# Table of Contents

1 Sampling and Data
   1.1 Sampling and Data .......................................................... 13
   1.2 Statistics ................................................................. 13
   1.3 Probability ............................................................... 15
   1.4 Key Terms ................................................................. 15
   1.5 Data ........................................................------------ 17
   1.6 Sampling ................................................................. 18
   1.7 Variation ................................................................. 24
   1.8 Answers and Rounding Off ............................................ 26
   1.9 Frequency ............................................................... 26
   1.10 Summary ................................................................. 31
   1.11 Practice: Sampling and Data ........................................ 32
   1.12 Homework ............................................................. 35
   1.13 Lab 1: Data Collection ................................................ 43
   1.14 Lab 2: Sampling Experiment ......................................... 45
   Solutions ................................................................. 48

2 Descriptive Statistics
   2.1 Descriptive Statistics .................................................. 51
   2.2 Displaying Data .......................................................... 51
   2.3 Stem and Leaf Graphs (Stemplots), Line Graphs and Bar Graphs . 52
   2.4 Histograms .............................................................. 55
   2.5 Box Plots ................................................................. 59
   2.6 Measures of the Location of the Data ............................... 62
   2.7 Measures of the Center of the Data ................................. 67
   2.8 Skewness and the Mean, Median, and Mode ........................ 70
   2.9 Measures of the Spread of the Data ................................. 72
   2.10 Summary of Formulas .................................................. 80
   2.11 Practice 1: Center of the Data ..................................... 81
   2.12 Practice 2: Spread of the Data ..................................... 84
   2.13 Homework ............................................................. 86
   2.14 Lab: Descriptive Statistics ........................................... 103
   Solutions ................................................................. 105

3 Probability Topics
   3.1 Probability Topics ..................................................... 113
   3.2 Terminology .............................................................. 114
   3.3 Independent and Mutually Exclusive Events ....................... 116
   3.4 Two Basic Rules of Probability ..................................... 119
   3.5 Contingency Tables .................................................... 123
   3.6 Venn Diagrams (optional) ............................................. 126
   3.7 Tree Diagrams (optional) .............................................. 127
   3.8 Summary of Formulas .................................................. 132
   3.9 Practice 1: Contingency Tables ..................................... 133

Additional Resources .................................................................. 5
Author Acknowledgements .......................................................... 9
Student Welcome Letter ............................................................. 11

Preface ......................................................................................... 1
3.10 Practice 2: Calculating Probabilities .............................................................. 135
3.11 Homework .................................................................................................... 136
3.12 Review .......................................................................................................... 147
3.13 Lab: Probability Topics ................................................................................. 149
   Solutions ........................................................................................................... 152
4 Discrete Random Variables
   4.1 Discrete Random Variables ....................................................................... 159
   4.2 Probability Distribution Function (PDF) for a Discrete Random Variable ...... 160
   4.3 Mean or Expected Value and Standard Deviation ...................................... 161
   4.4 Common Discrete Probability Distribution Functions ............................. 164
   4.5 Binomial ....................................................................................................... 164
   4.6 Geometric (optional) .................................................................................. 167
   4.7 Hypergeometric (optional) .......................................................................... 170
   4.8 Poisson ......................................................................................................... 172
   4.9 Summary of Functions .............................................................................. 175
   4.10 Practice 1: Discrete Distribution ............................................................... 177
   4.11 Practice 2: Binomial Distribution ............................................................... 178
   4.12 Practice 3: Poisson Distribution ............................................................... 180
   4.13 Practice 4: Geometric Distribution ........................................................... 181
   4.14 Practice 5: Hypergeometric Distribution .................................................. 183
   4.15 Homework .................................................................................................. 184
   4.16 Review ........................................................................................................ 194
   4.17 Lab 1: Discrete Distribution (Playing Card Experiment) ......................... 197
   4.18 Lab 2: Discrete Distribution (Lucky Dice Experiment) ............................. 201
      Solutions ....................................................................................................... 205
5 Continuous Random Variables
   5.1 Continuous Random Variables ................................................................ 213
   5.2 Continuous Probability Functions ............................................................. 215
   5.3 The Uniform Distribution ......................................................................... 218
   5.4 The Exponential Distribution .................................................................... 225
   5.5 Summary of the Uniform and Exponential Probability Distributions ....... 231
   5.6 Practice 1: Uniform Distribution ............................................................... 232
   5.7 Practice 2: Exponential Distribution ......................................................... 235
   5.8 Homework .................................................................................................. 237
   5.9 Review ........................................................................................................ 243
   5.10 Lab: Continuous Distribution ................................................................... 246
      Solutions ....................................................................................................... 249
6 The Normal Distribution
   6.1 The Normal Distribution ........................................................................... 253
   6.2 The Standard Normal Distribution ............................................................ 254
   6.3 Z-scores ........................................................................................................ 255
   6.4 Areas to the Left and Right of x .................................................................. 257
   6.5 Calculations of Probabilities ...................................................................... 257
   6.6 Summary of Formulas ............................................................................... 261
   6.7 Practice: The Normal Distribution ............................................................. 262
   6.8 Homework ................................................................................................... 264
   6.9 Review ........................................................................................................ 270
   6.10 Lab 1: Normal Distribution (Lap Times) .................................................. 272
   6.11 Lab 2: Normal Distribution (Pinkie Length) .............................................. 275
      Solutions ....................................................................................................... 277
7 The Central Limit Theorem
- 7.1 The Central Limit Theorem .................................................. 281
- 7.2 The Central Limit Theorem for Sample Means (Averages) ............... 282
- 7.3 The Central Limit Theorem for Sums ..................................... 285
- 7.4 Using the Central Limit Theorem ............................................. 286
- 7.5 Summary of Formulas ................................................................ 293
- 7.6 Practice: The Central Limit Theorem ....................................... 294
- 7.7 Homework ............................................................................. 297
- 7.8 Review ................................................................................... 304
- 7.9 Lab 1: Central Limit Theorem (Pocket Change) ............................. 306
- 7.10 Lab 2: Central Limit Theorem (Cookie Recipes) ......................... 310
- Solutions .................................................................................. 315

