**EXPERIMENT 5**

1. Consider that X is the time (in minutes) that a person has to wait in order to take a flight.

If each flight takes off each hour X ~ U(0, 60). Find the probability that

1. waiting time is more than 45 minutes, and

Sol.

punif(45, min=0,max=60,lower.tail=FALSE)



1. waiting time lies between 20 and 30 minutes.

Sol.

punif(30, min=0,max=60)-punif(20,min=0,max=60)

A number with blue text

Description automatically generated with medium confidence

2. The time (in hours) required to repair a machine is an exponential distributed random

variable with parameter λ = 1/2.

1. Find the value of density function at x = 3.

Sol.

dexp(3,0.5)

A number with numbers and symbols

Description automatically generated with medium confidence

1. Plot the graph of exponential probability distribution for 0 ≤ x ≤ 5.

Sol.

plot(dexp(seq(0,5,0.0001),0.5))

A graph with a line

Description automatically generated

1. Find the probability that a repair time takes at most 3 hours.

Sol.

pexp(3,0.5)

A number and a number

Description automatically generated with medium confidence

1. Plot the graph of cumulative exponential probabilities for 0 ≤ x ≤ 5.

Sol.

plot(pexp(seq(0,5,0.0001),0.5))

A line graph with numbers and a line

Description automatically generated

(e) Simulate 1000 exponential distributed random numbers with λ = 1⁄2 and plot the

simulated data.

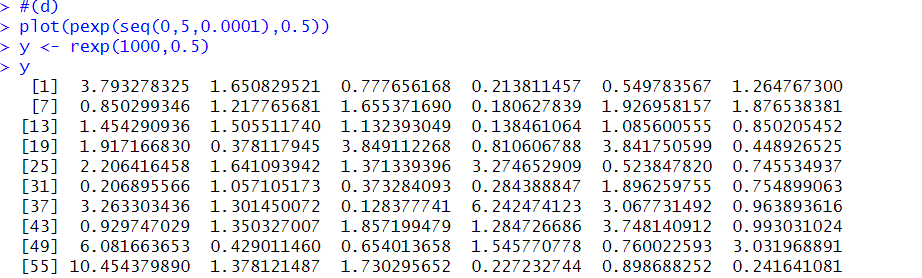
Sol.

y <- rexp(1000,0.5)

y

plot(density(y))

hist(y)



A graph with a bar graph

Description automatically generated

3. The lifetime of certain equipment is described by a random variable X that follows

Gamma distribution with parameters α = 2 and β = 1/3.

1. Find the probability that the lifetime of equipment is at least 1 unit of time.

Sol.

pgamma(1,shape=2,scale=1/3,lower.tail = FALSE)



1. What is the value of c, if P(X ≤ c) ≥ 0.70? (Hint: try quantile function qgamma())

Sol.

qgamma(0.7,shape=2,scale=1/3)

A blue text on a white background

Description automatically generated

**EXPERIMENT 6**

(1) The joint probability density of two random variables X and Y is

A math equations on a white background

Description automatically generated

Then write a R-code to

1. check that it is a joint density function or not? (Use integral2())

Sol.

install.packages('pracma')

library('pracma')

f<-function(x,y){

return (2\*(2\*x+3\*y)/5)

}

i<-integral2(f,xmin=0,xmax=1,ymin=0,ymax=1)

i$Q

A math equations on a white background

Description automatically generated

1. find marginal distribution g(x) at x = 1.

Sol.

g<-function(y){

return (f(1,y))

}

i=integral(g,0,1)

i

A white background with black text

Description automatically generated

1. find the marginal distribution h(y) at y = 0.

Sol.

h<-function(x){

f(x,0)

}

integral(h,0,1)

A white background with black text

Description automatically generated

1. find the expected value of g(x, y) = xy.

Sol.

m<-function(x,y){

return( x\*y\*(2\*(2\*x+3\*y))/5)

}

j=integral2(m,xmin=0,xmax=1,ymin=0,ymax=1)

j$Q

A white background with blue text

Description automatically generated

(2) The joint probability mass function of two random variables X and Y is

A math equations on a white background

Description automatically generated

Then write a R-code to

1. display the joint mass function in rectangular (matrix) form.

