

Collaborative Statistics (MT230 - Spring 2012)

Collection Editor:

Robert Gallagher

Collaborative Statistics (MT230 - Spring 2012)

Collection Editor:

Robert Gallagher

Authors:

Susan Dean

Barbara Illowsky, Ph.D.

Online:

< <http://cnx.org/content/col11403/1.1/> >

C O N N E X I O N S

Rice University, Houston, Texas

This selection and arrangement of content as a collection is copyrighted by Robert Gallagher. It is licensed under the Creative Commons Attribution 3.0 license (<http://creativecommons.org/licenses/by/3.0/>).

Collection structure revised: January 10, 2012

PDF generated: June 20, 2012

For copyright and attribution information for the modules contained in this collection, see p. 597.

Table of Contents

Preface	1
Additional Resources	5
Author Acknowledgements	9
Student Welcome Letter	11
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 Variation	18
1.7 Answers and Rounding Off	20
1.8 Frequency	21
1.9 Summary	25
1.10 Practice: Sampling and Data	26
1.11 Homework	29
1.12 Lab 1: Data Collection	37
1.13 Lab 2: Sampling Experiment	39
Solutions	42
2 Descriptive Statistics	
2.1 Descriptive Statistics	45
2.2 Displaying Data	45
2.3 Stem and Leaf Graphs (Stemplots), Line Graphs and Bar Graphs	46
2.4 Histograms	49
2.5 Measures of the Location of the Data	53
2.6 Measures of the Center of the Data	58
2.7 Skewness and the Mean, Median, and Mode	61
2.8 Measures of the Spread of the Data	63
2.9 Summary of Formulas	71
2.10 Practice 1: Center of the Data	72
2.11 Practice 2: Spread of the Data	75
2.12 Homework	76
2.13 Lab: Descriptive Statistics	93
Solutions	95
3 Linear Regression and Correlation	
3.1 Linear Regression and Correlation	103
3.2 Linear Equations	103
3.3 Slope and Y-Intercept of a Linear Equation	105
3.4 Scatter Plots	106
3.5 The Regression Equation	108
3.6 Correlation Coefficient and Coefficient of Determination	114
3.7 Testing the Significance of the Correlation Coefficient	116
3.8 Prediction	121
3.9 Outliers	121
3.10 95% Critical Values of the Sample Correlation Coefficient Table	128
3.11 Summary	131

3.12	Practice: Linear Regression	132
3.13	Homework	135
3.14	Lab 1: Regression (Distance from School)	150
3.15	Lab 2: Regression (Textbook Cost)	153
3.16	Lab 3: Regression (Fuel Efficiency)	155
	Solutions	158
4 Probability Topics		
4.1	Probability Topics	163
4.2	Terminology	164
4.3	Independent and Mutually Exclusive Events	166
4.4	Two Basic Rules of Probability	169
4.5	Contingency Tables	173
4.6	Venn Diagrams (optional)	176
4.7	Tree Diagrams (optional)	177
4.8	Summary of Formulas	182
4.9	Practice 1: Contingency Tables	183
4.10	Practice 2: Calculating Probabilities	185
4.11	Homework	186
4.12	Review	197
4.13	Lab: Probability Topics	199
	Solutions	202
5 Discrete Random Variables		
5.1	Discrete Random Variables	209
5.2	Probability Distribution Function (PDF) for a Discrete Random Variable	210
5.3	Mean or Expected Value and Standard Deviation	211
5.4	Common Discrete Probability Distribution Functions	214
5.5	Binomial	214
5.6	Summary of Functions	218
5.7	Practice 1: Discrete Distribution	220
5.8	Practice 2: Binomial Distribution	221
5.9	Homework	223
5.10	Review	233
5.11	Lab 1: Discrete Distribution (Playing Card Experiment)	236
5.12	Lab 2: Discrete Distribution (Lucky Dice Experiment)	240
	Solutions	244
6 The Normal Distribution		
6.1	The Normal Distribution	251
6.2	The Standard Normal Distribution	252
6.3	Z-scores	253
6.4	Areas to the Left and Right of x	255
6.5	Calculations of Probabilities	255
6.6	Summary of Formulas	259
6.7	Practice: The Normal Distribution	260
6.8	Homework	262
6.9	Review	268
6.10	Lab 1: Normal Distribution (Lap Times)	270
6.11	Lab 2: Normal Distribution (Pinkie Length)	273
	Solutions	275
7 The Central Limit Theorem		
7.1	The Central Limit Theorem	279

