Experiment-4

PROBABILITY DISTRIBUTION.

Experiment-A :— Sampeling from Distribution:

Binomial distribution.

Normal Distribution.

Binomial Distribution :—

Types of binomial distribution functions:

1. dbinom( ):-

It is used to find binomial distribution with carrying parameters like object [x, y], size of the object (20), probability if values respecting to object (1/6).

2. P binom( ):-

If an distribution with respect to probability.

3. R binom( ):-

Is used to take a random sample from a binomial distribution.

4. q binom( ):—

Is used to find quantities for a binomial distribution

d binom (x, size = 20, prob = 1/6).

P(x=3)

dbinom (x = 0:3, size = 20, prob = 1/6)

P(x<=3)

Sum (dbinom (x = 0:3, size = 20, prob = 1/6))

Pbinom (q=3, size = 20, prob = 1/6, lowertail = T)

Normal distribution :—

X - NN (μ = 75, σ^2 = 5^2)

P (X<=70)

pnorm (q=70, mean=75, sd=5, lower.tail=T)

qnorm (p=0.25, mean=75, sd=5, lower.tail=T)

rnorm (n=40, mean=75, sd=5)

dnorm (p=0.25, mean=75, sd=5)

Output:-

[1] 0.0000029973439

[1] 92.596 63.486 76.680

96% confident interval

-Inf 8.224

Experiment - 4b.

tTest, ztest, chi square test

tTest:

We Use t.test() to perform any data with carry. Parameters. [object(x), mu(μ) = (7), alternative(less), confidence level(0.08)]

Eg: t.test

data

t.test(Lunglap, mu = 8, alternative = "less", conf.level = 0.95)

t.test (data, mu=8, alter="two-sided", conf=0.95)

.attributes (TEST)

TEST & conf.int. → output

data : c(3:10)

t = −1.7321, df = 7, p-value = 0.0634

alternative hypothesis : true mean is less than 8

76 % percent confidence interval

-Inf 8.271898

sample estimates :

mean of x : 6.5.

Ztest :

We use prop.test function for calculating frequency

```
mydat <- read.csv("      .csv")
names(mydat)
tabel1 <- table(mydat$CN1, mydata$CN2)
tabel1
prop.test(tabel1, correct=FALSE)
prop.test(X=c(30,40), n=c(34+1350,
          40+217), correct=FALSE)
tabel2 = table(mydat$CN2, mydat$CN1)
table2
prop.test(table2, correct=FALSE).
```

(4)

chisquare-test :-

determine if two categorical variables have a significant correlation between them both must variables have same population.

```
chisq.test(df)
```

(2)
output :- data : df$c1 out of df$c2
X-squared = 753.64, df=90, p-value < 0.0000000000000022
alternative hypothesis : two.sided.
sample estimation : prop1 prop2 prop3 prop4
0.000000 0.1089109 0.1176471 0.1262136

(chi square)
output : data : df
X-squared = 365.46, df=90, p-value < 0.00000000022

Experiment - 6

Density functions

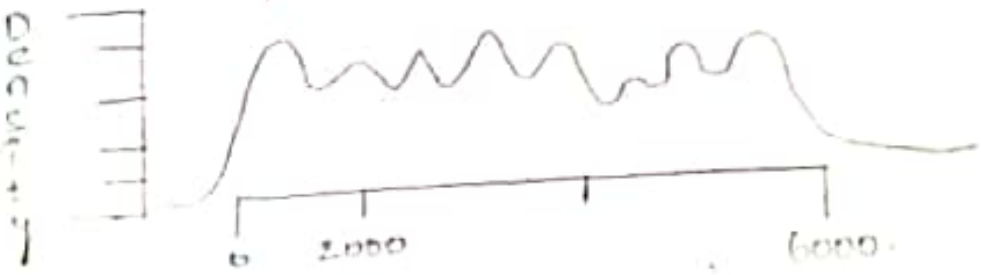
A density plot is a rep of the distribution of a numeric variable that uses a kernel density estimate to show the prob density function of variable use density().

```
library (read xl)
library ( ggplot2)
x<— read.csv(" Sample.csv")
den <— density ( x$CN)
Plot (den, frame =FALSE, col="blue", main=
    Density Plot").
```