8 Confidence Intervals
- 8.1 Confidence Intervals ............................................................... 319
- 8.2 Confidence Interval, Single Population Mean, Population Standard Deviation Known, Normal .............................................................. 321
- 8.3 Confidence Interval, Single Population Mean, Standard Deviation Unknown, Student’s-t ..................................................................... 328
- 8.4 Confidence Interval for a Population Proportion ............................. 331
- 8.5 Summary of Formulas ................................................................ 336
- 8.6 Practice 1: Confidence Intervals for Means, Known Population Standard Deviation ................................................................. 337
- 8.7 Practice 2: Confidence Intervals for Means, Unknown Population Standard Deviation ................................................................. 339
- 8.8 Practice 3: Confidence Intervals for Proportions ............................ 341
- 8.9 Homework ............................................................................. 343
- 8.10 Review ................................................................................... 353
- 8.11 Lab 1: Confidence Interval (Home Costs) ................................... 356
- 8.12 Lab 2: Confidence Interval (Place of Birth) ................................ 359
- 8.13 Lab 3: Confidence Interval (Womens’ Heights) ............................. 361
- Solutions .................................................................................. 363

9 Hypothesis Testing: Single Mean and Single Proportion
- 9.1 Hypothesis Testing: Single Mean and Single Proportion .................. 369
- 9.2 Null and Alternate Hypotheses .................................................. 370
- 9.3 Outcomes and the Type I and Type II Errors ................................. 371
- 9.4 Distribution Needed for Hypothesis Testing .................................. 372
- 9.5 Assumption ............................................................................ 373
- 9.6 Rare Events ............................................................................ 373
- 9.7 Using the Sample to Test the Null Hypothesis ............................... 374
- 9.8 Decision and Conclusion ......................................................... 375
- 9.9 Additional Information ............................................................. 375
- 9.10 Summary of the Hypothesis Test .............................................. 377
- 9.11 Examples ............................................................................... 377
- 9.12 Summary of Formulas .............................................................. 388
- 9.14 Practice 2: Single Mean, Unknown Population Standard Deviation .. 391
- 9.15 Practice 3: Single Proportion .................................................... 393
- 9.16 Homework ............................................................................. 395
- 9.17 Review ................................................................................... 408
- 9.18 Lab: Hypothesis Testing of a Single Mean and Single Proportion .... 411
- Solutions .................................................................................. 415
10 Hypothesis Testing: Two Means, Paired Data, Two Proportions
  10.1 Hypothesis Testing: Two Population Means and Two Population Proportions
  10.2 Comparing Two Independent Population Means with Unknown Population Standard Deviations
  10.3 Comparing Two Independent Population Means with Known Population Standard Deviations
  10.4 Comparing Two Independent Population Proportions
  10.5 Matched or Paired Samples
  10.6 Summary of Types of Hypothesis Tests
  10.7 Practice 1: Hypothesis Testing for Two Proportions
  10.8 Practice 2: Hypothesis Testing for Two Averages
  10.9 Homework
  10.10 Review
  10.11 Lab: Hypothesis Testing for Two Means and Two Proportions
Solutions
11 The Chi-Square Distribution
  11.1 The Chi-Square Distribution
  11.2 Notation
  11.3 Facts About the Chi-Square Distribution
  11.4 Goodness-of-Fit Test
  11.5 Test of Independence
  11.6 The Chi-Square Distribution: Test for Homogeneity
  11.7 The Chi-Square Distribution: Comparison Summary of the Chi-Square Tests
    Goodness-of-Fit, Independence and Homogeneity
  11.8 Test of a Single Variance
  11.9 Summary of Formulas
  11.10 Practice 1: Goodness-of-Fit Test
  11.11 Practice 2: Contingency Tables
  11.12 Practice 3: Test of a Single Variance
  11.13 Homework
  11.14 Review
  11.15 Lab 1: Chi-Square Goodness-of-Fit
  11.16 Lab 2: Chi-Square Test for Independence
Solutions
12 Linear Regression and Correlation
  12.1 Linear Regression and Correlation
  12.2 Linear Equations
  12.3 Slope and Y-Intercept of a Linear Equation
  12.4 Scatter Plots
  12.5 The Regression Equation
  12.6 Correlation Coefficient and Coefficient of Determination
  12.7 Testing the Significance of the Correlation Coefficient
  12.8 Prediction
  12.9 Outliers
  12.10 95% Critical Values of the Sample Correlation Coefficient Table
  12.11 Summary
  12.12 Practice: Linear Regression
  12.13 Homework
  12.14 Lab 1: Regression (Distance from School)
  12.15 Lab 2: Regression (Textbook Cost)
Preface

