R Textbook Companion for Introductory Statistics by Douglas S Shafer and Zhiyi Zhang[1]

Created by
Ananya K R
B.Tech
Electronics Engineering
Vellore Institute of Technology
Cross-Checked by
R TBC Team

August 11, 2020

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

Book Description

Title: Introductory Statistics

Author: Douglas S Shafer and Zhiyi Zhang

Publisher: Flat World Publishers

Edition: 1

Year: 2012

ISBN: 978-1-4533-4486-6

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

Introduction

R code Exa 1.1 Set Notation

```
#Page 18
age_of_students<-c
    (18,18,19,19,19,18,22,20,18,18,17,19,18,24,18,20,18,21,20,17,19)

print(age_of_students)
```

R code Exa 1.2 Data Frequency Table

```
#Page 19
dataset<-c
    (18,18,19,19,19,18,22,20,18,18,17,19,18,24,18,20,18,21,20,17,19)

library("plyr")
y<-count(dataset)
print(y)
```

Chapter 2

Descriptive Statistics

R code Exa 2.1 Operations on the Data Set

```r
#Page 39
data<-c(1,3,4)
a<-sum(data)
print(a)
b<-sum(data^2)
print(b)
c<-sum((data-1)^2)
print(c)
```

R code Exa 2.2 Mean of Sample Data

```r
#Page 39
dataset<-c(2,-1,0,2)
mean<-sum(dataset)/length(dataset)
print(mean)
```

R code Exa 2.3 Sample Mean of Students GPA

```
#Page 40
GPA<-c
   (1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33)

mean<-mean(GPA)
print(mean)
```

R code Exa 2.4 Measures of Variability

```
#Page 41
dataset<-c(rep(0,3),rep(1,6),rep(2,6),rep(3,3),rep
   (4,1))
library("plyr")
y<-count(dataset)
print(y)
print(dataset)
mean<-sum(dataset)/length(dataset)
print(mean)
```

R code Exa 2.5 Sample Median of Data Set

```
#Page 44
dataset<-c(2,-1,0,2)
median<-median(dataset)
print(median)
```

R code Exa 2.6 Sample Median of Students GPA

```
#Page 45
dataset<-c
   (1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33)

median<-median(dataset)
print(median)
```

R code Exa 2.7 Sample Median of Frequency Distribution Table

```
#Page 45
dataset<-c(rep(0,3),rep(1,6),rep(2,6),rep(3,3),rep
   (4,1))
median<-median(dataset)
print(median)
```

R code Exa 2.8 Sample Mode of Data Set

```
#Page 48
v<-c(-1,0,2,0)
library("modeest")
print(mfv(v))
```

R code Exa 2.9 Multiple Modes in Data Set

```
#Page 48
v<-c(rep(0,3),rep(1,6),rep(2,6),rep(3,3),rep(4,1))
library("modeest")
print(mfv(v))
```

R code Exa 2.10 Range of Data Set

```
#Page 58
dataset_1<-c(40,38,42,40,39,39,43,40,39,40)
dataset_2<-c(46,37,40,33,42,36,40,47,34,45)
findrange=function(v){
  range=max(v)-min(v)
  print(range)
}
findrange(dataset_1)
findrange(dataset_2)
```

R code Exa 2.11 Standard Deviation and Variance in a Data Set

```
#Page 60
dataset_2<-c(46,37,40,33,42,36,40,47,34,45)
findvariance<-function(v){
  mean<-sum(v)/length(v)
  x<-sum((v-mean)^2)
  variance<-x/(length(v)-1)
  print(variance)
}
findvariance(dataset_2)
std_deviation=sd(dataset_2)
print(std_deviation)
```

R code Exa 2.12 Standard Deviation and Variance of Students GPA

```
#Page 61
```

```
GPA<-c
   (1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33)
findvariance<-function(v){
  mean<-sum(v)/length(v)
  x<-sum((v-mean)^2)
  variance<-x/(length(v)-1)
  print(variance)
}
findvariance(GPA)
std_deviation=sd(GPA)
print(std_deviation)
```

R code Exa 2.13 Measures of Central Location

```
#Page 70
GPA<-c
   (1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33)
k<-sort(GPA)
print(k)
index<-match(c(3.33),k)
print(index)
percentile=(index/length(GPA))*100
print(percentile)
```

R code Exa 2.14 Quartiles in a Data Set

```
#Page 72
GPA<-c
   (1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33)
quartile_2<-median(GPA)
```

```
lower_subset<-print(GPA[GPA<median(GPA)])
upper_subset<-print(GPA[GPA>median(GPA)])
quartile_1<-median(lower_subset)
quartile_3<-median(upper_subset)
print(quartile_1)
print(quartile_2)
print(quartile_3)
```

R code Exa 2.15 Quartiles of Students GPA

```
#Page 73
GPA<-c(1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33,3.88)
quartile_2<-median(GPA)
lower_subset<-print(GPA[GPA<median(GPA)])
upper_subset<-print(GPA[GPA>median(GPA)])
quartile_1<-median(lower_subset)
quartile_3<-median(upper_subset)
print(quartile_1)
print(quartile_2)
print(quartile_3)
```

R code Exa 2.16 Box Plots and Interquartile Range

```
#Page 75
GPA<-c(1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33)
quartile_2<-median(GPA)
lower_subset<-GPA[GPA<median(GPA)]
upper_subset<-GPA[GPA>median(GPA)]
quartile_1<-median(lower_subset)
```

```
quartile_3<-median(upper_subset)
print(min(GPA))
print(quartile_1)
print(quartile_2)
print(quartile_3)
print(max(GPA))
boxplot(GPA)
IQR<-print(quartile_3-quartile_1)
```

R code Exa 2.17 Z Scores of Observations

```
#Page 78
data<-c
    (1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33)

mean<-mean(data)
sd<-sd(data)
zscores<-function(v){
  for(x in v){
    z<-((x-mean(data))/sd(data))
    print(z)
  }
}
zscores(data)

#The answer may slightly vary due to rounding off
    values.
```

R code Exa 2.18 Computing Data from given Standard Deviation and Variance

```
#Page 79
mean<-2.70
```

```
sd<-0.50
z1<--0.62
z2<-1.28

findGPA<-function(z){
  GPA<-mean+(sd)*(z)
  print(GPA)
}

findGPA(z1)
findGPA(z2)
```

Chapter 3

Basic Concepts of Probability

R code Exa 3.1 Sample Space for Tossing a Coin

```
#Page 112
Sample_space<-c('Heads','Tails')
library("sets")
print(as.set(Sample_space))
```

R code Exa 3.2 Sample Space for Rolling a Die and a few Cases

```
#Page 112
Sample_space<-c(1,2,3,4,5,6)
library("sets")
print(as.set(Sample_space))
event_1<-print(as.set(sort(Sample_space[Sample_space
   %%2==0])))
event_2<-print(as.set(sort(Sample_space[Sample_space
   >2])))
```

R code Exa 3.3.a Tossing two coins sample space

```
#Page 113
library("sets")
Sample_space_a<-print(as.set(c('hh','tt','d')))
```

R code Exa 3.3.b Tossing two Distinguishable coins

```
#Page 113
library("sets")
Sample_space_b<-print(as.set(c('hh','tt','ht','th'))
  )
```

R code Exa 3.4 Sample Space that describes all Three Child Families

```
#Page 114
sample_space<-c('b','g')
k<-expand.grid(sample_space,sample_space,sample_
  space)
print(k)

library("DiagrammeR")

nodes <- create_node_df(n=15,type = "number",
                        label = c("P","b","g","b","g
                          ","b","g","b","g","b","g"
                          ,"b","g","b","g"))

edges <- create_edge_df(from = c
  (1,1,2,2,3,3,4,4,5,5,6,6,7,7),
                        to = c
                          (2,3,4,5,6,7,8,9,10,11,12,13,14,15)
                          ,
```

```
                          rel = "leading to")

graph <- create_graph(nodes_df = nodes,
                      edges_df = edges,
                      attr_theme="lr")

print(render_graph(graph))
```

R code Exa 3.5 Probability for Tossing a Coin

```
#Page 116
sample_space<-c('H','T')
event_a<-c('H')
event_b<-c('T')
probablility_of_a<-(length(event_a)/length(sample_space))
probablility_of_b<-(length(event_b)/length(sample_space))
print(probablility_of_a)
print(probablility_of_b)
```

R code Exa 3.6 Probability of tossing a Single Die and a few cases

```
#Page 117
library("sets")
sample_space<-c(1,2,3,4,5,6)
Event_H<-print(as.set(sample_space[sample_space%%2==0]))
probability_of_H<-print(length(Event_H)/length(sample_space))
Event_T<-print(as.set(sample_space[sample_space>2]))
probability_of_T<-print(length(Event_T)/length(sample_space))
```

R code Exa 3.7 Probability of two tossed coins matching

```
#Page 118
library("sets")
both_heads<-c('hh')
both_tails<-c('tt')
head_and_tail<-c('ht')
tail_and_head<-c('th')
Sample_space<-as.set(c('hh','tt','ht','th'))
Probablity_of_both_heads<-print(length(both_heads)/
   length(Sample_space))
Probability_of_both_tails<-print(length(both_tails)/
   length(Sample_space))
Overall_Probability<-print(Probablity_of_both_heads+
   Probability_of_both_tails)
```

R code Exa 3.8.a Probability of occurrence of Event B

```
#Page 119
probability_of_white<-51/100
probability_of_black<-27/100
probability_of_hispanic<-11/10
probability_of_asian<-6/100
probability_of_others<-5/100
print(probability_of_black)
```

R code Exa 3.8.b Probability of occurrence of Event M

```
#Page 119
```

```
probability_of_white<-51/100
probability_of_black<-27/100
probability_of_hispanic<-11/10
probability_of_asian<-6/100
probability_of_others<-5/100
print(1-probability_of_white)
```

R code Exa 3.8.c Probability of occurrence of Event N

```
#Page 119
probability_of_white<-51/100
probability_of_black<-27/100
probability_of_hispanic<-11/10
probability_of_asian<-6/100
probability_of_others<-5/100
print(1-probability_of_black)
```

R code Exa 3.9.a Probability of occurrence of Event B

```
#Page 120
probability_of_white_male<-25/100
probability_of_white_female<-26/100
probability_of_black_male<-12/100
probability_of_black_female<-15/100
probability_of_hispanic_male<-6/10
probability_of_hispanic_female<-5/100
probability_of_asian_male<-3/100
probability_of_asian_female<-3/100
probability_of_others_male<-1/100
probability_of_others_female<-4/100

probability_of_black<-print(probability_of_black_male+probability_of_black_female)
```

R code Exa 3.9.b Probability of occurrence of Event MF

```
#Page 120
probability_of_white_male<-25/100
probability_of_white_female<-26/100
probability_of_black_male<-12/100
probability_of_black_female<-15/100
probability_of_hispanic_male<-6/10
probability_of_hispanic_female<-5/100
probability_of_asian_male<-3/100
probability_of_asian_female<-3/100
probability_of_others_male<-1/100
probability_of_others_female<-4/100

probability_of_minority_female<-print(probability_of
    _black_female+probability_of_hispanic_female+
    probability_of_asian_female+probability_of_others
    _female)
```

R code Exa 3.9.c Probability of occurrence of Event FN

```
#Page 120
probability_of_white_male<-25/100
probability_of_white_female<-26/100
probability_of_black_male<-12/100
probability_of_black_female<-15/100
probability_of_hispanic_male<-6/10
probability_of_hispanic_female<-5/100
probability_of_asian_male<-3/100
probability_of_asian_female<-3/100
probability_of_others_male<-1/100
```

```
probability_of_others_female<-4/100

probability_of_minority_female<-print(probability_of
    _white_female+probability_of_hispanic_female+
    probability_of_asian_female+probability_of_others
    _female)
```

R code Exa 3.10 Events concerned with the experiment of rolling a single die

```
#Page 131
library("sets")
sample_space<-c(1,2,3,4,5,6)
event_E<-sample_space[sample_space%%2==0]
print(as.set(event_E))
complement_of_E<-setdiff(sample_space,event_E)
print(as.set(complement_of_E))
event_T<-sample_space[sample_space>2]
print(as.set(event_T))
complement_of_T<-setdiff(sample_space,event_T)
print(as.set(complement_of_T))
```