Sol.

f<-function(x,y){

(x+y)/30

}

x<-c(0:3)

y<-c(0:2)

mat<-matrix(c(f(0,0:2),f(1,0:2),f(2,0:2),f(3,0:2)),nrow=4,ncol=3,byrow=TRUE)

print(mat)

A computer screen shot of numbers

Description automatically generated

1. check that it is joint mass function or not? (use: Sum())

Sol.

if(sum(mat)==1){

paste("It is joint mass fxn")

}else{

paste("not jms")

}

A white background with black text

Description automatically generated

1. find the marginal distribution g(x) for x = 0, 1, 2, 3. (Use:apply())

Sol.

g<-apply(mat,1,sum)

g

A close up of numbers

Description automatically generated

1. find the marginal distribution h(y) for y = 0, 1, 2. (Use:apply())

Sol.

h<-apply(mat,2,sum)

h

A number and numbers on a white background

Description automatically generated

1. find the conditional probability at x = 0 given y = 1.

Sol.

mat[1,2]/h[2]

A number on a white background

Description automatically generated

1. find E(x), E(y), E(xy), V ar(x), V ar(y), Cov(x, y) and its correlation coefficient.

Sol.

ex<-sum(x\*g)

ex

ey<-sum(y\*h)

ey

A math equations and numbers

Description automatically generated with medium confidence

f2<-function(x,y){

x\*y\*(x+y)/30

}

mat2<-matrix(c(f2(0,0:2),f2(1,0:2),f2(2,0:2),f2(3,0:2)),nrow=4,ncol=3,byrow=TRUE)

exy<-sum(mat2)

exy

ex2<-sum(x\*x\*g)

var\_x=(ex2-ex\*ex)

var\_x

A close-up of a number

Description automatically generated

ey2<-sum(y\*y\*h)

var\_y=ey2-ey\*ey

var\_y

cov<-exy-ex\*ey

cov

corr<-cov/(sqrt(var\_x\*var\_y))

corr

A screenshot of a computer

Description automatically generated

**EXPERIMENT 7**

1. Use the rt(n, df) function in r to investigate the t-distribution for n = 100 and df = n − 1 and plot the histogram for the same.

Sol.

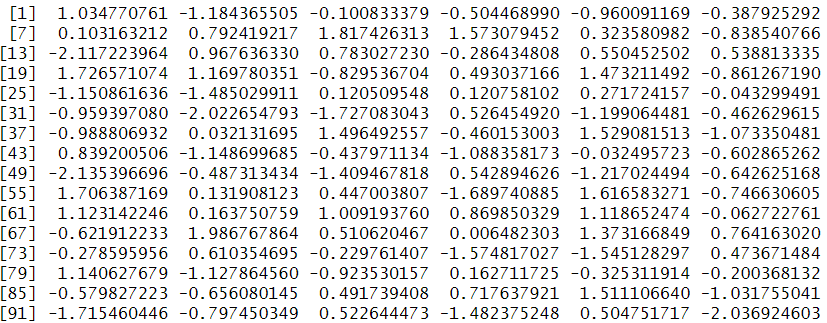
n<-100

df<-n-1

x<-rt(n,df)

x

hist(x)



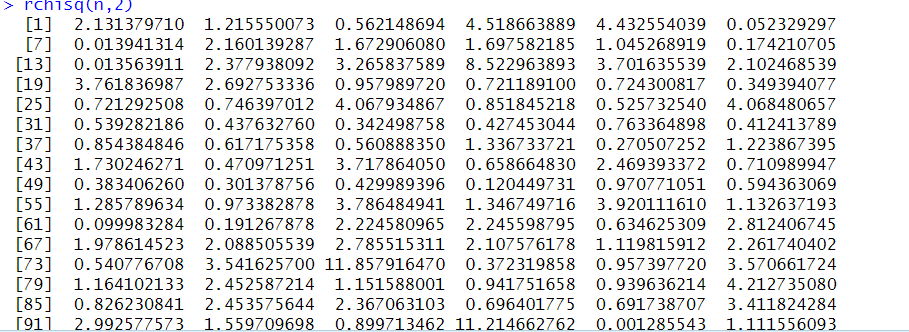
A graph of a graph

Description automatically generated

1. Use the rchisq(n, df) function in r to investigate the chi-square distribution with n = 100 and df = 2, 10, 25.

Sol.

rchisq(n,2)



rchisq(n,10)

A screenshot of a computer screen

Description automatically generated

rchisq(n,25)

A screenshot of a computer

Description automatically generated

(3) Generate a vector of 100 values between -6 and 6. Use the dt() function in r to find the values of at-distribution given a random variable x and degrees of freedom 1,4,10,30. Using these values plot the density function for students t-distribution with degrees of freedom 30. Also shows a comparison of probability density functions having different degrees of freedom (1,4,10,30).

