PRACTICAL FILE

Probability for Computing

DSC - 06, 2025

RAMANUJAN COLLEGE

Submitted by-

Name : Himanshu Kumar Jha

College roll no. : 24570026

University roll no. : 24020570043

Course : B.Sc.(Hons.)
Computer Science

Semester : 2

Submitted to-

Dr. Aakash

Assistant Professor
(Operational Research),
Department of Computer Science,
Ramanujan College,
University of Delhi,
CR Park Main Road,
Block H, Kalkaji,
New Delhi-110019

ACKNOWLEDGEMENT

I would like to express my heartfelt gratitude to my teacher **Dr. Aakash**, Assistant Professor (Operational Research), Ramanujan College, University of Delhi, for providing me with the opportunity to work on this practical file for Probability 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 motivation and encouragement, which has been a source of strength in completing this work successfully.

Himanshu Kumar Jha

Name : Himanshu Kumar Jha
College roll no. : 24570026
University roll no. : 24020570043

INDEX

PRACTICAL - 01

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

Binomial Distribution:

The binomial distribution is a discrete probability distribution. It describes the outcome of binary scenarios, e.g. toss of a coin.

Binomial Distribution Formula :

The binomial distribution formula is for any random variable X, given by:

$P(x:n,p) = {}_nC_x\ p_x\ (1-p)_{n-x}$

Or

$P(x:n,p) = {}_nC_x\ p_x\ (q)_{n-x}$

Where,
n = the number of experiments
x = 0, 1, 2, 3, 4, …
p = Probability of Success in a single experiment
q = Probability of Failure in a single experiment = 1 − p

The binomial distribution formula can also be written in the form of n-Bernoulli trials, where ${}_nC_x$ = n!/x!(n-x)!. Hence,

$P(x:n,p) = n!/[x!(n-x)!].p_x.(q)_{n-x}$

Binomial Distribution Mean and Variance :

For a binomial distribution, the mean, variance and standard deviation for the given number of success are represented using the formulas:

Mean, μ = np
Variance, σ_2 = npq
Standard Deviation σ= √(npq)

Where p is the probability of success, q is the probability of failure, where q = 1-p

Implementation in Excel:

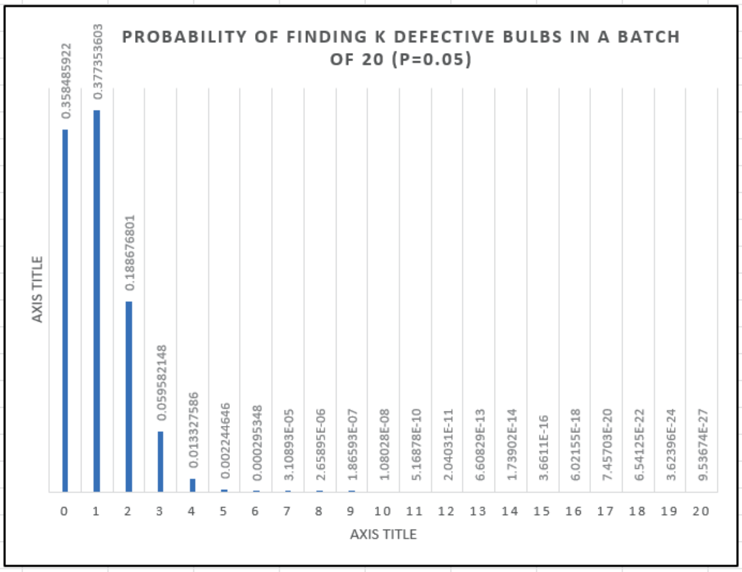
=BINOM.DIST(number_s, trials, probability_s, cumulative)

Where:
number_s: number of successes.
trials: total number of trials.
probability_s: probability of success on each trial.
cumulative: TRUE returns the cumulative probability; FALSE returns the exact probability

Scenario 1: Quality Control in Light Bulb Manufacturing

- **Situation**: A factory produces light bulbs. Historically, 5% of the bulbs are defective. A quality inspector randomly selects a batch of 20 bulbs. We want to know the probability of finding a specific number of defective bulbs in this batch.
- **Parameters:**
  - Number of trials (bulbs selected), **n = 20**
  - Probability of success (a bulb being defective), **p = 0.05**
  - Number of successes (defective bulbs), **k** (ranges from 0 to 20)

Number of Defective Bulbs (k)	Probability P(X=k)
0	0.358485922
1	0.377353603
2	0.188676801
3	0.059582148
4	0.013327586
5	0.002244646
6	0.000295348
7	3.10893E-05
8	2.65895E-06
9	1.86593E-07
10	1.08028E-08
11	5.16878E-10
12	2.04031E-11
13	6.60829E-13
14	1.73902E-14
15	3.6611E-16
16	6.02155E-18
17	7.45703E-20
18	6.54125E-22
19	3.62396E-24
20	9.53674E-27

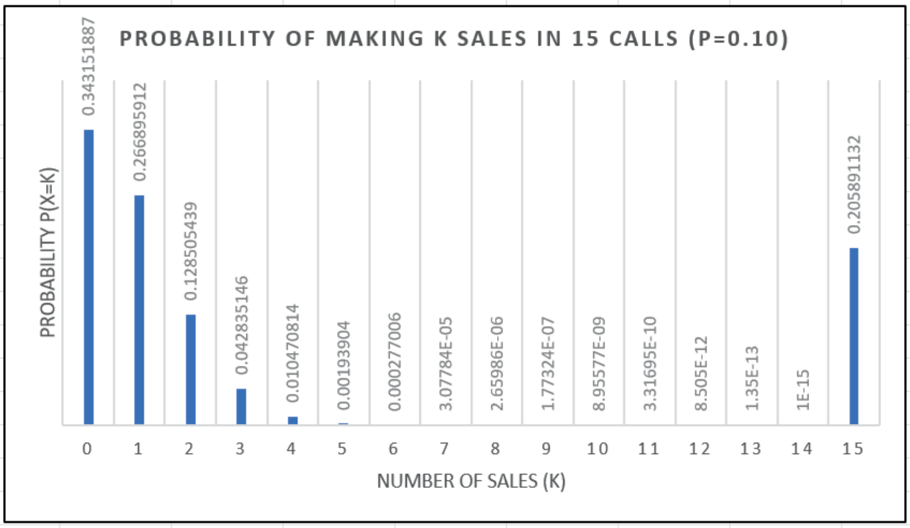


<u>Scenario 2</u>: Telemarketing Success Rate

- **Situation:** A telemarketer makes 15 calls per shift. The probability of making a sale on any single call is 10%. We want to find the probability of making a certain number of sales during a shift.