R code Exa 3.11 Probability Calculation

```
#Page 132
Event_0<-c('tails','tails','tails','tails','tails')
print(Event_0)
number_of_tosses<-5
possible_outcomes<-2
Total_outcomes<-possible_outcomes^number_of_tosses
Probability_of_all_tails<-1/Total_outcomes
print(Probability_of_all_tails)
```

```
Probability_of_atleast_one_head<-1-Probability_of_all_tails
print(Probability_of_atleast_one_head)
print(Probability_of_atleast_one_head*100)

#The answer may slightly vary due to rounding off values.
```

R code Exa 3.12 Finding the Intersection of two Events

```
#Page 133
library("sets")
sample_space<-c(1,2,3,4,5,6)
event_E<-sample_space[sample_space%%2==0]
print(as.set(event_E))
event_T<-sample_space[sample_space>2]
print(as.set(event_T))
intersection<-print(as.set(intersect(event_E,event_T)))
```

R code Exa 3.13.a Intersection of two events while rolling a fair die

```
#Page 134
library("sets")
sample_space<-c(1,2,3,4,5,6)
event_E<-sample_space[sample_space%%2==0]
probability_of_E<-length(event_E)/length(sample_space)
event_T<-sample_space[sample_space>2]
probability_of_T<-length(event_T)/length(sample_space)
intersection<-print(as.set(intersect(event_E,event_T)))
```

```
Probability_of_both_events_occuring<-(length(
   intersection)/length(sample_space))
print(Probability_of_both_events_occuring)
```

R code Exa 3.13.b Probability of intersection of two events

```
#Page 134
probability_of_1<-1/12
probability_of_6<-3/12
Sum_of_other_probabilities<-1-(probability_of_1+
   probability_of_6)
probability_of_2<-Sum_of_other_probabilities/4
probability_of_3<-Sum_of_other_probabilities/4
probability_of_4<-Sum_of_other_probabilities/4
probability_of_5<-Sum_of_other_probabilities/4
probability_of_both_events_occuring<-print(
   probability_of_4+probability_of_6)
```

R code Exa 3.14 Complement of an Event

```
#Page 135
library("sets")
library("gtools")
set<-as.set(c(1,2,3,4,5,6))
e<-c(2,4,6)
E<-as.set(e)
k<-set_complement(E,set)
possible_set_1<-print(combinations(n=3,r=3,v=c
   (1,3,5),repeats.allowed=F))
possible_set_2<-print(combinations(n=3,r=2,v=c
   (1,3,5),repeats.allowed=F))
possible_set_3<-print(combinations(n=3,r=1,v=c
   (1,3,5),repeats.allowed=F))
```

R code Exa 3.15 Finding the Union of two Events

```
#Page 137
library("sets")
sample_space<-c(1,2,3,4,5,6)
event_E<-sample_space[sample_space%%2==0]
print(as.set(event_E))
event_T<-sample_space[sample_space>2]
print(as.set(event_T))
union<-print(as.set(union(event_E,event_T)))
```

R code Exa 3.16 Finding the Union and Intersection of two Events

```
#Page 138
Sample_space<-c('bb','bg','gb','gg')
Sample_B<-c('bb','bg','gb')
Sample_D<-c('bg','gb')
Sample_M<-c('bb','gg')
B_union_D<-print(union(Sample_B,Sample_D))
B_union_M<-print(union(Sample_B,Sample_M))
```

R code Exa 3.17.a Probability that both dice show a 4 when two dice are shown

```
#Page 139
library("prob")
sample_space<-rolldie(2)
print(sample_space)
event<-expand.grid(x=4,y=4)
```

```
print(event)
probability<-(nrow(event)/nrow(sample_space))
print(probability)
```

R code Exa 3.17.b Finding probability of at least one die showing a 4 when two dice are thrown

```
#Page 139
library("prob")
sample_space<-rolldie(2)
print(sample_space)
Total_possibilitites<-nrow(sample_space)
num_Event_A<-6
num_Event_B<-6
num_intersection<-1
P_Event_A<-print(num_Event_A/Total_possibilitites)
P_Event_B<-print(num_Event_B/Total_possibilitites)
P_Intersection<-print(num_intersection/Total_possibilitites)
reqd_probability<-print((P_Event_A+P_Event_B)-P_Intersection)
```

R code Exa 3.18 Intersections and Unions of Probabilities

```
#page 140
Prob_of_maths<-63/100
Prob_of_eng<-34/100
Prob_of_eng_and_math<-27/100
Prob_of_eng_or_math<-print((Prob_of_maths+Prob_of_eng)-Prob_of_eng_and_math)
```

R code Exa 3.19.a Conditional Probability

```
#Page 141
c_s<-12
c_t<-1
c_total<-c_s+c_t
e_s<-4
e_t<-3
e_total<-e_s+e_t
m_s<-6
m_t<-2
m_total<-m_s+m_t
total_s<-c_s+e_s+m_s
total_t<-c_t+e_t+m_t
grand_total<-c_total+e_total+m_total
given_data<-matrix(c(c_s,c_t,c_total,e_s,e_t,e_total
    ,m_s,m_t,m_total,total_s,total_t,grand_total),
    ncol=3,byrow=TRUE)
colnames(given_data)<-c("S","T","Total")
rownames(given_data)<-c("C","E","M","Total")
given_data<-as.table(given_data)
print(given_data)
probability<-print(m_t/grand_total)
percentage<-print(probability*100)
```

R code Exa 3.19.b Conditional Probability

```
#Page 141
c_s<-12
c_t<-1
c_total<-c_s+c_t
e_s<-4
e_t<-3
e_total<-e_s+e_t
m_s<-6
```

```
m_t<-2
m_total<-m_s+m_t
total_s<-c_s+e_s+m_s
total_t<-c_t+e_t+m_t
grand_total<-c_total+e_total+m_total
given_data<-matrix(c(c_s,c_t,c_total,e_s,e_t,e_total
    ,m_s,m_t,m_total,total_s,total_t,grand_total),
    ncol=3,byrow=TRUE)
colnames(given_data)<-c("S","T","Total")
rownames(given_data)<-c("C","E","M","Total")
given_data<-as.table(given_data)
print(given_data)
probability_of_m<-print(m_total/grand_total)
probability_of_t<-print(total_t/grand_total)
probability_of_m_intersect_t<-print(m_t/grand_total)
final_probability<-print((probability_of_m+
    probability_of_t)-probability_of_m_intersect_t)
percentage<-final_probability*100
print(percentage)
```

R code Exa 3.19.c Conditional Probability

```
#Page 141
c_s<-12
c_t<-1
c_total<-c_s+c_t
e_s<-4
e_t<-3
e_total<-e_s+e_t
m_s<-6
m_t<-2
m_total<-m_s+m_t
total_s<-c_s+e_s+m_s
total_t<-c_t+e_t+m_t
grand_total<-c_total+e_total+m_total
```

```
given_data<-matrix(c(c_s,c_t,c_total,e_s,e_t,e_total
    ,m_s,m_t,m_total,total_s,total_t,grand_total),
    ncol=3,byrow=TRUE)
colnames(given_data)<-c("S","T","Total")
rownames(given_data)<-c("C","E","M","Total")
given_data<-as.table(given_data)
print(given_data)
probability_of_m<-print(m_total/grand_total)
reqd_probability<-print(1-probability_of_m)
```

R code Exa 3.20.a Fair Die

```
#Page 154
library('sets')
sample_space<-c(1,2,3,4,5,6)
event_A<-sample_space[sample_space==5]
event_B<-sample_space[(sample_space%%2)!=0]
Intersection<-intersect(event_A,event_B)
print(as.set(event_A))
print(as.set(event_B))
print(as.set(Intersection))
probability_of_intersection<-(length(Intersection)/
    length(sample_space))
probability_of_event_B<-(length(event_B)/length(
    sample_space))
reqd_probability<-probability_of_intersection/
    probability_of_event_B
print(reqd_probability)
```

R code Exa 3.20.b Fair Die

```
#Page 154
library("sets")
```

```
sample_space<-c(1,2,3,4,5,6)
event_A<-sample_space[sample_space==5]
event_B<-sample_space[(sample_space%%2)!=0]
Intersection<-intersect(event_A,event_B)
print(as.set(event_A))
print(as.set(event_B))
print(as.set(Intersection))
probability_of_intersection<-(length(Intersection)/length(sample_space))
probability_of_event_A<-(length(event_A)/length(sample_space))
reqd_probability<-probability_of_intersection/probability_of_event_A
print(reqd_probability)
```

R code Exa 3.21.a Conditional Probability and Independent Events

```
#Page 156
m_e<-43
m_w<-293
m_h<-114
m_total<-(m_e+m_w+m_h)
f_e<-82
f_w<-299
f_h<-71
f_total<-(f_e+f_w+f_h)
total_e<-(m_e+f_e)
total_w<-(m_w+f_w)
total_h<-(m_h+f_h)
grand_total<-(m_total+f_total)
given_data<-matrix(c(m_e,m_w,m_h,m_total,f_e,f_w,f_h,f_total,total_e,total_w,total_h,grand_total),ncol=4,byrow=TRUE)
colnames(given_data)<-c("E","W","H","Total")
rownames(given_data)<-c("M","F","Total")
```

```
given_data<-as.table(given_data)
print(given_data)
reqd_probability<-print(total_e/grand_total)
percentage<-reqd_probability*100
print(percentage)
```

R code Exa 3.21.b Conditional Probability and Independent Events

```
#Page 156
m_e<-43
m_w<-293
m_h<-114
m_total<-(m_e+m_w+m_h)
f_e<-82
f_w<-299
f_h<-71
f_total<-(f_e+f_w+f_h)
total_e<-(m_e+f_e)
total_w<-(m_w+f_w)
total_h<-(m_h+f_h)
grand_total<-(m_total+f_total)
given_data<-matrix(c(m_e,m_w,m_h,m_total,f_e,f_w,f_h
    ,f_total,total_e,total_w,total_h,grand_total),
    ncol=4,byrow=TRUE)
colnames(given_data)<-c("E","W","H","Total")
rownames(given_data)<-c("M","F","Total")
given_data<-as.table(given_data)
print(given_data)
reqd_probability<-print(m_e/m_total)
percentage<-reqd_probability*100
print(percentage)
```

R code Exa 3.22.a Conditional Probability and Independent Events

```
#Page 158
h_o<-0.09
h_not_o<-0.11
not_h_o<-0.02
not_h_not_o<-0.78
given_data<-matrix(c(h_o,h_not_o,not_h_o,not_h_not_o),ncol=2,byrow=TRUE)
colnames(given_data)<-c("Overweight","Not Overweight")
rownames(given_data)<-c("Hypertension","No Hypertension")
given_data<-as.table(given_data)
print(given_data)
reqd_probability<-print(h_o/(h_o+not_h_o))
```

R code Exa 3.22.b Conditional Probability and Independent Events

```
#Page 158
h_o<-0.09
h_not_o<-0.11
not_h_o<-0.02
not_h_not_o<-0.78
given_data<-matrix(c(h_o,h_not_o,not_h_o,not_h_not_o),ncol=2,byrow=TRUE)
colnames(given_data)<-c("Overweight","Not Overweight")
rownames(given_data)<-c("Hypertension","No Hypertension")
given_data<-as.table(given_data)
print(given_data)
reqd_probability<-print(h_not_o/(h_not_o+not_h_not_o))
```

R code Exa 3.22.c Conditional Probability and Independent Events

```
#Page 158
p1<-0.8181818
p2<-0.1235955
print(max(p1,p2))
factor<-print(p1/p2)
```

R code Exa 3.23 Test for Independent Events

```
#Page 160
library("sets")
sample_space<-c(1,2,3,4,5,6)
event_A<-c(3)
event_B<-c(1,3,5)
intersection<-intersect(event_A,event_B)
P_A<-length(event_A)/length(sample_space)
P_B<-length(event_B)/length(sample_space)
P_Intersection<-print(length(intersection)/length(
   sample_space))
product<-print(P_A*P_B)
print(product==P_Intersection)
```

R code Exa 3.24 Test for Independent Events

```
#Page 161
m_e<-43
m_w<-293
m_h<-114
m_total<-(m_e+m_w+m_h)
f_e<-82
f_w<-299
f_h<-71
```

```
f_total<-(f_e+f_w+f_h)
total_e<-(m_e+f_e)
total_w<-(m_w+f_w)
total_h<-(m_h+f_h)
grand_total<-(m_total+f_total)
given_data<-matrix(c(m_e,m_w,m_h,m_total,f_e,f_w,f_h
    ,f_total,total_e,total_w,total_h,grand_total),
    ncol=4,byrow=TRUE)
colnames(given_data)<-c("E","W","H","Total")
rownames(given_data)<-c("M","F","Total")
given_data<-as.table(given_data)
print(given_data)
P_F<-f_total/grand_total
P_E<-total_e/grand_total
P_intersection<-print(f_e/grand_total)
product<-print(P_F*P_E)
print(product==P_intersection)
```

