R Textbook Companion for Statistics for Management and Economics by Gerald Keller[1]

Created by
Sheeba Sekharan
M.Sc.
Mathematics
IIT Bombay
Cross-Checked by
R TBC Team

May 30, 2020

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

Book Description

Title: Statistics for Management and Economics

Author: Gerald Keller

Publisher: Cengage Learning USA

Edition: 9

Year: 2012

ISBN: 9788131528426

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

Exa Example (Solved example)

Eqn Equation (Particular equation of the above book)

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

Contents

List of R Codes

Chapter 2

Graphical Descriptive Techniques I

R code Exa 2.1 Work Status in the GSS 2008 Survey

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 2: Graphical Descriptive Techniques I
# Example 2.1 on Pg 18
# Work Status in the GSS 2008 Survey

# Complete dataset of 2021 observations could not be
     found on the
# website: https://www.cengage.com/cgi-wadsworth/
    course_products_wp.pl?fid=M20b&product_isbn_issn
    =9781285425450&template=nelson
# Partial data of 150 observations found in the book
     as given below:
data <- c(1, 1, 1, 1, 1, 7, 7, 1, 1, 5, 1, 5, 7, 1,
     1, 5, 7, 1, 5, 2, 5, 1, 5, 8, 1, 5, 7, 1, 4, 2,
    7, 1, 2,
            1, 1, 2, 1, 7, 1, 7, 1, 2, 1, 1, 1, 1, 1,
                 6, 5, 1, 1, 1, 1, 1, 2, 5, 2, 7, 2,
               7, 8, 1, 8, 1, 7, 1,
```

```r
              6, 7, 6, 1, 5, 1, 2, 2, 4, 1, 1, 1, 1, 1, 6, 5, 5, 3, 2, 1, 1, 8 ,1 ,5, 1, 1, 1, 1, 5, 5, 1, 5, 4,
              7, 1, 1, 1, 4, 5, 2, 5, 6, 7 ,7, 1, 4, 2, 1, 2, 6, 1, 1, 1, 1, 1, 1, 7, 4, 1, 1, 1, 7, 8, 1, 3, 1,
              1, 3, 1, 1, 1, 1, 1, 1, 2, 1, 5, 1, 1, 1, 1, 1, 2, 1)

# factor() function divides the dataset into its levels
f <- factor(data)

# levels() function used for renaming
levels(f) <- c('Working full-time', 'Working part-time', 'Temporary no work', 'laid off',
               'Retired', 'School', 'Keeping house', 'Other')

# table() function displays the frequency table
c <- table(f)
print(c) #gives frequencies
rel_c <- paste(round(prop.table(c)*100,2), "%", sep="") #gives relative frequencies in %
cbind(c, rel_c) #Table showing both frequencies and relative frequencies

# barplot() function plots the bar graph using the frequency table
barplot(c, main="Work Status", las=0)

# for pie-chart
# pie() function plots the pie chart using the frequency table
pct <- round(c/sum(c)*100) #computing percentages
lbls <- paste(levels(f), pct) #add percents to labels
lbls <- paste(lbls, "%", sep="") #add % to labels
```

```
pie(c, labels = lbls, main ="Pie Chart of Work Status")

#End
```

R code Exa 2.2 Energy Consumption in the United States in 2007

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 2: Graphical Descriptive Techniques I
# Example 2.2 on Pg 24
# Energy Consumption in the United States in 2007


data1_source <- c("Petroleum", "NaturalGas", "Coal", "Nuclear", "Hydroelec",
                        "Wood", "Biofuels", "Wind", "Waste", "Geotherm", "Solar")
data1_BTU <- c(39.77, 23.64, 22.8, 8.42, 2.45, 2.14, 1.02, 0.34, 0.43, 0.35, 0.08)

dev.off()

# the appropriate graphical technique, in describing the proportion of total energy consumption by all sources,
# is a pie-chart.
# pie-chart using pie() function
colors <- c("beige", "dodgerblue", "hotpink4", "navy", "lawngreen", "lightslategrey", "purple3", "red", "yellow", "black", "orange")
lbls <- paste(data1_BTU,"%", sep="")
pie_legend <- paste(data1_source, lbls)
pie(data1_BTU, main ="Pie Chart of Energy Consumption", cex=0.7, col=colors, labels = NA)
```

```
legend(x=0.75,y=0.75,legend =pie_legend, fill=colors, bty="n", cex=0.6)


#End
```

R code Exa 2.3 Per Capita Beer Consumption 10 Selected Countries

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 2: Graphical Descriptive Techniques I
# Example 2.3 on Pg 26
# Per Capita Beer Consumption (10 Selected Countries)


Country <- c("Australia","Austria","Belgium","Canada","Croatia","Czech Republic","Denmark","Finland","Germany",
             "Hungary","Ireland","Luxembourg","Netherlands","New Zealand","Poland","Portugal","Slovakia","Spain",
             "UK","USA")

Beer_consumption <- c(119.2,106.3,93,68.3,81.2,138.1,89.9,85,147.8,75.3,138.3,84.4, 79, 77, 69.1, 59.6,
                      84.1, 83.8, 96.8, 81.6)

#bar chart for beer consumption
barchart <- barplot(Beer_consumption, names.arg = Country, ylim=c(0,160),axisnames = FALSE,
                    main=" Per Capita Beer Consumption (10 Selected Countries)")
```

```
text(x = barchart, y = Beer_consumption, label =
    Beer_consumption, pos = 3, cex = 0.9, col = "red"
    )
text(x = barchart, y = par()$usr[3], label = Country
     ,srt = 45, adj = c(1.1,1.1), xpd = TRUE) #
    rotated x-axisnames

#End
```

Chapter 3

Graphical Descriptive Techniques II

Chapter 4

Numerical Descriptive Techniques

R code Exa 4.1 Mean Time Spent on the Internet

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 4: Numerical Descriptive Techniques
# Example 4.1 on Pg 99
# Mean Time Spent on the Internet

Internet_hours <- c(0, 7, 12, 5, 33, 14, 8, 0, 9,
    22)

#manually computing the sample mean of Internet
    hours
mean1 <- sum(Internet_hours)/length(Internet_hours)

#computing mean using function
mean2 <- mean(Internet_hours)

#Answer: sample mean is 11

#End
```

R code Exa 4.3 Median Time Spent on Internet

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 4: Numerical Descriptive Techniques
# Example 4.3 on Pg 100
# Median Time Spent on Internet

Internet_hours <- c(0, 7, 12, 5, 33, 14, 8, 0, 9,
   22)

#computing median using the function median()
median(Internet_hours)

#Answer: sample median is 8.5

#End
```

R code Exa 4.5 Mode Time Spent on Internet

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 4: Numerical Descriptive Techniques
# Example 4.5 on Pg 102
# Mode Time Spent on Internet

Internet_hours <- c(0, 7, 12, 5, 33, 14, 8, 0, 9,
   22)
```

```
# there is no inbuilt function for calculating Mode
# So, a function is written to compute Mode. It
    works if the data is numeric.
# It creates a frequency table using the function
    table() and gives the index of the value occuring
     maximum
# times using the function which.max().
Mode <- function(x)
{
  if (is.numeric(x))
  {
    x_table <- table(x)
    return(as.numeric(names(x_table)[which.max(x_
       table)]))
  }
}

Mode(Internet_hours)

#Answer: sample mode is 0

#End
```

R code Exa 4.7 Summer Jobs

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 4: Numerical Descriptive Techniques
# Example 4.7 on Pg 110
# Summer Jobs

summer_job <-  c(17, 15, 23, 7, 9, 13)
```

```
#Find the mean and variance of these data.

