\bar{x}$)

Where

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = \frac{\frac{\sum_{i=1}^{n} xy}{n} - (\bar{x}\bar{y})}{\frac{\sum_{i=1}^{n} x^2}{n} - (\bar{x})^2}$$

$$b_{yx} = -0.304$$

Hence the regression equation of y on x is

$$y - 2 = -0.304(x - 2.875)$$

(i.e) $y = 2.874 - 0.304\,x$

when x = 4 (measure of dietary component) the plasmid lipid level is

$$y = 2.874 - 0.304\,(4)$$

$$y = 1.658$$

The Regression equation of x on y is given by $x - \bar{x} = b_{xy}(y - \bar{y})$

Where $$b_{xy} = r\frac{\sigma_x}{\sigma_y} = \frac{\frac{\sum_{i=1}^{n} xy}{n} - (\bar{x}\bar{y})}{\frac{\sum_{i=1}^{n} y^2}{n} - (\bar{y})^2} \qquad b_{xy} = -0.278$$

Hence the regression equation of x on y is

$$x - 2.875 = -0.278(y - 2)$$

(i.e) $x = 3.431 - 0.278\,y$

(ii) The correlation coefficient between x and y is given by

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

$$r = \pm\sqrt{-0.304 \times -0.278} = \pm 0.291$$

3. From the data given below find (i) two regression lines (ii) coefficient of correlation between marks in Physics and marks in Chemistry (iii) most likely marks in Chemistry when marks in Physics is 78 (iv) most likely marks in Physics when marks in Chemistry is 92

Marks in Physics (X)	72	85	91	85	91	89	84	87	75	77
Marks in Chemistry (Y)	76	92	93	91	93	95	88	91	80	81

Solution:

(i)

X	**Y**	**X^2**	**Y^2**	**XY**
72	76	5184	5776	5472
85	92	7225	8464	7820
91	93	8281	8649	8463

85	91	7225	8281	7735
91	93	8281	8649	8463
89	95	7921	9025	8455
84	88	7056	7744	7395
87	91	7569	8281	7917
75	80	5625	6400	6000
77	81	5929	6561	6237
836	**880**	**70296**	**77830**	**73957**

Here n = 10 $\bar{x}$ = 83.6 $\bar{y}$ = 88

The Regression equation of y on x is given by $y - \bar{y} = b_{yx}(x - \bar{x})$

Where

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = \frac{\frac{\sum_{i=1}^{n} xy}{n} - (\bar{x}\bar{y})}{\frac{\sum_{i=1}^{n} x^2}{n} - (\bar{x})^2}$$

$$b_{yx} = 0.949$$

Hence the regression equation of y on x is

y – 88 = 0.949(x – 83.6)

(i.e) y = 8.6 + 0.949 x

The Regression equation of x on y is given by $x - \bar{x} = b_{xy}(y - \bar{y})$

Where

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} = \frac{\frac{\sum_{i=1}^{n} xy}{n} - (\bar{x}\bar{y})}{\frac{\sum_{i=1}^{n} y^2}{n} - (\bar{y})^2}$$

$$b_{xy} = 0.990$$

Hence the regression equation of x on y is

$$x - 83.6 = 0.990(y - 88)$$

(i.e) $x = -\ 3.5 + 0.990\ y$

(ii) The correlation coefficient between x and y is given by

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

$$r = \pm\sqrt{0.949 \times 0.990} = \pm 0.969$$

(iii) To find the most likely marks in Chemistry when marks in Physics is 78, we have to use the regression equation of y on x given by

$$y = 8.6 + 0.949\ x$$

Substituting the value of x as 78 in the above equation, we get,

$$y = 8.6 + 0.949\ (78)$$

$$y = 73.85$$

Hence the marks in Chemistry is 82.62

(iv) To find the most likely marks in Physics when marks in Chemistry is 92, we have to use the regression equation of x on y given by

$$x = -\ 3.5 + 0.990\ y$$

Substituting the value of y as 92 in the above equation, we get,

$$x = -3.5 + 0.990\ (92)$$

$$x = 87.58$$

Hence the marks in Physics is 87.58

4. For a given series of values, the following data were obtained, $\bar{x}$ =36, $\bar{y}$ =85, σ_x = 11, σ_y = 8 and r = 0.66. Find (i) two regression equations (ii) estimation of x when y = 75.

Solution:

We have $b_{yx} = r\dfrac{\sigma_y}{\sigma_x} = 0.66 \times \dfrac{8}{11} = 0.4799$

and $b_{xy} = r\dfrac{\sigma_x}{\sigma_y} = 0.66 \times \dfrac{11}{8} = 0.9075$

(i) The Regression equation of y on x is given by

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

$$y - 85 = 0.4799\,(x - 36)$$

(i.e.) $$y = -17.28 + 0.4799\,x$$

The Regression equation of x on y is given by

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

$$x - 36 = 0.9075(y - 85)$$

(i.e.) $$x = -41.35 + 0.9075\,y$$

(ii) To estimate the value of x when y = 75, we use the regression line of x on y

$$x = -41.35 + 0.9075\,y$$

Substituting y = 75, $x = -41.35 + 0.9075\,(75)$

Therefore $x = 29.9$

5. For a certain X and Y series which are correlated, the regression lines are $8x - 10y = -66$ and $40x - 18y = 214$. Find (i) the correlation coefficient between them and (ii) the mean of the two series.

Solution:

The given regression equations are

$$8x - 10y = -66 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$40x - 18y = 214 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

(i) Let us suppose that the equation (1) is the equation of line of regression of y on x and (2) as the equation of the line of regression of x on y, after rewriting (1) and (2), we get

$y = \dfrac{66}{10} + \dfrac{8}{10}x$ which gives the value of $b_{yx} = \dfrac{8}{10}$

x = $\frac{214}{40}$ + $\frac{18}{40}$y which gives the value of b_{xy} = $\frac{18}{40}$

Now $r = \pm\sqrt{b_{yx}b_{xy}} = \pm\sqrt{\frac{8}{10}\times\frac{18}{40}} = \pm 0.6$

(ii) Since both the lines of regression passes through the mean values $\bar{x}$ and $\bar{y}$, the point ($\bar{x}$, $\bar{y}$) must satisfy the given two regression lines.

Therefore, 8 $\bar{x}$ – 10 $\bar{y}$ = – 66

40 $\bar{x}$ – 18 $\bar{y}$ = 214

Solving the above two equations we get $\bar{x}$ = 13 and $\bar{y}$ = 17

Important Note: In the above problem in part (i), if we take equation (1) as the line of regression of x on y, we get, x = $-\frac{66}{8}$ + $\frac{10}{18}$y , and hence b_{xy} = $\frac{10}{8}$

and if we take equation (2) as the line of regression of y on x, we get,

y = $-\frac{214}{18}$ + $\frac{40}{18}$x and hence b_{yx} = $\frac{40}{18}$

Therefore, $r = \pm\sqrt{b_{yx}b_{xy}} = \pm\sqrt{\frac{10}{8}\times\frac{40}{18}} = \pm 1.67$

But the value of r cannot exceed unity. Hence the assumptions that line (1) is line of regression of x on y and the line (2) is line of regression of y on x are wrong.

Fitting curves by Method of Least Squares

Curve Fitting: Let (x_i, y_i); i = 1, 2, …, n be a given set of n pairs of values, X being independent variable and Y being the dependent variable. The general problem in curve fitting is to find, if possible, an analytic expression of the form y = f(x), for the functional relationship suggested by the given data.

Fitting a straight line

Let y = a + bx be the equation of the line to be fitted. To estimate the values of a and b we have, the following normal equations.

$$\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^{\,2}$$

Problems:

1. Fit a straight line to the following data:

X	1	2	3	4	6	8
Y	2.4	3	3.6	4	5	6

Solution : Let the straight line to be fitted is y = a + bx

X	**Y**	**XY**	**X^2**
1	2.4	2.0	1
2	3	6.0	4
3	3.6	10.8	9
4	4	16.0	16
6	5	30.0	36
8	6	48.0	64
24	**24**	**113.2**	**130**

Using the normal equations, $\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i$

$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^{\,2}$ we get,

24 = 6a + 24b and

113.2 = 24a + 130b

Solving above two equations, we get

a = 1.976 and b = 0.506

2. Fit a straight line of the form y = a + bx for the following data and estimate the value of y when x is 40

X	2	4	6	10	20	24
Y	6	8	13	12	35	42

Solution: Here n = 6

X	**Y**	**XY**	**X^2**
2	6	12	4
4	8	24	16
6	13	78	36
10	12	120	100
20	35	700	400
24	42	1008	576
66	**116**	**1950**	**1132**

Using the normal equations, $\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^{2}$$ we get,

116 = 6 a + 66 b and

1950 = 66 a + 1132

Solving the above two equations, we get

a = 1.073 and

b = 1.66

Now to estimate the value of y when x is 40, we substitute the value of x in the fitted equation

y = a + b x

(i.e.) y = 1.07 + 1.66 x

= 1.07 + 1.66 x 40

y = 67.47

Fitting a parabola

Let $y = a + bx + cx^2$ be the equation of the line to be fitted. To estimate the values of a and b and c, we have, the following normal equations.

$$\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} x_i^2$$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i^3$$

$$\sum_{i=1}^{n} x_i^2 y_i = a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^4$$

Problems:

1. Fit a parabola to the following data:

X	0	1	2	3	4
Y	1	1.8	1.3	2.5	6.3

Solution: Let $y = a + bx + cx^2$ be the second degree parabola to be fitted, n = 5

X	**Y**	**X^2**	**X^3**	**X^4**	**XY**	**X^2Y**
0	1.0	0	0	0	0	0
1	1.8	1	1	1	1.8	1.8
2	1.3	4	8	16	2.6	5.2
3	2.5	9	27	81	7.5	22.5
4	6.3	16	64	256	25.2	100.8
10	**12.9**	**30**	**100**	**354**	**37.1**	**130.3**

Using normal equations $\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} x_i^2$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^{\ 2} + c\sum_{i=1}^{n} x_i^{\ 3}$$

$$\sum_{i=1}^{n} x_i^{\ 2} y_i = a\sum_{i=1}^{n} x_i^{\ 2} + b\sum_{i=1}^{n} x_i^{\ 3} + c\sum_{i=1}^{n} x_i^{\ 4}$$ we get,

12.9 = 5a + 10b + 30c

37.1 = 10a + 30b + 100c

130.3 = 30a + 100b + 354c

Solving the above equations, we get

a = 1.42

b = – 1.07

c = 0.55.

Thus the required equation of parabola is $y = 1.42 – 1.07x + 0.55x^2$

2. Fit a parabola to the following data and estimate y when x is 6

X	1	3	4	5	7
Y	2	3	6	15	39

Solution: Let $y = a + bx + cx^2$ be the second degree parabola to be fitted, n = 5

X	**Y**	**X^2**	**X^3**	**X^4**	**XY**	**X^2Y**
1	2	1	1	1	2	2
3	3	9	27	81	9	27
4	6	16	64	256	24	96
5	15	25	125	625	75	375
7	39	49	343	2401	273	1911
20	**65**	**100**	**560**	**3364**	**383**	**2411**

Using normal equations $\sum_{i=1}^{n} y_i = na + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} x_i^{\ 2}$

$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i^3$$

$$\sum_{i=1}^{n} x_i^2 y_i = a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i^3 + c\sum_{i=1}^{n} x_i^4$$ we get,

65 = 5a + 20b + 100c

383 = 20a + 100b + 560c

2411 = 100a + 560b + 3364c

Solving the above equations, we get

a = 6.54

b = – 5.93

c = 1.51.

Thus the required equation of parabola is $y = 6.54 – 5.93x + 1.51x^2$

Now to estimate the value of y when x is 6, we substitute the value of x in the fitted equation

$y = a + bx + cx^2$

$= 6.54 – 5.93 \times 6 + 1.51 \times 6^2$

y = 25.32

EXERCISE PROBLEMS:

1. State Spearman's formula to find rank correlation coefficient for repeated ranks.

2. State any two properties of correlation coefficient.

3. For the given bivariate data, (a) Fit a regression line y on x and predict the value of y when x = 5. (b) Fit a regression line x on y and predict the value of x when y = 2.5 (c) Calculate Karl pearson's coefficient of correlation.

x	1	5	3	2	1	1	7	3
y	6	1	0	0	1	2	1	5

4. Ten competitors in a beauty contest are ranked by 3 judges in the following order

I Judge	1	6	5	10	3	2	4	9	7	8
II Judge	3	5	8	4	7	10	2	1	6	9
III Judge	6	4	9	8	1	2	3	10	5	7

Use the rank correlation coefficient to discuss which pair of judges have the Nearest approach to common taste of beauty.

5. Find the rank correlation coefficient for the data given below:-

x	70	75	80	60	70	71	82	83	85
y	60	62	61	70	61	72	75	61	59

6. Write down the distinguishing features of correlation and regression.

7. Find the most likely price in Bombay corresponding to the price of Rs. 70 at Calcutta from the following:

	Calcutta	Bombay
Average price	65	67
Standard deviation	2.5	3.5

Correlation coefficient between the prices of commodities in the two cities is 0.8.

8. Obtain the rank correlation coefficient for the following data:

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

9. Fit a parabola of second degree to the following data:

X :	1	2	3	4	6	8
Y :	2.4	3	3.6	4	5	6

10. In a partially destroyed laboratory record of an analysis of correlation data, the following results only are legible: Variance of X = 9. Regression equations: $8x - 10y + 66 = 0$, $40x - 18y = 214$. What were

(i) the mean values of X and Y .

(ii) the correlation coefficient between X and Y and

(iii) the standard deviation of Y?

11. Calculate the correlation coefficient for the following heights (in inches) of fathers(X) and their sons(Y):

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

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF MATHEMATICS

UNIT – III – PERMUTATIONS AND COMBINATIONS – SMTA1102

PERMUTATION AND COMBINATION

PERMUTATION

A Permutation is an arrangement of set of n objects in a definite order taken some or all at a time.

Example: 1.Three letters a,b,c can be arranged

abc, acb, bac, bca, cab, cba. We have taken all the three for arrangement.

2. Using the three letters a,b,c the total no. of arrangements or permutation taking two at a time.

ab, bc, ac, ba, cb, ca.

The no. of permutation of n objects taken r at a time is denoted by P(n,r) or nP_r and is defined as

$$nP_r = \frac{n!}{(n-r)!} \text{ where } r \le n.$$

Corollary

If r= n,

$$nP_n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$$

Permutation with repetition

Let $P(n; n_1, n_2, \ldots, n_r)$ denote the no. of permutation of n objects of which n_1 are alike n_2 are alike $\ldots n_r$ are alike then the formula is given by

$$P(n; n_1, n_2, n_3, \ldots n_r) = \frac{n!}{n_1! n_2! n_3! \ldots n_r!}$$

Circular Permutation

Arrangement of objects in a circle is called Circular Permutation. A circular Permutation of n different objects is (n-1)!

Solved Problems

1. Find the value of n if $nP_5 = 42nP_3$ where n>4

Solution

$$\frac{n!}{(n-5)!} = 42\frac{n!}{(n-3)!}$$
$$\frac{1}{(n-5)!} = 42\frac{1}{(n-3)(n-4)(n-5)!}$$
$(n-3)(n-4) = 42$
$n^2-7n-30 = 0$
$(n-10)(n+3) = 0$
n = 10, -3
Since n is positive, n = 10

2. How many four digit nos. can be formed by using the digits 1 to 9. If repetition of digits are not allowed.

Solution

$$9P_4 = \frac{9!}{5!} = \frac{9\times8\times7\times6\times5!}{5!}$$

$= 3024.$

3. Find the no. of permutations of the letters of the word ALLAHABAD.

Solution

There are 9 letters in this word. To form different words containing all these 9 letters is

$$= \frac{9!}{4!2!}$$

4. (i) A committee of 3 is to be chosen out of 5 English, 4 French, 3 Indians and the committee to contain one each. In how many ways can this be done? (ii) In how many arrangements one particular Indian can be chosen?

Solution

(i)One English member can be chosen in 5 ways
One French member can be chosen in 4 ways
One Indian member can be chosen in 3 ways
No of ways the committee can be formed = 5x4x3 = 60 ways.

(ii)Since the Indian member is fixed, we have to fill the remaining two places choosing

one from English and French each. This can be done in 5x4 = 20 ways.

5. There are 5 trains from Chennai to Delhi and back to Chennai. In how many ways a person go Chennai to Delhi and return to Chennai.

Solution

5 x 4 = 20.

6. There is a letter lock with three rings, each ring with 5 letters and the password is unknown. How many different useless attempts are made to open the lock.

Solution

Total no. of attempts = 5 x 5 x 5 = 5^3

Only one will unlock , so the total no. of useless attempts is (5^3-1) = 125 -1
= 124.

7. (i) Find the no. of arrangements of the letters of the word ELEVEN,(ii) How many of them begin and end with E. (iii) How many of them have three E's together. (iv) How many begin with E and end with N.

Solution

(i) $\frac{6!}{3!} = 6 \times 5 \times 4 = 120$ ways.

(ii) First and last places are fixed, the remaining 4 places are done in 4! ways.

(iii) Treat the 3 E’s as a single element.

Therefore, this single element along with L,V,N can be arranged in 4! ways.

(iv) $\frac{4!}{2!} = 4 \times 3 = 12.$

8. There are 6 different books on Physics, 3 on Chemistry, 2 on Mathematics. In how many ways can they be arranged on a shelf if the books of the same subject are always together?

Solution

Considering Physics books, Chemistry books, Mathematics books as three elements, three elements can be arranged in 3! ways. Also

Physics books can themselves be arranged in 6! Ways

Chemistry books can themselves be arranged in 3! Ways

Mathematics books can themselves be arranged in 2! Ways

No.of arrangements = 3! 6! 3! 2!

9. Find the no. of arrangements in which 6 boys and 4girls can be arranged in a line such that all the girls sit together and all the boys sit together.

Solution

The no. of arrangement with all the girls sit together and all the boys sit together is

2! 4! 6! ways.

10. Find the no. of ways in which 10 exam papers can be arranged so that 2 particular papers may not come together.

Solution

2 particular papers should not come together. The remaining 8 papers can be arranged in 8! ways.The 2 papers can be filled in 9 gaps in between these 8 papers in $9P_2$ ways.

11. In how many ways can an animal trainer arrange 5 lions and 4 tigers in a row so that no two lions are together?

Solution

The 5 lions should be arranged in the 5 places marked ‘L’.
This can be done in 5! ways.
The 4 tigers should be in the 4 places marked ‘T’.
This can be done in 4! ways.
Therefore, the lions and the tigers can be arranged in 5!*4!= 2880 ways

12. In how many ways 5 boys and 3 girls can be seated in a row so that no two girls are together?

Solution

5 boys can be seated in a row in 5! ways.

Also the girls can be seated in 3! ways

The 3 girls can be filled in the 6 gaps between the boys in $6P_3$ ways.

Total no of arrangements = 5! x 3! x $6P_3$ = 1440

13. There are 4 books on fairy tales, 5 novels and 3 plays. In how many ways can you arrange these so that books on fairy tales are together, novels are together and plays are together and in the order, books on fairy tales, novels and plays.

Solution

There are 4 books on fairy tales and they have to be put together.
They can be arranged in 4! ways.

Similarly, there are 5 novels.
They can be arranged in 5! ways.

And there are 3 plays.
They can be arranged in 3! ways.

So, by the counting principle all of them together can be arranged in 4!*5!*3!= 17280 ways

13. Suppose there are 4 books on fairy tales, 5 novels and 3 plays as in Example 5.3. They have to be arranged so that the books on fairy tales are together, novels are together and plays are together, but we no longer require that they should be in a specific order. In how many ways can this be done?

Solution

First, we consider the books on fairy tales, novels and plays as single objects.
These three objects can be arranged in 3!=6 ways.
Let us fix one of these 6 arrangements.
This may give us a specific order, say, novels -> fairy tales -> plays.

Given this order, the books on the same subject can be arranged as follows.
The 4 books on fairy tales can be arranged among themselves in 4!=24 ways.
The 5 novels can be arranged in 5!=120 ways.
The 3 plays can be arranged in 3!=6 ways.

For a given order, the books can be arranged in 24*120*6=17280 ways.
Therefore, for all the 6 possible orders the books can be arranged in 6*17280= 103680 ways.

14. In how many ways can 4 girls and 5 boys be arranged in a row so that all the four girls are together?

Solution

Let 4 girls be one unit and now there are 6 units in all.
They can be arranged in 6! ways.

In each of these arrangements 4 girls can be arranged in 4! ways.
=> Total number of arrangements in which girls are always together
=6!*4!=720*24= 17280.

15. How many arrangements of the letters of the word 'BENGALI' can be made

(i) If the vowels are never together.
(ii) If the vowels are to occupy only odd places.

Solution

There are 7 letters in the word 'Bengali; of these 3 are vowels and 4 consonants.

(i) Considering vowels a, e, i as one letter, we can arrange 4+1 letters in 5! ways in each of which vowels are together. These 3 vowels can be arranged among themselves in 3! ways.
=> Total number of words =5!*3!
=120*6=720
So there are total of 720 ways in which vowels are ALWAYS TOGEGHER.

Now,
Since there are no repeated letters, the total number of ways in which the letters of the word 'BENGALI' cab be arranged:
=7!=5040

So,
Total no. of arrangements in which vowels are never together:
=ALL the arrangements possible – arrangements in which vowels are ALWAYS TOGETHER
=5040–720=4320

ii) There are 4 odd places and 3 even places. 3 vowels can occupy 4 odd places in 4P3 ways and 4 constants can be arranged in 4P4 ways.
=> Number of words $=4P_3*4P_4= 576$.

16. In how many ways 5 gentlemen and 3 ladies can be arranged along a round table so that no 2 ladies are together.

Solution:

The 5 gentlemen can be arranged in a round table in (5-1)! = 4! ways.

Since no 2 ladies are together, they can occupy the 5 gaps in between the gentlemen in $5P_3$ ways.

Therefore, total no. of arrangements = $5P_3$ x 4!

COMBINATION

Let us consider the example of shirts and trousers as stated in the introduction. There you have 4 sets of shirts and trousers and you want to take 2 sets with you while going on a trip. In how many ways can you do it?

Let us denote the sets by S1,S2,S3,S4. Then you can choose two pairs in the following ways:

1. {S1,S2}	2. {S1,S3}	3. {S1,S4}
4. {S2,S3}	5. {S2,S4}	6. {S3,S4}

[Observe that {S1,S2} is the same as {S2,S1}. So, there are 6 ways of choosing the two sets that you want to take with you. Of course, if you had 10 pairs and you wanted to take 7 pairs, it will be much more difficult to work out the number of pairs in this way.