Welcome to Collaborative Statistics, presented by Connexions. The initial section below introduces you to Connexions. If you are familiar with Connexions, please skip to About "Collaborative Statistics." (Section: About Connexions)

About Connexions

Connexions Modular Content
Connexions (cnx.org) is an online, open access educational resource dedicated to providing high quality learning materials free online, free in printable PDF format, and at low cost in bound volumes through print-on-demand publishing. The Collaborative Statistics textbook is one of many collections available to Connexions users. Each collection is composed of a number of re-usable learning modules written in the Connexions XML markup language. Each module may also be re-used (or ‘re-purposed’) as part of other collections and may be used outside of Connexions. Including Collaborative Statistics, Connexions currently offers over 6500 modules and more than 350 collections.

The modules of Collaborative Statistics are derived from the original paper version of the textbook under the same title, Collaborative Statistics. Each module represents a self-contained concept from the original work. Together, the modules comprise the original textbook.

Re-use and Customization
The Creative Commons (CC) Attribution license applies to all Connexions modules. Under this license, any module in Connexions may be used or modified for any purpose as long as proper attribution to the original author(s) is maintained. Connexions’ authoring tools make re-use (or re-purposing) easy. Therefore, instructors anywhere are permitted to create customized versions of the Collaborative Statistics textbook by editing modules, deleting unneeded modules, and adding their own supplementary modules. Connexions’ authoring tools keep track of these changes and maintain the CC license’s required attribution to the original authors. This process creates a new collection that can be viewed online, downloaded as a single PDF file, or ordered in any quantity by instructors and students as a low-cost printed textbook. To start building custom collections, please visit the help page, “Create a Collection with Existing Modules”. For a guide to authoring modules, please look at the help page, “Create a Module in Minutes”.

Read the book online, print the PDF, or buy a copy of the book.
To browse the Collaborative Statistics textbook online, visit the collection home page at cnx.org/content/col10522/latest. You will then have three options.

---

1This content is available online at <http://cnx.org/content/m16026/1.16/>.  
2http://cnx.org/  
3http://creativecommons.org/licenses/by/2.0/  
4http://cnx.org/help/CreateCollection  
5http://cnx.org/help/ModuleInMinutes  
6Collaborative Statistics <http://cnx.org/content/col10522/latest/>
1. You may obtain a PDF of the entire textbook to print or view offline by clicking on the “Download PDF” link in the “Content Actions” box.
2. You may order a bound copy of the collection by clicking on the “Order Printed Copy” button.
3. You may view the collection modules online by clicking on the “Start” link, which takes you to the first module in the collection. You can then navigate through the subsequent modules by using their “Next” and “Previous” links to move forward and backward in the collection. You can jump to any module in the collection by clicking on that module’s title in the “Collection Contents” box on the left side of the window. If these contents are hidden, make them visible by clicking on “[show table of contents]”.

