THAKUR COLLEGE OF SCIENCE & COMMERCE

NAAC
Accredited
with Grade "A"
(3rd Cycle)

THAKUR
TRUSTS

ISO
9001 : 2015
Certified

Degree College

Computer Journal

CERTIFICATE

SEMESTER II **UID No.** 2020858

Class FYBSC CS **Roll No.** 1146 **Year** 2020-2021

This is to certify that the work entered in this journal is the work of Mst. / Ms. SHAIKH KAYSAN RAZAUDDIN

who has worked for the year 2020-2021 in the Computer Laboratory.

Teacher In-Charge

Head of Department

Date :

Examiner

Index

Practical No: 01

Aim: To demonstrate probability.

Q.1) A coin is tossed 10 times. What is the probability of getting exactly 6 heads?

x=pbinom(6,10,1/2)

x

Ans: [1] 0.2050781

Q.2) 60% of people who purchase sports cars are men. Of 10 sports car owners are randomly selected, find the probability that exactly 7 are men

x=dbinom(7,10,0.6)

x

Ans: [1] 0.2149908

Q.3) In a box of floppy dises it is known that 95% will work. A sample of three of the discs is selected at random. Find the probability that

(a) none (b) 1, (e) 2, (d) all 3 of the sample will work

x=dbinom(3,3,0.95)

x

y=dbinom(2,3,0.95)

y

z=dbinom(1,3,0.95)

z

p=dbinom(0,3,0.95)

p

Ans: a) [1] 0.000125

b) [1] 0.007125

c) [1] 0.135375

d) [1] 0.857375

Q.4) In a box of switches it is known 10% of the switches are faulty. A technician is wiring 30 circuits, each of which needs one switch. What is the probability that

a) all 30 work, b)at most 2 of the circuits do not work?

x=dbinom(30,30,0.9)

x

y=dbinom(28,30,0.9)

y

z=dbinom(29,30,0.9)

z

p=x+y+z

p

Ans: a) [1] 0.04239116

b) [1] 0.4113512

c) [1] 0.1875

d) [1] 0.4

Q.5) Find the probability of rolling exactly four even numbers in five rolls of a fair die. Find the probability of rolling exactly five even numbers in five rolls of a fair die. Hence find the probability of rolling four or more even numbers in five rolls of a fair die.

x=dbinom(4,5,0.5)

x

y=dbinom(5,5,0.5)

y

z=sum(dbinom(4,5,0.5)+dbinom(5,5,0.5))

z

Ans: a) [1] 0.15625

b) [1] 0.03125

c)[1] 0.1875

Q.6) Find eight random values from a sample of 150 With probability of 0.4.

x=rbinom(8,150,.4)

x

Ans: [1] 53 66 60 55 64 61 59 57

Q.7) For n=20 and, evaluate binomial probabilities and plot the graph of pmf and cdf. Before plotting round of the 3 decimals.

Ans: x=0:20

> x

[1] 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

> x=seq(1:20)

> x

[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

> y=dbinom(x,20,0.6)

> y

[1] 3.298535e-07 4.700412e-06 4.230371e-05 2.696862e-04 1.294494e-03

[6] 4.854351e-03 1.456305e-02 3.549744e-02 7.099488e-02 1.171416e-01

[11] 1.597385e-01 1.797058e-01 1.658823e-01 1.244117e-01 7.464702e-02

[16] 3.499079e-02 1.234969e-02 3.087423e-03 4.874878e-04 3.656158e-05

```
> z=round(y,3)
> z
 [1] 0.000 0.000 0.000 0.000 0.001 0.005 0.015 0.035 0.071 0.117 0.160 0.180
[13] 0.166 0.124 0.075 0.035 0.012 0.003 0.000 0.000
> y
 [1] 3.298535e-07 4.700412e-06 4.230371e-05 2.696862e-04 1.294494e-03
[6] 4.854351e-03 1.456305e-02 3.549744e-02 7.099488e-02 1.171416e-01
[11] 1.597385e-01 1.797058e-01 1.658823e-01 1.244117e-01 7.464702e-02
[16] 3.499079e-02 1.234969e-02 3.087423e-03 4.874878e-04 3.656158e-05
plot(x,z)
> plot(x,z,'l')
> plot(x,z,'h')
> plot(x,z,'h',col="green")
> plot(x,z,'h',col="green",main="plot of PMF")
> y1=pbinom(x,20,0.6)
> y1
 [1] 3.408486e-07 5.041261e-06 4.734497e-05 3.170311e-04 1.611525e-03
[6] 6.465875e-03 2.102893e-02 5.652637e-02 1.275212e-01 2.446628e-01
[11] 4.044013e-01 5.841071e-01 7.499893e-01 8.744010e-01 9.490480e-01
[16] 9.840388e-01 9.963885e-01 9.994760e-01 9.999634e-01 1.000000e+00
> z1=round(y1,3)
> z1
 [1] 0.000 0.000 0.000 0.000 0.002 0.006 0.021 0.057 0.128 0.245 0.404 0.584
[13] 0.750 0.874 0.949 0.984 0.996 0.999 1.000 1.000
> data=data.frame(x,z1)
```

> **Data:**

```
   x    z1
```

```
1   1 0.000
2   2 0.000
3   3 0.000
4   4 0.000
5   5 0.002
6   6 0.006
7   7 0.021
8   8 0.057
9   9 0.128
10 10 0.245
11 11 0.404
12 12 0.584
13 13 0.750
14 14 0.874
15 15 0.949
16 16 0.984
17 17 0.996
18 18 0.999
19 19 1.000
20 20 1.000
> plot(x,z1,"s")
> plot(x,z1,"s",main="cdf of data")
>  plot(x,z1,"s",main="cdf of data",col="yellow")
>  plot(x,z1,main="cdf of data",col="purple")
```

Q.8) Draw the random sample of size 10 from B(8,0.4) find mean and variance of the sample values.

```
n=8;p=0.4
x=rbinom(10,8,0.4)
```

x

m=mean(x)

m

v=var(x)

v

Ans: a)[1] 3.4

b)[1] 0.7734375

c)[1] 2.488889

Q.9) An biased coin is tossed 7 times . Calculate the probability of obtaining more head thentails.

x = pbinom(4,7,1/2)

x

Ans: 1/2

Q.10) At certain time one out of five telephone line is engaged in a conversation , what is probability that out of 10 telephones chosen at random only 2 are engaged.

n=10

probability=1/5 x=dbinom(2,10,0.2)

x

Ans: [1] 0.3019899

Q.11) It is known that during manufacturing , the probability of panel to be defective is 10%. Assume 18 solar panels are chosen find the probability of getting four of more sample to be defective with using

1)dbinom function

2) pbinom function

x=dbinom(4,18,0.1)

x

y=pbinom(4,18,0.1)

y

>1-sum(dbinom(0,18,0.1)+dbinom(1,18,0.1)+dbinom(2,18,0.1)

+dbinom(3,18,0.1))

Ans: [1] 0.09819684

2) > 1-pbinom(3,18,0.1)

Ans: [1] 0.09819684

Q.12) What are the 10th, 20th, and so forth quantiles of the bin(10, 1/3) distribution

x=qbinom(0.1, 10, 1/3)

x

Ans: [1] 0.7734375

y=qbinom(0.2, 10, 1/3)

y

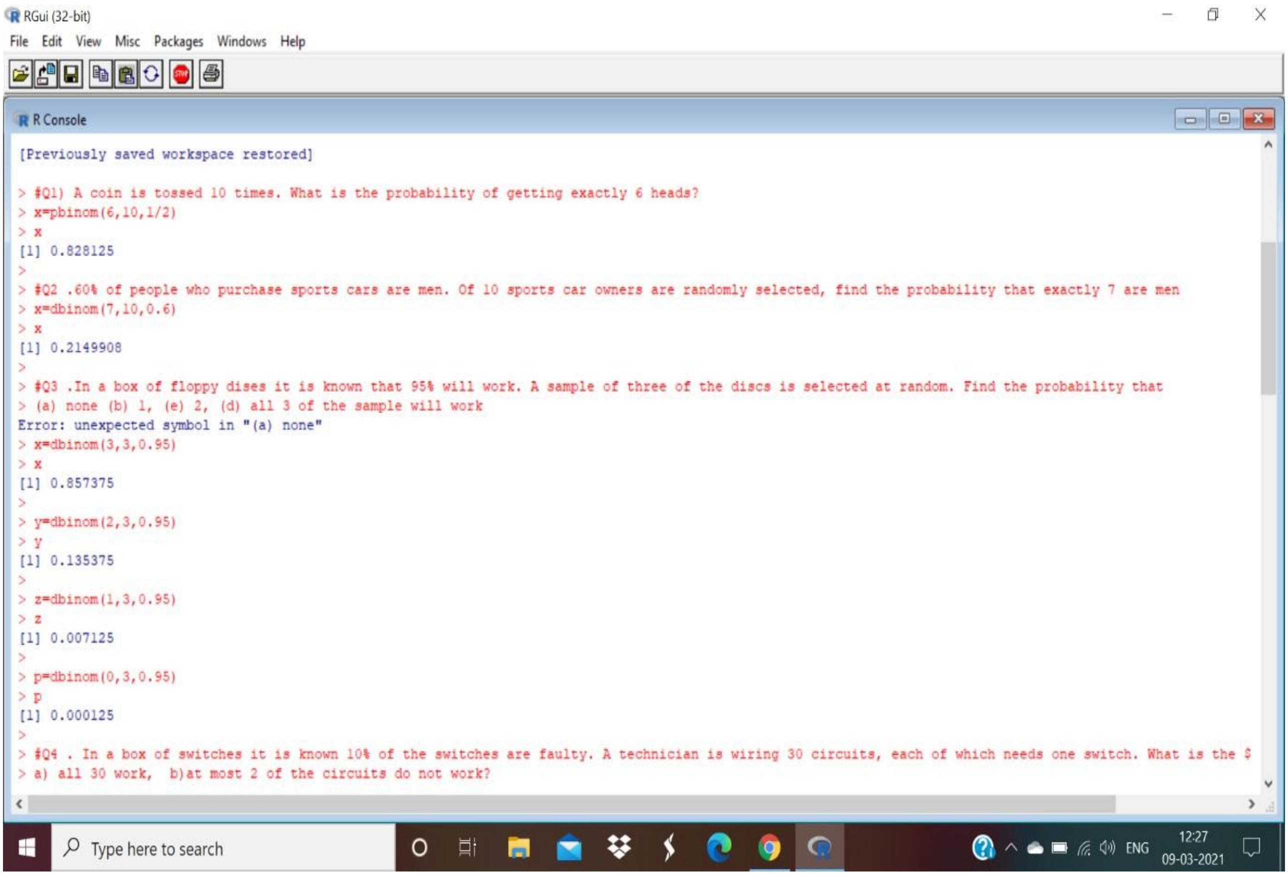
Ans [1] 0.9718061

and so forth, or all at once with z=qbinom(seq(0.1, 0.9, 0.1), 10

z

Ans: [1] 0.1875

Output:



RGui (32-bit)

File Edit View Misc Packages Windows Help

R Console

```
> #Q4 . In a box of switches it is known 10% of the switches are faulty. A technician is wiring 30 circuits, each of which needs one switch. What is the $
> a) all 30 work,  b)at most 2 of the circuits do not work?
Error: unexpected ')' in "a)"
> x=dbinom(30,30,0.9)
> x
[1] 0.04239116
>
> y=dbinom(28,30,0.9)
> y
[1] 0.2276562
> z=dbinom(29,30,0.9)
> z
[1] 0.1413039
> p=x+y+z
> p
[1] 0.4113512
>
> #Q5 .Find the probability of rolling exactly four even numbers in five rolls of a fair die. Find the probability of rolling exactly five even numbers i$
> x=dbinom(4,5,0.5)
> x
[1] 0.15625
>
> y=dbinom(5,5,0.5)
> y
[1] 0.03125
>
> z=sum(dbinom(4,5,0.5)+dbinom(5,5,0.5))
> z
[1] 0.1875
>
> #Q6 .Find eight random values from a sample of 150 With probability of 0.4.
> x=rbinom(8,150,.4)
> x
```

Type here to search   ENG 12:27 09-03-2021

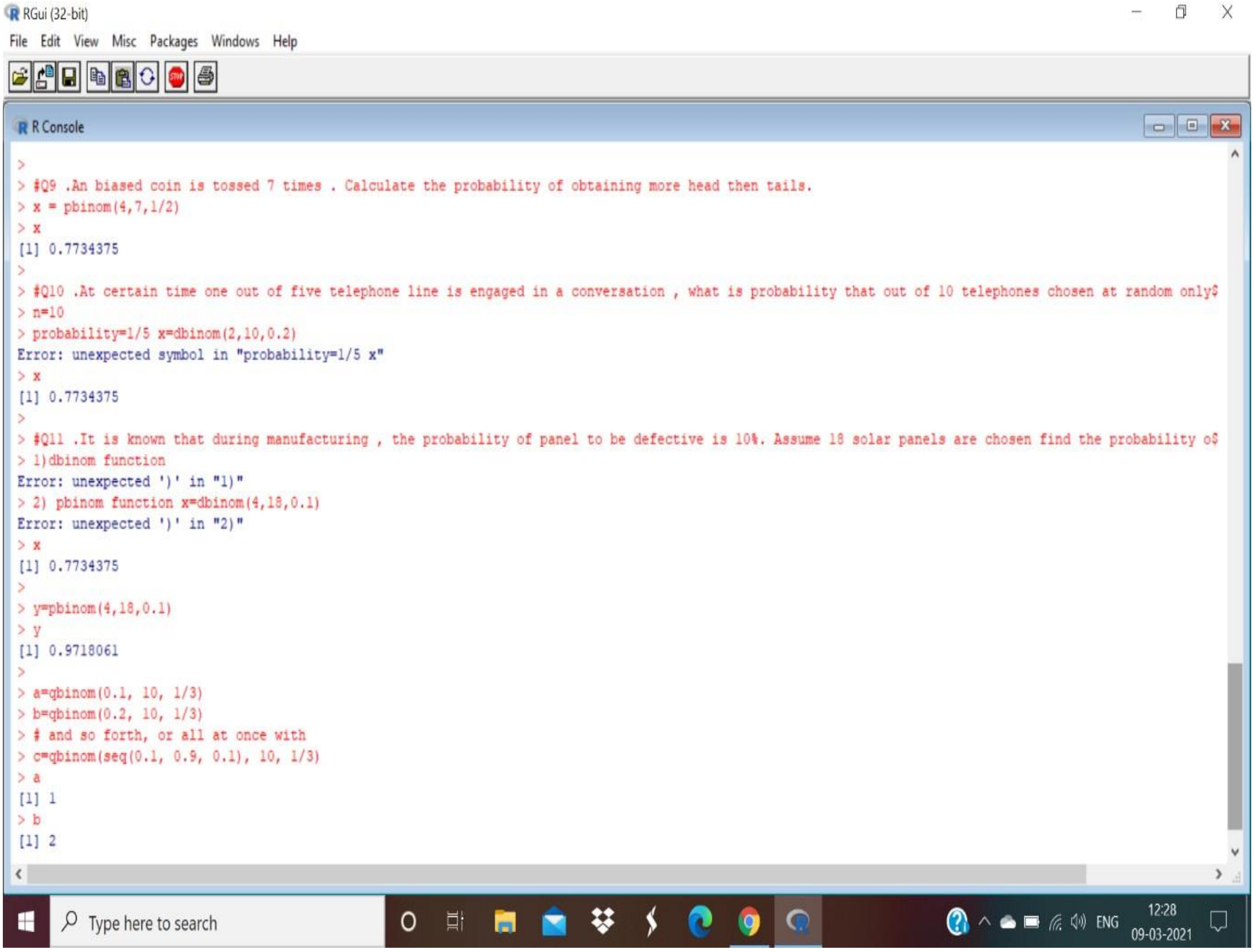
RGui (32-bit)

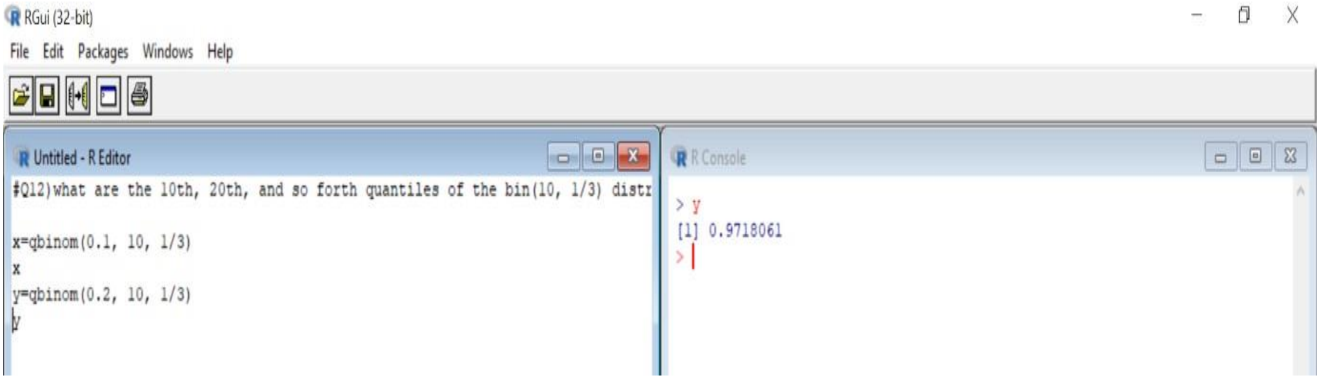
File Edit View Misc Packages Windows Help