- **Parameters:**
  - Number of trials (calls made), n = 15
  - Probability of success (making a sale), p = 0.10
  - Number of successes (sales made), k (ranges from 0 to 15)

Number of Sales (k)	Probability P(X=k)
0	0.343151887
1	0.266895912
2	0.128505439
3	0.042835146
4	0.010470814
5	0.00193904
6	0.000277006
7	3.07784E-05
8	2.65986E-06
9	1.77324E-07
10	8.95577E-09
11	3.31695E-10
12	8.505E-12
13	1.35E-13
14	1E-15
15	0.205891132

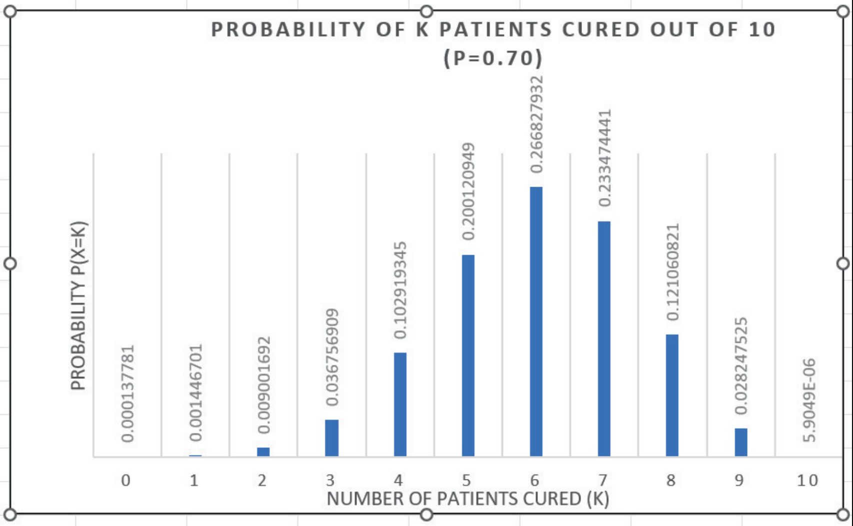


<u>Scenario 3:</u> Drug Effectiveness

- **Situation:** A new drug is claimed to be effective in curing a certain condition in 70% of patients. It is administered to 10 patients. What is the probability that a specific number of these patients are cured?

- **Parameters:**
  - Number of trials (patients treated), n = 10
  - Probability of success (patient is cured), p = 0.70
  - Number of successes (patients cured), k (ranges from 0 to 10)

Number of Patients Cured (k)	Probability P(X=k)
0	0.000137781
1	0.001446701
2	0.009001692
3	0.036756909
4	0.102919345
5	0.200120949
6	0.266827932
7	0.233474441
8	0.121060821
9	0.028247525
10	5.9049E-06



PRACTICAL - 02

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

Multinomial Distribution:

In probability theory, the multinomial distribution is a generalization of the binomial distribution. For example, it models the probability of counts for each side of a k-sided die rolled n times. For n independent trials each of which leads to a success for exactly one of k categories, with each category having a given fixed success probability, the multinomial distribution gives the probability of any particular combination of numbers of successes for the various categories.

$$p(X = x) = p(X_1 = x_1, \cdots, X_k = x_k) = \frac{n!}{x_1! x_2! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k}$$

Implementation in Excel:

Multinomial= MULTINOMIAL(x1, x2, x3)

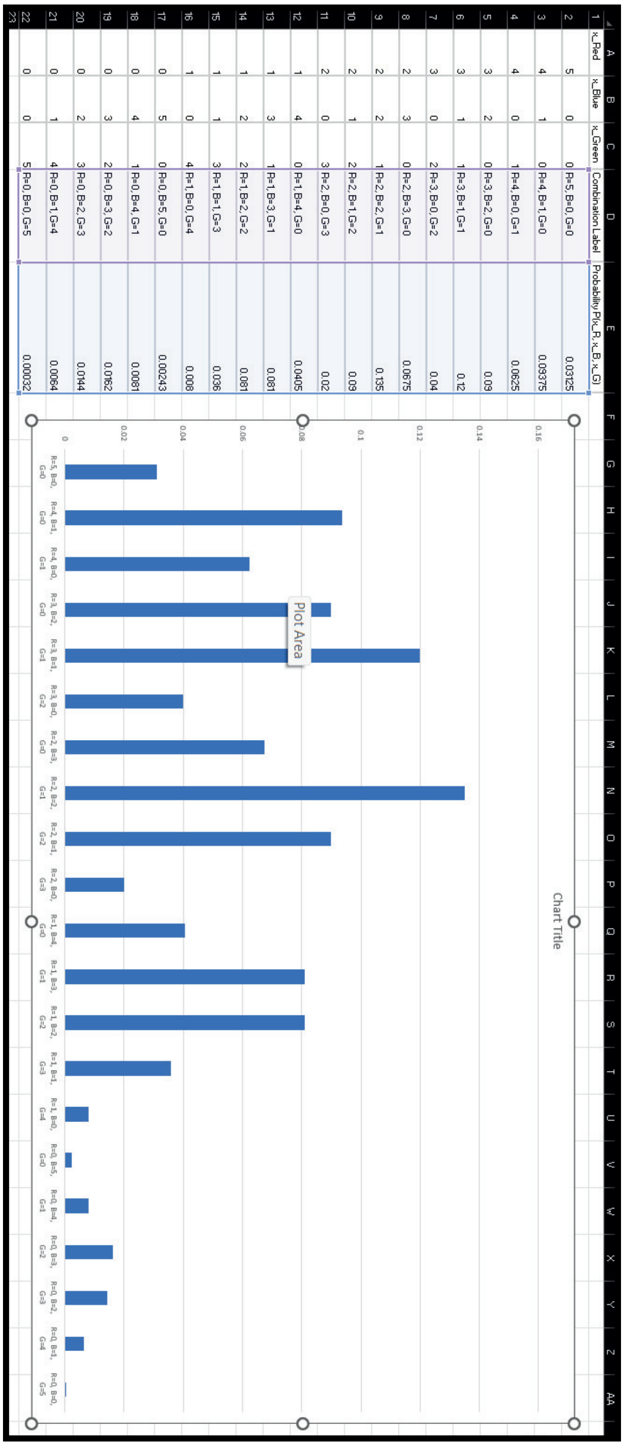
Probability = MULTINOMIAL*PRODUCT(p1^x1, p2^x2, p3^x3)

Scenario: Marble Colors

- **Situation:** Imagine a large bag filled with marbles where 50% are Red (R), 30% are Blue (B), and 20% are Green (G). You randomly draw 5 marbles, putting each one back before drawing the next (drawing with replacement). We want to calculate the probability of drawing each specific combination of colors (e.g., 2 Red, 2 Blue, 1 Green).

Parameters:

- Number of trials (marbles drawn), n = 5
- Number of outcomes (colors), k = 3
- Probabilities: p_Red = 0.50, p_Blue = 0.30, p_Green = 0.20 (Note: 0.5 + 0.3 + 0.2 = 1.0)
- Outcome counts: x_Red, x_Blue, x_Green, such that x_Red + x_Blue + x_Green = 5.



x_Red	x_Blue	x_Green	Combination Label	Probability P(x_R, x_B, x_G)
5	0	0	R=5, B=0, G=0	0.03125
4	1	0	R=4, B=1, G=0	0.09375
4	0	1	R=4, B=0, G=1	0.0625
3	2	0	R=3, B=2, G=0	0.09
3	1	1	R=3, B=1, G=1	0.12
3	0	2	R=3, B=0, G=2	0.04
2	3	0	R=2, B=3, G=0	0.0675
2	2	1	R=2, B=2, G=1	0.135
2	1	2	R=2, B=1, G=2	0.09
2	0	3	R=2, B=0, G=3	0.02
1	4	0	R=1, B=4, G=0	0.0405
1	3	1	R=1, B=3, G=1	0.081
1	2	2	R=1, B=2, G=2	0.081
1	1	3	R=1, B=1, G=3	0.036
1	0	4	R=1, B=0, G=4	0.008
0	5	0	R=0, B=5, G=0	0.00243
0	4	1	R=0, B=4, G=1	0.0081
0	3	2	R=0, B=3, G=2	0.0162
0	2	3	R=0, B=2, G=3	0.0144
0	1	4	R=0, B=1, G=4	0.0064
0	0	5	R=0, B=0, G=5	0.00032

PRACTICAL - 03

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

Poisson Distribution Definition :

The Poisson distribution is a discrete probability function that means the variable can only take specific values in a given list of numbers, probably infinite. A Poisson distribution measures how many times an event is likely to occur within “x” period of time. In other words, we can define it as the probability distribution that results from the Poisson experiment. A Poisson experiment is a statistical experiment that classifies the experiment into two categories, such as success or failure. Poisson distribution is a limiting process of the binomial distribution.

