An Introduction to Mathematical Statistics and Its Applications

Third Edition

Richard J. Larsen · Morris L. Marx

An Introduction to Mathematical Statistics and Its Applications

Third Edition

Richard J. Larsen
Vanderbilt University

Morris L. Marx
University of West Florida

Prentice Hall

PRENTICE HALL, Upper Saddle River, NJ 07458

Library of Congress Cataloging-in-Publication Data

Larsen, Richard J.
An Introduction to mathematical statistics and its applications / Richard J. Larsen,
Morris L. Marx,—3rd ed.
p. cm.
Includes bibliographical references and index
ISBN 0-13-922303-7
1. Mathematical statistics I. Marx, Morris L. II. Title.
QA276.L314 2001
519.5–dc.21 00-025176

Acquisitions Editor: *Kathleen Boothby Sestak*
Assistant Vice President of Production and Manufacturing: *David W. Riccardi*
Executive Managing Editor: *Kathleen Schiaparelli*
Senior Managing Editor: *Linda Mihatov Behrens*
Production Editor: *Betsy Williams*
Manufacturing Buyer: *Alan Fischer*
Manufacturing Manager: *Trudy Pisciotti*
Marketing Manager: *Angela Battle*
Marketing Assistant: *Vince Jansen*
Director of Marketing: *John Tweeddale*
Editorial Assistant/Supplements Editor: *Joanne Wendelken*
Art Director: *Jayne Conte*
PrePress: *Preparé Inc. / Emilcomp srl*

Prentice Hall

Upper Saddle River, NJ 07458

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

ISBN 0-13-922303-7

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall of Canada, Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Pearson Education Asia Pte. Ltd.
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*

Preface

Changes in this third edition have been primarily motivated by our own teaching experiences as well as by the comments of others who use the text. Technology, though, has also dictated certain revisions. The widespread use of statistical software packages has brought certain topics and concepts to the fore, while diminishing the relevance of others. All in all, we feel that this new edition has a sharper focus and that students will find it more accessible and easier to use.

Many of the major changes come in the middle third of the book, much of which has been rewritten. These are the chapters that make the critical transition from probability to statistics. We have taken a variety of steps to make that material come more alive, ranging from the addition of more helpful examples to the frequent use of computer simulations.

Chapter 4, for example, now addresses more fully the important question of *why* certain measurements are modeled by particular probability functions. Relationships that exist between pdfs are given more attention, and the connection between theoretical models and sample data is explored in greater depth. Chapter 5 has been restructured. In the new edition, methods of estimation come first and the underlying theory is taken up last. That arrangement makes it easier for instructors to adjust the amount of time spent on estimation to whatever suits their individual needs. In Chapter 6, the principles of decision-making are now introduced in the context of testing H_0: $\mu = \mu_0$ rather than H_0: $p = p_0$. The result is a more steamlined presentation that avoids the complications inherent in a test statistic whose pdf is discrete.

Positioned between Chapter 7, which deals with the normal distribution, and Chapters 9 through 14, where the various techniques for analyzing data are introduced, is a new chapter on experimental design. Chapter 8 profiles seven of the most frequently encountered "data models." The basic characteristics of each design are discussed as well as the types of questions each seeks to answer. By providing a framework and a theme, Chapter 8 brings cohesion and a sense of order to the chapters that follow.

Chapter 11 (*Regression*) has also been changed substantially. It now begins with curve-fitting, then introduces the linear model, and eventually concludes with the bivariate normal. Regression "diagnostics" have been added to the new edition, and the various inference procedures associated with the linear model have been explained and delineated more carefully.

Our overriding motivation in deciding which topics to present—and in what order—stem from our objective to write a book that emphasizes the interrelation between probability theory, mathematical statistics, and data analysis. We believe that integrating all three is vitally important, particularly for those students who take only one statistics course during their college careers. Our experience in the classroom has certainly strengthened our faith in this approach: Students can more clearly see the importance of each of the three when viewed in the context of the other two.

