Laxmi Charitable Trust's

Sheth L.U.J. & Sir M.V. College of Arts, Science & Commerce

Dr. S Radhakrishnan Marg, Andheri (E), Mumbai – 400 069.

Certificate

This is to certify that, Mr./Ms. Neeraj Appari

Seat No. F-129 *studying in* ***F.Y.B.Sc. SEM-II*** *Computer Science has satisfactorily completed the Practicals in the Subject of* Statistics *as prescribed by University of Mumbai, during academic year* ***2019-2020***.

Signature
Subject in charge
Date: -

Signature
Co-ordinator B.Sc. C.S
Date: -

Signature
External Examiner
Date: -

ALMP NO

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

Aptitude Alert

No	Date	Topic	Sign
1	7/1/2020	Solving problem based on binomial distribution	
2	7/1/2020	Solving problem based on normal distribution	14
3	14/1/2020	Property plotting of binomial distribution with expectation, variance and standard deviation	
4	14/1/2020	Property plotting of Pdf and Cdf	
5	21/1/2020	F test, normal test and t test	
6	28/1/2020	Property plotting of Pdf and Cdf	
7	7/02/2020	Analysis of variance (Anova)	
8	25/02/2020	Non parametric test - 1	
9	04/03/2020	Non parametric test - 2	
10	05/03/2020	Post of analysis One way Anova	

Scanned with CamScanner

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

Practica-1

Aim: Solving problem based on binomial distribution

1] It is observed that if students work hard then Chances of passing exam is very high ie 80%, random sample of such 10 students were selected what is the chance that :-

a] No students will pass the exam
b] 3 Students will pass the exam
c] greater than or equal to 8 students will pass the exam
d] All students will pass the exam

-> let x denote the result of student
let p denote ~~the possibility~~ probability of passing exam.
ie $p = 80\% = 0.8$
let q denote probability of students failing exam
ie $q = 1-p = 1-0.8 = 0.2$
formula :- $X \sim Bin(n,P)$
code :- (a) dbinom(0, size = 10, Prob = 0.8)

(b) dbinom(3, size = 10, prob=0.8)

c] dbinom (8:10, size=10, prob=0.8)

d) dbinom (10, size =10, prob= 0.8)

 ALNRAc

Vijay has started new retail outlet in mid of the market. In market there is business and competition. Therefore survival of a new outlet is very ~~ra~~ rare chance almost 4%. Vijay has started such 12 new retail outlet. Find out probability that a) atleast 3 shops will survive
b) exactly 7 shops will survive

let x denotes survival of retail outlets
let p denote probability of survival of outlets
i.e. P = 4% = 0.04
let q denote probability of failing survival
i.e. 1 - p = q = 1 - 0.04 = 0.96
let n denote the total 12 new retail outlets

Formula = dbinom (x, size = ___, prob ___)

a) dbinom (0:3, size = 12, prob = 0.04)
i.e. 1 - 0.3
b) dbinom (7, size = 12, prob = 0.04)
~~ie = 1 - 0.3~~

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

ALNRAO

Generate the random sample with $x = 9$, $n = 15$ & probability $= 0.3$

$x = 9$
size/n $= 15$
prob $= 0.3$

Formula := rbinom (x, size= ____ , prob = ____

rbinom (9, size= 15, prob = 0.3

14/1/20

```
[illegible].size=10,prob=0.8)
[illegible]1.024e-07
[illegible],size=10,prob=0.8)
[illegible].000786432
[illegible],size=10,prob=0.8)
[illegible]0.3019899 0.2684355 0.1073742
[illegible](10,size=10,prob=0.8)
[illegible].1073742
[illegible]9199+0.2684355+0.10737422
0.6777996

[illegible](0:12,size=12,prob=0.04)
[illegible] 9.780879e-03 9.141449e-04 6.094299e-05 2.962507e-06 1.058038e-07
[illegible] 2.755307e-09 5.102421e-11 6.378026e-13 4.831838e-15 1.677721e-17
[illegible],size=12,prob=0.04)
[illegible]58038e-07
[illegible],size=12,prob=[illegible])
[illegible].612709757 0.306354879 0.070206326 0.009750879
[illegible]+0.070206326+0.009750879
[illegible].9990218
[illegible](9,size=15,prob=0.3)
5 5 8 7 4 5 2 8 7
[illegible].612709757+0.306354879+0.070206326+0.009750879
0.9990218
[illegible]0.9990218
[illegible].0009782
[illegible](4400,[illegible],620)
[illegible].9486172
[illegible](3800,4400,620)
[illegible].9419844
[illegible],620) - pnorm(3800,4400,620)
[illegible].4266354
[illegible](66,[illegible],5.0)
0.1150697
[illegible]pnorm(72,69.8,3.0)
[illegible].2118554
```

ALNRAO

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

ALNRA

Practical-2

Aim :- Problem based on normal distribution

) The lion population in national park is approx. normally distributed with mean (μ) = 4400 and S.D (σ) = 620. Find the probability that

less than 5400

) more than 3300

) between 3500 and 4400

Given :- mean (μ) = 4400
S.D (σ) = 620

Formula :- Pnorm (x, μ, σ)

a) Pnorm (5400, 4400, 620)

b) 1 - Pnorm (3300, 4400, 620)

c) Pnorm (4400, 4400, 620) - Pnorm (3500, 4400, 620)

Scanned with CamScanner

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

Survey of height of Indian men was conducted, height are normally distributed with mean (μ) 69.6 & σ = 3.0 find the probability of height

a) for height less than 62 inch
b) for height more than 72 inch
c) for height in between 66 and 72 inch

→ Given : mean (μ) = 69.6 inch
S.D (σ) = 3.0

:- Pnorm (x, μ, σ)

→ Sol a) Pnorm (62, 69.6, 3.0)

b) 1 - Pnorm (72, 69.6, 3.0)

c) Pnorm (72, 69.6, 3.0) - Pnorm (62, 69.6, 3.0)

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

The weight of adult goat are normally distributed with mean (μ) = 25 kg and σ = 3kg. Find the probability of weight

a) less than 23 kg
b) more than 29 kg
c) between 20 kg and 27 kg
d) If 50 goats are randomly selected and with prob 0.25. Find the no of goats with weights 25 kg
e) Also if prob = 0.35 and 0.85 fore selected sample

mean (μ) = 25 kg & S.D (σ) = 3kg

a) Pnorm (x, μ, σ)
b) qnorm (Prob, μ, σ)

a) Pnorm (23, 25, 3)

b) 1 - Pnorm (29, 25, 3)

c) Pnorm (27, 25, 3) - Pnorm (20, 25, 3)

d) qnorm (0.25, 25, 3)

```
[1] 0.9990218
> [illegible]
[1] 0.0009782
> pnorm(5400,4400,620)
[1] 0.9466172
> 1-pnorm(3300,4400,620)
[1] 0.9619844
> pnorm(4400,4400,620)- pnorm(3500,4400,620)
[1] 0.4266954
> pnorm(66,69.6,3.0)
[1] 0.1150697

> 1-pnorm(72,69.6,3.0)
[1] 0.2118554

> pnorm(72,69.6,3.0)- pnorm(66,69.6,3.0)
[1] 0.6730749

> pnorm(23,25,3.0)
[1] 0.2524925

> 1-pnorm(29,25,3.0)
[1] 0.09121122

> pnorm(27,25,3.0)- pnorm(22,25,3.0)
[1] 0.6997171

> qnorm([illegible],25,[illegible]
[1] 22.97653

> qnorm([illegible]0.35,0.15),25,3)
Error: unexpected ',' in "qnorm((0.35,"
> qnorm(0.35,25,3)
[1] 23.84404
>
> qnorm(0.85,25,3)
[1] 28.1093
```

o

uary

ALN'RAO

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

ALNRAO

Practical -3

Aim - Property plotting of binomial distribution with expectation, variance and standard deviation

Q1) 40% of student passed in statistics in the survey of 65 students. Find the mean and variance

⇒ Probability of students passed : 0.4
Sample size of students - 65

Code -

```
Install.packages("distr6")
library(distr6)

x <- Binom(size=65, prob=0.4)
E(x)
var(x)
```

ALNDRAO

Output

$E(x) = 26$

$Var(x) = 15.6$

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

20% of the product produced are defective out of a lot of 30 products find its expectation, variance and standard deviation

Given - Probability of product being defective : 0.2
Sample size of Products = 30

Code :

```
library (distrEx)

x <- Binom (size = 30, prob = 0.2)
E (x)
Var (x)
Sd (x)
```

AKNIRAO

Output :

E (x) = 6

Var (x) = 4.8

Sd (x) = 2.19089

3 For the sample size of 65 and probability 0.25 find the expectation of Q

a) 2x+4

b) $\frac{2}{3}$ 2x+3

→ Given - probability : 0.25
Sample size : 65

a) Code :

```
library (distrEx)
x <- Binom (size = 65, prob = 0.25)
E (2*x + 4)
```

b) Code -

```
l
x <- Binom (size = 65, prob = 0.25)
E (2/3 * x + 3)
```

ALNRAO

Output

1) E (x) = 6
2) Var (x) = 4.8
3) sd (x) = 2.19089

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

Practical - 4

Aim · Property plotting of Pdf and cdf

1) For the following value of x & its corresponding prob. find mean (μ), variance and standard deviation
(x): 0, 1, 2, 3

-> Code:

i) x <- c (0,1,2,3)
f <- c(1/8), 3/8, 3/8, 1/8)
mu <- sum (x*f)
mu

ii) sigma2 <- sum((x-mu)^2*f)
sigma2

iii) sigma <- sqrt (sigma2)
sigma

Scanned with CamScanner

Output:

i) $E(2*x+4) = 56$

iii) $E(2/3*x+3)$: 20.3333

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

Q2. Find the probability for the following function where x lie between 0.14 and 2.3. $f = 3x^2$

-> Code - $0.14 < x < 2.3$ (given)
$f = 3x^2$

```
∴ f <- function (x) 3*x^2
integrate (f, lower = 0.14, upper = 2.3)
```

