

Statistics

Concepts and Applications



A m i r D . A c z e l



Statistics

Concepts and Applications

A m i r D. A c z e l

Bentley College

IRWIN

Chicago • Bogotá • Boston • Buenos Aires • Caracas
London • Madrid • Mexico City • Sydney • Toronto

© RICHARD D. IRWIN, INC., 1995

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

Senior sponsoring editor:	Richard T. Hercher, Jr.
Associate editor:	Colleen Tuscher
Development:	Burrston House
Project editor:	Stephanie M. Britt
Production supervisor:	Laurie Kersch
Interior designer:	Stuart Paterson; Imagehouse, Inc.
Cover designer:	Stuart Paterson; Imagehouse, Inc.
Cover photographer:	Art Resource
Art studio:	Electronic Publishing Services
Graphics supervisor:	Eurnice Harris
Compositor:	Carlisle Communications, Ltd.
Typeface:	10/12 Times Roman
Printer:	Von Hoffmann Press, Inc.

Library of Congress Cataloging-in-Publication Data

Aczel, Amir D.
Statistics : concepts and applications / Amir D. Aczel
p. cm.
Includes index.
ISBN 0-256-11935-X
1. Statistics. I. Title.
QA276. 12.A23 1995
519.5—dc20

94-41399
CP

Printed in the United States of America

2 3 4 5 6 7 8 9 0 VH 1 0 9 8 7 6 5

For Debra



Preface

We live in the information age. The half-life of human knowledge is currently ten years overall, and only two years in science. This means that new knowledge accumulates so fast that in ten years (and a mere two years in science), half of what we now know will no longer be valid and will be replaced with new knowledge. The methods for gathering, assessing, understanding and interpreting, and drawing inferences from data constitute the field of *statistics*.

Yet, despite the great importance of the field of statistics, it seems that we have not been able to teach this discipline well. In almost two decades of dealing with students of different levels, backgrounds, and prior statistical education, I have rarely met a student who—a semester or more after completing an introductory statistics course—could demonstrate a true understanding of basic statistical concepts. I believe that one of the culprits is poorly written statistics textbooks.

Most statistics textbooks have fallen into one of two categories. In the first group are books that are theoretically correct, but abstract, mathematical, and intimidating to the student. They have no direct relevance to, or bearing on, the real world. These books perpetuate the well-known “Oh no! Not statistics!” mentality among beginning students. In the second category are books designed to be very user friendly. These books aim at what their authors perceive as the level of understanding of today’s student—they focus on simple “cookbook” procedures and calculations. In my opinion, such books miss the boat completely. They tend to avoid the important concepts, their explanations tend to be superficial, and ultimately they do not really motivate

or teach statistics. Students who use such books come out of the introductory course with no lasting understanding of this important discipline or appreciation of its applications. Days after taking the final exam, students cannot remember what a p -value is, and in the future they are likely to misuse statistical methods and misinterpret results.

WHAT’S DIFFERENT, AND WHY?

The key features of this book that distinguish it from the competition are

- Superior treatment of the topics that usually cause the most difficulty for students, such as sampling, confidence intervals, and hypothesis testing.
- Optimal, research-based choice of topics to be covered.
- Careful selection of stimulating, relevant, and instructive examples that motivate the student and clarify key concepts.
- Wide variety of problems covering all levels and application areas.
- A flexible, modern introduction to statistical use of the computer, which stresses application rather than “syntax.”
- A unique cross-referencing system for key concepts.
- A friendly writing style and an approach that makes statistics understandable and interesting.

In writing this book, I felt that I could make statistics both interesting and useful to students—I have been doing this for many years, teaching stu-

dents of very different backgrounds and interests in many locations. The variety of students and their interests and abilities helped me develop methods for teaching statistics in a meaningful way to *anyone*. I have been aided as well by my academic research and publications, which have kept me on the leading edge of statistical knowledge. This book goes the extra mile in the sense that rather than shy away from the important and beautiful theory of statistics, the explanations of statistical ideas are thorough and relevant. I have tried to explain every concept in a clear way, often with a picture, to show the student what these important elements of statistics actually mean. Overall, reviewers have been uniformly and genuinely enthusiastic in their assessment of how well I was able to accomplish this task.