Pedagogical Enhancements

Other changes have been implemented throughout the book as well. New case studies and examples have been added; others have been updated, revised, or replaced. The number of exercises has been substantially expanded, a 50% increase in some sections. Many chapters have a "MINITAB Applications" Appendix. Included is the syntax for doing whatever procedures appear in that chapter, along with a discussion of the output. Answers to most odd-numbered exercises are given at the end of the book.

Supplements

Instructor's Solutions Manual. This resource contains worked-out solutions to all text exercises (0-13-922311-8).

Student Solutions Manual: Featuring complete solutions to odd-numbered exercises, this is a great tool for students as they study and work through the problem material (0-13-031015-8).

Acknowledgments

We would like to express our indebtedness to the editors at J.R. Geigy, Biometrika, and McGraw-Hill for letting us use tables that appear in the Appendix and to the many researchers whose data we have used for examples. And for their detailed comments, criticisms, and suggestions, we thank the reviewers:

Dane Wu, Pacific Lutheran University; Pantelis K. Vlachos, Carnegie Mellon University; Lee K. Jones, University of Massachusetts–Lowell; Daniel W. Chambers, Boston College; Josephine S. Hamer, Western Connecticut State University; Raisa E. Feldman, University of California–Santa Barbara.

Finally, we would like to convey our gratitude to our editor, Kathy Boothby Sestak, and production editor, Betsy Williams. Their counsel on matters of design, form, and style was sincerely appreciated.

Richard J. Larsen
Nashville, Tennessee

Morris L. Marx
Pensacola, Florida

Contents

Preface xi

1 INTRODUCTION **1**

1.1 A Brief History 2
1.2 Some Examples 11
1.3 A Chapter Summary 18

2 PROBABILITY **20**

2.1 Introduction 21
2.2 Sample Spaces and the Algebra of Sets 21
2.3 The Probability Function 29
2.4 Discrete Probability Functions 36
2.5 Continuous Probability Functions 42
2.6 Conditional Probability 51
2.7 Independence 70
2.8 Repeated Independent Trials 79
2.9 Combinatorics 86
2.10 Combinatorial Probability 109

3 RANDOM VARIABLES **119**

3.1 Introduction 120
3.2 The Probability Density Function 121
3.3 The Hypergeometric and Binomial Distributions 128

3.4 The Cumulative Distribution Function 144
3.5 Joint Densities 153
3.6 Independent Random Variables 167
3.7 Combining and Transforming Random Variables 170
3.8 Order Statistics 180
3.9 Conditional Densities 185
3.10 Expected Values 190
3.11 Properties of Expected Values 202
3.12 The Variance 218
3.13 Properties of Variances 222
3.14 Chebyshev's Inequality 229
3.15 Higher Moments 232
3.16 Moment-Generating Functions 236
Appendix 3.A.1 MINITAB Applications 243

4 SPECIAL DISTRIBUTIONS 245

4.1 Introduction 246
4.2 The Poisson Distribution 246
4.3 The Normal Distribution 263
4.4 The Geometric Distribution 286
4.5 The Negative Binomial Distribution 292
4.6 The Gamma Distribution 296
Appendix 4.A.1 MINITAB Applications 301
Appendix 4.A.2 A Proof of the Central Limit Theorem 305

5 ESTIMATION 307

5.1 Introduction 308
5.2 Estimating Parameters: the Method of Maximum Likelihood and the Method of Moments 310
5.3 Interval Estimation 323
5.4 Properties of Estimators 337
5.5 Minimum-Variance Estimators: The Cramer-Rao Lower Bound 350
5.6 Sufficiency 353
5.7 Consistency 358
Appendix 5.A.1 MINITAB Applications 361

6 HYPOTHESIS TESTING 364

6.1 Introduction 365
6.2 The Decision Rule 365
6.3 Testing Binomial Data—H_0: $p = p_0$ 375
6.4 Type I and Type II Errors 382
6.5 A Notion of Optimality: the Generalized Likelihood Ratio 397