A Poisson random variable “x” defines the number of successes in the experiment. This distribution occurs when there are events that do not occur as the outcomes of a definite number of outcomes. Poisson distribution is used under certain conditions. They are:

- The number of trials “n” tends to infinity
- Probability of success “p” tends to zero
- np = 1 is finite

Poisson Distribution Formula

The formula for the Poisson distribution function is given by:

$f(x) = (e^{-\lambda} \lambda^{x})/x!$

Where,

e is the base of the logarithm

x is a Poisson random variable

λ is an average rate of value

Implementation in Excel:

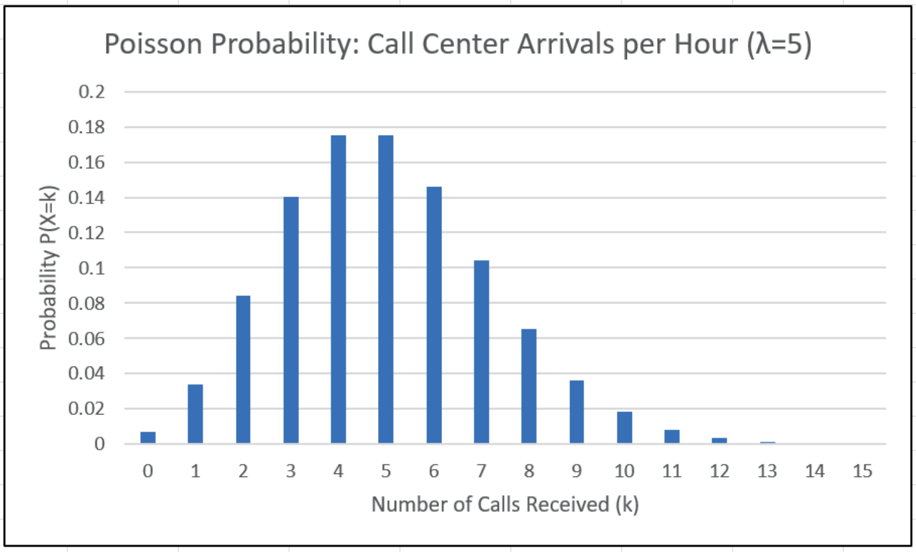
POISSON.DIST(number_s, average, cumulative)
POISSON.DIST(k, λ , FALSE)

Scenario: Call Center Arrivals

- Situation: A small customer service call center receives calls at an average rate of 5 calls per hour. We want to know the probability of receiving a specific number of calls (k) in any given hour.

- Parameters:
  - Average rate (mean), λ = 5 calls per hour.
  - Number of events (calls received), k (can theoretically be any non-negative integer 0, 1, 2, ...)

	A	B
1	Number of Calls (k)	Probability P(X=k)
2	0	0.006737947
3	1	0.033689735
4	2	0.084224337
5	3	0.140373896
6	4	0.17546737
7	5	0.17546737
8	6	0.146222808
9	7	0.104444863
10	8	0.065278039
11	9	0.036265577
12	10	0.018132789
13	11	0.008242177
14	12	0.00343424
15	13	0.001320862
16	14	0.000471736
17	15	0.000157245



PRACTICAL - 04

- **Plotting and fitting of Geometric distribution and graphical representation of probabilities**

Geometric distribution :

Geometric distribution is a probability distribution that defines the number of trials required to get the first success in a series of independent and identically distributed Bernoulli trials, where each trial has two possible outcomes: success or failure. The trials are conducted until the first success is observed, and the probability of success in each trial is constant.

The geometric distribution is commonly used in various real-life circumstances. In the financial industry, it is used to estimate the financial rewards of making a given decision in a cost-benefit analysis.

Geometric distributions are probability distributions that are based on three key assumptions.

- Trails are independent
- Each trial has one of two outcomes: success or failure
- For each trial, the probability of success, p, is constant across trials.

Geometric Distribution Formulas

PMF: $P(X = x) = (1 - p)^{x-1} \ (p)$

CDF: $P(x <= x) = 1 - (1 - p)^{x}$

Mean: $E(X) = \frac{1}{p}$

Variance: $Var(X) = \frac{(1\ p)}{p^2}$

Standard Deviation: $\sigma = \sqrt{\frac{1 - p}{p}}$

Implementation in Excel:

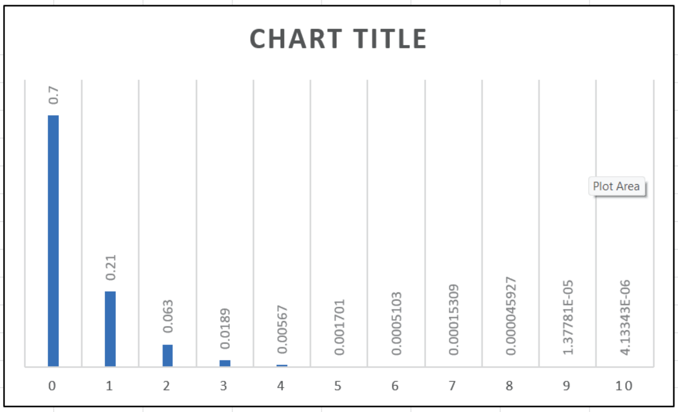
POISSON.DIST(number_s, average, cumulative)
POISSON.DIST(k, λ , FALSE)

Scenario: Basketball Free Throws

- **Situation**: A basketball player has a 70% chance of making any single free throw (this is their success probability). We want to know the probability that their first successful shot occurs immediately (0 failures), or after 1 miss, or after 2 misses, and so on.

- **Parameters**:
  - Probability of success (making the shot), $p = 0.70$.
  - Number of failures (missed shots) before the first success, k (can be 0, 1, 2, ...).

Number of Misses before 1st Success (k)	Probability P(Y=k)	Interpretation (First success on trial #)
0	0.7	1st trial
1	0.21	2nd trial
2	0.063	3rd trial
3	0.0189	4th trial
4	0.00567	5th trial
5	0.001701	6th trial
6	0.0005103	7th trial
7	0.00015309	8th trial
8	0.000045927	9th trial
9	1.37781E-05	10th trial
10	4.13343E-06	11th trial



PRACTICAL - 05

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

Uniform Distribution :

A Uniform Distribution is a type of probability distribution in which every outcome in a given range is equally likely to occur. That means there is no bias—no outcome is more likely than another within the specified set.

- It is also known as rectangular distribution (continuous uniform distribution).
- It has two parameters a and b: a = minimum and b = maximum. The distribution is written as U (a, b).