7.2	The Central Limit Theorem for Sample Means (Averages)	280
7.3	The Central Limit Theorem for Sums	283
7.4	Using the Central Limit Theorem	284
7.5	Summary of Formulas	292
7.6	Practice: The Central Limit Theorem	293
7.7	Homework	296
7.8	Review	303
7.9	Lab 1: Central Limit Theorem (Pocket Change)	305
7.10	Lab 2: Central Limit Theorem (Cookie Recipes)	309
	Solutions	314
8	Confidence Intervals	
8.1	Confidence Intervals	317
8.2	Confidence Interval, Single Population Mean, Population Standard Deviation Known, Normal	319
8.3	Confidence Interval, Single Population Mean, Standard Deviation Unknown, Student-T	326
8.4	Confidence Interval for a Population Proportion	329
8.5	Summary of Formulas	334
8.6	Practice 1: Confidence Intervals for Averages, Known Population Standard Deviation	335
8.7	Practice 2: Confidence Intervals for Averages, Unknown Population Standard Deviation	337
8.8	Practice 3: Confidence Intervals for Proportions	339
8.9	Homework	341
8.10	Review	351
8.11	Lab 1: Confidence Interval (Home Costs)	354
8.12	Lab 2: Confidence Interval (Place of Birth)	357
8.13	Lab 3: Confidence Interval (Womens' Heights)	359
	Solutions	361
9	Hypothesis Testing: Single Mean and Single Proportion	
9.1	Hypothesis Testing: Single Mean and Single Proportion	367
9.2	Null and Alternate Hypotheses	368
9.3	Outcomes and the Type I and Type II Errors	369
9.4	Distribution Needed for Hypothesis Testing	370
9.5	Assumption	371
9.6	Rare Events	371
9.7	Using the Sample to Support One of the Hypotheses	372
9.8	Decision and Conclusion	373
9.9	Additional Information	373
9.10	Summary of the Hypothesis Test	375
9.11	Examples	375
9.12	Summary of Formulas	386
9.13	Practice 1: Single Mean, Known Population Standard Deviation	387
9.14	Practice 2: Single Mean, Unknown Population Standard Deviation	389
9.15	Practice 3: Single Proportion	391
9.16	Homework	393
9.17	Review	406
9.18	Lab: Hypothesis Testing of a Single Mean and Single Proportion	409
	Solutions	413
10	Hypothesis Testing: Two Means, Paired Data, Two Proportions	

10.1 Hypothesis Testing: Two Population Means and Two Population Proportions	419
10.2 Comparing Two Independent Population Means with Unknown Population Standard Deviations	420
10.3 Comparing Two Independent Population Means with Known Population Stan- dard Deviations	423
10.4 Comparing Two Independent Population Proportions	425
10.5 Matched or Paired Samples	427
10.6 Summary of Types of Hypothesis Tests	432
10.7 Practice 1: Hypothesis Testing for Two Proportions	433
10.8 Practice 2: Hypothesis Testing for Two Averages	435
10.9 Homework	437
10.10 Review	449
10.11 Lab: Hypothesis Testing for Two Means and Two Proportions	451
Solutions	456
11 The Chi-Square Distribution	
11.1 The Chi-Square Distribution	461
11.2 Notation	462
11.3 Facts About the Chi-Square Distribution	462
11.4 Goodness-of-Fit Test	464
11.5 Test of Independence	471
11.6 Summary of Formulas	476
11.7 Practice 1: Goodness-of-Fit Test	477
11.8 Practice 2: Contingency Tables	479
11.9 Homework	481
11.10 Review	489
11.11 Lab 1: Chi-Square Goodness-of-Fit	493
11.12 Lab 2: Chi-Square Test for Independence	498
Solutions	500
12 F Distribution and ANOVA	
12.1 F Distribution and ANOVA	505
12.2 ANOVA	506
12.3 The F Distribution and the F Ratio	506
12.4 Facts About the F Distribution	508
12.5 Summary	513
12.6 Practice: ANOVA	514
12.7 Homework	516
12.8 Review	518
12.9 Lab: ANOVA	522
Solutions	524
13 Appendix	
13.1 Practice Final Exam 1	527
13.2 Practice Final Exam 2	536
13.3 Data Sets	545
13.4 Group Projects	548
13.5 Solution Sheets	559
13.6 English Phrases Written Mathematically	563
13.7 Symbols and their Meanings	564
13.8 Formulas	569
13.9 Notes for the TI-83, 83+, 84 Calculator	571
Solutions	581

14 Tables	585
Glossary	586
Index	593
Attributions	597

Preface¹

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.org²) 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 license³ 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"⁴. For a guide to authoring modules, please look at the help page, "Create a Module in Minutes"⁵.

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/latest⁶. You will then have three options.

¹This content is available online at <<http://cnx.org/content/m16026/1.16/>>.

²<http://cnx.org/>

³<http://creativecommons.org/licenses/by/2.0/>

⁴<http://cnx.org/help/CreateCollection>

⁵<http://cnx.org/help/ModuleInMinutes>

⁶*Collaborative 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/>⁷.
- For information on accessibility features specific to the Collaborative Statistics textbook, please visit <http://cnx.org/content/m17211/latest/>⁸.