R Console

```
> #Q6 .Find eight random values from a sample of 150 With probability of 0.4.
> x=rbinom(8,150,.4)
> x
[1] 56 70 61 63 56 65 59 61
>
> #Q7 .For n=20 and , evaluate binomial probabilities and plot the graph of pmf and cdf. Before plotting round of the 3 decimal.
>
>
> #Q8 .Draw the random sample of size 10 from B(8,0.4) find mean and variance of the sample values.
> n=8;p=0.4
> x=rbinom(10,8,0.4)
> x
 [1] 0 3 3 3 4 3 5 3 4 6
> m=mean(x)
> m
[1] 3.4
> v=var(x)
> v
[1] 2.488889
>
> #Q9 .An biased coin is tossed 7 times . Calculate the probability of obtaining more head then tails.
> x = pbinom(4,7,1/2)
> x
[1] 0.7734375
>
> #Q10 .At certain time one out of five telephone line is engaged in a conversation , what is probability that out of 10 telephones chosen at random only$
> n=10
> probability=1/5 x=dbinom(2,10,0.2)
Error: unexpected symbol in "probability=1/5 x"
> x
[1] 0.7734375
>
> #Q11 .It is known that during manufacturing , the probability of panel to be defective is 10%. Assume 18 solar panels are chosen find the probability o$
```

Type here to search   ENG 12:27 09-03-2021


RGui (32-bit)
File Edit View Misc Packages Windows Help
R Console
>
> #Q9 .An biased coin is tossed 7 times . Calculate the probability of obtaining more head then tails.
> x = pbinom(4,7,1/2)
> x
[1] 0.7734375
>
> #Q10 .At certain time one out of five telephone line is engaged in a conversation , what is probability that out of 10 telephones chosen at random only$
> n=10
> probability=1/5 x=dbinom(2,10,0.2)
Error: unexpected symbol in "probability=1/5 x"
> x
[1] 0.7734375
>
> #Q11 .It is known that during manufacturing , the probability of panel to be defective is 10%. Assume 18 solar panels are chosen find the probability o$
> 1)dbinom function
Error: unexpected ')' in "1)"
> 2) pbinom function x=dbinom(4,18,0.1)
Error: unexpected ')' in "2)"
> x
[1] 0.7734375
>
> y=pbinom(4,18,0.1)
> y
[1] 0.9718061
>
> a=qbinom(0.1, 10, 1/3)
> b=qbinom(0.2, 10, 1/3)
> # and so forth, or all at once with
> c=qbinom(seq(0.1, 0.9, 0.1), 10, 1/3)
> a
[1] 1
> b
[1] 2
Type here to search
ENG
12:28
09-03-2021


RGui (32-bit)
File Edit Packages Windows Help
Untitled - R Editor
#Q12)what are the 10th, 20th, and so forth quantiles of the bin(10, 1/3) distr
x=qbinom(0.1, 10, 1/3)
x
y=qbinom(0.2, 10, 1/3)
y
R Console
> y
[1] 0.9718061
>

Practical No: 02

Aim: Normal Distribution.

Q.1) Most graduate schools of business require applicants for admission to take the Graduate Management Admission Council's GMAT examination.Scores on the GMAT are roughly normally distributed with a mean of 527and a standard deviation of 112. What is the probability of an individual scoring above 500 on the GMAT?

ii)How high must an individual score on the GMAT in order to score in the highest 5%?

Ans: Mean=527s.d=112 x>500

i) >pnorm(500,527,112,lower.tail=F)

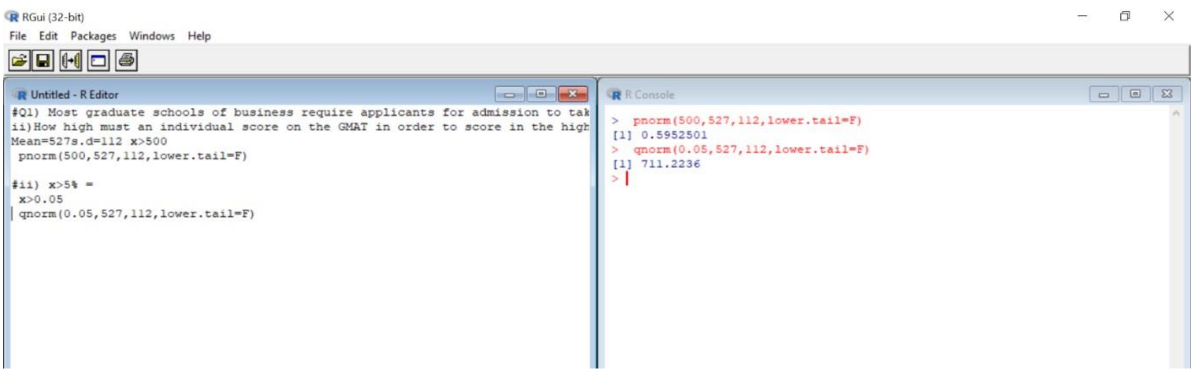
[1] 0.5952501

ii) x>5% =

> x>0.05
>qnorm(0.05,527,112,lower.tail=F)

[1] 711.2236

Output:



Q.2) The average number of acres burned by forest and range fires in a large New Mexico county is 4,300 acres per year, with a standard deviationof 750 acres. The distribution of the number of acres burned is normal.
What is the probability that between 2,500 and 4,200 acres will be burnedin any given year?

Ans: Mean=average=4300

s.d=750

2500<X<4200

X<2500:

> y=pnorm(2500,4300,750)

> y

[1] 0.008197536

X<4200:
> x=pnorm(4200,4300,750)-dnorm(4200,4300,750)

> x
[1] 0.4464377

2500<X<4200:

> x-y

[1] 0.4382401

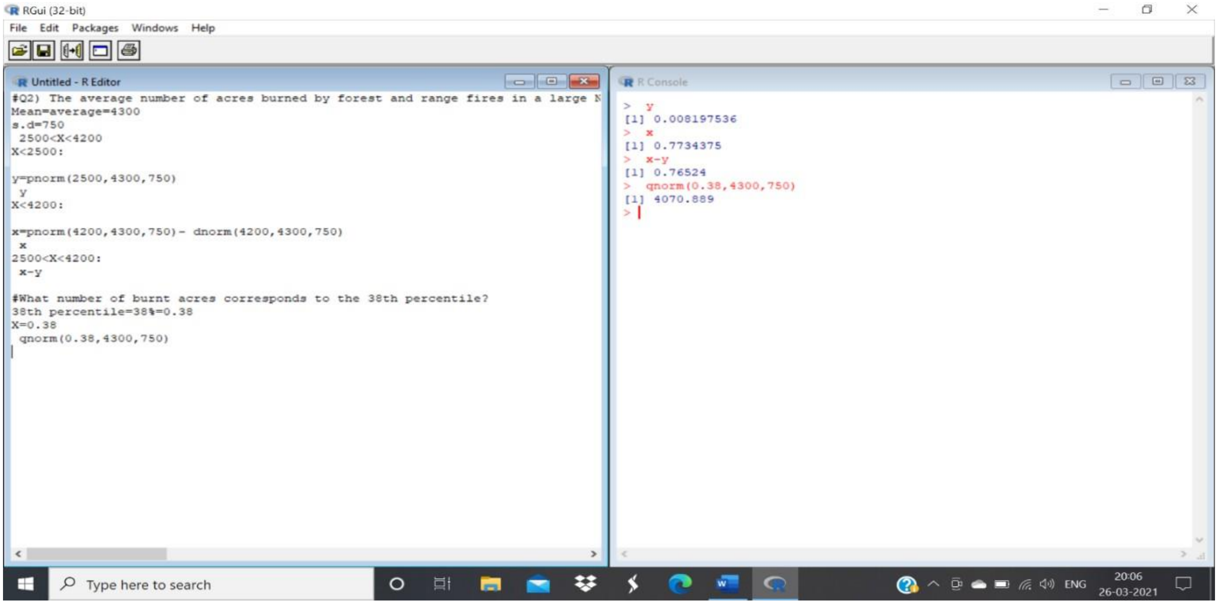
What number of burnt acres corresponds to the 38th percentile?

Ans: 38th percentile=38%=0.38

X=0.38
>qnorm(0.38,4300,750)

[1] 4070.889

Output:



Q.3) The Edwards's Theatre chain has studied its movie customers to determine how much money they spend on concessions. The study revealedthat the spending distribution is approximately normally distributed with amean of $4.11 and a standard deviation of $1.37. What percentage of customers will spend less than $3.00 on concessions?

What spending amount corresponds to the top 87th percentile?

Ans: Mean=40.11 s.d=1.37

i) X<3
>pnorm(3,4.11,1.37)

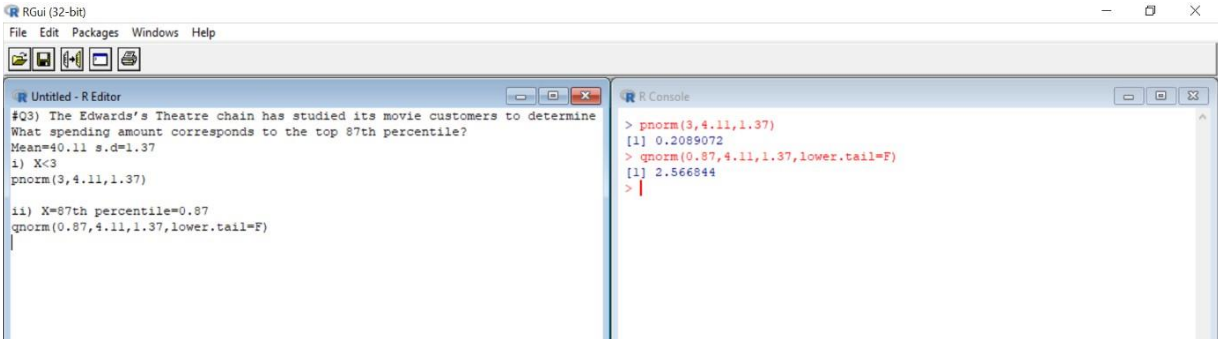
[1] 0.2089072

ii) X=87th percentile=0.87

>qnorm(0.87,4.11,1.37,lower.tail=F)

[1] 2.566844

Output:



Q.4) X is a normally distributed variable with mean $\mu = 30$ and standarddeviation $\sigma = 4$.

Find

a) $P(x < 40)$

b) $P(x > 21)$

c) $P(30 < x < 35)$

Ans: a) pnorm(40,30,4)-dnorm(40,30,4)[1]

0.9894083

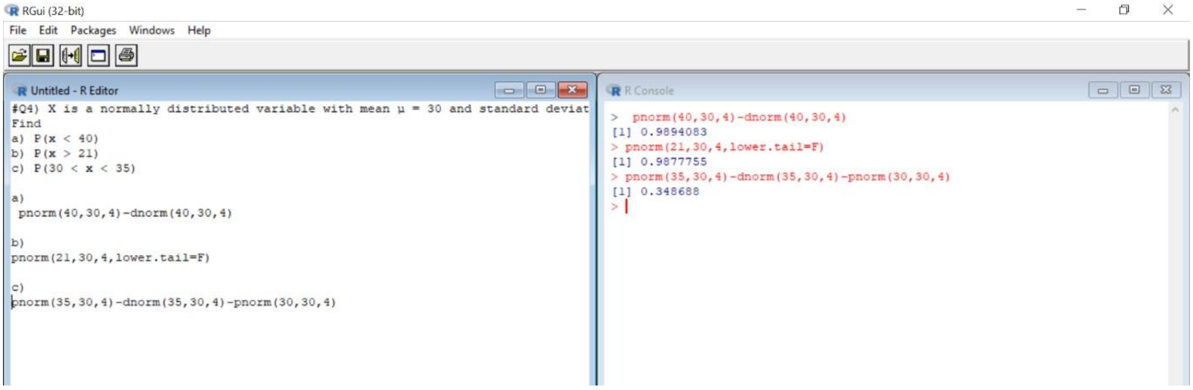
b) pnorm(21,30,4,lower.tail=F)[1]

0.9877755

c) pnorm(35,30,4)-dnorm(35,30,4)-pnorm(30,30,4)[1]

0.348688

Output:

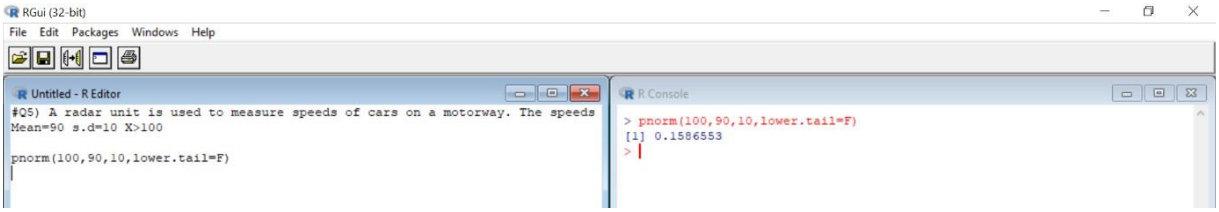


Q.5) A radar unit is used to measure speeds of cars on a motorway. The speeds are normally distributed with a mean of 90 km/hr and a standard deviation of 10 km/hr. What is the probability that a car picked at randomis travelling at more than 100 km/hr?

Ans: Mean=90 s.d=10 X>100
>pnorm(100,90,10,lower.tail=F)[1]

0.1586553

Output:

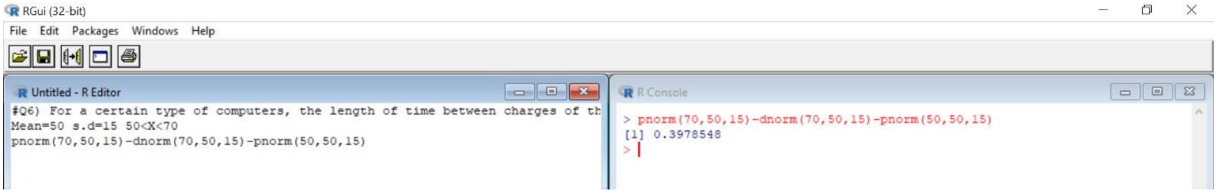


Q.6) For a certain type of computers, the length of time between charges of the battery is normally distributed with a mean of 50 hours and a standarddeviation of 15 hours. John owns one of these computers and wants to know the probability that the length of time will be between 50 and 70 hours ?

Ans: Mean=50 s.d=15 50<X<70
>pnorm(70,50,15)-dnorm(70,50,15)-pnorm(50,50,15)[1]

0.3978548

Output:



Q.7) Generate 10 random numbers from normal distribution with mean=12and standard deviation=4

Find mean of sample. Find standard deviation of sample.**Ans:** Mean=12

s.d=4

> a=rnorm(10,12,4)

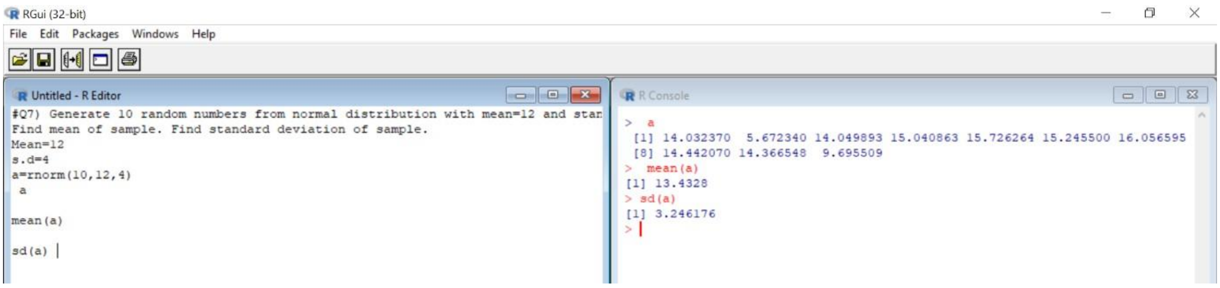
> a
[1] 9.272671 10.159778 8.067723 13.981327 14.903270 14.669195 15.819146
[8] 5.298671 7.179258 4.146990

> mean(a) [1] 10.3498

>sd(a)

[1] 4.254909

Output:



Q.8) Evaluate the probability for

***X* 50 100**

Ans: Mean=50variance=100 =>s.d=√variance = 10a)

>pnorm(70,50,10)

[1] 0.9772499

b) >pnorm(65,50,10,lower.tail=F)[1]

0.0668072

c) >pnorm(30,50,10)[1]

0.02275013

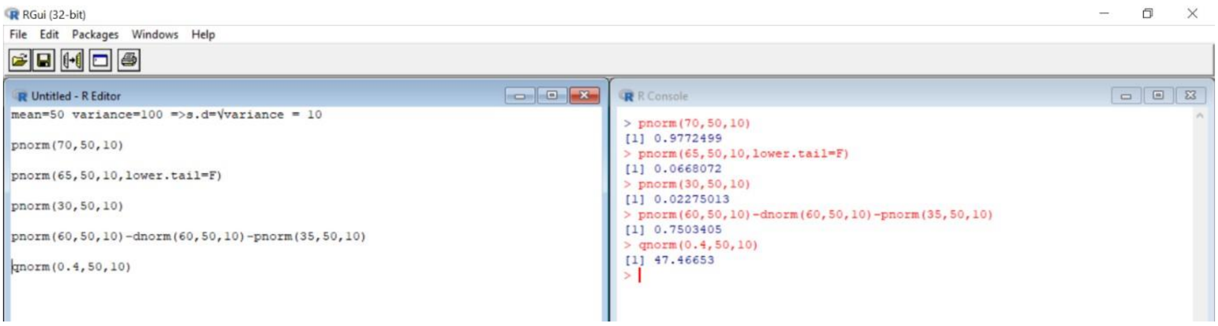
d) >pnorm(60,50,10)-dnorm(60,50,10)-pnorm(35,50,10)[1]

0.7503405

e) >qnorm(0.4,50,10)[1]

47.46653

<u>Output:</u>



Q.9) Generate 100 random numbers and evaluate its mean, median andvariance.