#Sample Mean
mean(summer_job)

#Sample Variance
var(summer_job)

#Answer: Mean is 14 & Variance is 33.2

#End
```

R code Exa 4.9 Using the Empirical Rule to Interpret Standard Deviation

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 4: NUMERICAL DESCRIPTIVE TECHNIQUES
# Example 4.9 on Pg. 113
# Using the Empirical Rule to Interpret Standard Deviation


population_mean <- 0.1
population_sd <- 0.08
sd1 <- 1
sd2 <- 2
sd3 <- 3

lower_bound_one_SD <- population_mean - population_sd * sd1 #Answer: 2%
upper_bound_one_SD <- population_mean + population_sd * sd1 #Answer: 18%
```

```
probability_within_bounds1 <- pnorm(upper_bound_one_SD, population_mean, population_sd) - pnorm(lower_bound_one_SD, population_mean, population_sd)
#Answer: 68%

lower_bound_two_SD <- population_mean - population_sd * sd2 #Answer: -6%
upper_bound_two_SD <- population_mean + population_sd * sd2 #Answer: 26%
probability_within_bounds2 <- pnorm(upper_bound_two_SD, population_mean, population_sd) - pnorm(lower_bound_two_SD, population_mean, population_sd)
#Answer: 95%

lower_bound_three_SD <- population_mean - population_sd * sd3 #Answer: -14%
upper_bound_three_SD <- population_mean + population_sd * sd3 #Answer: 34%
probability_within_bounds3 <- pnorm(upper_bound_three_SD, population_mean, population_sd) - pnorm(lower_bound_three_SD, population_mean, population_sd)
#Answer: 99.7%

cat("Given the histogram is bell shaped, we can apply the Empirical Rule and say that:")
cat("1. Approximately", paste(round((probability_within_bounds1)*100,digits=0),"%",sep=""),
  "of the returns on investment lie between", paste(round((lower_bound_one_SD)*100,digits=0),"%",sep=""),
  "and",
  paste(round((upper_bound_one_SD)*100,digits=0),"%",sep=""))

cat("2. Approximately", paste(round((probability_within_bounds2)*100,digits=0),"%",sep=""),
  "of the returns on investment lie between",
```

```
      paste(round((lower_bound_two_SD)*100,digits=0),"%",sep=""),
    "and",
    paste(round((upper_bound_two_SD)*100,digits=0),"%",sep=""))

cat("3. Approximately", paste(round((probability_within_bounds3)*100,digits=1),"%",sep=""),
    "of the returns on investment lie between", paste(round((lower_bound_three_SD)*100,digits=0),"%",sep=""),
    "and",
    paste(round((upper_bound_three_SD)*100,digits=0),"%",sep=""))

#End
```

R code Exa 4.10 Using Chebysheffs Theorem to Interpret Standard Deviation

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 4: NUMERICAL DESCRIPTIVE TECHNIQUES
# Example 4.10 on Pg. 114
# Using Chebysheff's Theorem to Interpret Standard Deviation


population_mean <- 28000
population_sd <- 3000

sd2 <- 2 #two standard deviations
sd3 <- 3 #three standard deviations
chebyshev_2 <- 1- 1/(sd2^2)
chebyshev_3 <- 1- 1/(sd3^2)
```

```
lower_bound_two_SD <- population_mean - population_
    sd * sd2 #Answer: $22,000
upper_bound_two_SD <- population_mean + population_
    sd * sd2 #Answer: $34,000

lower_bound_three_SD <- population_mean - population
    _sd * sd3 #Answer: $19,000
upper_bound_three_SD <- population_mean + population
    _sd * sd3 #Answer: $34,000

cat("Given the histogram is NOT bell shaped, we can
    only apply the Chebyshev's Thoerem and say that:"
    )

cat("1. Atleast", paste(round(chebyshev_2*100,digits
    =0),"%",sep=""),
     "of the returns on investment lie between",
        round(lower_bound_two_SD),
     "and",
     round(upper_bound_two_SD))

cat("2. Atleast", paste(round(chebyshev_3*100,digits
    =1),"%",sep=""),
     "of the returns on investment lie between",
        round(lower_bound_three_SD),
     "and",
     round(upper_bound_three_SD))

#End
```

R code Exa 4.11 Percentiles of Time Spent on Internet

```
# Statistics for Management and Economics by Gerald
    Keller
```

```
# Chapter 4: Numerical Descriptive Techniques
# Example 4.11 on Pg 118
# Percentiles of Time Spent on Internet

Internet_hours <- c(0, 7, 12, 5, 33, 14, 8, 0, 9,
   22)

quantile(Internet_hours, probs = c(.25, .50, .75),
   type=6)

#Answer: 25%    50%     75%
        #3.75   8.50   16.00

#End
```

R code Exa 4.16 Calculating the Coefficient of Correlation

```
# Statistics for Management and Economics by Gerald
   Keller
# Chapter 4: Numerical Descriptive Techniques
# Example 4.16 on Pg 129
# Calculating the Coefficient of Correlation

#Set 1
x1 <- c(2,6,7)
y1 <- c(13,20,27)
cor(x1,y1)
```

```
#Answer: Correlation coefficient for Set 1: 0.9449112

#Set 2
x2 <- c(2,6,7)
y2 <- c(27,20,13)
cor(x2,y2)
#Answer: Correlation coefficient for Set 2: -0.9449112

#Set 3
x3 <- c(2,6,7)
y3 <- c(20,27,13)
cor(x3,y3)
#Answer: Correlation coefficient for Set 3: -0.1889822

#End
```

Chapter 5

Data Collection and Sampling

R code Exa 5.1 Random Sample of Income Tax Returns

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 5: Data Collection and Sampling
# Example 5.1 on Pg. 168
# Random Sample of Income Tax Returns

sample(1:1000, 40, replace=TRUE) #random sample
    generation with replacement
sample(1:1000, 40, replace=FALSE) #random sample
    generation without replacement

#End
```

Chapter 6

Probability

R code Exa 6.1 Determinants of Success among Mutual Fund Managers Part 1

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 6: PROBABILITY
# Example 6.1 on Pg. 182
# Determinants of Success among Mutual Fund Managers
    -Part 1


#Denote:
#A1 = Fund manager graduated from a top-20 MBA
    program
#A2 = Fund manager did not graduate from a top-20
    MBA program
#B1 = Fund outperforms the market
#B2 = Fund does not outperform the market

#Given:
#P(A1 and B1) = 0.11
#P(A2 and B1) = 0.06
#P(A1 and B2) = 0.29
```

```
#P(A2 and B2) = 0.54

p_A1_B1 = 0.11
p_A2_B1 = 0.06
p_A1_B2 = 0.29
p_A2_B2 = 0.54

#P(A1) = P(A1 and B1) + P(A1 and B2)
p_A1 = p_A1_B1 + p_A1_B2
#Answer: P(A1) = 0.4

#P(A2) = P(A2 and B1) + P(A2 and B2)
p_A2 = p_A2_B1 + p_A2_B2
#Answer: P(A2) = 0.6

#P(B1) = P(A2 and B1) + P(A1 and B1)
p_B1 = p_A2_B1 + p_A1_B1
#Answer: P(B1) = 0.17

#P(B2) = P(A2 and B2) + P(A1 and B2)
p_B2 = p_A2_B2 + p_A1_B2
#Answer: P(B2) = 0.83


#End
```

R code Exa 6.2 Determinants of Success among Mutual Fund Managers Part 2

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 6: PROBABILITY
# Example 6.2 on Pg. 184
# Determinants of Success among Mutual Fund Managers-Part 2
```

```
#Denote:
#A1 = Fund manager graduated from a top-20 MBA program
#A2 = Fund manager did not graduate from a top-20 MBA program
#B1 = Fund outperforms the market
#B2 = Fund does not outperform the market

#Given:
#P(A1 and B1) = 0.11
#P(A2 and B1) = 0.06
#P(A1 and B2) = 0.29
#P(A2 and B2) = 0.54

p_A1_B1 = 0.11
p_A2_B1 = 0.06
p_A1_B2 = 0.29
p_A2_B2 = 0.54

#Find P(A1/B2)

p_A1_given_B2 =  p_A1_B2/ (p_A2_B2 + p_A1_B2)
#Answer: P(A1/B2) = 0.3494

cat("34.9% of all mutual funds that do not outperform the market are managed by top-20 MBA program graduates.")

#End
```