7 THE NORMAL DISTRIBUTION 401

7.1 Introduction 402
7.2 Point Estimates for μ and σ^2 402
7.3 The χ^2 Distribution; Inferences about σ^2 410
7.4 The F and t Distributions 421
7.5 Drawing Inferences about μ 430
Appendix 7.A.1 MINITAB Applications 451
Appendix 7.A.2 Some Distribution Results for $\bar{Y}$ and S^2 455
Appendix 7.A.3 A Proof of Theorem 7.3.5 457
Appendix 7.A.4 A Proof that the One-Sample t Test is a GLRT 459

8 TYPES OF DATA: A BRIEF OVERVIEW 462

8.1 Introduction 463
8.2 Classifying Data 468

9 TWO-SAMPLE PROBLEMS 484

9.1 Introduction 485
9.2 Testing H_0: $\mu_X = \mu_Y$—the Two-Sample t Test 486
9.3 Testing H_0: $\sigma_X^2 = \sigma_Y^2$—the F Test 497
9.4 Binomial Data: Testing H_0: $p_X = p_Y$ 505
9.5 Confidence Intervals for the Two-Sample Problem 510
Appendix 9.A.1 A Derivation of the Two-Sample t Test (A Proof of Theorem 9.2.2) 518
Appendix 9.A.2 Power Calculations for a Two-Sample t Test 520
Appendix 9.A.3 MINITAB Applications 523

10 **GOODNESS-OF-FIT TESTS** **525**

10.1 Introduction 526
10.2 The Multinomial Distribution 526
10.3 Goodness-of-Fit Tests: All Parameters Known 532
10.4 Goodness-of-Fit Tests: Parameters Unknown 540
10.5 Contingency Tables 548
Appendix 10.A.1 MINITAB Applications 556

11 **REGRESSION** **558**

11.1 Introduction 559
11.2 The Method of Least Squares 559
11.3 The Linear Model 584
11.4 Covariance and Correlation 606
11.5 The Bivariate Normal Distribution 620
Appendix 11.A.1 MINITAB Applications 627
Appendix 11.A.2 A Proof of Theorem 11.3.3 629

12 **THE ANALYSIS OF VARIANCE** **633**

12.1 Introduction 634
12.2 The F Test 636
12.3 Multiple Comparisons: Tukey's Method 646
12.4 Testing Subhypotheses with Orthogonal Contrasts 650
12.5 Data Transformations 656
Appendix 12.A.1 MINITAB Applications 659
Appendix 12.A.2 A Proof of Theorem 12.2.2 662
Appendix 12.A.3 The Distribution of $\dfrac{SSTR/(k-1)}{SSE/(n-k)}$ When H_1 is True 662

13 **RANDOMIZED BLOCK DESIGNS** **667**

13.1 Introduction 668
13.2 The F Test for a Randomized Block Design 669
13.3 The Paired t Test 682
Appendix 13.A.1 MINITAB Applications 685

14 NONPARAMETRIC STATISTICS 687

14.1 Introduction 688
14.2 The Sign Test 689
14.3 The Wilcoxon Signed Rank Test 692
14.4 The Kruskal-Wallis Test 703
14.5 The Friedman Test 707
Appendix 14.A.1 MINITAB Applications 710

Appendix: Statistical Tables 715

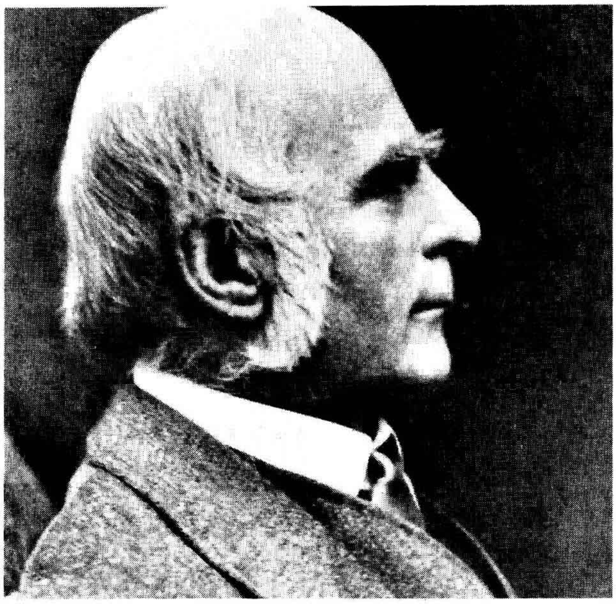
Answers to Selected Odd-Numbered Questions 742