R code Exa 3.25.a Conditional Probability and Independent Events

```
#Page 162
P_a<-92/100
reqd_probability<-print(P_a*P_a)
```

R code Exa 3.25.b Conditional Probability and Independent Events

```
#Page 162
P_a<-92/100
reqd_probability<-print(P_a+P_a-(P_a*P_a))
```

R code Exa 3.26.a Complement Rule of Probability

```
#Page 163
P_b<-89/100
reqd_probability<-print(1-P_b)
```

R code Exa 3.26.b Rule of Independence Events

```
#Page 163
P_b<-89/100
reqd_probability<-print((1-P_b)*(1-P_b))
```

R code Exa 3.27 Repetitive Probabilities of Independent Events

```
#Page 164
P_d<-90/100
P_not_d<-(1-P_d)
three_dogs_not_d<-print(P_not_d*P_not_d*P_not_d)
three_dogs_d<-print(1-three_dogs_not_d)
```

R code Exa 3.28.a Repetitive Probabilities Case 1

```
#Page 166
no_of_black_marbles<-7
no_of_white_marbles<-3
total_marbles<-no_of_white_marbles+no_of_black_
   marbles
P_first_black<-no_of_black_marbles/total_marbles
P_second_black<-(no_of_black_marbles-1)/(total_
   marbles-1)
```

```
probability_both_black<-print(P_first_black*P_second_black)
```

R code Exa 3.28.b Repetitive Probabilities Case 2

```
#Page 166
no_of_black_marbles<-7
no_of_white_marbles<-3
total_marbles<-no_of_white_marbles+no_of_black_marbles

P_first_black<-no_of_black_marbles/total_marbles
P_second_not_black<-no_of_white_marbles/(total_marbles-1)

P_first_not_black<-no_of_white_marbles/total_marbles
P_second_black<-no_of_black_marbles/(total_marbles-1)

reqd_probability<-print((P_first_black*P_second_not_black)+(P_first_not_black*P_second_black))
```

R code Exa 3.28.c Repetitive Probabilities Case 3

```
#Page 166
no_of_black_marbles<-7
no_of_white_marbles<-3
total_marbles<-no_of_white_marbles+no_of_black_marbles

P_first_white<-no_of_white_marbles/total_marbles
P_second_white<-(no_of_white_marbles-1)/(total_marbles-1)
```

```

P_both_white<-P_first_white*P_second_white

reqd_probability<-print(1-P_both_white)
```

Chapter 4

Discrete Random Variables

R code Exa 4.1.a Probability Distribution of Discrete Variables

```
#Page 186
library("sets")
sample_space<-c('hh','ht','th','tt')
Event_zero_heads<-as.set(c('tt'))
Event_one_head<-as.set(c('ht','th'))
Event_two_heads<-as.set(c('hh'))
x<-c(0,1,2)
Prob<-c((length(Event_zero_heads)/length(sample_
    space)),(length(Event_one_head)/length(sample_
    space)),(length(Event_two_heads)/length(sample_
    space)))
prob.dist<-cbind(x,Prob)
print(prob.dist)
```

R code Exa 4.1.b Probability Distribution of Discrete Variables

```
#Page 186
library("sets")
```

```
library("ggplot2")
sample_space<-c('hh','ht','th','tt')
Event_zero_heads<-as.set(c('tt'))
Event_one_head<-as.set(c('ht','th'))
Event_two_heads<-as.set(c('hh'))

x<-c(0,1,2)
Prob<-c((length(Event_zero_heads)/length(sample_space)),(length(Event_one_head)/length(sample_space)),(length(Event_two_heads)/length(sample_space)))
prob.dist<-cbind(x,Prob)
print(prob.dist)

df<-data.frame(x = x, y = Prob)
p <- ggplot(df,aes(x=x,after_stat(Prob))) + geom_histogram(binwidth=1, fill="#69b3a2", color="#e9ecef")
print(p)

reqd_probability<-sum(Prob[2],Prob[3])
print(reqd_probability)
```

R code Exa 4.2.a Probability Distribution of Discrete Variables

```
#Page 188
r1<-c(2,3,4,5,6,7)
r2<-c(3,4,5,6,7,8)
r3<-c(4,5,6,7,8,9)
r4<-c(5,6,7,8,9,10)
r5<-c(6,7,8,9,10,11)
r6<-c(7,8,9,10,11,12)
dice<-matrix(c(r1,r2,r3,r4,r5,r6),nrow=6,byrow=TRUE)
print(dice)
sum<-c(2,3,4,5,6,7,8,9,10,11,12)
```

```
prob<-c(1/36,2/36,3/36,4/36,5/36,6/36,5/36,4/36,3/
    36,2/36,1/36)
prob.dist<-cbind(sum,prob)
print(prob.dist)
```

R code Exa 4.2.b Probability of Discrete Random Variable

```
#Page 188
r1<-c(2,3,4,5,6,7)
r2<-c(3,4,5,6,7,8)
r3<-c(4,5,6,7,8,9)
r4<-c(5,6,7,8,9,10)
r5<-c(6,7,8,9,10,11)
r6<-c(7,8,9,10,11,12)
dice<-matrix(c(r1,r2,r3,r4,r5,r6),nrow=6,byrow=TRUE)

sum<-c(2,3,4,5,6,7,8,9,10,11,12)
prob<-c(1/36,2/36,3/36,4/36,5/36,6/36,5/36,4/36,3/
    36,2/36,1/36)
prob.dist<-cbind(sum,prob)

reqd_probability<-print(sum(prob[8],prob[9],prob
    [10],prob[11]))
```

R code Exa 4.2.c Probability Distribution for Tossing Two Fair Dice

```
#Page 188
library("ggplot2")
r1<-c(2,3,4,5,6,7)
r2<-c(3,4,5,6,7,8)
r3<-c(4,5,6,7,8,9)
r4<-c(5,6,7,8,9,10)
r5<-c(6,7,8,9,10,11)
```

```
r6<-c(7,8,9,10,11,12)
dice<-matrix(c(r1,r2,r3,r4,r5,r6),nrow=6,byrow=TRUE)

sum<-c(2,3,4,5,6,7,8,9,10,11,12)
prob<-c(1/36,2/36,3/36,4/36,5/36,6/36,5/36,4/36,3/36,2/36,1/36)
prob.dist<-cbind(sum,prob)

reqd_probability<-print(sum(prob[1],prob[3],prob[5],prob[7],prob[9],prob[11]))

df<-data.frame(x = sum, y = prob)

p <- ggplot(df,aes(x=sum,after_stat(prob))) + geom_histogram(binwidth=1, fill="#69b3a2", color="#e9ecef")
print(p)
```

R code Exa 4.3 Mean of Discrete Random Variable

```
#Page 190
x<-c(-2,1,2,3.5)
prob<-c(0.21,0.34,0.24,0.21)
prob.dist<-cbind(x,prob)
print(prob.dist)
product <- function(n) {
  k<-x[n]*prob[n]
}
mean<-print(sum(product(1),product(2),product(3),product(4)))
```

R code Exa 4.4.a Probability Distribution of Discrete Random Variable

```
#Page 191
total<-1000
fp<-300
sp<-200
tp<-100
los<-0

x<-fp-1
x1<-1
P_of_x<-x1/total

y<-sp-1
y1<-1
P_of_y<-y1/total

z<-tp-1
z1<-1
P_of_z<-z1/total

lose<-total-(x1+y1+z1)
l<-los-1
P_of_l<-lose/total

k<-c(x,y,z,l)
prob<-c(P_of_x,P_of_y,P_of_z,P_of_l)
prob.dist<-cbind(k,prob)
print(prob.dist)
```

R code Exa 4.4.b Probability of Discrete Random Variable Case

```
#Page 191
total<-1000
fp<-300
sp<-200
tp<-100
```

```
los<-0

x<-fp-1
x1<-1
P_of_x<-x1/total

y<-sp-1
y1<-1
P_of_y<-y1/total

z<-tp-1
z1<-1
P_of_z<-z1/total

lose<-total-(x1+y1+z1)
l<-los-1
P_of_l<-lose/total

k<-c(x,y,z,l)
prob<-c(P_of_x,P_of_y,P_of_z,P_of_l)

reqd_probability<-print(sum(prob[1],prob[2],prob[3])
   )
```

R code Exa 4.4.c Expected Values of Discrete Variables

```
#Page 191
total<-1000
fp<-300
sp<-200
tp<-100
los<-0

x<-fp-1
x1<-1
```

```
P_of_x<-x1/total

y<-sp-1
y1<-1
P_of_y<-y1/total

z<-tp-1
z1<-1
P_of_z<-z1/total

lose<-total-(x1+y1+z1)
l<-los-1
P_of_l<-lose/total

k<-c(x,y,z,l)
prob<-c(P_of_x,P_of_y,P_of_z,P_of_l)

product <- function(n){
  h<-k[n]*prob[n]
}
mean<-print(sum(product(1),product(2),product(3),
   product(4)))
```

R code Exa 4.5 Expected Values of Discrete Variable

```
#Page 192
x<-c(195,-199805)
prob<-c(99.7/100,0.03/100)
prob.dist<-cbind(x,prob)
print(prob.dist)
product <- function(n){
  k<-x[n]*prob[n]
}
mean<-print(sum(product(1),product(2)))

```

```
#The answer may slightly vary due to rounding off
   values.
```

R code Exa 4.6.a Finding missing values in Probability Distribution of Discrete Variable

```
#Page 194
x<-c(-1,0,1,4)
prob<-c(0.2,0.5,NA,0.1)
prob.dist<-cbind(x,prob)
print(prob.dist)
a<-print(1-(prob[1]+prob[2]+prob[4]))
```

R code Exa 4.6.b Probability of Discrete Random Variable from Probability Distribution

```
#Page 194
x<-c(-1,0,1,4)
prob<-c(0.2,0.5,NA,0.1)
prob.dist<-cbind(x,prob)

print(prob[2])
```

R code Exa 4.6.c Finding Probability from Probability Distribution of Discrete Variables

```
#Page 194
x<-c(-1,0,1,4)
prob<-c(0.2,0.5,0.2,0.1)
prob.dist<-cbind(x,prob)
```

```
P_of_zero<-print(sum(prob[3],prob[4]))
```

R code Exa 4.6.d Finding Probability from Probability Distribution of Discrete Variables

```
#Page 194
x<-c(-1,0,1,4)
prob<-c(0.2,0.5,0.2,0.1)
prob.dist<-cbind(x,prob)

P_of_zero<-print(sum(prob[2],prob[3],prob[4]))
```

R code Exa 4.6.e Finding Probability from Probability Distribution of Discrete Variables

```
#Page 194
x<-c(-1,0,1,4)
prob<-c(0.2,0.5,0.2,0.1)
prob.dist<-as.data.frame(cbind(x,prob))
P<-print(prob[which(prob.dist$x<=-2)])
```

R code Exa 4.6.f Finding Mean from Probability Distribution of Discrete Variables

```
#Page 194
x<-c(-1,0,1,4)
prob<-c(0.2,0.5,0.2,0.1)
prob.dist<-cbind(x,prob)

```

```
product <- function(n){
  k<-x[n]*prob[n]
}
mean<-print(sum(product(1),product(2),product(3),product(4)))
```

R code Exa 4.6.g Finding Variance from Probability Distribution of Discrete Variables

```
#Page 194
x<-c(-1,0,1,4)
prob<-c(0.2,0.5,0.2,0.1)
prob.dist<-cbind(x,prob)

product <- function(n){
  k<-x[n]*prob[n]
}
mean<-sum(product(1),product(2),product(3),product(4))
func<-function(n){
  l<-((x[n]-mean)**2)*(prob[n])
}
variance<-print(sum(func(1),func(2),func(3),func(4)))
```

R code Exa 4.6.h Finding SD from Probability Distribution of Discrete Variables

```
#Page 194
x<-c(-1,0,1,4)
prob<-c(0.2,0.5,0.2,0.1)
prob.dist<-cbind(x,prob)

