

R Textbook Companion for  
Statistics for Psychology  
by Arthur Aron, Elliot J. Coups and Elaine N.  
Aron<sup>1</sup>

Created by  
Shivam Agrawal  
B.Tech.  
Computer Science and Engineering  
Techno India NJR Institute of Technology  
Cross-Checked by  
R TBC Team

May 23, 2021

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

# Book Description

**Title:** Statistics for Psychology

**Author:** Arthur Aron, Elliot J. Coups and Elaine N. Aron

**Publisher:** Pearson, USA

**Edition:** 6

**Year:** 2013

**ISBN:** 9780205923922

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 Displaying the Order in a Group of Numbers Using Tables and Graphs	5
2 Central Tendency and Variability	13
3 Some Key Ingredients for Inferential Statistics	19
4 Introduction to Hypothesis Testing	27
5 Hypothesis Tests with Means of Samples	32
6 Making Sense of Statistical Significance	41
7 Introduction to t Tests	47
8 The t Test for Independent Means	59
9 Introduction to the Analysis of Variance	76
10 Factorial Analysis of Variance	107
11 Correlation	118
12 Prediction	128
13 Chi Square Tests	145

14 Strategies When Population Distributions Are Not Normal

158

# List of R Codes

Exa 1.1a	Frequency Tables an Example . . . . .	5
Exa 1.1c	Frequency Tables Another Example . . . . .	6
Exa 1.1d	Grouped Frequency Tables . . . . .	6
Exa 1.1e	Grouped Frequency Tables Another Example . . . . .	7
Exa 1.2a	Histogram for Frequency Table . . . . .	8
Exa 1.2b	Histogram for Grouped Frequency Table . . . . .	9
Exa 1.2c	Histogram for Grouped Frequency Table Example 2 . . . . .	9
Exa 1.2d	Histogram for Nominal Variables . . . . .	10
Exa 1.2e	How are you doing . . . . .	11
Exa 1.4a	Worked Out Examples . . . . .	12
Exa 2.1a	Mean Example 1 . . . . .	13
Exa 2.1b	Mean Example 2 . . . . .	13
Exa 2.1c	Mean Example 3 . . . . .	14
Exa 2.1d	How are you doing . . . . .	15
Exa 2.2a	Variance and Standard Deviation Example 1 . . . . .	16
Exa 2.2b	Variance and Standard Deviation Example 2 . . . . .	16
Exa 2.2c	How are you doing . . . . .	17
Exa 2.3a	Worked Out Examples . . . . .	17
Exa 3.1a	Formula to change Raw Score to Z Score and Vice Versa . . . . .	19
Exa 3.1b	Additional Example 1 to change Raw Score to Z Score and Vice Versa . . . . .	20
Exa 3.1c	Additional Example 2 to change Raw Score to Z Score and Vice Versa . . . . .	20
Exa 3.1d	How are you doing . . . . .	21
Exa 3.1e	How are you doing . . . . .	22
Exa 3.2a	Examples for finding Percentage from Z Scores and Raw Scores using Normal Curve Table . . . . .	23

Exa 3.2b	Examples for finding Z Scores and Raw Scores from Percentage using Normal Curve Table . . . . .	23
Exa 3.3a	Worked Out Examples 1 . . . . .	24
Exa 3.3b	Worked Out Examples 2 . . . . .	25
Exa 3.3c	Worked Out Examples 3 . . . . .	26
Exa 4.1a	The Hypothesis Testing Process . . . . .	27
Exa 4.1b	The Second Example . . . . .	28
Exa 4.1c	How are you doing . . . . .	28
Exa 4.2a	Example of Hypothesis Testing with a Two Tailed Test	29
Exa 4.2b	How are you doing . . . . .	30
Exa 4.3a	Worked Out Examples . . . . .	31
Exa 5.1a	Determining the Characteristics of a Distribution of Means	32
Exa 5.1b	Example of Determining the Characteristics of a Distribution of Means . . . . .	32
Exa 5.1c	How are you doing . . . . .	33
Exa 5.2a	Hypothesis testing with a distribution of means The Z Test . . . . .	34
Exa 5.2b	Example 1 for Hypothesis testing with a distribution of means The Z Test . . . . .	34
Exa 5.2c	Example 2 for Hypothesis testing with a distribution of means The Z Test . . . . .	35
Exa 5.2d	How are you doing . . . . .	36
Exa 5.3a	Confidence interval . . . . .	36
Exa 5.3b	How are you doing . . . . .	37
Exa 5.4a	Worked out examples 1 . . . . .	38
Exa 5.4b	Worked out examples 2 . . . . .	38
Exa 5.4c	Worked out examples 3 . . . . .	39
Exa 6.1a	Figuring effect size . . . . .	41
Exa 6.1b	How are you doing . . . . .	42
Exa 6.2a	Determining power from predicted effect size . . . . .	42
Exa 6.2b	How are you doing . . . . .	43
Exa 6.3a	Advanced topic Figuring Statistical Power . . . . .	43
Exa 6.3b	How are you doing . . . . .	44
Exa 6.4a	Worked out examples 1 . . . . .	44
Exa 6.4b	Worked out examples 2 . . . . .	45
Exa 6.4c	Worked out examples 3 . . . . .	45
Exa 7.1a	Basic Principle of the t Test Estimating the Population Variance from the Sample Scores . . . . .	47

Exa 7.1b	The Standard Deviation of the Distribution of Means .	47
Exa 7.1c	The Sample means score on the comparison distribution the t score . . . . .	48
Exa 7.1d	Another example of a t test for a single sample . . . . .	48
Exa 7.1e	How are you doing . . . . .	49
Exa 7.2a	Example of a t test for dependent means . . . . .	51
Exa 7.2b	A second example of a t test for dependent means . . . . .	52
Exa 7.2c	How are you doing . . . . .	53
Exa 7.3a	Effect size for the t test for dependent means . . . . .	55
Exa 7.3b	How are you doing . . . . .	55
Exa 7.4a	Worked out examples 1 . . . . .	56
Exa 7.4b	Worked out examples 2 . . . . .	57
Exa 8.1a	Estimating the population variance . . . . .	59
Exa 8.1b	Figuring the variance of each of the two distributions of means . . . . .	60
Exa 8.1c	The variance and standard deviation of the distribution of differences between means . . . . .	60
Exa 8.1d	The t score for the difference between the two actual means . . . . .	61
Exa 8.1e	How are you doing . . . . .	61
Exa 8.2a	Example of a T test for independent means . . . . .	63
Exa 8.2b	A second example of a T test for independent means . . . . .	65
Exa 8.2c	How are you doing . . . . .	67
Exa 8.3a	Effect size for the t test for independent means . . . . .	69
Exa 8.4a	Harmonic mean . . . . .	70
Exa 8.5a	Worked out examples 1 . . . . .	71
Exa 8.5b	Worked out examples 2 . . . . .	72
Exa 8.5c	Worked out examples 3 . . . . .	74
Exa 9.1a	Carrying out an Analysis of Variance . . . . .	76
Exa 9.1b	How are you doing . . . . .	80
Exa 9.1c	Another example . . . . .	82
Exa 9.1d	How are you doing . . . . .	84
Exa 9.2a	Planned contrast an example . . . . .	86
Exa 9.2b	Planned contrast a second example . . . . .	87
Exa 9.2c	How are you doing . . . . .	89
Exa 9.3a	Scheffe Test . . . . .	91
Exa 9.4a	Effect size for the analysis of variance . . . . .	92
Exa 9.4b	How are you doing . . . . .	94



Exa 9.5a	Advanced Topic The Structural Model in the Analysis of Variance an example . . . . .	95
Exa 9.6a	Worked out examples 1 . . . . .	97
Exa 9.6b	Worked out examples 2 . . . . .	99
Exa 9.6c	Worked out examples 3 . . . . .	101
Exa 9.6d	Worked out examples 4 . . . . .	102
Exa 9.6e	Worked out examples 5 . . . . .	103
Exa 9.6f	Worked out examples 6 . . . . .	104
Exa 10.1a	Table for a Two way Analysis of Variance . . . . .	107
Exa 10.1b	How are you doing . . . . .	110
Exa 10.2a	Advanced Topic Effect Size in the Factorial Analysis of Variance . . . . .	112
Exa 10.2b	How are you doing . . . . .	114
Exa 10.3a	worked out examples 1 . . . . .	114
Exa 11.1a	Graphing Correlations The Scatter Diagram An example	118
Exa 11.1b	How are you doing . . . . .	119
Exa 11.2a	Figuring the correlation coefficient an example . . . . .	120
Exa 11.2b	Figuring the correlation coefficient a second example . . . . .	121
Exa 11.2c	How are you doing . . . . .	122
Exa 11.3a	Significance of a correlation coefficient . . . . .	122
Exa 11.3b	Significance of a correlation coefficient an example . . . . .	123
Exa 11.3c	How are you doing . . . . .	124
Exa 11.4a	Worked out examples 1 . . . . .	125
Exa 11.4b	Worked out examples 2 . . . . .	125
Exa 11.4c	Worked out examples 3 . . . . .	126
Exa 12.1a	The linear prediction rule an example . . . . .	128
Exa 12.1b	How are you doing . . . . .	129
Exa 12.2a	Another example of drawing the regression line . . . . .	130
Exa 12.3a	Finding the best linear prediction rule . . . . .	131
Exa 12.3b	The least squared error principle . . . . .	132
Exa 12.3c	Finding a and b for the least squares linear prediction rule . . . . .	133
Exa 12.3d	How are you doing . . . . .	134
Exa 12.4a	The standardized regression coefficient . . . . .	136
Exa 12.4b	How are you doing . . . . .	136
Exa 12.5a	Multiple regression . . . . .	137
Exa 12.5b	How are you doing . . . . .	137
Exa 12.6a	Proportionate reduction in error . . . . .	138

Exa 12.6b	How are you doing . . . . .	139
Exa 12.7a	Worked out examples 1 . . . . .	140
Exa 12.7b	Worked out examples 2 . . . . .	142
Exa 12.7c	Worked out examples 4 . . . . .	143
Exa 12.7d	Worked out examples 5 . . . . .	143
Exa 12.7e	Worked out examples 6 . . . . .	144
Exa 13.1a	An example . . . . .	145
Exa 13.1b	The chi square statistic and the chi square test for goodness of fit . . . . .	145
Exa 13.1c	Another example . . . . .	146
Exa 13.1d	How are you doing . . . . .	147
Exa 13.2a	The chi square test for independence . . . . .	148
Exa 13.2b	Another example . . . . .	150
Exa 13.2c	How are you doing . . . . .	151
Exa 13.3a	Effect Size for Chi Square Tests for Independence . . . . .	153
Exa 13.3b	How are you doing . . . . .	153
Exa 13.4a	Worked out examples 1 . . . . .	154
Exa 13.4b	Worked out examples 2 . . . . .	155
Exa 13.4c	Worked out examples 3 and 4 . . . . .	157
Exa 14.1a	An example of a data transformation . . . . .	158
Exa 14.2a	An example of a rank order test . . . . .	160
Exa 14.2b	Using parametric tests with rank transformed data . . . . .	161
Exa 14.2c	How are you doing . . . . .	162
Exa 14.3a	Worked out examples 1 and 2 . . . . .	162

# Chapter 1

## Displaying the Order in a Group of Numbers Using Tables and Graphs

R code Exa 1.1a Frequency Tables an Example

```
1 # Page no. : 7
2
3 stress_rating <- c(8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3,
4                   7, 6, 5, 0, 9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7,
5                   10, 8, 8)
6
7 frequency_table <- data.frame(table(stress_rating))
8
9 percent <- round((prop.table(frequency_table$Freq) *
10                  100), 1)
11
12 frequency_table <- cbind(frequency_table, percent)
13
14 View(frequency_table)
```

---

### R code Exa 1.1c Frequency Tables Another Example

```
1 # Page no. : 8 - 9
2
3 social_interaction <- c(48, 15, 33, 3, 21, 19, 17,
4     16, 44, 25, 30, 3, 5, 9, 35, 32, 26, 13, 14,
5     14, 47, 47, 18, 11, 5, 19,
6     24, 17, 6, 25, 8, 18, 29,
7     1, 18, 22, 3, 22,
8     29, 2, 6, 10, 29, 10, 29,
9     21, 38, 41, 16, 17, 8,
10    40, 8, 10, 18, 7, 4, 4,
11    8, 11, 3, 23, 10, 19, 21,
12    13, 12, 10, 4, 17, 11,
13    21, 9, 8, 7, 5, 3, 22,
14    14, 25, 4, 11, 10, 18, 1,
15    28, 27, 19, 24, 35, 9,
16    30, 8, 26)
17
18 breaks <- seq(0, 49, by = 1)
19
20 social_interaction <- cut(social_interaction, breaks
21     , right = F)
22
23 frequency_table <- data.frame(table(social.
24     interaction))
25
26 View(frequency_table)
```

---

### R code Exa 1.1d Grouped Frequency Tables

```
1 # Page no. : 9
2
3 stress_rating <- c(8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3,
4     7, 6, 5, 0, 9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7,
```

```

    10, 8, 8)
4
5 breaks <- seq(-1, 11, by = 2)
6
7 stress.rating <- cut(stress_rating, breaks, right =
  T)
8
9 grouped_frequency_table <- data.frame(table(stress.
  rating))
10
11 percent <- round((prop.table(grouped_frequency_table
  $Freq) * 100), 1)
12
13 grouped_frequency_table <- cbind(grouped_frequency_
  table, percent)
14
15 View(grouped_frequency_table)

```

---

### R code Exa 1.1e Grouped Frequency Tables Another Example

```

1 # Page no. : 10
2
3 social_interaction <- c(48, 15, 33, 3, 21, 19, 17,
  16, 44, 25, 30, 3, 5, 9, 35, 32, 26, 13, 14,
4
5      14, 47, 47, 18, 11, 5, 19,
6      24, 17, 6, 25, 8, 18, 29,
7      1, 18, 22, 3, 22,
8      29, 2, 6, 10, 29, 10, 29,
9      21, 38, 41, 16, 17, 8,
10     40, 8, 10, 18, 7, 4, 4,
11     8, 11, 3, 23, 10, 19, 21,
12     13, 12, 10, 4, 17, 11,
13     21, 9, 8, 7, 5, 3, 22,
14     14, 25, 4, 11, 10, 18, 1,
15     28, 27, 19, 24, 35, 9,

```

```

                                30, 8, 26)
8
9 breaks <- seq(0, 50, by = 5)
10
11 social.interaction <- cut(social_interaction, breaks
    , right = F)
12
13 grouped_frequency_table <- data.frame(table(social.
    interaction))
14
15 percent <- round((prop.table(grouped_frequency_table
    $Freq) * 100), 1)
16
17 grouped_frequency_table <- cbind(grouped_frequency_
    table, percent)
18
19 View(grouped_frequency_table)

```

---

### R code Exa 1.2a Histogram for Frequency Table

```

1 # Page no. : 11
2
3 stress_rating <- c(8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3,
    7, 6, 5, 0, 9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7,
    10, 8, 8)
4
5 frequency_table <- data.frame(table(stress_rating))
6
7 library(ggplot2)
8
9 ggplot(data = frequency_table, aes(x = stress_rating
    , y = Freq)) +
10   geom_bar(stat = "identity", fill = "violet") +
11   labs(title = "Histogram for frequency table", x =
    "Stress Rating", y = "Frequency") +

```

```
12 theme_bw()
```

---

### R code Exa 1.2b Histogram for Grouped Frequency Table

```
1 # Page no. : 11
2
3 stress_rating <- c(8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3,
4                   7, 6, 5, 0, 9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7,
5                   10, 8, 8)
6
7 frequency_table <- data.frame(table(stress_rating))
8
9 View(frequency_table)
10
11 library(ggplot2)
12
13 ggplot(data = frequency_table, aes(x = stress_rating
14   , y = Freq)) +
15   geom_bar(stat = "identity", fill = "violet") +
16   labs(title = "Histogram for grouped frequency
17     table", x = "Stress Rating", y = "Frequency") +
18   theme_bw()
```

---

### R code Exa 1.2c Histogram for Grouped Frequency Table Example 2

```
1 # Page no. : 12
2
3 social_interaction <- c(48, 15, 33, 3, 21, 19, 17,
4   16, 44, 25, 30, 3, 5, 9, 35, 32, 26, 13, 14,
5   14, 47, 47, 18, 11, 5, 19,
6   24, 17, 6, 25, 8, 18, 29,
7   1, 18, 22, 3, 22,
```

```

5           29, 2, 6, 10, 29, 10, 29,
           21, 38, 41, 16, 17, 8,
           40, 8, 10, 18, 7, 4, 4,
6           8, 11, 3, 23, 10, 19, 21,
           13, 12, 10, 4, 17, 11,
           21, 9, 8, 7, 5, 3, 22,
7           14, 25, 4, 11, 10, 18, 1,
           28, 27, 19, 24, 35, 9,
           30, 8, 26)
8
9 breaks <- seq(0, 50, by = 5)
10
11 mid_values <- seq(0+5/2,50-5/2, 5)
12
13 social.interaction <- cut(social_interaction, breaks
    , right = F)
14
15 grouped_frequency_table <- data.frame(table(social.
    interaction))
16 grouped_frequency_table$social_interacton_mid <- mid
    _values
17
18 View(grouped_frequency_table)
19
20 library(ggplot2)
21
22 ggplot(data = grouped_frequency_table, aes(x =
    factor(social_interacton_mid), y = Freq)) +
23   geom_bar(stat = "identity", fill = "violet") +
24   labs(title = "Histogram for grouped frequency
    table", x = "Social Interaction Mid Values",
25         y = "Frequency") +
26   theme_bw()

```

---

**R code Exa 1.2d** Histogram for Nominal Variables



```

1 # Page no. : 14
2
3 closest_person <- c("Family member", "Nonromantic
  friend", "Romantic friend", "Other")
4
5 frequency <- c(33, 76, 92, 7)
6
7 frequency_table <- data.frame(closest_person,
  frequency)
8
9 library(ggplot2)
10
11 ggplot(data = frequency_table, aes(x = closest_
  person, y = frequency)) +
12   geom_bar(stat = "identity", fill = "violet") +
13   labs(title = "Histogram for nominal variables", x
    = "Closest Person", y = "Frequency") +
14   theme_bw()

```

---

#### R code Exa 1.2e How are you doing

```

1 # Page no. : 15
2
3 value <- c(1, 2, 3, 4, 5)
4
5 frequency <- c(3, 4, 8, 5, 2)
6
7 frequency_table <- data.frame(value, frequency)
8
9 library(ggplot2)
10
11 ggplot(data = frequency_table, aes(x = value, y =
  frequency)) +
12   geom_bar(stat = "identity", fill = "violet") +
13   labs(title = "Histogram", x = "Value", y = "

```

```
    Frequency") +  
14   theme_bw()
```

---

### R code Exa 1.4a Worked Out Examples

```
1 # Page no. : 25  
2  
3 Interest <- c(2, 4, 5, 5, 1, 3, 6, 3, 6, 6)  
4  
5 frequency_table <- data.frame(table(Interest))  
6  
7 Percent <- round((prop.table(frequency_table$Freq) *  
8   100), 0)  
9  
10 frequency_table <- cbind(frequency_table, Percent)  
11  
12 View(frequency_table)  
13  
14 library(ggplot2)  
15 ggplot(data = frequency_table, aes(x = Interest, y =  
16   Freq)) +  
17   geom_bar(stat = "identity", fill = "violet") +  
   labs(title = "Histogram", x = "Interest", y = "  
     Frequency")
```