Now as you may want to know the number of ways of wearing 2 out of 4 sets for two days, say Monday and Tuesday, and the order of wearing is also important to you. We know that it can be done in 4P4=12 ways. But note that each choice of 2 sets gives us two ways of wearing 2 sets out of 4 sets as shown below:

1. {S1,S2} -> S1 on Monday and S2 on Tuesday or S2 on Monday and S1 on Tuesday
2. {S1,S3} -> S1 on Monday and S3 on Tuesday or S3 on Monday and S1 on Tuesday
3. {S1,S4} -> S1 on Monday and S4 on Tuesday or S4 on Monday and S1 on Tuesday
4. {S2,S3} -> S2 on Monday and S3 on Tuesday or S3 on Monday and S2 on Tuesday
5. {**S2**,S4} -> S2 on Monday and S4 on Tuesday or S4 on Monday and S2 on Tuesday
6. {S3,S4} -> S3 on Monday and S4 on Tuesday or S4 on Monday and S3 on Tuesday

Thus, there are 12 ways of wearing 2 out of 4 pairs.
This argument holds good in general as we can see from the following theorem.

Theorem

Let $n \geq 1$ be an integer and $r \leq n$. Let us denote the number of ways of choosing r objects out of n objects by nCr. Then
$nCr = \frac{nP_r}{r!}$.

17. Find the number of subsets of the set {1,2,3,4,5,6,7,8,9,10,11} having 4 elements.

Solution

Here the order of choosing the elements doesn't matter and this is a problem in combinations.

We have to find the number of ways of choosing 4 elements of this set which has 11 elements.

$11C_4 = \frac{11\times10\times9\times8}{1\times2\times3\times4} = 330$

18. 12 points lie on a circle. How many cyclic quadrilaterals can be drawn by using these points?

Solution

For any set of 4 points we get a cyclic quadrilateral. Number of ways of choosing 4 points out of 12 points is $12C_4$=495.

Therefore, we can draw 495 quadrilaterals.

19. In a box, there are 5 black pens, 3 white pens and 4 red pens. In how many ways can 2 black pens, 2 white pens and 2 red pens can be chosen?

Solution

Number of ways of choosing 2 black pens from 5 black pens
=$5C_2$= $\frac{5P_2}{2!} = \frac{5\times 4}{1\times 2} = 10$

Number of ways of choosing 2 white pens from 3 white pens
=$3C_2$= $\frac{3P_2}{2!} = \frac{3\times 2}{1\times 2} = 3$

Number of ways of choosing 2 red pens from 4 red pens
=$4C_2$= $\frac{4P_2}{2!} = \frac{4\times 3}{1\times 2} = 6$

=> By the Counting Principle, 2 black pens, 2 white pens, and 2 red pens can be chosen in 10*3*6= 180ways.

20. A question paper consists of 10 questions divided into two parts A and B. Each part contains five questions. A candidate is required to attempt six questions in all of which at least 2 should be from part A and at least 2 from part B. In how many ways can the candidate select the questions if he can answer all questions equally well?

Solution

The candidate has to select six questions in all of which at least two should be from Part A and two should be from Part B. He can select questions in any of the following ways:

Part A	**Part B**
(i) 2	4
(ii) 3	3
(iii) 4	2

If the candidate follows choice (i), the number of ways in which he can do so is:
$5C_2$*$5C_4$=10*5=50

If the candidate follows choice (ii), the number of ways in which he can do so is: $5C_3*5C_3=10*10=100$

Similarly, if the candidate follows choice (iii), then the number of ways in which he can do so is:$5C_4*5C_2=5*10=50$

Therefore, the candidate can select the question in 50+100+50= 200 ways**.**

21. A committee of 5 persons is to be formed from 6 men and 4 women. In how many ways can this be done when:**(i)** At least 2 women are included?**(ii)** At most 2 women are included?

Solution

(i) When at least 2 women are included.
The committee may consist of

3 women, 2 men: It can be done in $4C*6C_2$ ways
Or, 4 women, 1 man: It can be done in $4C_4*6C_1$ ways
or, 2 women, 3 men: It can be done in $4C_2*6C_3$ ways

=> Total number of ways of forming the committee:
$=4C_3*6C_2+4C_4*6C_1+4C_2*6C_3$= 186 ways

(ii) When at most 2 women are included
The committee may consist of

2 women, 3 men: It can be done in $4C_2*6C_3$ ways
Or, 1 women, 4 men: It can be done in $4C_1*6C_4$ ways
Or, 5 men: It can be done in $6C_5$ ways

=> Total number of ways of forming the committee:
$=4C_2*6C_3+4C_1*6C_4+6C_5$= 186 ways

22. The Indian Cricket team consists of 16 players. It includes 2 wicket keepers and 5 bowlers. In how many ways can a cricket eleven be selected if we have to select 1 wicket keeper and at least 4 bowlers?

Solution

We are to choose 11 players including 1 wicket keeper and 4 bowlers
or, 1 wicket keeper and 5 bowlers.

Number of ways of selecting 1 wicket keeper, 4 bowlers and 6 other players
$=2C_1*5C_4*9C_6=840$

Number of ways of selecting 1 wicket keeper, 5 bowlers and 5 other players
$=2C_1*5C_5*9C_5=252$

=> Total number of ways of selecting the team:
=840+252= 1092

23. There are 5 novels and 4 biographies. In how many ways can 4 novels and 2 biographies can be arranged on a shelf?

Solution

4 novels can be selected out of 5 in $5C_4$ ways.
2 biographies can be selected out of4 in $4C_2$ ways.
Number of ways of arranging novels and biographies:
$=5C_4*4C_2=30$

After selecting any 6 books (4 novels and 2 biographies) in one of the 30 ways, they can be arranged on the shelf in 6!=720 ways.
By the Counting Principle, the total number of arrangements =30*720= **21600**

24. From 5 consonants and 4 vowels, how many words can be formed using 3 consonants and 2 vowels?

Solution

From 5 consonants, 3 consonants can be selected in $5C_3$ ways.
From 4 vowels, 2 vowels can be selected in $4C_2$ ways.
Now with every selection, number of ways of arranging 5 letters is $5P_5$

Total number of words $=5C_3*4C_2*5P_5= 7200$.

Binomial Theorem

$(a+b)^n = nC_0a^n+ nC_1a^{n-1}b+\ldots+ nC_nb^n$

25. Find the coefficient of the independent term of x in expansion of $(3x - (2/x^2))^{15}$.

Solution

The general term of $(3x - (2/x^2))^{15}$ is written, as $T_{r+1} = {}^{15}C_r (3x)^{15-r} (-2/x^2)^r$. It is independent of x if,

$15 - r - 2r = 0 \Rightarrow r = 5$

$\therefore\ T_6 = {}^{15}C_5(3)^{10}(-2)^5 = - {}^{16}C_5\ 3^{10}\ 2^5$.

26. Find the value of the greatest term in the expansion of $\sqrt{3}(1+(1/\sqrt{3}))^{20}$.

Solution

Let T_{r+1} be the greatest term, then $T_r < T_{r+1} > T_{r+2}$

Consider : $T_{r+1} > T_r$

$\Rightarrow {}^{20}C_r\ (1/\sqrt{3})^r > 20C_{r-1}(1/\sqrt{3})^{r-1}$

$\Rightarrow ((20)!/(20-r)!r!)\ (1/(\sqrt{3})^r) > ((20)!/(21-r)!(r-1)!)\ (1/(\sqrt{3})^{r-1})$

$\Rightarrow r < 21/(\sqrt{3}+1)$

$\Rightarrow r < 7.686$ .... (i)

Similarly, considering $T_{r+1} > T_{r+2}$

$\Rightarrow r > 6.69$ .... (ii)

From (i) and (ii), we get

$r = 7$

Hence greatest term = $T_8 = 25840/9$

27. Find the coefficient of x^{50} in the expansion of $(1+x)^{1000} + 2x(1+x)^{999} + 3x^2 (1+x)^{998} + ... + 1001x^{1000}$.

Solution

Let $S = (1 + x)^{1000} + 2x(1+x)^{999} + ... + 1000x^{999} (1+x) + 1001 x^{1000}$

This is an Arithmetic Geometric Series with $r = x/(1+x)$ and $d = 1$.

Now $(x/(1+x)) S = x(1 + x)^{999} + 2x^2 (1 + x)^{998} + ... + 1000x^{1000} + 1000x^{1001}/(1+x)$

Subtracting we get,

$(1 - (x/(x+1))) S =(1+x)^{1000} + x(1+x)^{999} + ... + x^{1000} - 1001x^{1000}/(1+x)$

or $S = (1+x)^{1001} + x(1+x)^{1000} + x^2(1+x)^{999} + ... + x^{1000} (1+x)-1001x^{1001}$

This is G.P. and sum is

$S = (1+x)^{1002} - x^{1002} - 1002x^{1001}$

So the coeff. of x^{50} is = ${}^{1002}C_{50}$

SATHYABAMA
INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited "A" Grade by NAAC | 12B Status by UGC | Approved by AICTE
www.sathyabama.ac.in

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF MATHEMATICS

UNIT – III –INTRODUCTION TO PROBABILITY– SMTA1102

Theory of Probability

Introduction

If an experiment is repeated under essential homogeneous and similar conditions we generally come across two types of situations:

(i) The result or what is usually known as the 'outcome' is unique or certain.

(ii) The result is not unique but may be one of the several possible outcomes.

The phenomena covered by (i) are known as deterministic. For example, for a perfect gas, PV = constant.

The phenomena covered by (ii) are known as probabilistic. For example, in tossing a coin we are not sure if a head or tail will be obtained.

In the study of statistics we are concerned basically with the presentation and interpretation of chance outcomes that occur in a planned study or scientific investigation.

Definition of various terms

Trial and event: Consider an experiment which, though repeated under essentially identical conditions, does not give unique results but may result in any one of the several possible outcomes. The experiment is known as a trial and outcomes are known as events or cases. For example, throwing of a die is a trial and getting 1(or 2 or … 6) is an event.

Exhaustive events: The total number of possible outcomes in any trial is known as exhaustive events or exhaustive cases. For example, in tossing of a coin there are two exhaustive case, viz.: Head and Tail(the possibility of the coin standing on an edge being ignored)

Favourable events or cases: The number of cases favourable to an event in a trial is the number of outcomes which entail the happening of the event. For example, in throwing of two dice, the number of cases favourable to getting the sum 3 is: (1,2) and (2,1)

Mutually exclusive events: Events are said to be mutually exclusive or incompatible if the happening of any one of them precludes the happening of all the others, that is if no two or more of them can happen simultaneously in the same trial. For example, in tossing a coin the events head and tail are mutually exclusive.

Equally likely events: Outcomes of a trial are said to be equally likely, if taking into consideration all the relevant evidences, there is no reason to expect one in preference to the others. For example, in throwing an unbiased die, all the six faces are equally likely to come.

Sample Space: Consider an experiment whose outcome is not predictable with certainty. However, although the outcome of the experiment will not be known in advance, let us suppose that the set of all possible outcomes is known. This set of all possible outcomes of an experiment is known as the **sample space** of the experiment and is denoted by S.

Some examples follow.

1. If the outcome of an experiment consists in the determination of the sex of a newborn child, then

$$S = \{ g,b\}$$

where the outcome g means that the child is a girl and b that it is a boy.

2. If the experiment consists of flipping two coins, then the sample space consists of the following four points:

$$S = \{(H,H), (H,T), (T,H), (T,T)\}$$

The outcome will be (H,H) if both coins are heads, (H,T) if the first coin is heads and the second tails, (T,H) if the first is tails and the second heads, and (T,T) if both coins are tails.

3. If the experiment consists of tossing two dice, then the sample space consists if the 36 points

$$S = \{ (i,j):i,j = 1, 2, 3, 4, 5,$$

$$= \{ (1,1)\text{------}(1,6)\text{-----}(6,1)\text{-----}(6,6) \}$$

where the outcome (i,j) is said to occur if i appears on the leftmost die and j on the other die.

3.2. Definitions of Probability

1. Mathematical or Classical or a priori probability:

If a trial results in n exhaustive, mutually exclusive and equally likely cases and m of them are favourable to the happening of an event E, then the probability 'p' of happening of E is given by,

$$p = P(E) = \frac{\text{Favourable number of cases}}{\text{Exhaustive number of cases}} = \frac{m}{n}$$

2. Statistical or empirical probability:

If a trial is repeated a number of times under essentially homogenous and identical conditions, then the limiting value of the number of times the event happens to the number of trials, as the number of trials become indefinitely large is called the probability of happening of the event. Symbolically, if in n trials an event E happens m times, then the probability ‘p’ of the happening of E is given by,

$$P = P(E) = \lim_{n \to \infty} \frac{m}{n}$$

3. Axiomatic Definition:

Consider an experiment whose sample space is S. For each event E of the sample space S, we assume that a number P(E) is defined and satisfies the following three axioms.

Axiom 1: $0 \le P(E) \le 1$

Axiom 2: P(S) = 1

Axiom 3: For any sequence of mutually exclusive events, E_1, E_2, … (that is, events for which $E_i E_j = \Phi$, when $i \neq j$),

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

Some Important Formulas

1. If A and B are any two events, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

This rule is known as additive rule on probability.

For three events A, B and C, we have,

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

2. If A and B are mutually exclusive events, then

$$P(A \cup B) = P(A) + P(B)$$

In general, if $A_1, A_2, \dots, A_n$ are mutually exclusive, then

$$P(A_1 \cup A_2 \cup A_3 \cup \dots \cup A_n) = P(A_1) + P(A_2) + \dots + P(A_n)$$

3. If A and A^c are complementary events, then

$P(A) + P(A^c) = 1$

4. $P(S) = 1$

5. $P(\Phi) = 0$

6. If A and B are any two events, then

$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$

7. If A and B are independent events, then

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

Glossary of Probability terms:

Statement	**Meaning in terms of Set theory**
1. At least one of the events A or B occurs	$\omega \in A \cup B$
2. Both the events A and B occur	$\omega \in A \cap B$
3. Neither A nor B occurs	$\omega \in \overline{A} \cap \overline{B}$
4. Event A occurs and B does not occur	$\omega \in A \cap \overline{B}$
5. Exactly one of the events A or B occurs	$\omega \in A \Delta B$
6. If event A occurs, so does B	$A \subset B$
7. Events A and B are mutually exclusive	$A \cap B = \Phi$
8. Complementary event of A	$\overline{A}$

9. **Sample space**	**Universal set S**

Example 1: Find the probability of getting a head in tossing a coin.

Solution: When a coin is tossed, we have the sample space $\{$Head, Tail$\}$

Therefore, the total number of possible outcomes is 2

The favourable number of outcomes is 1, that is the head.

$\therefore$The required probability is ½.

Example 2: Find the probability of getting two tails in two tosses of a coin.

Solution: When two coins are tossed, we have the sample space $\{$HH, HT, TH, TT$\}$

Where H represents the outcome Head and T represents the outcome Tail.

The total number of possible outcomes is 4.

The favourable number of outcomes is 1, that is TT

$\therefore$The required probability is ¼.

Example 3: Find the probability of getting an even number when a die is thrown

Solution: When a die is thrown the sample space is $\{1, 2, 3, 4, 5, 6\}$

The total number of possible outcomes is 6

The favourable number of outcomes is 3, that is 2, 4 and 6

$\therefore$The required probability is= $\frac{3}{6} =$½.

Example 4: What is the chance that a leap year selected at random will contain 53 Sundays?

Solution: In a leap year(which consists of 366 days) there are 52 complete weeks and 2 days over. The following are the possible combinations for these two over days:

(i) Sunday and Monday (ii)Monday and Tuesday (iii)Tuesday and Wednesday (iv)Wednesday and Thursday (v)Thursday and Friday (vi)Friday and Saturday (vii)Saturday and Sunday.

In order that a leap year selected at random should contain 53 Sundays, one of the two over days must be Sunday. Since out of the above 7 possibilities, 2 viz. (i) and (ii)are favourable to this event, Required probability $= \frac{2}{7}$

Example 5: If two dice are rolled, what is the probability that the sum of the upturned faces will equal 7?

Solution: We shall solve this problem under the assumption that all of the 36 possible outcomes are equally likely. Since there are 6 possible outcomes – namely (1,6), (2,5), (3,4), (4,3), (5,2,), (6,1) – that result in the sum of the dice being equal to 7, the desired probability is $\frac{6}{36} = \frac{1}{6}$.

Example 6: A bag contains 3 Red, 6 White and 7 Blue balls. What is the probability that two balls drawn are white and blue?

Solution: Total number of balls = 3 + 6 + 7 = 16.

Out of 16 balls, 2 can be drawn in $16C_2$ ways.

Therefore exhaustive number of cases is 120.

Out of 6 white balls 1 ball can be drawn in $6C_1$ ways and out of 7 blue balls 1 ball can be drawn in $7C_1$ ways. Since each of the former cases can be associated with each of the latter cases, total number of favourable cases is $6C_1$ x $7C_1$ = 6 x 7 = 42.

$\therefore$ The required probability is = $\frac{42}{120} = \frac{7}{20}$

Example 7: A lot consists of 10 good articles, 4 with minor defects and 2 with major defects. Two articles are chosen from the lot at random (without replacement). Find the probability that (i) both are good, (ii) both have major defects, (iii) at least 1 is good, (iv) at most 1 is good, (v)exactly 1 is good, (vi) neither has major defects and (vii) neither is good.

Solution: Although the articles may be drawn one after the other, we can consider that both articles are drawn simultaneously, as they are drawn without replacement.

(i) $P(both \text{ are good}) = \frac{\text{No. of ways drawing 2 good articles}}{\text{Total no. of ways of drawing 2 articles}}$

$= \frac{10C_2}{16C_2} = \frac{3}{8}$

(ii) $P(\text{both have major defects}) = \dfrac{\text{No. of ways of drawing 2 articles with major defects}}{\text{Total no. of ways}}$

$$= \frac{2C_2}{16C_2} = \frac{1}{120}$$

(iii) P(at least 1 is good) = P(exactly 1 is good or both are good)

=P(exactly 1 is good and 1 is bad or both are good)

$$= \frac{10C_1 x 6C_1 + 10C_2}{16C_2} = \frac{7}{8}$$

(iv) P(atmost 1 is good) =P(none is good or 1 is good and 1is bad)

$$= \frac{10C_0 x 6C_2 + 10C_1 x 6C_1}{16C_2} = \frac{5}{8}$$

(v) P(exactly 1is good) =P(1 is good and 1 is bad)

$$= \frac{10C_1 x 6C_1}{16C_2} = \frac{1}{2}$$

(vi) P(neither has major defects) = P(both are non-major defective articles)

$$= \frac{14C_2}{16C_2} = \frac{91}{120}$$

(vii) P(neither is good) = P(both are defective)

$$= \frac{6C_2}{16C_2} = \frac{1}{8}$$

Example 8: From 6 positive and 8 negative numbers, 4 numbers are chosen at random (without replacement) and multiplied. What is the probability that the product is positive?

Solution: If the product is to be positive, all the 4 numbers must be positive or all the 4 must be negative or 2 of them must be positive and the other 2 must be negative.

No. of ways of choosing 4 positive numbers= $6C_4$ =15.

No. of ways of choosing 4 negative numbers= $8C_4$ =70.

No.of ways of choosing 2 positive and 2 negative numbers

$$= 6C_2 x 8C_2 = 420.$$

Total no. of ways of choosing 4 numbers from all the 14 numbers

$$= \quad 14C_4 = 1001.$$

P(the product is positive)

$$= \frac{No.\text{ of ways by which the product is positive}}{\text{Total no. of ways}}$$

$$= \frac{15+70+420}{1001} = \frac{505}{1001}$$

Example 9: If 3 balls are "randomly drawn" from a bowl containing 6 white and 5 black balls, what is the probability that one of the drawn balls is white and the other two black?

Solution: If we regard the order in which the balls are selected as being relevant, then the sample space consists of 11· 10 · 9 = 990 outcomes. Furthermore, there are 6· 5· 4 = 120 outcomes in which the first ball selected is white and the other two black; 5 · 6· 4 = 120 outcomes in which the first is black, the second white and the third black; and 5· 4 · 6 = 120 in which the first two are black and the third white. Hence, assuming that "randomly drawn" means that each outcome in the sample space is equally likely to occur, we see that the desired probability is $\frac{120+120+120}{990} = \frac{4}{11}$

Example 10: In a large genetics study utilizing guinea pigs, *Cavia sp.,* 30% of the offspring produced had white fur and 40% had pink eyes. Two-thirds of the guinea pigs with white fur had pink eyes. What is the probability of a randomly selected offspring having both white fur and pink eyes?

Solution: P(W) = 0.30, P(Pi) = 0.40, and P(Pi W) = 0.67. Utilizing Formula 2.9,

P(Pi ∩ W) = P(Pi W). P(W) = 0.67. 0.30 = 0.20.

Twenty percent of all offspring are expected to have both white fur and pink eyes.

Example 11: Consider three gene loci in tomato, the first locus affects fruit shape with the oo genopyte causing oblate or flattened fruit and OO or Oo normal round fruit. The second locus affects fruit color with yy having yellow fruit and YY or Yy red fruit. The final locus affects leaf shape with pp having potato or smooth leaves and PP or Pp having the more typical cut leaves. Each of these loci is located on a different pair of chromosomes and, therefore, acts independently of the other loci. In the following cross OoYyPp × OoYypp, what is the probability that an offspring will have the dominant phenotype for each trait? What is the probability that it will be heterozygous for all three genes? What is the probability that it will have round, yellow fruit and potato leaves?

Solution: Genotypic array:

$$(\frac{1}{4}\text{OO} + \frac{2}{4}\text{Oo} + \frac{1}{4}\text{oo})\ (\frac{1}{4}\text{YY} + \frac{2}{4}\text{Yy} + \frac{1}{4}\text{yy})\ (\frac{1}{2}\text{pp})$$

Phenotypic array:

$$(\frac{3}{4}\text{O-} + \frac{1}{4}\text{oo})\ (\frac{3}{4}\text{Y-} + \frac{1}{4}\text{yy})\ (\frac{1}{2}\text{P} + \frac{1}{2}\text{pp})$$

The probabiltity of dominant phenotype for each trait from the phenotypic array above is

$$P(\text{O-Y-P-}) = P(\text{O-}) \times P(\text{Y-}) \times P(\text{P-}) = \frac{3}{4} \times \frac{3}{4} \times \frac{1}{2} = \frac{9}{32}.$$

The probability of heterozygous for all three genes from the genotypic array above is

$$P(\text{OoYyPp}) = P(\text{Oo}) \times P(\text{Yy}) \times P(\text{Pp}) = \frac{2}{4} \times \frac{2}{4} \times \frac{1}{2} = \frac{4}{32} = \frac{1}{8}.$$

The probability of a round, yellow-fruited plant with potato leaves from the phenotypic array above is

$$P(\text{O-yypp}) = P(\text{O-}) \times P(\text{yy}) \times P(\text{pp}) = \frac{3}{4} \times \frac{1}{4} \times \frac{1}{2} = \frac{3}{32}.$$

Each answer applies the probability rules for independent events to the separate gene loci.