```

```
product <- function(n){
  k<-x[n]*prob[n]
}
mean<-sum(product(1),product(2),product(3),product
   (4))
func<-function(n){
  l<-((x[n]-mean)**2)*(prob[n])
}
variance<-sum(func(1),func(2),func(3),func(4))
std_dev<-print(sqrt(variance))
```

R code Exa 4.7.a Constructing Probability Distribution from Binomial Values

```
#Page 210
n<-5
p<-0.17
q<-1-p

func<-function(n,x){
  P<-((factorial(n)/(factorial(x)*factorial(n-x)))*(
     p**x)*(q**(n-x)))
}
var<-c(0,1,2,3,4,5)
prob<-c(func(5,0),func(5,1),func(5,2),func(5,3),func
   (5,4),func(5,5))
prob.dist<-cbind(var,prob)
print(prob.dist)
```

R code Exa 4.7.b Maximum of the values

```
#Page 210
library("ggplot2")
```

```
n<-5
p<-0.17
q<-1-p

func<-function(n,x){
  P<-((factorial(n)/(factorial(x)*factorial(n-x)))*(p**x)*(q**(n-x)))
}
var<-c(0,1,2,3,4,5)
prob<-c(func(5,0),func(5,1),func(5,2),func(5,3),func(5,4),func(5,5))
prob.dist<-cbind(var,prob)

k<-which.max(prob)
print(var[k])

df<-data.frame(x = var, y = prob)
p <- ggplot(df,aes(x=var,after_stat(prob))) + geom_histogram(binwidth=1, fill="#69b3a2", color="#e9ecef")
print(p)
```

R code Exa 4.7.c Finding Mean from Probability Distribution of Discrete Variables

```
#Page 210
n<-5
p<-0.17
q<-1-p

func<-function(n,x){
  P<-((factorial(n)/(factorial(x)*factorial(n-x)))*(p**x)*(q**(n-x)))
}
var<-c(0,1,2,3,4,5)
```

```
prob<-c(func(5,0),func(5,1),func(5,2),func(5,3),func
   (5,4),func(5,5))
prob.dist<-cbind(var,prob)

product <- function(n){
  k<-var[n]*prob[n]
}
mean<-print(sum(product(1),product(2),product(3),
   product(4),product(5),product(6)))
```

R code Exa 4.8 Finding Mean and SD of the random variable

```
#Page 213
n<-5
p<-0.17
q<-1-p
mean<-print(n*p)
std_dev<-print(sqrt(n*p*q))
```

R code Exa 4.9.a Probability using Binomial Distribution

```
#Page 215
n<-10
p<-0.5
q<-1-p
func<-function(n,x){
  P<-((factorial(n)/(factorial(x)*factorial(n-x)))*(
     p**x)*(q**(n-x)))
}
P_of_k<-func(10,6)
print(P_of_k)
```

R code Exa 4.9.b Probability using Binomial Distribution

```
#Page 215
n<-10
p<-0.5
q<-1-p
func<-function(n,x){
  P<-((factorial(n)/(factorial(x)*factorial(n-x)))*(
     p**x)*(q**(n-x)))
}
reqd<-print(sum(func(10,6),func(10,7),func(10,8),
   func(10,9),func(10,10)))
```

Chapter 5

Continuous Random Variables

R code Exa 5.1.a Probability of Continuous Random Variable

```
#Page 231
x=seq(0,1,length=200)
y=dunif(x,min=0,max=1)
plot(x,y,type="l",xlim=c(0,3),ylim=c(0,2),lwd=2,col=
    "red",ylab="p")

x=seq(0.75,1,length=100)
y=dunif(x,min=0,max=1)
polygon(c(0.75,x,1),c(0,y,0),col="lightgray",border
    = NA)

full<-qunif(1,min=0,max=1)
limit<-qunif(0.75,min=0,max=1)
print(full-limit)
```

R code Exa 5.1.b Continuous Random Variables

```
#Page 231
```

```
x=seq(0,1,length=200)
y=dunif(x,min=0,max=1)
plot(x,y,type="l",xlim=c(0,3),ylim=c(0,2),lwd=2,col=
   "red",ylab="p")

x=seq(0,0.2,length=100)
y=dunif(x,min=0,max=1)
polygon(c(0,x,0.2),c(0,y,0),col="lightgray",border =
    NA)

print(qunif(0.2,min=0,max=1))
```

R code Exa 5.1.c Continuous Random Variables

```
#Page 231
x=seq(0,1,length=200)
y=dunif(x,min=0,max=1)
plot(x,y,type="l",xlim=c(0,3),ylim=c(0,2),lwd=2,col=
   "red",ylab="p")

x=seq(0.4,0.7,length=100)
y=dunif(x,min=0,max=1)
polygon(c(0.4,x,0.7),c(0,y,0),col="lightgray",border
    = NA)

l1<-qunif(0.7,min=0,max=1)
l2<-qunif(0.4,min=0,max=1)
print(l1-l2)
```

R code Exa 5.2 Computing Probability from Area of Density Function Curve

```
#Page 233
```

```
f<-1/30
y<-10
print(f*y)
```

R code Exa 5.3 Computing Probability from Area of Density Function Curve

```
#Page 237
x<-seq(40,100,by=0.1)
y<-dnorm(x,mean=69.75,sd=2.59)
plot(x,y,main="Normal Distribution")
```

R code Exa 5.4.a Probability of Standard Normal Variable

```
#Page 243

x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
print(pnorm(1.48,mean=0,sd=1))

x=seq(-4,1.5,length=200)
y=dnorm(x)
polygon(c(-4,x,1.48),c(0,y,0),col="gray")
```

R code Exa 5.4.b Probability of Standard Normal Variable

```
#Page 243

x=seq(-4,4,length=200)
```

```
y=dnorm(x,mean=0,sd=1)
print(pnorm(-0.25,mean=0,sd=1))
```

R code Exa 5.5.a Probability of Standard Normal Variable

```
#Page 244

x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-1-(pnorm(1.60,mean=0,sd=1))
print(reqd_probability)

x=seq(4,1.5,length=200)
y=dnorm(x)
polygon(c(-4,x,1.6),c(0,y,0),col="gray")
```

R code Exa 5.5.b Probability of Standard Normal Variable

```
#Page 244

x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
reqd_probability<-1-(pnorm(-1.02,mean=0,sd=1))
print(reqd_probability)
```

R code Exa 5.6.a Finding Probabilities of Normally Distributed Variable

```
#Page 246
x=seq(-6,6,length=200)
```

```
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(1.57,mean=0,sd=1)-pnorm(0.5,mean=0,sd=1))


x=seq(0.5,1.57,length=200)
y=dnorm(x)
polygon(c(0.5,x,1.57),c(0,y,0),col="gray")
```

R code Exa 5.6.b Finding Probabilities of Normally Distributed Variable

```
#Page 246
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(0.09,mean=0,sd=1)-pnorm(-2.55,mean=0,sd=1))

x=seq(-2.55,0.09,length=200)
y=dnorm(x)
polygon(c(-2.55,x,0.09),c(0,y,0),col="gray")
```

R code Exa 5.7.a Probability of Standard Normal Variable

```
#Page 248
x=seq(-5,5,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(4.16,mean=0,sd=1)-pnorm(1.13,mean=0,sd=1))

x=seq(1.13,4.16,length=200)
```

```
y=dnorm(x)
polygon(c(1.13,x,4.16),c(0,y,0),col="gray")
```

R code Exa 5.7.b Probability of Standard Normal Variable

```
#Page 248
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(2.15,mean=0,sd=1)-
   pnorm(-5.22,mean=0,sd=1))

x=seq(-5.22,2.15,length=200)
y=dnorm(x)
polygon(c(-5.22,x,2.15),c(0,y,0),col="gray")
```

R code Exa 5.8.a Probability of Standard Normal Variable

```
#Page 249
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(1,mean=0,sd=1)-pnorm
   (-1,mean=0,sd=1))

x=seq(-1,1,length=200)
y=dnorm(x)
polygon(c(-1,x,1),c(0,y,0),col="gray")
```

R code Exa 5.8.b Probability of Standard Normal Variable

```r
#Page 249
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(2,mean=0,sd=1)-pnorm
   (-2,mean=0,sd=1))

x=seq(-2,2,length=200)
y=dnorm(x)
polygon(c(-2,x,2),c(0,y,0),col="gray")
```

R code Exa 5.8.c Probability of Standard Normal Variable

```r
#Page 249
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(3,mean=0,sd=1)-pnorm
   (-3,mean=0,sd=1))

x=seq(-3,3,length=200)
y=dnorm(x)
polygon(c(-3,x,3),c(0,y,0),col="gray")
```

R code Exa 5.9.a Probability Computations for General Normal Random Variables

```r
#Page 256
mean<-10
sd<-2.5
k<-(14-mean)/sd
print(k)

```

```
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(k,mean=0,sd=1))

x=seq(-4,k,length=200)
y=dnorm(x)
polygon(c(-4,x,k),c(0,y,0),col="gray")
```

R code Exa 5.9.b Probability Computations for General Normal Random Variables

```
#Page 256
mean<-10
sd<-2.5
k<-(8-mean)/sd
l<-(14-mean)/sd
print(k)
print(l)
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(l,mean=0,sd=1)-pnorm(k
    ,mean=0,sd=1))

x=seq(k,l,length=500)
y=dnorm(x)
polygon(c(k,x,l),c(0,y,0),col="gray")
```

R code Exa 5.10 Finding Probability of General Normal Variable

```
#Page 258
mean<-37.5
```

```
sd<-4.5
k<-(30-mean)/sd
l<-(40-mean)/sd
print(k)
print(l)
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type=”l”,lwd=2,col=”red”)
reqd_probability<-print(pnorm(l,mean=0,sd=1)-pnorm(k,mean=0,sd=1))

x=seq(k,l,length=500)
y=dnorm(x)
polygon(c(k,x,l),c(0,y,0),col=”gray”)
```

R code Exa 5.11 Finding Probability of General Normal Variable

```
#Page 260

mean<-510
sd<-60
k<-(650-mean)/sd
print(k)

x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type=”l”,lwd=2,col=”red”)
reqd_probability<-print(1-(pnorm(k,mean=0,sd=1)))

x=seq(4,k,length=200)
y=dnorm(x)
polygon(c(4,x,k),c(0,y,0),col=”gray”)
```

R code Exa 5.12 Areas of Tails of Distributions

```
#Page 271
print(qnorm(0.0125,lower.tail = TRUE))
```

R code Exa 5.13 Areas of Tails of Distributions

```
#Page 272
k<-qnorm(0.025,lower.tail = FALSE)
print(k)

x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")

x=seq(4,k,length=200)
y=dnorm(x)
polygon(c(4,x,k),c(0,y,0),col="gray")
```

R code Exa 5.14 Right and Left tails of Area in the Standard Normal Distribution

```
#Page 274
print(qnorm(0.01,lower.tail = FALSE))
print(qnorm(1-0.01,lower.tail=FALSE))
```

R code Exa 5.15 Area of Tails of Distribution with Mean and SD

```
#Page 276
z<-qnorm(0.9332,lower.tail = TRUE)
```

```
mean<-10
sd<-2.5
k<-mean+(z)*(sd)
print(k)

x=seq(5,16,length=200)
y=dnorm(x,mean=10,sd=2)
plot(x,y,type="l",lwd=2,col="red")

x=seq(5,k,length=200)
y=dnorm(x,mean=10,sd=2)
polygon(c(5,x,k),c(0,y,0),col="gray")
```

R code Exa 5.16 Area of Tails of Distribution with Mean and SD

```
#Page 277
z<-qnorm(1-0.65,lower.tail = TRUE)
mean<-175
sd<-12
k1<-mean+(z)*(sd)
print(k1)

x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")

x=seq(z,4,length=500)
y=dnorm(x)
polygon(c(z,x,4),c(0,y,0),col="gray")
```

R code Exa 5.17 Area of Tails of Distribution with Mean and SD

```
#Page 278
```

```
z<-qnorm(1-0.05,lower.tail = TRUE)
mean<-510
sd<-60
x<-mean+(z)*(sd)
print(x)

x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type=”l”,lwd=2,col=”red”)

x=seq(z,4,length=500)
y=dnorm(x)
polygon(c(z,x,4),c(0,y,0),col=”gray”)
```

