

R Textbook Companion for
Introductory Statistics
by Douglas S Shafer and Zhiyi Zhang¹

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

August 11, 2020

¹Funded 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	4
1 Introduction	5
2 Descriptive Statistics	6
3 Basic Concepts of Probability	14
4 Discrete Random Variables	36
5 Continuous Random Variables	50
6 Sampling Distributions	62
7 Estimation	70
8 Testing Hypotheses	75
9 Two Sample Problems	87
10 Correlation and Regression	98
11 Chi Square Tests and F Tests	111

List of R Codes

Exa 1.1	Set Notation	5
Exa 1.2	Data Frequency Table	5
Exa 2.1	Operations on the Data Set	6
Exa 2.2	Mean of Sample Data	6
Exa 2.3	Sample Mean of Students GPA	7
Exa 2.4	Measures of Variability	7
Exa 2.5	Sample Median of Data Set	7
Exa 2.6	Sample Median of Students GPA	7
Exa 2.7	Sample Median of Frequency Distribution Table	8
Exa 2.8	Sample Mode of Data Set	8
Exa 2.9	Multiple Modes in Data Set	8
Exa 2.10	Range of Data Set	9
Exa 2.11	Standard Deviation and Variance in a Data Set	9
Exa 2.12	Standard Deviation and Variance of Students GPA	9
Exa 2.13	Measures of Central Location	10
Exa 2.14	Quartiles in a Data Set	10
Exa 2.15	Quartiles of Students GPA	11
Exa 2.16	Box Plots and Interquartile Range	11
Exa 2.17	Z Scores of Observations	12
Exa 2.18	Computing Data from given Standard Deviation and Variance	12
Exa 3.1	Sample Space for Tossing a Coin	14
Exa 3.2	Sample Space for Rolling a Die and a few Cases	14
Exa 3.3.a	Tossing two coins sample space	15
Exa 3.3.b	Tossing two Distinguishable coins	15
Exa 3.4	Sample Space that describes all Three Child Families	15
Exa 3.5	Probability for Tossing a Coin	16
Exa 3.6	Probability of tossing a Single Die and a few cases	16

Exa 3.7	Probability of two tossed coins matching	17
Exa 3.8.a	Probability of occurrence of Event B	17
Exa 3.8.b	Probability of occurrence of Event M	17
Exa 3.8.c	Probability of occurrence of Event N	18
Exa 3.9.a	Probability of occurrence of Event B	18
Exa 3.9.b	Probability of occurrence of Event MF	19
Exa 3.9.c	Probability of occurrence of Event FN	19
Exa 3.10	Events concerned with the experiment of rolling a single die	20
Exa 3.11	Probability Calculation	20
Exa 3.12	Finding the Intersection of two Events	21
Exa 3.13.a	Intersection of two events while rolling a fair die	21
Exa 3.13.b	Probability of intersection of two events	22
Exa 3.14	Complement of an Event	22
Exa 3.15	Finding the Union of two Events	23
Exa 3.16	Finding the Union and Intersection of two Events	23
Exa 3.17.a	Probability that both dice show a 4 when two dice are shown	23
Exa 3.17.b	Finding probability of at least one die showing a 4 when two dice are thrown	24
Exa 3.18	Intersections and Unions of Probabilities	24
Exa 3.19.a	Conditional Probability	25
Exa 3.19.b	Conditional Probability	25
Exa 3.19.c	Conditional Probability	26
Exa 3.20.a	Fair Die	27
Exa 3.20.b	Fair Die	27
Exa 3.21.a	Conditional Probability and Independent Events	28
Exa 3.21.b	Conditional Probability and Independent Events	29
Exa 3.22.a	Conditional Probability and Independent Events	29
Exa 3.22.b	Conditional Probability and Independent Events	30
Exa 3.22.c	Conditional Probability and Independent Events	31
Exa 3.23	Test for Independent Events	31
Exa 3.24	Test for Independent Events	31
Exa 3.25.a	Conditional Probability and Independent Events	32
Exa 3.25.b	Conditional Probability and Independent Events	32
Exa 3.26.a	Complement Rule of Probability	33
Exa 3.26.b	Rule of Independence Events	33
Exa 3.27	Repetitive Probabilities of Independent Events	33

Exa 3.28.a	Repetitive Probabilities Case 1	33
Exa 3.28.b	Repetitive Probabilities Case 2	34
Exa 3.28.c	Repetitive Probabilities Case 3	34
Exa 4.1.a	Probability Distribution of Discrete Variables	36
Exa 4.1.b	Probability Distribution of Discrete Variables	36
Exa 4.2.a	Probability Distribution of Discrete Variables	37
Exa 4.2.b	Probability of Discrete Random Variable	38
Exa 4.2.c	Probability Distribution for Tossing Two Fair Dice	38
Exa 4.3	Mean of Discrete Random Variable	39
Exa 4.4.a	Probability Distribution of Discrete Random Variable	39
Exa 4.4.b	Probability of Discrete Random Variable Case	40
Exa 4.4.c	Expected Values of Discrete Variables	41
Exa 4.5	Expected Values of Discrete Variable	42
Exa 4.6.a	Finding missing values in Probability Distribution of Discrete Variable	43
Exa 4.6.b	Probability of Discrete Random Variable from Probability Distribution	43
Exa 4.6.c	Finding Probability from Probability Distribution of Discrete Variables	43
Exa 4.6.d	Finding Probability from Probability Distribution of Discrete Variables	44
Exa 4.6.e	Finding Probability from Probability Distribution of Discrete Variables	44
Exa 4.6.f	Finding Mean from Probability Distribution of Discrete Variables	44
Exa 4.6.g	Finding Variance from Probability Distribution of Discrete Variables	45
Exa 4.6.h	Finding SD from Probability Distribution of Discrete Variables	45
Exa 4.7.a	Constructing Probability Distribution from Binomial Values	46
Exa 4.7.b	Maximum of the values	46
Exa 4.7.c	Finding Mean from Probability Distribution of Discrete Variables	47
Exa 4.8	Finding Mean and SD of the random variable	48
Exa 4.9.a	Probability using Binomial Distribution	48
Exa 4.9.b	Probability using Binomial Distribution	49
Exa 5.1.a	Probability of Continuous Random Variable	50

Exa 5.1.b	Continuous Random Variables	50
Exa 5.1.c	Continuous Random Variables	51
Exa 5.2	Computing Probability from Area of Density Function Curve	51
Exa 5.3	Computing Probability from Area of Density Function Curve	52
Exa 5.4.a	Probability of Standard Normal Variable	52
Exa 5.4.b	Probability of Standard Normal Variable	52
Exa 5.5.a	Probability of Standard Normal Variable	53
Exa 5.5.b	Probability of Standard Normal Variable	53
Exa 5.6.a	Finding Probabilities of Normally Distributed Variable	53
Exa 5.6.b	Finding Probabilities of Normally Distributed Variable	54
Exa 5.7.a	Probability of Standard Normal Variable	54
Exa 5.7.b	Probability of Standard Normal Variable	55
Exa 5.8.a	Probability of Standard Normal Variable	55
Exa 5.8.b	Probability of Standard Normal Variable	55
Exa 5.8.c	Probability of Standard Normal Variable	56
Exa 5.9.a	Probability Computations for General Normal Random Variables	56
Exa 5.9.b	Probability Computations for General Normal Random Variables	57
Exa 5.10	Finding Probability of General Normal Variable	57
Exa 5.11	Finding Probability of General Normal Variable	58
Exa 5.12	Areas of Tails of Distributions	59
Exa 5.13	Areas of Tails of Distributions	59
Exa 5.14	Right and Left tails of Area in the Standard Normal Distribution	59
Exa 5.15	Area of Tails of Distribution with Mean and SD	59
Exa 5.16	Area of Tails of Distribution with Mean and SD	60
Exa 5.17	Area of Tails of Distribution with Mean and SD	60
Exa 5.18	Range of Area of Tail Distributions	61
Exa 6.1	Finding Probability Distribution and Mean and SD of the sample	62
Exa 6.2	Finding Mean and SD	63
Exa 6.3.a	Finding Mean and SD of the Sample Distribution	63
Exa 6.3.b	Sampling Distribution Probabilities	64
Exa 6.3.c	Sampling Distribution Probabilities	64
Exa 6.4	Computing Mean from Sampling Distribution	65

Exa 6.5	Probability of Sample Mean of Sampling Distribution	65
Exa 6.6.a	Probability of Sampling Distribution	66
Exa 6.6.b	Probability of Sampling Distribution	66
Exa 6.7.a	Normal Distribution	67
Exa 6.7.b	The Sample Proportion	67
Exa 6.8.a	Computing Sample Proportion	68
Exa 6.8.b	Sampling Distributions	68
Exa 6.8.c	Finding Sample Proportion	69
Exa 7.1.a	Construction of Confidence Level	70
Exa 7.1.b	Construction of Confidence Level	70
Exa 7.2.a	Critical Values of Confidence Levels	70
Exa 7.2.b	Critical Values of Confidence Levels	71
Exa 7.3	Finding Critical Values	71
Exa 7.4	Construction of Confidence Level	71
Exa 7.5	Small Sample Estimation of a Population Mean	72
Exa 7.6	Observed Significance	72
Exa 7.7	Large Sample Estimation of a Population Proportion	72
Exa 7.8	Computing Minimal Sample Size with E	73
Exa 7.9	Estimation of Minimal Sample Size	73
Exa 7.10.a	Estimation of Minimal Sample Size	73
Exa 7.10.b	Estimation of Minimal Sample Size	74
Exa 7.11	Estimation of Minimal Sample Size	74
Exa 8.3	The Elements of Hypothesis Testing	75
Exa 8.4	Large Sample Tests for a Population Mean	76
Exa 8.5	Large Sample Tests for a Population Mean	76
Exa 8.6	Observed Significance of a Test	77
Exa 8.7	Observed Significance of a Test	78
Exa 8.8	Observed Significance of a Test	79
Exa 8.9	Observed Significance of a Test	80
Exa 8.10	Small Sample Tests for a Population Mean	81
Exa 8.11	Small Sample Tests for a Population Mean	81
Exa 8.12	Large Sample Tests for a Population Proportion	82
Exa 8.13	Large Sample Tests for a Population Proportion	83
Exa 8.14	p value approach	84
Exa 8.15	p value approach	85
Exa 9.1	Construction of a point estimate and confidence interval	87
Exa 9.2	Comparison of Two Population Means using critical value approach	87

Exa 9.3	p value approach	88
Exa 9.4	Construction of a point estimate and confidence interval	89
Exa 9.5	Critical value approach	90
Exa 9.7	Construction of a point estimate and confidence interval	91
Exa 9.8	Critical value approach	92
Exa 9.10	Construction of a point estimate and confidence interval	93
Exa 9.11	Critical value approach	93
Exa 9.12	p value approach	94
Exa 9.13	Sample Size Considerations	95
Exa 9.14	Sample Size Considerations	96
Exa 9.15.a	Finding minimum equal sample sizes	96
Exa 9.15.b	Finding minimum equal sample sizes	96
Exa 10.1	Linear Correlation Coefficient	98
Exa 10.2	The Least Squares Regression Line	98
Exa 10.3.a	Scatter Diagram	99
Exa 10.3.b	Linear Correlation Coefficient	99
Exa 10.3.c	Least Squares Regression Line	99
Exa 10.3.e	Correlation and Regression	100
Exa 10.3.f	Correlation and Regression	100
Exa 10.4.a	Sum of Squared Errors	100
Exa 10.4.b	Sum of Squared Errors	101
Exa 10.5	Sum of Squared Errors	102
Exa 10.6	Statistical Inferences About β_1	102
Exa 10.7	Statistical Inferences About β_1	103
Exa 10.8	Critical value test	104
Exa 10.9	Critical value test	105
Exa 10.10	Proportion of the variability	107
Exa 10.11	Coefficient of determination	107
Exa 10.12	Construction of Confidence Level	108
Exa 10.13	Construction of Confidence Level	109
Exa 11.1	Chi Square Test	111
Exa 11.2	Chi Square Test	112
Exa 11.3.a	F test	113
Exa 11.3.b	F test	114
Exa 11.4.a	F Test Critical Values	114
Exa 11.4.b	F Test Critical Values	114
Exa 11.4.c	F Test Critical Values	115
Exa 11.4.d	F Test Critical Values	115

Exa 11.5.a F Test Critical Values	115
Exa 11.5.b F Test Critical Values	116
Exa 11.6 F test	116
Exa 11.7 F test	117
Exa 11.8 F Test using MST and MSE	118
Exa 11.9 Chi Square Tests and F Tests	120

Chapter 1

Introduction

R code Exa 1.1 Set Notation

```
1 #Page 18
2 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)
3 print(age_of_students)
```

R code Exa 1.2 Data Frequency Table

```
1 #Page 19
2 dataset<-c
   (18,18,19,19,19,18,22,20,18,18,17,19,18,24,18,20,18,21,20,17,19)
3 library("plyr")
4 y<-count(dataset)
5 print(y)
```

