R Textbook Companion for Introductory Business Statistics by Alexander Holmes, Barbara Illowsky, Susan Dean[1]

Created by
Radha R
B.Sc.
Computer Science and Engineering
Shrimathi Devkunvar Nanalal Bhatt Vaishnav College for Women, Chennai
Cross-Checked by
R TBC Team

April 28, 2022

[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: Introductory Business Statistics

Author: Alexander Holmes, Barbara Illowsky, Susan Dean

Publisher: OpenStax

Edition: 2

Year: 2018

ISBN: 1-947172-47-6

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

Sampling and Data

R code Exa 1.14 Finding percentage of heights

```
#page 24
library(dplyr)
s_h<-data.frame(
                lbound=c(59.95,61.95,63.95,65.95,67.95,69.95,71.95,73.95),
                ubound=c(61.95,63.95,65.95,67.95,69.95,71.95,73.95,75.95),
                freq=c(5,3,15,40,17,12,7,1))
View(s_h)
total=sum(s_h$freq)
print(total)
s_h$r_freq=s_h$freq/total
View(s_h)
s_h$cum_freq=cumsum(s_h$r_freq)
percent=s_h %>%
  select(lbound,ubound,freq)%>%
  filter(ubound<=65.95)
p_6595=(sum(percent$freq)/total)*100
print(paste("the solution is =",p_6595,"%"))
```

R code Exa 1.15 Finding percentage of heights within a range

```
#page 25
library(dplyr)
s_h<-data.frame(
                lbound=c(59.95,61.95,63.95,65.95,67.95,69.95,71.95,73.95),
                ubound=c(61.95,63.95,65.95,67.95,69.95,71.95,73.95,75.95),
                freq=c(5,3,15,40,17,12,7,1))
View(s_h)
total=sum(s_h$freq)
print(total)
s_h$r_freq=s_h$freq/total
View(s_h)
s_h$cum_freq=cumsum(s_h$r_freq)
percent=s_h %>%
  select(lbound,ubound,freq,r_freq)%>%
  filter(lbound>=61.95 & ubound<=65.95)
p_6165=sum(percent$r_freq)
print(paste("the solution is =",p_6165))
```

R code Exa 1.16 Finding percentage of heights within a range

```
#page 25-26
library(dplyr)
s_h<-data.frame(
                lbound=c(59.95,61.95,63.95,65.95,67.95,69.95,71.95,73.95),
```

```
                ubound=c
                   (61.95,63.95,65.95,67.95,69.95,71.95,73.95,75.95)
                   ,
                freq=c(5,3,15,40,17,12,7,1))
View(s_h)
total=sum(s_h$freq)
print(total)
s_h$r_freq=s_h$freq/total
View(s_h)
s_h$cum_freq=cumsum(s_h$r_freq)
percent=s_h %>%
  dplyr::select(lbound,ubound,freq,r_freq)%>%
  filter(lbound>=67.95 & ubound<=71.95)
p_6771=(sum(percent$r_freq))*100
print(paste("solution of a is =",p_6771,"%"))
percent=s_h %>%
  dplyr::select(lbound,ubound,freq,r_freq)%>%
  filter(lbound>=67.95 & ubound<=73.95)
p_6773=(sum(percent$r_freq))*100
print(paste("solution of b is =",p_6773,"%"))
percent=s_h %>%
  dplyr::select(lbound,ubound,freq,r_freq)%>%
  filter(lbound>=65.95)
p_65=(sum(percent$r_freq))*100
print(paste("solution of c is =",p_65,"%"))
percent=s_h %>%
  dplyr::select(lbound,ubound,freq,r_freq)%>%
  filter(lbound>=61.95 & ubound<=71.95)
n_p=sum(percent$freq)
print(paste("solution of d is =",n_p))
```

R code Exa 1.17 Frequency table generation

```
#page 26-27
library(dplyr)
```

```
library(MASS)
miles<-data.frame(data=c(2,3,4,5,7,10,12,13,15,18,20),
                  freq=c(2,1,1,3,2,3,2,1,1,2,1))
View(miles)
total=sum(miles$freq)
print(total)
miles$r_freq=miles$freq/total
View(miles)
miles$cum_freq=cumsum(miles$r_freq)
filterdata<-filter(miles,data>=5 & data<=7)
ans1=sum(filterdata$freq)/total
print(paste("the answer c=",fractions(ans1)))
filterdata<-filter(miles,data>=12)
ans2=sum(filterdata$freq)/total
print(paste("the answer d is=",fractions(ans2)))
filterdata<-filter(miles,data<12)
ans3=sum(filterdata$freq)/total
print(paste("the answer d=",fractions(ans3)))
filterdata<-filter(miles,data>5 & data<13)
ans4=sum(filterdata$freq)/total
print(paste("the answer d=",fractions(ans4)))
```

R code Exa 1.18 Frequency table

```
#page 28
library(dplyr)
earthquake<-data.frame(year=c(2000,2001,2002,2003,2004,2005,2006,2007,2008,2009,2010,2011,2012
,
                  deaths=c(231,21357,11685,33819,228802,88003,6605,712,88
)
View(earthquake)
total=sum(earthquake$deaths)
```

```r
earthquake$r_freq=earthquake$deaths/total
View(earthquake)
earthquake$cum_freq=cumsum(earthquake$r_freq)
q_a=earthquake %>%
  dplyr::select(year,r_freq) %>%
  filter(year>=2006 & year<=2009)
View(q_a)
freq6_9=sum(q_a$r_freq)
p_6_9=freq6_9*100
cat("the answer (a) is = ",round(p_6_9,1),"%")
q_b=earthquake %>%
  dplyr::select(year,r_freq) %>%
  filter(year>2009)
View(q_b)
freq_9=sum(q_b$r_freq)
p_9=freq_9*100
cat("the answer % (b)is= ",round(p_9,1),"%")
q_c=earthquake %>%
  dplyr::select(year,r_freq) %>%
  filter(year<=2003)
View(q_c)
freq_3=sum(q_c$r_freq)
p_3=freq_3*100
cat("the answer (c) is=",round(p_3,1),"%")
q_d=earthquake %>%
  dplyr::select(year,r_freq) %>%
  filter(year==2004)
View(q_d)
freq_4=sum(q_d$r_freq)
p_4=freq_4*100
cat("the % (d) is= ",round(p_4,1),"%")
```

Chapter 2

Descriptive Statistics

R code Exa 2.1 Stem and leaf graph example1

```
# Page NO : 46
df<-data.frame(scores=c
   (33,42,49,49,53,55,55,61,63,67,68,68,69,69,72,
                  73,74,78,80,83,88,88,88,
                  90,92,94,94,94,94,96,100))
View(df)
stem(df$scores)
```

R code Exa 2.2 Stem and leaf graph example2

```
# Page NO : 47
df<-data.frame(dist=c
   (1.1,1.5,2.3,2.5,2.7,3.2,3.3,3.3,3.5,3.8,4.0,4.2,4.5,4.5,4.7,4.8,5

                      6.5,6.7,12.3))
View(df)
stem(df$dist,scale=3)
```

R code Exa 2.3 Stem and Leaf plot

```
# Page NO : 47-49
library("aplpack")
df<-data.frame(a_I=c(57,61,57,57,58,57,61,54,68,51,49,64,50,48,65,
                        52,56,46,54,49,51,47,55,55,54,42,51,56,55,51,
                        54,51,60,62,43,55,56,61,52,69,64,47,54,47))
df1<-data.frame(a_D=c(67,90,83,85,73,80,78,79,68,71,53,65,74,64,77,
                         56,66,63,70,49,56,71,67,71,58,60,72,67,57,60,
                         90,63,88,78,46,64,81,93,93))
View(df)
View(df1)
stem.leaf.backback(df$a_I,df1$a_D,m=1)
#The answer provided in the textbook is wrong.
```

R code Exa 2.4 Frequency curve example1

```
#page 49-50
library(dplyr)
remainder<-data.frame(no_times=c(0,1,2,3,4,5),
                      frequency=c(2,5,8,14,7,4))
View(remainder)
plot(remainder$no_times,remainder$frequency,xlab="no of times",
                          ylab="frequency ",type="o",col="blue")
```

R code Exa 2.5 Bar graph example1

```
#page 51
library(dplyr)
library(ggplot2)
dtabl<-data.frame(
          agegroups=c("13-25","26-44","45-64"),
          no_of_users=c(66082280, 53300200, 27885100),
          proportion=c(45,36, 19))
View(dtabl)
ggplot(dtabl,aes(x=agegroups,y=proportion,color=cut))+
  geom_bar(color="blue",stat="identity",width=0.5)+
  labs(x="agegroups",y="%")
```

R code Exa 2.6 bar graph example2

```
#page 52-53
library(dplyr)
library(ggplot2)
dtabl<-data.frame(
          race=c("1","2","3","4","5","6"),
          examinee=c(10.3,9.0,17.0,0.6,57.1,6.0),
          overall=c(5.7,14.7, 17.6,1.1,59.2,1.7))
View(dtabl)
ggplot(dtabl,aes(x=race,y=examinee),fill=race)+
  geom_bar(color="blue",stat="identity",width=0.5)+
  theme(legend.position="bottom")+
  theme(legend.direction="vertical")+
  theme(axis.text.x=element_text(angle=0))+
  geom_text(aes(label=examinee))
```

R code Exa 2.7 Two way table

```
#page 54
library(dplyr)
pets<-data.frame(dogs=c(4,4,8),
                 cats=c(2,6,8),
                 fish=c(2,2,4),
                 total=c(8,12,20))
rownames(pets)<-c("men","women","total")
last=ncol(pets)
total=0
for(c in 1:(last-1)){
  x=(pets[1,c]/pets[1,last])
  print(paste("number owning ",colnames(pets)[c],x))
  total=total+x
  }
print(paste("sum =  ", total))
```

R code Exa 2.8 Historgram graph example1

```
#page 55-56
library(ggplot2)
dtabl<-data.frame(height=c
   (59.95,60,60.5,61,61,61.5,63.5,63.5,63.5,64,64,64,64,64,64,64,

         64.5,64.5,64.5,64.5,64.5,64.5,64.5,64.5,
         66,66,66,66,66,66,66,66,66,66,
         66.5,66.5,66.5,66.5,66.5,66.5,66.5,66.5,66.5,66.5,66.5,

         67,67,67,67,67,67,67,67,67,67,67,67,
         67.5,67.5,67.5,67.5,67.5,67.5,67.5,
```

```
            68,68,
            69,69,69,69,69,69,69,69,69,69,
            69.5,69.5,69.5,69.5,69.5,
            70,70,70,70,70,70,
            70.5,70.5,70.5,
            71,71,71,
            72,72,72,
            72.5,72.5,
            73,
            73.5,
            74,75.95))
View(dtabl)
breaks=seq(59.95,75.95,by=2)
dcut=cut(dtabl$height,breaks,right=FALSE)
dout=as.data.frame(table(dcut))
dout<-transform(dout,cumFreq=cumsum(Freq),relative=
    round(prop.table(Freq),2))
ggplot(dout,aes(dcut,relative))+
  geom_col(fill="purple")+
  geom_text(aes(label=relative))+
   xlab("height")+ylab("r.freq")
#The answer may slightly vary due to rounding off
    values.
```

R code Exa 2.9 Historgram graph example2

```
#page 56-57
library(ggplot2)
dtabl<-data.frame(height=c(1,1,1,1,1,1,1,1,1,1,1,
                           2,2,2,2,2,2,2,2,2,2,
                           3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,3,

                           4,4,4,4,4,4,
                           5,5,5,5,5,
                           6,6))
```

```
View(dtabl)
breaks=seq(0.5,6.5,by=1)
dcut=cut(dtabl$height,breaks,right=FALSE)
dout=as.data.frame(table(dcut))
dout<-transform(dout,cumFreq=cumsum(Freq),relative=
   round(prop.table(Freq),2))
nobars=(6.5-0.5)/1
print(paste("the number of bars =",nobars))
ggplot(dout,aes(dcut,Freq))+
  geom_col(fill="purple")+
  geom_text(aes(label=Freq))+
         xlab("no.of books")+ylab("freq")
```

R code Exa 2.10 Historgram graph example3

```
#page 57−58
library(dplyr)
library(ggplot2)
dtabl<-data.frame(hours=c(9.95,10,2.25,16.75,0,
                          19.5,22.5,7.5,15,12.75,
                          5.5,11,10,20.75,17.5,
                          23,21.9,24,23.75,18,
                          20,15,22.9,18.8,20.5))
bins=seq(-0.01,25,by=5)
scores=cut(dtabl$hours,bins)
table(scores)
View(dtabl)
View(scores)
freq_table=transform(table(scores))
xx=transform(freq_table,Rel_freq=prop.table(Freq),
   cumfreq=cumsum(Freq))
View(xx)
ggplot(data=xx,
        aes(x=scores,y=Freq))+
    geom_col(color="red",fill="blue")+
```

```
    labs(y="no. of students",x="hours spent")+
  scale_y_continuous(breaks=seq(0,10,1))
```

R code Exa 2.11 Frequency curve example2

```
#page no 59
library(ggplot2)
df<-data.frame(lbound=c
   (39.5,49.5,59.5,69.5,79.5,89.5,99.5),
               ubound=c
                  (49.5,59.5,69.5,79.5,89.5,99.5,109.5)
                  ,
               freq=c(0,5,10,30,40,15,0),
               cfreq=c(0,5,15,45,85,100,0))
View(df)
df$mid=(df$lbound+df$ubound)/2

ggplot(df,aes(x=mid,y=freq))+
 geom_line(aes(x=mid,y=freq),size=1)+
  geom_point(aes(mid,size=0.5),shape=18,show.legend=
     FALSE)+
  scale_x_continuous(breaks=seq(from=44.5,to=104.5,
     by=10))+
  xlab("result")+ylab("number of students")
```

R code Exa 2.12 Frequency curve example3

```
#page no: 60-61
library(ggplot2)
df<-data.frame(lbound=c
   (39.5,49.5,59.5,69.5,79.5,89.5,99.5),
```

```
                    ubound=c(49.5,59.5,69.5,79.5,89.5,99.5,109.5),
                    freq=c(0,5,10,30,40,15,0),
                    cfreq=c(0,5,15,45,85,100,0))
View(df)
df$mid=(df$lbound+df$ubound)/2
df2=data.frame(lbound=c(39.5,49.5,59.5,69.5,79.5,89.5,99.5),
                    ubound=c(49.5,59.5,69.5,79.5,89.5,99.5,109.5),
                    freq=c(0,10,10,30,45,5,0),
                    cfreq=c(0,10,20,50,95,100,0))
df2$mid=(df2$lbound+df2$ubound)/2
View(df2)
ggplot(df,aes(x=mid,y=freq))+
 geom_line(aes(x=mid,y=freq),color="pink",size=1,show.legend=FALSE)+
  geom_point(aes(mid,size=0.5),shape=18,show.legend=FALSE)+
    geom_line(data=df2,aes(x=mid,y=freq),color="blue",size=1,show.legend=FALSE)+
  geom_point(aes(mid,size=0.5),shape=18,show.legend=FALSE)+
  scale_x_continuous(breaks=seq(44.5,104.5,10))+
  scale_y_continuous(breaks=seq(0,50,5))+
   xlab("result")+ylab("number of students")
```

R code Exa 2.13 Time series graph

```
#page 62-63
library(ggplot2)
consumer<-data.frame(year=c(2003,2004,2005,2006,2007,2008,2009,2010,2011,2012)
```

```
,
        jan=c
          (181.7,185.2,190.7,198.3,202.416,211.080,211.143,216.687
          ,
        feb=c
          (183.1,186.2,191.8,198.7,203.499,211.693,212.193,216.741
          ,
        mar=c
          (184.2,187.4,193.3,199.8,205.352,213.528,212.709,217.631
          ,
        apr=c
          (183.8,188,194.6,201.5,206.686,214.823,213.24,218.009,22
          ,
        may=c
          (183.5,189.1,194.4,202.5,207.949,216.632,213.856,218.178
          ,
        jun=c
          (183.7,189.7,194.5,202.9,208.352,218.815,215.693,217.965
          ,
        jul=c
          (183.9,189.4,195.4,203.5,208.299,219.964,215.351,218.011
          ,
        aug=c
          (184.6,189.5,196.4,203.9,207.917,219.086,215.834,218.312
          ,
        sep=c
          (185.2,189.9,198.8,202.9,208.49,218.783,215.969,218.439,
          ,
        oct=c
          (185,190.9,199.2,201.8,208.936,216.573,216.177,218.711,2
          ,
        nov=c
          (184.5,191,197.6,201.5,210.177,212.425,216.33,218.803,22
          ,
        dec=c
          (184.3,190.3,196.8,201.8,210.036,210.228,215.949,219.179
          ,
        annual=c
```

```
            (184,188.9,195.3,201.6,207.342,215.303,214.537,218.056,2
            )
View(consumer)
min<-as.Date("2003-1-1")
max<-as.Date("2012-12-30")
yrs<-as.Date(ISOdate(consumer$year,1,1))
ggplot(data=consumer,aes(x=yrs,y=annual))+
  geom_line(size=1)+
  scale_x_date(limits=c(min,max))+
  xlab("year")+ylab("annual price")
```

R code Exa 2.14 Measures of Location of Data Example1

```
#page no : 65-66
library(base)
library(dplyr)
price<-data.frame(
        dollars=c
           (389950,230500,158000,479000,639000,114950,5500000,387000

          575000,488800,1095000))
View(price)
M=median(price$dollars)
print(paste("median = ",M))
price1=price %>%
  select(dollars) %>%
  arrange(dollars)
View(price1)
rowcount=nrow(price1)
midindex=(rowcount+1)/2
if(floor(midindex)!=midindex)
{
  l<-(midindex-1)/2
  u<-(midindex+1)/2
} else
```

```r
{
  l<-(midindex-1)
  u<-(midindex+1)
}
q1=median(price1$dollars[1:l])
print(paste("Q1 = ",q1))
q3=median(price1$dollars[u:rowcount])
print(paste("Q3 = ",q3))
IQR1=q3-q1
print(paste("IQR = ",IQR1))
o_IQR1=(1.5)*IQR1
loutlier=q1-o_IQR1
print(paste("lower outlier =",loutlier))
uoutlier=q3+o_IQR1
print(paste("upper outlier =",uoutlier))
if(min(price1$dollars)>loutlier)
{
  print("No outliers in the lower end of price")
}  else
{
  print("Outliers in the lower end of price")
  print(paste("the outlier value is = ",min(price1$dollars)))
}
if(max(price1$dollars)<uoutlier)
{
  print("No outliers in the upper end of price")
}  else
{
  print("Outliers in the upper end of price")
  print(paste("the outlier value is = ",max(price1$dollars)))
}
```

