PRACTICALS OF PROBABILITY FOR COMPUTING

RAMANUJAN COLLEGE

UNIVERSITY OF DELHI

DISCIPLINE SPECIFIC CORE COURSE-06

PROBABILITY FOR COMPUTING

SEMESTER-02

SESSION (2024-25)

SUBMITTED BY

KOMAL

Roll No. 24570032
(Operational Research)

BSC. Computer Science (Honors)

SUBMITTED TO

DR. AAKASH

ASSISTANT PROFESSOR

Department of Computer Science

Ramanujan College

University of Delhi

CR Park Main Road,

Block – H, Kalkaji, N.D. - 19

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to Dr. Aakash, Assistant Professor (Operational Research) Ramanujan College, University of Delhi for providing me the opportunity to work on this practical file of Mathematics for Computing, his guidance and encouragement have been invaluable throughout the preparation of this file.

I am also thankful to my peers and the faculty members of the Computer Science Department for their continuous support and valuable insights, which have greatly enriched my understanding of the subject.

Finally, I extend my sincere appreciation to my family and friends for their constant support and encouragement which has been a source of strength in completing this work successfully!

KOMAL

ROLL NO. 24570032

UNIVERSITY ROLL NO. :- 24020570037

INDEX

S no.	TOPIC	PAGE NO.	SIGN.
1.	Plotting and fitting of Binomial distribution and graphical representation of probabilities.	4	
2.	Plotting and fitting of Multinomial distribution and graphical representation of probabilities.	5	
3.	Plotting and fitting of Poisson distribution and graphical representation of probabilities.	6	
4.	Plotting and fitting of Poisson distribution and graphical representation of probabilities.	7	
5.	Plotting and fitting of Uniform distribution and graphical representation of probabilities.	8	
6.	Plotting and fitting of Exponential distribution and graphical representation of probabilities.	9	
7.	Plotting and fitting of Normal distribution and graphical representation of probabilities.	10	
8.	Calculation of cumulative distribution functions for Exponential and Normal distribution.	11	
9.	Given data from two distributions, find the distance between the distributions.	12	
10.	Application problems based on the Binomial distribution.	13	
11.	Application problems based on the Poisson distribution	14	
12.	Application problems based on the Normal distribution.	15	
13.	Presentation of bivariate data through scatter-plot diagrams and calculations of covariance.	16	
14.	Calculation of Karl Pearson's correlation coefficients.	17	
15.	To find the correlation coefficient for a bivariate frequency distribution.	18	
16.	Generating Random numbers from discrete (Bernoulli, Binomial, Poisson) distributions.	19	
17.	Generating Random numbers from continuous (Uniform, Normal) distributions.	20	
18.	Find the entropy from the given data set.	21	

PRACTICAL-1

Plotting and fitting of Binomial distribution and graphical representation of probabilities.

What is binomial distribution?

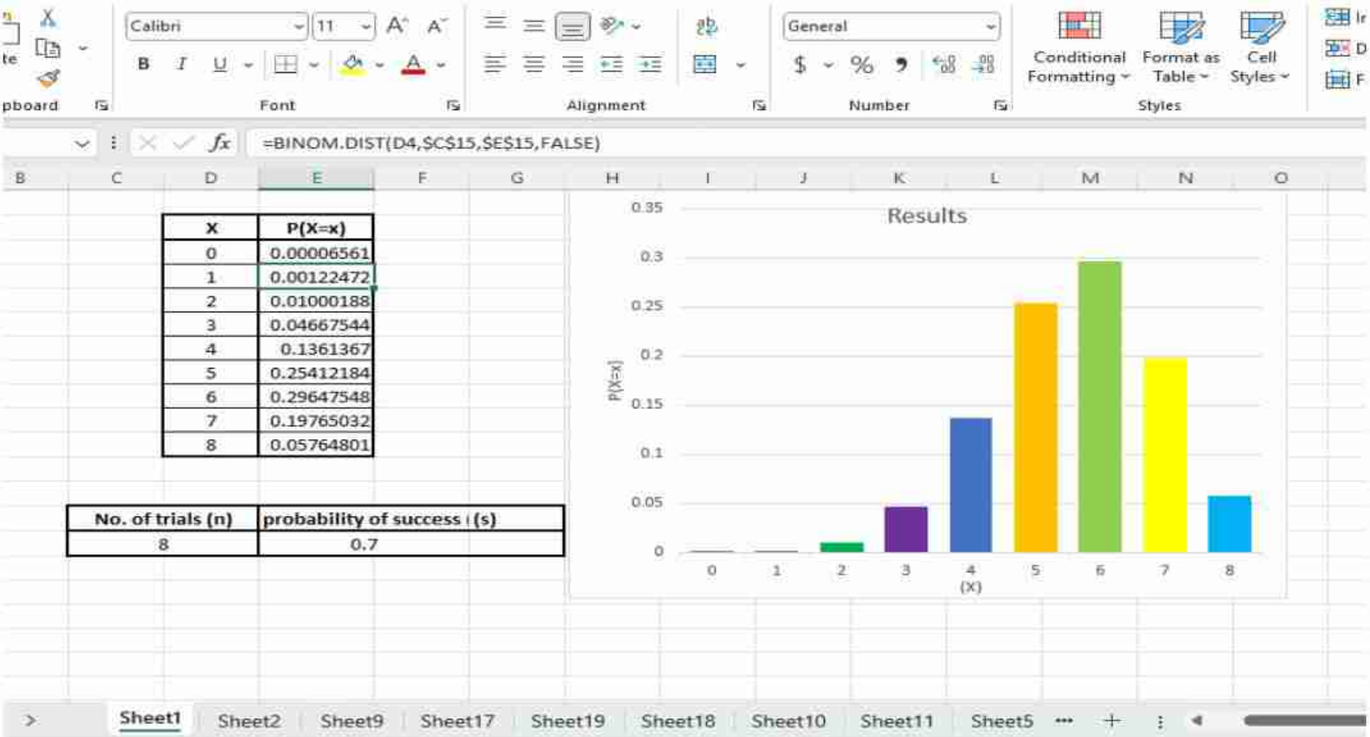
A binomial distribution is a discrete probability distribution that models the number of successes in a fixed number of independent experiments, where each experiment has only two possible outcomes: success or failure**.**

Binomial Probability Formula

$$P(X=x) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$$

<u>**Formula used in excel: -**</u> =BINOM.DIST(D3, \$C\$15, \$E\$15, FALSE)

E.g. A basketball player has a **70% chance** of making a free throw. She takes **8 shots** during practice. Each shot is independent of the others. Construct a binomial table for all her trial shots.



X	P(X=x)
0	0.00006561
1	0.00122472
2	0.01000188
3	0.04667544
4	0.1361367
5	0.25412184
6	0.29647548
7	0.19765032
8	0.05764801

No. of trials (n)	probability of success (s)
8	0.7

PRACTICAL-2

Plotting and fitting of Multinomial distribution and graphical representation of probabilities.

The **multinomial distribution** is a generalization of the **binomial distribution**. While the binomial distribution is used when there are only two possible outcomes (like success/failure), the multinomial distribution is used when there are **more than two possible outcomes** for each trial.