Chapter 2

Descriptive Statistics

R code Exa 2.1 Operations on the Data Set

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

R code Exa 2.2 Mean of Sample Data

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

R code Exa 2.3 Sample Mean of Students GPA

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

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

R code Exa 2.4 Measures of Variability

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

R code Exa 2.5 Sample Median of Data Set

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

R code Exa 2.6 Sample Median of Students GPA

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

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

R code Exa 2.7 Sample Median of Frequency Distribution Table

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

R code Exa 2.8 Sample Mode of Data Set

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

R code Exa 2.9 Multiple Modes in Data Set

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

R code Exa 2.10 Range of Data Set

```
1 #Page 58
2 dataset_1<-c(40,38,42,40,39,39,43,40,39,40)
3 dataset_2<-c(46,37,40,33,42,36,40,47,34,45)
4 findrange=function(v){
5   range=max(v)-min(v)
6   print(range)
7 }
8 findrange(dataset_1)
9 findrange(dataset_2)
```

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

```
1 #Page 60
2 dataset_2<-c(46,37,40,33,42,36,40,47,34,45)
3 findvariance<-function(v){
4   mean<-sum(v)/length(v)
5   x<-sum((v-mean)^2)
6   variance<-x/(length(v)-1)
7   print(variance)
8 }
9 findvariance(dataset_2)
10 std_deviation=sd(dataset_2)
11 print(std_deviation)
```

R code Exa 2.12 Standard Deviation and Variance of Students GPA

```
1 #Page 61
```



```

2 GPA<-c
   (1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33)

3 findvariance<-function(v){
4   mean<-sum(v)/length(v)
5   x<-sum((v-mean)^2)
6   variance<-x/(length(v)-1)
7   print(variance)
8 }
9 findvariance(GPA)
10 std_deviation=sd(GPA)
11 print(std_deviation)

```

R code Exa 2.13 Measures of Central Location

```

1 #Page 70
2 GPA<-c
   (1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33)

3 k<-sort(GPA)
4 print(k)
5 index<-match(c(3.33),k)
6 print(index)
7 percentile=(index/length(GPA))*100
8 print(percentile)

```

R code Exa 2.14 Quartiles in a Data Set

```

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

3 quartile_2<-median(GPA)

```

```
4 lower_subset<-print(GPA[GPA<median(GPA)])
5 upper_subset<-print(GPA[GPA>median(GPA)])
6 quartile_1<-median(lower_subset)
7 quartile_3<-median(upper_subset)
8 print(quartile_1)
9 print(quartile_2)
10 print(quartile_3)
```

R code Exa 2.15 Quartiles of Students GPA

```
1 #Page 73
2 GPA<-c
   (1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33,3.88)

3 quartile_2<-median(GPA)
4 lower_subset<-print(GPA[GPA<median(GPA)])
5 upper_subset<-print(GPA[GPA>median(GPA)])
6 quartile_1<-median(lower_subset)
7 quartile_3<-median(upper_subset)
8 print(quartile_1)
9 print(quartile_2)
10 print(quartile_3)
```

R code Exa 2.16 Box Plots and Interquartile Range

```
1 #Page 75
2 GPA<-c
   (1.90,3.00,2.53,3.71,2.12,1.76,2.71,1.39,4.00,3.33)

3 quartile_2<-median(GPA)
4 lower_subset<-GPA[GPA<median(GPA)]
5 upper_subset<-GPA[GPA>median(GPA)]
6 quartile_1<-median(lower_subset)
```

```

7 quartile_3<-median(upper_subset)
8 print(min(GPA))
9 print(quartile_1)
10 print(quartile_2)
11 print(quartile_3)
12 print(max(GPA))
13 boxplot(GPA)
14 IQR<-print(quartile_3-quartile_1)

```

R code Exa 2.17 Z Scores of Observations

```

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

3 mean<-mean(data)
4 sd<-sd(data)
5 zscores<-function(v){
6   for(x in v){
7     z<-((x-mean(data))/sd(data))
8     print(z)
9   }
10 }
11 zscores(data)
12
13 #The answer may slightly vary due to rounding off
   values.

```

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

```

1 #Page 79
2 mean<-2.70

```

```
3 sd<-0.50
4 z1<- -0.62
5 z2<-1.28
6
7 findGPA<-function(z){
8   GPA<-mean+(sd)*(z)
9   print(GPA)
10 }
11
12 findGPA(z1)
13 findGPA(z2)
```

Chapter 3

Basic Concepts of Probability

R code Exa 3.1 Sample Space for Tossing a Coin

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

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

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

R code Exa 3.3.a Tossing two coins sample space

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

R code Exa 3.3.b Tossing two Distinguishable coins

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

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

```
1 #Page 114
2 sample_space<-c('b','g')
3 k<-expand.grid(sample_space, sample_space, sample_
  space)
4 print(k)
5
6 library("DiagrammeR")
7
8 nodes <- create_node_df(n=15, type = "number",
9   label = c("P", "b", "g", "b", "g",
10    "b", "g", "b", "g", "b", "g",
11    "b", "g", "b", "g"))
12 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)
  ,
```

```

13             rel = "leading to")
14
15 graph <- create_graph(nodes_df = nodes,
16                       edges_df = edges,
17                       attr_theme="lr")
18
19 print(render_graph(graph))

```

R code Exa 3.5 Probability for Tossing a Coin

```

1 #Page 116
2 sample_space<-c('H', 'T')
3 event_a<-c('H')
4 event_b<-c('T')
5 probability_of_a<-(length(event_a)/length(sample_
   space))
6 probability_of_b<-(length(event_b)/length(sample_
   space))
7 print(probability_of_a)
8 print(probability_of_b)

```

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

```

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

```

R code Exa 3.7 Probability of two tossed coins matching

```
1 #Page 118
2 library("sets")
3 both_heads<-c('hh')
4 both_tails<-c('tt')
5 head_and_tail<-c('ht')
6 tail_and_head<-c('th')
7 Sample_space<-as.set(c('hh','tt','ht','th'))
8 Probability_of_both_heads<-print(length(both_heads)/
  length(Sample_space))
9 Probability_of_both_tails<-print(length(both_tails)/
  length(Sample_space))
10 Overall_Probability<-print(Probability_of_both_heads+
  Probability_of_both_tails)
```

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

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

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

```
1 #Page 119
```



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

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

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

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

```
1 #Page 120
2 probability_of_white_male<-25/100
3 probability_of_white_female<-26/100
4 probability_of_black_male<-12/100
5 probability_of_black_female<-15/100
6 probability_of_hispanic_male<-6/10
7 probability_of_hispanic_female<-5/100
8 probability_of_asian_male<-3/100
9 probability_of_asian_female<-3/100
10 probability_of_others_male<-1/100
11 probability_of_others_female<-4/100
12
13 probability_of_black<-print(probability_of_black_
    male+probability_of_black_female)
```

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

```
1 #Page 120
2 probability_of_white_male<-25/100
3 probability_of_white_female<-26/100
4 probability_of_black_male<-12/100
5 probability_of_black_female<-15/100
6 probability_of_hispanic_male<-6/10
7 probability_of_hispanic_female<-5/100
8 probability_of_asian_male<-3/100
9 probability_of_asian_female<-3/100
10 probability_of_others_male<-1/100
11 probability_of_others_female<-4/100
12
13 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

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

```

11 probability_of_others_female<-4/100
12
13 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

```

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

```

R code Exa 3.11 Probability Calculation

```

1 #Page 132
2 Event_0<-c('tails','tails','tails','tails','tails')
3 print(Event_0)
4 number_of_tosses<-5
5 possible_outcomes<-2
6 Total_outcomes<-possible_outcomes^number_of_tosses
7 Probability_of_all_tails<-1/Total_outcomes
8 print(Probability_of_all_tails)

```

```

9 Probability_of_atleast_one_head<-1-Probability_of_
  all_tails
10 print(Probability_of_atleast_one_head)
11 print(Probability_of_atleast_one_head*100)
12
13 #The answer may slightly vary due to rounding off
  values.

```

R code Exa 3.12 Finding the Intersection of two Events

```

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

```

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

```

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

```

```

9 Probability_of_both_events_occurring<-(length(
  intersection)/length(sample_space))
10 print(Probability_of_both_events_occurring)

```

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

```

1 #Page 134
2 probability_of_1<-1/12
3 probability_of_6<-3/12
4 Sum_of_other_probabilities<-1-(probability_of_1+
  probability_of_6)
5 probability_of_2<-Sum_of_other_probabilities/4
6 probability_of_3<-Sum_of_other_probabilities/4
7 probability_of_4<-Sum_of_other_probabilities/4
8 probability_of_5<-Sum_of_other_probabilities/4
9 probability_of_both_events_occurring<-print(
  probability_of_4+probability_of_6)

```

R code Exa 3.14 Complement of an Event

```

1 #Page 135
2 library("sets")
3 library("gtools")
4 set<-as.set(c(1,2,3,4,5,6))
5 e<-c(2,4,6)
6 E<-as.set(e)
7 k<-set_complement(E, set)
8 possible_set_1<-print(combinations(n=3,r=3,v=c
  (1,3,5),repeats.allowed=F))
9 possible_set_2<-print(combinations(n=3,r=2,v=c
  (1,3,5),repeats.allowed=F))
10 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

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

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

```
1 #Page 138
2 Sample_space<-c('bb','bg','gb','gg')
3 Sample_B<-c('bb','bg','gb')
4 Sample_D<-c('bg','gb')
5 Sample_M<-c('bb','gg')
6 B_union_D<-print(union(Sample_B,Sample_D))
7 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

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

```

6 print(event)
7 probability<-(nrow(event)/nrow(sample_space))
8 print(probability)

```

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

```

1 #Page 139
2 library("prob")
3 sample_space<-rolldie(2)
4 print(sample_space)
5 Total_possibilitites<-nrow(sample_space)
6 num_Event_A<-6
7 num_Event_B<-6
8 num_intersection<-1
9 P_Event_A<-print(num_Event_A/Total_possibilitites)
10 P_Event_B<-print(num_Event_B/Total_possibilitites)
11 P_Intersection<-print(num_intersection/Total_
    possibilitites)
12 reqd_probability<-print((P_Event_A+P_Event_B)-P_
    Intersection)

```

R code Exa 3.18 Intersections and Unions of Probabilities

```

1 #page 140
2 Prob_of_maths<-63/100
3 Prob_of_eng<-34/100
4 Prob_of_eng_and_math<-27/100
5 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

```
1 #Page 141
2 c_s<-12
3 c_t<-1
4 c_total<-c_s+c_t
5 e_s<-4
6 e_t<-3
7 e_total<-e_s+e_t
8 m_s<-6
9 m_t<-2
10 m_total<-m_s+m_t
11 total_s<-c_s+e_s+m_s
12 total_t<-c_t+e_t+m_t
13 grand_total<-c_total+e_total+m_total
14 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)
15 colnames(given_data)<-c("S","T","Total")
16 rownames(given_data)<-c("C","E","M","Total")
17 given_data<-as.table(given_data)
18 print(given_data)
19 probability<-print(m_t/grand_total)
20 percentage<-print(probability*100)
```

R code Exa 3.19.b Conditional Probability

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



```

9 m_t<-2
10 m_total<-m_s+m_t
11 total_s<-c_s+e_s+m_s
12 total_t<-c_t+e_t+m_t
13 grand_total<-c_total+e_total+m_total
14 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)
15 colnames(given_data)<-c("S","T","Total")
16 rownames(given_data)<-c("C","E","M","Total")
17 given_data<-as.table(given_data)
18 print(given_data)
19 probability_of_m<-print(m_total/grand_total)
20 probability_of_t<-print(total_t/grand_total)
21 probability_of_m_intersect_t<-print(m_t/grand_total)
22 final_probability<-print((probability_of_m+
    probability_of_t)-probability_of_m_intersect_t)
23 percentage<-final_probability*100
24 print(percentage)

```

R code Exa 3.19.c Conditional Probability

```

1 #Page 141
2 c_s<-12
3 c_t<-1
4 c_total<-c_s+c_t
5 e_s<-4
6 e_t<-3
7 e_total<-e_s+e_t
8 m_s<-6
9 m_t<-2
10 m_total<-m_s+m_t
11 total_s<-c_s+e_s+m_s
12 total_t<-c_t+e_t+m_t
13 grand_total<-c_total+e_total+m_total

```

```

14 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)
15 colnames(given_data)<-c("S","T","Total")
16 rownames(given_data)<-c("C","E","M","Total")
17 given_data<-as.table(given_data)
18 print(given_data)
19 probability_of_m<-print(m_total/grand_total)
20 reqd_probability<-print(1-probability_of_m)

```

R code Exa 3.20.a Fair Die

```

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

```

R code Exa 3.20.b Fair Die

```

1 #Page 154
2 library("sets")

```

```

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

```

R code Exa 3.21.a Conditional Probability and Independent Events

```

1 #Page 156
2 m_e<-43
3 m_w<-293
4 m_h<-114
5 m_total<-(m_e+m_w+m_h)
6 f_e<-82
7 f_w<-299
8 f_h<-71
9 f_total<-(f_e+f_w+f_h)
10 total_e<-(m_e+f_e)
11 total_w<-(m_w+f_w)
12 total_h<-(m_h+f_h)
13 grand_total<-(m_total+f_total)
14 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)
15 colnames(given_data)<-c("E","W","H","Total")
16 rownames(given_data)<-c("M","F","Total")

```

```

17 given_data<-as.table(given_data)
18 print(given_data)
19 reqd_probability<-print(total_e/grand_total)
20 percentage<-reqd_probability*100
21 print(percentage)

```

R code Exa 3.21.b Conditional Probability and Independent Events

