# Index

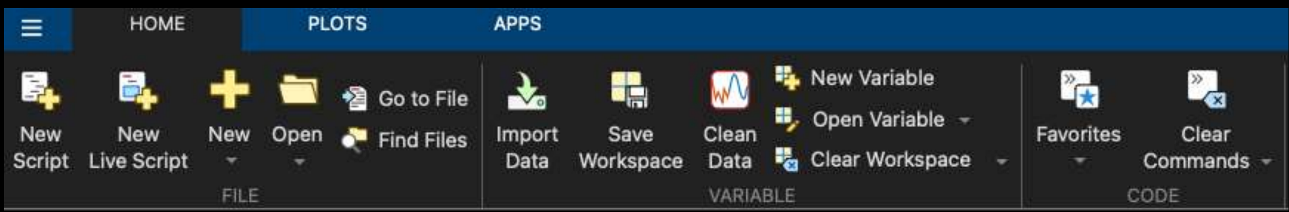
| S.No | Experiment Title | Date | Sign |
|---|---|---|---|
| 1 | Installation of MATLAB and demonstration of simple programming concepts like matrix multiplication (scalar and vector), loop, conditional statements and plotting. | | |
| 2 | Fitting of binomial distributions for given n and p. And Plot the Binomial and Geometric Distribution (PDF and CDF) curve with an example | | |
| 3 | Understanding Hypergeometric Distribution and its approximation to binomial distribution | | |
| 4 | Fitting of binomial distributions after computing mean and variance | | |
| 5 | Program to plot normal distributions and exponential distributions for various parametric values. | | |
| 6 | Fitting of Poisson distributions for given value of lambda. | | |
| 7 | Fitting of Poisson distributions after computing mean. | | |
| 8 | Fitting of normal distribution when parameters are given. | | |

# EXPERIMENT 1

**Objective: Installation of MATLAB and demonstration of simple programming concepts like matrix multiplication (scalar and vector), loop, conditional statements and plotting.**

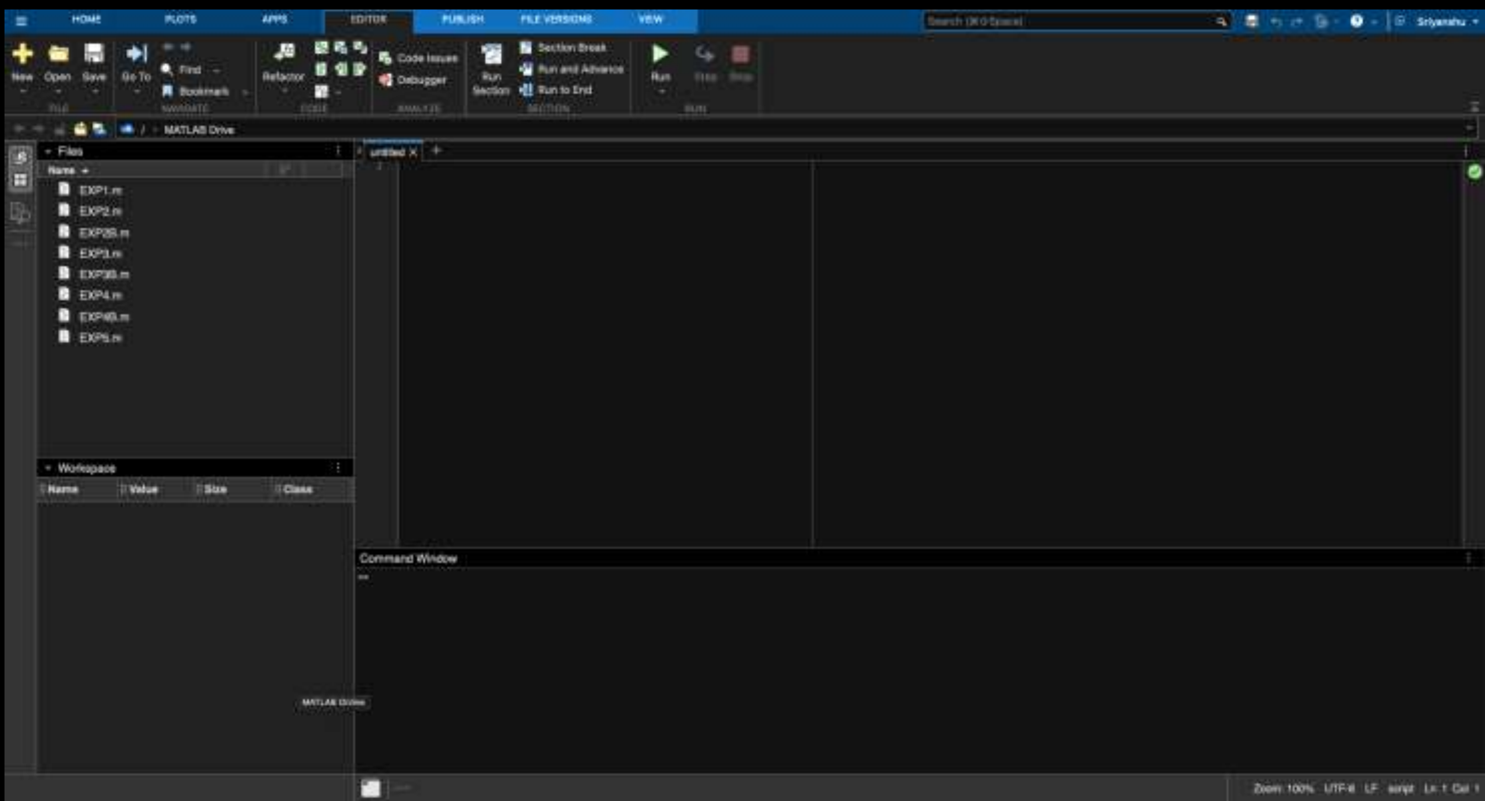
**Formulation & method:-**

1) Open the MATLAB software.




2) Go to command window to write the program or

3) Open a new script & save it



# A. Matrix Multiplication - Scalar and Vector

Matrix initialisation

```
Command Window
>> A = [1 2 3; 4 5 6; 7 8 9]

A =

     1     2     3
     4     5     6
     7     8     9

>> B = [0 1 2; 3 4 5; 6 7 8]

B =

     0     1     2
     3     4     5
     6     7     8
```

Scalar- Multiplication (.*)

```
>> C= A.*B

C =

     0     2     6
    12    20    30
    42    56    72
```

Vector Multiplication (*)

```
>> D = A*B

D =

    24    30    36
    51    66    81
    78   102   126
```

# B. Loop and Conditional Statement

```
EXP1.m
/MATLAB Drive/EXP1.m
for i = 1:20
    if mod(i,2) == 0
        disp(['Number ',num2str(i),' is Even']);
    else
        disp(['Number ',num2str(i),' is Odd']);
    end
end
```

Problem: Display numbers ranging from 1 to 20 are Even or Odd using MATLAB

Command Window

```
>> EXP1
Number 1 is Odd
Number 2 is Even
Number 3 is Odd
Number 4 is Even
Number 5 is Odd
Number 6 is Even
Number 7 is Odd
Number 8 is Even
Number 9 is Odd
Number 10 is Even
Number 11 is Odd
Number 12 is Even
Number 13 is Odd
Number 14 is Even
Number 15 is Odd
Number 16 is Even
Number 17 is Odd
Number 18 is Even
Number 19 is Odd
Number 20 is Even
>>
```

## C. Plotting of Parabola

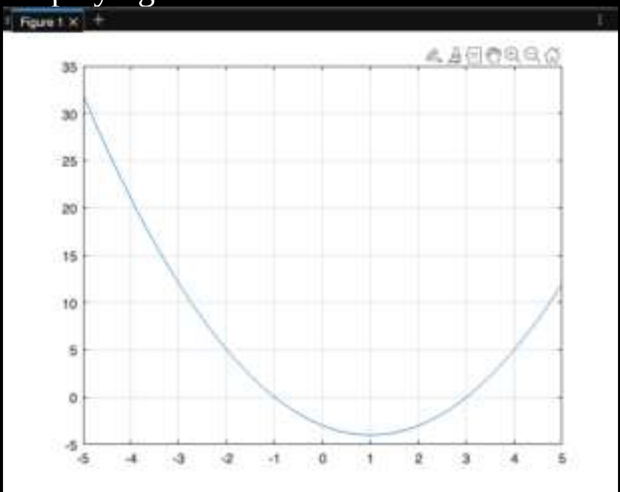
Feeding Data

EXP1.m

/MATLAB Drive/EXP1.m

```
X = -5:0.1:5;
Y = X.*X -2.*X -3;
plot(X,Y);
```

Displaying the Plotted Data



## Learning Outcomes:

This experiment helped to understand the functioning of MATLAB and its various components. MATLAB was used to

- Perform Matrix Multiplication in Scalar and Vector form
- Use Loop and Conditional statement to find even and odd number
- Plotting of a parabola graph using plot function.