Example 12: (a) Two cards are drawn at random from a well shuffled pack of 52 playing cards. Find the chance of drawing two aces.

(b) From a pack of 52 cards, three are drawn at random. Find the chance that they are a king, a queen and a knave.

(c) Four cards are drawn from a pack of cards. Find the probability that (i) all are diamond (ii) there is one card of each suit (iii) there are two spades and two hearts.

Solution: (a) From a pack of 52 cards 2 can be drawn in $52C_2$ ways, all being equally likely. $\therefore$ Exhaustive number of cases is $52C_2$.

In a pack there are 4 aces and therefore 2 aces can be drawn in $4C_2$ ways.

$\therefore$ Required probability = $\frac{4C_2}{52C_2} = \frac{1}{221}$

(b) Exhaustive number of cases = $52C_3$

A pack of cards contains 4 kings, 4 queens and 4 knaves. A king, a queen and a knave can each be drawn in $4C_1$ ways and since each way of drawing a king can be associated with each of the ways of drawing a queen and a knave, the total number of favrourable cases = $4C_1 \times 4C_1 \times 4C_1$.

$\therefore$ Required probability = $\frac{4C_1 \times 4C_1 \times 4C_1}{52C_3}$ = $\frac{16}{5525}$

(c) Exhaustive number of cases $52C_4$

(i) Required probability = $\frac{13C_4}{52C_4}$

(ii) Required probability = $\frac{13C_1 \times 13C_1 \times 13C_1 \times 13C_1}{52C_4}$

(iv) Required probability = $\frac{13C_2 \times 13C_2}{52C_4}$

Example 13: What is the probability of getting 9 cards of the same suit in one hand at a game of bridge?

Solution: One hand in a game of bridge consists of 13 cards.

$\therefore$ Exhaustive number of cases $52C_{13}$

Number of ways in which, in one hand, a particular player gets 9 cards of one suit are $13C_9$ and the number of ways in which the remaining 4 cards are of some other suit are $39C_4$. Since there are 4 suits in a pack of cards, total number of favourable cases is $4 \times 13C_9 \times 39C_4$.

$\therefore$ Required probability = $\frac{4 \times 13C_9 \times 39C_4}{52C_{13}}$

Example 14: A committee of 4 people is to be appointed from 3 officers of the production department, 4 officers of the purchase department, two officers of the sales department and 1 chartered accountant. Find the probability of forming the committee in the following manner:

(i) There must be one from each category
(ii) It should have at least one from the purchase department
(iii) The chartered accountant must be in the committee.

Solution: There are 3 + 4 + 2 + 1 = 10 persons in all and a committee of 4 people can be formed out of them in $10C_4$ ways. Hence exhaustive number of cases is $10C_4$ = 210

(i) Favourable number of cases for the committee to consist of 4 members, one from each category is $4C_1$ $\times 3C_1 \times 2C_1 \times 1 = 24$

$\therefore$ Required probability = $\frac{24}{120}$

(ii) P(Committee has at least one purchase officer) = 1 – P(Committee has no purchase Officer)

In order that the committee has no purchase officer, all the four members are to be selected amongst officers of production department, sales department and chartered accountant, that is out of 3 + 2 + 1 = 6 members and this can be done in $5C_4$ = 15 ways. Hence,

P(Committee has no purchase officer) = $\frac{15}{210} = \frac{1}{14}$

$\therefore$ P(Committee has at least one purchase officer) = 1 – $\frac{1}{14}$ = $\frac{13}{14}$

(iii) Favourable number of cases that the committee consists of a chartered accountant as a member and three others are:

$1 \times 9C_3$ = 84 ways.

Since a chartered accountant can be selected out of one chartered accountant in only 1 way and the remaining 3 members can be selected out of the remaining 10 – 1 persons in $9C_3$ ways. Hence the required probability = $\frac{84}{210} = \frac{2}{5}$.

Example 15: A box contains 6 red, 4 white and 5 black balls. A persons draws 4 balls from the box at random. Find the probability that among the balls drawn there is at least one ball of each colour.

Solution: The required event E that in a draw of 4 balls from the box at random there is at least one ball of each colour can materialize in the following mutually disjoint ways:

(i) 1 Red, 1 White and 2 Black balls
(ii) 2 Red, 1 White and 1 Black balls
(iii) 1 Red, 2 White and 1 Black balls

Hence by addition rule of probability, the required probability is given by,

P(E) = P(i) + P(ii) + P(iii)

$$= \frac{6C_1 \times 4C_1 \times 5C_2}{15C_4} + \frac{6C_2 \times 4C_1 \times 5C_1}{15C_4} + \frac{6C_1 \times 4C_2 \times 5C_1}{15C_4}$$

= 0.5275

Example 16: A problem in Statistics is given to the three students A, B and C whose chances of solving it are 1/2, 3/4 and 1/4 respectively. What is the probability that the problem will be solved if all of them try independently?

Solution: Let A, B and C denote the events that the problem is solved by the students A, B and C respectively. Then

P(A) = 1/2 P(B) = 3/4 P(C) = 1/4

P($\overline{A}$) = 1 – 1/2 = 1/2 P($\overline{B}$) = 1 – 3/4 = 1/4 P($\overline{C}$) = 1 – 1/4 = 3/4

P(Problem solved) = P(At least one of them solves the problem)

= 1 – P(None of them solve the problem)

= 1 – P($\overline{A \cup B \cup C}$)

= 1 – P($\overline{A} \cap \overline{B} \cap \overline{C}$)

= 1 – P($\overline{A}$) P($\overline{B}$) P($\overline{C}$)

$$= 1 - \frac{1}{2} \times \frac{1}{4} \times \frac{3}{4}$$

$$= \frac{29}{32}$$

Example 17: Three groups of children contain respectively 3 girls and 1 boy, 2 girls and 2 boys and 1 girl and 3 boys. One child is selected at random from each group. Find the probability that the three selected consist of 1 girl and 2 boys.

Solution: The required event of getting 1 girl and 2 boys among the three selected children can materialize in the following three mutually exclusive cases:

Group No. →	**I**	**II**	**III**
(i)	**Girl**	**Boy**	**Boy**
(ii)	**Boy**	**Girl**	**Boy**
(iii)	**Boy**	**Boy**	**Girl**

By addition rule of probability,

Required probability = P(i) + P(ii) + P(iii)

Since the probability of selecting a girl from the first group is 3/4, of selecting a boy from the second is 2/4, and of selecting a boy from the third group is ¾, and since these three events of selecting children from the three groups are independent of each other, we have,

P(i) = $\frac{3}{4} \times \frac{2}{4} \times \frac{3}{4} = \frac{9}{32}$

P(ii) = $\frac{1}{4} \times \frac{2}{4} \times \frac{3}{4} = \frac{3}{32}$

P(iii) = $\frac{1}{4} \times \frac{2}{4} \times \frac{1}{4} = \frac{1}{32}$

Hence the required probability = $\frac{9}{32} + \frac{3}{32} + \frac{1}{32} = \frac{13}{32}$

Conditional Probability and Baye's Theorem

Conditional Probability and Multiplication Law

For two events A and B

$P(A \cap B) = P(A) \,.\, P(B/A), \; P(A) > 0$

$= P(B) \,.\, P(A/B), \; P(B) > 0$

where P(B/A) represents the conditional probability of occurrence of B when the event A has already happened and P(A/B) is the conditional probability of occurrence of A when the event B has already happened.

Theorem of Total Probability:

If $B_1, B_2, \ldots, B_n$ be a set of exhaustive and mutually exclusive events, and A is another event associated with (or caused by) B_i, then

$$P(A) = \sum_{i=1}^{n} P(B_i) P(A/B_i)$$

Example 18 : A box contains 4 bad and 6 good tubes. Two are drawn out from the box at a time. One of them is tested and found to be good. What is the probability that the other one is also good?

Solution: Let A = one of the tubes drawn is good and B = the other tube is good.

P(A∩B) = P(both tubes drawn are good)

$$= \frac{6C_2}{10C_2} = \frac{1}{3}$$

Knowing that one tube is good, the conditional probability that the other tube is also good is required, i.e., P(B/A) is required.

By definition,

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{1/3}{6/10} = \frac{5}{9}$$

Example 19: A bolt is manufactured by 3 machines A, B and C. A turns out twice as many items as B, and machines B and C produce equal number of items. 2% of bolts produced by A and B are defective and 4% of bolts produced by C are defective. All bolts are put into 1 stock pile and chosen from this pile. What is the probability that it is defective?

Solution: Let A = the event in which the item has been produced by machine A, and so on.

Let D = the event of the item being defective.

$$P(A) = \frac{1}{2}, \quad P(B) = P(C) = \frac{1}{4}$$

P(D/A) = P(an item is defective, given that A has produced it)

$$= \frac{2}{100} = P(D/B)$$

$$P(D/C) = \frac{4}{100}$$

By theorem of total probability,

P(D) = P(A)× P(D/A) + P(B) × P(D/B) + P(C) ×P(D/c)

$$= \frac{1}{2} \times \frac{2}{100} + \frac{1}{4} \times \frac{2}{100} + \frac{1}{4} \times \frac{4}{100}$$

$$= \frac{1}{40}$$

Example 20: In a coin tossing experiment, if the coin shows head, one die is thrown and the result is recorded. But if the coin shows tail, 2 dice are thrown and their sum is recorded. What is the probability that the recorded number will be 2?

Solution: When a single die is thrown, P(2) = 1/6

When 2 dice are thrown, the sum will be 2 only if each dice shows 1.

$\therefore$ P(getting 2 as sum with 2 dice) = $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$ (since independence)

By theorem of total probability,

$$P(2) = P(H) \times P(2/H) + P(T) \times P(2/T)$$

$$= \frac{1}{2} \times \frac{1}{6} + \frac{1}{2} \times \frac{1}{36} = \frac{7}{72}$$

Example 21: An urn contains 10 white and 3 black balls. Another urn contains 3 white and 5 black balls. Two balls are drawn at random from the first urn and place in the second urn and then one ball is taken at random from the latter. What is the probability that it is a white ball?

Solution: The two balls transferred may be both white or both black or one white and one black.

Let B_1 = event of drawing 2 white balls from the first urn, B_2 = event of drawing 2 black balls from it and B_3 = event of drawing one white and one black ball from it.

Clearly B_1, B_2 and B_3 are exhaustive and mutually exclusive events.

Let A = event of drawing a white ball from the second urn after transfer.

$$P(B_1) = \frac{10C_2}{13C_2} = \frac{15}{26}$$

$$P(B_2) = \frac{3C_2}{13C_2} = \frac{1}{26}$$

$$P(B_3) = \frac{10 \times 3}{13C_2} = \frac{10}{26}$$

$P(A/B_1)$ = P(drawing a white ball / 2 white balls have been transferred)

= P(drawing a white ball / urn II contains 5 white and 5 black balls)

$= \frac{5}{10}$

Similarly, $P(A/B_2) = \frac{3}{10}$ and $P(A/B_3) = \frac{4}{10}$

By theorem of total probability,

$$P(A) = P(B_1) \times P(A/B_1) + P(B_2) \times P(A/B_2) + P(B_3) \times P(A/B_3)$$

$$= \frac{15}{26} \times \frac{5}{10} + \frac{1}{26} \times \frac{3}{10} + \frac{10}{26} \times \frac{4}{10} = \frac{59}{130}$$

Example 22: In 1989 there were three candidates for the position of principal – Mr.Chatterji, Mr. Ayangar and Mr. Singh – whose chances of getting the appointment are in the proportion 4:2:3 respectively. The probability that Mr. Chatterji if selected would introduce co-education in the college is 0.3. The probabilities of Mr. Ayangar and Mr.Singh doing the same are respectively 0.5 and 0.8. What is the proabability that there will be co-education in the college?

Solution: Let the events and probabilities be defined as follows:

A: Introduction of co-education

E_1: Mr.Chatterji is selected as principal

E_2: Mr.Ayangar is selected as principal

E_3: Mr.Singh is selected as principal

Then,

$P(E_1) = \frac{4}{9}$ $\qquad$ $P(E_2) = \frac{2}{9}$ $\qquad$ $P(E_3) = \frac{3}{9}$

$P(A/E_1) = 0.3$ $\qquad$ $P(A/E_2) = 0.5$ $\qquad$ $P(A/E_3) = 0.8$

$$P(A) = P[(A \cap E_1) \cup (A \cap E_2) \cup (A \cap E_3)]$$

$$= P[(A \cap E_1) + (A \cap E_2) + (A \cap E_3)]$$

$$= P(E_1)\,P(A/E_1) + P(E_2)\,P(A/E_2) + P(E_3)\,P(A/E_3)$$

$$= \frac{4}{9} \times \frac{3}{10} + \frac{2}{9} \times \frac{5}{10} + \frac{3}{9} \times \frac{8}{10} = \frac{23}{45}$$

3.3.4. Baye's theorem

If $E_1, E_2, \ldots, E_n$ are mutually disjoint events with $P(E_i) \neq 0$, (i = 1,2, … , n) then for any arbitrary event A which is a subset of $\bigcup_{i=1}^{n} E_i$ such that P(A) > 0, we have,

$$P(E_i/A) = \frac{P(E_i)P(A/E_i)}{\sum_{i=1}^{n} P(E_i)P(A/E_i)}, \quad i = 1, 2, \ldots, n$$

3.3.5. Solved Examples

Example 23. A bag contains 5 balls and it is not known how many of them are white. Two balls are drawn at random from the bag and they are noted to be white. What is the chance that all the balls in the bag are white?

Solution: Since 2 white balls have been drawn out, the bag must have contained 2, 3, 4 or 5 white balls.

Let B_1 = Event of the bag containing 2 white balls, B_2 = Events of the bag containing 3 white balls, B_3 = Event of the bag containing 4 white balls and B_4 = Event of the bag containing 5 white balls.

Let A = Event of drawing 2 white balls.

$P(A/B_1) = \frac{2C_2}{5C_2} = \frac{1}{10}$ $\quad$ $P(A/B_2) = \frac{3C_2}{5C_2} = \frac{3}{10}$

$P(A/B_3) = \frac{4C_2}{5C_2} = \frac{4}{10}$ $\quad$ $P(A/B_4) = \frac{5C_2}{5C_2} = 1$

Since the number of white balls in the bag is not known, B_i's are equally likely.

$$P(B_1) = P(B_2) = P(B_3) = P(B_4) = \frac{1}{4}$$

By Baye's theorem,

$$P(B_4/A) = \frac{P(B_4)\times P(A/B_4)}{\sum_{i=1}^{4} P(B_i)\times P(A/B_i)}$$

$$= \frac{\frac{1}{4}\times 1}{\frac{1}{4}\times\left(\frac{1}{10}+\frac{3}{10}+\frac{3}{5}+1\right)} = \frac{1}{2}$$

Example 24: There are 3 true coins and 1 false coin with 'head' on both sides. A coin is chosen at random and tossed 4 times. If 'head' occurs all the 4 times, what is the probability that the false coin has beeb chosen and used?

Solution:

$$P(T) = P(\text{the coin is a true coin}) = \frac{3}{4}$$

$$P(F) = P(\text{the coin is a false coin}) = \frac{1}{4}$$

Let A = Event of getting all heads in 4 tosses

Then P(A/T) = $\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}\times\frac{1}{2}=\frac{1}{16}$ and P(A/F) = 1

By Baye's theorem

P(F/A) = $\frac{P(F)\times P(A/F)}{P(F)\times P(A/F)+P(T)\times P(A/T)}$

$$= \frac{\frac{1}{4}\times 1}{\frac{1}{4}\times 1+\frac{3}{4}\times\frac{1}{16}}=\frac{16}{19}$$

Example 25: The contents of urns I, Ii and III are as follows:

1 white, 2 black and 3 red balls

2 white, 1 black and 1 red balls

4 white, 5 black and 3 red balls

One urn is chosen at random and two balls are drawn. They happen to be white and red. What is the probability that they come from urns I, II or III?

Solution: Let E_1, E_2 and E_3 denote the events that the urn I, II and III is chosen, respectively, and let A be the event that the two balls taken from the selected urn are white and red. Then

P(E_1) = P(E_2) = P(E_3) = $\frac{1}{3}$

P(A/E_1) = $\frac{1\times 3}{6C_2}=\frac{1}{5}$

P(A/E_2) = $\frac{2\times 1}{4C_2}=\frac{1}{3}$

P(A/E_3) = $\frac{4\times 3}{12C_2}=\frac{2}{11}$

Hence $P(E_2/A) = \dfrac{P(E_2)P(A/E_2)}{\sum_{i=1}^{3} P(E_i)P(A/E_i)}$

$$= \frac{\frac{1}{3} \times \frac{1}{3}}{\frac{1}{3} \times \frac{1}{5} + \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{2}{11}} = \frac{55}{118}$$

Similarly, $P(E_3/A) = \dfrac{\frac{1}{3} \times \frac{2}{11}}{\frac{1}{3} \times \frac{1}{5} + \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{1}{11}} = \dfrac{30}{118}$

Therefore $P(E_1/A) = 1 - \dfrac{55}{118} - \dfrac{30}{118} = \dfrac{33}{118}$

SCHOOL OF SCIENCE AND HUMANITIES

DEPARTMENT OF MATHEMATICS

UNIT – V – RANDOM VARIABLES- SMTA1102

RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

1. DISCRETE RANDOM VARIABLES

1.1. **Definition of a Discrete Random Variable.** A random variable X is said to be *discrete* if it can assume only a finite or countable infinite number of distinct values. A discrete random variable can be defined on both a countable or uncountable sample space.

1.2. **Probability for a discrete random variable.** The probability that X takes on the value x, P(X=x), is defined as the sum of the probabilities of all sample points in Ω that are assigned the value x. We may denote P(X=x) by p(x). The expression p(x) is a function that assigns probabilities to each possible value x; thus it is often called the probability function for X.

1.3. **Probability distribution for a discrete random variable.** The probability distribution for a discrete random variable X can be represented by a formula, a table, or a graph, which provides p(x) = P(X=x) for all x. The probability distribution for a discrete random variable assigns nonzero probabilities to only a countable number of distinct x values. Any value x not explicitly assigned a positive probability is understood to be such that P(X=x) = 0.

The function f(x) p(x)= P(X=x) for each x within the range of X is called the *probability distribution* of X. It is often called the probability mass function for the discrete random variable X.

1.4. **Properties of the probability distribution for a discrete random variable.** A function can serve as the probability distribution for a discrete random variable X if and only if it s values, f(x), satisfy the conditions:

a: f(x) $\geq$ 0 for each value within its domain
b: $\sum_x \; f(x) \; = \; 1$, where the summation extends over all the values within its domain

1.5. **Examples of probability mass functions.**

1.5.1. *Example 1.* Find a formula for the probability distribution of the total number of heads obtained in four tosses of a balanced coin.

The sample space, probabilities and the value of the random variable are given in table 1.
From the table we can determine the probabilities as

$$P(X=0) \; = \frac{1}{16}\, , \; P(X=1) \; = \frac{4}{16}\, , \; P(X=2) \; = \frac{6}{16}\, , \; P(X=3) \; = \frac{4}{16}\, , \; P(X=4) \; = \frac{1}{16} \qquad (1)$$

Notice that the denominators of the five fractions are the same and the numerators of the five fractions are 1, 4, 6, 4, 1. The numbers in the numerators is a set of binomial coefficients.

$$\frac{1}{16} \; = \; \binom{4}{0}\frac{4}{16} \; = \; \binom{4}{1}\frac{6}{16} \; = \; \binom{4}{2}\frac{4}{16} \; = \; \binom{4}{3}\frac{1}{16} \; = \; \binom{4}{4}$$

We can then write the probability mass function as

Date: August 20, 2004.

TABLE 1. **Probability of a Function of the Number of Heads from Tossing a Coin Four Times.**

Table R.1 Tossing a Coin Four Times		
Element of sample space	Probability	Value of random variable X (x)
HHHH	1/16	4
HHHT	1/16	3
HHTH	1/16	3
HTHH	1/16	3
THHH	1/16	3
HHTT	1/16	2
HTHT	1/16	2
HTTH	1/16	2
THHT	1/16	2
THTH	1/16	2
TTHH	1/16	2
HTTT	1/16	1
THTT	1/16	1
TTHT	1/16	1
TTTH	1/16	1
TTTT	1/16	0

$$f(x) = \frac{\binom{4}{x}}{16} \; for \; x = 0, 1, 2, 3, 4 \tag{2}$$

Note that all the probabilities are positive and that they sum to one.

1.5.2. *Example 2.* Roll a red die and a green die. Let the random variable be the larger of the two numbers if they are different and the common value if they are the same. There are 36 points in the sample space. In table 2 the outcomes are listed along with the value of the random variable associated with each outcome.

The probability that X = 1, P(X=1) = P[(1, 1)] = 1/36. The probability that X = 2, P(X=2) = P[(1, 2), (2,1), (2, 2)] = 3/36. Continuing we obtain

$$P(X=1) = \binom{1}{36}, \; P(X=2) = \binom{3}{36} \; P(X=3) = \binom{5}{36}$$
$$P(X=4) = \binom{7}{36}, \; P(X=5) = \binom{9}{36}, \; P(X=6) = \binom{11}{36}$$

We can then write the probability mass function as

$$f(x) = P(X=x) = \frac{2x-1}{36} \; for \; x = 1, 2, 3, 4, 5, 6$$

Note that all the probabilities are positive and that they sum to one.

1.6. **Cumulative Distribution Functions.**

TABLE 2. Possible Outcomes of Rolling a Red Die and a Green Die – First Number in Pair is Number on Red Die

Green (A)	1	2	3	4	5	6
Red (D)						
1	1 1 *1*	1 2 2	1 3 *3*	1 4 *4*	1 5 *5*	1 6 *6*
2	2 1 *2*	2 2 2	2 3 *3*	2 4 *4*	2 5 *5*	2 6 *6*
3	3 1 *3*	3 2 *3*	3 3 *3*	3 4 *4*	3 5 *5*	3 6 6
4	4 1 *4*	4 2 *4*	4 3 *4*	4 4 *4*	4 5 *5*	4 6 *6*
5	5 1 *5*	5 2 *5*	5 3 *5*	5 4 *5*	5 5 *5*	5 6 *6*
6	6 1 *6*	6 2 *6*	6 3 *6*	6 4 *6*	6 5 *6*	6 6 *6*

1.6.1. *Definition of a Cumulative Distribution Function.* If X is a discrete random variable, the function given by

$$F(x) \;=\; P(x \;\leq\; X) \;=\; \sum_{t \,\leq\, x} f(t) \; for \; - \; \infty \;\leq\; x \;\leq\; \infty \tag{3}$$

where f(t) is the value of the probability distribution of X at t, is called the *cumulative distribution function* of X. The function F(x) is also called the *distribution function* of X.

