R Textbook Companion for Probability and Statistics for Engineers and Scientists by Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye[1]

Created by
Dubey Varshit Shailendra
B.Tech.
Electronics Engineering
College of Engineering Pune
Cross-Checked by
R TBC Team

May 25, 2020

[1]Funded by a grant from the National Mission on Education through ICT - http://spoken-tutorial.org/NMEICT-Intro. This Textbook Companion and R codes written in it can be downloaded from the "Textbook Companion Project" section at the website - https://r.fossee.in.

Book Description

Title: Probability and Statistics for Engineers and Scientists

Author: Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye

Publisher: Pearson Education, Boston USA

Edition: 9

Year: 2016

ISBN: 978-0-321-62911-1

R numbering policy used in this document and the relation to the above book.

Exa Example (Solved example)

Eqn Equation (Particular equation of the above book)

For example, Exa 3.51 means solved example 3.51 of this book. Sec 2.3 means an R code whose theory is explained in Section 2.3 of the book.

Contents

List of R Codes

Chapter 1

Introduction To Statistics And Data Analysis

R code Exa 1.2 Dotplot

```
# Chapter 1
# Example 1.2 page no, 4 from the Pdf..
# To plot the Dotplot for the above data..
# package used "ggplot2" if not installed can be done using install.packages("ggplot2")

library(ggplot2)

obs <- c(0.32,0.53,0.28,0.37,0.47,0.43,0.36,0.42,0.38,0.43,0.26,0.43,0.47

cat <- c(rep("no_nit",10),rep("nit",10))

data1 <- data.frame(obs,cat)  # making it a data frame.

data1$f <- factor(data1$obs)   # adding another variable to the data frame.
```

```
# Plot..
ggplot(data1,aes(x = f, y = obs, fill = cat)) + geom
    _dotplot(binaxis = 'y',stackdir = 'center')
```

R code Exa 1.3 Plot for Corrosion Data

```
# Chapter 1
# Example 1.3 page no.9 from the pdf..
#To find the plot for the Corrosion data in example
    1.3
# package used "ggplot2" if not installed can be
    done using install.packages("ggplot2")


library(ggplot2)    # library for making visuals

humidity <- c(20,80,20,80)

average_corrosion <- c(975,350,1750,1550)

coating <- c("Uncoated","Uncoated","Chemical
    corrosion","Chemical corrosion")

a <- data.frame(coating,humidity,average_corrosion)
      #making it a dataframe

# final PLOT..
ggplot(data= a, aes(x= humidity,y= average_corrosion
    ,colour= coating))+geom_line() + geom_point() +
    xlab("Humidity(in %)")+ ggtitle("Corrosion
    Results for Example 1.3")
```

R code Exa 1.4 Mean Variance and Standard Deviation

```r
# Chapter 1
# Example 1.4 page no. 15 from the pdf..
# To find the mean, sample variance, and standard deviation.


a <- c(7.00, 7.07,7.10,6.97,7.00,7.03,7.01,7.01,6.98,7.08)

cat("The sample mean is ",mean(a))

cat("The sample variance is ",var(a))

cat("The sample standard deviation is ", sd(a))
```

R code Exa 1.5 Box and Whisker Plot

```r

# Example 1.5
# To make Boxplot of Nicotine Data..
# package used ggplot2, if not installed you can install it using install.packages("ggplot2")


library(ggplot2)

nicotine <- c(1.09,1.92,2.31,1.79, 2.28,1.74,1.47,1.97,0.85,1.24,1.58,2.03,1.70,2.17,2.55,2.11,1.86,

content <- data.frame(nicotine) #making it dataframe for making plots

```

```
colnames(content) <- c("content1") #changing the
    column name of the dataframe

# Final Plot..
ggplot(data = content,aes(x=1,y= content1))+ geom_
    boxplot(fill = "white", colour = "#3366FF")+
    coord_flip()+ylab("Nicotine Content")+ ggtitle("
    Nicotine Content in Cigarettes")
```

Chapter 2

Probability

R code Exa 2.2 Sample Space

```

# Chapter 2
# Example 2.2 page no. 36 from the pdf..
# To Find the Sample Space..

a <- c(rep("H",2),rep("T",6))

b <- c("H","T",seq(1,6))

paste(a,b)  #concatening both vectores to make a
   sample space
```

R code Exa 2.3 Find Sample Space example 2

```
# Chapter 2
# Example 2.3 page no. 37 from the pdf..
# to find the sample space.

```

```
a <- c(rep("D",2),rep("N",2))

b<- c("D", "N")

c <- c(rep("D",4),rep("N",4))

# final answer
print("The sample space is: ")
paste(c,rep(paste(a,rep(b))))      #merging in
   specific order to find the sample space.
```

R code Exa 2.8 To find Intersection

```
# Chapter 2
# Example 2.8 page no. 40 from the Pdf..
# To find the Intersection of the two sets

a <- c("a","e","i","o","u")

b <- c("l","r","s","t")

intersect(a,b) #displays the common elements in a
   and b. If there is no element common then denotes
    a vector of 0 length
```

R code Exa 2.10 To find Union

```

# Chapter 2
# Example 2.10 page no. 40 from the pdf..
# To find the Union of the given two sets

a <- c("a","b","c")
```

```
b <- c("b","c","d","e")

union(a,b)  #Displays the union of two sets..
```

R code Exa 2.15 Generalized Multiplication Rule

```

# Chapter 2
# Example 2.15 page no. 46 from the pdf..
# Generalized Multiplication rule..

# answer..
cat("We can elect the chair and treasurer in ",22*21, "ways")  #Using multiplication rule
```

R code Exa 2.18 Permutation Problem

```
# Chapter 2
# Example 2.18 page no. 48, from the pdf..
# Permutation Problem..


# Final answer..
cat("The total number of sample points is",factorial(25)/factorial(22))
```

R code Exa 2.19.a Permutation Problem Example 19a

```

```

```
# Chapter 2
# Example 2.19a page no.48 from the pdf..
# Permuation Problem
# number of ways to select president and treasurer
    from 50 students without restriction


cat("The number of choices of officers without
    restriction are",factorial(50)/factorial(48))


# It is simply permutation of 50 with 2
```

R code Exa 2.19.b Permutation Problem Example 19b

```
# Chapter 2
# example 2.19b page no. 48 from the pdf..
# Permuation Problem..


# to select the number of ways to elect president
    and treasurer from 50 people if A(a particular
    student) will serve ,if he is president
# 2 cases first A becomes president and in that case

cat("If A becomes president number of ways to select
     treasurer are ", 49)

a <- 49

# Second case if officers elected without A

cat("The number of possible ways to select without A
     are" ,factorial(49)/factorial(47))

```

```
b <- factorial(49)/factorial(47)

cat("The total number of ways to select the officers in this case are", a+b)
```

R code Exa 2.19.c Permutation Problem Example 19c

```

# Chapter 2
# Example 2.19c page no. 48 from the pdf..
# Permutation Problem..

# NUmber of ways to elect two officers from 50 people if B and C (2 individuals from 50 ) will only work together
# 2 cases . First , if B and C are elected

cat("The number of ways to select B and C both from 50 are",2) # if B treasurer and C president and vice versa..

a <- 2

# Second case: if B and C are not chosen so election done from rest 48 people.

cat("The number of selections when B and C are not chosen are",factorial(48)/factorial(46)) #48P2

b <- factorial(48)/factorial(46)

cat("The total number of choices in this case is",a+b)
```

R code Exa 2.19.d Permutation Problem Example 19d

```r

# Chapter 2.
# Example 2.19d page no. 48 from the pdf..
# Permutation Problem..

# Number of ways to select officers from 50 students if D and E(2 individuals from 50 students) will not work together
cat("The number of choices of officers without restriction are",factorial(50)/factorial(48))

a <- factorial(50)/factorial(48)

cat("Number of ways in which D and E are selected are",2) # if D is president and E is treasurer and vice versa

b <- 2

cat("Number of ways in which we can choose officers in this case are",a-b)
```

R code Exa 2.21 Partition Problem

```r
# Chapter 2
# Example 2.21 page no. 50 from the pdf..
# Partition Problem..


```

```
# number of ways, 7 students are assigned 1 triple
   and 2 double rooms..

# answer..
cat("The number of ways in which this partition can
   be done are",(factorial(7)/(factorial(3)*
   factorial(2)*factorial(2))))
```

R code Exa 2.22 Combination Problem

```

# Chapter 2
# Example 2.22 page no. 50 from the pdf..
# Combination Problem..

# number of ways to select 3 arcade and 2 sports
   games from 10 arcade and 5 sports game
cat("The number of ways to select 3 arcade from 10
   is",choose(10,3))

a <- choose(10,3)

cat("The number of ways of selecting 2 catridges
   from 5 is",choose(5,2))

b <- choose(5,2)

#using multiplication rule..
cat("The total number of ways of selecting 3 arcade
   and 2 sports games is",a*b)

# example 2.23 same as the above 2 questions 2.22
   and 2.21..
```

R code Exa 2.24 Probability Calculation Using Formula And Law of Large Numbers

```r
# Chapter 2
# Example 2.24 page no.53 from the pdf..
# Probability Using Normal Method and Law of Large numbers..

# to find probability for atleast one head if a coin is tossed twice..

# "dplyr" package is used, if not installed can be installed using
# install.packages("dplyr") , remove comment if you already have..
library(dplyr)      # this package is for data manipulation

a <- c("H","H","T","T")

b <- c("H","T","H","T")

a <- data.frame(a,b)     # making data frame which consists of 2 columns i.e first column and 2nd column

c <- a %>% filter(a=="H"| b=="H")

cat("The probability that we get atleast one head in the tosses is",nrow(c)/nrow(a))

# another method to solve the same problem, by using law of large numbers

```

```
d <- c("HH","TT","TH","HT")

e <- sample(d,10000,replace = T)

cat("The approximate probability of getting atleast one head in the two tosses is",sum(e=="HH"|e=="HT"|e=="TH")/10000)

# the answer here is near to 0.75 but not exactly 0.75, if we increase the number of trials say 200000 it will be more close to 0.75.
# if we make this say tending to infinity then it will be exactly 0.75..
```

R code Exa 2.25 Probability Calculation Using Law of Large Numbers

```
# Chapter 2
# Example 2.25 page no. 53 from the pdf..
# Probability Using Law of Large Numbers..

# Probability can also be calculated using direct formula(ratio) but the code will be very straight forward..
# Let's look at another way to achieve the same result.
# To find the probability of getting less than 4 in rolling a die, if even number has twice the probability than odd number..

a <- c(1,2,2,3,4,4,5,6,6)

b <- sample(a,10000,replace=T)    # using law of large numbers

```

```
cat("The probability of getting a number less than 4 is approximately",sum(b < 4)/10000)

#this answer is slighty different than the answer in textbook which is 0.444. If we increase the number of trials then the probability will be more close to the answer in textbook
```

R code Exa 2.30 Probability Calculation Using Law of Large Numbers And Union Property

```
# Chapter 2
# Example 2.30 page no. 57 from the pdf..
# Probability Using Law of Large Numbers and using union property..


#to find the probability the sum of die roll (ROLLED TWICE) is 7 or 11
#Using law of large numbers..

a <- c(1:6)

b <- sample(a,200000,replace=T) + sample(a,200000,replace=T)    # b is a sample of sum of 2 die rolls , with 10^6 times trials


cat("The approximate value of probability of getting the sum as either 7 or 11 is",sum(b==7|b==11)/200000)



# This is approximate as the textbook and standard
```

```
answer is 2/9, if we increase the number of trials then we would get very close to the final answer..
# A word of note- I have skipped some problems as you may have seen, as many problems are of same type just the numbers are changed, the concept is same, so I solved the relevant questions.
```

R code Exa 2.34 Conditional Probability Problem

```
# Chapter 2.
# Example 2.34 page no.63 from the pdf..
# Conditional Probability Problem..


p_departs_ontime <-  0.83
P_arrives_ontime <-  0.82
p_departs_arrives_ontime <- 0.78

# to find probability that it arrives on time given it departs in time
# Using conditional probability ..

cat("The probability that the plane arrives on time given that it departs on time is",p_departs_arrives_ontime/p_departs_ontime)

# to find probability that it departs on time given that it arrives on time..

cat("The probability that the plane departs on time given that it arrives on time is",p_departs_arrives_ontime/P_arrives_ontime)

```

```

#The answer in the textbook is approximated to 0.94
    and 0.95 respectively ..
```

R code Exa 2.37 Independence and Multiplicative Rule

```

# Chapter 2
# Example 2.37 page no. 66 from the pdf..
# Independence and Multiplication Rule..


# bag1 = 4 white and 3 black balls
#bag2= 3 white and 5 black balls
# to dind probability of ball drawn from 2nd bag is
    black ? given a ball drawn from first is unseen
    and dropped in second bag

p_b1 <- 3/7
p_w1 <- 4/7

p_b2_b1 <- 6/9  # if first black ball dropped in
    second bag
p_b2_w1 <- 5/9  # if first white ball drawn is
    dropped into second bag..

# Using independence and multiplicative rule

cat("The probability that the second ball drawn is
    black in this case is",p_b1*p_b2_b1+p_w1*p_b2_w1)
```

R code Exa 2.43 Bayes Rule

```r
# Chapter 2.
# Example 2.43 from the pdf..
# Bayes Rule..

# given different parameters to find which plan was most likely used and thus responsible..
# to find P(Pj\D) for j = 1,2,3..

# a vector of P(Pj) for j = 1,2,3..
a <- c(0.30,0.20,0.50)

# a vector of P(D|Pj) for j = 1,2,3..
b <- c(0.01,0.03,0.02)

# Bayes Rule Formula..
f <- function(x){

  a[x]*b[x]/sum(a*b)

}

cat("The value of P(Pj|D) for j = 1,2,3 is",f(1:3))

cat("We can see that for plan 3 has the highest conditional probability, hence a defective for a random product is most likely the result of the use of plan 3")


# Some Problems are not solved because they are one and the same of what I have solved just numbers are changed and they are easy too..
```

Chapter 3

Random Variables And Probability Distributions

R code Exa 3.8 Discrete Probability Distribution Problem

```

# Chapter 3
# Example 3.8 page no. 84 from the pdf..
# Discrete Probability Problem..

# From 3.1 to 3.8 the problems are too easy and
    theoretical..
# to find the probability distribution of number if
    defective laptops if 2 are drawn at a time from
    20.

#making a general function for the above
    distribution..
a <- function(x){

  choose(3,x)*choose(17,2-x)/choose(20,2)

}

```

```
# The Distribution..
cat("The probability distribution for x=0 is",a(0))

cat("The probability distribution value for x=1 is",
    a(1))

cat("The probability distribution for x=2 is",a(2))
```

R code Exa 3.9 Discrete Probability Distribution Problem 9

```
# Chapter 3
# Example 3.9 page no. 85 from the pdf..
# Discrete Probability Distribution Problem..

# to find probability distribution of no. of cars
    with side airbags among next 4 cars, given prob.
    of selling is 0.5
# for ex. to sell 3 cars with air sidebags ,
    partition 4 into two - with 3 bags and with 1 bag
# done in 4 choose 3 ways

# generalise
cat("The probability distribution in this problem is
    :"," ")

for(i in 0:4){

  cat("  ",choose(4,i)/16)

  }
```

R code Exa 3.10 Cumulative Distribution Function Problem

```
# Chapter 3
# Example 3.10 page no. 86 from the pdf..
# Cumulative Distribution function..

# to find cummulative distribution of the example
    3.9
# refer Example 3.9 page no. 85 from the pdf.
# the probability distribution of the example is
    choose(4,x)/16 where x is from 0 to 4
a <- function(x){

  choose(4,x)/16
}


cat("The cummulative probability distribution of the
     above problem is "," ")

# ANswer..
for(i in 0:4){

  cat("  ", sum(a(0:i)))
}
```

R code Exa 3.11 Probability Density Function Problem 11

```
# Chapter 3
# Example 3.11 Page no. 89 from the pdf..
# Probability Density FUnction Problem..

# A part of the Problem..
# to check the probability density function

```

```
a <- function(x){

  x*x/3
}

# to check the conditions of probability density
  function


cat("the Integration is",integrate(a,-1,2)$value,"
  nearly 1 with very small margin of error we can
  say that it is a prob. density function")

# B Part of the Problem..
# to find P(0<x<=1)

cat("The value of P(0<x<=1) is",integrate(a,0,1)$
  value)

#this answer matches with the answer in  the
  textbook.
```

R code Exa 3.12 Cumulative Distribution Function of Continuous Random Variable

```

# Chapter 3
# Example 3.12 Page no. 90 from the pdf..
# Cumulative Distribution Function of Continuous
  Random Variable

# given a density function find the value of P(0<X
  <=1)

# the Density Function..
```

```
f <- function(t){

  t*t/3

}

cat("The value of P(0<X<=1) is",integrate(f,0,1)$value)
```

R code Exa 3.14 Joint Probability Distribution of Discrete Random Variable

```

# Chapter 3
# Example 3.14 Page no. 95 from the pdf..
# Joint Probability Distribution of Discrete Random Variable..

# Note - Example 3.13 is theoretical..

# the formula for joint Prob. distribution..
a <- function(x,y){

  choose(3,x)*choose(2,y)*choose(3,2-x-y)/choose(8,2)

}

# to find probability that it will fall in x+y<=1


cat("The probability that x,y fall in region x+y<=1 is",a(0,0)+a(0,1)+a(1,0))
```

R code Exa 3.15 Joint Density Function of Continuous Random Variable

```
# Chapter 3
# Example 3.15 page no.96 from the pdf..
# Joint Density Function of Continuous Random
    Variable..

# A part of the Problem 3.15
# to verify the conditions of joint density function
     and find the probabilities of any intervals

# the joint density function..
a <- function(x,y){

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


}

print("The value of the integral over whole region
    is:")

integrate(function(y){

  sapply(y,function(y){

    integrate(function(x)a(x,y),0,1)$value
  })
},0,1)

# the integral is 1 then it is a joint density
    function

```

```
# Part B of the Problem 3.15
# to find the probability of a region A:{(x,y)| 0<x
    <0.5, 0.25<y<0.5}

cat("The probability is "," ")

integrate(function(y){

  sapply(y,function(y){

    integrate(function(x)a(x,y),0,0.5)$value

  })

},0.25,0.5)
```