# EXPERIMENT 2

**Objective: Fitting of binomial distributions for given n and p. And Plot the Binomial and Geometric Distribution (PDF and CDF) curve with an example using MATLAB**

## A. Fitting of Binomial Distribution for given n and p

Problem: Each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. i) Find the probability that in the next 18 samples, exactly 2 contain the pollutant.

ii) Determine the probability that at least four samples contain the pollutant.

Using MATLAB: binopdf(x,n,p) and binocdf(x,n,p)

EXP2.m ×

/MATLAB Drive/EXP2.m

```
x=2; % Num of pollutant
n=18; % Num of samples
p=0.1; % P(sample contains pollutant)
Px1 = binopdf(x,n,p); % P(exactly 2 samples contain pollutant)
plot(0:18,binopdf(0:18,n,p));
title("BINOMIAL PROBABILITY DISTRIBUTION");
xlabel('Number of Pollutants --->');
ylabel('Probability --->')
Px2 = binocdf(3,n,p);
r = 1-Px2; % P(at least 4 samples contain pollutant)
```

Output:

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| Px1 | 0.2835 | 1x1 | double |
| Px2 | 0.9018 | 1x1 | double |
| n | 18 | 1x1 | double |
| p | 0.1000 | 1x1 | double |
| r | 0.0982 | 1x1 | double |
| x | 2 | 1x1 | double |

x= number of pollutant = 2;

n= number of samples = 18;

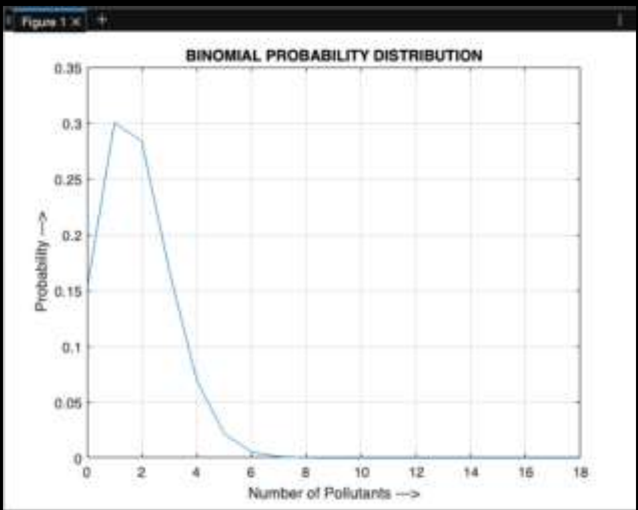
p=probability that sample contains pollutant = 0.1;

***SOLUTION i)** Px1= 0.2835 = Binomial probability distribution = Probability that exactly 2 samples contain the pollutant

Px2 = 0.9018 = Binomial Cumulative distribution from 0 to 3

***SOLUTION ii)** r= 0.0982 = Probability that at least 4 samples contain pollutant

Plotting Binomial Distribution curve




## B. Geometric Distribution

Problem: The probability that a wafer contains a large particle of contamination is 0.01. If it is assumed that the wafers are independent, what is the probability that exactly 125 wafers need to be analyzed before a large particle is detected?

Using MATLAB: geopdf(x,p)

EXP2B.m

/MATLAB Drive/EXP2B.m

```
x = 125; % samples analysed until a large particle is detected
p = 0.01; % P(wafer conatins a large particle of contamination
Px = geopdf(x,p);
stem(0:25:200,geopdf(0:25:200,geopdf(0:25:200,p)));
title('GEOMETRIC PROBABILITY DISTRIBUTION');
xlabel('Samples analysed --->');
ylabel('Probability --->');
```

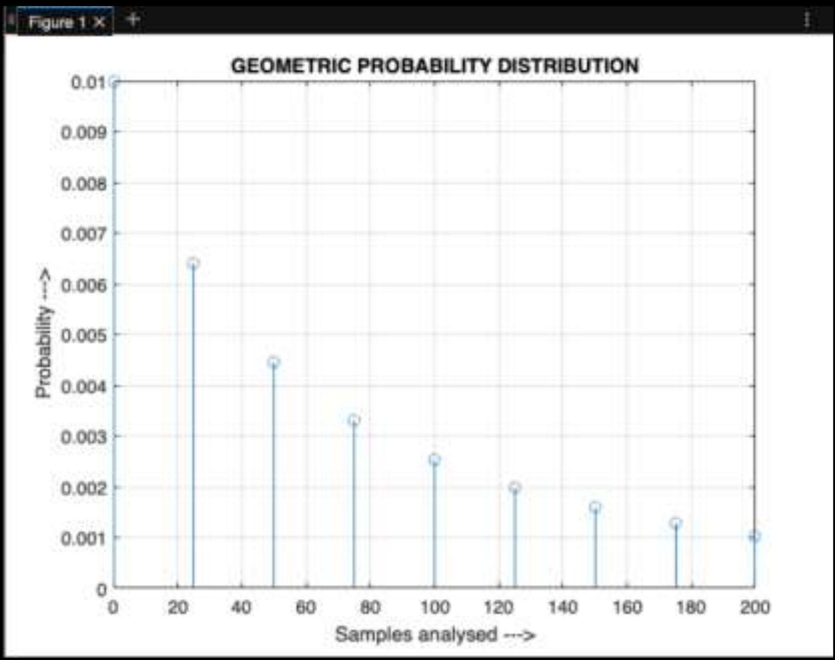
Output:

| Name | Value | Size | Class |
|---|---|---|---|
| Px | 0.0028 | 1x1 | double |
| p | 0.0100 | 1x1 | double |
| x | 125 | 1x1 | double |

x= samples analysed until a large particle is detected = 125

p= probability that water contains a large particle of contamination = 0.01

***SOLUTION** Px= 0.0028 = geometric probability distribution = probability that exactly 125 wafers need to be analysed before a large particle is detected




## Learning Outcomes:

This experiment helped to understand

- The concept and formulation of Binomial distribution and using binopdf(x,n,p) and binocdf(x,n,p) for probability distribution and cumulative distribution respectively along with the plotting of the given data.
- The concept and formulation of Geometric distribution and using geopdf(x,p) to solve a given problem and then plotting its graph.

# EXPERIMENT 3

## Objective: Understanding Hypergeometric Distribution and its approximation to binomial distribution using MATLAB

## A. Hypergeometric Distribution

Problem: A batch of parts contains 100 parts from a local supplier of tubing and 200 parts from a supplier of tubing in the next state. If four parts are selected randomly and without replacement, i) what is the probability that they are all from the same supplier?

ii) what is the probability that two or more parts in the sample are from the local supplier?