Uniform Distribution Formula

A random variable X is said to be uniformly distributed over the interval $-\infty < a < b < \infty$. Formulae for uniform distribution:

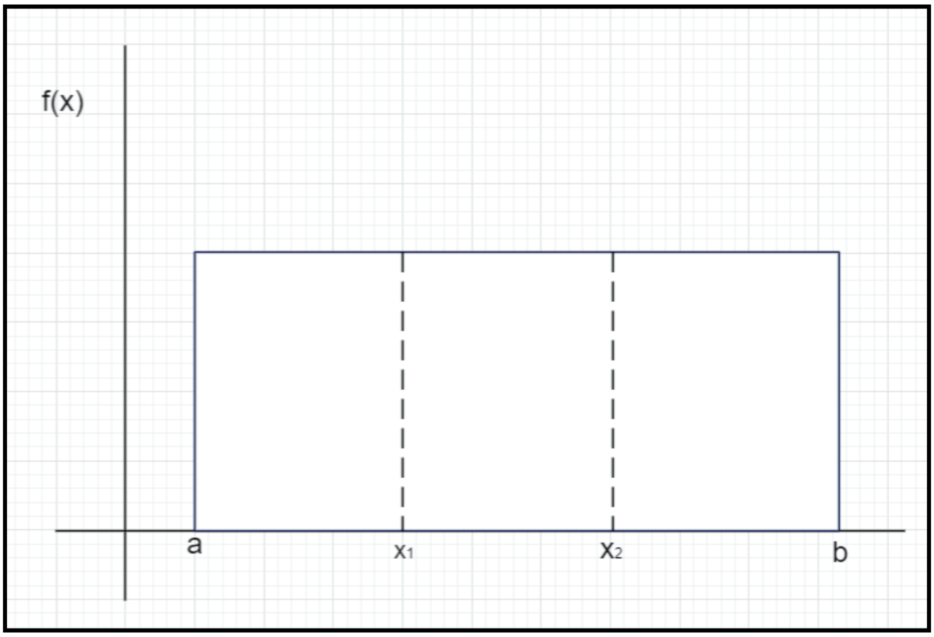
Probability density function(pdf)	$f(x) = 1/(b - a),\ a \le x \le b$
Mean(μ)	$\int_a^b x.f(x)\,dx = \frac{1}{b-a}[\frac{x^2}{2}]_a^b$ $= (a + b)/2$
Variance (σ^2)	$\int_a^b x.f(x)\,dx = \frac{1}{b-a}[\frac{x^2}{2}]_a^b$ $= \mu_2' - \mu^2 = \int_a^b x^2.\frac{1}{b-a}dx - (\frac{a+b}{2})^2$ $= (b - a)^2/12$
Standard Deviation (σ)	$= \sqrt{\frac{(b-a)^2}{12}}$
Cumulative Distribution function (CDF)	$= (x - a)/(b - a)$ for $x \in [a , b]$
Median	$= (a + b)/2$
For the conditional probability = P(c < x < d)	$= (d - c) \times f(x)$ $= (d - c)/(b - a)$

For calculating probability, we need:

a: minimum value in the distribution
b: maximum value in the distribution
x_1: the minimum value you're interested in
x_2: the maximum value you're interested in

$$P = (x_2 - x_1) / (b - a)$$

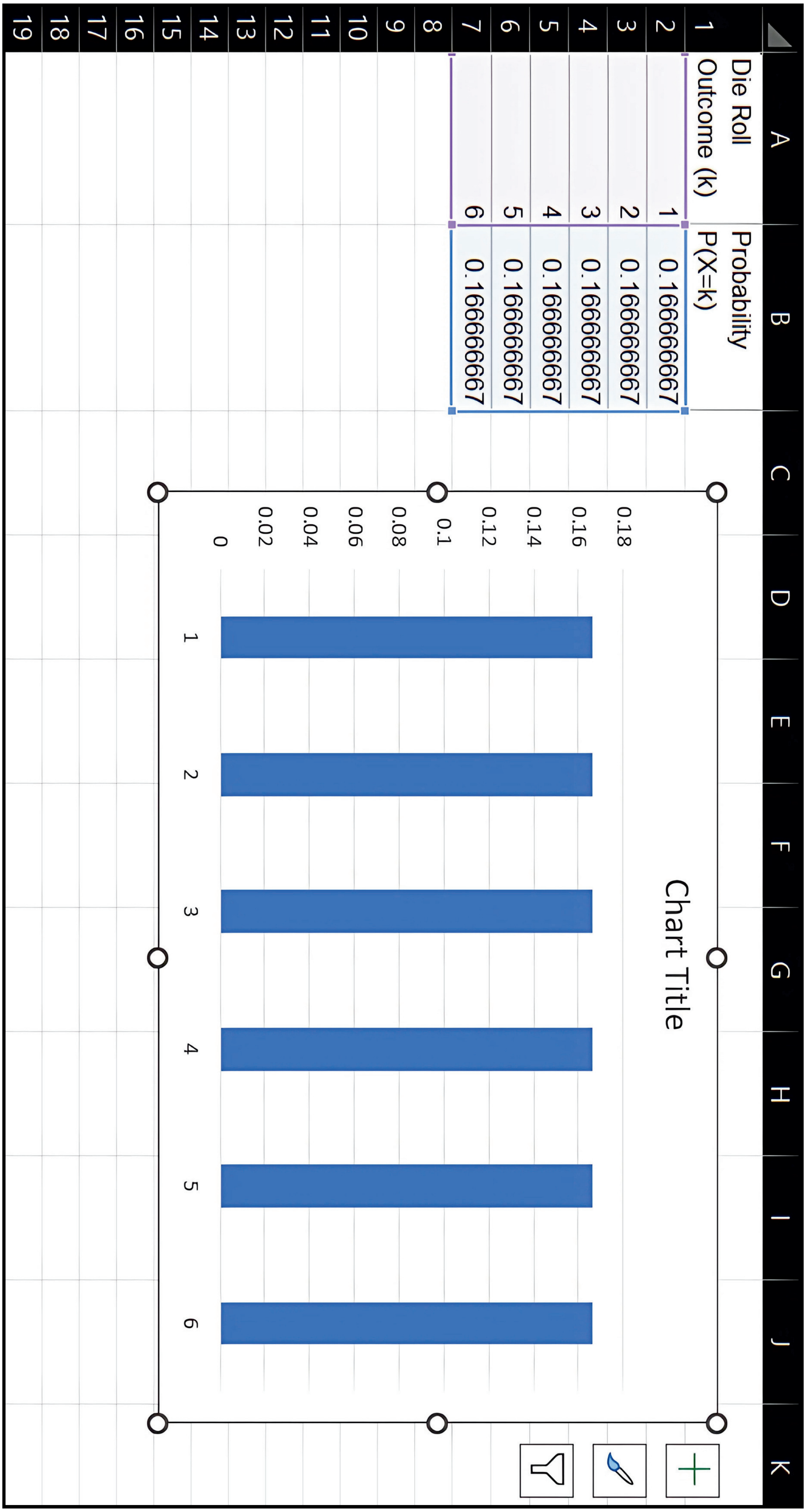
Example graph



Scenario: Rolling a Fair Six-Sided Die

- Situation: You roll a standard, fair, six-sided die. Each number from 1 to 6 has an equal chance of appearing on the top face.

- **Parameters:**
  - Possible Outcomes: {1, 2, 3, 4, 5, 6}
  - Number of possible outcomes, N = 6
  - Probability of each outcome: Since each outcome is equally likely, the probability is P(X=k) = 1 / N = 1 / 6.