```

1 #Page 156
2 m_e<-43
3 m_w<-293
4 m_h<-114
5 m_total<-(m_e+m_w+m_h)
6 f_e<-82
7 f_w<-299
8 f_h<-71
9 f_total<-(f_e+f_w+f_h)
10 total_e<-(m_e+f_e)
11 total_w<-(m_w+f_w)
12 total_h<-(m_h+f_h)
13 grand_total<-(m_total+f_total)
14 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)
15 colnames(given_data)<-c("E","W","H","Total")
16 rownames(given_data)<-c("M","F","Total")
17 given_data<-as.table(given_data)
18 print(given_data)
19 reqd_probability<-print(m_e/m_total)
20 percentage<-reqd_probability*100
21 print(percentage)

```

R code Exa 3.22.a Conditional Probability and Independent Events

```

1 #Page 158
2 h_o<-0.09
3 h_not_o<-0.11
4 not_h_o<-0.02
5 not_h_not_o<-0.78
6 given_data<-matrix(c(h_o,h_not_o,not_h_o,not_h_not_o
   ),ncol=2,byrow=TRUE)
7 colnames(given_data)<-c("Overweight","Not Overweight
   ")
8 rownames(given_data)<-c("Hypertension","No
   Hypertension")
9 given_data<-as.table(given_data)
10 print(given_data)
11 reqd_probability<-print(h_o/(h_o+not_h_o))

```

R code Exa 3.22.b Conditional Probability and Independent Events

```

1 #Page 158
2 h_o<-0.09
3 h_not_o<-0.11
4 not_h_o<-0.02
5 not_h_not_o<-0.78
6 given_data<-matrix(c(h_o,h_not_o,not_h_o,not_h_not_o
   ),ncol=2,byrow=TRUE)
7 colnames(given_data)<-c("Overweight","Not Overweight
   ")
8 rownames(given_data)<-c("Hypertension","No
   Hypertension")
9 given_data<-as.table(given_data)
10 print(given_data)
11 reqd_probability<-print(h_not_o/(h_not_o+not_h_not_o
   ))

```

R code Exa 3.22.c Conditional Probability and Independent Events

```
1 #Page 158
2 p1<-0.8181818
3 p2<-0.1235955
4 print(max(p1,p2))
5 factor<-print(p1/p2)
```

R code Exa 3.23 Test for Independent Events

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

R code Exa 3.24 Test for Independent Events

```
1 #Page 161
2 m_e<-43
3 m_w<-293
4 m_h<-114
5 m_total<-(m_e+m_w+m_h)
6 f_e<-82
7 f_w<-299
8 f_h<-71
```

```

9 f_total<-(f_e+f_w+f_h)
10 total_e<-(m_e+f_e)
11 total_w<-(m_w+f_w)
12 total_h<-(m_h+f_h)
13 grand_total<-(m_total+f_total)
14 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)
15 colnames(given_data)<-c("E","W","H","Total")
16 rownames(given_data)<-c("M","F","Total")
17 given_data<-as.table(given_data)
18 print(given_data)
19 P_F<-f_total/grand_total
20 P_E<-total_e/grand_total
21 P_intersection<-print(f_e/grand_total)
22 product<-print(P_F*P_E)
23 print(product==P_intersection)

```

R code Exa 3.25.a Conditional Probability and Independent Events

```

1 #Page 162
2 P_a<-92/100
3 reqd_probability<-print(P_a*P_a)

```

R code Exa 3.25.b Conditional Probability and Independent Events

```

1 #Page 162
2 P_a<-92/100
3 reqd_probability<-print(P_a+P_a-(P_a*P_a))

```

R code Exa 3.26.a Complement Rule of Probability

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

R code Exa 3.26.b Rule of Independence Events

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

R code Exa 3.27 Repetitive Probabilities of Independent Events

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

R code Exa 3.28.a Repetitive Probabilities Case 1

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



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

R code Exa 3.28.b Repetitive Probabilities Case 2

```
1 #Page 166
2 no_of_black_marbles<-7
3 no_of_white_marbles<-3
4 total_marbles<-no_of_white_marbles+no_of_black_marbles
5
6 P_first_black<-no_of_black_marbles/total_marbles
7 P_second_not_black<-no_of_white_marbles/(total_marbles-1)
8
9 P_first_not_black<-no_of_white_marbles/total_marbles
10 P_second_black<-no_of_black_marbles/(total_marbles-1)
11
12 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

```
1 #Page 166
2 no_of_black_marbles<-7
3 no_of_white_marbles<-3
4 total_marbles<-no_of_white_marbles+no_of_black_marbles
5
6 P_first_white<-no_of_white_marbles/total_marbles
7 P_second_white<-(no_of_white_marbles-1)/(total_marbles-1)
```

```
8
9 P_both_white<-P_first_white*P_second_white
10
11 reqd_probability<-print(1-P_both_white)
```

Chapter 4

Discrete Random Variables

R code Exa 4.1.a Probability Distribution of Discrete Variables

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

R code Exa 4.1.b Probability Distribution of Discrete Variables

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

```

3 library("ggplot2")
4 sample_space<-c('hh','ht','th','tt')
5 Event_zero_heads<-as.set(c('tt'))
6 Event_one_head<-as.set(c('ht','th'))
7 Event_two_heads<-as.set(c('hh'))
8
9 x<-c(0,1,2)
10 Prob<-c((length(Event_zero_heads)/length(sample_
        space)),(length(Event_one_head)/length(sample_
        space)),(length(Event_two_heads)/length(sample_
        space)))
11 prob.dist<-cbind(x,Prob)
12 print(prob.dist)
13
14 df<-data.frame(x = x, y = Prob)
15 p <- ggplot(df,aes(x=x,after_stat(Prob))) + geom_
        histogram(binwidth=1, fill="#69b3a2", color="#
        e9ecef")
16 print(p)
17
18 reqd_probability<-sum(Prob[2],Prob[3])
19 print(reqd_probability)

```

R code Exa 4.2.a Probability Distribution of Discrete Variables

```

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

```

```

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

```

R code Exa 4.2.b Probability of Discrete Random Variable

```

1 #Page 188
2 r1<-c(2,3,4,5,6,7)
3 r2<-c(3,4,5,6,7,8)
4 r3<-c(4,5,6,7,8,9)
5 r4<-c(5,6,7,8,9,10)
6 r5<-c(6,7,8,9,10,11)
7 r6<-c(7,8,9,10,11,12)
8 dice<-matrix(c(r1,r2,r3,r4,r5,r6),nrow=6,byrow=TRUE)
9
10 sum<-c(2,3,4,5,6,7,8,9,10,11,12)
11 prob<-c(1/36,2/36,3/36,4/36,5/36,6/36,5/36,4/36,3/
    36,2/36,1/36)
12 prob.dist<-cbind(sum,prob)
13
14 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

```

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

```

```

8 r6<-c(7,8,9,10,11,12)
9 dice<-matrix(c(r1,r2,r3,r4,r5,r6),nrow=6,byrow=TRUE)
10
11 sum<-c(2,3,4,5,6,7,8,9,10,11,12)
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)
13 prob.dist<-cbind(sum,prob)
14
15 reqd_probability<-print(sum(prob[1],prob[3],prob[5],
      prob[7],prob[9],prob[11]))
16
17 df<-data.frame(x = sum, y = prob)
18
19 p <- ggplot(df,aes(x=sum,after_stat(prob))) + geom_
      histogram(binwidth=1, fill="#69b3a2", color="#
      e9ecef")
20 print(p)

```

R code Exa 4.3 Mean of Discrete Random Variable

```

1 #Page 190
2 x<-c(-2,1,2,3.5)
3 prob<-c(0.21,0.34,0.24,0.21)
4 prob.dist<-cbind(x,prob)
5 print(prob.dist)
6 product <- function(n) {
7   k<-x[n]*prob[n]
8 }
9 mean<-print(sum(product(1),product(2),product(3),
      product(4)))

```

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

```

1 #Page 191
2 total<-1000
3 fp<-300
4 sp<-200
5 tp<-100
6 los<-0
7
8 x<-fp-1
9 x1<-1
10 P_of_x<-x1/total
11
12 y<-sp-1
13 y1<-1
14 P_of_y<-y1/total
15
16 z<-tp-1
17 z1<-1
18 P_of_z<-z1/total
19
20 lose<-total-(x1+y1+z1)
21 l<-los-1
22 P_of_l<-lose/total
23
24 k<-c(x,y,z,l)
25 prob<-c(P_of_x,P_of_y,P_of_z,P_of_l)
26 prob.dist<-cbind(k,prob)
27 print(prob.dist)

```

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

```

1 #Page 191
2 total<-1000
3 fp<-300
4 sp<-200
5 tp<-100

```

```

6 los<-0
7
8 x<-fp-1
9 x1<-1
10 P_of_x<-x1/total
11
12 y<-sp-1
13 y1<-1
14 P_of_y<-y1/total
15
16 z<-tp-1
17 z1<-1
18 P_of_z<-z1/total
19
20 lose<-total-(x1+y1+z1)
21 l<-lose-1
22 P_of_l<-lose/total
23
24 k<-c(x,y,z,l)
25 prob<-c(P_of_x,P_of_y,P_of_z,P_of_l)
26
27 reqd_probability<-print(sum(prob[1],prob[2],prob[3])
  )

```

R code Exa 4.4.c Expected Values of Discrete Variables

```

1 #Page 191
2 total<-1000
3 fp<-300
4 sp<-200
5 tp<-100
6 los<-0
7
8 x<-fp-1
9 x1<-1

```



```

10 P_of_x<-x1/total
11
12 y<-sp-1
13 y1<-1
14 P_of_y<-y1/total
15
16 z<-tp-1
17 z1<-1
18 P_of_z<-z1/total
19
20 lose<-total-(x1+y1+z1)
21 l<-los-1
22 P_of_l<-lose/total
23
24 k<-c(x,y,z,l)
25 prob<-c(P_of_x,P_of_y,P_of_z,P_of_l)
26
27 product <- function(n){
28   h<-k[n]*prob[n]
29 }
30 mean<-print(sum(product(1),product(2),product(3),
   product(4)))

```

R code Exa 4.5 Expected Values of Discrete Variable

```

1 #Page 192
2 x<-c(195,-199805)
3 prob<-c(99.7/100,0.03/100)
4 prob.dist<-cbind(x,prob)
5 print(prob.dist)
6 product <- function(n){
7   k<-x[n]*prob[n]
8 }
9 mean<-print(sum(product(1),product(2)))
10

```

11 #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

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

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

```
1 #Page 194
2 x<-c(-1,0,1,4)
3 prob<-c(0.2,0.5,NA,0.1)
4 prob.dist<-cbind(x,prob)
5
6 print(prob[2])
```

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

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

```
5
6 P_of_zero<-print(sum(prob[3],prob[4]))
```

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

```
1 #Page 194
2 x<-c(-1,0,1,4)
3 prob<-c(0.2,0.5,0.2,0.1)
4 prob.dist<-cbind(x,prob)
5
6 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

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

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

```
1 #Page 194
2 x<-c(-1,0,1,4)
3 prob<-c(0.2,0.5,0.2,0.1)
4 prob.dist<-cbind(x,prob)
5
```

```

6 product <- function(n){
7   k<-x[n]*prob[n]
8 }
9 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

```

1 #Page 194
2 x<-c(-1,0,1,4)
3 prob<-c(0.2,0.5,0.2,0.1)
4 prob.dist<-cbind(x,prob)
5
6 product <- function(n){
7   k<-x[n]*prob[n]
8 }
9 mean<-sum(product(1),product(2),product(3),product
  (4))
10 func<-function(n){
11   l<-((x[n]-mean)**2)*(prob[n])
12 }
13 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

```

1 #Page 194
2 x<-c(-1,0,1,4)
3 prob<-c(0.2,0.5,0.2,0.1)
4 prob.dist<-cbind(x,prob)
5

```

```

6 product <- function(n){
7   k<-x[n]*prob[n]
8 }
9 mean<-sum(product(1),product(2),product(3),product
  (4))
10 func<-function(n){
11   l<-((x[n]-mean)**2)*(prob[n])
12 }
13 variance<-sum(func(1),func(2),func(3),func(4))
14 std_dev<-print(sqrt(variance))

```

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

```

1 #Page 210
2 n<-5
3 p<-0.17
4 q<-1-p
5
6 func<-function(n,x){
7   P<-((factorial(n)/(factorial(x)*factorial(n-x)))*(
  p**x)*(q**(n-x)))
8 }
9 var<-c(0,1,2,3,4,5)
10 prob<-c(func(5,0),func(5,1),func(5,2),func(5,3),func
  (5,4),func(5,5))
11 prob.dist<-cbind(var,prob)
12 print(prob.dist)

```

R code Exa 4.7.b Maximum of the values

```

1 #Page 210
2 library("ggplot2")

```

```

3 n<-5
4 p<-0.17
5 q<-1-p
6
7 func<-function(n,x){
8   P<-((factorial(n)/(factorial(x)*factorial(n-x)))*(
9     p**x)*(q**(n-x)))
10 }
11 var<-c(0,1,2,3,4,5)
12 prob<-c(func(5,0),func(5,1),func(5,2),func(5,3),func
13   (5,4),func(5,5))
14 prob.dist<-cbind(var,prob)
15
16 k<-which.max(prob)
17 print(var[k])
18
19 df<-data.frame(x = var, y = prob)
20 p <- ggplot(df,aes(x=var,after_stat(prob))) + geom_
21   histogram(binwidth=1, fill="#69b3a2", color="#
22     e9ecef")
23 print(p)