Using MATLAB: hygepdf(x,m,k,n) and hygecdf(x,m,k,n)

EXP3.m × EXP3B.m × untitled * ×

/MATLAB Drive/EXP3.m

```
x = 0:4; % number of parts from local supplier
m = 300; % Total number of parts
k = 100; % Total parts from local supplier
n = 4; % number of parts selected
local = hygepdf(4,m,k,n); % all parts from local supplier
r=hygecdf(4,m,k,n)-hygecdf(1,m,k,n); %two or more parts from local supplier
Px = hygepdf(x,m,k,n); %P(parts from local supplier)
plot(x,Px,'r');
title("Hypergeometric Probability Distribution");
```

Output:

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| Px | [0.1955,0.3970,0.2978,0.0978,0.0119] | 1x5 | double |
| k | 100 | 1x1 | double |
| local | 0.0119 | 1x1 | double |
| m | 300 | 1x1 | double |
| n | 4 | 1x1 | double |
| r | 0.4074 | 1x1 | double |
| x | [0,1,2,3,4] | 1x5 | double |

x= number of parts from local supplier

m= total number of parts = 300

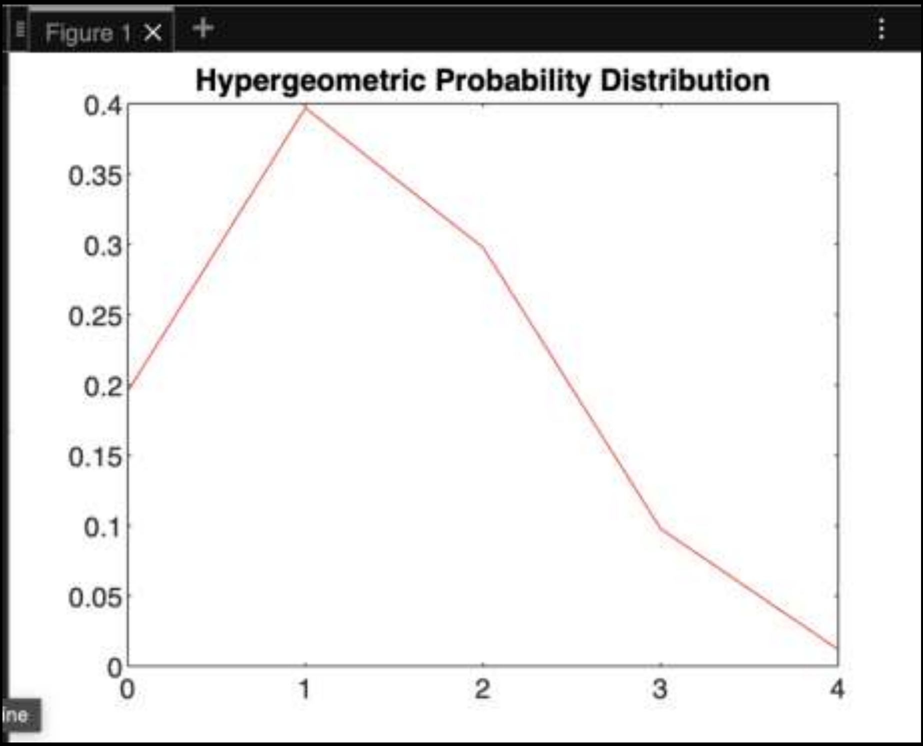
k= total parts from local supplier = 100

n= number of parts selected = 4

***SOLUTION** i) local= Probability that all parts from local supplier = 0.0119

***SOLUTION** ii) r= probability that two or more parts from local supplier = 0.4074

Px= hypergeometric probability distribution for parts from 0 to 4




## B. Hypergeometric Distribution and approximation to binomial distribution

Formulation and Method:

The curves of the Hypergeometric distribution and the Binomial distribution are similar because under certain conditions, the Hypergeometric distribution can be approximated by the Binomial distribution.

The conditions for this approximation are:

1. The population size is much larger than the number of items drawn in each trial (m>>n)
2. The number of successes in the population is much smaller than the population size
3. (k<<m)

x= samples of one type; m= total population; n=items drawn in each trial; k=total number of required samples

Px and Px3= hypergeometric distribution;

Px2 and Px4= binomial distribution;

EXP3_B.m

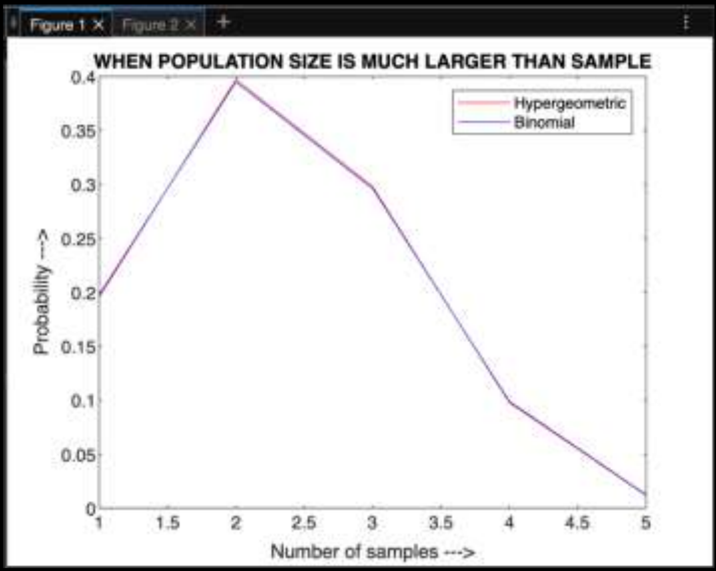
/MATLAB Drive/EXP3_B.m

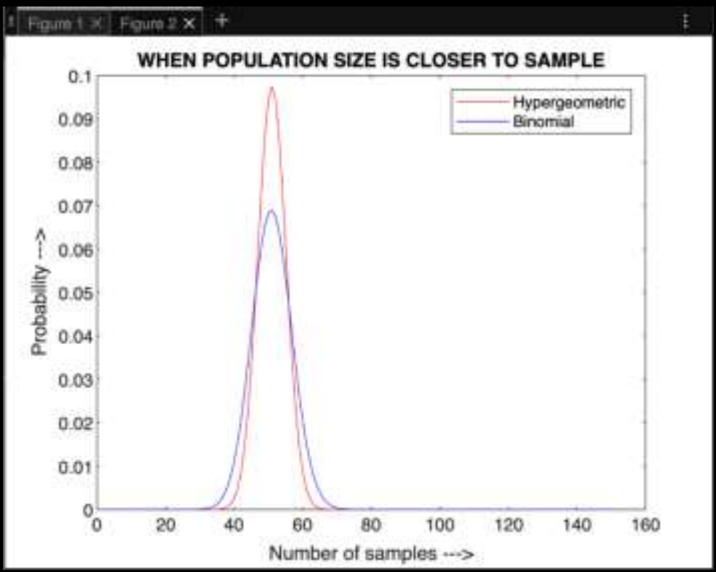
```
% When population size is much larger than number of samples drawn n>>m
x = 0:4;
m= 300;
k= 100;
n= 4;
Px= hygepdf(x,m,k,n);
p= k/m;
Px2= binopdf(x,n,p);
plot(Px,'r');
hold
plot(Px2,'b');
title('WHEN POPULATION SIZE IS MUCH LARGER THAN SAMPLE');
legend('Hypergeometric','Binomial');
xlabel('Number of samples --->');
ylabel('Probability --->');

% When Population size is near to the number of samples drawn
figure;
x2= 0:150;
m2=300;
k2=100;
n2=150;
Px3= hygepdf(x2,m2,k2,n2);
p2= k2/m2;
Px4= binopdf(x2,n2,p2);
plot(Px3,'r');
hold
plot(Px4,'b');
title('WHEN POPULATION SIZE IS CLOSER TO SAMPLE');
legend('Hypergeometric','Binomial');
xlabel('Number of samples --->');
ylabel('Probability --->');
```

OUTPUT:

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| Px | [0.1955,0.3970,0.2978,0.0978,0.0119] | 1x5 | double |
| Px2 | [0.1975,0.3951,0.2963,0.0988,0.0123] | 1x5 | double |
| Px3 | 1x151 double | 1x151 | double |
| Px4 | 1x151 double | 1x151 | double |
| k | 100 | 1x1 | double |
| k2 | 100 | 1x1 | double |
| m | 300 | 1x1 | double |
| m2 | 300 | 1x1 | double |
| n | 4 | 1x1 | double |
| n2 | 150 | 1x1 | double |
| p | 0.3333 | 1x1 | double |
| p2 | 0.3333 | 1x1 | double |
| x | [0,1,2,3,4] | 1x5 | double |
| x2 | 1x151 double | 1x151 | double |




## LEARNING OUTCOMES:

This experiment helped to understand

- The concept of hypergeometric distribution and solving problems in MATLAB with hygepdf(x,m,k,n) and hygecdf(x,m,k,n)
- How the curves of hypergeometric probability distribution and binomial probability distribution vary in different provided conditions based on the previous example. The curves vary when

i) n<<m i.e The population size is much larger than the number of items drawn in each trial

ii) k<<m i.e The number of successes in the population is much smaller than the population size

# EXPERIMENT 4

## Objective: Fitting of binomial distributions after computing mean and variance using MATLAB

Problem: Find mean, P(X) and expected frequency

| X | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Frequency | 36 | 40 | 22 | 2 |

Using MATLAB: binopdf(x,n.p)

EXP7.m

/MATLAB Drive/EXP7.m

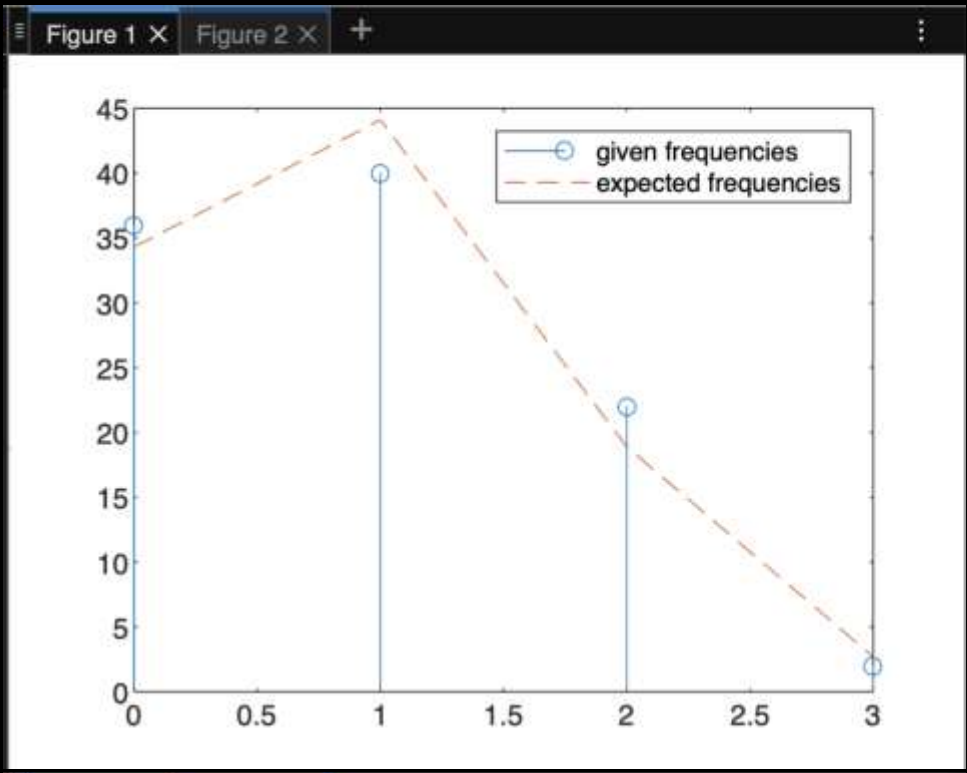
```
x = 0:3; % probable outcomes
freq = [36,40,22,2]; % frequencies of each outcome
n = 3; % number of trials
mean = x*freq'/sum(freq);
p=mean/n;
var=n*p*(1-p);

Px = binopdf(x,n,p); %P(X=x)
Exp = Px .* sum(freq); % expected frequencies

stem(x,freq);
hold;
plot(x,Exp,'--');
legend('given frequencies','expected frequencies');

figure();
plot(x,Px);
title("BINOMIAL PROBABILITY DISTRIBUTION");
```

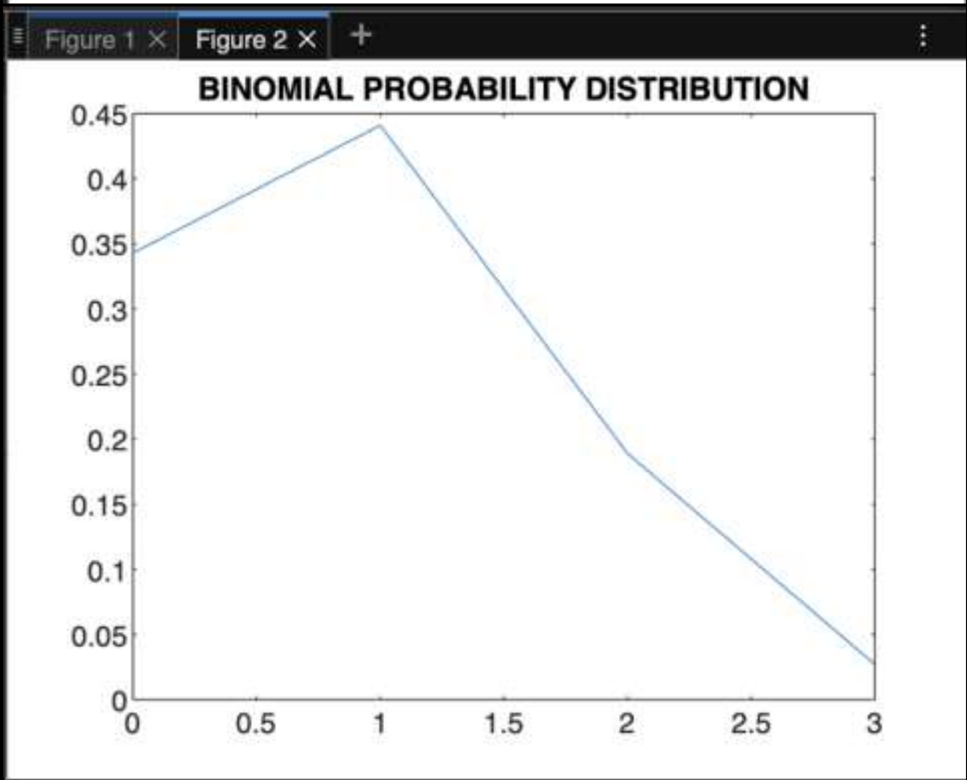
Output:

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| Exp | [34.3000,44.1000,18.9000,2.7000] | 1x4 | double |
| Px | [0.3430,0.4410,0.1890,0.0270] | 1x4 | double |
| freq | [36,40,22,2] | 1x4 | double |
| mean | 0.9000 | 1x1 | double |
| n | 3 | 1x1 | double |
| p | 0.3000 | 1x1 | double |
| var | 0.6300 | 1x1 | double |
| x | [0,1,2,3] | 1x4 | double |

| X | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Frequency | 36 | 40 | 22 | 2 |
| P(X) | 0.3430 | 0.4410 | 0.1890 | 0.0270 |
| Expected Frequency | 34.3000 | 44.1000 | 18.9000 | 2.7000 |




## Learning Outcomes:

In this experiment we first found out the missing values like mean (mu) and variance using the given data. And then the BINOMIAL probability distribution was calculated using binopdf(x,n,p). Finally, the expected frequencies using this method was obtained.

mean= 0.9000

Variance = 0.6300

The comparisons between given and expected frequencies is made by plotting their respective graphs.