Bibliography 776

Index 785

CHAPTER 1

Introduction



Francis Galton

"Some people hate the very name of statistics, but I find them full of beauty and interest. Whenever they are not brutalized, but delicately handled by the higher methods, and are warily interpreted, their power of dealing with complicated phenomena is extraordinary. They are the only tools by which an opening can be cut through the formidable thicket of difficulties that bars the path of those who pursue the Science of man."

1.1 A BRIEF HISTORY

Statistics is the science of sampling. How one set of measurements differs from another and what the implications of those differences might be are its primary concerns. Conceptually, the subject is rooted in the mathematics of probability, but its applications are everywhere. Statisticians are as likely to be found in a research lab or a field station as they are in a government office, an advertising firm, or a college classroom.

Properly applied, statistical techniques can be enormously effective in clarifying and quantifying natural phenomena. Figure 1.1.1 illustrates a case in point. Pictured at the top is a facsimile of the kind of data routinely recorded by a seismograph—listed chronologically are the occurrence times and Richter magnitudes for a series of earthquakes. Viewed in that format, the numbers are largely meaningless: No patterns are evident, nor is there any obvious connection between the frequencies of tremors and their severities.

Episode number	Date	Time	Severity (Richter scale)
⋮	⋮	⋮	⋮
217	6/19	4:53 P.M.	2.7
218	7/2	6:07 A.M.	3.1
219	7/4	8:19 A.M.	2.0
220	8/7	1:10 A.M.	4.1
221	8/7	10:46 P.M.	3.6
⋮	⋮	⋮	⋮

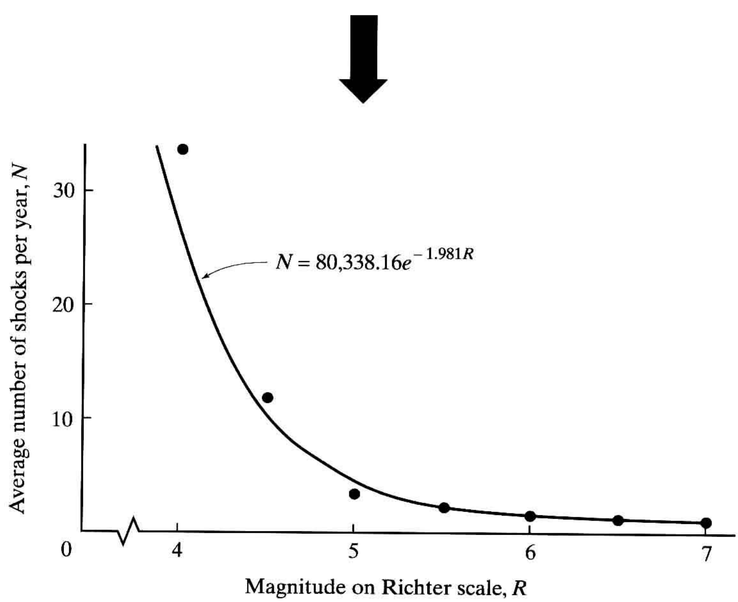


FIGURE 1.1.1

By way of contrast, the bottom of Figure 1.1.1 shows a statistical summary (using some of the regression techniques we will learn later) of a set of seismograph data recorded in southern California (59). Plotted above the Richter (R) value of 4.0, for example, is the average number (N) of earthquakes occurring per year in that region having magnitudes in the range 3.75 to 4.25. Similar points are included for R-values centered at 4.5, 5.0, 5.5, 6.0, 6.5, and 7.0. Now we can see that the two variables *are* related: Describing the (N, R)'s exceptionally well is the equation $N = 80{,}338.16e^{-1.981R}$.