1.6.2. *Properties of a Cumulative Distribution Function.* The values F(X) of the distribution function of a discrete random variable X satisfy the conditions

1: F(-∞) = 0 and F(∞) =1;
2: If a $<$ b, then F(a) $\leq$ F(b) for any real numbers a and b

1.6.3. *First example of a cumulative distribution function.* Consider tossing a coin four times. The possible outcomes are contained in table 1 and the values of f in equation 1. From this we can determine the cumulative distribution function as follows.

$$\begin{aligned}
F(0) \;&= f(0) \;\; = \frac{1}{16} \\
F(1) \;&= f(0) \; + \; f(1) \;\; = \frac{1}{16} \; + \frac{4}{16} \;\; = \frac{5}{16} \\
F(2) \;&= f(0) \; + \; f(1) \; + \; f(2) \;\; = \frac{1}{16} \; + \frac{4}{16} \; + \frac{6}{16} \;\; = \frac{11}{16} \\
F(3) \;&= f(0) \; + \; f(1) \; + \; f(2) \; + \; f(3) \;\; = \frac{1}{16} \; + \frac{4}{16} \; + \frac{6}{16} \; + \; \frac{4}{6} = \frac{15}{16} \\
F(4) \;&= f(0) \; + \; f(1) \; + \; f(2) \; + \; f(3) \; + \; f(4) \;\; = \frac{1}{16} \; + \frac{4}{16} \; + \frac{6}{16} \; + \; \frac{4}{6} \; + \frac{1}{16} \;\; = \frac{16}{16}
\end{aligned}$$

We can write this in an alternative fashion as

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

1.6.4. *Second example of a cumulative distribution function.* Consider a group of N individuals, M of whom are female. Then N-M are male. Now pick n individuals from this population without replacement. Let x be the number of females chosen. There are $\binom{M}{x}$ ways of choosing x females from the M in the population and $\binom{N-M}{n-x}$ ways of choosing n-x of the N - M males. Therefore, there are $\binom{M}{x} \times \binom{N-M}{n-x}$ ways of choosing x females and n-x males. Because there are $\binom{N}{n}$ ways of choosing n of the N elements in the set, and because we will assume that they all are equally likely the probability of x females in a sample of size n is given by

$$f(x) = P(X = x) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \begin{array}{l} for\ x = 0, 1, 2, 3, \cdots, n \\ and\ x \leq M,\ and\ n - x \leq N - M. \end{array} \tag{4}$$

For this discrete distribution we compute the cumulative density by adding up the appropriate terms of the probability mass function.

$$\begin{aligned} F(0) &= f(0) \\ F(1) &= f(0) + f(1) \\ F(2) &= f(0) + f(1) + f(2) \\ F(3) &= f(0) + f(1) + f(2) + f(3) \\ &\vdots \\ F(n) &= f(0) + f(1) + f(2) + f(3) + \cdots + f(n) \end{aligned} \tag{5}$$

Consider a population with four individuals, three of whom are female, denoted respectively by A, B, C, D where A is a male and the others are females. Then consider drawing two from this population. Based on equation 4 there should be $\binom{4}{2}$ = 6 elements in the sample space. The sample space is given by

TABLE 3. **Drawing Two Individuals from a Population of Four where Order Does Not Matter (no replacement)**

Element of sample space	Probability	Value of random variable X
AB	1/6	1
AC	1/6	1
AD	1/6	1
BC	1/6	2
BD	1/6	2
CD	1/6	2

We can see that the probability of 2 females is $\frac{1}{2}$. We can also obtain this using the formula as follows.

$$f(2) \;=\; P(X \,=\, 2) \;=\; \frac{\binom{3}{2}\,\binom{1}{0}}{\binom{4}{2}} = \frac{(3)(1)}{6} = \frac{1}{2} \tag{6}$$

Similarly

$$f(1) \;=\; P(X \,=\, 1) \;=\; \frac{\binom{3}{1}\,\binom{1}{1}}{\binom{4}{2}} = \frac{(3)(1)}{6} = \frac{1}{2} \tag{7}$$

We cannot use the formula to compute f(0) because (2 - 0) $\not\leq$ (4 - 3). f(0) is then equal to 0. We can then compute the cumulative distribution function as

$$\begin{aligned} F(0) \;&= f(0) \;=\; 0 \\ F(1) \;&= f(0) \;+\; f(1) \;=\frac{1}{2} \\ F(2) \;&= f(0) \;+\; f(1) \;+\; f(2) \;=\; 1 \end{aligned} \tag{8}$$

1.7. **Expected value.**

1.7.1. *Definition of expected value.* Let X be a discrete random variable with probability function p(x). Then the *expected value* of X, E(X), is defined to be

$$E(X) \;=\; \sum_x \; x \; p(x) \tag{9}$$

if it exists. The expected value exists if

$$\sum_x \; |\,x\,|\; p(x) \;<\; \infty \tag{10}$$

The expected value is kind of a weighted average. It is also sometimes referred to as the population mean of the random variable and denoted μ_X.

1.7.2. *First example computing an expected value.* Toss a die that has six sides. Observe the number that comes up. The probability mass or frequency function is given by

$$p\,(x) \;=\; P(X \;=\; x) \;=\begin{cases} \frac{1}{6} & for \; x \;=\; 1,\, 2,\, 3,\, 4,\, 5,\, 6 \\ 0 & otherwise \end{cases} \tag{11}$$

We compute the expected value as

$$\begin{aligned} E(X) \;&=\; \sum_{x \,\epsilon\, X} \; x \; p_X\,(x) \\ &= \sum_{i\,=\,1}^{6} i \; \left(\frac{1}{6}\right) \\ &= \frac{1 + \; 2 \;+\; 3 \;+\; 4 \;+\; 5 \;+\; 6}{6} \\ &= \frac{21}{6} \;=\; 3\;\frac{1}{2} \end{aligned} \tag{12}$$

1.7.3. *Second example computing an expected value.* Consider a group of 12 television sets, two of which have white cords and ten which have black cords. Suppose three of them are chosen at random and shipped to a care center. What are the probabilities that zero, one, or two of the sets with white cords are shipped? What is the expected number with white cords that will be shipped?

It is clear that x of the two sets with white cords and 3-x of the ten sets with black cords can be chosen in $\binom{2}{x} \times \binom{10}{3-x}$ ways. The three sets can be chosen in $\binom{12}{3}$ ways. So we have a probability mass function as follows.

$$f(x) \;=\; P(\,X \;=\; x) \;=\; \frac{\binom{2}{x}\,\binom{10}{3-x}}{\binom{12}{3}} \;\; for\; x \;=\; 0\,,\,1\,,\,2 \tag{13}$$

For example

$$f(x) \;=\; P(X \;=\; x) \;=\; \frac{\binom{2}{0}\,\binom{10}{3-0}}{\binom{12}{3}} \;=\; \frac{(1)\,(120)}{220} \;=\; \frac{6}{11} \tag{14}$$

We collect this information as in table 4.

TABLE 4. **Probabilities for Television Problem**

x	0	1	2
f(x)	6/11	9/22	1/22
F(x)	6/11	21/22	1

We compute the expected value as

$$\begin{aligned} E(X) \;&=\; \sum_{x\,\epsilon\,X} x\; p_X\,(x) \\ &= (0)\,\left(\frac{6}{11}\right) \,+\, (1)\,\left(\frac{9}{22}\right) \,+\, (2)\,\left(\frac{1}{22}\right) \;=\; \frac{11}{22} \;=\; \frac{1}{2} \end{aligned} \tag{15}$$

1.8. **Expected value of a function of a random variable.**

Theorem 1. *Let X be a discrete random variable with probability mass function p(x) and g(X) be a real-valued function of X. Then the expected value of g(X) is given by*

$$E[g(X)] \;=\; \sum_{x} \; g(x)\; p(x)\,. \tag{16}$$

Proof for case of finite values of X. Consider the case where the random variable X takes on a finite number of values $x_1, x_2, x_3, \cdots x_n$. The function g(x) may not be one-to-one (the different values of x_i may yield the same value of $g(x_i)$). Suppose that g(X) takes on m different values ($m \leq n$). It follows that g(X) is also a random variable with possible values g_1, g_2, g_2, . . . g_m and probability distribution

$$P[g(X) \;=\; g_i] \;=\; \sum_{\substack{\forall\, x_j such\; that \\ g(x_j)\,=\,g_i}} p(x_j) \;=\; p^*\,(g_i) \tag{17}$$

for all i = 1, 2, . . . m. Here $p*(g_i)$ is the probability that the experiment results in a value for the function f of the initial random variable of g_i. Using the definition of expected value in equation we obtain

$$E[g(X)] \ = \sum_{i\,=\,1}^{m} \ g_i \, p^*(g_i). \tag{18}$$

Now substitute in to obtain

$$\begin{aligned} E[g(X)] \ &= \sum_{i\,=\,1}^{m} \ g_i \, p^*(g_i)\,. \\ &= \sum_{i\,=\,1}^{m} \ g_i \left[\sum_{\substack{\forall\, x_j \,\ni \\ g\,(\,x_j\,)\,=\,g_i}} \quad p\,(\,x_j\,) \right] \\ &= \sum_{i\,=\,1}^{m} \left[\sum_{\substack{\forall\, x_j \,\ni \\ g\,(\,x_j\,)\,=\,g_i}} \quad g_i \ p\,(\,x_j\,) \right] \\ &= \sum_{j\,=\,1}^{n} \ \ g\,(x_j\,) \ \ p(\,x_j\,). \end{aligned} \tag{19}$$

□

1.9. **Properties of mathematical expectation.**

1.9.1. *Constants.*

Theorem 2. *Let X be a discrete random variable with probability function p(x) and c be a constant. Then E(c) = c.*

Proof. Consider the function g(X) = c. Then by theorem 1

$$E[c] \ \equiv \ \sum_{x} \ c\,p(x) \ = \ c \ \sum_{x} \ p(x) \tag{20}$$

But by property 1.4b, we have

$$\sum_{x} \ p(x) \ = \ 1$$

and hence

$$E\,(c) = c \cdot (1) = c. \tag{21}$$

□

1.9.2. *Constants multiplied by functions of random variables.*

Theorem 3. *Let X be a discrete random variable with probability function p(x), g(X) be a function of X, and let c be a constant. Then*

$$E\,[\,c\;g\,(\,X\,)\,]\;\equiv\;c\;E\;[\,(g\,(\,X\,)\,]\tag{22}$$

Proof. By theorem 1 we have

$$\begin{aligned} E[c\;g(X)] \;&\equiv \sum_x \;c\;g(x)\,p(x) \\ &= c\sum_x \;g(x)\;p(x) \\ &= c\;E[g(X)] \end{aligned}\tag{23}$$

□

1.9.3. *Sums of functions of random variables.*

Theorem 4. *Let X be a discrete random variable with probability function p(x),* $g_1(X), g_2(X), g_3(X), \cdots, g_k(X)$ *be k functions of X. Then*

$$E\,[g_1(X)+g_2(X)+g_3(X)+\cdots+g_k(X)]\;\equiv\;E[g_1(X)]+E[g_2(X)]+\cdots+E[g_k(X)]\tag{24}$$

Proof for the case of k = 2. By theorem 1 we have we have

$$\begin{aligned} E\,[g_1(X)\;+\;g_2\,(X)\,]\;&\equiv\sum_x\;[g_1\,(x)\;+\;g_2\,(x)\,]\;\;p(x) \\ &\equiv \sum_x\;g_1\,(x)\,p(x)\;+\;\sum_x\;g_2\,(x)\;p(x) \\ &= E\;[g_1\,(X)\,]\;+\;\;E\;[\,g_2\,(X)]\,, \end{aligned}\tag{25}$$

□

1.10. **Variance of a random variable.**

1.10.1. *Definition of variance.* The variance of a random variable X is defined to be the expected value of $(X-\mu)^2$. That is

$$V(X)\;=\;E\;\left[(\,X\;-\;\mu\,)^2\right]\tag{26}$$

The standard deviation of X is the positive square root of V(X).

1.10.2. *Example 1.* Consider a random variable with the following probability distribution.

TABLE 5. Probability Distribution for X

x	p(x)
0	1/8
1	1/4
2	3/8
3	1/4

We can compute the expected value as

$$\begin{aligned} \mu \;=\; E(X) \;&=\; \sum_{x=0}^{3} x\, p_X\,(x) \\ &= (0)\left(\frac{1}{8}\right) \;+\; (1)\left(\frac{1}{4}\right) \;+\; (2)\left(\frac{3}{8}\right) \;+\; (3)\left(\frac{1}{4}\right) \;= 1\,\frac{3}{4} \end{aligned} \tag{27}$$

We compute the variance as

$$\begin{aligned} \sigma^2 \;&= E[X \;-\; \mu)^2] \;=\; \Sigma_{x=0}^{3}\,(x \;-\; \mu)^2\, p_X(x) \\ &= (0 \;-\; 1.75)^2\left(\frac{1}{8}\right) \;+\; (1 \;-\; 1.75)^2\left(\frac{1}{4}\right) \;+\; (2 \;-\; 1.75)^2\left(\frac{3}{8}\right) \;+\; (3 \;-\; 1.75)^2\left(\frac{1}{4}\right) \\ &= .9375 \end{aligned}$$

and the standard deviation as

$$\begin{aligned} \sigma^2 \;&= 0.9375 \\ \sigma \;&=\; +\,\sqrt{\sigma}^2 \;=\; \sqrt{.9375} \;=\; 0.97. \end{aligned}$$

1.10.3. *Alternative formula for the variance.*

Theorem 5. *Let X be a discrete random variable with probability function $p_X\,(x)$; then*

$$V\,(X) \;\equiv\; \sigma^2 \;=\; E\left[(X \;-\; \mu\,)^2\right] \;=\; E\,\left(X^2\right) \;-\; \mu^2 \tag{28}$$

Proof. First write out the first part of equation 28 as follows

$$\begin{aligned} V\,(X) \;&\equiv \sigma^2 \;=\; E\left[(X \;-\; \mu\,)^2\right] \;=\; E\left(X^2 \;-\; 2\,\mu\,X \;+\; \mu^2\,\right) \\ &= E\left(X^2\right) \;-\; E\,(2\,\mu\,X) \;+\; E\left(\mu^2\right) \end{aligned} \tag{29}$$

where the last step follows from theorem 4. Note that μ is a constant then apply theorems 3 and 2 to the second and third terms in equation 28 to obtain

$$V\,(X) \;\equiv \sigma^2 \;=\; E\left[(\,X \;-\; \mu\,)^2\,\right] \;=\; E\left(X^2\right) \;-\; 2\,\mu\,E\,(X) \;+\; \mu^2 \tag{30}$$

Then making the substitution that E(X) = μ, we obtain

$$V\,(X) \;\equiv \sigma^2 \;=\; E\left(X^2\right) \;-\; \mu^2 \tag{31}$$

□

1.10.4. *Example 2.* Die toss.

Toss a die that has six sides. Observe the number that comes up. The probability mass or frequency function is given by

$$p(x) \;=\; P(X \;=\; x) \;=\; \begin{cases} \frac{1}{6} & for\;x \;=\; 1,\,2,\,3,\,4,\,5,\,6 \\ 0 & otherwise \end{cases}. \tag{32}$$

We compute the expected value as

$$\begin{aligned}
E(X) &= \sum_{x \,\epsilon\, X} x \, p_X\,(x) \\
&= \sum_{i\,=\,1}^{6} i \, \left(\frac{1}{6}\right) \\
&= \frac{1+\,2\,+\,3\,+\,4\,+\,5\,+\,6}{6} \\
&= \frac{21}{6} \;=\; 3\,\frac{1}{2}
\end{aligned} \tag{33}$$

We compute the variance by then computing the $E(X^2)$ as follows

$$\begin{aligned}
E(X^2) &= \sum_{x \,\epsilon\, X} x^2 \; p_X\,(x) \\
&= \sum_{i\,=\,1}^{6} i^2 \, \left(\frac{1}{6}\right) \\
&= \frac{1+\,4\,+\,9\,+\,16\,+\,2\,+\,36}{6} \\
&= \frac{91}{6} \;=\; 15\,\frac{1}{6}
\end{aligned} \tag{34}$$

We can then compute the variance using the formula Var(X) = E(X^2) - E^2(X) and the fact the E(X) = 21/6 from equation 33.

$$\begin{aligned}
Var(X) &= E\,(X^2)\;-\;E^2(X) \\
&= \frac{91}{6}\;-\;\left(\frac{21}{6}\right)^2 \\
&= \frac{91}{6}\;-\;\left(\frac{441}{36}\right) \\
&= \frac{546}{36}\;-\;\frac{441}{36} \\
&= \frac{105}{36}\;=\;\frac{35}{12}\;=\;2.916\overline{6}
\end{aligned} \tag{35}$$

2. The "Distribution" of Random Variables in General

2.1. Cumulative distribution function. The cumulative distribution function (cdf) of a random variable X, denoted by $F_X(\cdot)$, is defined to be the function with domain the real line and range the interval [0,1], which satisfies $F_X(x) = P_X[X \leq x] = P[\{\omega : X(\omega) \leq x\}]$ for every real number x. F has the following properties:

$$F_X(-\infty) = \lim_{x \to -\infty} F_X(x) = 0, \; F_X(+\infty) = \lim_{x \to +\infty} F_X(x) = 1, \tag{36a}$$

$$F_X(a) \leq F_X(b) \; for \; a < b, \; nondecreasing \; function \; of \; x, \tag{36b}$$

$$\lim_{0<h \to 0} F_X(x+h) = F_X(x), \; continuous \; from \; the \; right, \tag{36c}$$

2.2. Example of a cumulative distribution function. Consider the following function

$$F_X(x) = \frac{1}{1 + e^{-x}} \tag{37}$$

Check condition 36a as follows.

$$\begin{aligned} \lim_{x \to -\infty} F_X(x) &= \lim_{x \to -\infty} \frac{1}{1 + e^{-x}} = \lim_{x \to \infty} \frac{1}{1 + e^{x}} = 0 \\ \lim_{x \to \infty} F_X(x) &= \lim_{x \to \infty} \frac{1}{1 + e^{-x}} = 1 \end{aligned} \tag{38}$$

To check condition 36b differentiate the cdf as follows

$$\begin{aligned} \frac{d F_X(x)}{dx} &= \frac{d\left(\frac{1}{1+e^{-x}}\right)}{dx} \\ &= \frac{e^{-x}}{(1 + e^{-x})^2} > 0 \end{aligned} \tag{39}$$

Condition 36c is satisfied because $F_X(x)$ is a continuous function.

2.3. Discrete and continuous random variables.

2.3.1. *Discrete random variable.* A random variable X will be said to be discrete if the range of X is countable, that is if it can assume only a finite or countably infinite number of values. Alternatively, a random variable is discrete if $F_X(x)$ is a step function of x.

2.3.2. *Continuous random variable.* A random variable X is continuous if $F_X(x)$ is a continuous function of x.

2.4. Frequency (probability mass) function of a discrete random variable.

2.4.1. *Definition of a frequency (discrete density) function.* If X is a discrete random variable with the distinct values, $x_1, x_2, \cdots, x_n, \cdots$, then the function denoted by $p(\cdot)$ and defined by

$$p(x) = \begin{cases} P[X = x_j] & x = x_j, \;\; j = 1, 2, ..., n, ... \\ 0 & x \neq x_j \end{cases} \tag{40}$$

is defined to be the frequency, discrete density, or probability mass function of X. We will often write f(x) for p(x) to denote frequency as compared to probability.

A discrete probability distribution on R^k is a probability measure P such that

$$\sum_{i=1}^{\infty} P(\{x_i\}) = 1 \tag{41}$$

for some sequence of points in R^k, i.e. the sequence of points that occur as an outcome of the experiment. Given the definition of the frequency function in equation 40, we can also say that any non-negative function p on R^k that vanishes except on a sequence $x_1, x_2, \cdots, x_n, \cdots$ of vectors and that satisfies

$$\sum_{i=1}^{\infty} p(x_i) = 1$$

defines a unique probability distribution by the relation

$$P(A) = \sum_{x_i \,\epsilon\, A} p(x_i) \tag{42}$$

2.4.2. *Properties of discrete density functions.* As defined in section 1.4, a probability mass function must satisfy

$$p(x_j) > 0 \; for \; j = 1, 2, ... \tag{43a}$$

$$p(x) = 0 \; for x \neq x_j \, ; \; j = 1, 2, ..., \tag{43b}$$

$$\sum_j p(x)_j = 1 \tag{43c}$$

2.4.3. *Example 1 of a discrete density function.* Consider a probability model where there are two possible outcomes to a single action (say heads and tails) and consider repeating this action several times until one of the outcomes occurs. Let the random variable be the number of actions required to obtain a particular outcome (say heads). Let p be the probability that outcome is a head and (1-p) the probability of a tail. Then to obtain the first head on the xth toss, we need to have a tail on the previous x-1 tosses. So the probability of the first had occurring on the xth toss is given by

$$p(x) = P(X = x) = \begin{cases} (1 - p)^{x-1} p & for \; x = 1, 2, \dots \\ 0 & otherwise \end{cases} \tag{44}$$

For example the probability that it takes 4 tosses to get a head is 1/16 while the probability it takes 2 tosses is 1/4.