Ans: > x=rnorm(100)

> x
[1] 1.232222210 1.500724335 2.090438380 -0.149592228 -3.346933833
[6] 0.401185921 0.573559239 -2.086953499 0.123957777 0.067068317

[11] 0.021964374 0.770019513 -1.266550206 -0.660191227 0.330483566

[16] 1.082429774 0.144791408 -0.943401653 0.010352382 2.285993849

[21] 0.343372245 -1.319544951 -0.783149297 -1.139114121 -0.120975906

[26] 1.195447083 0.393645794 1.148945788 1.204047652 0.341601738

[31] -0.578237071 1.615074630 0.088088444 0.017289279 -0.455454435

[36] -0.254001133 0.195612798 -1.936312460 1.623942865 -1.850870135

[41] -0.245507817 -0.657558454 2.602248609 0.267342393 -1.033037633

[46] -0.554263489 0.282647990 -1.325653437 0.239224815 0.244266004

[51] 1.821649582 1.446218738 -0.904503376

[56] -0.825714637 0.221470758 1.125328757

[61] -1.758681027 -2.264325688 2.824811142

[66] 0.403954243 0.064935726
0.451119150 -0.651334164 1.901316242
[71] -0.524882078 -1.169000712
0.791408984 -0.589027622 -0.089571003

[76] 0.223716888 0.617878030
2.517468604 -0.300281266 -1.058761752
[81] -0.776891551 -0.579105018
0.061670553 -0.446324860 0.779235151

[86] 0.769324163 -0.008693308 -
0.664543679 -0.533860719 0.932028646
[91] 1.437132916 0.077423599 -
0.393383943 0.930456492 1.612780258

[96] -0.390976581 0.986148480
2.084471886 1.195879564 0.908026864

> y=pnorm(x,mean(x),sd(x))

> y
[1] 0.8272644930 0.8816300823
0.9563459946 0.3857133753 0.0008286165
[6] 0.5797756242 0.6387965226
0.0216649629 0.4815606532 0.4613482399

[11] 0.4453912387 0.7021786984
0.0988869566 0.2276950916 0.5549526526
[16] 0.7909293737 0.4889773354
0.1588107517 0.4412963449 0.9702443823

[21] 0.5594968197 0.0908985027
0.1959278357 0.1201728727 0.3955218255
[26] 0.8187393556 0.5771416224
0.8075892968 0.8207563604 0.5588730417

[31] 0.2503857434 0.9006538811
0.4688077084 0.4437418707 0.2865012506
[36] 0.3505843854 0.5070803607
0.0296476345 0.9020299899 0.0351711466

[41] 0.3533993673 0.2284055633
0.9848750845 0.5325975908 0.1402047557
[46] 0.2572429513 0.5380286954
0.0900093994 0.5226051291 0.5243979022

[51] 0.9291854921 0.8717081954
0.1673669476 0.6413971695 0.5694012003
[56] 0.1855896494 0.5162879400
0.8017692410 0.6075995356 0.3187285996

[61] 0.0420498967 0.0146694072
0.9909982323 0.7499065293 0.8940685943
[66] 0.5807417961 0.4605921566
0.5971234324 0.2300901734 0.9383284240

[71] 0.2657781959 0.1149122838
0.7087639012 0.2473312227 0.4063625433

[76] 0.5170874097 0.6535118282 0.9817403411 0.3353928155 0.1351502624
[81] 0.1974766348 0.2501393131 0.4594348076 0.2892822881 0.7050242544

[86] 0.7019634849 0.4345937709 0.2265232860 0.2631547649 0.7502721497
[91] 0.8699996859 0.4650215120 0.3056537517 0.7498260737 0.9002955670

[96] 0.3064078737 0.7653669993 0.9558509403 0.8188411192 0.7434162424

> mean(x)

[1] 0.1757367

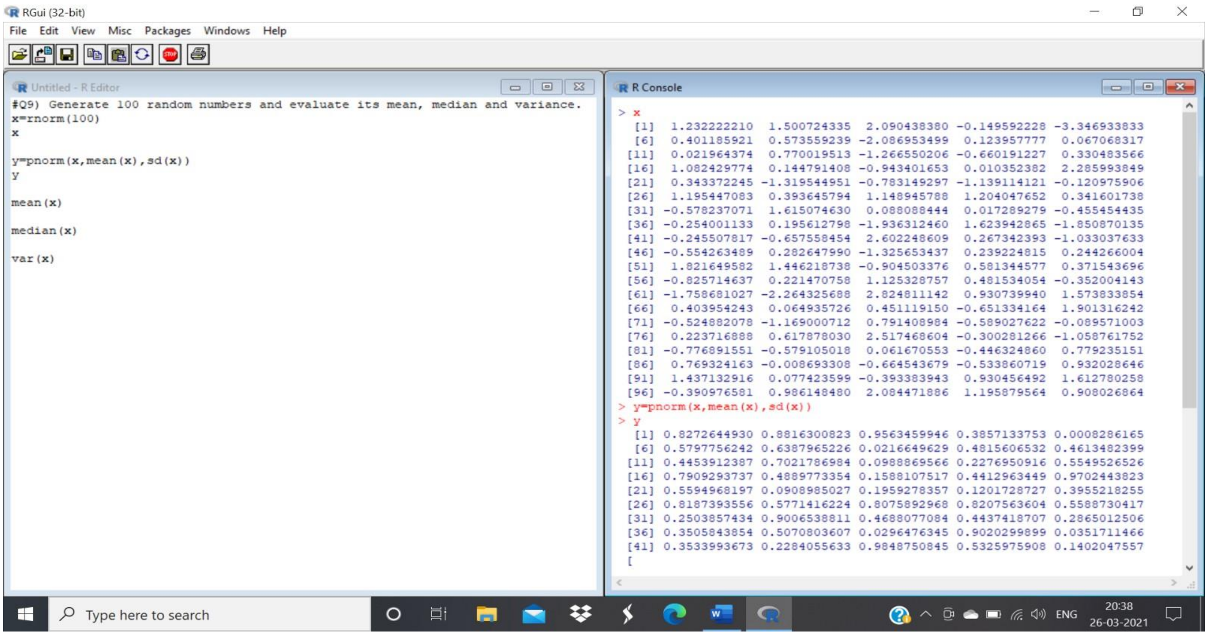
> median(x)

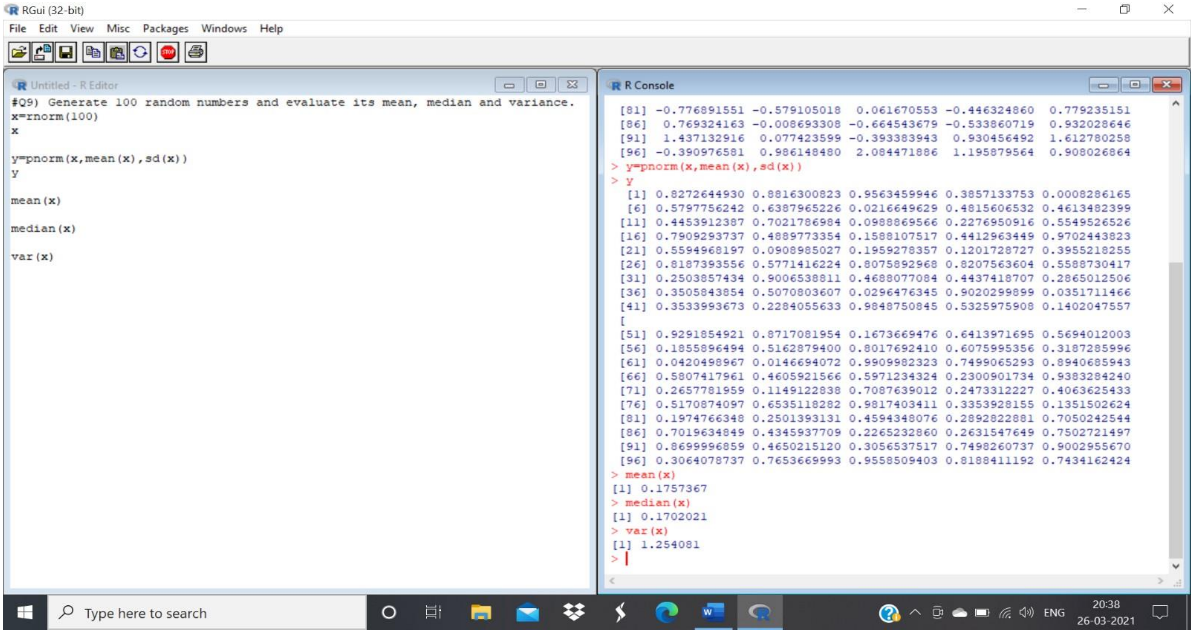
[1] 0.1702021

>var(x)

[1] 1.254081

<u>Output:</u>





Q.10) Plot a Standard normal distribution curve by taking a sequence starting from -4 and end at 4 with difference 0.1

Also plot cumulative distribution function for the same with proper labels.

Ans: > x=seq(-4,4,0.1)

> x

[1] -4.0 -3.9 -3.8 -3.7 -3.6 -3.5 -3.4 -3.3 -3.2 -3.1 -3.0 2.9 -2.8 -2.7 -2.6 [16] -2.5 -2.4 -2.3 -2.2 -2.1 -2.0 -1.9 -1.8 -1.7 -1.6 -1.5 1.4 -1.3 -1.2 -1.1[31] -1.0 -0.9 -0.8 -0.7 -0.6 -0.5 -0.4 -0.3 -0.2 -0.1 0.0 0.1 0.2 0.3 0.4 [46] 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6 1.7 1.8 1.9 [61] 2.0 2.1 2.2 2.3 2.4 2.5 2.6 2.7 2.8 2.9 3.0 3.1 3.2 3.3 3.4 [76] 3.5 3.6 3.7 3.8 3.9 4.0

> y=pnorm(x,mean(x),sd(x))

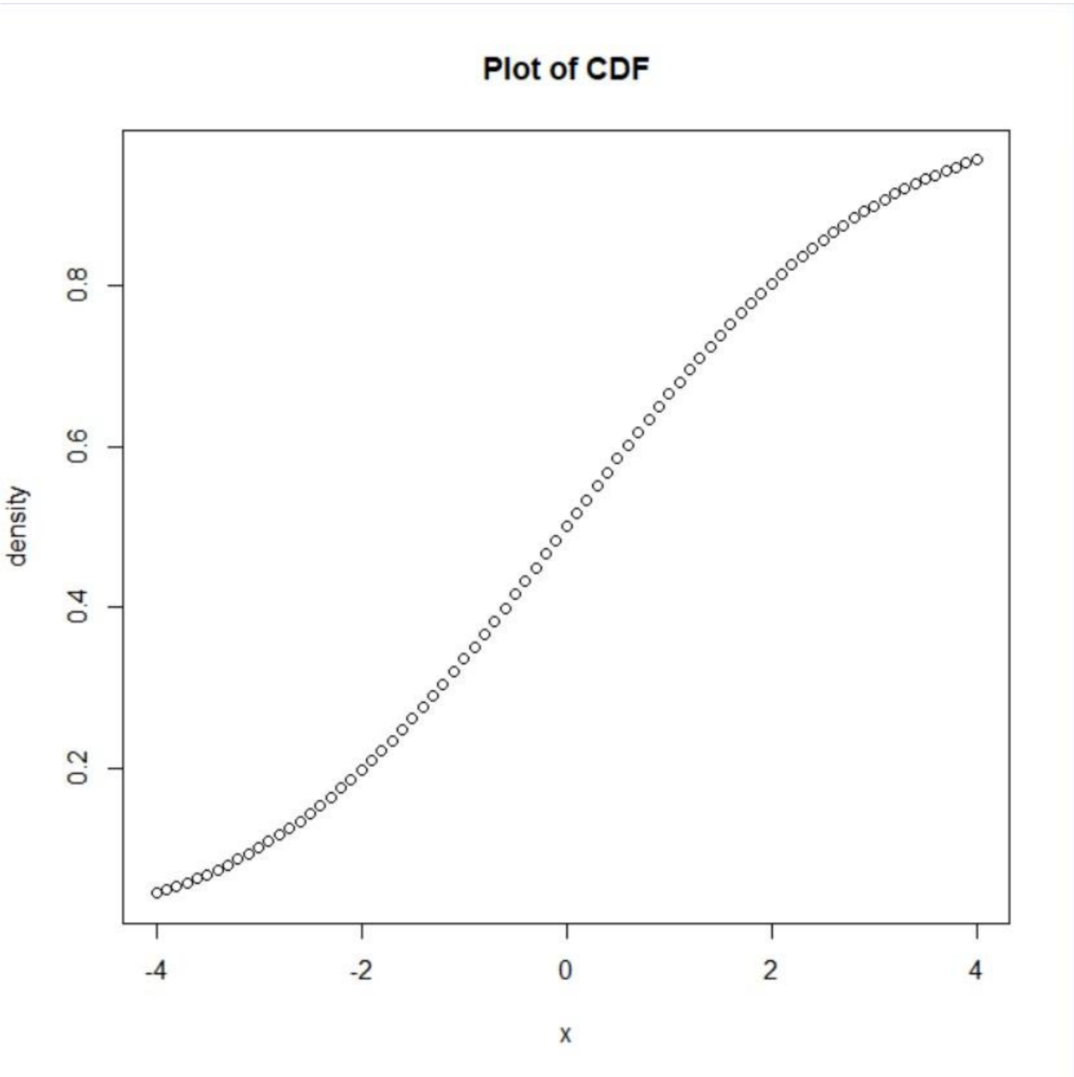
> y
[1] 0.04454623 0.04868907 0.05313431 0.05789540 0.06298558

0.06841778 [7] 0.07420453 0.08035782 0.08688907 0.09380899 0.10112747 0.10885350

[13] 0.11699506 0.12555900 0.13455099 0.14397537 0.15383513 0.16413174
[19] 0.17486516 0.18603373 0.19763411 0.20966126 0.22210837 0.23496685

```
[25] 0.24822633 0.26187463 0.27589781 0.29028015 0.30500423 0.32005098
[31] 0.33539970 0.35102820 0.36691286 0.38302874 0.39934970 0.41584852
[37] 0.43249702 0.44926627 0.46612663 0.48304802 0.50000000 0.51695198
[43] 0.53387337 0.55073373 0.56750298 0.58415148 0.60065030 0.61697126
[49] 0.63308714 0.64897180 0.66460030 0.67994902 0.69499577 0.70971985
[55] 0.72410219 0.73812537 0.75177367 0.76503315 0.77789163 0.79033874
[61] 0.80236589 0.81396627 0.82513484 0.83586826 0.84616487 0.85602463
[67] 0.86544901 0.87444100 0.88300494 0.89114650 0.89887253 0.90619101
[73] 0.91311093 0.91964218 0.92579547 0.93158222 0.93701442 0.94210460
[79] 0.94686569 0.95131093 0.95545377
> plot(x,y,main="Plot of CDF",ylab="density")
```

Output:



Practical No: 03

Aim: To demonstrate Large Sample test.

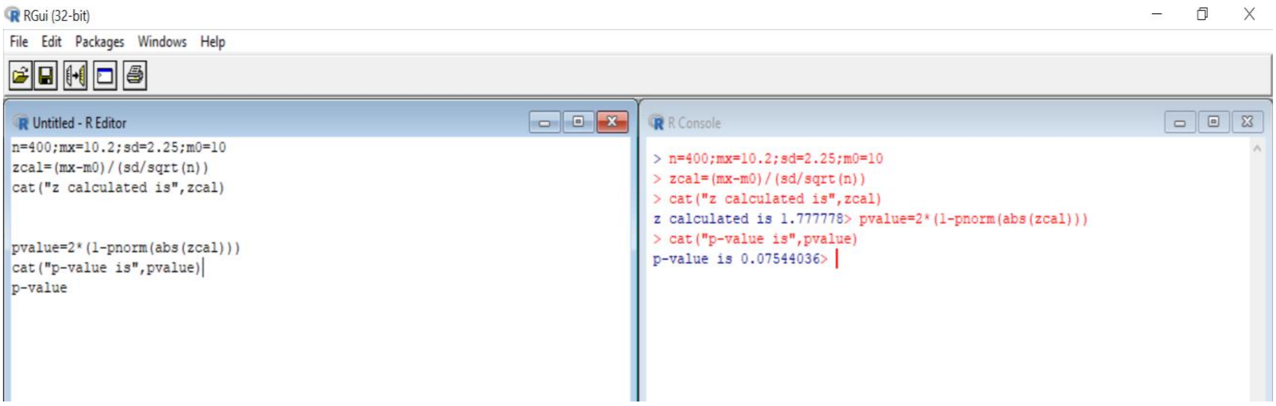
Example1:

Test the hypothesis H0: μ = 10 against H1: μ ≠ 10. A random sample of size 400 is drawn gives mean 10.2 and standard deviation 2.25. Use LOS = 5%

Solution:

```
> n=400;mx=10.2;sd=2.25;m0=10
> zcal=(mx-m0)/(sd/sqrt(n))
> cat("z calculated is",zcal)
z calculated is 1.777778
> pvalue=2*(1-pnorm(abs(zcal)))
> cat("p-value is",pvalue)
p-value is 0.07544036>
```

Output:



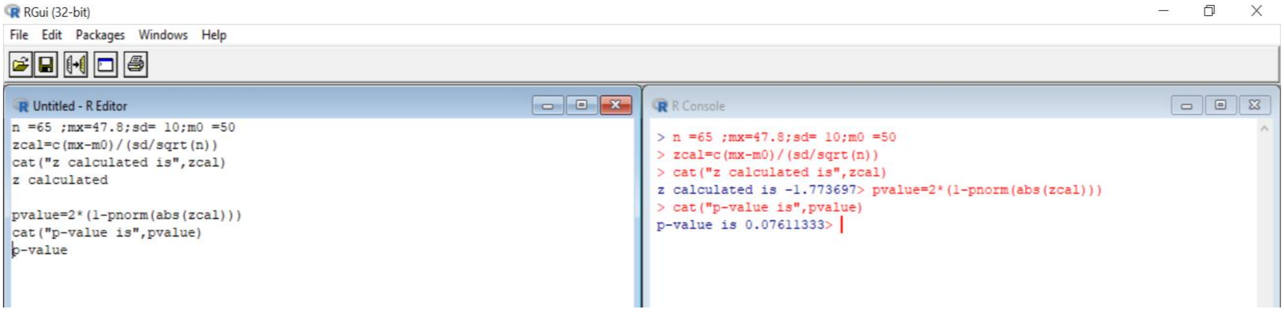
Example 2:

Test the hypothesis H0: μ ≥ 50 against H1: μ < 50. A random sample of size 65 is drawn gives mean 47.8 and standard deviation 10. Use LOS = 5%.