Creation of histogram & density plot for same frame

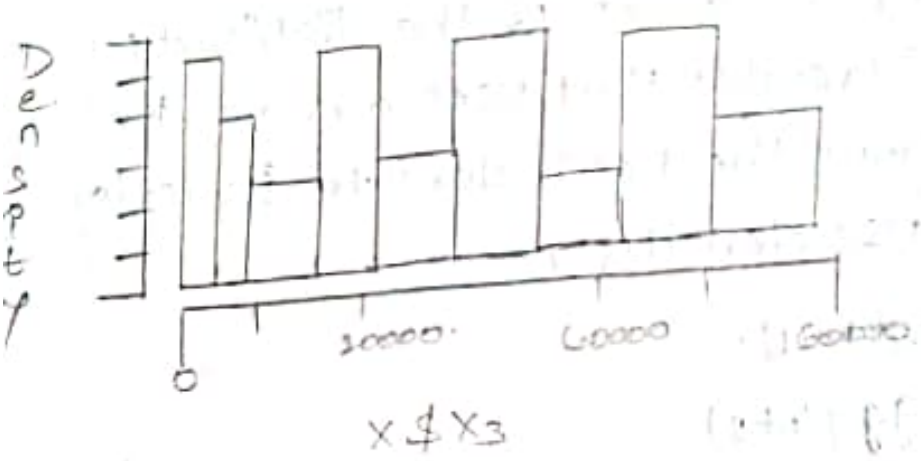
```
hist (x$CN , col="green", border="black")
lines ( density (x$CN), lwd=2, col="Red").
```

Density Plot

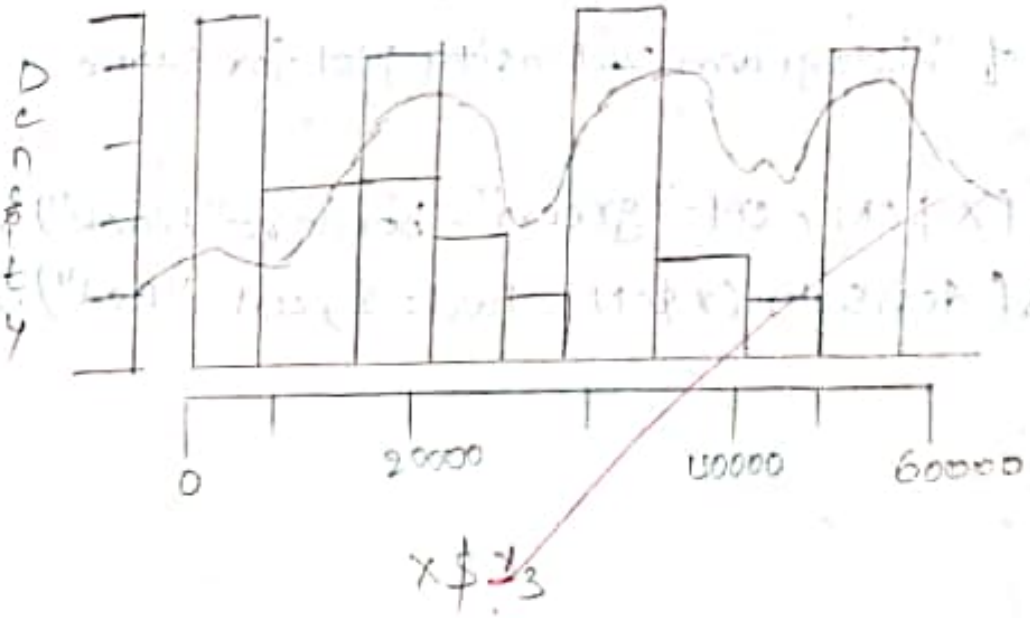


N = 4999 Bandwith = 263.1.