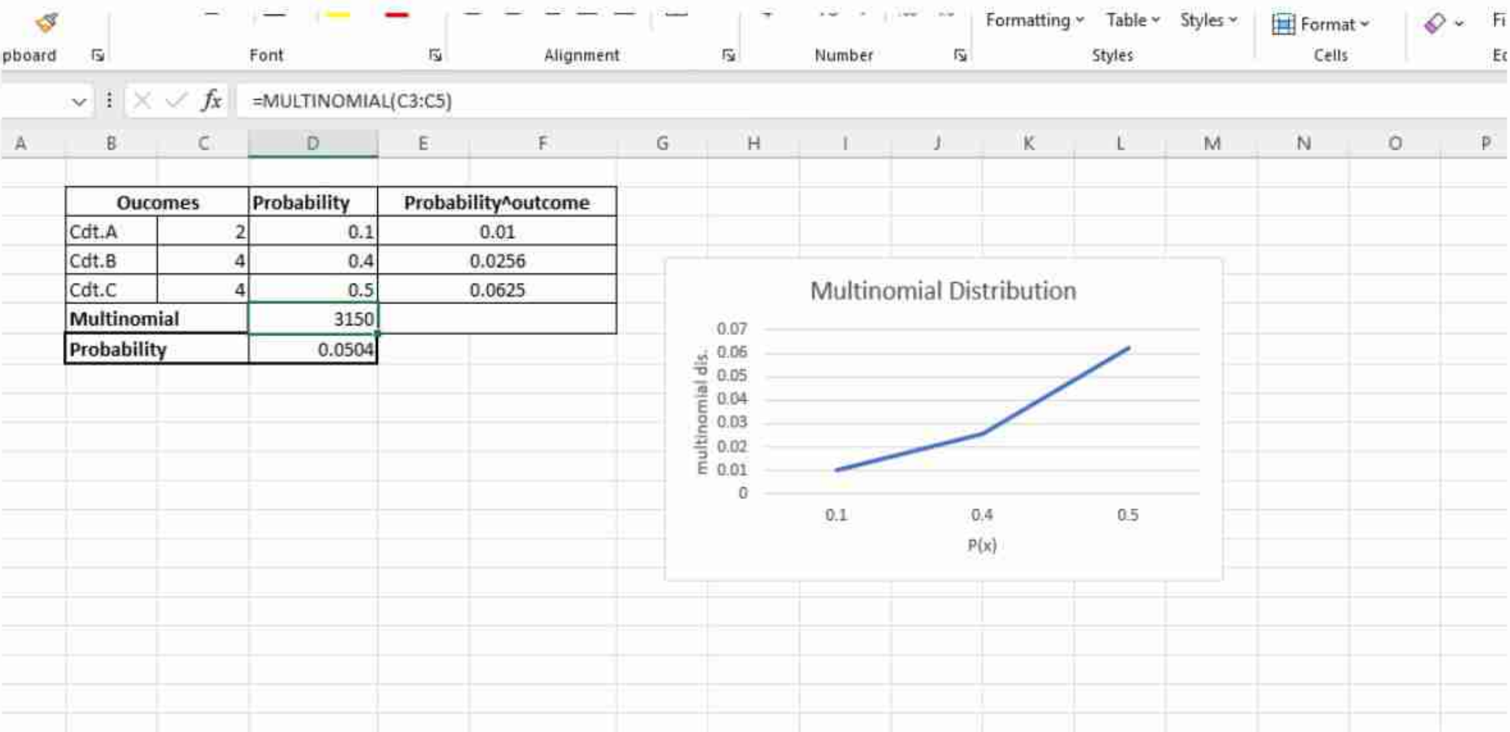
- x_1 outcomes of type 1,
- x_2 outcomes of type 2,
- X_k outcomes of type k

where $x_1+x_2+\cdots+x_k=n$, is given by the **multinomial probability mass function**:

$$P(X=x_1, x_2 \ldots x_k) = \frac{n!}{x1!.x2.....xk!} . p1^{x1.} . p2^{x2} \ldots\ldots\ldots pk^{xk}$$

E.g. In a three-way election for mayor, candidate A receives 10% of the votes, candidate B receives 40% of the votes, and candidate C receives 50% of the votes. If we select a random sample of 10 voters, what is the probability that 2 voted for candidate A, 4 voted for candidate B, and 4 voted for candidate C?

Formula used in excel: - =MULTINOMIAL(C3:C5)



Oucomes		Probability	Probability^outcome
Cdt.A	2	0.1	0.01
Cdt.B	4	0.4	0.0256
Cdt.C	4	0.5	0.0625
Multinomial		3150	
Probability		0.0504	

PRACTICAL-3

Plotting and fitting of Poisson distribution and graphical representation of probabilities.

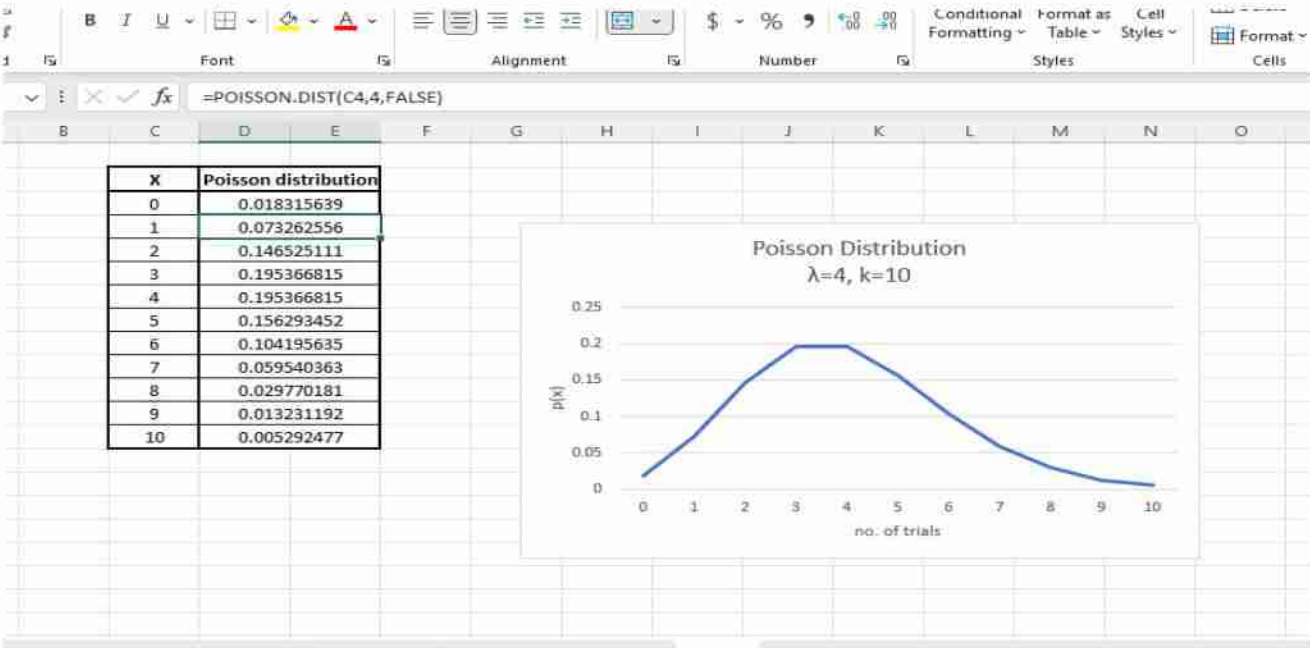
What is Poisson distribution?