Die Roll Outcome (k)	Probability P(X=k)
1	0.166666667
2	0.166666667
3	0.166666667
4	0.166666667
5	0.166666667
6	0.166666667

PRACTICAL - 06

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

What is Exponential Distribution?

In Probability theory and statistics, the exponential distribution is a continuous probability distribution that often concerns the amount of time until some specific event happens. It is a process in which events happen continuously and independently at a constant average rate. The exponential distribution has the key property of being memoryless. The exponential random variable can be either more small values or fewer larger variables. For example, the amount of money spent by the customer on one trip to the supermarket follows an exponential distribution.

Exponential Distribution Formula

The continuous random variable, say X is said to have an exponential distribution, if it has the following probability density function:

$$f_X(x|\lambda) = \begin{cases} \lambda e^{-\lambda x} & for\ x > 0 \\ 0 & for\ x \le 0 \end{cases}$$

Where

λ is called the distribution rate.

Implementation in Excel:

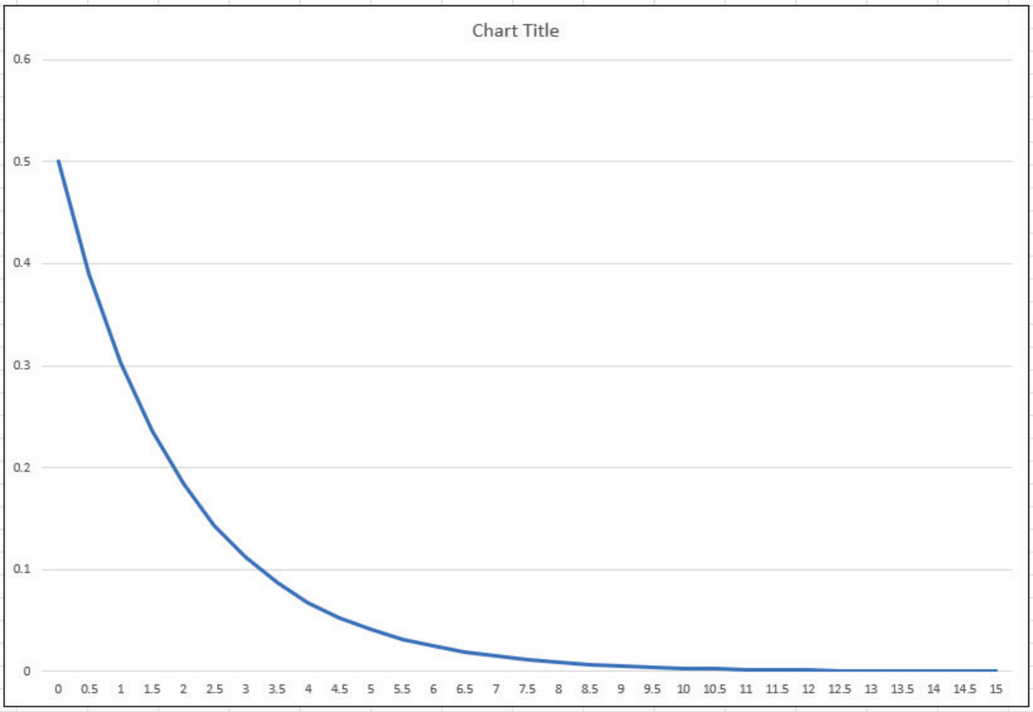
EXPON.DIST(X, lambda, FALSE)

Scenario: Time Between Customer Arrivals

- Situation: Customers arrive at a checkout counter independently and at a constant average rate. The average time between customer arrivals is 2 minutes. We want to visualize the probability density of this waiting time between customers.

- Parameters:
  - Average time between arrivals, Mean = $1/\lambda=2$ minutes.
  - Rate parameter, $\lambda=1/\text{Mean}=1/2=0.5$ arrivals per minute.
  - Time between arrivals, x (measured in minutes, $x \geq 0$).

Arrivals (x minutes)	Probability Density f(x)
0	0.5
0.5	0.389400392
1	0.30326533
1.5	0.236183276
2	0.183939721
2.5	0.143252398
3	0.11156508
3.5	0.086886972
4	0.067667642
4.5	0.052699612
5	0.041042499
5.5	0.031963931
6	0.024893534
6.5	0.019387104
7	0.015098692
7.5	0.011758873
8	0.009157819
8.5	0.007132117
9	0.005554498
9.5	0.004325848
10	0.003368973
10.5	0.002623759
11	0.002043386
11.5	0.00159139
12	0.001239376
12.5	0.000965227
13	0.00075172
13.5	0.00058544
14	0.000455941
14.5	0.000355087
15	0.000276542

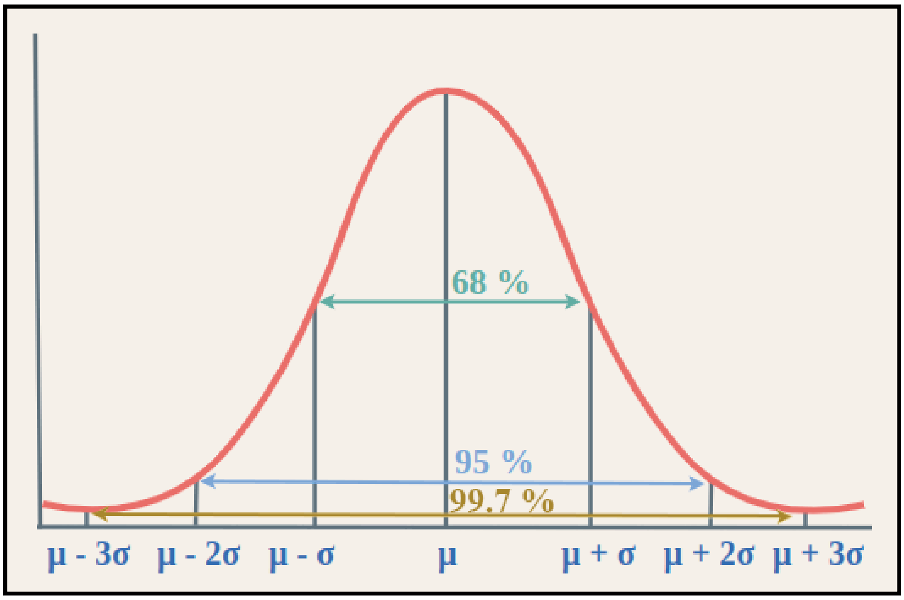


PRACTICAL - 07

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

What is Normal Distribution?
Normal distribution, also known as the Gaussian distribution, is a continuous probability distribution that is symmetric about the mean, depicting that data near the mean are more frequent in occurrence than data far from the mean.
We define Normal Distribution as the probability density function of any continuous random variable for any given system. Now for defining Normal Distribution suppose we take f(x) as the probability density function for any random variable X.
Also, the function is integrated between the interval, (x, {x + dx}) then, $f(x) \geq 0 \quad x \in (-\infty,+\infty)$, $\int_{-\infty}^{+\infty} f(x) = 1$

We observe that the curve traced by the upper values of the Normal Distribution is in the shape of a Bell, hence Normal Distribution is also called the "**Bell Curve**".