In general, statistical techniques are employed either to (1) describe what *did* happen or (2) predict what *might* happen. The graph at the bottom of Figure 1.1.1 does both. Having "fit" the model $N = \beta_0 e^{-\beta_1 R}$ to the observed set of minor tremors (and finding that $\beta_0 = 80{,}338.16$ and $\beta_1 = -1.981$), we can then use that same equation to predict the likelihood of events *not* represented in the data set. If $R = 8.0$, for example, we would expect N to equal 0.01:

$$N = 80{,}338.16e^{-1.981(8.0)}$$

$$= 0.01$$

(which implies that Californians can expect catastrophic earthquakes registering on the order of 8.0 on the Richter scale to occur, on the average, once every 100 years).

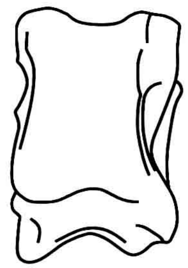
It is unarguably true that the interplay between description and prediction—similar to what we see in Figure 1.1.1—is the single most important theme in statistics. Additional examples highlighting other aspects of that connection will be discussed in Section 1.2. To set the stage for the rest of the text, though, we will conclude Section 1.1 with brief histories of probability and statistics. Both are interesting stories, replete with large casts of unusual characters and plots that have more than a few unexpected twists and turns.

Probability: The Early Years (Optional)

No one knows where or when the notion of chance first arose; it fades into our prehistory. Nevertheless, evidence linking early humans with devices for generating random events is plentiful: Archaeological digs, for example, throughout the ancient world consistently turn up a curious overabundance of *astragali*, the heel bones of sheep and other vertebrates. Why should the frequencies of these bones be so disproportionately high? One could hypothesize that our forebearers were fanatical foot fetishists, but two other explanations seem more plausible: The bones were used for religious ceremonies *and for gambling*.

Astragali have six sides but are not symmetrical (see Figure 1.1.2). Those found in excavations typically have their sides numbered or engraved. For many ancient civilizations, astragali were the primary mechanism through which oracles solicited the opinions of their gods. In Asia Minor, for example, it was customary in divination rites to roll, or *cast*, five astragali. Each possible configuration was associated with the name of a god and carried with it the sought-after advice. An outcome of

FIGURE 1.1.2



Sheep astragalus

(1, 3, 3, 4, 4), for instance, was said to be the throw of the savior Zeus, and its appearance was taken as a sign of encouragement (34):

> One one, two threes, two fours
> The deed which thou meditatest, go do it boldly.
> Put thy hand to it. The gods have given thee
> favorable omens
> Shrink not from it in thy mind, for no evil
> shall befall thee.

A (4, 4, 4, 6, 6), on the other hand, the throw of the child-eating Cronos, would send everyone scurrying for cover:

> Three fours and two sixes. God speaks as follows.
> Abide in thy house, nor go elsewhere,
> Lest a ravening and destroying beast come nigh thee.
> For I see not that this business is safe. But bide
> thy time.

Gradually, over thousands of years, astragali were replaced by dice, and the latter became the most common means for generating random events. Pottery dice have been found in Egyptian tombs built before 2000 B.C.; by the time the Greek civilization was in full flower, dice were everywhere. (*Loaded* dice have also been found. Mastering the mathematics of probability would prove to be a formidable task for our ancestors, but they quickly learned how to cheat!)

The lack of historical records blurs the distinction initially drawn between divination ceremonies and recreational gaming. Among more recent societies, though, gambling emerged as a distinct entity, and its popularity was irrefutable. The Greeks and Romans were consummate gamblers, as were the early Christians (82).

Rules for many of the Greek and Roman games have been lost, but we can recognize the lineage of certain modern diversions in what was played during the Middle Ages. The most popular dice game of that period was called *hazard*, the name deriving from the Arabic *al zhar*, which means "a die." Hazard is thought to have been brought to Europe by soldiers returning from the Crusades; its rules are much like those of our modern-day craps. Cards were first introduced in the fourteenth century and immediately gave rise to a game known as *Primero*, an early form of poker. Board games, such as backgammon, were also popular during this period.