R code Exa 2.15 Measures of Location of Data example2

```r
#page no : 66
library(dplyr)
t_s<-data.frame(
            minimum=c(32,25.5),
            q1=c(56,78),
            median=c(74.5,81),
            q3=c(82.5,89),
            maximum=c(99,98)
)
rownames(t_s)<-c("day","night")
View(t_s)
t_s$IQR=t_s$q3-t_s$q1
t_s$outlierq1=t_s$q1-(t_s$IQR)*1.5
t_s$outlierq3=t_s$q3+(t_s$IQR)*1.5
result1=t_s %>% select(q1,q3,minimum,maximum,outlierq1,outlierq3) %>%
              filter(row.names(t_s)=="day")
if(result1$minimum>result1$outlierq1)
{
  print("No outliers in the lower end of day")
}  else
{
  print("Outliers in the lower end of day")
  print(paste("any value < than", result1$outlierq1," is outlier"))
  print(paste("the value=",result1$minimum,"is outlier"))
}
if(result1$maximum<result1$outlierq3)
{
  print("No outliers in the upper end of day")
}  else
{
  print("Outliers in the upper end of day")
  print(paste("any value > than", result1$outlierq3," is outlier"))
  print(paste("the value=",result1$maximum,"is outlier"))
```

```
}
result2=t_s %>% select(q1,q3,minimum,maximum,
    outlierq1,outlierq3) %>%
  filter(row.names(t_s)=="night")
if(result2$minimum>result2$outlierq1)
{
  print("No outliers in the lower end of night")
}  else
{
  print("Outliers in the lower end of night")
  print(paste("any value < than", result2$outlierq1,
     " is outlier"))
  print(paste("the value=",result2$minimum,"is
     outlier"))
}
if(result2$maximum<result2$outlierq3)
{
  print("No outliers in the upper end of night")
}  else
{
  print("Outliers in the upper end of night")
  print(paste("any value > than", result2$outlierq3,
     " is outlier"))
  print(paste("the value=",result2$maximum,"is
     outlier"))
}
```

R code Exa 2.16 Measures of Location of Data example3

```
#page no : 66-67
library(base)
dtabl=c
    (4,4,5,5,5,5,5,6,6,6,6,6,6,6,7,7,7,7,7,7,7,7,7,7,7,7,8,8,8,8,8,8,8

                  9,9,9,9,9,9,9,10,10,10)
```

```
dtabl1<-data.frame(
        hours=c(4,5,6,7,8,9,10),
        freq=c(2,5,7,12,14,7,3))
total=sum(dtabl1$freq)
dtabl1$rfreq=dtabl1$freq/total
dtabl1$cum_freq=cumsum(dtabl1$rfreq)
View(dtabl1)
print(paste("28% is",quantile(dtabl,0.28)))
print(paste("median is",quantile(dtabl,0.50)))
print(paste("75% is",quantile(dtabl,0.75)))
#The answer may slightly vary due to rounding off
   values.
```

R code Exa 2.17 Measures of Location of Data example4

```
#page no : 68
library(base)
dtabl=c
   (4,4,5,5,5,5,5,6,6,6,6,6,6,6,7,7,7,7,7,7,7,7,7,7,7,7,8,8,8,8,8,8,8

                9,9,9,9,9,9,9,10,10,10)
dtabl1<-data.frame(
        hours=c(4,5,6,7,8,9,10),
        freq=c(2,5,7,12,14,7,3))
total=sum(dtabl1$freq)
dtabl1$rfreq=dtabl1$freq/total
dtabl1$cum_freq=cumsum(dtabl1$rfreq)
View(dtabl1)
dtabl_len=length(dtabl)
find_percentile<-function(val,dtabl_len)
{
  i=(val/100)*(dtabl_len+1)
  lvalue=floor(i)
  uvalue=ceiling(i)
  x=c(lvalue,uvalue)
```

```
  return(x)
}
test1=find_percentile(80,dtabl_len)
per_80=(dtabl[test1[1]]+dtabl[test1[2]])/2
test2=find_percentile(90,dtabl_len)
per_90=(dtabl[test2[1]]+dtabl[test2[2]])/2
test3=find_percentile(25,dtabl_len)
per_25=(dtabl[test3[1]]+dtabl[test3[2]])/2
print(paste("80th percentile is",per_80))
print(paste("90th percentile is",per_90))
print(paste("first quartile is",per_25))
```

R code Exa 2.18 Measures of Location of Data example5

```
#page no : 68-69
library(base)
dtabl=c
   (18,21,22,25,26,27,29,30,31,33,36,37,41,42,47,52,55,57,58,62,64,6
        71,72,73,74,76,77)
dtabl_len=length(dtabl)
find_percentile<-function(val,dtabl_len)
{
  i=(val/100)*(dtabl_len+1)
  lvalue=floor(i)
  uvalue=ceiling(i)
  x=c(lvalue,uvalue)
  return(x)
}
test1=find_percentile(70,dtabl_len)
per_70=(dtabl[test1[1]]+dtabl[test1[2]])/2
test2=find_percentile(83,dtabl_len)
per_83=(dtabl[test2[1]]+dtabl[test2[2]])/2
print(paste("70th percentile is",per_70))
print(paste("83rd percentile is",per_83))
```

R code Exa 2.19 Measures of Location of Data example6

```
#page no : 69
library(base)
dtabl=c
   (18,21,22,25,26,27,29,30,31,33,36,37,41,42,47,52,55,57,58,62,64,67

        71,72,73,74,76,77)
dtabl_len=length(dtabl)
find_range<-function(val)
{
count=0
for(i in 1:dtabl_len)
{
  if(dtabl[i]<val)
  {
    count=count+1
  }
  else
  {
    break
  }
}
return(count)
}
find_percentile<-function(val)
{
 x=find_range(val)
   y=1
  i=((x+0.5*y)/(dtabl_len))*100
  return(round(i))
}
per_58=find_percentile(58)
per_25=find_percentile(25)
```

```
print(paste("58 is",per_58,"th percentile"))
print(paste("25 is",per_25,"th percentile "))
```

R code Exa 2.23 survey for buying gym

```
#page no : 71
library(dplyr)
minutes<-data.frame(
        x=c
           (0,40,60,30,60,10,45,30,300,90,30,120,60,0,20)
           )
View(minutes)
M=median(minutes$x)
minutes1=minutes %>%
  select(x) %>%
  arrange(x)
View(minutes1)
rowcount=nrow(minutes1)
midindex=(rowcount+1)/2
if(floor(midindex)!=midindex)
{
  l<-(midindex-1)/2
  u<-(midindex+1)/2
} else
{
  l<-(midindex-1)
  u<-(midindex+1)
}
q1=median(minutes1$x[1:l])
print(paste("minimum=",min(minutes)))
print(paste(" (Q1) = ",q1," minutes"))
print(paste("median = ",M))
q3=median(minutes1$x[u:rowcount])
print(paste(" (Q3) = ",q3, " minutes"))
IQR1=q3-q1
```

```
print(paste("maximum=",max(minutes)))
o_IQR1=(1.5)*IQR1
loutlier=q1-o_IQR1
uoutlier=q3+o_IQR1
minutes3<-data.frame(
  x=c(0,40,60,30,60,10,45,30,90,30,120,60,0,20))
View(minutes3)
M=median(minutes3$x)
minutes4=minutes3 %>%
  select(x) %>%
  arrange(x)
View(minutes4)
rowcount=nrow(minutes4)
if(rowcount%%2==0)
{
  midindex=rowcount/2
}else
{
midindex=(rowcount+1)/2
}
if(floor(midindex)!=midindex)
{
  l<-(midindex-1)/2
  u<-(midindex+1)/2
} else
{
  l<-(midindex-1)
  u<-(midindex+1)
}
q1=median(minutes4$x[1:l])
print(paste("minimum=",min(minutes3)))
print(paste(" (Q1) = ",q1," minutes"))
q3=median(minutes4$x[u:rowcount])
print(paste(" (Q3) = ",q3, " minutes"))
IQR1=q3-q1
print(paste("maximum=",max(minutes3)))
#The answer provided in the textbook is wrong.
```

R code Exa 2.24 Measure of Centre

```
#page no : 72-73
library(dplyr)
aids<-c
    (3,4,8,8,10,11,12,13,14,15,15,16,16,17,17,18,21,22,22,24,24,25,26
          27,27,29,29,31,32,33,33,34,34,35,37,40,44,44,47)
aids_mean=mean(aids)
aids_median=median(aids)
print(paste("mean = ",round(aids_mean,1)))
print(paste("median = ",aids_median))
```

R code Exa 2.25 Measures of the Center of the Data example2

```
#page no : 73
earning<-c()
for(i in 1:49)
{
  earning[i]=30000
}
earning[50]=5000000
earning_mean=mean(earning)
earning_median=median(earning)
print(paste("the mean is ",earning_mean))
print(paste("the median is ",earning_median))
```

R code Exa 2.26 Measures of the Center of the Data example3

```r
#page no : 73
library(dplyr)
student<-c
  (50,53,59,59,63,63,72,72,72,72,7276,78,81,83,84,84,84,90,93)

mode<-function(x)
{
  ux<-unique(x)
  tab<-tabulate(match(x,ux))
  return(ux[tab==max(tab)])
}
s_mode=mode(student)
print(paste("the MODE is ",s_mode))
```

R code Exa 2.28 Calculating the Arithmetic Mean of Grouped Frequecy Tables example1

```r
#page no : 74-75
s_test<-data.frame(
                grade_l=c
                  (50,56.5,62.5,68.5,74.5,80.5,86.5,92.5)
                  ,
                grade_u=c
                  (56.5,62.5,68.5,74.5,80.5,86.5,92.5,98.5)
                  ,
                s_no=c(1,0,4,4,2,3,4,1))
s_test$midpoint=(s_test$grade_u+s_test$grade_l)/2
s_test$fm=s_test$midpoint*s_test$s_no
sum_fm=sum(s_test$fm)
sum_s_no=sum(s_test$s_no)
mean_value=sum_fm/sum_s_no
print(paste("the sum of product is ",round(sum_fm,2)
  ))
print(paste("the mean is ",round(mean_value,2)))
```

R code Exa 2.29 Measures of the Spread of the Data Example1

```
#page no : 81-82
library(dplyr)
avg_ages<-data.frame(ages=c(9,9.5,9.5,10,10,10,10,10.5,10.5,10.5,10.5,
                            11,11,11,11,11,11,11.5,11.5,11.5))
View(avg_ages)
ages1=avg_ages %>%
    group_by(ages)%>%
    mutate(frequency=n()) %>%
    select(ages,frequency)%>%
    distinct()
View(ages1)
total=sum(ages1$frequency)
ages1$val1=ages1$ages*ages1$frequency
xbar=sum(ages1$val1)/total
View(ages1)
ages1$dev1=ages1$ages-xbar
ages1$devsq=ages1$dev1*ages1$dev1
ages1$fd=ages1$frequency*ages1$devsq
View(ages1)
sv=sum(ages1$fd)/(total-1)
sd=sqrt(sv)
print(paste("the sample mean is =",round(xbar,3)))
print(paste("the sample variance is =",round(sv,4)))
print(paste("the sample sd is =",round(sd,2)))
```

R code Exa 2.30 Measures of the Spread of the Data example2

```
#page no : 83-84
```

```r
library(dplyr)
calculus<-data.frame(marks=c
  (33,42,49,49,53,55,55,61,63,67,68,68,69,69,72,73,74,78,
            80,83,88,88,88,90,92,94,94,94,94,96,100)
               )
View(calculus)
calc1=calculus %>%
    group_by(marks)%>%
    mutate(frequency=n()) %>%
    select(marks,frequency)%>%
    distinct()
View(calc1)
total=sum(calc1$frequency)
calc1$r_freq=round((calc1$frequency/total),3)
calc1$c_freq=round((cumsum(calc1$r_freq)),3)
View(calc1)
calc1$val1=calc1$marks*calc1$frequency
xbar=sum(calc1$val1)/total
View(calc1)
calc1$dev1=calc1$marks-xbar
calc1$devsq=calc1$dev1*calc1$dev1
calc1$fd=calc1$frequency*calc1$devsq
View(calc1)
sv=sum(calc1$fd)/(total-1)
sd=sqrt(sv)
rowcount=nrow(calculus)
midindex=(rowcount+1)/2
if(floor(midindex)!=midindex)
{
  l<-(midindex-1)/2
  u<-(midindex+1)/2
} else
{
  l<-(midindex-1)
  u<-(midindex+1)
}
q1=median(calculus$marks[1:l])
```

```
q3=median(calculus$marks[u:rowcount])
IQR1=q3-q1
mean1=mean(calculus$marks)
median1=median(calculus$marks)
print(paste("the sample mean is =",round(mean1,1)))
print(paste("the sample sd is =",round(sd,1)))
print(paste("the median is =",median1))
print(paste("the 1st quartile is =",q1))
print(paste("the 3rd quartile is =",q3))
print(paste("the IQR is =",IQR1))
```

R code Exa 2.31 Standard deviation of grouped frequency tables

```
#page no : 84-85
dtabl<-data.frame(lvalue=c(0,3,6,9,12),
                  uvalue=c(2,5,8,11,14),
                  f=c(1,6,10,7,0))

dtabl$m=(dtabl$lvalue+dtabl$uvalue)/2
total=sum(dtabl$f)
dtabl$fm=dtabl$f*dtabl$m
xbar=sum(dtabl$fm)/total
View(dtabl)
dtabl$dev1=dtabl$m-xbar
dtabl$devsq=dtabl$dev1*dtabl$dev1
dtabl$fmxbar=dtabl$f*dtabl$devsq
View(dtabl)
sv=sum(dtabl$fmxbar)/(total-1)
sd=sqrt(sv)
print(paste("the variance is =",round(sv,2)))
print(paste("the sample sd is =",round(sd,2)))
#The answer provided in the textbook is wrong.
```

R code Exa 2.32 Comparing values from different data sets

```
#page no : 85
dtabl<-data.frame(name=c("John","Ali"),
                  GPA=c(2.85,77),
                  meangpa=c(3.0,80),
                  schoolsd=c(0.7,10))
dtabl$z1=(dtabl$GPA-dtabl$meangpa)/dtabl$schoolsd
View(dtabl)
no=nrow(dtabl)
for(i in 1:no)
{
print(paste("sd of = ",dtabl$name[i],"is",round(
   dtabl$z1[i],2)))
}
if(dtabl$z1[1]>dtabl$z1[2])
  {
print(paste(dtabl$name[1],"has more GPA"))
} else
{
  print(paste(dtabl$name[2],"has more GPA"))
}
```

Chapter 3

Probability Topics

R code Exa 3.1 Finding probabilities example1

```
#page no :135
library(MASS)
S=c(1:19)
A=c(2,4,6,8,10,12,14,16,18)
B=c(14,15,16,17,18,19)
cat("a. S= ",S)
cat("b. A= ",A)
cat("b. B= ",B)
P_A=length(A)/length(S)
print(paste(" c. P(A) =",fractions(P_A)))
P_B=length(B)/length(S)
print(paste("c. P(B) =",fractions(P_B)))
ABint=intersect(A,B)
ABunion=union(A,B)
cat("d. AintB =",ABint)
cat("d. AUB=",ABunion)
P_AintB=length(ABint)/length(S)
print(paste("e. P(AintB) =",fractions(P_AintB)))
P_AUB=length(ABunion)/length(S)
print(paste("e. P(AUB) =",fractions(P_AUB)))
Acomp=setdiff(S,A)
```

```
cat("f. A' =",Acomp)
P_Acomp=length(Acomp)/length(S)
print(paste("f. P(A') =",fractions(P_Acomp)))
print(paste("g. P(A)+P(A') =",P_A+P_Acomp))
P_A_B=P_AintB/P_B
P_B_A=P_AintB/P_A
print(paste("h. P(A|B) = ",fractions(P_A_B)))
print(paste("h. P(B|A) = ",fractions(P_B_A)))
if(P_A_B==P_B_A){
  print("h. P(A|B) is equal to P(B|A)")
}else {
  print("h. P(A|B) not equal to P(B|A)")
}
```

R code Exa 3.2 Finding probabilities example2

```
#page no :136-137
library(MASS)
S=c(1:6)
T=c(2)
A=c(2,4,6)
B=c(1,2,3)
P_T=length(T)/length(S)
cat("a. T =",T)
print(paste("a. P(T) =",fractions(P_T)))
P_A=length(A)/length(S)
cat("b. A =",A)
print(paste("b. P(A) =",fractions(P_A)))
P_B=length(B)/length(S)
cat("c. B =",B)
print(paste("c. P(B) =",fractions(P_B)))
Acomp=setdiff(S,A)
cat("d. A' =",Acomp)
P_Acomp=length(Acomp)/length(S)
print(paste("d. P(A') =",fractions(P_Acomp)))
```

```
ABint=intersect(A,B)
ABunion=union(A,B)
P_AintB=length(ABint)/length(S)
P_A_B=P_AintB/P_B
P_B_A=P_AintB/P_A
print(paste("e.P(A|B) =",fractions(P_A_B)))
print(paste("f.P(B|A) =",fractions(P_B_A)))
cat("g. AintB =",ABint)
print(paste("g. P(AintB) =",fractions(P_AintB)))
P_AUB=length(ABunion)/length(S)
cat("h. AUB =",ABunion)
print(paste("h. P(AUB) =",fractions(P_AUB)))
Bcomp=setdiff(S,B)
ABunion=union(A,Bcomp)
cat("i. AUB'=",ABunion)
PABunion=length(ABunion)/length(S)
print(paste("i. P(AUB')=",fractions(PABunion)))
N=c(2,3,5)
PN=length(N)/length(S)
cat("j. N=",N)
print(paste("j. P(N)=",fractions(PN)))
print("k. P(7)=0")
```