Version Change History and Errata

- For a list of modifications, updates, and corrections, please visit <http://cnx.org/content/m17360/latest/>⁹.

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/>¹⁰.

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/>¹¹, and supplemental course materials including additional problem sets and video lectures are available at <http://cnx.org/content/col10586/latest/>¹². The on-line text for each of these 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 Resources¹³

Additional Resources Currently Available

- Glossary (Glossary, p. 5)
- View or Download This Textbook Online (View or Download This Textbook Online, p. 5)
- Collaborative Statistics Teacher's Guide (Collaborative Statistics Teacher's Guide, p. 5)
- Supplemental Materials (Supplemental Materials, p. 5)
- 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. 6)
- Accessibility and Section 508 Compliance (Accessibility and Section 508 Compliance, p. 6)

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/>¹⁴) 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/>¹⁵.

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/>¹⁶. 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 here¹⁷.

Collaborative Statistics Teacher's Guide

A complementary Teacher's Guide for Collaborative statistics is available through Connexions at <http://cnx.org/content/col10547/latest/>¹⁸. 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 here¹⁹.

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 course²⁰, also by textbook

¹³This content is available online at <http://cnx.org/content/m18746/1.6/>.

¹⁴<http://cnx.org/>

¹⁵"Collaborative Statistics: Glossary" <http://cnx.org/content/m16129/latest/>

¹⁶*Collaborative Statistics* <http://cnx.org/content/col10522/latest/>

¹⁷<http://my.qoop.com/store/7064943342106149/7781159220340>

¹⁸*Collaborative Statistics Teacher's Guide* <http://cnx.org/content/col10547/latest/>

¹⁹<http://my.qoop.com/store/7064943342106149/8791310589747>

²⁰<http://sofia.fhda.edu/gallery/statistics/index.html>

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/>²¹.

Video Lectures

- Video Lecture 1: Sampling and Data²²
- Video Lecture 2: Descriptive Statistics²³
- Video Lecture 3: Probability Topics²⁴
- Video Lecture 4: Discrete Distributions²⁵
- Video Lecture 5: Continuous Random Variables²⁶
- Video Lecture 6: The Normal Distribution²⁷
- Video Lecture 7: The Central Limit Theorem²⁸
- Video Lecture 8: Confidence Intervals²⁹
- Video Lecture 9: Hypothesis Testing with a Single Mean³⁰
- Video Lecture 10: Hypothesis Testing with Two Means³¹
- Video Lecture 11: The Chi-Square Distribution³²
- Video Lecture 12: Linear Regression and Correlation³³

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/>³⁴.

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/>³⁵.

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

²¹ *Collaborative Statistics: Supplemental Course Materials* <<http://cnx.org/content/col10586/latest/>>

²² "Elementary Statistics: Video Lecture - Sampling and Data" <<http://cnx.org/content/m17561/latest/>>

²³ "Elementary Statistics: Video Lecture - Descriptive Statistics" <<http://cnx.org/content/m17562/latest/>>

²⁴ "Elementary Statistics: Video Lecture - Probability Topics" <<http://cnx.org/content/m17563/latest/>>

²⁵ "Elementary Statistics: Video Lecture - Discrete Distributions" <<http://cnx.org/content/m17565/latest/>>

²⁶ "Elementary Statistics: Video Lecture - Continuous Random Variables" <<http://cnx.org/content/m17566/latest/>>

²⁷ "Elementary Statistics: Video Lecture - The Normal Distribution" <<http://cnx.org/content/m17567/latest/>>

²⁸ "Elementary Statistics: Video Lecture - The Central Limit Theorem" <<http://cnx.org/content/m17568/latest/>>

²⁹ "Elementary Statistics: Video Lecture - Confidence Intervals" <<http://cnx.org/content/m17569/latest/>>

³⁰ "Elementary Statistics: Video Lecture - Hypothesis Testing with a Single Mean" <<http://cnx.org/content/m17570/latest/>>

³¹ "Elementary Statistics: Video Lecture - Hypothesis Testing with Two Means" <<http://cnx.org/content/m17577/latest/>>

³² "Elementary Statistics: Video Lecture - The Chi-Square Distribution" <<http://cnx.org/content/m17571/latest/>>

³³ "Elementary Statistics: Video Lecture - Linear Regression and Correlation" <<http://cnx.org/content/m17572/latest/>>

³⁴ "Collaborative Statistics: Change History" <<http://cnx.org/content/m17360/latest/>>

³⁵ "Collaborative Statistics: Adoption and Usage" <<http://cnx.org/content/m18261/latest/>>

- For information on general Connexions accessibility features, please visit <http://cnx.org/content/m17212/latest/>³⁶.
- For information on accessibility features specific to the Collaborative Statistics textbook, please visit <http://cnx.org/content/m17211/latest/>³⁷.

³⁶"Accessibility Features of Connexions" <<http://cnx.org/content/m17212/latest/>>

³⁷"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