Histogram Plot of X $ X3



Lines of X $ X3



Experiment - D. (7)

Data Visualisation using ggplot

ggplot to :- Termed as grammer of graphics. Used for data visualizations mainly. It includes several layers.

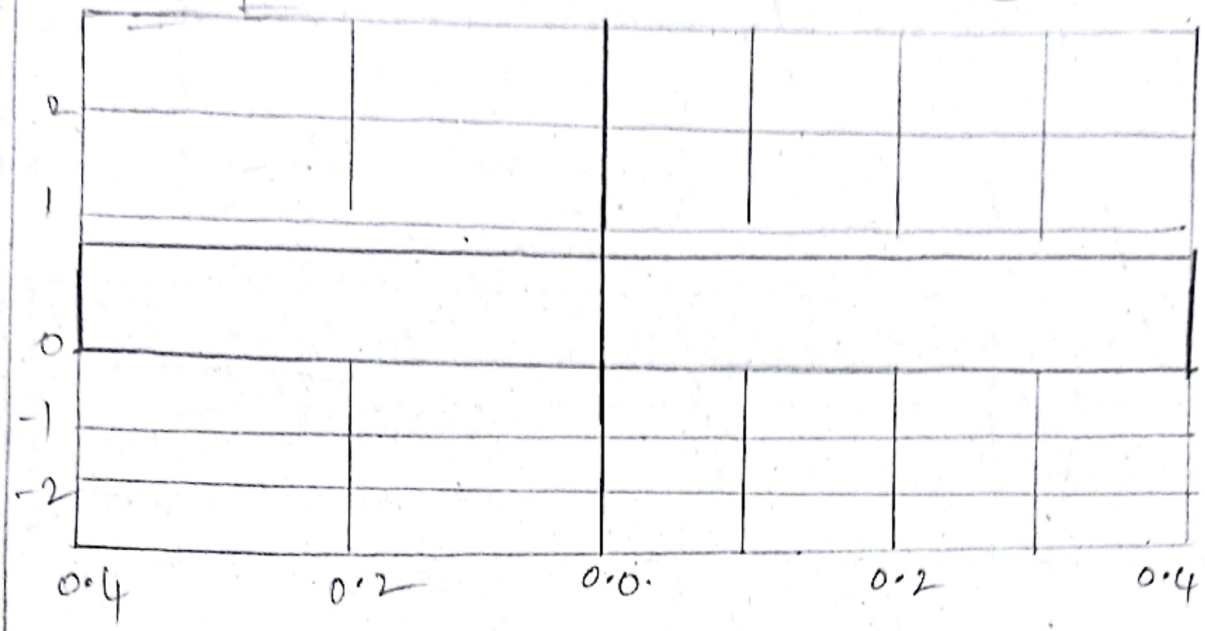
1. Data Layer :- The element is the data set itself.
2. Aesthetics :- The data is to map onto, x-axis, y-axis, colour, fill, size, labels, alpha, shape, line-width.
3. Geometric Layer :- How our data be displayed using point, line, histogram, bar, boxplot.
4. Facets Layer :- Displays at the subset of the data using columns & rows.
5. Statistics Layer :- Binning, smoothing, descriptive, intermediate.
6. Co-ordinate Layer :- The space b/w data and display using Partition, fixed, Polar, limits.
7. Team's Layer :- It is an non-data link related operation.