Q3 Find the value of x in the following function
$f = \frac{3}{x^2}$, limit 1 to ∞

∴ $1 < x < \infty$
$f = 3/x^2$

Code -
```
f <- function (x) 3/x^2
integrate (f, lower = 1, upper = Inf)
```

Q4 Obtain an Chi-square curve with 3 degrees of freedom $(0 < x < 20)$

-> Code

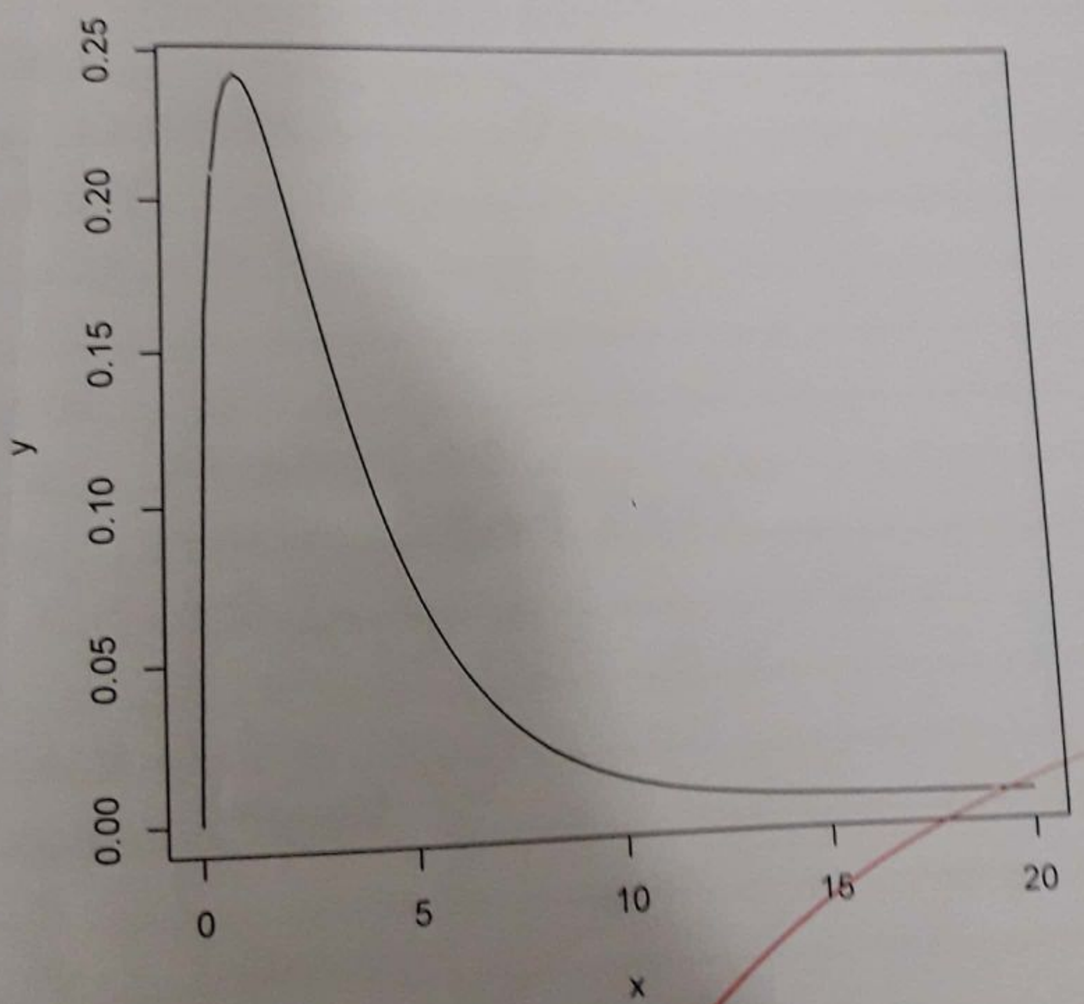
```
curve (dchisq (x, df = 3), from = 0, to = 20, ylab = "y").
```

ALMRAO

Output

mu=15
sigma2=0.75
sigma=0.8660254

ALNADO



ANURAG

Q2 12.16426 with obsolete error < 1.4e-13

Q3 3 with obsolete error < 3.3e-14

Q4

Practical - 05

Aim - T test, normal test and F test

Consider two random sample of size = 40 and 55 perform the F test for the 2 variables (x,y)

Code
~~x = rnorm ( )~~
Given
size = 40

Code
x = rnorm (40)
y = rnorm (~~50~~ (55)
~~z~~ t.test (x,y)

ttest = t.test (x,y)
names (ttest)

ttest $statistic

ttest [['statistic']]

 AUNRAU

For a random normal sample of size 50 and 30 with mean 0 and 1 and S.D 2 and 1 respectively. Perform the F test to compare their variance, also comment on the degrees of freedom

Given
Size = 50, 30
mean = 0, 1
SD = 2, 1

Code
```
x<-rnorm(50,0,2)
y<-rnorm(30,1,1)
var.test(x,y)

var.test(lm(x~1),lm(y~1))
```

AU

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science ALNRAO

3) Generate a random normal sample size = 100 and mean 0 calculate the weight with standard deviation file

->

Given
size = 100, mean = 0

Code

```
x<-rnorm (100, mean = 0)

shapiro.test(x)
```

CO2

```
y<-CO2[,5]
shapiro.test(Y)
```

ALNRAO

example1=output

```
> x=rnorm(40)
> y=rnorm(55)
> t.test(x,y)

        Welch Two Sample t-test

data:  x and y
t = 1.2159, df = 74.205, p-value = 0.2279
alternative hypothesis: true difference in means is
not equal to 0
95 percent confidence interval:
 -0.1536599  0.6348350
sample estimates:
  mean of x   mean of y
 0.05014231 -0.19044526

> ttest=t.test(x,y)
> names(ttest)
[1] "statistic"   "parameter"   "p.value"
"conf.int"    "estimate"    "null.value"
"alternative" "method"      "data.name"
> ttest$statistic
       t
1.215884
> ttest[['statistic']]
       t
1.215884
```

example2=output

```
> x<-rnorm(50,0,2)
> y<-rnorm(30,1,1)
> var.test(x,y)

        F test to compare two variances
```

ALNRAO

```
data:  x and y
F = 2.746, num df = 49, denom df = 29, p-value =
0.004702
alternative hypothesis: true ratio of variances is not
equal to 1
95 percent confidence interval:
 1.379631 5.166283
sample estimates:
ratio of variances
          2.745954

> var.test(lm(x~1),lm(y~1))

        F test to compare two variances

data:  lm(x ~ 1) and lm(y ~ 1)
F = 2.746, num df = 49, denom df = 29, p-value =
0.004702
alternative hypothesis: true ratio of variances is not
equal to 1
95 percent confidence interval:
 1.379631 5.166283
sample estimates:
ratio of variances
          2.745954
```

example3=output

```
> x<-rnorm(150,mean=0)
> shapiro.test(x)

        Shapiro-Wilk normality test

data:  x
W = 0.98352, p-value = 0.06997

> CO2
   Plant        Type  Treatment conc uptake
1    Qn1      Quebec nonchilled   95   16.0
2    Qn1      Quebec nonchilled  175   30.4
```

3	Qn1	Quebec	nonchilled	250	34.8
4	Qn1	Quebec	nonchilled	350	37.2
5	Qn1	Quebec	nonchilled	500	35.3
6	Qn1	Quebec	nonchilled	675	39.2
7	Qn1	Quebec	nonchilled	1000	39.7
8	Qn2	Quebec	nonchilled	95	13.6
9	Qn2	Quebec	nonchilled	175	27.3
10	Qn2	Quebec	nonchilled	250	37.1
11	Qn2	Quebec	nonchilled	350	41.8
12	Qn2	Quebec	nonchilled	500	40.6
13	Qn2	Quebec	nonchilled	675	41.4
14	Qn2	Quebec	nonchilled	1000	44.3
15	Qn3	Quebec	nonchilled	95	16.2
16	Qn3	Quebec	nonchilled	175	32.4
17	Qn3	Quebec	nonchilled	250	40.3
18	Qn3	Quebec	nonchilled	350	42.1
19	Qn3	Quebec	nonchilled	500	42.9
20	Qn3	Quebec	nonchilled	675	43.9
21	Qn3	Quebec	nonchilled	1000	45.5
22	Qc1	Quebec	chilled	95	14.2
23	Qc1	Quebec	chilled	175	24.1
24	Qc1	Quebec	chilled	250	30.3
25	Qc1	Quebec	chilled	350	34.6
26	Qc1	Quebec	chilled	500	32.5
27	Qc1	Quebec	chilled	675	35.4
28	Qc1	Quebec	chilled	1000	38.7
29	Qc2	Quebec	chilled	95	9.3
30	Qc2	Quebec	chilled	175	27.3
31	Qc2	Quebec	chilled	250	35.0
32	Qc2	Quebec	chilled	350	38.8
33	Qc2	Quebec	chilled	500	38.6
34	Qc2	Quebec	chilled	675	37.5
35	Qc2	Quebec	chilled	1000	42.4
36	Qc3	Quebec	chilled	95	15.1
37	Qc3	Quebec	chilled	175	21.0
38	Qc3	Quebec	chilled	250	38.1
39	Qc3	Quebec	chilled	350	34.0
40	Qc3	Quebec	chilled	500	38.9
41	Qc3	Quebec	chilled	675	39.6
42	Qc3	Quebec	chilled	1000	41.4