The **Poisson distribution** is a **discrete probability distribution** that models the **number of times an event happens** in a **fixed interval of time or space**, **given a known constant mean rate** and **independent occurrences**.

$$P(X=k) = \frac{e^{-\lambda}\lambda}{k!}$$

E.g. A customer service center receives an average of **4 calls per hour**. Assuming the number of calls follows a **Poisson distribution**, what is the probability that exactly **10 calls** are received in a given hour.

Formula used in excel:- =POISSON.DIST(C4,4,FALSE)



X	Poisson distribution
0	0.018315639
1	0.073262556
2	0.146525111
3	0.195366815
4	0.195366815
5	0.156293452
6	0.104195635
7	0.059540363
8	0.029770181
9	0.013231192
10	0.005292477

PRACTICAL-4

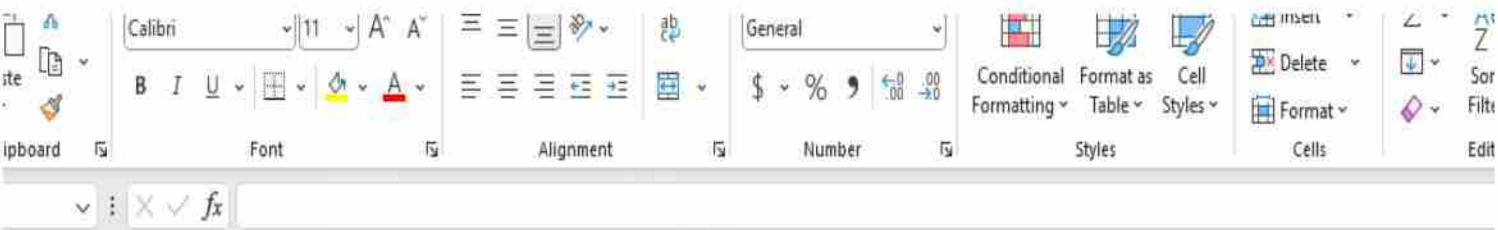
Plotting and fitting of Geometric distribution and graphical representation of probabilities.

What is geometric mean?

The **geometric distribution** is a discrete probability distribution that models the number of **trials** needed to get the **first success** in a sequence of independent Bernoulli trials (where each trial has only two outcomes: success or failure).

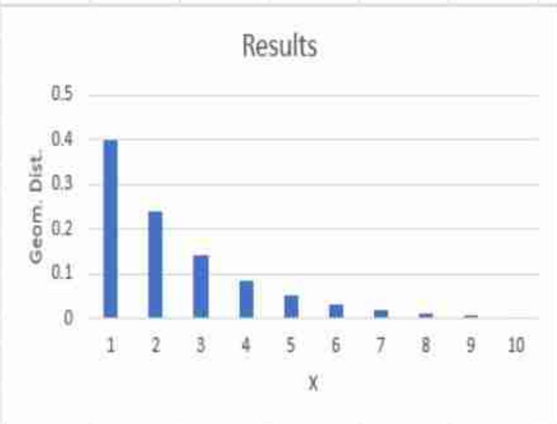
$P(X=k) = (1-p)^k p$

Formula used in excel:- =POWER(1-0.4,C5-1)*0.4



A(X=Call Number)	Geom. Distribution
1	0.4
2	0.24
3	0.144
4	0.0864
5	0.05184
6	0.031104
7	0.0186624
8	0.01119744
9	0.006718464
10	0.004031078

p	0.4



PRACTICAL-5

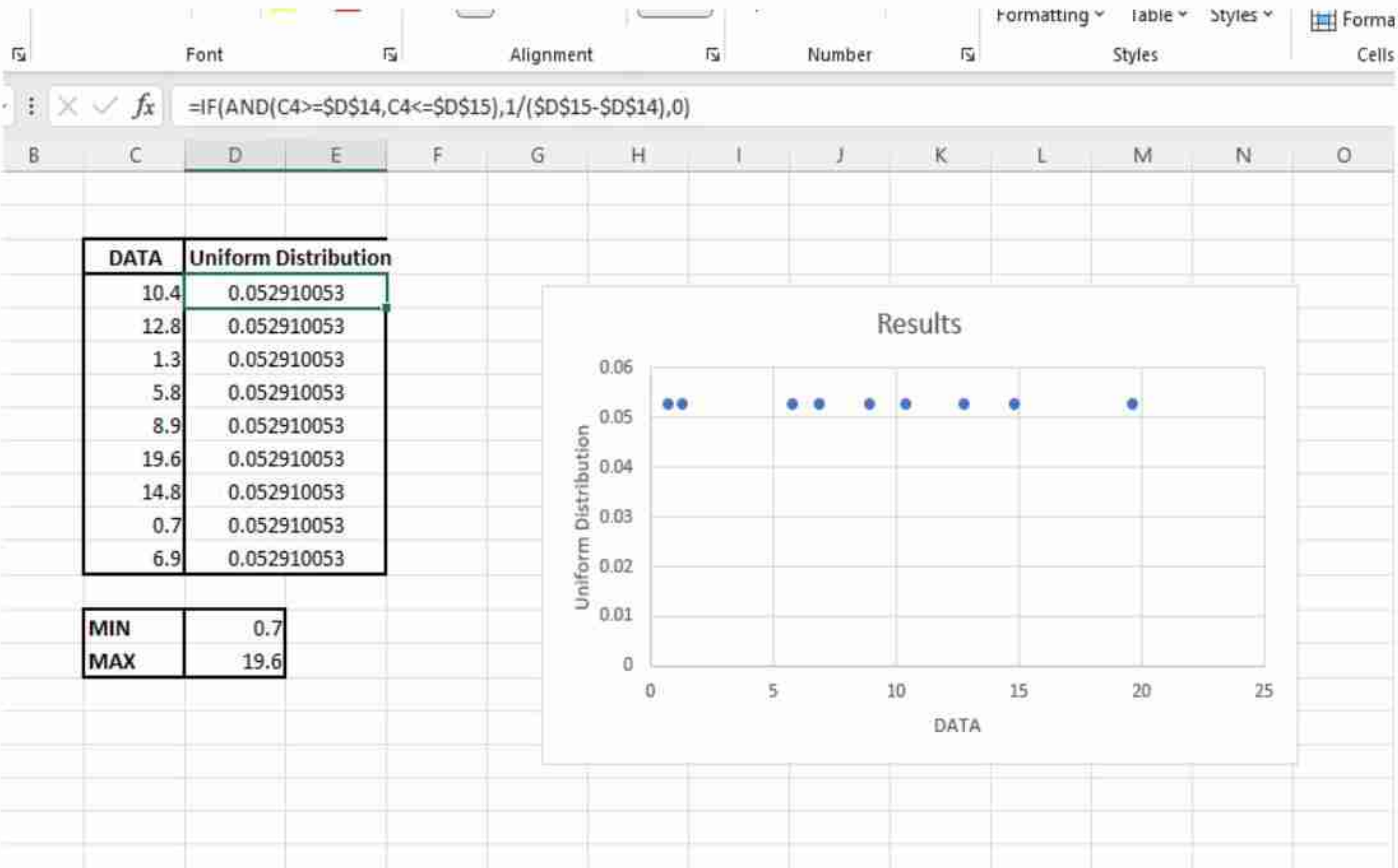
Plotting and fitting of Uniform distribution and graphical representation of probabilities.

What is uniform distribution?

A **uniform distribution** is a probability distribution in which **all outcomes are equally likely**. Each value in the distribution has the same probability of occurring.