Solution:

```
n =65 ;mx=47.8;sd= 10;m0 =50
> zcal=(mx-m0)/(sd/sqrt(n))
> cat("z calculated is",zcal)
z calculated is -1.773697> pvalue=2*(1-pnorm(abs(zcal)))
> cat("p-value is",pvalue)
p-value is 0.07611333>
```

Output:



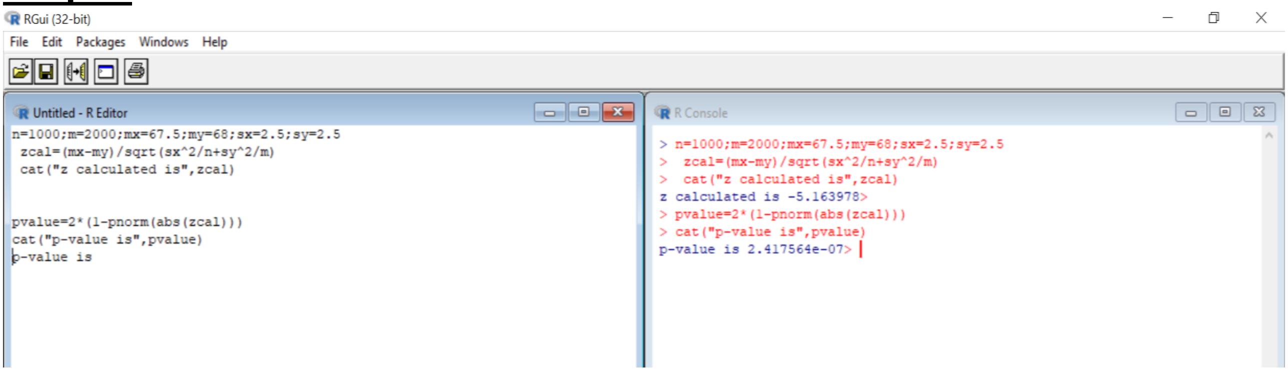
Example 3:

Two random samples of sizes 1000 and 2000 are drawn from two populations with same standard deviation 2.5 gives means 67.5 and 68 respectively. Test the hypothesis H0: µ1 = µ2 against H1: µ1 ≠ µ2. Use 5% LOS.

Solution:

n=1000;m=2000;mx=67.5;my=68;sx=2.5;sy=2.5
> zcal=(mx-my)/sqrt(sx^2/n+sy^2/m)
> cat("z calculated is",zcal)
z calculated is -5.163978> pvalue=2*(1-pnorm(abs(zcal)))
> cat("p-value is",pvalue)
p-value is 2.417564e-07
p-value is approximately zero
Conclusion: Since p-value<LOS,therefore reject H0 at 5% LOS

Output:



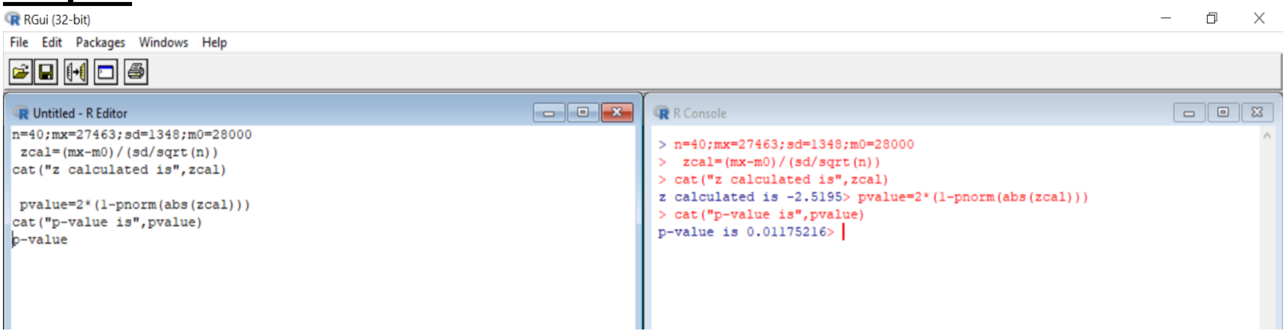
Example 4:

A trucking firm suspects that that the average lifetime of 28,000 miles claimed for certain tyre is too high. To check this claim the firm puts 40 of these tyres on trucks and gets a mean lifetime of 27,563 miles and a standard deviation of 1,348 miles. What will the trucking firm conclude at 0.01 level of significance if it tests the null hypothesis μ = 28,000 miles against an appropriate alternative? Assume Normal distribution. Find 'p' value and interpret the value. H0: μ = 28000 against H1: μ < 28000,

Solution:

n=40;mx=27463;sd=1348;m0=28000
> zcal=(mx-m0)/(sd/sqrt(n))
> cat("z calculated is",zcal)
z calculated is -2.5195> pvalue=2*(1-pnorm(abs(zcal)))
> cat("p-value is",pvalue)
p-value is 0.01175216>
Reject H0
There is sufficient evidence to doubt the trucking firm's claim

<u>Output:</u>



Example 5:

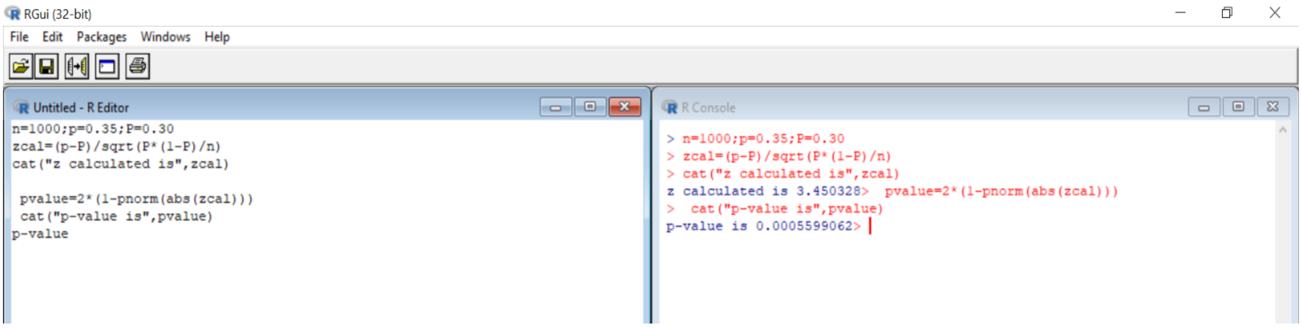
Out of 1000 residents in a certain area 350 were found to be earthquake affected. Can we accept the claim that there are less than 30% earthquake affected residents? Use 5% L.O.S.

Solution:

> n=1000;p=0.35;P=0.30
> zcal=(p-P)/sqrt(P*(1-P)/n)
> cat("z calculated is",zcal)
z calculated is 3.450328> pvalue=2*(1-pnorm(abs(zcal)))
> cat("p-value is",pvalue)

p-value is 0.0005599062
5% L.O.S ie 0.05
Since p-value<LOS, therefore reject H0 at 5% L.O.S.

Output:



Example 6:

From each of two consignments of apples, a sample of size 200 is drawn, and number of rotten apples counted. Test whether the proportion of rotten apples in the two consignments are significantly different?

	Sample Size	No. of rotten apples
Consignment A	200	44
Consignment B	200	30

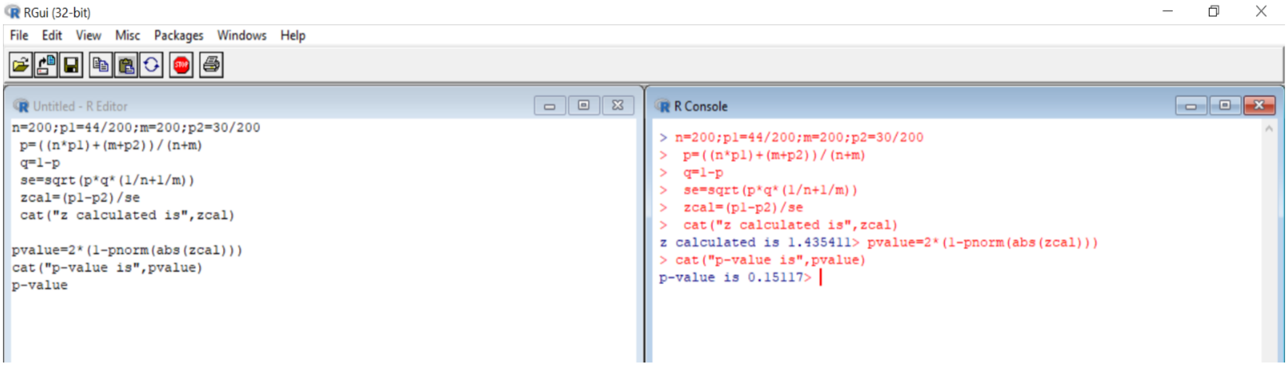
N1=200, p1=44/200, n2=200, p2=30/200, alpha=0.05
H0:P1 = P2 against H1: P1 ≠ P2

Solution:

```
> n=200;p1=44/200;m=200;p2=30/200
> p=((n*p1)+(m+p2))/(n+m)
> q=1-p
> se=sqrt(p*q*(1/n+1/m))
> zcal=(p1-p2)/se
> cat("z calculated is",zcal)
z calculated is 1.435411
> pvalue=2*(1-pnorm(abs(zcal)))
> cat("p-value is",pvalue)
p-value is 0.15117
```

Since p-value>LOS, therefore do not reject H0 at 5% LOS

Output:



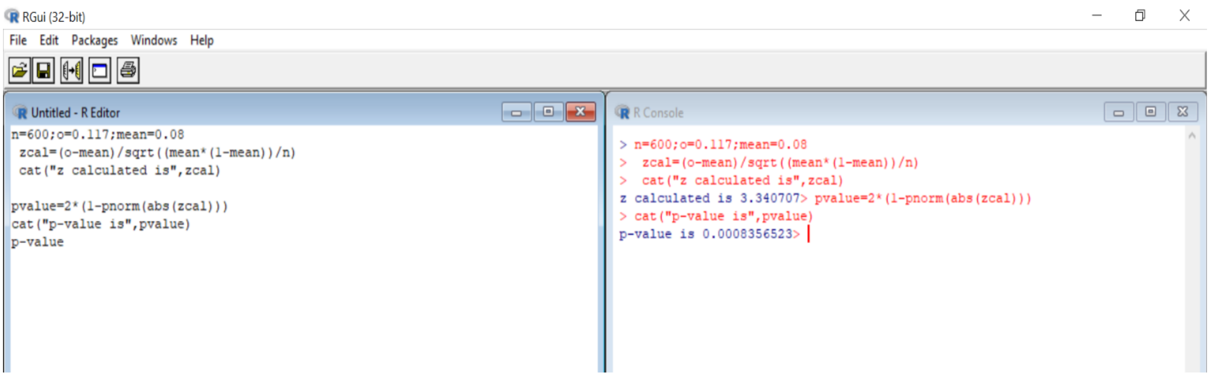
Example 7:

In a battery factory, 8% of all batteries made are assumed to be defective. Technical trouble with production line, however, has raised concern percent defective has increased in past few weeks. Of n = 600 batteries chosen at random, 70 600 ths 70 600 ≈ 0.117 of them are found to be defective. Does data support concern about defective batteries at α = 0.05

Solution:

n=600;o=0.117;mean=0.08
> zcal=(o-mean)/sqrt((mean*(1-mean))/n)
> cat("z calculated is",zcal)
z calculated is 3.340707> pvalue=2*(1-pnorm(abs(zcal)))
> cat("p-value is",pvalue)
p-value is 0.0008356523
Since p-value> LOS
Reject H0 at 5% LOS

Output:



Practical No: 04

Aim: To demonstrate t-test.

Q.1) Write R command for the following data to test the hypothesis (i) H0: μ = 3400 against H1: μ ≠
3400 (ii) H0: μ = 3400 against H1: μ < 3400 (iii) H0: μ = 3400 against H1: μ > 3400 at
5% L.O.S. 3366, 3337, 3361, 3410, 3316, 3357, 3348, 3356, 3376, 3382, 3377, 3355, 3408, 3401, 3390, 3424, 3383, 3374, 3384, 3390

Output:

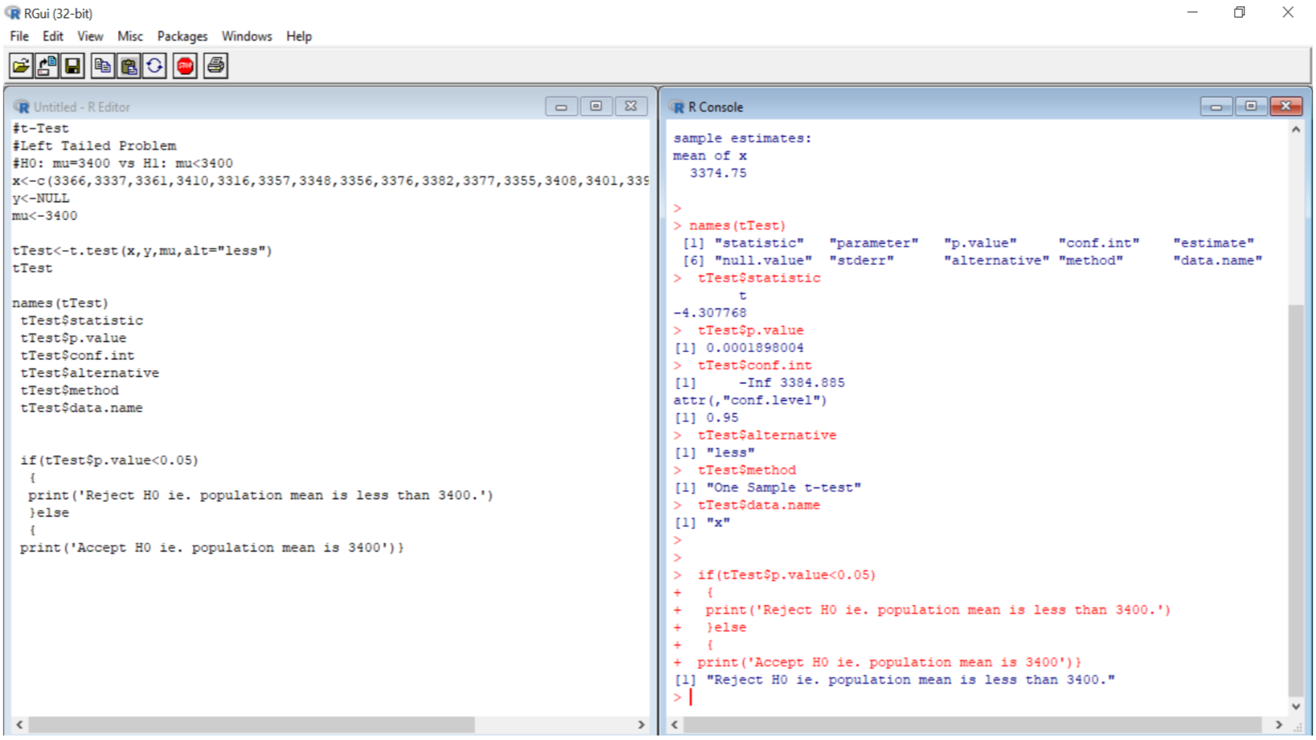
```
#t-Test
> #Left Tailed Problem
> #H0: mu=3400 vs H1: mu<3400
> x<-
c(3366,3337,3361,3410,3316,3357,3348,3356,3376,3382,3377,3355,3408,3401,3
390,3424,3383,3374,3384,3390)
> y<-NULL
> mu<-3400
>
> tTest<-t.test(x,y,mu,alt="less")
> tTest

        One Sample t-test

Data:  x
t = -4.3078, df = 19, p-value = 0.0001898

alternative hypothesis: true mean is less than

3400 95 percent confidence interval:

    -Inf 3384.885

sample estimates:

mean of x
  3374.75

>
> names(tTest)
```

```
[1] "statistic"   "parameter"   "p.value"     "conf.int"    "estimate"    "null.value"
[7] "stderr"      "alternative" "method"      "data.name"
```

```
> tTest$statistic
        t

-4.307768  >

tTest$parameter

df

19
> tTest$p.value
[1] 0.0001898004
> tTest$conf.int

[1]     -Inf

3384.885

attr(,"conf.level")

[1] 0.95 >

tTest$estimate

mean of x

  3374.75  >

tTest$null.value

mean

3400
> tTest$alternative
[1] "less"
> tTest$method
[1] "One Sample t-test"
> tTest$data.name
[1] "x"
>
> if(tTest$p.value<0.05)
```

```
+ {
+ print('Reject H0 ie. population mean is less than 3400.')
+ }else
+ {
+ print('Accept H0 ie. population mean is 3400')
+ }
[1] "Reject H0 ie. population mean is less than 3400."
```