HOW THE BOOK WAS DEVELOPED

More than any other statistics book, this text was designed specifically for the needs of the student and the instructor in an introductory course. Richard D. Irwin collaborated with Burrston House, Ltd., a publishing house specializing in developing the highest-quality textbooks, for the purpose of researching the market to define the optimal content for a text designed to meet the objectives of this course.

In the first phase of this exhaustive research effort, a survey was sent to a large random sample of instructors of statistics. Respondents were asked a variety of questions about their courses, the composition of the class in terms of majors and interests, and which topics students found to be most difficult or confusing. The information and insights gleaned from the survey became the cornerstone of this book.

In the next phase, a prospectus, table of contents and detailed questionnaire were sent out for review by a total of 72 experienced instructors. Their detailed feedback was used in writing the initial draft. Next, a group of eight instructors carefully reviewed and critiqued the entire manuscript. This process was followed by a conference with first-draft reviewers at the Joint Statistical Meetings in San Francisco in 1993. Based on this input, the second draft was composed. A panel of 15 instructors evaluated each chapter in the revised manuscript. Finally, in the third draft, reviewers' comments were used in yet another revision of the text.

Here is a summary of some of what we learned from this research and reviewing process.

- The three topics causing the most difficulty for students in this course are probability, hypothesis testing (including p -values), and sampling distributions. All three topics as covered in this book received extremely high marks from reviewers. Most said that the treatment of these critical topics was superior to that in any other book they have used.
- Very few instructors (less than 15%) teach the Poisson distribution at this level; the same holds true for multiple regression, two-way ANOVA, and nonparametrics. Congruent with this virtual consensus, we do not cover these topics.
- Reviewers responded favorably to the integrated inclusion of the topics of sampling distributions and confidence intervals within the same chapter, and found the combination especially successful in teaching both topics more effectively.
- Reviewers were uniformly impressed with and enthusiastic about the novel use of contemporary news articles and other excerpts in teaching statistics in the examples and problems.

This book contains the following additional features, all endorsed by the majority of reviewers, and very important for effective teaching of modern statistics.

Chapter Opening Stories. Statistics is more than a science, and more than an art. One new approach to teaching statistics is to view it as a liberal art. Every chapter in the book opens with what I believe is a fascinating story. I have collected these stories over years of research, often going to the exact geographical location—an island in the Mediterranean, Westminster Abbey, or the Bancroft Library—in search of original material with statistical relevance and interest. Based on their use in my own teaching, I am confident that these stories will make the students more interested in the subject matter. Virtually all the reviewers of this book were complimentary about the chapter openers, describing them as a great motivating tool for the student.

Extensive Illustrations. Nowhere is the saying “a picture is worth a thousand words” truer than here, in trying to explain statistical concepts. This text has a number of photos as well as graphs that are unique. In the regression chapter, for example, I use a human eye looking at a data set from different angles to

demonstrate the idea of the variance of Y as compared with the variance of the data around the regression line. The figures aim to enhance the student's intuition for, and comfort with, data and their display and analysis.

Everyday-Life Examples from Current Newspapers and Magazines. Statistics surround us. One cannot open a newspaper without seeing the results of some poll showing that 70 percent of Americans do not exercise, or that 47 percent approve of the manner the president handled this or that issue; and what causes cancer? or heart attacks? is a Mediterranean diet healthier than others? Statistical issues—in fact, statistical *inference* in the essence—are everywhere. By reprinting interesting and relevant articles from *The New York Times* and other leading newspapers as well as leading national magazines, I was able to bring these issues from everyday life to the students' attention. I often ask the student to analyze the statistical issues raised in the articles. This feature, too, was highly praised as a key element in differentiating this book.