ALNRAO

43	Mn1	Mississippi	nonchilled	95	10.6
44	Mn1	Mississippi	nonchilled	175	19.2
45	Mn1	Mississippi	nonchilled	250	26.2
46	Mn1	Mississippi	nonchilled	350	30.0
47	Mn1	Mississippi	nonchilled	500	30.9
48	Mn1	Mississippi	nonchilled	675	32.4
49	Mn1	Mississippi	nonchilled	1000	35.5
50	Mn2	Mississippi	nonchilled	95	12.0
51	Mn2	Mississippi	nonchilled	175	22.0
52	Mn2	Mississippi	nonchilled	250	30.6
53	Mn2	Mississippi	nonchilled	350	31.8
54	Mn2	Mississippi	nonchilled	500	32.4
55	Mn2	Mississippi	nonchilled	675	31.1
56	Mn2	Mississippi	nonchilled	1000	31.5
57	Mn3	Mississippi	nonchilled	95	11.3
58	Mn3	Mississippi	nonchilled	175	19.4
59	Mn3	Mississippi	nonchilled	250	25.8
60	Mn3	Mississippi	nonchilled	350	27.9
61	Mn3	Mississippi	nonchilled	500	28.5
62	Mn3	Mississippi	nonchilled	675	28.1
63	Mn3	Mississippi	nonchilled	1000	27.8
64	Mc1	Mississippi	chilled	95	10.5
65	Mc1	Mississippi	chilled	175	14.9
66	Mc1	Mississippi	chilled	250	18.1
67	Mc1	Mississippi	chilled	350	18.9
68	Mc1	Mississippi	chilled	500	19.5
69	Mc1	Mississippi	chilled	675	22.2
70	Mc1	Mississippi	chilled	1000	21.9
71	Mc2	Mississippi	chilled	95	7.7
72	Mc2	Mississippi	chilled	175	11.4
73	Mc2	Mississippi	chilled	250	12.3
74	Mc2	Mississippi	chilled	350	13.0
75	Mc2	Mississippi	chilled	500	12.5
76	Mc2	Mississippi	chilled	675	13.7
77	Mc2	Mississippi	chilled	1000	14.4
78	Mc3	Mississippi	chilled	95	10.6
79	Mc3	Mississippi	chilled	175	18.0
80	Mc3	Mississippi	chilled	250	17.9
81	Mc3	Mississippi	chilled	350	17.9
82	Mc3	Mississippi	chilled	500	17.9

ALNRAO

```
83   Mc3 Mississippi    chilled  675    18.9
84   Mc3 Mississippi    chilled 1000    19.9
> y<-CO2
> y<-CO2[,5]
> shapiro.test(y)

        Shapiro-Wilk normality test

data:  y
W = 0.94105, p-value = 0.0007908

> y<-CO2[,5]
> shapiro.test(y)

        Shapiro-Wilk normality test

data:  y
W = 0.94105, p-value = 0.0007908
```

dbinom(k, size = 30, prob = 0.4)

dbinom(k, size = 30, prob = 0.15) ALNRAO

RAO

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science ALNRAO

Practical - 6

Aim: property plotting of pdf and cdf

The following examples show how to common the quantile of some common distribution for a given probability (or a number between 0 to 1)

Normal (0,1) distribution:

```
> y<-c(.01, 0.05, .1, .2, .5, .95, .99)
> qnorm(y, mean=0, sd=1)
```

Binomial (n,p) distribution

```
> y<-c(.01,.05, .1, .2, .5, .8, .95, .99)
> qbinom(y, size=30, prob=.2)
```

The foll examples illustrate how to generate 8 random sample from some of the well known probability distribution

Normal (μ, σ^2) distribution:

The first sample is from N(0,1) distribution and the next one from N(5,1) distribution

```
> z<-rnorm(10)
> z
```

dbinom(k, size = 30, prob = 0.4)

dbinom(k, size = 30, prob = 0.15) ALNRAO

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science ALNRAO

b) If you would like to see how the distribution of sample points looks like
-> w <- rnorm (1000, mean = 5, sd = 1)
> hist w

c) Binomial (n,p) distribution:
> k <- rbinom (20, size = 5, prob = .2)
> k

d) Plotting the probability density function (pdf) of a normal distribution
> x11 ( )
> x <- seq (-4.5, 4.5, .1)
> normdensity <- dnorm (x, mean = 0, sd = 1)
> plot (x, normdensity, type = "l")

e) Plotting the probability mass function (pmf) of a binomial distribution
> par (mfrow = c (2,1))
> k <- c (1:30)
> plot (k, dbinom (k, size = 30, prob = .15), type = "h")
> plot (k, dbinom (k, size = 30, prob = .4), type = "h")
> par (mfrow = c (1,1))

3) R has two different function than can be used for generating a (Q-Q) plot. use the function qqnorm for plotting sample quantities against theoritical quantities of standard normal random variable
> normsamp <- rnorm (100, mean = 0, sd = 1)

Scanned with CamScanner

A(NDPO

norm samp <- rnorm(100, mean = 5, sd = 1)
binorm samp <- rbinom(100, size = 20, prob = .25)

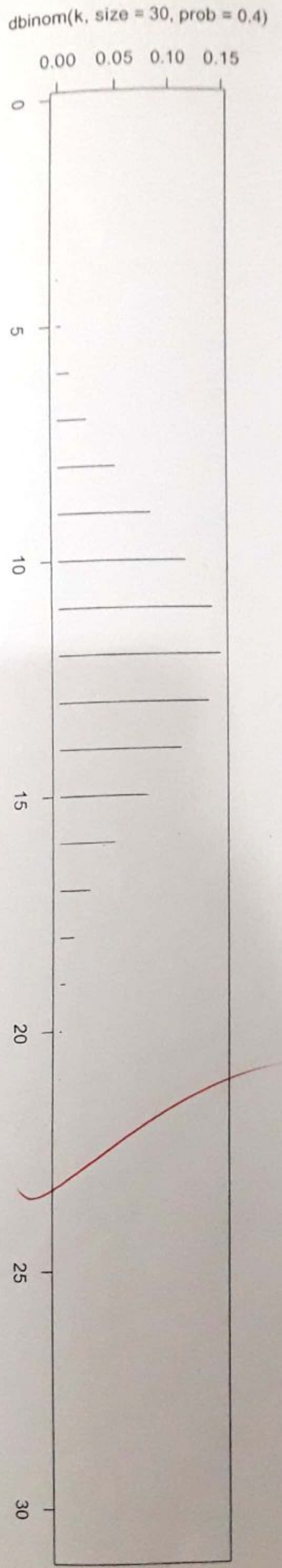
par (mfrow = c(2,2))

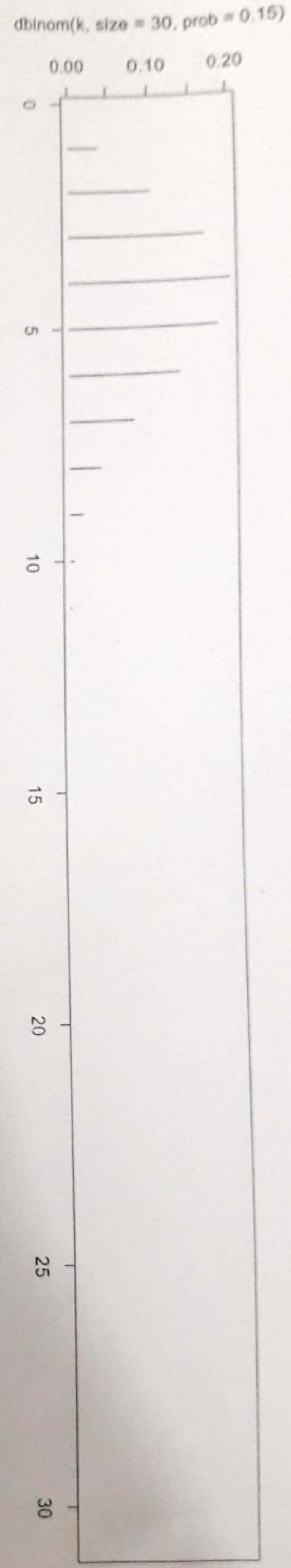
qqnorm(std norm samp, main = "Normal Q-Q plot : N(0,1) samples")
qqline (std norm samp, col = 2)
qqnorm (norm samp, main = "Normal Q-Q plot : N(5,1) samples")
qqline (norm samp, col = 2)
qqnorm (binom samp, main : "Normal Q-Q plot : Bin(20, .25) samples")
qqline (binom samp, col = 2)

Use of function qq plot for plotting sample quantiles for one sample against the sample quantiles of another sample

par (mfrow = c(2,1))
qqplot (std norm samp, norm samp, xlab = "sample quantiles N(0,1) sample",
ylab = "sample quantile : N(5,1) samples")
qq plot (std norm samp , binom samp, xlab = "sample quantiles N (0,1) Sample",
ylab = "sample quantiles : Bin (20, .25) samples")