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \le x \le b \\ 0, & \text{otherwise} \end{cases}$$

Formula used in excel:- =IF(AND(C4>=$D$14,C4<=$D$15),1/($D$15-$D$14),0)



DATA	Uniform Distribution
10.4	0.052910053
12.8	0.052910053
1.3	0.052910053
5.8	0.052910053
8.9	0.052910053
19.6	0.052910053
14.8	0.052910053
0.7	0.052910053
6.9	0.052910053

MIN	0.7
MAX	19.6

PRACTICAL-6

Plotting and fitting of Exponential distribution and graphical representation of probabilities.

What is exponential distribution?

The **exponential distribution** is a **continuous probability distribution** that models the time between events in a **Poisson process** — a process in which events occur **continuously and independently** at a constant average rate.

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Where:

- $\lambda > 0$ is the **rate parameter** (events per unit time)
- x is the time or distance until the event

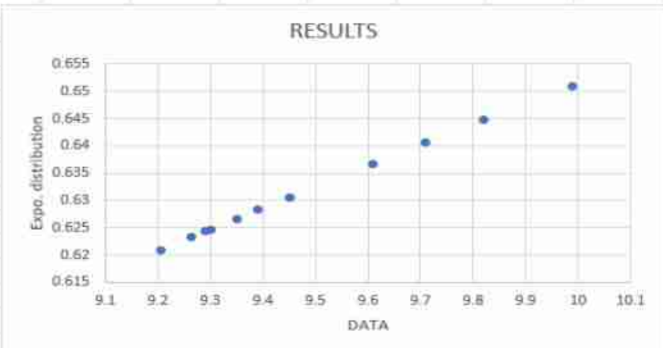
Formula used in excel:-

=EXPON.DIST(D4,$E$17,TRUE)

=EXPON.DIST(D5,$E$17,TRUE)

DATA	EXPONENTIAL DIST.
9.206	0.620980149
9.264	0.62328976
9.29	0.624320529
9.3	0.624716227
9.35	0.626688475
9.39	0.628258808
9.45	0.63060193
9.61	0.636778306
9.71	0.640585989
9.82	0.644728354
9.99	0.65103648

Mean	9.489091
λ	0.105384



PRACTICAL-7

Plotting and fitting of Normal distribution and graphical representation of probabilities.

What is normal distribution?

A **normal distribution** is a continuous probability distribution defined by the probability density function:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

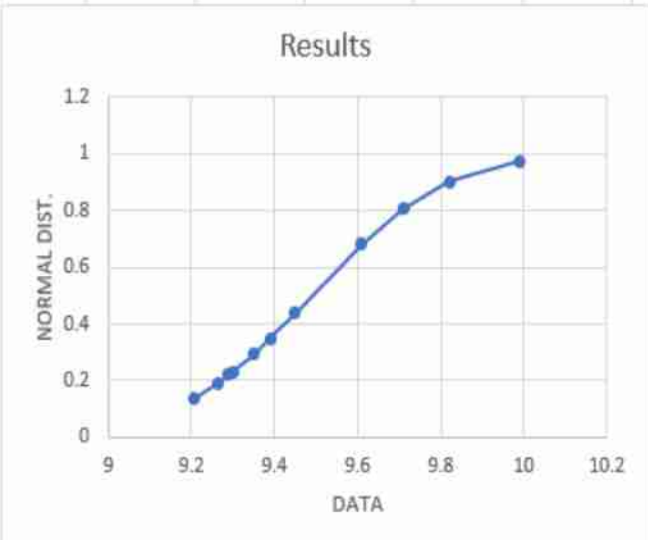
where:

- μ is the **mean** of the distribution,
- σ is the **standard deviation**,
- e is the base of the natural logarithm (approximately 2.718),
- π is the constant pi (approximately 3.1416),
- and x is the variable.

Formula used in excel:- =NORM.DIST(D4,$E$16,$E$17,TRUE)

DATA	NORMAL DIST.
9.206	0.135377207
9.264	0.190600453
9.29	0.219306871
9.3	0.230977159
9.35	0.294214379
9.39	0.349933586
9.45	0.43956275
9.61	0.680958389
9.71	0.804944734
9.82	0.901015066
9.99	0.974335507

MEAN	9.489091
STDEV	0.257045



PRACTICAL-8

Calculation of cumulative distribution functions for Exponential and Normal distribution.

CDF of Exponential Distribution:

For an **exponential distribution** with rate parameter λ, the CDF is:

$$F(x) = P(X \leq x) = 1 - e^{-\lambda x}, \quad \text{for } x \geq 0$$

CDF of Normal Distribution:

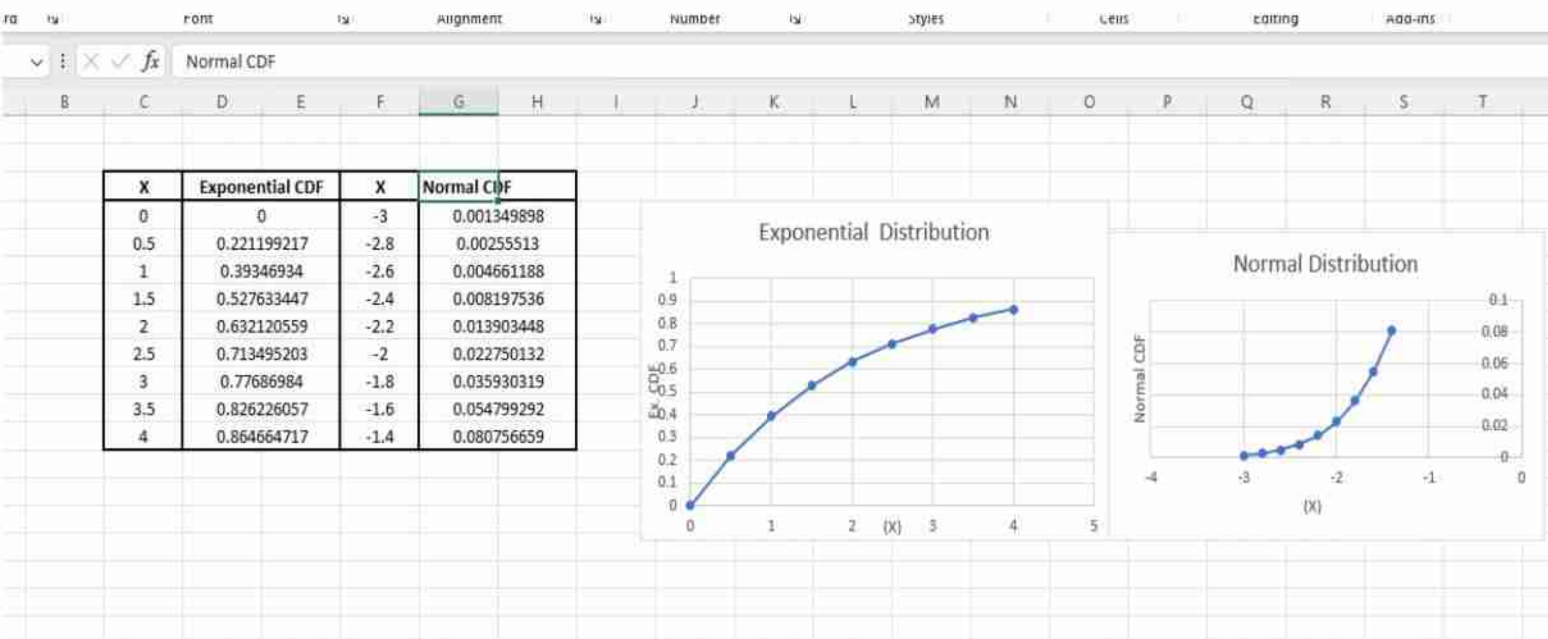
For a **normal distribution** with mean μ and standard deviation σ, the CDF is given by:

$$F(x) = P(X \leq x) = \frac{1}{2}\left[1 + \operatorname{erf}\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right]$$