Cross-Referencing System. For students to understand and use statistics, they need certain key concepts, or building blocks, of the subject (e.g. standard deviation, central limit theorem, or a random variable). To assist the students in this process I have included a unique cross-referencing system. Throughout this text, when students encounter an icon symbolizing links in a chain following a term, they should recognize it as a key concept. In later chapters, when this building block is used with other concepts, the icon will appear again, with the page reference to where the concept was originally developed directly underneath the icon. If the students need a quick review of the earlier topic to enhance their understanding of the current concept, they can easily do so. This study aid was ranked first (out of twelve) in utility by students in recent research conducted by the publisher.

Interviews. I was fortunate enough to have a group of leading experts and real-world practitioners in a number of fields agree to be interviewed for the book. The interviewees talk to the student about their professions and how statistics is important in their areas of expertise. For example, celebrated defense attorney and Harvard Professor Alan M. Dershowitz

discusses how the concepts of probability and statistical inference are used—and how, in his opinion, they *should* be used—by defense attorneys in criminal trials. One of this century's leading statisticians, Professor Bradley Efron of Stanford University, tells us how he used statistics in verifying Hubble's Law, as well as his view of the future of statistics and the role computing will have. Other interviews span a variety of fields and subjects, including television, the NFL, government, quality control, and medicine. Each interview is placed in a stand-alone module between chapters.

A Global Perspective. In today's world, a text that does not address global issues—population, economics, or the environment—does a disservice to the student. A highlight of this book is the international setting and character of many of the problems and examples. A graphical comparison of life in India and China, the risk of malaria in various countries, problems on exchange rates at Orly airport in Paris, the number of miles driven by a tourist in Europe, and using statistics in scientific efforts to save the Colosseum in Rome from collapse are just a few examples.

Comprehensive, Thoroughly Explained Examples. There are few things in an introductory textbook that are as important as the examples. I have chosen examples carefully from a large variety of fields. The situations are all realistic and often include the actual data. The student is taken through each example step by step, de-emphasizing formulas and concentrating on concepts and their application. The examples do not end with a numerical answer—interpretations are offered to show the student how to use and understand the results obtained from statistics.

Ample Real-World Problems. The problem sets in this book were highly praised during the review process. They were chosen from many fields—medicine, education, archaeology, social science, psychology, engineering, nursing, business, economics, and more. The problems are leveled: Starting with simple exercises in each section to build the student's confidence, and leading to more involved problems that make the student think not only about statistics, but about the problem area and its relevance. I have also included some problems that refer to large data sets at the end of the book, and some that ask for a written summary or report.

Integrated Computer Use. Since statistics is most often applied with the aid of computers, the computer applications are developed throughout the text, starting in Chapter 1. The package MINITAB was chosen as the example package for its ease of use. The commands are explained throughout the book as needed, so that the student does not need to consult a manual and can use the computer with ease early on. Certain problems are designated as Computer Problems and marked with an icon. These problems may also be solved by hand, but their preferred solution is by computer.

The Entire Package. The book is accompanied by a complete array of teaching and learning aids. These include:

- An Instructor's Solutions Manual.
- A Student's Solutions Guide.
- An Instructor Resource Guide.
- An Instructor's Transparency Master.
- A Data Disk.
- A Test Bank.
- A Computerized Test Bank.

In addition, there are several packaging options available from the publisher.

ACKNOWLEDGMENTS

I am deeply indebted to the following statistics instructors who have evaluated all ten chapters and with their comments, ideas, and advice have influenced every page of this text. Without their encouragement, willingness, and dedication to assist in producing a better text it would not have happened. Each of them will recognize their contributions and I'm hopeful they are pleased with the result. Special gratitude is extended to Patricia Buchanan of Pennsylvania State University, Chris Franklin of the University of Georgia, and Joseph Ibrahim of Harvard University, all of whom labored through two entire drafts. Chris and Joe also participated in the reviewer conference at the ASA meetings.

First Draft Reviewers

Patricia Buchanan, Pennsylvania State University
Shyam Chadha, University of Wisconsin-Eau Claire

Chris Franklin, University of Georgia
Brenda Gunderson, University of Michigan
Joseph Ibrahim, Harvard University
Y. Leon Maksoudian, California State Polytechnic University-San Louis Obispo
E. D. McCune, Stephen F. Austin State University
Mack Shelley, Iowa State University
Nola Tracey, McNeese State University

Second Draft Reviewers

William Beyer, University of Akron
Patricia Buchanan, Pennsylvania State University
Michael Doviak, Old Dominion University
Eugene Enneking, Portland State University
Chris Franklin, University of Georgia
Gavin Gregory, University of Texas-El Paso
Bernard Harris, University of Wisconsin-Madison
Jason Hsu, Ohio State University
Joseph Ibrahim, Harvard University
J. David Mason, University of Utah
Ronald McCuiston, Pensacola Junior College
Linda McDonald, University of Texas-Arlington
Wendy McGuire, Santa Fe Community College
John Unbehaun, University of Wisconsin-LaCrosse
John Whitesitt, Southern Oregon State College

As mentioned previously, the following instructors helped shape the scope, sequence, and depth of coverage in this text by completing an in-depth questionnaire, evaluating a preliminary table of contents, and critiquing a sample chapter. Their collective judgment was instrumental in putting this project on the right track from the outset.

Madeline A. Bradley, University of North Carolina-Greensboro
Shyam Chadha, University of Wisconsin-Eau Claire
Harold Jacobs, East Stroudsburg University
Jim Jacobs, Valencia Community College
Phillip E. Johnson, University of North Carolina-Charlotte
Don Loftsgarden, University of Montana
Gary Klingler, Grand Valley State University
Linda McDonald, University of Texas-Arlington

David K., Neal, Western Kentucky University
 John Whitesitt, Southern Oregon State College
 Shelemyahu Zacks, State University of New
 York-Binghamton

I also wish to thank the following instructors and students who read the interviews and gave me advice regarding the types of individuals and subjects they think students taking an introductory statistics course would find valuable and interesting:

E. Jacquelin Dietz, North Carolina State University
 Michael Doviak, Old Dominion University
 Chris Franklin, University of Georgia
 Susan K. Herring, Sonoma State University
 Joseph Ibrahim, Harvard University
 Barry Katz, St. Louis University
 J. David Mason, University of Utah

University of Georgia Students: Heather Adams, Matthew Freeman, Wood Pope, Trent Schueneman, Lisa Spitz. Thank you to Carl Mueller and Gene Enneking, who double checked all the calculations and solutions in the text. They were both extremely thorough and this text is much cleaner and more correct as a result. I take full responsibility for any errors that remain.

Thank you also to the interviewees for agreeing to provide their candid and detailed comments and information about their fields and the role statistics play in them. When we set out to do the interviews, we wanted to get the real, detailed, dramatic, inside scoop, and all of the interviewees were overwhelmingly supportive of that effort. In fact, they all graciously provided much more than we ended up using in the text. There were several goals we had in mind, that I believe are accomplished in the final versions included in the text. First, we wanted to provide the student with a personal insight into the people who participate in the field of statistics. Many students leave their course with the impression statistics is strictly numbers and mathematics. We tried therefore to include some biographical background, real personality, and some authority.

Second, we wanted the interviews to be 'deep' enough to be of real value to read, and honest. The interviewees didn't hesitate, and we hope that you and your students will appreciate the candor. Finally, we wanted to provide a broad range of application areas and show students how useful and important

statistics are in the world. A quick summary of who the interviewees are and why they are included is listed below.

Melissa Anderson was interviewed about statistics in education as well as issues in research and ethics. Her academic interest is in the study of education, and she has worked with the Acadia Institute.

Alan Dershowitz was interviewed about how statistics, especially probability, is used and not used in Law, particularly in criminal defense cases.

Arnold Zellner is included to provide background on the field of statistics, the Bayesian approach, and the interplay of economics and statistics.

Bud Goode talks about a popular field, sports, particularly the National Football League, and how he uses various statistics to analyze action on the gridiron.

Bradley Efron provides a research perspective on trends in the field of Statistics and on advances in the field based on improving uses of the computer.

Lee Wilkinson describes how his dual interests in statistics and computers led him into the role of entrepreneur, and founding Systat, Inc.

Toni Falvo provided an overview of the use of data and market research in programming and advertising decisions at WLS, a Chicago television station.

Nancy Kirkendall describes the information gathering and analysis used by the federal government, in this case, to monitor and plan for U.S. energy consumption.

Mary Guinan is active in the battle against AIDS and her interview describes some of the statistical methods and issues which are important to the CDC.

Phil Crosby is interviewed about the relationship between traditional statistics and the current trend toward Total Quality Management.

My warmest thanks also to the following friends and colleagues: Ray Ledoux, Art O'Leary, Professors Scott Callan, Richard Frese, Richard Fristensky, Steven Grubaugh, David Gulley, Charles Hadlock, Dominique Haughton, Jack Hegarty, Erl Sorensen, Andrew Stollar, Nicholas Teebagy, James Zeitler. I am also indebted to my friend Mario Demanuele of Malta for his kind help with pictures of the Mosta bomb.

Warm thanks to my editor, Richard T. Hercher Jr of Irwin, for his encouragement and direction throughout this entire project, and for his many

excellent ideas and suggestions. I thank Colleen Tuscher, the associate editor, for her enthusiasm and for her many contributions to the book. Many thanks to Stephanie Britt, the project editor, for her superb editing of the manuscript. Thanks also go to Stuart Paterson, Imagehouse, Inc., for the interior and cover design.

I want to express my deep gratitude to Glenn Turner of Burrston House for his wise suggestions,

direction, and encouragement, and for his countless insights in developing this book. Thanks also to Meg Turner for her extensive help on this project. I thank the staff at Burrston House for their efforts.

Finally, I thank my wife, Debra, for all her help and for some of the photographs.

Amir D. Aczel



Contents in Brief

CHAPTER 1 Information Everywhere 2

Interview with Melissa Anderson 67

CHAPTER 2 What Are the Chances? 70

Interview with Alan Dershowitz 107

CHAPTER 3 Chance Quantities 112

Interview with Arnold Zellner 151

CHAPTER 4 The Bell-Shaped Curve 154

Interview with Bud Goode 185

CHAPTER 5 Let's Take a Sample 190

Interview with Bradley Efron 239

CHAPTER 6 Trial by Probability 244

Interview with Lee Wilkinson 309

CHAPTER 7 Making Comparisons 314

Interview with Toni Falvo 351

CHAPTER 8 Making Extended
Comparisons 356

Interview with Nancy Kirkendall 401

CHAPTER 9 Is There a Relationship? 406

Interview with Mary Guinan 461

CHAPTER 10 The Quest for Quality 464

Interview with Phil Crosby 481



Contents

CHAPTER 1 Information Everywhere 2

- 1.1 Introduction 2
 - Samples and Populations 5
 - Data and Data Collection 5
 - Data Distributions 7
- 1.2 Methods of Displaying Data 9
 - Pie Charts 9
 - Bar Graphs 12
 - The Histogram 13
 - Stemplots (Stem-and-Leaf Displays) 15
 - Frequency Polygons and Ogives 19
 - A Caution about Graphs 21
- 1.3 Percentiles, Quartiles, and the Median 35
 - Percentiles 35
 - Quartiles 35
 - The Median 36
- 1.4 Box Plots 42
- 1.5 Measures of Centrality 45
- 1.6 Measures of Variability 50
 - The Empirical Rule 55
- 1.7 Caution, Care with Numbers, and Ethics 56
 - Ethics 57