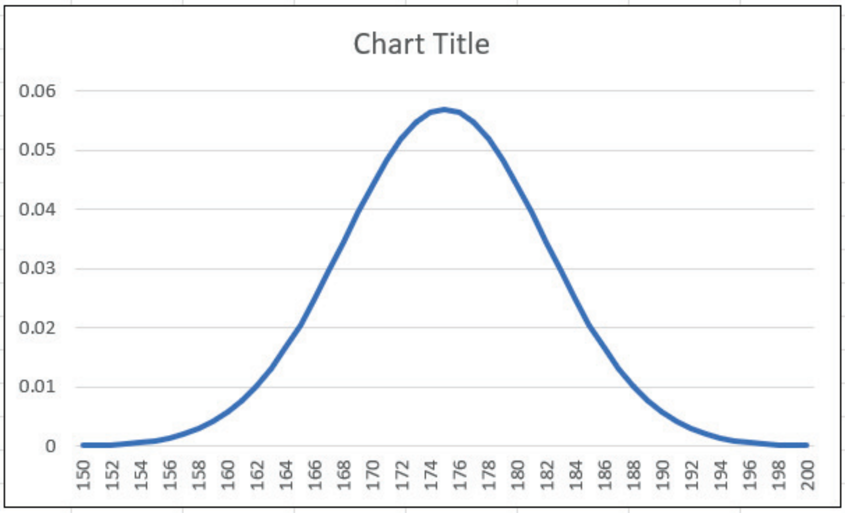
Implementation in Excel:

=NORMDIST(x, mean, standard_dev, cumulative)

Scenario: Heights of Adult Males

- Situation: Let's assume the heights of adult males in a particular population follow an approximately Normal distribution. The average height (mean) is 175 cm, and the standard deviation is 7 cm. We want to visualize the distribution of these heights.
- **Parameters:**
  - Mean, $\mu = 175$ cm
  - Standard Deviation, $\sigma = 7$ cm
  - Height, x (measured in cm)

Height (x cm)	Probability Density f(x)
150	9.68449E-05
151	0.00015967
152	0.000257934
153	0.000408253
154	0.000633121
155	0.000962014
156	0.001432231
157	0.002089206
158	0.002985977
159	0.004181465
160	0.005737297
161	0.007712995
162	0.010159577
163	0.013111882
164	0.016580258
165	0.020542552
166	0.024937582
167	0.029661365
168	0.034567246



PRACTICAL - 08

- **Calculation of cumulative distribution functions for Exponential and Normal distribution.**

Normal Distribution								Exponential Distribution				
Employee	incentive	Normal Distribution						No Of Days	Exponential Distribution			
EMP1	1000	0.572137292						300	0.527633447			
EMP2	2000	0.641935216		mean	5500			320	0.550671036			
EMP3	3000	0.707279533		standand Deviation	3027.650354			350	0.58313798		λ	0.0025
EMP4	4000	0.766470549						400	0.632120559			
EMP5	5000	0.81834893						420	0.650062251			
EMP6	6000	0.862343557						450	0.675347533			
EMP7	7000	0.898442582						500	0.713495203			
EMP8	8000	0.92710243						520	0.727468207			
EMP9	9000	0.949118248						550	0.747160404			
EMP10	10000	0.965481826						600	0.77686984			

PRACTICAL - 09

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

Euclidean distance

Euclidean distance is the distance between two real distinct value .It is calculated by the square root of the sum of the squared difference elements in
two vectors.

$$\textbf{Euclidean Distance} = |X - Y| = \sqrt{\sum_{i=1}^{i=n} (x_i - y_i)^2}$$

X: Array or vector X

Y: Array or vector Y

x_i: Values of horizontal axis in the coordinate plane

y_i: Values of vertical axis in the coordinate plane

n: Number of observations

Implementation in Excel:

= SORT(SUM X MYZ(array_X,array_Y))

Trials	Binomial Distribution	Poisson Distribution		distance			
1	0.08957952	0.268128018		0.178548		n	0.4
2	0.20901888	0.053625604		0.155393		k	8
3	0.27869184	0.00715008		0.271542		λ	0.4
4	0.2322432	0.000715008		0.231528			
5	0.12386304	5.72006E-05		0.123806			
6	0.04128768	3.81338E-06		0.041284			
7	0.00786432	2.17907E-07		0.007864			
8	0.00065536	1.08954E-08		0.000655			

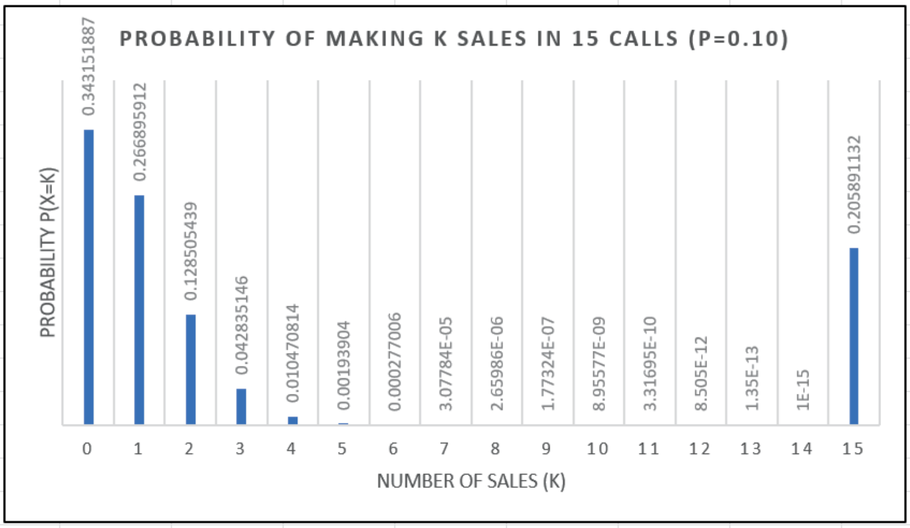
PRACTICAL - 10

● **Application problems based on the Binomial distribution.**

Scenario 2: Telemarketing Success Rate

- **Situation:** A telemarketer makes 15 calls per shift. The probability of making a sale on any single call is 10%. We want to find the probability of making a certain number of sales during a shift.
- **Parameters:**
  - Number of trials (calls made), n = 15
  - Probability of success (making a sale), p = 0.10
  - Number of successes (sales made), k (ranges from 0 to 15)

Number of Sales (k)	Probability P(X=k)
0	0.343151887
1	0.266895912
2	0.128505439
3	0.042835146
4	0.010470814
5	0.00193904
6	0.000277006
7	3.07784E-05
8	2.65986E-06
9	1.77324E-07
10	8.95577E-09
11	3.31695E-10
12	8.505E-12
13	1.35E-13
14	1E-15
15	0.205891132



PRACTICAL - 11

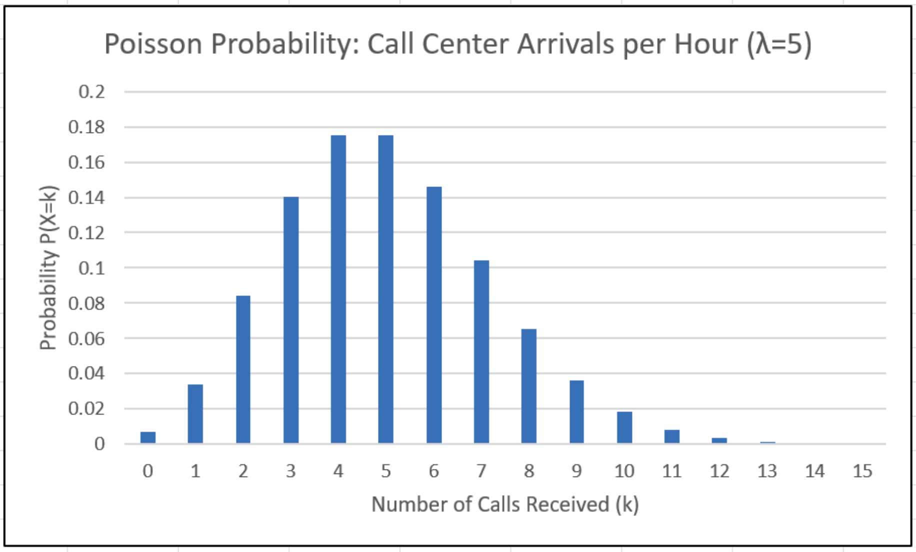
● Application problems based on the Poisson distribution.

