ENGINEERING STATISTICS WITH DESIGN OF EXPERIMENTS

PROF. DARIUSZ WANATOWSKI
万岱历 教授
PRO-DEAN OF THE SWJTU-LEEDS JOINT SCHOOL
西南交通大学-利兹学院 副院长
EMAIL: DWANAT@GMAIL.COM
OFFICE: X30530

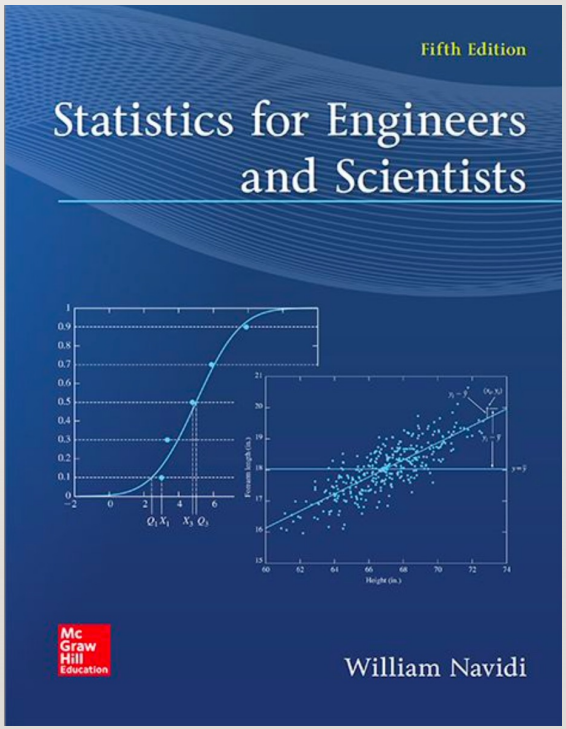
OVERVIEW

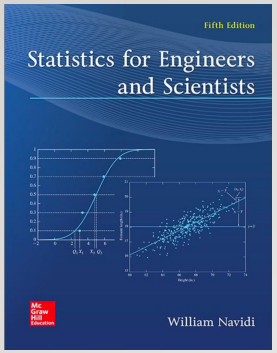
- **Class time:** Monday and/or Thursday **19:30-21:55** (Weeks 1-14, the **Final Exam in Week 14**)
- **Classroom:** X30404
- **Course description:**

  An introductory course to random variables and basic probability distributions, estimation, confidence intervals, hypothesis testing, basic analysis of variance, factorial arrangement of treatments and fractional factorial experiments, elementary quality control.

REQUIRED TEXTBOOK (**YOU MUST HAVE IT**)

- **Statistics for Engineers and Scientists**, 5th edition, by William Navidi





Statistics for Engineers and Scientists, 5th edition, by William Navidi

KEY FEATURES

- The book is flexible in its presentation of probability
- The book contains many examples that feature real, contemporary data sets to show connections to industry and scientific research.
- The book provides extensive coverage of propagation of error.
- The book presents a solid introduction to simulation methods including applications to verifying normality assumptions, computing probabilities, estimating bias, computing confidence intervals, and testing hypotheses.
- The book covers the standard introductory topics, including descriptive statistics, probability, confidence intervals, hypothesis tests, linear regression, factorial experiments, and statistical quality control.