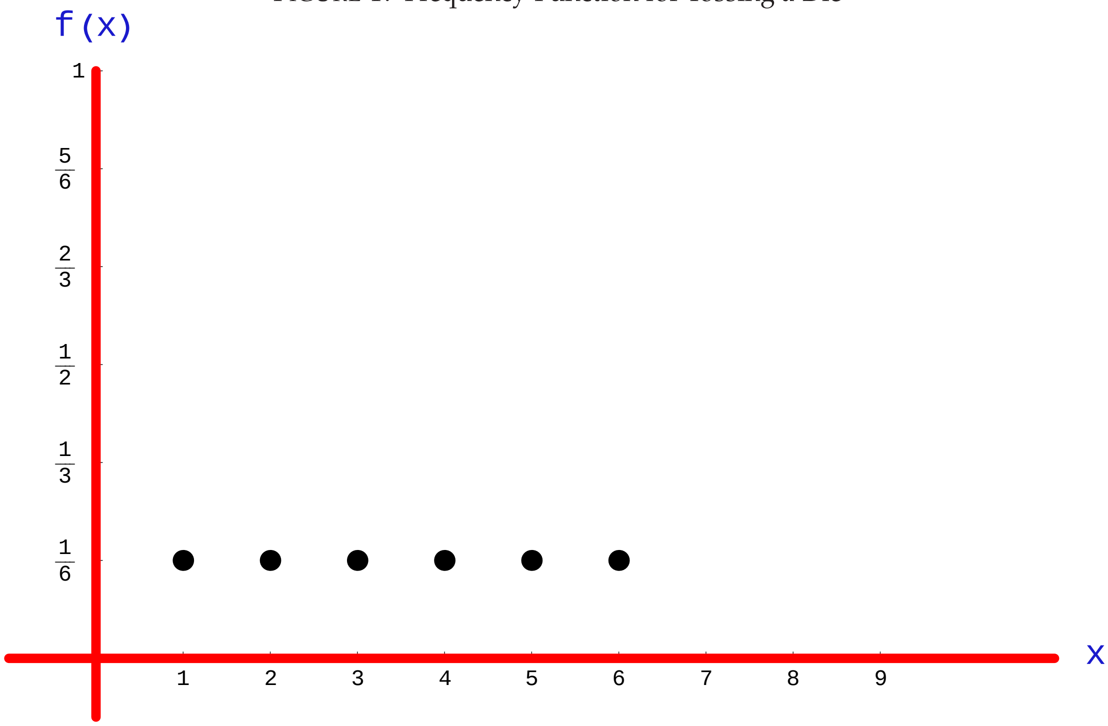
2.4.4. *Example 2 of a discrete density function.* Consider tossing a die. The sample space is {1, 2, 3, 4, 5, 6}. The elements are {1}, {2}, ... The frequency function is given by

$$p(x) = P(X = x) = \begin{cases} \frac{1}{6} & for \; x = 1, 2, 3, 4, 5, 6 \\ 0 & otherwise \end{cases} . \tag{45}$$

The density function is represented in figure 1.

2.5. **Probability density function of a continuous random variable.**

FIGURE 1. Frequency Function for Tossing a Die



2.5.1. *Alternative definition of continuous random variable.* In section 2.3.2, we defined a random variable to be continuous if $F_X(x)$ is a continuous function of x. We also say that a random variable X is continuous if there exists a function f(·) such that

$$F_X(x) \; = \; \int_{-\infty}^{x} f(u)\,du \tag{46}$$

for every real number x. The integral in equation 46 is a Riemann integral evaluated from -∞ to a real number x.

2.5.2. *Definition of a probability density frequency function (pdf).* The probability density function, $f_X(x)$, of a continuous random variable X is the function $f(\cdot)$ that satisfies

$$F_X(x) \; = \; \int_{-\infty}^{x} f_X(u)\,du \tag{47}$$

2.5.3. *Properties of continuous density functions.*

$$f(x) \; \geq \; 0 \; \forall x \tag{48a}$$

$$\int_{-\infty}^{\infty} f(x)\,dx \; = \; 1, \tag{48b}$$

Analogous to equation 42, we can write in the continuous case

$$P(X \,\epsilon\, A) \; = \; \int_{A} f_X(x) \; dx \tag{49}$$

where the integral is interpreted in the sense of Lebesgue.

Theorem 6. *For a density function f(x) defined over the set of all real numbers the following holds*

$$P(a \leq X \leq b) = \int_a^b f_X(x)\, dx \tag{50}$$

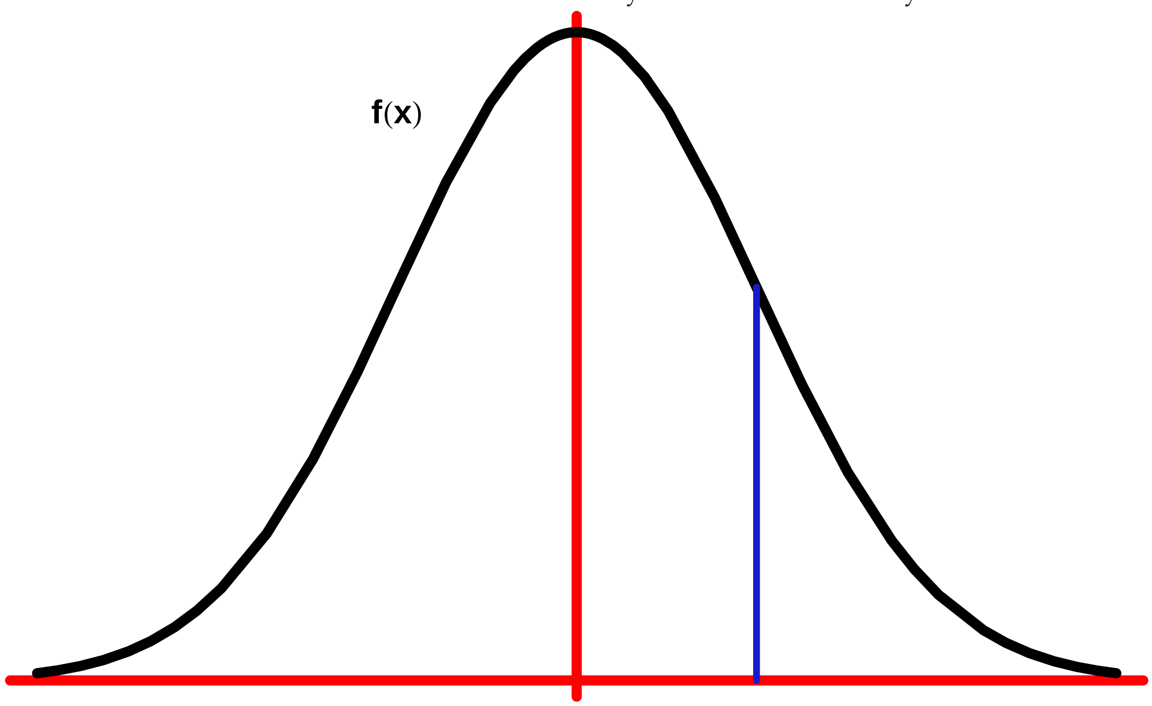
for any real constants a and b with $a \leq b$.

Also note that for a continuous random variable X the following are equivalent

$$P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b) \tag{51}$$

Note that we can obtain the various probabilities by integrating the area under the density function as seen in figure 2.

FIGURE 2. Area under the Density Function as Probability



2.5.4. *Example 1 of a continuous density function.* Consider the following function

$$f(x) = \begin{cases} k \cdot e^{-3x} & for\ x > 0 \\ 0 & elsewhere \end{cases}. \tag{52}$$

First we must find the value of k that makes this a valid density function?
Given the condition in equation 48b we must have that

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_0^{\infty} k \cdot e^{-3x}\, dx = 1 \tag{53}$$

Integrate the second term to obtain

$$\int_0^{\infty} k \cdot e^{-3x}\, dx = k \cdot \lim_{t \to \infty} \frac{e^{-3x}}{-3} \Big|_0^t = \frac{k}{3} \tag{54}$$

Given that this must be equal to one we obtain

$$\begin{aligned} \frac{k}{3} &= 1 \\ \Rightarrow k &= 3 \end{aligned} \tag{55}$$

The density is then given by

$$f(x) = \begin{cases} 3 \cdot e^{-3x} & for\ x > 0 \\ 0 & elsewhere \end{cases}. \tag{56}$$

Now find the probability that ($1 \leq X \leq 2$).

$$\begin{aligned} P(1 \leq X \leq 2) &= \int_1^2 3 \cdot e^{-3x}\, dx \\ &= -e^{-3x} \Big|_1^2 \\ &= -e^{-6} + e^{-3} \\ &= -0.00247875 + 0.049787 \\ &= 0.047308 \end{aligned} \tag{57}$$

2.5.5. *Example 2 of a continuous density function.* Let X have p.d.f.

$$f(x) = \begin{cases} x \cdot e^{-x} & for\ x \leq x \leq \infty \\ 0 & elsewhere \end{cases}. \tag{58}$$

This density function is shown in figure 3.

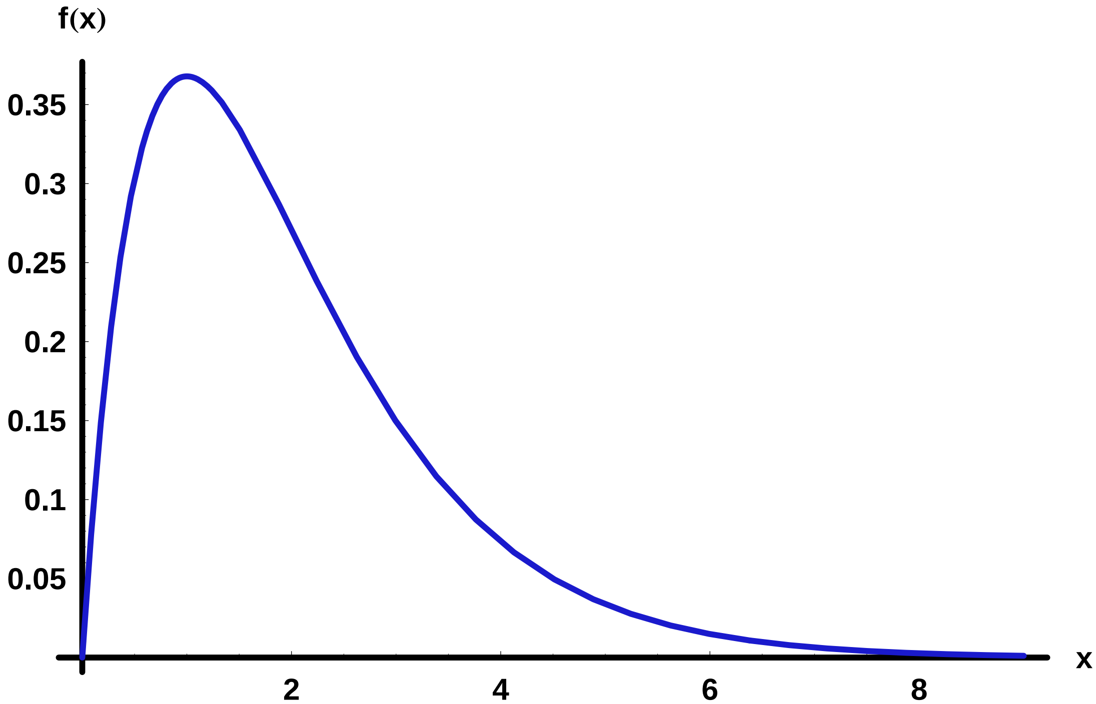
We can find the probability that ($1 \leq X \leq 2$) by integration

$$P(1 \leq X \leq 2) = \int_1^2 x \cdot e^{-x}\, dx \tag{59}$$

First integrate the expression on the right by parts letting u = x and dv = e^{-x} dx. Then du = dx and v = - e^{-x}dx. We then have

$$\begin{aligned} P(1 \leq X \leq 2) &= -x\,e^{-x} \Big|_1^2 - \int_1^2 -e^{-x}\, dx \\ &= -2\,e^{-2} + e^{-1} - \left[e^{-x} \Big|_1^2 \right] \\ &= -2\,e^{-2} + e^{-1} - e^{-2} + e^{-1} \\ &= -3\,e^{-2} + 2\,e^{-1} \\ &= \frac{-3}{e^2} + \frac{2}{e} \\ &= -0.406 + 0.73575 \\ &= 0.32975 \end{aligned} \tag{60}$$

FIGURE 3. Graph of Density Function $x\, e^{-x}$



This is represented by the area between the lines in figure 4.
We can also find the distribution function in this case.

$$F(x) \;=\; \int_0^x \; t \cdot e^{-t}\, dt \tag{61}$$

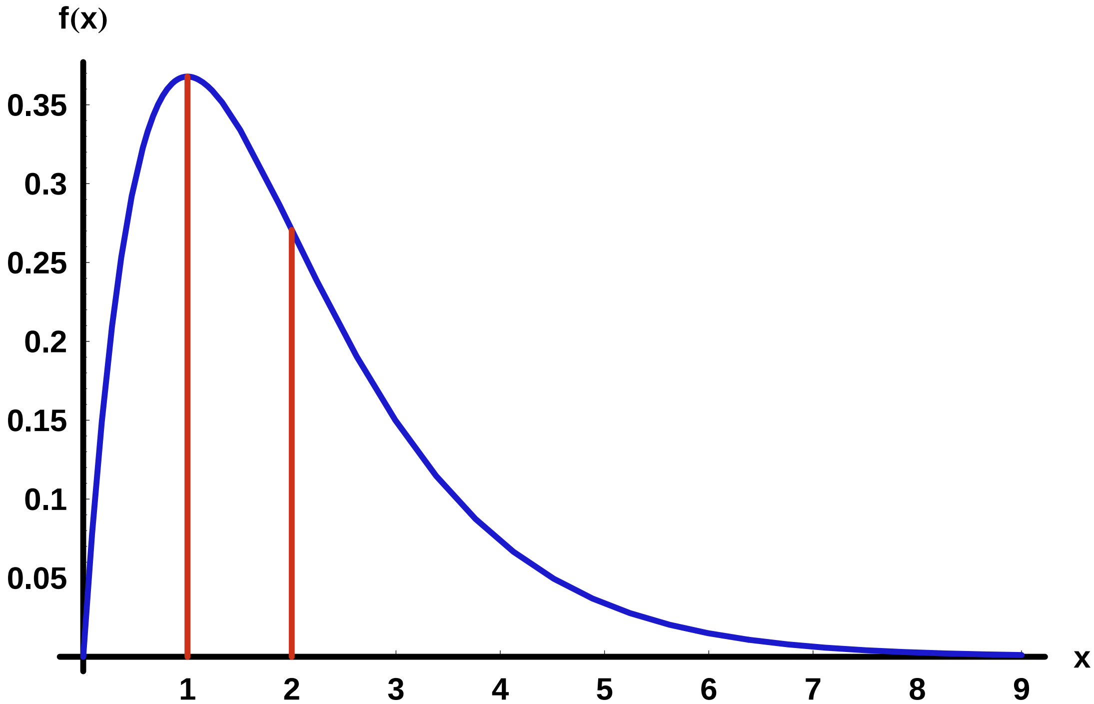
Make the u dv substitution as before to obtain

$$\begin{aligned} F(x) \;&=\; -\,t\,e^{-t}\,|_0^x \;-\; \int_0^x \; -e^{-t}\, dt \\ &=\; -\,t\,e^{-t}\,|_0^x \;-\; e^{-t}|_0^x \\ &= e^{-t}\,(-1 \;-\; t)|_0^x \\ &= e^{-x}\,(-1 \;-\; x) \;-\; e^{-0}\,(-1 \;-\; 0) \\ &= e^{-x}\,(-1 \;-\; x) \;+\; 1 \\ &= 1 \;-\; e^{-x}\,(1 \;+\; x) \end{aligned} \tag{62}$$

The distribution function is shown in figure 5.

Now consider the probability that $(1 \leq \mathrm{X} \leq 2)$

FIGURE 4. $P(1 \leq X \leq 2)$



$$
\begin{aligned}
P(1 \leq X \leq 2) &= F(2) - F(1) \\
&= 1 - e^{-2}(1 + 2) - 1 + e^{-1}(1 + 1) \\
&= 2e^{-1} - 3e^{-2} \\
&= 0.73575 - 0.406 \\
&= 0.32975
\end{aligned} \tag{63}
$$

We can see this as the difference in the values of F(x) at 1 and at 2 in figure 6

2.5.6. *Example 3 of a continuous density function.* Consider the normal density function given by

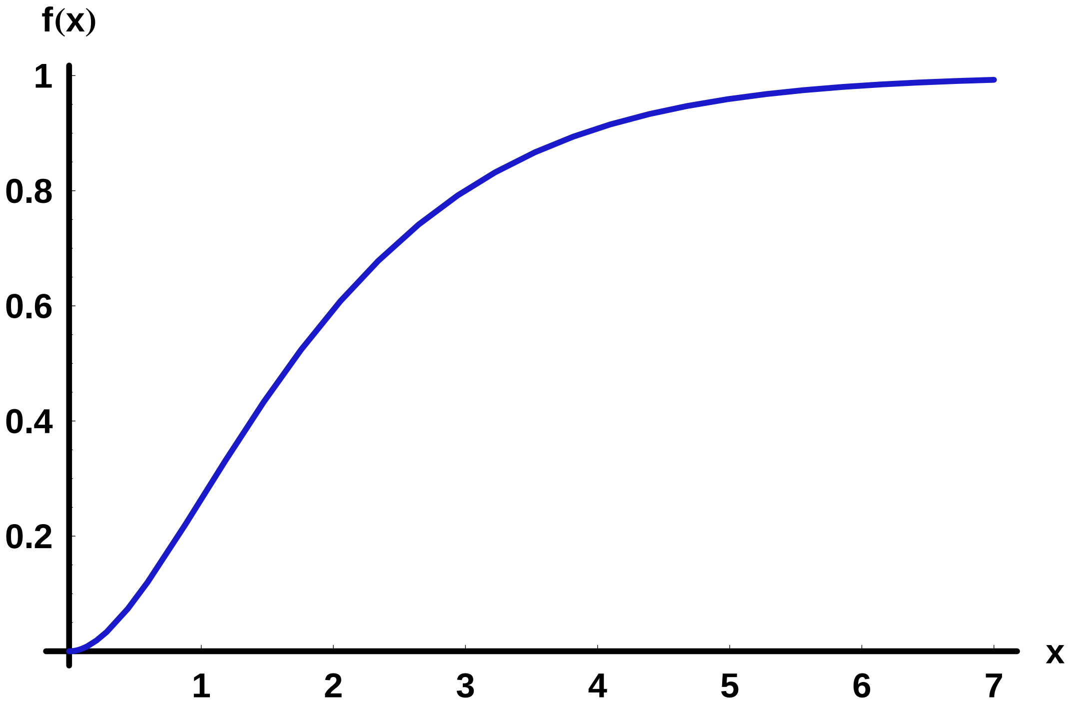
$$
f(x : \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-1}{2}\left(\frac{x - \mu}{\sigma}\right)^2} \tag{64}
$$

where μ and σ are parameters of the function. The shape and location of the density function depends on the parameters μ and σ. In figure 7 the diagram the density is drawn for μ = 0, and σ = 1 and σ = 2.

2.5.7. *Example 4 of a continuous density function.* Consider a random variable with density function given by

$$
f(x) = \begin{cases} (p+1)x^p & 0 \leq x \leq 1 \\ 0 & otherwise \end{cases} \tag{65}
$$

FIGURE 5. Graph of Distribution Function of Density Function $x\,e^{-x}$



where p is greater than -1. For example, if p = 0, then f(x) = 1, if p = 1, then f(x) = 2x and so on. The density function with p = 2 is shown in figure 8.

The distribution function with p = 2 is shown in figure 9.

2.6. **Expected value.**

2.6.1. *Expectation of a single random variable.* Let X be a random variable with density f(x). The expected value of the random variable, denoted E(X), is defined to be

$$E(X) \;=\; \begin{cases} \int_{-\infty}^{\infty} x\, f(x)\, dx & if\ X\ is\ continuous \\ \sum_{x\,\epsilon\,X}\ x\, p_X\,(x) & if\ X\ is\ discrete \end{cases}. \tag{66}$$

provided the sum or integral is defined. The expected value is kind of a weighted average. It is also sometimes referred to as the population mean of the random variable and denoted μ_X.

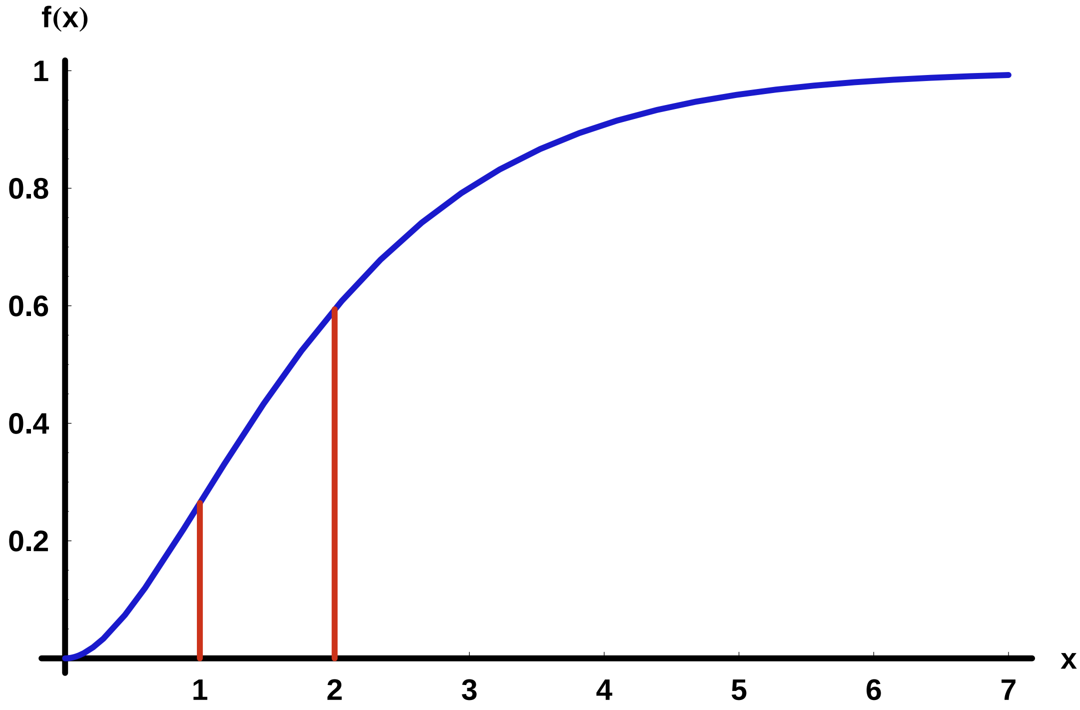
2.6.2. *Expectation of a function of a single random variable.* Let X be a random variable with density f(X). The expected value of a function g(·) of the random variable, denoted E(g(X)), is defined to be

$$E(g(X)) \;=\; \int_{-\infty}^{\infty} g(x)\, f\,(x) dx \tag{67}$$

if the integral is defined.