Scenario: Call Center Arrivals

- Situation: A small customer service call center receives calls at an average rate of 5 calls per hour. We want to know the probability of receiving a specific number of calls (k) in any given hour.

- Parameters:
    - Average rate (mean), λ = 5 calls per hour.
    - Number of events (calls received), k (can theoretically be any non-negative integer 0, 1, 2, ...)

Number of Calls (k)	Probability P(X=k)
0	0.006737947
1	0.033689735
2	0.084224337
3	0.140373896
4	0.17546737
5	0.17546737
6	0.146222808
7	0.104444863
8	0.065278039
9	0.036265577
10	0.018132789
11	0.008242177
12	0.00343424
13	0.001320862
14	0.000471736
15	0.000157245

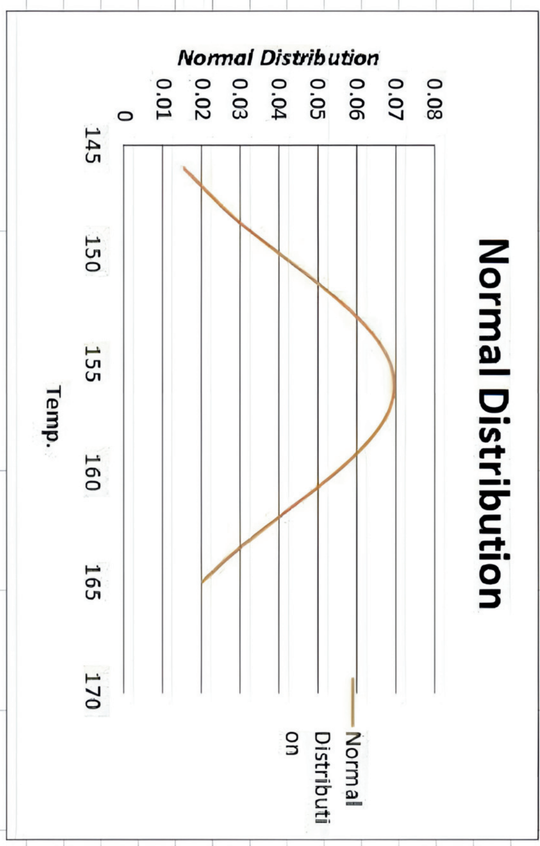


PRACTICAL - 12

● Application problems based on the Normal distribution.

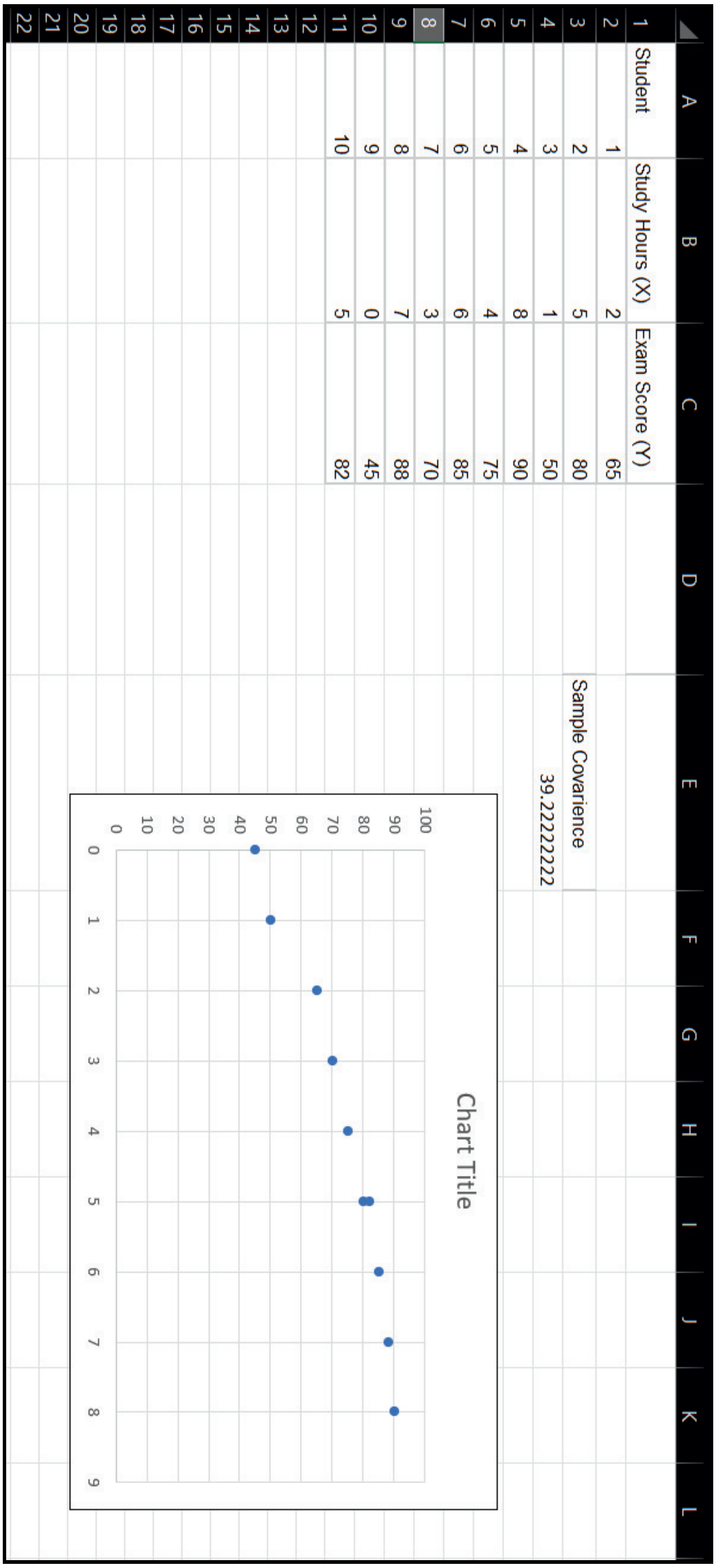
Temp.	Normal Distribution	Temp.	Normal Distribution
146	0.015374765	166	0.014888007
148	0.026566325	167	0.010783701
149	0.033356691	168	0.007575753
150	0.04062198	169	0.00516191
151	0.047980623	171	0.002186551
152	0.054966405	173	0.000819625
153	0.06107387		
154	0.065817321		
155	0.068794167		
156	0.069741244		
157	0.068573201		
158	0.065395191		
159	0.060487252		
160	0.054263594		
161	0.047214991		
162	0.039845376		
163	0.03261389		
164	0.025891304		
165	0.019935719		

mean	155.947
standand Deviation	5.7008



PRACTICAL - 13

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



Student	Study Hours (X)	Exam Score (Y)
1	2	65
2	5	80
3	1	50
4	8	90
5	4	75
6	6	85
7	3	70
8	7	88
9	0	45
10	5	82