BRIEF CONTENTS

Preface xi

1 Sampling and Descriptive Statistics 1

2 Probability 48

3 Propagation of Error 164

4 Commonly Used Distributions 200

5 Confidence Intervals 323

6 Hypothesis Testing 402

7 Correlation and Simple Linear Regression 514

8 Multiple Regression 603

9 Factorial Experiments 669

10 Statistical Quality Control 772

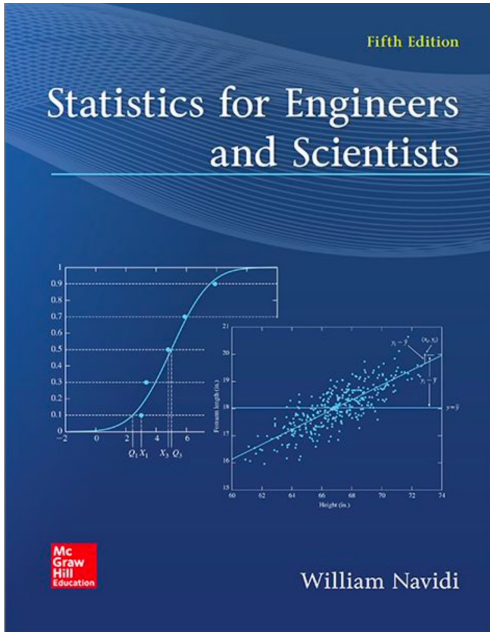
Appendix A: Tables 811

Appendix B: Partial Derivatives 836

Appendix C: Bibliography 838

Answers to Odd-Numbered Exercises 841

Index 912



CONTENTS

Preface xi

Chapter 1

Sampling and Descriptive Statistics 1

Introduction 1

1.1 Sampling 3

1.2 Summary Statistics 13

1.3 Graphical Summaries 25

Chapter 2

Probability 48

Introduction 48

2.1 Basic Ideas 48

2.2 Counting Methods 62

2.3 Conditional Probability and Independence 69

2.4 Random Variables 90

2.5 Linear Functions of Random Variables 116

2.6 Jointly Distributed Random Variables 127

Chapter 3

Propagation of Error 164

Introduction 164

3.1 Measurement Error 164

3.2 Linear Combinations of Measurements 170

3.3 Uncertainties for Functions of One Measurement 180

3.4 Uncertainties for Functions of Several Measurements 186

Chapter 4

Commonly Used Distributions 200

Introduction 200

4.1 The Bernoulli Distribution 200

4.2 The Binomial Distribution 203

4.3 The Poisson Distribution 215

4.4 Some Other Discrete Distributions 230

4.5 The Normal Distribution 241

4.6 The Lognormal Distribution 256

4.7 The Exponential Distribution 262

4.8 Some Other Continuous Distributions 272

4.9 Some Principles of Point Estimation 280

4.10 Probability Plots 285

4.11 The Central Limit Theorem 290

4.12 Simulation 303

Chapter 5

Confidence Intervals 323

Introduction 323

5.1 Large-Sample Confidence Intervals for a Population Mean 324

5.2 Confidence Intervals for Proportions 339

5.3 Small-Sample Confidence Intervals for a Population Mean 345

5.4 Confidence Intervals for the Difference Between Two Means 355

5.5 Confidence Intervals for the Difference Between Two Proportions 359

5.6 Small-Sample Confidence Intervals for the Difference Between Two Means 364

5.7 Confidence Intervals with Paired Data 371

5.8 Confidence Intervals for the Variance and Standard Deviation of a Normal Population 376

5.9 Prediction Intervals and Tolerance Intervals 381

Chapter 6

Hypothesis Testing 402

Introduction 402

6.1 Large-Sample Tests for a Population Mean 402

6.2 Drawing Conclusions from the Results of Hypothesis Tests 412

6.3 Tests for a Population Proportion 420

6.4 Small-Sample Tests for a Population Mean 425

6.5 Large-Sample Tests for the Difference Between Two Means 431

6.6 Tests for the Difference Between Two Proportions 437

6.7 Small-Sample Tests for the Difference Between Two Means 443

6.8 Tests with Paired Data 452

6.9 Distribution-Free Tests 458

6.10 Tests with Categorical Data 467

6.11 Tests for Variances of Normal Populations 477

6.12 Fixed-Level Testing 483

6.13 Power 488

6.14 Multiple Tests 497

6.15 Using Simulation to Perform Hypothesis Tests 501

Chapter 7

Correlation and Simple Linear Regression 514

Introduction 514

7.1 Correlation 514

7.2 The Least-Squares Line 532