R code Exa 3.16 Marginal Distribution Problem

```

# Chapter 3
# Example 3.16 page no. 98 from the pdf..
# Marginal Distribution Problem

# to find marignal Distribution Of X alone and Y
    alone..

c <- as.data.frame(matrix(c(3/28,3/14,1/28,9/28,3/
    14,0,3/28,0,0),ncol = 3))

colnames(c) <- c("x0","x1","x2")

rownames(c) <- c("y1","y2","y3")

cat("The marginal distribution of X alone is",apply(
    as.matrix(c),2,sum),"and the marginal
```

```
    distributions of Y alone is",apply(as.matrix(c)
    ,1,sum))

# Example 3.17 is theoretical..
```

R code Exa 3.18 Conditional Distribution Of a Random Variable

```

# Chapter 3
# Example 3.18 Page no.99 from the pdf..
# Conditional Distribution Of a Random Variable.

# to find the conditional joint probability
    distribution of the given function given the
    specific conditions
# Refer Example 3.14 from the pdf..
a <- function(x,y){

  choose(3,x)*choose(2,y)*choose(3,2-x-y)/choose
     (8,2)

}

# given y=1 to find P(x=0|y=1)

# P(x=0|y=1)= a(x,y)/h(y), h(y)>0

h1 <- sum(a(0:2,1))

# to find a(x,1)= a(x,1)/h1

b <- function(x){

  a(x,1)/h1

```

```
}

cat("The conditional probability function under the condition y=1 is"," ")

c <- c(b(0),b(1),b(2))

d <- t(data.frame(0:2,c))

rownames(d) <- c("X","f(x|1)")

d
```

R code Exa 3.20 Marginal and Conditional Densities Evaluation

```
# chapter 3
# Example 3.19 and 3.20 both are same in concept, just the function is changed, I am solving 3.20, so you can do the same for 3.19..
# Chapter 3.20 page no 100 from the pdf..
# Mariginal and Conditional Densities Evaluation..

# given the joint distribution function
a <- function(x,y){

  x*(1+3*y*y)/4

}

# with limits 0<x<2, 0<y<1

# given marginal density
 b <- function(x){

   x/2
```

```
 }

c <- function(y){

  (1+3*y*y)/2

}

# to find a(x|y)
# to find P(0.25<x<0.5|y=1/3)

d <- function(x){

  a(x,1/3)/c(1/3)

}

cat("The value of conditional Probability P(0.25<X
   <0.5|Y = 1/3) is",integrate(d,0.25,0.5)$value)
```

R code Exa 3.22 Statistical Independent Problem

```

# Chapter 3
# Example 3.22 page 103 from the pdf..
# Statistical Independent Problem..

# Note- Example 3.20 is theoretical, we need to just
    prove the variable is not statistical
   Independent, code involves just basic
   multiplication..

# package installed pracma just for simplicity for
   evaluating integrals numerically..
```

```
# given a density function and x1 ,x2, x3 represent 3 of these containers selected independently..
# to find P(x1<2,1<x2<3,x3>2)
# if the package is already installed remove the installed.packages("pracma") from your code..

install.packages("pracma")
library("pracma")

f <- function(x,y,z){

  exp(-x-y-z)

}


# Answer..
cat("The value of P(x1<2,1<x2<3,x3>2) is",integral3(f,0,2,1,3,2,22))
```

Chapter 4

Mathematical Expectation

R code Exa 4.1 Expected Value Evaluation Problem 1

```r

# Chapter 4
# Example 4.1 Page no. 113 from the pdf..
# Expected Value Evaluation Problem

# to find the expectation value
# out of 7, 4 good and 3 defective components, and
    sample of 3 is taken, find the expectation value

a <- function(x){

  choose(4,x)*choose(3,3-x)/choose(7,3)

}


b <- a(0:3) # to create a series of vector for X..

cat("The expectation value of this problem is",
    weighted.mean(0:3,b))
```

R code Exa 4.2 Expected Value Evaluation Problem 2

```

# Chapter 4
# Example 4.2 Page no. 113 from the pdf..
# Expected Value Evaluation Problem

# To find the Expected commission..

p <- 0.7

q <- 0.4

dat <- c(0,1000,1500,2500)

# Distribution..
weights <- c((1-p)*(1-q),p*(1-0.4),(1-p)*q,p*q)

cat("The expected commission is",weighted.mean(dat,
   weights))
```

R code Exa 4.3 Expected Value of probability density function

```

# Chapter 4
# Example 4.3 page no. 114 from the pdf..
# Expected Value of probability density function

# given that a(x) if x>100 else 0

a <- function(x){

```

```
  20000/(x*x*x)

}

a1 <- function(x){

  x*a(x)

}

cat("The Expected life of this type of device is:",
   integrate(a1,lower = 100,upper = Inf)$value)
```

R code Exa 4.4 Expected Value of a new Random Variable

```

# Chapter 4
# Example 4.4 Page no. 115 from the pdf..
# Expected Value of a new Random Variable..

# given a probability distribution find a expected
   value of a new random variable
# new random variable 2x-1

# P(X=x)
px <- c(1/12,1/12,1/4,1/4,1/6,1/6)

# x
x <- 4:9

# g(x)=2x-1

cat("The Attendant can expect to receive :",weighted
   .mean(2*x-1,px))
```

R code Exa 4.5 Expected Value of a new Random Variable

```
# Chapter 4
# example 4.5 page no. 115 from the pdf.
# Expected Value of New Random Variable

# to find expected value given a random variable f(x
    )= x^2/3 of g(X)= 4X +3

f <- function(x){

  (4*x+3)*x*x/3

}

cat("The expected value is",integrate(f,-1,2)$value)
```

R code Exa 4.6 Expected value of Discrete Joint Probability Distribution

```
# Chapter 4
# Example 4.6, page no. 116 from the pdf..
# Expected value of Discrete Joint Probability
    Distribution


# given 2 random variables with joint prob.
    distribution,
# find the expected value g(X,Y) = XY.

f <- matrix(c(3/28,3/14,1/28,9/28,3/14,0,3/28,0,0),
    ncol = 3)

```

```
ans <- 0

for(x in 0:2){

  for(y in 0:2){

    ans <- ans + x*y*f[x+1,y+1]

  }
}

cat("The expected value of g(X,Y)= XY is",ans)
```

R code Exa 4.7 Expected value of Joint Density Function

```

# Chapter 4
# Example 4.7 page no. 116 from the pdf..
# Expected Value of Joint Density Function


# to find the expected value of density function

# given function

a <- function(x,y){

  (x*(1+3*y*y))/4

}

# making a expectation value function

b <- function(x,y){
```

```
  a(x,y)*y

}

# using the expression to find the expectation value

print("The Value of E(Y/X) is")

integrate(function(y){

  sapply(y,function(y){

    integrate(function(x)b(x,y),0,2)$value
  })
},0,1)
```

R code Exa 4.8 Variance Of Discrete Distribution

```

# Chapter 4
# Example 4.8 page no. 120 from the pdf..
# Variance Of Discrete Distribution..

# given two distribution finding out which set has
   higher variance

# first distribution
a <- c(0.3,0.4,0.3)

b <- weighted.mean(1:3,a)

c1 <- function(x){

  (x-b)*(x-b)

```

```
}

d <- sum(c1(1:3)*a)

# second Distribution..
a1 <- c(0.2,0.1,0.3,0.3,0.1)

b1 <- weighted.mean(0:4,a1)  # since b and b1 are same..


d1 <- sum(c1(0:4)*a1)

cat("Since Variance of B is",d1,"and Variance of A is",d,"Variance of B is greater than A")
```

R code Exa 4.9 Variance using Expectation Formula

```

# Chapter 4
# Example 4.9 page no. 121 from the pdf..
# Variance using Expectation formula..


# calculate variance using expectation formula


# given data

a <- c(0.51,0.38,0.10,0.01)

# using E[x^2]-(E[X])^2

varia <- weighted.mean(0:3*0:3,a)- (weighted.mean(0:3,a))^2
```

```r

cat("The value of Variance is:",varia)
```

R code Exa 4.10 Mean And Variance of Continuous Random Variable

```r
# Chapter 4
# example 4.10 page no. 121 from the pdf..
# Mean And Variance of Continuous Random Variable..

# given a random variable X having the prob. density
# find the mean and variance of X..

f <- function(x){

  2*(x-1)

}

cat("The mean is mu = E(X) =",integrate(function(x){
    x*f(x)},1,2)$value)

cat("And the Variance is E(X^2)-[E(x)]^2 = ",
    integrate(function(x){x*x*f(x)},1,2)$value - (
    integrate(function(x){x*f(x)},1,2)$value)^2)
```

R code Exa 4.11 Variance of a new Random variable

```r

# Chapter 4
# Example 4.11 page no. 122 from the pdf..
# Variance of a new Random variable


```

```
# given a random variable find the variance of
    another random variable

#pdf for X
a <- c(1/4,1/8,1/2,1/8)

# finding variance for 2x+3

# E[(2x+3-6)^2]

varia <- weighted.mean(4*0:3*0:3-12*0:3+9,a)

cat("The variance of g(X)=2X+3 is",varia)
```

R code Exa 4.12 Variance of a new Random variable given the Density Function

```

# Chapter 4
#example 4.12 page no. 123 from the pdf..
# Variance of a new Random variable given the
    Density Function..

# given a random variable find the variance of g(X)
    =4X+3

f <- function(x){

  x*x/3

}

# using theorem 4.3 as given on page no. 122

cat("The variance of random variable g(4X+3) is",
```

```
integrate(function(x){f(x)*((4*x+3-8)^2)},-1,2)$
value)
```

R code Exa 4.13 Covariance of 2 Random Variables

```

# Chapter 4
# Example 4.13 page no. 124
# Covariance of 2 Random Variables..


# to find the covariance between two variables X and
    Y

a <- matrix(c(3/28,3/14,1/28,9/28,3/14,0,3/28,0,0),
    nrow = 3)

# individual column sum
ax <- apply(a,2,sum)

# indivdual row sum
ay <- apply(a,1,sum)

# E[X]..

ex <- sum(0:2*ax)

# E[Y]..

ey <- sum(0:2*ay)

# given E[XY]=3/14

# using covariance formula
# E[XY]-E[X]*E[Y]
```

```
covar <- 3/14-ex*ey

cat("The covariance between the random variable X
    and Y is",covar)
```

R code Exa 4.14 Covariance Of Two Random Variables given Joint Density Function

```

# Chapter 4
# Example 4.14 page no. 125 from the pdf..
# Covariance Of Two Random Variables given Joint
    Density Function..


# given a joint density function of 2 random
    variables X and Y
# find the covariance of X and Y..

f <- function(x,y){

  8*x*y

}

g <- function(x){

  4*(x^3)

}

h <- function(y){

  4*y*(1-y*y)
```

```
}

# finding E(X) and E(Y) from marginal densities..

mu_x <- integrate(function(x){x*g(x)},0,1)$value

mu_y <- integrate(function(y){y*h(y)},0,1)$value

# E(XY) from joint probability distributon..
E_XY <- integrate(function(y){

  sapply(y,function(y){

    integrate(function(x){f(x,y)*x*y},y,1)$value
  })},0,1)

cat("The covariance of X and Y is",E_XY$value-mu_x*
    mu_y)
```

R code Exa 4.15 Correlation Coefficient Evaluation

```

# Chapter 4
# Example 4.15 page no. 126 from the pdf..
# Correlation Coefficient Evaluation

# to find the correlation coefficient between two
    variables x and y..
# Refer Example 4.13 page 124 from the pdf..

a <- matrix(c(3/28,3/14,1/28,9/28,3/14,0,3/28,0,0),
    nrow = 3)

# individual column sum
```

```
ax <- apply(a,2,sum)

# indivdual row sum
ay <- apply(a,1,sum)

# E[X]..

ex <- sum(0:2*ax)

# E[Y]..

ey <- sum(0:2*ay)

# variance(x)..

variax <- sum(0:2*0:2*ax)-ex^2

# variance(y)..

variay <- sum(0:2*0:2*ay)-ey^2

# covariance..

covaria <- 3/14-ex*ey
# correlation coefficient

corr <- covaria/(sqrt(variax)*sqrt(variay))

cat("The correlation coefficient in this case is",
  corr)
```

R code Exa 4.16 Correlation Coefficient Of 2 Random Variables Given Joint Density Function

```
# Chapter 4
```

```
# Example 4.16 page no. 126 from the pdf..
# Correlation Coefficient Of 2 Random Variables Given Joint Density Function..


# given a joint density function of 2 random variables X and Y
# find the correlation coefficient of X and Y..

f <- function(x,y){

  8*x*y

}

g <- function(x){

  4*(x^3)

}

h <- function(y){

  4*y*(1-y*y)

}

# finding E(X) and E(Y) from marginal densities..
mu_x <- integrate(function(x){x*g(x)},0,1)$value

mu_y <- integrate(function(y){y*h(y)},0,1)$value

# finding sigma_x and sigma_y from marginal densities..
sig_x <- integrate(function(x){x*x*g(x)},0,1)$value - mu_x^2

sig_y <- integrate(function(y){y*y*h(y)},0,1)$value
```

```
    - mu_y^2

# E(XY) from joint probability distributon..
E_XY <- integrate(function(y){

  sapply(y,function(y){

    integrate(function(x){f(x,y)*x*y},y,1)$value
  })},0,1)

# finding sigma_xy
sig_xy <- E_XY$value-mu_x*mu_y

cat("The correlation coefficient of X and Y is",sig_
    xy/sqrt(sig_x*sig_y))
```

R code Exa 4.17 Expectation Value of Linear Combination Of Random Variables

```

# Chapter 4
# Example 4.17 page no. 128 as given in the pdf..
# Expectation Value of Linear Combination Of Random
    Variables..

# given a random variable and its distribution find
    the value of E(2X-1)
# Refer example 4.4 on page 115 from the pdf..

px <- c(1/12,1/12,1/4,1/4,1/6,1/6)

x <- 4:9

cat("The value of E(2X-1) is",weighted.mean(2*x-1,px
    ))
```

R code Exa 4.18 Expectation Value of Linear Combination Of Continuous Random Variables

```
# Chapter 4
# Example 4_18 page no. 129 from the pdf..
# Expectation Value of Linear Combination Of Random
    Variables

# given X a random variable find the expected value
    of another random variable 4X+3

f <- function(x){

  x*x/3

}

# using property E(aX+b)= aE(X)+b


cat("The value of E(4X+3) is",4*integrate(function(x
   ){x*f(x)},-1,2)$value + 3)
```

R code Exa 4.19 Expectation Value of Linear Combination Of Random Variables Using Properties

```
# Chapter 4
# Example 4.19 page no. 129 from the pdf..
# Expectation Value of Linear Combination Of Random
    Variables..
```

```
# finding expected value using properties..


a <- c(1/3,1/2,0,1/6)

# find E[(X-1)^2]

# E[X^2-2X+1]

#E[X]
ex <- weighted.mean(0:3,a)

#E[X^2]
ex2 <- weighted.mean(0:3*0:3,a)

# using properties

cat("The Expected Value of Y =(X-1)^2 is:",1*ex2 -2*
    ex +1)
```

R code Exa 4.20 Expectation Value Evaluation Using Properties Given a Density Function

```

# Chapter 4
# example 4.20 page no. 130 from the pdf..
# Expectation Value Evaluation Using Properties
    Given a Density Function..


# given a density function of a random variable find
     the value of E(X^2+X-2) using properties..

```

```
f <- function(x){

  2*(x-1)

}

# using properties..

cat("The value of E(X^2+X-2) is",integrate(function(
    x){x*x*f(x)},1,2)$value + integrate(function(x){x
    *f(x)},1,2)$value - integrate(function(x){2*f(x)
    },1,2)$value)
```

R code Exa 4.21 Expectation Value of 2 Independent Random Variable

```

# Chapter 4
# example 4.21 on page no. 131 from the pdf..
# Expected Value of 2 Independent Random Variable..

# to prove with example for 2 independent random
    variables E(XY) = E(x)*E(Y)

# XY function
f <- function(x,y){

  x*(1+3*y*y)/4

}

E_XY <- integrate(function(y){

  sapply(y,function(y){

    integrate(function(x){f(x,y)*x*y},0,2)$value
```

```
  })},0,1)

# finding E(X) and E(Y) value
E_X <- integrate(function(x){x*x/2},0,2)$value

E_Y <- integrate(function(y){y*(1+3*y*y)/2},0,1)$value

cat("The value of E(XY) is",E_XY$value,"and the value of E(X)*E(Y) is",E_X*E_Y,"Hence proved.")
```

Chapter 5

Some Discrete Probability Distribution

R code Exa 5.1 Binomial Distribution Problem

```

# Chapter 5
# Example 5.1 Page no. 145 from the pdf..
# Binomial Distribution Problem..

# to find the prob. that exactly 2 in 4 components
    test survive..

# given prob. of component surviving test 0.75

cat("The probability that exactly 2 will surivie in
    this test out of 4 is",dbinom(2,4,prob = 0.75))
```

R code Exa 5.2.a Binomial Distribution Problem

```
# Chapter 5
# Example 5.2a page no. 146 from the pdf..
# Binomial Distribution Problem..

# to find probability that atleast 10 survive in a sample of 15 people from a rare blood disease which have the probability of recovering from it 0.4.

cat("The probability of atleast 10 will survive from a sample of 15 people is ", 1-pbinom(9,size = 15,prob= 0.4))

#this problem can also be solved by dbinom function but we need to do it from 0 to 9 so pbinom function is simple to use in this example..

#The answer in the book is approximated to 0.0338.
```

R code Exa 5.2.b.c Binomial Distribution Problem

```

# Chapter 5
# Example 5.2b and 5.2c page no. 146 from the pdf..
# Binomial Distribution Problem..

# to find probability that from 3 to 8 survive in a sample of 15 people from a rare blood disease which have the probability of recovering from it 0.4.

cat("The probability that from 3 to 8 survive is",
    pbinom(8,15,0.4)-pbinom(2,15,0.4))

# the answer in the textbook is 0.8779 which is
```

```
rounded off to the answer I got here..

#to find probability that exactly 5 survive in a sample of 15 people from a rare blood disease which have the probability of recovering from it 0.4.

cat("The probabilty that exactly 5 survive is", dbinom(5,15,0.4))

# the answer in the textbook is rounded off to 0.1859 and the deviation from the answer in both cases is less than 2%..
```

R code Exa 5.3 Binomial Distribution Problem 3

```

# Chapter 5
# problem 5.3 on page 146 of pdf..
# Binomial Distribution Problem..

# given - p = 0.03, n = 20

cat("The probability that there will be at least one defective item among these 20 is",1- dbinom(0,20,0.03))

# problem 5.3 b part..
# testing of each shipment viewed as a bernoulli trial with p = 0.4562 from previous part..

cat("The probability that there will be exactly 3 shipments each containing at least one defective device among the 20 that are selected and tested from the shipment is",dbinom(3,10,0.4562))
```