# EXPERIMENT 5

**Objective: Program to plot normal distributions and exponential distributions for various parametric values.**

## A. Normal Distribution Curve with different values of sigma

Using MATLAB:

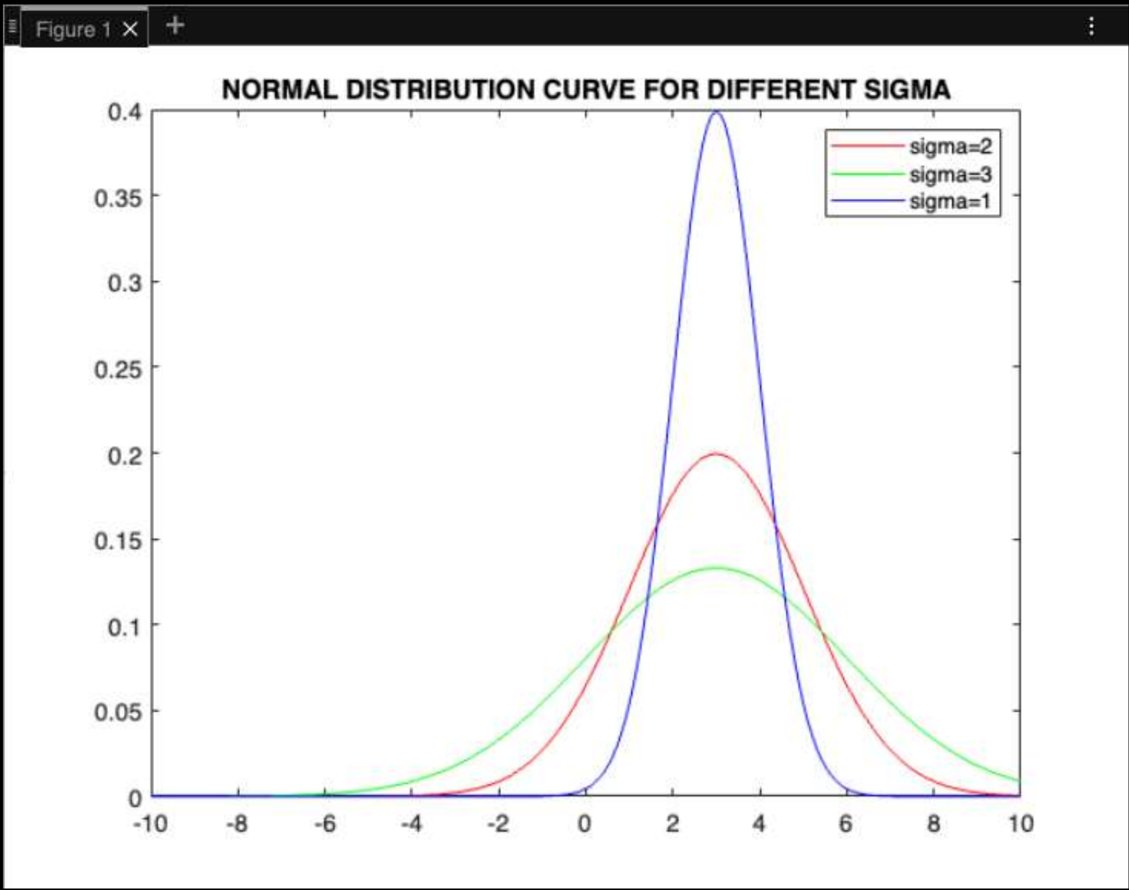
EXP4.m

/MATLAB Drive/EXP4.m

```
x = -10:0.1:10;
mu = 3;
sigma1 = 2;
sigma2 = 3;
sigma3 = 1;

y1 = normpdf(x,mu,sigma1);
y2 = normpdf(x,mu,sigma2);
y3 = normpdf(x,mu,sigma3);

plot(x,y1,'r');
hold on;
plot(x,y2,'g');
plot(x,y3,'b');
title('NORMAL DISTRIBUTION CURVE FOR DIFFERENT SIGMA');
legend('sigma=2','sigma=3','sigma=1');
```

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| mu | 3 | 1x1 | double |
| sigma1 | 2 | 1x1 | double |
| sigma2 | 3 | 1x1 | double |
| sigma3 | 1 | 1x1 | double |
| x | 1x201 double | 1x201 | double |
| y1 | 1x201 double | 1x201 | double |
| y2 | 1x201 double | 1x201 | double |
| y3 | 1x201 double | 1x201 | double |




**B. Exponential Distribution Curve with different values of lambda**

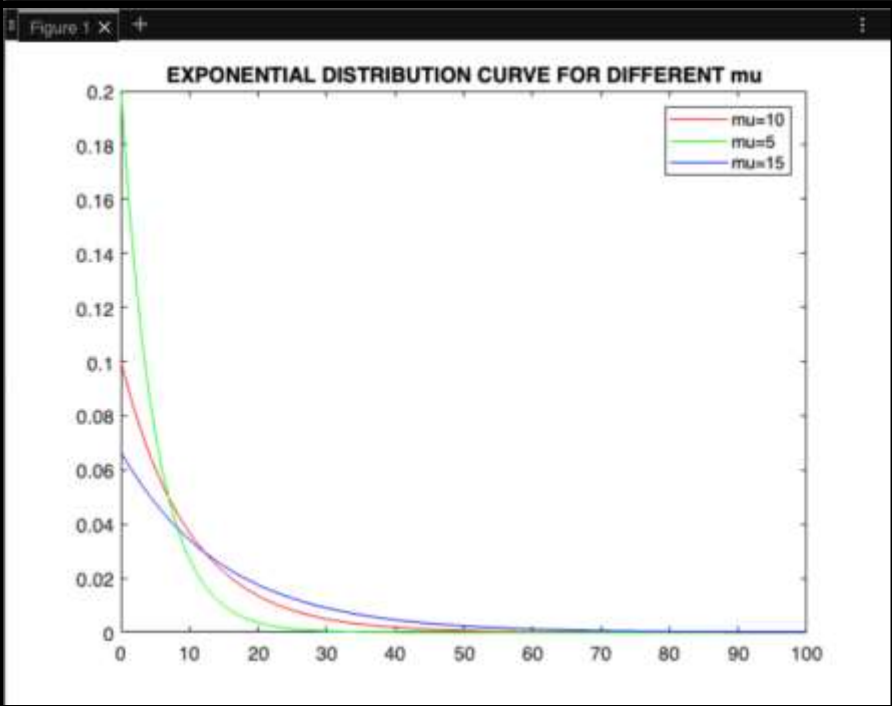
EXP4B.m

/MATLAB Drive/EXP4B.m

```
x = 0:100;
mu1 = 10;
mu2 = 5;
mu3 = 15;

y1 = exppdf(x,mu1);
y2 = exppdf(x,mu2);
y3 = exppdf(x,mu3);

plot(x,y1,'r');
hold on;
plot(x,y2,'g');
plot(x,y3,'b');
title('EXPONENTIAL DISTRIBUTION CURVE FOR DIFFERENT mu');
legend('mu=10','mu=5','mu=15');
```

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| mu1 | 10 | 1x1 | double |
| mu2 | 5 | 1x1 | double |
| mu3 | 15 | 1x1 | double |
| x | 1x101 double | 1x101 | double |
| y1 | 1x101 double | 1x101 | double |
| y2 | 1x101 double | 1x101 | double |
| y3 | 1x101 double | 1x101 | double |




## Learning Outcomes:

This experiment helped to understand

- How the values of normal distribution can vary for different values of sigma by plotting and comparing the bell curve for the given values of sigma
- How the values of exponential distribution can vary for different values of mean (mu) by plotting and comparing the exponential distribution curve of the given values of mu.

# EXPERIMENT 6

## Objective: Fitting of Poisson distributions for given value of lambda using MATLAB

Problem: On an average, 20 red blood cells are found in a fixed volume of blood for a normal person. Determine the probability that the blood sample of a normal person will contain less than 5 red blood cells.

