

Contents

- 1 Introduction: Biostatistics and R** 1
 - 1.1 Purpose of This Text..... 1
 - 1.2 Development of Biostatistics..... 2
 - 1.3 Development of R 3
 - 1.4 How R is Used in This Text 4
- 2 Data Exploration, Descriptive Statistics, and Measures of Central Tendency** 5
 - 2.1 Background on This Lesson..... 5
 - 2.1.1 Description of the Data..... 5
 - 2.1.2 Null Hypothesis (Ho) 7
 - 2.2 Data Import of a .csv Spreadsheet-Type Data File into R 7
 - 2.3 Organize the Data and Display the Code Book 9
 - 2.4 Conduct a Visual Data Check 9
 - 2.5 Descriptive Analysis of the Data..... 10
 - 2.6 Summary 13
 - 2.7 Addendum: Specialized External Packages and Functions 13
 - 2.8 Prepare to Exit, Save, and Later Retrieve This R Session 15
- 3 Student’s t-Test for Independent Samples**..... 17
 - 3.1 Background on This Lesson..... 17
 - 3.1.1 Description of the Data..... 17
 - 3.1.2 Null Hypothesis (Ho) 18
 - 3.2 Data Import of a .csv Spreadsheet-Type Data File into R 19
 - 3.3 Organize the Data and Display the Code Book 20
 - 3.4 Conduct a Visual Data Check 23
 - 3.5 Descriptive Analysis of the Data..... 34
 - 3.6 Conduct the Statistical Analysis 40
 - 3.7 Summary 42
 - 3.8 Addendum: t-Statistic v z-Statistic 43
 - 3.8.1 Create the Enumerated Dataset 44

| | | |
|----------|--|-----------|
| 3.8.2 | Calculate the t-Statistic | 44 |
| 3.8.3 | Calculate the z-Statistic | 45 |
| 3.9 | Prepare to Exit, Save, and Later Retrieve This R Session | 45 |
| 4 | Student's t-Test for Matched Pairs | 47 |
| 4.1 | Background on This Lesson | 47 |
| 4.1.1 | Description of the Data | 47 |
| 4.1.2 | Null Hypothesis (Ho) | 49 |
| 4.1.3 | Unstacked Data and Stacked Data | 49 |
| 4.2 | Data Import of a .csv Spreadsheet-Type Data File into R | 51 |
| 4.3 | Organize the Data and Display the Code Book | 52 |
| 4.4 | Conduct a Visual Data Check | 54 |
| 4.5 | Descriptive Analysis of the Data | 60 |
| 4.6 | Conduct the Statistical Analysis | 63 |
| 4.7 | Summary | 65 |
| 4.8 | Addendum 1: Stacked Data and Student's t-Test for Matched Pairs | 66 |
| 4.9 | Addendum 2: The Impact of N on Student's t-Test | 70 |
| 4.10 | Prepare to Exit, Save, and Later Retrieve This R Session | 72 |
| 5 | Oneway Analysis of Variance (ANOVA) | 73 |
| 5.1 | Background on This Lesson | 73 |
| 5.1.1 | Description of the Data | 73 |
| 5.1.2 | Null Hypothesis (Ho) | 75 |
| 5.2 | Data Import of a .csv Spreadsheet-Type Data File into R | 75 |
| 5.3 | Organize the Data and Display the Code Book | 77 |
| 5.4 | Conduct a Visual Data Check | 82 |
| 5.5 | Descriptive Analysis of the Data | 87 |
| 5.6 | Conduct the Statistical Analysis | 89 |
| 5.6.1 | Exploratory Oneway ANOVA | 90 |
| 5.6.2 | Oneway ANOVA Method 1: lm() and anova() Functions ... | 91 |
| 5.6.3 | Oneway ANOVA Method 2: aov() and TukeyHSD() Functions | 92 |
| 5.7 | Summary | 93 |
| 5.8 | Addendum: Other Packages for Display of Oneway ANOVA | 96 |
| 5.9 | Prepare to Exit, Save, and Later Retrieve This R Session | 97 |
| 6 | Twoway Analysis of Variance (ANOVA) | 99 |
| 6.1 | Background on This Lesson | 99 |
| 6.1.1 | Description of the Data | 99 |
| 6.1.2 | Null Hypothesis (Ho) | 100 |
| 6.2 | Data Import of a .csv Spreadsheet-Type Data File into R | 100 |
| 6.3 | Organize the Data and Display the Code Book | 101 |
| 6.4 | Conduct a Visual Data Check | 104 |
| 6.5 | Descriptive Analysis of the Data | 111 |
| 6.6 | Conduct the Statistical Analysis | 117 |

| | | |
|----------|--|------------|
| 6.7 | Summary | 122 |
| 6.8 | Addendum: Other Packages for Display of Twoway ANOVA | 124 |
| 6.9 | Prepare to Exit, Save, and Later Retrieve This R Session | 126 |
| 7 | Correlation and Linear Regression | 129 |
| 7.1 | Background on This Lesson..... | 129 |
| 7.1.1 | Description of the Data..... | 129 |
| 7.1.2 | Null Hypothesis (Ho) | 130 |
| 7.2 | Data Import of a .csv Spreadsheet-Type Data File into R | 131 |
| 7.3 | Organize the Data and Display the Code Book | 132 |
| 7.4 | Conduct a Visual Data Check | 135 |
| 7.5 | Descriptive Analysis of the Data..... | 140 |
| 7.6 | Conduct the Statistical Analysis | 142 |
| 7.6.1 | Correlation Using Pearson's r | 142 |
| 7.6.2 | Linear Regression | 150 |
| 7.7 | Summary | 154 |
| 7.8 | Addendum: Multiple Regression | 155 |
| 7.8.1 | Hand-Calculate Multiple Regression | 156 |
| 7.8.2 | Minimal Adequate Model (MAM) for Regression | 158 |
| 7.8.3 | Stepwise Regression | 160 |
| 7.9 | Prepare to Exit, Save, and Later Retrieve This R Session | 163 |
| 8 | Future Actions and Next Steps | 165 |
| 8.1 | Use of This Text | 165 |
| 8.2 | Future Use of R for Biostatistics..... | 166 |
| 8.3 | External Resources | 167 |
| 8.4 | Contact the Author..... | 167 |

Introduction to Data Analysis and Graphical
Presentation in Biostatistics with R
Statistics in the Large

MacFarland, Th.W.

2014, VII, 167 p. 16 illus., 14 illus. in color., Softcover

ISBN: 978-3-319-02531-5