R code Exa 5.4 Binomial Distribution Problem 4

```r

# Chapter 5
# example 5.4 page no. 148 from pdf..
# Binomial Distribution problem

# given p = 0.3, n = 10, binomial distribution

cat("The probability that exactly 3 wells have the
    impurity assuming that the conjecture is correct
    is",dbinom(3,10,0.3))

# example 5.4 b part..

cat("The probability that more than 3 wells are
    impure is",pbinom(3,10,0.3,lower.tail = F))
# the answer in the T.B is approx. to 0.3504..
```

R code Exa 5.6 Binomial Distribution Problem 6

```r

# Chapter 5
# example 5.6, page no. 148 from the pdf..
# Binomial Distribution Problem..

# Note- Example 5.5 not solved as the solution
    involves only the multiplication of 2 numbers..

# given - p = 0.3, n = 10, binomial distribution

```

```
cat("The probability that more than 6 are found to
    contain the impurity from 10 wells is",1-pbinom
    (5,10,0.3))

print("As the prob. is 0.0473 approx., this casts
    considerable doubt on conjecture and suggests
    that the impurity problem is more severe.")
```

R code Exa 5.7 Multinomial Distribution Problem

```

# Chapter 5
# Example 5.7 page no. 150 from the pdf..
# Multinomial Distribution..

# to find the prob. of 6 randomly arriving planes
    which are distributed in 3 runways and each have
    a ideal arriving prob.

p_runways <- c(2/9,1/6,11/18)

# prob. of runway 1 :2
        #  runway 2: 1
        #  runway 3 : 3

x <- c(2,1,3)   # a particular case of the
    distrbution of 6 randomly arriving, written in a
    vactor

# this is the case of multinomial distribution

cat("The probability of this type of distribution of
     6 randomly arriving airplanes is",dmultinom(x,6,
    p_runways))
```

R code Exa 5.8 Hypergeometric Distribution In Acceptance Sampling

```
# Chapter 5
# example 5.8 page no. 153 from pdf..
# Hypergeometric Distribution In Acceptance Sampling
   ..

# radom sampling, testing 3 of the parts out of 10
# condition - if none of the 3 is defective then the
    lot is accepted

# assume lot is truly unacceptable..
cat("The probability that the sampling plan finds
   the lot acceptable is",dhyper(0,m = 2,n = 8,k =
   3))

cat("Thus if the lot is truly unacceptable with 2
   defective parts, this sampling plan will allow
   acceptance roughly",dhyper(0,2,8,3)*100,"% of the
    time, so this plan should be considered faulty."
   )
```

R code Exa 5.9 Hypergeometric Distribution In Acceptance Sampling

```
# Chapter 5
# Example 5.9 Page no. 154 from the pdf..
# Hypergeometric Distribution Problem..

```

```
# to find prob. that exactly 1 defective if there
    are 3 in the entire lot of 40 and 5 are selected
    randomly

# this is the problem of hypergeometric distribution
    ..

cat("The probability that exactly 1 is found
    defective from 5 randomly selected components is"
    ,dhyper(1,3,37,5))
```

R code Exa 5.10 Hypergeometric Probability Distribution Problem

```

# Chapter 5
# example 5.10, page no. 154 from the pdf..
# Hypergeometric Probability function..

# given - lot of 100 items in which 12 are defective
    ..
# to find the prob. that in a sample of 10, 3 are
    defective..
# answer can be found using hypergeometric
    probability function

cat("The probability that in a sample of 10 =, 3 are
     found to be defective is",dhyper(3,m = 12,n =
    88,k = 10))
```

R code Exa 5.11 Mean And Variance Of Hypergeometric Experiment

```

# Chapter 5
```

```
# Example 5.11 page no. 155 from the pdf..
# Mean And Variance Of Hypergeometric Experiment..

# to find the mean and variance of the random
    variable which has hypergeometric distribution..

# mean..
me <- sum(0:3*dhyper(0:3,3,37,5))

# variance..
v <- sum(0:3*0:3*dhyper(0:3,3,37,5))-me*me

cat("The mean of this experiment is",me,"And
    Variance is",v)
```

R code Exa 5.12 Relation Between Hypergeometric and Binomial Distribution

```

# Chapter 5
# example 5.12, page no. 155 from the pdf..
# Relation Between Hypergeometric and Binomial
    Distribution..

# given - out of 5000, 1000 are slightly blemished..
# to find the prob. that exactly 3 are blemished if
    one purchases 10 tires at random

# since the N is large (5000) relative to sample
    size 10 we can do approx. to binomial
    distribution..


cat("The probability of obtaining exactly 3
    blemished tires from 10 randomly purchased
```

```
samples is",dbinom(3,10,0.2))
```

R code Exa 5.13 Multivariate Hypergeometric Distribution Problem

```

# Chapter 5
# Example 5.13 page no. 156 from the pdf..
# Multivariate Hypergeometric Distribution..

# to find the multinomial hypergeometric
    distribution of the following problem..
# Package Used - "extraDistr" , referance- Internet
    ..

# If you have already installed it remove the
    command below..
install.packages("extraDistr") # this package
    contain functions for various types of prob.
    distribution

library("extraDistr")    # to use this package..

#to find the prob. of 1 out 3 , 2 out 0f 4 and 1 out
     of 3, if 5 people are randomly selected from 10
    people

cat("The probability that the 5 randomly selected
    from 10 has the above distribution is",dmvhyper(
    matrix(c(1,2,2),ncol = 3),c(3,4,3),5))
```

R code Exa 5.14 Negative Binomial Distribution Problem

```

```

```
# Chapter 5
# example 5.14 page no. 159 from the pdf..
# Negative Binomial Distribution Problem..

# given A has 0.55 prob. of winning over B

# prob. that A will win in 6 games given that if any team win 4 out 7 wins then that team winner

# using negative binomial distribution we get..

cat("The probability that A will win the series in 6 games is",dnbinom(2,4,0.55))

cat("The probability that team A will win the series is",pnbinom(3,4,0.55))
# for this atleast 4 matches to be won I can also use dnbinom but I have to sum all from size 4 to 7


# now A and B are playing and winning series decided by 3 wins out of 5 games

cat("The probability that A will win the playoff is",pnbinom(2,3,0.55))

#The answer in the textbook is approximated to 4 decimal places..
```

R code Exa 5.15 Geometric Distribution Problem

```

# Chapter 5
# example 5.15 page no. 160 from the pdf..
```

```
# Geometric Distribution Problem..

# find prob. that 5th item inspected is first
    defective found given 1 in every 100 items is
    defective..

# using geometric distribution..

cat("The probability that 5th item inspected is
   found to be defective is",dgeom(4,0.01))

# the answer in textbook is approximated, 4 digits
    to the decimal..
```

R code Exa 5.16 Geometric Distribution Problem

```

# Chapter 5
# example 5.16, page no. 160 from the pdf..
# Geometic Distribution Problem..

# p = 0.05 of a connection during busy time..
# to find the prob. that 5 attempts are necessary
    for a successful call.

cat("The probability that 5 attempts are necessary
    for a successful call is",dgeom(4,0.05))
```

R code Exa 5.17 Poisson Distribution Problem

```

# Chapter 5
# Example 5.17 page no. 162 from the pdf..
```

```r
# Poisson Distribution Problem..

# given - avg no. of radioactive particle entering in 1msec is 4,

# to find- prob. of 6 particles entering in a given msec..

# using poisson distribution..

cat("The probability of 6 particles entering in a given millisecond is",dpois(6,4))

# the answer in the textbook is 0.1042 approx. to 4 decimal places..
```

R code Exa 5.18 Poisson Distribution Problem

```r

# Chapter 5
# example 5.18, page no. 162 from the pdf..
# Poisson Distribution Problem..

# avg no. of oil takers arriving each day = 10,
# facilities at port can handle 15 tankers per day.
# to find the prob. that on a given day tankers have to be turned away..


cat("The probability that on a given day tankers have to be turned away is",ppois(15,10,lower.tail = F))
```

R code Exa 5.19 Approximation of Binomial By a Poisson Distribution

```
# Chapter 5
# ex. 5.19 page no. 164 from the pdf..
# Approximation of Binomial By a Poisson
    Distribution..

# given P(accident on given day)=0.005 and
    independent of each other..

# to find a. prob. of accident on 1 day from 400
    days..

# this is the case of binomial distribution but in
    this case we can use poisson approx.

# lambda= np=2 in this case..

# answer using binomial ..
cat("The probability that there will be accident on
    1 day in a given period of 400 days is",dbinom
    (1,400,0.005))

# answer using poisson..
cat("Using poisson distribution the above answer
    becomes",dpois(1,2))
# so when n is very large and p is very small we can
     convert binomial to poisson..
# the margin of error is 1.2e-06 which negligible so
     we can use this conversion

# to find prob. of accident for atmost 3 days from
    400 days..

# again we can use poisson here..

cat("The probability that there are atmost 3 days
```

```
    with an accident is",ppois(3,2))

# the answer in the T.B is approx. to 3 decimal
    places 0.857.
```

R code Exa 5.20 Approximation of Binomial By a Poisson Distribution

```

# Chapter 5
# example 5.20 page no. 164 from the pdf..
# Approximation of Binomial By a Poisson
    Distribution..

# 1 in every 100 items produced has one or more
    bubbles means p = 0.001..
# to find the prob. that a random sample of 8000
    will yield fewer than 7 items processing bubbles,
# in short find P(X < 7)

# since p is very close to 0 and n is very large we
    can approx. it with Poisson distribution..

cat("The prob. that the random sample of 8000 will
    yield fewer than 7 items possessing bubbles is",
    ppois(6,8))
```

Chapter 6

Some Continuous Probability Distributions

R code Exa 6.1 Uniform Distribution Problem

```

# Chapter 6
# Example 6.1 page no. 171 from the pdf..
# Uniform Distribution Problem..

# to find the prob. density function of uniform
    distribution over 0 to 4

a <- function(x){

  if(x>=0 & x<=4){

    dunif(x,0,4)

  }

}

# P(X>=3)
```

```r

cat("The probability that any conference last at
   least 3 hours is",1-punif(3,0,4))
```

R code Exa 6.2 Normal Distribution Problem

```r

# Chapter 6
# Example 6.2 page no.178 from the pdf..
# NOrmal Distribution Problem..

# given standard normal distribution..


cat("The area under the curve to the right of 1.84
   is",pnorm(1.84,lower.tail = F))


cat("The area under the curve that lies between z
   =-1.97 to z=0.86 is",pnorm(0.86)-pnorm(-1.97))


# the answer in the textbook is approx.to 4 decimal
   places..
```

R code Exa 6.3 Find The Value of k given Normal Distribution

```r

# Chapter 6
# Example 6.3 page no. 179 from the pdf..
# Normal Distribution..

# to find the k such that..
```

```

cat("The value of k such that P(Z>k)= 0.3015 is", qnorm(0.3015,lower.tail = F))

cat("The value of k such that P(k<z<-0.18)=0.4197 is ",qnorm(pnorm(-0.18)-0.4197))
```

R code Exa 6.4 Find area under the Normal Distribution curve

```

# Chapter 6
# Example 6.4 page no. 179 from the pdf..
# Find area under the Normal Distribution..

# given popukation mean= 50 ,sd= 10

cat("The probability that random variable X assumes a value between 45 and 62 is",pnorm(62,50,10)-pnorm(45,50,10))
```

R code Exa 6.5 Find the Probability given Normal Distribution

```

# Chapter 6
# Example 6.5 page no. 180 from the pdf..
# Find the Probability given Normal Distribution..

# X normal distribution ,
# mu= 300, sigma= 50
# find P(X > 362)

```

```
cat("The probability that X assumes a value greater than 362 is",pnorm(362,mean = 300,sd = 50,lower.tail = F))
```

R code Exa 6.6 Using Normal Curve In Reverse

```

# Chapter 6
# Example 6.6 page no. 181 from the pdf..
# Using Normal Curve In Reverse..


# given population mean= 40, sd= 6

cat("The value of x that has 45% of the area to the left is",qnorm(0.45,40,6))

# the deviation of this value from the T.B is less than 2%.

cat("The value of x that has 14% of the area to the right is",qnorm(0.14,40,6,lower.tail = F))

#The answer in T.B is approximated to 2 decimal places..
```

R code Exa 6.7 Applications of Normal Distribution

```

# CHapter 6
# Example 6.7 page no. 182 from the pdf..
# Applications of Normal Distribution..

```

```r
# given - average battery life- 3 years
# standars deviation - 0.5
# to find P(X<2.3)

cat("The probability that a given battery will last
    less than 2.3 years is",pnorm(2.3,mean = 3,sd=
    0.5))
```

R code Exa 6.8 Applications of Normal Distribution

```r

# Chapter 6
# example 6.8 page no. 182 from the pdf..
# Applications of Normal Distribution..


# given distribution of life of bulb - normal, mean=
     800, sd=40hrs

cat("The probability that a bulb burns between 778
    and 834 hours is",pnorm(834,800,40)-pnorm
    (778,800,40))


# making a visual of distribution
# Optional not compulsory..
x <- seq(680,920,length=200)
y <- dnorm(x,800,40)
plot(x,y,type="l")
abline(v = c(778,834))
```

R code Exa 6.9 Applications of Normal Distribution

```r

# Chapter 6
# Example 6.9 page no. 182 from the pdf..
# Applications of Normal Distribution..


# given normal, mean= 3,sd= 0.005

# specification on diameter of ball bearing 3+0.01 and 3-.01

cat("On average",2*pnorm(2.99,3,0.005)*100,"% of ball bearings will be scrapped")
```

R code Exa 6.11 Applications of Normal Distribution

```r

# Chapter 6
# Example 6.11 page no. 184 from the pdf..
# Applications of Normal Distribution..

# Note - We just have to calculate d in example 6.10, also many questions of same type have been solved previously..

# given mean resistance= 40 and sd= 2 , normal distribution


#  to find %age of resistors exceeding 43 ohms

cat("The %age of resistors having resistance exceeding 43ohms is",100*pnorm(43,40,2,lower.tail = F),"%")
```

R code Exa 6.12 Applications of Normal Distribution

```
# Chapter 6
# Example 6.12 page no.184 from the pdf..
# Applications of Normal Distribution..

# to find the % of resistance exceeding 43 ohms from example 6.11, if resistance is measured to the nearest ohm.

# mean resistance = 40, sd= 2, normal distribution..

# here we will assign measurement of 43 ohms to all resistors whose value are greater than 43 and lesser than 43.5


cat("The difference between resistance that exceed 43 ohms and that exceeds beyond 43 is",-pnorm(43.5,mean = 40,sd=2,lower.tail = F)+pnorm(43,mean = 40,sd=2,lower.tail = F),"so this value in %age represents all resistance grater than 43 and less than 43.5 that are now being recorded as 43 ohms")
```

R code Exa 6.13 Applications of Normal Distribution

```
# Chapter 6
# Example 6.13 page no. 185 from the pdf..
# Applications of Normal Distribution..
```

```
# given - avg score of the class- 74,sd=7..
# given - 12% score "A"s and grades followsa normal
    distribution

# to find the lowest possible score for A and
    highest for B..


cat("The percentile for 0.12 in this distribution is
    ",qnorm(0.12,74,7,lower.tail = F),"marks")

cat("The lowest score for A is 83","and The highest
    score for B is 82")
```

R code Exa 6.14 Applications of Normal Distribution

```

# Chapter 6
# example 6.14 page no.185 from the pdf..
# Applications of Normal Distribution..

# given - mean 74, sd - 7
# to find the sixth decile..

cat("The sixth decile is",qnorm(0.6,mean = 74,sd =
    7))


# the answer varies slightly due to approximation
    used in the T.B
```

R code Exa 6.15 Normal Approximation To the Binomial

```

# Chapter 6
# Example 6.15 Page no. 191 from the pdf..
# Normal Approximation To the Binomial..

# given- P(patient recovers from disease)=0.4
# to find prob. that less than 30 survive if 100
   have prone to this diesease

# since n=100, this binomial variable can be approx.
    to normal model
p_survive=0.4

mu <- 100*p_survive

sig <- sqrt(100*p_survive*(1-p_survive))

# no. of people are integers , so we need to use
   approx , and so to obtain prob , we need to find
   area to the left of 29.5

cat("The probability that fewer than 30 of the 100
   patients survive is given by",pnorm(29.5,mu,sig))
```

R code Exa 6.16 Normal Approximation To the Binomial

```

# Chapter 6
# Example 6.16 page no. 192 from the pdf..
# Normal Approximation To the Binomial..

# to find the prob. that sheer guesswork yields from
    25 to 30 correct answer for 80 of the 200
# problem about which student has no knowledge..

```

```
# since the sample size is large we can do normal
    approx.

p <- 0.25    # prob. of guessing a correct answer.

n <- 80

# now using normal approx. mu = p*n and sd <- sqrt(n
    *p*(1-p))

mu = p*n

sd <- sqrt(n*p*(1-p))


cat("The prob. of correctly guessing from 25 to 30
    questions is",pnorm(30.5,mu,sd)-pnorm(24.5,mu,sd)
    )
```

R code Exa 6.17 Exponential Distribution Problem

```

# Chapter 6
# Example 6.17 page no. 197 from the pdf..
# Exponential Distribution Problem..

# given - T- time of failure
# T is exponential distribution beta=5
# find prob. that atleast 2 are functioning at the
    end of 8 years, if 5 of these components are
    taken..


bet <- 5

```

```
# first to find probability of a component functioning after 8 years


a <- pexp(8,rate=1/bet,lower.tail = F)

# finding total probability..

cat("The probability that atleast 2 are still functioning from 5 components after 8 years is",1-pbinom(1,5,a))

# the answer in the textbook is approximated to 0.2(P(T>8)) so there is also small error in the final probab.

# using computation we can reduce such error by a large extent..
```

R code Exa 6.18 Gamma Distribution Problem

```

# Chapter 6
# Example 6.18 page no. 178.
# Gamma Distribution Problem..

# given - beta- 1/5, alpha- 2

# to find the probability up to a minute will elapse by the time 2 calls have come in to the switchboard..

cat("The probability that up to a minute will elapse by the time 2 calls have come in to the switchboard is",pgamma(1,shape = 2,scale = 1/5))
```

```
# the answer in T.B is approx. to 0.96
```

R code Exa 6.19 Application of Gamma Distribution

```
# Chapter 6
# Example 6.19 page no.198 from the pdf..
# Application of Gamma Distribution..

# given - X is survival time..has gamma distribution
    with
# alpha = 5, beta =10

# find P(rat survives <=60) or P(X<=60)

# using gamma distribution..


cat("The probability that the rat survives no longer
    than 60 days is",pgamma(60,shape=5,scale= 10))

# The answer in the T.B is approx. to 3 decimal
  places..
```

R code Exa 6.20 Application of Gamma Distribution

```
# Chapter 6
# ex. 6.20 page no. 199 from the pdf..
# Application of Gamma Distribution..