R code Exa 5.18 Range of Area of Tail Distributions

```
#Page 280
mean<-29
sd<-2
alpha<-1-0.75
z1<-qnorm(1-alpha/2,lower.tail = FALSE)
z2<-qnorm(1-alpha/2,lower.tail = TRUE)
x_fast<-mean+(z1)*(sd)
x_slow<-mean+(z2)*(sd)
print(x_fast)
print(x_slow)

x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type=”l”,lwd=2,col=”red”)

x=seq(z1,z2,length=500)
y=dnorm(x)
polygon(c(z1,x,z2),c(0,y,0),col=”gray”)
```

Chapter 6

Sampling Distributions

R code Exa 6.1 Finding Probability Distribution and Mean and SD of the sample

```
#Page 289
set<-c(152,156,160,164)
l<-print.table(combn(set,2))
mean<-function(n){
  m<-(l[1,n]+l[2,n])/2
  a<-l[1,n]
  b<-l[2,n]
  print(paste("Mean of",a,"and",b,"is",m))
}
mean(1)
mean(2)
mean(3)
mean(4)
mean(5)
mean(6)

x<-c(152,154,156,158,160,162,164)
prob<-c(1/16,2/16,3/16,4/16,3/16,2/16,1/16)
prob.dist<-cbind(x,prob)
print(prob.dist)
```

```
product<- function(i){
  k<-x[i]*prob[i]
}
mean_of_dist<-sum(product(1),product(2),product(3),
    product(4),product(5),product(6),product(7))
print(mean_of_dist)

func<-function(i){
  k<-x[i]*x[i]*prob[i]
}
variance<-sum(func(1),func(2),func(3),func(4),func
    (5),func(6),func(7))
print(variance)

sd<-sqrt(variance-(mean_of_dist**2))
print(sd)
```

R code Exa 6.2 Finding Mean and SD

```
#Page 291
mean<-13525
sd<-4180
n<-100
print(mean)
print(sd/sqrt(n))
```

R code Exa 6.3.a Finding Mean and SD of the Sample Distribution

```
#Page 298
n<-50
mean<-112
sd<-40
```

```
print(mean)
print(sd/sqrt(n))
```

R code Exa 6.3.b Sampling Distribution Probabilities

```
#Page 298
n<-50
mean<-112
sd<-40
std_new<-sd/sqrt(n)
k<-(110-mean)/std_new
l<-(114-mean)/std_new
print(k)
print(l)
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(l,mean=0,sd=1)-pnorm(k
    ,mean=0,sd=1))

x=seq(k,l,length=500)
y=dnorm(x)
polygon(c(k,x,l),c(0,y,0),col="gray")
```

R code Exa 6.3.c Sampling Distribution Probabilities

```
#Page 298
n<-50
mean<-112
sd<-40
std_new<-sd/sqrt(n)
k<-(113-mean)/std_new
print(k)
```

```
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
reqd_probability<-print(1-(pnorm(k,mean=0,sd=1)))

#The answer may slightly vary due to rounding off values.
```

R code Exa 6.4 Computing Mean from Sampling Distribution

```
#Page 300
n<-100
mean<-2.61
sd<-0.5
std_new<-sd/sqrt(n)
k<-(2.51-mean)/std_new
l<-(2.71-mean)/std_new
print(k)
print(l)
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="red")
reqd_probability<-print(pnorm(l,mean=0,sd=1)-pnorm(k,mean=0,sd=1))

x=seq(k,l,length=500)
y=dnorm(x)
polygon(c(k,x,l),c(0,y,0),col="gray")
```

R code Exa 6.5 Probability of Sample Mean of Sampling Distribution

```
#Page 302
n<-5
mean<-38.5
```

```
sd<-2.5
std_new<-sd/sqrt(n)
k<-(36-mean)/std_new
print(k)
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
reqd_probability<-print(pnorm(k,mean=0,sd=1))

#The answer may slightly vary due to rounding off
   values.
```

R code Exa 6.6.a Probability of Sampling Distribution

```
#Page 303

n<-5
mean<-50
sd<-6
k<-(48-mean)/sd
print(k)
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
reqd_probability<-print(pnorm(k,mean=0,sd=1))

#The answer may slightly vary due to rounding off
   values.
```

R code Exa 6.6.b Probability of Sampling Distribution

```
#Page 303

n<-36
mean<-50
```

```
sd<-6
std_new<-sd/sqrt(n)
k<-(48-mean)/std_new
print(k)
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
reqd_probability<-print(pnorm(k,mean=0,sd=1))
```

R code Exa 6.7.a Normal Distribution

```
#Page 314
n<-900
p<-0.38
q<-1-p
mean<-p
print(mean)
std_dev<-sqrt((p*q)/n)
print(std_dev)
l<-3*std_dev
print(l)
upper_limit<-p+l
lower_limit<-p-l
print(upper_limit)
print(lower_limit)
print("Normally Distributed")
```

R code Exa 6.7.b The Sample Proportion

```
#Page 314

p<-0.38
q<-1-p
n<-900
```

```
mean<-0.38
sd<-sqrt((p*q)/n)
k<-((0.38-0.05)-mean)/sd
l<-((0.38+0.05)-mean)/sd
print(k)
print(l)
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
reqd_probability<-print(pnorm(l,mean=0,sd=1)-pnorm(k
    ,mean=0,sd=1))
```

R code Exa 6.8.a Computing Sample Proportion

```
#Page 316
x<-102
n<-121
sample_proportion<-x/n
print(sample_proportion)
```

R code Exa 6.8.b Sampling Distributions

```
#Page 316
n<-121
p<-0.90
q<-1-p
mean<-p
print(mean)
std_dev<-sqrt((p*q)/n)
print(std_dev)
l<-3*std_dev
print(l)
upper_limit<-p+l
lower_limit<-p-l
```

```
print(upper_limit)
print(lower_limit)
print("Normally Distributed")
```

R code Exa 6.8.c Finding Sample Proportion

```
#Page 316
x<-102
n<-121
mean<-0.90
sd<-0.027
sample_proportion<-x/n
k<-(sample_proportion-mean)/sd
print(k)
x=seq(-4,4,length=200)
y=dnorm(x,mean=0,sd=1)
reqd_probability<-print(pnorm(k,mean=0,sd=1))

#The answer may slightly vary due to rounding off
    values.
```

Chapter 7

Estimation

R code Exa 7.1.a Construction of Confidence Level

```
#Page 330
alpha<-1-0.90
print(qnorm(alpha/2,lower.tail=FALSE))
```

R code Exa 7.1.b Construction of Confidence Level

```
#Page 330
alpha<-1-0.99
print(qnorm(alpha/2,lower.tail=FALSE))
```

R code Exa 7.2.a Critical Values of Confidence Levels

```
#Page 331
alpha<-1-0.90
print(qnorm(alpha/2,lower.tail=FALSE))
```

R code Exa 7.2.b Critical Values of Confidence Levels

```
#Page 331
alpha<-1-0.99
print(qnorm(alpha/2,lower.tail=FALSE))
```

R code Exa 7.3 Finding Critical Values

```
#Page 332
n<-49
mean<-35
sd<-14
alpha<-1-0.98
k<-qnorm(alpha/2,lower.tail = FALSE)
Lower_interval<-print(mean-(((k)*(sd))/sqrt(n)))
Upper_interval<-print(mean+(((k)*(sd))/sqrt(n)))
```

R code Exa 7.4 Construction of Confidence Level

```
#Page 333
n<-120
mean<-2.71
sd<-0.51
alpha<-1-0.90
k<-qnorm(alpha/2,lower.tail = FALSE)
Lower_interval<-print(mean-(((k)*(sd))/sqrt(n)))
Upper_interval<-print(mean+(((k)*(sd))/sqrt(n)))
```

R code Exa 7.5 Small Sample Estimation of a Population Mean

```
#Page 343
n<-15
mean<-35
sd<-14
alpha<-1-0.95
df<-n-2
k<-qt(alpha/2,df)
Lower_interval<-print(mean+(((k)*(sd))/sqrt(n)))
Upper_interval<-print(mean-(((k)*(sd))/sqrt(n)))
```

R code Exa 7.6 Observed Significance

```
#Page 344
n<-12
mean<-2.71
sd<-0.51
alpha<-1-0.90
df<-n-2
k<-qt(alpha/2,df)
Lower_interval<-print(mean+(((k)*(sd))/sqrt(n)))
Upper_interval<-print(mean-(((k)*(sd))/sqrt(n)))
```

R code Exa 7.7 Large Sample Estimation of a Population Proportion

```
#Page 352
n<-120
f<-69
alpha<-1-0.90
prop_of_f<-f/n
sd<-sqrt((prop_of_f)*(1-prop_of_f))
k<-qnorm(alpha/2,lower.tail = FALSE)
```

```
Lower_interval<-print(prop_of_f-(((k)*(sd))/sqrt(n))
   )
Upper_interval<-print(prop_of_f+(((k)*(sd))/sqrt(n))
   )
```

R code Exa 7.8 Computing Minimal Sample Size with E

```
#Page 363
alpha<-1-0.99
sd<-1.3
E<-0.2
z<-qnorm(alpha/2,lower.tail = FALSE)
n<-print((z**2)*(sd**2)/(E**2))

#The answer may slightly vary due to rounding off
   values.
```

R code Exa 7.9 Estimation of Minimal Sample Size

```
#Page 364
alpha<-1-0.95
z<-qnorm(alpha/2,lower.tail = FALSE)
E<-1000
sd<-4000
n<-print((z**2)*(sd**2)/(E**2))
```

R code Exa 7.10.a Estimation of Minimal Sample Size

```
#Page 366
alpha<-1-0.98
```

```
E<-0.05
z<-qnorm(alpha/2,lower.tail = FALSE)
p<-0.5
n<-((z**2)*p*(1-p))/(E**2)
print(ceiling(n))
```

R code Exa 7.10.b Estimation of Minimal Sample Size

```
#Page 366
alpha<-1-0.98
E<-0.05
z<-qnorm(alpha/2,lower.tail = FALSE)
p<-0.1
n<-((z**2)*p*(1-p))/(E**2)
print(ceiling(n))
```

R code Exa 7.11 Estimation of Minimal Sample Size

```
#Page 367
alpha<-1-0.90
z<-qnorm(alpha/2,lower.tail = FALSE)
p<-0.5
E<-0.03
n<-((z**2)*p*(1-p))/(E**2)
print(ceiling(n))
```

Chapter 8

Testing Hypotheses

R code Exa 8.3 The Elements of Hypothesis Testing

```
#Page 380
mean<-8
sd<-0.15
n<-5
sd_x<-sd/sqrt(n)
alpha<-0.10
z<-qnorm(alpha/2,lower.tail = FALSE)
lower_critical_value<-print(mean-(z*sd_x))
upper_critical_value<-print(mean+(z*sd_x))

x=seq(7.5,8.5,length=100)
y=dnorm(x,mean=mean,sd=sd)
plot(x,y,type="l",lwd=2,col="black")

x=seq(7.5,lower_critical_value,length=100)
y=dnorm(x,mean = mean,sd=sd)
polygon(c(7.5,x,lower_critical_value),c(0,y,0),col="
    gray")

x=seq(upper_critical_value,8.5,length=100)
y=dnorm(x,mean=mean,sd=sd)
```

```
polygon(c(upper_critical_value,x,8.5),c(0,y,0),col="
    gray")
```

R code Exa 8.4 Large Sample Tests for a Population Mean

```
#Page 390
mean<-3.5
mean_exp<-3.1
sd<-1.5
alpha<-0.05
n<-50
critical_value<-qnorm(alpha,lower.tail = TRUE)
z<-print((mean_exp-mean)/(sd/sqrt(n)))

x=seq(-4,4,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(-4,critical_value,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(-4,x,critical_value),c(0,y,0),col="gray")
points(z,0,pch=19,col="red",cex=1.5)

if(z<critical_value){
  print("REJECT NULL HYPOTHESIS")
}else{
  print("ACCEPT NULL HYPOTHESIS")
}
```

R code Exa 8.5 Large Sample Tests for a Population Mean

```
#Page 392
mean<-8.1
```

```
mean_exp<-8.2
sd<-0.22
n<-30
alpha<-0.01
z0<-qnorm(alpha/2,lower.tail = TRUE)
z1<-qnorm(alpha/2,lower.tail = FALSE)
z<-print((mean_exp-mean)/(sd/sqrt(n)))

x=seq(-3,3,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(-3,z0,length=500)
y=dnorm(x,mean =0,sd=1)
polygon(c(-3,x,z0),c(0,y,0),col="gray")

x=seq(z1,3,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(z1,x,3),c(0,y,0),col="gray")

points(z,0,pch=19,col="red",cex=1.5)

if(z<z1 & z>z0){
  print("ACCEPT NULL HYPOTHESIS")
}else{
  print("REJECT NULL HYPOTHESIS")
}
```