The expectation of a random variable can also be defined using the Riemann-Stieltjes integral where F is a monotonically increasing function of X. Specifically

FIGURE 6. $P(1 \leq X \leq 2)$ using the Distribution Function



$$E(X) \;=\; \int_{-\infty}^{\infty} x \, dF(x) \;=\; \int_{-\infty}^{\infty} x \, dF \tag{68}$$

2.7. **Properties of expectation.**

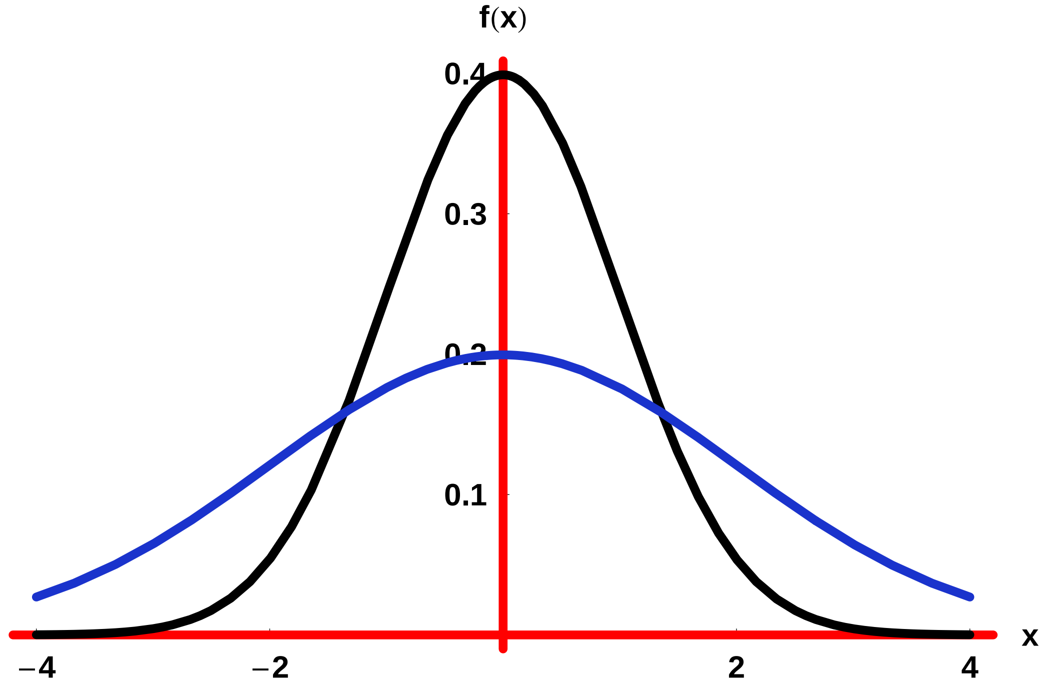
2.7.1. *Constants.*

$$\begin{aligned} E[a] \;&\equiv \int_{-\infty}^{\infty} a\, f(x) dx \\ &\equiv a \int_{-\infty}^{\infty} f(x) dx \\ &\equiv a \end{aligned} \tag{69}$$

2.7.2. *Constants multiplied by a random variable.*

$$\begin{aligned} E[a\, X] \;&\equiv \int_{-\infty}^{\infty} a\, x\, f(x) dx \\ &\equiv a \int_{-\infty}^{\infty} x\, f(x) dx \\ &\equiv a\, E[X] \end{aligned} \tag{70}$$

FIGURE 7. Normal Density Function



2.7.3. *Constants multiplied by a function of a random variable.*

$$\begin{aligned} E[a\,g(X)] &\equiv \int_{-\infty}^{\infty} a\ g(x)\ f(x)dx \\ &\equiv a\ \int_{-\infty}^{\infty} g(x)\, f(x)dx \\ &\equiv a\ E[g(X)] \end{aligned} \tag{71}$$

2.7.4. *Sums of expected values.* Let X be a continuous random variable with density function f(x) and let $g_1\,(X), g_2\,(X), g_3\,(X), \cdots, g_k\,(X)$ be k functions of X. Also let $c_1,\ c_2,\ c_3,\ \cdots c_k$ be k constants. Then

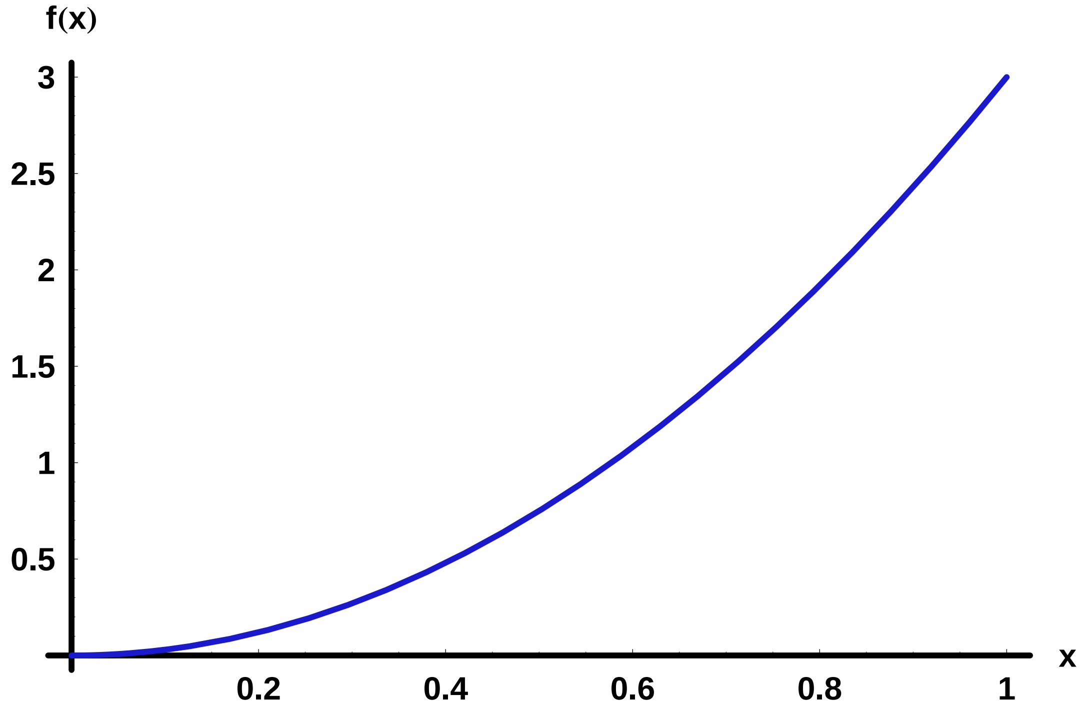
$$E\left[c_1\, g_1(X) + c_2\, g_2(X) + \cdots + c_k\, g_k(X)\ \right] \equiv\ E\left[c_1\, g_1(X)\right] + E\left[c_2\, g_2(X)\right]\ +\ \cdots\ +\ E\left[c_k\, g_k(X)\right] \tag{72}$$

2.8. **Example 1.** Consider the density function

$$f(x)\ = \begin{cases} (p+1)x^p & 0\ \leq\ x\ \leq\ 1 \\ 0 & otherwise \end{cases} \tag{73}$$

where p is greater than -1. We can compute the E(X) as follows.

FIGURE 8. Density Function $(p+1)\,x^p$



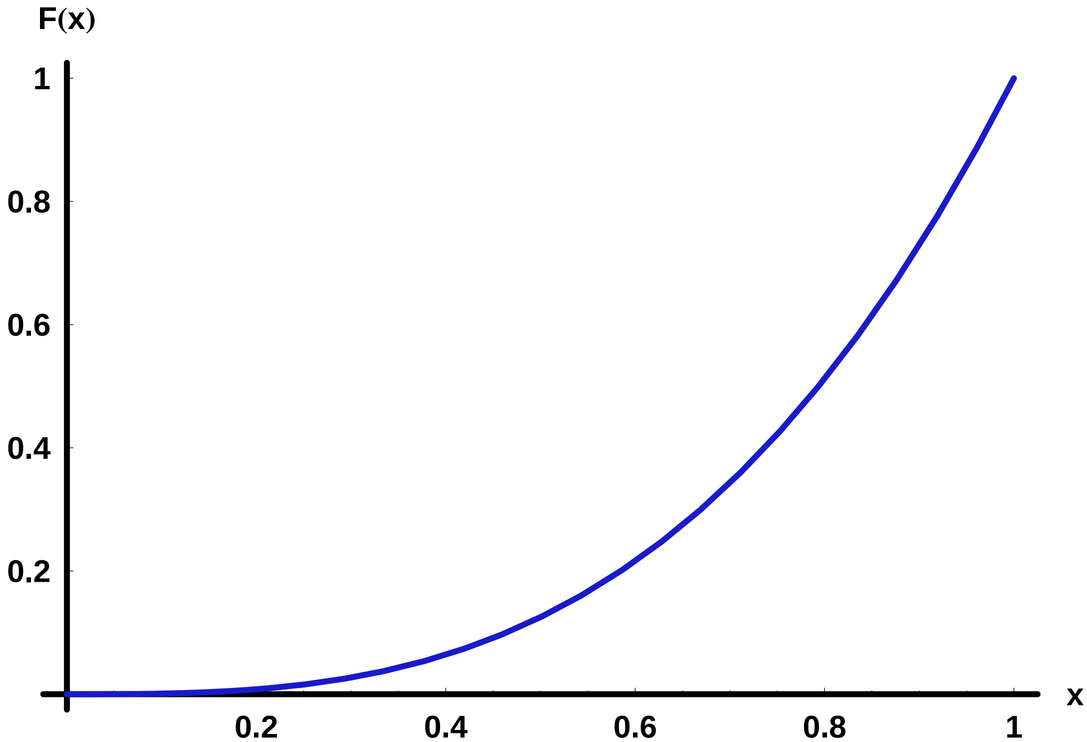
$$
\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x\, f(x) dx \\
&= \int_0^1 x(p+1)x^p dx \\
&= \int_0^1 x^{(p+1)}(p+1) dx \\
&= \frac{x^{(p+2)}(p+1)}{(p+2)} \mid {}_0^1 \\
&= \frac{p+1}{p+2}
\end{aligned} \tag{74}
$$

2.9. **Example 2.** Consider the exponential distribution which has density function

$$
f(x) \;=\; \frac{1}{\lambda}\, e^{\frac{-x}{\lambda}} \; 0 \;\leq\; x \;\leq\; \infty\,,\; \lambda \;>\; 0 \tag{75}
$$

We can compute the E(X) as follows.

FIGURE 9. Density Function $(p = 1)\, x^p$



$$
\begin{aligned}
E(X) &= \int_0^\infty x \tfrac{1}{\lambda}\, e^{\frac{-x}{\lambda}}\, dx \\
&= -\, x\, e^{\frac{-x}{\lambda}}\, |_0^\infty \; + \; \int_0^\infty e^{\frac{-x}{\lambda}}\, dx \; \left(u \;=\; \tfrac{x}{\lambda},\; du \;=\; \tfrac{1}{\lambda}\, dx,\; v = -\,\lambda\, e^{\frac{-x}{\lambda}},\; dv \;=\; e^{\frac{-x}{\lambda}}\, dx \right) \\
&= 0 \; + \; \int_0^\infty e^{\frac{-x}{\lambda}}\, dx \\
&= -\,\lambda\, e^{\frac{-x}{\lambda}}\, |_0^\infty \\
&= \lambda
\end{aligned}
\tag{76}
$$

2.10. **Variance.**

2.10.1. *Definition of variance.* The variance of a single random variable X with mean μ is given by

$$
\begin{aligned}
Var(X) \;\equiv\; \sigma^2 \;\equiv\; E\left[(X \;-\; E\,(X))^2 \right] \\
\equiv E\, \left[(\, X \;-\; \mu)^2 \right] \\
\equiv \int_{-\infty}^{\infty} \, (x \;-\; \mu)^2\, f(x) dx
\end{aligned}
\tag{77}
$$

We can write this in a different fashion by expanding the last term in equation 77.

$$\begin{aligned}
Var(X) &\equiv \int_{-\infty}^{\infty} (x - \mu)^2 f(x)dx \\
&\equiv \int_{-\infty}^{\infty} (x^2 - 2\mu x + \mu^2) f(x)dx \\
&\equiv \int_{-\infty}^{\infty} x^2 f(x)\, dx - 2\mu \int_{-\infty}^{\infty} x\, f(x)\, dx + \mu^2 \int_{-\infty}^{\infty} f(x)\, dx \\
&= E\left[X^2\right] - 2\mu E\,[X] + \mu^2 \\
&= E\left[X^2\right] - 2\mu^2 + \mu^2 \\
&= E\left[X^2\right] - \mu^2 \\
&\equiv \int_{-\infty}^{\infty} x^2 f(x)dx - \left[\int_{-\infty}^{\infty} x\, f(x)dx\right]^2
\end{aligned} \tag{78}$$

The variance is a measure of the dispersion of the random variable about the mean.

2.10.2. *Variance example 1.* Consider the density function

$$f(x) = \begin{cases} (p+1)x^p & 0 \le x \le 1 \\ 0 & otherwise \end{cases} \tag{79}$$

where p is greater than -1. We can compute the Var(X) as follows.

$$\begin{aligned}
E(X) &= \int_{-\infty}^{\infty} x\, f(x)dx \\
&= \int_0^1 x(p+1)x^p dx \\
&= \frac{x^{(p+2)}(p+1)}{(p+2)}\Big|_0^1 \\
&= \frac{p+1}{p+2} \\
E(X^2) &= \int_0^1 x^2\,(p+1)x^p\, dx \\
&= \frac{x^{(p+3)}(p+1)}{(p+3)}\Big|_0^1 \\
&= \frac{p+1}{p+3} \\
Var(X) &= E(X^2) - E^2(X) \\
&= \frac{p+1}{p+3} - \left(\frac{p+1}{p+2}\right)^2 \\
&= \frac{p+1}{(p+2)^2\,(p+3)}
\end{aligned} \tag{80}$$

The values of the mean and variances for various values of p are given in table 6.

TABLE 6. Mean and Variance for Distribution $f(x) = (p+1)\, x^p$ for alternative values of p

p	-.5	0	1	2	∞
E(x)	0.333	0.5	0.66667	0.75	1
Var(x)	0.08888	0.833333	0.277778	0.00047	0

2.10.3. *Variance example 2.* Consider the exponential distribution which has density function

$$f(x) = \frac{1}{\lambda}\, e^{\frac{-x}{\lambda}} \; 0 \le x \le \infty\,,\; \lambda > 0 \tag{81}$$

We can compute the $E(X^2)$ as follows

$$\begin{aligned}
E(X^2) &= \int_0^\infty x^2\, \frac{1}{\lambda}\, e^{\frac{-x}{\lambda}}\, dx \\
&= -x^2 e^{\frac{-x}{\lambda}}\, |_0^\infty + 2 \int_0^\infty x\, e^{\frac{-x}{\lambda}}\, dx \; \left(u = \frac{x^2}{\lambda}\,, du = \frac{2\,x}{\lambda}\, dx, v = -\,\lambda\, e^{\frac{-x}{\lambda}}, dv = e^{\frac{-x}{\lambda}}\, dx \right) \\
&= 0 + 2 \int_0^\infty x\, e^{\frac{-x}{\lambda}}\, dx \\
&= -\,2\,\lambda\, x\, e^{\frac{-x}{\lambda}}\, |_0^\infty + 2 \int_0^\infty \lambda\, e^{\frac{-x}{\lambda}}\, dx \; \left(u = 2\,x\,,\; du = 2\,dx,\; v = -\,\lambda\, e^{\frac{-x}{\lambda}}, dv = e^{\frac{-x}{\lambda}}\, dx \right) \\
&= 0 \;+\; 2\,\lambda \int_0^\infty e^{\frac{-x}{\lambda}}\, dx \\
&= (2\,\lambda) \left(-\,\lambda\, e^{\frac{-x}{\lambda}}\, |_0^\infty \right) \\
&= (2\,\lambda)\,(\,\lambda\,) \\
&= 2\,\lambda^2
\end{aligned} \tag{82}$$

We can then compute the variance as

$$\begin{aligned}
Var(X) &= E\,(X^2) \;-\; E^2\,(X) \\
&= 2\,\lambda^2 \;-\; \lambda^2 \\
&= \lambda^2
\end{aligned} \tag{83}$$

3. MOMENTS AND MOMENT GENERATING FUNCTIONS

3.1. **Moments.**

3.1.1. *Moments about the origin (raw moments).* The rth moment about the origin of a random variable X, denoted by μ'_r, is the expected value of X^r; symbolically,

$$\begin{aligned}
\mu'_r &= E(X^r) \\
&= \sum_x x^r\, f(x)
\end{aligned} \tag{84}$$

for r = 0, 1, 2, . . . when X is discrete and

$$\begin{aligned} \mu_r' &= E(\,X^r\,) \\ &= \int_{-\infty}^{\infty} x^r f(x)\,dx \end{aligned} \tag{85}$$

when X is continuous. The rth moment about the origin is only defined if $E[X^r]$ exists. A moment about the origin is sometimes called a raw moment. Note that $\mu_1' = E(X) = \mu_X$, the mean of the distribution of X, or simply the mean of X. The rth moment is sometimes written as function of θ where θ is a vector of parameters that characterize the distribution of X.

3.1.2. *Central moments.* The rth moment about the mean of a random variable X, denoted by μ_r, is the expected value of $(X - \mu_X)^r$ symbolically,

$$\begin{aligned} \mu_r \; &= E[(X - \mu_X)^r] \\ &= \sum_x \, (x - \mu_X)^r f(x) \end{aligned} \tag{86}$$

for r = 0, 1, 2, . . . when X is discrete and

$$\begin{aligned} \mu_r \; &= E[(X - \mu_X)^r] \\ &= \int_{-\infty}^{\infty} (x - \mu_X\,)^r f(x)\,dx \end{aligned} \tag{87}$$

when X is continuous. The rth moment about the mean is only defined if $E[(X - \mu_X)^r]$ exists. The rth moment about the mean of a random variable X is sometimes called the rth central moment of X. The rth central moment of X about a is defined as $E[(X - a)^r]$. If a = μ_X, we have the rth central moment of X about μ_X. Note that μ_1 = $E[(X - \mu_X)] = 0$ and μ_2 = $E[(X - \mu_X)^2]$ = Var[X]. Also note that all odd moments of X around its mean are zero for symmetrical distributions, provided such moments exist.

3.1.3. *Alternative formula for the variance.*

Theorem 7.

$$\sigma_X^2 \;=\; \mu_2' \;-\; \mu_X^2 \tag{88}$$

Proof.

$$\begin{aligned} Var(X) \;\equiv\; \sigma_X^2 \;&\equiv\; E\left[(X \;-\; E(X)\,)^2\right] \\ &\equiv E\left[(X \;-\; \mu_X)^2\right] \\ &\equiv E\left[X^2 \;-\; 2\,\mu_X\,X \;+\; \mu_X^2\right] \\ &= E\left[X^2\right] \;-\; 2\,\mu_X E\;[X\,] \;+\; \mu^2 \\ &= E\left[X^2\right] \;-\; 2\,\mu_X^2 \;+\; \mu_X^2 \\ &= E\left[X^2\right] \;-\; \mu_X^2 \\ &= \mu_2' \;-\; \mu_X^2 \end{aligned} \tag{89}$$

□

3.2. **Moment generating functions.**

3.2.1. *Definition of a moment generating function.* The moment generating function of a random variable X is given by

$$M_X(t) \;=\; E\, e^{t\,X} \tag{90}$$

provided that the expectation exists for t in some neighborhood of 0. That is, there is an $h > 0$ such that, for all t in $-h < t < h$, $E\, e^{tX}$ exists. We can write $M_X(t)$ as

$$M_X(\,t\,) \;=\; \begin{cases} \int_{-\infty}^{\infty} \, e^{t\,x}\, f_X\,(x)\; dx & if \;X\; is\; continuous \\ \sum_x \, e^{t\,x}\, P\;(X \;=\; x) & if\; X\; is\; discrete \end{cases}. \tag{91}$$

To understand why we call this a moment generating function consider first the discrete case. We can write e^{tx} in an alternative way using a Maclaurin series expansion. The Maclaurin series of a function f(t) is given by

$$\begin{aligned} f(t) \;&=\; \sum_{n\,=\,0}^{\infty} \frac{f^{(n)}\,(0)}{n\,!}\, t^n \;=\; \sum_{n\,=\,0}^{\infty} f^{(n)}\,(0)\, \frac{t^n}{n\,!} \\ &= f\;(\,0\,) \;+\; \frac{f^{(1)}\;(0)}{1\,!}\; t \;+\; \frac{f^{\,(\,2\,)}\;(0)}{2\,!}\, t^2 \;+\; \frac{f^{(3)}\,(0)}{3\,!}\; t^3 \;+\; \cdots \;+ \\ &= f\;(\,0\,) \;+\; f^{(1)}(0)\,\frac{t}{1\,!} \;+\; f^{\,(\,2\,)}\,(0)\,\frac{t^2}{2\,!} \;+\; f^{(3)}(0)\,\frac{t^3}{3\,!} \;+\; \cdots \;+ \end{aligned} \tag{92}$$

where $f^{(n)}$ is the nth derivative of the function with respect to t and $f^{(n)}(0)$ is the nth derivative of f with respect to t evaluated at t = 0. For the function e^{tx}, the requisite derivatives are

$$\begin{aligned} \frac{d\,e^{t\,x}}{d\,t} \;&=\; x\,e^{t\,x}\,, \quad \left.\frac{d\,e^{t\,x}}{d\,t}\right]_{\,t\,=\,0} \;=\; x \\ \frac{d^2\,e^{t\,x}}{d\,t^2} \;&=\; x^2\,e^{t\,x}\,, \quad \left.\frac{d^2\,e^{t\,x}}{d\,t^2}\right]_{\,t\,=\,0} \;=\; x^2 \\ \frac{d^3\,e^{t\,x}}{d\,t^3} \;&=\; x^3\,e^{t\,x}\,, \quad \left.\frac{d^3\,e^{t\,x}}{d\,t^3}\right]_{\,t\,=\,0} \;=\; x^3 \\ &\;\;\vdots \\ \frac{d^j\,e^{t\,x}}{d\,t^j} \;&=\; x^j\,e^{t\,x}\,, \quad \left.\frac{d^j\,e^{t\,x}}{d\,t^j}\right]_{\,t\,=\,0} \;=\; x^j \end{aligned} \tag{93}$$

We can then write the Maclaurin series as

$$\begin{aligned} e^{t\,x} \;&=\; \sum_{n\,=\,0}^{\infty} \frac{d^n e^{t\,x}}{d\,t^n}\,(0)\;\frac{t^n}{n\,!} \\ &=\; \sum_{n\,=\,0}^{\infty} x^n\,\frac{t^n}{n\,!} \\ &= 1 \;+\; t\,x \;+\; \frac{t^2\,x^2}{2\,!} \;+\; \frac{t^3\,x^3}{3\,!} \;+\; \cdots \;+\; \frac{t^r\,x^r}{r\,!} \;+\; \cdots \end{aligned} \tag{94}$$

We can then compute $E(e^{tx}) = M_X\,(t)$ as

$$
\begin{aligned}
E\left[e^{t\,x}\right] &= M_X(t) \;=\; \sum_x e^{t\,x}\; f(x) \\
&= \sum_x \left[1 + t\,x + \frac{t^2\,x^2}{2\,!} + \frac{t^3\,x^3}{3!} \;+\; \cdots + \frac{t^r\,x^r}{r!} + \cdots\right] f(x) \\
&= \sum_x f(x) + t\sum_x x f(x) + \frac{t^2}{2!}\sum_x x^2 f(x) + \frac{t^3}{3!}\sum_x x^3 f(x) + \cdots + \frac{t^r}{r!}\sum_x x^r f(x) + \cdots \\
&= 1 + \mu t + \mu_2'\frac{t^2}{2!} + \mu_3'\frac{t^3}{3!} + \cdots + \mu_r'\frac{t^r}{r!} \;+\; \cdots
\end{aligned}
\tag{95}
$$

In the expansion, the coefficient of $\frac{t^r}{t!}$ is μ_r', the rth moment about the origin of the random variable X.