```

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

```

1 #Page 210
2 n<-5
3 p<-0.17
4 q<-1-p
5
6 func<-function(n,x){
7   P<-((factorial(n)/(factorial(x)*factorial(n-x)))*(
8     p**x)*(q**(n-x)))
9 }
10
11 var<-c(0,1,2,3,4,5)

```

```

10 prob<-c(func(5,0),func(5,1),func(5,2),func(5,3),func
    (5,4),func(5,5))
11 prob.dist<-cbind(var,prob)
12
13 product <- function(n){
14   k<-var[n]*prob[n]
15 }
16 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

```

1 #Page 213
2 n<-5
3 p<-0.17
4 q<-1-p
5 mean<-print(n*p)
6 std_dev<-print(sqrt(n*p*q))

```

R code Exa 4.9.a Probability using Binomial Distribution

```

1 #Page 215
2 n<-10
3 p<-0.5
4 q<-1-p
5 func<-function(n,x){
6   P<-((factorial(n)/(factorial(x)*factorial(n-x)))*(
    p**x)*(q**(n-x)))
7 }
8 P_of_k<-func(10,6)
9 print(P_of_k)

```

R code Exa 4.9.b Probability using Binomial Distribution

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

Chapter 5

Continuous Random Variables

R code Exa 5.1.a Probability of Continuous Random Variable

```
1 #Page 231
2 x=seq(0,1,length=200)
3 y=dunif(x,min=0,max=1)
4 plot(x,y,type="l",xlim=c(0,3),ylim=c(0,2),lwd=2,col="red",ylab="p")
5
6 x=seq(0.75,1,length=100)
7 y=dunif(x,min=0,max=1)
8 polygon(c(0.75,x,1),c(0,y,0),col="lightgray",border=NA)
9
10 full<-qunif(1,min=0,max=1)
11 limit<-qunif(0.75,min=0,max=1)
12 print(full-limit)
```

R code Exa 5.1.b Continuous Random Variables

```
1 #Page 231
```

```

2 x=seq(0,1,length=200)
3 y=dunif(x,min=0,max=1)
4 plot(x,y,type="l",xlim=c(0,3),ylim=c(0,2),lwd=2,col="
  "red",ylab="p")
5
6 x=seq(0,0.2,length=100)
7 y=dunif(x,min=0,max=1)
8 polygon(c(0,x,0.2),c(0,y,0),col="lightgray",border =
  NA)
9
10 print(qunif(0.2,min=0,max=1))

```

R code Exa 5.1.c Continuous Random Variables

```

1 #Page 231
2 x=seq(0,1,length=200)
3 y=dunif(x,min=0,max=1)
4 plot(x,y,type="l",xlim=c(0,3),ylim=c(0,2),lwd=2,col="
  "red",ylab="p")
5
6 x=seq(0.4,0.7,length=100)
7 y=dunif(x,min=0,max=1)
8 polygon(c(0.4,x,0.7),c(0,y,0),col="lightgray",border
  = NA)
9
10 l1<-qunif(0.7,min=0,max=1)
11 l2<-qunif(0.4,min=0,max=1)
12 print(l1-l2)

```

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

```

1 #Page 233

```

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

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

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

R code Exa 5.4.a Probability of Standard Normal Variable

```
1 #Page 243
2
3 x=seq(-4,4,length=200)
4 y=dnorm(x,mean=0,sd=1)
5 plot(x,y,type="l",lwd=2,col="red")
6 print(pnorm(1.48,mean=0,sd=1))
7
8 x=seq(-4,1.5,length=200)
9 y=dnorm(x)
10 polygon(c(-4,x,1.48),c(0,y,0),col="gray")
```

R code Exa 5.4.b Probability of Standard Normal Variable

```
1 #Page 243
2
3 x=seq(-4,4,length=200)
```

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

R code Exa 5.5.a Probability of Standard Normal Variable

```
1 #Page 244
2
3 x=seq(-4,4, length=200)
4 y=dnorm(x, mean=0, sd=1)
5 plot(x,y, type="l", lwd=2, col="red")
6 reqd_probability<-1-(pnorm(1.60, mean=0, sd=1))
7 print(reqd_probability)
8
9 x=seq(4,1.5, length=200)
10 y=dnorm(x)
11 polygon(c(-4,x,1.6), c(0,y,0), col="gray")
```

R code Exa 5.5.b Probability of Standard Normal Variable

```
1 #Page 244
2
3 x=seq(-4,4, length=200)
4 y=dnorm(x, mean=0, sd=1)
5 reqd_probability<-1-(pnorm(-1.02, mean=0, sd=1))
6 print(reqd_probability)
```

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

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

```

3 y=dnorm(x, mean=0, sd=1)
4 plot(x,y, type="l", lwd=2, col="red")
5 reqd_probability<-print(pnorm(1.57, mean=0, sd=1) -
      pnorm(0.5, mean=0, sd=1))
6
7
8 x=seq(0.5, 1.57, length=200)
9 y=dnorm(x)
10 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

```

1 #Page 246
2 x=seq(-4, 4, length=200)
3 y=dnorm(x, mean=0, sd=1)
4 plot(x,y, type="l", lwd=2, col="red")
5 reqd_probability<-print(pnorm(0.09, mean=0, sd=1) -
      pnorm(-2.55, mean=0, sd=1))
6
7 x=seq(-2.55, 0.09, length=200)
8 y=dnorm(x)
9 polygon(c(-2.55, x, 0.09), c(0, y, 0), col="gray")

```

R code Exa 5.7.a Probability of Standard Normal Variable

```

1 #Page 248
2 x=seq(-5, 5, length=200)
3 y=dnorm(x, mean=0, sd=1)
4 plot(x,y, type="l", lwd=2, col="red")
5 reqd_probability<-print(pnorm(4.16, mean=0, sd=1) -
      pnorm(1.13, mean=0, sd=1))
6
7 x=seq(1.13, 4.16, length=200)

```

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

R code Exa 5.7.b Probability of Standard Normal Variable

```
1 #Page 248
2 x=seq(-4,4,length=200)
3 y=dnorm(x,mean=0,sd=1)
4 plot(x,y,type="l",lwd=2,col="red")
5 reqd_probability<-print(pnorm(2.15,mean=0,sd=1)-
  pnorm(-5.22,mean=0,sd=1))
6
7 x=seq(-5.22,2.15,length=200)
8 y=dnorm(x)
9 polygon(c(-5.22,x,2.15),c(0,y,0),col="gray")
```

R code Exa 5.8.a Probability of Standard Normal Variable

```
1 #Page 249
2 x=seq(-4,4,length=200)
3 y=dnorm(x,mean=0,sd=1)
4 plot(x,y,type="l",lwd=2,col="red")
5 reqd_probability<-print(pnorm(1,mean=0,sd=1)-pnorm
  (-1,mean=0,sd=1))
6
7 x=seq(-1,1,length=200)
8 y=dnorm(x)
9 polygon(c(-1,x,1),c(0,y,0),col="gray")
```

R code Exa 5.8.b Probability of Standard Normal Variable

```

1 #Page 249
2 x=seq(-4,4,length=200)
3 y=dnorm(x,mean=0,sd=1)
4 plot(x,y,type="l",lwd=2,col="red")
5 reqd_probability<-print(pnorm(2,mean=0,sd=1)-pnorm
      (-2,mean=0,sd=1))
6
7 x=seq(-2,2,length=200)
8 y=dnorm(x)
9 polygon(c(-2,x,2),c(0,y,0),col="gray")

```

R code Exa 5.8.c Probability of Standard Normal Variable

```

1 #Page 249
2 x=seq(-4,4,length=200)
3 y=dnorm(x,mean=0,sd=1)
4 plot(x,y,type="l",lwd=2,col="red")
5 reqd_probability<-print(pnorm(3,mean=0,sd=1)-pnorm
      (-3,mean=0,sd=1))
6
7 x=seq(-3,3,length=200)
8 y=dnorm(x)
9 polygon(c(-3,x,3),c(0,y,0),col="gray")

```

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

```

1 #Page 256
2 mean<-10
3 sd<-2.5
4 k<-(14-mean)/sd
5 print(k)
6

```

```

7 x=seq(-4,4,length=200)
8 y=dnorm(x,mean=0,sd=1)
9 plot(x,y,type="l",lwd=2,col="red")
10 reqd_probability<-print(pnorm(k,mean=0,sd=1))
11
12 x=seq(-4,k,length=200)
13 y=dnorm(x)
14 polygon(c(-4,x,k),c(0,y,0),col="gray")

```

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

```

1 #Page 256
2 mean<-10
3 sd<-2.5
4 k<-(8-mean)/sd
5 l<-(14-mean)/sd
6 print(k)
7 print(l)
8 x=seq(-4,4,length=200)
9 y=dnorm(x,mean=0,sd=1)
10 plot(x,y,type="l",lwd=2,col="red")
11 reqd_probability<-print(pnorm(l,mean=0,sd=1)-pnorm(k
    ,mean=0,sd=1))
12
13 x=seq(k,l,length=500)
14 y=dnorm(x)
15 polygon(c(k,x,l),c(0,y,0),col="gray")

```

R code Exa 5.10 Finding Probability of General Normal Variable

```

1 #Page 258
2 mean<-37.5

```



```

3 sd<-4.5
4 k<-(30-mean)/sd
5 l<-(40-mean)/sd
6 print(k)
7 print(l)
8 x=seq(-4,4,length=200)
9 y=dnorm(x,mean=0,sd=1)
10 plot(x,y,type="l",lwd=2,col="red")
11 reqd_probability<-print(pnorm(l,mean=0,sd=1)-pnorm(k
    ,mean=0,sd=1))
12
13 x=seq(k,l,length=500)
14 y=dnorm(x)
15 polygon(c(k,x,l),c(0,y,0),col="gray")

```

R code Exa 5.11 Finding Probability of General Normal Variable

```

1 #Page 260
2
3 mean<-510
4 sd<-60
5 k<-(650-mean)/sd
6 print(k)
7
8 x=seq(-4,4,length=200)
9 y=dnorm(x,mean=0,sd=1)
10 plot(x,y,type="l",lwd=2,col="red")
11 reqd_probability<-print(1-(pnorm(k,mean=0,sd=1)))
12
13 x=seq(4,k,length=200)
14 y=dnorm(x)
15 polygon(c(4,x,k),c(0,y,0),col="gray")

```

R code Exa 5.12 Areas of Tails of Distributions

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

R code Exa 5.13 Areas of Tails of Distributions

```
1 #Page 272
2 k<-qnorm(0.025, lower.tail = FALSE)
3 print(k)
4
5 x=seq(-4,4, length=200)
6 y=dnorm(x, mean=0, sd=1)
7 plot(x,y, type="l", lwd=2, col="red")
8
9 x=seq(4, k, length=200)
10 y=dnorm(x)
11 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

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

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

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

```

3 mean<-10
4 sd<-2.5
5 k<-mean+(z)*(sd)
6 print(k)
7
8 x=seq(5,16,length=200)
9 y=dnorm(x,mean=10,sd=2)
10 plot(x,y,type="l",lwd=2,col="red")
11
12 x=seq(5,k,length=200)
13 y=dnorm(x,mean=10,sd=2)
14 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

```

1 #Page 277
2 z<-qnorm(1-0.65,lower.tail = TRUE)
3 mean<-175
4 sd<-12
5 k1<-mean+(z)*(sd)
6 print(k1)
7
8 x=seq(-4,4,length=200)
9 y=dnorm(x,mean=0,sd=1)
10 plot(x,y,type="l",lwd=2,col="red")
11
12 x=seq(z,4,length=500)
13 y=dnorm(x)
14 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

```

1 #Page 278

```

```

2 z<-qnorm(1-0.05,lower.tail = TRUE)
3 mean<-510
4 sd<-60
5 x<-mean+(z)*(sd)
6 print(x)
7
8 x=seq(-4,4,length=200)
9 y=dnorm(x,mean=0,sd=1)
10 plot(x,y,type="l",lwd=2,col="red")
11
12 x=seq(z,4,length=500)
13 y=dnorm(x)
14 polygon(c(z,x,4),c(0,y,0),col="gray")

```

R code Exa 5.18 Range of Area of Tail Distributions