Output:



Q.2) Below are given the gain in weights (in lbs) of pigs fed on two diets A and B Gain in weight Diet A: 25, 32, 30, 43, 24, 14, 32, 24, 31, 31, 35, 25 Diet B: 44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21, 35, 29, 22 Test if the two diets differ significantly as regards their effect on increase in weight. Use L.O.S. 5%

Output:

```
#t-Test for double mean
#Two Tailed Problem
#H0: No significant difference between means x and y vs
#H1: Significant difference between means of x and y
x<-c(25, 32, 30, 43, 24, 14, 32, 24, 31, 31, 35, 25 )
y<-c(44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21, 35,29, 22)
```

```
ttest<-t.test(x,y,var.equal=T)
tTest

 if(tTest$p.value<0.05)
{
print('Reject H0 ie. there is significant difference b /n means.')
}else
{
print('Accept H0 ie. there is no significant difference b /n means.')}
```

One Sample t-test

Data: x
t = -4.3078, df = 19, p-value = 0.0001898

alternative hypothesis: true mean is less than

3400 95 percent confidence interval:

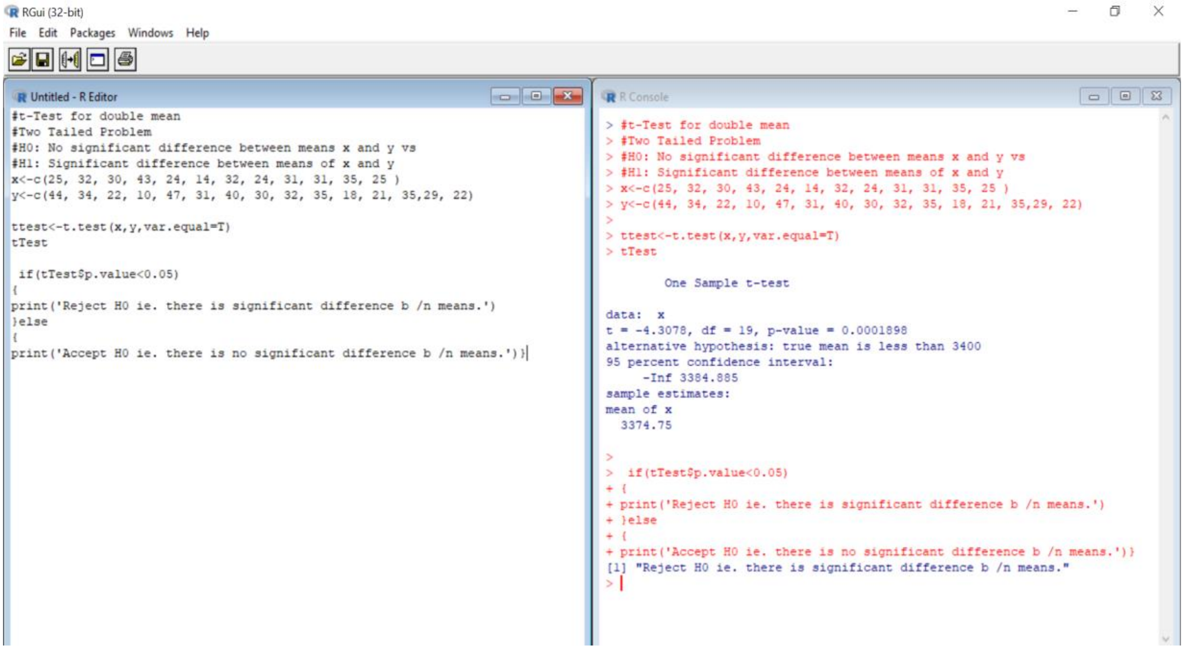
-Inf 3384.885

sample estimates:

mean of x
3374.75

[1] "Reject H0 ie. there is significant difference b /n means."

Output:



Q.3) Eleven school boys were given attest in mathematics. They were given a month's tuition and a second test was held at the end of it. Do the marks give evidence that the student's have benefited

by the coaching? Use LOS 1%. Marks in test 1: 23, 20, 19, 21, 18, 20, 18, 17, 23, 16, 19 Marks in test
2: 24, 19, 22, 18, 20, 22, 20, 20, 23, 20, 17

Output:

```
#Paired t-Test
#Two Tailed Problem
#H0: No significant difference between x and y vs
#H1: Significant difference between x and y
x<-c(23, 20, 19, 21, 18, 20, 18, 17, 23, 16, 19 )
y<-c(24, 19, 22, 18, 20, 22, 20, 20, 23, 20, 17)
tTest<-t.test(x,y,paired=T)
tTest
```

Paired t-test

Data: x and y
t = -1.4832, df = 10, p-value = 0.1688

alternative hypothesis: true difference in means is not

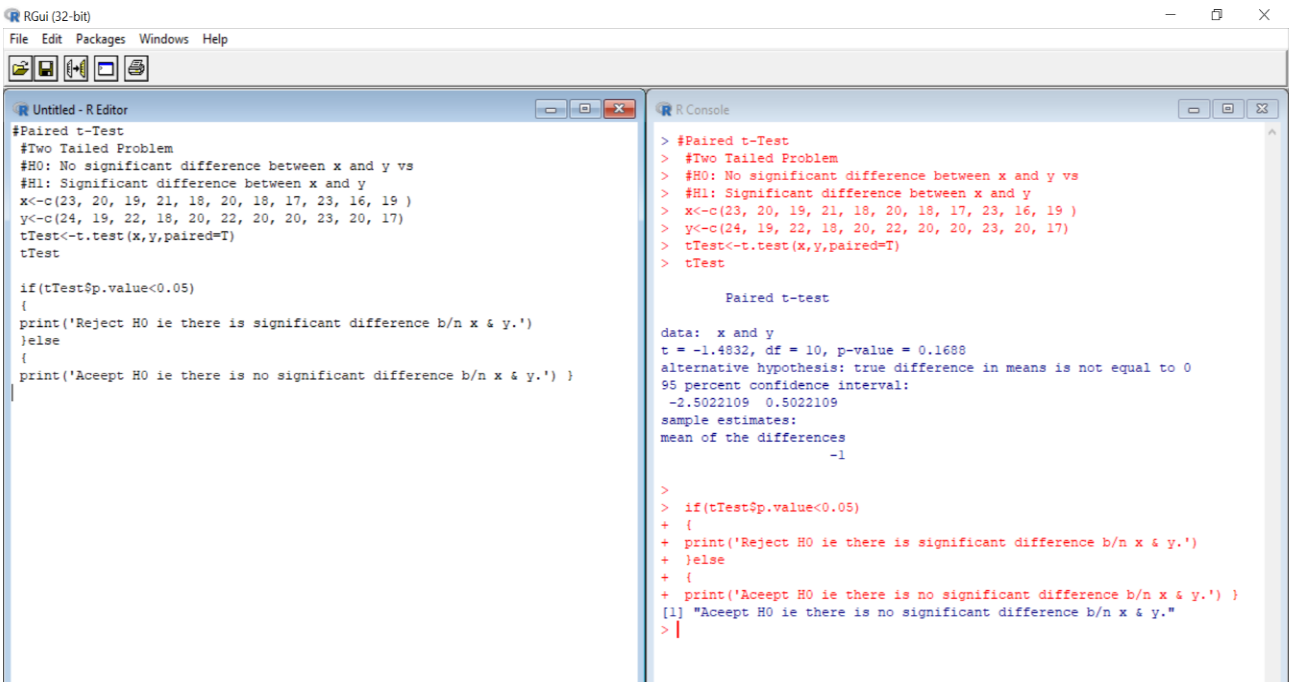
equal to 0 95 percent confidence interval: -2.5022109

0.5022109 sample estimates: mean of the differences

-1

```
>
> if(tTest$p.value<0.05)
+ {
+ print('Reject H0 ie there is significant difference b/n x & y.')
+ }else
+ {
+ print('Aceept H0 ie there is no significant difference b/n x & y.') }
[1] "Aceept H0 ie there is no significant difference b/n x & y."
```

Output:



Q.4) Twelve cars were equipped with radial tires and driven over a test course. Then the same 12 cars
(with the same drivers) were equipped with regular belted tires and driven over the same course.
After each run, the cars' gas economy (in km/l) was measured. Is there evidence that radial
tires produce better fuel economy? (Assume normality of data, and use □ = .05.)

Gas	1	2	3	4	5	6	7	8	9	10	11	12

Radial	4.2	4.7	6.6	7	6.7	4.5	5.7	6	7.4	4.9	6.1	5.2
Belted	4.1	4.9	6.2	6.9	6.8	4.4	5.7	5.8	6.9	4.7	6	4.9

Output:

```
#Paired t-Test
#Two Tailed Problem
#H0: Cars equipped with radial tires give better fuel company
#H1: Cars equipped with belted tires give better fuel company

x<-c(4.2,4.7,6.6,7,6.7,4.5,5.7,6,7.4,4.9,6.1,5.2)
y<-c(4.1,4.9,6.2,6.9,6.8,4.4,5.7,5.8,6.9,4.7,6,4.9)

tTest<-t.test(x,y,paired=T)
tTest
      Paired t-test
```

Data: x and y
t = 2.4845, df = 11, p-value = 0.03033
alternative hypothesis: true difference in means is not

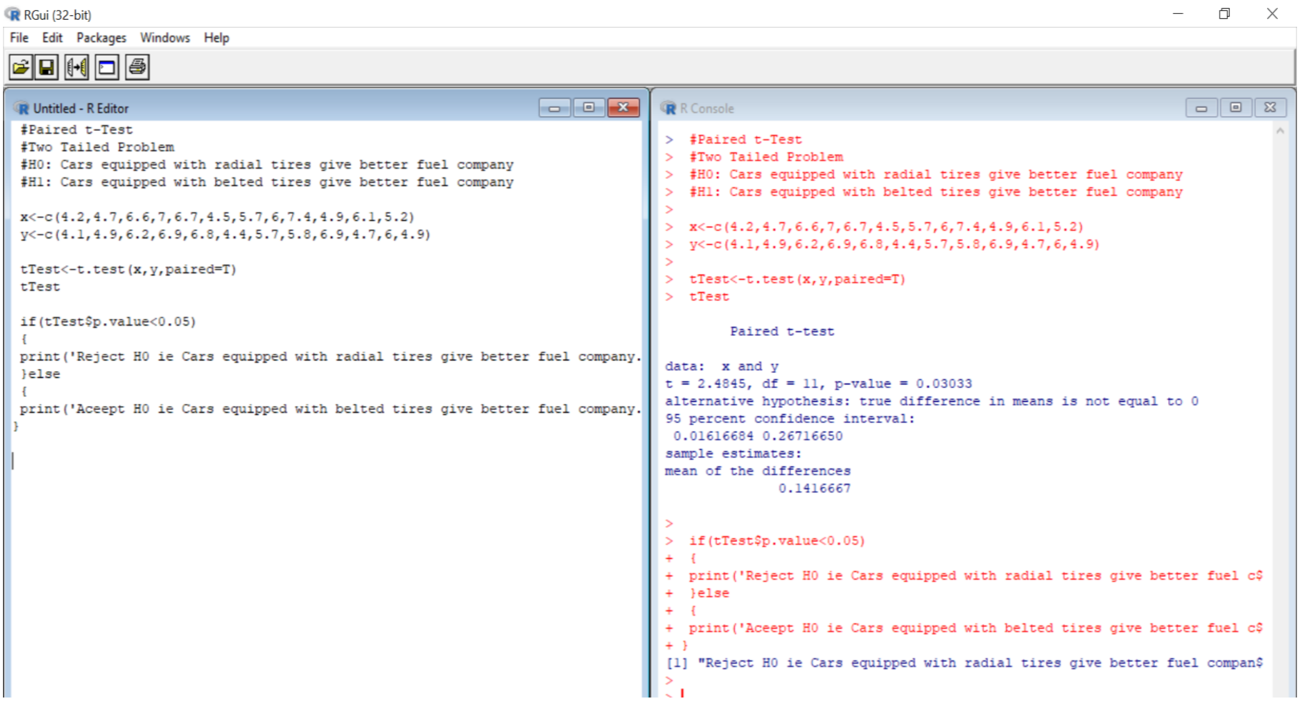
equal to 0 95 percent confidence interval: 0.01616684

0.26716650 sample estimates: mean of the differences

0.1416667

```
if(tTest$p.value<0.05)
{
print('Reject H0 ie Cars equipped with radial tires give better fuel company.')
}else
{
print('Aceept H0 ie Cars equipped with belted tires give better fuel company.')
}
[1] "Reject H0 ie Cars equipped with radial tires give better fuel company."
```

Output:



Q.5) To test the hypothesis that eating fish makes one smarter, a random sample of 12 persons take a fish oil supplement for one year and then are given an IQ test. Here are the results: 116 111 101 120 99 94 106 115 107 101 110 92

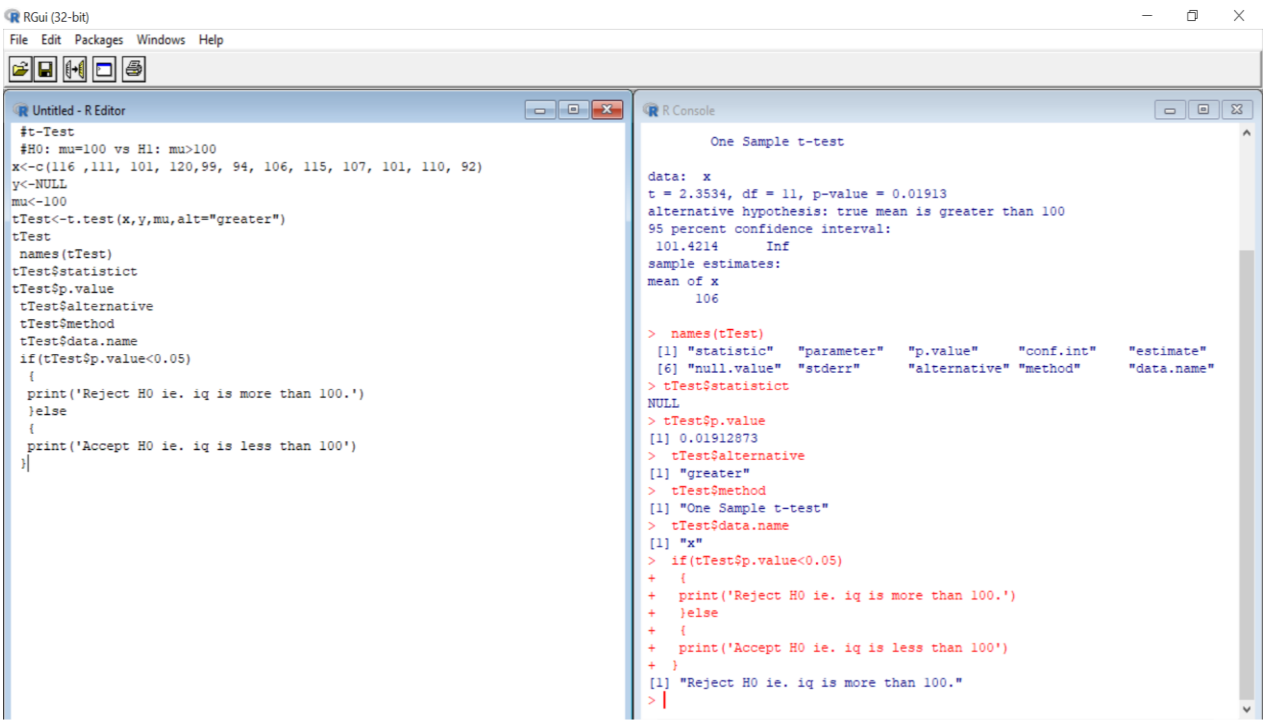
Output:

```
#t-Test

#H0: mu=100 vs H1: mu>100
x<-c(116 ,111, 101, 120,99, 94, 106, 115, 107, 101, 110, 92)

y<-NULL
mu<-100

tTest<-t.test(x,y,mu,alt="greater")
tTest


        One Sample t-test


data:  x
t = 2.3534, df = 11, p-value = 0.01913
alternative hypothesis: true mean is greater than
```

```
100 95 percent confidence interval:

 101.4214     Inf

sample estimates:

mean of x
      106


> names(tTest)
 [1] "statistic"   "parameter"   "p.value"     "conf.int"    "estimate"    "null.value"
 [7] "stderr"      "alternative" "method"      "data.name"
>
> tTest$statistic
      t
2.353394
> tTest$parameter

df

11
> tTest$p.value
[1] 0.01912873
> tTest$conf.int

[1] 101.4214

Inf

attr(,"conf.level")

[1] 0.95 >

tTest$estimate

mean of x



 106
> tTest$null.value

mean

 100
> tTest$alternative
[1] "greater"
> tTest$method
```

```
[1] "One Sample t-test"
> tTest$data.name
[1] "x"
>
> if(tTest$p.value<0.05)
+  {
+  print('Reject H0 ie. iq is more than 100.')
+  }else
+  {
+  print('Accept H0 ie. iq is less than 100')
+  }
[1] "Reject H0 ie. iq is more than 100."
```