Formula used in excel:-

exponential dist. CDF:- =EXPON.DIST(C5, 0.5, TRUE)

Normal dist. CDF :-=NORM.DIST(F5,0,1,TRUE)



X	Exponential CDF	X	Normal CDF
0	0	-3	0.001349898
0.5	0.221199217	-2.8	0.00255513
1	0.39346934	-2.6	0.004661188
1.5	0.527633447	-2.4	0.008197536
2	0.632120559	-2.2	0.013903448
2.5	0.713495203	-2	0.022750132
3	0.77686984	-1.8	0.035930319
3.5	0.826226057	-1.6	0.054799292
4	0.864664717	-1.4	0.080756659

PRACTICAL-9

Given data from two distributions, find the distance between the distributions.

Euclidean Distance:

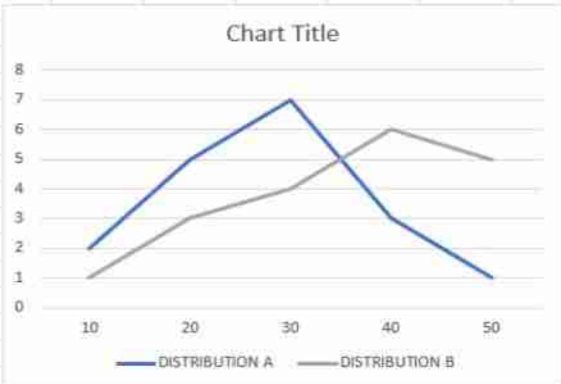
The **Euclidean distance** between two points $P = (p_1, p_2, ..., p_n)$ and $Q = (q_1, q_2, ..., q_n)$ is given by the formula:

$$d(P, Q) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}$$

For distributions, we consider the probability density functions (PDFs) of each distribution over a range of values, say $x_1, x_2, ..., x_n$, and compute the distance between the values of their PDFs.

Formula used in excel:- =SQRT(SUM(G4:G8))

SCORE	DISTRIBUTION A	DISTRIBUTION B	(C-E)^2
10	2	1	1
20	5	3	4
30	7	4	9
40	3	6	9
50	1	5	16
EUCLIDEAN DISTANCE=		6.244998	



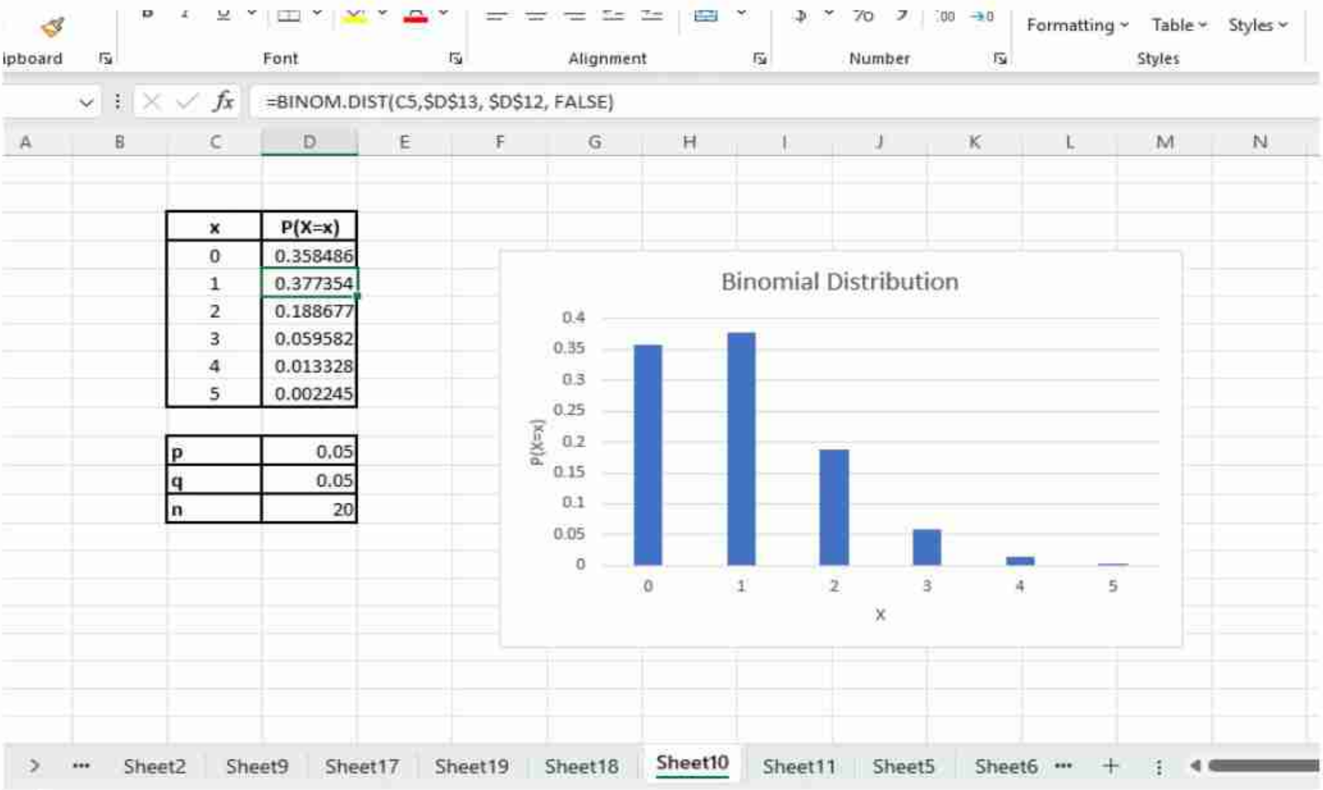
PRACTICAL-10

Application problems based on the Binomial distribution.

A factory produces bulbs, and 5% of them are defective. If a quality control inspector selects 20 bulbs at random, what is the probability that exactly 2 bulbs are defective?

- n = 20 (no. of trials)
- p = 0.05 (probability of success- defective bulbs)
- q = 0.95(probability of non-defective bulbs)
- x = 0 to 5(we'll calculate 0 to 5 bulbs)

Formula to solve in excel :- =BINOM.DIST(C4,$D$13, $D$12, FALSE)



x	P(X=x)
0	0.358486
1	0.377354
2	0.188677
3	0.059582
4	0.013328
5	0.002245

p	0.05
q	0.05
n	20

PRACTICAL-11

Application problems based on the Poisson distribution.