R code Exa 3.3 Finding probabilities example3

```
#page no :137-138
library(MASS)
S=data.frame(sex=c('M','F'),
             RH=c(43,44),
             LH=c(9,4))
S$rtotal=S$RH+S$LH
View(S)
total=sum(S$rtotal)
P_M=S$rtotal[1]/total
P_F=S$rtotal[2]/total
```

```
P_R=sum(S$RH)/total
P_L=sum(S$LH)/total
table=round(prop.table(S[,2:3]),2)
table=as.matrix(table)
table
RintM=table[1]
FintL=table[4]
LintF=table[4]
MunionF=sum(table)
MunionR=table[1]+table[2]+table[3]
FunionL=table[3]+table[4]+table[2]
csum=apply(table[,1:2],2,sum)
rsum=apply(table[,1:2],1,sum)
P_Mdash=sum(table)-rsum[1]
PRintM=round(RintM/P_M,4)
PFintL=round(FintL/P_L,4)
PLintF=round(LintF/P_F,4)
print(paste(" a. P(M) =",P_M))
print(paste(" b. P(F) =",P_F))
print(paste(" c. P(R) =",P_R))
print(paste(" d. P(L) =",P_L))
print(paste(" e. P(M int R) =",RintM))
print(paste(" f. P(F int L) =",FintL))
print(paste(" g. P(M U F) =",MunionF))
print(paste(" h. P(M U R) =",MunionR))
print(paste(" i. P(F U L) =",FunionL))
print(paste(" j. P(M') =",P_Mdash))
print(paste(" k. P(R|M) =",PRintM))
print(paste(" l. P(F|L) =",PFintL))
print(paste(" m. P(L|F) =",PLintF))
```

R code Exa 3.6 Mutually Exclusive Events of coins

```
# page no : 140
library(prob)
```

```
library(MASS)
coin=tosscoin(2)
A=subset(coin,toss1=="H" | toss2=="H")
P_A=nrow(A)/nrow(coin)
print(paste(" one tail =",fractions(P_A)))
B=subset(coin,toss1=="T" & toss2=="T")
P_B=1-P_A
print(paste(" all tail =",fractions(P_B)))
C=subset(coin,toss1=="H" & toss2=="H")
P_C=nrow(C)/nrow(coin)
print(paste(" all heads =",fractions(P_C)))
BintC=intersect(B,C)
P_BintC=nrow(BintC)/nrow(coin)
D=subset(coin,toss1=="T" & toss2=="T")
P_D=nrow(D)/nrow(coin)
print(paste("more than one tail =",fractions(P_D)))
E=subset(coin,toss1=="H" & (toss1=="H"|toss2=="T"))
P_E=nrow(E)/nrow(coin)
print(paste("head-first roll =",fractions(P_E)))
F=subset(coin,toss1=="T" |toss2=="T")
P_F=nrow(F)/nrow(coin)
print(paste("one tail-two filps =",fractions(P_F)))
```

R code Exa 3.7 Fliping of two coins and finding the probability

```
# page no : 141
library(prob)
library(MASS)
coin=tosscoin(2)
F=subset(coin,toss1=="T" | toss2=="T")
P_F=nrow(F)/nrow(coin)
print(paste("a. P(F)=",fractions(P_F)))
G=subset(coin,(toss1=="T" & toss2=="T") |(toss1=="H"
    & toss2=="H"))
P_G=nrow(G)/nrow(coin)
```

```
print(paste("b. P(G) =",fractions(P_G)))
H=subset(coin,toss1=="H" & (toss1=="H"|toss2=="T"))
P_H=nrow(H)/nrow(coin)
print(paste("c.P(H) =",fractions(P_H)))
FintG=intersect(F,G)
P_FintG=nrow(FintG)/nrow(coin)
if(P_FintG==0)
{
  print("d. F,G—mutually exclusive")
}else
{
  print("d. F,G—not mutually exclusive")
}
J=subset(coin,toss1=="T" & toss2=="T")
JintH=intersect(J,H)
P_JintH=nrow(JintH)/nrow(coin)
if(P_JintH==0)
{
  print("e.J,H—mutually exclusive")
}else
{
  print("e.J,H—not mutually exclusive")
}
```

R code Exa 3.8 Rolling of six sided dice

```
# page no : 141-142
library(prob)
library(MASS)
library(sets)
S=rolldie(1)
A=subset(S,(X1 %% 2!=0))
B=subset(S,(X1 %% 2==0))
P_A=nrow(A)/nrow(S)
print(paste("P(A) =",fractions(P_A)))
```

```
P_B=nrow(B)/nrow(S)
print(paste("P(B) =",fractions(P_B)))
C=subset(S,(X1 %% 2!=0) & (X1>2))
D=subset(S,(X1 %% 2==0) & (X1<5))
CintD=intersect(C,D)
P_CintD=length(CintD)/nrow(S)
if(P_CintD==0)
{
  print("C,D-mutually exclusive")
}else
{
  print("C,D-not mutually exclusive")
}
E=subset(S,(X1 < 5))
CintE=intersect(C,E)
P_CintE=length(CintE)/nrow(S)
if(P_CintE==0)
{
  print("C,E-mutually exclusive")
}else
{
  print("C,E-not mutually exclusive")
}
CintA=intersect(C,A)
P_CintA=length(CintA)/nrow(S)
P_C_A=P_CintA/P_B
print(paste(" P(C|A) = ",fractions(P_C_A)))
```

R code Exa 3.9 Mutually Exclusive Events example1

```
#page no:142-143
P_G=0.6
P_H=0.5
PGintH=0.3
P_G_H=PGintH/P_H
```

```
if(P_G_H==P_G)
{
  print(paste("a. P(G|H) = P(G)"))
}else
  {
  print(paste("a. P(G|H) != P(G)"))
  }
PGH=P_G*P_H
if(PGintH==PGH)
{
  print(paste("b. P(G int H) = P(G)P(H)"))
}else
{
  print(paste("b. P(G int H) != P(G)P(H)"))
}
```

R code Exa 3.10 Mutually Exclusive Events example2

```
#page no:143
P_C=0.75
P_D=0.3
P_C_D=0.75
PCintD=0.225
if(P_C_D==P_C)
{
  print(paste('independent ,P(C|D)=P(C)'))
}else
  {
      print(paste('dependent P(C|D)!=P(C)'))
  }
if(PCintD==0)
{
  print(paste('mutually exclusive P(C int D)=0'))
}else
{
```

```
  print(paste('not mutually exclusive P(C int D)!=0'
    ))
}
P_D_C=PCintD/P_C
print(paste("c. P(D|C) =", P_D_C))
```

R code Exa 3.11 Drawing of red and blue cards from a box

```
# page no : 144
library(prob)
library(MASS)
S<-c('R1','R2','R3','B1','B2','B3','B4','B5')
R1<-c('R1','R2','R3')
B1<-c('B1','B2','B3','B4','B5')
E1<-c('R2','B2','B4')
G1<-c('B4','B5')
H1<-c('B1','B2','B3','B4')
P_R=length(R1)/length(S)
print(paste("P(R) =",fractions(P_R)))
P_B=length(B1)/length(S)
print(paste("P(B) =",fractions(P_B)))
P_E=length(E1)/length(S)
print(paste("P(E) =",fractions(P_E)))
EintB=intersect(E1,B1)
P_EintB=length(EintB)/length(S)
P_E_B=P_EintB/P_B
print(paste("The P(E|B) =",fractions(P_E_B)))
BintE=intersect(B1,E1)
P_BintE=length(BintE)/length(S)
P_B_E=P_BintE/P_E
print(paste("The P(B|E) =",fractions(P_B_E)))
RintB=intersect(R1,B1)
P_RintB=length(RintB)/length(S)
if(P_RintB==0)
{
```

```
  print("R,B—mutually exclusive")
}else
{
  print("R,B—not mutually exclusive")
}
P_G=length(G1)/length(S)
print(paste("P(G)=",fractions(P_G)))
P_H=length(H1)/length(S)
print(paste("P(H)=",fractions(P_H)))
GintH=intersect(G1,H1)
P_GintH=length(GintH)/length(S)
P_G_H=P_GintH/P_H
print(paste("The P(G|H) =",fractions(P_G_H)))
if(P_G==P_G_H)
{
  print("G,H—independent")
}else
{ print(" G,H—not independent")
  }
```

R code Exa 3.12 Mutually Exclusive Events example3

```
#page no:145
P_F=0.60
P_L=0.5
PFintL=0.45
P_L_F=0.75
PFL=P_F*P_L
if(PFintL==PFL)
{
  print(paste("a. female,long hair—independent"))
}else
  {
    print(paste("a. female,long hair—dependent"))
  }
```

```
if(P_L_F==P_L)
{
  print(paste("b. female,long hair-independent"))
}else
{
  print(paste("b. female,long hair-dependent"))
}
```

R code Exa 3.13 Tossing of fair coin

```
# page no : 145-146
library(prob)
library(MASS)
S=tosscoin(1)
cat(paste("a. solution =",S$toss1[1] ,"and ",S$toss1
   [2],";",nrow(S)))
S1=rolldie(1)
writeLines("\n b.")
print(S1$X1)
print(paste("b.",nrow(S1)))
print(paste("c.", nrow(S1)*2))
cat(paste("d. ="))
count=0
for(val1 in 1:nrow(S1))
  {
  for(val in 1:nrow(S))
  {
    count=count+1
    cat(paste(S$toss1[val],S1$X1[val1])," ")
  }
}
A=c('H2','H4','H6')
P_A=length(A)/count
writeLines("\n")
print(paste("e. P(A) =", fractions(P_A)))
```

```
B=c('H3')
P_B=length(B)/count
print(paste("f. P(B) =", fractions(P_B)))
AintB=intersect(A,B)
P_AintB=length(AintB)/count
if(P_AintB==0)
{
  print("g. A,B-mutually exclusive")
}else
{
  print("g. A,B-not mutually exclusive")
}
righteq=P_A*P_B
lefteq=P_AintB

if(lefteq==righteq)
{
  print("h. A,B-independent")
}else
{ print("h. A,B-dependent")
}
```

R code Exa 3.14 Two basic rules of Probability example1

```
#page no:147
P_A=0.60
P_B=0.35
PAintB=0
PUB=P_A+P_B
PUBcomp=1-PUB
print(paste("PUB=",PUB))
print(paste("1-PUB =", PUBcomp))
```

R code Exa 3.15 Two basic rules of Probability example2

```
#page no:147-148
P_A=0.65
P_B=0.65
P_B_A=0.90
PAintB=P_B_A*P_A
print(paste("a. solution =",PAintB))
PUB=P_A+P_B-PAintB
print(paste("b. solution = ",PUB))
PBA=P_A*P_B
if(PAintB==PBA)
{
  print(paste("c. A,B-independent"))
}else
{
  print(paste("c. A,B-not independent"))
}
if(PAintB==0)
{
  print(paste("d. A,B-mutually exclusive"))
}else
{
  print(paste("d. A,B-not mutually exclusive"))
}
```

R code Exa 3.16 Two basic rules of Probability example3

```
#page no:148-149
library(MASS)
N=150
A_swim=75
A_swim1=40
I_swim=47
I_swim1=30
```

```
N_swim=N-(A_swim+I_swim)
N_swim1=10
P_N_swim=N_swim/N
print(paste("a. solution = ",N_swim,"/",N))
P_swim_4=(A_swim1+I_swim1+N_swim1)/N
print(paste("b. solution = ",A_swim1+I_swim1+N_swim1
    ,"/",N))
P_A_swim1=A_swim1/N
print(paste("c. solution = ",A_swim1,"/",N))
PAintI=0
PNint4=0.0667
P=P_N_swim*P_swim_4
if(PAintI==0)
{
  print(paste("d. A and B are mutually exclusive"))
}else
{
  print(paste("d. A and B are not mutually exclusive
    "))
}
if(PNint4==P)
{
  print(paste("e. A And B are independent"))
}else
{
  print(paste("e.A And B are not independent"))
}
```

R code Exa 3.17 Finding probabilities example4

```
#page no: 149
P_M=0.2
P_S=0.65
P_M_S=0.25
PMintS=P_M_S*P_S
```

```
print(paste("a. solution=",PMintS))
PMunionS=P_M+P_S-PMintS
print(paste("b. solution=",PMunionS))
if(P_M_S==P_M){
  print("c. M and S are independent")
}else
{
  print("c. M and S are not independent")
}
if(PMintS==0){
  print("d. M and S are mutually exclusive")
}else
{
  print("d. M and S are not mutually exclusive")
}
```

R code Exa 3.18 Finding probabilities example5

```
#page no:150
P_B=14.3/100
P_N=85/100
P_N_B=2/100
PBintN=P_B*P_N_B
PBunionN=P_B+P_N-PBintN
print(paste(" a. P(B) =",P_B))
print(paste(" a. P(N) =",P_N))
print(paste(" b. P(N|B) =",P_N_B))
print(paste(" c. P(B int N) =",round(PBintN,4)))
print(paste(" d. P(B U N) =",round(PBunionN,4)))
if(P_N==P_N_B){
  print("e. B and N are independent")
}else
{
  print("e. B and N are not independent")
}
```

```
if(PBintN==0){
  print("f. B and N are mutually exclusive")
}else
{
  print("f. B and N are not mutually exclusive")
}
```

R code Exa 3.19 Finding probabilities example6

```
#page no:151
P_B=14.3/100
P_N=85/100
P_N_B=2/100
PBintN=P_B*P_N_B
PBunionN=P_B+P_N-PBintN
P_P_B=1-P_N_B
PBintP=P_P_B*P_B
P_Bcomp=1-P_B
P_P=1-P_N
print(paste(" a. P(P|B) =",P_P_B))
print(paste(" b. P(B int P) =",round(PBintP,4)))
print(paste(" c. P(B') =",P_Bcomp))
print(paste(" d. P(P) =",P_P))
```

R code Exa 3.20 Contingency tables example1

```
#page no:152-153
library(MASS)
df=data.frame(violation=c(25,45,70),
              noviolation=c(280,405,685),
              rtotal=c(305,450,755))
rownames(df)=c("cell","nocell","ctotal")
View(df)
```

```
P_D_C=df$rtotal[1]/df$rtotal[3]
P_D_NV=df$noviolation[3]/df$rtotal[3]
PNVintC=df$noviolation[1]/df$rtotal[3]
e=df$violation[1]/df$violation[3]
f=df$noviolation[2]/df$noviolation[3]
term1=(df$rtotal[1]/df$rtotal[3])+(df$noviolation[3]
   /df$rtotal[3])
term2=(df$noviolation[1]/df$rtotal[3])
PC_U_NV=term1-term2
print(paste("a.solution=",df$rtotal[1],"/",df$rtotal
   [3]))
print(paste("b.solution=",df$noviolation[3],"/",df$
   rtotal[3]))
print(paste("c.solution=",df$noviolation[1],"/",df$
   rtotal[3]))
print(paste("d.solution=",fractions(PC_U_NV)))
print(paste("e.solution=",df$violation[1],"/",df$
   violation[3]))
print(paste("f.solution=",df$noviolation[2],"/",df$
   rtotal[2]))
#"The answer may vary due to difference in
   representation."
```

R code Exa 3.21 Contingency tables example2

```
#page no:153-155
library(MASS)
df=data.frame(coast=c(18,16,34),
              lakes=c(16,25,41),
              peaks=c(11,14,25),
              rtotal=c(45,55,100))
rownames(df)=c("F","M","ctotal")
View(df)
PFintC=df$coast[1]/df$rtotal[3]
P_F=df$rtotal[1]/df$rtotal[3]
```

```
P_C=df$coast[3]/df$rtotal[3]
PFC=P_F*P_C
print(paste("b.1 P(FintC) =",PFintC))
print(paste("b.2 P(F)P(C) =",PFC))
if(PFintC==PFC){
  print("b. F,C-independent")
}else
{
  print("b. F,C-not independent")
}
P_M_L=df[[2]][2]/df$lakes[3]
print(paste("c.2  P(M|L)= ",fractions(P_M_L)))
P_P=df[[3]][[3]]/df$rtotal[3]
print(paste("d.1  P(F)= ",fractions(P_F)))
print(paste("d.2  P(P)= ",fractions(P_P)))
PFintP=df$peaks[1]/df$rtotal[3]
print(paste("d.3  P(F int P)= ",fractions(PFintP)))
FUP=(P_F+P_P)-PFintP
print(paste("d.4  P(FUP)= ",fractions(FUP)))
#The answer may vary due to difference in
   representation.
```

R code Exa 3.22 Contingency tables example4

```
#page no:155-156
library(MASS)
df=data.frame(D1=c(1/15,4/15),
              D2=c(1/12,3/12),
              D3=c(1/6,1/6))
rs <- rowSums(df)
df <- cbind(df,rs)
cs <- colSums(df)
df_s <- sum(df)
df <- rbind(df,c(cs,df_s))
rownames(df)=c("caught","Ncought","ctotal")
```

```
View(df)
b=fractions(df$rtotal[2])
print(paste("b. solution= ",fractions(b)))
P_CD1=df$D1[1]
P_CD2=df$D2[1]
c = P_CD1 + P_CD2
print(paste("c. solution= ",fractions(c)))
#The answer provided in the textbook is wrong.
```

R code Exa 3.23 Contingency tables example3

```
#page no:156-157
library(MASS)
df=data.frame(robbery=c(145.7,133.1,119.3,113.7),
              burg=c(732.1,717.7,701,702.2),
              rape=c(29.7,29.1,27.7,26.8),
              vehicle=c(314.7,259.2,239.1,229.6))
rownames(df)=c("2008","2009","2010","2011")
View(df)
df$rtotal=rowSums(df)
ctotal<-c(colSums(df[]))
total=ctotal[5]
P_8=df$rtotal[1]/total
P_9=df$rtotal[2]/total
P_10=df$rtotal[3]/total
P_11=df$rtotal[4]/total
P_R=ctotal[1]/total
P_B=ctotal[2]/total
P_Ra=ctotal[3]/total
P_V=ctotal[4]/total
P9intR=df$robbery[2]/total
P10intB=df$burg[3]/total
P11intRa=df$rape[4]/total
P8intV=df$vehicle[1]/total
P10UB=(P_10+P_B)-P10intB
```

```
P11_Ra=P11intRa/P_Ra
PV_8=P8intV/P_8
print(paste("a.solution=",round(P9intR,4)))
print(paste("b.solution=",round(P10intB,4)))
print(paste("c.solution=",round(P10UB,4)))
print(paste("d.solution=",round(P11_Ra,4)))
print(paste("e.solution=",round(PV_8,4)))
```