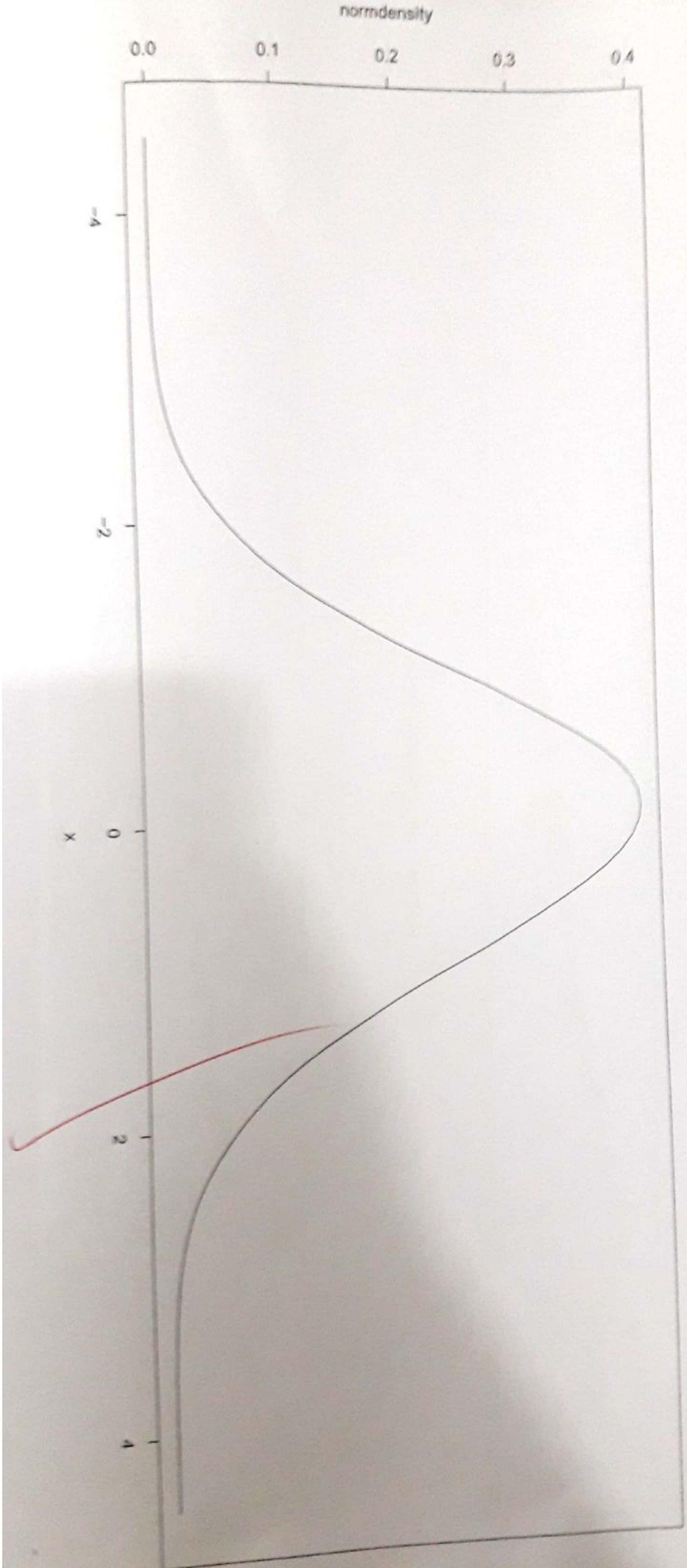
18/2/2020

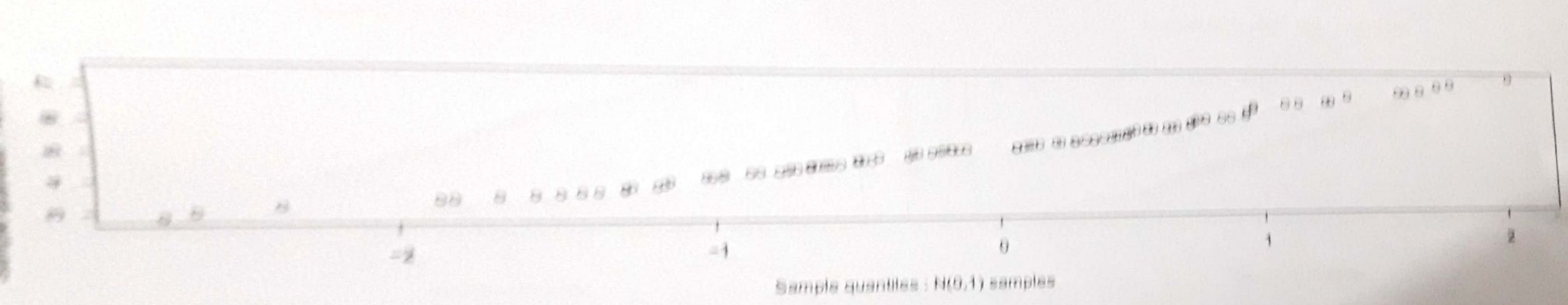


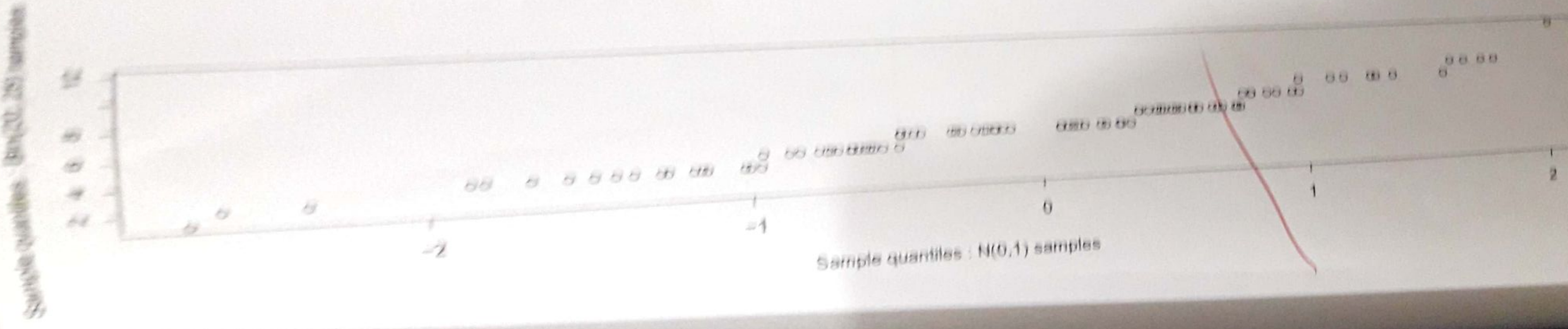


Scanned with CamScanner

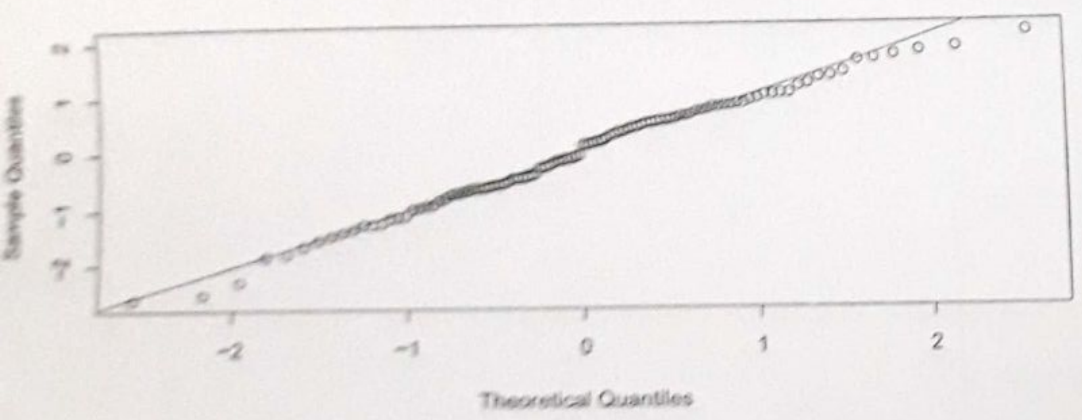
ALMRDO



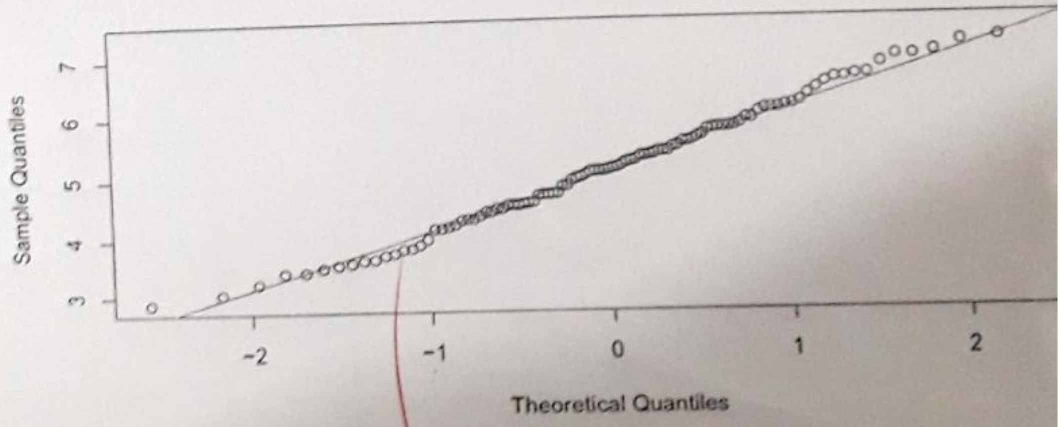

Sample quantiles : N(5,1) samples
3 4 5 6 7
-2 -1 0 1 2
Sample quantiles : N(0,1) samples


Sample quantiles : Bin(20, 25) samples
2 4 6 8 12
-2 -1 0 1 2
Sample quantiles : N(0,1) samples