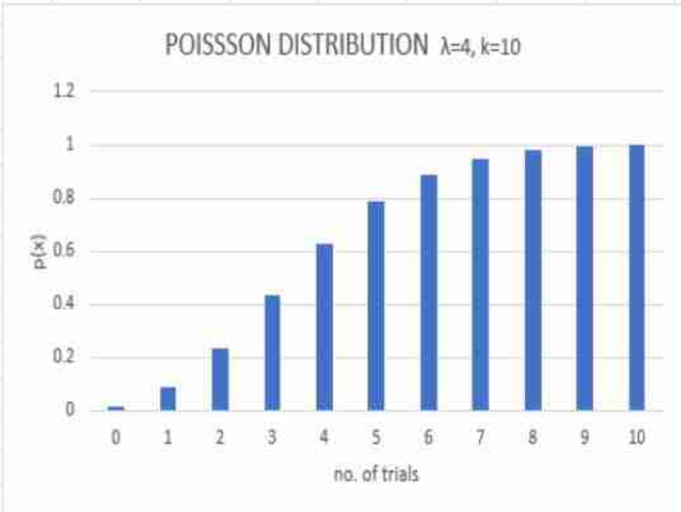
A call center receives an average of 4 calls per minute. What is the probability of receiving 0 to 10 calls in a minute?

- Λ = 4 (average rate of calls)

Formula used in excel :- =POISSON.DIST(C4, 4, TRUE)

=POISSON.DIST(C5, 4, TRUE)

X	p(X)
0	0.018316
1	0.091578
2	0.238103
3	0.43347
4	0.628837
5	0.78513
6	0.889326
7	0.948866
8	0.978637
9	0.991868
10	0.99716



PRACTICAL-12

Application problems based on the Normal distribution.

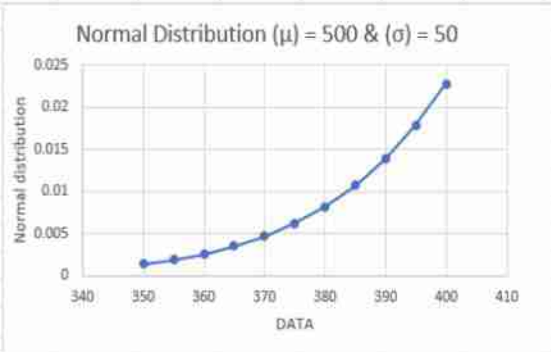
The average daily demand for a product n a store is 500 units, with standard deviation of 50 units. The demand follows a normal distribution. Create a graph showing the normal distribution curve for daily demand from 350-400 units.

- Mean (μ) = 500
- Standard Deviation (σ) = 50

Formula used in excel :- =NORM.DIST(E4, 500,50,TRUE)

DATA	NORMAL DIST.
350	0.001349898
355	0.001865813
360	0.00255513
365	0.003466974
370	0.004661188
375	0.006209665
380	0.008197536
385	0.01072411
390	0.013903448
395	0.017864421
400	0.022750132

stdv (σ)	50
mean(μ)	500

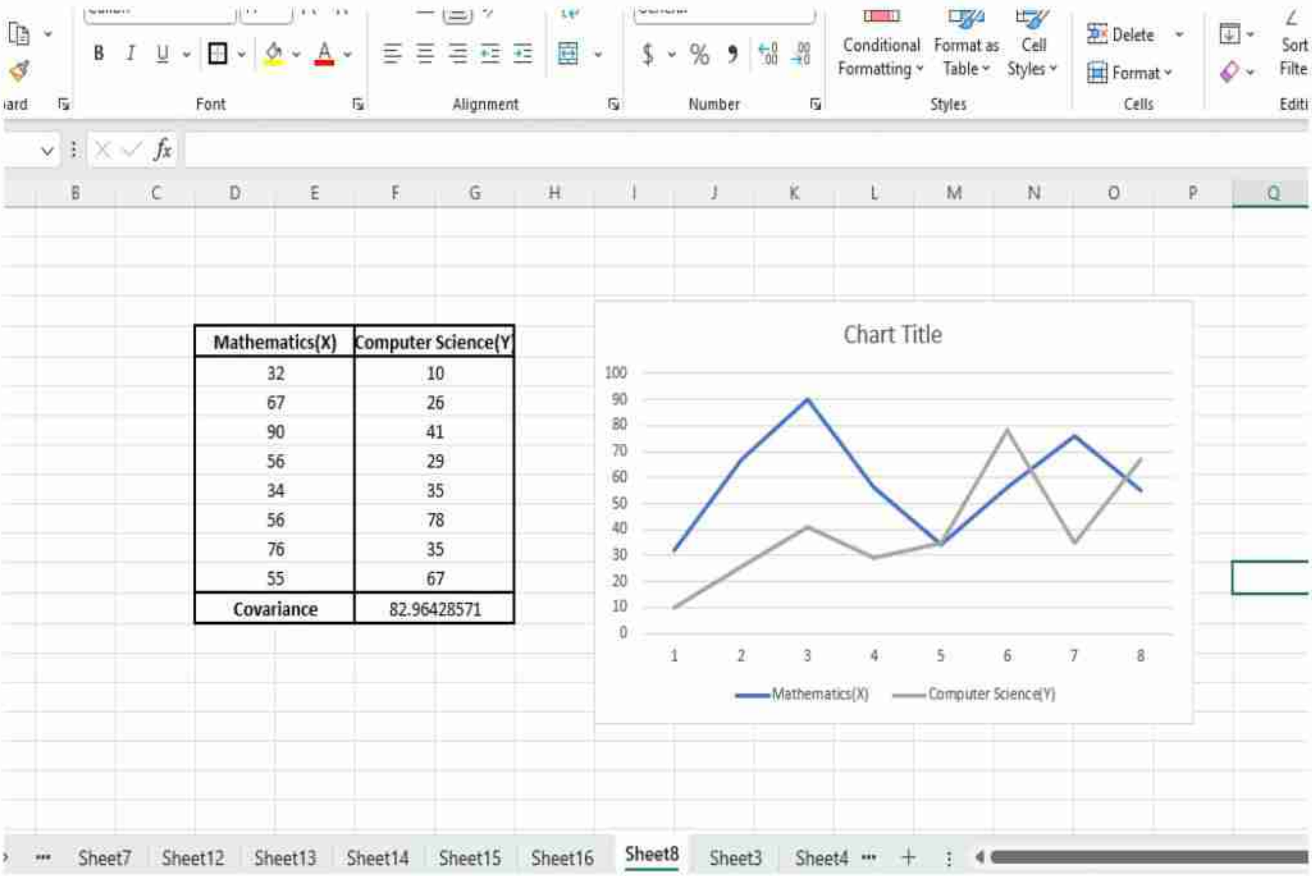


PRACTICAL-13

Presentation of bivariate data through scatter-plot diagrams and calculations of covariance.

The following data shows the marks obtained by 8 students in two different subjects: **Mathematics (X)** and **Computer Science(Y).** Calculate the covariance between the marks of Mathematics and Computer Science.