Accessibility and Section 508 Compliance

- For information on general Connexions accessibility features, please visit http://cnx.org/content/m17212/latest/.
- For information on accessibility features specific to the Collaborative Statistics textbook, please visit http://cnx.org/content/m17211/latest/.

Version Change History and Errata

- For a list of modifications, updates, and corrections, please visit http://cnx.org/content/m17360/latest/.

Adoption and Usage

- The Collaborative Statistics collection has been adopted and customized by a number of professors and educators for use in their classes. For a list of known versions and adopters, please visit http://cnx.org/content/m18261/latest/.

About “Collaborative Statistics”

Collaborative Statistics was written by Barbara Illowsky and Susan Dean, faculty members at De Anza College in Cupertino, California. The textbook was developed over several years and has been used in regular and honors-level classroom settings and in distance learning classes. Courses using this textbook have been articulated by the University of California for transfer of credit. The textbook contains full materials for course offerings, including expository text, examples, labs, homework, and projects. A Teacher’s Guide is currently available in print form and on the Connexions site at http://cnx.org/content/col10547/latest/, and supplemental course materials including additional problem sets and video lectures are available at http://cnx.org/content/col10586/latest/. The on-line text for each of these collections collections will meet the Section 508 standards for accessibility.

An on-line course based on the textbook was also developed by Illowsky and Dean. It has won an award as the best on-line California community college course. The on-line course will be available at a later date as a collection in Connexions, and each lesson in the on-line course will be linked to the on-line textbook chapter. The on-line course will include, in addition to expository text and examples, videos of course lectures in captioned and non-captioned format.

The original preface to the book as written by professors Illowsky and Dean, now follows:

---

7 “Accessibility Features of Connexions” <http://cnx.org/content/m17212/latest/>
8 “Collaborative Statistics: Accessibility” <http://cnx.org/content/m17211/latest/>
9 “Collaborative Statistics: Change History” <http://cnx.org/content/m17360/latest/>
10 “Collaborative Statistics: Adoption and Usage” <http://cnx.org/content/m18261/latest/>
11 Collaborative Statistics Teacher’s Guide <http://cnx.org/content/col10547/latest/>
12 Collaborative Statistics: Supplemental Course Materials <http://cnx.org/content/col10586/latest/>
This book is intended for introductory statistics courses being taken by students at two- and four-year colleges who are majoring in fields other than math or engineering. Intermediate algebra is the only prerequisite. The book focuses on applications of statistical knowledge rather than the theory behind it. The text is named *Collaborative Statistics* because students learn best by **doing**. In fact, they learn best by working in small groups. The old saying “two heads are better than one” truly applies here.

**Our emphasis in this text is on four main concepts:**

- thinking statistically
- incorporating technology
- working collaboratively
- writing thoughtfully

These concepts are integral to our course. Students learn the best by actively participating, not by just watching and listening. Teaching should be highly interactive. Students need to be thoroughly engaged in the learning process in order to make sense of statistical concepts. *Collaborative Statistics* provides techniques for students to write across the curriculum, to collaborate with their peers, to think statistically, and to incorporate technology.

This book takes students step by step. The text is interactive. Therefore, students can immediately apply what they read. Once students have completed the process of problem solving, they can tackle interesting and challenging problems relevant to today’s world. The problems require the students to apply their newly found skills. In addition, technology (TI-83 graphing calculators are highlighted) is incorporated throughout the text and the problems, as well as in the special group activities and projects. The book also contains labs that use real data and practices that lead students step by step through the problem solving process.

At De Anza, along with hundreds of other colleges across the country, the college audience involves a large number of ESL students as well as students from many disciplines. The ESL students, as well as the non-ESL students, have been especially appreciative of this text. They find it extremely readable and understandable. *Collaborative Statistics* has been used in classes that range from 20 to 120 students, and in regular, honor, and distance learning classes.

Susan Dean

Barbara Illowsky
Additional Resources

Additional Resources Currently Available

- Glossary (Glossary, p. 5)
- View or Download This Textbook Online (View or Download This Textbook Online, p. 5)
- Supplemental Materials (Supplemental Materials, p. 6)
- Video Lectures (Video Lectures, p. 6)
- Version History (Version History, p. 6)
- Textbook Adoption and Usage (Textbook Adoption and Usage, p. 6)
- Additional Technologies and Notes (Additional Technologies, p. 7)
- Accessibility and Section 508 Compliance (Accessibility and Section 508 Compliance, p. 7)

The following section describes some additional resources for learners and educators. These modules and collections are all available on the Connexions website (http://cnx.org/) and can be viewed online, downloaded, printed, or ordered as appropriate.