Example

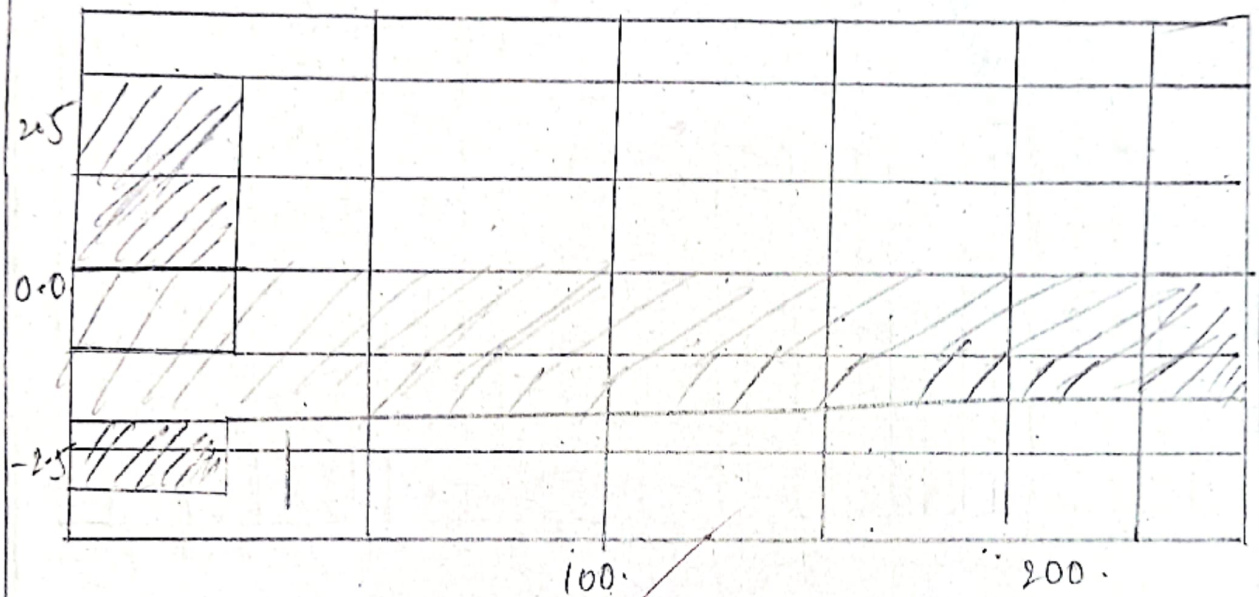
```
#Box plot
set.seed(3)
y <- rnorm(300)
dataf <- data.frame(y)
ggplot(dataf, aes(y=y)) +geom_boxplot(fill=2, alpha=0.5
```

```
color=1, outlier.color=2)
# Histogram
ggplot(dataf, aes(y=y)) + geom_histogram
    (binwidth=3, fill="magenta", color="green")
# scatter plotter
ggplot(Iris, aes(x=Speal.Length, y=Speal.Width)) +
    geom_point(col="magneta")
# linechart
ggplot(Iris, aes(x=Speal.length, y=Speal.width)).
    + geom_line(col="Blue")
# Barchart
ggplot(Iris, aes(x=Speal.Length, y=Speal.width)).
    + geom_bar(stat="identify", col="magneta")
# Heatmaps.
ggplot(Iris, aes(x=Speal.length, y=Speal.width))
    + geom_tile(col="real").
```