```

1 #Page 280
2 mean<-29
3 sd<-2
4 alpha<-1-0.75
5 z1<-qnorm(1-alpha/2,lower.tail = FALSE)
6 z2<-qnorm(1-alpha/2,lower.tail = TRUE)
7 x_fast<-mean+(z1)*(sd)
8 x_slow<-mean+(z2)*(sd)
9 print(x_fast)
10 print(x_slow)
11
12 x=seq(-4,4,length=200)
13 y=dnorm(x,mean=0,sd=1)
14 plot(x,y,type="l",lwd=2,col="red")
15
16 x=seq(z1,z2,length=500)
17 y=dnorm(x)
18 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

```
1 #Page 289
2 set<-c(152,156,160,164)
3 l<-print.table(combn(set,2))
4 mean<-function(n){
5   m<-(l[1,n]+l[2,n])/2
6   a<-l[1,n]
7   b<-l[2,n]
8   print(paste("Mean of",a,"and",b," is ",m))
9 }
10 mean(1)
11 mean(2)
12 mean(3)
13 mean(4)
14 mean(5)
15 mean(6)
16
17 x<-c(152,154,156,158,160,162,164)
18 prob<-c(1/16,2/16,3/16,4/16,3/16,2/16,1/16)
19 prob.dist<-cbind(x,prob)
20 print(prob.dist)
```

```

21
22 product<- function(i){
23   k<-x[i]*prob[i]
24 }
25 mean_of_dist<-sum(product(1),product(2),product(3),
26   product(4),product(5),product(6),product(7))
27 print(mean_of_dist)
28
29 func<-function(i){
30   k<-x[i]*x[i]*prob[i]
31 }
32 variance<-sum(func(1),func(2),func(3),func(4),func
33   (5),func(6),func(7))
34 print(variance)
35
36 sd<-sqrt(variance-(mean_of_dist**2))
37 print(sd)

```

R code Exa 6.2 Finding Mean and SD

```

1 #Page 291
2 mean<-13525
3 sd<-4180
4 n<-100
5 print(mean)
6 print(sd/sqrt(n))

```

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

```

1 #Page 298
2 n<-50
3 mean<-112
4 sd<-40

```

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

R code Exa 6.3.b Sampling Distribution Probabilities

```
1 #Page 298
2 n<-50
3 mean<-112
4 sd<-40
5 std_new<-sd/sqrt(n)
6 k<-(110-mean)/std_new
7 l<-(114-mean)/std_new
8 print(k)
9 print(l)
10 x=seq(-4,4,length=200)
11 y=dnorm(x,mean=0,sd=1)
12 plot(x,y,type="l",lwd=2,col="red")
13 reqd_probability<-print(pnorm(l,mean=0,sd=1)-pnorm(k
    ,mean=0,sd=1))
14
15 x=seq(k,l,length=500)
16 y=dnorm(x)
17 polygon(c(k,x,l),c(0,y,0),col="gray")
```

R code Exa 6.3.c Sampling Distribution Probabilities

```
1 #Page 298
2 n<-50
3 mean<-112
4 sd<-40
5 std_new<-sd/sqrt(n)
6 k<-(113-mean)/std_new
7 print(k)
```

```

8 x=seq(-4,4,length=200)
9 y=dnorm(x,mean=0,sd=1)
10 reqd_probability<-print(1-(pnorm(k,mean=0,sd=1)))
11
12 #The answer may slightly vary due to rounding off
    values.

```

R code Exa 6.4 Computing Mean from Sampling Distribution

```

1 #Page 300
2 n<-100
3 mean<-2.61
4 sd<-0.5
5 std_new<-sd/sqrt(n)
6 k<-(2.51-mean)/std_new
7 l<-(2.71-mean)/std_new
8 print(k)
9 print(l)
10 x=seq(-4,4,length=200)
11 y=dnorm(x,mean=0,sd=1)
12 plot(x,y,type="l",lwd=2,col="red")
13 reqd_probability<-print(pnorm(l,mean=0,sd=1)-pnorm(k
    ,mean=0,sd=1))
14
15 x=seq(k,l,length=500)
16 y=dnorm(x)
17 polygon(c(k,x,l),c(0,y,0),col="gray")

```

R code Exa 6.5 Probability of Sample Mean of Sampling Distribution

```

1 #Page 302
2 n<-5
3 mean<-38.5

```



```

4 sd<-2.5
5 std_new<-sd/sqrt(n)
6 k<-(36-mean)/std_new
7 print(k)
8 x=seq(-4,4,length=200)
9 y=dnorm(x,mean=0,sd=1)
10 reqd_probability<-print(pnorm(k,mean=0,sd=1))
11
12 #The answer may slightly vary due to rounding off
    values.

```

R code Exa 6.6.a Probability of Sampling Distribution

```

1 #Page 303
2
3 n<-5
4 mean<-50
5 sd<-6
6 k<-(48-mean)/sd
7 print(k)
8 x=seq(-4,4,length=200)
9 y=dnorm(x,mean=0,sd=1)
10 reqd_probability<-print(pnorm(k,mean=0,sd=1))
11
12 #The answer may slightly vary due to rounding off
    values.

```

R code Exa 6.6.b Probability of Sampling Distribution

```

1 #Page 303
2
3 n<-36
4 mean<-50

```

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

R code Exa 6.7.a Normal Distribution

```
1 #Page 314
2 n<-900
3 p<-0.38
4 q<-1-p
5 mean<-p
6 print(mean)
7 std_dev<-sqrt((p*q)/n)
8 print(std_dev)
9 l<-3*std_dev
10 print(l)
11 upper_limit<-p+l
12 lower_limit<-p-l
13 print(upper_limit)
14 print(lower_limit)
15 print("Normally Distributed")
```

R code Exa 6.7.b The Sample Proportion

```
1 #Page 314
2
3 p<-0.38
4 q<-1-p
5 n<-900
```

```

6 mean<-0.38
7 sd<-sqrt((p*q)/n)
8 k<-((0.38-0.05)-mean)/sd
9 l<-((0.38+0.05)-mean)/sd
10 print(k)
11 print(l)
12 x=seq(-4,4,length=200)
13 y=dnorm(x,mean=0,sd=1)
14 reqd_probability<-print(pnorm(l,mean=0,sd=1)-pnorm(k
,mean=0,sd=1))

```

R code Exa 6.8.a Computing Sample Proportion

```

1 #Page 316
2 x<-102
3 n<-121
4 sample_proportion<-x/n
5 print(sample_proportion)

```

R code Exa 6.8.b Sampling Distributions

```

1 #Page 316
2 n<-121
3 p<-0.90
4 q<-1-p
5 mean<-p
6 print(mean)
7 std_dev<-sqrt((p*q)/n)
8 print(std_dev)
9 l<-3*std_dev
10 print(l)
11 upper_limit<-p+l
12 lower_limit<-p-l

```

```
13 print(upper_limit)
14 print(lower_limit)
15 print("Normally Distributed")
```

R code Exa 6.8.c Finding Sample Proportion

```
1 #Page 316
2 x<-102
3 n<-121
4 mean<-0.90
5 sd<-0.027
6 sample_proportion<-x/n
7 k<-(sample_proportion-mean)/sd
8 print(k)
9 x=seq(-4,4,length=200)
10 y=dnorm(x,mean=0,sd=1)
11 reqd_probability<-print(pnorm(k,mean=0,sd=1))
12
13 #The answer may slightly vary due to rounding off
    values.
```

Chapter 7

Estimation

R code Exa 7.1.a Construction of Confidence Level

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

R code Exa 7.1.b Construction of Confidence Level

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

R code Exa 7.2.a Critical Values of Confidence Levels

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

R code Exa 7.2.b Critical Values of Confidence Levels

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

R code Exa 7.3 Finding Critical Values

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

R code Exa 7.4 Construction of Confidence Level

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

R code Exa 7.5 Small Sample Estimation of a Population Mean

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

R code Exa 7.6 Observed Significance

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

R code Exa 7.7 Large Sample Estimation of a Population Proportion

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

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

R code Exa 7.8 Computing Minimal Sample Size with E

```
1 #Page 363
2 alpha<-1-0.99
3 sd<-1.3
4 E<-0.2
5 z<-qnorm(alpha/2,lower.tail = FALSE)
6 n<-print((z**2)*(sd**2)/(E**2))
7
8 #The answer may slightly vary due to rounding off
  values.
```

R code Exa 7.9 Estimation of Minimal Sample Size

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

R code Exa 7.10.a Estimation of Minimal Sample Size

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



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

R code Exa 7.10.b Estimation of Minimal Sample Size

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

R code Exa 7.11 Estimation of Minimal Sample Size

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

Chapter 8

Testing Hypotheses

R code Exa 8.3 The Elements of Hypothesis Testing

```
1 #Page 380
2 mean<-8
3 sd<-0.15
4 n<-5
5 sd_x<-sd/sqrt(n)
6 alpha<-0.10
7 z<-qnorm(alpha/2,lower.tail = FALSE)
8 lower_critical_value<-print(mean-(z*sd_x))
9 upper_critical_value<-print(mean+(z*sd_x))
10
11 x=seq(7.5,8.5,length=100)
12 y=dnorm(x,mean=mean,sd=sd)
13 plot(x,y,type="l",lwd=2,col="black")
14
15 x=seq(7.5,lower_critical_value,length=100)
16 y=dnorm(x,mean = mean,sd=sd)
17 polygon(c(7.5,x,lower_critical_value),c(0,y,0),col="
    gray")
18
19 x=seq(upper_critical_value,8.5,length=100)
20 y=dnorm(x,mean=mean,sd=sd)
```

```
21 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

```
1 #Page 390
2 mean<-3.5
3 mean_exp<-3.1
4 sd<-1.5
5 alpha<-0.05
6 n<-50
7 critical_value<-qnorm(alpha,lower.tail = TRUE)
8 z<-print((mean_exp-mean)/(sd/sqrt(n)))
9
10 x=seq(-4,4,length=500)
11 y=dnorm(x,mean=0,sd=1)
12 plot(x,y,type="l",lwd=2,col="black")
13
14 x=seq(-4,critical_value,length=500)
15 y=dnorm(x,mean=0,sd=1)
16 polygon(c(-4,x,critical_value),c(0,y,0),col="gray")
17 points(z,0,pch=19,col="red",cex=1.5)
18
19 if(z<critical_value){
20   print("REJECT NULL HYPOTHESIS")
21 }else{
22   print("ACCEPT NULL HYPOTHESIS")
23 }
```

R code Exa 8.5 Large Sample Tests for a Population Mean

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

```

3 mean_exp<-8.2
4 sd<-0.22
5 n<-30
6 alpha<-0.01
7 z0<-qnorm(alpha/2,lower.tail = TRUE)
8 z1<-qnorm(alpha/2,lower.tail = FALSE)
9 z<-print((mean_exp-mean)/(sd/sqrt(n)))
10
11 x=seq(-3,3,length=500)
12 y=dnorm(x,mean=0,sd=1)
13 plot(x,y,type="l",lwd=2,col="black")
14
15 x=seq(-3,z0,length=500)
16 y=dnorm(x,mean=0,sd=1)
17 polygon(c(-3,x,z0),c(0,y,0),col="gray")
18
19 x=seq(z1,3,length=500)
20 y=dnorm(x,mean=0,sd=1)
21 polygon(c(z1,x,3),c(0,y,0),col="gray")
22
23 points(z,0,pch=19,col="red",cex=1.5)
24
25 if(z<z1 & z>z0){
26   print("ACCEPT NULL HYPOTHESIS")
27 }else{
28   print("REJECT NULL HYPOTHESIS")
29 }

```

R code Exa 8.6 Observed Significance of a Test

```

1 #Page 401
2 alpha<-0.01
3 z<-qnorm(alpha/2,lower.tail = FALSE)
4
5 x=seq(-3,3,length=500)

```

```

6 y=dnorm(x, mean=0, sd=1)
7 plot(x,y, type="l", lwd=2, col="black")
8
9 x=seq(z,3, length=500)
10 y=dnorm(x, mean =0, sd=1)
11 polygon(c(z,x,3), c(0,y,0), col="gray")
12
13 points(z,0, pch=19, col="red", cex=1)
14
15 auc<-pnorm(Inf)-pnorm(z)
16 significance<-2*auc
17 print(significance)
18
19 #The answer may slightly vary due to rounding off
    values.

```

R code Exa 8.7 Observed Significance of a Test

```

1 #Page 403
2 mean<-202.5
3 mean_exp<-199.2
4 sd<-19.63
5 alpha<-0.05
6 n<-85
7 z<-print((mean_exp-mean)/(sd/sqrt(n)))
8
9 p<-pnorm(z)-pnorm(-Inf)
10 print(p)
11
12 x=seq(-3,3, length=500)
13 y=dnorm(x, mean=0, sd=1)
14 plot(x,y, type="l", lwd=2, col="black")
15
16 x=seq(-3,z, length=500)
17 y=dnorm(x, mean=0, sd=1)

```

```

18 polygon(c(-3,x,z),c(0,y,0),col="gray")
19 points(z,0,pch=19,col="black",cex=1.5)
20
21
22 if(p<alpha){
23   print("REJECT NULL HYPOTHESIS")
24 }else{
25   print("ACCEPT NULL HYPOTHESIS")
26 }

```

R code Exa 8.8 Observed Significance of a Test

```

1 #Page 405
2 mean<-67
3 mean_exp<-69.4
4 sd<-6.1
5 alpha<-0.01
6 n<-64
7 z<-print((mean_exp-mean)/(sd/sqrt(n)))
8
9 p<-pnorm(Inf)-pnorm(z)
10 print(p)
11
12 x=seq(-3,3.5,length=100)
13 y=dnorm(x,mean=0,sd=1)
14 plot(x,y,type="l",lwd=2,col="black")
15
16 x=seq(z,3.5,length=100)
17 y=dnorm(x,mean=0,sd=1)
18 polygon(c(z,x,3.5),c(0,y,0),col="gray")
19 points(z,0,pch=19,col="black",cex=1.5)
20
21
22 if(p<alpha){
23   print("REJECT NULL HYPOTHESIS")

```