Output:



Q.6) The water diet requires you to drink 2 cups of water every half hour from when you get up until
you go to bed but eat anything you want. Four adult volunteers agreed to test this diet. They are
weighed prior to beginning the diet and 6 weeks after. Their weights in pounds are

Person	1	2	3	4	mean	s.d
Weight	180	125	240	150	173.75	49.56

before						
Weight after	170	130	215	152	166.75	36.09
Difference	10	-5	25	-2	7	13.64

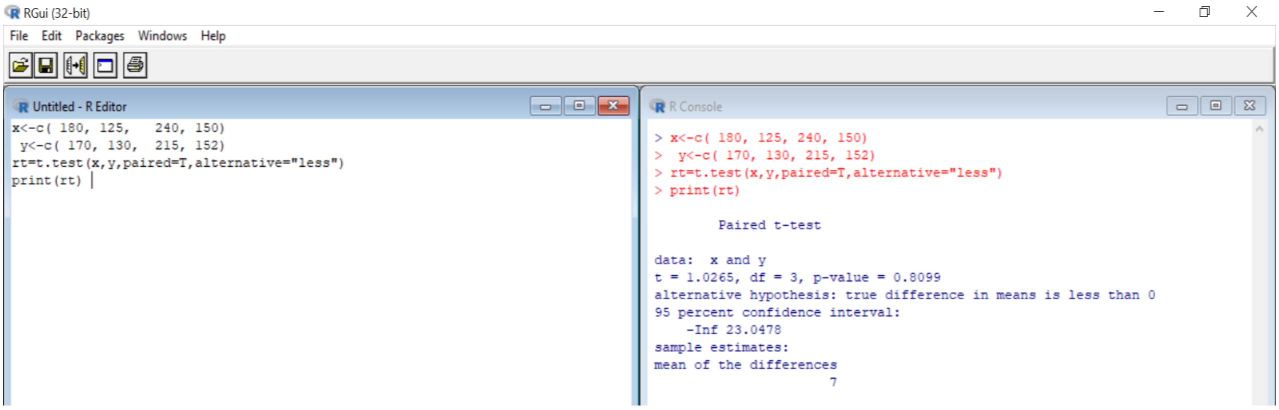
Solution:

```
x<-c( 180, 125,     240, 150)

 y<-c( 170, 130,     215, 152)

rt=t.test(x,y,paired=T,alternative="less")

print(rt)
```

Output:

```
> x<-c( 180,125,240,150)
> y<-c( 170,130,215,152)
> rt=t.test(x,y,paired=T,alternative="less")
> print(rt)


        Paired t-test



Data: x and y
t = 1.0265, df = 3, p-value = 0.8099
alternative hypothesis: true difference in means is less

than 0 95 percent confidence interval:

   -Inf 23.0478

sample estimates:

mean of the

differences

                  7
```

Output:



Q.7) Two different alloys are being considered for making lead-free solder used in the wave soldering process for printed circuit boards. A crucial characteristic of solder is its melting point, which is known to follow a Normal distribution. A study was conducted using a random sample of 21 pieces of solder made from each of the two alloys. In each sample, the temperature at which each of the 21 pieces melted was determined. The mean and standard deviation of the sample for Alloy 1 were x1 =218.9ºC and s1 = 2.7ºC; for Alloy 2 the results were x2 = 215.5ºC and s2 = 3.6ºC. If we were to test H0: µ1 = µ2 against Ha: µ1 ≠ µ2 , what would be the value of the test statistic?

Output:

```
> x1=rnorm(21,218.9,2.7)
> x1
 [1] 218.5438 218.5582 219.8610 214.9482 213.9162 219.0880 221.9167

217.4941  [9] 216.4643 220.5143 218.4316 225.2999 215.8468 218.1905

225.2801 223.1773

[17] 223.3622 217.0137 219.5962 218.5765 218.0556
> x2=rnorm(21,215.5,3.6)
> x2
 [1] 218.5797 214.3995 213.9013 216.2542 219.0522 210.0699 213.3623

217.2148  [9] 220.6031 211.7275 218.4122 210.2430 215.1707 215.6227

218.9657 220.0053

[17] 218.0041 212.2484 216.5479 214.8700 215.4511
>
> #Paired t-Test
```

```
> #Two Tailed Problem
> #H0: mu1 equal to mu2
> #H1: mu1 not equal to mu2
>
>
> tTest<-t.test(x1,x2,paired=T)
> tTest
```

Paired t-test

Data: x1 and x2
t = 3.4241, df = 20, p-value = 0.002688
alternative hypothesis: true difference in means is not

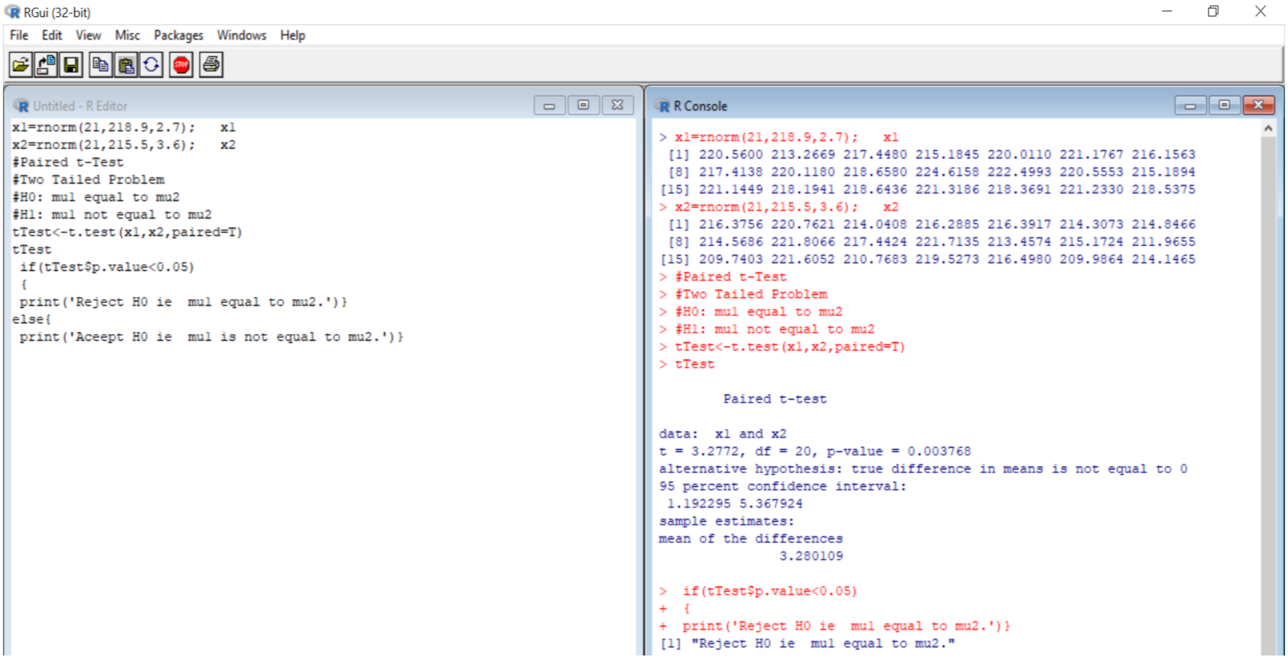
equal to 0 95 percent confidence interval: 1.366499

5.626816 sample estimates: mean of the differences

3.496657

```
>
> if(tTest$p.value<0.05)
+ {
+ print('Reject H0 ie  mu1 equal to mu2.')
+ }else
+ {
+ print('Aceept H0 ie  mu1 is not equal to mu2.')
+ }
[1] "Reject H0 ie  mu1 equal to mu2."
```

Output:



Q.8) A U.S. magazine, Consumer Reports, carried out a survey of the calorie and sodium content of a number of different brands of hotdog. There were two types of hotdog: beef,
'meat' (mainly pork and beef but can contain up to 15% poultry) and poultry. The results below are the calorie content of the different brands of beef and poultry hotdogs.
Beef hotdogs: 186, 181, 176, 149, 184, 190, 158, 139, 175, 148, 152, 111, 141, 153, 190, 157, 131,149, 135, 132
Poultry hotdogs: 129, 132, 102, 106, 94, 102, 87, 99, 170, 113, 135, 142, 86, 143, 152, 146, 144.
Is, there is strong evidence that the calorie content of poultry hotdogs is lower than the calorie content of beef hotdog

Output:

beef<-c(186, 181, 176, 149, 184, 190, 158, 139, 175, 148, 152, 111, 141, 153, 190, 157, 131,
149, 135, 132)
poultry<-c(129, 132, 102, 106, 94, 102, 87, 99, 170, 113, 135, 142, 86, 143, 152,

146, 144) t.test(beef,poultry)

t.test(beef,poultry,alt='greater',var.equal=T)

Output:

> beef<-c(186, 181, 176, 149, 184, 190, 158, 139, 175, 148, 152, 111, 141, 153, 190, 157, 131,
+ 149, 135, 132)
> poultry<-c(129, 132, 102, 106, 94, 102, 87, 99, 170, 113, 135, 142, 86, 143,

152, 146, 144) > t.test(beef,poultry)

Welch Two Sample t-test

Data: beef and poultry
t = 4.3031, df = 32.394, p-value = 0.0001455
alternative hypothesis: true difference in means is not equal

to 0 95 percent confidence interval:

18.11306

50.64577 sample

estimates: mean of

x mean of y

156.8500 122.4706

> t.test(beef,poultry,alt='greater',var.equal=T)

Two Sample t-test

Data: beef and poultry
t = 4.3455, df = 35, p-value = 5.683e-05
alternative hypothesis: true difference in means is greater

than 0 95 percent confidence interval:

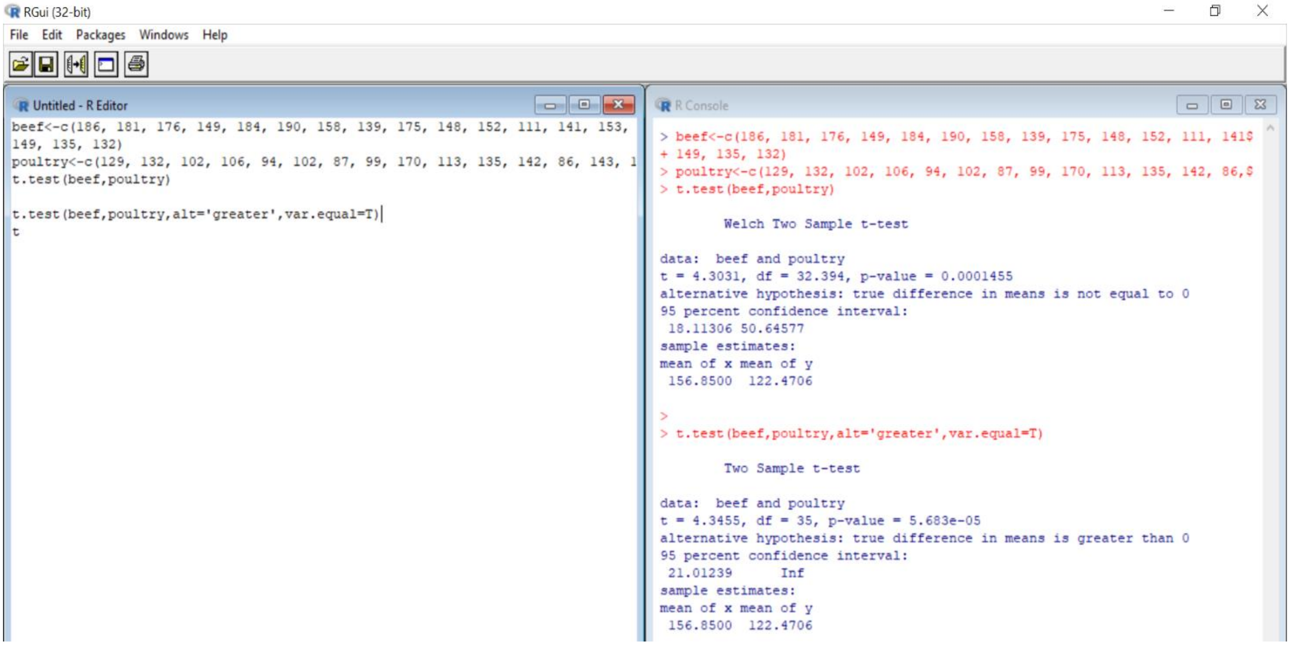
21.01239 Inf

sample estimates:

mean of x mean of y

156.8500 122.4706

Output:



Practical No: 05

Aim: To demonstrate f-test.

Q.1) Two random samples are drawn from two normal

populationsSample-I 20,16,26,27,22,23,18,24,19,25

Sample-II 27,33,42,35,32,34,38,28,41,43,30,47. Obtains the estimated variance of the population and test at 5% level of significance whether the two populations have the same variance.

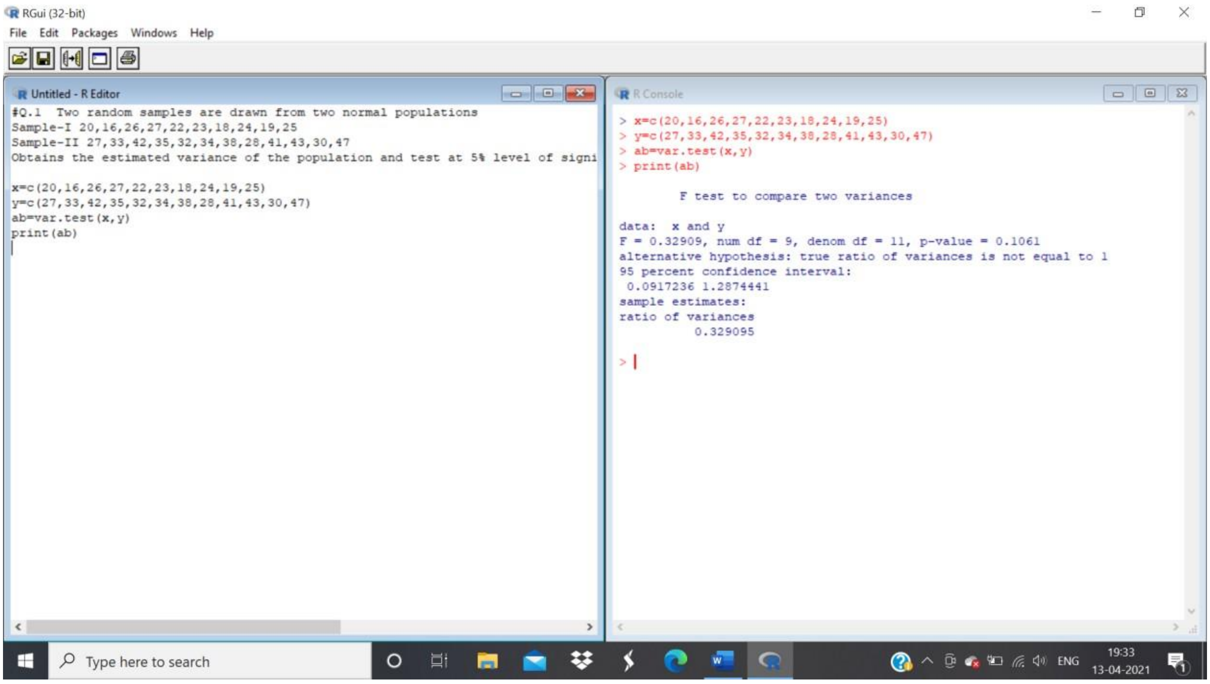
x=c(20,16,26,27,22,23,18,24,19,25)

y=c(27,33,42,35,32,34,38,28,41,43,30,47)

ab=var.test(x,y)

print(ab)

Output:



Ans: The p-value of F-test is p=0.1.61 which is greater than the alpha level 0.05. In conclusion, there is no difference between the two samples.