Sample Covarience
39.22222222

PRACTICAL - 14

- **Calculation of Karl Pearson's correlation coefficients.**

X	Y1	Y2	Y3
2	38	63	10
3	56	59	35
8	74	68	15
11	33	79	24
10	40	85	12
Karl pearson's correlation	XY1	-0.14493936	
	XY2	0.889803569	
	XY3	-0.140953713	

PRACTICAL - 15

- **To find the correlation coefficient for a bivariate frequency distribution.**

Marks	16_18	18_20	20-22	22_24	total
10_20	5	1	1	0	7
20_30	2	0	4	2	8
30_40	1	3	5	3	12
40_50	2	2	2	1	7
50_60	3	0	3	1	7
60_70	0	1	0	5	6
total	13	7	15	12	

margerial frequency distribution of X:

marks	total
15	7
25	8
34	12
33	7
53	7
42	6

correlation coffecient along X axis

-0.10896

corelation cofficient along y -axis

0.150279

age in years	total
18	13
12	7
11	15
14	12

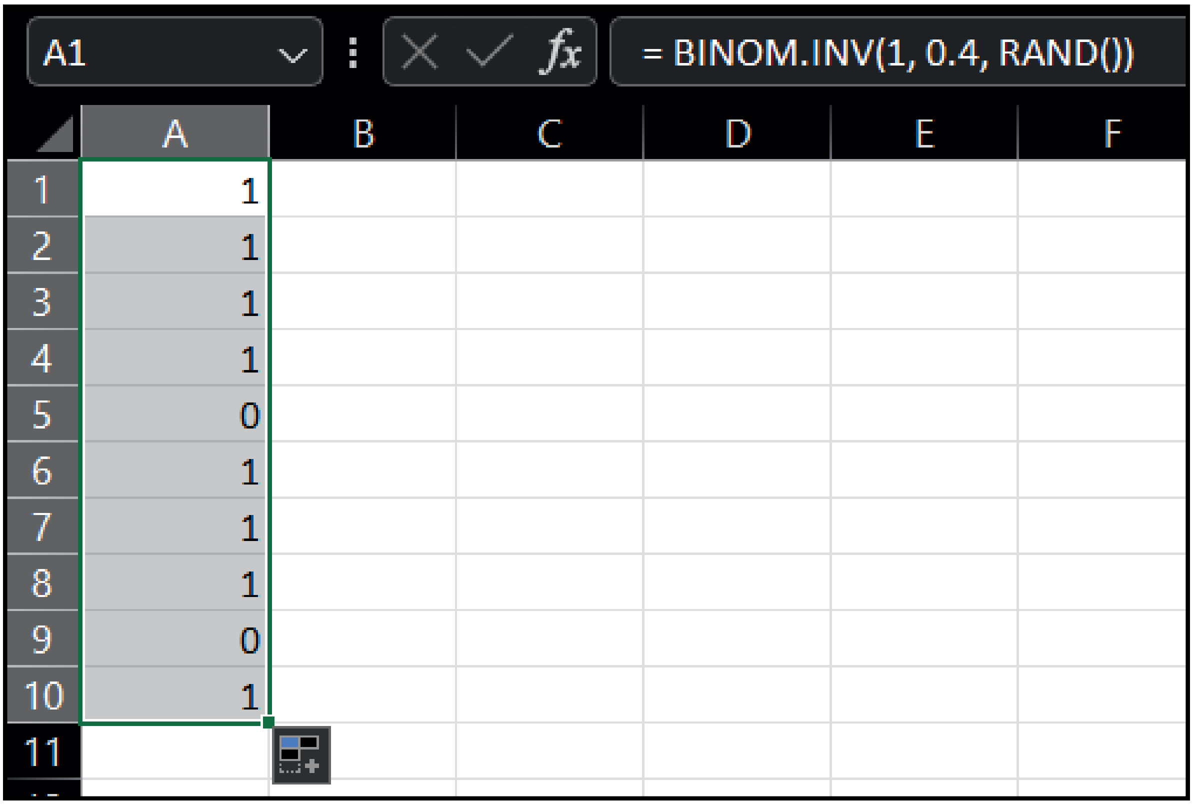
PRACTICAL - 16

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

Implementation in Excel:

= BINOM.INV (1,P, RAND())

- **Random 1 or 0 generated :**


A1 = BINOM.INV(1, 0.4, RAND())

	A	B	C	D	E	F
1	1					
2	1					
3	1					
4	1					
5	0					
6	1					
7	1					
8	1					
9	0					
10	1					
11						

PRACTICAL - 17

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

1. Generating Random Numbers from a Continuous Uniform Distribution

This distribution means any value within a defined range is equally likely.

- Basic Tool: Excel's RAND() function. By itself, =RAND() gives you a random decimal number between 0 and 1 (exclusive of 1, so the range is [0, 1)).
- Generating within a Specific Range [a, b]: To get a random number uniformly distributed between a lower bound a and an upper bound b, you can scale the output of RAND(). The formula is:

= a + (b - a) * RAND()

Generating Random Numbers from a Continuous Uniform Distribution
77.77653786
86.7902862
63.25755433
93.94880212
73.76844481
79.28125132
80.46067926
51.22600428
81.25465432
95.39432653

2. Generating Random Numbers from a Normal Distribution

This distribution follows the classic "bell curve," defined by a mean (average) and a standard deviation (spread). Values near the mean are more probable.

- Tool: We combine RAND() with the NORM.INV function. NORM.INV(probability, mean, standard_dev) finds the value on the normal curve such that the cumulative probability up to that value is equal to probability.
- Generating Random Normal Numbers: By feeding a random probability (generated by RAND()) into NORM.INV, we get a random value sampled from the desired normal distribution. The formula is:

= NORM.INV(RAND(), mean, standard_dev)

Example: Let's generate random heights assuming they follow a normal distribution with a mean (μ) of 175 cm and a standard deviation (σ) of 7 cm.

- mean = 175
- standard_dev = 7
- The formula in Excel would be: = NORM.INV(RAND(), 175, 7)

Generating Random Numbers from a Normal Distribution
171.0679188
167.4402332
165.3302835
172.2057014
169.7931832
180.4641115
185.2195172
179.9859472
161.1703432
192.8604768

PRACTICAL - 18

- **Find the entropy from the given data set.**

What is entropy of random variables?

The entropy of a random variable quantifies the average level of uncertainty or information associated with the variable's potential states or possible outcomes. This measures the expected amount of information needed to describe the state of the variable, considering the distribution of probabilities across all potential states. Given a discrete random variable X, which may be any member x within the set X and is distributed according to p:X→[0,1], the entropy is

$$\mathrm{H}(X) := -\sum_{x\in\mathcal{X}} p(x)\log p(x)$$

	P(x)	log2p(x)	H(x)
HH	1	0	0
HT	0.5	-1	0.5
TH	0.5	-1	0.5
TT	0	0	0

P(x)	H(x)
1	0
0.5	1
0	0