---

## Chapter 2

# Central Tendency and Variability

R code Exa 2.1a Mean Example 1

```
1 # Page no. : 36 - 37
2
3 x <- c(7, 8, 8, 7, 3, 1, 6, 9, 3, 8)
4
5 value <- mean(x)
6
7 cat("Mean of x is", value)
```

---

R code Exa 2.1b Mean Example 2

```
1 # Page no. : 37
2
3 x <- c(8, 7, 4, 10, 8, 6, 8, 9, 9, 7, 3, 7, 6, 5, 0,
4       9, 10, 7, 7, 3, 6, 7, 5, 2, 1, 6, 7, 10, 8, 8)
5 value <- mean(x)
```

```

6
7 cat("Mean of x is", round(value,2))
8
9 library(ggplot2)
10
11 ggplot(data = data.frame(x), aes(x = x)) +
12   geom_bar(stat = "count", fill = "violet") +
13   geom_vline(xintercept = mean(x), col = "red", lwd
14     = 0.8) +
15   labs(title = "Histogram with mean value", x = "
16     Stress Rating", y = "Frequency") +
17   theme_bw()

```

---

### R code Exa 2.1c Mean Example 3

```

1 # Page no. : 38
2
3 x <- c(48, 15, 33, 3, 21, 19, 17, 16, 44, 25, 30, 3,
4       5, 9, 35, 32, 26, 13, 14, 14, 47, 47, 18,
5       11, 5, 19, 24, 17, 6, 25, 8, 18, 29, 1, 18,
6       22, 3, 22, 29, 2, 6, 10, 29, 10, 29, 21,
7       38,
8       41, 16, 17, 8, 40, 8, 10, 18, 7, 4, 4, 8, 11,
9       3, 23, 10, 19, 21, 13, 12, 10, 4, 17, 11,
10      21, 9, 8, 7, 5, 3, 22, 14, 25, 4, 11, 10, 18,
11      1, 28, 27, 19, 24, 35, 9, 30, 8, 26)
12
13 value <- mean(x)
14
15 cat("Mean of x is", round(value, 2))
16
17 library(ggplot2)
18
19 ggplot(data = data.frame(table(x)), aes(x = x, y =
20   Freq)) +

```

```
15 geom_bar(stat = "identity", fill = "violet") +
16 geom_vline(xintercept = mean(x), col = "red", lwd
    = 0.8) +
17 labs(title = "Histogram with mean value", x = "
    Social Interaction", y = "Frequency") +
18 theme_bw()
```

---

### R code Exa 2.1d How are you doing

```
1 # Page no. : 43
2
3 scores <- c(2, 3, 3, 6, 6)
4
5 value <- mean(scores)
6
7 cat("Mean value of scores is", value)
8
9 scores <- c(5, 3, 2, 13, 2)
10
11 value <- mean(scores)
12
13 cat("\nMean value of scores is", value)
14
15 mode <- function(v)
16 {
17   x <- unique(v)
18   x[which.max(tabulate(match(v, x)))]
19 }
20
21 value <- mode(scores)
22
23 cat("\nMode of scores is", value)
24
25 value <- median(scores)
26
```

```
27 cat("\nMedian of scores is", value)
```

---

### R code Exa 2.2a Variance and Standard Deviation Example 1

```
1 # Page no. : 48
2
3 scores <- c(7, 8, 8, 7, 3, 1, 6, 9, 3, 8)
4
5 library(rafalib)
6
7 variance <- popvar(scores)
8
9 standard_deviation <- popsd(scores)
10
11 cat("Variance is", round(variance, 2), "and Standard
    Deviation is", round(standard_deviation, 2))
```

---

### R code Exa 2.2b Variance and Standard Deviation Example 2

```
1 # Page no. : 49
2
3 interactions <- c(48, 15, 33, 3, 21, 19, 17, 16, 44,
4     25, 30, 3, 5, 9, 35, 32, 26, 13, 14,
5     14, 47, 47, 18, 11, 5, 19, 24, 17,
6     6, 25, 8, 18, 29, 1, 18, 22,
7     3, 22,
8     29, 2, 6, 10, 29, 10, 29, 21, 38,
9     41, 16, 17, 8, 40, 8, 10, 18,
10    7, 4, 4,
11    8, 11, 3, 23, 10, 19, 21, 13, 12,
12    10, 4, 17, 11, 21, 9, 8, 7, 5,
13    3, 22,
```

```

7             14, 25, 4, 11, 10, 18, 1, 28, 27,
              19, 24, 35, 9, 30, 8, 26)
8
9  library(rafalib)
10
11 variance <- popvar(interactions)
12
13 standard_deviation <- popsd(interactions)
14
15 cat("Variance is", round(variance, 2), "and Standard
      Deviation is", round(standard_deviation, 2))

```

---

**R code Exa 2.2c** How are you doing

```

1 # Page no. : 51 - 52
2
3 scores <- c(2, 4, 3, 7)
4
5 library(rafalib)
6
7 variance <- popvar(scores)
8
9 standard_deviation <- popsd(scores)
10
11 cat("Variance is", round(variance, 2), "and Standard
      Deviation is", round(standard_deviation, 2))

```

---

**R code Exa 2.3a** Worked Out Examples

```

1 # Page no. : 58 - 60
2
3 scores <- c(8, 6, 6, 9, 6, 5, 6, 2)
4

```

```
5 value <- mean(scores)
6
7 cat("Mean is", value)
8
9 scores <- c(1, 7, 4, 2, 3, 6, 2, 9, 7)
10
11 value <- median(scores)
12
13 cat("\nMedian is", value)
14
15 scores <- c(8, 6, 6, 9, 6, 5, 6, 2)
16
17 value <- sum((scores - mean(scores))**2)
18
19 cat("\nSum of Squares is", value)
20
21 library(rafalib)
22
23 variance <- popvar(scores)
24
25 cat("\nVariance is", variance)
26
27 standard_deviation <- popsd(scores)
28
29 cat("\nStandard Deviation is", round(standard_
    deviation, 2))
```

---



# Chapter 3

## Some Key Ingredients for Inferential Statistics

**R code Exa 3.1a** Formula to change Raw Score to Z Score and Vice Versa

```
1 # Page no. : 71
2
3 x <- 8
4 m <- 12
5 sd <- 4
6
7 Z <- (x - m) / sd
8
9 cat("Z Score is", Z)
10
11 Z <- 1.5
12 sd <- 4
13 m <- 12
14
15 X <- (Z * sd) + m
16
17 cat("\nRaw Score is", X)
```

---

**R code Exa 3.1b** Additional Example 1 to change Raw Score to Z Score and Vice Versa

```
1 # Page no. : 72
2
3 m <- 3.40
4 sd <- 1.47
5 x <- 6
6
7 Z <- (x - m) / sd
8
9 cat("Z Score is", round(Z, 2))
10
11 Z <- -1.63
12
13 X <- (Z * sd) + m
14
15 cat("\nRaw Score is", round(X, 2))
```

---

**R code Exa 3.1c** Additional Example 2 to change Raw Score to Z Score and Vice Versa

```
1 # Page no. : 72 - 73
2
3 m <- 6.43
4 sd <- 2.56
5 x <- 10
6
7 Z <- (x - m) / sd
8
9 cat("Z Score is", round(Z, 2))
10
```

```
11 Z <- -1.73
12
13 X <- (Z * sd) + m
14
15 cat("\nRaw Score is", round(X, 2))
```

---

**R code Exa 3.1d** How are you doing

```
1 # Page no. : 73
2
3 m <- 20
4 sd <- 5
5
6 x1 <- 30
7
8 z1 <- (x1 - m) / sd
9
10 cat("Z Score is", round(z1, 2))
11
12 x2 <- 15
13
14 z2 <- (x2 - m) / sd
15
16 cat("\nZ Score is", round(z2, 2))
17
18 x3 <- 20
19
20 z3 <- (x3 - m) / sd
21
22 cat("\nZ Score is", round(z3, 2))
23
24 x4 <- 22.5
25
26 z4 <- (x4 - m) / sd
27
```

```
28 cat("\nZ Score is", round(z4, 2))
```

---

**R code Exa 3.1e** How are you doing

```
1 # Page no. : 73
2
3 m <- 10
4 sd <- 2
5
6 z1 <- 2
7
8 x1 <- (z1 * sd) + m
9
10 cat("Raw Score is", round(x1, 2))
11
12 z2 <- 0.5
13
14 x2 <- (z2 * sd) + m
15
16 cat("\nRaw Score is", round(x2, 2))
17
18 z3 <- 0
19
20 x3 <- (z3 * sd) + m
21
22 cat("\nRaw Score is", round(x3, 2))
23
24 z4 <- -3
25
26 x4 <- (z4 * sd) + m
27
28 cat("\nRaw Score is", round(x4, 2))
```

---

**R code Exa 3.2a** Examples for finding Percentage from Z Scores and Raw Scores using Normal Curve Table

```
1 # Page no. : 79 - 81
2
3 m <- 100
4 sd <- 15
5
6 IQ1 <- 125
7
8 Z1 <- (IQ1 - m) / sd
9
10 cat("Z Score is", round(Z1, 2))
11
12 percent <- pnorm(Z1, lower.tail = F) * 100
13
14 cat("\nPercentage of IQ level is", round(percent, 2)
15     )
16
17 IQ2 <- 95
18
19 Z2 <- (IQ2 - m) / sd
20
21 cat("\nZ Score is", round(Z2, 2))
22
23 percent <- pnorm(Z2, lower.tail = F) * 100
24
25 cat("\nPercentage of IQ level is", round(percent, 2)
26     )
```

---

**R code Exa 3.2b** Examples for finding Z Scores and Raw Scores from Percentage using Normal Curve Table

```
1 # Page no. : 81 - 84
2
```

```

3 m <- 100
4 sd <- 15
5
6 percent <- 0.05
7
8 Z1 <- qnorm(percent, lower.tail = F)
9
10 cat("Z Score is", round(Z1, 2))
11
12 IQ1 <- (sd * Z1) + m
13
14 cat("\nRaw Score is", round(IQ1, 2))
15
16 percent <- 0.55
17
18 Z2 <- qnorm(percent, lower.tail = F)
19
20 cat("\nZ Score is", round(Z2, 2))
21
22 IQ2 <- (sd * Z2) + m
23
24 cat("\nRaw Score is", round(IQ2, 2))
25
26 percent <- 0.95
27
28 Z3 <- qnorm(0.975, lower.tail = T)
29
30 cat("\nZ Score is", round(Z3, 2))
31
32 IQ3 <- (sd * Z3) + m
33
34 cat("\nRaw Score is", round(IQ3, 2))

```

---

**R code Exa 3.3a** Worked Out Examples 1

```

1 # Page no. : 100
2
3 m <- 80
4 sd <- 20
5 x <- 65
6
7 Z <- (x - m) / sd
8
9 cat("Z Score is", Z)
10
11 m <- 200
12 sd <- 50
13 z <- 1.26
14
15 X <- (z * sd) + m
16
17 cat("\nRaw Score is", X)

```

---

### R code Exa 3.3b Worked Out Examples 2

```

1 # Page no. : 100 - 102
2
3 m <- 20
4 sd <- 3
5 x <- 24
6
7 Z <- (x - m) / sd
8
9 cat("Z Score is", round(Z,2))
10
11 percent <- pnorm(Z, lower.tail = F) * 100
12
13 cat("\nPercentage of people have scores above 24 is"
14     , round(percent, 2))

```

```
15 m <- 20
16 sd <- 3
17 percent <- 0.75
18
19 Z <- qnorm(percent, lower.tail = F)
20
21 cat("\nZ Score is", round(Z, 2))
22
23 X <- (sd * Z) + m
24
25 cat("\nRaw Score is", round(X, 2))
```

---

### R code Exa 3.3c Worked Out Examples 3

```
1 # Page no. : 102 – 103
2
3 apple <- 20
4 strawberry <- 20
5 cherry <- 5
6 grape <- 5
7
8 total <- apple + strawberry + cherry + grape
9
10 case <- cherry + grape
11
12 probability <- case / total
13
14 cat("The required probability is", probability)
```

---



# Chapter 4

## Introduction to Hypothesis Testing

R code Exa 4.1a The Hypothesis Testing Process

```
1 # Page no. : 111 - 115
2
3 m <- 14
4 sd <- 3
5 x <- 6
6 alpha <- 0.01
7
8 cutoff <- qnorm(alpha)
9 cutoff <- round(cutoff, 2)
10
11 z <- (x - m) / sd
12
13 cat("\nCutoff Sample Z Score is ", cutoff)
14 cat("\nActual Sample Z Score is", round(z, 2))
15
16 if(z < cutoff)
17 {
18   cat("\nWe reject null hypothesis.")
19 }else{
```

```
20   cat("\nWe cannot reject null hypothesis.")
21 }
```

---

#### R code Exa 4.1b The Second Example

```
1 # Page no. : 116 - 118
2
3 m <- 70
4 sd <- 10
5 x <- 80
6 alpha <- 0.05
7
8 cutoff <- qnorm(1 - alpha)
9 cutoff <- round(cutoff, 2)
10
11 z <- (x - m) / sd
12
13 if(z > cutoff)
14 {
15   cat("We reject null hypothesis.")
16 }else{
17   cat("We cannot reject null hypothesis.")
18 }
```

---

#### R code Exa 4.1c How are you doing

```
1 # Page no. : 119 - 120
2
3 m <- 30
4 sd <- 4
5 x <- 40
6 alpha <- 0.05
7
```

```

8 cutoff <- qnorm(1 - alpha)
9 cutoff <- round(cutoff, 2)
10
11 z <- (x - m) / sd
12
13 cat("Cutoff Sample Z Score is ", cutoff)
14 cat("\nActual Sample Z Score is", z)
15
16 if(z > cutoff)
17 {
18   cat("\nWe reject null hypothesis.")
19 }else{
20   cat("\nWe cannot reject null hypothesis.")
21 }

```

---

**R code Exa 4.2a** Example of Hypothesis Testing with a Two Tailed Test

```

1 # Page no. : 123 - 124
2
3 m <- 69.5
4 sd <- 14.1
5 x <- 41
6 alpha <- 0.05
7
8 cutoff1 <- qnorm(0.025)
9 cutoff1 <- round(cutoff1, 2)
10
11 cutoff2 <- qnorm(0.975)
12 cutoff2 <- round(cutoff2, 2)
13
14 z <- (x - m) / sd
15
16 cat("Cutoff Z scores are", cutoff1, "and", cutoff2)
17 cat("\nActual Z score is", round(z, 2))
18

```

```
19 if(z > cutoff1 & z < cutoff2)
20 {
21   cat("\nWe cannot reject null hypothesis.")
22 }else{
23   cat("\nWe can reject null hypothesis.")
24 }
```

---

**R code Exa 4.2b** How are you doing

```
1 # Page no. : 125
2
3 m <- 500
4 sd <- 40
5 x <- 400
6 alpha <- 0.01
7
8 cutoff1 <- qnorm(0.005)
9 cutoff1 <- round(cutoff1, 2)
10
11 cutoff2 <- qnorm(0.995)
12 cutoff2 <- round(cutoff2, 2)
13
14 z <- (x - m) / sd
15
16 cat("Cutoff Z scores are", cutoff1, "and", cutoff2)
17 cat("\nActual Z score is", round(z, 2))
18
19 if(z > cutoff1 & z < cutoff2)
20 {
21   cat("\nWe cannot reject null hypothesis.")
22 }else{
23   cat("\nWe can reject null hypothesis..")
24 }
```

---

### R code Exa 4.3a Worked Out Examples

```
1 # Page no. : 131 - 132
2
3 m <- 19
4 sd <- 4
5 x <- 27
6 alpha <- 0.05
7
8 cutoff1 <- qnorm(0.025)
9 cutoff1 <- round(cutoff1, 2)
10
11 cutoff2 <- qnorm(0.975)
12 cutoff2 <- round(cutoff2, 2)
13
14 z <- (x - m) / sd
15
16 cat("Cutoff Z scores are", cutoff1, "and", cutoff2)
17 cat("\nActual Z score is", round(z, 2))
18
19 if(z > cutoff1 & z < cutoff2)
20 {
21   cat("\nWe cannot reject null hypothesis.")
22 }else{
23   cat("\nWe can reject null hypothesis.")
24 }
```

---

# Chapter 5

## Hypothesis Tests with Means of Samples

R code Exa 5.1a Determining the Characteristics of a Distribution of Means

```
1 # Page no. : 144
2
3 grade <- 6.67
4 N <- 2
5
6 ans <- grade / N
7
8 cat(" Answer is", round(ans,2))
9
10 pv <- 400
11 N <- 25
12
13 ans <- pv / N
14
15 cat("\nAnswer is", round(ans, 2))
```

---

**R code Exa 5.1b** Example of Determining the Characteristics of a Distribution of Means

```
1 # Page no. : 146 - 147
2
3 mean <- 200
4 sd <- 48
5 N <- 64
6
7 var <- sd ** 2
8 ans <- var / N
9
10 cat("Answer is", ans)
11
12 sd2 <- sqrt(ans)
13
14 cat("\nAnswer is", sd2)
```

---

**R code Exa 5.1c** How are you doing

```
1 # Page no. : 148
2
3 mean <- 60
4 sd <- 10
5 N <- 4
6
7 var <- sd ** 2
8 ans <- var / N
9
10 cat("Answer is", ans)
11
12 var2 <- sqrt(ans)
13
14 cat("\nAnswer is", var2)
```

---

**R code Exa 5.2a** Hypothesis testing with a distribution of means The Z Test

```
1 # Page no. : 148 - 149
2
3 mean <- 18
4 d_mean <- 10
5 sd <- 4
6
7 z <- (mean - d_mean) / sd
8
9 cat("Z score is", z)
```

---

**R code Exa 5.2b** Example 1 for Hypothesis testing with a distribution of means The Z Test

```
1 # Page no. : 150 - 152
2
3 n <- 64
4 mean <- 220
5 d_mean <- 200
6 sd <- 48
7 var <- sd ** 2
8 d_var <- var / n
9 d_sd <- sqrt(d_var)
10 alpha <- 0.05
11 zval <- qnorm(alpha, lower.tail = F)
12
13 z <- (mean - d_mean) / d_sd
14
15 cat("The value of z-score is", z)
16
```



```

17 if(z > zval)
18 {
19   cat("\nReject null hypothesis")
20 } else
21 {
22   cat("\nWe cannot reject null hypothesis")
23 }

```

---

**R code Exa 5.2c** Example 2 for Hypothesis testing with a distribution of means The Z Test

```

1 # Page no. : 152 - 154
2
3 n <- 25
4 mean <- 48
5 d_mean <- 53
6 sd <- 7
7 var <- sd ** 2
8 d_var <- var / n
9 d_sd <- sqrt(d_var)
10 alpha <- 0.01
11
12 zval <- qnorm(1 - alpha/2)
13 zval <- round(c(zval, -zval),2)
14
15 z <- (mean - d_mean) / d_sd
16
17 cat("The value of z-score is", z)
18
19 if(z > zval[1] || z < zval[2])
20 {
21   cat("\nReject null hypothesis")
22 } else
23 {
24   cat("\nWe cannot reject null hypothesis")

```

25 }

---

**R code Exa 5.2d** How are you doing

```
1 # Page no. : 154 - 155
2
3 n <- 36
4 mean <- 70
5 d_mean <- 75
6 sd <- 12
7 var <- sd ** 2
8 d_var <- var / n
9 d_sd <- sqrt(d_var)
10 alpha <- 0.05
11
12 zval <- qnorm(1 - alpha/2)
13 zval <- round(c(zval, -zval),2)
14
15 z <- (mean - d_mean) / d_sd
16
17 cat("The value of z-score is", z)
18
19 if(z > zval[1] || z < zval[2])
20 {
21   cat("\nReject null hypothesis")
22 } else
23 {
24   cat("\nWe cannot reject null hypothesis")
25 }
```