Q.2) The following data refer to yield of wheat in quintals on equal area in two agricultureblock A and B

	No. of Plots	Mean	Variance
Block A	8	60	50
Block B	6	51	40

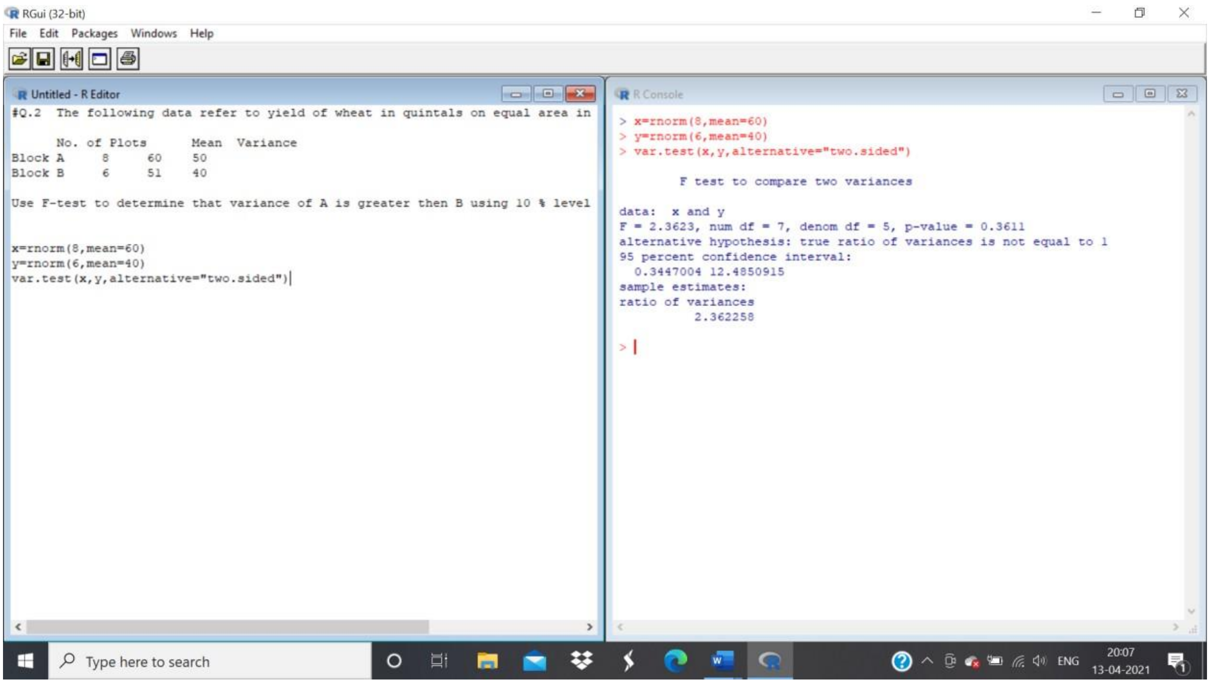
Use F-test to determine that variance of A is greater then B using 10 % level of significance.

x=rnorm(8,mean=60)

y=rnorm(6,mean=40)

var.test(x,y,alternative="two.sided")

Output:



Ans: The p-value of F-test is 0.5418 is less than the level of significance

Q.3**)** The following are price (in Rs) of the commodity in the sample of the shops selected atrandomfrom different cities

City A 74.10,77.70,75.35, 74.00,73.80, 79.30,75.80,76.80,77.10,76.40

CityB 70.80,74.90,76.20,72.80,78.10,74.70,69.80,81.20

Is it reasonable to say variances of the price in both cities are same(use F-test) . Is it reasonable to say average price in both cities are same(use T-test)

x=c(74.10,77.70,75.35, 74.00,73.80, 79.30,75.80,76.80)

y=c(70.80,74.90,76.20,72.80,78.10,74.70,69.80,81.20)

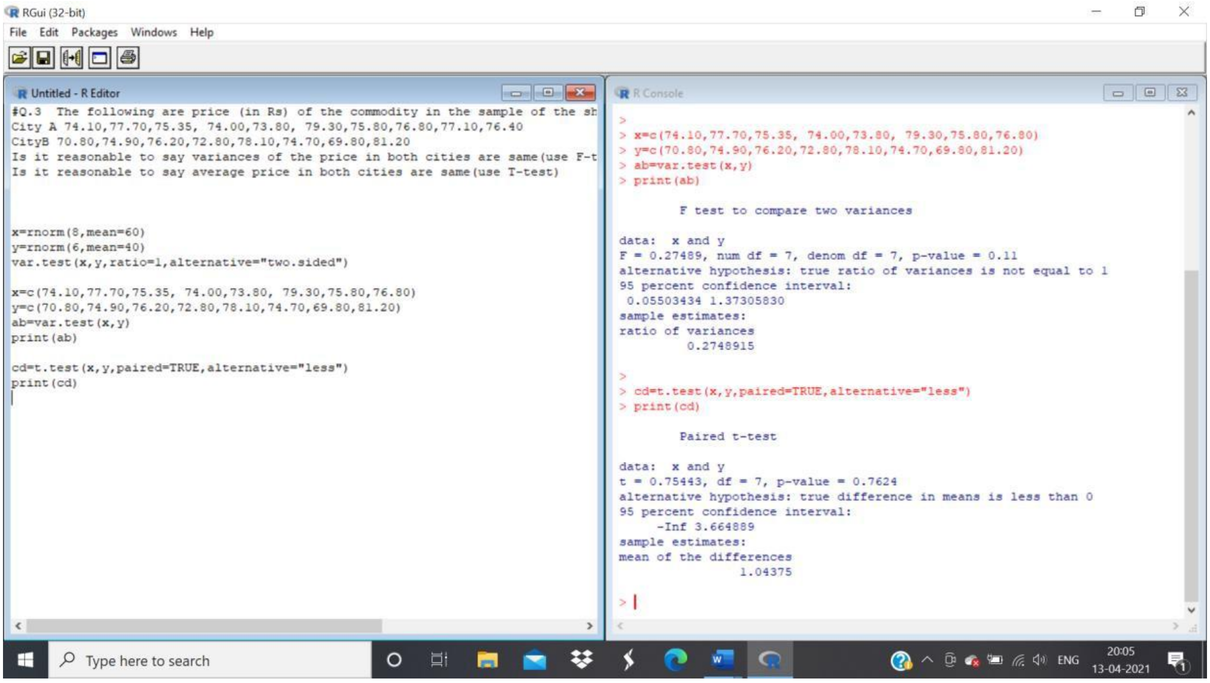
ab=var.test(x,y)

print(ab)

cd=t.test(x,y,paired=TRUE,alternative="less")

print(cd)

<u>Output:</u>



Ans: The p-value of F-Test 0.11 is greater than 0.05, we can conclude that the two variance of the price in both cities are same

The p-value of T-test i=0.7624 is greater than 0.05, we can conclude that the average price in both the cities are same

Q.4) y1=45, 87, 123, 120, 70)

y2= 51, 71, 42, 37, 51, 78, 51, 49, 56, 47, 58

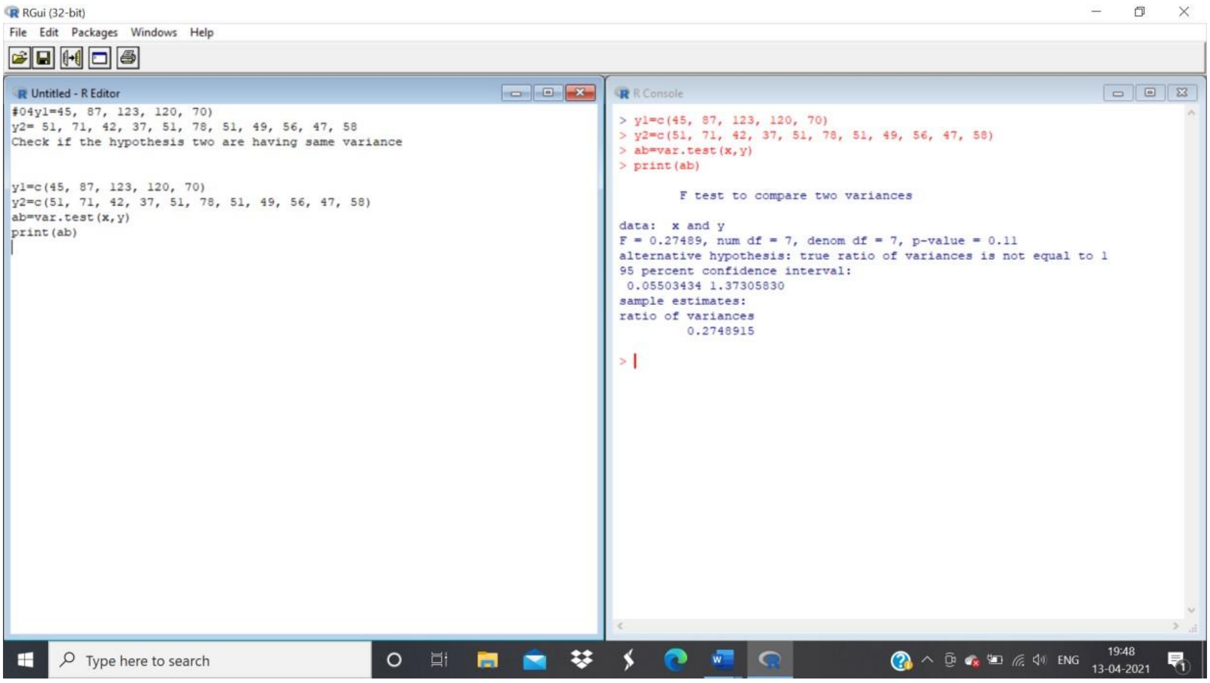
Check if the hypothesis two are having same variance y1=c(45, 87, 123, 120, 70)

y2=c(51, 71, 42, 37, 51, 78, 51, 49, 56, 47, 58)

ab=var.test(x,y)

print(ab)

Output:



Ans: The p-value of F-test 0.11 is greater than the level of significance 0.05. The conclusion is that the variance is the same.

Q.5) Generate two random sample one of size 300 and mean 45 and other of size 65 and mean 44. Check the variance of two data are equal

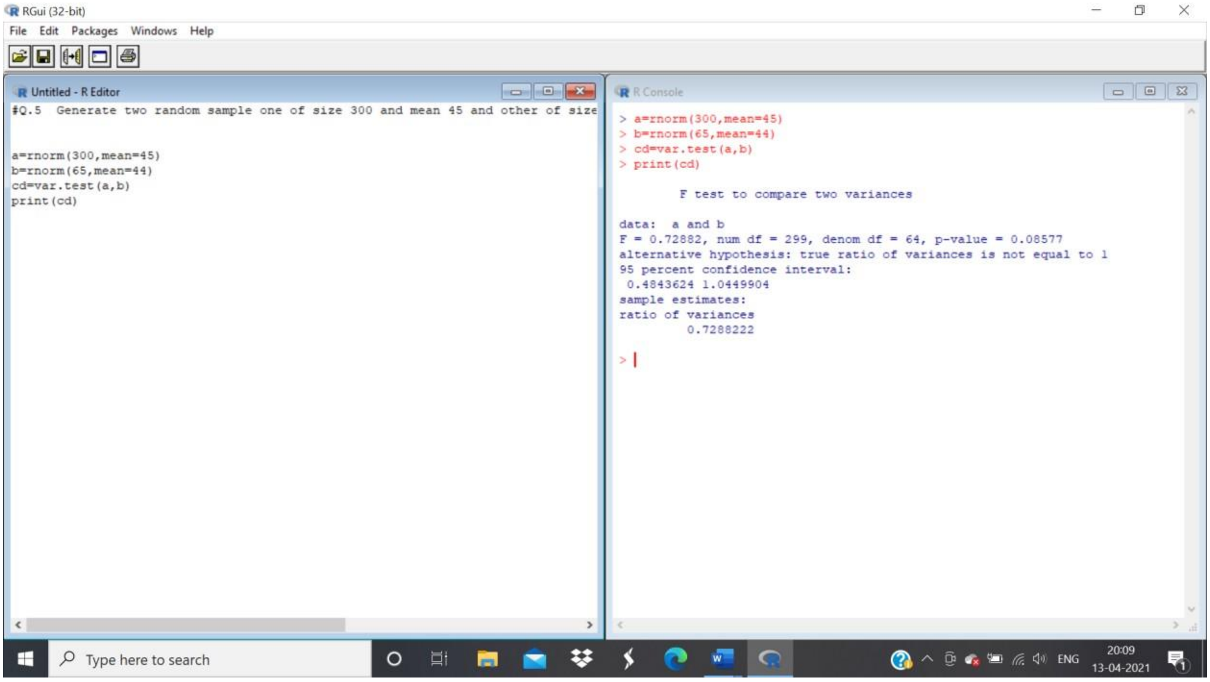
a=rnorm(300,mean=45)

b=rnorm(65,mean=44)

cd=var.test(a,b)

print(cd)

Output:



Ans: The p-value of F-test 0.3616 is greater than the significance level 0.05. The variance of two data are equal.

Q.6) Below are given the gain in weights (in lbs) of pigs fed on two diets A and B Gain in weight Diet A: 25, 32, 30, 43, 24, 14, 32, 24, 31, 31, 35, 25

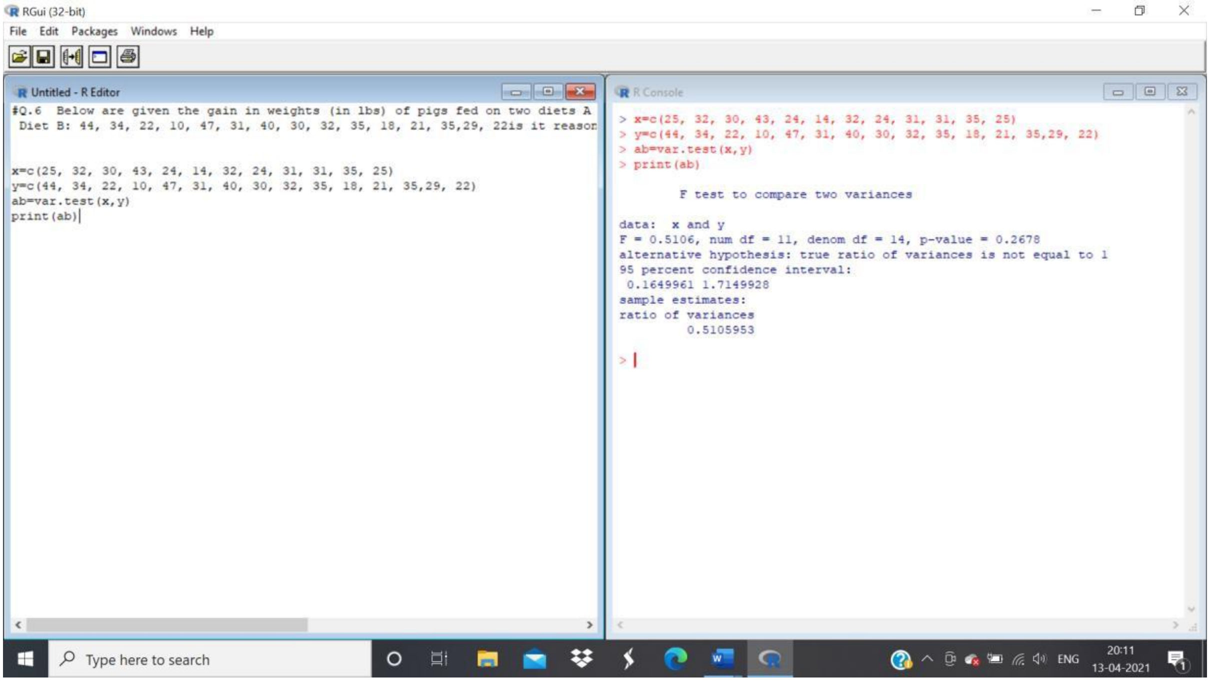
Diet B: 44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21, 35,29, 22is it reasonable to say two variance are equal.

x=c(25, 32, 30, 43, 24, 14, 32, 24, 31, 31, 35, 25)

y=c(44, 34, 22, 10, 47, 31, 40, 30, 32, 35, 18, 21, 35,29, 22)

ab=var.test(x,y)

print(ab)



Ans: The p-value of F-test 0.278 is greater than the level of significance 0.05. It concludesthat that the vairance of the weights is equals.

Q.7) Use the following data to test whether the attributes condition of home and condition ofchild are independent.