```

```

# given - beta= 4, alpha= 2,
# changes were made to tighten quality control
    requirement ,
# after these 20 months were passed , so we are asked
     to find that if the quality control tightening
    was effective ..


cat("The p-value for such time is",pgamma(20,shape =
     2,scale = 4,lower.tail = F))

print("Since the p- value is less so conditions of
    gamma distributions with alpha=2,beta=4 are not
    supported by data , hence we can conclude that
    quality control work was effective.")
```

R code Exa 6.21 Cumulative Distribution For Exponential Distribution Problem

```

# Chapter 6
# Example 6.21 page 199 from the pdf..
# Function Which resembles Exponential Distribution
    ..

# given a density function which resembles
    exponential function
# with beta= 4
# to find P(Y > 6)


cat("The value of P(Y > 6) is",pexp(6,rate = 1/4,
    lower.tail = F))

```

```r
cat("the above statement can interpret as the prob. that the washing machine wil require major repair after year 6 is",pexp(6,rate = 1/4,lower.tail = F))

cat("The prob. that a major repair is necessary in the first year is",pexp(1,rate = 1/4))
```

R code Exa 6.22 Lognormal Distribution Problem

```r

# Chapter 6
# Example 6.22 page no. 202 from the pdf..
# Lognormal Distribution Problem..

# given X= pollutant concentration,follows log normal distribution
# parameters of X- mu= 3.2 and sd = 1

# to find P(X>8)..?

#Using lognormal distribution

cat("The probability that the concentration exceeds 8 parts per million is",plnorm(8,meanlog = 3.2, sdlog = 1))

# the error in the T.B and computed answer is 0.0001621 which is very small and can be neglected..
```

R code Exa 6.23 Find Percentile From Lognormal DIstribution

```
# Chapter 6
# example 6.23 page number - 203 from pdf..
# Lognormal Distribution Problem..

# given a lognormal distribution with mu = 5.149 and sigma = 0.737
# to find the 5th percentile of the life of electronic control...


cat("The 5th percentile of such an distribution is", qlnorm(0.05,meanlog = 5.149, sdlog = 0.737))
```

R code Exa 6.24 Weibull Distribution Problem

```
# Chapter 6
# Example 6.24 on page no. 204 of pdf..
# Weibull Distribution Problem..

# given X is weibull distribution with alpha = 0.01 and beta = 2
# to find P(X < 8)

cat("The probability that life of X falls before 8 hours of usge is",pweibull(8,shape = 2,scale = sqrt(100)))

# Note - Chapter 7 contains only theoretical problems, so I am not solving that chapter, the chapter contains 7 problems total and all are theoretical with nothing or very less to compute anything..
# I will start with 8th Chapter..
```

Chapter 8

Fundamental Sampling Distributions And Data Descriptions

R code Exa 8.1 Sample Mode Evaluation

```

# Chapter 8
# Example 8.1 page no. 228 from the pdf..
# Sample Mode Evaluation..

# to find the mode of the given data
mod <- function(v){

  a <- unique(v)
  a[which.max(tabulate(match(v,a)))]


}
set <- c
   (0.32,0.53,0.28,0.37,0.47,0.43,0.36,0.42,0.38,0.43)

```

```
cat("The sample mode is",mod(set))
```

R code Exa 8.2 Sample mean and Sample Variance

```

# Chapter 8
# Example 8.2 page no. 229 from the pdf..
# Sample mean and Sample Variance..

# to find mean and variance..

set <- c(12,15,17,20)

cat("The mean and variance are",mean(set),"and",var(
    set),"respectively")
```

R code Exa 8.4 Central Limit Theorem

```

# Chapter 8
# Example 8.4 page no. 234 from the pdf..
# Central Limit Theorem..

# given - bulb life normally distributed , mean= 800
    ,,sd= 40

# prob. that average life is less than 775hrs for a
    random sample of 16 bulbs..

sd <- 40

n <- 16

```

```
# Using central limit theorem..

# sampling distribution..

sam_mean <- 800

sam_sd <- sd/sqrt(n)

cat("The probability that the random sample of 16
   have average lifetime less than 775hrs is",pnorm
   (775,sam_mean,sam_sd))
```

R code Exa 8.5 Central limit theorem

```

# Chapter 8
# Example 8.5 page no. 237 from the pdf..
# Central limit theorem..

# mu= 28, sd= 5,n= 40, to find prob. that
   averagetransport time greater than 30 min..

# P(Xbar > 30)


cat("The probability that the average transport time
    was more than 30 min. is",pnorm(30+0.5,mean =
   28,sd= 5/sqrt(40),lower.tail = F))
```

R code Exa 8.6 Sampling Distribution of the difference of 2 means

```

# Chapter 8
```

```
# Example 8.6 page no. 240 from the pdf..
# Sampling Distribution of the difference of 2 means ..

# given information about 2 manufacturers


A <- c(6.5,0.9,36)


B <- c(6,0.8,49)


# prob. 36 tubes from A have mean lifetime at least 1 more than mean of sample of 49 from B??

diff_mean <- A[1]-B[1]

diff_sd <- sqrt((A[2]*A[2]/A[3])+ (B[2]*B[2]/B[3]))

cat("The probability that the mean lifetime for 36 tubes from A will be at least 1 year longer than the mean lifetime for 49 tubes from B is",pnorm(1,diff_mean,diff_sd,lower.tail = F))
```

R code Exa 8.7 The Chi squared Distribution Problem

```

# Chapter 8
# Example 8.7 page no. 245 from the pdf..
# The Chi-squared Distribution Problem..

# given mu=3,sd=1,data of 5 batteries,
# to show whether manufacturer should be convinced that batteries have a sd of 1..
```

```r

sd = 1

obs <- c(1.9,2.4,3.0,3.5,4.2)

chi.sq <- var(obs)*(length(obs)-1)/(sd^2)

cat("The value of chi^2 is",chi.sq,"with 4 degrees of freedom")

cat("Since 95% of the chisq values with 4 degrees of freedom fall between",qchisq(0.025,df=4),"and",qchisq(0.975,df= 4),"the computed value with sig^2= 1 is reasonable and therefore manufacturer has no reason to suspect that the sd is other than 1")
```

R code Exa 8.8 The t distribution Problem

```r

# Chapter 8
# Example 8.8 page no. 249 from the pdf..
# The t-distribution Problem..

# to find the t value with left tail area 0.025 and degrees of freedom =14

cat("The t value with degrees of freedom 14 and leaves an area of 0.025 to the left is",qt(0.025,df= 14))

# Note- In example 8.9 the degrees of freedom is not mentioned so I am not solving example 8.9, I think the solution is wrong..
```

R code Exa 8.10 The t distribution Problem

```
# Chapter 8
# Example 8.9 page no. 249 from the pdf..
# The t- distribution Problem


# find k such that P(k<T<-1.761)= 0.045, sample size
   = 15,normal distribution

n <- 15

cat("The value of k such that P(k<T<-1.761)= 0.045
   is",qt(pt(-1.761,n-1)-0.045,n-1))
```

R code Exa 8.11 Applications of t distribution

```
# Chapter 8
# Example 8.11 page no. 250 from the pdf..
# Applications of t-distribution..

# mu 500, xbar= 518 , sample sd = 40,
# if t value between t(-0.05) and t(0.05) then the
   engineer is satisfied.

xbar <- 518

mu <- 500

sd <- 40
```

```
n <- 25

t_value = (xbar-mu)/(sd/sqrt(n))

cat("The engineer is satisfied if a sample of 25
    batches yields a t value between",qt(0.05,df=24),
    "and",-qt(0.05,df=24))

cat("The t value is",t_value,"which is above the
    levels mentioned above so the p value is",pt(t_
    value,df=24,lower.tail = F),"so we conclde that
    the process produces a better product than the
    engineer thought.")


# the answer varies slightly due to approximation in
     book.
```

R code Exa 8.12 To Construct a Normal quantile quantile plot

```

# Chapter 8
# Example 8.12 page no. 258 from the pdf..
# To Construct a Normal quantile-quantile plot..

# to plot the normal quantile-quantile plot of two
    data samples and draw conclusion whether they are
     from same normal distribution.

y <-c
    (5030,13700,10730,11400,860,2200,4250,15040,4980,11910,8130,26850

b <- c
```

```
    (2800,4670,6890,7720,7030,7330,2810,1330,3320,1230,2130,2190)


qqnorm(y,ylim = range(y,b),col = "blue")
par(new= TRUE)
qqnorm(b,ylim=range(y,b),col = "red")


# The clustering of observation would make it seem
    unlikely that two sample came from common normal
    distribution.
```

Chapter 9

One And Two Sample Estimation Problems

R code Exa 9.2 Confidence Interval Evaluation

```

# Chapter 9
# Example 9.2 page no. 271 from the pdf..
# COnfidence Interval Evaluation..

# to find the confidence interval given mu, sigma

n <- 36

mu <- 2.6

sig <- 0.3

a <- qnorm(0.975)*sig/sqrt(n)

b <- qnorm(0.995)*sig/sqrt(n)

cat("The 95% confidence interval for the above
    parameters is",mu-a,mu+a,"milliliter")
```

```
cat("The 99% confidence interval for the above parameters is",mu-b,mu+b,"milliliter")


# the similar question can be addressed as shown below
set.seed(100)
v <- rnorm(36,2.6,0.3)
t.test(v,conf.level = 0.95)$conf.int

# but as we can see each time different ci are there , if n becomes very large then it will be very close to the standard answer..
```

R code Exa 9.3 To Find the sample size

```

# Chapter 9
# Example 9.3 page no. 273 from the pdf..
# To Find the sample size..


# to find the samaple size if we want to be 95% confident that our estimate of mu in ex. 9.2 is off by less than 0.05

sd <- 0.3

n <- ((qnorm(0.975)*sd)/0.05)^2

cat("We can be 95% confident that a random sample of size",as.integer(n)+1,"will provide an estimate xbar differing from mu by an amount less than 0.05")
```

R code Exa 9.4 To find 95 Percent bound for mean

```
# Chapter 9
# example 9.4 page no. 274
# To find 95% bound for mean..

# to find the 95% bound for the mean reaction time

avg_t <- 6.2

n <- 25

s_d <- 2

cat("The 95% bound is given by",avg_t+qnorm(0.95)*s_
   d/sqrt(n))
```

R code Exa 9.5 To Find 95 Percent Confidence Interval for mean

```
# Chapter 9
# Example 9.5 page no. 275 from the pdf..
# To Find 95% Confidence Interval for mean..


# to calculate the 95% confidence interval for the
   contents of 7 containers of sulphuric acid


a <- c(9.8,10.2,10.4,9.8,10.0,10.2,9.6)
```

```
cat("The confidence interval for the above data
    sample is"," ")

t.test(a,conf.level = 0.95)$conf.int
```

R code Exa 9.6 To Find 99 Percent Confidence Interval

```

# Chapter 9
# Example 9.6 page no. 276 from the pdf..
# To Find 99% Confidence Interval..

# to find 99% confidence interval on the mean SAT
    maths score.
# Since random sample is large 500 we can use normal
     approximation..

mu <- 501

sd <- 112

n <- 500

sd1 <- sd/sqrt(n)

cat("The 99% confidence interval for the above
    example is",mu+qnorm(0.005)*sd1, mu-qnorm(0.005)*
    sd1)
```

R code Exa 9.7 To find 95 Percent Prediction Interval

```
# Chapter 9
# Example 9.7 page no. 278 from the pdf..
# To find 95% Prediction Interval..

# to find the 95% prediction interval for the loan
    amount.
# given ..
mu <- 257300

sd <- 25000

n <- 50

c <- sqrt(1+1/n)

cat("The 95% prediction interval for the future loan
     amount is",mu-qnorm(0.9750,lower.tail = T)*sd*c,
    mu+qnorm(0.9750,lower.tail = T)*sd*c)
```

R code Exa 9.8 To find 99 Percent prediction interval

```
# Chapter 9
# Example 9.8 page no. 279 from the pdf..
# to find 99% prediction interval..

# sample size = 30,mean = 96.2%, sd = 0.8

mu <- 96.2

sd <- 0.8

n <- 30

```

```
c <- sqrt(1+1/n)

t_val <- abs(qt(0.005,df= 29))

cat("The prediction interval for a new observation
   is",mu-t_val*sd*c,"and",mu+t_val*sd*c)
```

R code Exa 9.10 Confidence Interval for the difference between 2 means

```

# Chapter 9
# Example 9.10 page no. 286 from the pdf..
# Confidence Interval for the difference between 2
   means..

# to find 96% confidence interval for the given
   information
# given..
a_mu <- 36

b_mu <- 42

n_a <- 50

n_b <- 75

sd_a <- 6

sd_b <- 8

c <- qnorm(0.98)*sqrt((sd_a*sd_a/n_a)+(sd_b*sd_b/n_b
   ))

d <- b_mu - a_mu

```

```
cat("The 96% confidence interval of the difference
    in average gas mileage is",d-c,d+c)
```

R code Exa 9.11 Confidence Interval Evaluation Using Pooled Estimate of Variance

```

# Chapter 9
# Example 9.11 page no. 288 from the pdf..
# COnfidence Interval Evaluation Using Pooled
    Estimate of Variance..

# to find the 90% confidence interval of the given
    data

# the variance of both sampling stations are equal
    and and independent
n1 <- 12

x1 <- 3.11

s1 <- 0.771

n2 <- 10

x2 <- 2.04

s2 <- 0.448

# WE NEED to find pooled estimate for equal
    variances..

sp <- sqrt(((n1-1)*s1*s1+(n2-1)*s2*s2)/(n1+n2-2))

dfs <- n1+n2-2
```

```
c <- qt(0.95,dfs)*sp*sqrt((1/n1)+(1/n2))

cat("The 90% confidence interval for mu1-mu2 is",x1-
   x2-c,x1-x2+c)

# the answer in the textbook is approximated to 3
   decimal places..
```

R code Exa 9.12 Confidence Interval Evaluation For difference of 2 means when their variance is not equal and Unknown

```

# Chapter 9
# Example 9.12 page no. 290 from the pdf..
# Confidence Interval Evaluation For difference of 2
   means when their variance is not equal and
   Unknown..

# given unequal variances,independent samples,approx
   . normal

a <- c(3.84,3.07,15)  # parameters for 1st station

b <- c(1.49,0.80,12)  #parameters for second station

# to find the 95% ci for mu1-mu2

v <- (((a[2]*a[2]/a[3] + b[2]*b[2]/b[3])^2)/((((a[2]
   *a[2]/a[3])^2)/(a[3]-1))+(((b[2]*b[2]/b[3])^2)/(b
   [3]-1))))

c <- qt(0.975,v)*sqrt((a[2]*a[2]/a[3]) + (b[2]*b[2]/
   b[3]))

```

```r
cat("The 95% confidence interval for mu1-mu2 is",a[1]-b[1]-c,a[1]-b[1]+c,"milligram")


# The answer in T.B is rounded to 2 decimal places.
```

R code Exa 9.13 To find Confidence Intervals for Paired Observations

```r

# Chapter 9
# Example 9.13 page no. 293 from the pdf..
# to find Confidence Intervals for Paired Observations..

levels_plasma <- c(2.5,3.1,2.1,3.5,3.1,1.8,6.0,3.0,36.0,4.7,6.9,3.3,4.6,1.6,7.2,1.8

levels_fat <- c(4.9,5.9,4.4,6.9,7.0,4.2,10.0,5.5,41.0,4.4,7.0,2.9,4.6,1.4,7.7,1.


cat("The 95% confidence interval for paired observation is", " ")

t.test(x= levels_plasma,y= levels_fat,paired=TRUE,aternative="two.sided")$conf.int
```

R code Exa 9.14 Estimating a Proportion for Single Sample

```r

# Chapter 9
```

```
# Example 9.14 page no. 297 from the pdf..
# Estimating a Proportion for Single Sample..


# to calculate the 95% confidence interval
# given
n <- 500   # size of random sample

x <- 340   # number subscribed to HBO from te size
    of random sample

cat("The confidence interval for the actual
    proportion of families with TV sets in the city
    subscribed to HBO is"," ")

prop.test(340,500,0.68)$conf.int
```

R code Exa 9.15 Evaluate sample size for estimating a Proportion for Single Sample

```

# Chapter 9
# Example 9.15 page no. 299 from the pdf..
# Evaluate sample size for estimating a Proportion
    for Single Sample..

# to find the sample size required to be 95%
    confident that our estimate of p in ex9.14 is
    within 0.02 of the true value..
# Refer example 9.14 page no. 297 from the pdf..

n <- 500

x <- 340

```

```
p_hat <- x/n

e <- 0.02

sample_size <- qnorm(0.975)^2*p_hat*(1-p_hat)/(e^2)

cat("If we base our estimate of p on a random sample of size",as.integer(sample_size)+1,"we can be 95% confident that our sample proportion will not differ from the true proportion by more than 0.02")
```

R code Exa 9.16 Evaluate sample size for estimating a Proportion for Single Sample

```

# CHapter 9
# Example 9.16 page no. 299 from the pdf..
# Evaluate sample size for estimating a Proportion for Single Sample..

# to find a sample size if we want atleast 95% confidence interval


e <- 0.02      # error not exceeding this value

cat("The sample size if we want atleast 95% confidence interval that our estimate of p is within 0.02 of the true value is",round(qnorm(0.025)*qnorm(0.025)/(4*e*e)))


# In Example 9.17 all we have to do is put the values in the formula mentioned in the pdf and
```

```
get the answer ,I have solved some questions based
 on this but I am not going to solve , that 's
understood from the questions itself and will be
a very straight forward code..
```

R code Exa 9.18 To obtain Confidence Interval of Variance

```

# Chapter 9
# Example 9.18 page no. 304 from the pdf..
# to obtain confidence interval of variance

a <- c
  (46.4,46.1,45.8,47.0,46.1,45.9,45.8,46.9,45.2,46.0)


c <- qchisq(0.025,df = length(a)-1)
b <- qchisq(0.975,df= 9)

cat("The 95% confidence interval for variance is",
  var(a)*(length(a)-1)/b,var(a)*(length(a)-1)/c)
```