---

**R code Exa 5.3a** Confidence interval

```
1 # Page no. : 161
```

```

2
3 c_i <- 0.995
4 n <- 25
5 d_mean <- 53
6 mean <- 48
7 sd <- 7
8 var <- sd ** 2
9
10 se <- sqrt(var/n)
11
12 cat("Standard error is", se)
13
14 r_s <- round(qnorm(c_i), 2)
15
16 x <- se * r_s
17
18 y <- mean + c(-x, x)
19
20 cat("\nThe margin of error is given by", x)
21 cat("\nThe 99% confidence interval is given by", y)

```

---

**R code Exa 5.3b** How are you doing

```

1 # Page no. : 163
2
3 n <- 36
4 mean <- 70
5 d_mean <- 75
6 sd <- 12
7 var <- sd ** 2
8 c_i <- 0.975
9
10 se <- sqrt(var/n)
11
12 cat("Standard error is", se)

```

```
13
14 r_s <- round(qnorm(c_i), 2)
15
16 x <- se * r_s
17
18 y <- mean + c(-x, x)
19
20 cat("\nThe margin of error is given by", x)
21 cat("\nThe 99% confidence interval is given by", y)
```

---

#### R code Exa 5.4a Worked out examples 1

```
1 # Page no. : 167
2
3 sd <- 13
4 var <- sd ** 2
5 n <- 20
6
7 d_sd <- sqrt(var/n)
8
9 cat("Standard deviation for distribution of means is
    ", round(d_sd, 2))
```

---

#### R code Exa 5.4b Worked out examples 2

```
1 # Page no. : 167 - 168
2
3 n <- 75
4 mean <- 16
5 d_mean <- 15
6 sd <- 5
7 var <- sd ** 2
8 d_var <- var / n
```

```

 9 d_sd <- sqrt(d_var)
10 alpha <- 0.05
11
12 zval <- qnorm(1 - alpha/2)
13 zval <- round(c(zval, -zval),2)
14
15 z <- (mean - d_mean) / d_sd
16
17 cat("Z Score is", z)
18
19 if(z > zval[1] || z < zval[2])
20 {
21   cat("\nCan reject null hypothesis")
22 } else
23 {
24   cat("\nCannot reject null hypothesis")
25 }

```

---

### R code Exa 5.4c Worked out examples 3

```

1 # Page no. : 169
2
3 n <- 75
4 mean <- 16
5 d_mean <- 15
6 sd <- 5
7 var <- sd ** 2
8 d_var <- var / n
9 d_sd <- sqrt(d_var)
10 c_i <- 0.995
11
12 se <- d_sd
13
14 cat("Standard error is", se)
15

```

```
16 r_s <- round(qnorm(c_i), 2)
17
18 x <- se * r_s
19
20 y <- mean + c(-x, x)
21
22 cat("\nThe margin of error is given by", x)
23 cat("\nThe 99% confidence interval is given by", y)
```

---

# Chapter 6

## Making Sense of Statistical Significance

R code Exa 6.1a Figuring effect size

```
1 # Page no. : 184 - 185
2
3 m1 <- 220
4 m2 <- 200
5 sd <- 48
6
7 d1 <- (m1 - m2) / sd
8
9 cat("Estimated effect size is", round(d1, 2))
10
11 m1 <- 210
12
13 d2 <- (m1 - m2) / sd
14
15 cat("\nEstimated effect size is", round(d2, 2))
16
17 m1 <- 170
18
19 d3 <- (m1 - m2) / sd
```

20

```
21 cat("\nEstimated effect size is", round(d3, 2))
```

---

**R code Exa 6.1b** How are you doing

```
1 # Page no. : 188 - 189
2
3 m1 <- 540
4 m2 <- 500
5 sd <- 100
6
7 d1 <- (m1 - m2) / sd
8
9 cat("Estimated effect size is", round(d1, 2))
```

---

**R code Exa 6.2a** Determining power from predicted effect size

```
1 # Page no. : 197
2
3 m2 <- 200
4 d <- .20
5 sd <- 48
6
7 m <- m2 + (d * sd)
8
9 cat("Predicted value is", round(m, 2))
10
11 d <- .50
12
13 m <- m2 + (d * sd)
14
15 cat("\nPredicted value is", round(m, 2))
```

---



**R code Exa 6.2b** How are you doing

```
1 # Page no. : 204
2
3 m2 <- 500
4 d <- .80
5 sd <- 100
6
7 m <- m2 + (d * sd)
8
9 cat("Predicted value is", round(m, 2))
```

---

**R code Exa 6.3a** Advanced topic Figuring Statistical Power

```
1 # Page no. : 214 - 215
2
3 m <- 200
4 sd <- 48
5 var <- sd ** 2
6 n <- 64
7
8 d_sd <- sqrt(var/n)
9
10 m2 <- 208
11
12 alpha <- 0.05
13 zval <- qnorm(alpha, lower.tail = F)
14
15 r_s <- round((zval * d_sd) + m, 2)
16
17 z <- (r_s - m2) / d_sd
18
```

```
19 cat("Z is", round(z, 2))
```

---

**R code Exa 6.3b** How are you doing

```
1 # Page no. : 216 - 217
2
3 m <- 500
4 sd <- 100
5 var <- sd ** 2
6 n <- 60
7
8 d_sd <- sqrt(var/n)
9
10 m2 <- 540
11
12 alpha <- 0.05
13 zval <- qnorm(alpha, lower.tail = F)
14
15 r_s <- round((zval * d_sd) + m, 2)
16
17 z <- (r_s - m2) / d_sd
18
19 cat("Z is", round(z, 2))
```

---

**R code Exa 6.4a** Worked out examples 1

```
1 # Page no. : 218
2
3 m1 <- 37
4 m2 <- 40
5 sd <- 10
6
7 d1 <- (m1 - m2) / sd
```

```
8
9 cat(" Estimated effect size is", round(d1, 2))
```

---

#### R code Exa 6.4b Worked out examples 2

```
1 # Page no. : 218 - 219
2
3 m2 <- 40
4 d <- -0.20
5 sd <- 10
6
7 m <- m2 + (d * sd)
8
9 cat(" Predicted effect size is", round(m, 2))
```

---

#### R code Exa 6.4c Worked out examples 3

```
1 # Page no. : 219
2
3 m <- 40
4 sd <- 10
5 var <- sd ** 2
6 n <- 25
7
8 d_sd <- sqrt(var/n)
9
10 m2 <- 49
11
12 alpha <- 0.01
13 zval <- qnorm(alpha, lower.tail = F)
14
15 r_s <- round((zval * d_sd) + m, 2)
16
```

```
17 z <- (r_s - m2) / d_sd
18
19 cat("Z is", round(z, 2))
```

---

# Chapter 7

## Introduction to t Tests

**R code Exa 7.1a** Basic Principle of the t Test Estimating the Population Variance from the Sample Scores

```
1 # Page no. : 229 - 230
2
3 n <- 16
4 ss <- 694
5
6 var <- (ss)/(n - 1)
7
8 cat("Estimated population variance is", round(var,
9     2))
```

---

**R code Exa 7.1b** The Standard Deviation of the Distribution of Means

```
1 # Page no. : 231
2
3 s <- 16
4 var <- 46.27
5
```

```
6 d_var <- var / s
7
8 cat("Variance is", round(d_var, 2))
9
10 d_sd <- sqrt(d_var)
11
12 cat("\nStandard deviation is", round(d_sd, 2))
```

---

**R code Exa 7.1c** The Sample means score on the comparison distribution  
the t score

```
1 # Page no. : 234
2
3 m1 <- 21
4 m2 <- 17
5 sd <- 1.70
6
7 t <- (m1 - m2) / sd
8
9 cat("The t score is", round(t, 2))
```

---

**R code Exa 7.1d** Another example of a t test for a single sample

```
1 # Page no. : 235 – 236
2
3 rating <- c(5, 3, 6, 2, 7, 6, 7, 4, 2, 5)
4 n <- 10
5
6 mean <- mean(rating)
7
8 diff <- rating - mean
9
10 sq_diff <- diff ** 2
```

```

11
12 data_frame <- data.frame(rating, diff, sq_diff)
13
14 View(data_frame)
15
16 df <- n - 1
17
18 mu <- 4
19
20 var <- sum(data_frame$sq_diff) / df
21
22 d_var <- var / n
23
24 d_sd <- sqrt(d_var)
25
26 alpha <- 0.01
27
28 tval <- round(qt(1 - alpha/2, df), 2)
29 tval <- c(tval, -tval)
30
31 t <- (mean - mu) / d_sd
32
33 cat("\nValue of t is", t)
34
35 if(t > tval[1] || t < tval[2])
36 {
37   cat("\nReject null hypothesis")
38 } else
39 {
40   cat("\nCannot reject null hypothesis")
41 }

```

---

**R code Exa 7.1e** How are you doing

1 # Page no. : 238 – 240

```

2
3 n <- 4
4 df <- n - 1
5
6 m1 <- 23
7
8 scores <- c(20, 22, 22, 20)
9
10 mean <- mean(scores)
11
12 diff <- scores - mean
13
14 sq_diff <- diff ** 2
15
16 alpha <- 0.05
17
18 var <- sum(sq_diff) / df
19
20 d_var <- var / n
21
22 d_sd <- sqrt(d_var)
23
24 tval <- round(qt(alpha, df, lower.tail = T), 2)
25
26 t <- (mean - m1) / d_sd
27
28 cat("Value of t is", t)
29
30 if(t < tval)
31 {
32   cat("\nReject null hypothesis")
33 } else
34 {
35   cat("\nCannot reject null hypothesis")
36 }

```

---



**R code Exa 7.2a** Example of a t test for dependent means

```
1 # Page no. : 241 - 242
2
3 Husband <- LETTERS[1:19]
4
5 n <- 19
6
7 df <- n - 1
8
9 before <- c(126, 133, 126, 115, 108, 109, 124, 98,
             95, 120, 118, 126, 121, 116, 94, 105, 123, 125,
             128)
10 after <- c(115, 125, 96, 115, 119, 82, 93, 109, 72,
             104, 107, 118, 102, 115, 83, 87, 121, 100, 118)
11
12 diff <- after - before
13 mean <- mean(diff)
14
15 deviation <- round(diff - mean, 2)
16
17 sq_dev <- round(deviation ** 2, 2)
18
19 data_frame <- data.frame(Husband, before, after,
                           diff, deviation, sq_dev)
20
21 View(data_frame)
22
23 mu <- 0
24
25 var <- sum(data_frame$sq_dev) / df
26
27 d_var <- var / n
28
```

```

29 d_sd <- sqrt(d_var)
30
31 alpha <- 0.05
32
33 tval <- round(qt(1 - alpha/2, df), 2)
34 tval <- c(tval, -tval)
35
36 t <- (mean - mu) / d_sd
37
38 cat("\nValue of t is", t)
39
40 if(t > tval[1] || t < tval[2])
41 {
42   cat("\nReject null hypothesis")
43 } else
44 {
45   cat("\nCannot reject null hypothesis")
46 }

```

---

**R code Exa 7.2b** A second example of a t test for dependent means

```

1 # Page no. : 245 - 246
2
3 student <- c(1:10)
4
5 n <- 10
6
7 df <- n - 1
8
9 b_p <- c(1487.8, 1329.4, 1407.9, 1236.1, 1299.8,
10         1447.2, 1354.1, 1204.6, 1322.3, 1388.5)
11 c_p <- c(1487.2, 1328.1, 1405.9, 1234.0, 1298.2,
12         1444.7, 1354.3, 1203.7, 1320.8, 1386.8)

```

```

13 mean <- mean(diff)
14
15 deviation <- round(diff - mean, 2)
16
17 sq_dev <- round(deviation ** 2, 2)
18
19 data_frame <- data.frame(student, b_p, c_p, diff,
    deviation, sq_dev)
20
21 View(data_frame)
22
23 mu <- 0
24
25 var <- sum(data_frame$sq_dev) / df
26
27 d_var <- var / n
28
29 d_sd <- sqrt(d_var)
30
31 alpha <- 0.05
32
33 tval <- round(qt(alpha, df, lower.tail = F), 2)
34
35 t <- (mean - mu) / d_sd
36
37 cat("\nValue of t is", t)
38
39 if(t > tval)
40 {
41   cat("\nReject null hypothesis")
42 } else
43 {
44   cat("\nCannot reject null hypothesis")
45 }

```

---

### R code Exa 7.2c How are you doing

```
1 # Page no. : 249 - 250
2
3 person <- c(1:5)
4 n <- 5
5 df <- n - 1
6
7 before <- c(20, 30, 20, 40, 30)
8 after <- c(30, 50, 10, 30, 40)
9
10 diff <- after - before
11 mean <- mean(diff)
12
13 deviation <- round(diff - mean, 2)
14
15 sq_dev <- round(deviation ** 2, 2)
16
17 data_frame <- data.frame(person, before, after, diff
18   , deviation, sq_dev)
19 View(data_frame)
20
21 mu <- 0
22
23 var <- sum(data_frame$sq_dev) / df
24
25 d_var <- var / n
26
27 d_sd <- sqrt(d_var)
28
29 alpha <- 0.05
30
31 tval <- round(qt(1 - alpha/2, df), 2)
32 tval <- c(tval, -tval)
33
34 t <- (mean - mu) / d_sd
35
```

```
36 cat("\nValue of t is", t)
37
38 if(t > tval[1] || t < tval[2])
39 {
40   cat("\nReject null hypothesis")
41 } else
42 {
43   cat("\nCannot reject null hypothesis")
44 }
```

---

**R code Exa 7.3a** Effect size for the t test for dependent means

```
1 # Page no. : 252
2
3 mu1 <- 4
4 mu2 <- 0
5 sd <- 8
6
7 d <- (mu1 - mu2) / sd
8
9 cat("Effect size is", d)
10
11 m_d <- -12.05
12 var <- 153.49
13 sd <- round(sqrt(var), 2)
14
15 d <- round((m_d - mu2) / sd, 2)
16
17 cat("\nEffect size is", d)
```

---

**R code Exa 7.3b** How are you doing

```
1 # Page no. : 254
```

```

2
3 mu1 <- 40
4 mu2 <- 0
5 sd <- 80
6 alpha <- 0.05
7
8 d <- round((mu1 - mu2) / sd, 2)
9
10 cat("Effect size is", d)

```

---

#### R code Exa 7.4a Worked out examples 1

```

1 # Page no. : 259
2
3 scores <- c(14, 8, 6, 5, 13, 10, 10, 6)
4 n <- length(scores)
5 df <- n - 1
6 mean <- 6
7 alpha <- 0.05
8
9 m <- mean(scores)
10
11 diff <- scores - m
12
13 sq_diff <- diff ** 2
14
15 var <- sum(sq_diff) / df
16
17 d_var <- var / n
18
19 d_sd <- sqrt(d_var)
20
21 tval <- round(qt(1 - alpha/2, df), 2)
22 tval <- c(tval, -tval)
23

```

```

24 t <- (m - mean) / d_sd
25
26 cat(" Value of t is", t)
27
28 if(t > tval[1] || t < tval[2])
29 {
30   cat("\nReject null hypothesis")
31 } else
32 {
33   cat("\nCannot reject null hypothesis")
34 }

```

---

#### R code Exa 7.4b Worked out examples 2

```

1 # Page no. : 259 - 261
2
3 participant <- c(1:10)
4 n <- 10
5 df <- n - 1
6
7 before <- c(10.4, 12.6, 11.2, 10.9, 14.3, 13.2, 9.7,
8            11.5, 10.8, 13.1)
9 after <- c(10.8, 12.1, 12.1, 11.4, 13.9, 13.5, 10.9,
10          11.5, 10.4, 12.5)
11
12 diff <- after - before
13 mean <- mean(diff)
14
15 deviation <- round(diff - mean, 2)
16
17 sq_dev <- round(deviation ** 2, 2)
18
19 data_frame <- data.frame(participant, before, after,
20                          diff, deviation, sq_dev)

```

```

19 View(data_frame)
20
21 mu <- 0
22
23 var <- sum(data_frame$sq_dev) / df
24
25 d_var <- var / n
26
27 d_sd <- sqrt(d_var)
28
29 alpha <- 0.05
30
31 tval <- round(qt(1 - alpha/2, df), 2)
32 tval <- c(tval, -tval)
33
34 t <- (mean - mu) / d_sd
35
36 cat("\nValue of t is", t)
37
38 if(t > tval[1] || t < tval[2])
39 {
40   cat("\nReject null hypothesis")
41 } else
42 {
43   cat("\nCannot reject null hypothesis")
44 }

```

---



# Chapter 8

## The t Test for Independent Means

R code Exa 8.1a Estimating the population variance

```
1                                     # Page no. : 278
2                                     - 279
3 # Estimating the population variance
4
5 n1 <- 11   # Group 1 participants
6 df1 <- n1 - 1 # Degree of freedom for group 1
7 var1 <- 60  # Variance
8
9 n2 <- 31   # Group 2 participants
10 df2 <- n2 - 1 # Degree of freedom for group 2
11 var2 <- 80  # Variance
12
13 df <- df1 + df2 # Total degrees of freedom
14
15 p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)
16         # Pooled variance
17 cat("Pooled variance is", p_var)
```

---

**R code Exa 8.1b** Figuring the variance of each of the two distributions of means

```
1                                     # Page no. : 279 -
                                     280
2
3 # Figuring the variance of each of the two
   distributions of means
4
5 n1 <- 11   # Group 1 participants
6 n2 <- 31   # Group 2 participants
7 p_var <- 75 # Pooled variance
8
9 var1 <- p_var / n1 # Variance for group 1
10
11 cat("Variance for group 1 is", round(var1, 2))
12
13 var2 <- p_var / n2 # Variance for group 1
14
15 cat("Variance for group 2 is", round(var2, 2))
```

---

**R code Exa 8.1c** The variance and standard deviation of the distribution of differences between means

```
1                                     # Page no. :
                                     280
2
3 # The variance and standard deviation of the
   distribution of differences between means
4
5 var1 <- 6.82 # Variance of the distribution of
   means of group 1
```

```

6 var2 <- 2.42 # Variance of the distribution of
  means of group 2
7
8 diff_var <- var1 + var2 # Variance difference
9
10 cat("Variance difference is", diff_var)
11
12 diff_sd <- sqrt(diff_var) # Standard deviation
  difference
13
14 cat("Standard deviation difference is", round(diff_
  sd, 2))

```

---

**R code Exa 8.1d** The t score for the difference between the two actual means

```

1 # Page no. : 281
  - 282
2
3 # The t score for the difference between the two
  actual means
4
5 m1 <- 198 # Mean of the first sample
6 m2 <- 190 # Mean of the second sample
7
8 diff_sd <- 3.04 # Standard deviation difference
9
10 t <- (m1 - m2) / diff_sd # t score
11
12 cat("t score is", round(t, 2))

```

---

**R code Exa 8.1e** How are you doing

```
1
2
3 # How are you doing?
4
5 # 5th Question (a) part
6
7 n1 <- 21 # Sample 1 participants
8 df1 <- n1 - 1 # Degree of freedom 1
9 var1 <- 100 # Variance
10
11 n2 <- 31 # Sample 1 participants
12 df2 <- n2 - 1 # Degree of freedom 1
13 var2 <- 200 # Variance
14
15 df <- df1 + df2 # Total degrees of freedom
16
17 p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)
18 # Pooled variance
19
20 cat("Pooled variance is", p_var)
21
22 d_var1 <- p_var / n1 # Variance of the
23 # distribution of means of sample 1
24
25 cat("Variance of the distribution of means of sample
26 # 1 is", round(d_var1, 2))
27
28 d_var2 <- p_var / n2 # Variance of the
29 # distribution of means of sample 2
30
31 cat("Variance of the distribution of means of sample
32 # 2 is", round(d_var2, 2))
33
34 diff_var <- d_var1 + d_var2 # Variance difference
35
36 cat("Variance difference is", round(diff_var, 2))
37
```

```

33 diff_sd <- sqrt(diff_var) # Standard deviation
    difference
34
35 cat("Standard deviation difference is", round(diff_
    sd, 2))

```

---

**R code Exa 8.2a** Example of a T test for independent means

```

1 # # Page no. :
    284
2 #
3 # # Example of a T test for independent means
4
5 # Expressive Writing Group
6
7 score1 <- c(77, 88, 77, 90, 68, 74, 62, 93, 82, 79)
8 n1 <- length(score1) # Sample size 1
9 df1 <- n1 - 1 # Degree of freedom
10 m1 <- mean(score1) # Mean
11 deviation1 <- score1 - m1
12 sq_dev1 <- deviation1 ** 2 # Squared deviation
13
14 data_frame1 <- data.frame(score1, deviation1, sq_
    dev1)
15 View(data_frame1)
16
17 # Control Writing Group
18
19 score2 <- c(87, 77, 71, 70, 63, 50, 58, 63, 76, 65)
20 n2 <- length(score2) # Sample size 2
21 df2 <- n2 - 1 # Degree of freedom
22 m2 <- mean(score2) # Mean
23 deviation2 <- score2 - m2
24 sq_dev2 <- deviation2 ** 2 # Squared deviation
25

```

```

26 data_frame2 <- data.frame(score2, deviation2, sq_
    dev2)
27 View(data_frame2)
28
29 var1 <- sum(data_frame1$sq_dev1) / df1 # Variance
    1
30 var2 <- sum(data_frame2$sq_dev2) / df2 # Variance
    2
31
32 df <- df1 + df2 # Total degrees of freedom
33
34 p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)
    # Pooled variance
35
36 cat("Pooled variance is", round(p_var, 2))
37
38 d_var1 <- p_var / n1 # Variance of the
    distribution of means of sample 1
39
40 cat("# Variance of the distribution of means of
    sample 1 is", round(d_var1, 2))
41
42 d_var2 <- p_var / n2 # Variance of the
    distribution of means of sample 2
43
44 cat("# Variance of the distribution of means of
    sample 2 is", round(d_var2, 2))
45
46 diff_var <- d_var1 + d_var2 # Variance difference
47
48 cat("Variance difference is", diff_var)
49
50 diff_sd <- sqrt(diff_var) # Standard deviation
    difference
51
52 cat("Standard deviation difference is", round(diff_
    sd, 2))
53

```

```

54 # Two - Tailed
55
56 alpha <- 0.05 # 5% significance level
57
58 tval <- qt(1-alpha/2, df)
59 tval <- c(tval, -tval)
60
61 t <- round((m1 - m2) / diff_sd, 2) # t score
62
63 cat("t score is", t)
64
65 if(t > tval[1] || t < tval[2])
66 {
67   cat("Reject null hypothesis")
68 } else
69 {
70   cat("Cannot reject null hypothesis")
71 }

```

---

**R code Exa 8.2b** A second example of a T test for independent means

```

1
2
3 # A second example of a T test for independent means
4
5 # Experimental Group (Receiving Special Program)
6
7 score1 <- c(6, 4, 9, 7, 7, 3)
8 n1 <- length(score1) # Sample size 1
9 df1 <- n1 - 1 # Degree of freedom
10 m1 <- mean(score1) # Mean
11 deviation1 <- score1 - m1
12 sq_dev1 <- deviation1 ** 2 # Squared deviation
13

```

```

14 data_frame1 <- data.frame(score1, deviation1, sq_
    dev1)
15 View(data_frame1)
16
17 # Control Group (Receiving Ordinary Program)
18
19 score2 <- c(6, 1, 5, 3, 1, 1, 4, 3)
20 n2 <- length(score2) # Sample size 2
21 df2 <- n2 - 1 # Degree of freedom
22 m2 <- mean(score2) # Mean
23 deviation2 <- score2 - m2
24 sq_dev2 <- deviation2 ** 2 # Squared deviation
25
26 data_frame2 <- data.frame(score2, deviation2, sq_
    dev2)
27 View(data_frame2)
28
29 var1 <- sum(data_frame1$sq_dev1) / df1 # Variance
    1
30 var2 <- sum(data_frame2$sq_dev2) / df2 # Variance
    2
31
32 df <- df1 + df2 # Total degrees of freedom
33
34 p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)
    # Pooled variance
35
36 cat("Pooled variance is", round(p_var, 2))
37
38 d_var1 <- p_var / n1 # Variance of the
    distribution of means of sample 1
39
40 cat("# Variance of the distribution of means of
    sample 1 is", round(d_var1, 2))
41
42 d_var2 <- p_var / n2 # Variance of the
    distribution of means of sample 2
43

```



```

44 cat("# Variance of the distribution of means of
      sample 2 is", round(d_var2, 2))
45
46 diff_var <- d_var1 + d_var2 # Variance difference
47
48 cat("Variance difference is", diff_var)
49
50 diff_sd <- sqrt(diff_var) # Standard deviation
      difference
51
52 cat("Standard deviation difference is", round(diff_
      sd, 2))
53
54 # Two - Tailed
55
56 alpha <- 0.05 # 5% significance level
57
58 tval <- qt(1-alpha/2, df)
59 tval <- c(tval, -tval)
60
61 t <- round((m1 - m2) / diff_sd, 2) # t score
62
63 cat("t score is", t)
64
65 if(t > tval[1] || t < tval[2])
66 {
67   cat("Reject null hypothesis")
68 } else
69 {
70   cat("Cannot reject null hypothesis")
71 }

```

---

**R code Exa 8.2c** How are you doing

```

1                                     # Page no. : 289 -
                                     291
2
3 # How are you doing?
4
5 # 2nd Question (a) part
6
7 alpha <- 0.05   # 5% significance level
8
9 n1 <- 26   # Sample size 1
10 df1 <- n1 - 1   # Degree of freedom
11 m1 <- 5   # Mean 1
12 var1 <- 10   # Variance 1
13
14 n2 <- 36   # Sample size 2
15 df2 <- n2 - 1   # Degree of freedom
16 m2 <- 8   # Mean 2
17 var2 <- 12   # Variance 2
18
19 df <- df1 + df2   # Total degrees of freedom
20
21 p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)
    # Pooled variance
22
23 cat("Pooled variance is \n", round(p_var, 2))
24
25 d_var1 <- p_var / n1   # Variance of the
    distribution of means of sample 1
26
27 cat("Variance of the distribution of means of sample
    1 is \n", round(d_var1, 2))
28
29 d_var2 <- p_var / n2   # Variance of the
    distribution of means of sample 2
30
31 cat("Variance of the distribution of means of sample
    2 is \n", round(d_var2, 2))
32

```

```

33 diff_var <- d_var1 + d_var2 # Variance difference
34
35 cat("Variance difference is \n", round(diff_var, 2))
36
37 diff_sd <- sqrt(diff_var) # Standard deviation
   difference
38
39 cat("Standard deviation difference is \n", round(
   diff_sd, 2))
40
41 # One - Tailed
42
43 tval <- qt(alpha, df)
44
45 t <- round((m1 - m2) / diff_sd, 2) # t score
46
47 cat("t score is \n", t)
48
49 if(t < tval)
50 {
51   cat("Reject null hypothesis")
52 } else
53 {
54   cat("Cannot reject null hypothesis")
55 }

```

---

**R code Exa 8.3a** Effect size for the t test for independent means

```

1 # Page no. :
   293 - 294
2
3 # Effect size for the t test for independent means
4
5 mu1 <- 29 # Mean 1
6 mu2 <- 21 # Mean 2

```

```

7 sd <- 10 # Standard deviation
8
9 d <- (mu1 - mu2) / sd # Effect size
10
11 cat("Effect size is \n", d)
12
13 m1 <- 33.10 # Mean 1
14 m2 <- 27.00 # Mean 2
15 p_sd <- 12.99 # Pooled standard deviation
16
17 e_d <- round((m1 - m2) / p_sd, 2) # Estimated
    effect size
18
19 cat("Estimated effect size is \n", e_d)

```

---

#### R code Exa 8.4a Harmonic mean

```

1 # Page no. : 300
2
3 # Harmonic mean
4
5 n1 <- 11 # Sample 1
6 n2 <- 31 # Sample 2
7
8 total <- n1 + n2
9
10 h_m <- round((2 * n1 * n2) / total, 2) # Harmonic
    mean
11
12 cat("Harmonic mean is \n", h_m)
13
14 # How are you doing?
15
16 # 1st Question (Only harmonic mean)
17

```

```

18 n1 <- 6   # Sample 1
19 n2 <- 34  # Sample 2
20
21 total <- n1 + n2
22
23 h_m <- round((2 * n1 * n2) / total, 2)  # Harmonic
    mean
24
25 cat("Harmonic mean is \n", h_m)

```

---

### R code Exa 8.5a Worked out examples 1

```

1
2
3 # Worked out examples 1
4
5 n1 <- 40   # Sample 1
6 df1 <- n1 - 1 # Degree of freedom
7 var1 <- 15  # Variance 1
8
9 n2 <- 60   # Sample 2
10 df2 <- n2 - 1 # Degree of freedom
11 var2 <- 12  # Variance 2
12
13 df <- df1 + df2 # Total degrees of freedom
14
15 p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)
    # Pooled variance
16
17 cat("Pooled variance is \n", round(p_var, 2))
18
19 d_var1 <- p_var / n1 # Variance of the
    distribution of means of sample 1
20

```

```

21 cat(" Variance of the distribution of means of sample
      1 is \n", round(d_var1, 2))
22
23 d_var2 <- p_var / n2 # Variance of the
      distribution of means of sample 2
24
25 cat(" Variance of the distribution of means of sample
      2 is \n", round(d_var2, 2))
26
27 diff_var <- d_var1 + d_var2 # Variance difference
28
29 cat(" Variance difference is \n", round(diff_var, 2))
30
31 diff_sd <- sqrt(diff_var) # Standard deviation
      difference
32
33 cat(" Standard deviation difference is \n", round(
      diff_sd, 2))

```

---

### R code Exa 8.5b Worked out examples 2

```

1 # Page no. : 302 -
      303
2
3 # Worked out examples 2
4
5 score1 <- c(7, 6, 9, 7, 6)
6 n1 <- length(score1) # Sample size 1
7 df1 <- n1 - 1 # Degree of freedom
8 m1 <- mean(score1) # Mean
9 deviation1 <- score1 - m1
10 sq_dev1 <- deviation1 ** 2 # Squared deviation
11
12 data_frame1 <- data.frame(score1, deviation1, sq_
      dev1)

```

```

13 View(data_frame1)
14
15 score2 <- c(5, 2, 4, 3, 6)
16 n2 <- length(score2) # Sample size 2
17 df2 <- n2 - 1 # Degree of freedom
18 m2 <- mean(score2) # Mean
19 deviation2 <- score2 - m2
20 sq_dev2 <- deviation2 ** 2 # Squared deviation
21
22 data_frame2 <- data.frame(score2, deviation2, sq_
    dev2)
23 View(data_frame2)
24
25 var1 <- sum(data_frame1$sq_dev1) / df1 # Variance
    1
26 var2 <- sum(data_frame2$sq_dev2) / df2 # Variance
    2
27
28 df <- df1 + df2 # Total degrees of freedom
29
30 p_var <- ((df1 / df) * var1) + ((df2 / df) * var2)
    # Pooled variance
31
32 cat("Pooled variance is \n", round(p_var, 2))
33
34 d_var1 <- p_var / n1 # Variance of the
    distribution of means of sample 1
35
36 cat("Variance of the distribution of means of sample
    1 is \n", round(d_var1, 2))
37
38 d_var2 <- p_var / n2 # Variance of the
    distribution of means of sample 2
39
40 cat("Variance of the distribution of means of sample
    2 is \n", round(d_var2, 2))
41
42 diff_var <- d_var1 + d_var2 # Variance difference

```

```

43
44 cat("Variance difference is \n", diff_var)
45
46 diff_sd <- sqrt(diff_var) # Standard deviation
    difference
47
48 cat("Standard deviation difference is \n", round(
    diff_sd, 2))
49
50 # Two - Tailed
51
52 alpha <- 0.05 # 5% significance level
53
54 tval <- qt(1-alpha/2, df)
55 tval <- c(tval, -tval)
56
57 t <- round((m1 - m2) / diff_sd, 2) # t score
58
59 cat("t score is \n", t)
60
61 if(t > tval[1] || t < tval[2])
62 {
63   cat("Reject null hypothesis")
64 } else
65 {
66   cat("Cannot reject null hypothesis")
67 }

```

---

**R code Exa 8.5c** Worked out examples 3

```

1 # Page no. :
    304
2
3 # Worked out examples 3
4

```



```
5 n1 <- 22 # Sample 1
6 n2 <- 51 # Sample 2
7
8 total <- n1 + n2
9
10 h_m <- round((2 * n1 * n2) / total, 2) # Harmonic
    mean
11
12 cat("Harmonic mean is", h_m)
```

---

## Chapter 9

# Introduction to the Analysis of Variance

R code Exa 9.1a Carrying out an Analysis of Variance

```
1                                     # Page no. : 325
                                     - 331
2
3 # Carrying out an Analysis of Variance
4
5 # Page no. : 325 - 326
6
7 # Criminal Record Group
8
9 rating1 <- c(10, 7, 5, 10, 8)
10 n1 <- length(rating1) # Sample 1 size
11 df1 <- n1 - 1 # Degree of freedom
12 m1 <- mean(rating1) # Mean of rating 1
13 deviation1 <- rating1 - m1 # Deviation from mean 1
14 sq_dev1 <- deviation1 ** 2 # Squared deviation
    from mean 1
15
16 data_frame1 <- data.frame(rating1, deviation1, sq_
    dev1)
```

```

17 View(data_frame1)
18
19 # Clean Record Group
20
21 rating2 <- c(5, 1, 3, 7, 4)
22 n2 <- length(rating2) # Sample 2 size
23 df2 <- n2 - 1 # Degree of freedom
24 m2 <- mean(rating2) # Mean of rating 2
25 deviation2 <- rating2 - m2 # Deviation from mean 2
26 sq_dev2 <- deviation2 ** 2 # Squared deviation
    from mean 2
27
28 data_frame2 <- data.frame(rating2, deviation2, sq_
    dev2)
29 View(data_frame2)
30
31 # No Information Group
32
33 rating3 <- c(4, 6, 9, 3, 3)
34 n3 <- length(rating1) # Sample 3 size
35 df3 <- n3 - 1 # Degree of freedom
36 m3 <- mean(rating3) # Mean of rating 3
37 deviation3 <- rating3 - m3 # Deviation from mean 3
38 sq_dev3 <- deviation3 ** 2 # Squared deviation
    from mean 3
39
40 data_frame3 <- data.frame(rating3, deviation3, sq_
    dev3)
41 View(data_frame3)
42
43 var1 <- var(data_frame1$rating1) # Variance for
    rating 1
44 var2 <- var(data_frame2$rating2) # Variance for
    rating 2
45 var3 <- var(data_frame3$rating3) # Variance for
    rating 3
46
47 cat("Variance for rating 1 is", var1)

```

```

48 cat(" Variance for rating 2 is", var2)
49 cat(" Variance for rating 3 is", var3)
50
51 # Figuring the within – groups estimate of the
    population variance
52
53 # Page no. : 326 – 327
54
55 n_g <- 3 # Number of groups
56 df_within <- (n1 + n2 + n3) - n_g # Degree of
    freedom within-groups
57
58 var_within <- round((var1 + var2 + var3) / n_g, 2)
    # Variance within-groups
59
60 cat(" Variance within-groups is", var_within)
61
62 # Figuring the between – groups estimate of the
    population variance
63
64 # Page no. : 327 – 329
65
66 df_between <- n_g - 1 # Degree of freedom between-
    groups
67 gm <- round((m1 + m2 + m3) / n_g, 2) # Grand mean
68
69 gm_dev <- c((m1 - gm), (m2 - gm), (m3 - gm)) #
    Deviation from the grand mean
70 sq_gm_dev <- round(gm_dev ** 2, 2) # Squared
    deviation from the grand mean
71
72 d_var <- round(sum(sq_gm_dev) / df_between, 2) #
    Variance for the distribution of means
73
74 cat(" Variance for the distribution of means is", d_
    var)
75
76 N <- 5 # 5 elements in each rating

```

```

77 var_between <- d_var * N # Variance between-groups
78
79 cat("Variance between-groups is", var_between)
80
81 # Figuring the F ratio
82
83 # Page no. : 329
84
85 fvalue <- round(var_between / var_within, 2) # F
      ratio
86 cat("F ratio is", fvalue)
87
88 # Direct method to find F ratio
89
90 DF <- data.frame(rating1, rating2, rating3)
91 x <- c(t(as.matrix(DF)))
92 f <- c("r_1", "r_2", "r_3")
93 tm <- gl(3, 1, 5*3, factor(f))
94 result <- anova(lm(x ~ tm)) # Similar to aov(x ~
      tm)
95
96 cat("F ratio is", round(result$'F value'[1], 2))
97
98 # F value
99
100 # Page no. : 331
101
102 alpha <- 0.05 # 5% Sgnificance level
103
104 fval <- qf(0.95, df_between, df_within) # 1 - alpha
      = 1 - 0.05 = 0.95
105 fval <- round(fval, 2)
106
107 if(fvalue >= fval)
108 {
109   cat("Reject Null Hypothesis")
110 } else {
111   cat("Cannot reject Null Hypothesis")

```

---

**R code Exa 9.1b** How are you doing

```
1 # Page no. : 331 - 333
2
3 # How are you doing?
4
5 # Data given
6
7 scores_A <- c(5, 7)
8 m1 <- 6 # Mean 1
9 n1 <- 2 # Sample size
10
11 scores_B <- c(6, 9)
12 m2 <- 7.5 # Mean 2
13 n2 <- 2 # Sample size
14
15 scores_c <- c(8, 9)
16 m3 <- 8.5 # Mean 3
17 n3 <- 2 # Sample size
18
19 # 1st Question (c) part
20
21 n_s <- 3 # Number of samples
22 df_within <- (n1 + n2 + n3) - n_s # Degree of
    freedom within-groups
23
24 var_within <- round((var(scores_A) + var(scores_B) +
    var(scores_c)) / n_s, 2)
25 # Variance within-groups
26
27
28 cat("Variance within-groups is", var_within)
29
```

```

30 # 2nd and 3rd Question (c) parts
31
32 df_between <- n_s - 1 # Degree of freedom between-
    groups
33 gm <- round((m1 + m2 + m3) / n_s, 2) # Grand mean
34
35 gm_dev <- c((m1 - gm), (m2 - gm), (m3 - gm)) #
    Deviation from the grand mean
36 sq_gm_dev <- round(gm_dev ** 2, 2) # Squared
    deviation from the grand mean
37
38 d_var <- round(sum(sq_gm_dev) / df_between, 2) #
    Variance for the distribution of means
39
40 cat("Variance for the distribution of means is", d_
    var)
41
42 N <- 2 # 2 elements in each rating
43 var_between <- d_var * N # Variance between-groups
44
45 cat("Variance between-groups is", var_between)
46
47 # 4th Question (c) part
48
49 fvalue <- round(var_between / var_within, 2) # F
    ratio
50 cat("F ratio is", fvalue)
51
52 # 5th and 6th Question (c) parts
53
54 cat("Between-groups degree of freedom is", df_
    between)
55 cat("Within-groups degree of freedom is", df_within)
56
57 alpha <- 0.05 # 5% Sgnificance level
58
59 fval <- qf(0.95, df_between, df_within) # 1 - alpha
    = 1 - 0.05 = 0.95

```

```
60 fval <- round(fval, 2)
61
62 cat("Cutoff F for the 0.05 significance level is",
      fval)
```

---

### R code Exa 9.1c Another example

```
1
2
3 # Another example
4
5 # Attachment Styl – Secure
6
7 n1 <- 10 # Sample size
8 df1 <- n1 - 1 # Degree of freedom
9 m1 <- 2.10 # Mean
10 sd1 <- 1.66 # Standard deviation
11 var1 <- 2.76 # Variance
12
13 # Attachment Styl – Avoidant
14
15 n2 <- 10 # Sample size
16 df2 <- n2 - 1 # Degree of freedom
17 m2 <- 3.70 # Mean
18 sd2 <- 1.89 # Standard deviation
19 var2 <- 3.57 # Variance
20
21 # Attachment Styl – Anxious–Ambivalent
22
23 n3 <- 10 # Sample size
24 df3 <- n3 - 1 # Degree of freedom
25 m3 <- 4.20 # Mean
26 sd3 <- 1.93 # Standard deviation
27 var3 <- 3.72 # Variance
```

# Page no. :  
334 – 335



```

28
29 n_s <- 3 # Number of samples
30 df_within <- (n1 + n2 + n3) - n_s # Degree of
    freedom within-groups
31 df_between <- n_s - 1 # Degree of freedom between-
    groups
32
33 alpha <- 0.05 # 5% Sgnificance level
34
35 fval <- qf(0.95, df_between, df_within) # 1 - alpha
    = 1 - 0.05 = 0.95
36 fval <- round(fval, 2)
37
38 cat("Cutoff F for the 0.05 significance level is",
    fval)
39
40 var_within <- round((var1 + var2 + var3) / n_s, 2)
    # Variance within-groups
41
42 cat("Variance within-groups is", var_within)
43
44 gm <- round((m1 + m2 + m3) / n_s, 2) # Grand mean
45
46 gm_dev <- c((m1 - gm), (m2 - gm), (m3 - gm)) #
    Deviation from the grand mean
47 sq_gm_dev <- round(gm_dev ** 2, 2) # Squared
    deviation from the grand mean
48
49 d_var <- round(sum(sq_gm_dev) / df_between, 2) #
    Variance for the distribution of means
50
51 cat("Variance for the distribution of means is", d_
    var)
52
53 N <- 10 # 10 elements in each rating
54 var_between <- d_var * N # Variance between-groups
55
56 cat("Variance between-groups is", var_between)

```

```

57
58 fvalue <- round(var_between / var_within, 2) # F
      ratio
59 cat("F ratio is", fvalue)
60
61 if(fvalue >= fval)
62 {
63   cat("Reject Null Hypothesis")
64 } else {
65   cat("Cannot reject Null Hypothesis")
66 }

```

---

**R code Exa 9.1d** How are you doing

```

1
2
3 # How are you doing?
4
5 # 1st Question (a) part
6
7 N <- 16 # Participants in each group
8 df <- N - 1 # Degree of freedom
9
10 # Group 1
11
12 m1 <- 20 # Mean
13 var1 <- 8 # Variance
14
15 # Group 2
16
17 m2 <- 22 # Mean
18 var2 <- 9 # Variance
19
20 # Group 3

```

```

21
22 m3 <- 18 # Mean
23 var3 <- 7 # Variance
24
25 alpha <- 0.01 # 1% Sgnificance level
26
27 n_s <- 3 # Number of samples
28 df_within <- (3 * N) - n_s # Degree of freedom
    within-groups
29 df_between <- n_s - 1 # Degree of freedom between-
    groups
30
31
32 fval <- qf(0.99,df_between, df_within) # 1 - alpha
    = 1 - 0.01 = 0.99
33 fval <- round(fval, 2)
34
35 cat("Cutoff F for the 0.05 significance level is",
    fval)
36
37 var_within <- round((var1 + var2 + var3) / n_s, 2)
    # Variance within-groups
38
39 cat("Variance within-groups is", var_within)
40
41 gm <- round((m1 + m2 + m3) / n_s, 2) # Grand mean
42
43 gm_dev <- c((m1 - gm), (m2 - gm), (m3 - gm)) #
    Deviation from the grand mean
44 sq_gm_dev <- round(gm_dev ** 2, 2) # Squared
    deviation from the grand mean
45
46 d_var <- round(sum(sq_gm_dev) / df_between, 2) #
    Variance for the distribution of means
47
48 cat("Variance for the distribution of means is", d_
    var)
49

```

```

50 var_between <- d_var * N # Variance between-groups
51
52 cat("Variance between-groups is", var_between)
53
54 fvalue <- round(var_between / var_within, 2) # F
    ratio
55 cat("F ratio is", fvalue)
56
57 if(fvalue >= fval)
58 {
59   cat("Reject Null Hypothesis")
60 } else {
61   cat("Cannot reject Null Hypothesis")
62 }

```

---

### R code Exa 9.2a Planned contrast an example

```

1 # Page no. :
    340 - 341
2
3 # Planned contrast an example
4
5 m1 <- 8 # Mean 1
6 m2 <- 5 # Mean 2
7 n_s <- 2 # Number of samples for contrast
8 N <- 5 # 5 elements in each sample
9
10 var_within <- 5.33 # Variance within-groups
11
12 gm <- (m1 + m2) / n_s # Grand mean
13 gm_dev <- c((m1 - gm), (m2 - gm)) # Deviation from
    the grand mean
14 sq_gm_dev <- round(gm_dev ** 2, 2) # Squared
    deviation from the grand mean
15 df_between <- n_s - 1 # Degree of freedom between-

```

```

    groups (2 samples)
16 df_within <- (3 * N) - (n_s + 1) # Degree of
    freedom within-groups (3 samples)
17
18 d_var <- round(sum(sq_gm_dev) / df_between, 2) #
    Variance for the distribution of means
19
20 cat("Variance for the distribution of means is", d_
    var)
21
22 var_between <- d_var * N # Variance between-groups
23
24 cat("Variance between-groups is", var_between)
25
26 fvalue <- round(var_between / var_within, 2) # F
    ratio
27 cat("F ratio is", fvalue)
28
29 alpha <- 0.05 # 5% Sgnificance level
30
31 fval <- qf(0.95, df_between, df_within) # 1 - alpha
    = 1 - 0.05 = 0.95
32 fval <- round(fval, 2)
33
34 cat("Cutoff F for the 0.05 significance level is",
    fval)
35
36 if(fvalue >= fval)
37 {
38   cat("Reject null hypothesis (Significant)")
39 } else {
40   cat("Cannot reject null hypothesis (not
    significant)")
41 }

```

---

## R code Exa 9.2b Planned contrast a second example

```
1                                     # Page no. :
                                     341
2
3 # Planned contrast a second example
4
5 m1 <- 8   # Mean 1
6 m2 <- 4   # Mean 2
7 n_s <- 2  # Number of samples for contrast
8 N <- 5    # 5 elements in each sample
9
10 var_within <- 5.33 # Variance within-groups
11
12 gm <- (m1 + m2) / n_s # Grand mean
13 gm_dev <- c((m1 - gm), (m2 - gm)) # Deviation from
   the grand mean
14 sq_gm_dev <- round(gm_dev ** 2, 2) # Squared
   deviation from the grand mean
15 df_between <- n_s - 1 # Degree of freedom between-
   groups (2 samples)
16 df_within <- (3 * N) - (n_s + 1) # Degree of
   freedom within-groups (3 samples)
17
18 d_var <- round(sum(sq_gm_dev) / df_between, 2) #
   Variance for the distribution of means
19
20 cat("Variance for the distribution of means is", d_
   var)
21
22 var_between <- d_var * N # Variance between-groups
23
24 cat("Variance between-groups is", var_between)
25
26 fvalue <- round(var_between / var_within, 2) # F
   ratio
27 cat("F ratio is", fvalue)
28
```

```

29 alpha <- 0.05 # 5% Sgnificance level
30
31 fval <- qf(0.95,df_between, df_within) # 1 - alpha
    = 1 - 0.05 = 0.95
32 fval <- round(fval, 2)
33
34 cat("Cutoff F for the 0.05 significance level is",
    fval)
35
36 if(fvalue >= fval)
37 {
38   cat("Reject null hypothesis (Significant)")
39 } else {
40   cat("Cannot reject null hypothesis (not
    significant)")
41 }

```

---

### R code Exa 9.2c How are you doing

```

1 # Page no. : 342
    - 343
2
3 # How are you doing?
4
5 # 3rd Question
6
7 N <- 25 # 25 elements in each sample
8 var_within <- 100 # Variance within-groups
9
10 m1 <- 10 # Mean 1
11 m2 <- 16 # Mean 2
12
13 n_s <- 2 # Number of samples for contrast
14
15 gm <- (m1 + m2) / n_s # Grand mean

```

```

16 gm_dev <- c((m1 - gm), (m2 - gm)) # Deviation from
    the grand mean
17 sq_gm_dev <- round(gm_dev ** 2, 2) # Squared
    deviation from the grand mean
18 df_between <- n_s - 1 # Degree of freedom between-
    groups (2 samples)
19 df_within <- (3 * N) - (n_s + 1) # Degree of
    freedom within-groups (3 samples)
20
21 d_var <- round(sum(sq_gm_dev) / df_between, 2) #
    Variance for the distribution of means
22
23 cat("Variance for the distribution of means is", d_
    var)
24
25 var_between <- d_var * N # Variance between-groups
26
27 cat("Variance between-groups is", var_between)
28
29 fvalue <- round(var_between / var_within, 2) # F
    ratio
30 cat("F ratio is", fvalue)
31
32 alpha <- 0.05 # 5% Sgnificance level
33
34 fval <- qf(0.95, df_between, df_within) # 1 - alpha
    = 1 - 0.05 = 0.95
35 fval <- round(fval, 2)
36
37 cat("Cutoff F for the 0.05 significance level is",
    fval)
38
39 if(fvalue >= fval)
40 {
41   cat("Reject null hypothesis (Significant)")
42 } else {
43   cat("Cannot reject null hypothesis (not
    significant)")

```



44 }

---

### R code Exa 9.3a Scheffe Test

```
1                                     # Page no. :
                                     344 - 345
2
3 # Scheffe Test
4
5 N <- 5   # 5 elements in each sample
6 f <- 4.22 # F value (overall)
7
8 df_between <- 2   # Degree of freedom between-groups
9 df_within <- (3 * N) - (df_between + 1) # Degree
   of freedom within-groups (3 samples)
10 fvalue <- f / df_between # F value (contrast)
11
12 cat("F value (contrast) is", fvalue)
13
14 alpha <- 0.05 # 5% Sgnificance level
15
16 fval <- qf(0.95, df_between, df_within) # 1 - alpha
   = 1 - 0.05 = 0.95
17 fval <- round(fval, 2)
18
19 cat("Cutoff F for the 0.05 significance level is",
   fval)
20
21 if(fvalue >= fval)
22 {
23   cat("Reject null hypothesis (Significant)")
24 } else {
25   cat("Cannot reject null hypothesis (not
   significant)")
26 }
```

```

27 # How are you doing?
28
29 # 5th Question
30
31 N <- 50 # 50 elements in each sample
32 f <- 12.60 # F value (overall)
33
34 df_between <- 3 # Degree of freedom between-groups
35 df_within <- (4 * N) - (df_between + 1) # Degree
  of freedom within-groups (4 samples)
36 fvalue <- f / df_between # F value (contrast)
37
38 cat("F value (contrast) is", fvalue)
39
40 alpha <- 0.05 # 5% Sgnificance level
41
42 fval <- qf(0.95, df_between, df_within) # 1 - alpha
  = 1 - 0.05 = 0.95
43 fval <- round(fval, 2)
44
45 cat("Cutoff F for the 0.05 significance level is",
  fval)
46
47 if(fvalue >= fval)
48 {
49   cat("Reject null hypothesis (Significant)")
50 } else {
51   cat("Cannot reject null hypothesis (not
  significant)")
52 }

```

---

**R code Exa 9.4a** Effect size for the analysis of variance

1

# Page no. : 345  
- 346

```

2
3 # Effect size for the analysis of variance
4
5 var_within <- 5.33 # Variance within-groups
6 var_between <- 21.70 # Variance within-groups
7
8 df_between <- 2 # Degree of freedom between-groups
9 df_within <- 12 # Degree of freedom within-groups
10
11 r_sq <- round((var_between * df_between)/((var_
    between * df_between)+(var_within * df_within))
    ,2)
12
#
Proportion
of
variance
accounted
for

13
14 cat("Proportion of variance accounted for is", r_sq)
15
16 # Another approach
17
18 f <- 4.07 # F value (Ratio)
19
20 r_sq <- round((f * df_between)/((f * df_between) + (
    df_within)),2)
21
# Proportion
of
variance
accounted
for

```

```
22
23 cat("Proportion of variance accounted for is", r_sq)
```

---

**R code Exa 9.4b** How are you doing

```
1                                     # Page no. :
                                     348
2
3 # How are you doing?
4
5 # 2nd Question (d) part
6
7 var_within <- 7.20 # Variance within-groups
8 var_between <- 12.22 # Variance within-groups
9
10 df_between <- 2 # Degree of freedom between-groups
11 df_within <- 8 # Degree of freedom within-groups
12
13 r_sq <- round((var_between * df_between)/((var_
    between * df_between)+(var_within * df_within))
    ,2)
14                                     #
                                     Proportion
                                     of
                                     variance
                                     accounted
                                     for
15
16 cat("Proportion of variance accounted for is", r_sq)
17
18 # 3rd Question (c) part
19
20 N <- 18 # 18 participants in each group
21 t_g <- 3 # Total groups
```

```

22 f <- 4.50 # F value (Ratio)
23
24 df_between <- t_g - 1 # Degree of freedom between-
    groups
25 df_within <- (t_g * N) - t_g # Degree of freedom
    within-groups
26
27 r_sq <- round((f * df_between)/((f * df_between) + (
    df_within)),2)
28
    # Proportion
    of
    variance
    accounted
    for
29
30 cat("Proportion of variance accounted for is", r_sq)

```

---

**R code Exa 9.5a** Advanced Topic The Structural Model in the Analysis of Variance an example

```

1
    # Page no. : 351
    - 354
2
3 # Advanced Topic: The Structural Model in the
    Analysis of Variance an example
4
5 # Criminal Record Group
6
7 rating1 <- c(10, 7, 5, 10, 8)
8
9 # Clean Record Group
10
11 rating2 <- c(5, 1, 3, 7, 4)
12
13 # No Information Group

```

```

14
15 rating3 <- c(4, 6, 9, 3, 3)
16
17 DF <- data.frame(rating1, rating2, rating3)
18
19 # Direct Method
20
21 k <- ncol(DF) # Number of Treatments
22 n <- nrow(DF) # Number of Observations for each
    Treatment
23 N <- n * k # Total Observations
24
25 x <- c(t(as.matrix(DF)))
26 f <- c("r_1", "r_2", "r_3")
27 tm <- gl(k, 1, n*k, factor(f))
28 result <- anova(lm(x ~ tm)) # Similar to aov(x ~
    tm)
29
30 result
31
32 ss_total <- round(sum(result$'Sum Sq'), 2) # Total
    sum of squared deviation
33
34 cat("Total sum of squared deviation is", ss_total)
35
36 ss_within <- round(result$'Sum Sq'[2], 2) # Sum of
    squared deviation within-groups
37
38 cat("Sum of squared deviation within-groups is", ss_
    within)
39
40 ss_between <- round(result$'Sum Sq'[1], 2) # Sum
    of squared deviation between-groups
41
42 cat("Sum of squared deviation between-groups is", ss
    _between)
43
44 df_total <- sum(result$Df) # Total degrees of

```

```

    freedom
45
46 cat("Total degrees of freedom is", df_total)
47
48 df_within <- result$Df[2]    # Degree of freedom
    within-groups
49
50 cat("Degree of freedom within-groups is", df_within)
51
52 df_between <- result$Df[1]   # Degree of freedom
    between-groups
53
54 cat("Degree of freedom between-groups is", df_
    between)
55
56 var_within <- round(result$'Mean Sq'[2], 2)    #
    Variance within-groups
57
58 cat("Variance within-groups is", var_within)
59
60 var_between <- round(result$'Mean Sq'[1], 2)    #
    Variance between-groups
61
62 cat("Variance between-groups is", var_between)
63
64 f <- round(result$'F value'[1], 2)    # F value (
    Ratio)
65
66 cat("F value (Ratio) is", f)

```

---

**R code Exa 9.6a** Worked out examples 1

```

1                                     # Page no. : 359
2                                     - 360

```

```

3 # Worked out examples 1
4
5 N <- 20 # 20 participants in each group
6 n_g <- 4 # Numbr of groups
7
8 m1 <- 15 # Mean 1
9 var1 <- 20 # Variance 1
10
11 m2 <- 12 # Mean 2
12 var2 <- 25 # Variance 2
13
14 m3 <- 18 # Mean 3
15 var3 <- 14 # Variance 3
16
17 m4 <- 15 # Mean 4
18 var4 <- 27 # Variance 4
19
20 alpha <- 0.05 # 5% significant level
21
22 df_between <- n_g - 1 # Degree of freedom between-
    groups
23 df_within <- (4 * N) - (n_g) # Degree of freedom
    within-groups
24
25 var_within <- round((var1 + var2 + var3 + var4) / n_
    g, 2) # Variance within-groups
26 cat("Variance within-groups is", var_within)
27
28 gm <- (m1 + m2 + m3 + m4) / n_g # Grand mean
29 gm_dev <- c((m1 - gm), (m2 - gm), (m3 - gm), (m4 -
    gm)) # Deviation from the grand mean
30 sq_gm_dev <- round(gm_dev ** 2, 2) # Squared
    deviation from the grand mean
31
32 d_var <- round(sum(sq_gm_dev) / df_between, 2) #
    Variance for the distribution of means
33
34 cat("Variance for the distribution of means is", d_

```



```

    var)
35
36 var_between <- d_var * N # Variance between-groups
37
38 cat("Variance between-groups is", var_between)
39
40 fvalue <- round(var_between / var_within, 2) # F
    ratio
41 cat("F ratio is", fvalue)
42
43 fval <- qf(0.95, df_between, df_within) # 1 - alpha
    = 1 - 0.05 = 0.95
44 fval <- round(fval, 2)
45
46 if(fvalue >= fval)
47 {
48   cat("Reject Null Hypothesis")
49 } else {
50   cat("Cannot reject Null Hypothesis")
51 }

```

---

### R code Exa 9.6b Worked out examples 2

```

1 # Page no. :
    360 - 361
2
3 # Worked out examples 2
4
5 m1 <- 12 # Mean 1
6 var1 <- 25 # Variance 1
7
8 m2 <- 18 # Mean 2
9 var2 <- 14 # Variance 2
10
11 alpha <- 0.01 # 1% significant level

```

```

12 n_s <- 2 # Number of samples for contrast
13 N <- 20 # 20 elements in each sample
14
15 var_within <- 21.5 # Variance within-groups (
    overall)
16
17 df_between <- n_s - 1 # Degree of freedom between
    -groups (2 samples)
18 df_within <- (4 * N) - (n_s + 2) # Degree of
    freedom within-groups (4 samples)
19
20 gm <- (m1 + m2) / n_s # Grand mean
21 gm_dev <- c((m1 - gm), (m2 - gm)) # Deviation from
    the grand mean
22 sq_gm_dev <- round(gm_dev ** 2, 2) # Squared
    deviation from the grand mean
23
24 d_var <- round(sum(sq_gm_dev) / df_between, 2) #
    Variance for the distribution of means
25
26 cat("Variance for the distribution of means is", d_
    var)
27
28 var_between <- d_var * N # Variance between-groups
29
30 cat("Variance between-groups is", var_between)
31
32 fvalue <- round(var_between / var_within, 2) # F
    ratio
33 cat("F ratio is", fvalue)
34
35 fval <- qf(0.95, df_between, df_within) # 1 - alpha
    = 1 - 0.05 = 0.95
36 fval <- round(fval, 2)
37
38 cat("Cutoff F for the 0.05 significance level is",
    fval)
39

```

```

40 if(fvalue >= fval)
41 {
42   cat("Reject null hypothesis (Significant)")
43 } else {
44   cat("Cannot reject null hypothesis (not
      significant)")
45 }

```

---

### R code Exa 9.6c Worked out examples 3

```

1
2
3 # Worked out examples 3
4
5 n <- 6 # Six planned contrasts
6 alpha <- 0.05 # 5% significance level
7
8 b_c_s <- round(alpha / n, 4) # Bonferroni
   corrected significance level
9
10 cat("Bonferroni corrected significance level is", b_c
     _s)
11
12
13
14 # The Bonferroni Procedure
15
16 n <- 2 # Two planned contrasts
17 alpha <- 0.05 # 5% significance level
18
19 b_c_s <- round(alpha / n, 4) # Bonferroni
   corrected significance level
20

```

```

21 cat(" Bonferroni corrected significance level is", b_c
    _s)
22
23 n <- 3 # Two planned contrasts
24 alpha <- 0.05 # 5% significance level
25
26 b_c_s <- round(alpha / n, 4) # Bonferroni
    corrected significance level
27
28 cat(" Bonferroni corrected significance level is", b_c
    _s)
29
30 n <- 3 # Two planned contrasts
31 alpha <- 0.01 # 1% significance level
32
33 b_c_s <- round(alpha / n, 4) # Bonferroni corrected
    significance level
34
35 cat(" Bonferroni corrected significance level is", b_c
    _s)

```

---

#### R code Exa 9.6d Worked out examples 4

```

1 # Page no. :
    361
2
3 # Worked out examples 4
4
5 N <- 10 # 10 elements in each sample
6 f <- 11.21 # F value (overall)
7 n_g <- 5 # Number of groups
8
9 df_between <- n_g - 1 # Degree of freedom between-
    groups
10 df_within <- (5 * N) - (n_g) # Degree of freedom

```

```

    within-groups (5 samples)
11 fvalue <- f / df_between # F value (contrast)
12
13 cat("F value (contrast) is", fvalue)
14
15 alpha <- 0.01 # 1% Sgnificance level
16
17 fval <- qf(0.99, df_between, df_within) # 1 - alpha
    = 1 - 0.01 = 0.99
18 fval <- round(fval, 2)
19
20 cat("Cutoff F for the 0.01 significance level is",
    fval)
21
22 if(fvalue >= fval)
23 {
24   cat("Reject null hypothesis (Significant)")
25 } else {
26   cat("Cannot reject null hypothesis (not
    significant)")
27 }

```

---

### R code Exa 9.6e Worked out examples 5

```

1
    # Page no. :
    361
2
3 # Worked out examples 5
4
5 N <- 20 # 20 participants in each group
6 n_g <- 4 # Number of groups
7
8 df_between <- n_g - 1 # Degree of freedom between-
    groups
9 df_within <- (4 * N) - (n_g) # Degree of freedom

```

```

    within-groups
10
11 var_between <- 120 # Variance between-groups
12 var_within <- 21.5 # Variance within-groups
13
14 r_sq <- round((var_between * df_between)/((var_
    between * df_between)+(var_within * df_within))
    ,2)
15 #
# Proportion
# of
# variance
# accounted
# for
16
17 cat("Proportion of variance accounted for is", r_sq)

```

---

### R code Exa 9.6f Worked out examples 6

```

1 # Page no. :
2 362
3 # Worked out examples 6
4
5 # Treatment A
6
7 x1 <- c(8, 13, 10, 9)
8
9 # Treatment B

```

```

10
11 x2 <- c(7, 3, 8)
12
13 # Treatment C
14
15 x3 <- c(6, 4, 2)
16
17 DF <- data.frame( Y=c(x1, x2, x3), Treatment =factor
18                   (rep(c("x1", "x2", "x3"),
19                       times=c(length(x1), length(x2), length(
20                               x3))))))
21
22 View(DF)
23
24 res <- aov(Y~Treatment, data=DF)
25 result <- anova(res)
26 result
27
28 df_total <- sum(result$Df) # Total degree of
29 freedom
30
31 cat("Total degrees of freedom is", df_total)
32
33 df_within <- result$Df[2] # Degree of freedom
34 within-groups
35
36 cat("Degree of freedom within-groups is", df_within)
37
38 df_between <- result$Df[1] # Degree of freedom
39 between-groups
40
41 cat("Degree of freedom between-groups is", df_
42     between)
43
44 alpha <- 0.05 # 5% significance level
45
46 f <- round(result$`F value`[1], 2) # F value (
47     Ratio)

```

```

41
42 cat("F value (Ratio) is", f)
43
44 ss_total <- round(sum(result$'Sum Sq'), 2) # Total
      sum of squared deviation
45
46 cat("Total sum of squared deviation is", ss_total)
47
48 ss_within <- round(result$'Sum Sq'[2], 2) # Sum of
      squared deviation within-groups
49
50 cat("Sum of squared deviation within-groups is", ss_
      within)
51
52 ss_between <- round(result$'Sum Sq'[1], 2) # Sum
      of squared deviation between-groups
53
54 cat("Sum of squared deviation between-groups is", ss
      _between)
55
56 fval <- qf(0.95, df_between, df_within) # 1 -
      alpha = 1 - 0.05 = 0.95
57 fval <- round(fval, 2)
58
59 cat("Cutoff F for the 0.05 significance level is",
      fval)
60
61 if(f >= fval)
62 {
63   cat("Reject null hypothesis")
64 } else {
65   cat("Cannot reject null hypothesis")
66 }

```

---



# Chapter 10

## Factorial Analysis of Variance

R code Exa 10.1a Table for a Two way Analysis of Variance

```
1                                     # Page no. : 406
2                                     - 407
3 # Table for a Two-way Analysis of Variance
4
5 DF <- data.frame(Sensitivity = c(rep("Not High", 10)
6   , rep("High", 10)),
7   Test_Difficulty = c(rep(c("Easy", "
8     Hard"), 10)),
9   Value = c(2.63, 2.69, 2.53, 2.31,
10    2.25, 2.45, 2.22, 2.80, 2.52,
11    2.55, 2.06,
12    3.21, 2.32, 3.21, 2.04,
13    2.77, 2.31, 2.83,
14    2.22, 3.03)
15 )
16
17 View(DF)      # Table no. 10-10   (Table is different
18   as compared to book)
19
20 res <- aov(Value ~ Test_Difficulty*Sensitivity, data
```

```

      = DF)
14 result <- round(anova(res), 2)
15
16 # Table no. 10-11
17
18 result # Values differ with book's values (table
      values)
19
20 # Below results are in table no. 10-10
21
22 ss_total <- round(sum(result$'Sum Sq'), 2) # Total
      sum of squared deviation
23
24 cat("Total sum of squared deviation is \n", ss_total
      )
25
26 ss_within <- round(result$'Sum Sq'[4], 2) # Sum of
      squared deviation within-groups
27
28 cat("Sum of squared deviation within-groups is \n",
      ss_within)
29
30 ss_interaction <- round(result$'Sum Sq'[3], 2) #
      Sum of squared deviation of interaction
31
32 cat("Sum of squared deviation of interaction is \n",
      ss_interaction)
33
34 ss_rows <- round(result$'Sum Sq'[2], 2) # Sum of
      squared deviation of rows
35
36 cat("Sum of squared deviation of rows is \n", ss_
      rows)
37
38 ss_columns <- round(result$'Sum Sq'[1], 2) # Sum
      of squared deviation of columns
39
40 cat("Sum of squared deviation of columns is \n", ss_

```

```

        columns)
41
42 alpha <- 0.05 # 5% significance level
43 df_numerator <- 1 # 2 - 1 = 1 (2 groups)
44 df_denomenator <- result$Df[4]
45
46 fval <- qf(0.95, df_numerator, df_denomenator)
47 fval <- round(fval, 2)
48
49 # Note that fval is same for sensitivity , test
    difficulty and for interaction effect
50
51 cat("F value is \n", fval)
52
53 f1 <- result$'F value'[1] # F value for test
    difficulty
54 f2 <- result$'F value'[2] # F value for
    sensitivity
55 f3 <- result$'F value'[3] # F value for test
    difficulty : sensitivity
56
57 if(f1 >= fval)
58 {
59   cat("Reject null hypothesis for test difficulty")
60 } else {
61   cat("Cannot reject null hypothesis for test
        difficulty")
62 }
63
64 if(f2 >= fval)
65 {
66   cat("Reject null hypothesis for sensitivity")
67 } else {
68   cat("Cannot reject null hypothesis for sensitivity
        ")
69 }
70
71 if(f3 >= fval)

```

```

72 {
73   cat("Reject null hypothesis for test difficulty :
       sensitivity")
74 } else {
75   cat("Cannot reject null hypothesis for test
       difficulty : sensitivity")
76 }

```

---

### R code Exa 10.1b How are you doing

```

1                                     # Page no. :
                                     411 - 413
2
3 # How are you doing?
4
5 # 5th Question
6
7 DF <- data.frame(Participant = LETTERS[1:8],
8                  Length_condition = c(rep("Short",
9                  4), rep("Long", 4)),
9                  Vividness_condition = c(rep(c("Low"
10                 , "High"), 4)),
10                 Number_recalled = c(6, 9, 4, 5, 2,
11                 1, 4, 1)
11                 )
12
13 View(DF)   # Book's table is little bit different
14
15 res <- aov(Number_recalled ~ Vividness_condition*
16            Length_condition, data = DF)
17 result <- anova(res)
18 result
19 ss_total <- round(sum(result$'Sum Sq'), 2)   # Total
       sum of squared deviation

```

```

20
21 cat("Total sum of squared deviation is", ss_total)
22
23 ss_within <- round(result$'Sum Sq'[4], 2) # Sum of
      squared deviation within-groups
24
25 cat("Sum of squared deviation within-groups is", ss_
      within)
26
27 ss_interaction <- round(result$'Sum Sq'[3], 2) #
      Sum of squared deviation of interaction
28
29 cat("Sum of squared deviation of interaction is", ss
      _interaction)
30
31 ss_rows <- round(result$'Sum Sq'[2], 2) # Sum of
      squared deviation of rows
32
33 cat("Sum of squared deviation of rows is", ss_rows)
34
35 ss_columns <- round(result$'Sum Sq'[1], 2) # Sum
      of squared deviation of columns
36
37 cat("Sum of squared deviation of columns is", ss_
      columns)
38
39 alpha <- 0.05 # 5% significance level
40 df_numerator <- 1 # 2 - 1 = 1 (2 groups)
41 df_denominator <- result$Df[4]
42
43 fval <- qf(0.95, df_numerator, df_denominator)
44 fval <- round(fval, 2)
45
46 f1 <- result$'F value'[1] # F value for vividness
      condition
47 f2 <- result$'F value'[2] # F value for length
      condition
48 f3 <- result$'F value'[3] # F value for vividness

```

```

        condition : length condition
49
50 if(f1 >= fval)
51 {
52   cat("Reject null hypothesis")
53 } else {
54   cat("Cannot reject null hypothesis")
55 }
56
57 if(f2 >= fval)
58 {
59   cat("Reject null hypothesis")
60 } else {
61   cat("Cannot reject null hypothesis")
62 }
63
64 if(f3 >= fval)
65 {
66   cat("Reject null hypothesis")
67 } else {
68   cat("Cannot reject null hypothesis")
69 }

```

---

**R code Exa 10.2a** Advanced Topic Effect Size in the Factorial Analysis of Variance

```

1                                     # Page no. : 413 -
                                     414
2
3 # Advanced Topic : Effect Size in the Factorial
  Analysis of Variance
4
5 var_col <- 1.20 # Variance (columns)
6 df_col <- 1   # Degree of freedom (columns)
7

```

```

8 var_within <- 0.03 # Variance within-groups
9 df_within <- 16 # Degree of freedom within-groups
10
11 r_sq1 <- round((var_col * df_col) / ((var_col * df_
    col) + (var_within * df_within)), 2)
12 # The proportion of
    Variance announced
    for (test
    difficulty)
13
14 cat("The proportion of Variance announced for (test
    difficulty) is", r_sq1)
15
16 var_rows <- 0.00 # Variance (rows)
17 df_rows <- 1 # Degree of freedom (rows)
18
19 r_sq2 <- round((var_rows * df_rows) / ((var_rows *
    df_rows) + (var_within * df_within)), 2)
20 # The proportion of Variance announced for (
    sensitivity)
21
22 cat("The proportion of Variance announced for (
    sensitivity) is", r_sq2)
23
24 var_int <- 0.60 # Variance (interaction)
25 df_int <- 1 # Degree of freedom (interaction)
26
27 var_within <- 0.03 # Variance within-groups
28 df_within <- 16 # Degree of freedom within-groups
29
30 r_sq3 <- round((var_int * df_int) / ((var_int * df_
    int) + (var_within * df_within)), 2)
31 # The proportion of Variance announced for (test
    difficulty)
32
33 cat("The proportion of Variance announced for (test
    difficulty) is", r_sq3)

```

---

**R code Exa 10.2b** How are you doing

```
1                                     # Page no. :
                                     416 - 417
2
3 # How are you doing?
4
5 # 1st Question (c) part
6
7 var_rows <- 9.45 # Variance (rows)
8 df_rows <- 1    # Degree of freedom (rows)
9
10 var_within <- 3.67 # Variance within-groups
11 df_within <- 36   # Degree of freedom within-groups
12
13 r_sq2 <- round((var_rows * df_rows) / ((var_rows *
    df_rows) + (var_within * df_within)), 2)
14 # The proportion of Variance announced for
15
16 cat("The proportion of Variance announced for is", r
    _sq2)
```

---

**R code Exa 10.3a** worked out examples 1

```
1                                     # Page no. :
                                     419 - 421
2
3 # worked out examples 1
4
5 # Only (b) and (c) part
6
```



```

7 DF <- data.frame(Variable_A = c(rep("Level1", 8),
  rep("Level2", 8), rep("Level3", 8)),
8     Variable_B = c(rep(c("Level1", "
  Level2"), 12)),
9     Values = c(25, 19, 20, 24, 23, 21,
  24, 20, 22, 24, 19, 18, 22, 22,
10     21, 20,
  16, 18, 19, 21, 13, 16,
  16, 17)
11 )
12
13 View(DF) # Book's table is little bit different
14
15 # install.packages(ggplot2)
16 library(ggplot2)
17
18 ggplot(DF, aes(x = Variable_A, y = Values, fill =
  Variable_B)) + geom_bar(position = "dodge",
19
20 # Litte bit different barplot as compare with book's
  barplot
21
22 res <- aov(Values ~ Variable_B*Variable_A, data = DF
  )
23 result <- anova(res)
24 result
25
26 ss_total <- round(sum(result$'Sum Sq'), 2) # Total
  sum of squared deviation
27

```

```

28 cat("Total sum of squared deviation is", ss_total)
29
30 ss_within <- round(result$'Sum Sq'[4], 2) # Sum of
      squared deviation within-groups
31
32 cat("Sum of squared deviation within-groups is", ss_
      within)
33
34 ss_interaction <- round(result$'Sum Sq'[3], 2) #
      Sum of squared deviation of interaction
35
36 cat("Sum of squared deviation of interaction is", ss_
      _interaction)
37
38 ss_rows <- round(result$'Sum Sq'[2], 2) # Sum of
      squared deviation of rows
39
40 cat("Sum of squared deviation of rows is", ss_rows)
41
42 ss_columns <- round(result$'Sum Sq'[1], 2) # Sum
      of squared deviation of columns
43
44 cat("Sum of squared deviation of columns is", ss_
      columns)
45
46 alpha <- 0.05 # 5% significance level
47 df_numerator <- 2 # 3 - 1 = 1 (2 groups)
48 df_denominator <- result$Df[4]
49
50 fval <- qf(0.95, df_numerator, df_denominator)
51 fval <- round(fval, 2)
52
53 f1 <- result$'F value'[1] # F value for variable B
54 f2 <- result$'F value'[2] # F value for variable A
55 f3 <- result$'F value'[3] # F value for variable B
      : variable A
56
57 if(f1 >= fval)

```

```
58 {
59   cat("Reject null hypothesis")
60 } else {
61   cat("Cannot reject null hypothesis")
62 }
63
64 if(f2 >= fval)
65 {
66   cat("Reject null hypothesis")
67 } else {
68   cat("Cannot reject null hypothesis")
69 }
70
71 if(f3 >= fval)
72 {
73   cat("Reject null hypothesis")
74 } else {
75   cat("Cannot reject null hypothesis")
76 }
```

---

# Chapter 11

## Correlation

**R code Exa 11.1a** Graphing Correlations The Scatter Diagram An example

```
1                                     # Page no. : 441 -
                                     442
2
3 # Graphing Correlations : The Scatter Diagram An
  example
4
5 Hours_slept <- c(5, 7, 8, 6, 6, 10)
6 Happy_mood <- c(2, 4, 7, 2, 3, 6)
7
8 DF <- data.frame(Hours_slept, Happy_mood)
9 View(DF)
10
11 # Install Library if not install
12
13 # install.packages("ggplot2")
14
15 # Import Library
16
17 library(ggplot2)
18
```

```

19 ggplot(DF, aes(x = Hours_slept, y = Happy_mood)) +
    geom_point() + labs(title =
20     "Scatter plot", x = "Hours slept last night"
    , y = "Happy mood") + theme_bw() +
21   scale_x_continuous(limits=c(1, 12), breaks = c
    (1:12)) +
22   scale_y_continuous(limits=c(1, 8), breaks = c(1:8)
    )

```

---

### R code Exa 11.1b How are you doing

```

1
# Page no. :
# 443
2
3 # How are you doing?
4
5 # 3rd Question
6
7 person <- LETTERS[1:4]
8 x <- c(3, 6, 1, 4)
9 y <- c(4, 7, 2, 6)
10
11 DF <- data.frame(person, x, y)
12 View(DF)
13
14 # Install Library if not install
15
16 # install.packages("ggplot2")
17
18 # Import Library
19
20 library(ggplot2)
21
22 ggplot(DF, aes(x = x, y = y)) + geom_point() + labs(
    title = "Scatter plot", x = "X", y = "Y") +

```

```

23   theme_bw() + scale_x_continuous(limits=c(0, 6),
      breaks = c(0:6)) +
24   scale_y_continuous(limits=c(0, 7), breaks = c(0:7)
      )

```

---

### R code Exa 11.2a Figuring the correlation coefficient an example

```

1                                     # Page no. : 455
                                     - 456
2
3 # Figuring the correlation coefficient an example
4
5 Hours_slept <- c(5, 7, 8, 6, 6, 10)
6 Happy_mood <- c(2, 4, 7, 2, 3, 6)
7
8 # Book Method
9
10 m1 <- mean(Hours_slept) # Mean
11 m2 <- mean(Happy_mood) # Mean
12
13 deviation1 <- Hours_slept - m1
14 sq_dev1 <- deviation1 ** 2 # Squared deviation
15 sd1 <- round(sqrt(sum(sq_dev1) / 6), 2) # Standard
      deviation
16
17 z_score1 <- round(deviation1 / sd1, 2)
18
19 deviation2 <- Happy_mood - m2
20 sq_dev2 <- deviation2 ** 2 # Squared deviation
21 sd2 <- round(sqrt(sum(sq_dev2) / 6), 2) # Standard
      deviation
22
23 z_score2 <- round(deviation2 / sd2, 2)
24
25 c_p <- round(z_score1 * z_score2, 2) # Cross

```

```

    Product
26
27 DF <- data.frame(Hours_slept, deviation1, sq_dev1, z
    _score1, Happy_mood, deviation2, sq_dev2,
28                 z_score2, c_p)
29 View(DF)
30
31 r <- round(sum(DF$c_p) / nrow(DF), 2)  #
    Correlation Coefficient
32
33 cat(" Correlation coefficient is", r)
34
35 # Direct method
36
37 r2 <- round(cor(Happy_mood,Hours_slept), 2)
38 cat(" Correlation coefficient is", r2)

```

---

**R code Exa 11.2b** Figuring the correlation coefficient a second example

```

1                                     # Page no. : 456
                                     - 457
2
3 # Figuring the correlation coefficient a second
    example
4
5 No_of_exposures <- c(1:8)
6 No_of_words_recalled <- c(3, 2, 6, 4, 5, 5, 6, 9)
7
8 # Direct method
9
10 r <- round(cor(No_of_exposures, No_of_words_recalled)
    , 2)  # Correlation Coefficient
11 cat(" Correlation coefficient is", r)

```

---

**R code Exa 11.2c** How are you doing

```
1                                     # Page no. : 457
                                     - 458
2
3 # How are you doing?
4
5 # 4th Question
6
7 person <- c("K", "L", "M")
8 z_score1 <- c(0.5, -1.4, 0.9)
9 z_score2 <- c(-0.7, -0.8, 1.5)
10
11 c_p <- round(z_score1 * z_score2, 2) # Cross
    product
12
13 r <- round(sum(c_p) / 3, 2) # Correlation
    coefficient
14
15 cat("Correlation coefficient is", r)
```

---

**R code Exa 11.3a** Significance of a correlation coefficient

```
1                                     # Page no. : 458 -
                                     459
2
3 # Significance of a correlation coefficient
4
5 N <- 3 # No. of perons
6 r <- 0.24 # Correlation coefficient
7
```



```

8 t <- round(r / sqrt((1 - (r ** 2)) / (N - 2)), 2)
  # t value
9
10 cat("t value is", t)

```

---

**R code Exa 11.3b** Significance of a correlation coefficient an example

```

1                                     # Page no. : 459
2
3 # Significance of a correlation coefficient an
  example
4
5 N <- 6   # Sample size
6 r <- 0.85 # Correlation coefficient
7 alpha <- 0.05 # 5% significance level
8
9 df <- N - 2 # Degree of freedom
10
11 # Two - tailed
12
13 tval <- qt(1-alpha/2, df)
14 tval <- c(tval, -tval)
15
16 t <- round(r / sqrt((1 - (r ** 2)) / (N - 2)), 2)
  # t value
17
18 cat("t value is", t)
19
20 if(t > tval[1] || t < tval[2])
21 {
22   cat("Reject null hypothesis")
23 } else
24 {
25   cat("Cannot reject null hypothesis")
26 }

```

---

**R code Exa 11.3c** How are you doing

```
1                                     # Page no. : 461
2                                     - 462
3 # How are you doing?
4
5 # 3rd Question
6
7 N <- 60    # Sample size
8 r <- -0.31 # Correlation coefficient
9 alpha <- 0.05 # 5% significance level
10
11 df <- N - 2 # Degree of freedom
12
13 # Two - tailed
14
15 tval <- qt(1-alpha/2, df)
16 tval <- c(tval, -tval)
17
18 t <- round(r / sqrt((1 - (r ** 2)) / (N - 2)), 2)
19     # t value
20
21 cat("t value is", t)
22
23 if(t > tval[1] || t < tval[2])
24 {
25   cat("Reject null hypothesis")
26 } else
27 {
28   cat("Cannot reject null hypothesis")
29 }
```

---

### R code Exa 11.4a Worked out examples 1

```
1                                     # Page no. : 477
                                     - 478
2
3 # Worked out examples 1
4
5 Elementry_School <- c("Main Street", "Casat", "
   Lakeland", "Shady Grove", "Jefferson")
6 Class_size <- c(25, 14, 33, 28, 20)
7 Achievement <- c(80, 98, 50, 82, 90)
8
9 DF <- data.frame(Elementry_School, Class_size,
   Achievement)
10 View(DF)
11
12 # Install Library if not install
13
14 # install.packages("ggplot2")
15
16 # Import Library
17
18 library(ggplot2)
19
20 ggplot(DF, aes(x = Class_size, y = Achievement)) +
   geom_point() +
21   labs(title = "Scatter plot", x = "Class size", y =
   "Achievement test score") +
22   theme_bw()
```

---

### R code Exa 11.4b Worked out examples 2

```

1                                     # Page no. : 478 -
                                     479
2
3 # Worked out examples 2
4
5 Elementry_School <- c("Main Street", "Casat", "
   Lakeland", "Shady Grove", "Jefferson")
6 Class_size <- c(25, 14, 33, 28, 20)
7 Achievement <- c(80, 98, 50, 82, 90)
8
9 # Direct method
10
11 r <- round(cor(Class_size, Achievement), 2) #
   Correlation Coefficient
12 cat("Correlation coefficient is", r)

```

---

#### R code Exa 11.4c Worked out examples 3

```

1                                     # Page no. :
                                     479 - 480
2
3 # Worked out examples 3
4
5 N <- 5 # Sample size
6 r <- -0.90 # Correlation coefficient
7 alpha <- 0.05 # 5% significance level
8
9 df <- N - 2 # Degree of freedom
10
11 # Two - tailed
12
13 tval <- qt(1-alpha/2, df)
14 tval <- c(tval, -tval)
15
16 t <- round(r / sqrt((1 - (r ** 2)) / (N - 2)), 2)

```

```
    # t value
17
18 cat("t value is", t)
19
20 if(t > tval[1] || t < tval[2])
21 {
22   cat("Reject null hypothesis")
23 } else
24 {
25   cat("Cannot reject null hypothesis")
26 }
```

---

# Chapter 12

## Prediction

R code Exa 12.1a The linear prediction rule an example

```
1                                     # Page no. : 494
2                                     #         - 496
3 # The linear prediction rule an example
4
5 a <- 0.3   # Regression constant
6 b <- 0.004 # Regression coefficient
7 x <- 700   # SAT score
8
9 y_cap <- a + (b * x) # Predicted GPA (Linear
   predictor)
10
11 cat("Predicted GPA (Linear predictor) is", y_cap)
12
13 # Another example
14
15 a <- -3   # Regression constant
16 b <- 1    # Regression coefficient
17 x <- 9    # Hours of sleep
18
19 y_cap <- a + (b * x) # Predicted mood (Linear
```

```
    predictor)
20
21 cat(" Predicted mood (Linear predictor) is", y_cap)
```

---

**R code Exa 12.1b** How are you doing

```
1                                     # Page no. :
                                     497
2
3 # How are you doing?
4
5 # 4th Question
6
7 a <- -1.23   # Regression constant
8 b <- 6.11   # Regression coefficient
9
10 # (a)
11
12 x <- 2.00   # Predictor variable
13
14 y_cap <- a + (b * x)   # Predicted score (Linear
    predictor)
15
16 cat(" Predicted score (Linear predictor) is", y_cap)
17
18 # (b)
19
20 x <- 4.87   # Predictor variable
21
22 y_cap <- a + (b * x)   # Predicted score (Linear
    predictor)
23
24 cat(" Predicted score (Linear predictor) is", round(y
    _cap, 2))
25
```

```

26 # (a)
27
28 x <- -1.92 # Predictor variable
29
30 y_cap <- a + (b * x) # Predicted score (Linear
    predictor)
31
32 cat("Predicted score (Linear predictor) is", round(y
    _cap, 2))

```

---

**R code Exa 12.2a** Another example of drawing the regression line

```

1 # Page no. : 501
2
3 # Another example of drawing the regression line
4
5 # Data on page no. : 442
6
7 Hours_slept <- c(5, 7, 8, 6, 6, 10)
8 Happy_mood <- c(2, 4, 7, 2, 3, 6)
9
10 DF <- data.frame(Hours_slept, Happy_mood)
11 View(DF)
12
13 # Import Library
14
15 library(ggplot2)
16
17 ggplot(DF, aes(x = Hours_slept, y = Happy_mood)) +
    geom_point() +
18   labs(title = "Scatter plot", x = "Hours slept
    last night", y = "Happy mood") + theme_bw() +
19   scale_x_continuous(limits=c(1, 12), breaks = c
    (1:12)) +
20   scale_y_continuous(limits=c(1, 8), breaks = c(1:8))

```



```
    ) +  
21 geom_smooth(method='lm', se = F)
```

---

**R code Exa 12.3a** Finding the best linear prediction rule

```
1 # Page no. : 502 - 504  
2  
3 # Finding the best linear prediction rule  
4  
5 # Rule 1 ==> y_cap <- 8 - (.18)*X  
6 # Rule 2 ==> y_cap <- 4 + (0)*X  
7 # Rule 3 ==> y_cap <- -2.5 + (1)*X  
8 # Rule 4 ==> y_cap <- -3 + (1)*X  
9  
10 Hours_slept <- c(5, 6, 6, 7, 8, 10)  
11 Actual_mood <- c(2, 2, 3, 4, 7, 6)  
12  
13 # Rule 1  
14  
15 y_cap1 <- 8 - (.18) * Hours_slept  
16  
17 # Rule 2  
18  
19 y_cap2 <- 4 + (0) * Hours_slept  
20  
21 # Rule 3  
22  
23 y_cap3 <- -2.5 + (1) * Hours_slept  
24  
25 # Rule 4  
26  
27 y_cap4 <- -3 + (1) * Hours_slept  
28  
29 DF <- data.frame(Hours_slept, Actual_mood, y_cap1, y  
    _cap2, y_cap3, y_cap4)
```

---

**R code Exa 12.3b** The least squared error principle

```
1 # Page no. : 504 - 505
2
3 # The least squared error principle
4
5 # Rule 1 ==> y_cap <- 8 - (.18)*X
6 # Rule 2 ==> y_cap <- 4 + (0)*X
7 # Rule 3 ==> y_cap <- -2.5 + (1)*X
8 # Rule 4 ==> y_cap <- -3 + (1)*X
9
10 Hours_slept <- c(5, 6, 6, 7, 8, 10)
11 Actual_mood <- c(2, 2, 3, 4, 7, 6)
12
13 # Rule 1
14
15 y_cap1 <- 8 - (.18) * Hours_slept
16
17 error1 <- Actual_mood - y_cap1
18
19 error1_sq <- round(error1 ** 2, 2)
20
21 # Rule 2
22
23 y_cap2 <- 4 + (0) * Hours_slept
24
25 error2 <- Actual_mood - y_cap2
26
27 error2_sq <- error2 ** 2
28
29 # Rule 3
30
31 y_cap3 <- -2.5 + (1) * Hours_slept
```

```

32
33 error3 <- Actual_mood - y_cap3
34
35 error3_sq <- error3 ** 2
36
37 # Rule 4
38
39 y_cap4 <- -3 + (1) * Hours_slept
40
41 error4 <- Actual_mood - y_cap4
42
43 error4_sq <- error4 ** 2
44
45 DF <- data.frame(Hours_slept, Actual_mood, y_cap1,
46                 error1, error1_sq, y_cap4, error4, error4_sq)
47 View(DF)
48
49 s1 <- sum(DF$error1_sq)
50
51 cat("Rule 1 sum of squared errors is", s1)
52
53 s2 <- sum(error2_sq)
54
55 cat("Rule 2 sum of squared errors is", s2)
56
57 s3 <- sum(error3_sq)
58
59 cat("Rule 3 sum of squared errors is", s3)
60
61 s4 <- sum(DF$error4_sq)
62
63 cat("Rule 4 sum of squared errors is", s4)

```

---

**R code Exa 12.3c** Finding a and b for the least squares linear prediction rule

```

1                                     # Page no. :
                                     505 - 507
2
3 # Finding a and b for the least squares linear
  prediction rule
4
5 Hours_slept <- c(5, 7, 8, 6, 6, 10)
6 Happy_mood <- c(2, 4, 7, 2, 3, 6)
7
8 DF <- data.frame(Hours_slept, Happy_mood)
9 View(DF)
10
11 # Direct method
12
13 regressor <- lm(Happy_mood ~ Hours_slept, data = DF)
14 res <- summary(regressor)
15
16 res
17
18 b <- res$coefficients[[2]]
19
20 a <- res$coefficients[[1]]
21
22 cat("Linear Prediction Rule is y_cap =", a, "+", b, "x")

```

---

**R code Exa 12.3d** How are you doing

```

1 # Page no. : 507 - 508
2
3 # How are you doing?
4
5 # 4th Question
6
7 x <- c(4, 6, 7, 3)
8 y <- c(6, 8, 3, 7)

```

```

9
10 DF <- data.frame(x, y)
11 View(DF)
12
13 # (a) part
14
15 # Direct method
16
17 regressor <- lm(y ~ x, data = DF)
18 res <- summary(regressor)
19
20 res
21
22 b <- res$coefficients[[2]]
23
24 a <- res$coefficients[[1]]
25
26 cat("Linear Prediction Rule is y_cap =", a, "+", b, "x")
27
28 # (b) part
29
30 y_cap1 <- a + b * x
31
32 error1 <- y - y_cap1
33
34 error1_sq <- round(error1 ** 2, 2)
35
36 s1 <- sum(error1_sq)
37
38 cat("Sum of squared errors is", s1)
39
40 # (c) part
41
42 y_cap2 <- 9 - (0.7) * x
43
44 error2 <- y - y_cap2
45
46 error2_sq <- round(error2 ** 2, 2)

```

```
47
48 s2 <- sum(error2_sq)
49
50 cat("Sum of squared errors is", s2)
```

---

**R code Exa 12.4a** The standardized regression coefficient

```
1 # Page no. : 509 - 510
2
3 # The standardized regression coefficient
4
5 ss_x <- 16
6 ss_y <- 22
7 b <- 1
8
9 beta <- round(b * (sqrt(ss_x) / sqrt(ss_y)), 2)
10
11 cat("Standardized regression coefficient is", beta)
```

---

**R code Exa 12.4b** How are you doing

```
1 # Page no. : 511 - 512
2
3 # How are you doing?
4
5 # 2nd Question (b) part
6
7 ss_x <- 2.57
8 ss_y <- 7.21
9 b <- -1.21
10
11 beta <- round(b * (sqrt(ss_x) / sqrt(ss_y)), 2)
12
```

```
13 cat(" Standardized regression coefficient is", beta)
```

---

### R code Exa 12.5a Multiple regression

```
1                                     # Page no. : 512
2                                     - 513
3 # Multiple regression
4
5 a <- -3.78
6
7 b1 <- 0.87
8 x1 <- 7
9
10 b2 <- 0.33
11 x2 <- 3
12
13 b3 <- 0.20
14 x3 <- 1
15
16 y_cap <- round(a + (b1 * x1) + (b2 * x2) + (b3 * x3)
17           , 2) # Predicted mood (multiple regression)
18 cat(" Predicted mood (multiple regression) is", y_cap
19     )
```

---

### R code Exa 12.5b How are you doing

```
1                                     # Page no. :
2                                     514 - 515
3 # How are you doing?
4
```

```

5 # 3rd Question
6
7 a <- 2.19
8
9 b1 <- -3.16
10
11 b2 <- 0.99
12
13 # (a) part
14
15 x1 <- 0.40
16 x2 <- 10.50
17
18 y_cap1 <- round(a + (b1 * x1) + (b2 * x2), 2) #
    Predicted score (multiple regression)
19
20 cat("Predicted score (multiple regression) is", y_
    cap1)
21
22 # (b) part
23
24 x1 <- 0.15
25 x2 <- 5.50
26
27 y_cap2 <- round(a + (b1 * x1) + (b2 * x2), 2) #
    Predicted score (multiple regression)
28
29 cat("Predicted score (multiple regression) is", y_
    cap2)

```

---

**R code Exa 12.6a** Proportionate reduction in error

```

1
2
# Page no. :
520 - 522

```



```

3 # Proportionate reduction in error
4
5 Hours_slept <- c(5, 6, 6, 7, 8, 10)
6 Actual_mood <- c(2, 2, 3, 4, 7, 6)
7
8 DF <- data.frame(Hours_slept, Actual_mood)
9 View(DF)
10
11 # Direct method
12
13 regressor <- lm(Actual_mood ~ Hours_slept, data = DF
14 )
15
16 res <- summary(regressor)
17
18 p_r <- round(res$r.squared, 2) # Proportionate
19 reduction in error
20 cat("Proportionate reduction in error is", p_r)

```

---

**R code Exa 12.6b** How are you doing

```

1
2
3 # How are you doing?
4
5 # 4th Question
6
7 score <- c(6, 4, 2, 8)
8 predicted_score <- c(5.7, 4.3, 2.9, 7.1)
9
10 m <- mean(score) # mean
11

```

```

12 error <- score - m
13 error_sq <- error ** 2 # Error square
14
15 error2 <- score - predicted_score
16 error2_sq <- error2 ** 2 # Error square
17
18 DF <- data.frame(score, error, error_sq, predicted_
    score, error2, error2_sq)
19 View(DF)
20
21 ss_total <- sum(DF$error_sq)
22
23 ss_error <- sum(DF$error2_sq)
24
25 p_r <- round((ss_total - ss_error) / ss_total, 2) #
    Proportionate reduction in error
26
27 cat("Proportionate reduction in error is", p_r)
28
29 r <- round(sqrt(p_r), 2) # Correlation coefficient
30
31 cat("Correlation coefficient is", r)

```

---

### R code Exa 12.7a Worked out examples 1

```

1 # Page no. :
    526
2
3 # Worked out examples 1
4
5 # This example includes solution of worked out
    examples 3
6
7 Elementry_School <- c("Main Street", "Casat", "
    Lakeland", "Shady Grove", "Jefferson")

```

```

8 Class_size <- c(25, 14, 33, 28, 20)
9 Achievement <- c(80, 98, 50, 82, 90)
10
11 DF <- data.frame(Elementary_School, Class_size,
12   Achievement)
13 View(DF)
14 # Direct method
15
16 regressor <- lm(Achievement ~ Class_size, data = DF)
17 res <- summary(regressor)
18
19 res
20
21 b <- round(res$coefficients[[2]], 2)
22
23 a <- round(res$coefficients[[1]], 2)
24
25 cat("Linear Prediction Rule is y_cap =", a, "+", b, "x")
26
27 # (a) part
28
29 x <- 23 # Class size
30
31 y_cap <- a + (b * x) # Predicted achievement
32
33 cat("Predicted achievement is", y_cap)
34
35 # (b) part
36
37 x <- 14 # Class size
38
39 y_cap <- a + (b * x) # Predicted achievement
40
41 cat("Predicted achievement is", y_cap)

```

---

## R code Exa 12.7b Worked out examples 2

```
1                                     # Page no. :
2                                     526
3 # Worked out examples 2
4
5 Elementry_School <- c("Main Street", "Casat", "
   Lakeland", "Shady Grove", "Jefferson")
6 Class_size <- c(25, 14, 33, 28, 20)
7 Achievement <- c(80, 98, 50, 82, 90)
8
9 DF <- data.frame(Elementry_School, Class_size,
   Achievement)
10 View(DF)
11
12 # Install Library if not install
13
14 # install.packages("ggplot2")
15
16 # Import Library
17
18 library(ggplot2)
19
20 ggplot(DF, aes(x = Class_size, y = Achievement)) +
   geom_point() +
21   labs(title = "Scatter plot", x = "Class size", y =
   "Achievement test score") +
22   theme_bw() + geom_smooth(method = "lm", se = F)
23
24 # Book's figure is in page 527
```

---

**R code Exa 12.7c** Worked out examples 4

```
1                                     # Page no. :
2                                     528
3 # Worked out examples 4
4
5 ss_x <- 214
6 ss_y <- 1328
7 b <- -2.25
8
9 beta <- round(b * (sqrt(ss_x) / sqrt(ss_y)), 2) #
   Standardized regression coefficient
10
11 cat("Standardized regression coefficient is", beta)
```

---

**R code Exa 12.7d** Worked out examples 5

```
1                                     # Page no. :
2                                     528 -
3                                     529
4
5 # Worked out examples 5
6
7 a <- 2.13
8
9
10 b1 <- 1.32
11 x1 <- 4
12
13 b2 <- 1.21
14 x2 <- 5
15
16 b3 <- 1.41
17 x3 <- 3
18
```

```

16 y_cap <- round(a + (b1 * x1) + (b2 * x2) + (b3 * x3)
    , 2) # Predicted talkativeness
17
18 cat("Predicted talkativeness is", y_cap)

```

---

### R code Exa 12.7e Worked out examples 6

```

1                                     # Page no. : 529
                                     - 530
2
3 # Worked out examples 6
4
5 Elementry_School <- c("Main Street", "Casat", "
    Lakeland", "Shady Grove", "Jefferson")
6 Class_size <- c(25, 14, 33, 28, 20)
7 Achievement <- c(80, 98, 50, 82, 90)
8
9 DF <- data.frame(Elementry_School, Class_size,
    Achievement)
10 View(DF)
11
12 # Direct method
13
14 regressor <- lm(Achievement ~ Class_size, data = DF)
15 res <- summary(regressor)
16
17 res
18
19 p_r <- round(res$r.squared, 2) # Proportionate
    reduction in error
20
21 cat("Proportionate reduction in error is", p_r)

```

---

# Chapter 13

## Chi Square Tests

**R code Exa 13.1a** An example

```
1 # Page no. : 543 -544
2
3 # An example
4
5 gender <- c("Male", "Female")
6 o <- c(996, 390) # Observed frequency
7 e <- c(693, 693) # Expected frequency
8
9 diff <- o - e
10 diff_sq <- diff ** 2
11
12 ans <- round(diff_sq / e, 2) # Difference squared
   weighted by expected frequency
13
14 DF <- data.frame(gender, o, e, diff, diff_sq, ans)
15 View(DF)
```

---

**R code Exa 13.1b** The chi square statistic and the chi square test for goodness of fit

```

1 # Page no. : 545 -546
2
3 # The chi - square statistic and the chi - square
  test for goodness of fit
4
5 o <- c(996, 390) # Observed frequency
6 e <- c(693, 693) # Expected frequency
7
8 diff <- o - e
9 diff_sq <- diff ** 2
10 ans <- round(diff_sq / e, 2)
11
12 chi_sq <- sum(ans)
13
14 cat("Value of chi-square is", chi_sq)

```

---

#### R code Exa 13.1c Another example

```

1 # Page no. : 549 -550
2
3 # Another example
4
5 condition <- c("Anxiety disorder", "Alcohol or drug
  abuse", "Mood disorder",
6               "Impulse - control disorder", "None
  of these conditions")
7
8 observed <- c(138, 99, 123, 111, 529)
9
10 expected <- c(146, 80, 110, 128, 536)
11
12 DF <- data.frame(condition, observed, expected)
13 View(DF)
14
15 alpha <- 0.05

```



```

16 df <- nrow(DF) - 1
17
18 chi_sq_val <- round(qchisq(alpha,df,lower.tail = F),
    2)
19
20 diff <- DF$observed - DF$expected
21 diff_sq <- diff ** 2
22 ans <- round(diff_sq / DF$expected, 2)
23
24 chi_sq <- sum(ans)
25
26 cat("Value of chi-square is", chi_sq)
27
28 if(chi_sq > chi_sq_val)
29 {
30   cat("Reject null hypothesis")
31 } else
32 {
33   cat("Cannot reject null hypothesis")
34 }

```

---

**R code Exa 13.1d** How are you doing

```

1 # Page no. : 551 - 553
2
3 # How are you doing?
4
5 # 5th Question (a) part
6
7 alpha <- 0.05
8 categories <- 2
9
10 observed <- c(15, 35)
11 N <- sum(observed)
12 expected <- c(0.6 * N, 0.4 * N)

```

```

13
14 df <- categories - 1
15
16 chi_sq_val <- round(qchisq(alpha,df,lower.tail = F),
    2)
17
18 diff <- observed - expected
19 diff_sq <- diff ** 2
20 ans <- round(diff_sq / expected, 2)
21
22 chi_sq <- sum(ans)
23
24 cat("Value of chi-square is", chi_sq)
25
26 if(chi_sq > chi_sq_val)
27 {
28   cat("Reject null hypothesis")
29 } else
30 {
31   cat("Cannot reject null hypothesis")
32 }

```

---

**R code Exa 13.2a** The chi square test for independence

```

1 # Page no. : 553 - 557
2
3 # The chi - square test for independence
4
5 # Contingency Table
6
7 gender <- c(rep("Male", 28), rep("Female", 30), rep(
    "Male", 125), rep("Female", 39))
8 age <- c(rep("Child", 58), rep("Adult", 164))
9
10 c_t <- table(age, gender)

```

```

11 c_t
12
13 # Direct method
14
15 result <- chisq.test(c_t)
16 result
17
18 # Expected values
19
20 expected <- round(result$expected, 2)
21 expected
22
23 # Figuring chi - square
24
25 chi_sq <- round(result$statistic, 2)
26
27 cat("Chi-square value is", chi_sq)
28 # Degrees of freedom
29
30 df <- result$parameter
31
32 cat("Degrees of freedom is", df)
33
34 # Hypothesis testing
35
36 alpha <- 0.05
37
38 chi_sq_val <- round(qchisq(alpha,df,lower.tail = F),
39                       2)
40 if(chi_sq > chi_sq_val)
41 {
42   cat("Reject null hypothesis")
43 } else
44 {
45   cat("Cannot reject null hypothesis")
46 }

```

---

### R code Exa 13.2b Another example

```
1 # Page no. : 558 - 559
2
3 # Another example
4
5 # Contingency Table
6
7 generation <- c(rep("First", 73), rep("Other", 89),
8               rep("First", 657), rep("Other", 1226))
9 drop <- c(rep("Dropped out", 162), rep("Did not drop
10         out", 1883))
11
12 c_t <- table(drop, generation)
13
14 # Direct method
15
16 result <- chisq.test(c_t)
17 result
18
19 # Expected values
20
21 expected <- round(result$expected, 2)
22 expected
23
24 # Figuring chi - square
25
26 chi_sq <- round(result$statistic, 2)
27
28 cat("Chi-square value is", chi_sq)
29
30 # Degrees of freedom
```

```

31
32 df <- result$parameter
33
34 cat("Degrees of freedom is", df)
35
36 # Hypothesis testing
37
38 alpha <- 0.01
39
40 chi_sq_val <- round(qchisq(alpha,df,lower.tail = F),
41                     2)
42 if(chi_sq > chi_sq_val)
43 {
44   cat("Reject null hypothesis")
45 } else
46 {
47   cat("Cannot reject null hypothesis")
48 }
49 # Book's answer differ with our answer

```

---

**R code Exa 13.2c** How are you doing

```

1 # Page no. : 560 – 561
2
3 # How are you doing?
4
5 # 4th Question (a) part
6
7 nominal_variable_A <- c(rep("Category1", 10), rep("
8   Category2", 10), rep("Category1", 50),
9   rep("Category2", 10), rep("
10  Category1", 10), rep("
11  Category2", 10))
12
13 nominal_variable_B <- c(rep("Category1", 20), rep("

```

```

        Category2", 60), rep("Category3", 20))
10
11 c_t <- table(nominal_variable_B, nominal_variable_A)
12
13 c_t
14
15 # Direct method
16
17 result <- chisq.test(c_t)
18 result
19
20 # Expected values
21
22 expected <- round(result$expected, 2)
23 expected
24
25 # Figuring chi - square
26
27 chi_sq <- round(result$statistic, 2)
28
29 cat("Chi-square value is", chi_sq)
30
31 # Degrees of freedom
32
33 df <- result$parameter
34
35 cat("Degrees of freedom is", df)
36
37 # Hypothesis testing
38
39 alpha <- 0.10
40
41 chi_sq_val <- round(qchisq(alpha, df, lower.tail = F),
42                     2)
43 if(chi_sq > chi_sq_val)
44 {
45   cat("Reject null hypothesis")

```

```
46 } else
47 {
48   cat("Cannot reject null hypothesis")
49 }
```

---

### R code Exa 13.3a Effect Size for Chi Square Tests for Independence

```
1 # Page no. : 562 - 563
2
3 # Effect Size for Chi-Square Tests for Independence
4
5 chi_sq <- 6.73
6 N <- 2045
7
8 phi_coeff <- round(sqrt(chi_sq / N), 2)
9
10 cat("Phi coefficient is", phi_coeff)
11
12
13 chi_sq <- 12.70
14 N <- 100
15 df_smaller <- 1
16
17 c_phi_coeff <- round(sqrt(chi_sq / (N * df_smaller))
18   , 2)
19 cat("Cramer's phi coefficient is", c_phi_coeff) #
    Book's answer is wrong
```

---

### R code Exa 13.3b How are you doing

```
1 # Page no. : 565 - 566
2
```

```

3 # How are you doing?
4
5 # 2nd Question (d) part
6
7 chi_sq <- 12
8 N <- 100
9
10 phi_coeff <- round(sqrt(chi_sq / N), 2)
11
12 cat("Phi coefficient is", phi_coeff)
13
14 # 3rd Question (d) part
15
16 chi_sq <- 20
17 N <- 200
18 df_smaller <- 4 - 1
19
20 c_phi_coeff <- round(sqrt(chi_sq / (N * df_smaller))
21   , 2)
22 cat("Cramer's phi coefficient is", c_phi_coeff)

```

---

#### R code Exa 13.4a Worked out examples 1

```

1 # Page no. : 569 - 570
2
3 # Worked out examples 1
4
5 grade <- LETTERS[1:5]
6 observed <- c(10, 34, 140, 10, 6)
7 expected <- c(5, 28, 134, 28, 5)
8
9 DF <- data.frame(grade, observed, expected)
10 View(DF)
11

```



```

12 alpha <- 0.01
13 df <- nrow(DF) - 1
14
15 chi_sq_val <- round(qchisq(alpha,df,lower.tail = F),
2)
16
17 diff <- DF$observed - DF$expected
18 diff_sq <- diff ** 2
19 ans <- round(diff_sq / DF$expected, 2)
20
21 chi_sq <- sum(ans)
22
23 cat("Value of chi-square is", chi_sq)
24
25 if(chi_sq > chi_sq_val)
26 {
27   cat("Reject null hypothesis")
28 } else
29 {
30   cat("Cannot reject null hypothesis")
31 }

```

---

#### R code Exa 13.4b Worked out examples 2

```

1 # Page no. : 571 - 572
2
3 # Worked out examples 2
4
5 # Contingency Table
6
7 participant_gender <- c(rep("Men", 29), rep("Women",
17), rep("Men", 4), rep("Women", 14),
8 rep("Men", 26), rep("Women",
28))
9 comparision <- c(rep("Same sex", 46), rep("Opposite

```

```

    sex", 18), rep("Both sexes", 54))
10
11 c_t <- table(comparision, participant_gender)
12
13 c_t
14
15 # Direct method
16
17 result <- chisq.test(c_t)
18 result
19
20 # Expected values
21
22 expected <- round(result$expected, 2)
23 expected
24
25 # Figuring chi - square
26
27 chi_sq <- round(result$statistic, 2)
28
29 cat("Chi-square value is", chi_sq)
30
31 # Degrees of freedom
32
33 df <- result$parameter
34
35 cat("Degrees of freedom is", df)
36
37 # Hypothesis testing
38
39 alpha <- 0.05
40
41 chi_sq_val <- round(qchisq(alpha, df, lower.tail = F),
42                       2)
43 if(chi_sq > chi_sq_val)
44 {
45   cat("Reject null hypothesis")

```

```
46 } else
47 {
48   cat("Cannot reject null hypothesis")
49 }
```

---

**R code Exa 13.4c** Worked out examples 3 and 4

```
1                                     # Page no. :
                                     572
2
3 # Worked out examples 3 and 4
4
5 chi_sq <- 14.41 # Chi-square value
6 N <- 85 # No. of people
7
8 phi_coeff <- round(sqrt(chi_sq / N), 2) # Phi
   coefficient
9
10 cat("Phi coefficient is", phi_coeff)
11
12 chi_sq <- 8.78 # Chi-square value
13 N <- 118 # No. of people
14 df_smaller <- 2 - 1 # Degree of freedom on smaller
   side (2 x 2 contingency table)
15
16 c_phi_coeff <- round(sqrt(chi_sq / (N * df_smaller))
   , 2) # Cramer's phi coefficient
17
18 cat("Cramer's phi coefficient is", c_phi_coeff) #
   Book's answer is wrong
```

---

# Chapter 14

## Strategies When Population Distributions Are Not Normal

R code Exa 14.1a An example of a data transformation

```
1 # Page no. : 592 – 593
2
3 # An example of a data transformation
4
5 highly_sensitive_no <- c(0, 3, 10, 22)
6 highly_sensitive_yes <- c(17, 36, 45, 75)
7
8 m1 <- mean(highly_sensitive_no)
9 var1 <- round(var(highly_sensitive_no), 2)
10
11 m2 <- mean(highly_sensitive_yes)
12 var2 <- round(var(highly_sensitive_yes), 2)
13
14 sq_root1 <- round(sqrt(highly_sensitive_no), 2)
15 sq_root2 <- round(sqrt(highly_sensitive_yes), 2)
16
17 DF <- data.frame(highly_sensitive_no, sq_root1,
18                 highly_sensitive_yes, sq_root2)
18 View(DF)
```

```

19
20 alpha <- 0.05
21
22 df1 <- length(highly_sensitive_no) - 1
23 df2 <- length(highly_sensitive_yes) - 1
24
25 df_total <- df1 + df2
26
27 # One - tailed
28
29 tval <- round(qt(alpha, df_total, lower.tail = T),
30              2)
31 result <- t.test(DF$sq_root1, DF$sq_root2)
32 result
33
34 mean1 <- round(result$estimate[1], 2)
35 var_1 <- round(var(DF$sq_root1), 2)
36
37 mean2 <- round(result$estimate[2], 2)
38 var_2 <- round(var(DF$sq_root2), 2)
39
40 p_var <- round(((df1 / df_total) * var_1) + ((df2 /
41              df_total) * var_2), 2)
42
43 cat("Pooled variance is", p_var)
44
45 t <- round(result$statistic[[1]], 2) # t value
46
47 cat("t value is", t)
48
49 if(t < tval)
50 {
51   cat("Reject null hypothesis")
52 } else
53 {
54   cat("Cannot reject null hypothesis")
55 }

```

---

**R code Exa 14.2a** An example of a rank order test

```
1                                     # Page no. :
                                     595 – 596
2
3 # An example of a rank – order test
4
5 highly_sensitive_no <- c(0, 3, 10, 22)
6 highly_sensitive_yes <- c(17, 36, 45, 75)
7
8 DF <- data.frame( highly_sensitive = rep(c("No", "
   Yes"), each = 4),
9                   Values = c(highly_sensitive_no,
   highly_sensitive_yes)
10                  )
11
12 Rank <- rank(DF$Values)
13 DF <- cbind(DF, Rank)
14
15 View(DF)
16
17 alpha <- 0.05 # 5% significance level
18
19 result <- wilcox.test(highly_sensitive_no, highly_
   sensitive_yes, paired = T)
20
21 result
22
23 p_value <- result$p.value
24
25 if(p_value > alpha)
26 {
27   cat("Reject null hypothesis")
28 } else
```

```
29 {
30   cat("Cannot reject null hypothesis")
31 }
```

---

**R code Exa 14.2b** Using parametric tests with rank transformed data

```
1 # Page no. : 596 - 597
2
3 # Using parametric tests with rank - transformed
  data
4
5 highly_sensitive_no <- c(1, 2, 3, 5)
6 highly_sensitive_yes <- c(4, 6, 7, 8)
7
8 m1 <- mean(highly_sensitive_no)
9 var1 <- round(var(highly_sensitive_no), 2)
10
11 m2 <- mean(highly_sensitive_yes)
12 var2 <- round(var(highly_sensitive_yes), 2)
13
14 df1 <- length(highly_sensitive_no) - 1
15 df2 <- length(highly_sensitive_yes) - 1
16 df_total <- df1 + df2
17 # One - tailed
18
19 alpha <- 0.05 # 5% significance level
20
21 tval <- round(qt(alpha, df_total, lower.tail = T),
  2)
22
23 p_var <- round(((df1 / df_total) * var1) + ((df2 /
  df_total) * var2), 2) # Pooled variance
24
25 cat("Pooled variance is", p_var)
26
```

```

27 result <-t.test(highly_sensitive_no, highly_
    sensitive_yes)
28 result
29
30 t <- round(result$statistic[[1]], 2) # t value
31
32 cat("t value is", t)
33
34 if(t < tval)
35 {
36   cat("Reject null hypothesis")
37 } else
38 {
39   cat("Cannot reject null hypothesis")
40 }

```

---

**R code Exa 14.2c** How are you doing

```

1 # Page no. : 598 - 599
2
3 # How are you doing?
4
5 # 2nd Question
6
7 scores <- c(5, 18, 3, 9, 2)
8
9 rank <- rank(scores)
10
11 DF <- data.frame(scores, rank)
12
13 View(DF)

```

---

**R code Exa 14.3a** Worked out examples 1 and 2



```
1
# Page no. :
605
2
3 # Worked out examples 1 and 2
4
5 group_A <- c(15, 4, 12, 14)
6 group_B <- c(21, 16, 49, 17)
7 group_C <- c(18, 19, 11, 22)
8
9 sq_root1 <- round(sqrt(group_A), 2) # Square root
of group A
10 sq_root2 <- round(sqrt(group_B), 2) # Square root
of group B
11 sq_root3 <- round(sqrt(group_C), 2) # Square root
of group C
12
13 DF <- data.frame(sq_root1, sq_root2, sq_root3)
14 View(DF)
15
16 DF2 <- data.frame( Groups = rep(c("Group A", "Group
B", "Group C"), each = 4),
17 Values = c(group_A, group_B, group
_C)
18 )
19
20 Rank <- rank(DF2$Values)
21 DF2 <- cbind(DF2, Rank)
22
23 View(DF2)
```

---