3.2.2. *Example derivation of a moment generating function.* Find the moment-generating function of the random variable whose probability density is given by

$$
f(x) \;=\begin{cases} e^{-x} & for\ x\ >\ 0 \\ 0 & elsewhere \end{cases}
\tag{96}
$$

and use it to find an expression for μ_r'. By definition

$$
\begin{aligned}
M_X(t) \;=\; E\left(e^{tX}\right) &= \int_{-\infty}^{\infty} e^{t\,x}\cdot e^{-x}\,dx \\
&= \int_{o}^{\infty} e^{-x\,(1\,-\,t)}dx \\
&= \frac{-\,1}{t\;-\;1}\;e^{-x\,(1\,-\,t)}\,|_0^{\infty} \\
&= 0\;-\;\left[\frac{-1}{1\;-\;t}\right] \\
&= \frac{1}{1\;-\;t}\;for\ t\;<\;1
\end{aligned}
\tag{97}
$$

As is well known, when $|t| < 1$ the Maclaurin's series for $\frac{1}{1\,-\,t}$ is given by

$$
\begin{aligned}
M_x(t) \;=\; \frac{1}{1\;-\;t} &= 1\;+\;t\;+\;t^2\;+\;t^3\;+\;\cdots\;+\;t^r + \cdots \\
&= 1\;+\;1!\;\cdot\frac{t}{1!}\;+\;2!\;\cdot\;\frac{t^2}{2!}\;+\;3!\;\cdot\;\frac{t^3}{3}!\;+\;\cdots\;+\;r!\;\cdot\;\frac{t^r}{r!}\;+\;\cdots
\end{aligned}
\tag{98}
$$

or we can derive it directly using equation 92. To derive it directly utilizing the Maclaurin series we need the all derivatives of the function $\frac{1}{1\,-\,t}$ evaluated at 0. The derivatives are as follows

$$
\begin{aligned}
f(t) \ = \frac{1}{1 \ - \ t} & = (1 \ - \ t)^{-1} \\
f^{(1)} & = (1 \ - \ t)^{-2} \\
f^{(2)} & = 2 \ (1 \ - \ t)^{-3} \\
f^{(3)} & = 6 \ (1 \ - \ t)^{-4} \\
f^{(4)} & = 24 \ (1 \ - \ t)^{-5} \\
f^{(5)} & = 120 \ (1 \ - \ t)^{-6} \\
& \vdots \\
f^{(n)} & = n\,! \ (1 \ - \ t)^{(n \ + \ 1)} \\
& \vdots
\end{aligned}
\tag{99}
$$

Evaluating them at zero gives

$$
\begin{aligned}
f(0) \ = \frac{1}{1 \ - \ 0} & = (1 \ - \ 0)^{-1} \ = \ 1 \\
f^{(1)} & = (1 \ - \ 0)^{-2} \ = \ 1 \ = \ 1! \\
f^{(2)} & = 2 \ (1 \ - \ 0)^{-3} \ = \ 2 \ = \ 2! \\
f^{(3)} & = 6 \ (1 \ - \ 0)^{-4} \ = \ 6 \ = \ 3! \\
f^{(4)} & = 24 \ (1 \ - \ 0)^{-5} \ = \ 24 \ = \ 4! \\
f^{(5)} & = 120 \ (1 \ - \ 0)^{-6} \ = \ 120 \ = \ 5! \\
& \vdots \\
f^{(n)} & = n! \ (1 \ - \ 0)^{-\,(n \ + \ 1)} \ = \ n! \\
& \vdots
\end{aligned}
\tag{100}
$$

Now substituting in appropriate values for the derivatives of the function $f(t) \ = \frac{1}{1 \ - \ t}$ we obtain

$$
\begin{aligned}
f(t) \ & = \sum_{n \ = \ 0}^{\infty} \frac{f^{(n)} \, (0)}{n!} \ t^n \\
& = f \ (0) \ + \ \frac{f^{(1)} \, (0)}{1!} \ t \ + \ \frac{f^{(2)} \, (0)}{2!} \ t^2 \ + \ \frac{f^{(3)} \, (0)}{3!} \ t^3 \ + \ \cdots \ + \\
& = 1 \ + \ \frac{1!}{1!} \ t \ + \ \frac{2!}{2!} \ t^2 \ + \ \frac{3!}{3!} \ t^3 \ + \ \cdots \ + \\
& = 1 \ + \ t \ + \ t^2 \ + \ t^3 \ + \ \cdots \ +
\end{aligned}
\tag{101}
$$

A further issue is to determine the radius of convergence for this particular function. Consider an arbitrary series where the nth term is denoted by a_n. The ratio test says that

$$If \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = L < 1 \,,\; then\; the\; series\; is\; absolutely\; convergent \tag{102a}$$

$$\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = L > 1 \;\; or \;\; \lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = \infty \,,\; then\; the\; series\; is\; divergent \tag{102b}$$

Now consider the nth term and the (n+1)th term of the Maclaurin series expansion of $\frac{1}{1 - t}$.

$$\begin{aligned} a_n &= t^n \\ \lim_{n \to \infty} \left| \frac{t^{n+1}}{t^n} \right| &= \lim_{n \to \infty} |t| = L \end{aligned} \tag{103}$$

The only way for this to be less than one in absolute value is for the absolute value of t to be less than one, i.e., $|t| < 1$. Now writing out the Maclaurin series as in equation 98 and remembering that in the expansion, the coefficient of $\frac{t^r}{r!}$ is μ_r', the rth moment about the origin of the random variable X

$$\begin{aligned} M_x(t) &= \frac{1}{1 - t} = 1 + t + t^2 + t^3 + \cdots + t^r + \cdots \\ &= 1 + 1! \cdot \frac{t}{1!} + 2! \cdot \frac{t^2}{2!} + 3! \cdot \frac{t^3}{3}! + \cdots + r! \cdot \frac{t^r}{r!} + \cdots \end{aligned} \tag{104}$$

it is clear that μ_r' = r! for r = 0, 1, 2, ... For this density function E[X] = 1 because the coefficient of $\frac{t^1}{1!}$ is 1. We can verify this by finding E[X] directly by integrating.

$$E(X) = \int_0^\infty x \cdot e^{-x} \, dx \tag{105}$$

To do so we need to integrate by parts with u = x and dv = $e^{-x} dx$. Then du = dx and v = $-e^{-x} \, dx$. We then have

$$\begin{aligned} E(X) &= \int_0^\infty x \cdot e^{-x} \, dx,\; u = x,\; du = dx\,,\; v = -e^{-x}\,,\; dv = e^{-x} \, dx \\ &= -x e^{-x} \big|_0^\infty - \int_0^\infty -e^{-x} \, dx \\ &= [0 - 0] - \left[e^{-x} \big|_0^\infty \right] \\ &= 0 - [0 - 1] = 1 \end{aligned} \tag{106}$$

3.2.3. *Moment property of the moment generating functions for discrete random variables.*

Theorem 8. *If $M_X(t)$ exists, then for any positive integer k,*

$$\left. \frac{d^k M_X(t)}{dt^k} \right]_{t=0} = M_X^{(k)}(0) = \mu_k' \,. \tag{107}$$

In other words, if you find the kth derivative of M_X(t) with respect to t and then set t = 0, the result will be μ_k' .

Proof. $\frac{d^k M_X(t)}{dt^k}$, or $M_X^{(k)}(t)$, is the kth derivative of $M_X(t)$ with respect to t. From equation 95 we know that

$$M_X(t) \;=\; E\;\left(e^{t\,X}\right) \;=\; 1 \;+\; t\;\mu_1' \;+\; \frac{t^2}{2!}\;\mu_2' \;+\; \frac{t^3}{3!}\;\mu_3' \;+\;\cdots \tag{108}$$

It then follows that

$$M_X^{(1)}\,(t) \;= \mu_1' \;+\; \frac{2\,t}{2!}\;\mu_2' \;+\; \frac{3\,t^2}{3!}\;\mu_3' \;+\;\cdots \tag{109a}$$

$$M_X^{(2)}\,(t) \;= \mu_2' \;+\; \frac{2\,t}{2!}\;\mu_3' \;+\; \frac{3\,t^2}{3!}\;\mu_4' \;+\;\cdots \tag{109b}$$

where we note that $\frac{n}{n!} \;=\; \frac{1}{(n\;-\;1\,)\,!}$. In general we find that

$$M_X^{(\,k\,)}\,(\,t\,) \;=\; \mu_k' \;+\; \frac{2\,t}{2\,!}\;\mu_{k\,+\,1}' \;+\; \frac{3\,t^2}{3\,!}\;\mu_{k\,+\,2}' \;+\;\cdots\;. \tag{110}$$

Setting t = 0 in each of the above derivatives, we obtain

$$M_X^{(1)}\,(0) \;= \mu_1' \tag{111a}$$

$$M_X^{(2)}\,(0) \;= \mu_2' \tag{111b}$$

and, in general,

$$M_X^{(k)}\,(0) \;=\; \mu_k' \tag{112}$$

□

These operations involve interchanging derivatives and infinite sums, which can be justified if $M_X\;(t)$ exists.

3.2.4. *Moment property of the moment generating functions for continuous random variables.*

Theorem 9. *If X has mgf $M_X\;(t)$, then*

$$E\;X^n \;=\; M_X^{(n)}\;(0)\,, \tag{113}$$

where we define

$$M_X^{(n)}\,(0) \;= \frac{d^n}{dt^n}\;M_X\;(t)\,|_{t\;=\,0} \tag{114}$$

The nth moment of the distribution is equal to the nth derivative of $M_X\;(t)$ evaluated at t = 0.

Proof. We will assume that we can differentiate under the integral sign and differentiate equation 91.

$$\begin{aligned}
\frac{d}{dt}\;M_X\;(t) \;&= \frac{d}{dt}\;\int_{-\infty}^{\infty}\;e^{t\,x}\,f_X\;(x\,)\;dx \\
&= \int_{-\infty}^{\infty}\;\left(\frac{d}{dt}\;\;e^{t\,x}\right)\;\;f_X\;(x\,)\;dx \\
&= \int_{-\infty}^{\infty}\;\left(x\,e^{t\,x}\right)\;\;f_X\;(x\,)\;dx \\
&= E\;\left(X\,e^{\,t\,X}\right)
\end{aligned} \tag{115}$$

Now evaluate equation 115 at t = 0.

$$\frac{d}{dt} M_X(t)\,|_{t=0} \;=\; E\left(X\,e^{t\,X}\right)\,|t \;=\; 0 \;=\; E\;X \tag{116}$$

We can proceed in a similar fashion for other derivatives. We illustrate for n = 2.

$$\begin{aligned}\frac{d^2}{dt^2} M_X(t) &= \frac{d^2}{dt^2}\int_{-\infty}^{\infty} e^{t\,x} f_X(x)\,dx\\ &= \int_{-\infty}^{\infty}\left(\frac{d^2}{dt^2}\;e^{t\,x}\right) f_X(x)\,dx\\ &= \int_{-\infty}^{\infty}\left(\frac{d}{dt}\;x\,e^{t\,x}\right) f_X(x)\,dx\\ &= \int_{-\infty}^{\infty}\left(x^2\,e^{t\,x}\right) f_X(x)\,dx\\ &= E\,\left(X^2\,e^{t\,X}\right)\end{aligned} \tag{117}$$

Now evaluate equation 117 at t = 0.

$$\frac{d^2}{dt^2} M_X(t)\,|_{t=0} \;=\; E\left(X^2\,e^{t\,X}\right)\,|t \;=\; 0 \;=\; E\;X^2 \tag{118}$$

$\square$

3.3. **Some properties of moment generating functions.** If a and b are constants, then

$$M_{X+a}(t) = E\left(e^{(X\,+\,a)\,t}\right) \;=\; e^{a\,t}\cdot M_X(t) \tag{119a}$$

$$M_{bX}(t) = E\left(e^{b\,X\,t}\right) \;=\; M_X(b\,t) \tag{119b}$$

$$M_{\frac{X\,+\,a}{b}}(t) = E\left(e^{\left(\frac{X+a}{b}\right)t}\right) \;=\; e^{\frac{a}{b}\,t}\cdot M_X\left(\frac{t}{b}\right) \tag{119c}$$

3.4. **Examples of moment generating functions.**

3.4.1. *Example 1.* Consider a random variable with two possible values, 0 and 1, and corresponding probabilities f(1) = p, f(0) = 1-p. For this distribution

$$\begin{aligned}M_X(t) &= E\left(e^{t\,X}\right)\\ &= e^{t\,\cdot\,1} f(1) \;+\; e^{t\,\cdot\,0} f(0)\\ &= e^t\,p \;+\; e^0\,(1\;-\;p)\\ &= e^0\,(1\;-\;p) \;+\; e^t\,p\\ &= 1\;-\;p\;+\;e^t\,p\\ &= 1\;+\;p\,\left(e^t\;-\;1\right)\end{aligned} \tag{120}$$

The derivatives are

$$\begin{aligned} M_X^{(1)}(t) &= p\,e^t \\ M_X^{(2)}(t) &= p\,e^t \\ M_X^{(3)}(t) &= p\,e^t \\ &\vdots \\ M_X^{(k)}(t) &= p\,e^t \\ &\vdots \end{aligned} \tag{121}$$

Thus

$$E\left[X^k\right] \;=\; M_X^{(k)}(0) \;=\; p\,e^0 \;=\; p \tag{122}$$

We can also find this by expanding $M_X(t)$ using the Maclaurin series for the moment generating function for this problem

$$\begin{aligned} M_X(t) &= E\left(e^{t\,X}\right) \\ &= 1 \,+\, p\left(e^t \,-\, 1\right) \end{aligned} \tag{123}$$

To obtain this we first need the series expansion of e^t. All derivatives of e^t are equal to e^t. The expansion is then given by

$$\begin{aligned} e^t &= \sum_{n=0}^{\infty} \frac{d^n\,e^t}{d\,t^n}(0)\,\frac{t^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{t^n}{n!} \\ &= 1 \,+\, t \,+\, \frac{t^2}{2!} \,+\, \frac{t^3}{3!} \,+\, \cdots \,+\, \frac{t^r}{r!} \,+\, \cdots \end{aligned} \tag{124}$$

Substituting equation 124 into equation 123 we obtain

$$\begin{aligned} M_X(t) &= 1 \,+\, p\,e^t \,-\, p \\ &= 1 \,+\, p\left[1 \,+\, t \,+\, \frac{t^2}{2!} \,+\, \frac{t^3}{3!} \,+\, \cdots \,+\, \frac{t^r}{r\,!} \,+\, \cdots\right] \,-\, p \\ &= 1 \,+\, p \,+\, p\,t \,+\, p\,\frac{t^2}{2!} \,+\, p\,\frac{t^3}{3!} \,+\, \cdots \,+\, p\,\frac{t^r}{r!} \,+\, \cdots \,-\, p \\ &= 1 \,+\, p\,t \,+\, p\,\frac{t^2}{2!} \,+\, p\,\frac{t^3}{3!} \,+\, \cdots \,+\, p\,\frac{t^r}{r!} \,+\, \cdots \end{aligned} \tag{125}$$

We can then see that all moments are equal to p. This is also clear by direct computation

$$\begin{aligned}
E(X) &= (1)\,p + (0)\,(1-p) = p \\
E\left(X^2\right) &= (1^2)\,p + (0^2)\,(1-p) = p \\
E\left(X^3\right) &= (1^3)\,p + (0^3)\,(1-p) = p \\
&\vdots \\
E\left(X^k\right) &= (1^k)\,p + (0^k)\,(1-p) = p \\
&\vdots
\end{aligned} \tag{126}$$

3.4.2. *Example 2.* Consider the exponential distribution which has a density function given by

$$f(x) = \frac{1}{\lambda}\, e^{\frac{-x}{\lambda}} \; 0 \le x \le \infty,\; \lambda > 0 \tag{127}$$

For $\lambda\, t < 1$, we have

$$\begin{aligned}
M_X(t) &= \int_0^\infty e^{t\,x}\, \frac{1}{\lambda}\, e^{\frac{-x}{\lambda}}\, dx \\
&= \frac{1}{\lambda} \int_0^\infty e^{-\left(\frac{1}{\lambda} - t\right) x}\, dx \\
&= \frac{1}{\lambda} \int_0^\infty e^{-\left(\frac{1 - \lambda t}{\lambda}\right) x}\, dx \\
&= \frac{1}{\lambda} \left[\frac{-\lambda}{1 - \lambda t}\right] e^{-\left(\frac{1 - \lambda t}{\lambda}\right) x} \Big|_0^\infty \\
&= \left[\frac{-1}{1 - \lambda t}\right] e^{-\left(\frac{1 - \lambda t}{\lambda}\right) x} \Big|_0^\infty \\
&= 0 - \left[\frac{-1}{1 - \lambda t}\right] e^0 \\
&= \frac{1}{1 - \lambda t}
\end{aligned} \tag{128}$$

We can then find the moments by differentiation. The first moment is

$$\begin{aligned}
E(X) &= \frac{d}{dt}\,(1 - \lambda t)^{-1} \,|_{t=0} \\
&= \lambda\,(1 - \lambda t)^{-2} \,|_{t=0} \\
&= \lambda
\end{aligned} \tag{129}$$

The second moment is

$$\begin{aligned} E(X^2) &= \frac{d^2}{dt^2}\ (1 - \lambda t)^{-1} \,|_{t=0} \\ &= \frac{d}{dt}\ \left(\lambda\,(1 - \lambda t)^{-2}\right)\,|_{t=0} \\ &= 2\,\lambda^2\,(1 - \lambda t)^{-3}\,|_{t=0} \\ &= 2\,\lambda^2 \end{aligned} \tag{130}$$

3.4.3. *Example 3.* Consider the normal distribution which has a density function given by

$$f(x\,;\ \mu,\,\sigma^2) \ = \ \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{131}$$

Let g(x) = X - μ, where X is a normally distributed random variable with mean μ and variance σ^2. Find the moment-generating function for (X - μ). This is the moment generating function for central moments of the normal distribution.

$$M_X(t) \ = \ E[e^{t\,(X\,-\,\mu)}] \ = \ \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{t\,(x\,-\,\mu)}\ e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \tag{132}$$

To integrate, let u = x - μ. Then du = dx and

$$\begin{aligned} M_X(t) \ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tu} e^{\frac{-u^2}{2\sigma^2}}\ du \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\left[t\,u\,-\,\frac{u^2}{2\,\sigma^2}\right]}\ du \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\left[\frac{1}{2\,\sigma^2}\ \left(2\,\sigma^2\,t\,u\,-\,u^2\right)\right]}\ du \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\left(\frac{-1}{2\,\sigma^2}\right)\ (u^2\ -\ 2\sigma^2\,t\,u\,)\right]\ du \end{aligned} \tag{133}$$

To simplify the integral, complete the square in the exponent of e. That is, write the second term in brackets as

$$\left(u^2\ -\ 2\sigma^2\,t\,u\right) \ = \ \left(u^2\ -\ 2\sigma^2\,t\,u\ +\ \sigma^4\,t^2\ -\ \sigma^4\,t^2\right) \tag{134}$$

This then will give

$$\begin{aligned} \exp\left[\left(\frac{-1}{2\sigma^2}\right)(u^2 - 2\sigma^2 tu)\right] &= \exp\left[\left(\frac{-1}{2\sigma^2}\right)(u^2 - 2\sigma^2 tu + \sigma^4 t^2 - \sigma^4 t^2)\right] \\ &= \exp\left[\left(\frac{-1}{2\,\sigma^2}\right)\ (u^2\ -\ 2\sigma^2\,t\,u + \sigma^4\,t^2)\right] \cdot \exp\ \left[\left(\frac{-1}{2\sigma^2}\right)(-\,\sigma^4 t^2)\right] \\ &= \exp\ \left[\left(\frac{-1}{2\sigma^2}\right)(u^2 - 2\sigma^2 tu + \sigma^4\,t^2)\right] \cdot \exp\left[\frac{\sigma^2 t^2}{2}\right] \end{aligned} \tag{135}$$

Now substitute equation 135 into equation 133 and simplify. We begin by making the substitution and factoring out the term $\exp\left[\frac{\sigma^2 t^2}{2}\right]$.

$$\begin{aligned} M_X(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\left(\frac{-1}{2\,\sigma^2}\right) (u^2 - 2\sigma^2\,t\,u)\right] du \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[\left(\frac{-1}{2\,\sigma^2}\right) (u^2 \;-\; 2\sigma^2\,t\,u +\; \sigma^4\,t^2)\right] \cdot \exp\left[\frac{\sigma^2\,t^2}{2}\right] du \\ &= \exp\left[\frac{\sigma^2\,t^2}{2}\right] \left[\frac{1}{\sigma\sqrt{2\pi}}\right] \int_{-\infty}^{\infty} \exp\left[\left(\frac{-1}{2\,\sigma^2}\right)(u^2 \;-\; 2\sigma^2\,t\,u +\; \sigma^4\,t^2)\right] du \end{aligned} \tag{136}$$

Now move $\left[\frac{1}{\sigma\sqrt{2\pi}}\right]$ inside the integral sign, take the square root of $(u^2 \;-\; 2\sigma^2\,t\,u +\; \sigma^4\,t^2)$ and simplify

$$\begin{aligned} M_X(t) &= \exp\left[\frac{\sigma^2\,t^2}{2}\right] \int_{-\infty}^{\infty} \frac{\exp\left[\left(\frac{-1}{2\,\sigma^2}\right) (u^2 - 2\sigma^2\,t\,u + \sigma^4\,t^2)\right]}{\sigma\sqrt{2\pi}} du \\ &= \exp\left[\frac{\sigma^2\,t^2}{2}\right] \int_{-\infty}^{\infty} \frac{\exp\left[\left(\frac{-1}{2\,\sigma^2}\right) (u - \sigma^2\,t\,)^2\right]}{\sigma\sqrt{2\pi}} du \\ &= e^{\frac{t^2\,\sigma^2}{2}} \int_{-\infty}^{\infty} \frac{e^{\frac{-1}{2}\left[\frac{u-\sigma^2 t}{\sigma}\right]^2}}{\sigma\sqrt{2\pi}} du \end{aligned} \tag{137}$$