R code Exa 6.3 Determinants of Success among Mutual Fund Managers Part 3

```
# Statistics for Management and Economics by Gerald
```

```
Keller
# Chapter 6: PROBABILITY
# Example 6.3 on Pg. 185
# Determinants of Success among Mutual Fund Managers -Part 3


#Denote:
#A1 = Fund manager graduated from a top-20 MBA program
#A2 = Fund manager did not graduate from a top-20 MBA program
#B1 = Fund outperforms the market
#B2 = Fund does not outperform the market

#Given:
#P(A1 and B1) = 0.11
#P(A2 and B1) = 0.06
#P(A1 and B2) = 0.29
#P(A2 and B2) = 0.54

p_A1_B1 = 0.11
p_A2_B1 = 0.06
p_A1_B2 = 0.29
p_A2_B2 = 0.54

#determine whether A1 and B1 are independent

p_A1_given_B1 =  p_A1_B1/ (p_A2_B1 + p_A1_B1)
p_A1 = p_A1_B1 + p_A1_B2

cat("P(A1/B1) =", p_A1_given_B1)
cat("P(A1) =", p_A1)

if(p_A1 == p_A1_given_B1)
{cat("A1 and B1 are independent since P(A1/B1) and P(A1) have same value")}else
   {cat("A1 and B1 are not independent since P(A1/
```

```
B1) and P(A1) do not have same value")}

#End
```

R code Exa 6.4 Determinants of Success among Mutual Fund Managers Part 4

```
# Statistics for Management and Economics by Gerald
  Keller
# Chapter 6: PROBABILITY
# Example 6.4 on Pg. 186
# Determinants of Success among Mutual Fund Managers
  -Part 4


#Denote:
#A1 = Fund manager graduated from a top-20 MBA
  program
#A2 = Fund manager did not graduate from a top-20
  MBA program
#B1 = Fund outperforms the market
#B2 = Fund does not outperform the market

#Given:
#P(A1 and B1) = 0.11
#P(A2 and B1) = 0.06
#P(A1 and B2) = 0.29
#P(A2 and B2) = 0.54

p_A1_B1 = 0.11
p_A2_B1 = 0.06
p_A1_B2 = 0.29
p_A2_B2 = 0.54

#Find P(A1 or B1) i.e., P(A1 union B1)
```

```
#P(A1 or B1) = 1 - P(A2 and B2)

p_A1_or_B1 = 1 - p_A2_B2
#Answer: 0.46

cat("Thus,", paste(round(p_A1_or_B1*100), "%", sep=""),"of mutual funds either outperform the market or are managed by a top-20 MBA program graduate
  or have both characteristics. ")

#End
```

R code Exa 6.5 Selecting Two Students without Replacement

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 6: PROBABILITY
# Example 6.5 on Pg. 192
# Selecting Two Students without Replacement

#A is the event that the first student chosen is female
#B is the event that the second student chosen is also female.

#Find P(A and B) without replacement

#Given:
number_of_males = 7
number_of_females = 3

p_A = number_of_females/(number_of_females + number_of_males)
p_B_given_A = (number_of_females-1)/((number_of_females + number_of_males)-1) #without
```

```
replacement

p_A_and_B = p_A * p_B_given_A
#Answer: 0.06666667

cat("Probability that the two students chosen are female:", p_A_and_B)

#End
```

R code Exa 6.6 Selecting Two Students with Replacement

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 6: PROBABILITY
# Example 6.6 on Pg. 193
# Selecting Two Students with Replacement

#A is the event that the first student chosen is female
#B is the event that the second student chosen is also female.

#Find P(A and B) with replacement

#Given:
number_of_males = 7
number_of_females = 3

p_A = number_of_females/(number_of_females + number_of_males)
p_B = number_of_females/(number_of_females + number_of_males) #with replacement

p_A_and_B = p_A * p_B
```

```
#Answer: 0.09

cat("Probability that the two students chosen are
    female:", p_A_and_B)

#End
```

R code Exa 6.7 Applying the Addition Rule

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 6: PROBABILITY
# Example 6.7 on Pg. 194
# Applying the Addition Rule

#A = the household subscribes to the Sun
#B = the household subscribes to the Post

#Given P(A) = 0.22, P(B) = 0.35 and P(A and B) =
    0.06
#Find P(A union B) i.e., P(A or B)

p_A = 0.22
p_B = 0.35
p_A_and_B = 0.06

#Addition rule: P(A union B) = P(A) + P(B) - P(A and
    B)
p_A_or_B = p_A + p_B - p_A_and_B
#Answer: 0.51

cat("The probability that a randomly selected
    household subscribes to either newspaper is", p_A
    _or_B)

```

```
#End
```

R code Exa 6.8 Probability of Passing the Bar Exam

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 6: PROBABILITY
# Example 6.8 on Pg. 196
# Probability of Passing the Bar Exam

#Given:
#P(pass rate for first-time Bar Exam takers) = 0.72
#P(pass rate for second-time Bar Exam takers who
    failed first time) = 0.88

pass_1 = 0.72
fail_1 = 1-pass_1

pass2_Given_fail1 = 0.88
#fail_and_pass = P(Fail [on first exam] and Pass [on
    second exam])

fail1_and_pass2 = pass2_Given_fail1 * fail_1
#Answer: P(Fail [on first exam] and Pass [on second
    exam]) = 0.2464

#We need probability that a randomly selected law
    school graduate becomes a lawyer i.e.,
#we need to find probability of passing the first or
    second exam.

pass = pass_1 + fail1_and_pass2

cat("probability that a randomly selected law school
    graduate becomes a lawyer:", pass)
```

```

#End
```

R code Exa 6.9 Should an MBA Applicant Take a Preparatory Course

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 6: PROBABILITY
# Example 6.9 on Pg. 199
# Should an MBA Applicant Take a Preparatory Course?

#A1 = GMAT score is 650 or more
#A2 = GMAT score less than 650
#B = Take preparatory course

#Given:
#P(B given A1) = .52
#P(A1) = p_A1 = 0.1
#P(B given A2) = .23

#Find P(A1/B)

p_A1 = 0.1
p_A2 = 1 - p_A1
p_B_given_A1 = 0.52
p_B_given_A2 = 0.23

#BAYE'S Rule:
#P(A1 given B) =  P(A1)*P(B given A1) / (P(A1)*P(B
    given A1) + P(A2)*P(B given A2))

p_A1_given_B = (p_A1*p_B_given_A1) / (p_A1*p_B_given
    _A1 + p_A2*p_B_given_A2)
#Answer: 0.2007722

```