7.3 Uncertainties in the Least-Squares Coefficients 549

7.4 Checking Assumptions and Transforming Data 570

Chapter 8

Multiple Regression 603

Introduction 603

8.1 The Multiple Regression Model 603

8.2 Confounding and Collinearity 621

8.3 Model Selection 630

Chapter 9

Factorial Experiments 669

Introduction 669

9.1 One-Factor Experiments 669

9.2 Pairwise Comparisons in One-Factor Experiments 694

9.3 Two-Factor Experiments 707

9.4 Randomized Complete Block Designs 732

9.5 2^p Factorial Experiments 742

Chapter 10

Statistical Quality Control 772

Introduction 772

10.1 Basic Ideas 772

10.2 Control Charts for Variables 775

10.3 Control Charts for Attributes 795

10.4 The CUSUM Chart 800

10.5 Process Capability 804

Appendix A: Tables 811

Appendix B: Partial Derivatives 836

Appendix C: Bibliography 838

Answers to Odd-Numbered Exercises 841

Exercises for Section 1.1

1. Each of the following processes involves sampling from a population. Define the population, and state whether it is tangible or conceptual.

a. A chemical process is run 15 times, and the yield is measured each time.

b. A pollster samples 1000 registered voters in a certain state and asks them which candidate they support for governor.

c. In a clinical trial to test a new drug that is designed to lower cholesterol, 100 people with high cholesterol levels are recruited to try the new drug.

d. Eight concrete specimens are constructed from a new formulation, and the compressive strength of each is measured.

e. A quality engineer needs to estimate the percentage of bolts manufactured on a certain day that meet a strength specification. At 3:00 in the afternoon he samples the last 100 bolts to be manufactured.

2. If you wanted to estimate the mean height of all the students at a university, which one of the following sampling strategies would be best? Why? Note that none of the methods are true simple random samples.

i. Measure the heights of 50 students found in the gym during basketball intramurals.

ii. Measure the heights of all engineering majors.

iii. Measure the heights of the students selected by choosing the first name on each page of the campus phone book.

Answers to Odd-Numbered Exercises

Section 1.1

1. (a) The population consists of all the times the process could be run. It is conceptual.

(b) The population consists of all the registered voters in the state. It is tangible.

(c) The population consists of all people with high cholesterol levels. It is tangible.

(d) The population consists of all concrete specimens that could be made from the new formulation. It is conceptual.

(e) The population consists of all bolts manufactured that day. It is tangible.

3. (a) False

(b) True

5. (a) No. What is important is the population proportion of defectives; the sample percentage is only an approximation. The population proportion for the new process may in fact be greater or less than that of the old process.

(b) No. The population proportion for the new process may be 12% or more, even though the sample proportion was only 11%.

(c) Finding two defective circuits in the sample.

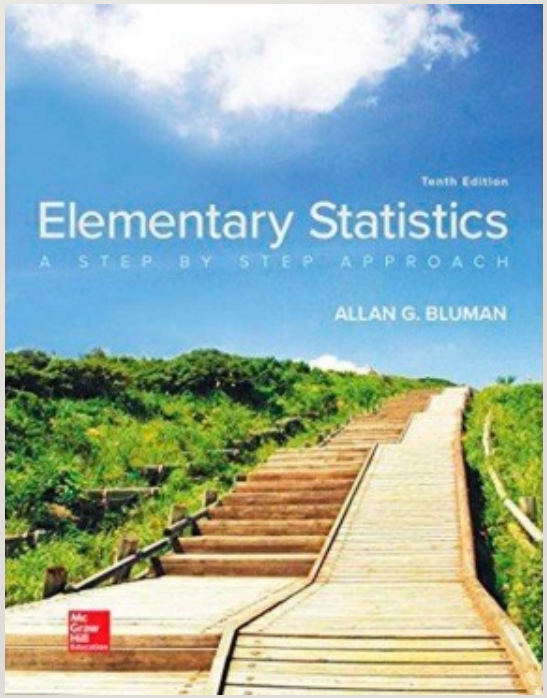
7. A good knowledge of the process that generated the data.