R code Exa 8.6 Observed Significance of a Test

```
#Page 401
alpha<-0.01
z<-qnorm(alpha/2,lower.tail = FALSE)

x=seq(-3,3,length=500)
```

```
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(z,3,length=500)
y=dnorm(x,mean =0,sd=1)
polygon(c(z,x,3),c(0,y,0),col="gray")

points(z,0,pch=19,col="red",cex=1)

auc<-pnorm(Inf)-pnorm(z)
significance<-2*auc
print(significance)

#The answer may slightly vary due to rounding off values.
```

R code Exa 8.7 Observed Significance of a Test

```
#Page 403
mean<-202.5
mean_exp<-199.2
sd<-19.63
alpha<-0.05
n<-85
z<-print((mean_exp-mean)/(sd/sqrt(n)))

p<-pnorm(z)-pnorm(-Inf)
print(p)

x=seq(-3,3,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(-3,z,length=500)
y=dnorm(x,mean=0,sd=1)
```

```
polygon(c(-3,x,z),c(0,y,0),col="gray")
points(z,0,pch=19,col="black",cex=1.5)


if(p<alpha){
  print("REJECT NULL HYPOTHESIS")
}else{
  print("ACCEPT NULL HYPOTHESIS")
}
```

R code Exa 8.8 Observed Significance of a Test

```
#Page 405
mean<-67
mean_exp<-69.4
sd<-6.1
alpha<-0.01
n<-64
z<-print((mean_exp-mean)/(sd/sqrt(n)))

p<-pnorm(Inf)-pnorm(z)
print(p)

x=seq(-3,3.5,length=100)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(z,3.5,length=100)
y=dnorm(x,mean=0,sd=1)
polygon(c(z,x,3.5),c(0,y,0),col="gray")
points(z,0,pch=19,col="black",cex=1.5)


if(p<alpha){
  print("REJECT NULL HYPOTHESIS")
```

```
}else{
  print("ACCEPT NULL HYPOTHESIS")
}
```

R code Exa 8.9 Observed Significance of a Test

```
#Page 407

mean<-7.4
mean_exp<-7.7
sd<-0.5
alpha<-0.01
n<-30
z<-print((mean_exp-mean)/(sd/sqrt(n)))

p<-2*(pnorm(Inf)-pnorm(z))
print(p)

x=seq(-5,5,length=100)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(z,5,length=100)
y=dnorm(x,mean=0,sd=1)
polygon(c(z,x,5),c(0,y,0),col="gray")

x=seq(-5,-z,length=100)
y=dnorm(x,mean=0,sd=1)
polygon(c(-5,x,-z),c(0,y,0),col="gray")

points(z,0,pch=19,col="black",cex=1.5)


if(p<alpha){
  print("REJECT NULL HYPOTHESIS")
```

```
}else{
  print("ACCEPT NULL HYPOTHESIS")
}
```

R code Exa 8.10 Small Sample Tests for a Population Mean

```
#Page 416
data<-c(155,179,175,175,161)
mean_exp<-mean(data)
mean<-179
sd<-sd(data)
n<-length(data)
df<-n-1
alpha<-0.05
t<-print((mean_exp-mean)/(sd/sqrt(n)))
t_alpha<-print(qt(alpha,df))

x=seq(-3,3,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(-3,t,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(-3,x,t),c(0,y,0),col="gray")
points(t,0,pch=19,col="red",cex=1)

if(t_alpha>t){
  print("REJECT NULL HYPOTHESIS")
}else{
  print("ACCEPT NULL HYPOTHESIS")
}
```

R code Exa 8.11 Small Sample Tests for a Population Mean

```
#Page 418
data<-c(0.021,0.019,0.023,0.020)
mean_exp<-mean(data)
mean<-0.02
sd<-sd(data)
n<-length(data)
df<-n-1
alpha<-0.01
t<-print((mean_exp-mean)/(sd/sqrt(n)))
t_a1<-print(qt(alpha/2,df))
t_a2<-print(-t_a1)

x=seq(-8,8,length=100)
y=dnorm(x,mean=0,sd=2.5)
plot(x,y,type="l",lwd=2,col="black")

x=seq(-8,t_a1,length=100)
y=dnorm(x,mean=0,sd=2.5)
polygon(c(-8,x,t_a1),c(0,y,0),col="gray")

x=seq(t_a2,8,length=100)
y=dnorm(x,mean=0,sd=2.5)
polygon(c(t_a2,x,8),c(0,y,0),col="gray")

points(t,0,pch=17,col="red",cex=1.5)

if(t>t_a1 & t<t_a2){
  print("ACCEPT NULL HYPOTHESIS")
}else{
  print("REJECT NULL HYPOTHESIS")
}
```

R code Exa 8.12 Large Sample Tests for a Population Proportion

```
#Page 429
```

```
n<-500
soft_drink_maker<-270
competitor<-211
non_deciders<-19
alpha<-0.05
p<-soft_drink_maker/500
p_0<-0.5
q_0<-1-p_0
z<-print((p-p_0)/(sqrt((p_0*q_0)/n)))
z0<-qnorm(alpha,lower.tail=FALSE)
print(z0)

x=seq(-3,3,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(z0,3,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(z0,x,3),c(0,y,0),col="gray")

points(z,0,pch=19,col="red",cex=1.5)

if(z0<z){
  print("REJECT NULL HYPOTHESIS")
}else{
  print("ACCEPT NULL HYPOTHESIS")
}
```

R code Exa 8.13 Large Sample Tests for a Population Proportion

```
#Page 432
n<-5000
alpha<-0.05
p<-0.5255
p_0<-0.5146
```

```
q_0<-1-p_0
z<-print((p-p_0)/(sqrt((p_0*q_0)/n)))
z0<-qnorm(alpha,lower.tail=FALSE)
z1<-qnorm(alpha,lower.tail=TRUE)
print(z0)
print(z1)

x=seq(-3,3,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(-3,z1,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(-3,x,z1),c(0,y,0),col="gray")

x=seq(z0,3,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(z0,x,3),c(0,y,0),col="gray")

points(z,0,pch=19,col="red",cex=1.5)

if(z>z1 & z<z0){
  print("ACCEPT NULL HYPOTHESIS")
}else{
  print("REJECT NULL HYPOTHESIS")
}
```

R code Exa 8.14 p value approach

```
#Page 435
n<-500
soft_drink_maker<-270
competitor<-211
non_deciders<-19
alpha<-0.05
```

```
p<-soft_drink_maker/500
p_0<-0.5
q_0<-1-p_0
z<-print((p-p_0)/(sqrt((p_0*q_0)/n)))

p<-pnorm(Inf)-pnorm(z)
print(p)

x=seq(-3,3,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(z,3,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(z,x,3),c(0,y,0),col="gray")

points(z,0,pch=19,col="black",cex=1.5)

if(p<alpha){
  print("REJECT NULL HYPOTHESIS")
}else{
  print("ACCEPT NULL HYPOTHESIS")
}
```

R code Exa 8.15 p value approach

```
#Page 436
n<-5000
alpha<-0.05
p<-0.5255
p_0<-0.5146
q_0<-1-p_0
z<-print((p-p_0)/(sqrt((p_0*q_0)/n)))

p<-2*(pnorm(Inf)-pnorm(z))
```

```
print(p)

x=seq(-3,3,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(z,3,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(z,x,3),c(0,y,0),col="gray")

points(z,0,pch=19,col="black",cex=1.5)

if(p<alpha){
  print("REJECT NULL HYPOTHESIS")
}else{
  print("ACCEPT NULL HYPOTHESIS")
}
```

Chapter 9

Two Sample Problems

R code Exa 9.1 Construction of a point estimate and confidence interval

```
#Page 447
n1<-174
n2<-355
x1<-3.51
x2<-3.24
s1<-0.51
s2<-0.52
pe<-x1-x2
alpha<-1-0.99
z<-qnorm(alpha/2,lower.tail=FALSE)
Lower_interval<-print(pe-z*(sqrt((s1**2)/(n1))+(s2**
    2)/(n2)))
Upper_interval<-print(pe+z*(sqrt((s1**2)/(n1))+(s2**
    2)/(n2)))
#The answer may slightly vary due to rounding off
    values.
```

R code Exa 9.2 Comparison of Two Population Means using critical value approach

```
#Page 450
n1<-174
n2<-355
x1<-3.51
x2<-3.24
s1<-0.51
s2<-0.52
alpha<-0.01
m<-(((s1**2)/n1)+((s2**2)/n2))
z<-((x1-x2)/sqrt(m))
print(z)
z0<-qnorm(alpha,lower.tail = FALSE)
print(z0)
x=seq(-6,6,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(z0,6,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(z0,x,6),c(0,y,0),col="gray")

points(z,0,pch=19,col="red",cex=1.5)

if(z0<z){
  print("REJECT NULL HYPOTHESIS")
}else{
  print("ACCEPT NULL HYPOTHESIS")
}
```

R code Exa 9.3 p value approach

```
#Page 452

n1<-174
n2<-355
```

```
x1<-3.51
x2<-3.24
s1<-0.51
s2<-0.52
alpha<-0.01
m<-(((s1**2)/n1)+((s2**2)/n2))
z<-((x1-x2)/sqrt(m))
print(z)

p<-pnorm(Inf)-pnorm(z)
print(p)

x=seq(-6,6,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(z,6,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(z,x,6),c(0,y,0),col="gray")

points(z,0,pch=19,col="red",cex=1.5)

if(p<alpha){
  print("REJECT NULL HYPOTHESIS")
}else{
  print("ACCEPT NULL HYPOTHESIS")
}
```

R code Exa 9.4 Construction of a point estimate and confidence interval

```
#Page 468

n1<-11
x1<-52
sd1<-12
```

```
n2<-6
x2<-46
sd2<-10
alpha<-0.05
df<-n1+n2-2
t<-qt(alpha/2,df)
n<-(((n1-1)*(sd1**2))+((n2-1)*(sd2**2)))
d<-(n1+n2-2)

s_p_sq<-n/d
print(s_p_sq)

lower_interval<-print((x1-x2)+t*sqrt(s_p_sq*(1/n1+1/
    n2)))
upper_interval<-print((x1-x2)-t*sqrt(s_p_sq*(1/n1+1/
    n2)))
```

R code Exa 9.5 Critical value approach

```
#Page 470

n1<-11
x1<-52
sd1<-12
n2<-6
x2<-46
sd2<-10
alpha<-0.005
df<-n1+n2-2
n<-(((n1-1)*(sd1**2))+((n2-1)*(sd2**2)))
d<-(n1+n2-2)
s_p_sq<-n/d

t<-(x1-x2)/(sqrt(s_p_sq*((1/n1)+(1/n2))))
t0<-qt(alpha,df)
```

```
t1<--t0
print(t)
print(t0)
print(t1)

x=seq(-4,4,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(t1,4,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(t1,x,4),c(0,y,0),col="gray")

x=seq(-4,t0,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(-4,x,t0),c(0,y,0),col="gray")

points(t,0,pch=19,col="red",cex=1.5)

if(t<t1 & t>t0){
  print("ACCEPT NULL HYPOTHESIS")
}else{
  print("REJECT NULL HYPOTHESIS")
}
```

R code Exa 9.7 Construction of a point estimate and confidence interval

```
#Page 484

car_1<-c(17,13.2,35.3,13.6,32.7,18.4,22.5,26.8,15.1)
car_2<-c(17,12.9,35.4,13.2,32.5,18.1,22.5,26.7,15.0)

diff<-c(car_1-car_2)
diff_sq<-c(diff**2)
n<-length(car_1)
```

```
d<-sum(diff)/n
sd<-(sqrt(sum(diff_sq)-((1/n)*(sum(diff)**2))))/(
   sqrt(n-1))

df<-n-1
alpha<-0.05
t<-qt(alpha/2,df,lower.tail = FALSE)

lower_interval<-print(d-(t*sd/sqrt(n)))
upper_interval<-print(d+(t*sd/sqrt(n)))
```