		Condition of Home	
		Clean	Dirty
Condition of child	Clean	70	50
	Fairly Clean	80	20
	Dirty	35	45

(Use chi-square test to check the independence)

cat("Enter observed frequencies

column wise" \n)

x=c(70,80,35,50,20,45)

y=matrix(x,nr

ow=3,ncol=2

)

ab=chisq.test

(y,correct=F)

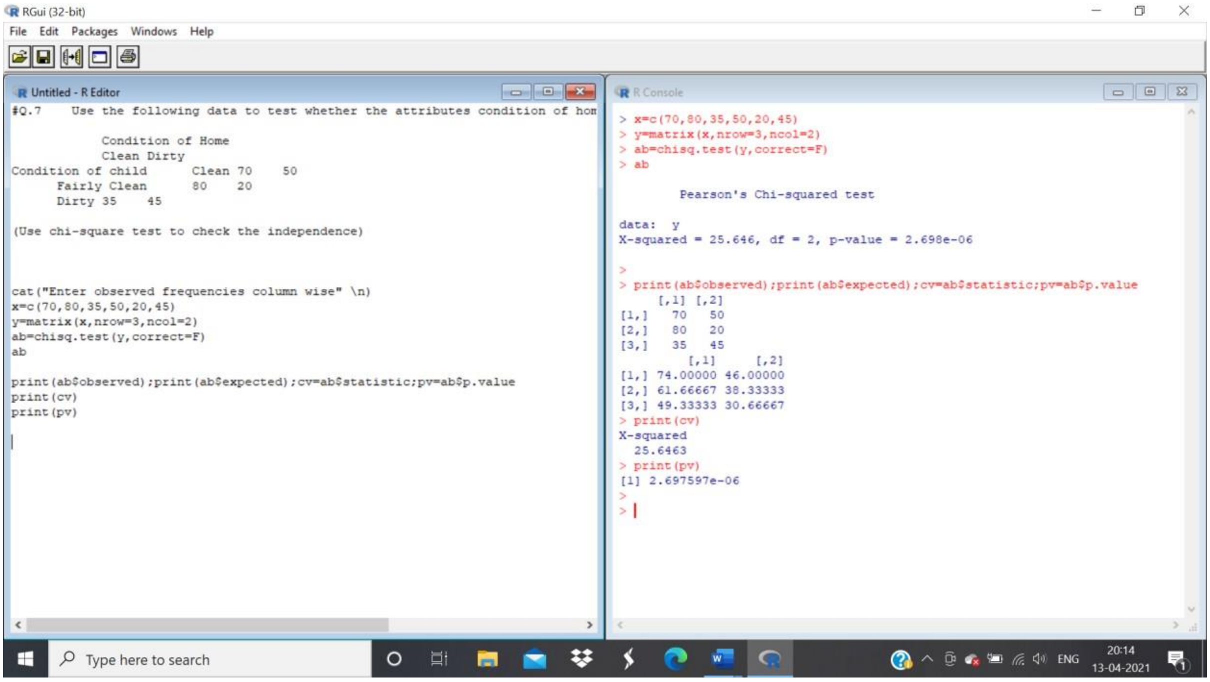
ab

print(ab$observed);print(ab$expected);cv=ab$statistic;pv=ab$

p.value print(cv)

print(pv)

Output:



Practical No: 06

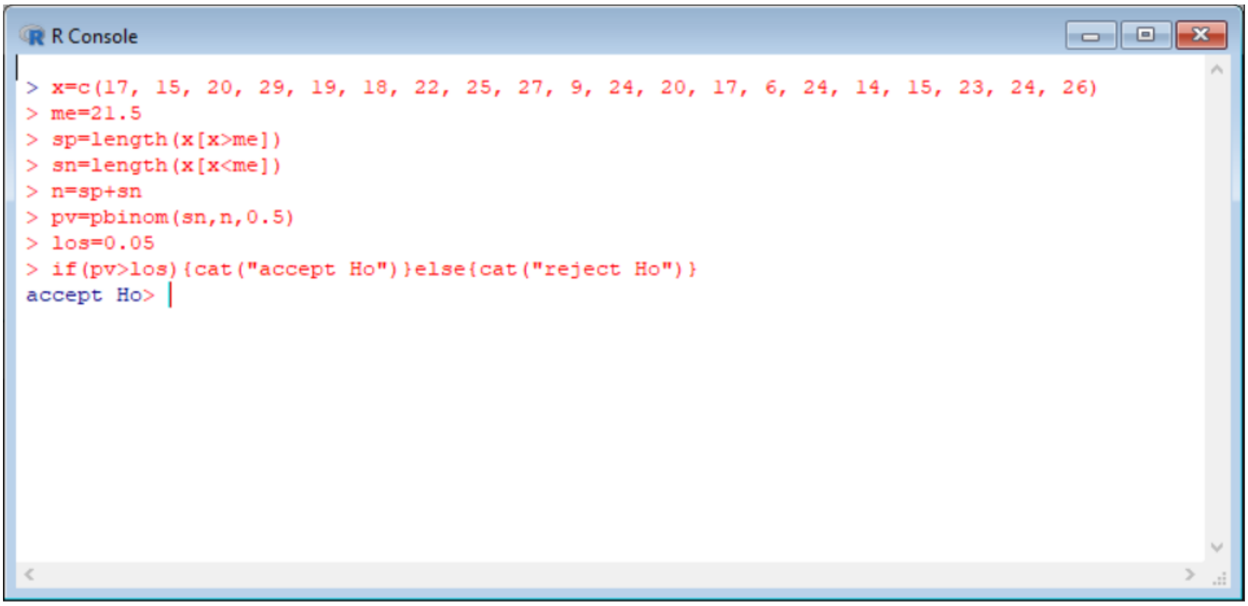
Aim: To demonstrate Sign Test

Q.1) Following are the amounts of Sulphur oxides (x) (in tons) emitted by large industrial plant in 20 days. 17, 15, 20, 29, 19, 18, 22, 25, 27, 9, 24, 20, 17, 6, 24, 14, 15, 23, 24, 26. Apply sign test to test the hypothesis that population median of X is 21.5 against the alternative hypothesis that is less than 21.5 at 0.05 level of significance.

Solution:

```
x=c(17, 15, 20, 29, 19, 18, 22, 25, 27, 9, 24, 20, 17, 6, 24, 14, 15, 23, 24, 26)
me=21.5
sp=length(x[x>me])
sn=length(x[x<me])
n=sp+sn
pv=pbinom(sn,n,0.5)
los=0.05
if(pv>los){cat("accept Ho")}else{cat("reject Ho")}
```

Output:

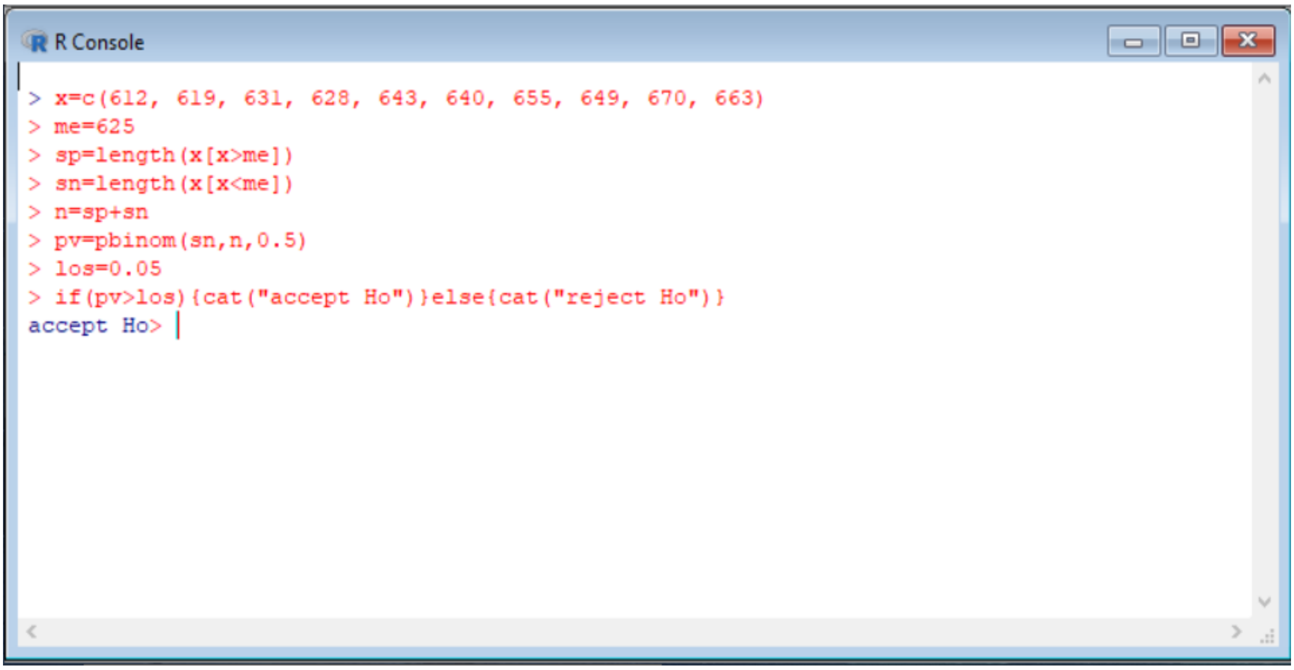


Q.2) Following are data on ten randomly selected specimen of a certain material subjected to stress and the fatigue lives(in kilocycles) 612, 619, 631, 628, 643, 640, 655, 649, 670, 663. Apply sign test to test the hypothesis that population median fatigue life is 625 against the alternative hypothesis that it is greater than 625 at 5% level of significance.

Solution:

```
x=c(612, 619, 631, 628, 643, 640, 655, 649, 670, 663)

me=625

sp=length(x[x>me])

sn=length(x[x<me])

n=sp+sn

pv=pbinom(sn,n,0.5)

los=0.05

if(pv>los){cat("accept Ho")}else{cat("reject Ho")}
```

Output:



Q.3) The following are the measurements of the breaking strength (X) (in pounds) of a certain kind of 2-inch cotton ribbon. 163, 165, 160, 189, 161, 171, 158, 151, 169, 162, 163, 139, 172, 165, 148, 166,172, 163, 187, 173. Test the null hypothesis that population median of X is 160 against the alternative that it is greater than 160 at 0.05 level of significance using Wilcoxon signed rank test.

Solution:

x=c(163,165,160,189,161,171,158,151,169,162,163,139,172,165,148,166,172, 163,187,173)

wilcox.test(x,alter="greater",mu=160)

Output:

```
R Console
> x=c(163,165,160,189,161,171,158,151,169,162,163,139,172,165,148,166,172,163,187,173)
> wilcox.test(x,alter="greater",mu=160)

        Wilcoxon signed rank test with continuity correction

data:  x
V = 146, p-value = 0.02095
alternative hypothesis: true location is greater than 160

Warning messages:
1: In wilcox.test.default(x, alter = "greater", mu = 160) :
  cannot compute exact p-value with ties
2: In wilcox.test.default(x, alter = "greater", mu = 160) :
  cannot compute exact p-value with zeroes
>
```

Conclusion:

Here,los is 0.05 which is greater than p value=0.02095 here we reject Ho.

Q.4) An I.Q test was administered to 5 persons before and after they were trained. The results are given below. Candidate I.Q before training I.Q after training 1 110 120 2 120 118 3 123 125 4 132 136 5125 121. Use sign test to test whether there is increase in I.Q after the training programme at 5% level of significance.

Candidate	IQ before training	IQ after training
1	110	120
2	120	118
3	123	125
4	132	136
5	125	121

Solution:

```
x=c(110,120,123,132,125)
y=c(120,118,125,136,121)
d=x-y
wilcox.test(d,mu=0,alter="less")
wilcox.test(x,y,paired=T,alter="less")
```

Output:

```
R Console
> x=c(110,120,123,132,125)
> y=c(120,118,125,136,121)
> d=x-y
> wilcox.test(d,mu=0,alter="less")

        Wilcoxon signed rank test with continuity correction

data:  d
V = 5, p-value = 0.2931
alternative hypothesis: true location is less than 0

Warning message:
In wilcox.test.default(d, mu = 0, alter = "less") :
  cannot compute exact p-value with ties
> wilcox.test(x,y,paired=T,alter="less")

        Wilcoxon signed rank test with continuity correction

data:  x and y
V = 5, p-value = 0.2931
alternative hypothesis: true location shift is less than 0

Warning message:
In wilcox.test.default(x, y, paired = T, alter = "less") :
  cannot compute exact p-value with ties
> 
```

Conclusion: Since p>0.05. We can accept Ho.

Q.5) The following are the weights in pounds of 16 persons,before and after a certain weight reducing diet programme of four weeks. Use Wilcoxon's signed rank test to test whether the weight reducing diet is effective at 0.01 level of significance.

Person	**Weight Before**	**Weight After**	**Person**	**Weight Before**	**Weight After**
1	147	137.9	9	147.7	149
2	183.5	176.2	10	208.1	195.4
3	232.1	219	11	166.8	158.5
4	161.6	163.8	12	131.9	134.4
5	197.5	193.5	13	150.3	149.3
6	206.3	201.4	14	197.2	189.1
7	177	180.6	15	159.8	159.1
8	215.4	203.2	16	171.7	173.2

Solution:

```
x=c(147,183.5,232.1,161.6,197.5,206.3,177,215.4,147.7,208.1,166.8,131.9,150.3,197.2,159.8,171.7)

y=c(137.9,176.2,219,163.8,193.5,201.4,180.6,203.2,149,195.4,158.5,134.4,149.3,189.1,159.1,173.2)

d=x-y

wilcox.test(d,mu=0,alter="greater")

wilcox.test(x,y,paired=T,alter="greater")
```

Output:

```
R Console
> x=c(147,183.5,232.1,161.6,197.5,206.3,177,215.4,147.7,208.1,166.8,131.9,150.3,197.2,159.8,171.7)
> y=c(137.9,176.2,219,163.8,193.5,201.4,180.6,203.2,149,195.4,158.5,134.4,149.3,189.1,159.1,173.2)
> d=x-y
> wilcox.test(d,mu=0,alter="greater")

        Wilcoxon signed rank exact test

data:  d
V = 111, p-value = 0.01248
alternative hypothesis: true location is greater than 0

> wilcox.test(x,y,paired=T,alter="greater")

        Wilcoxon signed rank exact test

data:  x and y
V = 111, p-value = 0.01248
alternative hypothesis: true location shift is greater than 0

>
```

Conclusion: Since p is 0.0124 and los is 0.01, We reject Ho.

Practical No: 07

Aim: To demonstrate Anova.

When you have more than two samples to compare you would usually attempt to use analysis of variance. However, if the data are not normally distributed (i.e. not parametric) then an alternative must be sought. This is where the Kruskal-Wallis test comes in. It is designed to test for significant differences in population medians when you have more than two samples. K-W test is a non-parametric version of one-way anova.

Q.1 Write this data in excel save as csv file and import apply Kruskal-Wallis test.

Growth	Sugar
75	C
72	C
73	C
61	F
67	F
64	F
62	S
63	S

Solution:

```
c=c(75,72,73)
f=c(61,67,64)
s=c(62,63)
kruskal.test(list(c,f,s))
```

Output:

```
c=c(75,72,73)
f=c(61,67,64)
s=c(62,63)
kruskal.test(list(c,f,s))
```

R Console

```
> c=c(75,72,73)
> f=c(61,67,64)
> s=c(62,63)
> kruskal.test(list(c,f,s))

        Kruskal-Wallis rank sum test

data:  list(c, f, s)
Kruskal-Wallis chi-squared = 5.1389, df = 2, p-value = 0.07658

>
```

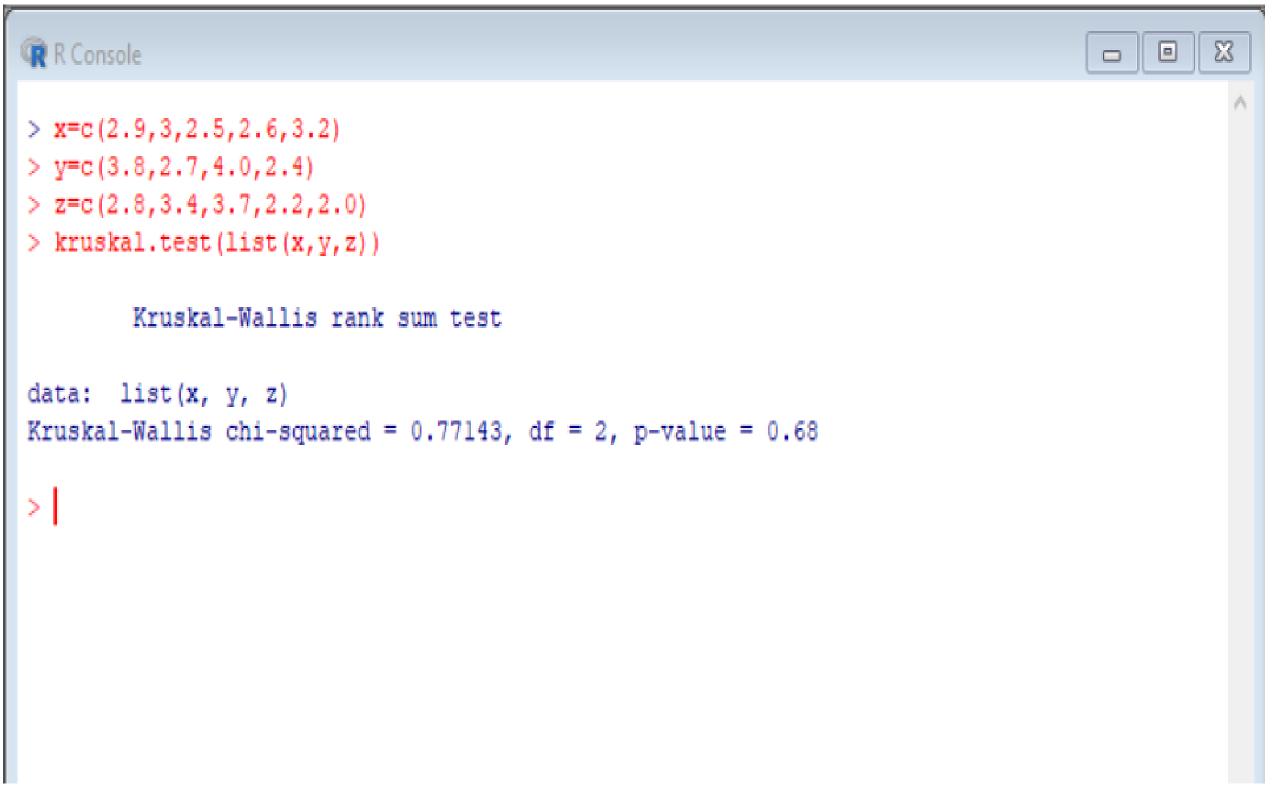
Q.2) The time taken to complete job on three machines are noted test the hypothesis that there is no significant difference between average time taken on these machines to complete the job

Machine	Time taken in hrs
X	2.9,3,2.5,2.6,3.2
Y	3.8,2.7,4.0,2.4
Z	2.8,3.4,3.7,2.2,2.0

Solution:

```
x=c(2.9,3,2.5,2.6,3.2)
y=c(3.8,2.7,4.0,2.4)
z=c(2.8,3.4,3.7,2.2,2.0)
kruskal.test(list(x,y,z))
```

Output:



Q.3) Carry out the analysis of variance for the following data

Varieties	Observations
A	50,52
B	53,55,53
C	60,58,57,56
D	52,54,54,55

Solution:

a=c(20,52)

b=c(53,55,53)

c=c(60,58,57,56)

d=c(52,54,54,55)

kruskal.test(list(a,b,c,d))

Output: