

R Textbook Companion for  
Probability and Statistics for Engineers and  
Scientists  
by Ronald E. Walpole, Raymond H. Myers,  
Sharon L. Myers, Keying Ye<sup>1</sup>

Created by  
Dubey Varshit Shailendra  
B.Tech.  
Electronics Engineering  
College of Engineering Pune  
Cross-Checked by  
R TBC Team

May 26, 2020

<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:** Probability and Statistics for Engineers and Scientists

**Author:** Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, Keying Ye

**Publisher:** Pearson Education, Boston USA

**Edition:** 9

**Year:** 2016

**ISBN:** 978-0-321-62911-1

R numbering policy used in this document and the relation to the above book.

**Exa** Example (Solved example)

**Eqn** Equation (Particular equation of the above book)

For example, Exa 3.51 means solved example 3.51 of this book. Sec 2.3 means an R code whose theory is explained in Section 2.3 of the book.

# Contents

List of R Codes	4
1 Introduction To Statistics And Data Analysis	5
2 Probability	9
3 Random Variables And Probability Distributions	22
4 Mathematical Expectation	33
5 Some Discrete Probability Distribution	52
6 Some Continuous Probability Distributions	66
8 Fundamental Sampling Distributions And Data Descriptions	82
9 One And Two Sample Estimation Problems	90
10 One And Two Sample Tests Of Hypothesis	104
11 Simple Linear Regression And Correlation	116
12 Multiple Linear Regression And Certain Non Linear Regression Models	127
13 One Factor Experiment General	145
14 Factorial Experiments Two Or More Factors	153

15	2 to the power k Factorial Experiments and Fractions	161
16	Non Parametric Statistics	168
17	Statistical Quality Control	176
18	Bayesian Statistics	181

# List of R Codes

Exa 1.2	Dotplot . . . . .	5
Exa 1.3	Plot for Corrosion Data . . . . .	6
Exa 1.4	Mean Variance and Standard Deviation . . . . .	7
Exa 1.5	Box and Whisker Plot . . . . .	7
Exa 2.2	Sample Space . . . . .	9
Exa 2.3	Find Sample Space example 2 . . . . .	9
Exa 2.8	To find Intersection . . . . .	10
Exa 2.10	To find Union . . . . .	10
Exa 2.15	Generalized Multiplication Rule . . . . .	11
Exa 2.18	Permutation Problem . . . . .	11
Exa 2.19.a	Permutation Problem Example 19a . . . . .	11
Exa 2.19.b	Permutation Problem Example 19b . . . . .	12
Exa 2.19.c	Permutation Problem Example 19c . . . . .	13
Exa 2.19.d	Permutation Problem Example 19d . . . . .	14
Exa 2.21	Partition Problem . . . . .	14
Exa 2.22	Combination Problem . . . . .	15
Exa 2.24	Probability Calculation Using Formula And Law of Large Numbers . . . . .	16
Exa 2.25	Probability Calculation Using Law of Large Numbers .	17
Exa 2.30	Probability Calculation Using Law of Large Numbers And Union Property . . . . .	18
Exa 2.34	Conditional Probability Problem . . . . .	19
Exa 2.37	Independence and Multiplicative Rule . . . . .	20
Exa 2.43	Bayes Rule . . . . .	20
Exa 3.8	Discrete Probability Distribution Problem . . . . .	22
Exa 3.9	Discrete Probability Distribution Problem 9 . . . . .	23
Exa 3.10	Cumulative Distribution Function Problem . . . . .	23
Exa 3.11	Probability Density Function Problem 11 . . . . .	24

Exa 3.12	Cumulative Distribution Function of Continuous Random Variable . . . . .	25
Exa 3.14	Joint Probability Distribution of Discrete Random Variable . . . . .	26
Exa 3.15	Joint Density Function of Continuous Random Variable	27
Exa 3.16	Marginal Distribution Problem . . . . .	28
Exa 3.18	Conditional Distribution Of a Random Variable . . . . .	29
Exa 3.20	Marginal and Conditional Densities Evaluation . . . . .	30
Exa 3.22	Statistical Independent Problem . . . . .	31
Exa 4.1	Expected Value Evaluation Problem 1 . . . . .	33
Exa 4.2	Expected Value Evaluation Problem 2 . . . . .	34
Exa 4.3	Expected Value of probability density function . . . . .	34
Exa 4.4	Expected Value of a new Random Variable . . . . .	35
Exa 4.5	Expected Value of a new Random Variable . . . . .	36
Exa 4.6	Expected value of Discrete Joint Probability Distribution	36
Exa 4.7	Expected value of Joint Density Function . . . . .	37
Exa 4.8	Variance Of Discrete Distribution . . . . .	38
Exa 4.9	Variance using Expectation Formula . . . . .	39
Exa 4.10	Mean And Variance of Continuous Random Variable .	40
Exa 4.11	Variance of a new Random variable . . . . .	40
Exa 4.12	Variance of a new Random variable given the Density Function . . . . .	41
Exa 4.13	Covariance of 2 Random Variables . . . . .	42
Exa 4.14	Covariance Of Two Random Variables given Joint Density Function . . . . .	43
Exa 4.15	Correlation Coefficient Evaluation . . . . .	44
Exa 4.16	Correlation Coefficient Of 2 Random Variables Given Joint Density Function . . . . .	45
Exa 4.17	Expectation Value of Linear Combination Of Random Variables . . . . .	47
Exa 4.18	Expectation Value of Linear Combination Of Continuous Random Variables . . . . .	48
Exa 4.19	Expectation Value of Linear Combination Of Random Variables Using Properties . . . . .	48
Exa 4.20	Expectation Value Evaluation Using Properties Given a Density Function . . . . .	49
Exa 4.21	Expectation Value of 2 Independent Random Variable	50
Exa 5.1	Binomial Distribution Problem . . . . .	52

Exa 5.2.a	Binomial Distribution Problem . . . . .	52
Exa 5.2.b.c	Binomial Distribution Problem . . . . .	53
Exa 5.3	Binomial Distribution Problem 3 . . . . .	54
Exa 5.4	Binomial Distribution Problem 4 . . . . .	55
Exa 5.6	Binomial Distribution Problem 6 . . . . .	55
Exa 5.7	Multinomial Distribution Problem . . . . .	56
Exa 5.8	Hypergeometric Distribution In Acceptance Sampling	57
Exa 5.9	Hypergeometric Distribution In Acceptance Sampling	57
Exa 5.10	Hypergeometric Probability Distribution Problem . . .	58
Exa 5.11	Mean And Variance Of Hypergeometric Experiment .	58
Exa 5.12	Relation Between Hypergeometric and Binomial Distribution . . . . .	59
Exa 5.13	Multivariate Hypergeometric Distribution Problem . .	60
Exa 5.14	Negative Binomial Distribution Problem . . . . .	60
Exa 5.15	Geometric Distribution Problem . . . . .	61
Exa 5.16	Geometric Distribution Problem . . . . .	62
Exa 5.17	Poisson Distribution Problem . . . . .	62
Exa 5.18	Poisson Distribution Problem . . . . .	63
Exa 5.19	Approximation of Binomial By a Poisson Distribution	64
Exa 5.20	Approximation of Binomial By a Poisson Distribution	65
Exa 6.1	Uniform Distribution Problem . . . . .	66
Exa 6.2	Normal Distribution Problem . . . . .	67
Exa 6.3	Find The Value of k given Normal Distribution . . . .	67
Exa 6.4	Find area under the Normal Distribution curve . . . .	68
Exa 6.5	Find the Probability given Normal Distribution . . . .	68
Exa 6.6	Using Normal Curve In Reverse . . . . .	69
Exa 6.7	Applications of Normal Distribution . . . . .	69
Exa 6.8	Applications of Normal Distribution . . . . .	70
Exa 6.9	Applications of Normal Distribution . . . . .	70
Exa 6.11	Applications of Normal Distribution . . . . .	71
Exa 6.12	Applications of Normal Distribution . . . . .	72
Exa 6.13	Applications of Normal Distribution . . . . .	72
Exa 6.14	Applications of Normal Distribution . . . . .	73
Exa 6.15	Normal Approximation To the Binomial . . . . .	73
Exa 6.16	Normal Approximation To the Binomial . . . . .	74
Exa 6.17	Exponential Distribution Problem . . . . .	75
Exa 6.18	Gamma Distribution Problem . . . . .	76
Exa 6.19	Application of Gamma Distribution . . . . .	77



Exa 6.20	Application of Gamma Distribution . . . . .	77
Exa 6.21	Cumulative Distribution For Exponential Distribution Problem . . . . .	78
Exa 6.22	Lognormal Distribution Problem . . . . .	79
Exa 6.23	Find Percentile From Lognormal Distribution . . . . .	79
Exa 6.24	Weibull Distribution Problem . . . . .	80
Exa 8.1	Sample Mode Evaluation . . . . .	82
Exa 8.2	Sample mean and Sample Variance . . . . .	83
Exa 8.4	Central Limit Theorem . . . . .	83
Exa 8.5	Central limit theorem . . . . .	84
Exa 8.6	Sampling Distribution of the difference of 2 means . . . . .	84
Exa 8.7	The Chi squared Distribution Problem . . . . .	85
Exa 8.8	The t distribution Problem . . . . .	86
Exa 8.10	The t distribution Problem . . . . .	87
Exa 8.11	Applications of t distribution . . . . .	87
Exa 8.12	To Construct a Normal quantile quantile plot . . . . .	88
Exa 9.2	Confidence Interval Evaluation . . . . .	90
Exa 9.3	To Find the sample size . . . . .	91
Exa 9.4	To find 95 Percent bound for mean . . . . .	92
Exa 9.5	To Find 95 Percent Confidence Interval for mean . . . . .	92
Exa 9.6	To Find 99 Percent Confidence Interval . . . . .	93
Exa 9.7	To find 95 Percent Prediction Interval . . . . .	93
Exa 9.8	To find 99 Percent prediction interval . . . . .	94
Exa 9.10	Confidence Interval for the difference between 2 means . . . . .	95
Exa 9.11	Confidence Interval Evaluation Using Pooled Estimate of Variance . . . . .	96
Exa 9.12	Confidence Interval Evaluation For difference of 2 means when their variance is not equal and Unknown . . . . .	97
Exa 9.13	To find Confidence Intervals for Paired Observations . . . . .	98
Exa 9.14	Estimating a Proportion for Single Sample . . . . .	98
Exa 9.15	Evaluate sample size for estimating a Proportion for Single Sample . . . . .	99
Exa 9.16	Evaluate sample size for estimating a Proportion for Single Sample . . . . .	100
Exa 9.18	To obtain Confidence Interval of Variance . . . . .	101
Exa 9.19	To find Confidence Interval For the Ratio of 2 Variances . . . . .	101
Exa 9.22	Maximum Likelihood Estimation Problem . . . . .	102
Exa 10.3	Test Concerning a Single Mean for Single Sample . . . . .	104

Exa 10.4	Test Concerning a Single Mean for Single Sample . . .	105
Exa 10.5	The t statistic for a test on a single mean when Variance is Unknown . . . . .	106
Exa 10.6	The 2 Sample Pooled t test . . . . .	107
Exa 10.7	Choice of Sample Size for Testing Mean . . . . .	108
Exa 10.9	One Sample Test on a Single Proportion . . . . .	109
Exa 10.10	One Sample Test on a Single Proportion . . . . .	110
Exa 10.11	Two Sample Tests On 2 Proportions . . . . .	111
Exa 10.12	One and Two Sample Tests Concerning Variance . . .	112
Exa 10.13	To do Hypothesis Testing for 2 variances . . . . .	113
Exa 10.14	Test For Homogeneity . . . . .	114
Exa 10.15	Testing for several proportion . . . . .	115
Exa 11.1	To Estimate the Regression Line . . . . .	116
Exa 11.2	Confidence Interval for Slope of Regression Line . . .	117
Exa 11.3	Hypothesis Testing On Slope Of Regression Line . . .	118
Exa 11.4	Confidence Interval For Intercept Of Regression Line .	119
Exa 11.5	Hypothesis Testing on Intercept of Regression line . .	120
Exa 11.6	To find the confidence interval of the mean response .	121
Exa 11.7	To Find Prediction Interval . . . . .	122
Exa 11.8	Computation of Lack of Fit Sum of Squares . . . . .	123
Exa 11.9	Transformation To Linear Regression Model . . . . .	124
Exa 11.10	To find Correlation Coefficient . . . . .	125
Exa 11.11	To do Hypothesis Testing of linear association between two variables . . . . .	125
Exa 12.1	Estimation Of Equation Of Regression Line and predict values . . . . .	127
Exa 12.2	To find the estimate of polynomial regression . . . . .	128
Exa 12.3	Estimate the regression coefficients in polynomial model	129
Exa 12.4	Linear regression model using matrices . . . . .	130
Exa 12.5	To test the hypothesis on the slope of the regression model . . . . .	131
Exa 12.6	To Construct Confidence Interval for Mean Response .	134
Exa 12.7	Prediction interval Evaluation . . . . .	135
Exa 12.8	To show Analysis of Variance Table for Grain Radius Data . . . . .	136
Exa 12.9	To find the model for data which have 3 levels in a variable . . . . .	137
Exa 12.10	To predict all possible regression lines . . . . .	138

Exa 12.11	To model the data set using forward selection . . . . .	139
Exa 12.12	Cp Statistic . . . . .	141
Exa 12.13	Logistic regression model . . . . .	143
Exa 13.1	One way ANOVA . . . . .	145
Exa 13.2	One way ANOVA . . . . .	146
Exa 13.3	Bartlett Test . . . . .	146
Exa 13.4	Contrasts Sum Of Squares Corresponding to Orthogonal Contrasts . . . . .	147
Exa 13.5	Dunnett Test . . . . .	149
Exa 13.6	Randomized Complete Block Diagram . . . . .	150
Exa 13.7	Random Effects Model . . . . .	151
Exa 14.1	Two Factor Analysis of Variance . . . . .	153
Exa 14.2	Single Degree Of Freedom Sum Of Squares . . . . .	154
Exa 14.3	Evaluate Analysis of Variance table and derive conclu- sion also test on main effects . . . . .	156
Exa 14.4	Anova for 3 factor experiment . . . . .	157
Exa 14.5	Pooling in Multi factor Model . . . . .	158
Exa 14.6	Factorial Experiments for Random Effects . . . . .	159
Exa 15.1	2 squared Factorial Experiment . . . . .	161
Exa 15.2	Factorial Experiment in a Regression Setting . . . . .	162
Exa 15.4	Standard Errors of the least Squares Regression Coeffi- cients . . . . .	163
Exa 15.6	Use a half replicate to study the effects of five factors each at 2 levels on some response . . . . .	163
Exa 15.7	To construct a 2 level screening design with 6 variables containing 12 design points . . . . .	165
Exa 15.8	Response Surface Analysis . . . . .	165
Exa 16.1	Use sign test to test the hypothesis . . . . .	168
Exa 16.2	Use Sign test to test the hypothesis . . . . .	169
Exa 16.3	Use signed rank test to test the hypothesis . . . . .	170
Exa 16.4	Use signed rank test to test the hypothesis . . . . .	170
Exa 16.5	To perform Wilcoxin rank sum test . . . . .	171
Exa 16.6	Kruskal Wallis Test . . . . .	172
Exa 16.7	Runs Test . . . . .	173
Exa 16.8.9	Rank Correlation Coefficient . . . . .	174
Exa 17.1	Expected Value of samples required to detect the shift	176
Exa 17.2	Xbar and S chart . . . . .	176
Exa 17.3	UCL and LCL for preliminary control chart value . . .	178

Exa 17.4	Find Sample Size per subgroup . . . . .	179
Exa 17.5	Control Charts for Defects C chart . . . . .	179
Exa 18.1	To find the Posterior probability distribution . . . . .	181
Exa 18.4	Posterior Mean and Posterior Mode . . . . .	182
Exa 18.6	To find 95 percent Bayesian interval . . . . .	183
Exa 18.7	To find the 95 percent Bayesian interval for $\mu$ given the various parameter . . . . .	184
Exa 18.8	To find Bayes estimates of $p$ . . . . .	184
Exa 18.9	To find Bayes estimates of $p$ . . . . .	185
Exa 18.11	To find Bayes estimates of $p$ under Absolute Error loss	186

# Chapter 1

## Introduction To Statistics And Data Analysis

R code Exa 1.2 Dotplot

```
1 # Chapter 1
2 # Example 1.2 page no, 4 from the Pdf..
3 # To plot the Dotplot for the above data..
4 # package used "ggplot2" if not installed can be
   done using install.packages("ggplot2")
5
6 library(ggplot2)
7
8 obs <- c
   (0.32,0.53,0.28,0.37,0.47,0.43,0.36,0.42,0.38,0.43,0.26,0.43,0.47
9
10 cat <- c(rep("no_nit",10),rep("nit",10))
11
12 data1 <- data.frame(obs,cat) # making it a data
   frame.
13
14 data1$f <- factor(data1$obs) # adding another
   variable to the data frame.
```

```

15
16 # Plot..
17 ggplot(data1,aes(x = f, y = obs, fill = cat)) + geom
    _dotplot(binaxis = 'y',stackdir = 'center')

```

---

### R code Exa 1.3 Plot for Corrosion Data

```

1 # Chapter 1
2 # Example 1.3 page no.9 from the pdf..
3 #To find the plot for the Corrosion data in example
  1.3
4 # package used "ggplot2" if not installed can be
  done using install.packages("ggplot2")
5
6
7 library(ggplot2)    # library for making visuals
8
9 humidity <- c(20,80,20,80)
10
11 average_corrosion <- c(975,350,1750,1550)
12
13 coating <- c("Uncoated","Uncoated","Chemical
  corrosion","Chemical corrosion")
14
15 a <- data.frame(coating,humidity,average_corrosion)
    #making it a dataframe
16
17 # final PLOT..
18 ggplot(data= a, aes(x= humidity,y= average_corrosion
  ,colour= coating))+geom_line() + geom_point() +
  xlab("Humidity(in %)")+ ggtitle("Corrosion
  Results for Example 1.3")

```

---

### R code Exa 1.4 Mean Variance and Standard Deviation

```
1 # Chapter 1
2 # Example 1.4 page no. 15 from the pdf..
3 # To find the mean, sample variance, and standard
  deviation.
4
5
6 a <- c(7.00,
        7.07,7.10,6.97,7.00,7.03,7.01,7.01,6.98,7.08)
7
8 cat("The sample mean is ",mean(a))
9
10 cat("The sample variance is ",var(a))
11
12 cat("The sample standard deviation is ", sd(a))
```

---

### R code Exa 1.5 Box and Whisker Plot

```
1
2 # Example 1.5
3 # To make Boxplot of Nicotine Data..
4 # package used ggplot2, if not installed you can
  install it using install.packages("ggplot2")
5
6
7 library(ggplot2)
8
9 nicotine <- c(1.09,1.92,2.31,1.79,
               2.28,1.74,1.47,1.97,0.85,1.24,1.58,2.03,1.70,2.17,2.55,2.11,1.86,
10
11 content <- data.frame(nicotine) #making it dataframe
  for making plots
12
```

```
13 colnames(content) <- c("content1") #changing the
    column name of the dataframe
14
15 # Final Plot..
16 ggplot(data = content, aes(x=1, y= content1))+ geom_
    boxplot(fill = "white", colour = "#3366FF")+
    coord_flip()+ylab("Nicotine Content")+ ggtitle("
    Nicotine Content in Cigarettes")
```

---



# Chapter 2

## Probability

**R code Exa 2.2** Sample Space

```
1
2 # Chapter 2
3 # Example 2.2 page no. 36 from the pdf..
4 # To Find the Sample Space..
5
6 a <- c(rep("H",2),rep("T",6))
7
8 b <- c("H","T",seq(1,6))
9
10 paste(a,b) #concatening both vectores to make a
    sample space
```

---

**R code Exa 2.3** Find Sample Space example 2

```
1 # Chapter 2
2 # Example 2.3 page no. 37 from the pdf..
3 # to find the sample space.
4
```

```

5 a <- c(rep("D",2),rep("N",2))
6
7 b<- c("D", "N")
8
9 c <- c(rep("D",4),rep("N",4))
10
11 # final answer
12 print("The sample space is: ")
13 paste(c,rep(paste(a,rep(b)))) #merging in
    specific order to find the sample space.

```

---

**R code Exa 2.8** To find Intersection

```

1 # Chapter 2
2 # Example 2.8 page no. 40 from the Pdf..
3 # To find the Intersection of the two sets
4
5 a <- c("a","e","i","o","u")
6
7 b <- c("l","r","s","t")
8
9 intersect(a,b) #displays the common elements in a
    and b. If there is no element common then denotes
    a vector of 0 length

```

---

**R code Exa 2.10** To find Union

```

1
2 # Chapter 2
3 # Example 2.10 page no. 40 from the pdf..
4 # To find the Union of the given two sets
5
6 a <- c("a","b","c")

```

```

7
8 b <- c("b","c","d","e")
9
10 union(a,b) #Displays the union of two sets..

```

---

#### R code Exa 2.15 Generalized Multiplication Rule

```

1
2 # Chapter 2
3 # Example 2.15 page no. 46 from the pdf..
4 # Generalized Multiplication rule..
5
6 # answer..
7 cat("We can elect the chair and treasurer in ",22*
    21, " ways") #Using multiplication rule

```

---

#### R code Exa 2.18 Permutation Problem

```

1 # Chapter 2
2 # Example 2.18 page no. 48, from the pdf..
3 # Permutation Problem..
4
5
6 # Final answer..
7 cat("The total number of sample points is",factorial
    (25)/factorial(22))

```

---

#### R code Exa 2.19.a Permutation Problem Example 19a

```

1

```

```

2 # Chapter 2
3 # Example 2.19a page no.48 from the pdf..
4 # Permuation Problem
5 # number of ways to select president and treasurer
  from 50 students without restriction
6
7
8 cat("The number of choices of officers without
  restriction are",factorial(50)/factorial(48))
9
10
11 # It is simply permutation of 50 with 2

```

---

#### **R code Exa 2.19.b** Permutation Problem Example 19b

```

1 # Chapter 2
2 # example 2.19b page no. 48 from the pdf..
3 # Permuation Problem..
4
5
6 # to select the number of ways to elect president
  and treasurer from 50 people if A(a particular
  student) will serve ,if he is president
7 # 2 cases first A becomes president and in that case
8
9 cat("If A becomes president number of ways to select
  treasurer are ", 49)
10
11 a <- 49
12
13 # Second case if officers elected without A
14
15 cat("The number of possible ways to select without A
  are" ,factorial(49)/factorial(47))
16

```

```

17 b <- factorial(49)/factorial(47)
18
19 cat("The total number of ways to select the officers
      in this case are", a+b)

```

---

**R code Exa 2.19.c** Permutation Problem Example 19c

```

1
2 # Chapter 2
3 # Example 2.19c page no. 48 from the pdf..
4 # Permutation Problem..
5
6 # Number of ways to elect two officers from 50
   people if B and C (2 individuals from 50 ) will
   only work together
7 # 2 cases . First , if B and C are elected
8
9 cat("The number of ways to select B and C both from
      50 are",2) # if B treasurer and C president and
      vice versa..
10
11 a <- 2
12
13 # Second case: if B and C are not chosen so election
      done from rest 48 people.
14
15 cat("The number of selections when B and C are not
      chosen are",factorial(48)/factorial(46)) #48P2
16
17 b <- factorial(48)/factorial(46)
18
19 cat("The total number of choices in this case is",a+
      b)

```

---

### R code Exa 2.19.d Permutation Problem Example 19d

```
1
2 # Chapter 2.
3 # Example 2.19d page no. 48 from the pdf..
4 # Permutation Problem..
5
6 # Number of ways to select officers from 50 students
   if D and E(2 individuals from 50 students) will
   not work together
7 cat("The number of choices of officers without
   restriction are",factorial(50)/factorial(48))
8
9 a <- factorial(50)/factorial(48)
10
11 cat("Number of ways in which D and E are selected
   are",2) # if D is president and E is treasurer
   and vice versa
12
13 b <- 2
14
15 cat("Number of ways in which we can choose officers
   in this case are",a-b)
```

---

### R code Exa 2.21 Partition Problem

```
1 # Chapter 2
2 # Example 2.21 page no. 50 from the pdf..
3 # Partition Problem..
4
5
```

```

6 # number of ways, 7 students are assigned 1 triple
  and 2 double rooms..
7
8 # answer..
9 cat("The number of ways in which this partition can
  be done are", (factorial(7)/(factorial(3)*
  factorial(2)*factorial(2))))

```

---

### R code Exa 2.22 Combination Problem

```

1
2 # Chapter 2
3 # Example 2.22 page no. 50 from the pdf..
4 # Combination Problem..
5
6 # number of ways to select 3 arcade and 2 sports
  games from 10 arcade and 5 sports game
7 cat("The number of ways to select 3 arcade from 10
  is", choose(10,3))
8
9 a <- choose(10,3)
10
11 cat("The number of ways of selecting 2 catridges
  from 5 is", choose(5,2))
12
13 b <- choose(5,2)
14
15 #using multiplication rule..
16 cat("The total number of ways of selecting 3 arcade
  and 2 sports games is", a*b)
17
18 # example 2.23 same as the above 2 questions 2.22
  and 2.21..

```

---

**R code Exa 2.24** Probability Calculation Using Formula And Law of Large Numbers

```
1 # Chapter 2
2 # Example 2.24 page no.53 from the pdf..
3 # Probability Using Normal Method and Law of Large
  numbers..
4
5 # to find probability for atleast one head if a coin
  is tossed twice..
6
7 # "dplyr" package is used, if not installed can be
  installed using
8 # install.packages("dplyr") , remove comment if you
  already have..
9 library(dplyr)      # this package is for data
  manipulation
10
11 a <- c("H", "H", "T", "T")
12
13 b <- c("H", "T", "H", "T")
14
15 a <- data.frame(a,b)      # making data frame which
  consists of 2 columns i.e first column and 2nd
  column
16
17 c <- a %>% filter(a=="H" | b=="H")
18
19 cat("The probability that we get atleast one head in
  the tosses is",nrow(c)/nrow(a))
20
21 # another method to solve the same problem, by using
  law of large numbers
22
```



```

23
24 d <- c("HH", "TT", "TH", "HT")
25
26 e <- sample(d, 10000, replace = T)
27
28 cat("The approximate probability of getting atleast
    one head in the two tosses is", sum(e=="HH" | e=="HT"
    " | e=="TH") / 10000)
29
30 # the answer here is near to 0.75 but not exactly
    0.75, if we increase the number of trials say
    200000 it will be more close to 0.75.
31 # if we make this say tending to infinity then it
    will be exactly 0.75..

```

---

#### **R code Exa 2.25** Probability Calculation Using Law of Large Numbers

```

1 # Chapter 2
2 # Example 2.25 page no. 53 from the pdf..
3 # Probability Using Law of Large Numbers..
4
5 # Probability can also be calculated using direct
    formula(ratio) but the code will be very straight
    forward..
6 # Let's look at another way to achieve the same
    result.
7 # To find the probability of getting less than 4 in
    rolling a die, if even number has twice the
    probability than odd number..
8
9 a <- c(1, 2, 2, 3, 4, 4, 5, 6, 6)
10
11 b <- sample(a, 10000, replace=T)    # using law of
    large numbers
12

```

```

13 cat("The probability of getting a number less than 4
      is approximately",sum(b < 4)/10000)
14
15 #this answer is slightly different than the answer in
      textbook which is 0.444. If we increase the
      number of trials then the probability will be
      more close to the answer in textbook

```

---

**R code Exa 2.30** Probability Calculation Using Law of Large Numbers  
And Union Property

```

1 # Chapter 2
2 # Example 2.30 page no. 57 from the pdf..
3 # Probability Using Law of Large Numbers and using
      union property..
4
5
6 #to find the probability the sum of die roll (ROLLED
      TWICE) is 7 or 11
7 #Using law of large numbers..
8
9 a <- c(1:6)
10
11 b <- sample(a,200000,replace=T) + sample(a,200000,
      replace=T) # b is a sample of sum of 2 die
      rolls, with 10^6 times trials
12
13
14 cat("The approximate value of probability of getting
      the sum as either 7 or 11 is",sum(b==7|b==11)/
      200000)
15
16
17
18 # This is approximate as the textbook and standard

```

```
    answer is 2/9, if we increase the number of
    trials then we would get very close to the final
    answer..
19 # A word of note– I have skipped some problems as
    you may have seen , as many problems are of same
    type just the numbers are changed, the concept is
    same, so I solved the relevant questions.
```

---

### **R code Exa 2.34** Conditional Probability Problem

```
1
2 # Chapter 2.
3 # Example 2.34 page no.63 from the pdf..
4 # Conditional Probability Problem..
5
6
7 p_depends_ontime <- 0.83
8 P_arrives_ontime <- 0.82
9 p_depends_arrives_ontime <- 0.78
10
11 # to find probability that it arrives on time given
    it departs in time
12 # Using conditional probability ..
13
14 cat("The probability that the plane arrives on time
    given that it departs on time is",p_depends_
    arrives_ontime/p_depends_ontime)
15
16 # to find probability that it departs on time given
    that it arrives on time..
17
18 cat("The probability that the plane departs on time
    given that it arrives on time is",p_depends_
    arrives_ontime/P_arrives_ontime)
19
```

```
20
21
22 #The answer in the textbook is approximated to 0.94
    and 0.95 respectively ..
```

---

### R code Exa 2.37 Independence and Multiplicative Rule

```
1
2 # Chapter 2
3 # Example 2.37 page no. 66 from the pdf..
4 # Independence and Multiplication Rule..
5
6
7 # bag1 = 4 white and 3 black balls
8 #bag2= 3 white and 5 black balls
9 # to find probability of ball drawn from 2nd bag is
    black ? given a ball drawn from first is unseen
    and dropped in second bag
10
11 p_b1 <- 3/7
12 p_w1 <- 4/7
13
14 p_b2_b1 <- 6/9 # if first black ball dropped in
    second bag
15 p_b2_w1 <- 5/9 # if first white ball drawn is
    dropped into second bag..
16
17 # Using independence and multiplicative rule
18
19 cat("The probability that the second ball drawn is
    black in this case is",p_b1*p_b2_b1+p_w1*p_b2_w1)
```

---

### R code Exa 2.43 Bayes Rule

```

1 # Chapter 2.
2 # Example 2.43 from the pdf..
3 # Bayes Rule..
4
5 # given different parameters to find which plan was
  most likely used and thus responsible..
6 # to find  $P(P_j|D)$  for  $j = 1, 2, 3..$ 
7
8 # a vector of  $P(P_j)$  for  $j = 1, 2, 3..$ 
9 a <- c(0.30, 0.20, 0.50)
10
11 # a vector of  $P(D|P_j)$  for  $j = 1, 2, 3..$ 
12 b <- c(0.01, 0.03, 0.02)
13
14 # Bayes Rule Formula..
15 f <- function(x){
16
17   a[x]*b[x]/sum(a*b)
18
19 }
20
21 cat("The value of  $P(P_j|D)$  for  $j = 1, 2, 3$  is", f(1:3))
22
23 cat("We can see that for plan 3 has the highest
  conditional probability, hence a defective for a
  random product is most likely the result of the
  use of plan 3")
24
25
26 # Some Problems are not solved because they are one
  and the same of what I have solved just numbers
  are changed and they are easy too..

```

---

# Chapter 3

## Random Variables And Probability Distributions

**R code Exa 3.8** Discrete Probability Distribution Problem

```
1
2 # Chapter 3
3 # Example 3.8 page no. 84 from the pdf..
4 # Discrete Probability Problem..
5
6 # From 3.1 to 3.8 the problems are too easy and
   theoretical..
7 # to find the probability distribution of number of
   defective laptops if 2 are drawn at a time from
   20.
8
9 #making a general function for the above
   distribution..
10 a <- function(x){
11
12   choose(3,x)*choose(17,2-x)/choose(20,2)
13
14 }
15
```

```

16 # The Distribution..
17 cat("The probability distribution for x=0 is",a(0))
18
19 cat("The probability distribution value for x=1 is",
    a(1))
20
21 cat("The probability distribution for x=2 is",a(2))

```

---

### R code Exa 3.9 Discrete Probability Distribution Problem 9

```

1 # Chapter 3
2 # Example 3.9 page no. 85 from the pdf..
3 # Discrete Probability Distribution Problem..
4
5 # to find probability distribution of no. of cars
  with side airbags among next 4 cars , given prob.
  of selling is 0.5
6 # for ex. to sell 3 cars with air sidebags ,
  partition 4 into two – with 3 bags and with 1 bag
7 # done in 4 choose 3 ways
8
9 # generalise
10 cat("The probability distribution in this problem is
    :", " ")
11
12 for(i in 0:4){
13
14   cat("    ",choose(4,i)/16)
15
16   }

```

---

### R code Exa 3.10 Cumulative Distribution Function Problem

```

1
2 # Chapter 3
3 # Example 3.10 page no. 86 from the pdf..
4 # Cumulative Distribution function..
5
6 # to find cummulative distribution of the example
  3.9
7 # refer Example 3.9 page no. 85 from the pdf.
8 # the probability distribution of the example is
  choose(4,x)/16 where x is from 0 to 4
9 a <- function(x){
10
11   choose(4,x)/16
12 }
13
14
15 cat("The cummulative probability distribution of the
  above problem is "," ")
16
17 # ANswer..
18 for(i in 0:4){
19
20   cat("  ", sum(a(0:i)))
21 }

```

---

### **R code Exa 3.11** Probability Density Function Problem 11

```

1
2 # Chapter 3
3 # Example 3.11 Page no. 89 from the pdf..
4 # Probability Density FUnction Problem..
5
6 # A part of the Problem..
7 # to check the probability density function
8

```



```

9  a <- function(x){
10
11    x*x/3
12  }
13
14  # to check the conditions of probability density
    function
15
16
17  cat("the Integration is",integrate(a,-1,2)$value,"
    nearly 1 with very small margin of error we can
    say that it is a prob. density function")
18
19  # B Part of the Problem..
20  # to find P(0<x<=1)
21
22  cat("The value of P(0<x<=1) is",integrate(a,0,1)$
    value)
23
24  #this answer matches with the answer in the
    textbook.

```

---

**R code Exa 3.12** Cumulative Distribution Function of Continuous Random Variable

```

1
2  # Chapter 3
3  # Example 3.12 Page no. 90 from the pdf..
4  # Cumulative Distribution Function of Continuous
    Random Variable
5
6  # given a density function find the value of P(0<X
    <=1)
7
8  # the Density Function..

```

```

9 f <- function(t){
10
11   t*t/3
12
13 }
14
15 cat("The value of P(0<X<=1) is",integrate(f,0,1)$
    value)

```

---

**R code Exa 3.14** Joint Probability Distribution of Discrete Random Variable

```

1
2 # Chapter 3
3 # Example 3.14 Page no. 95 from the pdf..
4 # Joint Probability Distribution of Discrete Random
   Variable..
5
6 # Note – Example 3.13 is theoretical..
7
8 # the formula for joint Prob. distribution..
9 a <- function(x,y){
10
11   choose(3,x)*choose(2,y)*choose(3,2-x-y)/choose
     (8,2)
12
13 }
14
15 # to find probability that it will fall in x+y<=1
16
17
18 cat("The probability that x,y fall in region x+y<=1
    is",a(0,0)+a(0,1)+a(1,0))

```

---

**R code Exa 3.15** Joint Density Function of Continuous Random Variable

```
1
2 # Chapter 3
3 # Example 3.15 page no.96 from the pdf..
4 # Joint Density Function of Continuous Random
  Variable..
5
6 # A part of the Problem 3.15
7 # to verify the conditions of joint density function
  and find the probabilities of any intervals
8
9 # the joint density function..
10 a <- function(x,y){
11
12   (2*(2*x+3*y))/5
13
14
15 }
16
17 print("The value of the integral over whole region
  is:")
18
19 integrate(function(y){
20
21   sapply(y,function(y){
22
23     integrate(function(x)a(x,y),0,1)$value
24   })
25 },0,1)
26
27 # the integral is 1 then it is a joint density
  function
28
```

```

29 # Part B of the Problem 3.15
30 # to find the probability of a region A:  $\{(x,y) \mid 0 < x < 0.5, 0.25 < y < 0.5\}$ 
31
32 cat("The probability is ", " ")
33
34 integrate(function(y){
35
36     sapply(y,function(y){
37
38         integrate(function(x)a(x,y),0,0.5)$value
39
40     })
41
42 },0.25,0.5)

```

---

#### R code Exa 3.16 Marginal Distribution Problem

```

1
2 # Chapter 3
3 # Example 3.16 page no. 98 from the pdf..
4 # Marginal Distribution Problem
5
6 # to find marignal Distribution Of X alone and Y
  alone..
7
8 c <- as.data.frame(matrix(c(3/28,3/14,1/28,9/28,3/
  14,0,3/28,0,0),ncol = 3))
9
10 colnames(c) <- c("x0","x1","x2")
11
12 rownames(c) <- c("y1","y2","y3")
13
14 cat("The marginal distribution of X alone is",apply(
  as.matrix(c),2,sum),"and the marginal

```

```

        distributions of Y alone is",apply(as.matrix(c)
        ,1,sum))
15
16 # Example 3.17 is theoretical..

```

---

### R code Exa 3.18 Conditional Distribution Of a Random Variable

```

1
2 # Chapter 3
3 # Example 3.18 Page no.99 from the pdf..
4 # Conditional Distribution Of a Random Variable.
5
6 # to find the conditional joint probability
   distribution of the given function given the
   specific conditions
7 # Refer Example 3.14 from the pdf..
8 a <- function(x,y){
9
10    choose(3,x)*choose(2,y)*choose(3,2-x-y)/choose
      (8,2)
11
12 }
13
14 # given y=1 to find P(x=0|y=1)
15
16 #  $P(x=0|y=1) = a(x,y)/h(y)$ ,  $h(y) > 0$ 
17
18 h1 <- sum(a(0:2,1))
19
20 # to find  $a(x,1) = a(x,1)/h1$ 
21
22 b <- function(x){
23
24    a(x,1)/h1
25

```

```

26 }
27
28 cat("The conditional probability function under the
      condition y=1 is", " ")
29
30 c <- c(b(0),b(1),b(2))
31
32 d <- t(data.frame(0:2,c))
33
34 rownames(d) <- c("X", "f(x|1)")
35
36 d

```

---

### R code Exa 3.20 Marginal and Conditional Densities Evaluation

```

1 # chapter 3
2 # Example 3.19 and 3.20 both are same in concept,
   just the function is changed, I am solving 3.20,
   so you can do the same for 3.19..
3 # Chapter 3.20 page no 100 from the pdf..
4 # Mariginal and Conditional Densities Evaluation..
5
6 # given the joint distribution function
7 a <- function(x,y){
8
9   x*(1+3*y*y)/4
10
11 }
12
13 # with limits 0<x<2, 0<y<1
14
15 # given marginal density
16 b <- function(x){
17
18   x/2

```

```

19
20 }
21
22 c <- function(y){
23
24   (1+3*y*y)/2
25
26 }
27
28 # to find a(x|y)
29 # to find P(0.25<x<0.5|y=1/3)
30
31 d <- function(x){
32
33   a(x,1/3)/c(1/3)
34
35 }
36
37 cat("The value of conditional Probability P(0.25<X
    <0.5|Y = 1/3) is",integrate(d,0.25,0.5)$value)

```

---

### R code Exa 3.22 Statistical Independent Problem

```

1
2 # Chapter 3
3 # Example 3.22 page 103 from the pdf..
4 # Statistical Independent Problem..
5
6 # Note— Example 3.20 is theoretical, we need to just
   prove the variable is not statistical
   Independent, code involves just basic
   multiplication..
7
8 # package installed pracma just for simplicity for
   evaluating integrals numerically..

```

```

9 # given a density function and x1 ,x2, x3 represent
   3 of these containers selected independently..
10 # to find P(x1<2,1<x2<3,x3>2)
11 # if the package is already installed remove the
   installed.packages("pracma") from your code..
12
13 install.packages("pracma")
14 library("pracma")
15
16 f <- function(x,y,z){
17
18     exp(-x-y-z)
19
20 }
21
22
23 # Answer..
24 cat("The value of P(x1<2,1<x2<3,x3>2) is",integral3(
   f,0,2,1,3,2,22))

```

---



## Chapter 4

# Mathematical Expectation

**R code Exa 4.1** Expected Value Evaluation Problem 1

```
1
2 # Chapter 4
3 # Example 4.1 Page no. 113 from the pdf..
4 # Expected Value Evaluation Problem
5
6 # to find the expectation value
7 # out of 7, 4 good and 3 defective components, and
   sample of 3 is taken, find the expectation value
8
9 a <- function(x){
10
11   choose(4,x)*choose(3,3-x)/choose(7,3)
12
13 }
14
15
16 b <- a(0:3) # to create a series of vector for X..
17
18 cat("The expectation value of this problem is",
   weighted.mean(0:3,b))
```

---

### R code Exa 4.2 Expected Value Evaluation Problem 2

```
1
2 # Chapter 4
3 # Example 4.2 Page no. 113 from the pdf..
4 # Expected Value Evaluation Problem
5
6 # To find the Expected commission..
7
8 p <- 0.7
9
10 q <- 0.4
11
12 dat <- c(0,1000,1500,2500)
13
14 # Distribution..
15 weights <- c((1-p)*(1-q),p*(1-0.4),(1-p)*q,p*q)
16
17 cat("The expected commission is",weighted.mean(dat,
18       weights))
```

---

### R code Exa 4.3 Expected Value of probability density function

```
1
2 # Chapter 4
3 # Example 4.3 page no. 114 from the pdf..
4 # Expected Value of probability density function
5
6 # given that a(x) if x>100 else 0
7
8 a <- function(x){
9
```

```

10    20000/(x*x*x)
11
12 }
13
14 a1 <- function(x){
15
16     x*a(x)
17
18 }
19
20 cat("The Expected life of this type of device is:",
      integrate(a1,lower = 100,upper = Inf)$value)

```

---

#### **R code Exa 4.4** Expected Value of a new Random Variable

```

1
2 # Chapter 4
3 # Example 4.4 Page no. 115 from the pdf..
4 # Expected Value of a new Random Variable..
5
6 # given a probability distribution find a expected
   value of a new random variable
7 # new random variable 2x-1
8
9 # P(X=x)
10 px <- c(1/12,1/12,1/4,1/4,1/6,1/6)
11
12 # x
13 x <- 4:9
14
15 # g(x)=2x-1
16
17 cat("The Attendant can expect to receive :",weighted
      .mean(2*x-1,px))

```

---

**R code Exa 4.5** Expected Value of a new Random Variable

```
1 # Chapter 4
2 # example 4.5 page no. 115 from the pdf.
3 # Expected Value of New Random Variable
4
5 # to find expected value given a random variable f(x
   )= x^2/3 of g(X)= 4X +3
6
7 f <- function(x){
8
9   (4*x+3)*x*x/3
10
11 }
12
13 cat("The expected value is",integrate(f,-1,2)$value)
```

---

**R code Exa 4.6** Expected value of Discrete Joint Probability Distribution

```
1 # Chapter 4
2 # Example 4.6, page no. 116 from the pdf..
3 # Expected value of Discrete Joint Probability
   Distribution
4
5
6 # given 2 random variables with joint prob.
   distribution ,
7 # find the expected value g(X,Y) = XY.
8
9 f <- matrix(c(3/28,3/14,1/28,9/28,3/14,0,3/28,0,0) ,
   ncol = 3)
10
```

```

11
12 ans <- 0
13
14 for(x in 0:2){
15
16     for(y in 0:2){
17
18         ans <- ans + x*y*f[x+1,y+1]
19
20     }
21 }
22
23 cat("The expected value of  $g(X,Y)=XY$  is",ans)

```

---

#### R code Exa 4.7 Expected value of Joint Density Function

```

1
2 # Chapter 4
3 # Example 4.7 page no. 116 from the pdf..
4 # Expected Value of Joint Density Function
5
6
7 # to find the expected value of density function
8
9 # given function
10
11 a <- function(x,y){
12
13     (x*(1+3*y*y))/4
14
15 }
16
17 # making a expectation value function
18
19 b <- function(x,y){

```

```

20     a(x,y)*y
21
22 }
23
24 # using the expression to find the expectation value
25
26 print("The Value of E(Y/X) is")
27
28 integrate(function(y){
29     supply(y,function(y){
30         integrate(function(x)b(x,y),0,2)$value
31     })
32 },0,1)

```

---

#### R code Exa 4.8 Variance Of Discrete Distribution

```

1
2 # Chapter 4
3 # Example 4.8 page no. 120 from the pdf..
4 # Variance Of Discrete Distribution..
5
6 # given two distribution finding out which set has
   higher variance
7
8 # first distribution
9 a <- c(0.3,0.4,0.3)
10
11 b <- weighted.mean(1:3,a)
12
13 c1 <- function(x){
14     (x-b)*(x-b)
15 }
16

```

```

17 }
18
19 d <- sum(c1(1:3)*a)
20
21 # second Distribution..
22 a1 <- c(0.2,0.1,0.3,0.3,0.1)
23
24 b1 <- weighted.mean(0:4,a1) # since b and b1 are
    same..
25
26
27 d1 <- sum(c1(0:4)*a1)
28
29 cat("Since Variance of B is",d1,"and Variance of A
    is",d,"Variance of B is greater than A")

```

---

#### R code Exa 4.9 Variance using Expectation Formula

```

1
2 # Chapter 4
3 # Example 4.9 page no. 121 from the pdf..
4 # Variance using Expectation formula..
5
6
7 # calculate variance using expectation formula
8
9
10 # given data
11
12 a <- c(0.51,0.38,0.10,0.01)
13
14 # using  $E[x^2] - (E[X])^2$ 
15
16 varia <- weighted.mean(0:3*0:3,a) - (weighted.mean
    (0:3,a))^2

```

```
17
18 cat("The value of Variance is:", varia)
```

---

**R code Exa 4.10** Mean And Variance of Continuous Random Variable

```
1 # Chapter 4
2 # example 4.10 page no. 121 from the pdf..
3 # Mean And Variance of Continuous Random Variable..
4
5 # given a random variable X having the prob. density
6 # find the mean and variance of X..
7
8 f <- function(x){
9
10     2*(x-1)
11
12 }
13
14 cat("The mean is mu = E(X) =", integrate(function(x){
15     x*f(x)}, 1, 2)$value)
16
17 cat("And the Variance is  $E(X^2) - [E(x)]^2 =$ ",
18     integrate(function(x){x*x*f(x)}, 1, 2)$value - (
19     integrate(function(x){x*f(x)}, 1, 2)$value)^2)
```

---

**R code Exa 4.11** Variance of a new Random variable

```
1
2 # Chapter 4
3 # Example 4.11 page no. 122 from the pdf..
4 # Variance of a new Random variable
5
6
```



```

7 # given a random variable find the variance of
  another random variable
8
9 #pdf for X
10 a <- c(1/4,1/8,1/2,1/8)
11
12 # finding variance for 2x+3
13
14 # E[(2x+3-6)^2]
15
16 varia <- weighted.mean(4*0:3*0:3-12*0:3+9,a)
17
18 cat("The variance of g(X)=2X+3 is",varia)

```

---

**R code Exa 4.12** Variance of a new Random variable given the Density Function

```

1
2 # Chapter 4
3 #example 4.12 page no. 123 from the pdf..
4 # Variance of a new Random variable given the
  Density Function..
5
6 # given a random variable find the variance of g(X)
  =4X+3
7
8 f <- function(x){
9
10   x*x/3
11
12 }
13
14 # using theorem 4.3 as given on page no. 122
15
16 cat("The variance of random variable g(4X+3) is",

```

```
integrate(function(x){f(x)*((4*x+3-8)^2)},-1,2)$  
value)
```

---

### R code Exa 4.13 Covariance of 2 Random Variables

```
1  
2 # Chapter 4  
3 # Example 4.13 page no. 124  
4 # Covariance of 2 Random Variables..  
5  
6  
7 # to find the covariance between two variables X and  
8   Y  
9 a <- matrix(c(3/28,3/14,1/28,9/28,3/14,0,3/28,0,0),  
10             nrow = 3)  
11 # individual column sum  
12 ax <- apply(a,2,sum)  
13  
14 # individual row sum  
15 ay <- apply(a,1,sum)  
16  
17 # E[X]..  
18  
19 ex <- sum(0:2*ax)  
20  
21 # E[Y]..  
22  
23 ey <- sum(0:2*ay)  
24  
25 # given E[XY]=3/14  
26  
27 # using covariance formula  
28 # E[XY]-E[X]*E[Y]
```

```

29
30 covar <- 3/14-ex*ey
31
32 cat("The covariance between the random variable X
    and Y is",covar)

```

---

**R code Exa 4.14** Covariance Of Two Random Variables given Joint Density Function

```

1
2 # Chapter 4
3 # Example 4.14 page no. 125 from the pdf..
4 # Covariance Of Two Random Variables given Joint
    Density Function..
5
6
7 # given a joint density function of 2 random
    variables X and Y
8 # find the covariance of X and Y..
9
10 f <- function(x,y){
11
12     8*x*y
13
14 }
15
16 g <- function(x){
17
18     4*(x^3)
19
20 }
21
22 h <- function(y){
23
24     4*y*(1-y*y)

```

```

25
26 }
27
28 # finding E(X) and E(Y) from marginal densities..
29
30 mu_x <- integrate(function(x){x*g(x)},0,1)$value
31
32 mu_y <- integrate(function(y){y*h(y)},0,1)$value
33
34 # E(XY) from joint probability distributon..
35 E_XY <- integrate(function(y){
36
37     sapply(y,function(y){
38
39         integrate(function(x){f(x,y)*x*y},y,1)$value
40     })},0,1)
41
42 cat("The covariance of X and Y is",E_XY$value-mu_x*
    mu_y)

```

---

#### R code Exa 4.15 Correlation Coefficient Evaluation

```

1
2 # Chapter 4
3 # Example 4.15 page no. 126 from the pdf..
4 # Correlation Coefficient Evaluation
5
6 # to find the correlation coefficient between two
  variables x and y..
7 # Refer Example 4.13 page 124 from the pdf..
8
9 a <- matrix(c(3/28,3/14,1/28,9/28,3/14,0,3/28,0,0),
  nrow = 3)
10
11 # individual column sum

```

```

12 ax <- apply(a,2,sum)
13
14 # individual row sum
15 ay <- apply(a,1,sum)
16
17 # E[X]..
18
19 ex <- sum(0:2*ax)
20
21 # E[Y]..
22
23 ey <- sum(0:2*ay)
24
25 # variance(x)..
26
27 variax <- sum(0:2*0:2*ax)-ex^2
28
29 # variance(y)..
30
31 variay <- sum(0:2*0:2*ay)-ey^2
32
33 # covariance..
34
35 covaria <- 3/14-ex*ey
36 # correlation coefficient
37
38 corr <- covaria/(sqrt(variax)*sqrt(variay))
39
40 cat("The correlation coefficient in this case is",
      corr)

```

---

**R code Exa 4.16** Correlation Coefficient Of 2 Random Variables Given Joint Density Function

```
1 # Chapter 4
```

```

2 # Example 4.16 page no. 126 from the pdf..
3 # Correlation Coefficient Of 2 Random Variables
  Given Joint Density Function..
4
5
6 # given a joint density function of 2 random
  variables X and Y
7 # find the correlation coefficient of X and Y..
8
9 f <- function(x,y){
10
11   8*x*y
12
13 }
14
15 g <- function(x){
16
17   4*(x^3)
18
19 }
20
21 h <- function(y){
22
23   4*y*(1-y*y)
24
25 }
26
27 # finding E(X) and E(Y) from marginal densities..
28 mu_x <- integrate(function(x){x*g(x)},0,1)$value
29
30 mu_y <- integrate(function(y){y*h(y)},0,1)$value
31
32 # finding sigma_x and sigma_y from marginal
  densities..
33 sig_x <- integrate(function(x){x*x*g(x)},0,1)$value
  - mu_x^2
34
35 sig_y <- integrate(function(y){y*y*h(y)},0,1)$value

```

```

      - mu_y^2
36
37 # E(XY) from joint probability distributon..
38 E_XY <- integrate(function(y){
39
40     sapply(y,function(y){
41
42         integrate(function(x){f(x,y)*x*y},y,1)$value
43     })},0,1)
44
45 # finding sigma_xy
46 sig_xy <- E_XY$value-mu_x*mu_y
47
48 cat("The correlation coefficient of X and Y is",sig_
      xy/sqrt(sig_x*sig_y))

```

---

**R code Exa 4.17** Expectation Value of Linear Combination Of Random Variables

```

1
2 # Chapter 4
3 # Example 4.17 page no. 128 as given in the pdf..
4 # Expectation Value of Linear Combination Of Random
  Variables..
5
6 # given a random variable and its distribution find
  the value of E(2X-1)
7 # Refer example 4.4 on page 115 from the pdf..
8
9 px <- c(1/12,1/12,1/4,1/4,1/6,1/6)
10
11 x <- 4:9
12
13 cat("The value of E(2X-1) is",weighted.mean(2*x-1,px
  ))

```

---

**R code Exa 4.18** Expectation Value of Linear Combination Of Continuous Random Variables

```
1
2 # Chapter 4
3 # Example 4_18 page no. 129 from the pdf..
4 # Expectation Value of Linear Combination Of Random
  Variables
5
6 # given X a random variable find the expected value
  of another random variable 4X+3
7
8 f <- function(x){
9
10   x*x/3
11
12 }
13
14 # using property  $E(aX+b) = aE(X)+b$ 
15
16
17 cat("The value of  $E(4X+3)$  is", 4*integrate(function(x)
  {x*f(x)},-1,2)$value + 3)
```

---

**R code Exa 4.19** Expectation Value of Linear Combination Of Random Variables Using Properties

```
1
2 # Chapter 4
3 # Example 4.19 page no. 129 from the pdf..
4 # Expectation Value of Linear Combination Of Random
  Variables..
```



```

5
6
7 # finding expected value using properties..
8
9
10 a <- c(1/3,1/2,0,1/6)
11
12 # find  $E[(X-1)^2]$ 
13
14 #  $E[X^2-2X+1]$ 
15
16 # $E[X]$ 
17 ex <- weighted.mean(0:3,a)
18
19 # $E[X^2]$ 
20 ex2 <- weighted.mean(0:3*0:3,a)
21
22 # using properties
23
24 cat("The Expected Value of  $Y=(X-1)^2$  is :",1*ex2 -2*
      ex +1)

```

---

**R code Exa 4.20** Expectation Value Evaluation Using Properties Given a Density Function

```

1
2 # Chapter 4
3 # example 4.20 page no. 130 from the pdf..
4 # Expectation Value Evaluation Using Properties
   Given a Density Function..
5
6
7 # given a density function of a random variable find
   the value of  $E(X^2+X-2)$  using properties..
8

```

```

9 f <- function(x){
10
11   2*(x-1)
12
13 }
14
15 # using properties..
16
17 cat("The value of E(X^2+X-2) is",integrate(function(
    x){x*x*f(x)},1,2)$value + integrate(function(x){x
    *f(x)},1,2)$value - integrate(function(x){2*f(x)
    },1,2)$value)

```

---

**R code Exa 4.21** Expectation Value of 2 Independent Random Variable

```

1
2 # Chapter 4
3 # example 4.21 on page no. 131 from the pdf..
4 # Expected Value of 2 Independent Random Variable..
5
6 # to prove with example for 2 independent random
   variables  $E(XY) = E(x)*E(Y)$ 
7
8 # XY function
9 f <- function(x,y){
10
11   x*(1+3*y*y)/4
12
13 }
14
15 E_XY <- integrate(function(y){
16
17   sapply(y,function(y){
18
19     integrate(function(x){f(x,y)*x*y},0,2)$value

```

```

20   })),0,1)
21
22 # finding E(X) and E(Y) value
23 E_X <- integrate(function(x){x*x/2},0,2)$value
24
25 E_Y <- integrate(function(y){y*(1+3*y*y)/2},0,1)$
      value
26
27 cat("The value of E(XY) is",E_XY$value,"and the
      value of E(X)*E(Y) is",E_X*E_Y,"Hence proved.")

```

---

## Chapter 5

# Some Discrete Probability Distribution

**R code Exa 5.1** Binomial Distribution Problem

```
1
2 # Chapter 5
3 # Example 5.1 Page no. 145 from the pdf..
4 # Binomial Distribution Problem..
5
6 # to find the prob. that exactly 2 in 4 components
   test survive..
7
8 # given prob. of component surviving test 0.75
9
10 cat("The probability that exactly 2 will survive in
   this test out of 4 is",dbinom(2,4,prob = 0.75))
```

---

**R code Exa 5.2.a** Binomial Distribution Problem

```
1
```

```

2 # Chapter 5
3 # Example 5.2a page no. 146 from the pdf..
4 # Binomial Distribution Problem..
5
6 # to find probability that atleast 10 survive in a
   sample of 15 people from a rare blood disease
   which have the probability of recovering from it
   0.4.
7
8 cat("The probability of atleast 10 will survive from
   a sample of 15 people is ", 1-pbinom(9,size =
   15,prob= 0.4))
9
10 #this problem can also be solved by dbinom function
   but we need to do it from 0 to 9 so pbinom
   function is simple to use in this example..
11
12 #The answer in the book is approximated to 0.0338.

```

---

#### R code Exa 5.2.b.c Binomial Distribution Problem

```

1
2 # Chapter 5
3 # Example 5.2b and 5.2c page no. 146 from the pdf..
4 # Binomial Distribution Problem..
5
6 # to find probability that from 3 to 8 survive in a
   sample of 15 people from a rare blood disease
   which have the probability of recovering from it
   0.4.
7
8 cat("The probability that from 3 to 8 survive is",
   pbinom(8,15,0.4)-pbinom(2,15,0.4))
9
10 # the answer in the textbook is 0.8779 which is

```

```

    rounded off to the answer I got here..
11
12 #to find probability that exactly 5 survive in a
    sample of 15 people from a rare blood disease
    which have the probability of recovering from it
    0.4.
13
14 cat("The probabilty that exactly 5 survive is",
    dbinom(5,15,0.4))
15
16 # the answer in the textbook is rounded off to
    0.1859 and the deviation from the answer in both
    cases is less than 2%..

```

---

### R code Exa 5.3 Binomial Distribution Problem 3

```

1
2 # Chapter 5
3 # problem 5.3 on page 146 of pdf..
4 # Binomial Distribution Problem..
5
6 # given - p = 0.03, n = 20
7
8 cat("The probability that there will be at least one
    defective item among these 20 is",1- dbinom
    (0,20,0.03))
9
10 # problem 5.3 b part..
11 # testing of each shipment viewed as a bernoulli
    trial with p = 0.4562 from previous part..
12
13 cat("The probability that there will be exactly 3
    shipments each containing at least one defective
    device among the 20 that are selected and tested
    from the shipment is",dbinom(3,10,0.4562))

```

---

**R code Exa 5.4** Binomial Distribution Problem 4

```
1
2 # Chapter 5
3 # example 5.4 page no. 148 from pdf..
4 # Binomial Distribution problem
5
6 # given  $p = 0.3$ ,  $n = 10$ , binomial distribution
7
8 cat("The probability that exactly 3 wells have the
    impurity assuming that the conjecture is correct
    is",dbinom(3,10,0.3))
9
10 # example 5.4 b part..
11
12 cat("The probability that more than 3 wells are
    impure is",pbinom(3,10,0.3,lower.tail = F))
13 # the answer in the T.B is approx. to 0.3504..
```

---

**R code Exa 5.6** Binomial Distribution Problem 6

```
1
2 # Chapter 5
3 # example 5.6, page no. 148 from the pdf..
4 # Binomial Distribution Problem..
5
6 # Note- Example 5.5 not solved as the solution
    involves only the multiplication of 2 numbers..
7
8 # given -  $p = 0.3$ ,  $n = 10$ , binomial distribution
9
```

```

10
11 cat("The probability that more than 6 are found to
    contain the impurity from 10 wells is",1-pbinom
    (5,10,0.3))
12
13 print("As the prob. is 0.0473 approx., this casts
    considerable doubt on conjecture and suggests
    that the impurity problem is more severe.")

```

---

### R code Exa 5.7 Multinomial Distribution Problem

```

1
2 # Chapter 5
3 # Example 5.7 page no. 150 from the pdf..
4 # Multinomial Distribution..
5
6 # to find the prob. of 6 randomly arriving planes
    which are distributed in 3 runways and each have
    a ideal arriving prob.
7
8 p_runways <- c(2/9,1/6,11/18)
9
10 # prob. of runway 1 :2
11         # runway 2: 1
12         # runway 3 : 3
13
14 x <- c(2,1,3) # a particular case of the
    distribution of 6 randomly arriving , written in a
    vactor
15
16 # this is the case of multinomial distribution
17
18 cat("The probability of this type of distribution of
    6 randomly arriving airplanes is",dmultinom(x,6,
    p_runways))

```



---

**R code Exa 5.8** Hypergeometric Distribution In Acceptance Sampling

```
1
2 # Chapter 5
3 # example 5.8 page no. 153 from pdf..
4 # Hypergeometric Distribution In Acceptance Sampling
5   ..
6 # radom sampling , testing 3 of the parts out of 10
7 # condition – if none of the 3 is defective then the
8   lot is accepted
9
10 # assume lot is truly unacceptable..
11
12 cat("The probability that the sampling plan finds
    the lot acceptable is",dhyper(0,m = 2,n = 8,k =
    3))
11
12 cat("Thus if the lot is truly unacceptable with 2
    defective parts , this sampling plan will allow
    acceptance roughly",dhyper(0,2,8,3)*100,"% of the
    time, so this plan should be considered faulty."
    )
```

---

**R code Exa 5.9** Hypergeometric Distribution In Acceptance Sampling

```
1
2 # Chapter 5
3 # Example 5.9 Page no. 154 from the pdf..
4 # Hypergeometric Distribution Problem..
5
```

```

6 # to find prob. that exactly 1 defective if there
   are 3 in the entire lot of 40 and 5 are selected
   randomly
7
8 # this is the problem of hypergeometric distribution
   ..
9
10 cat("The probability that exactly 1 is found
      defective from 5 randomly selected components is"
      ,dhyper(1,3,37,5))

```

---

#### **R code Exa 5.10** Hypergeometric Probability Distribution Problem

```

1
2 # Chapter 5
3 # example 5.10, page no. 154 from the pdf..
4 # Hypergeometric Probability function..
5
6 # given – lot of 100 items in which 12 are defective
   ..
7 # to find the prob. that in a sample of 10, 3 are
   defective..
8 # answer can be found using hypergeometric
   probability function
9
10 cat("The probability that in a sample of 10 =, 3 are
      found to be defective is",dhyper(3,m = 12,n =
      88,k = 10))

```

---

#### **R code Exa 5.11** Mean And Variance Of Hypergeometric Experiment

```

1
2 # Chapter 5

```

```

3 # Example 5.11 page no. 155 from the pdf..
4 # Mean And Variance Of Hypergeometric Experiment..
5
6 # to find the mean and variance of the random
  variable which has hypergeometric distribution..
7
8 # mean..
9 me <- sum(0:3*dhyper(0:3,3,37,5))
10
11 # variance..
12 v <- sum(0:3*0:3*dhyper(0:3,3,37,5))-me*me
13
14 cat("The mean of this experiment is",me,"And
  Variance is",v)

```

---

**R code Exa 5.12** Relation Between Hypergeometric and Binomial Distribution

```

1
2 # Chapter 5
3 # example 5.12, page no. 155 from the pdf..
4 # Relation Between Hypergeometric and Binomial
  Distribution..
5
6 # given – out of 5000, 1000 are slightly blemished..
7 # to find the prob. that exactly 3 are blemished if
  one purchases 10 tires at random
8
9 # since the N is large (5000) relative to sample
  size 10 we can do approx. to binomial
  distribution..
10
11
12 cat("The probability of obtaining exactly 3
  blemished tires from 10 randomly purchased

```

```
samples is",dbinom(3,10,0.2))
```

---

**R code Exa 5.13** Multivariate Hypergeometric Distribution Problem

```
1
2 # Chapter 5
3 # Example 5.13 page no. 156 from the pdf..
4 # Multivariate Hypergeometric Distribution..
5
6 # to find the multinomial hypergeometric
  distribution of the following problem..
7 # Package Used – "extraDistr" , reference– Internet
  ..
8
9 # If you have already installed it remove the
  command below..
10 install.packages("extraDistr") # this package
  contain functions for various types of prob.
  distribution
11
12 library("extraDistr")    # to use this package..
13
14 #to find the prob. of 1 out 3 , 2 out 0f 4 and 1 out
  of 3, if 5 people are randomly selected from 10
  people
15
16 cat("The probability that the 5 randomly selected
  from 10 has the above distribution is",dmvhyper(
  matrix(c(1,2,2),ncol = 3),c(3,4,3),5))
```

---

**R code Exa 5.14** Negative Binomial Distribution Problem

```
1
```

```

2 # Chapter 5
3 # example 5.14 page no. 159 from the pdf..
4 # Negative Binomial Distribution Problem..
5
6 # given A has 0.55 prob. of winning over B
7
8 # prob. that A will win in 6 games given that if any
   team win 4 out 7 wins then that team winner
9
10 # using negative binomial distribution we get..
11
12 cat("The probability that A will win the series in 6
   games is",dnbinom(2,4,0.55))
13
14 cat("The probability that team A will win the series
   is",pnbinom(3,4,0.55))
15 # for this atleast 4 matches to be won I can also
   use dnbinom but I have to sum all from size 4 to
   7
16
17
18 # now A and B are playing and winning series decided
   by 3 wins out of 5 games
19
20 cat("The probability that A will win the playoff is"
   ,pnbinom(2,3,0.55))
21
22 #The answer in the textbook is approximated to 4
   decimal places..

```

---

#### **R code Exa 5.15** Geometric Distribution Problem

```

1
2 # Chapter 5
3 # example 5.15 page no. 160 from the pdf..

```

```

4 # Geometric Distribution Problem..
5
6 # find prob. that 5th item inspected is first
  defective found given 1 in every 100 items is
  defective..
7
8 # using geometric distribution..
9
10 cat("The probability that 5th item inspected is
    found to be defective is",dgeom(4,0.01))
11
12 # the answer in textbook is approximated, 4 digits
    to the decimal..

```

---

#### **R code Exa 5.16** Geometric Distribution Problem

```

1
2 # Chapter 5
3 # example 5.16, page no. 160 from the pdf..
4 # Geometric Distribution Problem..
5
6 # p = 0.05 of a connection during busy time..
7 # to find the prob. that 5 attempts are necessary
  for a successful call.
8
9 cat("The probability that 5 attempts are necessary
    for a successful call is",dgeom(4,0.05))

```

---

#### **R code Exa 5.17** Poisson Distribution Problem

```

1
2 # Chapter 5
3 # Example 5.17 page no. 162 from the pdf..

```

```

4 # Poisson Distribution Problem..
5
6 # given – avg no. of radioactive particle entering
  in 1msec is 4,
7
8 # to find– prob. of 6 particles entering in a given
  msec..
9
10 # using poisson distribution..
11
12 cat("The probability of 6 particles entering in a
  given millisecond is",dpois(6,4))
13
14 # the answer in the textbook is 0.1042 approx. to 4
  decimal places..

```

---

#### R code Exa 5.18 Poisson Distribution Problem

```

1
2 # Chapter 5
3 # example 5.18, page no. 162 from the pdf..
4 # Poisson Distribution Problem..
5
6 # avg no. of oil takers arriving each day = 10,
7 # facilities at port can handle 15 tankers per day.
8 # to find the prob. that on a given day tankers have
  to be turned away..
9
10
11 cat("The probability that on a given day tankers
  have to be turned away is",ppois(15,10,lower.tail
    = F))

```

---

### R code Exa 5.19 Approximation of Binomial By a Poisson Distribution

```
1
2 # Chapter 5
3 # ex. 5.19 page no. 164 from the pdf..
4 # Approximation of Binomial By a Poisson
  Distribution..
5
6 # given  $P(\text{accident on given day})=0.005$  and
  independent of each other..
7
8 # to find a. prob. of accident on 1 day from 400
  days..
9
10 # this is the case of binomial distribution but in
  this case we can use poisson approx.
11
12 #  $\lambda = np=2$  in this case..
13
14 # answer using binomial ..
15 cat("The probability that there will be accident on
  1 day in a given period of 400 days is",dbinom
  (1,400,0.005))
16
17 # answer using poisson..
18 cat("Using poisson distribution the above answer
  becomes",dpois(1,2))
19 # so when n is very large and p is very small we can
  convert binomial to poisson..
20 # the margin of error is  $1.2e-06$  which negligible so
  we can use this conversion
21
22 # to find prob. of accident for atmost 3 days from
  400 days..
23
24 # again we can use poisson here..
25
26 cat("The probability that there are atmost 3 days
```



```

    with an accident is",ppois(3,2))
27
28 # the answer in the T.B is approx. to 3 decimal
    places 0.857.

```

---

**R code Exa 5.20** Approximation of Binomial By a Poisson Distribution

```

1
2 # Chapter 5
3 # example 5.20 page no. 164 from the pdf..
4 # Approximation of Binomial By a Poisson
    Distribution..
5
6 # 1 in every 100 items produced has one or more
    bubbles means p = 0.001..
7 # to find the prob. that a random sample of 8000
    will yield fewer than 7 items processing bubbles ,
8 # in short find  $P(X < 7)$ 
9
10 # since p is very close to 0 and n is very large we
    can approx. it with Poisson distribution..
11
12 cat("The prob. that the random sample of 8000 will
    yield fewer than 7 items possessing bubbles is",
    ppois(6,8))

```

---

## Chapter 6

# Some Continuous Probability Distributions

**R code Exa 6.1** Uniform Distribution Problem

```
1
2 # Chapter 6
3 # Example 6.1 page no. 171 from the pdf..
4 # Uniform Distribution Problem..
5
6 # to find the prob. density function of uniform
   distribution over 0 to 4
7
8 a <- function(x){
9
10   if(x>=0 & x<=4){
11
12     dunif(x,0,4)
13
14   }
15
16 }
17
18 # P(X>=3)
```

```
19
20 cat("The probability that any conference last at
    least 3 hours is",1-punif(3,0,4))
```

---

### **R code Exa 6.2** Normal Distribution Problem

```
1
2 # Chapter 6
3 # Example 6.2 page no.178 from the pdf..
4 # NOrmal Distribution Problem..
5
6 # given standard normal distribution..
7
8
9 cat("The area under the curve to the right of 1.84
    is",pnorm(1.84,lower.tail = F))
10
11
12 cat("The area under the curve that lies between z
    =-1.97 to z=0.86 is",pnorm(0.86)-pnorm(-1.97))
13
14
15 # the answer in the textbook is approx.to 4 decimal
    places..
```

---

### **R code Exa 6.3** Find The Value of k given Normal Distribution

```
1
2 # Chapter 6
3 # Example 6.3 page no. 179 from the pdf..
4 # Normal Distribution..
5
6 # to find the k such that..
```

```

7
8 cat("The value of k such that P(Z>k)= 0.3015 is",
      qnorm(0.3015,lower.tail = F))
9
10 cat("The value of k such that P(k<z<-0.18)=0.4197 is
      ",qnorm(pnorm(-0.18)-0.4197))

```

---

**R code Exa 6.4** Find area under the Normal Distribution curve

```

1
2 # Chapter 6
3 # Example 6.4 page no. 179 from the pdf..
4 # Find area under the Normal Distribution..
5
6 # given popukation mean= 50 ,sd= 10
7
8 cat("The probability that random variable X assumes
      a value between 45 and 62 is",pnorm(62,50,10)-
      pnorm(45,50,10))

```

---

**R code Exa 6.5** Find the Probability given Normal Distribution

```

1
2 # Chapter 6
3 # Example 6.5 page no. 180 from the pdf..
4 # Find the Probability given Normal Distribution..
5
6 # X normal distribution ,
7 # mu= 300, sigma= 50
8 # find P(X > 362)
9

```

```
10 cat("The probability that X assumes a value greater
    than 362 is",pnorm(362,mean = 300,sd = 50,lower.
    tail = F))
```

---

#### **R code Exa 6.6** Using Normal Curve In Reverse

```
1
2 # Chapter 6
3 # Example 6.6 page no. 181 from the pdf..
4 # Using Normal Curve In Reverse..
5
6
7 # given population mean= 40, sd= 6
8
9 cat("The value of x that has 45% of the area to the
    left is",qnorm(0.45,40,6))
10
11 # the deviation of this value from the T.B is less
    than 2%.
12
13 cat("The value of x that has 14% of the area to the
    right is",qnorm(0.14,40,6,lower.tail = F))
14
15 #The answer in T.B is approximated to 2 decimal
    places..
```

---

#### **R code Exa 6.7** Applications of Normal Distribution

```
1
2 # CHapter 6
3 # Example 6.7 page no. 182 from the pdf..
4 # Applications of Normal Distribution..
5
```

```

6 # given – average battery life – 3 years
7 # standars deviation – 0.5
8 # to find P(X<2.3)
9
10 cat("The probability that a given battery will last
    less than 2.3 years is",pnorm(2.3,mean = 3,sd=
    0.5))

```

---

#### R code Exa 6.8 Applications of Normal Distribution

```

1
2 # Chapter 6
3 # example 6.8 page no. 182 from the pdf..
4 # Applications of Normal Distribution..
5
6
7 # given distribution of life of bulb – normal, mean=
    800, sd=40hrs
8
9 cat("The probability that a bulb burns between 778
    and 834 hours is",pnorm(834,800,40)-pnorm
    (778,800,40))
10
11
12 # making a visual of distribution
13 # Optional not compulsory..
14 x <- seq(680,920,length=200)
15 y <- dnorm(x,800,40)
16 plot(x,y,type="l")
17 abline(v = c(778,834))

```

---

#### R code Exa 6.9 Applications of Normal Distribution

```

1
2 # Chapter 6
3 # Example 6.9 page no. 182 from the pdf..
4 # Applications of Normal Distribution..
5
6
7 # given normal, mean= 3,sd= 0.005
8
9 # specification on diameter of ball bearing 3+0.01
  and 3-.01
10
11 cat("On average",2*pnorm(2.99,3,0.005)*100,"% of
    ball bearings will be scrapped")

```

---

#### R code Exa 6.11 Applications of Normal Distribution

```

1
2 # Chapter 6
3 # Example 6.11 page no. 184 from the pdf..
4 # Applications of Normal Distribution..
5
6 # Note – We just have to calculate d in example
  6.10, also many questions of same type have been
  solved previously..
7
8 # given mean resistance= 40 and sd= 2 , normal
  distribution
9
10
11 # to find %age of resistors exceeding 43 ohms
12
13 cat("The %age of resistors having resistance
    exceeding 43ohms is",100*pnorm(43,40,2,lower.tail
    = F),"%")

```

---

### R code Exa 6.12 Applications of Normal Distribution

```
1
2 # Chapter 6
3 # Example 6.12 page no.184 from the pdf..
4 # Applications of Normal Distribution..
5
6 # to find the % of resistance exceeding 43 ohms from
   example 6.11, if resistance is measured to the
   nearest ohm.
7
8 # mean resistance = 40, sd= 2, normal distribution..
9
10 # here we will assign measurement of 43 ohms to all
    resistors whose value are greater than 43 and
    lesser than 43.5
11
12
13 cat("The difference between resistance that exceed
    43 ohms and that exceeds beyond 43 is",-pnorm
    (43.5,mean = 40,sd=2,lower.tail = F)+pnorm(43,
    mean = 40,sd=2,lower.tail = F),"so this value in
    %age represents all resistance grater than 43 and
    less than 43.5 that are now being recorded as 43
    ohms")
```

---

### R code Exa 6.13 Applications of Normal Distribution

```
1
2 # Chapter 6
3 # Example 6.13 page no. 185 from the pdf..
4 # Applications of Normal Distribution..
```



```

5
6 # given - avg score of the class - 74, sd = 7..
7 # given - 12% score "A"s and grades follow a normal
  distribution
8
9 # to find the lowest possible score for A and
  highest for B..
10
11
12 cat("The percentile for 0.12 in this distribution is
    ", qnorm(0.12, 74, 7, lower.tail = F), "marks")
13
14 cat("The lowest score for A is 83", "and The highest
    score for B is 82")

```

---

#### R code Exa 6.14 Applications of Normal Distribution

```

1
2 # Chapter 6
3 # example 6.14 page no.185 from the pdf..
4 # Applications of Normal Distribution..
5
6 # given - mean 74, sd - 7
7 # to find the sixth decile..
8
9 cat("The sixth decile is", qnorm(0.6, mean = 74, sd =
    7))
10
11
12 # the answer varies slightly due to approximation
    used in the T.B

```

---

#### R code Exa 6.15 Normal Approximation To the Binomial

```

1
2 # Chapter 6
3 # Example 6.15 Page no. 191 from the pdf..
4 # Normal Approximation To the Binomial..
5
6 # given- P(patient recovers from disease)=0.4
7 # to find prob. that less than 30 survive if 100
   have prone to this disease
8
9 # since n=100, this binomial variable can be approx.
   to normal model
10 p_survive=0.4
11
12 mu <- 100*p_survive
13
14 sig <- sqrt(100*p_survive*(1-p_survive))
15
16 # no. of people are integers , so we need to use
   approx, and so to obtain prob, we need to find
   area to the left of 29.5
17
18 cat("The probability that fewer than 30 of the 100
   patients survive is given by",pnorm(29.5,mu,sig))

```

---

#### **R code Exa 6.16** Normal Approximation To the Binomial

```

1
2 # Chapter 6
3 # Example 6.16 page no. 192 from the pdf..
4 # Normal Approximation To the Binomial..
5
6 # to find the prob. that sheer guesswork yields from
   25 to 30 correct answer for 80 of the 200
7 # problem about which student has no knowledge..
8

```

```

9 # since the sample size is large we can do normal
  approx.
10
11 p <- 0.25      # prob. of guessing a correct answer.
12
13 n <- 80
14
15 # now using normal approx. mu = p*n and sd <- sqrt(n
    *p*(1-p))
16
17 mu = p*n
18
19 sd <- sqrt(n*p*(1-p))
20
21
22 cat("The prob. of correctly guessing from 25 to 30
    questions is",pnorm(30.5,mu,sd)-pnorm(24.5,mu,sd)
    )

```

---

#### R code Exa 6.17 Exponential Distribution Problem

```

1
2 # Chapter 6
3 # Example 6.17 page no. 197 from the pdf..
4 # Exponential Distribution Problem..
5
6 # given - T- time of failure
7 # T is exponential distribution beta=5
8 # find prob. that atleast 2 are functioning at the
    end of 8 years, if 5 of these components are
    taken..
9
10
11 bet <- 5
12

```

```

13 # first to find probability of a component
    functioning after 8 years
14
15
16 a <- pexp(8,rate=1/bet,lower.tail = F)
17
18 # finding total probability..
19
20 cat("The probability that atleast 2 are still
    functioning from 5 components after 8 years is"
    ,1-pbinom(1,5,a))
21
22 # the answer in the textbook is approximated to 0.2(
    P(T>8)) so there is also small error in the final
    probab.
23
24 # using computation we can reduce such error by a
    large extent..

```

---

#### R code Exa 6.18 Gamma Distribution Problem

```

1
2 # Chapter 6
3 # Example 6.18 page no. 178.
4 # Gamma Distribution Problem..
5
6 # given - beta= 1/5, alpha= 2
7
8 # to find the probability up to a minute will elapse
    by the time 2 calls have come in to the
    switchboard..
9
10 cat("The probability that up to a minute will elapse
    by the time 2 calls have come in to the
    switchboard is",pgamma(1,shape = 2,scale = 1/5))

```

```
11
12
13 # the answer in T.B is approx. to 0.96
```

---

**R code Exa 6.19** Application of Gamma Distribution

```
1
2 # Chapter 6
3 # Example 6.19 page no.198 from the pdf..
4 # Application of Gamma Distribution..
5
6 # given - X is survival time..has gamma distribution
  with
7 # alpha = 5, beta =10
8
9 # find P(rat survives <=60) or P(X<=60)
10
11 # using gamma distribution..
12
13
14 cat("The probability that the rat survives no longer
    than 60 days is", pgamma(60, shape=5, scale= 10))
15
16 # The answer in the T.B is approx. to 3 decimal
    places..
```

---

**R code Exa 6.20** Application of Gamma Distribution

```
1
2 # Chapter 6
3 # ex. 6.20 page no. 199 from the pdf..
4 # Application of Gamma Distribution..
5
```

```

6
7 # given - beta= 4, alpha= 2,
8 # changes were made to tighten quality control
  requirement,
9 # after these 20 months were passed, so we are asked
  to find that if the quality control tightening
  was effective..
10
11
12 cat("The p-value for such time is",pgamma(20,shape =
  2,scale = 4,lower.tail = F))
13
14 print("Since the p- value is less so conditions of
  gamma distributions with alpha=2,beta=4 are not
  supported by data, hence we can conclude that
  quality control work was effective.")

```

---

**R code Exa 6.21** Cumulative Distribution For Exponential Distribution Problem

```

1
2 # Chapter 6
3 # Example 6.21 page 199 from the pdf..
4 # Function Which resembles Exponential Distribution
  ..
5
6 # given a density function which resembles
  exponential function
7 # with beta= 4
8 # to find  $P(Y > 6)$ 
9
10
11 cat("The value of  $P(Y > 6)$  is",pexp(6,rate = 1/4,
  lower.tail = F))
12

```

```

13 cat("the above statement can interpret as the prob.
    that the washing machine wil require major repair
    after year 6 is",pexp(6,rate = 1/4,lower.tail =
    F))
14
15 cat("The prob. that a major repair is necessary in
    the first year is",pexp(1,rate = 1/4))

```

---

#### R code Exa 6.22 Lognormal Distribution Problem

```

1
2 # Chapter 6
3 # Example 6.22 page no. 202 from the pdf..
4 # Lognormal Distribution Problem..
5
6 # given X= pollutant concentration ,follows log
    normal distribution
7 # parameters of X- mu= 3.2 and sd = 1
8
9 # to find P(X>8)..?
10
11 #Using lognormal distribution
12
13 cat("The probability that the concentration exceeds
    8 parts per million is",plnorm(8,meanlog = 3.2,
    sdlog = 1))
14
15 # the error in the T.B and computed answer is
    0.0001621 which is very small and can be
    neglected..

```

---

#### R code Exa 6.23 Find Percentile From Lognormal Distribution

```

1
2 # Chapter 6
3 # example 6.23 page number – 203 from pdf..
4 # Lognormal Distribution Problem..
5
6 # given a lognormal distribution with mu = 5.149 and
   sigma = 0.737
7 # to find the 5th percentile of the life of
   electronic control...
8
9
10 cat("The 5th percentile of such an distribution is",
      qlnorm(0.05,meanlog = 5.149, sdlog = 0.737))

```

---

#### R code Exa 6.24 Weibull Distribution Problem

```

1
2 # Chapter 6
3 # Example 6.24 on page no. 204 of pdf..
4 # Weibull Distribution Problem..
5
6 # given X is weibull distribution with alpha = 0.01
   and beta = 2
7 # to find P(X < 8)
8
9 cat("The probability that life of X falls before 8
   hours of usge is",pweibull(8,shape = 2,scale =
   sqrt(100)))
10
11 # Note – Chapter 7 contains only theoretical
   problems, so I am not solving that chapter, the
   chapter contains 7 problems total and all are
   theoretical with nothing or very less to compute
   anything..
12 # I will start with 8th Chapter..

```





## Chapter 8

# Fundamental Sampling Distributions And Data Descriptions

**R code Exa 8.1** Sample Mode Evaluation

```
1
2 # Chapter 8
3 # Example 8.1 page no. 228 from the pdf..
4 # Sample Mode Evaluation..
5
6 # to find the mode of the given data
7 mod <- function(v){
8
9   a <- unique(v)
10  a[which.max(tabulate(match(v,a)))]
11
12
13 }
14 set <- c
    (0.32,0.53,0.28,0.37,0.47,0.43,0.36,0.42,0.38,0.43)
15
```

```
16 cat("The sample mode is",mod(set))
```

---

#### **R code Exa 8.2** Sample mean and Sample Variance

```
1
2 # Chapter 8
3 # Example 8.2 page no. 229 from the pdf..
4 # Sample mean and Sample Variance..
5
6 # to find mean and variance..
7
8 set <- c(12,15,17,20)
9
10 cat("The mean and variance are",mean(set),"and",var(
    set),"respectively")
```

---

#### **R code Exa 8.4** Central Limit Theorem

```
1
2 # Chapter 8
3 # Example 8.4 page no. 234 from the pdf..
4 # Central Limit Theorem..
5
6 # given – bulb life normally distributed , mean= 800
    ,sd= 40
7
8 # prob. that average life is less than 775hrs for a
    random sample of 16 bulbs..
9
10 sd <- 40
11
12 n <- 16
13
```

```

14 # Using central limit theorem..
15
16 # sampling distribution..
17
18 sam_mean <- 800
19
20 sam_sd <- sd/sqrt(n)
21
22 cat("The probability that the random sample of 16
    have average lifetime less than 775hrs is",pnorm
    (775,sam_mean,sam_sd))

```

---

#### R code Exa 8.5 Central limit theorem

```

1
2 # Chapter 8
3 # Example 8.5 page no. 237 from the pdf..
4 # Central limit theorem..
5
6 # mu= 28, sd= 5,n= 40, to find prob. that
    averagetransport time greater than 30 min..
7
8 # P(Xbar > 30)
9
10
11 cat("The probability that the average transport time
    was more than 30 min. is",pnorm(30+0.5,mean =
    28,sd= 5/sqrt(40),lower.tail = F))

```

---

#### R code Exa 8.6 Sampling Distribution of the difference of 2 means

```

1
2 # Chapter 8

```

```

3 # Example 8.6 page no. 240 from the pdf..
4 # Sampling Distribution of the difference of 2 means
5   ..
6 # given information about 2 manufacturers
7
8
9 A <- c(6.5,0.9,36)
10
11
12 B <- c(6,0.8,49)
13
14
15 # prob. 36 tubes from A have mean lifetime at least
16   1 more than mean of sample of 49 from B??
17 diff_mean <- A[1]-B[1]
18
19 diff_sd <- sqrt((A[2]*A[2]/A[3]) + (B[2]*B[2]/B[3]))
20
21 cat("The probability that the mean lifetime for 36
22     tubes from A will be at least 1 year longer than
23     the mean lifetime for 49 tubes from B is",pnorm
24     (1,diff_mean,diff_sd,lower.tail = F))

```

---

### **R code Exa 8.7** The Chi squared Distribution Problem

```

1
2 # Chapter 8
3 # Example 8.7 page no. 245 from the pdf..
4 # The Chi-squared Distribution Problem..
5
6 # given mu=3,sd=1,data of 5 batteries ,
7 # to show whether manufacturer should be convinced
8   that batteries have a sd of 1..

```

```

8
9 sd = 1
10
11 obs <- c(1.9,2.4,3.0,3.5,4.2)
12
13 chi.sq <- var(obs)*(length(obs)-1)/(sd^2)
14
15 cat("The value of chi^2 is",chi.sq,"with 4 degrees
    of freedom")
16
17 cat("Since 95% of the chisq values with 4 degrees of
    freedom fall between",qchisq(0.025,df=4),"and",
    qchisq(0.975,df= 4),"the computed value with sig
    ^2= 1 is reasonable and therefore manufacturer
    has no reason to suspect that the sd is other
    than 1")

```

---

#### R code Exa 8.8 The t distribution Problem

```

1
2 # Chapter 8
3 # Example 8.8 page no. 249 from the pdf..
4 # The t-distribution Problem..
5
6 # to find the t value with left tail area 0.025 and
    degrees of freedom =14
7
8 cat("The t value with degrees of freedom 14 and
    leaves an area of 0.025 to the left is",qt(0.025,
    df= 14))
9
10 # Note- In example 8.9 the degrees of freedom is not
    mentioned so I am not solving example 8.9, I
    think the solution is wrong..

```

---

**R code Exa 8.10** The t distribution Problem

```
1
2 # Chapter 8
3 # Example 8.9 page no. 249 from the pdf..
4 # The t- distribution Problem
5
6
7 # find k such that  $P(k < T < -1.761) = 0.045$ , sample size
  = 15, normal distribution
8
9 n <- 15
10
11 cat("The value of k such that  $P(k < T < -1.761) = 0.045$ 
    is ", qt(pt(-1.761, n-1) - 0.045, n-1))
```

---

**R code Exa 8.11** Applications of t distribution

```
1
2 # Chapter 8
3 # Example 8.11 page no. 250 from the pdf..
4 # Applications of t-distribution..
5
6 # mu 500, xbar= 518 , sample sd = 40,
7 # if t value between  $t(-0.05)$  and  $t(0.05)$  then the
  engineer is satisfied.
8
9 xbar <- 518
10
11 mu <- 500
12
13 sd <- 40
```

```

14
15 n <- 25
16
17 t_value = (xbar-mu)/(sd/sqrt(n))
18
19 cat("The engineer is satisfied if a sample of 25
    batches yields a t value between",qt(0.05,df=24),
    "and",-qt(0.05,df=24))
20
21 cat("The t value is",t_value,"which is above the
    levels mentioned above so the p value is",pt(t_
    value,df=24,lower.tail = F),"so we conclde that
    the process produces a better product than the
    engineer thought.")
22
23
24 # the answer varies slightly due to approximation in
    book.

```

---

**R code Exa 8.12** To Construct a Normal quantile quantile plot

```

1
2 # Chapter 8
3 # Example 8.12 page no. 258 from the pdf..
4 # To Construct a Normal quantile-quantile plot..
5
6 # to plot the normal quantile-quantile plot of two
    data samples and draw conclusion whether they are
    from same normal distribution.
7
8 y <-c
    (5030,13700,10730,11400,860,2200,4250,15040,4980,11910,8130,26850
9
10 b <- c

```



```
(2800,4670,6890,7720,7030,7330,2810,1330,3320,1230,2130,2190)

11
12
13 qqnorm(y,ylim = range(y,b),col = "blue")
14 par(new= TRUE)
15 qqnorm(b,ylim=range(y,b),col = "red")
16
17
18 # The clustering of observation would make it seem
    unlikely that two sample came from common normal
    distribution.
```

---

# Chapter 9

## One And Two Sample Estimation Problems

**R code Exa 9.2** Confidence Interval Evaluation

```
1
2 # Chapter 9
3 # Example 9.2 page no. 271 from the pdf..
4 # COnfidence Interval Evaluation..
5
6 # to find the confidence interval given mu, sigma
7
8 n <- 36
9
10 mu <- 2.6
11
12 sig <- 0.3
13
14 a <- qnorm(0.975)*sig/sqrt(n)
15
16 b <- qnorm(0.995)*sig/sqrt(n)
17
18 cat("The 95% confidence interval for the above
      parameters is",mu-a,mu+a," milliliter")
```

```

19
20 cat("The 99% confidence interval for the above
    parameters is",mu-b,mu+b," milliliter")
21
22
23 # the similar question can be addressed as shown
    below
24 set.seed(100)
25 v <- rnorm(36,2.6,0.3)
26 t.test(v,conf.level = 0.95)$conf.int
27
28 # but as we can see each time different ci are there
    , if n becomes very large then it will be very
    close to the standard answer..

```

---

### R code Exa 9.3 To Find the sample size

```

1
2 # Chapter 9
3 # Example 9.3 page no. 273 from the pdf..
4 # To Find the sample size..
5
6
7 # to find the samaple size if we want to be 95%
    confident that our estimate of mu in ex. 9.2 is
    off by less than 0.05
8
9 sd <- 0.3
10
11 n <- ((qnorm(0.975)*sd)/0.05)^2
12
13 cat("We can be 95% confident that a random sample of
    size",as.integer(n)+1," will provide an estimate
    xbar differing from mu by an amount less than
    0.05")

```

---

**R code Exa 9.4** To find 95 Percent bound for mean

```
1
2 # Chapter 9
3 # example 9.4 page no. 274
4 # To find 95% bound for mean..
5
6 # to find the 95% bound for the mean reaction time
7
8 avg_t <- 6.2
9
10 n <- 25
11
12 s_d <- 2
13
14 cat("The 95% bound is given by", avg_t + qnorm(0.95) * s_d / sqrt(n))
```

---

**R code Exa 9.5** To Find 95 Percent Confidence Interval for mean

```
1
2 # Chapter 9
3 # Example 9.5 page no. 275 from the pdf..
4 # To Find 95% Confidence Interval for mean..
5
6
7 # to calculate the 95% confidence interval for the
  contents of 7 containers of sulphuric acid
8
9
10 a <- c(9.8, 10.2, 10.4, 9.8, 10.0, 10.2, 9.6)
```

```

11
12
13 cat("The confidence interval for the above data
      sample is", " ")
14
15 t.test(a, conf.level = 0.95)$conf.int

```

---

#### R code Exa 9.6 To Find 99 Percent Confidence Interval

```

1
2 # Chapter 9
3 # Example 9.6 page no. 276 from the pdf..
4 # To Find 99% Confidence Interval..
5
6 # to find 99% confidence interval on the mean SAT
  maths score.
7 # Since random sample is large 500 we can use normal
  approximation..
8
9 mu <- 501
10
11 sd <- 112
12
13 n <- 500
14
15 sd1 <- sd/sqrt(n)
16
17 cat("The 99% confidence interval for the above
      example is", mu+qnorm(0.005)*sd1, mu-qnorm(0.005)*
      sd1)

```

---

#### R code Exa 9.7 To find 95 Percent Prediction Interval

```

1
2 # Chapter 9
3 # Example 9.7 page no. 278 from the pdf..
4 # To find 95% Prediction Interval..
5
6 # to find the 95% prediction interval for the loan
  amount.
7 # given ..
8 mu <- 257300
9
10 sd <- 25000
11
12 n <- 50
13
14 c <- sqrt(1+1/n)
15
16 cat("The 95% prediction interval for the future loan
  amount is",mu-qnorm(0.9750,lower.tail = T)*sd*c,
  mu+qnorm(0.9750,lower.tail = T)*sd*c)

```

---

**R code Exa 9.8** To find 99 Percent prediction interval

```

1
2 # Chapter 9
3 # Example 9.8 page no. 279 from the pdf..
4 # to find 99% prediction interval..
5
6 # sample size = 30,mean = 96.2%, sd = 0.8
7
8 mu <- 96.2
9
10 sd <- 0.8
11
12 n <- 30
13

```

```

14 c <- sqrt(1+1/n)
15
16 t_val <- abs(qt(0.005,df= 29))
17
18 cat("The prediction interval for a new observation
      is ",mu-t_val*sd*c,"and",mu+t_val*sd*c)

```

---

### R code Exa 9.10 Confidence Interval for the difference between 2 means

```

1
2 # Chapter 9
3 # Example 9.10 page no. 286 from the pdf..
4 # Confidence Interval for the difference between 2
  means..
5
6 # to find 96% confidence interval for the given
  information
7 # given..
8 a_mu <- 36
9
10 b_mu <- 42
11
12 n_a <- 50
13
14 n_b <- 75
15
16 sd_a <- 6
17
18 sd_b <- 8
19
20 c <- qnorm(0.98)*sqrt((sd_a*sd_a/n_a)+(sd_b*sd_b/n_b
  ))
21
22 d <- b_mu - a_mu
23

```

```
24 cat("The 96% confidence interval of the difference
    in average gas mileage is",d-c,d+c)
```

---

**R code Exa 9.11** Confidence Interval Evaluation Using Pooled Estimate of Variance

```
1
2 # Chapter 9
3 # Example 9.11 page no. 288 from the pdf..
4 # COncidence Interval Evaluation Using Pooled
  Estimate of Variance..
5
6 # to find the 90% confidence interval of the given
  data
7
8 # the variance of both sampling stations are equal
  and and independent
9 n1 <- 12
10
11 x1 <- 3.11
12
13 s1 <- 0.771
14
15 n2 <- 10
16
17 x2 <- 2.04
18
19 s2 <- 0.448
20
21 # WE NEED to find pooled estimate for equal
  variances..
22
23 sp <- sqrt(((n1-1)*s1*s1+(n2-1)*s2*s2)/(n1+n2-2))
24
25 dfs <- n1+n2-2
```



```

26
27 c <- qt(0.95,dfs)*sp*sqrt((1/n1)+(1/n2))
28
29 cat("The 90% confidence interval for mu1-mu2 is",x1-
      x2-c,x1-x2+c)
30
31 # the answer in the textbook is approximated to 3
    decimal places..

```

---

**R code Exa 9.12** Confidence Interval Evaluation For difference of 2 means when their variance is not equal and Unknown

```

1
2 # Chapter 9
3 # Example 9.12 page no. 290 from the pdf..
4 # Confidence Interval Evaluation For difference of 2
    means when their variance is not equal and
    Unknown..
5
6 # given unequal variances ,independent samples ,approx
    . normal
7
8 a <- c(3.84,3.07,15) # parameters for 1st station
9
10 b <- c(1.49,0.80,12) #parameters for second station
11
12 # to find the 95% ci for mu1-mu2
13
14 v <- (((a[2]*a[2]/a[3] + b[2]*b[2]/b[3])^2)/((((a[2]
    *a[2]/a[3])^2)/(a[3]-1))+(((b[2]*b[2]/b[3])^2)/(b
    [3]-1))))
15
16 c <- qt(0.975,v)*sqrt((a[2]*a[2]/a[3]) + (b[2]*b[2]/
    b[3]))
17

```

```

18 cat("The 95% confidence interval for  $\mu_1 - \mu_2$  is", a
      [1] - b[1] - c, a[1] - b[1] + c, " milligram")
19
20
21 # The answer in T.B is rounded to 2 decimal places.

```

---

**R code Exa 9.13** To find Confidence Intervals for Paired Observations

```

1
2 # Chapter 9
3 # Example 9.13 page no. 293 from the pdf..
4 # to find Confidence Intervals for Paired
  Observations..
5
6 levels_plasma <- c
  (2.5,3.1,2.1,3.5,3.1,1.8,6.0,3.0,36.0,4.7,6.9,3.3,4.6,1.6,7.2,1.8
7
8 levels_fat <- c
  (4.9,5.9,4.4,6.9,7.0,4.2,10.0,5.5,41.0,4.4,7.0,2.9,4.6,1.4,7.7,1.
9
10
11 cat("The 95% confidence interval for paired
      observation is", " ")
12
13 t.test(x= levels_plasma, y= levels_fat, paired=TRUE,
        atalternative="two.sided")$conf.int

```

---

**R code Exa 9.14** Estimating a Proportion for Single Sample

```

1
2 # Chapter 9

```

```

3 # Example 9.14 page no. 297 from the pdf..
4 # Estimating a Proportion for Single Sample..
5
6
7 # to calculate the 95% confidence interval
8 # given
9 n <- 500 # size of random sample
10
11 x <- 340 # number subscribed to HBO from te size
      of random sample
12
13 cat("The confidence interval for the actual
      proportion of families with TV sets in the city
      subscribed to HBO is", " ")
14
15 prop.test(340,500,0.68)$conf.int

```

---

**R code Exa 9.15** Evaluate sample size for estimating a Proportion for Single Sample

```

1
2 # Chapter 9
3 # Example 9.15 page no. 299 from the pdf..
4 # Evaluate sample size for estimating a Proportion
      for Single Sample..
5
6 # to find the sample size required to be 95%
      confident that our estimate of p in ex9.14 is
      within 0.02 of the true value..
7 # Refer example 9.14 page no. 297 from the pdf..
8
9 n <- 500
10
11 x <- 340
12

```

```

13 p_hat <- x/n
14
15 e <- 0.02
16
17 sample_size <- qnorm(0.975)^2*p_hat*(1-p_hat)/(e^2)
18
19 cat("If we base our estimate of p on a random sample
      of size",as.integer(sample_size)+1,"we can be
      95% confident that our sample proportion will not
      differ from the true proportion by more than
      0.02")

```

---

**R code Exa 9.16** Evaluate sample size for estimating a Proportion for Single Sample

```

1
2 # CHapter 9
3 # Example 9.16 page no. 299 from the pdf..
4 # Evaluate sample size for estimating a Proportion
  for Single Sample..
5
6 # to find a sample size if we want atleast 95%
  confidence interval
7
8
9 e <- 0.02      # error not exceeding this value
10
11 cat("The sample size if we want atleast 95%
      confidence interval that our estimate of p is
      within 0.02 of the true value is",round(qnorm
      (0.025)*qnorm(0.025)/(4*e*e))
12
13
14 # In Example 9.17 all we have to do is put the
    values in the formula mentioned in the pdf and

```

get the answer,I have solved some questions based on this but I am not going to solve , that's understood from the questions itself and will be a very straight forward code..

---

**R code Exa 9.18** To obtain Confidence Interval of Variance

```
1
2 # Chapter 9
3 # Example 9.18 page no. 304 from the pdf..
4 # to obtain confidence interval of variance
5
6 a <- c
   (46.4,46.1,45.8,47.0,46.1,45.9,45.8,46.9,45.2,46.0)
7
8
9 c <- qchisq(0.025,df = length(a)-1)
10 b <- qchisq(0.975,df= 9)
11
12 cat("The 95% confidence interval for variance is",
     var(a)*(length(a)-1)/b, var(a)*(length(a)-1)/c)
```

---

**R code Exa 9.19** To find Confidence Interval For the Ratio of 2 Variances

```
1
2 # Chapter 9
3 # Example 9.19 page no. 306 from the pdf..
4 # To find Confidence Interval For the Ratio of 2
   Variances..
5
6 # to find 98% confidence interval of 2 independent
   sample variances..
```

```

7  n1 <- 15
8
9  n2 <- 12
10
11 s1 <- 3.07
12
13 s2 <- 0.80
14
15 f1 <- qf(0.99,14,11)
16
17 f2 <- qf(0.99,11,14)
18
19 cat("The 98% confidence interval for the ratio
      signal/sigma2 is",sqrt(s1*s1/(s2*s2*f1)),sqrt(s1*
      s1*f2/(s2*s2)))

```

---

#### R code Exa 9.22 Maximum Likelihood Estimation Problem

```

1
2 # Chapter 9
3 # example 9.22 page no. 311 from the pdf..
4 # Maximum Likelihood Estimation Problem..
5
6 # Note – Example 9.20 and 9.21 are theoretical with
  nothing to Compute..
7
8 # given survival time in months for 10 rats , to find
  the maximum estimate of the mean survival time..
9 # package used "stats4", if it is not installed on
  your studio then you need to execute the
  following instruction – install.packages("stats4
  ")
10
11 library(stats4)
12

```

```

13 y <- c(14,17,27,18,12,8,22,13,19,12)
14
15 f <- function(lambda){ -sum(dpois(y, lambda, log =
    TRUE))}
16
17 ans <- mle(f, start = list(lambda = 1), nobs = NROW(
    y))
18
19 print("The maximum likelihood estimate of the mean
    survival time is")
20
21 coef(ans)
22
23 # same is the concept for example 9.23 just we have
    to change the function and vector of observations
    ..

```

---

## Chapter 10

# One And Two Sample Tests Of Hypothesis

**R code Exa 10.3** Test Concerning a Single Mean for Single Sample

```
1
2 # Chapter 10
3 # Example 10.3 page no. 338 from the pdf..
4 # Test Concerning a Single Mean for Single Sample..
5
6 # Note – 10.1 and 10.2 are theoretical with nothing
  to Compute..
7
8 # given
9 m <- 71.8
10
11 n <- 100
12
13 sd <- 8.9
14
15 alpha <- 0.05
16
17 # Null hypothesis , mean = 70 years
18 # alternative mean > 70 years
```



```

19
20 sd1 <- sd/sqrt(n)
21
22 p_value <- 1 - pnorm((m-70)/sd1,0,1)
23
24 cat("The p-value is",p_value,"which is less than
      0.05 level of significance hence we conclude the
      evidence favour alternative hypothesis more than
      null")

```

---

#### **R code Exa 10.4** Test Concerning a Single Mean for Single Sample

```

1
2 # Chapter 10
3 # Example 10.4 page no. 338 from the pdf..
4 # Test Concerning a Single Mean for Single Sample..
5
6 # given
7 # H0 = 8kg
8 # H1 != 8kg
9
10 alpha <- 0.01
11
12
13 n <- 50
14
15 sd <- 0.5
16
17 m <- 7.8
18
19 sd1 <- sd/sqrt(n)
20
21 # two sided hypothesis test..
22
23 p_value <- 2*pnorm((m-8)/sd1)

```

```

24
25 if(p_value < 0.01){
26
27     print("Since the p_value is less than 0.01 we
        reject the null hypothesis that mu= 8kg")
28
29 }else{
30
31     print("Since the p_value is more than 0.01 there
        is not enough evidence to claim the alternative
        hypothesis , so evidence supports mu can be 8")
32
33 }

```

---

**R code Exa 10.5** The t statistic for a test on a single mean when Variance is Unknown

```

1
2 # Chapter 10
3 # Example 10.5 page no. 340 from the pdf..
4 # The t-statistic for a test on a single mean when
   Variance is Unknown..
5
6 # given H0- mean = 46kWh
7 # H1 - mean < 46kWh
8
9 alpha <- 0.05
10
11 n <- 12
12
13 # using t-staistic
14
15 m <- 42
16
17 sd <- 11.9

```

```

18
19 sd1 <- sd/sqrt(n)
20
21 # one sided testing
22 p_value <- pt((m-46)/sd1,n-1)
23
24
25 if(p_value < 0.05){
26
27     print("Since the p_value is less than 0.05 we
           reject the null hypothesis that mu= 46kWh")
28
29 }else{
30
31     print("Since the p_value is more than 0.05, we
           cannot reject H0 and conclude that the average
           no. of kWh used is not significantly less than
           46")
32
33 }

```

---

#### R code Exa 10.6 The 2 Sample Pooled t test

```

1
2 # Chapter 10
3 # Example 10.6 page no. 344 from the pdf..
4 # The 2 Sample Pooled t-test..
5
6 # 2 sample t- test
7 n1 <- 12
8
9 n2 <- 10
10
11 x1 <- 85
12

```

```

13 x2 <- 81
14
15 s1 <- 4
16
17 s2 <- 5
18
19 # Null H0-  $\mu_1 - \mu_2 = 2$ 
20 # alternate H1  $\mu_1 - \mu_2 > 2$ 
21 # using 2 sample pooled t-test..
22
23 deg_f <- n1+n2-2
24
25 sp <- sqrt((s1*s1*(n1-1)+s2*s2*(n2-1))/deg_f)
26
27 p_value <- 1-pt(((x1-x2-2)/(sp*sqrt(1/n1+1/n2))),deg
    _f)
28
29 if(p_value < 0.05){
30
31     print("Since the p_value is less than 0.05 we
        reject the null hypothesis that  $\mu_1 - \mu_2 = 2$ ")
32
33 }else{
34
35     print("Since the p_value is more than 0.05, we
        cannot reject H0 and conclude that material 1
        wear exceeds that of 2 by more than 2 units")
36
37 }

```

---

#### R code Exa 10.7 Choice of Sample Size for Testing Mean

```

1
2 # Chapter 10
3 # Example 10.7 page no. 351 from the pdf..

```

```

4 # Choice of Sample Size for Testing Mean..
5
6 # given the hypothesis test, using alpha= 0.05,
7 # to find the sample size if power of our test is to
   be 0.95 when true mean is 69 kg.
8 # Null - mu = 68 kg.
9 # alternative - mu > 68 kg.
10
11 mu = 68
12
13 mu1 = 69
14
15 alpha = 0.05
16
17 sig <- 5
18
19 delta <- mu1-mu
20
21 z_alpha <- abs(qnorm(0.05))
22
23 n <- ((z_alpha+z_alpha)^2)*(sig*sig)/(delta*delta)
24
25 cat("The sample size required if the test is to
   reject the null hypothesis 95% of the time when
   in fact, mu is as large as 69 kg is",n)
26
27 # example 10.8 is bit absurd all they have done is
   mere division and making conclusions based on
   that so I am not coding example 10.8

```

---

### R code Exa 10.9 One Sample Test on a Single Proportion

```

1
2 # Chapter 10
3 # Example 10.9 page no. 362 from the pdf..

```

```

4 # One Sample Test on a Single Proportion..
5
6 # given H0 - p=0.7
7 # H1 alternative - p!=0.7
8
9 # test staistic - binomial
10
11 x <- 8    # no. of heat pumps installed in sample
12
13 n <- 15   # no. of randomly selected homes..
14
15 # it is a binomial variable and 2 sided hypothesis
    test..
16 p_value <- pbinom(x,n,0.7)*2
17
18 if(p_value > 0.10){
19
20     cat("As the p value is",p_value,"there is
        insufficient reason to doubt the builder's
        claim so cannot reject null hypothesis")
21
22 }else{
23
24     print("The evidence supports alternative
        hypothesis and builder claim should be rejected
        ")
25
26 }

```

---

#### R code Exa 10.10 One Sample Test on a Single Proportion

```

1
2 # Chapter 10
3 # Example 10.10 page no. 362 from the pdf..
4 # One Sample Test on a Single Proportion..

```

```

5
6 # Null - p = 0.6
7 # alternative - p > 0.6
8 # alpha = 0.05.
9 # critical region - z>1.645
10
11 # using binomial approximation..
12
13
14 ans <- pnorm(0.7,mean = 0.6,sd=sqrt(0.6*0.4/100),
15           lower.tail = F)
16 cat("The p_value is",ans,"we reject null hypothesis
17     and conclude that the new drug is superior")

```

---

#### R code Exa 10.11 Two Sample Tests On 2 Proportions

```

1
2 # Chapter 10
3 # Example 10.11 page 364 from the pdf..
4 # Two Sample Tests On 2 Proportions..
5
6 # hypothesis testing on 2 proportion..using z test
7 # Null- H0 p1=p2
8 # alternate - H1 = p1>p2
9
10 # information -
11 x1 <- 120
12
13 x2 <- 240
14
15 n1 <- 200
16
17 n2 <- 500
18

```

```

19 # to find the estimate..
20 p_hat <- c(x1/n1,x2/n2)
21
22 pooled_p <- (x1+x2)/(n1+n2)
23
24 p_value <- 1-pnorm((p_hat[1]-p_hat[2])/(sqrt((pooled
    _p*0.49)*(1/n1+1/n2))))
25
26 cat("Since the p_value is",p_value,"we reject the
    null hypothesis and conclude that town voters
    favouring the proporsal are higher than county
    voters")
27
28 # this problem can also be solved directly using
    prop.test(c(120,240),c(200,500),alternative="
    greater"), but this test uses chi square test
    method and the T.B solved it with z-test..
29 # If you want to explore it write this code into
    your console -
30 # prop.test(c(120,240),c(200,500),alternative = "
    greater")

```

---

### R code Exa 10.12 One and Two Sample Tests Concerning Variance

```

1
2 # Chapter 10
3 # Example 10.12 page no. 366 from the pdf..
4 # One and Two Sample Tests Concerning Variance..
5
6 # null - sig^2 = 0.81
7 # alternaitve - sig^2 > 0.81
8 # alpha = 0.05
9
10 n <- 10
11

```



```

12 sig <- 0.9
13
14 s <- 1.2
15
16 chi_sq <- (n-1)*s*s/(sig*sig)
17
18 p_value <- pchisq(chi_sq,df=9 ,lower.tail = F)
19
20 cat("The p value is",p_value,"there is evidence that
      sig > 0.9")

```

---

**R code Exa 10.13** To do Hypothesis Testing for 2 variances

```

1
2 # Chapter 10
3 # example 10.13 page no. 368 from the pdf..
4 # To do Hypothesis Testing for 2 variances..
5
6 # refer example 10.6 for value of 2 variances
7 # Null –  $\sigma_1^2 = \sigma_2^2$ 
8 # alternative –  $\sigma_1^2 \neq \sigma_2^2$ 
9
10 s1_sq <- 16
11
12 s2_sq <- 25
13
14
15 # null reject if  $f < 0.34$  or  $f > 3.11$ ..
16 ifelse(s1_sq/s2_sq < qf(0.05,11,9) | s1_sq/s2_sq >
      qf(0.95,11,9),print("The null hypothesis is
      rejected and conclude there is sufficient
      evidence that variance differ"),print("Do not
      reject null hypothesis and conclude that there is
      insufficient evidence that the variances differ.
      "))

```

---

**R code Exa 10.14** Test For Homogeneity

```
1
2 # Chapter 10
3 # Example 10.14 page no. 377 from the pdf..
4 # Test For Homogeneity..
5
6 # given data test the hypothesis
7
8 # H0 – proportion of all political affiliation are
   same for each opinion
9
10 # H1 – for at least 1 opinion the proportion is not
    same
11
12 # using chi-sq statistic..
13
14 for1 <- c(82,70,62)
15
16 against <- c(93,62,67)
17
18
19 undecided <- c(25,18,21)
20
21 pol_affi <- as.data.frame(rbind(for1,against,
   undecided))
22
23 names(pol_affi) <- c('Democrat','Republican','
   Independent')
24
25 chisq.test(pol_affi)
26
27 cat("The p-value is",chisq.test(pol_affi)$p.value,"
   so we cannot reject null hypothesis")
```

28

29 # the answer to the  $\chi^2$  is approximated to 1.53 as  
computed 1.5274 by R..

---

**R code Exa 10.15** Testing for several proportion

```
1
2 # Chapter 10
3 # Example 10.15 page no. 378 from the pdf..
4 # testing for several proportion..
5
6 defectives <- c(45,55,70)
7
8 nondefectives <- c(905,890,870)
9
10 # making a data frame of Variables..
11 shifts <- as.data.frame(rbind(defectives,
12                                nondefectives))
13
14 # Giving names to the columns..
15 names(shifts) <- c('Day', 'Evening', 'Night')
16
17 chisq.test(shifts)
18
19 cat("Since the p-value is",chisq.test(shifts)$p.
20     value," grater than the significant value 0.025
21     we cannot reject null hypothesis but from the p-
22     value calculated we also cannot say that the
23     proportion of defectives is same ")
```

---

# Chapter 11

## Simple Linear Regression And Correlation

**R code Exa 11.1** To Estimate the Regression Line

```
1
2 # Chapter 11
3 # Example 11.1 page no. 396 from the pdf..
4 # To Estimate the Regression Line..
5
6
7 # to estimate the regression line from the given
  data..
8 x <- c
  (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,
9
10 y <- c
  (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36
11
12 pol <- data.frame(x,y)
13
14 line_eq <- lm(y~x,data=pol)
```

```

15
16 b <-coefficients(line_eq)
17
18 cat("The estimate of regression line is y =",b[1],"+"
      ",b[2],"*x")

```

---

### R code Exa 11.2 Confidence Interval for Slope of Regression Line

```

1
2 # Chapter 11
3 # Example 11.2 page no. 403 from the pdf..
4 # Confidence Interval for Slopes of Regression Line
5   ..
6 # to find the confidence interval of the given data
7   ..
8 x <- c
   (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,
9
10 y <- c
   (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36
11
12 pol <- data.frame(x,y)
13
14 line <- lm(y~x,data = pol)
15
16 cat("The 95% confidence interval for beta1(i.e slope
      ) in the regression line is",confint(line,'x',
      level=0.95))
17
18 # The answer matches with the answer in the T.B upto
      3 decimal places , as T.B answer is approximated

```

..