R code Exa 9.19 To find Confidence Interval For the Ratio of 2 Variances

```

# Chapter 9
# Example 9.19 page no. 306 from the pdf..
# To find Confidence Interval For the Ratio of 2
  Variances..

# to find 98% confidence interval of 2 independent
  sample variances..
```

```
n1 <- 15

n2 <- 12

s1 <- 3.07

s2 <- 0.80

f1 <- qf(0.99,14,11)

f2 <- qf(0.99,11,14)

cat("The 98% confidence interval for the ratio sigma1/sigma2 is",sqrt(s1*s1/(s2*s2*f1)),sqrt(s1*s1*f2/(s2*s2)))
```

R code Exa 9.22 Maximum Likelihood Estimation Problem

```

# Chapter 9
# example 9.22 page no. 311 from the pdf..
# Maximum Likelihood Estimation Problem..

# Note - Example 9.20 and 9.21 are theoretical with nothing to Compute..

# given survival time in months for 10 rats, to find the maximum estimate of the mean survival time..
# package used "stats4", if it is not installed on your studio then you need to execute the following instruction - install.packages("stats4")

library(stats4)

```

```
y <- c(14,17,27,18,12,8,22,13,19,12)

f <- function(lambda){ -sum(dpois(y, lambda, log =
    TRUE))}

ans <- mle(f, start = list(lambda = 1), nobs = NROW(
    y))

print("The maximum likelihood estimate of the mean
    survival time is")

coef(ans)

# same is the concept for example 9.23 just we have
    to change the function and vector of observations
    ..
```

Chapter 10

One And Two Sample Tests Of Hypothesis

R code Exa 10.3 Test Concerning a Single Mean for Single Sample

```

# Chapter 10
# Example 10.3 page no. 338 from the pdf..
# Test Concerning a Single Mean for Single Sample..

# Note - 10.1 and 10.2 are theoretical with nothing
   to Compute..

# given
m <- 71.8

n <- 100

sd <- 8.9

alpha <- 0.05

# Null hypothesis , mean = 70years
# alternative mean > 70 years
```

```
sd1 <- sd/sqrt(n)

p_value <- 1 - pnorm((m-70)/sd1,0,1)

cat("The p-value is",p_value,",which is less than
    0.05 level of significance hence we conclude the
    evidence favour alternative hypothesis more than
    null")
```

R code Exa 10.4 Test Concerning a Single Mean for Single Sample

```

# Chapter 10
# Example 10.4 page no. 338 from the pdf..
# Test Concerning a Single Mean for Single Sample..

# given
# H0 = 8kg
# H1 != 8kg

alpha <- 0.01


n <- 50

sd <- 0.5

m <- 7.8

sd1 <- sd/sqrt(n)

# two sided hypothesis test..

p_value <- 2*pnorm((m-8)/sd1)
```

```
if(p_value < 0.01){

  print("Since the p_value is less than 0.01 we
     reject the null hypothesis that mu= 8kg")

}else{

  print("Since the p_value is more than 0.01 there
     is not enough evidence to claim the alternative
      hypothesis , so evidence supports mu can be 8")

}
```

R code Exa 10.5 The t statistic for a test on a single mean when Variance is Unknown

```

# Chapter 10
# Example 10.5 page no. 340 from the pdf..
# The t-statistic for a test on a single mean when
   Variance is Unknown..

# given H0- mean = 46kWh
# H1 - mean < 46kWh

alpha <- 0.05

n <- 12

# using t-staistic

m <- 42

sd <- 11.9
```

```r

sd1 <- sd/sqrt(n)

# one sided testing
p_value <- pt((m-46)/sd1,n-1)


if(p_value < 0.05){

  print("Since the p_value is less than 0.05 we
     reject the null hypothesis that mu= 46kWh")

}else{

  print("Since the p_value is more than 0.05, we
     cannot reject H0 and conclude that the average
     no. of kWh used is not significantly less than
     46")

}
```

R code Exa 10.6 The 2 Sample Pooled t test

```r

# Chapter 10
# Example 10.6 page no. 344 from the pdf..
# The 2 Sample Pooled t-test..

# 2 sample t- test
n1 <- 12

n2 <- 10

x1 <- 85

```

```
x2 <- 81

s1 <- 4

s2 <- 5

# Null H0- mu1-mu2= 2
# alternate H1 mu1-mu2 >2
# using 2 sample pooled t-test..

deg_f <- n1+n2-2

sp <- sqrt((s1*s1*(n1-1)+s2*s2*(n2-1))/deg_f)

p_value <- 1-pt(((x1-x2-2)/(sp*sqrt(1/n1+1/n2))),deg_f)

if(p_value < 0.05){

  print("Since the p_value is less than 0.05 we reject the null hypothesis that mu1-mu2=2")

}else{

  print("Since the p_value is more than 0.05, we cannot reject H0 and conclude that material 1 wear exceeds that of 2 by more than 2 units")

}
```

R code Exa 10.7 Choice of Sample Size for Testing Mean

```

# Chapter 10
# Example 10.7 page no. 351 from the pdf..
```

```r
# Choice of Sample Size for Testing Mean..

# given the hypothesis test, using alpha= 0.05,
# to find the sample size if power of our test is to be 0.95 when true mean is 69 kg.
# Null - mu = 68 kg.
# alternative - mu > 68 kg.

mu = 68

mu1 = 69

alpha  = 0.05

sig <- 5

delta <- mu1-mu

z_alpha <- abs(qnorm(0.05))

n <- ((z_alpha+z_alpha)^2)*(sig*sig)/(delta*delta)

cat("The sample size required if the test is to reject the null hypothesis 95% of the time when in fact, mu is as large as 69 kg is",n)

# example 10.8 is bit absurd all they have done is mere division and making conclusions based on that so I an not coding example 10.8
```

R code Exa 10.9 One Sample Test on a Single Proportion

```r

# Chapter 10
# Example 10.9 page no. 362 from the pdf..
```

```
# One Sample Test on a Single Proportion..

# given H0 - p=0.7
# H1 alternative - p!=0.7

# test staistic - binomial

x <- 8   # no. of heat pumps installed in sample

n <- 15   # no. of randomly selected homes..

# it is a binomial variable and 2 sided hypothesis
    test..
p_value <- pbinom(x,n,0.7)*2

if(p_value > 0.10){

  cat("As the p value is",p_value,"there is
     insufficient reason to doubt the builder's
     claim so cannot reject null hypothesis")

}else{

  print("The evidence supports alternative
     hypothesis and builder claim should be rejected
     ")

}
```

R code Exa 10.10 One Sample Test on a Single Proportion

```

# Chapter 10
# Example 10.10 page no. 362 from the pdf..
# One Sample Test on a Single Proportion..
```

```
# Null − p = 0.6
# alternative − p > 0.6
# alpha = 0.05.
# critical region − z>1.645

# using binomial approximation..


ans <- pnorm(0.7,mean = 0.6,sd=sqrt(0.6*0.4/100),
    lower.tail = F)

cat("The p_value is",ans,"we reject null hypothesis
    and conclude that the new drug is superior")
```

R code Exa 10.11 Two Sample Tests On 2 Proportions

```

# Chapter 10
# Example 10.11 page 364 from the pdf..
# Two Sample Tests On 2 Proportions..

# hypothesis testing on 2 proportion..using z test
# Null− H0 p1=p2
# alternate − H1 = p1>p2

# information −
x1 <- 120

x2 <- 240

n1 <- 200

n2 <- 500

```

```
# to find the estimate..
p_hat <- c(x1/n1,x2/n2)

pooled_p <- (x1+x2)/(n1+n2)

p_value <- 1-pnorm((p_hat[1]-p_hat[2])/(sqrt((pooled
    _p*0.49)*(1/n1+1/n2))))

cat("Since the p_value is",p_value,"we reject the
    null hypothesis and conclude that town voters
    favouring the proporsal are higher than county
    voters")

# this problem can also be solved directly using
    prop.test(c(120,240),c(200,500),alternative="
    greater"), but this test uses chi square test
    method and the T.B solved it with z-test..
# If you want to explore it write this code into
    your console -
# prop.test(c(120,240),c(200,500),alternative = "
    greater")
```

R code Exa 10.12 One and Two Sample Tests Concerning Variance

```

# Chapter 10
# Example 10.12 page no. 366 from the pdf..
# One and Two Sample Tests Concerning Variance..

# null - sig^2 = 0.81
# alternaitve - sig^2 > 0.81
# alpha = 0.05

n <- 10

```

```
sig <- 0.9
s <- 1.2
chi.sq <- (n-1)*s*s/(sig*sig)
p_value <- pchisq(chi.sq,df=9 ,lower.tail = F)
cat("The p value is",p_value,"there is evidence that sig >0.9")
```

R code Exa 10.13 To do Hypothesis Testing for 2 variances

```
# Chapter 10
# example 10.13 page no. 368 from the pdf..
# To do Hypothesis Testing for 2 variances..
# refer example 10.6 for value of 2 variances
# Null - sig1^2 = sig2^2
#alternaitive - sig1^2 != sig2^2
s1_sq <- 16
s2_sq <- 25
# null reject if f <0.34 or f > 3.11..
ifelse(s1_sq/s2_sq < qf(0.05,11,9) | s1_sq/s2_sq > qf(0.95,11,9),print("The null hypothesis is rejected and conclude there is sufficient evidence that variance differ"),print("Do not reject null hypothesis and conclude that there is insufficient evidence that the variances differ."))
```

R code Exa 10.14 Test For Homogeneity

```
# Chapter 10
# Example 10.14 page no. 377 from the pdf..
# Test For Homogeneity..

# given data test the hypothesis

# H0 - proportion of all political affiliation are
    same for each opinion

# H1 - for at least 1 opinion the proportion is not
    same

# using chi-sq statistic..

for1 <- c(82,70,62)

against <- c(93,62,67)


undecided <- c(25,18,21)

pol_affi <- as.data.frame(rbind(for1,against,
    undecided))

names(pol_affi) <- c('Democrat','Republican','
    Independent')

chisq.test(pol_affi)

cat("The p-value is",chisq.test(pol_affi)$p.value,"
    so we cannot reject null hypothesis")
```

```
# the answer to the chi^2 is aprroximated to 1.53 as
    computed 1.5274 by R..
```

R code Exa 10.15 Testing for several proportion

```

# Chapter 10
# Example 10.15 page no. 378 from the pdf..
# testing for several proportion..

defectives <- c(45,55,70)

nondefectives <- c(905,890,870)

# making a data frame of Variables..
shifts <- as.data.frame(rbind(defectives,
    nondefectives))

# Giving names to the columns..
names(shifts) <- c('Day','Evening','Night')

chisq.test(shifts)

cat("Since the p-value is",chisq.test(shifts)$p.
    value," grater than the significant value 0.025
    we cannot reject null hypothesis but from the p-
    value calculated we also cannot say that the
    proportion of defectives is same ")
```

Chapter 11

Simple Linear Regression And Correlation

R code Exa 11.1 To Estimate the Regression Line

```

# Chapter 11
# Example 11.1 page no. 396 from the pdf..
# To Estimate the Regression Line..


# to estimate the regression line from the given data..
x <- c(3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,

y <- c(5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36

pol <- data.frame(x,y)

line_eq <- lm(y~x,data=pol)
```

```

b <-coefficients(line_eq)

cat("The estimate of regression line is y =",b[1],"+ ",b[2],"*x")
```

R code Exa 11.2 Confidence Interval for Slope of Regression Line

```

# Chapter 11
# Example 11.2 page no. 403 from the pdf..
# Confidence Interval for Slopes of Regression Line ..

# to find the confidence interval of the given data ..

x <- c(3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,

y <- c(5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36

pol <- data.frame(x,y)

line <- lm(y~x,data = pol)

cat("The 95% confidence interval for beta1(i.e slope) in the regression line is",confint(line,'x',level=0.95))

# The answer matches with the answer in the T.B upto 3 decimal places , as T.B answer is approximated
```

```
..
```

R code Exa 11.3 Hypothesis Testing On Slope Of Regression Line

```

# Chapter 11
# Example 11.3 page no. 404 from the pdf..
# Hypothesis Testing On Slope Of Regression Line..

# to test the hypothesis of beta1 slope of
   regression line..
# Null - beta1 = 1.0, alternative < 1.0

x <- c
   (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,

y <- c
   (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36

pol <- data.frame(x,y)

l <- lm(y~x,data = pol)

coefficients(l)

p_value <- function(reg_m,conum,val){

  coefi <- coef(summary(reg_m))
  t <- (coefi[conum,1]-val)/coefi[conum,2]
  pt(abs(t),reg_m$df.residual,lower.tail= F)

}
cat("Since the p value is",p_value(l,2,1)," less
```

```
    than 0.05 suggesting strong evidence that beta1 <
     1.0")


# an alternate method is to use the library car and
    the use linearHypothesis function, since this is
    one sided hypothesis
# we need to divide the result by 2 as the function
    does 2 sided hypothesis..
# If "car" package not installed can be installed by
     install.packages("car")

library(car)
linearHypothesis(l,hypothesis.matrix = c(0,1),rhs =
    1)/2
print("Using linearHypothesis function also we get
    the same result and we can derive the same
    inference as before.")
```

R code Exa 11.4 Confidence Interval For Intercept Of Regression Line

```

# Chapter 11
# Example 11.4 page no. 406 from the pdf..
# Confidence Interval For Intercept Of Regression
    Line..

# to find the confidence interval of the given data
    ..

x <- c
    (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,

y <- c
```

```
   (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36

pol <- data.frame(x,y)

line <- lm(y~x,data = pol)

cat("The 95% confidence interval for beta0(i.e
   intercept) in the regression line is",confint(
   line,'(Intercept)',level=0.95))

# The answer matches with the answer in the T.B upto
    2 decimal places , as T.B answer is approximated
   ..
```

R code Exa 11.5 Hypothesis Testing on Intercept of Regression line

```

# Chapter 11
# Example 11.5 page no. 407 from the pdf..
# Hypothesis Testing on Intercept of Regression line
   ..

# Null - beta0 =0
# alternate - beta0 != 0

x <- c
   (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,

y <- c
   (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36

pol <- data.frame(x,y)   # making data frame of
```

```
    observations..

line <- lm(y~x,data = pol)  # regression model..

a <- summary(line) # summary of the regression line

cat("The p-value is",a$coef[1,"Pr(>|t|)"],"which is
    less than 0.05 hence we conclude that beta0 != 0"
    )
```

R code Exa 11.6 To find the confidence interval of the mean response

```

# Chapter 11
# Example 11.6 page no. 409 from the pdf..
# To find the confidence interval of the mean
    response..

x <- c
    (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,

y <- c
    (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36

pol <- data.frame(x,y)

line <- lm(y~x,data = pol)

# here x0 <- 20%
data <- data.frame(x=20)

c <- predict(line,data, interval = "confidence")

```

```
cat("When solid reduction is 20%, the confidence
    interval of chemical oxygen demand is",c[2],c[3])

# the may vary slightly due to approximations in T.B
    .
```

R code Exa 11.7 To Find Prediction Interval

```

# Chapter 11
# Example 11.7 page no. 411 from the pdf..
# To Find Prediction Interval..

# To find the 95% prediction interval for y0 when x0
    is 20%

x <- c
    (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,

y <- c
    (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36

pol <- data.frame(x,y)

line <- lm(y~x,data = pol)


# making a new data frame to predict..

data <- data.frame(x=20)

d <- predict(line,data, interval = "prediction")

```

```
cat("The 95% prediction interval for y0 when x0=20%
    is",d[2],d[3])


# the answer may vary slightly due to approximation
    ..
```

R code Exa 11.8 Computation of Lack of Fit Sum of Squares

```

# Chapter 11
# Example 11.8 page no. 420 from the pdf..
# Computation of Lack of Fit Sum of Squares..

# to do lack of fit test

# to do lack of fit test I used the package "alr3",
    referance - Internet
# using only anova shows the regression and error,
    not lack of fit and pure error so I used this
    package

install.packages("alr3")   # Remove it if you have
    already Installed the package..

library("alr3")

y <- c
    (77.4,76.7,78.2,84.1,84.5,83.7,88.9,89.2,89.7,94.8,94.7,95.9)


x <- c
    (150,150,150,200,200,200,250,250,250,300,300,300)

dat <- data.frame(y,x)
```

```
c <- lm(y~x,data=dat)

cat("The following table shows analysis of variance
    on temperature field data")

pureErrorAnova(c)
```

R code Exa 11.9 Transformation To Linear Regression Model

```

# Chapter 11
# Example 11.9 page no. 426 from the pdf..
# Transformation To Linear Regression Model..

# given data on P and V find the constants of the
    equation PV^gamma = C

v <- c(50,60,70,90,100) #volume

p <- c(64.7,51.3,40.5,25.9,7.8) # pressure

# the model
l <- lm(log(p)~log(v))

# coefficients of the model..
co <- coefficients(l)

# we modelled the data as lnP = lnC - gamma*lnV +
    additive error.

cat("The value of C is",exp(co[1]),"and the vlaue of
    gamma is",-co[2])
```

R code Exa 11.10 To find Correlation Coefficient

```

# Chapter 11
# Example 11.10 page no. 433 from the pdf..
# To find Correlation Coefficient..

# to find the correlation coefficient between the 2 variables

spec_gra <- c(0.414,0.383,0.399,0.402,0.442,0.422,0.466,0.500,0.514,0.530,0.569

mod_of_rup <- c(29186,29266,26215,30162,38867,37831,44576,46097,59698,67705,66088

cat("The sample correlation coefficient is",cor(spec_gra,mod_of_rup))
```

R code Exa 11.11 To do Hypothesis Testing of linear association between two variables

```

# Chapter 11
# Example 11.11 page no. 434 from the pdf..
# To do Hypothesis Testing of linear association between two variables

# Null H0: po =0
# alternate H1 : p1 !=0
```

```
# alpha = 0.05

spec_gra <- c
   (0.414,0.383,0.399,0.402,0.442,0.422,0.466,0.500,0.514,0.530,0.56

mod_of_rup <- c
   (29186,29266,26215,30162,38867,37831,44576,46097,59698,67705,6608

print("The test is shown below as follows:")

cor.test(spec_gra,mod_of_rup)

cat("The p-value is very less then 0.05, we reject
   the null hypothesis of no linear association")

# example 11.12 is same as example 11.11 except now
   we have to do hypothesis as p0 = 0.9, just
   changed the value for testing, so It is same
   concept just numbers are changed, so I have not
   solved..
```