Formula used in excel:- =COVARIANCE.S(D6:E13,F6:G13)



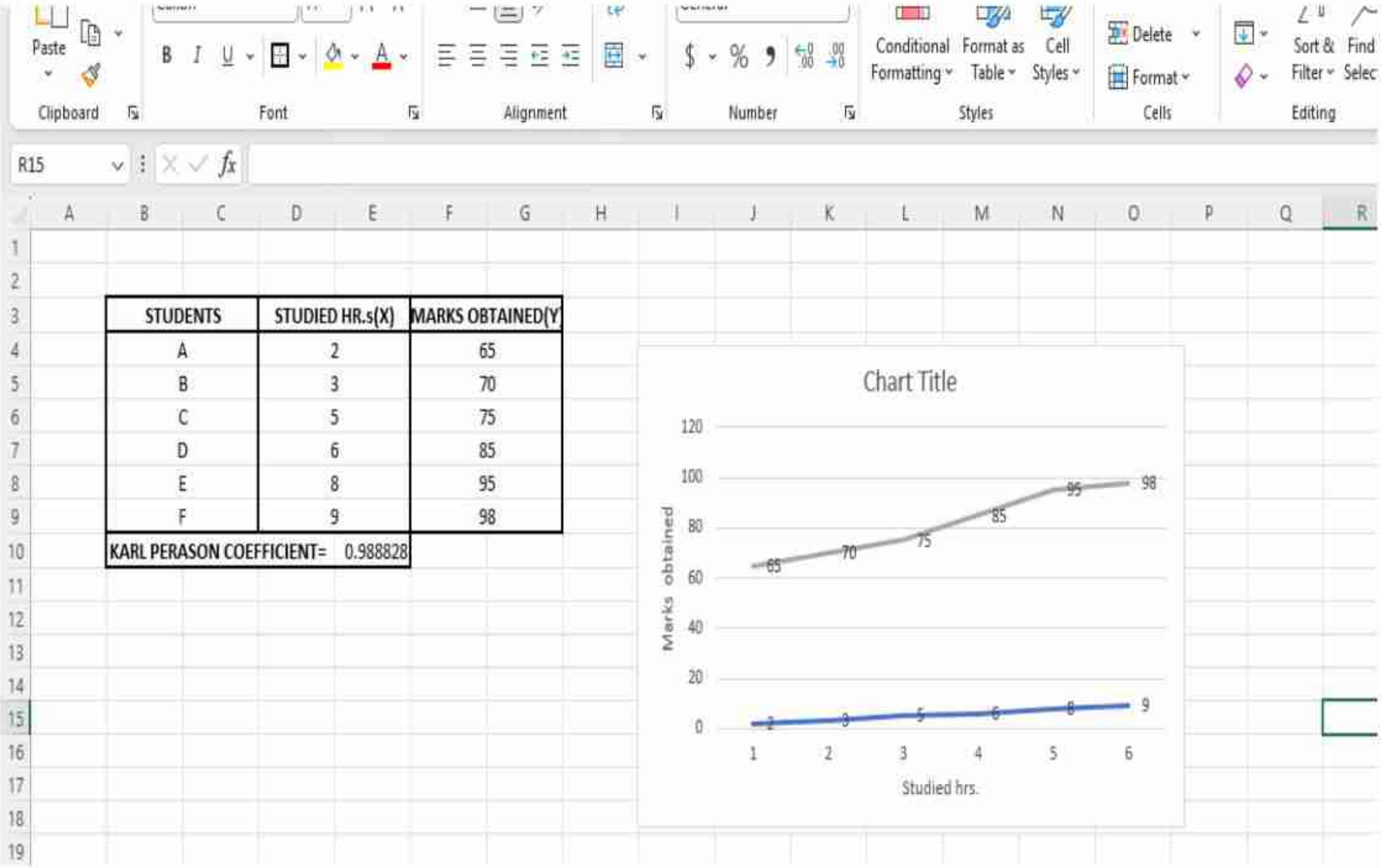
Mathematics(X)	Computer Science(Y
32	10
67	26
90	41
56	29
34	35
56	78
76	35
55	67
Covariance	82.96428571

PRACTICAL-14

Calculation of Karl Pearson's correlation coefficients.

The following table shows the number of hours studied and the marks obtained by six student in a test. Now calculate the Karl Pearson's Correlation Coefficient between hours studied and marks obtained.

Formula used in excel:- =PEARSON(D4:D9,F4:F9)



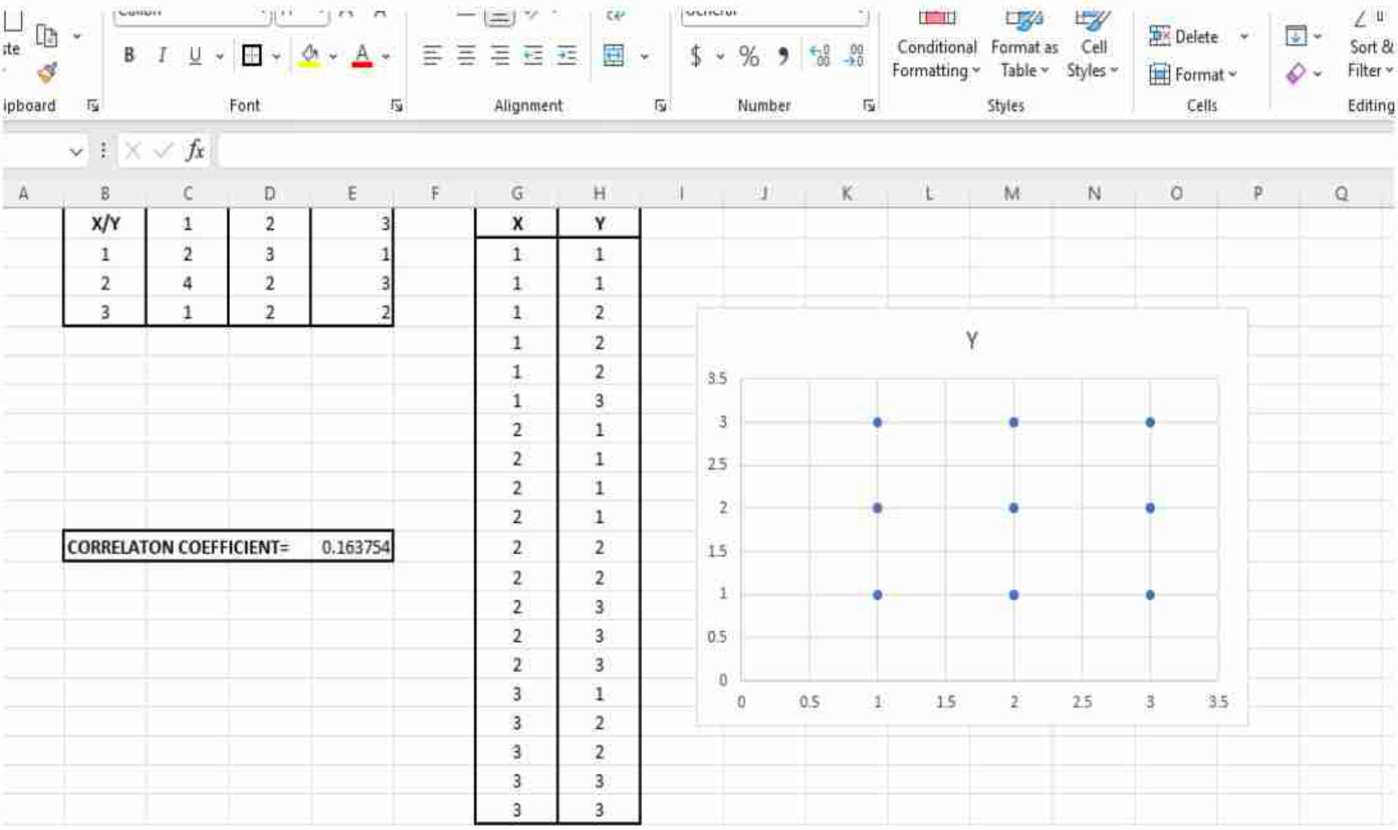
STUDENTS	STUDIED HR.s(X)	MARKS OBTAINED(Y)
A	2	65
B	3	70
C	5	75
D	6	85
E	8	95
F	9	98
KARL PERASON COEFFICIENT=	0.988828	

PRACTICAL-15

To find the correlation coefficient for a bivariate frequency distribution.