---

### R code Exa 11.3 Hypothesis Testing On Slope Of Regression Line

```
1
2 # Chapter 11
3 # Example 11.3 page no. 404 from the pdf..
4 # Hypothesis Testing On Slope Of Regression Line..
5
6 # to test the hypothesis of beta1 slope of
  regression line..
7 # Null – beta1 = 1.0, alternative < 1.0
8
9 x <- c
    (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,
10
11 y <- c
    (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36
12
13 pol <- data.frame(x,y)
14
15 l <- lm(y~x,data = pol)
16
17 coefficients(l)
18
19 p_value <- function(reg_m,conum,val){
20
21   coefi <- coef(summary(reg_m))
22   t <- (coefi[conum,1]-val)/coefi[conum,2]
23   pt(abs(t),reg_m$df.residual,lower.tail= F)
24
25 }
26 cat("Since the p value is",p_value(1,2,1)," less
```

```

    than 0.05 suggesting strong evidence that  $\beta_1 < 1.0$ ")
27
28
29 # an alternate method is to use the library car and
    the use linearHypothesis function, since this is
    one sided hypothesis
30 # we need to divide the result by 2 as the function
    does 2 sided hypothesis..
31 # If "car" package not installed can be installed by
    install.packages("car")
32
33 library(car)
34 linearHypothesis(l,hypothesis.matrix = c(0,1),rhs =
    1)/2
35 print("Using linearHypothesis function also we get
    the same result and we can derive the same
    inference as before.")

```

---

#### R code Exa 11.4 Confidence Interval For Intercept Of Regression Line

```

1
2 # Chapter 11
3 # Example 11.4 page no. 406 from the pdf..
4 # Confidence Interval For Intercept Of Regression
    Line..
5
6 # to find the confidence interval of the given data
    ..
7
8 x <- c
    (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,
9
10 y <- c

```

```

(5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36

11
12 pol <- data.frame(x,y)
13
14 line <- lm(y~x,data = pol)
15
16 cat("The 95% confidence interval for beta0(i.e
    intercept) in the regression line is",confint(
    line,'(Intercept)',level=0.95))
17
18 # The answer matches with the answer in the T.B upto
    2 decimal places , as T.B answer is approximated
    ..

```

---

#### R code Exa 11.5 Hypothesis Testing on Intercept of Regression line

```

1
2 # Chapter 11
3 # Example 11.5 page no. 407 from the pdf..
4 # Hypothesis Testing on Intercept of Regression line
    ..
5
6 # Null – beta0 =0
7 # alternate – beta0 != 0
8
9 x <- c
    (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,
10
11 y <- c
    (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36
12
13 pol <- data.frame(x,y)    # making data frame of

```



```

      observations..
14
15 line <- lm(y~x,data = pol) # regression model..
16
17 a <- summary(line) # summary of the regression line
18
19 cat("The p-value is",a$coef[1,"Pr(>|t|)"],"which is
      less than 0.05 hence we conclude that beta0 != 0"
      )

```

---

**R code Exa 11.6** To find the confidence interval of the mean response

```

1
2 # Chapter 11
3 # Example 11.6 page no. 409 from the pdf..
4 # To find the confidence interval of the mean
  response..
5
6 x <- c
    (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,
7
8 y <- c
    (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36
9
10 pol <- data.frame(x,y)
11
12 line <- lm(y~x,data = pol)
13
14 # here x0 <- 20%
15 data <- data.frame(x=20)
16
17 c <- predict(line,data, interval = "confidence")
18

```

```

19 cat("When solid reduction is 20%, the confidence
    interval of chemical oxygen demand is",c[2],c[3])
20
21 # the may vary slightly due to approximations in T.B
    .

```

---

### R code Exa 11.7 To Find Prediction Interval

```

1
2 # Chapter 11
3 # Example 11.7 page no. 411 from the pdf..
4 # To Find Prediction Interval..
5
6 # To find the 95% prediction interval for y0 when x0
    is 20%
7
8 x <- c
    (3,7,11,15,18,27,29,30,30,31,31,32,33,33,34,36,36,36,37,38,39,39,
9
10 y <- c
    (5,11,21,16,16,28,27,25,35,30,40,32,34,32,34,37,38,34,36,38,37,36
11
12 pol <- data.frame(x,y)
13
14 line <- lm(y~x,data = pol)
15
16
17 # making a new data frame to predict..
18
19 data <- data.frame(x=20)
20
21 d <- predict(line,data, interval = "prediction")
22

```

```

23 cat("The 95% prediction interval for y0 when x0=20%
    is",d[2],d[3])
24
25
26 # the answer may vary slightly due to approximation
    ..

```

---

### R code Exa 11.8 Computation of Lack of Fit Sum of Squares

```

1
2 # Chapter 11
3 # Example 11.8 page no. 420 from the pdf..
4 # Computation of Lack of Fit Sum of Squares..
5
6 # to do lack of fit test
7
8 # to do lack of fit test I used the package "alr3",
    referance - Internet
9 # using only anova shows the regression and error ,
    not lack of fit and pure error so I used this
    package
10
11 install.packages("alr3") # Remove it if you have
    already Installed the package..
12
13 library("alr3")
14
15 y <- c
    (77.4,76.7,78.2,84.1,84.5,83.7,88.9,89.2,89.7,94.8,94.7,95.9)
16
17 x <- c
    (150,150,150,200,200,200,250,250,250,300,300,300)
18
19 dat <- data.frame(y,x)

```

```

20
21 c <- lm(y~x,data=dat)
22
23 cat("The following table shows analysis of variance
      on temperature field data")
24
25 pureErrorAnova(c)

```

---

### R code Exa 11.9 Transformation To Linear Regression Model

```

1
2 # Chapter 11
3 # Example 11.9 page no. 426 from the pdf..
4 # Transformation To Linear Regression Model..
5
6 # given data on P and V find the constants of the
   equation  $PV^{\gamma} = C$ 
7
8 v <- c(50,60,70,90,100) #volume
9
10 p <- c(64.7,51.3,40.5,25.9,7.8) # pressure
11
12 # the model
13 l <- lm(log(p)~log(v))
14
15 # coefficients of the model..
16 co <- coefficients(l)
17
18 # we modelled the data as  $\ln P = \ln C - \gamma \ln V +$ 
   additive error.
19
20 cat("The value of C is",exp(co[1]),"and the vlaue of
      gamma is",-co[2])

```

---

**R code Exa 11.10** To find Correlation Coefficient

```
1
2 # Chapter 11
3 # Example 11.10 page no. 433 from the pdf..
4 # To find Correlation Coefficient..
5
6 # to find the correlation coefficient between the 2
  variables
7
8 spec_gra <- c
  (0.414,0.383,0.399,0.402,0.442,0.422,0.466,0.500,0.514,0.530,0.560)
9
10 mod_of_rup <- c
  (29186,29266,26215,30162,38867,37831,44576,46097,59698,67705,66080)
11
12 cat("The sample correlation coefficient is",cor(spec
  _gra,mod_of_rup))
```

---

**R code Exa 11.11** To do Hypothesis Testing of linear association between two variables

```
1
2 # Chapter 11
3 # Example 11.11 page no. 434 from the pdf..
4 # To do Hypothesis Testing of linear association
  between two variables
5
6 # Null H0:  $\rho = 0$ 
7 # alternate H1 :  $\rho \neq 0$ 
```

```

8 # alpha = 0.05
9
10 spec_gra <- c
    (0.414,0.383,0.399,0.402,0.442,0.422,0.466,0.500,0.514,0.530,0.563)
11
12 mod_of_rup <- c
    (29186,29266,26215,30162,38867,37831,44576,46097,59698,67705,66083)
13
14 print("The test is shown below as follows:")
15
16 cor.test(spec_gra,mod_of_rup)
17
18 cat("The p-value is very less then 0.05, we reject
    the null hypothesis of no linear association")
19
20 # example 11.12 is same as example 11.11 except now
    we have to do hypothesis as  $p_0 = 0.9$ , just
    changed the value for testing, so It is same
    concept just numbers are changed, so I have not
    solved..

```

---

## Chapter 12

# Multiple Linear Regression And Certain Non Linear Regression Models

**R code Exa 12.1** Estimation Of Equation Of Regression Line and predict values

```
1
2 # Chapter 12
3 # Example 12.1 page no. 445 from the pdf..
4 # Estimation Of Equation Of Regression Line..
5
6 y <- c
   (0.90,0.91,0.96,0.89,1.00,1.10,1.15,1.03,0.77,1.07,1.07,0.94,1.10
7
8 x1 <- c
   (72.4,41.6,34.3,35.1,10.7,12.9,8.3,20.1,72.2,24.0,23.2,47.4,31.5,
9
10 x2 <- c
    (76.3,70.3,77.1,68.0,79.0,67.4,66.8,76.9,77.7,67.7,76.8,86.6,76.9
```

```

11
12 x3 <- c
    (29.18,29.35,29.24,29.27,29.78,29.39,29.69,29.48,29.09,29.60,29.30)

13
14 dat <- data.frame(y,x1,x2,x3)
15
16 line <- lm(y~x1+x2+x3,data=dat)
17
18 c <- coefficients(line)
19
20 cat("The regression estimate of the above data is",c
    [1],"+",c[2],"*x1 +",c[3],"*x2 +",c[4],"*x3")
21
22 pre <- data.frame(x1 = 50,x2 = 76.0,x3 = 29.30)
23
24 # Prediction of values..
25 cat("The prediction of the line for this values is")
26
27 predict(line,pre)

```

---

**R code Exa 12.2** To find the estimate of polynomial regression

```

1
2 # Chapter 12
3 # Example 12.2 page no. 446 from the pdf..
4 # to find the estimate of polynomial regression..
5
6 x <- c(0,1,2,3,4,5,6,7,8,9)
7
8 y <- c(9.1,7.3,3.2,4.6,4.8,2.9,5.7,7.1,8.8,10.2)
9
10 pol <- lm(y~x+I(x^2))
11
12 c <- coefficients(pol)

```



```

13
14 dat <- data.frame(x=2)
15
16
17 cat("The fitted polynomial is",c[1],c[2],"*x +",c
    [3],"*x^2")
18
19 cat("When x=2, the estimate is",predict(pol,dat))

```

---

**R code Exa 12.3** Estimate the regression coefficients in polynomial model

```

1
2 # Chapter 12
3 # Example 12.3 page no. 447 from the pdf..
4 # Estimate the regression coefficients in polynomial
  model.
5
6 m1 <- c(14.05,14.93,16.56,15.85,22.41,21.66)
7
8 m2 <- c(10.55,9.48,13.63,11.75,18.55,17.98)
9
10 m3 <- c(7.55,6.59,9.23,8.78,15.93,16.44)
11
12 observation <- c(m1,m2,m3)
13
14 temp <- c(rep(75,6),rep(100,6),rep(125,6))
15
16 m <- c(rep(15,2),rep(20,2),rep(25,2))
17
18 ster_time <- rep(m,3)
19
20 dat <- data.frame(observation,temp,ster_time)
21
22 pol <- lm(observation ~ temp + ster_time + I(temp^2)
    + I(ster_time^2) + temp*ster_time,data = dat)

```

```

23
24 c <- coefficients(pol)
25
26
27 cat("The fitted polynomial is",c[1],c[2],"*x1",c[3],
      "*x2 +",c[4],"*x1^2 +",c[5],"*x2^2 +",c[6],"*x1*
      x2")

```

---

#### R code Exa 12.4 Linear regression model using matrices

```

1
2 # Chapter 12
3 # Example 12.3 page no. 449 from the pdf..
4 # Linear regression model using matrices..
5
6 # given - % survival(y) of certain type of animal
7 # combinations of concentration of 3 materials used
8 # estimate the linear regression model using the
9 # data by matrices form.
10 y <- c
11     (25.5,31.2,25.9,38.4,18.4,26.7,26.4,25.9,32.0,25.2,39.7,35.7,26.5)
12
13
14 x1 <- c
15     (1.74,6.32,6.22,10.52,1.19,1.22,4.10,6.32,4.08,4.15,10.15,1.72,1.7)
16
17
18 x2 <- c
19     (5.30,5.42,8.41,4.63,11.60,5.85,6.62,8.72,4.42,7.60,4.83,3.12,5.3)
20
21
22 x3 <- c

```

```

(10.80,9.40,7.20,8.50,9.40,9.90,8.00,9.10,8.70,9.20,9.40,7.60,8.20,
17
18 # now forming the matrix ,
19
20
21 xxdash <- matrix(c(length(x1),sum(x1),sum(x2),sum(x3
    ),sum(x1),
22                    sum(x1*x1),sum(x1*x2),sum(x1*x3),
23                    sum(x2),sum(x2*x1),sum(x2*x2),sum
    (x2*x3),
24                    sum(x3),sum(x3*x1),sum(x3*x2),sum
    (x3*x3)),byrow = T,nrow = 4)
25
26 # now forming the matrix xdash_y
27
28
29 xdash_y <- matrix(c(sum(y),sum(x1*y),sum(x2*y),sum(
    x3*y)),byrow = T,nrow = 4)
30
31 # now solving the matrix equation xxdash*b = xdash_y
    using solve function in R
32
33 sol <- solve(xxdash,xdash_y)
34
35 # final answer..
36 cat("The regression line is",sol[1,1],"+",sol[2,1],
    *x1",
37     sol[3,1],"*x2",sol[4,1],"*x3")

```

---

**R code Exa 12.5** To test the hypothesis on the slope of the regression model

```

1
2 # Chapter 12

```

```

3 # example 12.5 page no. 456 from the pdf..
4 # to test the hypothesis on the slope of the
   regression model..
5
6 # Null - beta2 = -2.5
7 # alternate > -2.5
8
9 # solving example 12.4 again..
10 # linear regression model using matrices..
11 # given - % survival(y) of certain type of animal
   semen after storage..
12 # combinations of concentration of 3 materials used
   to increase the chance of survival..
13 # estimate the linear regression model using the
   data by matrices form.
14
15 y <- c
   (25.5,31.2,25.9,38.4,18.4,26.7,26.4,25.9,32.0,25.2,39.7,35.7,26.5)
16
17 x1 <- c
   (1.74,6.32,6.22,10.52,1.19,1.22,4.10,6.32,4.08,4.15,10.15,1.72,1.7)
18
19 x2 <- c
   (5.30,5.42,8.41,4.63,11.60,5.85,6.62,8.72,4.42,7.60,4.83,3.12,5.3)
20
21 x3 <- c
   (10.80,9.40,7.20,8.50,9.40,9.90,8.00,9.10,8.70,9.20,9.40,7.60,8.2)
22
23 # now forming the matrix,
24
25
26 xxdash <- matrix(c(length(x1),sum(x1),sum(x2),sum(x3
   ),sum(x1),
27                   sum(x1*x1),sum(x1*x2),sum(x1*x3),

```

```

28             sum(x2),sum(x2*x1),sum(x2*x2),sum
                (x2*x3),
29             sum(x3),sum(x3*x1),sum(x3*x2),sum
                (x3*x3)),byrow = T,nrow = 4)
30
31 # now forming the matrix xdash_y
32
33
34 xdash_y <- matrix(c(sum(y),sum(x1*y),sum(x2*y),sum(
    x3*y)),byrow = T,nrow = 4)
35
36 # now solving the matrix equation xxdash*b = xdash_y
    using solve function in R
37
38 sol <- solve(xxdash,xdash_y)
39
40 unit <- matrix(c(c(1,0,0,0),c(0,1,0,0),c(0,0,1,0),c
    (0,0,0,1)),ncol = 4)
41
42
43 # finding inverse of the matrix xxdash..
44 inv_xxdash <- solve(xxdash,unit)
45
46 # final answer..
47 cat("The regression line is",sol[1,1],"+",sol[2,1],
    *x1",
48     sol[3,1],"*x2",sol[4,1],"*x3")
49
50 # the coefficient of x2 (i.e beta2) is -1.861649 as
    we can see from the solution..
51 # now testing our hypothesis..
52
53 t.score <- (sol[3,1]+2.5)/(2.073*sqrt(inv_xxdash
    [3,3]))
54
55 cat("The p-value is",pt(t.score,9,lower.tail = F),"
    we reject the null hypothesis and conclude beta2
    > -2.5")

```

```

56
57 # the answer in the T.B is double the value obtained
    by computation ,
58 #I think they took both sided alternative , but we
    have to take only one sided alternative as our
    alternative is  $\beta_2 > -2.5$ ,
59 #hence it is one sided and so T.B answer is wrong..

```

---

### **R code Exa 12.6** To Construct Confidence Interval for Mean Response

```

1
2 # Chapter 12
3 # Example 12.6 page no. 457 from the pdf..
4 # To Construct Confidence Interval for Mean Response
    ..
5
6 y <- c
    (25.5,31.2,25.9,38.4,18.4,26.7,26.4,25.9,32.0,25.2,39.7,35.7,26.5)
7
8 x1 <- c
    (1.74,6.32,6.22,10.52,1.19,1.22,4.10,6.32,4.08,4.15,10.15,1.72,1.7)
9
10 x2 <- c
    (5.30,5.42,8.41,4.63,11.60,5.85,6.62,8.72,4.42,7.60,4.83,3.12,5.3)
11
12 x3 <- c
    (10.80,9.40,7.20,8.50,9.40,9.90,8.00,9.10,8.70,9.20,9.40,9.60,8.2)
13
14 dat <- data.frame(y,x1,x2,x3)
15
16 c <- lm(y~x1+x2+x3,data= dat)

```

```

17
18 newdat <- data.frame(x1=3,x2=8,x3=9)
19
20 d <- predict(c,newdat,interval = "confidence",level
    = 0.95,type = "response")
21
22 cat("The 95% confidence interval for the mean
    response when x1=3%,x2=8% and x3=9% is",d[2],d
    [3])
23
24 # the answer may vary slightly from the T.B

```

---

#### R code Exa 12.7 Prediction interval Evaluation

```

1
2 # Chapter 12
3 # Example 12.7 Page no. 458 from the pdf..
4 # Prediction interval Evaluation..
5
6 y <- c
    (25.5,31.2,25.9,38.4,18.4,26.7,26.4,25.9,32.0,25.2,39.7,35.7,26.5)
7
8 x1 <- c
    (1.74,6.32,6.22,10.52,1.19,1.22,4.10,6.32,4.08,4.15,10.15,1.72,1.7)
9
10
11 x2 <- c
    (5.30,5.42,8.41,4.63,11.60,5.85,6.62,8.72,4.42,7.60,4.83,3.12,5.3)
12
13 x3 <- c
    (10.80,9.40,7.20,8.50,9.40,9.90,8.00,9.10,8.70,9.20,9.40,9.60,8.2)

```

```

14
15 dat <- data.frame(y,x1,x2,x3)
16
17 c <- lm(y~x1+x2+x3,data= dat) #modelling the
    variables of datasaet..
18
19
20 newdat <- data.frame(x1=3,x2=8,x3=9)
21
22 d <- predict(c,newdat,interval = "prediction")
23
24 cat("The 95% prediction interval for the individual
    response when x1=3%,x2=8% and x3=9% is",d[2],d
    [3])

```

---

**R code Exa 12.8** To show Analysis of Variance Table for Grain Radius Data

```

1
2 # Chapter 12
3 # Example 12.8 page no. 470 from the pdf..
4 # To show Analysis of Variance Table for Grain
    Radius Data
5
6 pow_temp <- c(150,190,150,150,190,190,150,190)
7 ext_rate <- c(12,12,24,12,24,12,24,24)
8 die_temp <- c(220,220,220,250,220,250,250,250)
9
10
11
12 gra_radius <- c(82,93,114,124,111,129,157,164)
13
14 dat <- data.frame(pow_temp,ext_rate,die_temp,gra_
    radius)
15

```



```

16 # converting into factors..
17 dat$pow_temp <- factor(dat$pow_temp)
18 dat$ext_rate <- factor(dat$ext_rate)
19 dat$die_temp <- factor(dat$die_temp)
20
21
22 l <- lm(gra_radius~pow_temp+ext_rate+die_temp,data =
      dat)
23
24 # analysis of variance table for grain radius data..
25 print("The analysis of variance table for this data
      is:")
26 summary(aov(l,data= dat))

```

---

**R code Exa 12.9** To find the model for data which have 3 levels in a variable

```

1
2 # Chapter 12
3 # Example 12.9 page no. 474 from the pdf..
4 # to find the model for data which have 3 levels in
  a variable..
5
6 y <- c
  (292,329,352,378,392,410,198,227,277,297,364,375,167,225,247,268,
7
8 x <- c
  (6.5,6.9,7.8,8.4,8.8,9.2,6.7,6.9,7.5,7.9,8.7,9.2,6.5,7.0,7.2,7.6,
9
10 p <- c(1,1,1,1,1,1,2,2,2,2,2,2,3,3,3,3,3,3)
11
12 dat <- data.frame(y,x,factor(p))
13

```

```

14 dat$factor.p. <- relevel(dat$factor.p.,ref=3)
15
16 c <- lm(y ~ x +factor.p.,data= dat)
17
18 cat("The model suggests")
19
20 coefficients(c) # Coefficients of the model..
21
22 summary(c) # Summary of the regression model..

```

---

**R code Exa 12.10** To predict all possible regression lines

```

1
2 # Chapter 12
3 # Example 12.10 page no. 477 from the pdf..
4 # to predict all possible regression lines
5
6 # refer ex10 page no. 478 on pdf.
7
8 # using leaps library if not installed can be done
  by
9 # install.packages("leaps")
10
11 # using sequential replacement method..
12
13 library(leaps)
14
15 y <- c(57.5,52.8,61.3,67.0,53.5,62.7,56.2,68.5,69.2)
16
17 x1 <- c(78,69,77,88,67,80,74,94,102)
18
19 x2 <- c
    (48.2,45.5,46.3,49.0,43.0,48.0,48.0,53.0,58.0)
20
21 x3 <- c

```

```

      (2.75,2.15,4.41,5.52,3.21,4.32,2.31,4.30,3.71)
22
23
24 x4 <- c
      (29.5,26.3,32.2,36.5,27.2,27.7,28.3,30.3,28.7)
25
26
27 dat <- data.frame(y,x1,x2,x3,x4)
28
29 s <- regsubsets(y~.,data = dat,method = "seqrep")
30
31 print("All possible regression line equation
      coefficients are")
32
33 coef(s,1:4)
34
35 # for final model we need to see adj r squared value
      ..this is just the coefficients for various
      variables in the model..
36
37 # as nothing specific is asked here so I have shown
      what would be the model for all combination of
      variables..