9. (a) A controlled experiment.

(b) Yes, because it is based on a controlled experiment rather than an observational study.

ADDITIONAL TEXTBOOK (YOU DO NOT NEED TO HAVE IT)

- **Statistics for Engineers and Scientists**, 10th edition, by Allan G. Bluman



GRADING POLICY

Homework 1	**10%**
Homework 2	10%
Mid-term Exam	30%
Final Exam	50%
Extra Credit	Max. 5%

FINAL GRADE SCALE

According to the agreement between SWJTU and OSU, the following scale will be used to assign the final grades:

- **A = 85-100%**
- B = 75-84%
- C = 60-74
- **F = 0-59%**

ACADEMIC INTEGRITY

- In this course, homework should represent your **individual effort**, unless explicitly stated in the assignment.
- You may talk with other students and tutors about assignments, but you should work through the computations and submit **your own answers**.

ATTENDANCE AND CLASS PARTICIPATION

- Attendance is **COMPULSORY**.
- Active participation is highly encouraged.
- It is **YOUR responsibility** to come to class prepared.
- You will be responsible for bringing your textbook **EVERY WEEK** and making your own notes in the lectures.
- The tentative timetable of the course is provided. Please make sure you have all necessary materials to participate in each lecture.

TENTATIVE TIMETABLE

Teaching week	Date	Teaching Contents	Requirement of students	Teaching time		Notes
				Lecture	Practice	
1	2nd September (Monday)	Ch 1: Sampling and Descriptive Statistics	Textbook, calculator	2	1	-----
2	9th September (Monday)	Ch 2: Probability	Textbook, calculator	2	1	-----
4	23rd September (Monday)	Ch 3: Propagation of Error	Textbook, calculator	2	1	**Homework 1 issued**
6	10th October (Thursday)	Ch 4: Commonly Used Distributions	Textbook, calculator	2	1	-----
7	14th October (Monday)	Ch 4: Commonly Used Distributions (continued)	Textbook, calculator	2	1	**Homework 1 due**

TENTATIVE TIMETABLE

8	21st October (Monday)	Ch 5: Confidence Intervals (continued)	Textbook, calculator	1	2	-----
8	**24th October (Thursday)**	**Mid-term Exam (Ch 1-5)**	**Calculator, dictionary**	**-----**	**2**	**2hr Midterm Exam**
9	28th October (Monday)	Ch 6: Hypothesis Testing	Textbook, calculator	2	1	-----
9	31st October (Thursday)	Ch 6: Hypothesis Testing (continued)	Textbook, calculator	2	1	-----
10	4th November (Monday)	Ch 7: Correlation and Simple Linear Regression	Textbook, calculator	2	1	-----

TENTATIVE TIMETABLE

11	11th November (Monday)	Ch 8: Multiple Regression	Textbook, calculator	1	2	**Homework 2 issued**
12	18th November (Monday)	Ch 9: Factorial Experiments	Textbook, calculator	2	1	-----
13	25th November (Monday)	Ch 10: Statistical Quality Control	Textbook, calculator	2	1	**Homework 2 due**
14	2nd December (Monday)	Ch 10: Statistical Quality Control (continued) and Revision	Textbook, calculator	2	1	-----
14	**5th December (Thursday)**	**FINAL EXAM (Ch 1-10)**	**Calculator, Dictionary**	**-----**	**2**	**2hr FINAL EXAM**

Note: For details of Chapters and Sections please refer to textbook *Statistics for Engineers and Scientists, 5th Edition, by W. Navidi.*

OFFICE HOURS

- By appointment
- Office: X30530
- You can also email me or post a question on the course Wechat group

HOW CAN YOU DO WELL IN THIS COURSE ???

- Come prepared to the class.
- Participate in the class.
- Get familiar with a new vocabulary, if necessary.
- Practice example problems from the textbook and other resources.
- Solve homework problems yourself and do not worry too much about homework grades.
- Revise well before exams.

QUESTIONS ???