R Textbook Companion for Miller and Freund's Probability and Statistics for Engineers by Richard A. Johnson[1]

Created by
Prakash Narayan Singh Rautela
B.Sc.
Computer Science and Engineering
Vivekanand Education Society's College of Arts Science & Commerce
Cross-Checked by
R TBC Team

August 28, 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: Miller and Freund's Probability and Statistics for Engineers

Author: Richard A. Johnson

Publisher: Pearson, USA

Edition: 9

Year: 2017

ISBN: 978-0-321-98624-5

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 2

Organization and Description of Data

R code Exa 2.1 Pareto Diagrams and Dot Diagrams

```
#The industrial engineer records the maximum
#amount of bacteria present along the production
   line , in the units Aerobic Plate
#Count per square inch (APC/in2), for n = 7 day
Data<-c(96.3,155.6,3408.0,333.3,122.2,38.9,58.0)#7
   day data
stripchart(Data,method ="stack",
           at = c(0.1),las=1,xlab = "Bacteria Count
              (APC/sq.in)",pch=20)
```

R code Exa 2.2 Pareto Diagrams and Dot Diagrams

```
#Two Production heats of welding Material
heat1<-c(0.27,0.35,0.37)
heat2<-c(0.23,0.15,0.25,0.24,0.30,0.33,0.26)
data<-list(heat1,heat2)
```

```
stripchart(data,
           xlab=”Copper Content”,
           col=c(”orange”,”red”),
           pch=16,
           method =”stack”,
           cex=1,
           xlim=c(0.15,0.40),
           las=1
)
```

R code Exa 2.3 Frequency Distributions

```
#class mark and class intervals
message(”Class marks are : ”)
message((205+245)/2,” ”,(245+285)/2,” ,”,(285+325)/
    2,” ,”,(325+365)/2,”, ”,(365+405)/2)
print(”Class Interval is”)
message(405-365)
```

R code Exa 2.4 Frequency Distributions

```
#Data of 50 nanopillars
Data<-c(245,333,296,304,276,336,289,234,253,292,
        366,323,309,284,310,338,297,314,305,330,
        266,391,315,305,290,300,292,311,272,312,
        315,355,346,337,303,265,278,276,373,271,
        308,276,364,390,298,290,308,221,274,343)
breaks = seq(205, 405, by=40)
Data.cut = cut(Data, breaks, right=TRUE)
Data.freq = table(Data.cut)
cumfreq = cumsum(Data.freq)
message(”Cumulative Frequency of data”)
cbind(cumfreq)
```

R code Exa 2.6 Graphs of Frequency Distributions

```
#Histogram of interrequest time data
data=c
  (2808,4201,3848,9112,2082,5913,1620,6719,21657,
      3072,2949,11768,4731,14211,1583,9853,78811,6655,
      1803,7012,1892,4227,6583,15147,4740,8528,10563,
      43003,16723,2613,26463,34867,4191,4030,2472,28840,
      24487,14001,15241,1643,5732,5419,28608,2487,995,
      3116,29508,11440,28336,3440
)
#It has an long-right data  tail
hist(data,ylab="Class frequency", xlab="time")
```

R code Exa 2.7 Graphs of Frequency Distributions

```
#density histogram
#58 specimens of new aluminum alloy undergoing
  development
data<-c
  (66.4,67.7,68.0,68.0,68.3,68.4,68.6,68.8,68.9,69.0,69.1,
      69.2,69.3,69.3,69.5,69.5,69.6,69.7,69.8,69.8,69.9,70.0,
      70.0,70.1,70.2,70.3,70.3,70.4,70.5,70.6,70.6,70.8,70.9,
      71.0,71.1,71.2,71.3,71.3,71.5,71.6,71.6,71.7,71.8,71.8,
```

```
        71.9,72.1,72.2,72.3,72.4,72.6,72.7,72.9,73.1,73.3,73.5,
        72.4,74.5,75.3
)
hist(data,freq=FALSE,xlab="Teansile Strength")
lines(density(data))
```

R code Exa 2.8 Descriptive Measures

```
#Mean and variance
data<-c(100,45,60,130,30)#this is a university
   students responded for time in social media
#Using function
mean(data)
median(data)
#Without function
xbar=sum(data)/length(data)
xbar
med=data[ceiling(length(data)/2)]
med
cat("mean and median of data is",xbar,'and',med)
```

R code Exa 2.9 Descriptive Measures

```
?stripchart()
data<-c(11,9,17,19,4,15)
mean(data)
median(data)
stripchart(data,method ="stack",
  at = c(0.1),
  cex=1,pch=20,
  xlim=c(0,20),
  las=1,
```

```
main="Data",xlab = "Email - Request")
```

R code Exa 2.11 Descriptive Measures

```
#Variance
time<-c(0.6,1.2,0.9,1.0,0.6,.8)
xbar=mean(time)
xbar#Mean
diff=time-xbar
diffsq=diff**2
diffsum=sum(diffsq)
n=length(time)
var1=diffsum/(n-1)
cat("Variance is",var1)
#using function
var(time)
```

R code Exa 2.12 Descriptive Measures

```
Time = c(0.6, 1.2,0.9, 1,0.6,0.8)#delay Time
sd=sd(Time)
message("standard deviation is : ",sd," min")
```

R code Exa 2.13 Descriptive Measures

```
#Coefficient of varience
#a)first micrometer observation
mean=3.92
sd=0.015
CV=sd/mean*100
```

```
CV
#a)second micrometer observation
mean2=1.54
sd2=0.008
CV2=sd2/mean2*100
CV2
if(CV2>CV)
  {
  print("measurment mead by first micrometer is more
      precise")
}else
{
  print("measurment mead by second micrometer is
     more precise")
}
```

R code Exa 2.14 Quartiles and Percentiles

```
#quantile and percentile
data<-c(136,143,147,151,158,160,
        161,163,165,167,173,174,
        181,181,185,188,190,205)
options(digits = 1)
data2=quantile(data,c(.25,.5,.75))
Q1=data2[1]
Q2=data2[2]
Q3=data2[3]
percentile=data[ceiling(0.1*18)]
cat("first quartile is",Q1)
cat("second quartile is",Q2)
cat("third quartile is",Q3)
cat("10% percentile  is",percentile)
```

R code Exa 2.15 Quartiles and Percentiles

```
#Range and interquartile range
Data<-c(136,143,147,151,158,160,
        161,163,165,167,173,174,
        181,181,185,188,190,205)
min=min(Data)
max=max(Data)
range=max-min
interrange=IQR(Data)
message("range : ",range," Mpa")
message("interquartile range : ",interrange,"Mpa")
```

R code Exa 2.16 Quartiles and Percentiles

```
time<-c
   (.021,.107,.179,.190,.196,.283,.580,.854,1.18,2.00,7.30)

boxplot(time, horizontal = TRUE,ylim=c(0,8),xlab='
   Time(s)')
quantile=quantile(time,probs = c(0,0.25,.5,.75,1))
#minimum
min=quantile[1]
min
#maimum
max=quantile[5]
max
#Q1
Q1=0.179
Q1
#Q3
Q3=1.18
Q3
#interquartile range is
inter=Q3-Q1
```

```
inter
```

R code Exa 2.18 The Calculation of x and s

```
#20 test runs performed on urban roads with an
    intermediate-size car.
data<-c(19.7,21.5,22.5,22.2,22.6
        ,21.9,20.5,19.3,19.9,21.7
        ,22.8,23.2,21.4,20.8,19.4
        ,22.0,23.0,21.1,20.9,21.3)
xbar=mean(data)
var=var(data)
message("mean : ",xbar," mpg ","standard deviation "
    ,var," mpg")
```

R code Exa 2.19 The Calculation of x and s

```
#Variance and mean of group data
xi=c(225,265,305,345,385)
options(digits = 2)
fi=c(3,11,23,9,4)
n=sum(fi)
mean=weighted.mean(xi,fi)
var=(sum(xi^2*fi)-sum(xi*fi)^2/n)/(n-1)
sd=sqrt(var)
cat("mean: ",mean," var: ",var," sd: ",sd)
```

Chapter 3

PROBABILITY

R code Exa 3.1 Sample Spaces and Events

```
library(sets)#To import the library for set
s=set(pair(0,0),pair(0,1),pair(1,0),pair(2,0),pair
   (1,1),pair(0,2))#Sample space
s
C=set(pair(1,0),pair(0,1.0))
C
D=set(pair(0,0),pair(0,1),pair(0,2))
D
E=set(pair(0,0),pair(1,1))
E
#Union of the Sets
set_union(C,E)
#Intersection of the set
cset_intersection(C,D)
#Complement of the set
set_complement(D,s)
```

R code Exa 3.5 Counting

```
#There are total 12 true−false question  in how many
    way the student can mark the question paper
#Since each question can answer in two ways,there
   are altogether
Possibility=2^12
message("The student can marks the question in :",
   Possibility," possible ways")
```

R code Exa 3.6 Counting

```
#(a)
n1=4#Total operator
n2=3#Total Machines
Pair=n1*n2
message("Total operator−machine pairs are:",Pair)
#(b)
n3=8#test speciment
message("Total test speciment required for the
   entire production is :",n3*Pair)
```

R code Exa 3.11 Probability

```
#Total no of ways to select ace from well−shuffled
   deck of card
n=52#Total no of cards
s=4#Total ace
message("Total probability of drawing an ace from
   deck of card is  : ",s/n)
```

R code Exa 3.13 Probability

```
n=300#Total insulator
t=294#Total insulater withstand the thermal shock
message("The Probability that any untested insulator
    will be able to withstand the thermal shock is :
    ",t/n)
```

R code Exa 3.14 The Axioms of Probability

```
#Permissibility of probability
#if
#1)0<=p(A)<=1 for each event A in S
#2)P(S)=1
#3)A and B are Mutually Exclusive in event
#Ans:-
#Given There is Three Mutually exClusive event A,B
   and c
#a)
permissibility<-function(c){
  Flag=TRUE
  i=1
  while(i<=length(c)){

    if(c[i]<0){
      Flag=FALSE
      return(" negative value not permissible")
    }
    i=i+1
  }
  if(Flag==TRUE && sum(c)==1){
    return("Permissibile")
  }else{
    return(" Not Permissibile")
  }
}
prob1=c(1/3,1/3,1/3)
```

```
prob2=c(0.64,.38,-.02)
prob3=c(0.35,0.52,0.26)
prob4=c(0.57,0.24,0.19)
permissibility(prob1)
permissibility(prob2)
permissibility(prob3)
permissibility(prob4)
```

R code Exa 3.15 Some Elementary Theorems

```
#Probabity that it Rating is
vp=0.07#Very Poor
P=0.12#Poor
fair=0.17
good=0.32
vgood=0.21
excellent=0.11
#(a) probabity that it will rate vp,p,fair,good
#The Possiblity are all mutually excluive
probabilty=vp+p+fair+good
message("The Probability is : ",probabilty)
#(a) probabity that it will rate good ,vgood,
    excellent
#The Possiblity are all mutually excluive
probabilty=good+vgood+excellent
message("The Probability is : ",probabilty)
```

R code Exa 3.16 Some Elementary Theorems

```
#find P(M1),P(P1),P(C3)ETC FROM diagram
PM1= 0.03+0.06+0.07+0.02+0.01+0.01# P(M1)
PP1=0.03 + 0.06 + 0.07 + 0.09 + 0.16 + 0.10 +
    0.05+0.05 + 0.14# P(P1)
```

```
PC3= 0.07 + 0.01 + 0.10 + 0.06 + 0.14 + 0.02# P(C3)
M1INTP1=0.03 + 0.06 + 0.07#M1 intersection with P1
M1IntC3=0.07 + 0.01#M1 intersection with C3
message(PM1," ",PP1," ",PC3," ",M1INTP1," ",M1IntC3)
```

R code Exa 3.17 Some Elementary Theorems

```
M1=0.20#P(M1)
C3=0.4#P(C3)
M1andC3=0.08#M1 intersect with C3
#By addtion Rule
M1ORC3=M1+C3-M1andC3
cat("Probability is",EORC)
```

R code Exa 3.18 Some Elementary Theorems

```
#Probability taht the car require repair on engine,
   drive train or both
P1=0.87
P2=0.36
P1INTP2=0.29
P1UNIONP2=P1+P2-P1INTP2
message("Proabaility that it require bothe kind of
   repair is : ",P1UNIONP2)
```

R code Exa 3.19 Some Elementary Theorems

```
M1=0.20#probability that car will  have low mileage
C3=0.4#Probability that car is expensive to operate
```

```
#(a)the probability that a used car will not have low mileage
M1bar=1-M1
cat('probability that lawn mower is not easy to opratable is ',M1bar)
#b)the probability that a used car will either not have low mileage or not be
#expensive to operate
M1andC3=0.08
AbarorBbar=1-M1andC3
cat('probability that lawn mower is not easy to opratable  or not have high coat',AbarorBbar)
```

R code Exa 3.20 Conditional Probability

```
#A=Communication system have high fideality
#B=Communication system have high selectivity
#Given p(A)=0.81,p(A ??? B)=0.18,find=p(B/A)=p(A ??? B)/p(A)
pA=0.81
inter=0.18
result=inter/pA
cat("Probability that system have high fideality will also have selectivity is",result)
```

R code Exa 3.21 Conditional Probability

```
#Conditional Probability
#Given
M1IntC3=0.08#E1 INTERSECTION C1
PC3=0.4#probability of C1
#PrObability is:
P=M1IntC3/PC3
```

```
message("There for the probability is : ",P)
```

R code Exa 3.22 Conditional Probability

```
N=20#tOTAL GROUP OF WORKER
n1=12#Worker favor the regulation
n2=8#workers against the regulation
#Find the probability that if two worker select are
    against the regulation
F1=(n2/N)#For the first selection
F2=(n2-1)/(N-1)#the second selection from n2−1 and N
    −1
P=F1*F2#PROBABILITY
print("The probablity that the two selected worker
    are against the regulation is,")
P
```

R code Exa 3.24 Conditional Probability

```
N=52#Total cards
k=2#two card are selected at rendom
T=4#Total aces in 52 cards
#Find the probability of getting aces
#(a) first card is replaced before second card
message("Probability is : ",(4/52)*(4/52))
#(b) first card is not replaced before second card
message("Probability is : ",(4/52)*(3/51))
```

R code Exa 3.25 Conditional Probability

```
#Find where C and D is Independ events
C=0.65#p(C)
D=0.4#P(D)
CandD=0.24
#if C and D are Independ events then CandD=p(C).P(D)
result=C*D
if(result==CandD){
  print("Events are independent")
}else{
  print("Events are dependent")
}
```

R code Exa 3.26 Conditional Probability

```
#A=the event that raw material available when needed
#B=The event that the machining time is less than 1
   hours
PA=0.8#P(A)
PB=0.7#P(B)
#Proababilty of A intersection B
AIntB=PA*PB
AIntB
```

R code Exa 3.28 Conditional Probability

```
#Many companies must moniter the effluent that is
   discharged gfrom there plants into river and
   waterway.
p=0.01 #XProbability that the measurement on a test
   specimen will exceed L(limit)
#a)Find the probability of fail for to test
q=1-p#Probability of fail
#Both process is independent
```

```
#Ther for by product rule
prob=q*q
message("The Probability is ",prob)
#b) probability that test is free of contantminant
   conducted for two year one day a week
#total weeks in two year is 104 ther for
prob=q^104
message("The Probability is ",prob)
```

R code Exa 3.29 Conditional Probability

```
#There is two ways of data transmissio 1st is one
   digit transmission 2nd is three digit
   transmission method
#a)Evaluate the probability that the transmitted 1
   will recevied 1 only by three digit scheme
#Where the probability are
p1=0.01
p2=0.02
p3=0.05
#if we send 111 there is probability for all three
   digit are (1-p)*(1-p)*(1-p)
#Also the Probability of receving 011 is p*(1-p)*(1-
   p)
#so the probability of 0 is 3*p*(1-p)*(1-p)
#By rule
prob1=(1-p1)^3+3*p1*(1-p1)*(1-p1)#Probability of
   getting correct
message("Probability is : ",prob1)
prob2=(1-p2)^3+3*p2*(1-p2)*(1-p2)#Probability of
   getting correct
message("Probability is : ",prob2)
prob3=(1-p3)^3+3*p3*(1-p3)*(1-p3)#Probability of
   getting correct
message("Probability is : ",prob3)
```

```
#check where  digit  1 and 0 are  receved correcly
   in the prob of 0.05
#in part on the probability under 0.05 is 0.99275
message("The Total probability is : ",prob3*prob3)
```

R code Exa 3.30 Bayes Theorem

```
#Baye's theorem
B1=0.2#Prabaility of janet
B2=0.6#Prabaility of tom
B3=0.15#Prabaility of georgin
B4=0.05#Prabaility of peter
AbyB1=1/20
AbyB2=1/10
AbyB3=1/10
AbyB4=1/20
#Prabability tah intial repaire made by janet
B1byA=(B1*AbyB1)/(B1*AbyB1+B2*AbyB2+B3*AbyB3+B4*
   AbyB4)
B1byA
```

R code Exa 3.31 Bayes Theorem

```
B1= 1700#message is spam
B2=3300#message is normal
#A is words in the list
M=5000#TOtal messages
#Probability for spam message
PB1=B1/M
message("The Probability is: ",PB1)
#Probability for normal message
PB2=B2/M
message("The Probability is: ",PB2)
```

```
#Among the spam messages,1343 contain words in the
    list
B1IntA=1343
# from normal messages only 297 contain words in the
    list
B2IntA=297
#Conditional Probability
PAB1=B1IntA/B1#P(A|B1)
PAB2=B2IntA/B2#P(A|B2)
#There for by Bayes' Theorem
PB1A=(PAB1*B1)/(PAB1*B1+PAB2*B2)#P(B1|A)=P(A|B1)*P(
    B1)/(P(A|B1)*P(B1)+P(A|B2)*P(B2))
PB1A
print("Since the probability is large there for the
    message is spam")
```

Chapter 4

PROBABILITY DISTRIBUTIONS

R code Exa 4.1 Random Variables

```
#Checking for probability distribution(PD)
#a)
probability <- function(f) {
  print(f)
  i=1;flag=FALSE
  while (i<length(x)){
    if(f[i]<0){
      flag=FALSE
      print("Nagative value is ",f[i])
      break

    }
    else{ flag=TRUE}
    i=i+1
  }
  if(flag==TRUE && sum(f)==1){
   return("it is PD")
  }
  else{
```

```
    return("it is not PD")
  }
}
x=seq(1,4,by=1)
f=(x-2)/2
print(probability(f))
y=seq(0,4,by=1)
g=(x**2/25)
print(probability(g))
```

R code Exa 4.3 The Binomial Distribution

```
#(a) List all possible outcomes in terms of success, S, repaired within one hour, and
#failure, F, not repaired within one hour.
list=c('FFF','FFS',' FSS',' SSS','FSF','SFS','SFF','SSF')
cat("The Possible outcome is: ",list)
#(b) Find the probability distribution of the number of successes, X, among the
#3 repairs.
#X=0
cat("The Probability is: ",0.1*0.1*0.1)
#X=1
cat("The Probability is: ",3 * 0.009)
#X=2
cat("The Probability is: ",3*(0.1 * 0.9 * 0.9 ) )
#X=3
cat("The Probability is: ",0.9* 0.9* 0.9)
```

R code Exa 4.4 The Binomial Distribution

```
#a)
```

```
n=5#no of trail
r=4#no of successor
p=0.6#probability of success
# Create the binomial distribution.
prob <- dbinom(r,n,p)#This function gives the
#probability density distribution at each point
cat("The probability  is",prob)
#b)
#at least four of five installations
prob=dbinom(4,5,p)+dbinom(5,n,p)
cat("The probability bill reduce by one third is",
   prob)
```

R code Exa 4.5 The Binomial Distribution

```
P=0.05#probability
n=16
#(a)at most two will fail
pbinom(2,16,P)
#(b) at least four will fail
1-pbinom(3,16,P)
```

R code Exa 4.6 The Binomial Distribution

```
p=0.65
n=15
#probability for 11 will be written by the algorithm
prob=dbinom(11, 15, 0.65)
message("probability is : ",prob)
#probability for at least 10 will be written by the
   algorithm
prob=1-pbinom(9,15,0.65 )
message("probability is : ",prob)
```

```
#probability between 8 and 11 inclusive will be
    written by the algorithm
prob=pbinom( 11, 15, 0.65 ) - pbinom( 7, 15, 0.65 )
message("probability is : ",prob)
```

R code Exa 4.7 The Binomial Distribution

```
#Claim:10% of his machine require repair within the
    warranty period of 12 month
p=0.10#Probability
n=20#Total fax machine
#P>5 probability for 5 or more machine require
    repairs
prob=1-pbinom(4,n,p)
message("Probability : ",prob)
message("Since The probability is to small there for
     we reject the claim")
```

R code Exa 4.8 The Hypergeometric Distribution

```
#hypergeometric distribution
#for sample without replacement
x=2#successer in n
n=10#sample size
N=20#lot Size (Population size)
a=5#succeser in N
h=(choose(a,x)*choose(N-a,n-x))/choose(N,n)
cat("probabilit is",h)
```

R code Exa 4.9 The Hypergeometric Distribution

```
#binomial distribution(bd) and hypergeometric
    distribution(hd)
n=10#sample
x=2#2 defective out of 10
#a)hd
a=25#total defective
N=100#Toale tape
h=(choose(a,x)*choose(N-a,n-x))/choose(N,n)
message("probabilit is ",h)
#b)using bd
p=0.25#probability
dbinom(2,10,p)
```

R code Exa 4.10 The Mean and the Variance of a Probability Distribution

```
s<-c('HHH','THH',' HTH','TTH','HHT',' THT','HTT','
   TTT')#Total sample space
x<-c(0,1,2,3)#Random variable for Head
p<-c(1/8,3/8,3/8,1/8)#Probability for 0,1,2,3 head
data.frame(x,p)
mean=sum(x*p)#mean=sum(x.f(x))
mean
```

R code Exa 4.11 The Mean and the Variance of a Probability Distribution

```
#Given
x<-c(0,1,2,3)
prob<-c(0.18,0.50,0.29,0.03)
#Mean of probabilites
mu=weighted.mean(x,prob)
message("Mean is :",mu)
```

R code Exa 4.12 The Mean and the Variance of a Probability Distribution

```
n=3#no of flips of coin
p=1/2#Probability of head
mu=n*p
message("There for mean is",mu)
```

R code Exa 4.13 The Mean and the Variance of a Probability Distribution

```
#Mean of hypergeometric distribution
n=10#random variable
a=5#Total deffective
N=20#Total taps
#mean
mu=n*a/N
cat("Mean is",mu)
```

R code Exa 4.14 The Mean and the Variance of a Probability Distribution

```
#Standard deviation for  Probability distribution
#(a)
x<-c(0,1,2,3,4)
Fx<-c(1/16,4/16,6/16,4/16,1/16)
mu=weighted.mean(x,Fx)
mu
var=sum(((x-2)^2)*Fx)
var
message("The standard deviation is ",sqrt(var))
```

R code Exa 4.15 The Mean and the Variance of a Probability Distribution

```
#variance using alternative computing formula
s=c(1,2,3,4,5,6)#sample space
p=1/6#Probability of die
mu1=sum(p*s)
mu1
mu2=sum(p*s**2)
var=mu2-mu1**2
var
cat("Variance is",var)
```

R code Exa 4.16 The Mean and the Variance of a Probability Distribution

```
x<-c(0,1,2,3,4)
Fx<-c(0.05,0.20,0.45,0.20,0.10)
xFx<-x*Fx
cbind(x,Fx,xFx)
mu=weighted.mean(x,Fx)
var=sum((x^2)*Fx)-(mu)^2
message("mean :",mu," variance: ",var)
```

R code Exa 4.17 The Mean and the Variance of a Probability Distribution

```
n=16
p=1/2
#the variance of a binomial distribution
q=1-p
var=n*p*q
cat("The Variance is: ",sqrt(var))
```

R code Exa 4.18 The Mean and the Variance of a Probability Distribution

```
#variance of hypergeometric distribution
N=20#Total tap recorder
n=10#small sample
a=5#defective tap
sigma_seq=(n*a*(N-a)*(N-n))/((N^2)*(N-1))
var=sigma_seq**0.5
cat("variance of hd is",var)
```

R code Exa 4.19 Chebyshevs Theorem

```
mu=18
sigma=2.5
#Probability for the we can assert that there will
    be customer be between 8 and 28
n=8
n2=28
K1=(n2-mu)/sigma
k2=(mu-n)/sigma
#There for the Probability is
Prob=1-1/K1^2
message(Prob)
```

R code Exa 4.20 Chebyshevs Theorem

```
#show that for 40000 flips of coin th probality of
    head is between 0.47 and 0.525
n=40000
```

```
prob=0.99
p=1/2#probability of success
q=1-p
mu=n*p
sigma=(n*p*q)^0.5
sigma
k=(1/(1-prob))^0.5
val1=mu-k*sigma
val2=mu+k*sigma
p1=val1/n
p2=val2/n
cat("Hence the probaility is at leat 0.99 that we
    have probabilty of getting had in range ")
cat(p1 ,'and' ,p2)
```

R code Exa 4.21 The Poisson Distribution and Rare Events

```
lambda=1.3#gamma particales per millisecond coming
    from radio active substain
#(a)probability for exactly one week
prob=dpois(1,1.3)#P(X=1)
prob
#(b)probability for one or more gamma parical i.e. P
    >1
prob=1 - ppois(0, 1.3)#1-P(X=0)
prob
#(c)the probability of at least two but no more than
     four particles
prob=ppois(4,lambda)-ppois(1,lambda)
prob
#(d)vaiance
print("For Poisson the variance equal to lambda i.e
    1.3")
```

R code Exa 4.22 The Poisson Distribution and Rare Events

```
#(a)the formula for the binomial distribution
x = 2
n = 100
p = 0.05
dbinom(x,n,p)
#(b)the Poisson approximation to the binomial
    distribution.
dpois(2,5)
```

R code Exa 4.24 Poisson Processes

```
alpha=6
#4 bad check for any one day
x=4
T=1
lambda=alpha*T
prob=dpois(x,lambda)
cat("THER FOR th e probability that 4 bad check on
    any day is",prob)
#For 10 bad check on  Two Consecutive dayes
T=2
x=10
lambda=alpha*T
prob=ppois(x,lambda) - ppois(x-1,lambda)
cat("THER FOR th e probability that 10 bad check on
    any two day  is",prob)
```

R code Exa 4.25 Poisson Processes

```
#Here in this example ppios() for Poisson
    distribution
alpha=0.2
#one interruption in 3 weeks
x=1
T=3
lambda=alpha*T
prob=ppois(x,lambda)-ppois(0,lambda)
prob
cat("probability is",prob)
#at least two interruptions in 5 weeks
T=5
x=1
lambda=alpha*T
prob=1-ppois(x,lambda)
cat("probability is",prob)
#For at most one imperfaction in 15 week
T=15
x=1
lambda=alpha*T
prob=ppois(x,lambda)
cat("probability is",prob)
```

R code Exa 4.26 The Geometric and Negative Binomial Distribution

```
#Geometric Distribution
#it deals with the number of trials required for a
    single success
#the geometric distribution is a negative binomial
    distribution
options(digits = 2)
p=0.05
x=6
```

```
#by function
dgeom(x, prob = p)
#by formula
g=p*(1-p)^(x-1)
g
cat("Probability that 6 measuring device show excessive brift is",g)
```

R code Exa 4.27 The Multinomial Distribution

```
#Multinomial Distribution
#(X1, . . . , Xk) ??? multinom(size = n, prob = pk1)
x<-c(2,5,1)
prob<-c(0.3,0.5,0.2)
dmultinom(x, size = 8, prob, log = FALSE)
```

Chapter 5

PROBABILITY DENSITIES

R code Exa 5.1 Continuous Random Variables

```
library(distr)#PDF
#f(x)=2e^-2x for x >0
f=function(x) 2*exp(1)^(-2*x)
#P(1<X<3)
integrate(f, lower = 1, upper = 3)
#P(X>0.5)
integrate(f, lower = 0.5, upper = Inf)
```

R code Exa 5.2 Continuous Random Variables

```
#Distribution function for x=1
f=function(x) 1-exp(1)^(-2*x)
f(1)
```

R code Exa 5.3 Continuous Random Variables

```
#f(x)=k*xe^-4*x^2,find K x>0
f=function(x) x*exp(1)^(-4*x^2)
k=1/(integrate(f, lower = 0, upper = Inf))$value
message("Value of K is",k)
```

R code Exa 5.4 Continuous Random Variables

```
#f(x)=2e^(-2x) for x>0#is PDF
#MEAN
#f(x)=xf(x)
options(digits = 3)
f=function(x) 2*x*exp(1)^(-2*x)
mean <- integrate(f, lower =0, upper =Inf)
cat("Mean for PDF is",mean$value)
#variance
#var=E(X^2)-E(X)^2
E=function(x) 2*(x^2)*exp(1)^(-2*x)#E(X^2)
E1=integrate(E, lower =0, upper =Inf)
var=E1$value-((mean$value)**2)
cat("Varience for PDF is",var)
```

R code Exa 5.5 The Normal Distribution

```
#Standered Normal Distribution
mean=0
sd=1
#a)area between 1.28  and 0.87
a=pnorm(1.28) -pnorm(0.87)
a
#b)area between  -0.34 and 0.62
b=pnorm(0.62) -pnorm(-0.34)
b
#c)greater then 0.85
```

```
c=pnorm(0.85,lower.tail = FALSE)
c
#c)greater then −0.65
d=pnorm(-0.65,lower.tail = FALSE)
d
```

R code Exa 5.6 The Normal Distribution

```
#Find value of z0.01 and z0.05
z0.01=qnorm(1-0.01)
z0.05=qnorm(1-0.05)
message("z0.01 is : ",z0.01," and z0.05 is : ",z0
    .05)
```

R code Exa 5.7 The Normal Distribution

```
mu= 10.1
sigma= 2.7
#a) For the next food or product, what is the
    probability that its maximum
#attenuation is between 8.5 dB and 13.0 dB?
a=pnorm((13-mu)/sigma) -pnorm((8.5-mu)/sigma)
a
#b)According to the normal model, what proportion of
     the products have
#maximum attenuation between 8.5 dB and 13.0 dB?
cat("o 0.5801 as the proportion having maximum
    attenuation between
8.5 and 13.0 dB")
#c)What proportion of the products have maximum
    attenuation greater than 15.1 dB
1-pnorm((15.1-mu)/sigma)
```

R code Exa 5.8 The Normal Distribution

```
weight<-c(72.2,67.8,78.0,64.4,76.3,72.3,73.1,71.7,66.2,63.3,85.4,67.4,
66.3,76.3,57.7,50.3,77.4,63.1,73.9,67.4,74.7,68.2,87.4,86.4,
69.4,58.0,63.3,72.7,73.6,68.8,63.3,63.3,73.0,64.8,73.1,70.9,
85.9,74.4,75.9,72.3,84.3,61.8,79.2,64.3,65.4,66.7,77.2,50.0,
70.3,90.4,63.9,62.1,68.2,55.1,52.6,68.5,55.2,73.5,53.7,61.7,
47.9,72.3,61.1,71.8,83.1,71.2,58.8,61.8,86.8,64.5,52.3,58.3,
65.9,80.2,75.1,59.9,62.3,48.8,64.3,75.4)
n = 80
mu=68.4 #pounds
sd = 9.6 #pounds
#(a) Find the probability of using 80 or more pounds of cheese.
1-pnorm((n-mu)/sd)
#Set a limit so that only 10 % of production runs have less than L pounds of cheese
Z0.1=qnorm(1-0.1)#probability closest to .1000
L = mu - sd*Z0.1
L #Pounds
#Determine a new mean for the distribution so that only 5 % of the runs have less than L pounds
Z0.05=1.645
mu = L + 9.6 * Z0.05
mu
cat("The mean must be increased by 3.5 pounds to decrease the percentage of units
```

```
below the limit L from 10% to 5%")
```

R code Exa 5.9 The Normal Distribution

```
mu=4.6
sigma=1.21
X=log(0.0015)
value=(X+(mu))/sigma**0.5
value
pnorm(value)
cord.x <- c(-3,seq(-3,value,0.01),value)
cord.y <- c(0,dnorm(seq(-3,value,0.01)),0)
curve(dnorm(x,0,1),xlim=c(-3,3),main='Standard
    Normal')
polygon(cord.x,cord.y,col='skyblue')
print('The Standard Normal Pobabability that obtain
    is smaller' )
```

R code Exa 5.10 The Normal Distribution

```
mu = 11.6
sigma = 3.3
#what is the probability that there will be at least
     8 outages in any one month?
#we take 7.5 instand of 8
Z=1-pnorm((7.5-mu)/sigma)
Z
```

R code Exa 5.11 The Normal Approximation to the Binomial Distribution

```r
#We take the probability p = 0.056 which is the value estimated from the physics experiment
p = 0.056
n=300
#(a) find the mean and standard deviation of the number which will disappear
mean=n*p
sd=sqrt(n*p*(1-p))
mean
sd
#(b) Approximate the probability that 12 or more will disappear.
1-pnorm((11.5-mean)/sd)
#(c) Approximate the probability of exactly 12
pnorm((12.5-mean)/sd)-pnorm((11.5-mean)/sd)
```

R code Exa 5.12 The Log Normal Distribution

```r
#Log-normal Distribution
alpha=2#mean
beta=0.1#variance
#Find the Probaility between 8.1 to 6.1
plnorm(8.2,2,0.1)-plnorm(6.1,2,0.1)
```

R code Exa 5.13 The Log Normal Distribution

```r
#Histogram of interrequest time data
data=c(2208,4201,3848,9112,2082,5913,1620,6719,21657,
      3072,2949,11768,4731,14211,1583,9853,78811,6655,

      1803,7012,1892,4227,6583,15147,4740,8528,10563,
```

```
        43003,16723,2613,26463,34867,4191,4030,2472,28840,
        24487,14001,15241,1643,5732,5419,28608,2487,995,
        3116,29508,11440,28336,3440
)
hist(data, freq = FALSE,xlab = "Time")
lines(density(data), col="blue", lwd=2)
```

R code Exa 5.14 The Log Normal Distribution

```
alpha=4
beta=0.3
a=33
#log norm probability
P=1-plnorm(a,alpha,beta)
P
```

R code Exa 5.15 The Log Normal Distribution

```
alpha=2
beta=0.1
#mean
mu=exp(1)^(alpha+((beta)^2)/2)
mu
cat("mean is",mu)
#variance
var=exp(1)^((2*alpha+beta^2))*(exp(1)^(beta^2)-1)
var
cat("variance is",var)
```

R code Exa 5.16 The Gamma Distribution

```
data<-c(0.894,0.991,0.061,0.186,0.311,0.817,2.267,0.091,0.139,0.083,
        0.235,0.424,0.216,0.579,0.429,0.612,0.143,0.055,0.752,0.188,
        0.071,0.195,0.082,1.653,2.010,0.158,0.527,1.033,2.863,0.365,
        0.459,0.431,0.092,0.830,1.718,0.099,0.162,0.076,0.107,0.278,
        0.100,0.919,0.900,0.093,0.041,0.712,0.994,0.149,0.866,0.054)
#decay Time is represented by histogram
h<-hist(data, breaks=10, col=”red”,xlab = ”Millisecond” ,prob = TRUE)
xfit<-seq(0,max(data),length=50)
yfit<-dnorm(xfit,mean=mean(data),sd=sd(data))
yfit <- yfit*diff(h$mids[1:2])*length(data)
lines(xfit, yfit, col=”blue”, lwd=2)
```

R code Exa 5.17 The Gamma Distribution

```
f=function(x) 3*exp(1)^(-3*x)
#Assuming that the arrival follow possion process with alpha=3 and beta=1/3
#a)less then 5 min
integrate(f,lower = 0,upper = 1/12)
#b)at least 45 min
integrate(f,lower = 3/4,upper = Inf)
```

R code Exa 5.18 The Beta Distribution

```
#beta distribution
#a)mean
alpha=3
beta=2
mu=alpha/(alpha+beta)
mu
cat("There for on average",mu*100,"% highway section
    require repaire")
#b)
F=pbeta(1/2,3,2)
plot(density(pbeta(seq(0,1/2,by=0.1),3,2)))
cat("The probability is",F)
```

R code Exa 5.19 The Weibull Distribution

```
#weibull distribution
scale=0.5
#mean lifetime
mu=(0.1)**(-2)*factorial((1+1/0.5)-1)
mu
#P(X>300)
g=function(x)(0.05)*x^(-0.5)*exp(1)^(-0.1*x^scale)
integrate(g, lower = 300, upper = Inf)
```

R code Exa 5.20 Joint Distributions Discrete and Continuous

```
#P(X1+X2>1)
P=0.2+0.1+0
message("Probability is: ",P)
#P(Xi=xi)
p0=0.1+0.2#p(X=0)
p1=0.4+0.2#p(X=1)
p3=0.1+0#p(X=2)
```

```
message("Probabilits is: ")
p0
p1
p3
```

R code Exa 5.21 Joint Distributions Discrete and Continuous

```
#Check the independency of X1 and X2
#Conditional probability
#X2=1
#f(0|1)
f1=0.2/0.4#f(0|1)/f2(1)
#f(1|1)
f2=0.2/0.4#f(1|1)/f2(1)
#f(2|1)
f3=0/0.4#f(2|1)/f2(1)
f1
f2
f3
message("since f(0|1)!=f1(0) there for it's
    dependent")
```

R code Exa 5.22 Joint Distributions Discrete and Continuous

```
llimy <- 2;  llimx=1
ulimy <-  3    ; ulimx=2
#the first random variable will take on a value
    between 1 and 2 and the second
#random variable will take on a value between 2 and
    3
f <- function(x1,x2)6*exp(1)^(-2*x1-3*x2)
integrate(function(x2) {
  sapply(x2, function(x2) {
```

```
    integrate(function(x1) f(x1,x2), llimx, ulimx)$value
  })
}, llimy, ulimy)
#(b) the first random variable will take on a value less than 2 and the second
#random variable will take on a value greater than 2
llimy <- 2;  llimx=0
ulimy <-  Inf    ; ulimx=2
f <- function(x1,x2)6*exp(1)^(-2*x1-3*x2)
integrate(function(x2) {
  sapply(x2, function(x2) {
    integrate(function(x1) f(x1,x2), llimx, ulimx)$value
  })
}, llimy, ulimy)
```

R code Exa 5.23 Joint Distributions Discrete and Continuous

```
F<-function(x,y){
  return (1 - exp(1)^-2*x)*(1 - exp(1)^-3*y)
}
F(1,1)
```

R code Exa 5.31 Joint Distributions Discrete and Continuous

```
EX1=4
EX2=-2
VX1=9
VX2=6
#Calcuation
cat("E( 2X1+X2−5 ) = ",2*EX1+EX2-5)
cat("Var(2X1+X2−5 ) = ",2^2*VX1+VX2)
```

R code Exa 5.42 Checking If the Data Are Normal

```
# Normal Scores plot of normalized data
Score=c(67, 48, 76, 81)
qqnorm(Score, ylab="Ordered observations", xlab="
   Normal Scores", main="")
```

R code Exa 5.43 Transforming Observations to Near Normality

```
#Histogram to Normal Scores plot of normalized data
data=c
   (2208,4201,3848,9112,2082,5913,1620,6719,21657,
  3072,2949,11768,4731,14211,1583,9853,78811,6655,
  1803,7012,1892,4227,6583,15147,4740,8528,10563,
  43003,16723,2613,26463,34867,4191,4030,2472,28840,
  24487,14001,15241,1643,5732,5419,28608,2487,995,
  3116,29508,11440,28336,3440
)
hist(data,ylab="Class frequency", xlab="time")
qqnorm(data, ylab="Interrquest time", xlab="Normal
   Scores", main="")
ln=log(data)
hist(ln,ylab="Class frequency", xlab="ln(time)")
qqnorm(ln, ylab="ln(Interrquest time)", xlab="Normal
    Scores", main="")
```

R code Exa 5.44 Simulation

```
Data<-c(0.57,0.74,0.26,0.77,0.12)
```

```
alpha = 0.05
beta = 2.0
((-1/alpha)*log(1-Data))^(1/beta)
```

R code Exa 5.45 Simulation

```
mu=50
sigma=5
u1=0.253
u2=0.531
#standard normal values
z1=(-2*log(u2))**0.5*cos(2*pi*u1)
z1
z2=(-2*log(u2))**0.5*sin(2*pi*u1)
z2
#normal value
x1=50+5*z1
x2=50+5*z2
x1
x2
```

Chapter 6

SAMPLING DISTRIBUTIONS

R code Exa 6.49 The Sampling Distribution of the Mean

```
#Find the population correction factor
#Given
n=10
N=1000
factor=(N-n)/(N-1)
message("There for the factor is : ",factor)
```

R code Exa 6.51 The Sampling Distribution of the Mean

```
#sampling distribution
options(digits = 3)
sd=1.20
mean=1.82#mean
n=40
Xbar<-c(1.65,2.04)
Xbar
z=(Xbar-mean)*(n**0.5)/(sd)
z
```

```
message("Probability is: ",pnorm(1.16)-pnorm(-0.896)
   )
```

R code Exa 6.52 The Sampling Distribution of the Mean Sigma Unknown

```
n = 20
mu=40 #mg/l
xbar = 46
s = 9.4 #mg/l
df=n-1
t0.01=qt(1-0.01,df)
t0.01
t=(xbar-mu)/(s/n^0.5)
message("t value is: ",t)
P=1 - pt(t,19)
message("probability is: ",P)
print("Sinces the t value exceed t0.01 there for we
   reject the claim")
```

R code Exa 6.53 The Sampling Distribution of the Variance

```
var=1.35#population variance
s=1.4#sample variance
n=20#total no of sample
chi=((n+2)*1.4*1.4)/1.2^2
ceiling(chi)
qchisq(1-0.05,19)
1 -pchisq(30.6, 19)
cat("There for the probability that a good shipmeant
 will erroneously be reject is less then 0.05")
```

R code Exa 5.54 The Sampling Distribution of the Variance

```
#Random sample size
n1=7
n2=13
#there for numerator and denominator degree of
   freedom is
v1=n1-1
v2=n2-1
1/df(0.05,v1,v2)
F0.05=qf(1-0.05,6,12)
F0.05
print("There for the probability is 0.05")
```

R code Exa 6.55 The Sampling Distribution of the Variance

```
#F distribution
#left-hand tail probability
#a)by first method
df1=20#first degree of freedom
df2=10#second degree of freedom
data=qf(0.95,df1,df2)#quantile
Area=1/data
cat("Value under th area 0.95 is",Area)
```

Chapter 7

Inferences Concerning a Mean

R code Exa 7.2 Point Estimation

```
Data<-c(136,143,147,151,158,160,
161,163,165,167,173,174,
181,181,185,188,190,205)
Xbar=mean(Data)
s=sd(Data)
#standard error is
E=s/sqrt(length(Data))
E
```

R code Exa 7.3 Point Estimation

```
n=150
sigma=6.2
Z0.05=2.575
E=sigma*Z0.05/sqrt(n)
E
message("Thus, the engineer can assert with
    probability 0.99 that his error will be at
most 1.30.")
```

R code Exa 7.4 Point Estimation

```
# 98% confidence about the maximum error?
n = 6
s = 1.14
t0.01 =qt(1-0.01,5)
E=t0.01*s/sqrt(n)
E
message("Thus the chemist can assert with 98%
    confidence that his figure for the melting point
of the aluminum alloy is off by at most 1.566049
    degrees")
```

R code Exa 7.5 Point Estimation

```
E = 0.50
sigma = 1.6
z0.025 = 1.96
n=(sigma*z0.025/E)^2
n
message("Thus, the research worker will have to time
     40 mechanics performing the task of
rotating the tires of a car")
```

R code Exa 7.6 Interval Estimation

```
n=100#random sample of size
sigma=5.1
xbar=21.6
```

```
z0.025=1.96
#ther for 95% confidence interval for the population
    mean is
Int1=xbar-z0.025*(sigma/sqrt(n))
Int2=xbar+z0.025*(sigma/sqrt(n))
cat("There for confidence interval mean is :",Int1,"
    < ",Int2)
```

R code Exa 7.7 Interval Estimation

```
n =50
x = 305.58#nm
var = 1366.86
s = 36.97#nm
#construct a 99% confidence interval for the
   population mean of all nanopillars
z0.005 = 2.575
Int1=x-z0.005*(s/sqrt(n))
Int1
Int2=x+z0.005*(s/sqrt(n))
Int2
message("There for the confidence interval is : ",
   Int1," < ",Int2)
```

R code Exa 7.8 Interval Estimation

```
n = 18
x = 22.6
s = 15.7
t0.025=qt(1-0.025,17)
Int1=x-t0.025*s/sqrt(n)
Int1
Int2=x+t0.025*s/sqrt(n)
```

```
Int2
cat("We are 95 % confident that the interval from 14.79 to 30.41 MJ/m3 contains the
mean toughness of all possible artificial fibers created by the current process.")
```

R code Exa 7.13 Maximum Likelihood Estimation

```
#Poisson distribution of defective hard drive for ten day
Data<-c(7,3,1,2,4,1,2,3,1,2)
T=sum(Data)
lamda=T/10
#the maximum likelihood estimate is
#P(X=0 or 1)
P=exp(1)^(-lamda)+(lamda*exp(1)^(-lamda))
P
cat("There  for there will be 1 or fewer defectives on just over one-quarter of the days.")
```

R code Exa 7.15 Maximum Likelihood Estimation

```
#The yields of chemical rection is as
op<-c(5.57,5.76,4.18,4.64,7.02,6.62,6.33,7.24,
  5.57,7.89,4.67,7.24,6.43,5.59,5.39)
#(a) Obtain the maximum likelihood estimates of the mean yield and the variance
mean=mean(op)
cat(mean,"gal")
var=(1/length(op))*sum((op-mean)^2)
var
#Obtain the maximum likelihood estimate of the coefficient of variation ??/??
```

```
cat("coefficient is : ",sqrt(var)/mean)
```

R code Exa 7.18 Hypotheses Concerning One Mean

```
xbar=3.9
mu=4.5
sigma=1.5
n=25
Z=(xbar-mu)/(sigma/sqrt(n))
Z
P=pnorm(Z)
P
cat("There the probability that the value of z is −2 is 0.0227")
message("There for the p value is: ",2*P)
```

R code Exa 7.19 Hypotheses Concerning One Mean

```
xbar = 68.45
s = 9.583
mu=71
#Null hypothesis: mu = 71 pounds
#Alternative hypothesis: mu < 71 pounds
z.alpha=qnorm(0.01)
#Criterion: Since the probability of a Type I error is greatest when mu = 71
#pounds, we proceed as if we were testing the null hypothesis mu = 71 pounds
#against the alternative hypothesis mu < 71 pounds at the 0.01 level of
#significance. Thus, the null hypothesis must be rejected if Z < −2.33, where
Z=(xbar-mu)/(s/sqrt(80))
```

```
Z
#Decision: Since Z = −2.38 is less than −2.33, the null hypothesis must be
#rejected at level of significance 0.01. In other words, the suspicion that mu < 71
#pounds is confirmed.
```

R code Exa 7.20 Hypotheses Concerning One Mean

```
#measurements of lead content( ?? g / L) are taken from twelve water specimens spiked with a known concentration
data<-c(2.4,2.9,2.7,2.6,2.9,2.0,2.8,2.2,2.4,2.4,2.0,2.5)
x = 2.483
s = 0.3129
#Null hypothesis: mu = 2.25 mig/L
#Alternative hypothesis: mu > 2.25 mig/L
mu=2.25
#Criterion: Reject the null hypothesis if t > 2.201, where 2.201 is the value of
#t0.025 for 12−1 = 11 degrees of freedom
t0.025=qt(1-0.025,11)
t.test(data,mu=2.25, alt='greater',conf=.95)
cat("Decision: Since t = 2.58 is greater than 2.201, the null hypothesis must
be rejected. In other words, the mean lead content is above
2.25 mig/L.")
```

R code Exa 7.22 Power Sample Size and Operating Characteristic Curves

```
sd=3.6#dB
```

```
mu0=75.2
#null hypothesis mu0 = 75.20
#alternative hypothesis mu0 > 75.20
n=15#there are 15 machines
mu=77#dB
#The test is one−sided
alpha = 0.05
z0.05 = qnorm(1-0.05)
Z=z0.05+sqrt(n)*((mu0-mu)/sd)
Z
P=pnorm(Z,lower.tail = FALSE)
P
cat("There for the probability for type 2 error is:
    ",1-P)
```

R code Exa 7.23 Power Sample Size and Operating Characteristic Curves

```
#The test is two−tailed
mean=2.000#cm
sd=0.050#cm
mu0=2
#FIND THE PROBABILITY OF TYPE 2 ERROR
z0.025 = qnorm(1-0.025)
mu= 2.010
n=30
Z1=z0.025+sqrt(n)*((mu0-mu)/sd)
Z1
Z2=-z0.025+sqrt(n)*((mu0-mu)/sd)
Z2
P=pnorm(Z1,lower.tail = FALSE)+pnorm(Z2)
P
cat("The probability is: ",1-P)
```

R code Exa 7.24 Power Sample Size and Operating Characteristic Curves

```
mu0 =1600
sigma = 192
alpha= 0.05
mu1=1680
z0.05 = qnorm(1-0.05)
beta= 0.10
z0.10 = 1.28
n=(sigma*(z0.05+z0.10)/(mu0-mu1))^2
n
cat("There for the sample size is: ",n)
```

Chapter 8

Comparing Two Treatments

R code Exa 8.3 Comparisons Two Independent Large Samples

```
#After changing the electrical pricing during peak
   hours for user havingair-conditioning
#and with-out it to study the variance ;
#n1=45 homes with Air-Conditioning
n1=45
mean1=204.4
var1=13825.3
#n2=55 home without it
n2=55
mean2=130.0
var2=8632.0
#Obtain a 95% confidence interval for delta = mu1 -
   mu2
alpha=0.05
z.alphahalf=qnorm(1-0.025)
#Hence the Confidence interval is,
Int1=mean1-mean2-z.alphahalf*sqrt((var1/n1)+(var2/n2
   ))
Int1
Int2=mean1-mean2+z.alphahalf*sqrt((var1/n1)+(var2/n2
   ))
```

```
Int2
cat("hence The mean on-peak usage for homes with air
   -conditioning
is higher than for homes without")
```

R code Exa 8.4 Comparisons Two Independent Large Samples

```
#Test for two types of Drivers
#Drivers with 0 Blood Alcohol
n1 = 54
xbar = 1.63
s1 = 0.177
#Drivers with 0.1 % Blood Alcohol
n2 = 54
ybar = 1.77
s2 = 0.183
z.alpha=qnorm(0.01)
z.alpha
#delta = mu1 - mu2
#Null hypothesis: delta  = 0 there is no difference
   in mean of drivers
#Alternative hypothesis: delta != 0 there is
   difference in means of driver
#Level of significance= 0.02
#Criterion: Reject the null hypothesis if Z < -2.33
   or Z > 2.33.
Z=(xbar-ybar)/sqrt((s1^2/n1)+(s2^2/n2))
Z
cat("Since z = -4.04 is less than -2.33, the null
   hypothesis must be
rejected at level of significance 0.02")
```

R code Exa 8.5 Comparisons Two Independent Large Samples

```
#We  testing the claim that the resistance of electric wire can be reduced by more than
#0.050 ohm by alloying
n1=n2=32
#For Standard wire
xbar = 0.136#ohm
s1 = 0.004 #ohm
#For alloyed wire
ybar = 0.083#ohm
s2 = 0.005#ohm
loc=0.05
#Null hypothesis: mu1-mu2 = 0.050
#Alternative hypothesis: mu1- mu2 > 0.050
z.alpha=qnorm(1-0.05)
z.alpha
#Criterion: Reject the null hypothesis if Z > 1.645, where Z is given as
Z=(xbar-ybar-loc)/sqrt((s1^2/n1)+(s2^2/n2))
Z
cat("Since z = 2.65 exceeds 1.645, the null hypothesis must be rejected;
that is, the data substantiate the claim")
```

R code Exa 8.6 Comparisons Two Independent Large Samples

```
deltabar=0.054
delta0=0.05
segma_square=0.000041
n=32
z.alpha=1.645
Z=z.alpha+sqrt(n)*(delta0-deltabar)/sqrt(segma_square)
P=pnorm(Z, lower.tail = FALSE)#P-value for P( Z > -1.889)
message("The Type II error probability is ",(1-P))
```

R code Exa 8.7 Comparisons Two Independent Small Samples

```
#recycled materials Data for two different location
Loc1<-c(707,632,604,652,669,674)
Loc2<-c(552,554,484,630,648,610)
#there is Measurements of  6 specimens of recycled materials
n1=n2=6
#delta=mu1-mu2 # where delta is difference in mean strength for materials from the two locations
#Null hypothesis: delta = 0#Same strength
#Alternative hypothesis: delta!= 0#differnt in strength
t0.025=qt(1-0.025,10)
t0.025
#Criterion: Reject the null hypothesis if t < -t0.025 or t > t0.025 where
#t0.025 = 2.228 for 6 + 6 - 2 = 10 degrees of freedom
t.test(Loc1,Loc2)
print("since the value of t is greater then t0.025 there for we reject the null hypothesis")
```

R code Exa 8.8 Comparisons Two Independent Small Samples

```
Alloy<-c(66.4,67.7,68.0,68.0,68.3,68.4,68.6,68.8,68.9,69.0,69.1,
69.2,69.3,69.3,69.5,69.5,69.6,69.7,69.8,69.8,69.9,70.0,
70.0,70.1,70.2,70.3,70.3,70.4,70.5,70.6,70.6,70.8,70.9,
```

```
71.0,71.1,71.2,71.3,71.3,71.5,71.6,71.6,71.7,71.8,71.8,
71.9,72.1,72.2,72.3,72.4,72.6,72.7,72.9,73.1,73.3,73.5,
74.2,74.5,75.3)
stem(Alloy)
Alloy2<-c
   (71.2,71.8,72.6,72.8,73.7,73.7,73.9,74.4,74.9,75.5,
          75.9,76.3,76.5,76.7,76.9,77.1,77.3,77.6,77.7,77.8,
          78.1,78.2,78.4,78.6,79.0,79.3,79.8)
stem(Alloy2)
n1=58#Alloy of 58 material
n2=27#Alloy of 27 material
t0.025 =qt(1-0.025,83) #for 83 degrees of freedom
print("T")
t.test(Alloy,Alloy2)
```

R code Exa 8.9 Comparisons Two Independent Small Samples

```
# 95% large sample confidence interval
Alloy<-c
   (66.4,67.7,68.0,68.0,68.3,68.4,68.6,68.8,68.9,69.0,69.1,
        69.2,69.3,69.3,69.5,69.5,69.6,69.7,69.8,69.8,69.9,70.0,
         70.0,70.1,70.2,70.3,70.3,70.4,70.5,70.6,70.6,70.8,70.9,
         71.0,71.1,71.2,71.3,71.3,71.5,71.6,71.6,71.7,71.8,71.8,
         71.9,72.1,72.2,72.3,72.4,72.6,72.7,72.9,73.1,73.3,73.5,
         74.2,74.5,75.3)

```

```
Alloy2<-c
   (71.2,71.8,72.6,72.8,73.7,73.7,73.9,74.4,74.9,75.5,
          75.9,76.3,76.5,76.7,76.9,77.1,77.3,77.6,77.7,77.8,
          78.1,78.2,78.4,78.6,79.0,79.3,79.8)
xbar=mean(Alloy)
ybar=mean(Alloy2)
s1=sd(Alloy)
s2=sd(Alloy2)
n1=58
n2=27
z.alpha=qnorm(1-0.025)
Int1=xbar-ybar-z.alpha*sqrt(s1^2/n1+s2^2/n2)
Int1
Int2=xbar-ybar+z.alpha*sqrt(s1^2/n1+s2^2/n2)
Int2
print("The 95% Confidence Interval is")
cat(Int1," < ",Int2)
```

R code Exa 8.10 Comparisons Two Independent Small Samples

```
#There are two types of catalyst for the chemical
   reaction
Catalyst1<-c
   (0.63,2.64,1.85,1.68,1.09,1.67,0.73,1.04,0.68)
Catalyst2<-c
   (3.71,4.09,4.11,3.75,3.49,3.27,3.72,3.49,4.26)
n1 = n2 = 9
loc=0.05
#Null hypothesis: mu1-mu2 = 0
#Alternative hypothesis: mu1-mu2!= 0
#We choose the Smith-Satterthwaite test statistic
   with delta0 = 0
#The null hypothesis will be rejected if t< -t0.025
```

```
or t > t0.025, but the value
#of t0.025 depends on the estimated degrees of freedom.
df=11
t0.025=qt(1-0.025,11)
t0.025
t.test(Catalyst1,Catalyst2,alternative ="g")
cat("Decision: Since t= −9.71 is less than −2.201, the null hypothesis must be
rejected at level of significance 0.05")
```

R code Exa 8.11 Comparisons Two Independent Small Samples

```
#There are two types of catalyst for the chemical reaction
Catalyst1<-c(0.63,2.64,1.85,1.68,1.09,1.67,0.73,1.04,0.68)
Catalyst2<-c(3.71,4.09,4.11,3.75,3.49,3.27,3.72,3.49,4.26)
# 95% confidence interval for delta = mu1 − mu2
x =mean(Catalyst1)
s1 = sd(Catalyst1)
y =mean(Catalyst2)
s2=sd(Catalyst2)
df=11
t0.025 =qt(1-0.025,df)
t.test(Catalyst1,Catalyst2)
cat("The mean product volume for the second catalyst is
greater than that of the first catalyst by 1.880 to 2.982 gallons")
```

R code Exa 8.12 Matched Pairs Comparisons

```
before<-c(45,73,46,124,33,57,83,34,26,17)
After<-c(36,60,44,119,35,51,77,29,24,11)
#Null Hypothesis:mu=0#There is no difference in mean
    of data there for the program is not effective
#Alternative hypothesis: mu>0 the program is
   effective
l.o.c=0.05
t0.05=qt(1-0.05,9)
t0.05
#Criterion:Reject Null Hypothesis if t>t0.05 or
   1.833 for df=9
#t calculation
t.test(before,After,paired=TRUE)
print("hence we reject Null hypothesis because t>t0
   .05")
print("There for the safety program is effective")
```

R code Exa 8.13 Matched Pairs Comparisons

```
#measurment of toxic level of chemical
data<-c(1.26,1.34,1.82,0.55,0.73,0.78,1.10)
n = 7 #Sample size
t.alpha=qt(1-0.025,6)
# the 95% confidence formula for mu becomes
t.test(data)
cat("We are 95% confident that the interval from
   0.68 to 1.49
contains the mean change in the red color component"
   )
```

R code Exa 8.14 Matched Pairs Comparisons

```
#11 observation of lab
```

```
Commercial_lab<-c(27,23,64,44,30,75,26,124,54,30,14)
State_lab<-c(15,13,22,29,31,64,30,64,56,20,21)
t0.025=qt(1-0.025,10)
t0.025
t.test(Commercial_lab,State_lab,paired=TRUE)
cat("This 95% confidence interval just covers 0, so no difference is indicated with this
small sample size")
```

Chapter 9

INFERENCES CONCERNING VARIANCES

R code Exa 9.1 The Estimation of Variances

```
#Data from chapter 8 and Example 7
location<-c(707,632,604,652,669,674)
n=6
d2=2.534#With refered to the table
min=604
max=707
R=max-min
sigma=R/d2
message("sigma for the resiliency modulus of
   recycled materials is: ",sigma)
```

R code Exa 9.2 The Estimation of Variances

```
#From Example 8, it is very reasonable to assume
   that the population is normal
weight<-c(72.2 ,67.8 ,78.0 ,64.4 ,76.3 ,72.3 ,73.1
   ,71.7 ,66.2 ,63.3 ,85.4 ,67.4,
```

```
            66.3 ,76.3 ,57.7 ,50.3 ,77.4 ,63.1 ,73.9 ,67.4 ,74.7 ,68.2 ,87.4 ,86.4,
            69.4 ,58.0 ,63.3 ,72.7 ,73.6 ,68.8 ,63.3 ,63.3 ,73.0 ,64.8 ,73.1 ,70.9,
            85.9 ,74.4 ,75.9 ,72.3 ,84.3 ,61.8 ,79.2 ,64.3 ,65.4 ,66.7 ,77.2 ,50.0,
            70.3 ,90.4 ,63.9 ,62.1 ,68.2 ,55.1 ,52.6 ,68.5 ,55.2 ,73.5 ,53.7 ,61.7,
            47.9 ,72.3 ,61.1 ,71.8 ,83.1 ,71.2 ,58.8 ,61.8 ,86.8 ,64.5 ,52.3 ,58.3,
            65.9,80.2,75.1,59.9,62.3,48.8,64.3,75.4)
n=80
df=n-1
chi1=qchisq(1-0.975,df)
chi2=qchisq(1-0.025,df)
s=sd(weight)
#Substituting into the formula for the confidence interval for sigma_square yields
Int1=df*(s^2)/chi2
Int2=df*(s^2)/chi1
message(Int1," < sigma_square <",Int2)
message(sqrt(Int1)," < sigma <",sqrt(Int2))
cat("This means we are 95% confident that the interval from 8.29 to 11.35 pounds
contains sigma")
```

R code Exa 9.3 Hypotheses Concerning One Variance

```
#sigma=0.5 null hypothesis thickness is equal to 0.5 mil
#sigma>0.5 alternative hypothesis
sigma=0.5
LOC=0.05#level of significance
X.chi=qchisq(1-0.05,14)
#Reject the null hypothesis if chi.seq>X.chi if
```

```
    degree if freedom is 14
n=15#total die cut
s=0.64#standard deviation
chi.seq=((n-1)*s**2)/sigma**2
chi.seq
if(chi.seq>X.chi){
  print("Reject null hypothesis")
}else{
  print("accept null hypothesis")
}
```

R code Exa 9.4 Hypotheses Concerning Two Variances

```
#Null Hypothesis:sigma_seq1 =  sigma_seq2
#Alternative Hypothesis:sigma_seq1 <  sigma_seq2
alpha=0.05#loc
F0.05=qf(1-0.05,11,11)#df1=11,df2=11
F0.05
#reject null hypothesis if F>F0.05 for 11 and 11
    degrees of freedom
s1=0.035
s2=0.062
F=s2**2/s1**2
F
print("hence the null hypothesis is rejectecd ")
#there for plating done by company 1 is less
    variable that done by company 2
```

R code Exa 9.5 Hypotheses Concerning Two Variances

```
#dealing with the strength of recycled materials for
    use in pavements
loca1<-c(707,632,604,652,669,674)
```

```
loca2<-c(552,554,484,630,648,610)
L.O.C=0.02
S1=var(loca1)
S2=var(loca2)
#Null Hypothesis:=sigma_seq1=sigma_seq2
#Alterative Hypothesis:=sigma_seq1 != sigma_seq2
F0.01=qf(1-0.01,5,5)
#reject null hypothesis if F>F0.01 for 4 and 5
    degrees of freedom
F=(S2)/(S1)
F
print("Since F doesn't exceed F0.01 there for null
    hypothesis can't be rejected")
```

R code Exa 9.6 Hypotheses Concerning Two Variances

```
n1=n2=9
s1=0.4548
s2=0.1089
#Obtain a 98% confidence interval for Sigma_square2/
    Sigma_square1
df1=8
df2=8
F=qf(1-0.01,df1,df2)
F0.99=1/F
Int1=F0.99*(s2/s1)
Int2=F*(s2/s1)
message("Interval 1: ",Int1," and Interval 2: ",Int2
    )
```

Chapter 10

INFERENCES CONCERNING PROPORTIONS

R code Exa 10.1 Estimation of Proportions

```
X=8
n=55
#The point estimate is
P=X/n
cat("The Point is: ",P)
# standard error is
E=sqrt((P*(1-P))/n)
cat("Error is : ",E)
```

R code Exa 10.2 Estimation of Proportions

```
#install.packages("binom") package binomial
    praportion
library(binom)
```

```
n=55#Wind turbines
r=8#Total noisy turbine
binom.confint( r, n, conf.level=0.95,methods="exact"
   )
cat("We are 95% confident that for proportion p of
   wind turbines that
are too noisy is in between 0.065 and 0.267")
```

R code Exa 10.3 Estimation of Proportions

```
x=36
n=100
#find 95%confidence interval
ratio=x/n
Z.alpha=1.96
Int1=ratio-Z.alpha*sqrt(ratio*(1-ratio)/n)
Int2=ratio+Z.alpha*sqrt(ratio*(1-ratio)/n)
cat(Int1," < p < ",Int2)
```

R code Exa 10.4 Estimation of Proportions

```
#Maximum error
n=400
x=136
Z.alpha=2.575
ratio=x/n
E=Z.alpha*sqrt((ratio)*(1-ratio)/n)
cat("Error is: ",E)
```

R code Exa 10.5 Estimation of Proportions

```
#95% confident that the error is at most 0.04
#(a) we have no idea what the true proportion might be
Z.alpha=qnorm(1-0.025)
E=0.04
n1=1/4*(Z.alpha/E)^2
message("The sample size is: ",ceiling(n1))
#(b)we know that the true proportion does not exceed 0.12
p=0.12
n2=p*(1-p)*(Z.alpha/E)^2
message("The sample size is: ",ceiling(n2))
```

R code Exa 10.6 Hypotheses Concerning One Proportion

```
#a)
#install.packages("binom") #package
library(binom)
binom.test(4, 34, conf.level=0.95, p=.3,alternative = 'less')
cat("The P-value = 0.0117 provides quite strong evidence against the null
hypothesis H0 : p = 0.3 and in favor of the alternative that the probability is
less than 0.3.")
#b)level of significance for the test
#Alternative test P<0.3
pbinom(5, 34,0.3)#P(X<= 5)
```

R code Exa 10.7 Hypotheses Concerning One Proportion

```
#Null hypothesis: p = 0.70
#Alternative hypothesis: p > 0.70
```

```
loc=0.05
z.alpha=qnorm(1-0.05)
z.alpha
#Criterion: Reject the null hypothesis if Z > 1.645, where
x = 48
n = 60
p0 = 0.70
Z=(x - n*p0)/sqrt(n*p0*( 1 - p0))
Z
cat("Since z = 1.69 is greater than 1.645, we reject the null hypothesis at
level 0.05. In other words, there is sufficient evidence to conclude that the
proportion of good transceivers that would be produced is greater than 0.70")
```

R code Exa 10.8 Hypotheses Concerning Several Proportions

```
#equality of three praportion
#null hypothesis:p1=p2=p3 probability of crumbling is the same for all three materials
#alternative hypothesis:p1,p2and p3 or not equal
Crumbled<-c(41,27,22)
Remained_intact<-c(79,53,78)
loc=0.05
n=3
df=n-1
x.chi=qchisq(1-0.05,df)
#reject null hypothesis if X>x.chi
table<-rbind(Crumbled,Remained_intact)
chi_sq=chisq.test(table)
chi_sq
cat("Since chi_sq = 4.575 does not exceed 5.991, the null hypothesis cannot
```

```
be rejected")
```

R code Exa 10.9 Hypotheses Concerning Several Proportions

```
#Null hypothesis: p1 = p2 = p3 = p4 no difference in the proportions of superconductors produced
#Alternative hypothesis: p1, p2, p3, and p4 are not all equal
Superconductors<-c(31,42,22,25)
Failures<-c(19,8,28,25)
loc=0.05
n=4
df=n-1
Chi_squr0.05=qchisq(1-0.05,df)
Chi_squr0.05
#Criterion: Reject the null hypothesis if Chi_squr > 7.815, the value of Chi_squr0.05
#for 4 − 1 = 3 degrees of freedom.
table<-rbind(Superconductors,Failures)
table
chisq.test(table)
cat("Since 19.50 greatly exceeds 7.815, we reject the null hypothesis of
equal proportions at the 5% level of significance")
```

R code Exa 10.10 Hypotheses Concerning Several Proportions

```
#Let p1 be the probability a visitor to the original page purchases and item and let p2
#be the probability for the modified page
#Hypothesis
#Null hypothesis= p1 = p2
#Alternative hypothesis= p1 < p2
```

```
#Level of significance:0.01
#Criterion: Reject the null hypothesis if Z < ??? 2.33
x1 = 77
n1 = 2841
x2 = 107
n2 = 2297
Pbar=(x1+x2)/(n1+n2)
Z=(x1/n1-x2/n2)/sqrt(Pbar*(1-Pbar)*(1/n1+1/n2))
Z
cat("Ther for we reject the null hypothesis")
cat("We conclude that the proportion of purchasers is higher for the modified page than the original page.")
```

R code Exa 10.11 Hypotheses Concerning Several Proportions

```
#95% confidence interval for p1 ??? p2
x1=77
n1=2841
x2=107
n2=2297
P1bar=x1/n1
P2bar=x2/n2
Int1=P1bar-P2bar-1.96*sqrt((P1bar*(1-P1bar)/n1)+(P2bar*(1-P2bar)/n2))
Int2=P1bar-P2bar+1.96*sqrt((P1bar*(1-P1bar)/n1)+(P2bar*(1-P2bar)/n2))
Int1
Int2
message("so the 95% confidence interval is  0.030 < p1 − p2 < 0.009")
```

R code Exa 10.12 Analysis of r x c Tables

```
#expected frequencies
column<-c("Shop1","Shop2","Shop3")
row<-c('Complete','Repair Adjustment','Incomplete')
Shop1<-c(78,56,54)
Shop2<-c(15,30,31)
Shop3<-c(7,14,15)
Total<-c(100,100,100)
Col4<-c(188,76,36)
GT=300#grand total
e11=e12=e13=100*Col4[1]/GT
e21=e22=e23=100*Col4[2]/GT
e31=e32=e33=100*Col4[3]/GT
#Frequency are
e11
e21
e31
```

R code Exa 10.13 Analysis of r x c Tables

```
table<-rbind(c(23,60,29),c(28,79,60),c(9,49,63))
rownames(table) <- c("Poor","Average","Very good")
colnames(table) <- c("Below Average","Average","Above Average")
table
#Null Hypothesis :Performance in training program and sucess in job are independent
#Alternative hypothesis : Performance in training program and sucess in job are dependent
l.O.c=0.01
df=4
chisq0.01=qchisq(1-0.01,df)
Xsq=chisq.test(table)
Xsq
```

```
Xsq$expected
#Criterion:Reject The Null Hypothesis if chisq>
    chisq0.01
#Calculation:
print("since the chisq value exceed the value of
    chisq0.01 there for we reject the null hypothesis
    ")
```

R code Exa 10.15 Goodness of Fit

```
#Null Hypothesis:Random variable has Poisson
    distribution with lambda=4.6
#Alternative Hypothesis:Random variable does not
    have Poisson distribution with lambda=4.6
L.o.c=0.01
n=10
m=4
df=n-m
chsqr0.01=qchisq(1-0.01,6)
chsqr0.01
#Reject Null hypothesis if chisq>chisq0.01 for df=6
#Calculating Chisq :-
observedf<-c(18,47,76,68,74,46,39,15,9,8)#observed
    frequency
expectedf<-c
    (22.4,42.8,65.2,74.8,69.2,52.8,34.8,20,10,8)#
    expected frequency
chisq.test(observedf, p = expectedf/sum(expectedf))
print("Since the value of chisq does't exceed chisq0
    .01 there for the distribution is poisson
    distribution")
```

Chapter 11

REGRESSION ANALYSIS

R code Exa 11.1 The Method of Least Squares

```
#least squares estimates and sum of squares error
   for the cooling rate
x<-c(0,1,2,2,4,4,5,6)
y<-c(25,20,30,40,45,50,60,50)
fit<-lm(y~x)
fit
#the least squares line is
message("y=",fit$coefficients[[1]],"+",fit$
   coefficients[[2]],"x")
anova(fit)
cat("forom Table  the SSE is 270")
```

R code Exa 11.2 The Method of Least Squares

```
#ormal Equation of precious example is,
x<-c(20,60,100,140,180,220,260,300,340,380)#
   AirVelocity(cm/sec)
y<-c
```

```
(0.18,0.37,0.35,0.78,0.56,0.75,1.18,1.36,1.17,1.650) #Evaporation coefficent (mm^2/sec)
table<-data.frame(x,y)
fit<-lm(y~x)
plot(x,y)
abline(fit)
message("regression equation is: ",fit$coefficients[[2]],'x + ',fit$coefficients[[1]])
summary(aov(y~x))
cat("from table SSE is 0.20238")
print("For x = 190, we predict that the evaporation coefficient will be")
Y<-data.frame(predict(fit,data.frame(x=c(190)), interval = "confidence"))
Y$fit
```

R code Exa 11.3 The Method of Least Squares

```
n = 50
x = 88.34
y = 305.58
Sxx = 7239.22
Sxy = 17840.1
Syy = 66976.2
#(a) Find the least squares line for predicting height from width.
beta=Sxy/Sxx
alpha=y-beta*x
cat('height = ',87.88,' + ',2.464,'width')
#(b) Find the least squares line for predicting width from heigh
beta=Sxy/Syy
alpha=88.34 -0.2664*305.58
cat('width = ',6.944,' + ',0.266,'height')
#(c) Make a scatter plot and show both lines.
```

```
Width<-c(77,80,83,73,79,100,93,92,101,87,96,99,94)
Height<-c(290,292,289,284,271,292,308,303,308,315,309,300,305)
plot(Width,Height)
Height=87.88 + 2.464*Width
lines(Width, Height)
Height=-26.11 + 3.759*Width
lines(Width, Height)
```

R code Exa 11.5 Inferences Based on the Least Squares Estimators

```
velocity<-c(20,60,100,140,180,220,260,300,340,380)
Evaporation<-c(0.18,0.37,0.35,0.78,0.56,0.75,1.18,1.36,1.17,1.65)
Dat<-data.frame(velocity,Evaporation)
fit<-lm(Evaporation~velocity)
summary(lm(Evaporation~velocity), data = Dat)
print("the 95% confidence limits are")
predict(fit,data.frame(velocity=c(0.1009737)), interval = "confidence",level = 0.95)
```

R code Exa 11.6 Inferences Based on the Least Squares Estimators

```
velocity<-c(20,60,100,140,180,220,260,300,340,380)
Evaporation<-c(0.18,0.37,0.35,0.78,0.56,0.75,1.18,1.36,1.17,1.65)
#Null hypothesis: beta= 0
#Alternative hypothesis: beta!= 0
loc=0.05
t=qt(1-0.025,8)
```

```
#Criterion: Reject the null hypothesis if t < −2.306 or t > 2.306, where 2.306
#is the value of t0.025
betabar=0.003829
beta=0
Sxx=132000
Se=0.1591
t=(betabar-beta)*sqrt(Sxx)/Se
t
cat("Decision: Since t = 8.744 exceeds 2.306, the null hypothesis must be rejected,
we conclude that there is a relationship between air velocity and the average")
```

R code Exa 11.7 Inferences Based on the Least Squares Estimators

```
#Peak load
y1<-c(8.6,8.9,9.1)
y2<-c(9.0,9.3,9.4)
y3<-c( 9.5,9.8,9.8)
y4<-c(10.2,10.2,10.3)
x<-c(0,3,6,12)#Prestrain
x=c(0,0,0,3,3,3,6,6,6,12,12,12)
y=c(y1,y2,y3,y4)
#a)
summary(lm(y~x))
summary(aov(y~x))
plot(x,y)
message("the P−value for the one−sided test is less than 0.0000 when rounded
down. This is very strong evidence that prestressing results in stronger
material.")
#b)
fit<-lm(y~x)
```

```
#The estimated regression line is
cat("y = ",fit$coefficients[[1]]," + ",fit$
   coefficients[[2]],"x")
x=1# percent prestrain
y=data.frame(predict(fit,data.frame(x=c(1)),interval
    = "confidence"))
message(y$fit," kN")
#The 95% confidence interval becomes
y
message("We are 95% confident that mean strength is
   between 8.860385 and 9.182154 kN for all
alloy sheets that could undergo a prestrain of 1
   percent")
#c)Note:-This is Part (b) of T.B
x=9# percent prestrain
y=data.frame(predict(fit,data.frame(x=c(9)),interval
    = "confidence"))
message(y$fit," kN")
#The 95% confidence interval becomes
y
message("We are 95% confident that mean strength is
   between 9.79 and 10.09 kN for all
alloy sheets that could undergo a prestrain of 9
   percent")

#d)Note:-This is Part (c) of T.B
cat(", there are no outliers in the residuals and
   the assumption of
normal errors appears reasonable. We can rely upon
   the
statistical conclusions above")
```

R code Exa 11.8 Inferences Based on the Least Squares Estimators

```
#95% prediction limit
```

```
#from example 7
y=9.938
t0.025=qt(1-0.025,11)
Se=0.180634
n=12
Sxx=236.250
x0=9
xbar=5.25
Int1=y-t0.025*Se*sqrt(1+1/n+((x0-xbar)/Sxx))
Int2=y+t0.025*Se*sqrt(1+1/n+((x0-xbar)/Sxx))
Int1
Int2
cat("We are 95% confident that the observed value of peak load for this new
sheet of aluminum lies between 9.51 and 10.37 kN")
```

R code Exa 11.9 Inferences Based on the Least Squares Estimators

```
x0 = 18
y1<-c(8.6,8.9,9.1)
y2<-c(9.0,9.3,9.4)
y3<-c( 9.5,9.8,9.8)
y4<-c(10.2,10.2,10.3)
x<-c(0,3,6,12)
y<-c(mean(y1),mean(y2),mean(y3),mean(y4))
predict(lm(y~x),data.frame(x=c(18)),interval = "confidence",level = 0.95)
cat("We are 95% confident that the observed value of peak load for this
new sheet of aluminum lies between 10.43 and 11.51 kN")
```

R code Exa 11.10 Curvilinear Regression

```
x<-c(2,3,6,10,15,20)
y<-c(164.7,156.1,142.5,133.8,114.6,107.1)
#a)plot of logy vs x
#the relation ship between logy and x is exponenetial curve
plot(x,log(y))
#b)Regression Equation
summary(lm(log(y)~ x))
fit=lm(log(y)~ x)
#logarithmic form is:
cat(”log(y) = ”,fit$coefficients[2],’x +’,fit$coefficients[1])
#exponential form is:
cat(”y = ”,exp(5.1257),’*e^(’,fit$coefficients[2],’)’)
#c)Using the logarithmic form which is more convenient, we predict
result<-predict(fit,data.frame(x=c(5)),interval = ”confidence”,level = 0.95)
result
```

R code Exa 11.11 Curvilinear Regression

```
x<-c(0,1,2,3,4,5,6,7,8)#Amount of varnish additive(grams)
y<-c(12,10.5,10,8,7,8,7.5,8.5,9.0)#Drying Time
#a)Scatter plot
plot(x,y)
#b)Fit a second-degree polynomial
fit<-lm(y~x+I(x^2))
fit
#Polynomial equestion is
cat(”y = ”,fit$coefficients[[1]],””,fit$coefficients[[2]],”x + ”,fit$coefficients[[3]],”x^2”)
summary(fit)
```

```
lines(smooth.spline(x,predict(fit)),col='blue',lwd
    =3)
#c)drying time of the varnish when 6.5 grams of the
    additive is being used
predict(fit,data.frame(x=c(6.5)),interval = "
    confidence")
```

R code Exa 11.12 Multiple Regression

```
y<-c
    (41,49,69,65,40,50,58,57,31,36,44,57,19,31,33,43)
x1<-c(1,2,3,4,1,2,3,4,1,2,3,4,1,2,3,4)
x2<-c(5,5,5,5,10,10,10,10,15,15,15,15,20,20,20,20)
fit<-lm(y~x1+x2)
summary(fit)
anova(lm(y~x1+x2))
print("The least squares regression plane is")
cat("Y=",fit$coefficients[[1]],'+',fit$coefficients
    [[2]],'x',fit$coefficients[[3]],'x^2')
cat("This equation estimates that the average number
     of twists required to break a
bar increases by 7.78 if the percent of element A is
     increased by 1% and x2
remains fixed.")

#For x1=2.5 and x2=12
data<-data.frame(predict(fit,data.frame(x1=c(2.5),x2
    =c(12)),interval ="confidence"))
data$fit
```

R code Exa 11.14 Correlation

```
#Time taken by 10 mechine to assemble in morning x and afternoon y
x<-c(11.1,10.3,12.0,15.1,13.7,18.5,17.3,14.2,14.8,15.3)
y<-c(10.9,14.2,13.8,21.5,13.2,21.1,16.4,19.3,17.4,19.0)
plot(x,y,xlim=c(0,20),ylim=c(0,25))
r=cor(x,y)
r
```

R code Exa 11.15 Correlation

```
influent<-c(250,290,270,100,300,410,110,130,1100)
effluent<-c(19,10,17,11,70,60,18,30,180)
#(a)make scatter plot
plot(influent,effluent)
#(b)log scatter plot
plot(log(influent),log(effluent))
#(c)
r1=cor(influent,effluent)
r1
r2=cor((influent),log(effluent))
r2
#d)
cat("hence r is not really appropriate for the original data ")
```

R code Exa 11.17 Correlation

```
#Null hypothesis=delta=0
#Alternative hypothesis= delta!=0
```

```
l.o.c=0.05
n=10
r=0.732
#reject null hypothesis if Z>1.96 or z<-1.96
#Z=sqrt(n-1)*Zbar
Zbar=1/2*(log((1+r)/(1-r)))
z=sqrt(n-3)*Zbar
Zbar
z
print("There for we reject the null hypothesis")
```

R code Exa 11.18 Correlation

```
#Cinfidence interval for delta normal population
r=0.70
n=30#number of student
z0.025=1.96
Zbar=1/2*log((1+r)/(1-r))
Zbar
Int1=Zbar-z0.025/sqrt(n)
Int2=Zbar+z0.025/sqrt(n)
message("the interval is: ",Int1 ," and ",Int2)
r1=(exp(1)^Int1-exp(1)^-Int1)/(exp(1)^Int1+exp(1)^-Int1)
r1
r2=(exp(1)^Int2-exp(1)^-Int2)/(exp(1)^Int2+exp(1)^-Int2)
r2
#there for the 95% confidence interval is
message(r1," < rho < ",r2)
```

R code Exa 11.19 Multiple Linear Regression Matrix Notation

```
#Multiple Linear Regrassion
sumx1=40
sumx2=200
sumx1squar=120
sumx2squar=3000
sumy=723
x1.y=1963
x2.y=8210
x1.x2=500
n=16
XbarX=matrix(c(n,sumx1,sumx2,sumx1,sumx1squar,x1.x2,
    sumx2,x1.x2,sumx2squar),ncol = 3,byrow = TRUE)
XbarX
Inverse=solve(XbarX)
Xbary=matrix(c(sumy,x1.y,x2.y))
Xbary
b=Inverse%*%Xbary
print("The least square estimation is ")
b
```

R code Exa 11.20 Multiple Linear Regression Matrix Notation

```
X<-matrix(c(0,1,2,3,4))
X
y<-matrix(c(8,9,4,3,1))
n=5
XbarX<-matrix(c(n,sum(X),sum(X),sum(X^2)),ncol = 2,
    byrow = TRUE)
XbarX
Inverse=solve(XbarX)
Xbary<-matrix(c(sum(y),sum(X*y)))
Xbary
b=Inverse%*%Xbary
b
message("The Fitted equestion is :  y = ",b[1],"",b
```

```
    [2],"x ")
XX=matrix(c(1,0,1,1,1,2,1,3,1,4),byrow =TRUE,ncol
    =2)
XX
ybar=XX%*%b
ybar
#So the vector of residuals
result=y-ybar
#The Residual sum of squares is
result2=t(result)
data=result2%*%result
Ssquar=data/(n-1-1)
message(Ssquar)
```

R code Exa 11.21 Multiple Linear Regression Matrix Notation

```
#Calculate estimated variances var(b0) and var(b1)
#With refering to previous example
Inverse=matrix(c(0.6,-0.2,-0.2,0.1),ncol =2,byrow =
    TRUE)
Ssquar=2
var=Inverse*Ssquar
print("The  estimated variances is")
var
varb0=var[1,1]
varb1=var[2,2]
print("The var(b0) is")
varb0
print("The var(b1) is")
varb1
```

Chapter 12

ANALYSIS OF VARIANCE

R code Exa 12.1 Completely Randomized Designs

```
#H0=the laboratories obtaining consistent result
#tin-coating weight for 12 disc in each liboratory is show as
LabA<-c(0.25,0.27,0.22,0.30,0.27,0.28,0.32,0.24,0.31,0.26,0.22,0.28) #Laboratory A Data of Tin-Coting weight
LabB<-c(0.18,0.28,0.21,0.23,0.25,0.20,0.27,0.19,0.24,0.22,0.29,0.16) #Laboratory B Data of Tin-Coting weight
LabC<-c(0.19,0.25,0.27,0.24,0.18,0.26,0.28,0.24,0.25,0.20,0.21,0.19) #Laboratory C Data of Tin-Coting weight
LabD<-c(0.23,0.30,0.28,0.28,0.24,0.34,0.20,0.18,0.24,0.28,0.22,0.21) #Laboratory d Data of Tin-Coting weight
#Null hypothesis: mu1 = mu2 = mu3 = mu4 #Data produce dy 4 diff laboratory has  same mean
#Alternative hypothesis: The mu's are not all equal
#mu=mean
#alpha=0.05
F0.05=qf(1-0.05,3,44)
```

```
cat("The F0.05 value is:",F0.05)
weights = c(LabA,LabB,LabC,LabD)#Single vactor of weights
Laboratory = rep(1:4,rep(12, 4))#Making the group in the single vextor
data = data.frame(weight =weights,Laboratory = factor(Laboratory))
fit = lm(weight ~ Laboratory , data)
anova(fit)
cat("Since the observed value of F exceeds value of F0.05, the null
hypothesis of equal mean weights is rejected at the 0.05 level of significance
We conclude that the laboratories are not obtaining consistent results")
```

R code Exa 12.2 Completely Randomized Designs

```
#Estimating The Parameter of onw-way classification
#tin-coating weight for 12 disc in each liboratory is show as
LabA<-c(0.25,0.27,0.22,0.30,0.27,0.28,0.32,0.24,0.31,0.26,0.22,0.28)#Laboratory A Data of Tin-Coting weight
LabB<-c(0.18,0.28,0.21,0.23,0.25,0.20,0.27,0.19,0.24,0.22,0.29,0.16)#Laboratory B Data of Tin-Coting weight
LabC<-c(0.19,0.25,0.27,0.24,0.18,0.26,0.28,0.24,0.25,0.20,0.21,0.19)#Laboratory C Data of Tin-Coting weight
LabD<-c(0.23,0.30,0.28,0.28,0.24,0.34,0.20,0.18,0.24,0.28,0.22,0.21)#Laboratory d Data of Tin-Coting weight
cbind(LabA,LabB,LabC,LabD)#Table View of data
k=4#Total laboratory
```

```
n=12#total no of data in each laboratory
mu=mean(c(LabA,LabB,LabC,LabD))
sigma1=(sum(LabA)/n)-mu
sigma2=(sum(LabB)/n)-mu
sigma3=(sum(LabC)/n)-mu
sigma4=(sum(LabD)/n)-mu
print("There for Parameters of one way
    classification are:")
message("mu:",mu,"sigma1:",sigma1,"sigma2:",sigma2,"
    sigma3:",sigma3,"sigma4:",sigma4)
```

R code Exa 12.3 Completely Randomized Designs

```
#EXAMPLE 3
#Confidence intervals
#Given
MSE=0.0234
dfE=12
n=5
MDmu=1.334
Edmu=0.964
PFmu=0.776
x=#The Three confidence intervals become
t0.025=-1*qt(0.025,dfE)
Interval<- function(y1,y2,n1,n2){
  int1=y1-y2+(t0.025*sqrt((MSE)*((1/n1)+(1/n2))))
  int2=y1-y2-t0.025*sqrt(MSE*((1/n1)+(1/n2)))
  interval<-c(int2," ",int1)
  return (interval)
}
MDED=Interval(MDmu,Edmu,n,n)
MDPF=Interval(MDmu,PFmu,n,n)
EDPF=Interval(Edmu,PFmu,n,n)
message(MDED)
message(MDPF)
```

```
message(EDPF)
message("The resin MD has a higher bond strength than the other two, which cannot be distinguished")
```

R code Exa 12.4 Completely Randomized Designs

```
#EXAMPLE 4
#There are 4 different types of paper,
#We measuring the strength of paper by repeating the process
Paper1<-c(2.8,0.75,3.70)
Paper2<-c(0.00,-0.1,3.45)
Paper3<-c(1.15,1.75,4.20)
Paper4<-c(1.88,2.65,2.70)
#Mean of different observation
x1=mean(Paper1)
x2=mean(Paper2)
x3=mean(Paper3)
x4=mean(Paper4)
#Null Hypothesis:-alpha1=alpha2=aplha3=alpha4=0 , There is no differnce in two method of testing paper strength
#Alternative Hypothesis:-alpha's not equal to zero
l.o.c=0.5
pf(1-0.5,3,8)
strength<-c(Paper1,Paper2,Paper3,Paper4)
group = rep(1:4,rep(3, 4))
data = data.frame(y = strength, group = factor(group))
fit = lm(y ~ group, data)
anova(fit)
print("Since F doesn't exceed the value of 1 there for we doesn't reject null hypothesis")
mean=mean(strength)
```

```
sd=sd(strength)
df3=length(strength)-1
t0.025=qt(1-0.025,11)
#95% confidence interval for mean is
Int1=mean+t0.025*sd/sqrt(12)
Int2=mean-t0.025*sd/sqrt(12)
cat("The Confidence Interval For Mean is:",c(Int2,Int1))
```

R code Exa 12.5 Completely Randomized Designs

```
#H0=the laboratories obtaining consistent result
#tin-coating weight for 12 disc in each liboratory is show as
LabA<-c(0.25,0.27,0.22,0.30,0.27,0.28,0.32,0.24,0.31,0.26,0.21,0.28) #Laboratory A Data of Tin-Coting weight
LabB<-c(0.18,0.28,0.21,0.23,0.25,0.20,0.27,0.19,0.24,0.22,0.29,0.16) #Laboratory B Data of Tin-Coting weight
LabC<-c(0.19,0.25,0.27,0.24,0.18,0.26,0.28,0.24,0.25,0.20,0.21,0.19) #Laboratory C Data of Tin-Coting weight
LabD<-c(0.23,0.30,0.28,0.28,0.24,0.34,0.20,0.18,0.24,0.28,0.22,0.21) #Laboratory d Data of Tin-Coting weight
cbind(LabA,LabB,LabC,LabD)#Table View of data
k=4
n=12
T. = 11.70#grand total
N=48
SumA=sum(LabA)
SumB=sum(LabB)
SumC=sum(LabC)
SumD=sum(LabD)
```

```
C=T.^2/N
SST=sum(LabA^2)+sum(LabB^2)+sum(LabC^2)+sum(LabD^2)-
    C
SSTr=0.0135
SSE=SST-SSTr
MSTr=SSTr/(k-1)
MSE=0.0015
F=MSTr/MSE
F
```

R code Exa 12.6 Randomized Block Designs

```
Source<-c('Treatments','Blocks','Error')
Treatment1<-c(13,8,9,6)
Treatment2<-c(7,3,6,4)
Treatment3<-c(13,7,12,8)
Temp<-c(Treatment1,Treatment2,Treatment3)
f = c("Item1", "Item2", "Item3", "Item4")
k = 4 # number of treatment levels
n = 3 # number of control blocks
tm = gl(k, 1, n*k, factor(f))   # matching treatment
tm
blk = gl(n, k, k*n)             # blocking factor
blk
av = aov(Temp ~blk+ tm)
summary(av)
F0.05=qf(1-0.05,2,6)
cat("The value of F0.05 with 2 and 6 degrees of
    freedom is 5.14, so we reject the null
hypothesis of equal mean particulate material
    removal. Blocking was important because
     we also reject the null hypothesis of equal
        block means")
```

R code Exa 12.7 Randomized Block Designs

```
#There are 4 different type of Detergentfor engen.
#There are 3 engine
#Null hypothesis:alpha1=alpha2=apha3=0 and beta1=beta2=beta3=0
#Alternative hypothesis:The alphas and betas not equal to zero
l.o.c=0.01
#Criteria:For treatment ,reject null hypothesis if F>9.78,the value of F0.01 with df is 3 and 6
#for block reject null hypothesis if F>10.92  for 2 and 6 degrees of freedom
#Calculation
Detergent_A<-c(45,43,51)
Detergent_B<-c(47,46,52)
Detergent_C<-c(48,50,55)
Detergent_D<-c(42,37,49)
data<-c(Detergent_A,Detergent_B,Detergent_C,Detergent_D)
f = c("Item1", "Item2", "Item3")
k = 3 # number of treatment levels
n = 4 # number of control blocks
Engines= gl(k, 1, n*k, factor(f))# matching treatment
Detergents= gl(n, k, k*n) # blocking factor
av = aov(data ~Detergents+ Engines)
summary(av)
print("Sice FDet exceed the value 9.78 there we Conclude that there is difference in effectivness in Detergent")
print("Since FEn exceed the value 10.92 there for we conclude that there is difference in result obtain by 3 machine ")
```

R code Exa 12.8 Randomized Block Designs

```
#Data from Example 7
Tem=c(45,43,51,47,46,52,48,50,55,42,37,49)
data<-matrix(c
   (45,43,51,139,47,46,52,145,48,50,55,153,42,37,49,128,182,176,207,
   ,ncol=4,byrow=TRUE)
data
rownames(data)<-c("Detergent A","Detergent B","
   Detergent C","Detergent D","Totals")
colnames(data)<-c("Engine 1","Engine 2","Engine 3","
   Totals")
data
a = 4
b = 3
T1. = 139
T2. = 145
T3. = 153
T4. = 128
T.1 = 182
T.2 = 176
T.3 = 207
T.. = 565
Sumy=26867
n=12
C=T..^2/n#factor
C
SST=sum(Tem^2)-C
SST
SSTr=((139^2+145^2+153^2+128^2)/3)-C
SSTr
SSB=((182^2+176^2+207^2)/4)-C
SSB
SSE=SST-SSTr-SSB
```

```
SSE
```

R code Exa 12.9 Multiple Comparisons

```
#Find 94% Bonferroni SimultaneousConfidence interval
MSE=0.0237
alpha=0.06
k=3
l.o.c=alpha/k*(k-1)
df=12
t0.01=qt(1-0.01,df=12)
t0.01
MDmu=1.334
Edmu=0.964
PFmu=0.776
n=5
Interval<- function(y1,y2,n1,n2){
  int1=y1-y2+t0.01*sqrt(MSE*((1/n1)+(1/n2)))
  int2=y1-y2-t0.01*sqrt(MSE*((1/n1)+(1/n2)))
  interval<-c(int2," ",int1)
  return (interval)
}
MDED=Interval(MDmu,Edmu,n,n)
MDPF=Interval(MDmu,PFmu,n,n)
EDED=Interval(Edmu,PFmu,n,n)
#There for the interval are as
message(MDED)
message(MDPF)
message(EDED)
```

R code Exa 12.10 Multiple Comparisons

```
#Find 94% Bonferroni SimultaneousConfidence interval
```

```
ED=c(0.99,1.19,0.79,0.95,0.90)
MD=c(1.11,1.53,1.37,1.24,1.42)
PF=c(0.83,0.68,0.94,0.86,0.57)
Strength<-c(ED,MD,PF)
Resin<-rep(1:3,rep(5,3))
Dat = data.frame(Strength=Strength, Resin = factor(
   Resin))
summary(fm1<-aov(lm(Strength~Resin, data=Dat)))
TukeyHSD(fm1,'Resin', conf.level=0.94,ordered = TRUE
   )
```

R code Exa 12.11 Analysis of Covariance

```
Tr<-c(1,1,1,1,2,2,2,2,3,3,3,3)
x<-c
   (0.90,0.95,1.05,0.80,0.50,0.40,0.15,0.25,0.20,0.55,0.30,0.40)
   #Original reflectivity
y<-c
   (1.05,0.95,1.15,0.85,1.10,1.00,0.90,0.80,0.75,1.05,0.95,0.90)
   #Final reflectivity
#Null hypothesis: aplha1 = alpha2 = alpha3 = 0
#Alternative hypothesis: The alpha's are not all
   equal to zero.
loc=0.05
F0.05=qf(1-0.05,2,8)
#criterion: Reject the null hypothesis if F > 4.46,
   the value of F0.05 for
#k - 1 = 3 - 1 = 2 and nk - k-1 = 4   3 -3 - 1 = 8
   degrees of freedom
Tr<-c(rep('A',4),rep('B',4),rep('C',4))
fit<-lm(y~x+factor(Tr))
anova(fit)
cat("Since F value of Treatments is 6.48 there for
   we reject the Null Hypothesis")
```

Chapter 13

Factorial Experimentation

R code Exa 13.1 Two Factor Experiments

```
#Null hypotheses: alpha1 = alpha2 = alpha3 = 0; beta1 = beta2 = 0
#( alphabeta ) terms are all equal to zero
#Alternative hypotheses: The alpha's are not all equal to zero; the beta's are not all
#equal to zero; the (alphabeta) terms are not all equal to zero
F0.01=qf(1-0.01,2,12)#df=2,12
F0.01
F0.01=qf(1-0.01,1,12)#df=1,12
F0.01
#Criteria:a)For replications reject the null hypotheses if F > 6.93, the value of
#F0.01 for 2,12 df.
#b)for the factor A, reject the null hypothesis if F > 9.33, the value of F0.01 for df=1,12
#c)for factor B, reject if F > 9.33, the value of F0.01 for df= 1,12
#d)for the interaction effect, reject if F > 6.93, the value of F0.01 for df=2,12
rep_1<-c(707,652,522,630,450,845)
```

```
rep_2<-c(632,669,554,648,545,810)
rep_3<-c(604,674,484,610,474,682)
r<-c(707,632,604,652,669,674,522,554,484,630,648,610,450,545,474,845,
A<-c(rep('MN',6),rep('CO',6),rep('TX',6))#Location
B<-c(rep('RCA',3),rep('RPA',3),rep('RCA',3),rep('RPA',3),rep('RCA',3),rep('RPA',3))#Type of Mat.
Dat<-data.frame(A,B,r)
av = lm(r~A*B,data=Dat)
anova(av)
cat('The F = 12.0 for Factor B exceeds F0.01 = 9.33 for 1 and 12 degrees of freedom and that
F = 36.1 for the AB interaction term exceeds F0.01 = 6.93 = for 2 and')
```

R code Exa 13.2 Multifactor Experiments

```
#Factor Level
#From previous example variance is Ssquare=0.25
Ssquare=2394
t0.025=qt(1-0.025,12)
t0.025
#There for the confidence intervals for difference in mean due to the a=3 level of oven width,Factor are
ybar1=656.3
ybar2=574.7
ybar3=634.3
b=2
r=3
levelss<-function(x,y,a){
  Int1=x-y+(t0.025*sqrt((Ssquare*2)/a*r))
  Int2=x-y-t0.025*sqrt((Ssquare*2)/a*r)
  return(c(Int1,Int2))
```

```
}
Level1<-levelss(ybar1,ybar2,b)
Level1
Level2<-levelss(ybar1,ybar3,b)
Level2
Level3<-levelss(ybar2,ybar3,b)
Level3
#Concluesion:
message("Because the interation was significant, we cannot interpret on differences of mean coking time as due to changing over wIdth along")
#Signgle diffeence in mean due to the b=2 flue tem is
mean=levelss(552.4,691.1,3)
mean
```

R code Exa 13.3 Multifactor Experiments

```
#To improve quality in production There are three thing to study that are
# 3 initiators (A), 2 booster charges (B), and 4 main charges(C)
A<-c(rep('Initiator 1',8),rep('Initiator 2',8),rep('Initiator 3',8))
B<-c('Powder','Pellet','Powder','Pellet','Powder','Pellet','Powder','Pellet','Powder','Pellet','Powder','Pellet','Powder','Pellet','Powder','Pellet','Powder','Pellet','Powder','Pellet','Powder','Pellet','Powder','Pellet')
C<-c('MC1','MC1','MC2','MC2','MC3','MC3','MC4','MC4','MC1','MC1','MC2','MC2','MC3','MC3','MC4','MC4','MC1','MC1','MC2','MC2','MC3','MC3','MC4','MC4')
DelayTime<-c(10.7,9.82,10.02,13.66,14.46,20.86,11.44,13.76,15.04,
```

```
16.02,27.26,21.42,20.82,14.46,24.56,36.48,18.42,18.62,

22.80,25.14,33.40,20.62,31.86,19.78,22.94,31.12,32.92,21.38,

27.92,59.86,31.94,28.32,7.14,7.98,24.32,10.26,
8.30,7.86,7.00,8.40,8.40,10.94,17.82,15.28,9.56,19.04,19.98,18.46)

TempC<-c(rep('MC1',4),rep('MC2',4),rep('MC3',4),rep(
    'MC4',4))
C<-c(rep(TempC,3))
TempB<-c(rep('Powder',2),rep('Pellet',2))
B<-c(rep(TempB,12))
length(B)
A<-c(rep('Initiator 1',16),rep('Initiator 2',16),rep
    ('Initiator 3',16))
length(DelayTime)
Dat<-data.frame(A,B,C,DelayTime)
fit<-lm(DelayTime~A*B*C,data=Dat)
#The Anova table is
anova(fit)
#the initiators and main charges are significant as,
with(Dat, tapply(DelayTime, list(A), mean))
with(Dat, tapply(DelayTime, list(C), mean))
```

R code Exa 13.4 Response Surface Analysis

```
#A compound is produced for a coating process.The
    Two Facctor Add i.e Addactivity and Tem to get
    best yield
Run<-c(1,2,3,4,5,6,7,8,9)
Addactive<-c(0,70,35,0,70,70,0,35,35)
Tem<-c(100,100,140,180,180,140,140,100,180)
yield<-c(81,65,92,50,75,75,68,90,77)
Dat<-data.frame(cons,Addactive,Tem,yield)
Dat
```

```
fit=lm(yield~Addactive+Tem+I(Addactive^2)+I(Tem^2)+Addactive*Tem,data=Dat)
fit
#Estimated Regression Coefficients for Yield
summary(fit)
# from fit estimated response surface is
cat("y =  60.2639 + 0.0417*x1 + 0.5354*x2 - 0.0141*x1^2 - 0.0033*x2^2 + 0.0073*x^1*x^2")
x1 = 33
x2 = 117
y =  60.2639 + 0.0417*x1 + 0.5354*x2 - 0.0141*x1^2 - 0.0033*x2^2 + 0.0073*x^1*x^2
y
cat("It is usually not reasonable to drop a linear term when the associated square term is
in the response surface model")
```

Chapter 14

NONPARAMETRIC TESTS

R code Exa 14.1 The Sign Test

```r
surment<-c
   (99.0,102.3,99.8,100.5,99.7,96.2,99.1,102.5,103.3,97.4,100.4,98.9
mu=98.0
#install.packages(BSDA)
library(BSDA)
#m0=98.0(p=1/2)
#m1>98.0(p>1/2)
SIGN.test(surment,NULL,alternative = 'g',mu)
print("the probability of X >= 12 is 0.0065")
cat("Since 0.0065 is less than 0.01, the null
   hypothesis must be rejected;
we conclude that the median octane rating of the
   given kind of gasoline
exceeds 98.0")
```

R code Exa 14.2 The Sign Test

```
#Following is the sverage weekly losses of work—
    hours due to accident in 10 industrial plants
    before and after
#safety program was put into operation
Before<-c(45,73,46,124,33,57,83,34,26,17)
After<-c(36,60,44,119,35,51,77,29,24,11)
l.O.c=0.05
data=Before-After
data
library(BSDA)
SIGN.test(data)
#Null Hypothesis:mu=0
#Alternative hypothesis:mu>0
#Criterion:reject null hypothesis if Probability of
    x is less then 0.05
sign=c()#This Store the sign for the data
for (i in 1:length(data)){
  if(data[i]>0){
    sign[i]='+'
  }else{
    sign[i]='—'
  }
}
message(sign)
X=length(sign[ sign %in% '+'])#x for no of  pluse
    sign
p=0.5#Probability
#P(X>9)
PX=1-pbinom(X-1,10,p)
message("P(X>9) = ",PX)
message("Since PX is less then 0.05 there for we
    reject the Null Hypothesis")
print("There for the safety program is effective")
```

R code Exa 14.3 Rank Sum Tests

```
#m0=Population are identical
#m1=population is not identical
Sand1<-c(0.63,0.17,0.35,0.49,0.18,0.43,0.12,0.20,
0.47,1.36,0.51,0.45,0.84,0.32,0.40)
Sand2<-c(1.13,0.54,0.96,0.26,0.39,0.88,0.92,0.53,
1.01,0.48,0.89,1.07,1.11,0.58)
Loc=0.01
z.alpha=qnorm(0.01)
#reject mo if Z<z.alpha or z>z.alpha
n1=15
n2=14
W1=162
U1=W1-((n1*(n1+1))/2)
U1
mu=n1*n2/2
mu
sigma=n2*n1*(n1+n2+1)/12
sigma
z=(U1-mu)/(sigma**0.5)
z
wilcox.test(Sand1,Sand2,exact = FALSE, correct =
    FALSE)
cat("Since z = ???2.75 is less than ???2.575, the
    null hypothesis must
be rejected at ?? = 0.01. The P-value = 0.0060 and
    we conclude that there is a
strong evidence of difference in the populations of
    grain size")
```

R code Exa 14.4 Rank Sum Tests

```
method_A<-c(54,66,67,71,74,77)
method_B<-c(41,52,59,60,62,64,65)
method_C<-c(47,49,52,56,69)
#Null Hypothesis=Population are identical
```

```
#Alternative Hypothesis=Population are not identical
Loc=0.05
H.chi=qchisq(1-0.05,2)
H.chi
#reject mo if Chi>h.chi
dati = list(g1=method_A, g2=method_B, g3=method_C)#
    getting the list of different method
dati
#Chi square using Kruskal test
kruskal.test(dati)
print("since Chi exceed H.chi there for m0 is
    rejected")
print("there for preventative method against
    corrosion are not equally effective")
```

R code Exa 14.5 Correlation Based on Ranks

```
#losses of worker-hours before and after safety
    programs in 10 industrial plants
Before<-c(45,73,46,12,43,57,83,34,26,17)
After<-c(36,60,44,19,35,51,77,29,24,11)
data<-cbind(Before,After)
#Ranks of the data
Ri<-rank(Before)
Si<-rank(After)
n=10
#Calculate the Rs(rho) value with text case
cor.test(Ri,Si,method=c("spearman"))
cat('This large positive value indicates strong
    association along an increasing curve')
```

R code Exa 14.6 Tests of Randomness

```
#randomness
#install.packages(tseries)
#library(tseries)
#Null hypothesis: Arrangement is random.
#Alternative hypothesis: Arrangement is not random
z=alpha=qnorm(1-0.01)
#Criterion: Reject the null hypothesis if Z < −2.32 or Z > 2.32
arr<-as.factor(c('n','n','n','n','n','d','d','d','d','n','n','n','n','n','n','n','n'
        ,'n','n','d','d','n','n','d','d','d','d'))
runs.test(arr)
cat("Since z = −3.20 is less than −2.32, the null hypothesis must be
rejected. We conclude that the arrangement is not random")
```

R code Exa 14.7 Tests of Randomness

```
#install.package(tseries)
library(tseries)
lathdata<-c(0.261,0.258,0.249,0.251,0.247,0.256,0.250,0.247,0.255,0.243,
            0.252,0.250,0.253,0.247,0.251,0.243,0.258,0.251,0.245,0.2
            0.248,0.252,0.254,0.250,0.247,0.253,0.251,0.246,0.249,0.2
            0.247,0.250,0.253,0.247,0.249,0.253,0.246,0.251,0.249,0.2
LOC=0.01#level of significance
#mo=Arragment of sample data is random(null hypothesis)
#m1=There is frequently alternating pattern(alternative hypothesis)
```

```
z.alpha=qnorm(1-0.01)
#reject mo if z>z.alpha
data<-c()
mu=0.25
lathdata=lathdata[!lathdata %in% mu]
for(i in 1:length(lathdata)){
  if(lathdata[i]>mu){
    data[i]='a'}
  if(lathdata[i]<mu){
    data[i]='b'}
}
runs.test(factor(data),alternative = "g")
cat(" Since z = 2.98 exceeds 2.33, the null
   hypothesis of randomness must
be rejected")
```

R code Exa 14.8 The Kolmogorov Smirnov and Anderson Darling Tests

```
#NUll hypothesis :The pinholes are uniformly
   distributed across the tin plate
#Alternative hypothesis :The pinholes are not
   uniformly distributed across the tin plate
Data<-c
   (4.8,14.8,28.2,23.1,4.4,28.7,19.5,2.4,25.0,6.2)#
   distance of  10 pinhole
y<-Data/30#Function F(X)=DATA/30
L.o.c=0.05
#Criterion : Reject the null hypothesis if D is
   large. where Dnis maximum difference between the
   empirical cumulative distribution
#From figure we have greatest value x=6.2
ks.test(y,"punif")# Kolmogorov-Smirnov test
message("There for we doesn't reject the null
   hypothesis")
```

R code Exa 14.9 The Kolmogorov Smirnov and Anderson Darling Tests

```
#install.packages("nortest")
library(nortest)
#distances in inches of 10 pinholes from one edge of
   a long strip
data<-c
  (2.4,4.4,4.8,6.2,14.8,19.5,23.1,25.0,28.2,28.7)
ad.test(data)
print(" there for According to the large sample
  critical value, we fail to reject the null
  hypothesis ")
```

Chapter 15

The Statistical Content of Quality Improvement Programs

R code Exa 15.2 Experimental Designs for Quality

```
LSL = 10.98
USL = 11.01
n=80
x = 10.991
s = 0.0035
#Estimating the process capability indices Cp and
    Cpk
Cp=(USL-LSL)/(6*s)
Cp
Cpk=min(x-LSL,USL-x)/(3*s)
Cpk
cat("This second index is substantially smaller than
    the first because the mean is
off-center.")
```

R code Exa 15.3 Tolerance Limits

```
n = 100
x = 1.507
s = 0.004
P=0.95#praportion
K = 2.335#With n=100 sample size and p=0.95
#The resulting tolerance interval is
Int1=x-K*s
Int2=x+K*s
L=c(Int1,Int2)
message("We assert, with 99% confidence, that at
    least 95 % of springs have free lengths from
1.497 to 1.517 inches")
```

R code Exa 15.4 Tolerance Limits

```
weight<-c(210,234,216,232,262,183,227,197,
248,218,256,218,244,259,263,185,
218,196,235,223,212,237,275,240,
217,263,240,247,253,269,231,254,
248,261,268,262,247,292,238,215)
n=40#Total Box
xbar=mean(weight)
s=sd(weight)
K = 2.125#for95% of confidence and n=40
L=xbar-K*s
L
cat("there for We are 95% confident that at least a
    proportion 0.95 of the population of burst
strengths, for cardboard boxes, is above 183 psi")
```

Chapter 16

Application to Reliability and Life Testing

R code Exa 16.1 Reliability

```
#A system consists of 5 independent components in
   series , each having a reliability of 0.970
n=5#No of Component
R=0.97# reliability of component
#The reliability of System wIth n component
message("System reliability is : ",R^5)
#Increasing system complexity to 10 components will
   decrease the system reliability to
message("System reliability of 10 component is",R
   ^10)
print("There for to make reliability equal we have
   to take reliability of components is 0.985")
```

R code Exa 16.2 Reliability

```
#Using The Diagram find the reliability
```

```
#reliability of parallel assembly
#1)reliability of component C, D, E as Cbar
Cbar=1 - ( 1 - 0.70 )^3
Cbar
#1)reliability of component C, D, E as Fbar
Fbar= 1 - ( 1 - 0.75 )^2
Fbar
#reliability of Series assembly
result=(0.95)*(0.99)*(0.973)*(0.9375)*(0.90)
cat("The Resultent  reliability  is : ",result)
```

R code Exa 16.3 The Exponential Model in Life Testing

```
n=50#Suppose that 50 units are placed on life test (
   without replacement)
r=10#failure
Time=c(65, 110, 380, 420, 505, 580, 650, 840,
   910,950)#Failure Time in Hourse
#Accumulated life to r failures
T10=sum(Time)+(n-r)*950
#mean life of the component as
mu=T10/r
#failure rate is
alpha=1/mu#failure per hour
Chi0.05=qchisq(1-0.05,20)
Chi0.95=qchisq(1-0.95,20)
#confidence interval for mean life is
Int1=(2*T10)/Chi0.05
Int2=(2*T10)/Chi0.95
cat("The mean life interval is : ",Int1," < ",Int2)
```

R code Exa 16.4 The Exponential Model in Life Testing

```
#Testing hypotheses concerning mean life for preceding example
T10=43410
alpha=0.4#failure rate is 0.40 failure per thousand hours
mu0=1000/alpha#1000 for thousand hourse
Chi0.05=qchisq(1-0.05,20)#degree of fredom =20
#Null hypothesis: mu = mu0
#Alternative hypothesis: mu > mu0 hours
#Level of significance: = 0.05
#Criterion: Reject the null hypothesis if f Tr >(mu0*Chi0.05)/2
#Calculating the value
result=(mu0*Chi0.05)/2
result
cat("Since T10 = 43,410 exceeds the critical value, we must reject the
null hypothesis, concluding that the mean lifetime exceeds 2,500 hours, or,
equivalently, that the failure rate is less than 0.40 failure per thousand hours")
```