Interview with Melissa Anderson 67

CHAPTER 2 What Are the Chances? 70

- 2.1 Introduction 72
- 2.2 Basic Definitions: Events, Sample Space, and Probabilities 75
- 2.3 Basic Rules for Probability 81
 - The Range of Values 81
 - The Rule of Complements 82
 - The Rule of Unions 83
 - Mutually Exclusive Events 83

- 2.4 Conditional Probability 87
- 2.5 Independence of Events 92
 - Product Rules for Independent Events 93
- 2.6 Bayes' Theorem 97

Interview with Alan Dershowitz 107

CHAPTER 3 Chance Quantities 112

- 3.1 Random Variables 115
 - Discrete and Continuous Random Variables 120
- 3.2 Expected Values of Discrete Random Variables 124
 - The Variance and the Standard Deviation of a Random Variable 127
 - Some Properties of Means and Variances 129
- 3.3 The Binomial Distribution 130
 - The Binomial Table 132
 - Mean, Variance, and Shape of the Binomial Distribution 135
 - Sampling with or without Replacement 137
- 3.4 Continuous Random Variables 140

Interview with Arnold Zellner 151

CHAPTER 4 The Bell-Shaped Curve 154

- 4.1 The Normal Probability Distribution 156
 - The Shape of the Normal Distribution 158
 - Notation 159
- 4.2 The Standard Normal Distribution 159
 - Finding Probabilities of the Standard Normal Distribution 160
 - Finding Values of Z Given a Probability 164

- 4.3 The Transformation of Normal Random Variables 167
 - Using the Normal Transformation 169
- 4.4 The Relationship between X and Z and the Use of the Inverse Transformation 173
- 4.5 The Normal Distribution as an Approximation to Other Probability Distributions 176
- 4.6 Normal Data 178

Interview with Bud Goode 185

CHAPTER 5 Let's Take a Sample 190

- 5.1 Statistical Inference 192
 - Sampling Methods and Statistical Inference 196
- 5.2 Sampling Distributions 199
- 5.3 The Sampling Distribution of the Sample Mean 200
- 5.4 Confidence Intervals 205
 - Confidence Interval for the Population Mean When the Population Standard Deviation Is Known 205
- 5.5 Confidence Intervals for μ When σ Is Unknown—the t Distribution 214
 - The t Distribution 214
- 5.6 The Sampling Distribution of the Sample Proportion 222
 - Large-Sample Confidence Intervals for the Population Proportion 223
- 5.7 Sample Size Determination 228
- 5.8 Epilogue: The Polls 234

Interview with Bradley Efron 239

CHAPTER 6 Trial by Probability 244

- 6.1 Introduction 247
- 6.2 Statistical Hypothesis Testing 257
- 6.3 Tests about the Population Mean, μ , When the Population Standard Deviation, σ , Is Known 257
 - Standardizing the Test 257
 - One-Tailed Tests 260
 - The p -Value 265
- 6.4 Tests about the Population mean, μ , When the Population Standard Deviation, σ , Is Unknown 274
- 6.5 Large-Sample Tests for the Population Proportion 284

- 6.6 How Hypothesis Testing Works 291
 - Some Comments and Caveats 294
- 6.7 The Probability of a Type II Error and the Power of the Test 297

Interview with Lee Wilkinson 309

CHAPTER 7 Making Comparisons 314

- 7.1 Introduction 316
- 7.2 Paired-Observations Comparisons 316
 - Experimental Design 316
 - Confidence Intervals 318
- 7.3 A Test for the Difference between Two Population Means, Using Independent Random Samples 322
 - Confidence Intervals 325
- 7.4 A Test for the Difference between Two Population Means, Assuming Equal Population Variances 330
 - Confidence Intervals 333
- 7.5 A Large-Sample Test for the Difference between Two Population Proportions 336
 - Confidence Intervals 338