Using MATLAB: poisscdf(x,mu)

EXP6.m *

/MATLAB Drive/EXP6.m

```
lambda = 20; % number found in fixed volume
x = 5; % number under study
Px = poisscdf(x,lambda); % P(sample is less than x)

plot(0:20,poisspdf(0:20,lambda));
title("POISSON DISTRIBUTION CURVE");
```

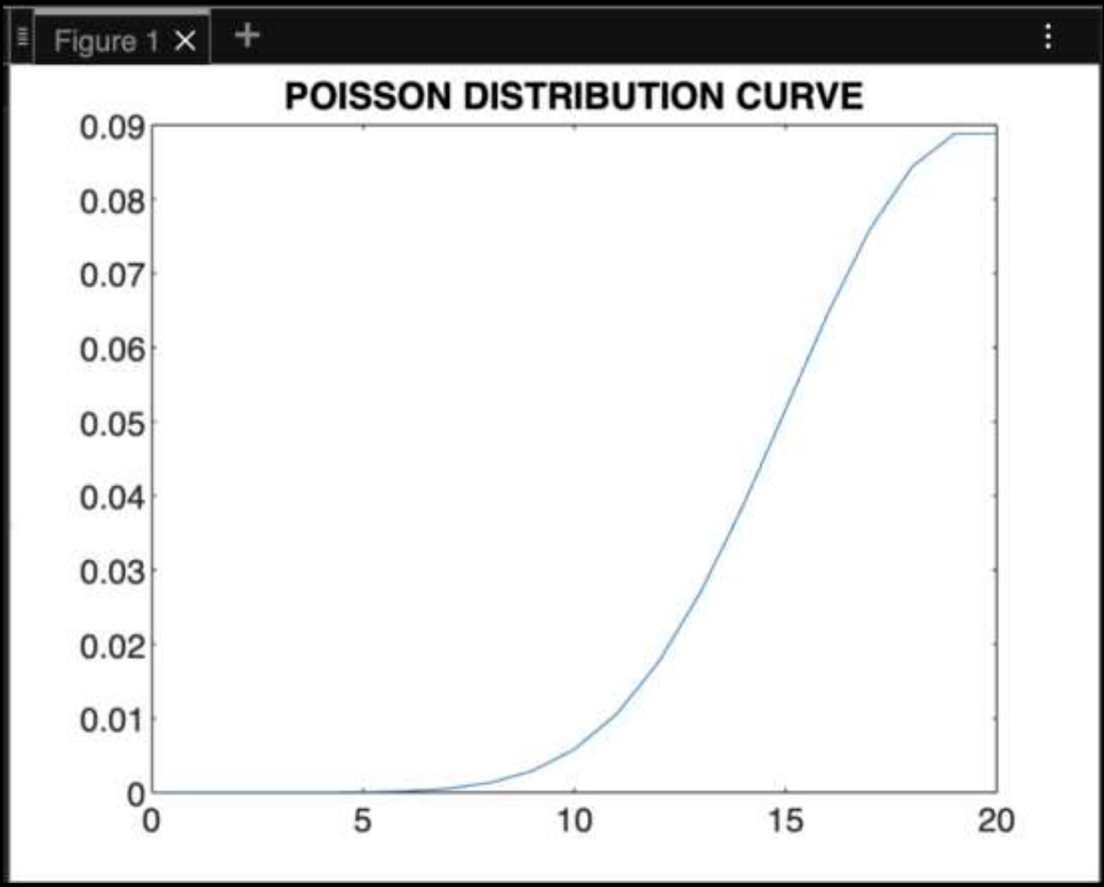
Output:

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| Px | 7.1909e-05 | 1x1 | double |
| lambda | 20 | 1x1 | double |
| x | 5 | 1x1 | double |

lambda (or mu)= rate parameter (number found in fixed volume) = 20; x = number that needs to be studied = 5

***SOLUTION** Px= probability that sample is less than x = $7.19 \times 10^{-5}$




## Learning Outcomes:

This experiment helped to understand the concept of poisson distribution and how MATLAB can be used to solve the given problem by using poisscdf(x, lambda) or poisspdf(x, lambda). The curve of poisson distribution is also studied.

# EXPERIMENT 7

## Objective: Fitting of Poisson distributions after computing mean using MATLAB

Problem: Find mean, P(X) and expected frequency

| X | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Frequency | 36 | 40 | 22 | 2 |

Using MATLAB:

EXP7.m

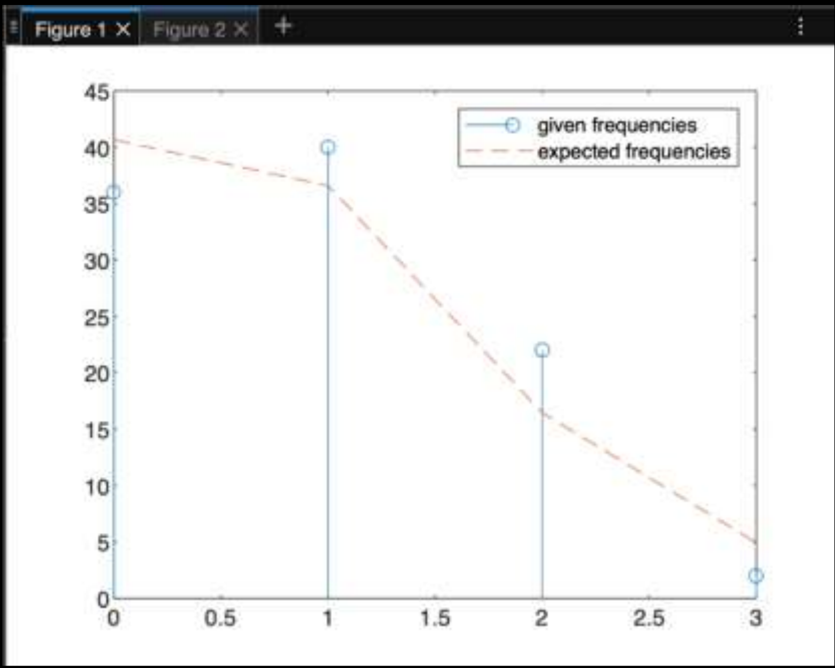
/MATLAB Drive/EXP7.m

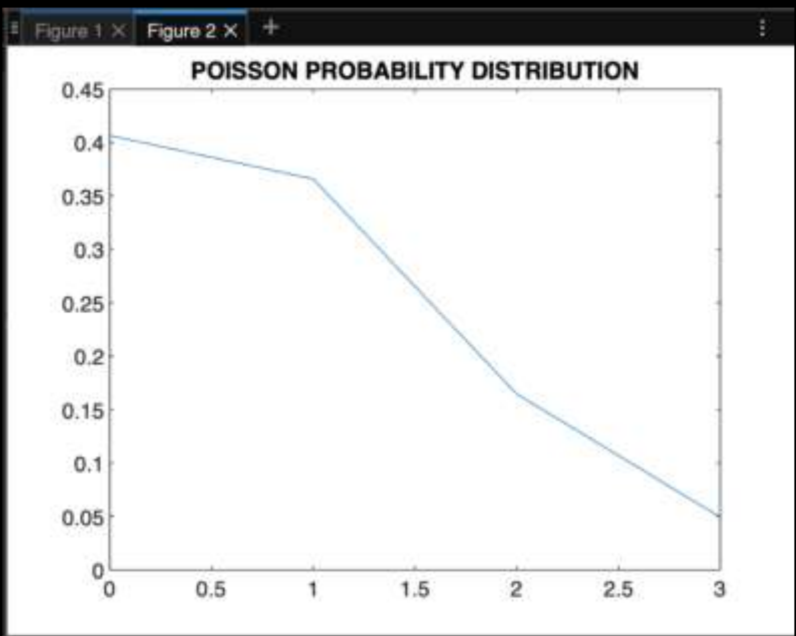
```
x = 0:3; % probable outcomes
freq = [36,40,22,2]; % frequencies of each outcome
n = 3; % number of trials
mean = x*freq'/sum(freq);


Px = poisspdf(x,mean); %P(X=x)
Exp = Px .* sum(freq); % expected frequencies

stem(x,freq);
hold;
plot(x,Exp,'--');
legend('given frequencies','expected frequencies');

figure();
plot(x,Px);
title("POISSON PROBABILITY DISTRIBUTION");