```
24 }else{
25   print("ACCEPT NULL HYPOTHESIS")
26 }
```

R code Exa 8.9 Observed Significance of a Test

```
1 #Page 407
2
3 mean<-7.4
4 mean_exp<-7.7
5 sd<-0.5
6 alpha<-0.01
7 n<-30
8 z<-print((mean_exp-mean)/(sd/sqrt(n)))
9
10 p<-2*(pnorm(Inf)-pnorm(z))
11 print(p)
12
13 x=seq(-5,5,length=100)
14 y=dnorm(x,mean=0,sd=1)
15 plot(x,y,type="l",lwd=2,col="black")
16
17 x=seq(z,5,length=100)
18 y=dnorm(x,mean=0,sd=1)
19 polygon(c(z,x,5),c(0,y,0),col="gray")
20
21 x=seq(-5,-z,length=100)
22 y=dnorm(x,mean=0,sd=1)
23 polygon(c(-5,x,-z),c(0,y,0),col="gray")
24
25 points(z,0,pch=19,col="black",cex=1.5)
26
27
28 if(p<alpha){
29   print("REJECT NULL HYPOTHESIS")
```

```
30 }else{
31   print("ACCEPT NULL HYPOTHESIS")
32 }
```

R code Exa 8.10 Small Sample Tests for a Population Mean

```
1 #Page 416
2 data<-c(155,179,175,175,161)
3 mean_exp<-mean(data)
4 mean<-179
5 sd<-sd(data)
6 n<-length(data)
7 df<-n-1
8 alpha<-0.05
9 t<-print((mean_exp-mean)/(sd/sqrt(n)))
10 t_alpha<-print(qt(alpha,df))
11
12 x=seq(-3,3,length=500)
13 y=dnorm(x,mean=0,sd=1)
14 plot(x,y,type="l",lwd=2,col="black")
15
16 x=seq(-3,t,length=500)
17 y=dnorm(x,mean=0,sd=1)
18 polygon(c(-3,x,t),c(0,y,0),col="gray")
19 points(t,0,pch=19,col="red",cex=1)
20
21 if(t_alpha>t){
22   print("REJECT NULL HYPOTHESIS")
23 }else{
24   print("ACCEPT NULL HYPOTHESIS")
25 }
```

R code Exa 8.11 Small Sample Tests for a Population Mean


```

1 #Page 418
2 data<-c(0.021,0.019,0.023,0.020)
3 mean_exp<-mean(data)
4 mean<-0.02
5 sd<-sd(data)
6 n<-length(data)
7 df<-n-1
8 alpha<-0.01
9 t<-print((mean_exp-mean)/(sd/sqrt(n)))
10 t_a1<-print(qt(alpha/2,df))
11 t_a2<-print(-t_a1)
12
13 x=seq(-8,8,length=100)
14 y=dnorm(x,mean=0,sd=2.5)
15 plot(x,y,type="l",lwd=2,col="black")
16
17 x=seq(-8,t_a1,length=100)
18 y=dnorm(x,mean=0,sd=2.5)
19 polygon(c(-8,x,t_a1),c(0,y,0),col="gray")
20
21 x=seq(t_a2,8,length=100)
22 y=dnorm(x,mean=0,sd=2.5)
23 polygon(c(t_a2,x,8),c(0,y,0),col="gray")
24
25 points(t,0,pch=17,col="red",cex=1.5)
26
27 if(t>t_a1 & t<t_a2){
28   print("ACCEPT NULL HYPOTHESIS")
29 }else{
30   print("REJECT NULL HYPOTHESIS")
31 }

```

R code Exa 8.12 Large Sample Tests for a Population Proportion

```
1 #Page 429
```

```

2 n<-500
3 soft_drink_maker<-270
4 competitor<-211
5 non_deciders<-19
6 alpha<-0.05
7 p<-soft_drink_maker/500
8 p_0<-0.5
9 q_0<-1-p_0
10 z<-pnorm((p-p_0)/(sqrt((p_0*q_0)/n)))
11 z0<-qnorm(alpha,lower.tail=FALSE)
12 print(z0)
13
14 x=seq(-3,3,length=500)
15 y=dnorm(x,mean=0,sd=1)
16 plot(x,y,type="l",lwd=2,col="black")
17
18 x=seq(z0,3,length=500)
19 y=dnorm(x,mean=0,sd=1)
20 polygon(c(z0,x,3),c(0,y,0),col="gray")
21
22 points(z,0,pch=19,col="red",cex=1.5)
23
24 if(z0<z){
25   print("REJECT NULL HYPOTHESIS")
26 }else{
27   print("ACCEPT NULL HYPOTHESIS")
28 }

```

R code Exa 8.13 Large Sample Tests for a Population Proportion

```

1 #Page 432
2 n<-5000
3 alpha<-0.05
4 p<-0.5255
5 p_0<-0.5146

```

```

6 q_0<-1-p_0
7 z<-print((p-p_0)/(sqrt((p_0*q_0)/n)))
8 z0<-qnorm(alpha,lower.tail=FALSE)
9 z1<-qnorm(alpha,lower.tail=TRUE)
10 print(z0)
11 print(z1)
12
13 x=seq(-3,3,length=500)
14 y=dnorm(x,mean=0,sd=1)
15 plot(x,y,type="l",lwd=2,col="black")
16
17 x=seq(-3,z1,length=500)
18 y=dnorm(x,mean=0,sd=1)
19 polygon(c(-3,x,z1),c(0,y,0),col="gray")
20
21 x=seq(z0,3,length=500)
22 y=dnorm(x,mean=0,sd=1)
23 polygon(c(z0,x,3),c(0,y,0),col="gray")
24
25 points(z,0,pch=19,col="red",cex=1.5)
26
27 if(z>z1 & z<z0){
28   print("ACCEPT NULL HYPOTHESIS")
29 }else{
30   print("REJECT NULL HYPOTHESIS")
31 }

```

R code Exa 8.14 p value approach

```

1 #Page 435
2 n<-500
3 soft_drink_maker<-270
4 competitor<-211
5 non_deciders<-19
6 alpha<-0.05

```

```

7 p<-soft_drink_maker/500
8 p_0<-0.5
9 q_0<-1-p_0
10 z<-print((p-p_0)/(sqrt((p_0*q_0)/n)))
11
12 p<-pnorm(Inf)-pnorm(z)
13 print(p)
14
15 x=seq(-3,3,length=500)
16 y=dnorm(x,mean=0,sd=1)
17 plot(x,y,type="l",lwd=2,col="black")
18
19 x=seq(z,3,length=500)
20 y=dnorm(x,mean=0,sd=1)
21 polygon(c(z,x,3),c(0,y,0),col="gray")
22
23 points(z,0,pch=19,col="black",cex=1.5)
24
25 if(p<alpha){
26   print("REJECT NULL HYPOTHESIS")
27 }else{
28   print("ACCEPT NULL HYPOTHESIS")
29 }

```

R code Exa 8.15 p value approach

```

1 #Page 436
2 n<-5000
3 alpha<-0.05
4 p<-0.5255
5 p_0<-0.5146
6 q_0<-1-p_0
7 z<-print((p-p_0)/(sqrt((p_0*q_0)/n)))
8
9 p<-2*(pnorm(Inf)-pnorm(z))

```

```
10 print(p)
11
12 x=seq(-3,3,length=500)
13 y=dnorm(x,mean=0,sd=1)
14 plot(x,y,type="l",lwd=2,col="black")
15
16 x=seq(z,3,length=500)
17 y=dnorm(x,mean=0,sd=1)
18 polygon(c(z,x,3),c(0,y,0),col="gray")
19
20 points(z,0,pch=19,col="black",cex=1.5)
21
22 if(p<alpha){
23   print("REJECT NULL HYPOTHESIS")
24 }else{
25   print("ACCEPT NULL HYPOTHESIS")
26 }
```

Chapter 9

Two Sample Problems

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

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

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

```

1 #Page 450
2 n1<-174
3 n2<-355
4 x1<-3.51
5 x2<-3.24
6 s1<-0.51
7 s2<-0.52
8 alpha<-0.01
9 m<-(((s1**2)/n1)+((s2**2)/n2))
10 z<-((x1-x2)/sqrt(m))
11 print(z)
12 z0<-qnorm(alpha,lower.tail = FALSE)
13 print(z0)
14 x=seq(-6,6,length=500)
15 y=dnorm(x,mean=0,sd=1)
16 plot(x,y,type="l",lwd=2,col="black")
17
18 x=seq(z0,6,length=500)
19 y=dnorm(x,mean=0,sd=1)
20 polygon(c(z0,x,6),c(0,y,0),col="gray")
21
22 points(z,0,pch=19,col="red",cex=1.5)
23
24 if(z0<z){
25   print("REJECT NULL HYPOTHESIS")
26 }else{
27   print("ACCEPT NULL HYPOTHESIS")
28 }

```

R code Exa 9.3 p value approach

```

1 #Page 452
2
3 n1<-174
4 n2<-355

```

```

5 x1<-3.51
6 x2<-3.24
7 s1<-0.51
8 s2<-0.52
9 alpha<-0.01
10 m<-(((s1**2)/n1)+((s2**2)/n2))
11 z<-((x1-x2)/sqrt(m))
12 print(z)
13
14 p<-pnorm(Inf)-pnorm(z)
15 print(p)
16
17 x=seq(-6,6,length=500)
18 y=dnorm(x,mean=0,sd=1)
19 plot(x,y,type="l",lwd=2,col="black")
20
21 x=seq(z,6,length=500)
22 y=dnorm(x,mean=0,sd=1)
23 polygon(c(z,x,6),c(0,y,0),col="gray")
24
25 points(z,0,pch=19,col="red",cex=1.5)
26
27 if(p<alpha){
28   print("REJECT NULL HYPOTHESIS")
29 }else{
30   print("ACCEPT NULL HYPOTHESIS")
31 }

```

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

```

1 #Page 468
2
3 n1<-11
4 x1<-52
5 sd1<-12

```



```

6 n2<-6
7 x2<-46
8 sd2<-10
9 alpha<-0.05
10 df<-n1+n2-2
11 t<-qt(alpha/2,df)
12 n<-(((n1-1)*(sd1**2))+((n2-1)*(sd2**2)))
13 d<-(n1+n2-2)
14
15 s_p_sq<-n/d
16 print(s_p_sq)
17
18 lower_interval<-print((x1-x2)+t*sqrt(s_p_sq*(1/n1+1/
    n2)))
19 upper_interval<-print((x1-x2)-t*sqrt(s_p_sq*(1/n1+1/
    n2)))

```

R code Exa 9.5 Critical value approach

```

1 #Page 470
2
3 n1<-11
4 x1<-52
5 sd1<-12
6 n2<-6
7 x2<-46
8 sd2<-10
9 alpha<-0.005
10 df<-n1+n2-2
11 n<-(((n1-1)*(sd1**2))+((n2-1)*(sd2**2)))
12 d<-(n1+n2-2)
13 s_p_sq<-n/d
14
15 t<-(x1-x2)/(sqrt(s_p_sq*((1/n1)+(1/n2))))
16 t0<-qt(alpha,df)

```

```

17 t1<--t0
18 print(t)
19 print(t0)
20 print(t1)
21
22 x=seq(-4,4,length=500)
23 y=dnorm(x,mean=0,sd=1)
24 plot(x,y,type="l",lwd=2,col="black")
25
26 x=seq(t1,4,length=500)
27 y=dnorm(x,mean=0,sd=1)
28 polygon(c(t1,x,4),c(0,y,0),col="gray")
29
30 x=seq(-4,t0,length=500)
31 y=dnorm(x,mean=0,sd=1)
32 polygon(c(-4,x,t0),c(0,y,0),col="gray")
33
34 points(t,0,pch=19,col="red",cex=1.5)
35
36 if(t<t1 & t>t0){
37   print("ACCEPT NULL HYPOTHESIS")
38 }else{
39   print("REJECT NULL HYPOTHESIS")
40 }

```

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

```

1 #Page 484
2
3 car_1<-c(17,13.2,35.3,13.6,32.7,18.4,22.5,26.8,15.1)
4 car_2<-c(17,12.9,35.4,13.2,32.5,18.1,22.5,26.7,15.0)
5
6 diff<-c(car_1-car_2)
7 diff_sq<-c(diff**2)
8 n<-length(car_1)

```

```

9 d<-sum(diff)/n
10 sd<-(sqrt(sum(diff_sq)-((1/n)*(sum(diff)**2))))/(
      sqrt(n-1))
11
12 df<-n-1
13 alpha<-0.05
14 t<-qt(alpha/2,df,lower.tail = FALSE)
15
16 lower_interval<-print(d-(t*sd/sqrt(n)))
17 upper_interval<-print(d+(t*sd/sqrt(n)))

```

R code Exa 9.8 Critical value approach

```

1 #Page 487
2
3 car_1<-c(17,13.2,35.3,13.6,32.7,18.4,22.5,26.8,15.1)
4 car_2<-c(17,12.9,35.4,13.2,32.5,18.1,22.5,26.7,15.0)
5
6 diff<-c(car_1-car_2)
7 diff_sq<-c(diff**2)
8 n<-length(car_1)
9 d<-sum(diff)/n
10
11 t<-d/(sd/sqrt(n))
12 print(t)
13
14 alpha<-0.05
15 df<-n-1
16 t0<-qt(alpha,df,lower.tail = FALSE)
17 print(t0)
18
19 x=seq(-4,4,length=500)
20 y=dnorm(x,mean=0,sd=1)
21 plot(x,y,type="l",lwd=2,col="black")
22

```