A teacher recorded the number of students scoring combinations of marks in two subjects: Math(X) and Science(Y). Calculate the correlation coefficient for a bivariate frequency distribution.

Formula used in excel :- =CORREL(G2:G21,H2:H21)



X/Y	1	2	3
1	2	3	1
2	4	2	3
3	1	2	2

CORRELATON COEFFICIENT=	0.163754

X	Y
1	1
1	1
1	2
1	2
1	2
1	3
2	1
2	1
2	1
2	1
2	2
2	2
2	3
2	3
2	3
3	1
3	2
3	2
3	3
3	3

PRACTICAL-16

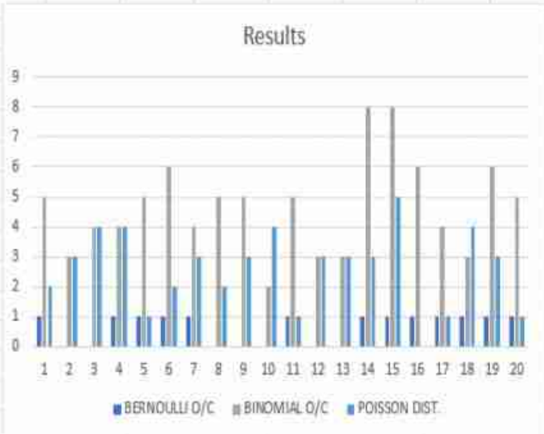
Generating Random numbers from discrete (Bernoulli, Binomial, Poisson) distributions.

Formula to generate random numbers from:-

- Bernoulli distribution:- =IF(RAND()<=0.6, 1, 0)
- Binomial distribution:- =BINOM.INV(10,0.5, RAND())
- Poisson distribution:- =POISSON.INV(RAND(), 3)

TRIAL	BERNOULLI O/C	BINOMIAL O/C	POISSON DIST.
1	1	5	2
2	0	3	3
3	0	4	4
4	1	4	4
5	1	5	1
6	1	6	2
7	1	4	3
8	0	5	2
9	0	5	3
10	0	2	4
11	1	5	1
12	0	3	3
13	0	3	3
14	1	8	3
15	1	8	5
16	1	6	0
17	1	4	1
18	1	3	4
19	1	6	3
20	1	5	1

no. of trials (n)	20
prob. Of success (p)	0.5
mean (λ)	3

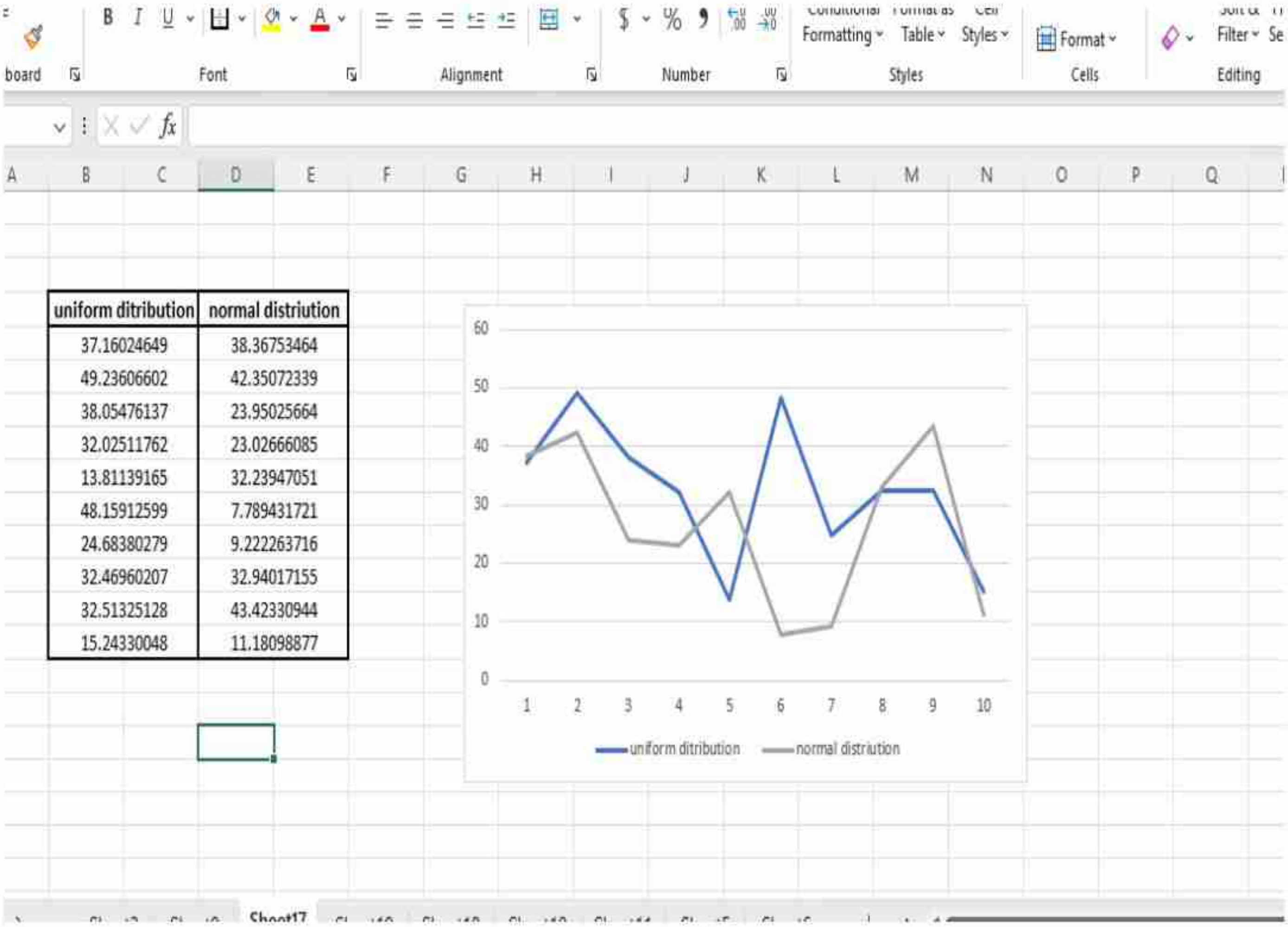


PRACTICAL-17

Generating Random numbers from continuous (Uniform, Normal) distributions.

Formula to find random numbers from continuous:-

- Uniform distribution:- =RAND()*(upper limit-lower limit)+lower limit
- Normal Distribution:- =NORMINV(RAND(),mean, standard deviation)



uniform ditribution	normal distriution
37.16024649	38.36753464
49.23606602	42.35072339
38.05476137	23.95025664
32.02511762	23.02666085
13.81139165	32.23947051
48.15912599	7.789431721
24.68380279	9.222263716
32.46960207	32.94017155
32.51325128	43.42330944
15.24330048	11.18098877

PRACTICAL-18

Find the entropy from the given data set.

Given the weather conditions and the corresponding decision to play or not, calculate the entropy of the "Play" decision.

Formula used in excel:- =(p*log2(p))*2

Weather	Play
sunny	0
sunny	0
overcast	1
rainy	1
rainy	0
rainy	1
overcast	1
sunny	0
sunny	1
rainy	1

yes=1	
no=0	

Class	Count	(p)	p*log2(p)	Entropy
yes	5	0.5	-0.5	
no	5	0.5	-0.5	1
total	10