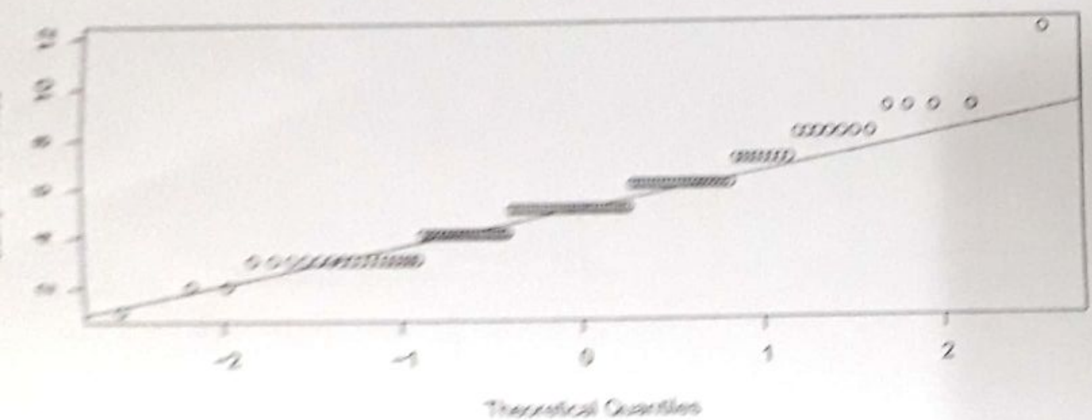
Normal Q–Q plot : N(0,1) samples



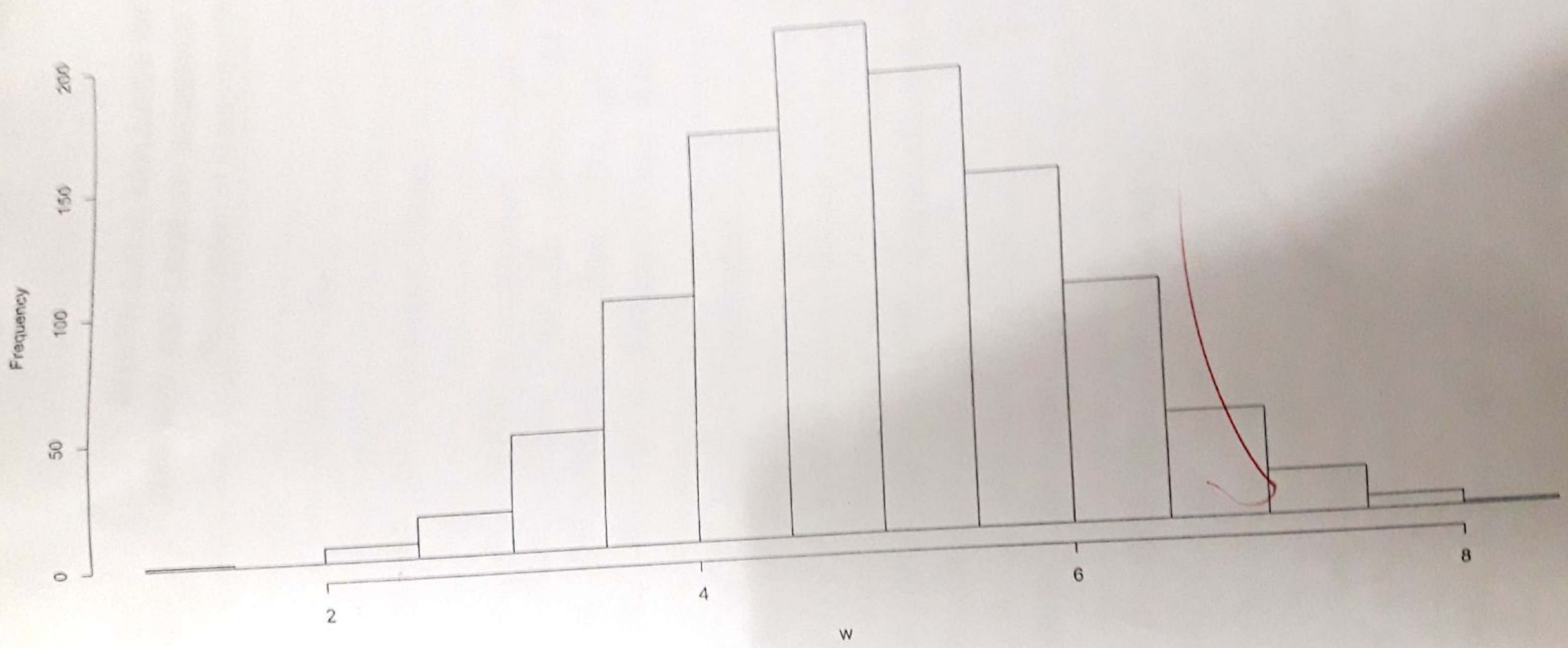
Normal Q–Q plot : N(5,1) samples



Normal Q–Q plot : Bin(20,.25) samples



Histogram of w



 ALNRAO

Practical No-7

Aim - Analysis of variance (Anova)

1) Suppose a company produce sugar and it wants to test the market on the basis of increasing demand prepare an observation from the selected population and comment on its variance - Also find its summary anova

->

field = read csv (file.choose())
field

your.aov = aov (field $growth ~ field $sugar)
your.aov

summary (your.aov)

1) b) field = read.csv (file.choose())
field

your.aov = aov (field $om ~ field $ YASH)
your.aov

summary (your.aov)

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

Ex-3) A media company wants to examine different combination media and marketing strategy the unit sales of mango juice for 3 different areas are collected using the hypothetical data provided below test whether which city has a proper marketing strategy. use 5% LOS

→

Area 1	Area 2	Area 3
492	676	577
712	626	538
559	682	581

~~use 95 los~~

$\bar{X}_j$ = 587.67 ~~565.33~~ 661.33 565.33 ~~587.67~~

N = Total Size = 9 $\bar{X}$ = 604.77

SSTR = $(x_{ij}(\bar{x}_{ij} - \bar{x})^2 = (3 \times (587.66 - 604.77)^2 + 3(661.33 - 604.77)^2 + 3 \times (565.33 - 604.77)^2$

= 878.25 + 9597.10 + 4660.54

= 15141.89

SSE = $\Sigma\Sigma(x_{ij} - \bar{x}_j) = (492 - 587.66)^2 + (712 - 587.66)^2 + (559 - 587.66)^2 + (676 - 661.33)^2 + (626 - 661.33)^2 + (682 - 661.33)^2 + (577 - 565.33)^2 + (538 - 565.33)^2 + (581 - 565.33)^2$

= 9150.33 + 15466.43 + 821.39 + 215.20 + 1248.20 + 427.24 + 69.38 + 746.92 + 245.54

= 28385.13

SST = SSTR + SSE

= 15141.89 + 28355.13

= 43497.02

18/2/2020

ALNRAO

$MSTR = \frac{SSTR}{C-1} = \frac{15141.89}{2} = 7570.945$

$MSE = \frac{SSE}{n-c} = \frac{28451.93}{6} = 4741.9883$

$F = \frac{MSTR}{MSE} = \frac{7570.945}{4741.9883} = 1.5965$

Anova Table for one way Analysis of variance

Source	Sum of Square	D.d.f
Between group	$SSTR = \Sigma r_j (\bar{X}_{ij} - \bar{X})^2 = 15141.89$	$C-1 = 3-1$
Within group	$SSE = \Sigma\Sigma (x_{ij} - \bar{X}_{i.}) = 28451.83$	$n-C = 9-3$
Total Correction	$SST = SSTR + SSE = 4{,}3593.52$	$n-1 = 9-1$

Mean Square	F ratio
$MSTR = \frac{SSTR}{(c-1)} = 7570.945$ $MSE = \frac{SSE}{(n-c)} = 4741.9883$	$F = \frac{MSTR}{MSE} = 1.5965$

28/2

ALNDAO

```
> field=read.csv(file.choose())
> field
   GROWTH SUGAR
1      75     A
2      72     B
3      73     D
4      69     B
5      68     C
6      75     A
7      65     A
8      76     B
9      98     C
10     77     A
11     54     D
12     45     D
13     65     D
14     45     A
15     78     C
16     55     B
17     74     A
18     23     C
19     43     D
20     54     A
21     45     A
22     34     B
23     22     C
24     43     B
25     89     D
26     56     C
```

```
27  33  A
28  65  A
29  34  C
30  11  C
31  32  C
32  67  C
33  86  D
34  97  A
35  45  A
36  64  D
37  23  C
38  33  A
39  54  B
40  66  B
41  75  A
42  54  C
43  23  A
44  43  D
45  55  C
46  65  B
47  76  A
48  11  A
49  86  D
```

```
> your.aov=aov(field$GROWTH~field$SUGAR)
> your.aov
Call:
   aov(formula = field$GROWTH ~ field$SUGAR)
```

ALMAD

Terms:

```
            field$SUGAR Residuals
Sum of Squares   1755.274 21246.849
Deg. of Freedom         3        45

Residual standard error: 21.72906
Estimated effects may be unbalanced
> summary(your.aov)
            Df Sum Sq Mean Sq F value Pr(>F)
field$SUGAR  3   1755   585.1   1.239  0.307
Residuals   45  21247   472.2
> field=read.csv(file.choose())
> field
   OM YASH
1  65  A
2  66  V
3  44  C
4  23  F
5  54  G
6  66  K
7  22  H
8  44  R
9  44  D
10 89  S
11 66  D
12 88  A
13 12  F
14 33  E
15 42  F
```

ALNRAO

16 55 C

17 32 V

18 13 A

19 53 R

20 11 F

21 78 S

22 54 E

23 12 A

24 43 F

25 66 F

```
> your.aov=aov(field$OM~field$YASH)

> your.aov

Call:

   aov(formula = field$OM ~ field$YASH)


Terms:

             field$YASH Residuals

Sum of Squares    5106.167  7857.833

Deg. of Freedom        10        14


Residual standard error: 23.69122

Estimated effects may be unbalanced

> summary(your.aov)

            Df Sum Sq Mean Sq F value Pr(>F)

field$YASH  10   5106   510.6    0.91   0.55

Residuals   14   7858   561.3
```

ALNRAO

Q1

```
>binom.test(5,18)

        Exact binomial test

data:  5 and 18
number of successes = 5, number of trials = 18, p-value = 0.09625
alternative hypothesis: true probability of success is not equal to 0.5
95 percent confidence interval:
 0.09694921 0.53480197
sample estimates:
probability of success
             0.2777778
```

Q2

```
library(MASS)
> head(immer)
  Loc Var    Y1    Y2
1  UF   M  81.0  80.7
2  UF   S 105.4  82.3
3  UF   V 119.7  80.4
4  UF   T 109.7  87.2
5  UF   P  98.3  84.2
6   W   M 146.6 100.4
wilcox.test(immer$Y1,immer$Y2,paired=TRUE)

        Wilcoxon signed rank test with continuity correction

data:  immer$Y1 and immer$Y2
V = 368.5, p-value = 0.005318
alternative hypothesis: true location shift is not equal to 0
```

Q3

```
library(BSDA)
Loading required package: lattice

Attaching package: 'BSDA'

The following object is masked from 'package:datasets':

    Orange
> x<-c(7.8,6.6,6.5,7.4,7.3,7.7,6.4,7.1,6.7,7.6,6.8)
> SIGN.test(x,md=6.5)

        One-sample Sign-Test

data:  x
s = 9, p-value = 0.02148
alternative hypothesis: true median is not equal to 6.5
95 percent confidence interval:
 6.571273 7.628727
```