Chapter 12

Multiple Linear Regression And Certain Non Linear Regression Models

R code Exa 12.1 Estimation Of Equation Of Regression Line and predict values

```

# Chapter 12
# Example 12.1 page no. 445 from the pdf..
# Estimation Of Equation Of Regression Line..

y <- c
   (0.90,0.91,0.96,0.89,1.00,1.10,1.15,1.03,0.77,1.07,1.07,0.94,1.10

x1 <- c
   (72.4,41.6,34.3,35.1,10.7,12.9,8.3,20.1,72.2,24.0,23.2,47.4,31.5,

x2 <- c
   (76.3,70.3,77.1,68.0,79.0,67.4,66.8,76.9,77.7,67.7,76.8,86.6,76.9
```

```
x3 <- c(29.18,29.35,29.24,29.27,29.78,29.39,29.69,29.48,29.09,29.60,29.3

dat <- data.frame(y,x1,x2,x3)

line <- lm(y~x1+x2+x3,data=dat)

c <- coefficients(line)

cat("The regression estimate of the above data is",c[1],"+",c[2],"*x1 +",c[3],"*x2 +",c[4],"*x3")

pre <- data.frame(x1 = 50,x2 = 76.0,x3 = 29.30)

# Prediction of values..
cat("The prediction of the line for this values is")

predict(line,pre)
```

R code Exa 12.2 To find the estimate of polynomial regression

```

# Chapter 12
# Example 12.2 page no. 446 from the pdf..
# to find the estimate of polynomial regression..

x <- c(0,1,2,3,4,5,6,7,8,9)

y <- c(9.1,7.3,3.2,4.6,4.8,2.9,5.7,7.1,8.8,10.2)

pol <- lm(y~x+I(x^2))

c <- coefficients(pol)
```

```
dat <- data.frame(x=2)


cat("The fitted polynomial is",c[1],c[2],"*x +",c
    [3],"*x^2")

cat("When x=2, the estimate is",predict(pol,dat))
```

R code Exa 12.3 Estimate the regression coefficients in polynomial model

```

# Chapter 12
# Example 12.3 page no. 447 from the pdf..
# Estimate the regression coefficients in polynomial
    model.

m1 <- c(14.05,14.93,16.56,15.85,22.41,21.66)

m2 <- c(10.55,9.48,13.63,11.75,18.55,17.98)

m3 <- c(7.55,6.59,9.23,8.78,15.93,16.44)

observation <- c(m1,m2,m3)

temp <- c(rep(75,6),rep(100,6),rep(125,6))

m <- c(rep(15,2),rep(20,2),rep(25,2))

ster_time <- rep(m,3)

dat <- data.frame(observation,temp,ster_time)

pol <- lm(observation ~ temp + ster_time + I(temp^2)
    + I(ster_time^2) + temp*ster_time,data = dat)
```

```
c <- coefficients(pol)


cat("The fitted polynomial is",c[1],c[2],"*x1",c[3],
   "*x2 +",c[4],"*x1^2 +",c[5],"*x2^2 +",c[6],"*x1*
   x2")
```

R code Exa 12.4 Linear regression model using matrices

```

# Chapter 12
# Example 12.3 page no. 449 from the pdf..
# Linear regression model using matrices..

# given - % survival(y) of certain type of animal
   semen after storage..
# combinations of concentration of 3 materials used
   to increase the chance of survival..
# estimate the linear regression  model using the
   data by matrices form.

y <- c
   (25.5,31.2,25.9,38.4,18.4,26.7,26.4,25.9,32.0,25.2,39.7,35.7,26.5

x1 <- c
   (1.74,6.32,6.22,10.52,1.19,1.22,4.10,6.32,4.08,4.15,10.15,1.72,1.

x2 <- c
   (5.30,5.42,8.41,4.63,11.60,5.85,6.62,8.72,4.42,7.60,4.83,3.12,5.3

x3 <- c
```

```
(10.80,9.40,7.20,8.50,9.40,9.90,8.00,9.10,8.70,9.20,9.40,7.60,8.20

# now forming the matrix,


xxdash <- matrix(c(length(x1),sum(x1),sum(x2),sum(x3),sum(x1),
                        sum(x1*x1),sum(x1*x2),sum(x1*x3),
                        sum(x2),sum(x2*x1),sum(x2*x2),sum(x2*x3),
                        sum(x3),sum(x3*x1),sum(x3*x2),sum(x3*x3)),byrow = T,nrow = 4)

# now forming the matrix xdash_y


xdash_y <- matrix(c(sum(y),sum(x1*y),sum(x2*y),sum(x3*y)),byrow = T,nrow = 4)

# now solving the matrix equation xxdash*b = xdash_y using solve function in R

sol <- solve(xxdash,xdash_y)

# final answer..
cat("The regression line is",sol[1,1],"+",sol[2,1],"*x1",
    sol[3,1],"*x2",sol[4,1],"*x3")
```

R code Exa 12.5 To test the hypothesis on the slope of the regression model

```

# Chapter 12
```

```
# example 12.5 page no. 456 from the pdf..
# to test the hypothesis on the slope of the regression model..

# Null - beta2 = -2.5
# alternate > -2.5

# solving example 12.4 again..
# linear regression model using matrices..
# given - % survival(y) of certain type of animal semen after storage..
# combinations of concentration of 3 materials used to increase the chance of survival..
# estimate the linear regression  model using the data by matrices form.

y <- c(25.5,31.2,25.9,38.4,18.4,26.7,26.4,25.9,32.0,25.2,39.7,35.7,26.5

x1 <- c(1.74,6.32,6.22,10.52,1.19,1.22,4.10,6.32,4.08,4.15,10.15,1.72,1.

x2 <- c(5.30,5.42,8.41,4.63,11.60,5.85,6.62,8.72,4.42,7.60,4.83,3.12,5.3

x3 <- c(10.80,9.40,7.20,8.50,9.40,9.90,8.00,9.10,8.70,9.20,9.40,7.60,8.2

# now forming the matrix,


xxdash <- matrix(c(length(x1),sum(x1),sum(x2),sum(x3),sum(x1),
                   sum(x1*x1),sum(x1*x2),sum(x1*x3),
```

```
                        sum(x2),sum(x2*x1),sum(x2*x2),sum(x2*x3),
                        sum(x3),sum(x3*x1),sum(x3*x2),sum(x3*x3)),byrow = T,nrow = 4)

# now forming the matrix xdash_y


xdash_y <- matrix(c(sum(y),sum(x1*y),sum(x2*y),sum(x3*y)),byrow = T,nrow = 4)

# now solving the matrix equation xxdash*b = xdash_y using solve function in R

sol <- solve(xxdash,xdash_y)

unit <- matrix(c(c(1,0,0,0),c(0,1,0,0),c(0,0,1,0),c(0,0,0,1)),ncol = 4)


# finding inverse of the matrix xxdash..
inv_xxdash <- solve(xxdash,unit)

# final answer..
cat("The regression line is",sol[1,1],"+",sol[2,1],"*x1",
    sol[3,1],"*x2",sol[4,1],"*x3")

# the coefficient of x2 (i.e beta2) is -1.861649 as we can see from the solution..
# now testing our hypothesis..

t.score <- (sol[3,1]+2.5)/(2.073*sqrt(inv_xxdash[3,3]))

cat("The p-value is",pt(t.score,9,lower.tail = F),"we reject the null hypothesis and conclude beta2 > -2.5")
```

```
# the answer in the T.B is double the value obtained by computation,
#I think they took both sided alternative, but we have to take only one sided alternative as our alternative is beta2 > −2.5,
#hence it is one sided and so T.B answer is wrong..
```

R code Exa 12.6 To Construct Confidence Interval for Mean Response

```

# Chapter 12
# Example 12.6 page no. 457 from the pdf..
# To Construct Confidence Interval for Mean Response ..

y <-c(25.5,31.2,25.9,38.4,18.4,26.7,26.4,25.9,32.0,25.2,39.7,35.7,26.5

x1 <- c(1.74,6.32,6.22,10.52,1.19,1.22,4.10,6.32,4.08,4.15,10.15,1.72,1.

x2 <- c(5.30,5.42,8.41,4.63,11.60,5.85,6.62,8.72,4.42,7.60,4.83,3.12,5.3

x3 <- c(10.80,9.40,7.20,8.50,9.40,9.90,8.00,9.10,8.70,9.20,9.40,9.60,8.2

dat <- data.frame(y,x1,x2,x3)

c <- lm(y~x1+x2+x3,data= dat)
```

```
newdat <- data.frame(x1=3,x2=8,x3=9)

d <- predict(c,newdat,interval = "confidence",level
   = 0.95,type = "response")

cat("The 95% confidence interval for the mean
   response when x1=3%,x2=8% and x3=9% is",d[2],d
   [3])

# the answer may vary slightly from the T.B
```

R code Exa 12.7 Prediction interval Evaluation

```

# Chapter 12
# Example 12.7 Page no. 458 from the pdf..
# Prediction interval Evaluation..

y <-c
   (25.5,31.2,25.9,38.4,18.4,26.7,26.4,25.9,32.0,25.2,39.7,35.7,26.5

x1 <- c
   (1.74,6.32,6.22,10.52,1.19,1.22,4.10,6.32,4.08,4.15,10.15,1.72,1.

x2 <- c
   (5.30,5.42,8.41,4.63,11.60,5.85,6.62,8.72,4.42,7.60,4.83,3.12,5.3

x3 <- c
   (10.80,9.40,7.20,8.50,9.40,9.90,8.00,9.10,8.70,9.20,9.40,9.60,8.2
```

```
dat <- data.frame(y,x1,x2,x3)

c <- lm(y~x1+x2+x3,data= dat)  #modelling the
    variables of datasaet..


newdat <- data.frame(x1=3,x2=8,x3=9)

d <- predict(c,newdat,interval = "prediction")

cat("The 95% prediction interval for the individual
    response when x1=3%,x2=8% and x3=9% is",d[2],d
    [3])
```

R code Exa 12.8 To show Analysis of Variance Table for Grain Radius Data

```

# Chapter 12
# Example 12.8 page no. 470 from the pdf..
# To show Analysis of Variance Table for Grain
    Radius Data

pow_temp <- c(150,190,150,150,190,190,150,190)
ext_rate <- c(12,12,24,12,24,12,24,24)
die_temp <- c(220,220,220,250,220,250,250,250)



gra_radius <- c(82,93,114,124,111,129,157,164)

dat <- data.frame(pow_temp,ext_rate,die_temp,gra_
    radius)

```

```
# converting into factors..
dat$pow_temp <- factor(dat$pow_temp)
dat$ext_rate <- factor(dat$ext_rate)
dat$die_temp <- factor(dat$die_temp)


l <- lm(gra_radius~pow_temp+ext_rate+die_temp,data =
    dat)

# analysis of variance table for grain radius data..
print("The analysis of variance table for this data
   is:")
summary(aov(l,data= dat))
```

R code Exa 12.9 To find the model for data which have 3 levels in a variable

```

# Chapter 12
# Example 12.9 page no. 474 from the pdf..
# to find the model for data which have 3 levels in
   a variable..

y <- c
   (292,329,352,378,392,410,198,227,277,297,364,375,167,225,247,268,

x <- c
   (6.5,6.9,7.8,8.4,8.8,9.2,6.7,6.9,7.5,7.9,8.7,9.2,6.5,7.0,7.2,7.6,

p <- c(1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3)

dat <- data.frame(y,x,factor(p))

```

```
dat$factor.p. <- relevel(dat$factor.p.,ref=3)

c <- lm(y ~ x +factor.p.,data= dat)

cat("The model suggests")

coefficients(c) # Coefficients of the model..

summary(c) # Summary of the regression model..
```

R code Exa 12.10 To predict all possible regression lines

```

# Chapter 12
# Example 12.10 page no. 477 from the pdf..
# to predict all possible regression lines

# refer ex10 page no. 478 on pdf.

# using leaps library if not installed can be done by
# install.packages("leaps")

# using sequential replacement method..

library(leaps)

y <- c(57.5,52.8,61.3,67.0,53.5,62.7,56.2,68.5,69.2)

x1 <- c(78,69,77,88,67,80,74,94,102)

x2 <- c(48.2,45.5,46.3,49.0,43.0,48.0,48.0,53.0,58.0)

x3 <- c
```

```
(2.75,2.15,4.41,5.52,3.21,4.32,2.31,4.30,3.71)


x4 <- c(29.5,26.3,32.2,36.5,27.2,27.7,28.3,30.3,28.7)


dat <- data.frame(y,x1,x2,x3,x4)

s <- regsubsets(y~.,data = dat,method = "seqrep")

print("All possible regression line equation coefficients are")

coef(s,1:4)

# for final model we need to see adj r squared value ..this is just the coefficients for various variables in the model..

# as nothing specific is asked here so I have shown what would be the model for all combination of variables..
```

R code Exa 12.11 To model the data set using forward selection

```

# Chapter 12
# Example 12.11 page no. 480 from the pdf..
# to model the data set using forward selection..

# the data set from example 10 is used for modelling ..
y <- c(57.5,52.8,61.3,67.0,53.5,62.7,56.2,68.5,69.2)

```

```
x1 <- c(78,69,77,88,67,80,74,94,102)

x2 <- c(48.2,45.5,46.3,49.0,43.0,48.0,48.0,53.0,58.0)

x3 <- c(2.75,2.15,4.41,5.52,3.21,4.32,2.31,4.30,3.71)


x4 <- c(29.5,26.3,32.2,36.5,27.2,27.7,28.3,30.3,28.7)


dat <- data.frame(y,x1,x2,x3,x4)

# model..using forward selection

line <- lm(y~x1,data=dat)
summary(line)$adj.r.sq

line1 <- lm(y~x2,data=dat)
summary(line1)$adj.r.sq

line2 <- lm(y~x3,data=dat)
summary(line2)$adj.r.sq

line3 <- lm(y~x4,data=dat)
summary(line3)$adj.r.sq
# by looking at 4 adj r sq. values x1 has max. increase so x1 in the model
line4 <- lm(y~x1 +x2,data=dat)
summary(line4)$adj.r.sq

line5 <- lm(y~x1 +x3,data=dat)
summary(line5)$adj.r.sq

line6 <- lm(y~x1 +x4,data=dat)
summary(line6)$adj.r.sq
```

```
# by looking at the adj. r .sq. x1 +x3 combination
   has the max. value so x1+x3 in the model.

line7 <- lm(y~x1+x3+x4,data = dat)
summary(line7)$adj.r.sq
# by looking at this estimate, adj. r. sq. is
   decreases so x4 is excluded from the final model
   ..

line8 <- lm(y~x1+x2+x3,data = dat)
summary(line8)$adj.r.sq
cat("The correlation coefficient between x1 and x3
   is",cor(x1,x2))
#although the adj. r .squared is max. for this model
   ,but it is high from x1+x3 by small amount,we
   will not include x2 as the correlation
   coefficient between x1 and x2 is high meaning
   they are dependent..

#final model..
cat("The final model is ")
lm(y~x1+x3,data = dat)
summary(lm(y~x1+x3,data = dat))
cat("The coefficients of the final model is",
   coefficients(lm(y~x1+x3,data = dat)))
# we can also solve the same by backward elimination
   ..
#So both answers are correct..
```

R code Exa 12.12 Cp Statistic

```
# Chapter 12
# Example 12.12 page no.492 from the pdf..
# Cp Statistic..

```

```r
# to find the relationship between sales for a particular year and factor that affect sales..

# Package "leaps" is used if not installed can be done using install.packages("leaps")

library(leaps)

x1 <- c(5.5,2.5,8.0,3.0,3.0,2.9,8.0,9.0,4.0,6.5,5.5,5.0,6.0,5.0,3.5)

x2 <- c(31,55,67,50,38,71,30,56,42,73,60,44,50,39,55)

x3 <- c(10,8,12,7,8,12,12,5,8,5,11,12,6,10,10)

x4 <- c(8,6,9,16,15,17,8,10,4,16,7,12,6,4,4)

y <- c(79.3,200.1,163.2,200.1,146.0,177.7,30.9,291.9,160.0,339.4,159.6,

dat <- data.frame(x1,x2,x3,x4,y)

# I am comparing on basis of Cp values..

print("The Cp values for all subsets is")

leaps(dat[,1:4],y= dat[,5],names = names(dat)[1:4], method ="Cp")

# if you want to calculate adjusted r-squared or r-squared the code will be same as above just we have to change method = "r2" or method = "adjr2"

print("From Cp values it seems that x1x2x3 model appears quite good and had lowest Cp value")
```

```
print("Also You can see Cp for full model = 5.0 from the table in output")

# to obatain PRESS statistic for each model , we can do something like this..

# this is only PRESS for full model..
print("The PRESS for full model x1x2x3x4 is")

model <- lm(y~.,data = dat)

sum((model$residuals/(1-hatvalues(model)))^2)
# like this we can take any model and calculate its PRESS statistic..
# to compare answers from T.B look for Cp stats table and PRESS valuesin the tables and match it , The answer is correct.
```

R code Exa 12.13 Logistic regression model

```

# Chapter 12
# Example 12.13 page no.499 from the pdf..
# Logistic regression model..

x <- c(0.10,0.15,0.20,0.30,0.50,0.70,0.95)

n <- c(47,53,55,52,46,54,52)

y <- c(8,14,24,32,38,50,50)

p <- c(17.0,26.4,43.6,61.5,82.6,92.6,96.2)/100

dat <- data.frame(x,n,y,p)
```

```

# making a logistic regression..

model1 <- glm(p~x,data = dat,family = binomial(link
    = "logit"))
summary(aov(model1))
c <- coefficients(model1)
cat("The beta0 and beta1 for the logistic regression
     function is",c[1],c[2])

# the answer may vary slightly..
```

Chapter 13

One Factor Experiment General

R code Exa 13.1 One way ANOVA

```

# Chapter 13
# Example 13.1 page no. 513 from the pdf..
# One way ANOVA..

# H0: mu1=mu2=mu3=mu4=mu5
# H1 : atleast two of means not equal..

a <- c
  (551,457,450,731,499,632,595,580,508,583,633,517,639,615,511,573,

b <- c(rep(1,6),rep(2,6),rep(3,6),rep(4,6),rep(5,6))

dat <- data.frame(a,b)

c <- aov(a~factor(b),data = dat)

summary(c)
```

```
cat("Since the p-value is 0.00875 we reject null hypothesis and and conclude aggregates don't have same mean absorption")
```