R code Exa 9.8 Critical value approach

```
#Page 487

car_1<-c(17,13.2,35.3,13.6,32.7,18.4,22.5,26.8,15.1)
car_2<-c(17,12.9,35.4,13.2,32.5,18.1,22.5,26.7,15.0)

diff<-c(car_1-car_2)
diff_sq<-c(diff**2)
n<-length(car_1)
d<-sum(diff)/n

t<-d/(sd/sqrt(n))
print(t)

alpha<-0.05
df<-n-1
t0<-qt(alpha,df,lower.tail = FALSE)
print(t0)

x=seq(-4,4,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

```

```
x=seq(t0,4,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(t0,x,4),c(0,y,0),col="gray")

points(t,0,pch=19,col="red",cex=1.5)

if(t<t0){
  print("ACCEPT NULL HYPOTHESIS")
}else{
  print("REJECT NULL HYPOTHESIS")
}
```

R code Exa 9.10 Construction of a point estimate and confidence interval

```
#Page 503

n1<-500
n2<-100
p1<-0.67
p2<-0.80

alpha<-1-0.9
z<-qnorm(alpha/2,lower.tail = FALSE)
m<-sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
lower_interval<-print((p1-p2)-(z*m))
upper_interval<-print((p1-p2)+(z*m))

#The answer may slightly vary due to rounding off
   values.
```

R code Exa 9.11 Critical value approach

```
#Page 507
```

```

n1<-500
n2<-100
p1<-0.67
p2<-0.80
d0<--0.05
alpha<-1-0.9
m<-sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
z<-(p1-p2-d0)/m
print(z)

z0<-qnorm(alpha,lower.tail = TRUE)
print(z0)

x=seq(-3,3,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(-3,z0,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(-3,x,z0),c(0,y,0),col="gray")

points(z,0,pch=19,col="red",cex=1.5)

if(z0<z){
  print("ACCEPT NULL HYPOTHESIS")
}else{
  print("REJECT NULL HYPOTHESIS")
}
```

R code Exa 9.12 p value approach

```
#Page 509

n1<-500
```

```
n2<-100
p1<-0.67
p2<-0.80
d0<--0.05
alpha<-1-0.9
m<-sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
z<-(p1-p2-d0)/m
print(z)

p<-1-(pnorm(Inf)-pnorm(z))
print(p)

if(p>alpha){
  print("ACCEPT NULL HYPOTHESIS")
} else {
  print("REJECT NULL HYPOTHESIS")
}
```

R code Exa 9.13 Sample Size Considerations

```
#Page 521

alpha<-1-0.995
k1<-0.75
k2<-1.15
z<-qnorm(alpha/2,lower.tail =FALSE)
E=0.5
n<-((z**2)*(k1**2+k2**2))/(E**2)
print(n)

#The answer may slightly vary due to rounding off
   values.
```

R code Exa 9.14 Sample Size Considerations

```
#Page 523

alpha<-1-0.999
k<-0.025
z<-qnorm(alpha/2,lower.tail =FALSE)
E=0.01
n<-(z**2)*(k**2)/(E**2)
print(n)

#The answer may slightly vary due to rounding off
    values.
```

R code Exa 9.15.a Finding minimum equal sample sizes

```
#Page 525

alpha<-1-0.98
p1<-0.5
p2<-0.5
z<-qnorm(alpha/2,lower.tail =FALSE)
E<-0.05
n<-(z**2)*(p1*(1-p1)+p2*(1-p2))/E**2
print(n)

#The answer may slightly vary due to rounding off
    values.
```

R code Exa 9.15.b Finding minimum equal sample sizes

```
#Page 525

```

```
alpha<-1-0.98
p1<-0.2
p2<-0.3
z<-qnorm(alpha/2,lower.tail=FALSE)
E<-0.05
n<-(z**2)*(p1*(1-p1)+p2*(1-p2))/E**2
print(n)

#The answer may slightly vary due to rounding off
   values.
```

Chapter 10

Correlation and Regression

R code Exa 10.1 Linear Correlation Coefficient

```
#Page 544
x<-c(68,69,70,70,71,72,72,72,73,73,74,75)
y<-c
   (151,146,157,164,171,160,163,180,170,175,178,188)
print(cor(x,y))
```

R code Exa 10.2 The Least Squares Regression Line

```
#Page 564
x<-c(2,2,6,8,10)
y<-c(0,1,2,3,3)
matrix<-cbind(x,y)
print(matrix)
plot(x, y)
fit <- lm(y~x)
print(fit)
abline(fit,col='blue')
```

R code Exa 10.3.a Scatter Diagram

```
#Page 566
x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

plot(x,y)
```

R code Exa 10.3.b Linear Correlation Coefficient

```
#Page 566
x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

print(cor(x,y))
```

R code Exa 10.3.c Least Squares Regression Line

```
#Page 566
x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

plot(x, y)
fit <- lm(y~x)
abline(fit,col='blue')
print(fit)
```

R code Exa 10.3.e Correlation and Regression

```
#Page 566
x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

k<-lm(y~x)
print(k)
b0<-coef(k)["(Intercept)"]
b1<-coef(k)["x"]
x_1<-4
y_1<-print((b1*x_1)+b0)
```

R code Exa 10.3.f Correlation and Regression

```
#Page 566
x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

k<-lm(y~x)
print(k)
b0<-coef(k)["(Intercept)"]
b1<-coef(k)["x"]
x_1<-20
y_1<-print((b1*x_1)+b0)
```

R code Exa 10.4.a Sum of Squared Errors

```
#Page 572
x<-c(2,2,6,8,10)
y<-c(0,1,2,3,3)
fit<-lm(y~x)
print(fit)
beta_0<-coef(fit)["(Intercept)"]
beta_1<-coef(fit)["x"]
find<-function(m){
  y<-c((beta_1*x)+beta_0)
}
k<-find(m)

library("Metrics")
print(sse(k,y))
```

R code Exa 10.4.b Sum of Squared Errors

```
#Page 572
x<-c(2,2,6,8,10)
y<-c(0,1,2,3,3)
sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5])
sum_of_y<-sum(y[1],y[2],y[3],y[4],y[5])
sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x
   [5]**2)
sum_of_y_sq<-sum(y[1]**2,y[2]**2,y[3]**2,y[4]**2,y
   [5]**2)
n<-length(x)
SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
SS_yy<-(sum_of_y_sq)-((1/n)*(sum_of_y)**2)
sum_of_xy<-sum(x[1]*y[1],x[2]*y[2],x[3]*y[3],x[4]*y
   [4],x[5]*y[5])
SS_xy<-(sum_of_xy)-((1/n)*(sum_of_x)*(sum_of_y))
beta_1<-SS_xy/SS_xx
SSE<-print(SS_yy-((beta_1)*SS_xy))
```

R code Exa 10.5 Sum of Squared Errors

```
#Page 573
x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c
  (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)
fit<-lm(y~x)
print(fit)
beta_0<-32.83
beta_1<--2.05
find<-function(x){
  y<-c((beta_1*x)+beta_0)
}
k<-find(x)
library("Metrics")
print(sse(k,y))
```

R code Exa 10.6 Statistical Inferences About beta1

```
#Page 583
x<-c(2,2,6,8,10)
y<-c(0,1,2,3,3)
fit<-lm(y~x)
print(fit)
beta_0<--0.125
beta_1<-0.3438
find<-function(x){
  y<-c((beta_1*x)+beta_0)
}
k<-find(x)
library("Metrics")
```

```
SSE<-sse(k,y)
sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5])
sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x[5]**2)
n<-length(x)
SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
S_e<-sqrt(SSE/(n-2))
alpha<-1-0.95
df<-n-2
t<-qt(alpha/2,df)
Upper_interval<-print(beta_1+((t)*(S_e)/sqrt(SS_xx)))
Lower_interval<-print(beta_1-((t)*(S_e)/sqrt(SS_xx)))
```

R code Exa 10.7 Statistical Inferences About beta1

```
#Page 584
x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c(28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)
fit<-lm(y~x)
print(fit)
beta_0<-32.83
beta_1<--2.05
find<-function(x){
  y<-c((beta_1*x)+beta_0)
}
k<-find(x)
library("Metrics")
SSE<-sse(k,y)
sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5],x[6],x[7],x[8],x[9],x[10])
sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x
```

```
[5]**2,x[6]**2,x[7]**2,x[8]**2,x[9]**2,x[10]**2)
n<-length(x)
SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
S_e<-sqrt(SSE/(n-2))
alpha<-1-0.90
df<-n-2
t<-qt(alpha/2,df)
Upper_interval<-print(beta_1+((t)*(S_e)/sqrt(SS_xx)))
Lower_interval<-print(beta_1-((t)*(S_e)/sqrt(SS_xx)))
```

R code Exa 10.8 Critical value test

```
#Page 586
x<-c(2,2,6,8,10)
y<-c(0,1,2,3,3)
fit<-lm(y~x)
print(fit)
beta_0<--0.1250
beta_1<-0.3438
alpha<-0.02
find<-function(x){
  y<-c((beta_1*x)+beta_0)
}
k<-find(x)
library("Metrics")
SSE<-sse(k,y)
sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5])
sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x[5]**2)
n<-length(x)
df<-n-2
SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
S_e<-sqrt(SSE/(n-2))
```

```
t<-beta_1/((S_e)/sqrt(SS_xx))
print(t)
t0<-qt(alpha/2,df)
t1<--t0

x=seq(-6,6,length=100)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(t1,6,length=100)
y=dnorm(x,mean=0,sd=1)
polygon(c(t1,x,6),c(0,y,0),col="gray")

x=seq(-6,t0,length=100)
y=dnorm(x,mean=0,sd=1)
polygon(c(-6,x,t0),c(0,y,0),col="gray")


points(t,0,pch=19,col="black",cex=1.5)

if(t>t0 & t<t1){
  print("ACCEPT NULL HYPOTHESIS")
}else{
  print("REJECT NULL HYPOTHESIS")
}
```

R code Exa 10.9 Critical value test

```
#Page 588
x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c(28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)
plot(x, y)
n<-length(x)
```

```
fit <- lm(y~x)
beta_1_init<--1.1
beta_0<-32.83
beta_1<--2.05
alpha<-0.05
df<-n-2
find<-function(x){
  y<-c((beta_1*x)+beta_0)
}
k<-find(x)
library("Metrics")
SSE<-sse(k,y)
sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5],x[6],x[7],x[8],x[9],x[10])
sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x[5]**2,x[6]**2,x[7]**2,x[8]**2,x[9]**2,x[10]**2)

SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
S_e<-sqrt(SSE/(n-2))

t<-(beta_1-beta_1_init)/((S_e)/sqrt(SS_xx))
print(t)

t0<-qt(alpha,df)

x=seq(-3,3,length=500)
y=dnorm(x,mean=0,sd=1)
plot(x,y,type="l",lwd=2,col="black")

x=seq(-3,t0,length=500)
y=dnorm(x,mean=0,sd=1)
polygon(c(-3,x,t0),c(0,y,0),col="gray")

points(t,0,pch=19,col="red",cex=1.5)

if(t0>t){
  print("REJECT NULL HYPOTHESIS")
}else{
```

```
  print("ACCEPT NULL HYPOTHESIS")
}
```

R code Exa 10.10 Proportion of the variability

```
#Page 599
x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)
r<-print(cor(x,y))
variability<-100*(r**2)
print(variability)
```

R code Exa 10.11 Coefficient of determination

```
#Page 600
x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)
fit <- lm(y~x)
n<-length(x)
beta_0<-32.83
beta_1<--2.05
alpha<-0.05
df<-n-2
find<-function(x){
  y<-c((beta_1*x)+beta_0)
}
k<-find(x)
library("Metrics")
SSE<-sse(k,y)
```

```
sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5],x[6],x[7],x
    [8],x[9],x[10])
sum_of_y<-sum(y[1],y[2],y[3],y[4],y[5],y[6],y[7],y
    [8],y[9],y[10])
sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x
    [5]**2,x[6]**2,x[7]**2,x[8]**2,x[9]**2,x[10]**2)
sum_of_y_sq<-sum(y[1]**2,y[2]**2,y[3]**2,y[4]**2,y
    [5]**2,y[6]**2,y[7]**2,y[8]**2,y[9]**2,y[10]**2)
SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
SS_yy<-(sum_of_y_sq)-((1/n)*(sum_of_y)**2)
S_e<-sqrt(SSE/(n-2))
sum_of_xy<-sum(x[1]*y[1],x[2]*y[2],x[3]*y[3],x[4]*y
    [4],x[5]*y[5],x[6]*y[6],x[7]*y[7],x[8]*y[8],x[9]*
    y[9],x[10]*y[10])
SS_xy<-(sum_of_xy)-((1/n)*(sum_of_x)*(sum_of_y))

r1_sq<-print((SS_yy-SSE)/SS_yy)
r2_sq<-print((SS_xy**2)/((SS_xx)*(SS_yy)))
r3_sq<-print(beta_1*(SS_xy/SS_yy))
```