```

23 x=seq(t0,4,length=500)
24 y=dnorm(x,mean=0,sd=1)
25 polygon(c(t0,x,4),c(0,y,0),col="gray")
26
27 points(t,0,pch=19,col="red",cex=1.5)
28
29 if(t<t0){
30   print("ACCEPT NULL HYPOTHESIS")
31 }else{
32   print("REJECT NULL HYPOTHESIS")
33 }

```

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

```

1 #Page 503
2
3 n1<-500
4 n2<-100
5 p1<-0.67
6 p2<-0.80
7
8 alpha<-1-0.9
9 z<-qnorm(alpha/2,lower.tail = FALSE)
10 m<-sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
11 lower_interval<-print((p1-p2)-(z*m))
12 upper_interval<-print((p1-p2)+(z*m))
13
14 #The answer may slightly vary due to rounding off
    values.

```

R code Exa 9.11 Critical value approach

```

1 #Page 507

```

```

2
3 n1<-500
4 n2<-100
5 p1<-0.67
6 p2<-0.80
7 d0<- -0.05
8 alpha<-1-0.9
9 m<-sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
10 z<-(p1-p2-d0)/m
11 print(z)
12
13 z0<-qnorm(alpha,lower.tail = TRUE)
14 print(z0)
15
16 x=seq(-3,3,length=500)
17 y=dnorm(x,mean=0,sd=1)
18 plot(x,y,type="l",lwd=2,col="black")
19
20 x=seq(-3,z0,length=500)
21 y=dnorm(x,mean=0,sd=1)
22 polygon(c(-3,x,z0),c(0,y,0),col="gray")
23
24 points(z,0,pch=19,col="red",cex=1.5)
25
26 if(z0<z){
27   print("ACCEPT NULL HYPOTHESIS")
28 }else{
29   print("REJECT NULL HYPOTHESIS")
30 }

```

R code Exa 9.12 p value approach

```

1 #Page 509
2
3 n1<-500

```

```

4 n2<-100
5 p1<-0.67
6 p2<-0.80
7 d0<- -0.05
8 alpha<-1-0.9
9 m<-sqrt((p1*(1-p1)/n1)+(p2*(1-p2)/n2))
10 z<-(p1-p2-d0)/m
11 print(z)
12
13 p<-1-(pnorm(Inf)-pnorm(z))
14 print(p)
15
16 if(p>alpha){
17   print("ACCEPT NULL HYPOTHESIS")
18 } else {
19   print("REJECT NULL HYPOTHESIS")
20 }

```

R code Exa 9.13 Sample Size Considerations

```

1 #Page 521
2
3 alpha<-1-0.995
4 k1<-0.75
5 k2<-1.15
6 z<-qnorm(alpha/2,lower.tail =FALSE)
7 E=0.5
8 n<-((z**2)*(k1**2+k2**2))/(E**2)
9 print(n)
10
11 #The answer may slightly vary due to rounding off
    values.

```

R code Exa 9.14 Sample Size Considerations

```
1 #Page 523
2
3 alpha<-1-0.999
4 k<-0.025
5 z<-qnorm(alpha/2,lower.tail =FALSE)
6 E=0.01
7 n<-(z**2)*(k**2)/(E**2)
8 print(n)
9
10 #The answer may slightly vary due to rounding off
    values.
```

R code Exa 9.15.a Finding minimum equal sample sizes

```
1 #Page 525
2
3 alpha<-1-0.98
4 p1<-0.5
5 p2<-0.5
6 z<-qnorm(alpha/2,lower.tail =FALSE)
7 E<-0.05
8 n<-(z**2)*(p1*(1-p1)+p2*(1-p2))/E**2
9 print(n)
10
11 #The answer may slightly vary due to rounding off
    values.
```

R code Exa 9.15.b Finding minimum equal sample sizes

```
1 #Page 525
2
```

```
3 alpha<-1-0.98
4 p1<-0.2
5 p2<-0.3
6 z<-qnorm(alpha/2,lower.tail=FALSE)
7 E<-0.05
8 n<-(z**2)*(p1*(1-p1)+p2*(1-p2))/E**2
9 print(n)
10
11 #The answer may slightly vary due to rounding off
    values.
```

Chapter 10

Correlation and Regression

R code Exa 10.1 Linear Correlation Coefficient

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

R code Exa 10.2 The Least Squares Regression Line

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

R code Exa 10.3.a Scatter Diagram

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

R code Exa 10.3.b Linear Correlation Coefficient

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

R code Exa 10.3.c Least Squares Regression Line

```
1 #Page 566
2 x<-c(2,3,3,3,4,4,5,5,5,6)
3 y<-c
  (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)
4 plot(x, y)
5 fit <- lm(y~x)
6 abline(fit,col='blue')
7 print(fit)
```

R code Exa 10.3.e Correlation and Regression

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

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

R code Exa 10.3.f Correlation and Regression

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

4 k<-lm(y~x)
5 print(k)
6 b0<-coef(k)[ "(Intercept)" ]
7 b1<-coef(k)[ "x" ]
8 x_1<-20
9 y_1<-print((b1*x_1)+b0)
```

R code Exa 10.4.a Sum of Squared Errors

```

1 #Page 572
2 x<-c(2,2,6,8,10)
3 y<-c(0,1,2,3,3)
4 fit<-lm(y~x)
5 print(fit)
6 beta_0<-coef(fit)["(Intercept)"]
7 beta_1<-coef(fit)["x"]
8 find<-function(m){
9   y<-c((beta_1*x)+beta_0)
10 }
11 k<-find(m)
12
13 library("Metrics")
14 print(sse(k,y))

```

R code Exa 10.4.b Sum of Squared Errors

```

1 #Page 572
2 x<-c(2,2,6,8,10)
3 y<-c(0,1,2,3,3)
4 sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5])
5 sum_of_y<-sum(y[1],y[2],y[3],y[4],y[5])
6 sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x
7   [5]**2)
8 sum_of_y_sq<-sum(y[1]**2,y[2]**2,y[3]**2,y[4]**2,y
9   [5]**2)
10 n<-length(x)
11 SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
12 SS_yy<-(sum_of_y_sq)-((1/n)*(sum_of_y)**2)
13 sum_of_xy<-sum(x[1]*y[1],x[2]*y[2],x[3]*y[3],x[4]*y
14   [4],x[5]*y[5])
15 SS_xy<-(sum_of_xy)-((1/n)*(sum_of_x)*(sum_of_y))
16 beta_1<-SS_xy/SS_xx
17 SSE<-print(SS_yy-((beta_1)*SS_xy))

```

R code Exa 10.5 Sum of Squared Errors

```
1 #Page 573
2 x<-c(2,3,3,3,4,4,5,5,5,6)
3 y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

4 fit<-lm(y~x)
5 print(fit)
6 beta_0<-32.83
7 beta_1<--2.05
8 find<-function(x){
9   y<-c((beta_1*x)+beta_0)
10 }
11 k<-find(x)
12 library("Metrics")
13 print(sse(k,y))
```

R code Exa 10.6 Statistical Inferences About beta1

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

```

13 SSE<-sse(k,y)
14 sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5])
15 sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x
      [5]**2)
16 n<-length(x)
17 SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
18 S_e<-sqrt(SSE/(n-2))
19 alpha<-1-0.95
20 df<-n-2
21 t<-qt(alpha/2,df)
22 Upper_interval<-print(beta_1+((t)*(S_e)/sqrt(SS_xx))
      )
23 Lower_interval<-print(beta_1-((t)*(S_e)/sqrt(SS_xx))
      )

```

R code Exa 10.7 Statistical Inferences About beta1

```

1 #Page 584
2 x<-c(2,3,3,3,4,4,5,5,5,6)
3 y<-c
      (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

4 fit<-lm(y~x)
5 print(fit)
6 beta_0<-32.83
7 beta_1<- -2.05
8 find<-function(x){
9   y<-c((beta_1*x)+beta_0)
10 }
11 k<-find(x)
12 library("Metrics")
13 SSE<-sse(k,y)
14 sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5],x[6],x[7],x
      [8],x[9],x[10])
15 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)
16 n<-length(x)
17 SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
18 S_e<-sqrt(SSE/(n-2))
19 alpha<-1-0.90
20 df<-n-2
21 t<-qt(alpha/2,df)
22 Upper_interval<-print(beta_1+((t)*(S_e)/sqrt(SS_xx))
    )
23 Lower_interval<-print(beta_1-((t)*(S_e)/sqrt(SS_xx))
    )

```

R code Exa 10.8 Critical value test

```

1 #Page 586
2 x<-c(2,2,6,8,10)
3 y<-c(0,1,2,3,3)
4 fit<-lm(y~x)
5 print(fit)
6 beta_0<- -0.1250
7 beta_1<-0.3438
8 alpha<-0.02
9 find<-function(x){
10   y<-c((beta_1*x)+beta_0)
11 }
12 k<-find(x)
13 library("Metrics")
14 SSE<-sse(k,y)
15 sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5])
16 sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x
    [5]**2)
17 n<-length(x)
18 df<-n-2
19 SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
20 S_e<-sqrt(SSE/(n-2))

```

```

21 t<-beta_1/((S_e)/sqrt(SS_xx))
22 print(t)
23 t0<-qt(alpha/2,df)
24 t1<--t0
25
26 x=seq(-6,6,length=100)
27 y=dnorm(x,mean=0,sd=1)
28 plot(x,y,type="l",lwd=2,col="black")
29
30 x=seq(t1,6,length=100)
31 y=dnorm(x,mean=0,sd=1)
32 polygon(c(t1,x,6),c(0,y,0),col="gray")
33
34 x=seq(-6,t0,length=100)
35 y=dnorm(x,mean=0,sd=1)
36 polygon(c(-6,x,t0),c(0,y,0),col="gray")
37
38
39 points(t,0,pch=19,col="black",cex=1.5)
40
41 if(t>t0 & t<t1){
42   print("ACCEPT NULL HYPOTHESIS")
43 }else{
44   print("REJECT NULL HYPOTHESIS")
45 }

```

R code Exa 10.9 Critical value test

```

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

4 plot(x, y)
5 n<-length(x)

```



```

6 fit <- lm(y~x)
7 beta_1_init<- -1.1
8 beta_0<-32.83
9 beta_1<- -2.05
10 alpha<-0.05
11 df<-n-2
12 find<-function(x){
13   y<-c((beta_1*x)+beta_0)
14 }
15 k<-find(x)
16 library("Metrics")
17 SSE<-sse(k,y)
18 sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5],x[6],x[7],x
19   [8],x[9],x[10])
20 sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x
21   [5]**2,x[6]**2,x[7]**2,x[8]**2,x[9]**2,x[10]**2)
22 SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
23 S_e<-sqrt(SSE/(n-2))
24 t<-(beta_1-beta_1_init)/((S_e)/sqrt(SS_xx))
25 print(t)
26
27 t0<-qt(alpha,df)
28
29 x=seq(-3,3,length=500)
30 y=dnorm(x,mean=0,sd=1)
31 plot(x,y,type="l",lwd=2,col="black")
32
33 x=seq(-3,t0,length=500)
34 y=dnorm(x,mean=0,sd=1)
35 polygon(c(-3,x,t0),c(0,y,0),col="gray")
36
37 points(t,0,pch=19,col="red",cex=1.5)
38
39 if(t0>t){
40   print("REJECT NULL HYPOTHESIS")
41 }else{

```

```
42   print("ACCEPT NULL HYPOTHESIS")
43 }
```

R code Exa 10.10 Proportion of the variability

```
1 #Page 599
2 x<-c(2,3,3,3,4,4,5,5,5,6)
3 y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

4 r<-print(cor(x,y))
5 variability<-100*(r**2)
6 print(variability)
```

R code Exa 10.11 Coefficient of determination

```
1 #Page 600
2 x<-c(2,3,3,3,4,4,5,5,5,6)
3 y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

4 fit <- lm(y~x)
5 n<-length(x)
6 beta_0<-32.83
7 beta_1<-2.05
8 alpha<-0.05
9 df<-n-2
10 find<-function(x){
11   y<-c((beta_1*x)+beta_0)
12 }
13 k<-find(x)
14 library("Metrics")
15 SSE<-sse(k,y)
```

```

16 sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5],x[6],x[7],x
    [8],x[9],x[10])
17 sum_of_y<-sum(y[1],y[2],y[3],y[4],y[5],y[6],y[7],y
    [8],y[9],y[10])
18 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)
19 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)
20 SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
21 SS_yy<-(sum_of_y_sq)-((1/n)*(sum_of_y)**2)
22 S_e<-sqrt(SSE/(n-2))
23 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])
24 SS_xy<-(sum_of_xy)-((1/n)*(sum_of_x)*(sum_of_y))
25
26 r1_sq<-print((SS_yy-SSE)/SS_yy)
27 r2_sq<-print((SS_xy**2)/((SS_xx)*(SS_yy)))
28 r3_sq<-print(beta_1*(SS_xy/SS_yy))

```

R code Exa 10.12 Construction of Confidence Level

```

1 #Page 609
2
3 x<-c(2,3,3,3,4,4,5,5,5,6)
4 y<-c
    (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)
5 fit <- lm(y~x)
6 n<-length(x)
7 beta_0<-32.83
8 beta_1<-2.05
9 alpha<-0.05
10 df<-n-2
11