Glossary
This module contains the entire glossary for the Collaborative Statistics textbook collection (col10522) since its initial release on 15 July 2008. The glossary is located at http://cnx.org/content/m16129/latest/.

Below are links to additional resources:
Link to the Statistics Glossary by Dr. Philip Stark, UC Berkeley
http://statistics.berkeley.edu/~stark/SticiGui/Text/gloss.htm

Link to Wikipedia
http://www.wikipedia.org/ (Search on “Glossary of probability and statistics.”)

View or Download This Textbook Online
The complete contents of this book are available at no cost on the Connexions website at http://cnx.org/content/col10522/latest/. Anybody can view this content free of charge either as an online e-book or a downloadable PDF file. A low-cost printed version of this textbook is also available here.

---

13This content is available online at <http://cnx.org/content/m18746/1.7/>.
14http://cnx.org/
15“Collaborative Statistics: Glossary” <http://cnx.org/content/m16129/latest/>
16http://statistics.berkeley.edu/~stark/SticiGui/Text/gloss.htm
17http://www.wikipedia.org/
18Collaborative Statistics <http://cnx.org/content/col10522/latest/>
19http://my.qoop.com/store/706493342106149/7781159220340
Collaborative Statistics Teacher’s Guide

A complementary Teacher’s Guide for Collaborative statistics is available through Connexions at http://cnx.org/content/col10547/latest/20. The Teacher’s Guide includes suggestions for presenting concepts found throughout the book as well as recommended homework assignments. A low-cost printed version of this textbook is also available here21.

Supplemental Materials

This companion to Collaborative Statistics provides a number of additional resources for use by students and instructors based on the award winning Elementary Statistics Sofia online course22, also by textbook authors Barbara Illowsky and Susan Dean. This content is designed to complement the textbook by providing video tutorials, course management materials, and sample problem sets. The Supplemental Materials collection can be found at http://cnx.org/content/col10586/latest/23.

Video Lectures

- Video Lecture 1: Sampling and Data24
- Video Lecture 2: Descriptive Statistics25
- Video Lecture 3: Probability Topics26
- Video Lecture 4: Discrete Distributions27
- Video Lecture 5: Continuous Random Variables28
- Video Lecture 6: The Normal Distribution29
- Video Lecture 7: The Central Limit Theorem30
- Video Lecture 8: Confidence Intervals31
- Video Lecture 9: Hypothesis Testing with a Single Mean32
- Video Lecture 10: Hypothesis Testing with Two Means33
- Video Lecture 11: The Chi-Square Distribution34
- Video Lecture 12: Linear Regression and Correlation35

Version History

This module contains a listing of changes, updates, and corrections made to the Collaborative Statistics textbook collection (col10522) since its initial release on 15 July 2008. The Version History is located at http://cnx.org/content/m17360/latest/36.

Textbook Adoption and Usage

This module is designed to track the various derivations of the Collaborative Statistics textbook and its various companion resources, as well as keep track of educators who have adopted various versions for their courses. New adopters are encouraged to provide their contact information and describe how they will use this book for their courses. The goal is to provide a list that will allow educators using this book
to collaborate, share ideas, and make suggestions for future development of this text. The Adoption and Usage module is located at http://cnx.org/content/m18261/latest/.

**Additional Technologies**

In order to provide the most flexible learning resources possible, we invite collaboration from all instructors wishing to create customized versions of this content for use with other technologies. For instance, you may be interested in creating a set of instructions similar to this collection’s calculator notes. If you would like to contribute to this collection, please use the contact the authors with any ideas or materials you have created.

**Accessibility and Section 508 Compliance**

- For information on general Connexions accessibility features, please visit http://cnx.org/content/m17212/latest/.
- For information on accessibility features specific to the Collaborative Statistics textbook, please visit http://cnx.org/content/m17211/latest/.

---

37 “Collaborative Statistics: Adoption and Usage” <http://cnx.org/content/m18261/latest/>
38 “Accessibility Features of Connexions” <http://cnx.org/content/m17212/latest/>
39 “Collaborative Statistics: Accessibility” <http://cnx.org/content/m17211/latest/>
Thank You for previewing this eBook

You can read the full version of this eBook in different formats:

- HTML (Free / Available to everyone)
- PDF / TXT (Available to V.I.P. members. Free Standard members can access up to 5 PDF/TXT eBooks per month each month)
- Epub & Mobipocket (Exclusive to V.I.P. members)

To download this full book, simply select the format you desire below

Free-eBooks.net