R code Exa 10.12 Construction of Confidence Level

```
#Page 609

x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c
    (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

fit <- lm(y~x)
n<-length(x)
beta_0<-32.83
beta_1<--2.05
alpha<-0.05
df<-n-2

```

```
find<-function(x){
  y<-c((beta_1*x)+beta_0)
}
k<-find(x)
library("Metrics")
SSE<-sse(k,y)
sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5],x[6],x[7],x
   [8],x[9],x[10])
sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x
   [5]**2,x[6]**2,x[7]**2,x[8]**2,x[9]**2,x[10]**2)
SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
S_e<-sqrt(SSE/(n-2))


x_p<-3.5
y_p<-(beta_1*x_p)+beta_0
t<-qt(alpha/2,df,lower.tail = FALSE)
m<-x_p-mean(x)
lower_interval<-print(y_p-(t*S_e)*sqrt((1/n)+(m**2/
   SS_xx)))
upper_interval<-print(y_p+(t*S_e)*sqrt((1/n)+(m**2/
   SS_xx)))
```

R code Exa 10.13 Construction of Confidence Level

```
#Page 610

x<-c(2,3,3,3,4,4,5,5,5,6)
y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)
fit <- lm(y~x)
n<-length(x)
beta_0<-32.83
beta_1<--2.05
```

```
alpha<-0.05
df<-n-2

find<-function(x){
  y<-c((beta_1*x)+beta_0)
}
k<-find(x)
library("Metrics")
SSE<-sse(k,y)
sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5],x[6],x[7],x
   [8],x[9],x[10])
sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x
   [5]**2,x[6]**2,x[7]**2,x[8]**2,x[9]**2,x[10]**2)
SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
S_e<-sqrt(SSE/(n-2))


x_p<-3.5
y_p<-(beta_1*x_p)+beta_0
t<-qt(alpha/2,df,lower.tail = FALSE)
m<-x_p-mean(x)
lower_interval<-print(y_p-(t*S_e)*sqrt(1+(1/n)+(m**2
   /SS_xx)))
upper_interval<-print(y_p+(t*S_e)*sqrt(1+(1/n)+(m**2
   /SS_xx)))
```

Chapter 11

Chi Square Tests and F Tests

R code Exa 11.1 Chi Square Test

```
#Page 640
col1<-c(35,6)
col2<-c(12,24)
col3<-c(5,18)
n<-100
R_T<-c(col1[1]+col2[1]+col3[1],col1[2]+col2[2]+col3[2])
C_T<-c(sum(col1),sum(col2),sum(col3))

alpha<-0.01
df<-(length(R_T)-1)*(length(C_T)-1)

find_E<-function(i,k){
  E<-(R_T[i]*C_T[k])/n
}
test_col1<-function(m,i,k){
  test<-((col1[m]-find_E(i,k))**2)/find_E(i,k)
}
test_col2<-function(m,i,k){
  test<-((col2[m]-find_E(i,k))**2)/find_E(i,k)
}
```

```
test_col3<-function(m,i,k){
  test<-((col3[m]-find_E(i,k))**2)/find_E(i,k)
}

chisq<-sum(test_col1(1,1,1),test_col1(2,2,1),test_
    col2(1,1,2),test_col2(2,2,2),test_col3(1,1,3),
    test_col3(2,2,3))
print(chisq)

crit_val<-print(qchisq(alpha,df,lower.tail = FALSE))

if(crit_val<chisq){
  print("REJECT NULL HYPOTHESIS")
} else {
  print("ACCEPT NULL HYPOTHESIS")
}

x=seq(7,33,length=100)
y=dchisq(x,df)
plot(x,y,type="l",lwd=2,col="black")

x=seq(crit_val,35,length=50)
y=dchisq(x,df)
polygon(c(crit_val,x,35),c(0,y,0),col="gray")

points(chisq,0,pch=19,col="red",cex=1.5)
```

R code Exa 11.2 Chi Square Test

```
#Page 655

alpha<-0.01

assumed_dist<-c(0.743,0.216,0.012,0.012,0.008,0.009)
observed_freq<-c(1732,538,32,42,133,23)
```

```
expected_freq<-c(2500*assumed_dist)
print(cbind(assumed_dist,observed_freq,expected_freq))

test<-function(i){
  test<-(((observed_freq[i]-expected_freq[i])**2)/(expected_freq[i]))
}
chisq<-sum(test(1),test(2),test(3),test(4),test(5),test(6))
print(chisq)


df<-length(assumed_dist)-1

crit_val<-print(qchisq(alpha,df,lower.tail = FALSE))

if(crit_val<chisq){
  print("REJECT NULL HYPOTHESIS")
} else {
  print("ACCEPT NULL HYPOTHESIS")
}

x=seq(10,35,length=1000)
y=dchisq(x,df)
plot(x,y,type="l",lwd=2,col="black")

x=seq(crit_val,35,length=50)
y=dchisq(x,df)
polygon(c(crit_val,x,35),c(0,y,0),col="gray")

points(chisq,0,pch=19,col="red",cex=1.5)
```

R code Exa 11.3.a F test

```
#Page 666
df1<-5
df2<-4
alpha<-0.10
print(qf(alpha,df1,df2,lower.tail = FALSE))
```

R code Exa 11.3.b F test

```
#Page 666
df1<-5
df2<-4
alpha<-0.95
print(qf(alpha,df1,df2,lower.tail = FALSE))
```

R code Exa 11.4.a F Test Critical Values

```
#Page 668

df1<-2
df2<-20
alpha<-0.05
print(qf(alpha,df1,df2,lower.tail = FALSE))
```

R code Exa 11.4.b F Test Critical Values

```
#Page 668

df1<-2
df2<-20
alpha<-0.05
print(qf(alpha/2,df1,df2,lower.tail = FALSE))
```

R code Exa 11.4.c F Test Critical Values

```
#Page 668

df1<-2
df2<-20
alpha<-0.05
print(qf(1-alpha,df1,df2,lower.tail = FALSE))
```

R code Exa 11.4.d F Test Critical Values

```
#Page 668

df1<-2
df2<-20
alpha<-0.05
print(qf(1-alpha/2,df1,df2,lower.tail = FALSE))

#The answer may slightly vary due to rounding off values.
```

R code Exa 11.5.a F Test Critical Values

```
#Page 671

df1<-13
df2<-8
alpha<-0.01
k<-qf(1-alpha,df2,df1,lower.tail = FALSE)
```

```
print(1/k)

#The answer may slightly vary due to rounding off values.
```

R code Exa 11.5.b F Test Critical Values

```
#Page 671

df1<-40
df2<-10
alpha<-0.975
k<-qf(1-alpha,df1,df2,lower.tail = FALSE)
print(1/k)


#The answer may slightly vary due to rounding off values.
```

R code Exa 11.6 F test

```
#Page 674

n1<-16
n2<-21
var1<-2.09
var2<-1.10
alpha<-0.10
df1<-n1-1
df2<-n2-1

f0<-print(qf(1-alpha/2,df1,df2,lower.tail=FALSE))
f1<-print(qf(alpha/2,df1,df2,lower.tail=FALSE))
```

```
t<-var1/var2
print(t)

x=seq(-1,4,length=1000)
y=df(x,df1,df2)
plot(x,y,type="l",lwd=2,col="black")
x=seq(-1,f0,length=50)
y=df(x,df1,df2)
polygon(c(-1,x,f0),c(0,y,0),col="gray")
x=seq(f1,4,length=50)
y=df(x,df1,df2)
polygon(c(f1,x,4),c(0,y,0),col="gray")
points(t,0,pch=19,col="red",cex=1.5)

if(t>f0 & t<f1){
  print("ACCEPT NULL HYPOTHESIS")
} else {
  print("REJECT NULL HYPOTHESIS")
}
```

R code Exa 11.7 F test

```
#Page 676

n1<-16
n2<-21
var1<-2.09
var2<-1.10
alpha<-0.10
df1<-n1-1
df2<-n2-1

f1<-print(qf(alpha,df1,df2,lower.tail=FALSE))

```

```
t<-var1/var2
print(t)

x=seq(-1,4,length=1000)
y=df(x,df1,df2)
plot(x,y,type="l",lwd=2,col="black")

x=seq(f1,4,length=50)
y=df(x,df1,df2)
polygon(c(f1,x,4),c(0,y,0),col="gray")
points(t,0,pch=17,col="red",cex=1.5)

if(t<f1){
  print("ACCEPT NULL HYPOTHESIS")
} else {
  print("REJECT NULL HYPOTHESIS")
}
```

R code Exa 11.8 F Test using MST and MSE

```
#Page 690
maths<-c(2.59,3.13,2.97,2.50,2.53,3.29,2.53,3.17,2.70,3.88,2.64)
english<-c(3.64,3.19,3.15,3.78,3.03,2.61,3.20,3.30,3.54,3.25,4.00)
education<-c(4.00,3.59,2.80,2.39,3.47,3.59,3.74,3.77,3.13,3.00,3.47)
biology<-c(2.78,3.51,2.65,3.16,2.94,2.32,2.58,3.21,3.23,3.57,3.22)

print(cbind(maths,english,education,biology))
```

```
n1<-length(maths)
n2<-length(english)
n3<-length(education)
n4<-length(biology)

x1<-mean(maths)
x2<-mean(english)
x3<-mean(education)
x4<-mean(biology)

var1<-var(maths)
var2<-var(english)
var3<-var(education)
var4<-var(biology)

n<-44
K<-4

x<-(x1*n1+x2*n2+x3*n3+x4*n4)/(n1+n2+n3+n4)

MST<-((n1*((x1-x)**2))+(n2*((x2-x)**2))+(n3*((x3-x)*
    *2))+(n4*((x4-x)**2)))/(K-1)
MSE<-(((n1-1)*(var1))+((n2-1)*(var2))+((n3-1)*(var3)
    )+((n4-1)*(var4)))/(n-K)

f0<-print(MST/MSE)

alpha<-0.05
df1<-K-1
df2<-n-K
f<-qf(alpha,df1,df2,lower.tail = FALSE)
print(f)

if(f<f0){
  print("REJECT NULL HYPOTHESIS")
} else {
  print("ACCEPT NULL HYPOTHESIS")
```

```
}

x=seq(-1,4,length=1000)
y=df(x,df1,df2)
plot(x,y,type="l",lwd=2,col="black")

x=seq(f,4,length=50)
y=df(x,df1,df2)
polygon(c(f,x,4),c(0,y,0),col="gray")

points(f0,0,pch=18,col="red",cex=2)
```

R code Exa 11.9 Chi Square Tests and F Tests

```
#Page 693

t1<-c(71,72,75,80,60,65,63,78,75,73,72,65,63,69,64,71)
t2<-c(77,67,79,78,81,72,71,84,91)
t3<-c(81,79,73,71,75,84,77,67)

n1<-length(t1)
n2<-length(t2)
n3<-length(t3)

x1<-mean(t1)
x2<-mean(t2)
x3<-mean(t3)

var1<-var(t1)
var2<-var(t2)
var3<-var(t3)

x<-(x1*n1+x2*n2+x3*n3)/(n1+n2+n3)
alpha<-0.01
```

```
n<-33
K<-3
df1<-K-1
df2<-n-K

MST<-((n1*((x1-x)**2))+(n2*((x2-x)**2))+(n3*((x3-x)**2)))/(K-1)
MSE<-(((n1-1)*(var1))+((n2-1)*(var2))+((n3-1)*(var3)))/(n-K)
f0<-print(MST/MSE)

f<-qf(alpha,df1,df2,lower.tail = FALSE)
print(f)

if(f<f0){
  print("REJECT NULL HYPOTHESIS")
} else {
  print("ACCEPT NULL HYPOTHESIS")
}
x=seq(0,8,length=100)
y=df(x,10,10)
plot(x,y,type="l",lwd=2,col="black")
x=seq(f,10,length=1000)
y=df(x,10,10)
polygon(c(f,x,f),c(0,y,0),col="gray")
points(f0,0,pch=19,col="red",cex=1.5)
```