```

```

12 find<-function(x){
13   y<-c((beta_1*x)+beta_0)
14 }
15 k<-find(x)
16 library("Metrics")
17 SSE<-sse(k,y)
18 sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5],x[6],x[7],x
   [8],x[9],x[10])
19 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)
20 SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
21 S_e<-sqrt(SSE/(n-2))
22
23
24 x_p<-3.5
25 y_p<-(beta_1*x_p)+beta_0
26 t<-qt(alpha/2,df,lower.tail = FALSE)
27 m<-x_p-mean(x)
28 lower_interval<-print(y_p-(t*S_e)*sqrt((1/n)+(m**2/
   SS_xx)))
29 upper_interval<-print(y_p+(t*S_e)*sqrt((1/n)+(m**2/
   SS_xx)))

```

R code Exa 10.13 Construction of Confidence Level

```

1 #Page 610
2
3 x<-c(2,3,3,3,4,4,5,5,5,6)
4 y<-c
   (28.7,24.8,26.0,30.5,23.8,24.6,23.8,20.4,21.6,22.1)

5 fit <- lm(y~x)
6 n<-length(x)
7 beta_0<-32.83
8 beta_1<- -2.05

```

```

9 alpha<-0.05
10 df<-n-2
11
12 find<-function(x){
13   y<-c((beta_1*x)+beta_0)
14 }
15 k<-find(x)
16 library("Metrics")
17 SSE<-sse(k,y)
18 sum_of_x<-sum(x[1],x[2],x[3],x[4],x[5],x[6],x[7],x
19   [8],x[9],x[10])
19 sum_of_x_sq<-sum(x[1]**2,x[2]**2,x[3]**2,x[4]**2,x
20   [5]**2,x[6]**2,x[7]**2,x[8]**2,x[9]**2,x[10]**2)
20 SS_xx<-(sum_of_x_sq)-((1/n)*(sum_of_x)**2)
21 S_e<-sqrt(SSE/(n-2))
22
23
24 x_p<-3.5
25 y_p<-(beta_1*x_p)+beta_0
26 t<-qt(alpha/2,df,lower.tail = FALSE)
27 m<-x_p-mean(x)
28 lower_interval<-print(y_p-(t*S_e)*sqrt(1+(1/n)+(m**2
29   /SS_xx)))
29 upper_interval<-print(y_p+(t*S_e)*sqrt(1+(1/n)+(m**2
30   /SS_xx)))

```

Chapter 11

Chi Square Tests and F Tests

R code Exa 11.1 Chi Square Test

```
1 #Page 640
2 col1<-c(35,6)
3 col2<-c(12,24)
4 col3<-c(5,18)
5 n<-100
6 R_T<-c(col1[1]+col2[1]+col3[1], col1[2]+col2[2]+col3
      [2])
7 C_T<-c(sum(col1), sum(col2), sum(col3))
8
9 alpha<-0.01
10 df<-(length(R_T)-1)*(length(C_T)-1)
11
12 find_E<-function(i,k){
13   E<-(R_T[i]*C_T[k])/n
14 }
15 test_col1<-function(m,i,k){
16   test<-((col1[m]-find_E(i,k))**2)/find_E(i,k)
17 }
18 test_col2<-function(m,i,k){
19   test<-((col2[m]-find_E(i,k))**2)/find_E(i,k)
20 }
```

```

21 test_col3<-function(m,i,k){
22   test<-((col3[m]-find_E(i,k))**2)/find_E(i,k)
23 }
24
25 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))
26 print(chisq)
27
28 crit_val<-print(qchisq(alpha,df,lower.tail = FALSE))
29
30 if(crit_val<chisq){
31   print("REJECT NULL HYPOTHESIS")
32 } else {
33   print("ACCEPT NULL HYPOTHESIS")
34 }
35
36 x=seq(7,33,length=100)
37 y=dchisq(x,df)
38 plot(x,y,type="l",lwd=2,col="black")
39
40 x=seq(crit_val,35,length=50)
41 y=dchisq(x,df)
42 polygon(c(crit_val,x,35),c(0,y,0),col="gray")
43
44 points(chisq,0,pch=19,col="red",cex=1.5)

```

R code Exa 11.2 Chi Square Test

```

1 #Page 655
2
3 alpha<-0.01
4
5 assumed_dist<-c(0.743,0.216,0.012,0.012,0.008,0.009)
6 observed_freq<-c(1732,538,32,42,133,23)

```

```

7 expected_freq<-c(2500*assumed_dist)
8 print(cbind(assumed_dist,observed_freq,expected_freq
  ))
9
10 test<-function(i){
11   test<-(((observed_freq[i]-expected_freq[i])**2)/(
     expected_freq[i]))
12 }
13 chisq<-sum(test(1),test(2),test(3),test(4),test(5),
  test(6))
14 print(chisq)
15
16
17 df<-length(assumed_dist)-1
18
19 crit_val<-print(qchisq(alpha,df,lower.tail = FALSE))
20
21 if(crit_val<chisq){
22   print("REJECT NULL HYPOTHESIS")
23 } else {
24   print("ACCEPT NULL HYPOTHESIS")
25 }
26
27 x=seq(10,35,length=1000)
28 y=dchisq(x,df)
29 plot(x,y,type="l",lwd=2,col="black")
30
31 x=seq(crit_val,35,length=50)
32 y=dchisq(x,df)
33 polygon(c(crit_val,x,35),c(0,y,0),col="gray")
34
35 points(chisq,0,pch=19,col="red",cex=1.5)

```

R code Exa 11.3.a F test


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

R code Exa 11.3.b F test

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

R code Exa 11.4.a F Test Critical Values

```
1 #Page 668
2
3 df1<-2
4 df2<-20
5 alpha<-0.05
6 print(qf(alpha,df1,df2,lower.tail = FALSE))
```

R code Exa 11.4.b F Test Critical Values

```
1 #Page 668
2
3 df1<-2
4 df2<-20
5 alpha<-0.05
6 print(qf(alpha/2,df1,df2,lower.tail = FALSE))
```

R code Exa 11.4.c F Test Critical Values

```
1 #Page 668
2
3 df1<-2
4 df2<-20
5 alpha<-0.05
6 print(qf(1-alpha,df1,df2,lower.tail = FALSE))
```

R code Exa 11.4.d F Test Critical Values

```
1 #Page 668
2
3 df1<-2
4 df2<-20
5 alpha<-0.05
6 print(qf(1-alpha/2,df1,df2,lower.tail = FALSE))
7
8 #The answer may slightly vary due to rounding off
   values.
```

R code Exa 11.5.a F Test Critical Values

```
1 #Page 671
2
3 df1<-13
4 df2<-8
5 alpha<-0.01
6 k<-qf(1-alpha,df2,df1,lower.tail = FALSE)
```

```
7 print(1/k)
8
9 #The answer may slightly vary due to rounding off
  values.
```

R code Exa 11.5.b F Test Critical Values

```
1 #Page 671
2
3 df1<-40
4 df2<-10
5 alpha<-0.975
6 k<-qf(1-alpha,df1,df2,lower.tail = FALSE)
7 print(1/k)
8
9
10 #The answer may slightly vary due to rounding off
    values.
```

R code Exa 11.6 F test

```
1 #Page 674
2
3 n1<-16
4 n2<-21
5 var1<-2.09
6 var2<-1.10
7 alpha<-0.10
8 df1<-n1-1
9 df2<-n2-1
10
11 f0<-print(qf(1-alpha/2,df1,df2,lower.tail=FALSE))
12 f1<-print(qf(alpha/2,df1,df2,lower.tail=FALSE))
```

```

13
14 t<-var1/var2
15 print(t)
16
17 x=seq(-1,4,length=1000)
18 y=df(x,df1,df2)
19 plot(x,y,type="l",lwd=2,col="black")
20 x=seq(-1,f0,length=50)
21 y=df(x,df1,df2)
22 polygon(c(-1,x,f0),c(0,y,0),col="gray")
23 x=seq(f1,4,length=50)
24 y=df(x,df1,df2)
25 polygon(c(f1,x,4),c(0,y,0),col="gray")
26 points(t,0,pch=19,col="red",cex=1.5)
27
28 if(t>f0 & t<f1){
29   print("ACCEPT NULL HYPOTHESIS")
30 } else {
31   print("REJECT NULL HYPOTHESIS")
32 }

```

R code Exa 11.7 F test

```

1 #Page 676
2
3 n1<-16
4 n2<-21
5 var1<-2.09
6 var2<-1.10
7 alpha<-0.10
8 df1<-n1-1
9 df2<-n2-1
10
11 f1<-print(qf(alpha,df1,df2,lower.tail=FALSE))
12

```

```

13 t<-var1/var2
14 print(t)
15
16 x=seq(-1,4,length=1000)
17 y=df(x,df1,df2)
18 plot(x,y,type="l",lwd=2,col="black")
19
20 x=seq(f1,4,length=50)
21 y=df(x,df1,df2)
22 polygon(c(f1,x,4),c(0,y,0),col="gray")
23 points(t,0,pch=17,col="red",cex=1.5)
24
25 if(t<f1){
26   print("ACCEPT NULL HYPOTHESIS")
27 } else {
28   print("REJECT NULL HYPOTHESIS")
29 }

```

R code Exa 11.8 F Test using MST and MSE

```

1 #Page 690
2 maths<-c
   (2.59,3.13,2.97,2.50,2.53,3.29,2.53,3.17,2.70,3.88,2.64)
3 english<-c
   (3.64,3.19,3.15,3.78,3.03,2.61,3.20,3.30,3.54,3.25,4.00)
4 education<-c
   (4.00,3.59,2.80,2.39,3.47,3.59,3.74,3.77,3.13,3.00,3.47)
5 biology<-c
   (2.78,3.51,2.65,3.16,2.94,2.32,2.58,3.21,3.23,3.57,3.22)
6
7 print(cbind(maths,english,education,biology))

```

```

8
9 n1<-length(maths)
10 n2<-length(english)
11 n3<-length(education)
12 n4<-length(biology)
13
14 x1<-mean(maths)
15 x2<-mean(english)
16 x3<-mean(education)
17 x4<-mean(biology)
18
19 var1<-var(maths)
20 var2<-var(english)
21 var3<-var(education)
22 var4<-var(biology)
23
24 n<-44
25 K<-4
26
27 x<-(x1*n1+x2*n2+x3*n3+x4*n4)/(n1+n2+n3+n4)
28
29 MST<-((n1*((x1-x)**2))+(n2*((x2-x)**2))+(n3*((x3-x)*
      *2))+(n4*((x4-x)**2)))/(K-1)
30 MSE<-(((n1-1)*(var1))+((n2-1)*(var2))+((n3-1)*(var3)
      )+((n4-1)*(var4)))/(n-K)
31
32 f0<-print(MST/MSE)
33
34 alpha<-0.05
35 df1<-K-1
36 df2<-n-K
37 f<-qf(alpha,df1,df2,lower.tail = FALSE)
38 print(f)
39
40 if(f<f0){
41   print("REJECT NULL HYPOTHESIS")
42 } else {
43   print("ACCEPT NULL HYPOTHESIS")

```

```

44 }
45
46 x=seq(-1,4,length=1000)
47 y=df(x,df1,df2)
48 plot(x,y,type="l",lwd=2,col="black")
49
50 x=seq(f,4,length=50)
51 y=df(x,df1,df2)
52 polygon(c(f,x,4),c(0,y,0),col="gray")
53
54 points(f0,0,pch=18,col="red",cex=2)

```

R code Exa 11.9 Chi Square Tests and F Tests

```

1 #Page 693
2
3 t1<-c
      (71,72,75,80,60,65,63,78,75,73,72,65,63,69,64,71)
4 t2<-c(77,67,79,78,81,72,71,84,91)
5 t3<-c(81,79,73,71,75,84,77,67)
6
7 n1<-length(t1)
8 n2<-length(t2)
9 n3<-length(t3)
10
11 x1<-mean(t1)
12 x2<-mean(t2)
13 x3<-mean(t3)
14
15 var1<-var(t1)
16 var2<-var(t2)
17 var3<-var(t3)
18
19 x<-(x1*n1+x2*n2+x3*n3)/(n1+n2+n3)
20 alpha<-0.01

```

```

21 n<-33
22 K<-3
23 df1<-K-1
24 df2<-n-K
25
26 MST<-((n1*((x1-x)**2))+n2*((x2-x)**2))+n3*((x3-x)*
      *2))/(K-1)
27 MSE<-(((n1-1)*(var1))+((n2-1)*(var2))+((n3-1)*(var3)
      ))/(n-K)
28 f0<-print(MST/MSE)
29
30 f<-qf(alpha,df1,df2,lower.tail = FALSE)
31 print(f)
32
33 if(f<f0){
34   print("REJECT NULL HYPOTHESIS")
35 } else {
36   print("ACCEPT NULL HYPOTHESIS")
37 }
38 x=seq(0,8,length=100)
39 y=df(x,10,10)
40 plot(x,y,type="l",lwd=2,col="black")
41 x=seq(f,10,length=1000)
42 y=df(x,10,10)
43 polygon(c(f,x,f),c(0,y,0),col="gray")
44 points(f,0,pch=19,col="red",cex=1.5)

```