Sol.

x<-seq(-6,6,length=100)

df1=dt(x,1)

df2=dt(x,4)

df3=dt(x,10)

df4=dt(x,30)

df1

df2

df3

df4

plot(x,df4)

color=c("red","green","yellow","blue")

df<-c(1,4,10,30)

for (i in 1:4) {

lines(x,dt(x,df[i]),col=color[i])

}

A graph of a function

Description automatically generated

(4) Write a r-code

(i) To find the 95th percentile of the F-distribution with (10, 20) degrees of freedom.

Sol.

x<-0.95

qf(x,10,20)

A number of numbers on a white background

Description automatically generated

(ii) To calculate the area under the curve for the interval [0, 1.5] and the interval [1.5, +∞) of

a F-curve with v1 = 10 and v2 = 20 (USE pf()).

Sol.

x<-1.5

df1<-10

df2<-20

pf(x,df1,df2)

pf(x,df1,df2,lower.tail = FALSE)

A computer code with numbers and letters

Description automatically generated

(iii) To calculate the quantile for a given area (= probability) under the curve for a F-curve

with v1 = 10 and v2 = 20 that corresponds to q = 0.25, 0.5, 0.75 and 0.999. (use the qf())

Sol.

v1<-10

v2<-20

qf(0.25,v1,v2)

qf(0.5,v1,v2)

qf(0.75,v1,v2)

qf(0.999,v1,v2)

A white background with black numbers and symbols

Description automatically generated

(iv) To generate 1000 random values from the F-distribution with v1 = 10 and v2 = 20 (use

rf())and plot a histogram.

Sol.

y<-rf(1000,v1,v2)

hist(y)

A graph with a bar graph

Description automatically generated

**EXPERIMENT 8**

A pipe manufacturing organization produces different kinds of pipes. We are given

the monthly data of the wall thickness of certain types of pipes (data is available on

LMS Clt-data.csv).

The organization has an analysis to perform and one of the basic assumption of that

analysis is that the data should be normally distributed.

You have the following tasks to do:

1. Import the csv data file in R.

Sol.

df<-read.csv("C:/Users/ExtremePc/Downloads/Clt-data.csv")



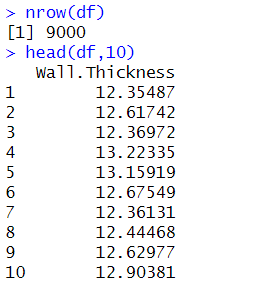
(b) Validate data for correctness by counting number of rows and viewing the top

ten rows of the dataset.

Sol.

nrow(df)

head(df,10)



1. Calculate the population mean and plot the observations by making a histogram.

Sol.

m<-mean(df$Wall.Thickness)

m

hist(df$Wall.Thickness)

A close up of a number

Description automatically generated

A graph of a graph

Description automatically generated

1. Mark the mean computed in last step by using the function abline.

Sol.

A graph of a graph

Description automatically generated

See the red vertical line in the histogram? That’s the population mean. Comment on

whether the data is normally distributed or not?

Now perform the following tasks:

(a) Draw sufficient samples of size 10, calculate their means, and plot them in R

by making histogram. Do you get a normal distribution.

s<-sample(df$Wall.Thickness,10)

s

mean(s)

hist(s)

A graph with a bar graph

Description automatically generated with medium confidence

v<-c()

for (i in 1:9000) {

v[i]=mean(sample(df$Wall.Thickness,10),replace=TRUE)

}

v

hist(v)

abline(v=m,col="red")

A graph of a graph

Description automatically generated

(b) Now repeat the same with sample size 50, 500 and 9000. Can you comment on

what you observe.

Here, we get a good bell-shaped curve and the sampling distribution approaches

normal distribution as the sample sizes increase. Therefore, we can recommend the

organization to use sampling distributions of mean for further analysis.

s1<-sample(df$Wall.Thickness,50)

m1<-mean(s1)

v<-c()

for (i in 1:9000) {

v[i]=mean(sample(df$Wall.Thickness,50),replace=TRUE)

}

v

hist(v)

#similarly with 500 9000 sample sizes

s2<-sample(df$Wall.Thickness,500)

m2<-mean(s2)

s3<-sample(df$Wall.Thickness,9000)

m3<-mean(s3)

s4<-c(m1,m2,m3)

s4

A graph of a person with a beard

Description automatically generated with medium confidence