R code Exa 13.2 One way ANOVA

```

# Chapter 13
# Example 13.2 page no. 514 from the pdf..
# One way ANOVA..

# NULL : H0: mu1=mu2=mu3=mu4
# alternate : H1: at least two are not equal


a <- c(49.20,44.54,45.80,95.84,30.10,36.50,82.30,87.85,105.00,95.22,97.

b <- c(rep(1,20),rep(2,9),rep(3,9),rep(4,7))

dat <- data.frame(a,b)

c <- aov(a~factor(b),data = dat)  # anova

summary(c)  # Analysis of Variance table

cat("Since the p-value is 0.022,we reject the null hypothesisand conclude alkaline levels for the four drug groups are not the same")
```

R code Exa 13.3 Bartlett Test

```

# Chapter 13
# Example 13.3 page no. 517 from the pdf..
# Bartlett Test..

# to test the following hypothesis
# NULL : H0: sig1^2=sig2^2=sig3^2=sig4^2
# alternate : H1: at least two are not equal

# using bartlett test..
a <- c(49.20,44.54,45.80,95.84,30.10,36.50,82.30,87.85,105.00,95.22,97.

b <- c(rep("I",20),rep("J",9),rep("K",9),rep("L",7))

dat <- data.frame(a,b)

bartlett.test(a~b,data = dat)  # display of Bartlett test..

cat("The p-value is",bartlett.test(a~b,data = dat)$p.value,"we do not reject the null hypothesis and conclude popn. variances are not significantly different")

# the T.B has done this by comparing the areas in the left tail of bartlett distribution ,using bartlett test function directly gives p-value in R, so it is more handy..
```

R code Exa 13.4 Contrasts Sum Of Squares Corresponding to Orthogonal Contrasts

```
# Chapter 13
# Example 13.4 page no. 522 from the pdf..
# Contrasts Sum Of Squares Corresponding to orthogonal Contrasts..

# from ex 13.1..
# w1 = mu1+mu2-mu3-mu5, w2 = mu1+mu2+mu3-4mu4+ mu5.
# w3 = mu1 -mu2 , w4 = mu3-mu5.

a <- c(551,457,450,731,499,632,595,580,508,583,633,517,639,615,511,573,

b <- c(rep(1,6),rep(2,6),rep(3,6),rep(4,6),rep(5,6))

# making a data frame and re-classifying the data.
dat <- data.frame(a,b)
dat$b <- as.factor(dat$b)

# the ANOVA..
l <- lm(a~b,data = dat)
anova(l)

# contrasts coefficients..
contrastmatrix <- cbind(c(1,1,-1,0,-1),c(1,1,1,-4,1),c(1,-1,0,0,0),
                        c(0,0,1,0,-1))
contrasts(dat$b) <- contrastmatrix


l_contrast <- aov(a~b,data = dat)


print("The Analysis of Variance Table Using ortogonal contrats is:")
summary(l_contrast, split = list(b = list("(1,2)vs(3,5)"=1,"1,2,3,5 vs 4"=2)))
```

R code Exa 13.5 Dunnett Test

```

# Chapter 13
# Example 13.5 page no. 529 from the pdf..
# Dunnett Test..

# To compare each catalyst with control using 2
    sided hypothesis..
# Package used "DescTools", if already installed
    delete the below line from the code..
install.packages("DescTools")

library(DescTools)

co <- c(50.7,51.5,49.2,53.1,52.7)

cat1 <- c(54.1,53.8,53.1,52.5,54.0)

cat2 <- c(52.7,53.9,57.0,54.1,52.5)

cat3 <- c(51.2,50.8,49.7,48.0,47.2)

DunnettTest(list(co,cat1,cat2,cat3))

# The value of Diff column in the table varies
    slightly from the T.B, due to approximatons..
# Like d1 and d3  in T.B are 2.14 and -2.14, and
    from computation coming out to be 2.06 and -2.06
    so that's due to approximations or different
    method to approach the same problem but final
    inferences are same..

print("From the Table looking at p-value we conclude
```

```
that catalyst 2 is significantly different from the mean yield of the reaction using the control ..")
```

R code Exa 13.6 Randomized Complete Block Diagram

```

# Chapter 13
# Example 13.8 page no. 537 from the pdf..
# Randomized Complete Block Diagram..

# to to test hypothesis(0.05 level) that machines perform at the same mean rate of speed
# 6 different operators used in randomized block experiment to compare 4 machines..

m1 <- c(42.5,39.3,39.6,39.9,42.9,43.6)

m2 <- c(39.8,40.1,40.5,42.3,42.5,43.1)

m3 <- c(40.2,40.5,41.3,43.4,44.9,45.1)

m4 <- c(41.3,42.2,43.5,44.2,45.9,42.3)

dat <- rbind(m1,m2,m3,m4)   # combining rows to make matrix..

a <- c(t(as.matrix(dat)))   # concatenate different rows into a vector..

b <- c("o1","o2","o3","o4","o5","o6") # treatment levels

n_tr <- 6  # no. of treatment levels

```

```
n_cont <- 4 # no. of control blocks..

operator <- gl(n_tr,1,n_cont*n_tr,factor(b))  # vector of treatment factors corresponding to each element of vector a..

machines <- gl(n_cont,n_tr,n_tr*n_cont) # vector of blocking factors corresponding to each element in vector a..

print("The Analysis of Variance table is:")

summary(aov(a~operator+machines)) # anova table display..
```

R code Exa 13.7 Random Effects Model

```

# Chapter 13
# Example 13.7 page no. 549 from the pdf..
# Random Effects Model..
# to find the batch variance component from the data ..

m1 <- c(9.7,5.6,8.4,7.9,8.2,7.7,8.1)

m2 <- c(10.4,9.6,7.3,6.8,8.8,9.2,7.6)

m3 <- c(15.9,14.4,8.3,12.8,7.9,11.6,9.8)

m4 <- c(8.6,11.1,10.7,7.6,6.4,5.9,8.1)

m5 <- c(9.7,12.8,8.7,13.4,8.3,11.7,10.7)

dat <- cbind(m1,m2,m3,m4,m5)
```

```
b <- c(t(as.matrix(dat)))

a <- c("b1","b2","b3","b4","b5")

n_batch <- 5  # n. of treatment groups

n_row <- 7  # no. of rows..

batch <- gl(n_batch,1,n_row*n_batch,factor(a))  # vector of treatment factors corresponding to each element of vector b..

summary(c <- aov(b~batch))

cat("From the table we can calculate the estimated variance from the mean sq. values which comes out to be",(18.149-4.068)/n_row)
```

Chapter 14

Factorial Experiments Two Or More Factors

R code Exa 14.1 Two Factor Analysis of Variance

```

# Chapter 14
# Example 14.1 page no. 571 from the pdf..
# Two Factor Analysis of Variance..

# H0`: alpha1=alpha2=alpha3=0
#H0``:beta1=beta2=beta3=beta4=0
#H0```:(alpha*beta)11=(alpha*beta)12...=(alpha*beta)
    34=0
# H1`: atleast one of alpha non zero
# H1``: atleast one of the beta non zero
#H1```:atleast one of the (alpha*beta) non zero..

v1 <- c(34.0,32.7,32.0,33.2,28.4,29.3)

v2 <- c(30.1,32.8,30.2,29.8,27.3,28.9)

v3 <- c(29.8,26.7,28.7,28.1,29.7,27.3)

```

```
v4 <- c(29.0,28.9,27.6,27.8,28.8,29.1)

observations <- c(v1,v2,v3,v4)

prop_type <- c(rep(1,6),rep(2,6),rep(3,6),rep(4,6))

a <- c(rep(1,2),rep(2,2),rep(3,2))

missile_sys <- c(rep(a,4))

dat <- data.frame(observations,prop_type,missile_sys)

d <- aov(observations~factor(missile_sys)*factor(prop_type),data = dat)

summary(d) # analysis of variance table..

cat("We reject H0` and conclude that different missile have different propellent rates as p-value is 0.0169")

cat("We reject H0``and conclude that mean propellent rates are not same for 4 propellent types as p-value is 0.0010")

cat("p-value is approx. 0.0513, so interaction is barely significant.")
```

R code Exa 14.2 Single Degree Of Freedom Sum Of Squares

```

# Chapter 14
# Example 14.2 page no. 571 from the pdf..
# Single Degree Of Freedom Sum Of Squares..
```

```
# to choose 2 orthogonal contrasts to partition the sum of sqaures for missile systems into
# single-degree-of-freedom components to be used in comparing systems 1 and 2 versus 3 and system 1 with system 2.
v1 <- c(34.0,32.7,32.0,33.2,28.4,29.3)
v2 <- c(30.1,32.8,30.2,29.8,27.3,28.9)
v3 <- c(29.8,26.7,28.7,28.1,29.7,27.3)
v4 <- c(29.0,28.9,27.6,27.8,28.8,29.1)
observations <- c(v1,v2,v3,v4)
prop_type <- c(rep(1,6),rep(2,6),rep(3,6),rep(4,6))
a <- c(rep(1,2),rep(2,2),rep(3,2))
missile_sys <- c(rep(a,4))
dat <- data.frame(observations,prop_type,missile_sys)
dat$missile_sys <- as.factor(dat$missile_sys)
contrastmatrix <- cbind(c(1,1,-2),c(1,-1,0))
contrasts(dat$missile_sys) <- contrastmatrix
missile_contrast <- aov(observations~missile_sys, data = dat)
print("The Sum of Squares for missiles systems to be
```

```
    used in comparing systems 1 and 2 versus 3 and
   system 1 versus 2 can be seen from the analysis
   of variance table below.")
summary(missile_contrast, split = list(missile_sys =
                                    list("(1,2) vs 3"
                                      =1,"1 vs 2"=2)
                                      ))
```

R code Exa 14.3 Evaluate Analysis of Variance table and derive conclusion also test on main effects

```

# Chapter 14
# Example 14.3 page no. 574 from thr pdf..
# Evaluate Analysis of Variance table and derive
   conclusion,also test on main effects..

m1 <- c(288,488,670,360,465,720)

m2 <- c(385,482,692,411,521,724)

m3 <- c(488,595,761,462,612,801)

obs <- c(m1,m2,m3)

power_supp <- c(rep(1:3,6))


flow_rate <- c(rep(1,6),rep(2,6),rep(3,6))

dat <- data.frame(obs,power_supp,flow_rate)

d <- aov(obs~factor(flow_rate)*factor(power_supp),
   data = dat)

```

```
summary(d)

cat("p-value for test of interaction is 0.4484 so we
    conclude there is no significant interaction")

# duncan test can be done by duncan.test , in
   agricolae library , I am having some problems with
    the package it is not installing the package and
    showing some error , so I cannot use duncan.test
   function ..
```

R code Exa 14.4 Anova for 3 factor experiment

```

# Chapter 14
# Example 14.4 page no. 581 from the pdf..
# Anova for 3 factor experiment..

m1 <- c
   (10.7,10.3,11.2,10.9,10.5,12.2,10.8,10.2,11.6,12.1,11.1,11.7,11.3

m2 <- c
   (11.4,10.2,10.7,9.8,12.6,10.8,11.8,10.9,10.5,11.3,7.5,10.2,11.5,1

m3 <- c
   (13.6,12.0,11.1,10.7,10.2,11.9,14.1,11.6,11.0,11.7,11.5,11.6,14.5



obs <- c(m1,m2,m3)

operator <- c(rep(1,18),rep(2,18),rep(3,18))

```

```
catalyst <- c(rep(1:3,18))

a <- c(rep("15 min",3),rep("20 min",3))

washing_time <- c(rep(a,9))

dat <- data.frame(obs,operator,catalyst,washing_time)

d <- aov(obs~(factor(operator)*factor(catalyst)*factor(washing_time)),data = dat)

print("The Analysis of Variance Table is shown as follows:")

summary(d)
```

R code Exa 14.5 Pooling in Multi factor Model

```

# Chapter 14
# Example 14.5 page no. 584 from the pdf..
# Pooling in Multifactor Model..

# to find the anova table from the following data set by removing some interactions and then look at the effect..

m1 <- c(43,49,44,47)

m2 <- c(64,68,97,102)

m3 <- c(49,57,51,55)

m4 <- c(70,76,103,106)
```

```
obs <- c(m1,m2,m3,m4)

a <- c(rep("L",4),rep("H",4))
temp <- c(rep(a,2))

batch <- c(rep(1,8),rep(2,8))

b <- c(rep("Low",2),rep("high",2))
string_rate <- c(rep(b,4))

pressure <- c(rep(c("Low","High"),8))

dat <- data.frame(obs,batch,temp,string_rate,
    pressure)  # making data frame

cat("The anova table after removing particular
    interactions is")

summary(aov(obs~factor(batch)+factor(temp)*factor(
    pressure)*factor(string_rate),data = dat))  #
    taking the elements of interest for interaction..

# we can also use "update" function by writitng the
    full model of 4 variables and subtracting the
    undesired interactions,but it will be very
    tedious..

# the answer may vary slightly..
```

R code Exa 14.6 Factorial Experiments for Random Effects

```

# Chapter 14
# Example 14.6 page no. 589 from the pdf..
```

```
# Factorial Experiments for Random Effects..

# to determine which are the important sources of
    variation in an industrial process

op1 <- c
    (66.9,68.3,69.0,69.3,68.1,67.4,69.8,70.9,67.2,67.7,67.5,71.4)


op2 <- c
    (66.3,68.1,69.7,69.4,65.4,66.9,68.8,69.6,65.8,67.6,69.2,70.0)


op3 <- c
    (65.6,66.0,67.1,67.9,66.3,66.9,66.2,68.4,65.2,67.3,67.4,68.7)



obs <- c(op1,op2,op3)

operator <- c(rep(1,12),rep(2,12),rep(3,12))

batch <- c(rep(1:4,9))


dat <- data.frame(obs,operator, batch)

print("The analysis of variance table for this
    example is:")

summary(aov(obs~factor(operator)+ factor(batch) +
    factor(batch)*factor(operator)),data = dat)
# interaction in this case is the product of
    operator and batch components.
```

Chapter 15

2 to the power k Factorial Experiments and Fractions

R code Exa 15.1 2 squared Factorial Experiment

```
# Chapter 15
# Example 15.1 page no. 601 from the pdf..
# 2^2 factorial Experiment..

# 2^2 factorial with no interaction
obs <- c(50,70,80,100)

a_no <- c(rep(1,2),rep(-1,2))

b_no <- c(rep(c(-1,1),2))

cat("The main effects are A =",(sum(obs[1:2])-sum(
   obs[3:4]))/2)

cat("B = ",(obs[4]+obs[2])/2-(obs[1]+obs[3])/2)

cat("The interaciton effect is",(obs[1]+obs[4])/2-(
   obs[2]+obs[3])/2)
```

```

obs1 <- c(50,70,80,40)

cat("The interaction effect in this case is",(obs1[2]+obs1[3])/2-(obs1[1]+obs1[4])/2)
```

R code Exa 15.2 Factorial Experiment in a Regression Setting

```

# Chapter 15
# example 15.2 page no. 613 from the pdf
# Factorial Experiment in a Regression Setting..

hold_tim <- c(0.5,0.8,0.5,0.8)

flex_time <- c(0.10,0.10,0.20,0.20)

yield <- c(28,39,32,46)

dat <- data.frame(hold_tim,flex_time,yield)

dat$hold_tim <- factor(dat$hold_tim)
dat$flex_time <- factor(dat$flex_time)


f <- function(x){

  dat$yield[x]

}

levels(dat$hold_tim) <- c(-1,1)
levels(dat$flex_time) <- c(-1,1)

cat("The regression equation is",(f(1)+f(2)+f(3)+f
```

```
(4))/4,"+",(f(2)+f(4)-f(1)-f(3))/4,"*x1 +",(f(3)+
    f(4)-f(1)-f(2))/4,"*x2")
```

R code Exa 15.4 Standard Errors of the least Squares Regression Coefficients

```

# Chapter 15
# example 15.4 page no. 619
# Standard Errors of the least Sqaures Regression
    Coefficients

# given data and anova table , to find the standard
    errors of the least squares regression
    coefficients..

# Note 15.3 — Theoretical with nothing to compute..

# standard errors of all coefficients for the 2^k
    factorial are equal so..
# from anova table given s^2 = 2.46

cat("The standard error of the least squares
    regression coefficients are",sqrt(2.46/(2^4*2)))
```

R code Exa 15.6 Use a half replicate to study the effects of five factors each at 2 levels on some response

```

# Chapter 15
# example 15.6 page no.633 from the pdf..
# use a half replicate to study the effects of five
    factors each at 2 levels on some response..
```

```
# to perform analysis of variance on the data below
    testing all main effects for the significance at
    the 0.05 level..

response <- c
    (11.3,15.6,12.7,10.4,9.2,11.0,8.9,9.6,14.1,14.2,11.7,9.4,16.2,13.9
treatment <- c("a","b","c","d","e","abc","abd","acd"
    ,"bcd","abe","ace","ade","bce","bde","cde","abcde
    ")
cat("The sum of squares and effects for the main
    effects are")

a <- sum(response)-2*sum(response[which(treatment %
    in% c("b","c","d","e","bcd","bce","bde","cde"))])

b <- sum(response)-2*sum(response[which(treatment %
    in% c("a","c","d","e","acd","ace","ade","cde"))])

c <- sum(response)-2*sum(response[which(treatment %
    in% c("b","a","d","e","abd","abe","ade","bde"))])

d <- sum(response)-2*sum(response[which(treatment %
    in% c("b","c","a","e","abc","abe","ace","bce"))])

e <- sum(response)-2*sum(response[which(treatment %
    in% c("b","c","d","a","abc","abd","acd","bcd"))])

cat("The sum of squares and effects for the main
    effects are")

cat("SSA=",a^2/16,"and A=",a/8)

cat("SSB=",b^2/16,"and B=",b/8)

cat("SSC=",c^2/16,"and C=",c/8)

cat("SSD=",d^2/16,"and D=",d/8)
```

```
cat("SSE=",e^2/16,"and E=",e/8)
```

R code Exa 15.7 To construct a 2 level screening design with 6 variables containing 12 design points

```

# Chapter 15
# example 15.7 page no. 639
# to construct a 2 level screening design with 6
    variables containing 12 design points..
# used package FrF2, if not installed can be
    installed using install.packages("FrF2")

library(FrF2)

print("The 2 level scrrening design with 6 variables
     containing 12 design points is")

pb(12,6,randomize = F,default.levels = c("-","+"))

# the answer may differ but the concept is same, the
     T.B as well as the software answers are correct,
     R has selected randomly the design,
#if we set randomize= T, then every time the
    instruction is executed then every time a
    diferent design will appear, so that's why the
    design is correct.
```