```
#End
```

R code Exa 6.10 Probability of Prostate Cancer

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 6: PROBABILITY
# Example 6.10 on Pg. 203
# Probability of Prostate Cancer

#Given:
#Prior: P(Has Prostrate Cancer) = .010
#Given Likelihood probabilities
#True negative:   P(Negative test GIVEN No Prostrate Cancer) = 1 - .135 = .865
#False positive:  P(Positive test GIVEN No Prostrate Cancer) = .135
#True positive:   P(Positive test GIVEN Prostrate Cancer) = 1 - .300 = .700
#False negative:  P(Negative test GIVEN Prostrate Cancer) = .300


#Function 'bayes_probability_tree' that creates a Probability Tree using Bayes Theorem

install.packages("DiagrammeR")
library(DiagrammeR)

bayes_probability_tree <- function(prior, true_positive, true_negative) {

  if (!all(c(prior, true_positive, true_negative) > 0) && !all(c(prior, true_positive, true_negative) < 1)) {
```

```r
    stop("probabilities must be greater than 0 and less than 1.",
         call. = FALSE)
  }
  c_prior <- 1 - prior
  c_tp <- 1 - true_positive
  c_tn <- 1 - true_negative

  round4 <- purrr::partial(round, digits = 4)

  b1 <- round4(prior * true_positive)
  b2 <- round4(prior * c_tp)
  b3 <- round4(c_prior * c_tn)
  b4 <- round4(c_prior * true_negative)

  bp <-  round4(b1/(b1 + b3))

  labs <- c("Cancer", prior, c_prior, true_positive, c_tp, true_negative, c_tn, b1, b2, b4, b3)

  tree <-
    create_graph() %>%
    add_n_nodes(
      n = 11,
      type = "path",
      label = labs,
      node_aes = node_aes(
        shape = "circle",
        height = 1,
        width = 1,
        x = c(0, 3, 3, 6, 6, 6, 6, 8, 8, 8, 8),
        y = c(0, 2, -2, 3, 1, -3, -1, 3, 1, -3, -1))
      ) %>%
    add_edge(
      from = 1,
      to = 2,
      edge_aes = edge_aes(
        label = "Has Prostrate Cancer"
```

```
      )
    ) %>%
    add_edge(
      from = 1,
      to = 3,
      edge_aes = edge_aes(
        label = "Does not have Prostrate Cancer"
      )
    ) %>%
    add_edge(
      from = 2,
      to = 4,
      edge_aes = edge_aes(
        label = "True Positive: Positive test GIVEN
           Cancer"
      )
    ) %>%
    add_edge(
      from = 2,
      to = 5,
      edge_aes = edge_aes(
        label = "False Negative: Negative test GIVEN
           Cancer"
      )
    ) %>%
    add_edge(
      from = 3,
      to = 7,
      edge_aes = edge_aes(
        label = "False Positive: Positive test GIVEN
           NO Cancer "
      )
    ) %>%
    add_edge(
      from = 3,
      to = 6,
      edge_aes = edge_aes(
        label = "True Negative: Negative test GIVEN
```

```
NO Cancer"
        )
      ) %>%
      add_edge(
        from = 4,
        to = 8,
        edge_aes = edge_aes(
          label = "="
        )
      ) %>%
      add_edge(
        from = 5,
        to = 9,
        edge_aes = edge_aes(
          label = "="
        )
      ) %>%
      add_edge(
        from = 7,
        to = 11,
        edge_aes = edge_aes(
          label = "="
        )
      ) %>%
      add_edge(
        from = 6,
        to = 10,
        edge_aes = edge_aes(
          label = "="
        )
      )
  message(glue::glue("The probability that the man
     has prostate cancer given a positive test
     result is {bp}"))
  print(render_graph(tree))
  invisible(tree)
}
```

```
bayes_probability_tree(prior = 0.01, true_positive =
    0.7, true_negative = (1-0.135))

#End
```

Chapter 7

Random Variables and Discrete Probability Distributions

R code Exa 7.1 Probability Distribution of Persons per Household

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS
# Example 7.1 on Pg. 220
# Probability Distribution of Persons per Household

#X is used to denote the random variable, the number of persons per household.
#Develop the probability distribution of X.

Number_of_Persons <- c(1,2,3,4,5,6,7)
Number_of_Households <- c(31.1, 38.6, 18.8, 16.2, 7.2, 2.7, 1.4)

#we need Probability of X i.e., the relative frequency. Let it be denoted by P_X

P_X <- round(Number_of_Households/sum(Number_of_
```

```
    Households), digits=3)

#Answer: P(X): 0.268 0.333 0.162 0.140 0.062 0.023
    0.012

#End
```

R code Exa 7.2 Probability Distribution of the Number of Sales

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE
    PROBABILITY DISTRIBUTIONS
# Example 7.2 on Pg. 221
# Probability Distribution of the Number of Sales

# Denote:
# X = the number of sales
# prob = P(success) = 0.2
# q = P(failure) = 0.8
# three trials

ProbofSales <- function(q)
{
  p = pbinom(q, size = 3, prob = 0.2, lower.tail =
      TRUE)
  return(p)
}

#p_0 = P(X=0)
p_0 = ProbofSales(0)
#p_1 = P(X=1)
p_1 = ProbofSales(1) - p_0
#p_2 = P(X=2)
p_2 = ProbofSales(2) - ProbofSales(1)
```

```
#p_3 = P(X=3)
p_3 = ProbofSales(3) - ProbofSales(2)

cat("The Probability Distribution of number of Sales
   :")
cat("P(Number of Sales is 0):", p_0) #Answer: 0.512
cat("P(Number of Sales is 1):", p_1) #Answer: 0.384
cat("P(Number of Sales is 2):", p_2) #Answer: 0.096
cat("P(Number of Sales is 3):", p_3) #Answer: 0.008

#End
```

R code Exa 7.3 Describing the Population of the Number of Persons per Household

```
# Statistics for Management and Economics by Gerald
   Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE
   PROBABILITY DISTRIBUTIONS
# Example 7.3 on Pg. 224
# Describing the Population of the Number of Persons
    per Household


#X is used to denote the random variable, the number
    of persons per household.
#Find the mean, variance, and standard deviation for
    the population of the number of persons per
   household

Number_of_Persons <- c(1,2,3,4,5,6,7)
Number_of_Households <- c(31.1, 38.6, 18.8, 16.2,
   7.2, 2.7, 1.4)

#we need Probability of X i.e., the relative
```

```
    frequency. Let it be denoted by P_X
P_X <- round(Number_of_Households/sum(Number_of_
    Households), digits=3)

E_X <- sum(P_X*Number_of_Persons)
V_X <- sum(((Number_of_Persons-E_X)^2)*P_X)
STDEV <- sqrt(V_X)

#Answer: E(X) = 2.512
        #Var(X) = 1.9539
        #Std deviation (X) = 1.3978


#End
```

R code Exa 7.4 Describing the Population of Monthly Profits

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE
    PROBABILITY DISTRIBUTIONS
# Example 7.4 on Pg. 225
# Describing the Population of Monthly Profits

#Given:
mean_sales = 25000  #mean of monthly sales at a
    computer store
stdev_sales = 4000 #standard deviation of monthly
    sales at a computer store

#Given fixed cost:
fc = 6000

#Laws of Expected Value: E(c) = c; E(X + c) = E(X) +
     c; E(cX) = c*E(X)
```

```
#Laws of Variance: V(X + c) = V(X); V(cX) = c^2*V(X); V(c)=0

#Given: Profit = 0.3*Sales - fixed cost.

#Applying the laws of expected value, E(Profit) = 0.3*E(Sales) - 6000
#Applying the laws of variance, V(Profit) = V(0.30(Sales) - 6,000) = 0.09V(Sales)

expected_profit = 0.3*mean_sales - fc
#Answer: 1500
stdev_profit = sqrt(0.09*stdev_sales^2)
#Answer: 1200

#End
```

R code Exa 7.5 Bivariate Distribution of the Number of House Sales

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS
# Example 7.5 on Pg. 230
# Bivariate Distribution of the Number of House Sales


# X = number of houses that Xavier will sell in a month
# Y = number of houses Yvette will sell in a month.

# bivariate probability distribution of X & Y
matr=matrix(c(0.12, 0.21, 0.07, 0.42, 0.06, 0.02, 0.06, 0.03, 0.01),3,3)
```

```
#Marginal probabilities of Y
Y_marginal <- margin.table(matr, 1)
Y_marginaltable <- matrix(c(0,1,2, Y_marginal),3,2)
colnames(Y_marginaltable) <- c('Y', 'P(Y)')
rownames(Y_marginaltable) <- c('', '', '')
Y_marginaltable

#Expected value of Y, E(Y):
Expected_Y = X_marginaltable[1]*Y_marginaltable[4] + Y_marginaltable[2]*Y_marginaltable[5] +
             Y_marginaltable[3]*Y_marginaltable[6]
Expected_Y
#Answer: 0.5

#Variance(Y):
Var_Y = (Y_marginaltable[1]-Expected_Y)^2*Y_marginaltable[4] +
        (Y_marginaltable[2]-Expected_Y)^2*Y_marginaltable[5] +
        (Y_marginaltable[3]-Expected_Y)^2*Y_marginaltable[6]
Var_Y
#Answer: 0.45

#Standard Deviation of Y
Std_Y = sqrt(Var_Y)
#Answer: 0.6708204

####################################

#Marginal probabilities of X
X_marginal <- margin.table(matr, 2)
X_marginaltable <- matrix(c(0,1,2, X_marginal),3,2)
colnames(X_marginaltable) <- c('X', 'P(X)')
rownames(X_marginaltable) <- c('', '', '')
X_marginaltable
```

```
#Expected value of X, E(X):
Expected_X = X_marginaltable[1]*X_marginaltable[4] + X_marginaltable[2]*X_marginaltable[5] +
  X_marginaltable[3]*X_marginaltable[6]
Expected_X
#Answer: 0.7

#Variance(X):
Var_X = (X_marginaltable[1]-Expected_X)^2*X_marginaltable[4] +
  (X_marginaltable[2]-Expected_X)^2*X_marginaltable[5] +
  (X_marginaltable[3]-Expected_X)^2*X_marginaltable[6]
Var_X
#Answer: 0.41

#Standard Deviation of X
Std_X = sqrt(Var_X)
#Answer: 0.6403124


#End
```

