

# Statistics Using Technology

Second Edition

By Kathryn Kozak



Photo taken by Richard Kozak at Parkes Observatory in Parkes, NSW, Australia



**Creative Commons Attribution Sharealike.** This license is considered to be some to be the most open license. It allows reuse, remixing, and distribution (including commercial), but requires any remixes use the same license as the original. This limits where the content can be remixed into, but on the other hand ensures that no-one can remix the content then put the remix under a more restrictive license.

2015 Kathryn Kozak

**ISBN:** 978-1-329-75725-7

## Table of Content:

<b>Preface</b>	<b>iii</b>
<b>Chapter 1: Statistical Basics</b>	<b>1</b>
Section 1.1: What is Statistics?	1
Section 1.2: Sampling Methods	8
Section 1.3: Experimental Design	14
Section 1.4: How Not to Do Statistics	19
<b>Chapter 2: Graphical Descriptions of Data</b>	<b>25</b>
Section 2.1: Qualitative Data	25
Section 2.2: Quantitative Data	37
Section 2.3: Other Graphical Representations of Data	59
<b>Chapter 3: Numerical Descriptions of Data</b>	<b>75</b>
Section 3.1: Measures of Center	75
Section 3.2: Measures of Spread	90
Section 3.3: Ranking	107
<b>Chapter 4: Probability</b>	<b>121</b>
Section 4.1: Empirical Probability	121
Section 4.2: Theoretical Probability	124
Section 4.3: Conditional Probability	140
Section 4.4: Counting Techniques	152
<b>Chapter 5: Discrete Probability Distributions</b>	<b>157</b>
Section 5.1: Basics of Probability Distributions	157
Section 5.2: Binomial Probability Distribution	167
Section 5.3: Mean and Standard Deviation of Binomial Distribution	181
<b>Chapter 6: Continuous Probability Distributions</b>	<b>187</b>
Section 6.1: Uniform Distribution	187
Section 6.2: Graphs of the Normal Distribution	190
Section 6.3: Finding Probabilities for the Normal Distribution	193
Section 6.4: Assessing Normality	203
Section 6.5: Sampling Distribution and the Central Limit Theorem	216

<b>Chapter 7: One-Sample Inference</b>	<b>229</b>
Section 7.1: Basics of Hypothesis Testing	229
Section 7.2: One-Sample Proportion Test	242
Section 7.3: One-Sample Test for the Mean	249
<b>Chapter 8: Estimation</b>	<b>263</b>
Section 8.1: Basics of Confidence Intervals	263
Section 8.2: One-Sample Interval for the Proportion	267
Section 8.3: One-Sample Interval for the Mean	272
<b>Chapter 9: Two-Sample Inference</b>	<b>283</b>
Section 9.1: Two Proportions	283
Section 9.2: Paired Samples for Two Means	293
Section 9.3: Independent Samples for Two Means	313
Section 9.4: Which Analysis Should You Conduct?	339
<b>Chapter 10: Regression and Correlation</b>	<b>343</b>
Section 10.1: Regression	343
Section 10.2: Correlation	363
Section 10.3: Inference for Regression and Correlation	371
<b>Chapter 11: Chi-Square and ANOVA Tests</b>	<b>393</b>
Section 11.1: Chi-Square Test for Independence	393
Section 11.2: Chi-Square Goodness of Fit	411
Section 11.3: Analysis of Variance (ANOVA)	419
<b>Appendix: Critical Value Tables</b>	<b>433</b>
Table A.1: Normal Critical Values for Confidence Levels	434
Table A.2: Critical Values for t-Interval	435
<b>Answers to Odd Questions</b>	<b>439</b>
<b>Index</b>	<b>447</b>

## Preface:

I hope you find this book useful in teaching statistics. When writing this book, I tried to follow the GAISE Standards (*GAISE recommendations*. (2014, January 05). Retrieved from [http://www.amstat.org/education/gaise/GAISECollege\\_Recommendations.pdf](http://www.amstat.org/education/gaise/GAISECollege_Recommendations.pdf)), which are

- 1.) Emphasis statistical literacy and develop statistical understanding.
- 2.) Use real data.
- 3.) Stress conceptual understanding, rather than mere knowledge of procedure.
- 4.) Foster active learning in the classroom.
- 5.) Use technology for developing concepts and analyzing data.

To this end, I ask students to interpret the results of their calculations. I incorporated the use of technology for most calculations. Because of that you will not find me using any of the computational formulas for standard deviations or correlation and regression since I prefer students understand the concept of these quantities. Also, because I utilize technology you will not find the standard normal table, Student's t-table, binomial table, chi-square distribution table, and F-distribution table in the book. The only tables I provided were for critical values for confidence intervals since they are more difficult to find using technology. Another difference between this book and other statistics books is the order of hypothesis testing and confidence intervals. Most books present confidence intervals first and then hypothesis tests. I find that presenting hypothesis testing first and then confidence intervals is more understandable for students. Lastly, I have de-emphasized the use of the z-test. In fact, I only use it to introduce hypothesis testing, and never utilize it again. You may also notice that when I introduced hypothesis testing and confidence intervals, proportions were introduced before means. However, when two sample tests and confidence intervals are introduced I switched this order. This is because usually many instructors do not discuss the proportions for two samples. However, you might try assigning problems for proportions without discussing it in class. After doing two samples for means, the proportions are similar. Lastly, to aid student understanding and interest, most of the homework and examples utilize real data. Again, I hope you find this book useful for your introductory statistics class.

I want to make a comment about the mathematical knowledge that I assumed the students possess. The course for which I wrote this book has a higher prerequisite than most introductory statistics books. However, I do feel that students can read and understand this book as long as they have had basic algebra and can substitute numbers into formulas. I do not show how to create most of the graphs, but most students should have been exposed to them in high school. So I hope the mathematical level is appropriate for your course.

The technology that I utilized for creating the graphs was Microsoft Excel, and I utilized the TI-83/84 graphing calculator for most calculations, including hypothesis testing, confidence intervals, and probability distributions. This is because these tools are readily available to my students. Please feel free to use any other technology that is more appropriate for your students. Do make sure that you use some technology.

**Acknowledgments:**

I would like to thank the following people for taking their valuable time to review the book. Their comments and insights improved this book immensely.

Jane Tanner, Onondaga Community College  
Rob Farinelli, College of Southern Maryland  
Carrie Kinnison, retired engineer  
Sean Simpson, Westchester Community College  
Kim Sonier, Coconino Community College  
Jim Ham, Delta College  
David Straayer, Tacoma Community College  
Kendra Feinstein, Tacoma Community College  
Students of Coconino Community College  
Students Tacoma Community College

I also want to thank Coconino Community College for granting me a sabbatical so that I would have the time to write the book. Lastly, I want to thank my husband Rich and my son Dylan for supporting me in this project. Without their love and support, I would not have been able to complete the book.

**New to the Second Edition:**

The additions to this edition mostly involve adding the commands to create graphs, compute descriptive statistics, finding probabilities, and computing inferential analysis using the open source software R. Another change involve adding an example at the end of chapter 3 that shows analyzing a data set using graphical and numerical descriptions. Another major change was adding a section 9.4 that gives some insight into which inferential analysis should be completed based on a series of questions that should be asked. Lastly, minor explanations were made and corrections were made where necessary.

On a personal note, I wanted to thank my brother, John Matic, his wife Jenelle, and their children Hannah and Eli for their hospitality when writing the first edition. In addition to allowing my family access to their home, John provided numerous examples and data sets for business applications in this book. I inadvertently left this thank you out of the first edition of the book, and for that I apologize. His help and his family's hospitality were invaluable to me.