R code Exa 15.8 Response Surface Analysis

```

# Chapter 15
```

```
# example 15.8 page no. 640 from the pdf..
# Response Surface Analysis..

# a central composite design is given on page no.641
    of the pdf,to determine the impact that x1 and
   x2 have on % conversion process..
#Package used - rsm if not installed can be done
   using install.packaged("rsm")
library(rsm)

x1 <- c(-1,1,-1,1,-1.414,1.414,0,0,0,0,0,0)

x2 <- c(-1,-1,1,1,0,0,-1.414,1.414,0,0,0,0)

y <- c(43,78,69,73,48,78,65,74,76,79,83,81)

dat <- data.frame(x1,x2,y)

res_model <- rsm(y~FO(x1,x2,x1^2,x2^2,x1*x2),data =
   dat)

print("The coefficients of the resulting second
   order response model is given in coded variables
   as")

coefficients(res_model)

print("The detailed response surface model is given
   below")

summary(res_model)


# the natural model..
e1 <- c
   (200,250,200,250,189.65,260.35,225,225,225,225,225,225)

```

```
e2 <- c(15,15,25,25,20,20,12.93,27.07,20,20,20,20)

dat1 <- data.frame(e1,e2,y)

print("The natural surface model is shown below:")

nat_model <- lm(y~e1+e2+I(e1^2)+I(e2^2)+I(e1*e2),
   data = dat1)

summary(nat_model)

print("The coefficients of the natural surface model
    is given as follows:")

coefficients(nat_model)
```

Chapter 16

Non Parametric Statistics

R code Exa 16.1 Use sign test to test the hypothesis

```

# Chapter 16
# Example 16.1 page no. 658 from the pdf..
# use sign test to test the hypothesis..

# Package used "BSDA", referance- Internet..

install.packages("BSDA")  # package for sign testing

library("BSDA")

#NULL H0:mu=1.8
#alternate: mu!= 8

dat <- c(1.5,2.2,0.9,1.3,2.0,1.6,1.8,1.5,2.0,1.2,1.7)


SIGN.test(dat,md=1.8)

cat("Since The p-value is",SIGN.test(dat,md=1.8)$p.
```

```
    value,"we do not reject the null hypothesis.")
```

R code Exa 16.2 Use Sign test to test the hypothesis

```

# Chapter 16
# Example 16.2 page no. 659 from the pdf..
# Hypothesis Testing Using Sign Test..
# to test the hypothesis with 0.05 level of
    significance

# Null :mu1-mu2 =0
#alternate:mu1-mu2 !=0


install.packages("BSDA") # package for sign testing
#if already installed comment it..


library("BSDA")

rad_tires <- c
    (4.2,4.7,6.6,7.0,6.7,4.5,5.7,6.0,7.4,4.9,6.1,5.2,5.7,6.9,6.8,4.9)

bel_tires <- c
    (4.1,4.9,6.2,6.9,6.8,4.4,5.7,5.8,6.9,4.9,6.0,4.9,5.3,6.5,7.1,4.8)

diff <- data.frame(rad_tires,bel_tires)

SIGN.test(rad_tires,bel_tires,alternative = "greater
    ",paired =T)

cat("Since the p-value is",SIGN.test(rad_tires,bel_
```

```
    tires,alternative = "greater",paired =T)$p.value,
    "we reject null, so on average radial tires do
    improve fuel economy")


#the answer is slightly different than in the T.B as
     in T.B normal approximation is used..
```

R code Exa 16.3 Use signed rank test to test the hypothesis

```

# Chapter 16
# Example 16 page no. 661 from the pdf..
#use signed rank test to test the hypothesis..

#NULL H0:mu=1.8
#alternate: mu!= 8

dat <- c
    (1.5,2.2,0.9,1.3,2.0,1.6,1.8,1.5,2.0,1.2,1.7)

wilcox.test(dat,alternative = "two.sided",mu=1.8)


cat("The p-vaue is more than 0.05,we don't reject
    null and conclude median operating time not
    significantly different from 1.8hrs")
```

R code Exa 16.4 Use signed rank test to test the hypothesis

```

# Chapter 16
# Example 16.4 page no. 662 from the pdf..
```

```
#use signed rank test to test the hypothesis..

# to do hypothesis testing on effect of sample questions on one's gre score..

# NUll H0: mu1-mu2=50
#alternate:mu1-mu2 <50

with_sample <- c(531,621,663,579,451,660,591,719,543,575)

without_sample <- c(509,540,688,502,424,683,568,748,530,524)

wilcox.test(with_sample,without_sample,mu=50,paired = T,alternative = "less")

cat("since the p-value is",wilcox.test(with_sample,without_sample,mu=50,paired = T,alternative = "less")$p.value,"we reject null and conclude that sample problems don't increase graduate record score by as much as 50 points on average..")
```

R code Exa 16.5 To perform Wilcoxin rank sum test

```

# Chapter 16
# Example 16.5 page no. 666 from the pdf..
# to perform wilcoxin rank sum test...

# Null H0 :mu1=mu2
#alternate: mu1!=mu2

brand_a <- c(2.1,4.0,6.3,5.4,4.8,3.7,6.1,3.3)

```

```
brand_b <- c
   (4.1,0.6,3.1,2.5,4.0,6.2,1.6,2.2,1.9,5.4)

wilcox.test(brand_a,brand_b,alternative = "two.sided
   ")

cat("Since the p-value is",wilcox.test(brand_a,brand
   _b,alternative = "two.sided")$p.value,"we do not
   reject null and conclude that there is no
   significant difference in median nicotine
   contents of the above two brand of cigarettes")
```

R code Exa 16.6 Kruskal Wallis Test

```

# Chapter 16
# Example 16.6 page no. 668 from the pdf..
# Kruskal-Wallis Test..

# to do hypothesis testing of propellent burning
   rates for 3 missile system using kruskal wallis
   test..

#Null: H0: mu1=mu2=mu3
#alternate : the means are not equal.
#alpha = 0.05

m1 <- c(24.0,16.7,22.8,19.8,18.9)

m2 <- c(23.2,19.8,18.1,17.6,20.2,17.8)

m3 <- c(18.4,19.1,17.3,17.3,19.7,18.9,18.8,19.3)

obs <- c(m1,m2,m3)

```

```
miss_sys <- c(rep(1,5),rep(2,6),rep(3,8))

kruskal.test(obs,factor(miss_sys))

cat("Since the p-value is",kruskal.test(obs,factor(miss_sys))$p.value,"we failed to reject null that propellent burning rates are same for the 3 missile system")
```

R code Exa 16.7 Runs Test

```

# Chapter 16
# Example 16.7 page no. 672 from the pdf..
# Runs Test..

# to do hypothesis testing of randomness of the sequence..

# Null :H0= sequence is random
# alternate: H1= sequence is is not random..

# package used - "randtests", referance - internet..
install.packages("randtests") # package for runs.test
# comment it if already installed..

library("randtests")

content <- c(3.6,3.9,4.1,3.6,3.8,3.7,3.4,4.0,3.8,4.1,3.9,4.0,3.8,4.2,4.1)

runs.test(content,alternative = "two.sided")

```

```
cat("The p−value is more than 0.1 we don't reject
   the null and conclude that sequence of measured
   values varies randomly")
```

R code Exa 16.8.9 Rank Correlation Coefficient

```

# Chapter 16
# Example 16.8  page no. 675 and example 16.9 page
   no.677 from the pdf..
# Rank Correlation Coefficient..

# to find the rank correlation coefficient between
   tar and nicotine content in cigarettes..
# I am solving ex 8 and 9 in this code only as in 8
   correlation coefficient is asked and in 9th
   hypothesis testing of the sam eproblem is done

# to do hypothesis testing of correlation between
   tar and nicotine content in cigarettes
# null H0:rho=0
# alternate H1: rho>0
# alpha =0.01

tar <- c(14,17,28,17,16,13,24,25,18,31)

nicotine_cont <- c
   (0.9,1.1,1.6,1.3,1.0,0.8,1.5,1.4,1.2,2.0)

cor.test(tar,nicotine_cont,method = "spearman",
   alternative = "greater")

cat("From the test we can see the value of rank
   correlation coefficint(rho) is 0.9665698
```

```
indicating high positive correlation between the
    two variables..")

cat("Since the p-value is very small,",cor.test(tar,
    nicotine_cont,method = "spearman",alternative = "
    greater")$p.value,"we reject null hypothesis and
    conclude that there is significant correlation
    between the two variables..")

# ignore the warning in each case..
```

Chapter 17

Statistical Quality Control

R code Exa 17.1 Expected Value of samples required to detect the shift

```

# Chapter 17
# Example 17.1 page no. 694 from the pdf..
# Expected Number of samples required..

# given n = 4 ,r = 1
# By graph we can see beta = 0.84

beta <- 0.84

cat("The mean of samples requires to detect the
   shift is",1/(1-beta))
```

R code Exa 17.2 Xbar and S chart

```

# Chapter 17
# Example 17.2 page no. 697 from the pdf..
```

```
# Plot Xbar and S chart..

# also the questions asks UCL and LCL caculation but
    their computation is easy relatively , invovles
    only multiplication and division..
# So I am plotting the S and Xbar Charts..

# 25 samples of size 5 each used to establish the
    quality control parameters.To plot the xbar and S
    control charts.
# library used - qicharts2.

install.packages("qicharts2")

library(qicharts2)

m1 <- c
    (62.255,62.187,62.421,62.301,62.400,62.372,62.297,62.325,62.327,62

m2 <- c
    (62.301,62.225,62.377,62.315,62.375,62.275,62.303,62.362,62.297,62

m3 <- c
    (62.289,62.337,62.257,62.293,62.295,62.315,62.337,62.351,62.318,62

m4 <- c
    (62.189,62.297,62.295,62.317,62.272,62.372,62.392,62.371,62.342,62

m5 <- c
    (62.311,62.307,62.222,62.409,62.372,62.302,62.344,62.397,62.318,62

obseravation <- c(m1,m2,m3,m4,m5)
```

```r
samples <- c(rep(1:25,5))


dat <- data.frame(obseravation,samples)


print("The xbar and S chart for the above data is:")


# Run the below two code individually..
#xbar chart
qic(obseravation,
    x= samples,
    data = dat,
    chart = 'xbar',
    xlab = 'Sample Number')

# S chart
qic(obseravation,
    x = samples,
    chart = 's',
    xlab = 'Sample Number',
    data = dat)
```

R code Exa 17.3 UCL and LCL for preliminary control chart value

```r
# Chapter 17
# Example 17.3 page no. 700 from the pdf..
# UCL and LCL for preliminary control chart values.

# given data on no. of defective components in
   sample sizes of 50.

def_comp <- c
   (8,6,5,7,2,5,3,8,4,4,3,1,5,4,4,2,3,5,6,3)
```

```r

samples <- c(1:20)

dat <- data.frame(samples,def_comp)

m <- mean(dat$def_comp/50)

u <- m + 3*sqrt(m*(1-m)/50)

l <- m- 3*sqrt(m*(1-m)/50)

cat("The LCL and UCL are",l,"and", u,"respectively")
```

R code Exa 17.4 Find Sample Size per subgroup

```r

# Chapter 17
# Example 17.4 page no. 701 from the pdf..
# Find Sample Size per subgroup..

# to find the sample size per subgroup producing a
    prob. of 0.5 that a process shift to p=p1= 0.05
    will be detected.
#given in control prob. of a defective= 0.01

p = 0.01

p1 = 0.05

cat("The appropriate sample size is",(9/(p1-p)^2)*p*
    (1-p))
```

R code Exa 17.5 Control Charts for Defects C chart

```

# Chapter 17
# Examples 17.5 page no.702 from the pdf..
# Control Charts for Defects C chart..

# given - no. of defects in 20 successive samples of sheet metal rolls each 100 feet long.
# to develop a control chart.

# package used - qicharts2
# if not installed can be installed by install.packages("qicharts2")

library(qicharts2)

lambd <- 5.95

num_def <- c(8,7,5,4,4,7,6,4,5,6,3,7,5,9,7,7,8,6,7,4)

samples <- c(1:20)

dat <- data.frame(samples, num_def)

# C control chart for the above preliminary data.
qic(num_def,
    x = samples,
    chart = 'c',
    xlab = 'Sample',
    ylab = 'Number of defects',
    data = dat)

cat("The UCL and LCL for the above preliminary data are",lambd+3*sqrt(lambd),"and",lambd-3*sqrt(lambd),"respectively.")
```

Chapter 18

Bayesian Statistics

R code Exa 18.1 To find the Posterior probability distribution

```

# Chapter 18
# Example 18.1 page no. 711 from the pdf..
# To find the posterior probability distribution..

b <- function(p,x){

  choose(2,x)*p^x*(1-p)^(2-x)

}

p1 <- c(0.1,0.2)

pi <- c(0.6,0.4)

m <- c(b(0.1,0)*pi[1]+b(0.2,0)*pi[2],b(0.1,1)*pi[1]+
    b(0.2,1)*pi[2],b(0.1,2)*pi[1]+b(0.2,2)*pi[2])

post_p1 <- c(b(0.1,0)*pi[1]/m[1],b(0.1,1)*pi[1]/m
    [2],b(0.1,2)*pi[1]/m[3])

```

```
post_p2  <- c(1,1,1)-post_p1

cat("The posterior for p=0.1 given x is",post_p1,"x
   ranges from 0 to 2")

cat("The posterior for p=0.2 given x is",post_p2,"x
   ranges from 0 to 2")
```

R code Exa 18.4 Posterior Mean and Posterior Mode

```

# Chapter 18
# Example 18.4 page no. 713 from the pdf..
# Posterior Mean and Posterior Mode..

# Note - Example 18.2 and 18.3 are theoretical with
   nothing to compute..

# given a distribution as on example 18.2 on page no
   . 712 of pdf
# to find the posterior mean and mode when x=1

# the posterior distribution function for 0<p<1
# package used "polynom"

install.packages("polynom")
library(polynom)

a <- function(x,p){

  3*choose(2,x)*(p^x)*((1-p)^(2-x))

}

cat("The posterior mean when x=1 is",integrate(
```

```
function(b){b*a(1,b)},0,1)$value)

a <- polynomial(coef = c(0,6,-6))

b <- deriv(a)

cat(”The posterior mode at x=1 occurs at p equal to”,Re(polyroot(b)))
```

R code Exa 18.6 To find 95 percent Bayesian interval

```

# Chapter 18
# example 18.6 page no. 715 from the pdf..
# to find 95% bayesian interval

# Note − Example 18.5 is theoretical with nothing to compute..

# given prior distribution uniform , for 0<p<1
# refer example 18.2 on page no.712

# from example 18.2 we can see that at x=0 the posterior distribution is 3(1−p)ˆ2
# package ”polynom” is used , if not installed can be done using install.packages(”polynom”)

library(polynom)

a <- polynomial(coef = c(3,-6,3))

b <- polynomial(c(1,-3,3,-1)) - 0.025

c <- integral(a)-0.025

```

```
cat("The 95% bayesian interval is",polyroot(c)[1],"
    and",polyroot(b)[1])
```

R code Exa 18.7 To find the 95 percent Bayesian interval for mu given the various parameter

```

# Chapter 18
# example 18.7 page no. 716, from the pdf..
# to find the 95% bayesian interval for mu..given
    the various parameters.

mu_o <- 800

sig_o <- 10

n <- 25

sig <- 100

x <- 780

mu <- (n*x*sig_o^2 + mu_o*sig^2)/(n*sig_o^2+sig^2)

sd <- sig_o*sig/sqrt(n*sig_o^2+sig^2)

cat("The 95% Bayesian interval for mu is given by",
    mu+qnorm(0.025)*sd,"to",mu-qnorm(0.025)*sd)
```

R code Exa 18.8 To find Bayes estimates of p

```

# Chapter 18
```

```r
# example 18.8 page no.717 from the pdf..
# to find bayes estimates of p, for all values of x
    in example 18.1

# refer example 18.1..

b <- function(p,x){

  choose(2,x)*p^x*(1-p)^(2-x)

}

p1 <- c(0.1,0.2)

pi <- c(0.6,0.4)

m <- c(b(0.1,0)*pi[1]+b(0.2,0)*pi[2],b(0.1,1)*pi[1]+
    b(0.2,1)*pi[2],b(0.1,2)*pi[1]+b(0.2,2)*pi[2])

post_p1 <- c(b(0.1,0)*pi[1]/m[1],b(0.1,1)*pi[1]/m
    [2],b(0.1,2)*pi[1]/m[3])

post_p2  <- c(1,1,1)-post_p1

cat("The bayes estimate of p for x = 0 is",p1[1]*
    post_p1[1]+p1[2]*post_p2[1])

cat("The bayes estimate of p for x = 1 is",p1[1]*
    post_p1[2]+p1[2]*post_p2[2])

cat("The bayes estimate of p for x = 2 is",p1[1]*
    post_p1[3]+p1[2]*post_p2[3])
```

R code Exa 18.9 To find Bayes estimates of p

```

# Chapter 18
# example 18.9 page no. 717 from the pdf..
# to find the bayes estimate of p for all values of x, for example 18.2..

# refer example 18.2
# refer section 6.8 on page 201...
# the posterior can also be written as B(x+1,3-x)

f <- function(x){

  3*choose(2,x)

}

g <- function(p,a){

  (p^(a+1))*((1-p)^(2-a))

}
cat("The bayes estimate for the above distribution for different values of x is")

for(x in 0:2){

  cat("for x =",x,"p star is",f(x)*integrate(function(d){g(d,x)},0,1)$value,"\n")

}
```

R code Exa 18.11 To find Bayes estimates of p under Absolute Error loss

```

# Chapter 18
```

```
# Example 18.11 page no. 718 from the pdf..
# to find the bayes estimate for example 18.9, under absolute error loss, when x = 1 is observed..

# Note - Example 18.10 is theoretical with nothing to compute..

# refer example 18.9..
# package used "polynom" if not installed can be installed by executing install.packages("polynom") on console..

library(polynom)

# we see that for for x = 1 the distribution becomes,
# 6x(1-x)

p <- polynomial(coef = c(0,6,-6))

eq <- integral(p)- 0.5

cat("The bayes estimate under absolute error loss for x = 1 is",Re(polyroot(eq)[1]))
```