The function inside the integral is a normal density function with mean and variance equal to $\sigma^2 t$ and σ^2, respectively. Hence the integral is equal to 1. Then

$$M_X(t) \;=\; e^{\frac{t^2\,\sigma^2}{2}}\,. \tag{138}$$

The moments of u = x - μ can be obtained from M_X(t) by differentiating. For example the first central moment is

$$\begin{aligned} E(X \;-\; \mu\,) &= \frac{d}{dt} \left(e^{\frac{t^2\,\sigma^2}{2}}\right) \;|_{t\,=\,0} \\ &= t\;\sigma^2 \;\left(e^{\frac{t^2\,\sigma^2}{2}}\right) \;|_{t\,=\,0} \\ &= 0 \end{aligned} \tag{139}$$

The second central moment is

$$\begin{aligned} E(X \;-\; \mu\,)^2 &= \frac{d^2}{dt^2} \;\left(e^{\frac{t^2\,\sigma^2}{2}}\right) \;|_{t\,=\,0} \\ &= \frac{d}{dt} \;\left(t\;\sigma^2 \;\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\right) \;|_{t\,=\,0} \\ &= \left(t^2\,\sigma^4 \;\left(e^{\frac{t^2\,\sigma^2}{2}}\right) \;+\; \sigma^2 \;\left(e^{\frac{t^2\,\sigma^2}{2}}\right)\right) \;|_{t\,=\,0} \\ &= \sigma^2 \end{aligned} \tag{140}$$

The third central moment is

$$\begin{aligned} E(X - \mu)^3 &= \frac{d^3}{dt^3}\left(e^{\frac{t^2\sigma^2}{2}}\right)|_{t=0} \\ &= \frac{d}{dt}\left(t^2\sigma^4\left(e^{\frac{t^2\sigma^2}{2}}\right) + \sigma^2\left(e^{\frac{t^2\sigma^2}{2}}\right)\right)|_{t=0} \\ &= \left(t^3\sigma^6\left(e^{\frac{t^2\sigma^2}{2}}\right) + 2\,t\,\sigma^4\left(e^{\frac{t^2\sigma^2}{2}}\right) + t\,\sigma^4\left(e^{\frac{t^2\sigma^2}{2}}\right)\right)|_{t=0} \\ &= \left(t^3\sigma^6\left(e^{\frac{t^2\sigma^2}{2}}\right) + 3\,t\,\sigma^4\left(e^{\frac{t^2\sigma^2}{2}}\right)\right)|_{t=0} \\ &= 0 \end{aligned} \tag{141}$$

The fourth central moment is

$$\begin{aligned} E(X - \mu)^4 &= \frac{d^4}{dt^4}\left(e^{\frac{t^2\sigma^2}{2}}\right)|_{t=0} \\ &= \frac{d}{dt}\left(t^3\sigma^6\left(e^{\frac{t^2\sigma^2}{2}}\right) + 3\,t\,\sigma^4\left(e^{\frac{t^2\sigma^2}{2}}\right)\right)|_{t=0} \\ &= \left(t^4\sigma^8\left(e^{\frac{t^2\sigma^2}{2}}\right) + 3\,t^2\sigma^6\left(e^{\frac{t^2\sigma^2}{2}}\right) + 3\,t^2\sigma^6\left(e^{\frac{t^2\sigma^2}{2}}\right) + 3\,\sigma^4\left(e^{\frac{t^2\sigma^2}{2}}\right)\right)|_{t=0} \\ &= \left(t^4\sigma^8\left(e^{\frac{t^2\sigma^2}{2}}\right) + 6\,t^2\sigma^6\left(e^{\frac{t^2\sigma^2}{2}}\right) + 3\,\sigma^4\left(e^{\frac{t^2\sigma^2}{2}}\right)\right)|_{t=0} \\ &= 3\,\sigma^4 \end{aligned} \tag{142}$$

3.4.4. *Example 4.* Now consider the raw moments of the normal distribution. The density function is given by

$$f(x\,;\,\mu,\,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{143}$$

To find the moment-generating function for X we integrate the following function.

$$M_X(t) = E[e^{tX}] = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{tx}\,e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,dx \tag{144}$$

First rewrite the integral as follows by putting the exponents over a common denominator.

$$\begin{aligned} M_X(t) = E[e^{tX}] &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{tx}\,e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}(x-\mu)^2 + tx}\,dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}(x-\mu)^2 + \frac{2\sigma^2 t x}{2\sigma^2}}\,dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left[(x-\mu)^2 - 2\sigma^2 t x\right]}\,dx \end{aligned} \tag{145}$$

Now square the term in the exponent and simplify

$$
\begin{aligned}
M_X(t) \;=\; E[e^{tX}] \;&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} \; e^{\frac{-1}{2\,\sigma^2}\left[x^2 \,-\, 2\,\mu\,x \,+\, \mu^2 \,-\, 2\,\sigma^2\,t\,x\right]}\,dx \\
&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} \; e^{\frac{-1}{2\,\sigma^2}\left[\,x^2 \,-\, 2\,x\left(\,\mu \,+\, \sigma^2\,t\,\right) \,+\, \mu^2\,\right]}\,dx
\end{aligned}
\tag{146}
$$

Now consider the exponent of e and complete the square for the portion in brackets as follows.

$$
\begin{aligned}
x^2 - 2x\,\left(\mu + \sigma^2\,t\right) + \mu^2 =& x^2 - 2\,x\,\left(\,\mu + \sigma^2\,t\right) + \mu^2 + \mathbf{2}\mu\sigma^{\mathbf{2}}\mathbf{t} + \sigma^{\mathbf{4}}\mathbf{t}^{\mathbf{2}} - 2\mu\sigma^2 t - \sigma^4 t^2 \\
=& \left(x^2 - (\mu + \sigma^2\,t)\right)^2 - 2\,\mu\,\sigma^2\,t - \sigma^4\,t^2
\end{aligned}
\tag{147}
$$

To simplify the integral, complete the square in the exponent of e by multiplying and dividing by

$$
\left[e^{\frac{2\,\mu\,\sigma^2\,t + \sigma^4\,t^2}{2\,\sigma^2}}\right]\,\left[e^{\frac{-2\,\mu\,\sigma^2\,t\,-\,\sigma^4\,t^2}{2\,\sigma^2}}\right] \;=\; 1
\tag{148}
$$

in the following manner

$$
\begin{aligned}
M_X\,(t) \;&=\; \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{\frac{-1}{2\,\sigma^2}\left[x^2 - 2x\left(\mu + \sigma^2\,t\,\right)\,+\,\mu^2\right]}\,dx \\
&= \left[e^{\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}}\right]\,\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left[x^2 - 2x\left(\mu + \sigma^2 t\right) + \mu^2\,\right]}\,\left[e^{\frac{-\,2\mu\sigma^2 t - \sigma^4 t^2}{2\,\sigma^2}}\right]\,dx \\
&= \left[\,e^{\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}}\right]\,\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left[x^2 - 2x\left(\,\mu + \sigma^2 t\right) + \mu^2 + 2\mu\sigma^2 t + \sigma^4 t^2\right]}\,dx
\end{aligned}
\tag{149}
$$

Now find the square root of

$$
x^2 - 2x\left(\mu + \sigma^2 t\right) + \mu^2 + 2\mu\sigma^2 t + \sigma^4 t^2
\tag{150}
$$

Given we would like to have $(x - \mathit{something})^2$, try squaring $x - (\mu + \sigma^2 t)$ as follows

$$
\begin{aligned}
\left[x - (\mu + \sigma^2 t)\right] \;&= x^2 \;-\; 2\left(x(\mu + \sigma^2 t)\right) \;+\; \left(\mu + \sigma^2 t\right)^2 \\
&= x^2 \;-\; 2x\left(\mu - \sigma^2 t\right) \;+\; \mu^2 \;+\; 2\mu\sigma^2 t \;+\; \sigma^4 t^2
\end{aligned}
\tag{151}
$$

So $\left[x - (\mu + \sigma^2 t)\right]$ is the square root of $x^2 \;-\; 2x\left(\mu - \sigma^2 t\right) \;+\; \mu^2 \;+\; 2\mu\sigma^2 t \;+\; \sigma^4 t^2$. Making the substitution in equation 149 we obtain

$$
\begin{aligned}
M_X\,(t) \;&=\; \left[e^{\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}}\right]\,\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left[x^2 - 2x\left(\mu + \sigma^2 t\right) + \mu^2 + 2\mu\sigma^2 t + \sigma^4 t^2\right]}\,dx \\
&=\; \left[e^{\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}}\right]\,\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{\frac{-1}{2\sigma^2}\left(\left[x - (\mu + \sigma^2 t)\right]\right)}
\end{aligned}
\tag{152}
$$

The expression to the right of $e^{\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\,\sigma^2}}$ is a normal density function with mean and variance equal to μ + $\sigma^2 t$ and σ^2, respectively. Hence the integral is equal to 1. Then

$$
\begin{aligned}
M_X(t) &= \left[e^{\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}} \right]. \\
&= e^{\mu t + \frac{t^2 \sigma^2}{2}}
\end{aligned}
\tag{153}
$$

The moments of X can be obtained from $M_X(t)$ by differentiating with respect to t. For example the first raw moment is

$$
\begin{aligned}
E(X) &= \frac{d}{dt} \left(e^{\mu\, t \, + \, \frac{t^2\, \sigma^2}{2}} \right) |_{t\,=\,0} \\
&= (\mu + t\, \sigma^2) \left(e^{\mu t + \frac{t^2\, \sigma^2}{2}} \right) |_{t\,=\,0} \\
&= \mu
\end{aligned}
\tag{154}
$$

The second raw moment is

$$
\begin{aligned}
E(x^2) &= \frac{d^2}{dt^2} \left(e^{\mu\, t \, + \, \frac{t^2\, \sigma^2}{2}} \right) |_{t=0} \\
&= \frac{d}{dt} \left((\mu + t\, \sigma^2) \left(e^{\mu\, t + \frac{t^2\, \sigma^2}{2}} \right) \right) |_{t=0} \\
&= \left((\mu \;+\; t\; \sigma^2)^2 \left(e^{\mu\, t + \frac{t^2\, \sigma^2}{2}} \right) + \sigma^2 \left(e^{\mu t + \frac{t^2\, \sigma^2}{2}} \right) \right) |_{t\,=\,0} \\
&= \mu^2 \;+\; \sigma^2
\end{aligned}
\tag{155}
$$

The third raw moment is

$$
\begin{aligned}
E(X^3) &= \frac{d^3}{dt^3} \left(e^{\mu t + \frac{t^2 \sigma^2}{2}} \right) |_{t=0} \\
&= \frac{d}{dt} \left((\mu + t\sigma^2)^2 \left(e^{\mu t + \frac{t^2\, \sigma^2}{2}} \right) + \sigma^2 \left(e^{\mu t + \frac{t^2\, \sigma^2}{2}} \right) \right) |_{t\,=\,0} \\
&= \left[(\mu + t\sigma^2)^3 \left(e^{\mu + \frac{t^2 \sigma^2}{2}} \right) + 2\, \sigma^2 \left(\mu + t\sigma^2 \right) \left(e^{\mu + \frac{t^2 \sigma^2}{2}} \right) + \sigma^2 \left(\mu + t\sigma^2 \right) \left(e^{\mu + \frac{t^2 \sigma^2}{2}} \right) \right] |_{t=0} \\
&= \left((\mu + t\, \sigma^2)^3 \left(e^{\mu + \frac{t^2 \sigma^2}{2}} \right) + 3\, \sigma^2 \left(\mu + t\sigma^2 \right) \left(e^{\mu + \frac{t^2 \sigma^2}{2}} \right) \right) |_{t\,=\,0} \\
&= \mu^3 \;+\; 3\sigma^2 \mu
\end{aligned}
\tag{156}
$$

The fourth raw moment is

$$
\begin{aligned}
E(X^4) &= \frac{d^4}{dt^4} \left(e^{\mu + \frac{t^2 \sigma^2}{2}} \right) |_{t=0} \\
&= \frac{d}{dt} \left((\mu + t\, \sigma^2)^3 \left(e^{\mu + \frac{t^2\, \sigma^2}{2}} \right) + 3\, \sigma^2 \left(\mu + t\, \sigma^2 \right) \left(e^{\mu + \frac{t^2\, \sigma^2}{2}} \right) \right) |_{t=0} \\
&= \left((\mu + t\, \sigma^2)^4 \left(e^{\mu + \frac{t^2\, \sigma^2}{2}} \right) + 3\sigma^2 \left(\mu + t\, \sigma^2 \right)^2 \left(e^{\mu + \frac{t^2\, \sigma^2}{2}} \right) \right) |_{t=0} \\
&+ \left(3\sigma^2 \left(\mu + t\, \sigma^2 \right)^2 \left(e^{\mu + \frac{t^2\, \sigma^2}{2}} \right) + 3\, \sigma^4 \left(e^{\mu + \frac{t^2\, \sigma^2}{2}} \right) \right) |_{t=0} \\
&= \left((\mu + t\, \sigma^2)^4 \left(e^{\mu + \frac{t^2\, \sigma^2}{2}} \right) + 6\, \sigma^2 \left(\mu + t\, \sigma^2 \right)^2 \left(e^{\mu + \frac{t^2\, \sigma^2}{2}} \right) + 3\, \sigma^4 \left(e^{\mu + \frac{t^2\, \sigma^2}{2}} \right) \right) |_{t=0} \\
&= \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4
\end{aligned}
\tag{157}
$$

4. Chebyshev's inequality

Chebyshev's inequality applies equally well to discrete and continuous random variables. We state it here as a theorem.

4.1. **A Theorem of Chebyshev.**

Theorem 10. *Let X be a random variable with mean μ and finite variance σ^2. Then, for any constant $k > 0$,*

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2} \qquad or \qquad P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}. \tag{158}$$

The result applies for any probability distribution, whether the probability histogram is bell-shaped or not. The results of the theorem are very conservative in the sense that the actual probability that X is in the interval $\mu \pm k\sigma$ usually exceeds the lower bound for the probability, $1 - 1/k^2$, by a considerable amount.

Chebyshev's theorem enables us to find bounds for probabilities that ordinarily would have to be obtained by tedious mathematical manipulations (integration or summation). We often can obtain estimates of the means and variances of random variables without specifying the distribution of the variable. In situations like these, Chcbyshev's inequality provides meaningful bounds for probabilities of interest.

Proof. Let f(x) denote the density function of X. Then

$$\begin{aligned} V(X) = \sigma^2 &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x)\, dx \\ &= \int_{-\infty}^{\mu - k\sigma} (x - \mu)^2 f(x)\, dx \\ &+ \int_{\mu - k\sigma}^{\mu + k\sigma} (x - \mu)^2 f(x)\, dx \\ &+ \int_{\mu + k\sigma}^{\infty} (x - \mu)^2 f(x)\, dx. \end{aligned} \tag{159}$$

The second integral is always greater than or equal to zero.

Now consider relationship between $(x - \mu)^2$ and $k\sigma^2$.

$$\begin{aligned} x &\leq \mu - k\sigma \\ \Rightarrow -x &\geq k\sigma - \mu \\ \Rightarrow \mu - x &\geq k\sigma \\ \Rightarrow (\mu - x)^2 &\geq k^2\sigma^2 \\ \Rightarrow (x - \mu)^2 &\geq k^2\sigma^2 \end{aligned} \tag{160}$$

And similarly,

$$\begin{aligned} x &\geq \mu + k\,\sigma \\ \Rightarrow\; x - \mu &\geq k\,\sigma \\ \Rightarrow\; (x - \mu)^2 &\geq k^2\,\sigma^2 \end{aligned} \tag{161}$$

Now replace $(x - \mu)^2$ with $k\sigma^2$ in the first and third integrals of equation 159 to obtain the inequality

$$V(X) \;=\; \sigma^2 \;\geq\; \int_{-\infty}^{\mu\,-\,k\,\sigma} k^2\,\sigma^2\; f(x)\,dx \;+\; \int_{\mu\,+\,k\,\sigma}^{\infty} k^2\,\sigma^2\; f(x)\,dx\;. \tag{162}$$

Then

$$\sigma^2 \;\geq\; k^2\,\sigma^2\,\left[\int_{-\infty}^{\mu\,-\,k\,\sigma} f(x)\,dx \;+\; \int_{\mu\,+\,k\,\sigma}^{+\infty} f(x)\,dx\right] \tag{163}$$

W can write this in the following useful manner

$$\begin{aligned} \sigma^2 &\geq k^2\,\sigma^2\{P\,(X \;\leq\; \mu \;-\; k\,\sigma) \;+\; P\,(X + \;\geq\; \mu \;+\; k\,\sigma)\} \\ &= k^2\,\sigma^2\,P\,(\,|\,X \;-\; \mu\,| \;\geq\; k\,\sigma). \end{aligned} \tag{164}$$

Dividing by $k^2\sigma^2$, we obtain

$$P\;(\,|\,X \;-\; \mu\,| \;\geq\; k\,\sigma\;) \;\leq \frac{1}{k^2}, \tag{165}$$

or, equivalently,

$$P\;(\,|\,X \;-\; \mu\,| \;<\; k\,\sigma\;) \;\geq\; 1 \;-\; \frac{1}{k^2}. \tag{166}$$

$\square$

4.2. **Example.** The number of accidents that occur during a given month at a particular intersection, X, tabulated by a group of Boy Scouts over a long time period is found to have a mean of 12 and a standard deviation of 2. The underlying distribution is not known. What is the probability that, next month, X will be greater than eight but less than sixteen. We thus want $P[8 < X < 16]$. We can write equation 158 in the following useful manner.

$$P\;[(\mu \;-\; k\,\sigma\,) \;<\; X \;<\; (\,\mu \;+\; k\,\mu)] \;\geq\; 1 \;-\; \frac{1}{k^2} \tag{167}$$

For this problem μ = 12 and σ = 2 so $\mu - k\sigma$ = 12 - 2k. We can solve this equation for the k that gives us the desired bounds on the probability.

$$
\begin{aligned}
\mu \; - \; k\,\mu \; &= 12 - (k)\,(2) \; = \; 8 \\
\Rightarrow \; 2\,k \; &= 4 \\
\Rightarrow \; k \; &= 2 \\
&and \\
12 \; + \; (\,k\,)\;(\,2\,) \; &= 16 \\
\Rightarrow \; 2\,k - 4 \\
\Rightarrow \; k \; &= 2
\end{aligned}
\tag{168}
$$

We then obtain

$$
P\;[(8)\; < \; X \; < \; (16)\,] \; \geq \; 1 \; - \frac{1}{2^2} \; = \; 1 \; - \; \frac{1}{4} \; = \frac{3}{4} \tag{169}
$$

Therefore the probability that X is between 8 and 16 is at least 3/4.

4.3. **Alternative statement of Chebyshev's inequality.**

Theorem 11. *Let X be a random variable and let g(x) be a non-negative function. Then for* $r > 0$*,*

$$
P\,[g(X) \; \geq \; r] \; \leq \; \frac{E\,g\,(X)}{r} \tag{170}
$$

Proof.

$$
\begin{aligned}
E\,g\,(X) \; &= \int_{-\,\infty}^{\infty} \, g\,(x) \; f_X\,(x) \; dx \\
&\geq \int_{[x\,:\,g\,(x)\,\geq\,r]} \, g\;(x) \; f_X\,(x) \; dx \qquad\qquad (g \; is \; nonnegative) \\
&\geq r \int_{[\,x\,:\,g(x)\,\geq\,r\,]} \, f_X\,(x) \; dx \qquad\qquad (g\,(x) \; \geq \; r) \\
&= r\,P\;[\,g\,(X) \; \geq \; r\,] \\
\Rightarrow \; P\;[g\,(X) \; \geq \; r\,] \; &\leq \; \frac{E\,g\,(X)}{r}
\end{aligned}
\tag{171}
$$

□

4.4. **Another version of Chebyshev's inequality as special case of general version.**

Corollary 1. Let X be a random variable with mean μ and variance σ^2. Then for any $k > 0$ or any $\varepsilon > 0$

$$
P\;[\,|X \; - \; \mu| \; \geq \; k\,\sigma\,] \; \leq \; \frac{1}{k^2} \tag{172a}
$$

$$
P\;[\,|X \; - \; \mu| \; \geq \; \varepsilon\,] \; \leq \; \frac{\sigma^2}{\varepsilon^2} \tag{172b}
$$

Proof. Let g(x) = $\frac{(x-\mu)^2}{\sigma^2}$, where μ = E(X) and σ^2 = Var (X). Then let r = k^2. Then

$$\begin{aligned} P\left[\frac{(X-\mu)^2}{\sigma^2} \geq k^2\right] &\leq \frac{1}{k^2} E\left(\frac{(X-\mu)^2}{\sigma^2}\right) \\ &= \frac{1}{k^2} \frac{E(X-\mu)^2}{\sigma^2} = \frac{1}{k^2} \end{aligned} \tag{173}$$

because $E(X-\mu)^2 = \sigma^2$. We can then rewrite equation 173 as follows

$$\begin{aligned} P\left[\frac{(X-\mu)^2}{\sigma^2} \geq k^2\right] &\leq \frac{1}{k^2} \\ \Rightarrow P\left[(X-\mu)^2 \geq k^2\sigma^2\right] &\leq \frac{1}{k^2} \\ \Rightarrow P\left[|X-\mu| \geq k\sigma\right] &\leq \frac{1}{k^2} \end{aligned} \tag{174}$$

□

References

[1] Amemiya, T. *Advanced Econometrics.* Cambridge: Harvard University Press, 1985.
[2] Bickel P.J., and K.A. Doksum. *Mathematical Statistics: Basic Ideas and Selected Topics, Vol 1).* 2^{nd} Edition. Upper Saddle River, NJ: Prentice Hall, 2001.
[3] Billingsley, P. *Probability and Measure.* 3rd edition. New York: Wiley, 1995.
[4] Casella, G. And R.L. Berger. *Statistical Inference*. Pacific Grove, CA: Duxbury, 2002.
[5] Cramer, H. *Mathematical Methods of Statistics.* Princeton: Princeton University Press, 1946.
[6] Goldberger, A.S. *Econometric Theory.* New York: Wiley, 1964.
[7] Lindgren, B.W. *Statistical Theory* 4^{th} edition. Boca Raton, FL: Chapman & Hall/CRC, 1993.
[8] Rao, C.R. *Linear Statistical Inference and its Applications.* 2nd edition. New York: Wiley, 1973.
[9] Theil, H. *Principles of Econometrics*. New York: Wiley, 1971.