```

---

**R code Exa 12.11** To model the data set using forward selection

```

1
2 # Chapter 12
3 # Example 12.11 page no. 480 from the pdf..
4 # to model the data set using forward selection..
5
6 # the data set from example 10 is used for modelling
      ..
7 y <- c(57.5,52.8,61.3,67.0,53.5,62.7,56.2,68.5,69.2)
8

```

```

9  x1 <- c(78,69,77,88,67,80,74,94,102)
10
11 x2 <- c
    (48.2,45.5,46.3,49.0,43.0,48.0,48.0,53.0,58.0)
12
13 x3 <- c
    (2.75,2.15,4.41,5.52,3.21,4.32,2.31,4.30,3.71)
14
15
16 x4 <- c
    (29.5,26.3,32.2,36.5,27.2,27.7,28.3,30.3,28.7)
17
18
19 dat <- data.frame(y,x1,x2,x3,x4)
20
21 # model..using forward selection
22
23 line <- lm(y~x1,data=dat)
24 summary(line)$adj.r.sq
25
26 line1 <- lm(y~x2,data=dat)
27 summary(line1)$adj.r.sq
28
29 line2 <- lm(y~x3,data=dat)
30 summary(line2)$adj.r.sq
31
32 line3 <- lm(y~x4,data=dat)
33 summary(line3)$adj.r.sq
34 # by looking at 4 adj r sq. values x1 has max.
    increase so x1 in the model
35 line4 <- lm(y~x1 +x2,data=dat)
36 summary(line4)$adj.r.sq
37
38 line5 <- lm(y~x1 +x3,data=dat)
39 summary(line5)$adj.r.sq
40
41 line6 <- lm(y~x1 +x4,data=dat)
42 summary(line6)$adj.r.sq

```

```

43 # by looking at the adj. r .sq. x1 +x3 combination
    has the max. value so x1+x3 in the model.
44
45 line7 <- lm(y~x1+x3+x4,data = dat)
46 summary(line7)$adj.r.sq
47 # by looking at this estimate , adj. r. sq. is
    decreases so x4 is excluded from the final model
    ..
48
49 line8 <- lm(y~x1+x2+x3,data = dat)
50 summary(line8)$adj.r.sq
51 cat("The correlation coefficient between x1 and x3
    is",cor(x1,x2))
52 #although the adj. r .squared is max. for this model
    ,but it is high from x1+x3 by small amount,we
    will not include x2 as the correlation
    coefficient between x1 and x2 is high meaning
    they are dependent..
53
54 #final model..
55 cat("The final model is ")
56 lm(y~x1+x3,data = dat)
57 summary(lm(y~x1+x3,data = dat))
58 cat("The coefficients of the final model is",
    coefficients(lm(y~x1+x3,data = dat)))
59 # we can also solve the same by backward elimination
    ..
60 #So both answers are correct..

```

---

## R code Exa 12.12 Cp Statistic

```

1 # Chapter 12
2 # Example 12.12 page no.492 from the pdf..
3 # Cp Statistic..
4

```

```

5 # to find the relationship between sales for a
  particular year and factor that affect sales..
6
7 # Package "leaps" is used if not installed can be
  done using install.packages("leaps")
8
9 library(leaps)
10
11 x1 <- c
    (5.5,2.5,8.0,3.0,3.0,2.9,8.0,9.0,4.0,6.5,5.5,5.0,6.0,5.0,3.5)
12
13 x2 <- c
    (31,55,67,50,38,71,30,56,42,73,60,44,50,39,55)
14
15 x3 <- c(10,8,12,7,8,12,12,5,8,5,11,12,6,10,10)
16
17 x4 <- c(8,6,9,16,15,17,8,10,4,16,7,12,6,4,4)
18
19 y <- c
    (79.3,200.1,163.2,200.1,146.0,177.7,30.9,291.9,160.0,339.4,159.6,
20
21 dat <- data.frame(x1,x2,x3,x4,y)
22
23 # I am comparing on basis of Cp values..
24
25 print("The Cp values for all subsets is")
26
27 leaps(dat[,1:4],y= dat[,5],names = names(dat)[1:4],
    method ="Cp")
28
29 # if you want to calculate adjusted r -squared or r-
    squared the code will be same as above just we
    have to change method = "r2" or method = "adjr2"
30
31 print("From Cp values it seems that x1x2x3 model
    appears quite good and had lowest Cp value")

```

```

32
33 print("Also You can see Cp for full model = 5.0 from
    the table in output")
34
35 # to obtain PRESS statistic for each model , we can
    do something like this..
36
37 # this is only PRESS for full model..
38 print("The PRESS for full model x1x2x3x4 is")
39
40 model <- lm(y~.,data = dat)
41
42 sum((model$residuals/(1-hatvalues(model)))^2)
43 # like this we can take any model and calculate its
    PRESS statistic..
44 # to compare answers from T.B look for Cp stats
    table and PRESS values in the tables and match it ,
    The answer is correct.

```

---

### R code Exa 12.13 Logistic regression model

```

1
2 # Chapter 12
3 # Example 12.13 page no.499 from the pdf..
4 # Logistic regression model..
5
6 x <- c(0.10,0.15,0.20,0.30,0.50,0.70,0.95)
7
8 n <- c(47,53,55,52,46,54,52)
9
10 y <- c(8,14,24,32,38,50,50)
11
12 p <- c(17.0,26.4,43.6,61.5,82.6,92.6,96.2)/100
13
14 dat <- data.frame(x,n,y,p)

```

```
15
16 # making a logistic regression..
17
18 model1 <- glm(p~x,data = dat,family = binomial(link
    = "logit"))
19 summary(aov(model1))
20 c <- coefficients(model1)
21 cat("The beta0 and beta1 for the logistic regression
    function is",c[1],c[2])
22
23 # the answer may vary slightly..
```

---



# Chapter 13

## One Factor Experiment General

R code Exa 13.1 One way ANOVA

```
1
2 # Chapter 13
3 # Example 13.1 page no. 513 from the pdf..
4 # One way ANOVA..
5
6 # H0:  $\mu_1=\mu_2=\mu_3=\mu_4=\mu_5$ 
7 # H1 : at least two of means not equal..
8
9 a <- c
    (551,457,450,731,499,632,595,580,508,583,633,517,639,615,511,573,
10
11 b <- c(rep(1,6),rep(2,6),rep(3,6),rep(4,6),rep(5,6))
12
13 dat <- data.frame(a,b)
14
15 c <- aov(a~factor(b),data = dat)
16
17 summary(c)
```

```

18
19 cat("Since the p-value is 0.00875 we reject null
      hypothesis and conclude aggregates don't have
      same mean absorption")

```

---

### R code Exa 13.2 One way ANOVA

```

1
2 # Chapter 13
3 # Example 13.2 page no. 514 from the pdf..
4 # One way ANOVA..
5
6 # NULL : H0: mu1=mu2=mu3=mu4
7 # alternate : H1: at least two are not equal
8
9
10 a <- c
      (49.20,44.54,45.80,95.84,30.10,36.50,82.30,87.85,105.00,95.22,97.1)
11
12 b <- c(rep(1,20),rep(2,9),rep(3,9),rep(4,7))
13
14 dat <- data.frame(a,b)
15
16 c <- aov(a~factor(b),data = dat) # anova
17
18 summary(c) # Analysis of Variance table
19
20 cat("Since the p-value is 0.022,we reject the null
      hypothesisand conclude alkaline levels for the
      four drug groups are not the same")

```

---

### R code Exa 13.3 Bartlett Test

```

1
2 # Chapter 13
3 # Example 13.3 page no. 517 from the pdf..
4 # Bartlett Test..
5
6 # to test the following hypothesis
7 # NULL : H0:  $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2$ 
8 # alternate : H1: at least two are not equal
9
10 # using bartlett test..
11 a <- c
    (49.20,44.54,45.80,95.84,30.10,36.50,82.30,87.85,105.00,95.22,97.5)
12
13 b <- c(rep("I",20),rep("J",9),rep("K",9),rep("L",7))
14
15 dat <- data.frame(a,b)
16
17 bartlett.test(a~b,data = dat) # display of Bartlett
    test..
18
19 cat("The p-value is",bartlett.test(a~b,data = dat)$p
    .value,"we do not reject the null hypothesis and
    conclude popn. variances are not significantly
    different")
20
21 # the T.B has done this by comparing the areas in
    the left tail of bartlett distribution,using
    bartlett test function directly gives p-value in
    R, so it is more handy..

```

---

**R code Exa 13.4** Contrasts Sum Of Squares Corresponding to Orthogonal Contrasts

1

```

2 # Chapter 13
3 # Example 13.4 page no. 522 from the pdf..
4 # Contrasts Sum Of Squares Corresponding to
  orthogonal Contrasts..
5
6 # from ex 13.1..
7 # w1 = mu1+mu2-mu3-mu5, w2 = mu1+mu2+mu3-4mu4+ mu5.
8 # w3 = mu1 -mu2 , w4 = mu3-mu5.
9
10 a <- c
    (551,457,450,731,499,632,595,580,508,583,633,517,639,615,511,573,
11
12 b <- c(rep(1,6),rep(2,6),rep(3,6),rep(4,6),rep(5,6))
13
14 # making a data frame and re-classifying the data.
15 dat <- data.frame(a,b)
16 dat$b <- as.factor(dat$b)
17
18 # the ANOVA..
19 l <- lm(a~b,data = dat)
20 anova(l)
21
22 # contrasts coefficients..
23 contrastmatrix <- cbind(c(1,1,-1,0,-1),c(1,1,1,-4,1)
    ,c(1,-1,0,0,0),
24                        c(0,0,1,0,-1))
25 contrasts(dat$b) <- contrastmatrix
26
27
28 l_contrast <- aov(a~b,data = dat)
29
30
31 print("The Analysis of Variance Table Using
    ortogonal contrats is:")
32 summary(l_contrast, split = list(b = list("(1,2) vs
    (3,5)"=1,"1,2,3,5 vs 4"=2)))

```

---

### R code Exa 13.5 Dunnett Test

```
1
2 # Chapter 13
3 # Example 13.5 page no. 529 from the pdf..
4 # Dunnett Test..
5
6 # To compare each catalyst with control using 2
   sided hypothesis..
7 # Package used "DescTools", if already installed
   delete the below line from the code..
8 install.packages("DescTools")
9
10 library(DescTools)
11
12 co <- c(50.7,51.5,49.2,53.1,52.7)
13
14 cat1 <- c(54.1,53.8,53.1,52.5,54.0)
15
16 cat2 <- c(52.7,53.9,57.0,54.1,52.5)
17
18 cat3 <- c(51.2,50.8,49.7,48.0,47.2)
19
20 DunnettTest(list(co,cat1,cat2,cat3))
21
22 # The value of Diff column in the table varies
   slightly from the T.B, due to approximations..
23 # Like d1 and d3 in T.B are 2.14 and -2.14, and
   from computation coming out to be 2.06 and -2.06
   so that's due to approximations or different
   method to approach the same problem but final
   inferences are same..
24
25 print("From the Table looking at p-value we conclude
```

that catalyst 2 is significantly different from  
the mean yield of the reaction using the control  
..”)

---

### R code Exa 13.6 Randomized Complete Block Diagram

```
1
2 # Chapter 13
3 # Example 13.8 page no. 537 from the pdf..
4 # Randomized Complete Block Diagram..
5
6 # to to test hypothesis(0.05 level) that machines
   perform at the same mean rate of speed
7 # 6 different operators used in randomized block
   experiment to compare 4 machines..
8
9 m1 <- c(42.5,39.3,39.6,39.9,42.9,43.6)
10
11 m2 <- c(39.8,40.1,40.5,42.3,42.5,43.1)
12
13 m3 <- c(40.2,40.5,41.3,43.4,44.9,45.1)
14
15 m4 <- c(41.3,42.2,43.5,44.2,45.9,42.3)
16
17 dat <- rbind(m1,m2,m3,m4)    # combining rows to make
   matrix..
18
19 a <- c(t(as.matrix(dat)))    # concatenate different
   rows into a vector..
20
21 b <- c("o1","o2","o3","o4","o5","o6") # treatment
   levels
22
23 n_tr <- 6 # no. of treatment levels
24
```

```

25 n_cont <- 4 # no. of control blocks..
26
27 operator <- gl(n_tr,1,n_cont*n_tr,factor(b)) #
    vector of treatment factors corresponding to each
    element of vector a..
28
29 machines <- gl(n_cont,n_tr,n_tr*n_cont) # vector of
    blocking factors corresponding to each element in
    vector a..
30
31 print("The Analysis of Variance table is:")
32
33 summary(aov(a~operator+machines)) # anova table
    display..

```

---

### R code Exa 13.7 Random Effects Model

```

1
2 # Chapter 13
3 # Example 13.7 page no. 549 from the pdf..
4 # Random Effects Model..
5 # to find the batch variance component from the data
    ..
6
7 m1 <- c(9.7,5.6,8.4,7.9,8.2,7.7,8.1)
8
9 m2 <- c(10.4,9.6,7.3,6.8,8.8,9.2,7.6)
10
11 m3 <- c(15.9,14.4,8.3,12.8,7.9,11.6,9.8)
12
13 m4 <- c(8.6,11.1,10.7,7.6,6.4,5.9,8.1)
14
15 m5 <- c(9.7,12.8,8.7,13.4,8.3,11.7,10.7)
16
17 dat <- cbind(m1,m2,m3,m4,m5)

```

```

18
19 b <- c(t(as.matrix(dat)))
20
21 a <- c("b1", "b2", "b3", "b4", "b5")
22
23 n_batch <- 5 # n. of treatment groups
24
25 n_row <- 7 # no. of rows..
26
27 batch <- gl(n_batch, 1, n_row*n_batch, factor(a)) #
      vector of treatment factors corresponding to each
      element of vector b..
28
29 summary(c <- aov(b~batch))
30
31 cat("From the table we can calculate the estimated
      variance from the mean sq. values which comes out
      to be", (18.149-4.068)/n_row)

```

---



# Chapter 14

## Factorial Experiments Two Or More Factors

**R code Exa 14.1** Two Factor Analysis of Variance

```
1
2 # Chapter 14
3 # Example 14.1 page no. 571 from the pdf..
4 # Two Factor Analysis of Variance..
5
6 # H0': alpha1=alpha2=alpha3=0
7 #H0'': beta1=beta2=beta3=beta4=0
8 #H0''': (alpha*beta)11=(alpha*beta)12...=(alpha*beta)
   34=0
9 # H1': atleast one of alpha non zero
10 # H1'': atleast one of the beta non zero
11 #H1''': atleast one of the (alpha*beta) non zero..
12
13 v1 <- c(34.0,32.7,32.0,33.2,28.4,29.3)
14
15 v2 <- c(30.1,32.8,30.2,29.8,27.3,28.9)
16
17 v3 <- c(29.8,26.7,28.7,28.1,29.7,27.3)
18
```

```

19 v4 <- c(29.0,28.9,27.6,27.8,28.8,29.1)
20
21 observations <- c(v1,v2,v3,v4)
22
23 prop_type <- c(rep(1,6),rep(2,6),rep(3,6),rep(4,6))
24
25 a <- c(rep(1,2),rep(2,2),rep(3,2))
26
27 missile_sys <- c(rep(a,4))
28
29 dat <- data.frame(observations,prop_type,missile_sys
30 )
31 d <- aov(observations~factor(missile_sys)*factor(
32   prop_type),data = dat)
33 summary(d) # analysis of variance table..
34
35 cat("We reject H0' and conclude that different
36   missile have different propellent rates as p-
37   value is 0.0169")
38
39 cat("We reject H0'' and conclude that mean propellent
40   rates are not same for 4 propellent types as p-
41   value is 0.0010")
42
43 cat("p-value is approx. 0.0513, so interaction is
44   barely significant.")

```

---

## R code Exa 14.2 Single Degree Of Freedom Sum Of Squares

```

1
2 # Chapter 14
3 # Example 14.2 page no. 571 from the pdf..
4 # Single Degree Of Freedom Sum Of Squares..

```

```

5
6 # to choose 2 orthogonal contrasts to partition the
  sum of squares for missile systems into
7 # single-degree-of-freedom components to be used in
  comparing systems 1 and 2 versus 3 and system 1
  with system 2.
8
9
10 v1 <- c(34.0,32.7,32.0,33.2,28.4,29.3)
11
12 v2 <- c(30.1,32.8,30.2,29.8,27.3,28.9)
13
14 v3 <- c(29.8,26.7,28.7,28.1,29.7,27.3)
15
16 v4 <- c(29.0,28.9,27.6,27.8,28.8,29.1)
17
18 observations <- c(v1,v2,v3,v4)
19
20 prop_type <- c(rep(1,6),rep(2,6),rep(3,6),rep(4,6))
21
22 a <- c(rep(1,2),rep(2,2),rep(3,2))
23
24 missile_sys <- c(rep(a,4))
25
26 dat <- data.frame(observations,prop_type,missile_sys
  )
27
28 dat$missile_sys <- as.factor(dat$missile_sys)
29
30 contrastmatrix <- cbind(c(1,1,-2),c(1,-1,0))
31
32 contrasts(dat$missile_sys) <- contrastmatrix
33
34 missile_contrast <- aov(observations~missile_sys,
  data = dat)
35
36
37 print("The Sum of Squares for missiles systems to be

```

```

        used in comparing systems 1 and 2 versus 3 and
        system 1 versus 2 can be seen from the analysis
        of variance table below.")
38 summary(missile_contrast, split = list(missile_sys =
39                                     list("(1,2) vs 3"
                                           =1,"1 vs 2"=2)
                                           ))

```

---

**R code Exa 14.3** Evaluate Analysis of Variance table and derive conclusion also test on main effects

```

1
2 # Chapter 14
3 # Example 14.3 page no. 574 from thr pdf..
4 # Evaluate Analysis of Variance table and derive
   conclusion ,also test on main effects ..
5
6 m1 <- c(288,488,670,360,465,720)
7
8 m2 <- c(385,482,692,411,521,724)
9
10 m3 <- c(488,595,761,462,612,801)
11
12 obs <- c(m1,m2,m3)
13
14 power_supp <- c(rep(1:3,6))
15
16
17 flow_rate <- c(rep(1,6),rep(2,6),rep(3,6))
18
19 dat <- data.frame(obs,power_supp,flow_rate)
20
21 d <- aov(obs~factor(flow_rate)*factor(power_supp),
   data = dat)
22

```

```

23 summary(d)
24
25 cat("p-value for test of interaction is 0.4484 so we
      conclude there is no significant interaction")
26
27 # duncan test can be done by duncan.test, in
    agricolae library, I am having some problems with
    the package it is not installing the package and
    showing some error, so I cannot use duncan.test
    function..

```

---

#### R code Exa 14.4 Anova for 3 factor experiment

```

1
2 # Chapter 14
3 # Example 14.4 page no. 581 from the pdf..
4 # Anova for 3 factor experiment..
5
6 m1 <- c
    (10.7,10.3,11.2,10.9,10.5,12.2,10.8,10.2,11.6,12.1,11.1,11.7,11.3
7
8 m2 <- c
    (11.4,10.2,10.7,9.8,12.6,10.8,11.8,10.9,10.5,11.3,7.5,10.2,11.5,1
9
10 m3 <- c
    (13.6,12.0,11.1,10.7,10.2,11.9,14.1,11.6,11.0,11.7,11.5,11.6,14.5
11
12
13 obs <- c(m1,m2,m3)
14
15 operator <- c(rep(1,18),rep(2,18),rep(3,18))
16

```

```

17 catalyst <- c(rep(1:3,18))
18
19 a <- c(rep("15 min",3),rep("20 min",3))
20
21 washing_time <- c(rep(a,9))
22
23 dat <- data.frame(obs,operator,catalyst,washing_time
24 )
25 d <- aov(obs~(factor(operator)*factor(catalyst)*
26   factor(washing_time)),data = dat)
27
28 print("The Analysis of Variance Table is shown as
29 follows:")

```

---

#### R code Exa 14.5 Pooling in Multi factor Model

```

1
2 # Chapter 14
3 # Example 14.5 page no. 584 from the pdf..
4 # Pooling in Multifactor Model..
5
6 # to find the anova table from the following data
7   set by removing some interactions and then look
8   at the effect..
9
10 m1 <- c(43,49,44,47)
11
12 m2 <- c(64,68,97,102)
13
14 m3 <- c(49,57,51,55)
15
16 m4 <- c(70,76,103,106)

```

```

15
16 obs <- c(m1,m2,m3,m4)
17
18 a <- c(rep("L",4),rep("H",4))
19 temp <- c(rep(a,2))
20
21 batch <- c(rep(1,8),rep(2,8))
22
23 b <- c(rep("Low",2),rep("high",2))
24 string_rate <- c(rep(b,4))
25
26 pressure <- c(rep(c("Low","High"),8))
27
28 dat <- data.frame(obs,batch,temp,string_rate,
29                   pressure) # making data frame
30
31 cat("The anova table after removing particular
32   interactions is")
33
34 cat(summary(aov(obs~factor(batch)+factor(temp)*factor(
35   pressure)*factor(string_rate),data = dat)) #
36   taking the elements of interest for interaction..
37
38 # we can also use "update" function by writitng the
39   full model of 4 variables and subtracting the
40   undesired interactions ,but it will be very
41   tedious..
42
43 # the answer may vary slightly..

```

---

#### R code Exa 14.6 Factorial Experiments for Random Effects

```

1
2 # Chapter 14
3 # Example 14.6 page no. 589 from the pdf..

```

```

4 # Factorial Experiments for Random Effects..
5
6 # to determine which are the important sources of
  variation in an industrial process
7
8 op1 <- c
  (66.9,68.3,69.0,69.3,68.1,67.4,69.8,70.9,67.2,67.7,67.5,71.4)
9
10 op2 <- c
  (66.3,68.1,69.7,69.4,65.4,66.9,68.8,69.6,65.8,67.6,69.2,70.0)
11
12 op3 <- c
  (65.6,66.0,67.1,67.9,66.3,66.9,66.2,68.4,65.2,67.3,67.4,68.7)
13
14
15 obs <- c(op1,op2,op3)
16
17 operator <- c(rep(1,12),rep(2,12),rep(3,12))
18
19 batch <- c(rep(1:4,9))
20
21
22 dat <- data.frame(obs,operator, batch)
23
24 print("The analysis of variance table for this
  example is:")
25
26 summary(aov(obs~factor(operator)+ factor(batch) +
  factor(batch)*factor(operator)),data = dat)
27 # interaction in this case is the product of
  operator and batch components.