```

Output:

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| Exp | [40.6570,36.5913,16.4661,4.9398] | 1x4 | double |
| Px | [0.4066,0.3659,0.1647,0.0494] | 1x4 | double |
| freq | [36,40,22,2] | 1x4 | double |
| mean | 0.9000 | 1x1 | double |
| n | 3 | 1x1 | double |
| x | [0,1,2,3] | 1x4 | double |



| X | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Frequency | 36 | 40 | 22 | 2 |
| P(X) | 0.4066 | 0.3659 | 0.1647 | 0.0494 |
| Expected Frequency | 40.6570 | 36.5913 | 16.4661 | 4.9398 |

## Learning Outcomes:

In this experiment we first found out the missing values like mean (mu) and variance using the given data. And then the POISSON probability distribution was calculated using poisspdf(x, mean). Finally, the expected frequencies using this method was obtained.

mean= 0.9000

The comparisons between given and expected frequencies is made by plotting their respective graphs.

# EXPERIMENT 8

## Objective: Fitting of normal distribution when parameters are given.

Problem: The amount of pollutant X released by an industry should lie between 30 and 35. Assume that X is normally distributed with mean X = 33 and S.D. σ = 3. The industry gets a profit of Rs. 100 when $30 < X < 35$; Rs. 50 when $25 < X \leq 30$ or $35 \leq X < 40$ and incurs a fine of Rs. 60 otherwise. Determine the expected profit for the industry.

Using MATLAB:

EXP4.m

/MATLAB Drive/EXP4.m

```
x1 = [30 35]; % pollutant released
x2 = [25 30];
x3 = [35 40];
x4 = [20 25];
x5 = [40 45];

mu = 33; % mean
sigma = 3; % standard deviation

Px1 = normspec(x1,mu,sigma); % probability of shaded area
Px2 = normspec(x2,mu,sigma);
Px3 = normspec(x3,mu,sigma);
Px4 = normspec(x4,mu,sigma);
Px5 = normspec(x5,mu,sigma);

Profit= 100*Px1 + 50*(Px2+Px3) - 60*(Px4+Px5); % Answer
```

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| Profit | 77.9439 | 1x1 | double |
| Px1 | 0.5889 | 1x1 | double |
| Px2 | 0.1548 | 1x1 | double |
| Px3 | 0.2427 | 1x1 | double |
| Px4 | 0.0038 | 1x1 | double |
| Px5 | 0.0098 | 1x1 | double |
| mu | 33 | 1x1 | double |
| sigma | 3 | 1x1 | double |
| x1 | [30,35] | 1x2 | double |
| x2 | [25,30] | 1x2 | double |
| x3 | [35,40] | 1x2 | double |
| x4 | [20,25] | 1x2 | double |
| x5 | [40,45] | 1x2 | double |