Box plot

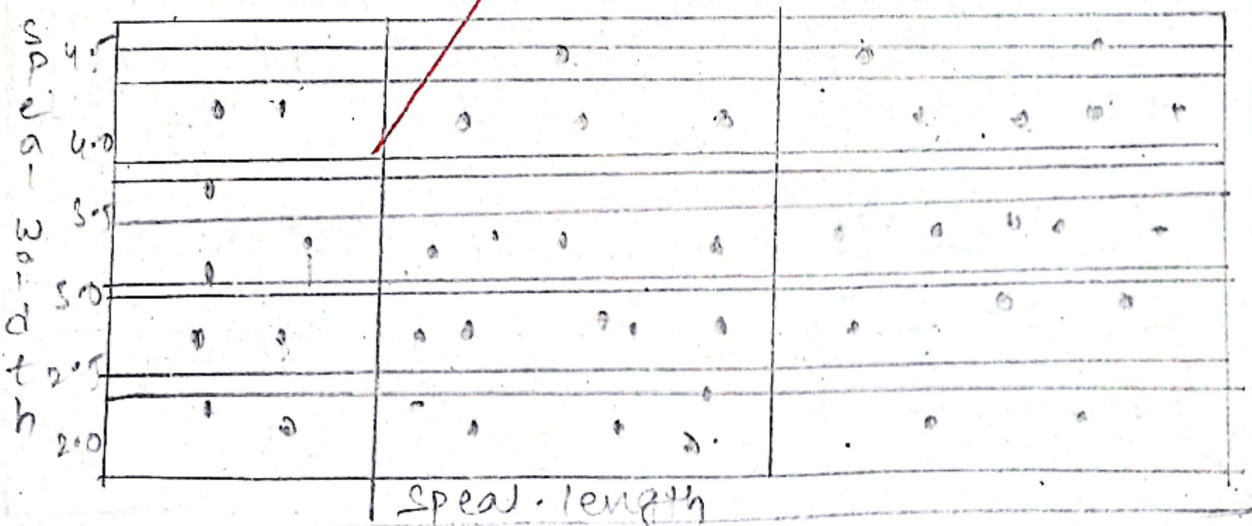


Histogram



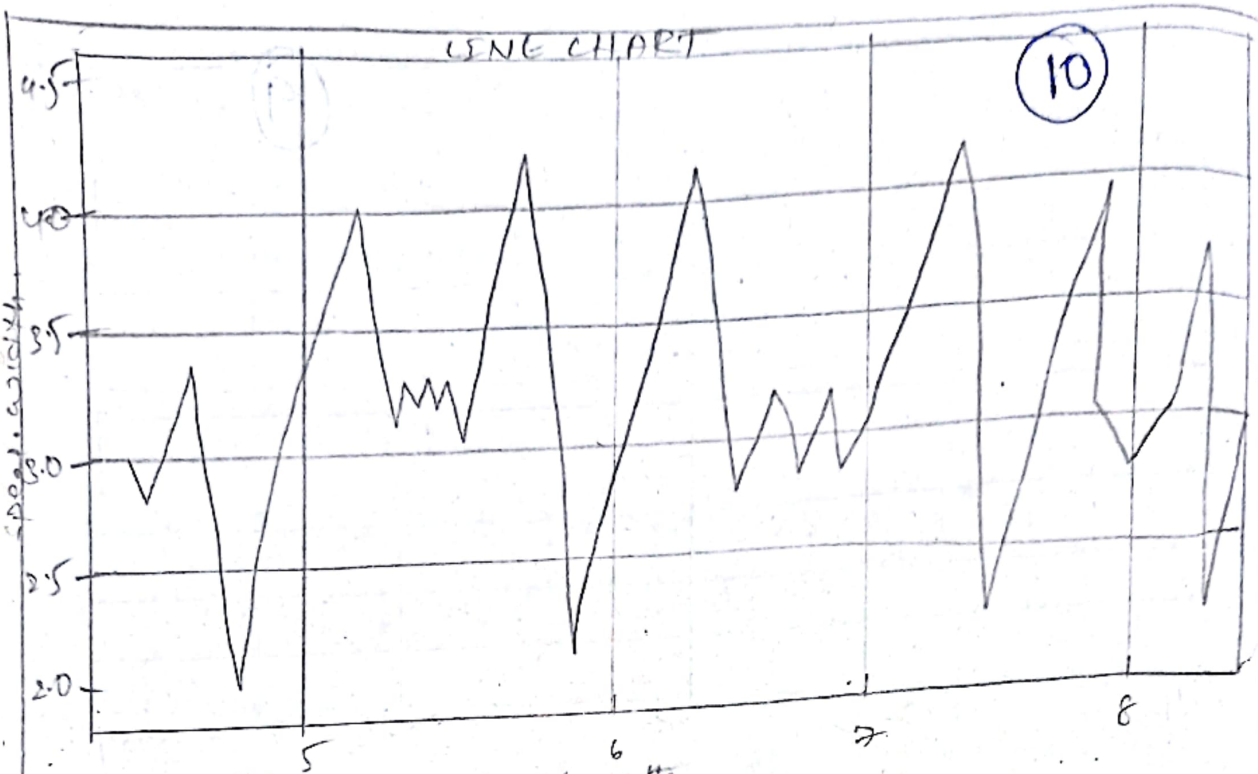
count.

Scatter plot.



LINE CHART

10



Speal. length
Bar chart.