BCMRAO

```
sample estimates:
median of x
      7.1

Achieved and Interpolated Confidence Intervals:

                  Conf.Level L.E.pt U.E.pt
Lower Achieved CI     0.9346 6.6000 7.6000
Interpolated CI       0.9500 6.5713 7.6287
Upper Achieved CI     0.9883 6.5000 7.7000
```

Q4

```
> Input =("
  Speaker              Rater     Likert
+ 'Maggie Simpson'     1         3
+ 'Maggie Simpson'     2         4
+ 'Maggie Simpson'     3         5
+ 'Maggie Simpson'     4         4
+ 'Maggie Simpson'     5         4
+ 'Maggie Simpson'     6         4
+ 'Maggie Simpson'     7         4
+ 'Maggie Simpson'     8         3
+ 'Maggie Simpson'     9         2
+ 'Maggie Simpson'     10        5
+
+ ")
> Data = read.table(textConnection(Input),header=TRUE)
> SIGN.test(Data$Likert,
+           md = 3)

        One-sample Sign-Test

data:  Data$Likert
s = 7, p-value = 0.07031
alternative hypothesis: true median is not equal to 3
95 percent confidence interval:
 3.000000 4.675556
sample estimates:
median of x
          4

Achieved and Interpolated Confidence Intervals:

                  Conf.Level L.E.pt U.E.pt
Lower Achieved CI     0.8906      3 4.0000
Interpolated CI       0.9500      3 4.6756
Upper Achieved CI     0.9785      3 5.0000
```

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

Practical - 8

Aim: Non parametric test

1) A softdrink company has invented a new drink and would like to find out if it will be as popular as the existing favourite drink for this purpose, its research department arranges 18 participants for taste testing. Each participants tries both drinks in random order before giving his or her opinion
> binom-test (5, 18)

2) In the built in data set named immer, the barley yield in years 1931 and 1932 of the same field are recorded. The yield data are presented in the data frame columns Y1 and Y2
> library (MASS)
> head (immer)
> wilcox. test (immer$Y1, immer $Y2, paired = TRUE)

3) > library (BSDA)
> x <- c (7.8, 6.6, 6.5, 7.4, 7.3, 7, 6.4, 7.1, 6.7, 7.6, 6.8)
> SIGN.test (x, md = 6.5)

ALNRAO

13) > Input = c("
+ speaker Rater likert
+ 'Maggie Simpson' 1 3
+ 'Maggie Simpson' 2 4
+ 'Maggie Simpson' 3 5
+ 'Maggie Simpson' 4 4
+ 'Maggie Simpson' 5 4
+ 'Maggie Simpson' 6 4
+ 'Maggie Simpson' 7 ~~3~~ 4
+ 'Maggie Simpson' 8 ~~2~~ 3
+ 'Maggie Simpson' 9 ~~5~~ 2
+ 'Maggie Simpson' 10 5

> Data = read.table(textconnection(Input), header = TRUE)
> SIGN.test (Data $likert, md=3)

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

ALNAAO

Practical No-9

Aim- Non-Parametric Test 2

Suppose a company produce sugar and it want to test the market on the basis of increasing demand prepare a observation from the scatered population find its variance & Tukey's method

field= read.csv (file.choose())
field

your.aov = aov (field $ sugar ~ field $ growth
your.aov

Tukey HSD (your.aov)

Test using Kruskals wallis random test.

attach (field)
field.kw = kruskal.test (growth, sugar)
field.kw

Prepare a box plot for Tukey test and Kruskal test

boxplot (field)

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

Q4 Generate random and Run a test

library (tseries)

a) x <- factor (sign (rnorm (100)))
runs.test

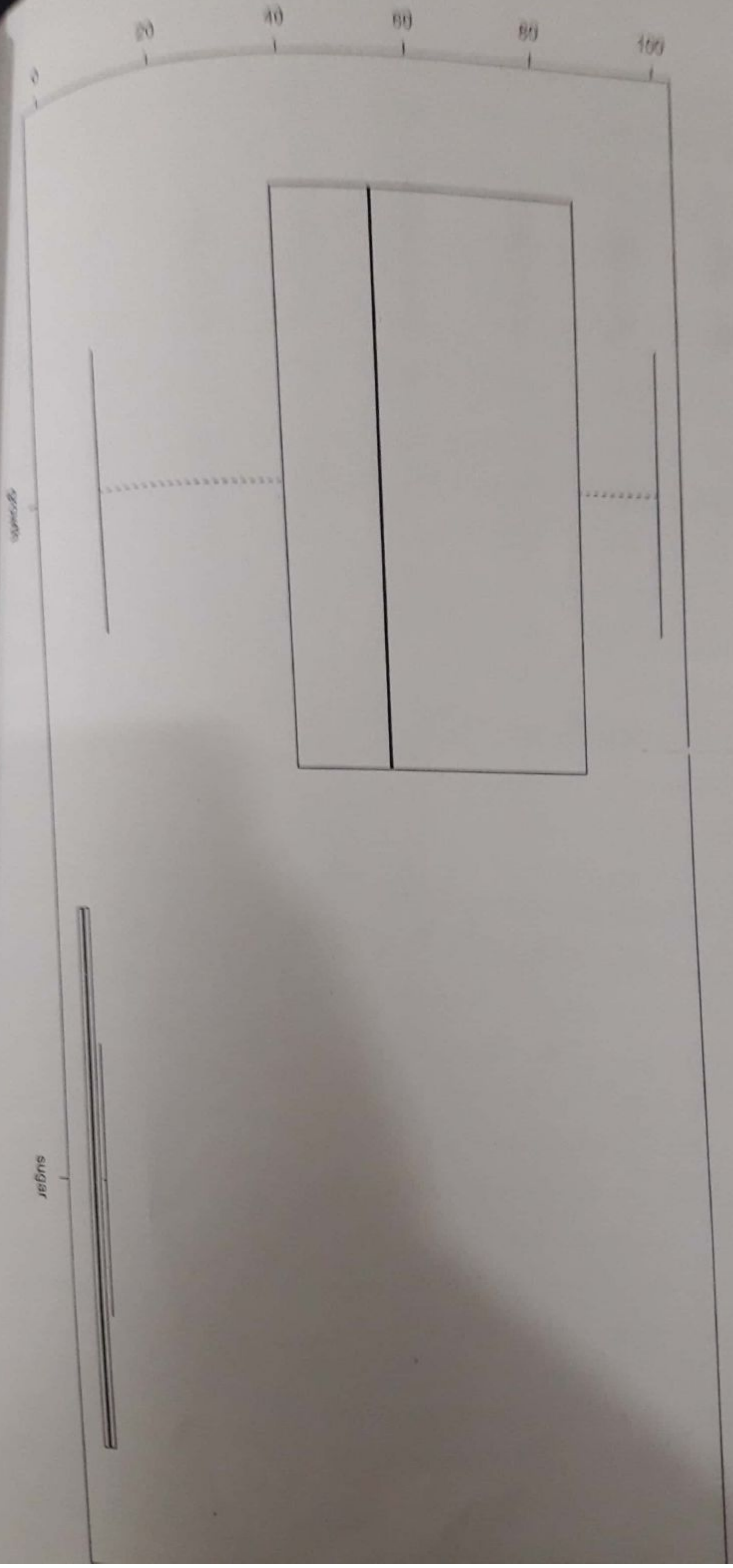
b) x <- factor (rep (c(-1,1), 50))
runs.test

ALNRAO

```
> field=read.csv(file.choose())
> field
   growth sugar
1      45     A
2      32     B
3      45     C
4      67     D
5      31     D
6      23     B
7      87     C
8      98     A
9      13     B
10     43     D
11     67     C
12     89     A
13      9     D
14     76     A
15     45     B
16     34     C
17     78     D
18     90     B
19     86     C
20     98     A
21     53     B
22     74     A
23     94     D
24     90     C
25     38     A
26     49     B
27     38     A
28     38     D
29     93     C
30     86     B
31     54     A
32     66     B
33     77     D
34     88     C
35     99     D
36     48     B
37     44     A
38     55     B
39     97     D
40     67     B
41     16     A
42     18     C
43     53     C
44     70     D
45     50     A
46     40     B
47     30     C
48     21     A
> your.aov=aov(field $growth ~ field $sugar)
> TukeyHSD(your.aov)
  Tukey multiple comparisons of means
```

ALNRAO

```
95% family-wise confidence level

Fit: aov(formula = field$growth ~ field$sugar)

$`field$sugar`
          diff       lwr      upr     p adj
B-A  -5.692308 -33.81074 22.42612 0.9485972
C-A   5.818182 -23.55056 35.18692 0.9515953
D-A   6.909091 -22.45965 36.27783 0.9224587
C-B  11.510490 -17.85825 40.87923 0.7232833
D-B  12.601399 -16.76734 41.97014 0.6636489
D-C   1.090909 -29.47705 31.65886 0.9996851

> attach(field)
> field.kw=kruskal.test(growth,sugar)
> field.kw

        Kruskal-Wallis rank sum test

data:  growth and sugar
Kruskal-Wallis chi-squared = 1.4936, df = 3, p-value = 0.6838

> boxplot(field)
>installpackages(tseries)
>library(tseries)
> x<-factor(sign(rnorm(100)))
> runs.test(x)

        Runs Test

data:  x
Standard Normal = 0.40606, p-value = 0.6847
alternative hypothesis: two.sided

> x<-factor(rep(c(-1,1),50))
> runs.test(x)

        Runs Test

data:  x
Standard Normal = 9.8499, p-value < 2.2e-16
alternative hypothesis: two.sided
```



BINDAU

SHETH L.U.J. COLLEGE OF ARTS &
SIR M.V. COLLEGE OF SCIENCE & COMMERCE
Department of Computer Science

Practical No 10

Aim: Post of analysis One-way-anova

Consider the source and test the post of analysis also comment on the output

Prepare a data and perform a pair wise test, Tukey comment on the plot produced using Tukey test and explain the summary anova