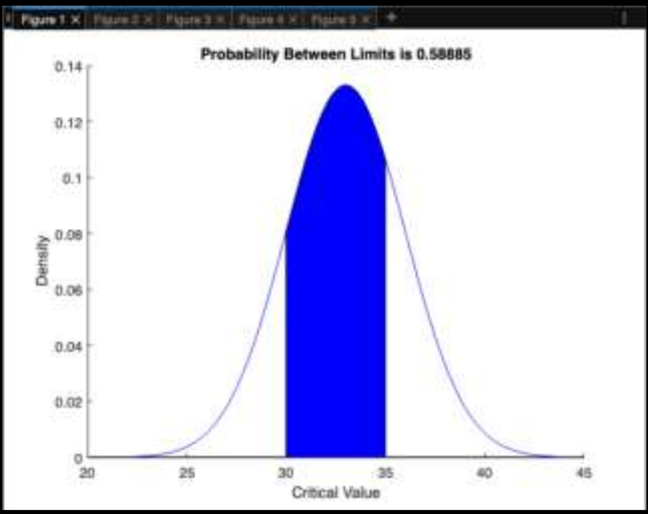


Figure 1
Probability Between Limits is 0.58885
Density
Critical Value

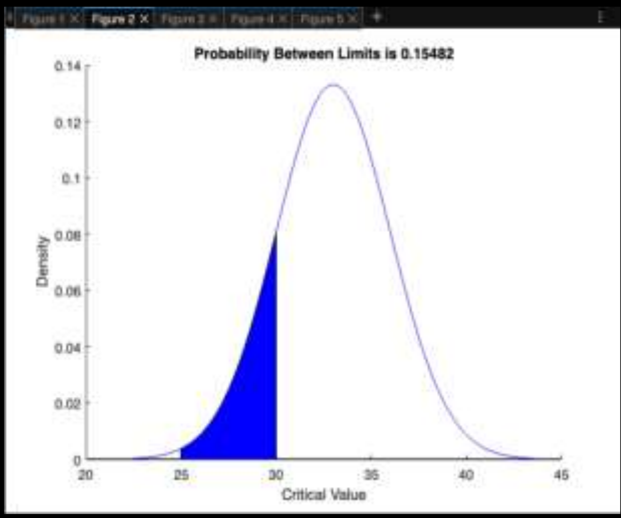


Figure 2
Probability Between Limits is 0.15482
Density
Critical Value

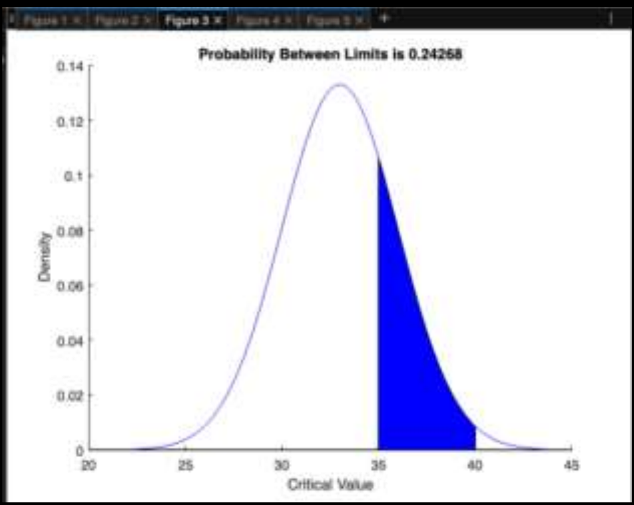


Figure 3
Probability Between Limits is 0.24268
Density
Critical Value

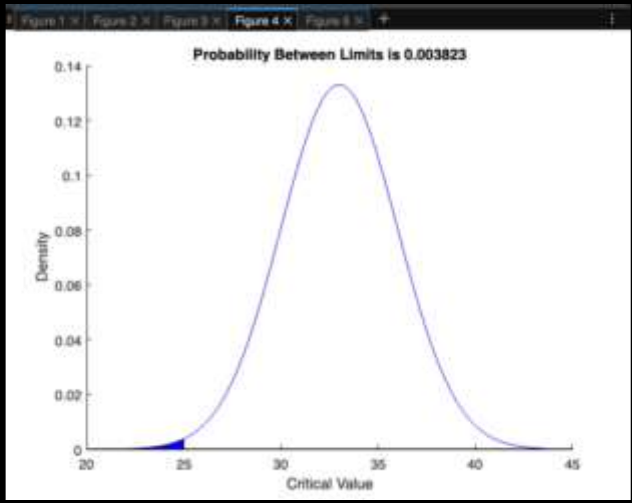


Figure 4
Probability Between Limits is 0.003823
Density
Critical Value

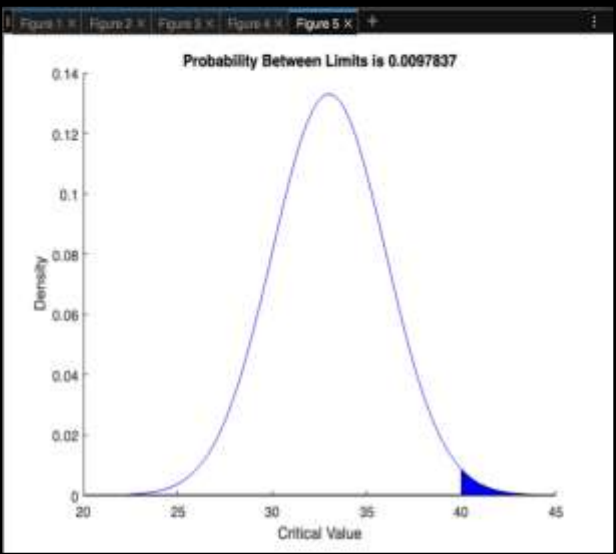


Figure 5
Probability Between Limits is 0.0097837
Density
Critical Value

## Learning Outcomes:

This experiment helps to understand the normal distribution and how the normal distribution curve can be used to find the probability or expected values by the normal distribution table.

The problem deals with different ranges of pollutant and for each range the probability was found which was then used to calculate the total profit or loss to the company.

Here,

x= different ranges of pollutant released

Mu = mean =33; sigma = standard deviation = 3;

Px = probability of shaded area using normal distribution

**Profit = total profit = 77.94, solution**

# EXPERIMENT 9

**Objective:** Fitting of linear regression line through given data set and testing of goodness of fit using mean error.

Problem: Fit the linear regression line for the following data

| X | | | | | | |
|---|---|---|---|---|---|---|
| Y | 1 | 22 | 50 | 62 | 97 | 110 |

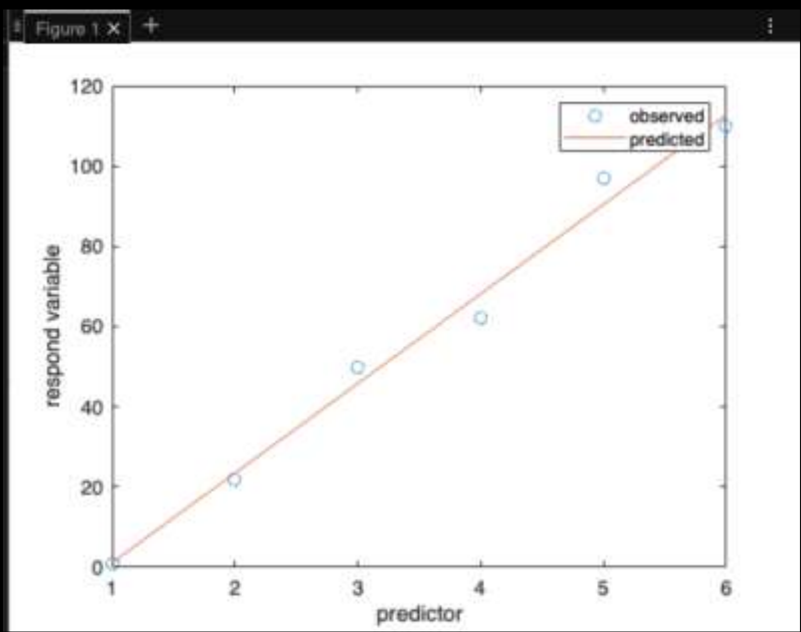
Using MATLAB:

exp9.m

/MATLAB Drive/exp9.m

```
X=1:6;
Y=[1 22 50 62 97 110];
plot(X,Y,'o');
hold on;
a=polyfit(X,Y,1);
Y_H=polyval(a,X);
plot(X,Y_H);
legend('observed','predicted');
xlabel('predictor');
ylabel('respond variable');
E=Y_H - Y;
MSE=sqrt(mean(E.^2));
% Best fitting line for the data is the line y=ax+b where a and b  are
% coefficients of a matrix.
display(MSE);
```

**Output:-**



Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| E | [0.1429,1.4857,-4.1714,6.1714,-6.4857,2.8571] | 1x6 | double |
| MSE | 4.2415 | 1x1 | double |
| X | [1,2,3,4,5,6] | 1x6 | double |
| Y | [1,22,50,62,97,110] | 1x6 | double |
| Y_H | [1.1429,23.4857,45.8286,68.1714,90.5143,112.8571] | 1x6 | double |
| a | [22.3429,-21.2000] | 1x2 | double |

X and Y are given data.

Y_H is the predicted data

a stores coefficients of linear polynomial

E stores difference between Y and Y_H

MSE stores Mean Squared Error between observed and predicted response

Command Window

```
>> exp9

MSE =

    4.2415

>>
```

BEST FITTING LINE FOR THE GIVEN DATA(Y on X) IS: **Y=22.3429X - 21.2000**

## Learning Outcomes:

This experiment helps to understand the concept of linear regression and the use of functions in MATLAB like ‘polyval’ and ‘polyfit’. The predicted response was calculated and plotted with the help of given data and then finally mean squared error(MSE) was found between the observed and predicted values.

# EXPERIMENT 10

**Objective:** Fitting of Multiple Linear Regression (MLR) curve through given data set and testing of goodness of fit using mean error.

Problem: Fit the multi linear regression line for the following data

| | | | | |
|---|---|---|---|---|
| X1 | | | | |
| X2 | 4 | 5 | 8 | 2 |
| Y | 1 | 6 | 8 | 12 |

Using MATLAB

/MATLAB Drive/exp9.m

```
X=[1 1 4;1 2 5;1 3 8; 1 4 2];
Y=[1;6;8;12];
X_t=transpose(X);
a_h=(inv(X_t * X)*X_t)*Y;
Y_h=X*a_h;
MSE = sqrt(mean((Y_h-Y).^2));
% Best fitting line for the data is the line y=a0 + a1x1 + a2x2 where
% a0,a1,a2 are coefficients of a_h matrix
display(MSE);
```

**Output:-**

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| E | [0.1429,1.4857,-4.1714,6.1714,-6.4857,2.8571] | 1x6 | double |
| MSE | 4.2415 | 1x1 | double |
| X | [1,2,3,4,5,6] | 1x6 | double |
| Y | [1,22,50,62,97,110] | 1x6 | double |
| Y_H | [1.1429,23.4857,45.8286,68.1714,90.5143,112.8571] | 1x6 | double |
| a | [22.3429,-21.2000] | 1x2 | double |

Workspace

| Name | Value | Size | Class |
|---|---|---|---|
| MSE | 0.6011 | 1x1 | double |
| X | *4x3 double* | 4x3 | double |
| X_t | *3x4 double* | 3x4 | double |
| Y | [1;6;8;12] | 4x1 | double |
| Y_h | [1.5656;4.9945;8.3142;12.1257] | 4x1 | double |
| a_h | [-1.6995;3.4836;-0.0546] | 3x1 | double |

X is matrix containing predictor variables

X_t is transpose of X

Y contains observed response variable values

a_h contains estimated coefficients of linear regression model; a_h = $(X^T * X)^{-1} * X^T * Y$

Y_h contains the predicted response variable values; Y_h = X * a_h

MSE is the mean squared error between the observed response values and predicted response values

BEST FITTING LINE FOR THE GIVEN DATA(Y on X1 and X2) IS: **y= -1.6995 + 3.4836 x1 - 0.0546 x2**

Command Window

```
>> exp10

MSE =

    0.6011

>>
```

## Learning Outcomes:

This experiment helps to understand the concept of multilinear regression and practice estimating regression coefficients for multiple predictors. This experiment demonstrated how the data can be converted to matrix and then using matrix operations the required predicted values were found. The Mean Squared Error (MSE) between the observed and predicted response values was also calculated.