Interview with Toni Falvo 351

CHAPTER 8 Making Extended Comparisons 356

- 8.1 Introduction 358
- 8.2 The Hypothesis Test of Analysis of Variance (ANOVA) 359
 - The Test Statistic and the F Distribution 361
- 8.3 The Theory and the Computations of ANOVA 365
 - The Sum of Squares Principle 369
 - The Degrees of Freedom 372
 - The Mean Squares 373
 - The Expected Values of the Statistics MSTR and MSE under the Null Hypothesis 373
 - The F Statistic 374
 - The ANOVA Table 374
 - More on Experimental Design 379
 - Randomized Complete Block Design 379
- 8.4 Contingency Table Analysis: A Chi-Square Test for Independence 384
 - The Chi-Square Distribution 386
 - A Chi-Square Test for Equality of Proportions 389

Interview with Nancy Kirkendall 401

CHAPTER 9 Is There a Relationship? 406

- 9.1 Introduction 408
- 9.2 Statistical Models 409
- 9.3 The Simple Linear Regression Model 410
- 9.4 Estimation: The Method of Least Squares 414
- 9.5 Error Variance and the Standard Errors of Regression Estimators 421
 - Confidence Intervals for the Regression Parameters 424
 - A Hypothesis Test for the Existence of a Linear Regression Relationship 425
- 9.6 How Good Is the Regression? 431
 - Analysis of Variance Table and an F Test of the Regression Model 435
- 9.7 Residual Analysis and Checking for Model Inadequacies 437
 - A Check for the Equality of Variance of the Errors 438
 - Testing for Missing Variables 438
 - Detecting a Curvilinear Relationship between Y and X 439
 - Detecting Deviations from the Normal Distribution Assumption 439
- 9.8 Use of the Regression Model for Prediction 444
 - Point Predictions 444
 - Prediction Intervals 445
 - A Confidence Interval for the Average Y , Given a Particular Value of X 446
- 9.9 Correlation 448

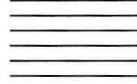
Interview with Mary Guinan 461

CHAPTER 10 The Quest for Quality 464

- 10.1 Control Charts 466
- 10.2 The \bar{x} Chart 468
- 10.3 The R Chart and the s Chart 472
 - The R Chart 472
 - The s Chart 473
- 10.4 The p Chart 474
- 10.4 The c Chart 476

Interview with Phil Crosby 481

GLOSSARY 483**BIBLIOGRAPHY 487****APPENDIX A: DATA TABLES 489****APPENDIX B: STATISTICAL TABLES 493****APPENDIX C: ANSWERS TO MOST ODD-NUMBERED PROBLEMS 519****INDEX 531**



Statistics

Concepts and Applications

CHAPTER

1

Information Everywhere

The Phaistos Disk



CHAPTER OUTLINE

1.1 Introduction	4	1.5 Measures of Centrality	45
1.2 Methods of Displaying Data	9	1.6 Measures of Variability	50
1.3 Percentiles, Quartiles, and the Median	35	1.7 Caution, Care with Numbers, and Ethics	52
1.4 Box Plots	42		



f the several archaeological studies begun in Crete in the 1800s, the most famous is Arthur Evans's excavation of King Minos's palace at Knossos. But farther south on Crete lies the site of the 3,500-year-old palace of Phaestos. Many artifacts were recovered from the site, and archaeologists were able to understand the purpose of most of the objects they discovered. Then they discovered the Phaestos disk. The disk is like no other object ever seen. It is a flat, light round disk of clay about 7 inches in diameter. Hieroglyphic-like characters, similar to the ones seen in Egypt, are impressed on it in a spiral. The disk was a compact and unusual device for storing and displaying large amounts of information in the ancient Minoan civilization. Yet the Phaestos disk remains undeciphered.

In this chapter we discuss information: its collection, display, and summarization. These are the first steps of any statistical analysis.