Code

```
source ("https://www.r-statistics.com/wp-content/uploads/2010/02/
        Friedman-Test-with-Post-Hoc.r.txt")
>Winetesting <- data.frame(
+Task = c(5.40, 5.50, 5.55,
+ 5.85, 5.70, 5.75,
+5.20, 5.60, 5.50,
+5.55, 5.50, 5.40,
+5.90, 5.85, 5.70,
+ 5.45, 5.55, 5.60,
+5.40, 5.40, 5.35,
+ 5.45, 5.50, 5.35,
+ 5.25, 5.15, 5.00,
+5.85, 5.80, 5.70,
+5.25, 5.20, 5.10,
+5.65, 5.55, 5.45,
+ 5.05, 5.00, 5.05,
```

```
+ 5.45, 5.55, 5.35
+ 5.50, 5.45, 5.25
+ 6.30, 6.30, 6.25),
+ Wine = factor (rep (c("Wine A", "Wine B", "Wine C"), 22))
+ Taster = factor (rep (1:22, rep (3, 22))))

> with (Wine Tasting, boxplot (Taste ~ Wine))
> friedman.test.with.post.hoc (Taste ~ Wine | Taster, WineTasting)
> field = read.csv (file.choose())
> field
> attach (field)
> group = as.factor (sugar)
> mm = tapply (growth, group, length), sdsd = tapply (growth, group, sd)
  > mm; sdsd

> aov.exl = aov (growth ~ group)
> summary (aov.exl, intercept = T)

3) > Pairwise.t.test (growth, group, p.adjust = "none", pool.sd = T)
> Pairwise.t.test (growth, group, p.adjust = "bonferroni", pool.sd = T)

> TukeyHSD (aov.exl, conf.level = 95)
> plot (TukeyHSD (aov (growth ~ group), conf.level = 95))

> summary (aov.exl)
```

ALMRAO


```
> source("https://www.r-statistics.com/wp-content/uploads/2010/02/Friedman-Test-with-Post-Hoc.r.txt")

> WineTasting <- data.frame(

+  Taste = c(5.40, 5.50, 5.55,

+ 5.85, 5.70, 5.75,

+ 5.20, 5.60, 5.50,

+ 5.55, 5.50, 5.40,

+ 5.90, 5.85, 5.70,

+ 5.45, 5.55, 5.60,

+ 5.40, 5.40, 5.35,

+ 5.45, 5.50, 5.35,

+ 5.25, 5.15, 5.00,

+ 5.85, 5.80, 5.70,

+ 5.25, 5.20, 5.10,

+ 5.65, 5.55, 5.45,

+ 5.60, 5.35, 5.45,

+ 5.05, 5.00, 4.95,

+ 5.50, 5.50, 5.40,

+ 5.45, 5.55, 5.50,

+ 5.55, 5.55, 5.35,

+ 5.45, 5.50, 5.55,

+ 5.50, 5.45, 5.25,

+ 5.65, 5.60, 5.40,

+ 5.70, 5.65, 5.55,

+ 6.30, 6.30, 6.25),

+ Wine = factor(rep(c("Wine A", "Wine B", "Wine C"), 22)),

+ Taster = factor(rep(1:22, rep(3, 22))))

>

> with(WineTasting , boxplot( Taste  ~ Wine ))

> friedman.test.with.post.hoc(Taste ~ Wine | Taster ,WineTasting)
```

ALMRDO

```
> field=read.csv(file.choose())
> field
   growth sugar
1      23     1
2      24     1
3      28     1
4      26     1
5      34     1
6       7     1
7       8     1
8       9     1
9      11     2
10     12     2
11     15     2
12     17     2
13     18     2
14     19     2
15     15     2
16     29     2
17     28     3
18     27     3
19     39     3
20     38     3
21     37     3
22     36     3
23     35     3
24     34     3
25     32     4
26     21     4
27     12     4
```

ALNIRAU

```
28 33 4
29 23 4
30 16 4
31 15 4
32 23 4
33 34 5
34 12 5
35 34 5
36 6 5
37 4 5
38 3 5
39 2 5
40 12 5
41 18 6
42 17 6
43 22 6
44 34 6
45 12 6
46 32 6
47 34 6
48 23 6
> attach(field)
> group=as.factor(sugar)
> mm=tapply(growth,group,mean)
> mm
  1 2 3 4 5 6
19.875 17.000 34.250 21.875 13.375 24.000
> nn=tapply(growth,group,length);sdsd=tapply(growth,group,sd);nn;sdsd
1 2 3 4 5 6
8 8 8 8 8 8
  1 2 3 4 5 6
```

ALNPAU

```
10.384570 5.580579 4.464143 7.642690 13.276591 8.434623
> aov.exl=aov(growth~group)
> summary(aov.exl,intercept=T)
            Df Sum Sq Mean Sq F value  Pr(>F)
(Intercept)  1  22664   22664  292.60 < 2e-16 ***
group        5   2060     412    5.32 0.00071 ***
Residuals   42   3253      77
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> pairwise.t.test(growth,group,p.adjust="none",pool.sd=T)

        Pairwise comparisons using t tests with pooled SD

data:  growth and group

  1       2       3       4       5
2 0.51710 -       -       -       -
3 0.00217 0.00032 -       -       -
4 0.65181 0.27423 0.00745 -       -
5 0.14710 0.41472 2.4e-05 0.06017 -
6 0.35391 0.11917 0.02473 0.63167 0.02019
P value adjustment method: none
pairwise.t.test(growth,group,p.adjust="bonferroni",pool.sd=T)

        Pairwise comparisons using t tests with pooled SD

data:  growth and group

  1       2       3       4       5
2 1.00000 -       -       -       -
3 0.03257 0.00481 -       -       -
```

ALNRRO

4 1.00000 1.00000 0.11173 - -

5 1.00000 1.00000 0.00037 0.90257 -

6 1.00000 1.00000 0.37088 1.00000 0.30285

P value adjustment method: Bonferroni

TukeyHSD(aov.ex1, conf.level=.95)

Tukey multiple comparisons of means

95% family-wise confidence level

Fit: aov(formula = growth ~ group)