```

---



# Chapter 15

## 2 to the power k Factorial Experiments and Fractions

R code Exa 15.1 2 squared Factorial Experiment

```
1
2 # Chapter 15
3 # Example 15.1 page no. 601 from the pdf..
4 # 2^2 factorial Experiment..
5
6 # 2^2 factorial with no interaction
7 obs <- c(50,70,80,100)
8
9 a_no <- c(rep(1,2),rep(-1,2))
10
11 b_no <- c(rep(c(-1,1),2))
12
13 cat("The main effects are A =", (sum(obs[1:2]) - sum(
    obs[3:4])) / 2)
14
15 cat("B = ", (obs[4] + obs[2]) / 2 - (obs[1] + obs[3]) / 2)
16
17 cat("The interaciton effect is", (obs[1] + obs[4]) / 2 - (
    obs[2] + obs[3]) / 2)
```

```

18
19 obs1 <- c(50,70,80,40)
20
21 cat("The interaction effect in this case is", (obs1
    [2]+obs1[3])/2-(obs1[1]+obs1[4])/2)

```

---

### R code Exa 15.2 Factorial Experiment in a Regression Setting

```

1
2 # Chapter 15
3 # example 15.2 page no. 613 from the pdf
4 # Factorial Experiment in a Regression Setting..
5
6 hold_tim <- c(0.5,0.8,0.5,0.8)
7
8 flex_time <- c(0.10,0.10,0.20,0.20)
9
10 yield <- c(28,39,32,46)
11
12 dat <- data.frame(hold_tim,flex_time,yield)
13
14 dat$hold_tim <- factor(dat$hold_tim)
15 dat$flex_time <- factor(dat$flex_time)
16
17
18 f <- function(x){
19
20     dat$yield[x]
21
22 }
23
24 levels(dat$hold_tim) <- c(-1,1)
25 levels(dat$flex_time) <- c(-1,1)
26
27 cat("The regression equation is", (f(1)+f(2)+f(3)+f

```

```
(4))/4,"+",(f(2)+f(4)-f(1)-f(3))/4,"*x1 +", (f(3)+
f(4)-f(1)-f(2))/4,"*x2")
```

---

**R code Exa 15.4** Standard Errors of the least Squares Regression Coefficients

```
1
2 # Chapter 15
3 # example 15.4 page no. 619
4 # Standard Errors of the least Squares Regression
  Coefficients
5
6 # given data and anova table , to find the standard
  errors of the least squares regression
  coefficients..
7
8 # Note 15.3 – Theoretical with nothing to compute..
9
10 # standard errors of all coefficients for the 2^k
   factorial are equal so..
11 # from anova table given s^2 = 2.46
12
13 cat("The standard error of the least squares
   regression coefficients are",sqrt(2.46/(2^4*2)))
```

---

**R code Exa 15.6** Use a half replicate to study the effects of five factors each at 2 levels on some response

```
1
2 # Chapter 15
3 # example 15.6 page no.633 from the pdf..
4 # use a half replicate to study the effects of five
   factors each at 2 levels on some response..
```

```

5 # to perform analysis of variance on the data below
  testing all main effects for the significance at
  the 0.05 level..
6
7 response <- c
  (11.3,15.6,12.7,10.4,9.2,11.0,8.9,9.6,14.1,14.2,11.7,9.4,16.2,13.0)
8 treatment <- c("a","b","c","d","e","abc","abd","acd"
  ,"bcd","abe","ace","ade","bce","bde","cde","abcde"
  ")
9 cat("The sum of squares and effects for the main
  effects are")
10
11 a <- sum(response)-2*sum(response[which(treatment %
  in% c("b","c","d","e","bcd","bce","bde","cde"))])
12
13 b <- sum(response)-2*sum(response[which(treatment %
  in% c("a","c","d","e","acd","ace","ade","cde"))])
14
15 c <- sum(response)-2*sum(response[which(treatment %
  in% c("b","a","d","e","abd","abe","ade","bde"))])
16
17 d <- sum(response)-2*sum(response[which(treatment %
  in% c("b","c","a","e","abc","abe","ace","bce"))])
18
19 e <- sum(response)-2*sum(response[which(treatment %
  in% c("b","c","d","a","abc","abd","acd","bcd"))])
20
21 cat("The sum of squares and effects for the main
  effects are")
22
23 cat("SSA=",a^2/16,"and A=",a/8)
24
25 cat("SSB=",b^2/16,"and B=",b/8)
26
27 cat("SSC=",c^2/16,"and C=",c/8)
28
29 cat("SSD=",d^2/16,"and D=",d/8)

```

```
30
31 cat("SSE=", e^2/16, " and E=", e/8)
```

---

**R code Exa 15.7** To construct a 2 level screening design with 6 variables containing 12 design points

```
1
2 # Chapter 15
3 # example 15.7 page no. 639
4 # to construct a 2 level screening design with 6
  variables containing 12 design points..
5 # used package FrF2, if not installed can be
  installed using install.packages("FrF2")
6
7 library(FrF2)
8
9 print("The 2 level screening design with 6 variables
  containing 12 design points is")
10
11 pb(12,6,randomize = F,default.levels = c("-", "+"))
12
13 # the answer may differ but the concept is same, the
  T.B as well as the software answers are correct,
  R has selected randomly the design,
14 #if we set randomize= T, then every time the
  instruction is executed then every time a
  diferent design will appear, so that's why the
  design is correct.
```

---

**R code Exa 15.8** Response Surface Analysis

```
1
2 # Chapter 15
```

```

3 # example 15.8 page no. 640 from the pdf..
4 # Response Surface Analysis..
5
6 # a central composite design is given on page no.641
  of the pdf,to determine the impact that x1 and
  x2 have on % conversion process..
7 #Package used – rsm if not installed can be done
  using install.packaged("rsm")
8 library(rsm)
9
10 x1 <- c(-1,1,-1,1,-1.414,1.414,0,0,0,0,0,0)
11
12 x2 <- c(-1,-1,1,1,0,0,-1.414,1.414,0,0,0,0)
13
14 y <- c(43,78,69,73,48,78,65,74,76,79,83,81)
15
16 dat <- data.frame(x1,x2,y)
17
18 res_model <- rsm(y~F0(x1,x2,x1^2,x2^2,x1*x2),data =
  dat)
19
20 print("The coefficients of the resulting second
  order response model is given in coded variables
  as")
21
22 coefficients(res_model)
23
24 print("The detailed response surface model is given
  below")
25
26 summary(res_model)
27
28
29 # the natural model..
30 e1 <- c
  (200,250,200,250,189.65,260.35,225,225,225,225,225,225)
31

```

```
32 e2 <- c(15,15,25,25,20,20,12.93,27.07,20,20,20,20)
33
34 dat1 <- data.frame(e1,e2,y)
35
36 print("The natural surface model is shown below:")
37
38 nat_model <- lm(y~e1+e2+I(e1^2)+I(e2^2)+I(e1*e2),
    data = dat1)
39
40 summary(nat_model)
41
42 print("The coefficients of the natural surface model
    is given as follows:")
43
44 coefficients(nat_model)
```

---

# Chapter 16

## Non Parametric Statistics

**R code Exa 16.1** Use sign test to test the hypothesis

```
1
2 # Chapter 16
3 # Example 16.1 page no. 658 from the pdf..
4 # use sign test to test the hypothesis..
5
6 # Package used "BSDA", reference - Internet..
7
8 install.packages("BSDA") # package for sign testing
9
10 library("BSDA")
11
12 #NULL H0:mu=1.8
13 #alternate: mu!= 8
14
15 dat <- c
16     (1.5,2.2,0.9,1.3,2.0,1.6,1.8,1.5,2.0,1.2,1.7)
17
18 SIGN.test(dat,md=1.8)
19
20 cat("Since The p-value is",SIGN.test(dat,md=1.8)$p.
```



```
value,"we do not reject the null hypothesis.")
```

---

**R code Exa 16.2** Use Sign test to test the hypothesis

```
1
2 # Chapter 16
3 # Example 16.2 page no. 659 from the pdf..
4 # Hypothesis Testing Using Sign Test..
5 # to test the hypothesis with 0.05 level of
  significance
6
7 # Null : $\mu_1 - \mu_2 = 0$ 
8 #alternate: $\mu_1 - \mu_2 \neq 0$ 
9
10
11 install.packages("BSDA") # package for sign testing
12 #if already installed comment it..
13
14
15 library("BSDA")
16
17 rad_tires <- c
  (4.2,4.7,6.6,7.0,6.7,4.5,5.7,6.0,7.4,4.9,6.1,5.2,5.7,6.9,6.8,4.9)
18
19 bel_tires <- c
  (4.1,4.9,6.2,6.9,6.8,4.4,5.7,5.8,6.9,4.9,6.0,4.9,5.3,6.5,7.1,4.8)
20
21 diff <- data.frame(rad_tires,bel_tires)
22
23 SIGN.test(rad_tires,bel_tires,alternative = "greater
  ",paired =T)
24
25 cat("Since the p-value is",SIGN.test(rad_tires,bel_
```

```

    tires,alternative = "greater",paired =T)$p.value,
    "we reject null, so on average radial tires do
    improve fuel economy")
26
27
28 #the answer is slightly different than in the T.B as
    in T.B normal approximation is used..

```

---

**R code Exa 16.3** Use signed rank test to test the hypothesis

```

1
2 # Chapter 16
3 # Example 16 page no. 661 from the pdf..
4 #use signed rank test to test the hypothesis..
5
6 #NULL H0:mu=1.8
7 #alternate: mu!= 8
8
9 dat <- c
    (1.5,2.2,0.9,1.3,2.0,1.6,1.8,1.5,2.0,1.2,1.7)
10
11 wilcox.test(dat,alternative = "two.sided",mu=1.8)
12
13
14 cat("The p-vauue is more than 0.05,we don't reject
    null and conclude median operating time not
    significantly different from 1.8hrs")

```

---

**R code Exa 16.4** Use signed rank test to test the hypothesis

```

1
2 # Chapter 16
3 # Example 16.4 page no. 662 from the pdf..

```

```

4 #use signed rank test to test the hypothesis..
5
6 # to do hypothesis testing on effect of sample
  questions on one's gre score..
7
8 # NULL H0:  $\mu_1 - \mu_2 = 50$ 
9 #alternate:  $\mu_1 - \mu_2 < 50$ 
10
11 with_sample <- c
    (531,621,663,579,451,660,591,719,543,575)
12
13 without_sample <- c
    (509,540,688,502,424,683,568,748,530,524)
14
15 wilcox.test(with_sample,without_sample,mu=50,paired
    = T,alternative = "less")
16
17 cat("since the p-value is",wilcox.test(with_sample,
    without_sample,mu=50,paired = T,alternative = "
    less")$p.value,"we reject null and conclude that
    sample problems don't increase graduate record
    score by as much as 50 points on average..")

```

---

**R code Exa 16.5** To perform Wilcoxin rank sum test

```

1
2 # Chapter 16
3 # Example 16.5 page no. 666 from the pdf..
4 # to perform wilcoxin rank sum test...
5
6 # Null H0 :  $\mu_1 = \mu_2$ 
7 #alternate:  $\mu_1 \neq \mu_2$ 
8
9 brand_a <- c(2.1,4.0,6.3,5.4,4.8,3.7,6.1,3.3)
10

```

```

11 brand_b <- c
    (4.1,0.6,3.1,2.5,4.0,6.2,1.6,2.2,1.9,5.4)
12
13 wilcox.test(brand_a,brand_b,alternative = "two.sided
    ")
14
15 cat("Since the p-value is",wilcox.test(brand_a,brand
    _b,alternative = "two.sided")$p.value,"we do not
    reject null and conclude that there is no
    significant difference in median nicotine
    contents of the above two brand of cigarettes")

```

---

#### R code Exa 16.6 Kruskal Wallis Test

```

1
2 # Chapter 16
3 # Example 16.6 page no. 668 from the pdf..
4 # Kruskal-Wallis Test..
5
6 # to do hypothesis testing of propellant burning
   rates for 3 missile system using kruskal wallis
   test..
7
8 #Null: H0: mu1=mu2=mu3
9 #alternate : the means are not equal.
10 #alpha = 0.05
11
12 m1 <- c(24.0,16.7,22.8,19.8,18.9)
13
14 m2 <- c(23.2,19.8,18.1,17.6,20.2,17.8)
15
16 m3 <- c(18.4,19.1,17.3,17.3,19.7,18.9,18.8,19.3)
17
18 obs <- c(m1,m2,m3)
19

```

```

20 miss_sys <- c(rep(1,5),rep(2,6),rep(3,8))
21
22 kruskal.test(obs,factor(miss_sys))
23
24 cat("Since the p-value is",kruskal.test(obs,factor(
    miss_sys))$p.value,"we failed to reject null that
    propellent burning rates are same for the 3
    missile system")

```

---

#### R code Exa 16.7 Runs Test

```

1
2 # Chapter 16
3 # Example 16.7 page no. 672 from the pdf..
4 # Runs Test..
5
6 # to do hypothesis testing of randomness of the
  sequence..
7
8 # Null :H0= sequence is random
9 # alternate: H1= sequence is is not random..
10
11 # package used – "randtests", reference – internet..
12 install.packages("randtests") # package for runs.
  test
13 # comment it if already installed..
14
15 library("randtests")
16
17 content <- c
  (3.6,3.9,4.1,3.6,3.8,3.7,3.4,4.0,3.8,4.1,3.9,4.0,3.8,4.2,4.1)
18
19 runs.test(content,alternative = "two.sided")
20

```

```

21
22 cat("The p-value is more than 0.1 we don't reject
    the null and conclude that sequence of measured
    values varies randomly")

```

---

### R code Exa 16.8.9 Rank Correlation Coefficient

```

1
2 # Chapter 16
3 # Example 16.8 page no. 675 and example 16.9 page
  no.677 from the pdf..
4 # Rank Correlation Coefficient..
5
6 # to find the rank correlation coefficient between
  tar and nicotine content in cigarettes..
7 # I am solving ex 8 and 9 in this code only as in 8
  correlation coefficient is asked and in 9th
  hypothesis testing of the sam eproblem is done
8
9 # to do hypothesis testing of correlation between
  tar and nicotine content in cigarettes
10 # null H0:rho=0
11 # alternate H1: rho>0
12 # alpha =0.01
13
14 tar <- c(14,17,28,17,16,13,24,25,18,31)
15
16 nicotine_cont <- c
    (0.9,1.1,1.6,1.3,1.0,0.8,1.5,1.4,1.2,2.0)
17
18 cor.test(tar,nicotine_cont,method = "spearman",
    alternative = "greater")
19
20 cat("From the test we can see the value of rank
    correlation coefficint(rho) is 0.9665698

```

```
    indicating high positive correlation between the
    two variables..")
21
22 cat("Since the p-value is very small,",cor.test(tar,
    nicotine_cont,method = "spearman",alternative = "
    greater")$p.value,"we reject null hypothesis and
    conclude that there is significant correlation
    between the two variables..")
23
24 # ignore the warning in each case..
```

---

# Chapter 17

## Statistical Quality Control

**R code Exa 17.1** Expected Value of samples required to detect the shift

```
1
2 # Chapter 17
3 # Example 17.1 page no. 694 from the pdf..
4 # Expected Number of samples required..
5
6 # given n = 4 ,r = 1
7 # By graph we can see beta = 0.84
8
9 beta <- 0.84
10
11 cat("The mean of samples requires to detect the
      shift is",1/(1-beta))
```

---

**R code Exa 17.2** Xbar and S chart

```
1
2 # Chapter 17
3 # Example 17.2 page no. 697 from the pdf..
```





```

28 samples <- c(rep(1:25,5))
29
30
31 dat <- data.frame(observavation,samples)
32
33
34 print("The xbar and S chart for the above data is:")
35
36
37 # Run the below two code individually..
38 #xbar chart
39 qic(observavation,
40     x= samples,
41     data = dat,
42     chart = 'xbar',
43     xlab = 'Sample Number')
44
45 # S chart
46 qic(observavation,
47     x = samples,
48     chart = 's',
49     xlab = 'Sample Number',
50     data = dat)

```

---

### R code Exa 17.3 UCL and LCL for preliminary control chart value

```

1 # Chapter 17
2 # Example 17.3 page no. 700 from the pdf..
3 # UCL and LCL for preliminary control chart values.
4
5 # given data on no. of defective components in
  sample sizes of 50.
6
7 def_comp <- c
  (8,6,5,7,2,5,3,8,4,4,3,1,5,4,4,2,3,5,6,3)

```

```

8
9 samples <- c(1:20)
10
11 dat <- data.frame(samples, def_comp)
12
13 m <- mean(dat$def_comp/50)
14
15 u <- m + 3*sqrt(m*(1-m)/50)
16
17 l <- m - 3*sqrt(m*(1-m)/50)
18
19 cat("The LCL and UCL are", l, "and", u, "respectively")

```

---

#### R code Exa 17.4 Find Sample Size per subgroup

```

1
2 # Chapter 17
3 # Example 17.4 page no. 701 from the pdf..
4 # Find Sample Size per subgroup..
5
6 # to find the sample size per subgroup producing a
   prob. of 0.5 that a process shift to p=p1= 0.05
   will be detected.
7 #given in control prob. of a defective= 0.01
8
9 p = 0.01
10
11 p1 = 0.05
12
13 cat("The appropriate sample size is", (9/(p1-p)^2)*p*
   (1-p))

```

---

#### R code Exa 17.5 Control Charts for Defects C chart

```

1
2 # Chapter 17
3 # Examples 17.5 page no.702 from the pdf..
4 # Control Charts for Defects C chart..
5
6 # given – no. of defects in 20 successive samples of
   sheet metal rolls each 100 feet long.
7 # to develop a control chart.
8
9 # package used – qicharts2
10 # if not installed can be installed by install.
   packages("qicharts2")
11
12 library(qicharts2)
13
14 lambd <- 5.95
15
16 num_def <- c
   (8,7,5,4,4,7,6,4,5,6,3,7,5,9,7,7,8,6,7,4)
17
18 samples <- c(1:20)
19
20 dat <- data.frame(samples, num_def)
21
22 # C control chart for the above preliminary data.
23 qic(num_def,
24     x = samples,
25     chart = 'c',
26     xlab = 'Sample',
27     ylab = 'Number of defects',
28     data = dat)
29
30 cat("The UCL and LCL for the above preliminary data
   are",lambd+3*sqrt(lambd),"and",lambd-3*sqrt(lambd
   ),"respectively.")

```

---

# Chapter 18

## Bayesian Statistics

**R code Exa 18.1** To find the Posterior probability distribution

```
1
2 # Chapter 18
3 # Example 18.1 page no. 711 from the pdf..
4 # To find the posterior probability distribution..
5
6 b <- function(p,x){
7
8   choose(2,x)*p^x*(1-p)^(2-x)
9
10 }
11
12 p1 <- c(0.1,0.2)
13
14 pi <- c(0.6,0.4)
15
16 m <- c(b(0.1,0)*pi[1]+b(0.2,0)*pi[2],b(0.1,1)*pi[1]+
17       b(0.2,1)*pi[2],b(0.1,2)*pi[1]+b(0.2,2)*pi[2])
18
19 post_p1 <- c(b(0.1,0)*pi[1]/m[1],b(0.1,1)*pi[1]/m
20             [2],b(0.1,2)*pi[1]/m[3])
```

```

20 post_p2 <- c(1,1,1)-post_p1
21
22 cat("The posterior for p=0.1 given x is",post_p1,"x
    ranges from 0 to 2")
23
24 cat("The posterior for p=0.2 given x is",post_p2,"x
    ranges from 0 to 2")

```

---

#### R code Exa 18.4 Posterior Mean and Posterior Mode

```

1
2 # Chapter 18
3 # Example 18.4 page no. 713 from the pdf..
4 # Posterior Mean and Posterior Mode..
5
6 # Note – Example 18.2 and 18.3 are theoretical with
    nothing to compute..
7
8 # given a distribution as on example 18.2 on page no
    . 712 of pdf
9 # to find the posterior mean and mode when x=1
10
11 # the posterior distribution function for 0<p<1
12 # package used "polynom"
13
14 install.packages("polynom")
15 library(polynom)
16
17 a <- function(x,p){
18
19     3*choose(2,x)*(p^x)*((1-p)^(2-x))
20
21 }
22
23 cat("The posterior mean when x=1 is",integrate(

```

```

      function(b){b*a(1,b)},0,1)$value)
24
25 a <- polynomial(coef = c(0,6,-6))
26
27 b <- deriv(a)
28
29 cat("The posterior mode at x=1 occurs at p equal to"
      ,Re(polyroot(b)))

```

---

**R code Exa 18.6** To find 95 percent Bayesian interval

```

1
2 # Chapter 18
3 # example 18.6 page no. 715 from the pdf..
4 # to find 95% bayesian interval
5
6 # Note – Example 18.5 is theoretical with nothing to
  compute..
7
8 # given prior distribution uniform , for 0<p<1
9 # refer example 18.2 on page no.712
10
11 # from example 18.2 we can see that at x=0 the
  posterior distribution is  $3(1-p)^2$ 
12 # package "polynom" is used, if not installed can be
  done using install.packages("polynom")
13
14 library(polynom)
15
16 a <- polynomial(coef = c(3,-6,3))
17
18 b <- polynomial(c(1,-3,3,-1)) - 0.025
19
20 c <- integral(a)-0.025
21

```

```
22 cat("The 95% bayesian interval is",polyroot(c)[1],"
    and",polyroot(b)[1])
```

---

**R code Exa 18.7** To find the 95 percent Bayesian interval for mu given the various parameter

```
1
2 # Chapter 18
3 # example 18.7 page no. 716, from the pdf..
4 # to find the 95% bayesian interval for mu..given
  the various parameters.
5
6 mu_o <- 800
7
8 sig_o <- 10
9
10 n <- 25
11
12 sig <- 100
13
14 x <- 780
15
16 mu <- (n*x*sig_o^2 + mu_o*sig^2)/(n*sig_o^2+sig^2)
17
18 sd <- sig_o*sig/sqrt(n*sig_o^2+sig^2)
19
20 cat("The 95% Bayesian interval for mu is given by",
    mu+qnorm(0.025)*sd," to",mu-qnorm(0.025)*sd)
```

---

**R code Exa 18.8** To find Bayes estimates of p

```
1
2 # Chapter 18
```



```

3 # example 18.8 page no.717 from the pdf..
4 # to find bayes estimates of p, for all values of x
  in example 18.1
5
6 # refer example 18.1..
7
8 b <- function(p,x){
9
10   choose(2,x)*p^x*(1-p)^(2-x)
11
12 }
13
14 p1 <- c(0.1,0.2)
15
16 pi <- c(0.6,0.4)
17
18 m <- c(b(0.1,0)*pi[1]+b(0.2,0)*pi[2],b(0.1,1)*pi[1]+
  b(0.2,1)*pi[2],b(0.1,2)*pi[1]+b(0.2,2)*pi[2])
19
20 post_p1 <- c(b(0.1,0)*pi[1]/m[1],b(0.1,1)*pi[1]/m
  [2],b(0.1,2)*pi[1]/m[3])
21
22 post_p2 <- c(1,1,1)-post_p1
23
24 cat("The bayes estimate of p for x = 0 is",p1[1]*
  post_p1[1]+p1[2]*post_p2[1])
25
26 cat("The bayes estimate of p for x = 1 is",p1[1]*
  post_p1[2]+p1[2]*post_p2[2])
27
28 cat("The bayes estimate of p for x = 2 is",p1[1]*
  post_p1[3]+p1[2]*post_p2[3])

```

---

**R code Exa 18.9** To find Bayes estimates of p

```

1
2 # Chapter 18
3 # example 18.9 page no. 717 from the pdf..
4 # to find the bayes estimate of p for all values of
   x, for example 18.2..
5
6 # refer example 18.2
7 # refer section 6.8 on page 201...
8 # the posterior can also be written as  $B(x+1,3-x)$ 
9
10 f <- function(x){
11
12     3*choose(2,x)
13
14 }
15
16 g <- function(p,a){
17
18     (p^(a+1))*((1-p)^(2-a))
19
20 }
21 cat("The bayes estimate for the above distribution
   for different values of x is")
22
23 for(x in 0:2){
24
25     cat("for x =",x,"p star is",f(x)*integrate(
       function(d){g(d,x)},0,1)$value,"\n")
26
27 }

```

---

**R code Exa 18.11** To find Bayes estimates of  $p$  under Absolute Error loss

```

1
2 # Chapter 18

```

```

3 # Example 18.11 page no. 718 from the pdf..
4 # to find the bayes estimate for example 18.9, under
  absolute error loss, when  $x = 1$  is observed..
5
6 # Note – Example 18.10 is theoretical with nothing
  to compute..
7
8 # refer example 18.9..
9 # package used "polynom" if not installed can be
  installed by executing install.packages("polynom
  ") on console..
10
11 library(polynom)
12
13 # we see that for  $x = 1$  the distribution becomes
  ,
14 #  $6x(1-x)$ 
15
16 p <- polynomial(coef = c(0,6,-6))
17
18 eq <- integral(p)- 0.5
19
20 cat("The bayes estimate under absolute error loss
  for  $x = 1$  is",Re(polyroot(eq)[1]))

```

---