R code Exa 7.6 Describing the Bivariate Distribution

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS
# Example 7.6 on Pg. 232
# Describing the Bivariate Distribution


# X = number of houses that Xavier will sell in a
```

```
month
# Y = number of houses Yvette will sell in a month.

# bivariate probability distribution of X & Y
matr=matrix(c(0.12, 0.21, 0.07, 0.42, 0.06, 0.02, 0.06, 0.03, 0.01),3,3)

#Marginal probabilities of Y
Y_marginal <- margin.table(matr, 1)
Y_marginaltable <- matrix(c(0,1,2, Y_marginal),3,2)
colnames(Y_marginaltable) <- c('Y', 'P(Y)')
rownames(Y_marginaltable) <- c('', '', '')
Y_marginaltable

#Expected value of Y, E(Y):
Expected_Y = X_marginaltable[1]*Y_marginaltable[4] + Y_marginaltable[2]*Y_marginaltable[5] +
  Y_marginaltable[3]*Y_marginaltable[6]
Expected_Y
#Answer: 0.5

#Variance(Y):
Var_Y = (Y_marginaltable[1]-Expected_Y)^2*Y_marginaltable[4] +
  (Y_marginaltable[2]-Expected_Y)^2*Y_marginaltable[5] +
  (Y_marginaltable[3]-Expected_Y)^2*Y_marginaltable[6]
Var_Y
#Answer: 0.45

#Standard Deviation of Y
Std_Y = sqrt(Var_Y)
#Answer: 0.6708204

#####################################

#Marginal probabilities of X
```

```
X_marginal <- margin.table(matr, 2)
X_marginaltable <- matrix(c(0,1,2, X_marginal),3,2)
colnames(X_marginaltable) <- c('X', 'P(X)')
rownames(X_marginaltable) <- c('', '', '')
X_marginaltable

#Expected value of X, E(X):
Expected_X = X_marginaltable[1]*X_marginaltable[4] + X_marginaltable[2]*X_marginaltable[5] +
  X_marginaltable[3]*X_marginaltable[6]
Expected_X
#Answer: 0.7

#Variance(X):
Var_X = (X_marginaltable[1]-Expected_X)^2*X_marginaltable[4] +
  (X_marginaltable[2]-Expected_X)^2*X_marginaltable[5] +
  (X_marginaltable[3]-Expected_X)^2*X_marginaltable[6]
Var_X
#Answer: 0.41

#Standard Deviation of X
Std_X = sqrt(Var_X)
#Answer: 0.6403124


#####################

#Covariance(X,Y):
cov_x_y = (Y_marginaltable[1]-Expected_Y)*(X_marginaltable[1]-Expected_X)*0.12+(Y_marginaltable[1]-Expected_Y)*(X_marginaltable[2]-Expected_X)*0.42+(Y_marginaltable[1]-Expected_Y)*(X_marginaltable[3]-Expected_X)*0.06+(Y_marginaltable[2]-Expected_Y)*(X_marginaltable[1]-Expected_X)*0.21+(Y_marginaltable[2]-Expected_Y)*
```

```
(X_marginaltable[2]-Expected_X)*0.06+(Y_marginaltable[2]-Expected_Y)*(X_marginaltable[3]-Expected_X)*0.03+(Y_marginaltable[3]-Expected_Y)*(X_marginaltable[1]-Expected_X)*0.07+(Y_marginaltable[3]-Expected_Y)*(X_marginaltable[2]-Expected_X)*0.02+(Y_marginaltable[3]-Expected_Y)*(X_marginaltable[3]-Expected_X)*0.01
cov_x_y
#Answer: -0.15


#Correlation(X,Y)
corr_x_y = cov_x_y/(Std_X*Std_Y)
corr_x_y
#Answer: -0.3492151


#End
```

R code Exa 7.7 Describing the Population of the Total Number of House Sales

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS
# Example 7.7 on Pg. 234
# Describing the Population of the Total Number of House Sales


# X = number of houses that Xavier will sell in a month
# Y = number of houses Yvette will sell in a month.

```

```r
# bivariate probability distribution of X & Y
matr=matrix(c(0.12, 0.21, 0.07, 0.42, 0.06, 0.02, 0.06, 0.03, 0.01),3,3)

#Marginal probabilities of Y
Y_marginal <- margin.table(matr, 1)
Y_marginaltable <- matrix(c(0,1,2, Y_marginal),3,2)
colnames(Y_marginaltable) <- c('Y', 'P(Y)')
rownames(Y_marginaltable) <- c('', '', '')
Y_marginaltable

#Expected value of Y, E(Y):
Expected_Y = X_marginaltable[1]*Y_marginaltable[4] + Y_marginaltable[2]*Y_marginaltable[5] +
  Y_marginaltable[3]*Y_marginaltable[6]
Expected_Y
#Answer: 0.5

#Variance(Y):
Var_Y = (Y_marginaltable[1]-Expected_Y)^2*Y_marginaltable[4] +
  (Y_marginaltable[2]-Expected_Y)^2*Y_marginaltable[5] +
  (Y_marginaltable[3]-Expected_Y)^2*Y_marginaltable[6]
Var_Y
#Answer: 0.45

#Standard Deviation of Y
Std_Y = sqrt(Var_Y)
#Answer: 0.6708204


#Marginal probabilities of X
X_marginal <- margin.table(matr, 2)
X_marginaltable <- matrix(c(0,1,2, X_marginal),3,2)
colnames(X_marginaltable) <- c('X', 'P(X)')
rownames(X_marginaltable) <- c('', '', '')
```

```
X_marginaltable

#Expected value of X, E(X):
Expected_X = X_marginaltable[1]*X_marginaltable[4] + X_marginaltable[2]*X_marginaltable[5] +
  X_marginaltable[3]*X_marginaltable[6]
Expected_X
#Answer: 0.7

#Variance(X):
Var_X = (X_marginaltable[1]-Expected_X)^2*X_marginaltable[4] +
  (X_marginaltable[2]-Expected_X)^2*X_marginaltable[5] +
  (X_marginaltable[3]-Expected_X)^2*X_marginaltable[6]
Var_X
#Answer: 0.41

#Standard Deviation of X
Std_X = sqrt(Var_X)
#Answer: 0.6403124

#Covariance(X,Y):
cov_x_y = (Y_marginaltable[1]-Expected_Y)*(X_marginaltable[1]-Expected_X)*0.12+(Y_marginaltable[1]-Expected_Y)*(X_marginaltable[2]-Expected_X)*0.42+(Y_marginaltable[1]-Expected_Y)*(X_marginaltable[3]-Expected_X)*0.06+(Y_marginaltable[2]-Expected_Y)*(X_marginaltable[1]-Expected_X)*0.21+(Y_marginaltable[2]-Expected_Y)*(X_marginaltable[2]-Expected_X)*0.06+(Y_marginaltable[2]-Expected_Y)*(X_marginaltable[3]-Expected_X)*0.03+(Y_marginaltable[3]-Expected_Y)*(X_marginaltable[1]-Expected_X)*0.07+(Y_marginaltable[3]-Expected_Y)*(X_marginaltable[2]-Expected_X)*0.02+(Y_marginaltable[3]-Expected_Y)*(X_marginaltable[3]-Expected_X)*0.01
```

```
cov_x_y
#Answer:  −0.15

###################################################
# Describing the Population of the Total Number of House Sales

# Laws of Expected Value: E(X + Y) = E(X) + E(Y)
# Laws of Variance: V(X + Y) = V(X) + V(Y) +2*Cov(X, Y)

#E(X+Y)
Exp_X_Y = Expected_X + Expected_Y
#Answer: 1.2

#Var(X+Y)
V_X_Y = Var_X + Var_Y + 2*cov_x_y
#Answer: 0.56


#End
```

R code Exa 7.8.a Describing the Population of the Returns on a Portfolio

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS
# Example 7.8a on Pg. 239
# Describing the Population of the Returns on a Portfolio

#Given w1, w2
w1 = .25
w2 = .75
```

```r

E_R1 = .08 #Expected value of McDonalds stock given
E_R2 = .15 #Expected value of Cisco stock
E_Rp = w1*E_R1 + w2*E_R2 #Expected return of the
    Portfolio
#Answer: 0.1325

#End
```

R code Exa 7.8b Describing the Population of the Returns on a Portfolio

```r
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE
    PROBABILITY DISTRIBUTIONS
# Example 7.8b on Pg. 239
# Describing the Population of the Returns on a
    Portfolio

#Given:

w1 = 0.25
w2 = 0.75
s1 = 0.12 #Standard Deviation of stock McD
s2 = 0.22 #Standard Deviation of stock Cisco

StandardDev <- function(Rho)
{
  return(sqrt(w1^2*s1^2 + w2^2*s2^2 + 2*w1*w2*Rho*s1
    *s2))
}

cat ("standard deviation of the returns on the
    portfolio, when the two stocks' returns are
    perfectly positively correlated, is:",
```

```
      StandardDev(1))
#Answer: 0.195

cat ("standard deviation of the returns on the portfolio, when the coefficient of correlation is 0.5, is:",
      StandardDev(0.5))
#Answer: 0.1819

cat ("standard deviation of the returns on the portfolio, when the two stocks' returns are uncorrelated, is:",
      StandardDev(0))
#Answer: 0.1677

#End
```

R code Exa 7.9.a Pat Statsdud and the Statistics Quiz

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS
# Example 7.9a on Pg. 246
# Pat Statsdud and the Statistics Quiz


# What is the probability that Pat gets no answers correct?
# n=10 iid trials. probability of each success is 1/5. Binomial distribution is apt.

#dbinom() function for Binomial
ans <- dbinom(0, 10, 0.2) #x=0, n=10, p=0.2

```

```
cat("P(Pat gets no answers correct) =", ans)

#Answer: 0.10737

#End
```

R code Exa 7.9.b Pat Statsdud and the Statistics Quiz

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS
# Example 7.9a on Pg. 246
# Pat Statsdud and the Statistics Quiz


# What is the probability that Pat gets two answers correct?
# n=10 iid trials. probability of each success is 1/5. Binomial distribution is apt.

#dbinom() function for Binomial
ans <- dbinom(2, 10, 0.2) #x=2, n=10, p=0.2

cat("P(Pat gets two answers correct) =", ans)

#Answer: 0.30199

#End
```

R code Exa 7.10 Will Pat Fail the Quiz

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS
# Example 7.10 on Pg. 247
# Will Pat Fail the Quiz?


# Find the probability that Pat fails the quiz. A mark is considered a failure if it is less than 50%
# n=10 iid trials. probability of each success is 1/5. Binomial distribution is apt.

#dbinom() function for Binomial
p0 <- dbinom(0, 10, 0.2) #x=0, n=10, p=0.2
p1 <- dbinom(1, 10, 0.2) #x=1, n=10, p=0.2
p2 <- dbinom(2, 10, 0.2) #x=2, n=10, p=0.2
p3 <- dbinom(3, 10, 0.2) #x=3, n=10, p=0.2
p4 <- dbinom(4, 10, 0.2) #x=4, n=10, p=0.2

cat("P(Pat fails the quiz) =", sum(p0,p1,p2,p3,p4))

#Answer: 0.96721

#End
```

R code Exa 7.11 Pat Statsdud Has Been Cloned

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS
# Example 7.11 on Pg. 249
# Pat Statsdud Has Been Cloned!
```

```
#mean n sd of a class with students like Pat?!

mean.function <- function(n,p)
{
  return(n*p)
}

sd.function <- function(n,p)
{
  return(sqrt(n*p*(1-p)))
}

#mean of binomial i.e., nxp
mean.function(10,0.2)

#variance of binomial i.e., nxpxq
sd.function(10,0.2)

#Answer: mean is 2
#        sd is 1.264911

#End
```

R code Exa 7.12 Probability of the Number of Typographical Errors in Textbooks

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE PROBABILITY DISTRIBUTIONS
# Example 7.12 on Pg. 252
# Probability of the Number of Typographical Errors in Textbooks

```

```

# Given the number of errors per 100 pages follows
    Poisson (1.5)

# P(there are no typographical errors in a sample of
     100 pages) is given as:

v <- dpois(0, 1.5)

cat("P(there are no typographical errors in a sample
     of 100 pages) =", v )

#Answer: 0.22313

#End
```

R code Exa 7.13.a Probability of the Number of Typographical Errors in 400 Pages

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE
    PROBABILITY DISTRIBUTIONS
# Example 7.13a on Pg. 253
# Probability of the Number of Typographical Errors
    in 400 Pages


# Given the number of errors per 100 pages follows
    Poisson (1.5).
# Probability of the Number of Typographical Errors
    in 400 Pages. Now, mean is 6 typos per 400 pages.

# P(there are no typographical errors in a sample of
     400 pages) is given as:
```

```
v <- dpois(0, 4*1.5)

#Answer: 0.0024788

cat("P(there are no typographical errors in a sample
    of 400 pages) =", v )

#End
```

R code Exa 7.13.b Probability of the Number of Typographical Errors in 400 Pages

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 7: RANDOM VARIABLES AND DISCRETE
    PROBABILITY DISTRIBUTIONS
# Example 7.13b on Pg. 253
# Probability of the Number of Typographical Errors
    in 400 Pages


# Given the number of errors per 100 pages follows
    Poisson (1.5).
# Probability of the Number of Typographical Errors
    in 400 Pages. Now, mean is 6 typos per 400 pages.

# P(there are five or fewer typos) is given as:

p0 <- dpois(0, 4*1.5)
p1 <- dpois(1, 4*1.5)
p2 <- dpois(2, 4*1.5)
p3 <- dpois(3, 4*1.5)
p4 <- dpois(4, 4*1.5)
p5 <- dpois(5, 4*1.5)
```

```
cat("P(X <= 5) = P(0) + P(1) + P(2) + P(3) + P(4) +
    P(5) =", sum(p0,p1,p2,p3,p4,p5))

#Answer: 0.44568

#End
```

Chapter 8

Continuous Probability Distributions

R code Exa 8.1.a Uniformly Distributed Gasoline Sales

```
# Statistics for Management and Economics by Gerald
    Keller
# CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.1a on Pg 267
# Uniformly Distributed Gasoline Sales

#Uniformly Distributed Gasoline Sales ~ U(2000,5000)

#U(2000,5000) graph
curve(dunif(x, min = 2000, max = 5000), from = 0, to
    = 6000, ylab = "f(x)", main = "Uniform Density f
    (x)")

#a. Find the probability that daily sales will fall
    between 2,500 and 3,000 gallons
#denote p1 = P(2500 <= X <= 3000) = P(X <= 3000) - P
    (X < 2500)
# punif() fives the probability of Uniform dist
    below a specified number
```

```

p1 <- punif(3000, min=2000, max=5000) - punif(2500, min=2000, max=5000)

#Answer: 0.16667

#End
```

R code Exa 8.1.b Uniformly Distributed Gasoline Sales

```
# Statistics for Management and Economics by Gerald Keller
# CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.1b on Pg 267
# Uniformly Distributed Gasoline Sales


#Uniformly Distributed Gasoline Sales ~ U(2000,5000)

# What is the probability that the service station will sell at least 4,000 gallons?
# denote p2 = P(X >= 4000) = 1 - P(X<4000)
# punif() fives the probability of Uniform dist below a specified number

p2 <- 1-punif(4000, min=2000, max=5000)

#Answer: 0.33333

#End
```

R code Exa 8.1.c Uniformly Distributed Gasoline Sales

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.1c on Pg 267
# Uniformly Distributed Gasoline Sales


#Uniformly Distributed Gasoline Sales ~ U(2000,5000)

#c. What is the probability that the station will sell exactly 2,500 gallons?
# punif() fives the probability of Uniform dist below a specified number

p3 <- punif(2500, min=2000, max=5000) - punif(2499.99999999999999999999, min=2000, max=5000)
#Answer: 0

#End
```

R code Exa 8.2 Normally Distributed Gasoline Sales

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.2 on Pg 272
# Normally Distributed Gasoline Sales


curve(dnorm(x,mean = 1000,sd=100), -1100, 2000)

#Given  daily demand for regular gasoline at another gas station ~ N(1000,100)
#Given mean=1000, sd=100

```

```
#Find P(X <= 1100). Let 'p' denote this required probability

p <- pnorm(1100, mean=1000, sd=100)
#Answer: 0.8413

#End
```

R code Exa 8.3.a Probability of a Negative Return on Investment

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.3a on Pg 277
# Probability of a Negative Return on Investment

#an ROI variable ~ N(10,5)

#Probability of losing money. Denote it by 'p'

p <- pnorm(0, mean=10, sd=5)
cat("The probability of losing money:", p)

#Answer: 0.02275

#End
```

R code Exa 8.3.b Probability of a Negative Return on Investment

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.3b on Pg 277
```

```
# Probability of a Negative Return on Investment


# Find the probability of losing money when the
    standard deviation is equal to 10%.


p <- pnorm(0, mean=10, sd=10)
cat("The probability of losing money when the
    standard deviation is equal to 10%:", p)

#Answer: 0.1586553

#End
```

R code Exa 8.4 Finding Z 05

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.4 on Pg 279
# Finding Z .05


# Find the value of a standard normal random
    variable such that the
# probability that the random variable is greater
    than it is 5%.

p <- qnorm(0.95)
cat("Z:", p)

#Answer: 1.644854

#End
```

R code Exa 8.5 Finding Z 05

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.5 on Pg 280
# Finding Z -.05


# Find the value of a standard normal random variable such that the
# probability that the random variable is less than it is 5%.

p <- qnorm(0.05)
cat("Z:", p)

#Answer: -1.644854

#End
```

R code Exa 8.6 Determining the Reorder Point

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.6 on Pg 283
# Determining the Reorder Point

mu = 200
sd = 50
```

```
Z_0.05 = qnorm(0.95)

reorderpoint = sd*Z_0.05 + mu
#Answer: 282.2427

#End
```

R code Exa 8.7.a Lifetimes of Alkaline Batteries

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.7a on Pg 288
# Lifetimes of Alkaline Batteries

#The lifetime of an alkaline battery is exp(0.05)
    distributed.
lambda = 0.05
#a.What is the mean and standard deviation of the
    battery's lifetime?

cat("Mean of battery's lifetime in hours:", 1/lambda
    )
cat("Standard Deviation of battery's lifetime in
    hours:", 1/lambda)

#Answer: 20 hours

#End
```

R code Exa 8.7.b Lifetimes of Alkaline Batteries

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.7b on Pg 288
# Lifetimes of Alkaline Batteries

#The lifetime of an alkaline battery is exp(0.05)
    distributed.
lambda = 0.05
#b. Find the probability that a battery will last
    between 10 and 15 hours.

p = pexp(15, rate=lambda) - pexp(10, rate=lambda)
cat("P(10 < battery lifetime < 15):",p)

#Answer: 0.1341641

#End
```

R code Exa 8.7.c Lifetimes of Alkaline Batteries

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.7c on Pg 288
# Lifetimes of Alkaline Batteries


#The lifetime of an alkaline battery is exp(0.05)
    distributed.
lambda = 0.05

#c. What is the probability that a battery will last
     for more than 20 hours?

```

```
p = 1- pexp(20, rate=lambda)
cat("P(battery lifetime > 20):",p)

#Answer: 0.3678794

#End
```

R code Exa 8.8.a Supermarket Checkout Counter

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.8a on Pg 290
# Supermarket Checkout Counter


#a.Find the probability of service is completed in
    fewer than 5 minutes
#the random variable, service process, X ~ exp(6/
    hour) i.e., X ~ exp(0.1/minute)
lambda = 0.1 #lambda = 0.1/minute

p = pexp(5, rate=lambda)

cat("P(X < 5):",p)

#Answer:0.3934693

#End
```

R code Exa 8.8.b Supermarket Checkout Counter

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.8b on Pg 290
# Supermarket Checkout Counter


#b.Find the probability of customer leaving checkout counter more than 10 minutes after arriving

#the random variable, service process, X ~ exp(6/hour) i.e., X ~ exp(0.1/minute)
lambda = 0.1 #lambda = 0.1/minute

p = 1 - pexp(10, rate=lambda)  #P(X > 10) = 1 - P(X < 10)

cat("P(X > 10):",p)

#Answer:0.367879

#End
```

R code Exa 8.8.c Supermarket Checkout Counter

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 8: CONTINUOUS PROBABILITY DISTRIBUTIONS
# Example 8.8b on Pg 290
# Supermarket Checkout Counter


#c.Find the probability of the service being completed in a time between 5 and 8 minutes

```

```
#the random variable, service process, X ~ exp(6/hour) i.e., X ~ exp(0.1/minute)
lambda = 0.1 #lambda = 0.1/minute

p = pexp(8, rate=lambda) - pexp(5, rate=lambda)  #P(5 < X < 8) = P(X < 8) - P(X < 5)

cat("P(5 < X < 8):",p)

#Answer: 0.1572017

#End
```

Chapter 9

Sampling Distributions

R code Exa 9.1.a Contents of a 32 Ounce Bottle

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 9: Sampling Distributions
# Example 9.1a on Pg 316
# Contents of a 32−Ounce Bottle

# random variable is amount of soda in each 32−ounce bottle denoted by X. X ~ N(32.2, 0.3)

#Given:
mu = 32.2
sd = 0.3

# probability that one bottle will contain more than 32 ounces. P(X > 32). Lets denote by 'p'
# pnorm() gives P(X < x) when X ~ Normal
p = 1- pnorm(32, mean=32.2, sd=0.3)
cat("P(X > 32):", p)

#Answer: 0.7475075

```

```
#Book's answer slightly different: 0.7486

#End
```

R code Exa 9.1.b Contents of a 32 Ounce Bottle

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 9: Sampling Distributions
# Example 9.1b on Pg 316
# Contents of a 32-Ounce Bottle


# random variable is amount of soda in each 32-ounce
    bottle denoted by X. X ~ N(32.2, 0.3)

#Given:
mu = 32.2
sd = 0.3

# what is the probability that the mean amount of
    the four bottles > 32 ounces.
# (X_bar > 32). Lets denote by 'p'
# pnorm() gives P(X < x) when X ~ Normal

p = 1 - pnorm(32, mean=32.2, sd=0.3/sqrt(4))
cat("P(X_bar > 32):", p)

#Answer: 0.9087888

#Book's answer slightly different: 0.9082

#End
```

R code Exa 9.2 Political Survey

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 9: Sampling Distributions
# Example 9.2 on Pg 326
# Political Survey


# Given number of respondents who would vote ~
    Binomial(300,0.52)
n = 300
p = 0.52

# what is the probability that the sample proportion
     is greater than 50% i.e., P(p^ > 0.5)
# We know that sample proportion ~ Normal(p, sd)
    where p = 0.52 and sd = sqrt(p*(1-p)/n)

sigma = sqrt(p*(1-p)/n)
#Answer: Sigma = 0.02884441

p1 = 1 - pnorm(0.5, mean=0.52, sd=sigma)
cat("P(p^ > 0.5):", p1)

#Answer:  0.755963

#Book's answer slightly different: 0.7549

#End
```

R code Exa 9.3 Starting Salaries of MBAs

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 9: Sampling Distributions
# Example 9.3 on Pg 328
# Starting Salaries of MBAs

# Given starting salaries of MBAs at WLU, X1 ~ Normal(62000,14500)
mu1 = 62000
sd1 = 14500
v1 = sd1^2
n1 = 50

# Given starting salaries of MBAs at UWO, X2 ~ Normal(60000,18300)
mu2 = 60000
sd2 = 18300
v2 = sd2^2
n2 = 60

# find probability that the sample mean starting salary of WLU graduates will exceed that of the UWO graduates
# i.e., find P (X1 - X2 > 0) denoted by 'p'
#we know X1-X2 ~ N(mu1-mu2, sqrt(v1/n1 + v2/n2))
p = 1 - pnorm(0, mean=mu1-mu2, sd=sqrt((v1/n1)+(v2/n2)))
cat("P(X1 - X2 > 0):", p)

#Answer:  0.7386917

#End
```

Chapter 10

Introduction to Estimation

R code Exa 10.1 Doll Computer Company

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 10: Introduction to Estimation
# Example 10.1 on Pg 342
# Doll Computer Company

data1 <- c(235, 374, 309, 499, 253, 421, 361, 514,
   462, 369, 394, 439,
           348, 344, 330, 261, 374, 302, 466, 535,
              386, 316, 296, 332, 334)
data1
mean1 <- mean(data1)
mean1
alpha = 0.05
library(stats)
std1 = 75
std2 <- sd(data1)
std2

ll <- mean1 - 1.96*75/(sqrt(25))
ul <- mean1 + 1.96*75/(sqrt(25))
```

```
cat("The 95% confidence interval is:","(",ll, ul,")"
   )

#End
```

Chapter 11

Introduction to Hypothesis Testing

Chapter 12

Inference About A Population

R code Exa 12.3 Consistency of a Container Filling Machine Part 1

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 12: INFERENCE ABOUT A POPULATION
# Example 12.3 on Pg 415
# Consistency of a Container-Filling Machine, Part 1

data1 <- c(999.6, 1000.7, 999.3, 1000.1, 999.5,
    1000.5, 999.7, 999.6, 999.1, 997.8,
           1001.3, 1000.7, 999.4, 1000.0, 998.3,
              999.5, 1000.1, 998.3, 999.2, 999.2,
           1000.4, 1000.1, 1000.1, 999.6, 999.9)
data1
mean1 <- mean(data1)
mean1
popmean = 1 #Null Hypothesis: H0: population mean =
    1 (sigma^2 =1)
n <- length(data1)
```

```r
n #sample size = 25
library(stats)
stdev1 <- sd(data1)
stdev1 #Answer: 0.7958
stdev1^2 #Answer: 0.6333

chistat <- (n-1)*stdev1^2/popmean
chistat #Answer: Chi-square test statistic = 15.20

#One-Sample Chi-Squared Test On Variance,  using varTest()
install.packages("EnvStats")
library(EnvStats)
result <- varTest(data1, alternative = "greater", conf.level = 0.95, sigma.squared = 1)

#Answer:

#Results of Hypothesis Test
#--------------------------

#Null Hypothesis:                 variance = 1
#Alternative Hypothesis:          True variance is greater than 1
#Test Name:                       Chi-Squared Test on Variance
#Estimated Parameter(s):          variance = 0.6333333
#Data:                            data1
#Test Statistic:                  Chi-Squared = 15.2
#Test Statistic Parameter:        df = 24
#P-value:                         0.9147699
#95% Confidence Interval:         LCL = 0.4174101
                                  UCL =       Inf

if(result$p.value > 0.05)
{
  print("there is NOT enough evidence to infer that
```

```
    the claim of sigmasquared = 1 is true.")
} else
{
  print("there is enough evidence to infer that the
    claim of sigmasquared = 1 is true.")
}

#End
```

R code Exa 12.4 Consistency of a Container Filling Machine Part 2

```
# Statistics for Management and Economics by Gerald
   Keller
# Chapter 12: INFERENCE ABOUT A POPULATION
# Example 12.4 on Pg. 418
# Consistency of a Container-Filling Machine, Part 2

data1 <- c(999.6, 1000.7, 999.3, 1000.1, 999.5,
   1000.5, 999.7, 999.6, 999.1, 997.8,
            1001.3, 1000.7, 999.4, 1000.0, 998.3,
               999.5, 1000.1, 998.3, 999.2, 999.2,
            1000.4, 1000.1, 1000.1, 999.6, 999.9)
data1
mean1 <- mean(data1)
mean1
popmean = 1 #Null Hypothesis: H0: population mean =
   1
n <- length(data1)
n
library(stats)
stdev1 <- sd(data1)
stdev1

chistat <- (n-1)*stdev1^2/popmean
chistat
```

```r

chisqalphaby2 <- qchisq(0.005, df=(n-1), lower.tail=FALSE)
chisq1minusalphaby2 <- qchisq(0.995, df=(n-1), lower.tail=FALSE)

lcl <- (n-1)*stdev1^2 / chisqalphaby2
lcl
ucl <- (n-1)*stdev1^2 / chisq1minusalphaby2
ucl

cat("The 99% confidence interval is:", "(", round(lcl,3), ",", round(ucl,3),")" )
#Answer: (0.333, 1.537)

#End
```

Chapter 13

Inference About Comparing Two Populations

Chapter 14

Analysis of Variance

Chapter 15

Chi Squared Tests

R code Exa 15.1 Testing Market Shares

```
# Statistics for Management and Economics by Gerald
    Keller
# Chapter 15: CHI-SQUARED TESTS
# Example 15.1 on Pg 598
# Testing Market Shares

#Null Hypothesis, Ho: p1 = .45, p2 = .40, p3 = .15
#Alternative Hypothesis, H1: At least one pi is not
    equal to its specified value

fabric <- c(102, 82, 16)
chi <- chisq.test(fabric, p = c(.45, .40, .15))
chi$statistic
chi$p.value #its less than 0.05 implying one can
    reject the Null hypothesis

tabchi <- qchisq(.95, df=2)

if(chi$statistic > tabchi)
{
  print("Advertising campaigns do have an effect.
```

```
      NUll Hypothesis is rejected.")
} else
{
  print("Advertising campaigns do NOT have an effect
      .")
}

#End
```

Chapter 16

Simple Linear Regression And Correlation

R code Exa 16.1 Annual Bonus and Years of Experience

```
# Statistics for Management and Economics by Gerald Keller
# Chapter 16: SIMPLE LINEAR REGRESSION AND CORRELATION
# Example 16.1 on Pg 638
# Annual Bonus and Years of Experience


years_of_exp <- c(1,2,3,4,5,6) #years of experience - Explanatory variable
annual_bonus <- c(6,1,9,5,17,12) #annual bonus in 1000s - Respone variable

#determine the straight line relationship between years of experience and annual bonus using least squares

regression_line <- lm(annual_bonus ~ years_of_exp) # gives regression line
```

```
summary(regression_line) #gives the Residuals , Std Error etc

plot(years_of_exp, annual_bonus) #scatter plot
abline(lm(annual_bonus ~ years_of_exp))

cat("The least squares or regression line is Y =",
    regression_line$coefficients[1], "+", regression_line$coefficients[2], "X",
    "where Y is Annual Bonus and X is years of job experience")

# The least squares line is Y = 0.934 + 2.114X

#End
```

Chapter 17

Multiple Regression