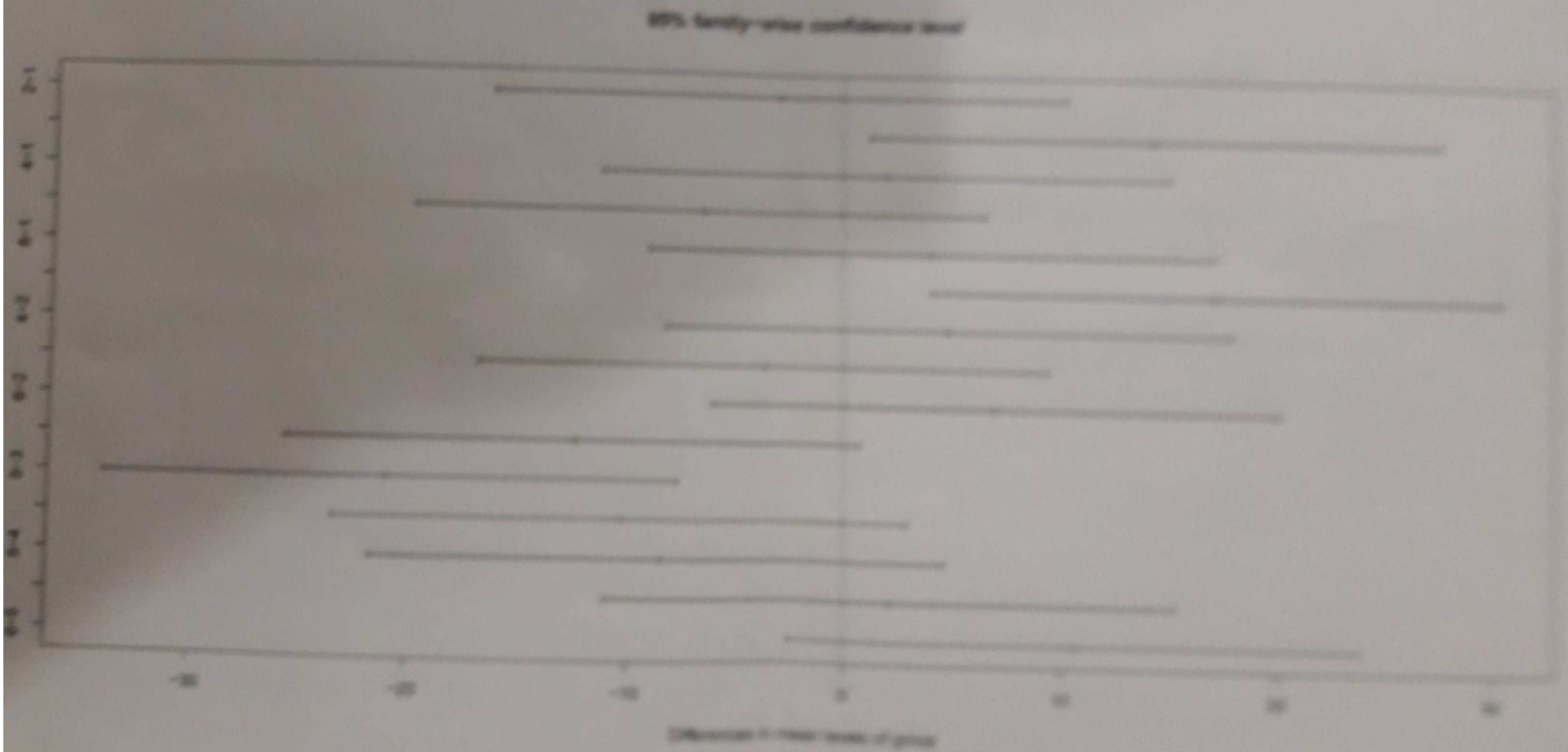
$group

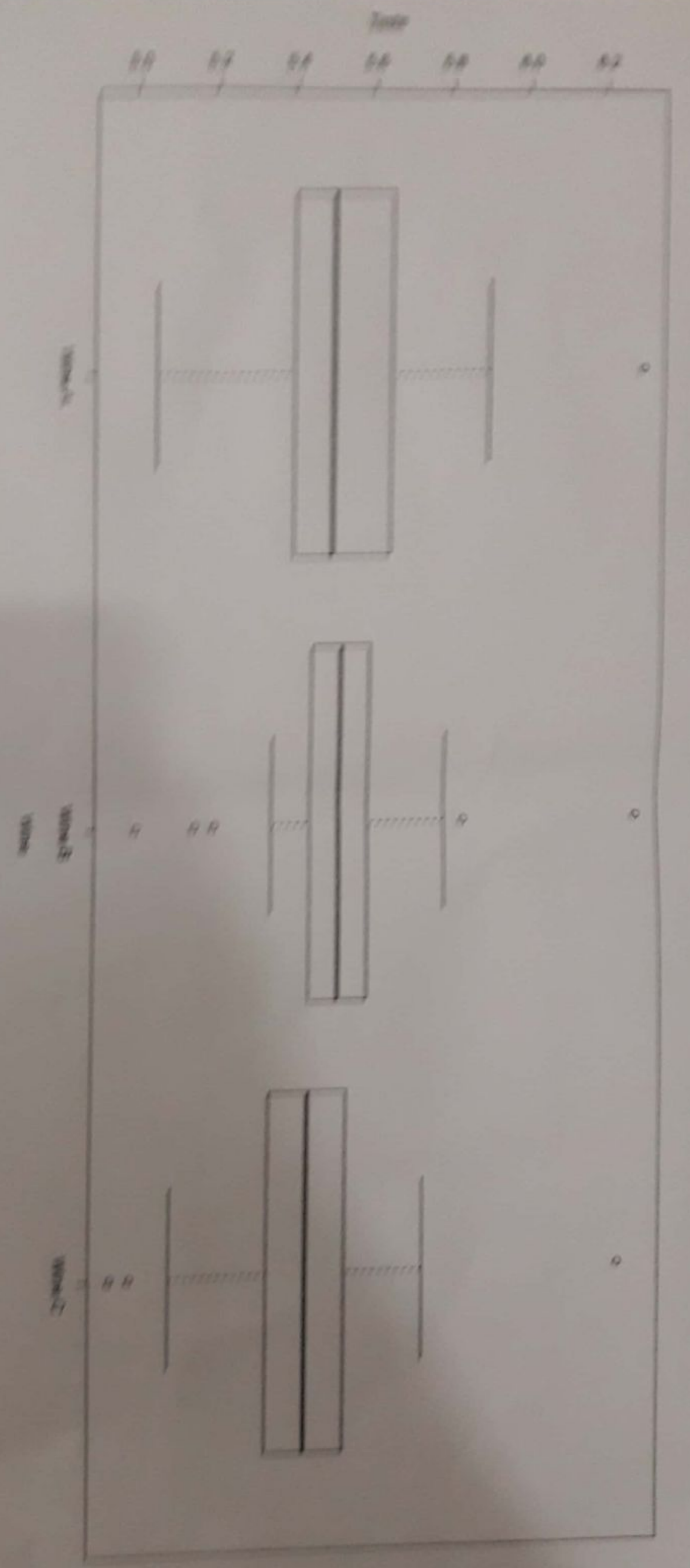
	diff	lwr	upr	p adj
2-1	-2.875	-16.011397	10.2613972	0.9859766
3-1	14.375	1.238603	27.5113972	0.0247157
4-1	2.000	-11.136397	15.1363972	0.9973920
5-1	-6.500	-19.636397	6.6363972	0.6802680
6-1	4.125	-9.011397	17.2613972	0.9343909
3-2	17.250	4.113603	30.3863972	0.0040719
4-2	4.875	-8.261397	18.0113972	0.8753930
5-2	-3.625	-16.761397	9.5113972	0.9614416
6-2	7.000	-6.136397	20.1363972	0.6089973
4-3	-12.375	-25.511397	0.7613972	0.0750699
5-3	-20.875	-34.011397	-7.7386028	0.0003326
6-3	-10.250	-23.386397	2.8863972	0.2054094
5-4	-8.500	-21.636397	4.6363972	0.3979603
6-4	2.125	-11.011397	15.2613972	0.9965252
6-5	10.625	-2.511397	23.7613972	0.1745835

plot(TukeyHSD(aov(growth~group), conf.level=.95))

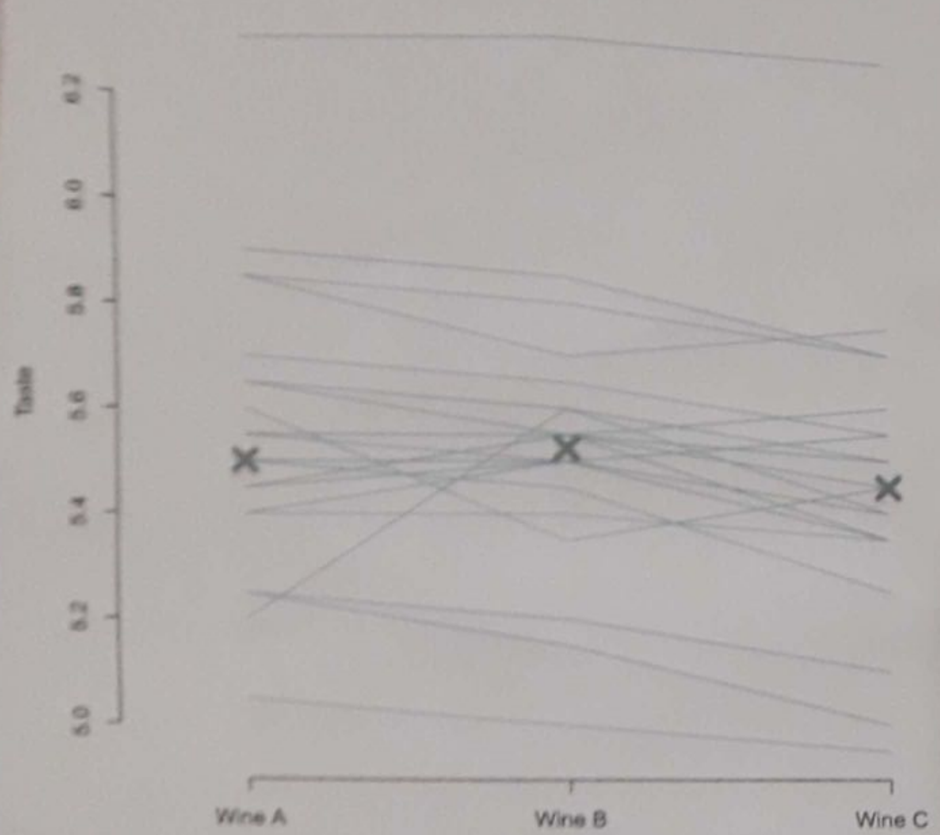
Alnrao

```
> summary(aov.exl)
          Df Sum Sq Mean Sq F value  Pr(>F)
group      5   2060   412.1    5.32 0.00071 ***
Residuals 42   3253    77.5
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

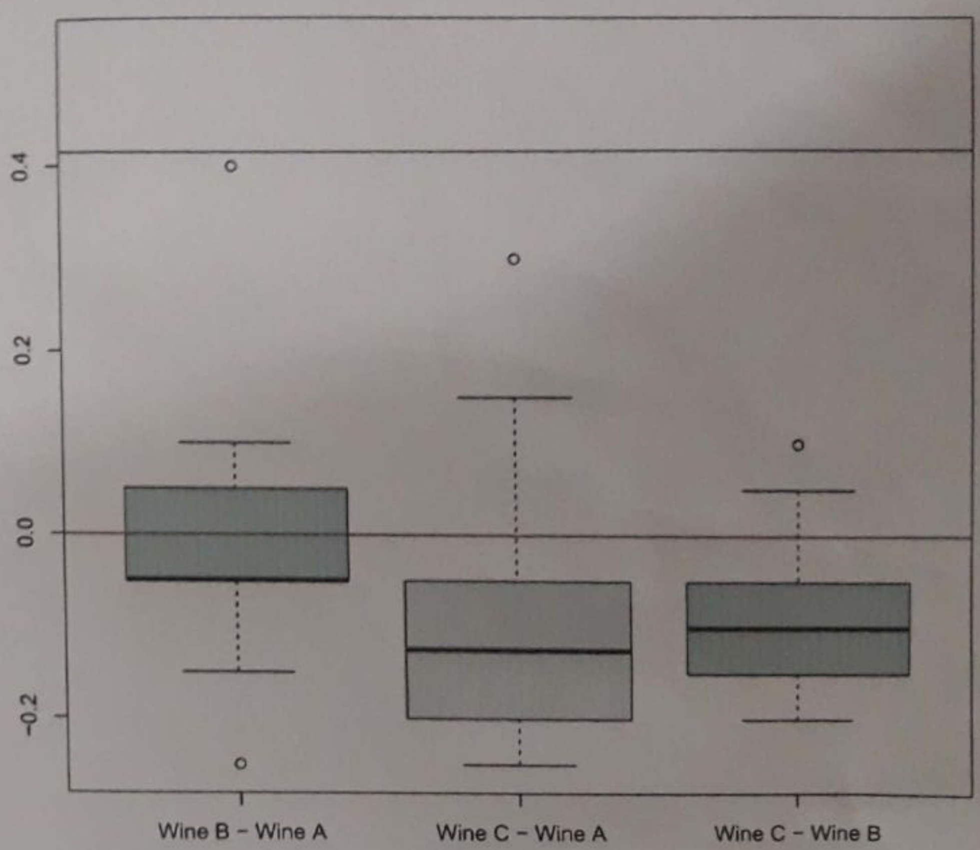

95% family-wise confidence level
-30
-20
-10
0
10
20
30
Differences in mean levels of group



DANP AU



Parallel coordinates plot
Taste
5.0
5.2
5.4
5.6
5.8
6.0
6.2
Wine A
Wine B
Wine C



Boxplots (of the differences)
-0.2
0.0
0.2
0.4
Wine B - Wine A
Wine C - Wine A
Wine C - Wine B