R code Exa 3.24 drawing of three colored balls from urn

```
#page no:157-159
library("DiagrammeR")
library(data.tree)
library("prob")
library("dplyr")
library("MASS")
x<-Node$new(" ")
    firstdraw_blue<-x$AddChild("8B")
       seconddraw_blue<-firstdraw_blue$AddChild("8B"
          )
       seconddraw_red<-firstdraw_blue$AddChild("3R")
           thirddraw_blue<-seconddraw_blue$AddChild(
              "64BB")
           thirddraw_red<-seconddraw_red$AddChild("
              24BR")
    firstdraw_red<-x$AddChild("3R")
       seconddraw_blue<-firstdraw_red$AddChild("8B")
       seconddraw_red<-firstdraw_red$AddChild("3R")
            thirddraw_blue<-seconddraw_blue$AddChild
               ("24RB")
            thirddraw_red<-seconddraw_red$AddChild("
               9RR")
print(x)
plot(x)
balls_R <- c(rep("R1",1),rep("R2",1),rep("R3",1))
```

```
balls_B <- c(rep("B1",1),rep("B2",1),rep("B3",1),rep("B4",1),rep("B5",1),rep("B6",1),rep("B7",1),rep("B8",1))
apply(as.data.frame(expand.grid(balls_B,balls_R)),1,function(x){paste(x,collapse = "")})
M<-unique(urnsamples(balls,size=2,replace=TRUE,ordered=TRUE))
N<-probspace(M)
P_RR=(3/11)*(3/11)
print(paste("b. P(RR)=",fractions(P_RR)))
P_RB_U_BR=((3/11)*(8/11))+((8/11)*(3/11))
print(paste("c. P(RR U BR)=",fractions(P_RB_U_BR)))
P_R1_int_B2=((3/11)*(8/11))
print(paste("d. P(R1 int B2)=",fractions(P_R1_int_B2)))
P_R2_int_B1=(24/(24+64))
print(paste("e. P(R2| B1)=",fractions(P_R2_int_B1)))
P_BB=64/(64+24+24+9)
print(paste("f. P(BB)=",fractions(P_BB)))
P_B2_int_R1=(24/(9+24))
print(paste("g.(B2 | R1)=",fractions(P_B2_int_R1)))
```

R code Exa 3.25 drawing of marbles without replacement

```
#page no:160-161
library(data.tree)
library("prob")
library("dplyr")
library("MASS")
x<-Node$new(" ")
    firstdraw_blue<-x$AddChild("B \n 8/11")
       seconddraw_blue<-firstdraw_blue$AddChild("B \n 7/10")
       seconddraw_red<-firstdraw_blue$AddChild("R \n 3/10")
```

```
            thirddraw_blue<-seconddraw_blue$AddChild("56/110 \n BB")
            thirddraw_red<-seconddraw_red$AddChild("24/110 \n BR")
    firstdraw_red<-x$AddChild("R \n 3/11")
        seconddraw_blue<-firstdraw_red$AddChild("B \n 8/10")
        seconddraw_red<-firstdraw_red$AddChild("R \n 2/10")
             thirddraw_blue<-seconddraw_blue$AddChild("24/110 \n RB")
             thirddraw_red<-seconddraw_red$AddChild("6/110 \n RR")

print(x)
plot(x)
P_RR=(3/11)*(2/10)
print(paste("a. P(RR) =",fractions(P_RR)))
P_RB_U_BR=((3/11)*(8/10))+((8/11)*(3/10))
print(paste("b. P(RR U BR)=",fractions(P_RB_U_BR)))
P_R2_int_B1=(3/10)
print(paste("c.P(R2 | B1)=",fractions(P_R2_int_B1)))
P_R1_int_B2=(3/11)*(8/10)
print(paste("d.P(R1 int B2))=",fractions(P_R1_int_B2)))
P_BB=(8/11)*(7/10)
print(paste("e. P(BB)=",fractions(P_BB)))
P_B2_int_R1=(8/10)
print(paste("f. P(B on 2nd | R on 1st)=",fractions(P_B2_int_R1)))
```

R code Exa 3.26 adoptiong of Tabby kittens and black kittens

```
#pageno 162-163
library(data.tree)
```

```
library("prob")
library("dplyr")
library("MASS")
x<-Node$new(" ")
    firstdraw_blue<-x$AddChild("T \n 4/9")
       seconddraw_blue<-firstdraw_blue$AddChild("T \n 3/8")
       seconddraw_red<-firstdraw_blue$AddChild("B \n 5/8")
           thirddraw_blue<-seconddraw_blue$AddChild("TT")
           thirddraw_red<-seconddraw_red$AddChild("TB")
    firstdraw_red<-x$AddChild("B \n 5/9")
       seconddraw_blue<-firstdraw_red$AddChild("T \n 4/8")
       seconddraw_red<-firstdraw_red$AddChild("B \n 4/8")
            thirddraw_blue<-seconddraw_blue$AddChild("BT")
            thirddraw_red<-seconddraw_red$AddChild("BB")
print(x)
plot(x)
P_TT=(4/9)*(3/8)
print(paste("a. =",fractions(P_TT)))
P_T_B=((4/9)*(5/8))+((5/9)*(4/8))
print(paste("b. =",fractions(P_T_B)))
P_B_T=(4/8)
print(paste("c.=",fractions(P_B_T)))
P_TT_BB=((4/9)*(3/8))+((5/9)*(4/8))
print(paste("d. =",fractions(P_TT_BB)))
```

R code Exa 3.28 drawing venn diagram for flipping of two fair coins

```
#page number : 164−165
library(RAM)
library(VennDiagram)
A=c("TT","TH")
B=c("TT","HT")
AintB=c("TT")
AUB=c("TH","TT","HT")
X=c("HH","HT","TH","TT")
v<-venn.diagram(list(A=A,B=B),
                fill=c("orange","blue"),
                alpha=c(0.5,0.5),cat.cex=1.5,cex
                   =1.5,
                filename=NULL)
grid.newpage()
grid.draw(v)
lapply(v,names)
lapply(v,function(i) i$label)
v[[5]]$label<-paste(setdiff(A,B),collapse="\n")
v[[6]]$label<-paste(setdiff(B,A ),collapse="\n")
v[[7]]$label<-paste(intersect(A,B),collapse="\n")
grid.newpage()
grid.text(x=c(0.5,0),y=c(0.1,0),label=paste(setdiff(
   X,c(A,B)),collapse=" "))
grid.draw(v)
```

R code Exa 3.30 Venn diagram for selection of dogs

```
#page no:166_167
library(eulerr)
library(grid)
VennDiag <- euler(c("C" = 0.40, "PT" = .50, "C&PT" =
    0.05))
p1<-plot(VennDiag, counts = TRUE, font=1, cex=1,
   alpha=0.5,
     fill=c("grey", "blue","red"))
```

```r
grid.text("C int PT",x=0.5,y=0.9)
gridExtra::grid.arrange(p1)
P_C=0.40
P_PT=0.50
print(paste(" P(C)=",P_C))
print(paste(" P(PT)=",P_PT))
CintPT=0.05
print(paste(" P(C\u2229PT)=",CintPT))
C_PT=CintPT/P_PT
print(paste("P(C|PT)=",C_PT))
CUPT=P_C+P_PT-CintPT
print(paste("P(C U PT)=",CUPT))
```

R code Exa 3.31 venn diagram for selection of dogs2

```r
#page no:168_169
library(eulerr)
library(grid)
VennDiag <- euler(c("Male" = 12, "Female" = 8, "Male&Female" = 0.00))
p1<-plot(VennDiag, counts = TRUE, font=1, cex=1, alpha=0.5,
    fill=c("grey", "blue","red"))
VennDiag <- euler(c("Male" = 12, "Brown" = 10, "Male&Brown" = 0.83))
p2<-plot(VennDiag, counts = TRUE, font=1, cex=1, alpha=0.5,
    fill=c("grey", "blue","red"))
grid.text("10",x=0.5,y=0.5)
VennDiag <- euler(c("Female" = 8, "White Fur" = 5, "Female&White Fur" = 0.625))
p3<-plot(VennDiag, counts = TRUE, font=1, cex=1, alpha=0.5,
    fill=c("blue", "white"))
gridExtra::grid.arrange(p1,p2,p3)
```

R code Exa 3.32 probability of selecting student as fresher or earning B grade

```r
#page no:170_172
library(eulerr)
library(grid)
library(multipanelfigure)
S=50
A=20
B=15
sopho=30
P_A=A/S
P_B=B/S
P_AintB=(5/S)
P_AUB=P_A+P_B-P_AintB
VennDiag <- euler(c("A" = A, "B" = B, "A&B" = 0.10))
p1<-plot(VennDiag, counts = TRUE, font=1, cex=1, alpha=0.5,
     fill=c("grey", "lightgray","red"))
VennDiag <- euler(c("A" = A, "B" = B, "A&B" = 0.10))
p2<-plot(VennDiag, counts = TRUE, font=1, cex=1, alpha=0.5,
     fill=c("blue", "lightgray","blue"))
VennDiag <- euler(c("A" = A, "B" = B, "A&B" = 0.10))
p3<-plot(VennDiag, counts = TRUE, font=1, cex=1, alpha=0.5,
         fill=c("lightgray","blue","blue"))
VennDiag <- euler(c("A" = A, "B" = B, "A&B" = 0.10))
p4<-plot(VennDiag, counts = TRUE, font=1, cex=1, alpha=0.5,
         fill=c("lightgray","lightgray","blue"))

gridExtra::grid.arrange(p1,p2,p3,p4,nrow=2)
grid.text("P(AUB)",x=0.25,y=.95)
```

```
grid.text(”=”,x=0.50,y=.75)
grid.text(”P(A)”,x=0.56,y=.95)
grid.text(”+”,x=0.99,y=.75)
grid.text(”P(B)”,x=0.35,y=.50)
grid.text(”−”,x=0.50,y=.25)
grid.text(”P(A\u2229B)”,x=0.75,y=.50)
```

R code Exa 3.33 Two basic rules of Probability example4

```
#page no:172−173
library(MASS)
df=data.frame(LHG=c(5,8),
              HG=c(15,12),
              C=c(40,30),
              CG=c(60,30))
rownames(df)=c(”M”,”F”)
View(df)
df$rtotal=rowSums(df)
ctotal<-c(colSums(df[]))
total=ctotal[5]
PCGintF=df$CG[2]/total
P_F=df$rtotal[2]/total
P_CG=ctotal[4]/total
PCGUF=(P_F+P_CG)-PCGintF
PHGintM=df$HG[1]/total
P_M=df$rtotal[1]/total
P_HG_M=PHGintM/P_M
P_HG=ctotal[2]/total
LHS1=PCGintF
RHS1=P_CG*P_F
print(paste(”a.solution=”,round(PCGintF,2)))
print(paste(”b.solution=”,round(PCGUF,2)))
print(paste(”c.solution=”,round(P_HG_M,3)))
if(LHS1!=RHS1)
{
```

```
  print("d. gender and education are not independent
    ")
}else
{
  print("d. gender and education are independent")
}
LHS2=P_HG_M
RHS2=P_HG
if(LHS2!=RHS2)
{
  print("d. gender and education are not independent
    ")
}else
{
  print("d. gender and education are independent")
}
```

Chapter 4

Discrete Random Variables

R code Exa 4.1 hypergeometric Distribution

```
#page no: 206
A=30
gd=20
x=5
k=10
factorial=function(x)
{
  fact=1
  for(i in 1:x)
  {
    fact=fact*i
  }
return(fact)
  }
binomial=function(a,x)
{
  b=factorial(a)/(factorial(x)*factorial(a-x))
    return(b)
}
term1=binomial(A,x)
term2=binomial(gd,k-x)
```

```
term3=binomial(A+gd,k)
num=term1*term2
P=num/term3
print(paste("1.solution =",round(P,3)))
P1=dhyper(x,A,gd,k)
print(paste("2. solution(buil-in fuction) =",round(
   P1,3)))
```

R code Exa 4.8 Geometric probability distribution function example1

```
#page no: 211-212
library(ggplot2)
p=0.02
x=7
P_7=((1-p)^(x-1))*p
print(paste("a. P(x=7)= ",round(P_7,4)))
m=1/p
print(paste("mean =",round(m,4)))
pp<-0
for(i in 1:51)
{
  pp[i]=((1-p)^(i-1))*p
}
variance=(1/p)*((1/p)-1)
sd=sqrt(variance)
print(paste("the variance=",variance))
print(paste("the sd=",round(sd,1)))
x1=seq(1,51,by=1)
barplot(pp,names.arg=x1,xlab="x",ylab="P(X=x)")
```

R code Exa 4.9 Geometric probability distribution function example2

```
#page no: 212-213
```

```
p=1.28/100
x=9
P_9=((1-p)^x)*p
print(paste("a. P(x=9)= ",round(P_9,4)))
x=20
P_20=((1-p)^(x-1))*p
print(paste("b. P(x=20)= ",round(P_20,2)))
m=(1-p)/p
print(paste("c. mean= ",round(m,2)))
sd=sqrt((1-p)/(p^2))
print(paste("d sd= ",round(sd,2)))
```

R code Exa 4.10 Geometric probability distribution function example3

```
#page no: 213
p=0.320
x=3
P_3=((1-p)^(x-1))*p
print(paste("a. P(x=3)= ",round(P_3,4)))
m=(1/p)
print(paste("c. mean= ",round(m,2)))
```

R code Exa 4.11 Geometric probability distribution function example4

```
#page no: 213
p=0.80
x=3
P_3=((1-p)^(x))*p
print(paste("a. P(x=3)= ",round(P_3,4)))
```

R code Exa 4.14 Poison Probability distribution example1

```
#page no :215-216
library(ggplot2)
mu=(1/8)*6
x=c(0,1,2,3,4,5)
xtot=0
p=mu
for(i in 1:length(x))
{
  xtot=xtot+ppois(x[i],lambda=mu,lower=FALSE)
}
y=ppois(1,lambda=mu,lower.tail=FALSE)
print(paste('P(x>1)',round(y,4)))
pp<-0
for(i in 1:6)
{
  pp[i]=dpois(x[i],lambda=mu)
}
x1=c(0:5)
barplot(pp,names.arg=x1,beside=TRUE,xlab="x",ylab="P
    (X=x)",space=0,ylim=range(pretty(c(0,pp))))
```

R code Exa 4.15 Poison Probability distribution example2

```
#page no :216
mu=7
e=2.718
P_2=dpois(2,lambda=mu)
print(paste('P(x=2)',round(P_2,3)))
x<-c(0,1,2)
y=0
for(i in 1:length(x))
{
  y=y+dpois(x[i],lambda=mu)
```

```
}
print(paste('P(x \u2264 2)',round(y,4)))
sd=sqrt(7)
print(paste('\u03c3 =',round(sd,2)))
```

R code Exa 4.16 Poison Probability distribution example3

```
#page no :216-217
muday=41.5
muhour=muday/24
e=2.718
print(paste('a. solution',round(muhour,4)))
factorial=function(v)
{
  fact=1
  if(v==0)
  {
    return(1)
  }else{
    for(i in 1:v)
  {
    fact=fact*i
  }
  return(fact)
  }
}
x=2
nu=((muhour^x)*(e^(-muhour)))
p2=nu/factorial(x)
print(paste('b. P(x=2)',round(p2,3)))
x=0
nu1=((muhour^x)*(e^(-muhour)))
p0=nu1/factorial(x)
x=1
nu2=((muhour^x)*(e^(-muhour)))
```

```
p1=nu2/factorial(x)
x=2
nu3=((muhour^x)*(e^(-muhour)))
p2=nu3/factorial(x)
px2=1-(p0+p1+p2)
print(paste('c. P(x > 2) ',round(px2,3)))
```

R code Exa 4.17 Poison Probability distribution example4

```
#page no :217
mu=200*0.0102
e=2.718
N=200
r=10
a=dbinom(10,size=200,prob=.0102)
sprintf("a. binomial distribution = %.6f",a)
factorial=function(v)
{
  fact=1
  if(v==0)
  {
    return(1)
  }else{
    for(i in 1:v)
    {
      fact=fact*i
    }
    return(fact)
  }
}

x=10
nu=((mu^x)*(e^(-mu)))
p2=nu/factorial(x)
sprintf('b. poison distribution = %.6f',p2)
```

R code Exa 4.18 Poison Probability distribution example5

```r
#page no :217-218
size=500
prob=0.01
mu=size*prob
x<-c(0,1,2)
binom1=0
for(i in 1:length(x))
{
  v1=dbinom(x[i],size,prob)
 binom1 = binom1 + v1
 print(paste("binomial distribution P(x=",x[i],")=",
    round(v1,5)))
 }
print(paste("binomial distribution P(0,1,2)",round(
   binom1,5)))
app1=1-binom1
print(paste("binomial solution =",round(app1,5)))
poisson1=0
for(i in 1:length(x))
{
  v2=dpois(x[i],mu)
  poisson1 = poisson1 +v2
  print(paste("poisson approx P(x=",x[i],")=",round(
     v2,4)))
}
print(paste("poisson approx P(0,1,2)=",round(
   poisson1,4)))
app2=1-poisson1
print(paste("poisson solution =",round(app2,4)))
```

Chapter 5

Continuous Random Variables

R code Exa 5.1 Continuous Probability Density functions example1

```
#page no: 243-246
library(visualize)
a=20
b=0
x<-seq(b,a,by=1)
fx1=dunif(x,min=0,max=20)
base=(2-0)
height=1/20
area=base*height
p_2=(2-0)*(1/20)
print(paste("a. p(0<x<2)=",round(p_2,4)))
par(mfcol=c(2,2))
visualize.unif(stat=c(0,2),a=0,b=20,section="bounded
   ")
base2=15-4
area2=base2*height
p_15=base2*(1/20)
print(paste("b. p(4 < x < 15)=",round(p_15,4)))
visualize.unif(stat=c(4,15),a=0,b=20,section="
   bounded")
base3=(12.7-2.3)
```

```
p_23_127=base3*height
print(paste("b. p(2.3 < x < 12.7)=",round(p_23_
    127,4)))
visualize.unif(stat=c(2.3,12.7),a=0,b=20,section="
    bounded")
```

R code Exa 5.2 Uniform Distribution example1

```
#page no: 247-248
library(visualize)
a=15
b=0
x<-seq(b,a,by=1)
fx1=dunif(x,min=0,max=12.5)
p_12=(12.5-0)*(1/15)
print(paste("a. p(x<12.5)=",round(p_12,4)))
mu=(a+b)/2
var=((b-a)^2)/12
sd=sqrt(((b-a)^2)/12)
print(paste("b. mean =",mu))
print(paste("b. sd =",round(sd,1)))
k=quantile(x,.90)
print(paste("c. 90th percentile =",k))
visualize.unif(stat=c(0,12.5),a=0,b=15,section="
    bounded")
```

R code Exa 5.3 The Exponential Distribution example1

```
#page no:250
library(ggplot2)
x<-seq(0,20,length.out=100)
mu=4
e=2.718
```

```
m=1/mu
px=dexp(x,rate=m)
df<-data.frame(x=x,y=px)
x=5
p_5=(m*(e^(-m*x)))
print(paste("a.P>5 =",round(p_5,3)))
qplot(x,y,data=df,geom="line")+
  geom_ribbon(data=subset(df,x>=5 ),aes(ymax=y),ymin
     =0,
              fill="red",colour=NA,alpha=0.5)+
  scale_x_continuous(labels=as.character(seq(0:20)),
     breaks=seq(0:20))+
  annotate(geom="text",x=4,y=0.25,label=paste("m
     =0.25"))+
  labs(x= "mu=4", y="f(x)")
```

R code Exa 5.4 The Exponential Distribution example2

```
#page no: 251
library(ggplot2)
e=2.718
mu=0.25
x<-seq(0,20,length.out=100)
px=dexp(x,rate=mu)
df<-data.frame(x=x,y=px)
x1=5
p_5=(1-(e^(-mu*x1)))
x2=4
p_4=(1-(e^(-mu*x2)))
p_4_5=p_5-p_4
qplot(x,y,data=df,geom="line")+
  geom_ribbon(data=subset(df,x>=4 & x<5),aes(ymax=y)
     ,ymin=0,
              fill="red",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq
```

```
      (0:20)),breaks=seq(0:20))+
  annotate(geom="text",x=10,y=0.05,label=paste("P(4<
      x<5)"))+
  labs( y="f(x)")
print(paste("P(x<5) =",round(p_5,4)))
print(paste("P(x<4) =",round(p_4,4)))
print(paste("P(4<x<5) =",round(p_4_5,4)))
```

R code Exa 5.5 The Exponential Distribution example3

```
#page no: 251-253
library(ggplot2)
require(gridExtra)
e=2.718
mu=10
m=1/10
x<-seq(0,15,length.out=200)
px=dexp(x,rate=m)
df<-data.frame(x=x,y=px)
x1=7
p_7=(1-(e^(-m*x1)))
p10=(1-p_7)
p1<-qplot(x,y,data=df,geom="line")+
  geom_ribbon(data=subset(df,x>=7),aes(ymax=y),ymin
      =0.02,
                fill="blue",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq
      (0:20)),breaks=seq(0:20))+
  annotate(geom="text",x=7,y=0.06,label=paste("P(x
      >7)"))+
  labs(x="mu=10", y="f(x)")
print(paste("P(x<7) =",round(p_7,4)))
print(paste("a. P(x > 7) =",round(p10,4)))
print(paste("b. solution =",5*mu," years"))
x2=9
```

```
p_9=(1-(e^(-m*x2)))
x3=11
p_11=(1-(e^(-m*x3)))
p_9_11=p_11-p_9
print(paste("d. P(9 < x < 11)=",round(p_9_11,4)))
p2<-qplot(x,y,data=df,geom="line")+
  geom_ribbon(data=subset(df,x>=9 & x<=11),aes(ymax=
     y),ymin=0.02,
              fill="blue",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq
     (0:20)),breaks=seq(0:20))+
  annotate(geom="text",x=10,y=0.05,label=paste("P(9<
     x<11)"))+
  labs(x="mu=10", y="f(x)")
grid.arrange(p1,p2,ncol=2)
```

R code Exa 5.6 The Exponential Distribution example4

```
#page no : 253
library(MASS)
m=1/12
e=2.718
mu=1/m
sd=12
x=5
p_5=(1-(e^(-m*x)))
p_gt_5=1-p_5
print(paste("a. m =",fractions(m)))
print(paste("b. mu =",fractions(mu)))
print(paste("c. sigma =",fractions(sd)))
print(paste("a. prob. more than 5=",round(p_gt_5,4))
   )
#The answer may slightly vary due to rounding off
   values.
```

R code Exa 5.7 The Exponential Distribution example5

```
#page no: 253-254
library(ggplot2)
require(gridExtra)
e=2.718
avg=30
hour=60
a=hour/avg
print(paste("a. solution=",a,"minutes"))
b=a*3
print(paste("b. solution=",b,"minutes"))
mu=2
m=1/mu
x<-seq(0,20,length.out=200)
px=dexp(x,rate=m)
df<-data.frame(x=x,y=px)
x1=1
p_1=(1-(e^(-m*x1)))
print(paste("c. P(x<1) =",round(p_1,4)))
p1<-qplot(x,y,data=df,geom="line")+
  geom_ribbon(data=subset(df,x<1),aes(ymax=y),ymin
     =0,
              fill="red",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq
     (0:20)),breaks=seq(0:20))+
  annotate(geom="text",x=7,y=0.4,label=paste("P
     =0.3935"))+
  labs(y="f(x)")
x2=5
p_5=(1-(e^(-m*x2)))
p_gt_5=1-p_5
print(paste("d. P(x>5)=",round(p_gt_5,4)))
p2<-qplot(x,y,data=df,geom="line")+
```

```
  geom_ribbon(data=subset(df,x>=5),aes(ymax=y),ymin
     =0,
              fill="red",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq
     (0:20)),breaks=seq(0:20))+
  annotate(geom="text",x=10,y=0.1,label=paste("P(x
     >5)=1-P(x<5)"))+
  labs( y="f(x)")
grid.arrange(p1,p2,ncol=2)
```

R code Exa 5.8 Exponential probability

```
#page no:255
library(ggplot2)
require(gridExtra)
x<-seq(0,20,length.out=100)
e=2.718
mu=4
m=1/mu
px=dexp(x,rate=m)
df<-data.frame(x=x,y=px)
x1=3
p_3=1-(e^(-m*x1))
print(paste("P(x<3)=",round(p_3,4)))
p_gt_3=1-p_3
print(paste("Answer P(x>3)=",round(p_gt_3,4)))
p1<-qplot(x,y,data=df,geom="line")+
  geom_ribbon(data=subset(df,x>3),aes(ymax=y),ymin
     =0,
              fill="red",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq
     (0:20)),breaks=seq(0:20))+
  annotate(geom="text",x=12,y=0.05,label=paste("(P
     >3)=0.4724"))+
  labs( y="f(x)")
```

```
grid.arrange(p1,ncol=1)
```

Chapter 6

The Normal Distribution

R code Exa 6.1 The Standard Normal Distribution example1

```
#page no: 280-281
mu=5
sd=6
x=17
z=(x-mu)/sd
if(z>0)
{
print(paste("x=",x,"is to right of mean"))
}else{
  print(paste("x=",x,"is to left of mean"))
  }
x=1
z=(x-mu)/sd
if(z>0)
{
  print(paste("x=",x,"is to right of mean"))
}else{
  print(paste("x=",x,"is to left of mean"))
}
```

R code Exa 6.2 Finding z scores for given mean and standard deviation

```
#page no: 281
mu=50
sd=6
x=68
x1=mu-sd
x2=mu+sd
z=(x-mu)/sd
z1=(x1-mu)/sd
z2=(x2-mu)/sd
print(paste("for the x value =",x1, "the z score is
   =",z1))
print(paste("for the x value =",x2, "the z score is
   =",z2))
sd1=2*sd
x3=mu-sd1
x4=mu+sd1
z3=(x3-mu)/sd
z4=(x4-mu)/sd
print(paste("for the x value =",x3, "the z score is
   =",z3))
print(paste("for the x value =",x4, "the z score is
   =",z4))
sd2=3*sd
x5=mu-sd2
x6=mu+sd2
z5=(x5-mu)/sd
z6=(x6-mu)/sd
print(paste("for the x value =",x5, "the z score is
   =",z5))
print(paste("for the x value =",x6, "the z score is
   =",z6))
```

R code Exa 6.3 Using the Normal Distribution example1

```r
#page no:284-285
library(ggplot2)
require(gridExtra)
x=65
mu=63
sd=5
p65<-pnorm(x,mean=mu,sd=sd)
pgt65=1-p65
z=(x-mu)/sd
p0=pnorm(63,mean=mu,sd=5)
p63_65=p65-p0
print(paste("a. z1 left area=",round(p63_65,4)))
print(paste("a. P(x>65) area=",round(pgt65,4)))
xx=seq(0,100,length.out=1000)
px=dnorm(xx,mean=mu,sd=sd)
df<-data.frame(x=xx,y=px)
p1<-qplot(x,y,data=df,geom="line")+
  geom_ribbon(data=subset(df, x>=65 ),aes(ymax=y),ymin=0,
              fill="blue",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq(0:100)),breaks=seq(0:100))+
  geom_vline(aes(xintercept=mu),color="blue",linetype="dashed")+
  annotate("text",x=64,y=0.01,size=3,label="0.1554")+
  annotate("text",x=80,y=0.01,size=3,label="0.3446")+
  annotate("text",x=63,y=0.00,size=3,label="mu=63")+
  annotate("text",x=65,y=0.00,size=3,label="x=65")
grid.arrange(p1,ncol=1)
x1=85
```

```
z1=(x1-mu)/sd
if(z1>4.0)
{
  print("b. solution = 1")
}
```

R code Exa 6.4 Using the Normal Distribution example2

```
#page no:285-287
library(ggplot2)
require(gridExtra)
mu=2
sd=0.5
x=1.8
p18<-pnorm(x,mean=mu,sd=sd)
x1=2.75
p275<-pnorm(x1,mean=mu,sd=sd)
p=p275-p18
print(paste("a. P(1.8<=x<=2.75)=",round(p,4)))
fz=sd-0.25
z=qnorm(0.25)
print(paste("a.  z score =",round(z,3)))
xvalue=(z*sd)+mu
print(paste("b. max =",round(xvalue,2),"hours"))
xx=seq(0,10,length.out=100)
px=dnorm(xx,mean=mu,sd=sd)
df<-data.frame(x=xx,y=px)
z1=1.8
z2=2.75
p1<-qplot(x,y,data=df,geom="line")+
  geom_ribbon(data=subset(df, x>=z1 & x<=z2 ),aes(
     ymax=y),ymin=0,
                fill="blue",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq
     (0:10)),breaks=seq(0:10))+
```

```
  geom_vline(aes(xintercept=mu),color="blue",
     linetype="dashed")+
annotate("text",x=2,y=0.00,size=3,label="mu=2")+
  annotate("text",x=1.8,y=-0.005,size=3,label="1.8")
     +
  annotate("text",x=2.75,y=0.00,size=3,label="2.75")
     +
  annotate("text",x=5,y=0.2,size=3,label="sd=0.5")
p2<-qplot(x,y,data=df,geom="line")+
  geom_ribbon(data=subset(df, x<1.66 ),aes(ymax=y),
     ymin=0,
              fill="blue",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq
     (0:10)),breaks=seq(0:10))+
  geom_vline(aes(xintercept=mu),color="blue",
     linetype="dashed")+
  annotate("text",x=0.5,y=0.4,size=3,label="P(x<k)
     =0.25")+
  annotate("text",x=4,y=0.2,size=3,label="P(x>k)
     =0.75")
grid.arrange(p1,p2,nrow=2)
```

R code Exa 6.5 Using the Normal Distribution example3

```
#page no:287
library(ggplot2)
mu=36.9
sd=13.9
x1=23
p23<-pnorm(x1,mean=mu,sd=sd)
x2=64.7
p64<-pnorm(x2,mean=mu,sd=sd)
p=p64-p23
print(paste("a. solution =",round(p,4)))
x3=50.8
```

```
p508=pnorm(x3,mean=mu,sd=sd)
print(paste("b. solution =",round(p508,4)))
```

R code Exa 6.6 Using the Normal Distribution example4

```
#page no:288
library(ggplot2)
mu=5.85
sd=0.24
x1=6.0
p6<-round(pnorm(x1,mean=mu,sd=sd),4)
pgt6=1-p6
z1=(x1-mu)/sd
print(paste("a. P(>=6.0)=",round(pgt6,4)))
print(paste("a. Z1=",round(z1,4)))
f_z=0.20/2
fz1=0.25
fz2=-0.25
x2=(fz1*sd)+mu
x3=(fz2*sd)+mu
print(paste("b. solution = ",x3," and ",x2))
xx=seq(0,7,length.out=100)
px=dnorm(xx,mean=mu,sd=sd)
df<-data.frame(x=xx,y=px)
p1<-qplot(x,y,data=df,geom="line")+
  geom_ribbon(data=subset(df, x>=6.0 & x<=7 ),aes(
     ymax=y),ymin=0,
               fill="blue",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq
     (0:7)),breaks=seq(0:7))+
  geom_vline(aes(xintercept=mu),color="blue",
     linetype="dashed")+
  geom_text(x=7,y=0.5,size=3.5,
             label=paste0("sd= ",sd))+
  annotate("text",x=5.85,y=-0.05,size=3,label="mu
```

```
      =5.85")+
  annotate("text",x=6.0,y=-0.05,size=3,label="6.0")
grid.arrange(p1,ncol=1)
```

R code Exa 6.7 Finding binomial and normal distribution

```
#page no:290-291
library(ggplot2)
library(gridExtra)
mu=10
sd=3
x=16
z=(x-mu)/sd
xx=0:100
n=100
px=dbinom(xx,size=n,prob=0.0228)
df<-data.frame(x=xx,y=px)
z1=15
z2=100
p1<-plot(df$x,df$y,type='h',lty=1,lwd=5,xlab="number
    of sucesses",ylab="P(X)",col=ifelse(df$x<15,'
   grey','blue'))
x_norm=seq(0:100)
p_norm=dnorm(x_norm,mean=mu,sd=sd)
df_norm<-data.frame(x=x_norm,y=p_norm)
p2<-qplot(x,y,data=df_norm,geom="line")+
  geom_ribbon(data=subset(df_norm, x>=15 ),aes(ymax=
     y),ymin=0,
              fill="blue",colour=NA,alpha=0.5)+
  scale_x_continuous("x",labels=as.character(seq
     (0:100)),breaks=seq(0:100))+
  geom_vline(aes(xintercept=mu),color="blue",
     linetype="dashed")
p2
x1=16.0
```

```
p16<-round(pnorm(x1,mean=mu,sd=sd),4)
pgt16=1-p16
print(paste("P(X>16) =",pgt16))
```

Chapter 7

The Central Limit Theorem

R code Exa 7.1 Finite Population Correction Factor example1

```
#page no: 320-321
N=4000
n=100
sd=10.37
mu=75.45
differ1=2
differ2=-2
term1=sd/sqrt(n)
term2=sqrt((N-n)/(N-1))
denom=term1*term2
z=differ1/denom
Z1=differ2/denom
zscore_r=0.4744
fz=zscore_r*differ1
print(paste("f(Z) = ",fz))
```

R code Exa 7.2 Finite Population Correction Factor example2

```
#page no: 321
N=3000
n=360
p=0.06
term1=sqrt((p*(1-p))/n)
term2=sqrt((N-n)/(N-1))
sdp1=term1*term2
p10=10/n
p20=20/n
znum1=(p10-p)
znum2=(p20-p)
z10=znum1/sdp1
z20=znum2/sdp1
zscore10=0.4969
zscore20=0.1480
ans=zscore10-zscore20
print(paste("sigma p' =",round(sdp1,4)))
print(paste("P1' =",round(p10,4)))
print(paste("P2' =",round(p20,4)))
print(paste("Z1' =",round(z10,2)))
print(paste("Z2' =",round(z20,2)))
print(paste("the solution=",round(ans,4)))
```

Chapter 8

Confidence Intervals

R code Exa 8.1 Calculating the Confidence Intervals example1

```
#page no:338
xbar=68
sd=3
n=36
conf_l=0.90
alpha=1-conf_l
zalpha=alpha/2
arearight=zalpha
arealeft=1-zalpha
zscore=round(qnorm(arealeft),3)
al=(1+conf_l)/2
EBM=zscore*(sd/sqrt(n))
interval1=xbar-EBM
interval2=xbar+EBM
print(paste("z score is=",zscore))
print(paste("90% confidence interval= (",interval1,"
    , ",interval2,")"))
```

R code Exa 8.2 Calculating the Confidence Intervals example2

```
#page no:339-340
library(ggplot2)
xbar=68
sd=3
n=36
conf_l=0.95
alpha=1-conf_l
x=seq(67,69,length=n)
y=dnorm(x,xbar,sd)
df=data.frame(x,y)
zalpha=alpha/2
arearight=zalpha
arealeft=1-zalpha
zscore=round(qnorm(arealeft),3)
al=(1+conf_l)/2
EBM=zscore*(sd/sqrt(n))
mu1=xbar-EBM
mu2=xbar+EBM
print(paste("z =",zscore))
print(paste("90% interval= (",mu1," , ",mu2,")"))
par(mfcol=c(2,1))
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=xbar)
abline(v=mu1,lwd=2,col='red')
abline(v=mu2,lwd=2,col='red')
polygon(c(x[x<=mu1],mu1 ),c(y[x<=mu1],0),col="red")
polygon(c(x[x>=mu2],mu2 ),c(y[x>=mu2],0),col="red")
text(mu1,0.126,"67.02")
text(mu2,0.126,"68.98")
text(68,0.126,"68")
x=seq(-3,3,length=100)
y <- dt(x,df=Inf)
df=data.frame(x,y)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=-1.96,lwd=2,col='red')
abline(v=1.96,lwd=2,col='red')
polygon(c(x[x<=-1.96],-1.96 ),c(y[x<=-1.96],0),col="
```

```
red")
polygon(c(x[x>=abs(1.96)],abs(1.96) ),c(y[x>=abs
   (1.96)],0),col="red")
text(-1.96,0,-1.96)
text(1.96,0,1.96)
text(0,0,0)
```

R code Exa 8.3 Calculating the Confidence Intervals example3

```
#page no:340-341
xbar=68
sd=3
n=100
conf_l=0.90
alpha=1-conf_l
zalpha=alpha/2
arearight=zalpha
arealeft=1-zalpha
zscore=round(qnorm(arealeft),3)
al=(1+conf_l)/2
EBM=zscore*(sd/sqrt(n))
interval1=xbar-EBM
interval2=xbar+EBM
print(paste("z score is=",zscore))
print(paste("solution A (",interval1," , ",interval2
   ,")"))
n=25
EBM=zscore*(sd/sqrt(n))
interval1=xbar-EBM
interval2=xbar+EBM
print(paste("Solution B (",interval1," , ",interval2
   ,")"))
```

R code Exa 8.4 Calculating the Confidence Intervals example4

```
#page no: 341-342
n=80
sd=369.34
xbar=593.84
conf_l=0.92
alpha=1-conf_l
zalpha=alpha/2
arearight=zalpha
arealeft=1-zalpha
zscore=round(qnorm(arealeft),2)
v1=round((zscore*(sd/sqrt(n))),2)
mu1=xbar-v1
mu2=xbar+v1
print(paste("z score =",zscore))
print(paste("Answer :(",round(mu1,2)," , ",round(mu2
    ,2),")"))
xseq=seq(mu1,mu2,length=n)
d=dnorm(xseq,xbar,sd)
par(mfcol=c(2,1))
plot(xseq,d,type='l',lwd=3,col='blue',xlab='x')
abline(v=xbar)
abline(v=mu1,lwd=2,col='red')
abline(v=mu2,lwd=2,col='red')
text(mu1,0.001060,"521.58")
text(mu2,0.001060,"666.10")
text(xbar,0.001060,"593.84")
x=seq(-3,3,length=100)
y <- dt(x,df=Inf)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=-1.75,lwd=2,col='red')
abline(v=1.75,lwd=2,col='red')
polygon(c(x[x<=-1.75],-1.75 ),c(y[x<=-1.75],0),col="
    red")
polygon(c(x[x>=abs(1.75)],abs(1.75) ),c(y[x>=abs
    (1.75)],0),col="red")
```

```
text(-1.75,0,-1.75)
text(1.75,0,1.75)
text(0,0,0)
text(-2.7,0.2,expression(frac(alpha,2)))
text(-2.5,0.2,zalpha)
text(2.3,0.2,expression(frac(alpha,2)))
text(2.5,0.2,zalpha)
```

R code Exa 8.5 Confidence interval for a small sample example1

```
#page no: 345-346
n=10
sd=0.395
xbar=1.851
conf_l=0.99
alpha=1-conf_l
zalpha=alpha/2
arearight=zalpha
arealeft=1-zalpha
zscore=round(qnorm(arealeft),2)
tscore=round(qt(arealeft,df=9),4)
v1=round((tscore*(sd/sqrt(n))),3)
mu1=xbar-v1
mu2=xbar+v1
print(paste("t score is=",tscore))
print(paste("Answer :",round(mu1,3)," \u2264 \u00B5
    \u2264" , round(mu2,3)))
x=seq(-4,4,length=100)
y <- dt(x,df=9)
prob1=-3.2498
par(mfcol=c(1,2))
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=prob1,lwd=2,col='red')
abline(v=abs(prob1),lwd=2,col='red')
```

```
polygon(c(x[x<=prob1],prob1 ),c(y[x<=prob1],0),col="red")
polygon(c(x[x>=abs(prob1)],abs(prob1) ),c(y[x>=abs(prob1)],0),col="red")
text(prob1,0,prob1)
text(abs(prob1),0,abs(prob1))
text(0,0,round(0,2))
x=seq(mu1-2,mu2+2,length=100)
y=dnorm(x,xbar,sd)
df=data.frame(x,y)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=xbar)
abline(v=mu1,lwd=2,col='red')
abline(v=mu2,lwd=2,col='red')
polygon(c(x[x<=abs(mu1)],abs(mu1) ),c(y[x<=abs(mu1)],0),col="red")
polygon(c(x[x>=mu2],mu2 ),c(y[x>=mu2],0),col="red")
text(mu1,0,round(mu1,2))
text(mu2,0,round(mu2,2))
text(xbar,0,round(xbar,2))
```

R code Exa 8.6 Confidence interval for a population proportion example1

```
#page no: 347-348
n=500
sample_x=421
p_dash=sample_x/n
q_dash=1-p_dash
conf_l=0.95
alpha=1-conf_l
zalpha=alpha/2
arearight=zalpha
arealeft=1-zalpha
zscore=round(qnorm(arealeft),2)
tscore=round(qt(arealeft,df=Inf),4)
```

```
v1=(zscore*(sqrt((p_dash*q_dash)/n)))
p1=p_dash-v1
p2=p_dash+v1
print(paste("p' =",p_dash))
print(paste("q' =",q_dash))
print(paste("z =",zscore))
print(paste(round(p1,3)," \u2264 p \u2264" ,round(p2
    ,3)))
print(paste("people -cell phone ", round(p1,3)*100,"
    %"," to ", round(p2,3)*100,"%"))
```

R code Exa 8.7 Confidence interval for a population proportion example2

```
#page no: 348-349
n=150
l_limit=0.08
u_limit=0.16
p_dash=(l_limit+u_limit)/2
p=1-p_dash
c_interval=u_limit-p_dash
v1=sqrt((p_dash*(1-p_dash))/n)
zscore=round(c_interval*(1/v1),2)
p_zvalue=0.4345
p_zvalue1=2*p_zvalue
print(paste("p' is=",p_dash))
print(paste("z score is=",zscore))
print(paste("confidence is ", round(p_zvalue1,3)*
    100,"%"))
```

R code Exa 8.8 Confidence interval for a population proportion example3

```
#page no: 349
n=500
```

```
x=300
p_dash=(x/n)
q_dash=1-p_dash
cl=0.90
alpha=1-cl
zalpha=alpha/2
arearight=zalpha
arealeft=1-zalpha
zscore=round(qnorm(arealeft),3)
v1=round((zscore*(sqrt((p_dash*q_dash)/n))),3)
p1=p_dash-v1
p2=p_dash+v1
print(paste("z score is=",zscore))
print(paste("90% confidence :",round(p1,3)," , ",
    round(p2,3)))
print(paste("solution : ", round(p1,3)*100,"%"," to
    ", round(p2,3)*100,"%"))
```

R code Exa 8.9 Calculating the sample size n continuous and binary random variables

```
#page no 351
aer=0.03
cl=0.90
alpha=1-cl
zalpha=alpha/2
arearight=zalpha
arealeft=1-zalpha
zscore=round(qnorm(arealeft),3)
p_dash=0.5
q_dash=1-p_dash
n=((zscore^2)*p_dash*q_dash)/(aer^2)
print(paste(" the sample size is = ",round(n,0)))
```

Chapter 9

Hypothesis Testing with One Sample

R code Exa 9.8 Full Hypothesis test example1

```
#page no 392-394
mu0=16.43
xbar=16
sd=0.8
n=15
t_stat=(xbar-mu0)/(sd/sqrt(n))
alpha=0.05
t_c_left=qnorm(alpha)
p_value=pnorm(t_stat)
print(paste('t static =',round(t_stat,4)))
if(t_stat <= t_c_left)
{
  print(paste("left tail : Reject H0"))
}else
  {
  print(paste("Left tail :Accept H0"))
  }
t_c_right=qnorm(1-alpha)
p_value=1-pnorm(t_stat)
```

```
  if(t_stat >= t_c_right)
  {
    print(paste("right tail : Reject H0"))
  }else
  {
    print(paste("right tail :Accept H0"))
  }
two_t_stat=abs((xbar-mu0))/(sd/sqrt(n))
two_t_critical=qnorm(1-alpha/2)
p_value=2*(1-pnorm(two_t_stat))
if(two_t_stat>=two_t_critical)
{
  print("reject H0")
}else
{
  print("accept H0")
}
dfs<-n-1
x<-seq(-3,3,0.1)
y<-dt(x,dfs)
t.val<-qt(0.95,df=dfs)
p1<-plot(x,y,type='l',lwd=3,col='blue',xlab='x')+
    abline(v=0)+
    abline(v=t.val,lwd=2,col="green")+
    polygon(c(x[x<=-1.76],-2.08),c(y[x<=-1.76],-2.08),col="white")+
    polygon(c(x[x<=-2.08],-2.08),c(y[x<=-2.08],0),col="blue")+
    text(t_stat,0,round(t_stat,2))+
    text(-t.val,0,"-1.76")+
    text(-2.3,0.1,"alpha=")+
    text(-2,0.1,alpha)
```

R code Exa 9.9 Full Hypothesis test example2

```r
#page no 395
library(ggplot2)
mu0=100
xbar=108
sd=12
n=16
dfs=n-1
t_stat=(xbar-mu0)/(sd/sqrt(n))
alpha=0.05
t_c_right=abs(qt(1-alpha,df=dfs))
p_value=1-pt(t_stat,df=dfs)
print(paste("critical value= ",round(t_c_right,3)))
print(paste("t-static =",round(t_stat,2)))
if(t_stat >= t_c_right)
{
  print(paste("right tail : Reject H0"))
}else
{
  print(paste("right tail :Accept H0"))
}
two_t_stat=abs((xbar-mu0))/(sd/sqrt(n))
two_t_critical=qnorm(1-alpha/2)
p_value=2*(1-pnorm(two_t_stat))
if(two_t_stat>=two_t_critical)
{
  print("reject H0")
}else
{
  print("accept H0")
}
dfs <- n-1
x <- seq(-4,4,0.1)
y <- dt(x,dfs)
t.val <- qt(0.95,df=dfs)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t.val,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
```

```
polygon(c(x[x>=1.75],1.75 ),c(y[x>=1.75],0),col="red
   ")
text(t_stat,0,round(t_stat,2))
text(t.val,0,round(t.val,2))
text(3.0,0.1,"alpha=")
text(3.6,0.1,alpha)
```

R code Exa 9.10 Full Hypothesis test example3

```
#page no 396-397
library(ggplot2)
mu0=8
xbar=7.91
sv=.03
sd=sqrt(sv)
n=35
dfs=n-1
t_stat=(xbar-mu0)/(sd/sqrt(n))
alpha=(1-0.99)/2
t_c_left=(qt(alpha,df=Inf))
p_value=pt(t_stat,df=Inf)
print(paste("critical value= ",round(t_c_left,3)))
print(paste("t-static =",round(t_stat,2)))
if(t_stat <= t_c_left)
{
  print(paste("left tail : Reject H0"))
}else
{
  print(paste("left tail :Accept H0"))
}
two_t_stat=abs((xbar-mu0))/(sd/sqrt(n))
two_t_critical=qnorm(1-alpha/2)
p_value=2*(1-pnorm(two_t_stat))
if(two_t_stat>=two_t_critical)
{
```

```
  print("reject H0")
}else
{
  print("accept H0")
}
x <- seq(-4,4,0.1)
y <- dt(x,df=Inf)
t.val <- qt(0.95,df=Inf)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t_c_left,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
abline(v=abs(t_c_left),lwd=2,col='green')
polygon(c(x[x>=abs(t_c_left)],abs(t_c_left) ),c(y[x
   >=abs(t_c_left)],0),col="red")
polygon(c(x[x<=t_c_left],t_c_left ),c(y[x<=t_c_left
   ],0),col="red")
text(t_c_left,0,round(t_c_left,3))
text(t_stat,0,round(t_stat,2))
text(abs(t_c_left),0,round(abs(t_c_left),3))
text(3.0,0.1,"alpha/2=")
text(3.6,0.1,alpha)
text(-3.6,0.1,"alpha/2=")
text(-3.0,0.1,alpha)
```

R code Exa 9.11 Hypothesis test for proportions example1

```
#page no 398-399
library(ggplot2)
n=100
p_dash=53/n
p0=50/n
q0=1-p0
n=100
t_stat=(p_dash-p0)/(sqrt((p0*q0)/n))
```

```
alpha=((5/n))/2
t_c_left=(qt(alpha,df=Inf))
p_value=pt(t_stat,df=Inf)
print(paste("critical value= ",round(t_c_left,2)))
print(paste("t-static =",round(t_stat,2)))
if(t_stat <= t_c_left)
{
  print(paste("left tail : Reject H0"))
}else
{
  print(paste("left tail :Accept H0"))
}
t_c_right=qnorm(1-alpha)
p_value=1-pnorm(t_stat)
if(t_stat >= t_c_right)
{
  print(paste("right tail : Reject H0"))
}else
{
  print(paste("right tail :Accept H0"))
}

two_t_stat=abs((p_dash-p0))/(sqrt((p0*q0)/n))
two_t_critical=qnorm(1-alpha/2)
p_value=2*(1-pnorm(two_t_stat))
if(two_t_stat>=two_t_critical)
{
  print("reject H0")
}else
{
  print("accept H0")
}
x <- seq(-4,4,0.1)
y <- dt(x,df=Inf)
t.val <- qt(0.95,df=Inf)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t_c_left,lwd=2,col='red')
```

```
abline(v=t_stat,lwd=2,col='green')
abline(v=abs(t_c_left),lwd=2,col='green')
polygon(c(x[x>=abs(t_c_left)],abs(t_c_left) ),c(y[x>=abs(t_c_left)],0),col="red")
polygon(c(x[x<=t_c_left],t_c_left ),c(y[x<=t_c_left],0),col="red")
text(t_c_left,0,round(t_c_left,2))
text(t_stat,0,round(t_stat,2))
text(abs(t_c_left),0,round(abs(t_c_left),2))
text(3.0,0.1,"alpha/2=")
text(3.6,0.1,alpha)
text(-3.6,0.1,"alpha/2=")
text(-3.0,0.1,alpha)
```

R code Exa 9.12 Hypothesis test for proportions example2

```
#page no 400
library(ggplot2)
n=150
x=43
p_dash=round(x/n,3)
p0=30/100
q0=1-p0
t_stat=abs((p_dash-p0)/(sqrt((p0*q0)/n)))
alpha=(1-0.90)/2
t_c_left=(qt(alpha,df=Inf))
p_value=pt(t_stat,df=Inf)
print(paste("critical value= ",round(t_c_left,2)))
print(paste("t-static =",round(t_stat,3)))
if(t_stat <= t_c_left)
{
  print(paste("left tail : Reject H0"))
}else
{
  print(paste("left tail :Accept H0"))
```

```
}
t_c_right=qnorm(1-alpha)
p_value=1-pnorm(t_stat)
if(t_stat >= t_c_right)
{
  print(paste("right tail : Reject H0"))
}else
{
  print(paste("right tail :Accept H0"))
}
two_t_stat=abs((p_dash-p0))/(sqrt((p0*q0)/n))
two_t_critical=qnorm(1-alpha/2)
p_value=2*(1-pnorm(two_t_stat))
if(two_t_stat>=two_t_critical)
{
  print("reject H0")
}else
{
  print("accept H0")
}
x <- seq(-4,4,0.1)
y <- dt(x,df=Inf)
t.val <- qt(0.95,df=Inf)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t_c_left,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
abline(v=abs(t_c_left),lwd=2,col='green')
polygon(c(x[x>=abs(t_c_left)],abs(t_c_left) ),c(y[x>=abs(t_c_left)],0),col="red")
polygon(c(x[x<=t_c_left],t_c_left ),c(y[x<=t_c_left],0),col="red")
text(t_c_left,0,round(t_c_left,2))
text(t_stat,0,round(t_stat,3))
text(abs(t_c_left),0,round(abs(t_c_left),2))
text(3.0,0.1,"alpha/2=")
text(3.6,0.1,alpha)
text(-3.6,0.1,"alpha/2=")
```

```
text(-3.0,0.1,alpha)
```

R code Exa 9.13 Hypothesis test for proportions example3

```
#page no 400-401
library(ggplot2)
glass=c
   (1.11,1.07,1.11,1.07,1.12,1.08,0.98,0.98,1.02,0.95,0.95)

n=length(glass)
xbar=mean(glass)
sd=sd(glass)
mu0=1
t_stat=abs((xbar-mu0)/(sd/sqrt(n)))
alpha=(1-0.90)/2
t_c_left=(qt(alpha,df=Inf))
p_value=pt(t_stat,df=Inf)
print(paste("critical value= ",round(t_c_left,2)))
print(paste("t-static =",round(t_stat,2)))
t_c_right=qnorm(1-alpha)
p_value=1-pnorm(t_stat)
if(t_stat >= t_c_right)
{
  print(paste("right tail : Reject H0"))
}else
{
  print(paste("right tail :Accept H0"))
}
two_t_stat=abs((xbar-mu0)/(sd/sqrt(n)))
two_t_critical=qnorm(1-alpha/2)
p_value=2*(1-pnorm(two_t_stat))
if(two_t_stat>=two_t_critical)
{
  print("reject H0")
}else
```

```
{
  print("accept H0")
}
x <- seq(-4,4,0.1)
y <- dt(x,df=Inf)
t.val <- qt(0.95,df=Inf)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t_c_left,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
abline(v=abs(t_c_left),lwd=2,col='green')
polygon(c(x[x>=abs(t_c_left)],abs(t_c_left) ),c(y[x
   >=abs(t_c_left)],0),col="red")
polygon(c(x[x<=t_c_left],t_c_left ),c(y[x<=t_c_left
   ],0),col="red")
text(t_c_left,0,round(t_c_left,2))
text(t_stat,-0.01,round(t_stat,3))
text(abs(t_c_left),0,round(abs(t_c_left),2))
text(3.0,0.1,"alpha/2=")
text(3.6,0.1,alpha)
text(-3.6,0.1,"alpha/2=")
text(-3.0,0.1,alpha)
```

R code Exa 9.14 Hypothesis test for proportions example4

```
#page no 401
library(ggplot2)
n=420019
x=172
p_dash=round(x/n,6)
p0=.0340/100
q0=1-p0
t_stat=abs((p_dash-p0)/(sqrt((p0*q0)/n)))
alpha=.005
t_c_left=(qt(alpha,df=Inf))
```

```
p_value=pt(t_stat,df=Inf)
print(paste("critical value= ",round(t_c_left,2)))
print(paste("t-static =",round(t_stat,2)))
if(t_stat <= t_c_left)
{
  print(paste("left tail : Reject H0"))
}else
{
  print(paste("left tail :Accept H0"))
}
t_c_right=qnorm(1-alpha)
p_value=1-pnorm(t_stat)
if(t_stat >= t_c_right)
{
  print(paste("right tail : Reject H0"))
}else
{
  print(paste("right tail :Accept H0"))
}
two_t_stat=abs((p_dash-p0))/(sqrt((p0*q0)/n))
two_t_critical=qnorm(1-alpha/2)
p_value=2*(1-pnorm(two_t_stat))
if(two_t_stat>=two_t_critical)
{
  print("reject H0")
}else
{
  print("accept H0")
}
x <- seq(-3,3,0.1)
y <- dt(x,df=Inf)
t.val <- qt(0.95,df=Inf)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t_c_left,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
abline(v=abs(t_c_left),lwd=2,col='green')
polygon(c(x[x>=abs(t_c_left)],abs(t_c_left) ),c(y[x
```

```
   >=abs(t_c_left)],0),col="red")
polygon(c(x[x<=t_c_left],t_c_left ),c(y[x<=t_c_left
   ],0),col="red")
text(t_c_left,0,round(t_c_left,2))
text(t_stat,0.05,round(t_stat,3))
text(abs(t_c_left),0,round(abs(t_c_left),2))
text(-3.0,0.1,"alpha/2=")
text(-2.8,0.1,alpha)
text(2.6,0.1,"alpha/2=")
text(2.8,0.1,alpha)
```

Chapter 10

Hypothesis Testing with Two Samples

R code Exa 10.1 Comparing two independent population means example1

```
#page no 422−424
library(ggplot2)
library(MASS)
s1=0.866
s2=1.00
n1=9
n2=16
numerator=(((s1)^2/n1)+((s2)^2/n2))^2
deno1=(1/(n1-1))*(((s1^2)/n1))^2
deno2=(1/(n2-1))*(((s2^2)/n2))^2
df=round(numerator/(deno1+deno2),0)
xbar1=2
xbar2=3.2
diffmu=0
num3=round((xbar1-xbar2)-diffmu,3)
term1=round((s1^2)/n1,3)
term2=round((s2^2)/n2,3)
deno3=round(sqrt(term1+term2),3)
t_stat=num3/deno3
```

```
alpha=.05
twotail_alpha=alpha/2
t_c_left=qt(twotail_alpha,df=df)
p_value=pnorm(t_stat)
if(t_stat <= t_c_left)
{
  print(paste("left tail : Reject H0"))
}else
{
  print(paste("Left tail :Accept H0"))
}
two_t_stat=abs(t_stat)
two_t_critical=qt(twotail_alpha,df=df)
p_value=2*(1-pnorm(two_t_stat))
if(two_t_stat>=two_t_critical)
{
  print("reject H0")
}else
{
  print("accept H0")
}
dfs <- df
x <- seq(-4,4,0.1)
y <- dt(x,dfs)
t.val <- qt(0.025,df=dfs)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t.val,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x>=abs(t_c_left)],abs(t_c_left) ),c(y[x
    >=abs(t_c_left)],0),col="red")
polygon(c(x[x<=t_c_left],t_c_left ),c(y[x<=t_c_left
    ],0),col="red")
text(t_c_left,0,round(t_c_left,2))
text(t_stat,0,round(t_stat,2))
text(abs(t_c_left),0,round(abs(t_c_left),2))
text(3.0,0.1,expression(frac(alpha,2)))
text(3.6,0.1,alpha/2)
```

```
text(-3.6,0.1,expression(frac(alpha,2)))
text(-3.0,0.1,alpha/2)
#The answer may slightly vary due to rounding off
   values
```

R code Exa 10.2 Comparing two independent population means example2

```
#page no 425-426
library(ggplot2)
s1=1.5
n1=11
xbar1=4
s2=1
n2=9
xbar2=3.5
diffmu=0
num3=round((xbar1-xbar2)-diffmu,3)
term1=round((s1^2)/n1,3)
term2=round((s2^2)/n2,3)
deno3=round(sqrt(term1+term2),3)
t_stat=num3/deno3
alpha=.01
df=10
t_c_right=qt(1-alpha,df=df)
p_value=1-pt(t_stat,df=df)
print(paste(" critical value= ",round(t_c_right,3)))
print(paste("g. t-static =",round(t_stat,2)))
if(t_stat >= t_c_right)
{
  print(paste("right tail : Reject H0"))
}else
{
  print(paste("right tail :Accept H0"))
}
dfs <- df
```

```
x <- seq(-4,4,0.1)
y <- dt(x,dfs)
t.val <- qt(1-alpha,df=dfs)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t.val,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x>=abs(t_c_right)],abs(t_c_right) ),c(y[
    x>=abs(t_c_right)],0),col="red")
text(t_c_right,0,round(t_c_right,2))
text(t_stat,0,round(t_stat,2))
text(abs(t_c_right),0,round(abs(t_c_right),2))
text(3.0,0.1,expression(alpha))
text(3.6,0.1,alpha)
text(-3.6,0.1,expression(alpha))
text(-3.0,0.1,alpha)
```

R code Exa 10.3 Comparing two independent population means example3

```
#page no 426-427
library(ggplot2)
s1=16
n1=35
xbar1=74
s2=9
n2=40
xbar2=76
diffmu=0
num3=round((xbar1-xbar2)-diffmu,3)
term1=round((s1^2)/n1,3)
term2=round((s2^2)/n2,3)
deno3=round(sqrt(term1+term2),3)
t_stat=num3/deno3
print(paste("t-static =",round(t_stat,2)))
alpha=0.05
```

```
twotail_alpha=alpha/2
df=n1+n2-2
two_t_stat=abs(num3)/deno3
two_t_critical=qnorm(1-twotail_alpha)
p_value=2*(1-pnorm(t_stat))
if(two_t_stat>=two_t_critical)
{
  print("Reject H0")
}else
{
  print("Accept H0")
}
dfs <- df
x <- seq(-4,4,0.1)
y <- dt(x,dfs)
t.val <- qt(1-alpha,df=dfs)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t.val,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
abline(v=two_t_critical,lwd=2,col='green')
polygon(c(x[x>=abs(two_t_critical)],abs(two_t_critical) ),c(y[x>=abs(two_t_critical)],0),col="red")
text(two_t_critical,0,round(two_t_critical,2))
text(t_stat,0,round(t_stat,2))
text(abs(two_t_critical),0,round(abs(two_t_critical),2))
text(3.0,0.1,expression(frac(alpha,2)))
text(3.6,0.1,twotail_alpha)
text(-3.6,0.1,expression(frac(alpha,2)))
text(-3.0,0.1,twotail_alpha)
```

R code Exa 10.4 Cohens standards for small medium and large effect sizes

```
#page no : 428
xbar1=4
s1=1.5
n1=11
xbar2=3.5
s2=1
n2=9
small=0.2
medium=0.5
large=0.8
num1=xbar1-xbar2
s_pool=sqrt((((n1-1)*s1^2)+((n2-1)*s2^2))/(n1+n2-2))
cohend=(num1/s_pool)
if((cohend>=small) & (cohend<=medium))
{
  print(paste("small- size =",round(cohend,3)))
}
if((cohend>=medium) & (cohend<=large))
{
  print(paste("medium-size =",round(cohend,3)))
}
if(cohend>=large)
{
  print(paste(" large-size is =",round(cohend,3)))
}
```

R code Exa 10.5 Test for differences in means

```
#page no : 428-429
xbar1=8
s1=5.4
n1=18
xbar2=4
s2=2.4
n2=11
```

```
alpha=0.05
num1=xbar1-xbar2
term1=(((n1-1)*s1^2)+((n2-1)*s2^2))/(n1+n2-2)
term2=((1/n1)+(1/n2))
df=n1+n2-2
t_stat=num1/sqrt(term1*term2)
t_critical=qt(alpha,df)
print(paste("t-static =",round(t_stat,2)))
print(paste("t-alpha =",df))
if(t_stat>t_critical)
{
   print("Reject H0")
}else
{
  print("Accept H0")
}
```

R code Exa 10.6 Comparing two independent population proportions example1

```
#page no : 431-432
xa=20
xb=12
na=200
nb=200
pc=round(((xa+xb)/(na+nb)),3)
p=1-pc
print(paste('1-pc =',p))
p_dash_a=xa/na
p_dash_b=xb/nb
ediff=p_dash_a-p_dash_b
print(paste('P\'A =',p_dash_a))
print(paste('P\'B =',p_dash_b))
print(paste('P\'A - P\'B =',ediff))
deno1=(pc*p*((1/na)+(1/nb)))
```

```
t_stat=(ediff/sqrt(deno1))
print(paste('Zc =',round(t_stat,3)))
alpha=.1/2
df=na+nb-2
t_critical=qnorm(1-alpha)
if(t_stat>t_critical)
{
   print("Reject H0")
}else
{
  print("Accept H0")
}
dfs <- na+nb-2
x <- seq(-4,4,0.1)
y <- dt(x,dfs)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t_critical,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x>=abs(t_critical)],abs(t_critical) ),c(
   y[x>=abs(t_critical)],0),col="red")
text(t_critical,0,round(t_critical,2))
text(t_stat,0,round(t_stat,2))
text(abs(t_critical),0,round(abs(t_critical),2))
text(3.0,0.1,expression(frac(alpha,2)))
text(3.6,0.1,alpha)
text(-3.6,0.1,expression(frac(alpha,2)))
text(-3.0,0.1,alpha)
#The answer provided in the textbook is wrong
```

R code Exa 10.7 Two population means with known standard deviations example1

```
#page no : 433-434
x1=3
```

```r
x2=2.9
psd1=0.33
psd2=0.36
n=20
mdiff=x1-x2
sd=sqrt((psd1^2/n)+(psd2^2/n))
t_stat=(mdiff/sd)
print(paste("mean diff= ",x1-x2))
print(paste("t stat= ",round(t_stat,1)))
alpha=.05
df=n+n-2
t_critical=round(qnorm(1-alpha),3)
print(paste("t critical= ",t_critical))
if(t_stat>t_critical)
{
   print("Reject H0")
}else
{
  print("Accept H0")
}
dfs <- n+n-2
x <- seq(-4,4,0.1)
y <- dt(x,dfs)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t_critical,lwd=2,col='red')
abline(v=mdiff,lwd=2,col='green')
polygon(c(x[x>=abs(t_critical)],abs(t_critical) ),c(y[x>=abs(t_critical)],0),col="red")
text(t_critical,0,round(t_critical,2))
text(mdiff,0,round(mdiff,2))
text(abs(t_critical),0,round(abs(t_critical),2))
text(3.0,0.1,expression(alpha))
text(3.6,0.1,alpha)
text(-3.6,0.1,expression(alpha))
text(-3.0,0.1,alpha)
#The answer provided in the textbook is wrong
```

R code Exa 10.8 Two population means with known standard deviations example2

```
#page no : 434-435
x1=61.675
x2=61.704
psd1=10.17
psd2=9.55
n1=30
n2=30
mdiff=abs(x1-x2)
sd=sqrt((psd1^2/60)+(psd2^2/60))
t_stat=((mdiff/sd))
alpha=.05
t_critical=qnorm(1-alpha)
print(paste("t critical=",round(t_critical,2)))
p_value=pnorm(t_stat)
if(t_stat>t_critical)
{
   print("Reject H0")
}else
{
  print("Accept H0")
}
if(alpha>p_value)
{
  print("Reject H0")
}else
{
  print("Accept H0")
}
dfs <- n1+n2-2
x <- seq(-3,3,0.01)
y <- dt(x,dfs)
```

```
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=round(t_stat,2),lwd=2,col='green')
polygon(c(x[x>=abs(t_critical)],abs(t_critical)),c(
    y[x>=abs(t_critical)],0),col="red")
text(t_stat,0,round(t_stat,2))
text(2.0,0.1,expression(alpha))
text(2.3,0.1,alpha)
text(-2.6,0.1,expression(alpha))
text(-2.0,0.1,alpha)
#The answer may slightly vary due to rounding off
    values.
```

R code Exa 10.9 Matched or Paired Samples

```
#page no : 436-437
xd=23.9
mud=20.4
sd=3.8
n=20
t_stat=round(((xd-mud)-0)/(sd/sqrt(n)),2)
alpha=0.1
t_critical1=qt(alpha/2,df=n-2)
t_critical2=qt(1-alpha/2,df=n-2)
p_value=pnorm(t_stat)
print(paste("t_static=",t_stat))
print(paste("t_critical1=",round(t_critical1,3)))
print(paste("t_critical2=",round(t_critical2,3)))
if(t_stat>t_critical2)
{
  print("Reject H0")
}else
{
  print("Accept H0")
}
```

```
dfs <- n-2
x <- seq(-5,5,0.1)
y <- dt(x,dfs)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t_critical2,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x>=abs(t_critical2)],abs(t_critical2) ),
   c(y[x>=abs(t_critical2)],0),col="red")
text(t_critical2,0,round(t_critical2,2))
text(t_stat,0,round(t_stat,2))
text(abs(t_critical2),0,round(abs(t_critical2),2))
text(3.0,0.1,expression(alpha))
text(3.6,0.1,alpha)
text(-3.6,0.1,expression(alpha))
text(-3.0,0.1,alpha)
```

R code Exa 10.10 Matched or Paired Samples example2

```
#page no : 437-439
study<-data.frame(subject=c('A','B','C','D','E','F',
   'G','H'),
                  before=c
                     (6.6,6.5,9.0,10.3,11.3,8.1,6.3,11.6)
                     ,
                  after=c
                     (6.8,2.4,7.4,8.5,8.1,6.1,3.4,2.0)
                     )
study$diff=study$after-study$before
xbar=round(mean(study$diff),3)
study$dbarsq=(study$diff-xbar)^2
View(study)
n=8
sumxbar=round(sum(study$dbarsq),2)
s=round(sqrt(sumxbar/(n-1)),2)
```

```
SE=round(s/sqrt(n),2)
df=n-1
t_stat=round((xbar-0)/SE,2)
alpha=0.05
t_critical=qt(alpha,df=n-1)
p_value=pnorm(t_stat)
if(t_stat<t_critical)
{
  print("Reject H0")
}else
{
  print("Accept H0")
}
dfs <- n-1
x <- seq(-5,5,0.1)
y <- dt(x,dfs)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t_critical,lwd=2,col='red')
abline(v=xbar,lwd=2,col='green')
polygon(c(x[x<=t_critical],t_critical ),c(y[x<=t_
   critical],0),col="red")
text(t_critical,0,round(t_critical,2))
text(xbar,0,round(xbar,2))
text(3.0,0.1,expression(alpha))
text(3.6,0.1,alpha)
text(-3.6,0.1,expression(alpha))
text(-3.0,0.1,alpha)
#The answer may slightly vary due to rounding off
   values.
```

R code Exa 10.11 Matched or Paired Samples example3

```
#page no : 439-440
w_lift<-data.frame(weight=c("play1","play2","play3",
```

```
    "play4"),
                         before=c(205,241,338,368),
                         after=c(295,252,330,360))
w_lift$diff=w_lift$after-w_lift$before
xbar=round(mean(w_lift$diff),3)
w_lift$dbarsq=(w_lift$diff-xbar)^2
View(w_lift)
n=4
sumxbar=round(sum(w_lift$dbarsq),2)
s=round(sqrt(sumxbar/(n-1)),2)
SE=round(s/sqrt(n),2)
df=n-1
t_stat=round((xbar-0)/SE,2)
alpha=0.05
t_critical=qt(1-alpha,df=n-1)
print(paste("t critical= ",round(t_critical,3)))
print(paste("t static= ",round(t_stat,3)))
p_value=pnorm(t_stat)
if(t_stat>t_critical)
{
  print("Reject H0")
}else
{
  print("Accept H0")
}
dfs <- n-1
x <- seq(-5,5,0.1)
y <- dt(x,dfs)
plot(x,y,type='l',lwd=3,col='blue',xlab='x')
abline(v=0)
abline(v=t_critical,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x>=t_critical],t_critical ),c(y[x>=t_
    critical],0),col="red")
text(t_critical,0,round(t_critical,2))
text(t_stat,0,round(t_stat,2))
text(3.0,0.1,expression(alpha))
text(3.6,0.1,alpha)
```

```
text(-3.6,0.1,expression(alpha))
text(-3.0,0.1,alpha)
```

Chapter 11

The Chi Square Distribution

R code Exa 11.2 Test of single variance example1

```
#page no : 467-468
library(ggplot2)
library(MASS)
options(scipen=999)
n=25
df=n-1
sigma=7.2
s_mean=3.5
t_stat=round(((n-1)*(s_mean)^2)/(sigma)^2,2)
alpha=0.05
p=round(pchisq(t_stat,df,lower.tail=TRUE),6)
print(paste0("df=",df))
print(paste0("tstatic=",t_stat))
if(alpha > p)
{
  print(paste("left tail : Reject H0"))
}else
{
  print(paste("Left tail :Accept H0"))
}
t_critical=13.85
```

```
curve(dchisq(x,df=df),from=0,to=100,
      ylab="density",
      lwd=2,
      col='steelblue')
x<-seq(0,100)
y<-dchisq(x,df=df)
abline(v=0)
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x<=abs(t_critical)],abs(t_critical) ),c(
   y[x<=abs(t_critical)],0),col="red")
text(t_stat,0,round(t_stat,2))
text(abs(t_critical),0,round(abs(t_critical),2))
text(60,0.04,expression(alpha))
text(65,0.04,alpha)
mtext(expression(chi^2),side=1,at=5.67,cex=1)
```

R code Exa 11.3 Test of single variance example2

```
#page no : 469
library(ggplot2)
library(MASS)
options(scipen=999)
n=24
df=n-1
sigma=0.2
s_mean=.11
t_stat_num=(n-1)*(s_mean)^2
t_stat_deno=sigma*sigma
t_stat=round(t_stat_num/t_stat_deno,2)
alpha=0.025
p=round(pchisq(t_stat,df),6)
print(paste0(" tstatic=",t_stat))
print(paste0("P(tstatic)=",p))
if(alpha > p)
{
```

```r
  print(paste("left tail : Reject H0"))
}else
{
  print(paste("Left tail :Accept H0"))
}
t_critical1=round(qchisq(alpha,df),2)
t_critical2=round(qchisq(1-alpha,df),2)
if((t_stat<=t_critical1) & (t_stat>=t_critical2))
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
curve(dchisq(x,df=df),from=0,to=100,
      ylab="density",
      lwd=2,
      col='steelblue')
x<-seq(0,100)
y<-dchisq(x,df=df)
abline(v=0)
abline(v=t_stat,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x<=abs(t_critical1)],abs(t_critical1) ),c(y[x<=abs(t_critical1)],0),col="red")
polygon(c(x[x>=abs(t_critical2)],abs(t_critical2) ),c(y[x>=abs(t_critical2)],0),col="red")
text(t_stat,0,round(t_stat,2))
text(t_critical1,0,t_critical1)
text(t_critical2,0,t_critical2)
text(60,0.04,expression(frac(alpha, 2)))
text(65,0.04,alpha)
mtext(expression(chi^2),side=1,at=t_stat,cex=1)
#The answer provided in the textbook is wrong.
```

R code Exa 11.4 Goodness of Fit test example1

```
#apge no: 470-471
students<-data.frame(lterm=c(0,3,6,9),
                     uterm=c(2,5,8,12),
                     E=c(50,30,12,8),
                     O=c(35,40,20,5))
View(students)
no=nrow(students)
df=no-1
print(paste("b. df =",df))
```

R code Exa 11.5 Goodness of Fit test example2

```
#apge no: 472-473
emp<-data.frame(E=c(12,12,12,12,12),
                O=c(15,12,9,9,15))
emp$OE=emp$O-emp$E
emp$OEsq=emp$OE^2
emp$OE2=emp$OEsq/emp$E
View(emp)
no=nrow(emp)
t_stat=sum(emp$OE2)
df=no-1
print(paste("df =",df))
print(paste(" t stat =",t_stat))
alpha=0.05
t_critical=qchisq(1-alpha,df=no-1)
if(t_stat<=t_critical)
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
```

```
curve(dchisq(x,df=df),from=0,to=15,
      ylab="density",
      lwd=2,
      col='steelblue')
x<-seq(0,15,0.1)
y<-dchisq(x,df=df)
abline(v=0)
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x>=t_critical],t_critical ),c(y[x>=t_
   critical],0),col="red")
text(t_stat,0,round(t_stat,2))
text(t_critical,0,round(t_critical,2))
text(60,0.04,expression(alpha))
text(65,0.04,alpha)
mtext(expression(chi^2),side=1,at=3,cex=1)
```

R code Exa 11.6 Goodness of Fit test example3

```
#apge no: 474-476
tv<-data.frame(no_TV=c(0,1,2,3,4),
                    E=c(60,96,330,66,48),
                     O=c(66,119,340,60,15))
tv$OE=tv$O-tv$E
tv$OEsq=tv$OE^2
tv$OE2=tv$OEsq/tv$E
View(tv)
no=nrow(tv)
t_stat=sum(tv$OE2)
df=no-1
print(paste("df =",df))
print(paste("t statistic =",round(t_stat,2)))
alpha=0.01
t_critical=qchisq(1-alpha,df=no-1)
print(paste("critical value =",round(t_critical,3)))
if(t_stat<=t_critical)
```

```
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
curve(dchisq(x,df=df),from=0,to=35,
      ylab="density",
      lwd=2,
      col='steelblue')
x<-seq(0,35,0.1)
y<-dchisq(x,df=df)
abline(v=0)
abline(v=t_critical,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x>=t_critical],t_critical ),c(y[x>=t_
   critical],0),col="blue")
text(20,0.05,expression(alpha ==0.01))
mtext(expression(chi^2),side=1,at=t_stat,cex=1)
mtext(round(t_stat,2),side=1,at=t_stat+2,cex=1)
mtext(round(t_critical,2),side=1,at=t_critical,cex
   =1)
```

R code Exa 11.7 Goodness of Fit test example4

```
#apge no: 476-477
library(GetoptLong)
coin<-data.frame(no_coin=c(0,1,2),
                 E=c(25,50,25),
                  O=c(20,57,23))
coin$OE=coin$O-coin$E
coin$OEsq=coin$OE^2
coin$OE2=coin$OEsq/coin$E
View(coin)
no=nrow(coin)
```

```
t_stat=sum(coin$OE2)
chisquare=t_stat
df=no-1
print(paste("df =",df))
print(paste("chi square =",chisquare))
alpha=0.05
t_critical=qchisq(1-alpha,df=no-1)
if(t_stat<=t_critical)
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
x<-seq(0,15,length=100)
curve(dchisq(x,df=df),from=0,to=15,n=10000,
      ylab="density",
      lwd=2,
      col='steelblue')
y<-dchisq(x,df=df)
abline(v=0)
abline(v=t_critical,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x>=t_critical],t_critical ),c(y[x>=t_critical],0),col="blue")
text(10,0.3,expression(alpha ==0.05))
mtext((expression(chi^2==2.14)),side=1,at=t_stat,cex=1)
mtext(round(t_critical,2),side=1,at=t_critical,cex=1)
```

R code Exa 11.8 Goodness of Fit test example5

```
#page no: 478
library(GetoptLong)
```

```
A=70
B=305
total_pop=755
P_A=A/total_pop
P_B=B/total_pop
P_A_int_B=P_A*P_B
y=P_A_int_B*total_pop
print(paste("answer  =",round(y,digits=1)))
```

R code Exa 11.9 Test of Independence example1

```
#apge no: 479-480
library(MASS)
obs_matrix<-matrix(c(111,96,91,96,133,150,48,61,53),ncol=3)
rownames(obs_matrix)<-c("comm_stud","four_y_stud","non_stud")
colnames(obs_matrix)<-c('hours13','hours46','hours79')
exp_matrix<-matrix(c(90.57,103.00,104.42,115.19,131.00,132.81,49.24,56.00,56.77),ncol=3)
rownames(exp_matrix)<-c("comm_stud","four_y_stud","non_stud")
colnames(exp_matrix)<-c('hours13','hours46','hours79')
no_r=nrow(obs_matrix)
no_c=ncol(obs_matrix)
df=(no_r-1)*(no_c-1)
t_stat=0
for(i in 1:dim(obs_matrix)[1])
    {
    for(j in 1:dim(obs_matrix)[2])
      {
      x=(obs_matrix[i,j]-exp_matrix[i,j])^2/exp_
```

```
            matrix[i,j]
        t_stat=t_stat+x
    }
}
print(paste("critical value",round(t_stat,2)))
print(paste("df =",df))
alpha=0.05
t_critical=qchisq(1-alpha,df=df)
if(t_stat<=t_critical)
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
x<-seq(0,15,length=100)
curve(dchisq(x,df=df),from=0,to=15,
      ylab="density",
      lwd=2,
      col='steelblue')
y<-dchisq(x,df=df)
abline(v=0)
abline(v=t_critical,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x>=t_critical],t_critical ),c(y[x>=t_
   critical],0),col="blue")
text(10,0.1,expression(alpha ==0.05))
mtext((expression(chi^2==12.99)),side=1,at=t_stat,
   cex=1)
mtext(round(t_critical,2),side=1,at=t_critical,cex
   =1)
```

R code Exa 11.10 Test of Independence example2

```
#apge no: 481-482
```

```
library(MASS)
obs_matrix<-matrix(c
  (35,18,4,42,48,5,53,63,11,15,33,15,10,31,17),ncol
  =5)
rownames(obs_matrix)<-c("high","medium","low")
colnames(obs_matrix)<-c('high_anx','med_hig_anx','
  med_anx','med_low_anx','low_anx')
View(obs_matrix)
H_Sum=0
for(i in 1:dim(obs_matrix)[1])
{
  j=1
  H_Sum=H_Sum+obs_matrix[i,j]
}
H_A_sum=0
for(j in 1:dim(obs_matrix)[2])
{
  i=1
  H_A_sum=H_A_sum+obs_matrix[i,j]
}
total=0
for(i in 1:dim(obs_matrix)[1])
  {
  for(j in 1:dim(obs_matrix)[2])
    {
    total=total+obs_matrix[i,j]
  }
}
E1=(H_Sum*H_A_sum)/total
print(paste("a. solution  : ", round(E1,2)))
L_Sum=0
for(i in 1:dim(obs_matrix)[1])
{
  j=4
  L_Sum=L_Sum+obs_matrix[i,j]
}
M_L_A_Sum=0
for(j in 1:dim(obs_matrix)[2])
```

```
{
  i=3
  M_L_A_Sum=M_L_A_Sum+obs_matrix[i,j]
}
E2=(L_Sum*M_L_A_Sum)/total
print(paste("b. solution : ", total))
print(paste("c. solution : ", round(E2,2)))
print(paste("d. solution  : ", round(E2,0)))
```

R code Exa 11.11 Test of Homogeneity

```
#page no: 483-484
library(MASS)
obs_matrix<-matrix(c(72,91,84,86,49,88,45,35),ncol=4)
rownames(obs_matrix)<-c("males","females")
colnames(obs_matrix)<-c('dormitory','apart','withparents','other')
View(obs_matrix)
obs_table<-as.table(obs_matrix)
View(obs_table)
x=chisq.test(obs_table)
t_stat=x$statistic
df=x$parameter
print(paste("\u03C7 = ",round(t_stat,2)))
print(paste("df =",df))
alpha=0.05
t_critical=round(qchisq(1-alpha,df=df),3)
if(t_stat<=t_critical)
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
```

```
print(paste(" critical value ",t_critical))
x<-seq(0,15,length=100)
curve(dchisq(x,df=df),from=0,to=15,
      ylab="density",
      lwd=2,
      col='steelblue')
y<-dchisq(x,df=df)
abline(v=0)
abline(v=t_critical,lwd=2,col='red')
abline(v=t_stat,lwd=2,col='green')
polygon(c(x[x>=t_critical],t_critical ),c(y[x>=t_critical],0),col="blue")
text(10,0.1,expression(alpha ==0.05))
mtext((expression(chi^2==10.13)),side=1,at=t_stat,cex=1)
mtext(t_critical,side=1,at=t_critical,cex=1)
```

Chapter 12

F Distribution and one way ANOVA

R code Exa 12.1 Test of two variances

```
#page no: 515-516
library(MASS)
n1=10
n2=10
s1sq=52.3
s2sq=89.9
f_stat=round((s2sq)/(s1sq),3)
alpha=0.01
df1=n1-1
df2=n2-1
t_critical=round(qf(1- alpha,df1=df1,df2=df2),3)
if(f_stat<=t_critical)
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
print(paste("the critical value ",t_critical))
```

```
print(paste("the f-static ",f_stat))
x<-seq(0,10,length=100)
curve(df(x,df1=df1,df2=df2),from=0,to=10,
      ylab="density",
      lwd=2,
      col='steelblue')
y<-df(x,df1=df1,df2=df2)
abline(v=0)
abline(v=t_critical,lwd=2,col='red')
abline(v=f_stat,lwd=2,col='green')
polygon(c(x[x>=t_critical],t_critical ),c(y[x>=t_
   critical],0),col="blue")
text(9,0.1,expression(alpha ==0.01))
mtext((expression(F[c]==1.719)),side=1,at=f_stat,cex
   =1)
mtext(round(t_critical,2),side=1,at=t_critical,cex
   =1)
```

R code Exa 12.2 The F distribution and the F Ratio example1

```
#page no : 519-521
plan1<-c(5,4.5,4,3)
plan2<-c(3.5,7,4.5)
plan3<-c(8,4,3.5)
plan4<-c(plan1,plan2,plan3)
n<-c(length(plan1),length(plan2),length(plan3))
s<-c(sum(plan1),sum(plan2),sum(plan3))
term1=0
term2=0
n_sum=0
n_count=length(n)
for(i in 1:n_count)
{
  term1=term1+(s[i]^2)/n[i]
  term2=term2+s[i]
```

```
  n_sum=n_sum+n[i]
}
ss_between=term1-((term2)^2/n_sum)
t_term1=0
t_term2=0
for(i in 1:length(plan4))
{
t_term1=t_term1+plan4[i]^2
t_term2=t_term2+plan4[i]
}
s_total=t_term1-((t_term2)^2/n_sum)
ss_within=s_total-ss_between
df_between=n_count-1
df_within=n_sum-n_count
df=n_sum-1
ms_between=ss_between/df_between
ms_within=ss_within/df_within
f_stat=ms_between/ms_within
print(paste("ss_between= ",round(ss_between,4)))
print(paste("ss_within= ",round(ss_within,4)))
print(paste("ss_total= ",round(s_total,1)))
print(paste("ms_between= ",round(ms_between,4)))
print(paste("ms_within= ",round(ms_within,4)))
print(paste("f_stat= ",round(f_stat,4)))
```

R code Exa 12.3 The F distribution and the F Ratio example2

```
#page no : 521-522
v1<-c(2625,2997,4915)
v2<-c(5348,5682,5482)
v3<-c(6583,8560,3830)
v4<-c(7285,6897,9230)
v5<-c( 6277,7818,8677)
v6<-c(v1,v2,v3,v4,v5)
n<-c(length(v1),length(v2),length(v3),length(v4),
```

```
    length(v5))
s<-c(sum(v1),sum(v2),sum(v3),sum(v4),sum(v5))
term1=0
term2=0
n_sum=0
n_count=length(n)
for(i in 1:n_count)
{
  term1=term1+(s[i]^2)/n[i]
  term2=term2+s[i]
  n_sum=n_sum+n[i]
}
ss_between=term1-((term2)^2/n_sum)
t_term1=0
t_term2=0
for(i in 1:length(v6))
{
t_term1=t_term1+v6[i]^2
t_term2=t_term2+v6[i]
}
s_total=t_term1-((t_term2)^2/n_sum)
ss_within=s_total-ss_between
df_between=n_count-1
df_within=n_sum-n_count
df=n_sum-1
ms_between=ss_between/df_between
ms_within=ss_within/df_within
f_stat=ms_between/ms_within
print(paste("ss_b= ",round(ss_between,0)))
print(paste("ss_w= ",round(ss_within,0)))
print(paste("ss_t= ",round(s_total,0)))
print(paste("ms_b= ",round(ms_between,0)))
print(paste("ms_w= ",round(ms_within,1)))
print(paste("f_stat= ",round(f_stat,4)))
alpha=0.05
df1=df_between
df2=df_within
p_value=round(pf(f_stat,df1=df1,df2=df2,lower.tail=
```

```
    FALSE),4)
print(paste("P(F>4.481)= ",p_value))
if(f_stat<=p_value)
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
if(alpha<=p_value)
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
x<-seq(0,5,length=100)
curve(df(x,df1=df1,df2=df2),from=0,to=5,
      ylab="density",
      lwd=2,
      col='steelblue')
y<-df(x,df1=df1,df2=df2)
abline(v=0)
abline(v=f_stat,lwd=2,col='green')
text(4,0.1,expression(alpha ==0.05))
mtext((expression(F==4.481)),side=1,at=f_stat,cex=1)
```

R code Exa 12.4 The F distribution and the F Ratio example3

```
#page no : 523-525
v1<-c(2.17,1.85,2.83,1.69,3.33)
v2<-c(2.63,1.77,3.25,1.86,2.21)
v3<-c(2.63,3.78,4.00,2.55,2.45)
v4<-c(3.79,3.45,3.08,2.26,3.18)
v5<-c(v1,v2,v3,v4)
```

```r
n<-c(length(v1),length(v2),length(v3),length(v4))
s<-c(sum(v1),sum(v2),sum(v3),sum(v4))
term1=0
term2=0
n_sum=0
n_count=length(n)
for(i in 1:n_count)
{
  term1=term1+(s[i]^2)/n[i]
  term2=term2+s[i]
  n_sum=n_sum+n[i]
}
ss_between=term1-((term2)^2/n_sum)
t_term1=0
t_term2=0
for(i in 1:length(v5))
{
t_term1=t_term1+v5[i]^2
t_term2=t_term2+v5[i]
}
s_total=t_term1-((t_term2)^2/n_sum)
ss_within=s_total-ss_between
df_between=n_count-1
df_within=n_sum-n_count
df=n_sum-1
ms_between=ss_between/df_between
ms_within=ss_within/df_within
f_stat=ms_between/ms_within
print(paste("ss_b= ",round(ss_between,0)))
print(paste("ss_w= ",round(ss_within,0)))
print(paste("ss_t= ",round(s_total,0)))
print(paste("ms_b= ",round(ms_between,0)))
print(paste("ms_w= ",round(ms_within,1)))
print(paste("f_stat= ",round(f_stat,2)))
alpha=0.01
df1=df_between
df2=df_within
p_value=round(pf(f_stat,df1=df1,df2=df2,lower.tail=
```

```
    FALSE),4)
print(paste("P(F>2.23) = ",round(p_value,4)))
if(alpha<=p_value)
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
x<-seq(0,5,length=100)
curve(df(x,df1=df1,df2=df2),from=0,to=5,
      ylab="density",
      lwd=2,
      col='steelblue')
y<-df(x,df1=df1,df2=df2)
abline(v=0)
abline(v=f_stat,lwd=2,col='green')
polygon(c(x[x>=f_stat],f_stat ),c(y[x>=f_stat],0),
    col="blue")
text(4,0.1,expression(italic(p) ==0.1241))
mtext((expression(F==2.23)),side=1,at=f_stat,cex=1)
```

R code Exa 12.5 The F distribution and the F Ratio example4

```
#page no : 525-526
v1<-c(24,21,23,30,23)
v2<-c(25,31,23,20,28)
v3<-c(23,27,22,30,20)
v4<-c(v1,v2,v3)
n<-c(length(v1),length(v2),length(v3))
s<-c(sum(v1),sum(v2),sum(v3))
n_group=length(n)
SM_v1=mean(v1)
SM_v2=mean(v2)
SM_v3=mean(v3)
```

```
SV_v1=var(v1)
SV_v2=var(v2)
SV_v3=var(v3)
v5=c(SM_v1,SM_v2,SM_v3)
VG_Means=var(v5)
n=length(v1)
ms_between=n*VG_Means
v6=c(SV_v1,SV_v2,SV_v3)
s_pooled=mean(v6)
ms_within=s_pooled
f_stat=round(ms_between/ms_within,3)
df_num=n_group-1
df_deno=length(v4)-n_group
print(paste("variance= ",round(VG_Means,3)))
print(paste("mean = ",round(s_pooled,3)))
print(paste("ms_b= ",round(ms_between,3)))
print(paste("ms_w= ",round(ms_within,3)))
print(paste("f_stat= ",round(f_stat,4)))
alpha=0.03
df1=df_num
df2=df_deno
p_value=round(pf(f_stat,df1=df1,df2=df2,lower.tail=
   FALSE),4)
print(paste("P(F>0.134) = ",round(p_value,4)))
if(alpha<=p_value)
{
  print("Accept H0")
}else
{
  print("Reject H0")
}
```

Chapter 13

Linear Regression and Correlation

R code Exa 13.2 Linear Equations example1

```
#page no : 556
x<-seq(-5,20,length=100)
y=-1+2*x
plot(x,y,type="l",ylab="y=-1+2x")
```

R code Exa 13.5 The Regression Equations example1

```
#page no: 571
x<-c(65,67,71,71,66,75,67,70,71,69,69)
y<-c(175,133,185,163,126,198,153,163,159,151,159)
plot(x,y,xlab="3rd exam",ylab="final exam")
```

R code Exa 13.6 The Regression Equations example2

```r
#page no: 576
x<-c(65,67,71,71,66,75,67,70,71,69,69)
y<-c(175,133,185,163,126,198,153,163,159,151,159)
ypredict=lm(y~x,data=faithful)
coeffs=coefficients(ypredict)
score1=66
eq1=round(coeffs[1],2)+round(coeffs[2],2)*score1
print(paste("a. solution =",eq1))
score2=90
eq2=round(coeffs[1],2)+round(coeffs[2],2)*score2
print(paste("b. solution =",eq2))
```