Given this rich tapestry of games and the obsession with gambling that characterized so much of the Western world, it may seem more than a little puzzling that a formal study of probability was not undertaken sooner than it was. As we will see shortly, the first instance of anyone *conceptualizing* probability, in terms of a mathematical model, occurred in the sixteenth century. That means that more than 2000 years of dice games, card games, and board games passed by before someone finally had the insight to write down even the simplest of probabilistic abstractions.

Historians generally agree that, as a subject, probability got off to a rocky start because of its incompatibility with two of the most dominant forces in the evolution of our Western culture, Greek philosophy and early Christian theology. The Greeks were comfortable with the notion of chance (something the Christians were not), but it went against their nature to suppose that random events could be quantified in any useful fashion. They believed that any attempt to reconcile mathematically what *did* happen with what *should have* happened was, in their phraseology, an improper juxtaposition of the "earthly plane" with the "heavenly plane."

Making matters worse was the antiempiricism that permeated Greek thinking. Knowledge, to them, was not something that should be derived by experimentation. It was better to reason out a question logically than to search for its explanation in a set of numerical observations. Together, these two attitudes had a deadening effect: The Greeks had no motivation to think about probability in any abstract sense, nor were they faced with the problems of interpreting data that might have pointed them in the direction of a probability calculus.

If the prospects for the study of probability were dim under the Greeks, they became even worse when Christianity broadened its sphere of influence. The Greeks and Romans at least accepted the *existence* of chance. They believed their gods to be either unable or unwilling to get involved in matters so mundane as the outcome of the roll of a die. Cicero writes:

> Nothing is so uncertain as a cast of dice, and yet there is no one who plays often who does not make a Venus-throw[1] and occasionally twice and thrice in succession. Then are we, like fools, to prefer to say that it happened by the direction of Venus rather than by chance?

For the early Christians, though, there was no such thing as chance: Every event that happened, no matter how trivial, was perceived to be a direct manifestation of God's deliberate intervention. In the words of St. Augustine:

> Nos eas causas quae dicuntur fortuitae ... non dicimus
> nullas, sed latentes; easque tribuimus vel veri Dei ...
> (We say that those causes that are said to be by chance
> are not non-existent but are hidden, and we attribute
> them to the will of the true God ...)

Taking Augustine's position makes the study of probability moot, and it makes a probabilist a heretic. Not surprisingly, nothing of significance was accomplished in the subject for the next fifteen hundred years.

---

[1] When rolling four astragali, each of which is numbered on *four* sides, a Venus-throw was having each of the four numbers appear.

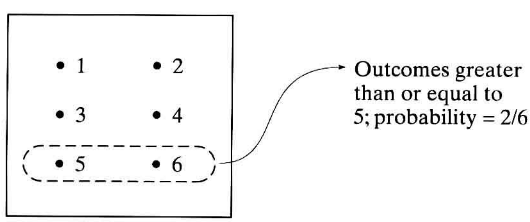
It was in the sixteenth century that probability, like a mathematical Lazarus, arose from the dead. Orchestrating its resurrection was one of the most eccentric figures in the entire history of mathematics, Gerolamo Cardano. By his own admission, Cardano personified the worst and the best—the Jekyll and the Hyde—of the Renaissance man. He was born in 1501 in Pavia. Facts about his personal life are difficult to verify. He wrote an autobiography, but his penchant for lying raises doubts about much of what he says. Whether true or not, though, his "one-sentence" self-assessment paints an interesting portrait (117):

> Nature has made me capable in all manual work, it has given me the spirit of a philosopher and ability in the sciences, taste and good manners, voluptuousness, gaiety, it has made me pious, faithful, fond of wisdom, meditative, inventive, courageous, fond of learning and teaching, eager to equal the best, to discover new things and make independent progress, of modest character, a student of medicine, interested in curiosities and discoveries, cunning, crafty, sarcastic, an initiate in the mysterious lore, industrious, diligent, ingenious, living only from day to day, impertinent, contemptuous of religion, grudging, envious, sad, treacherous, magician and sorcerer, miserable, hateful, lascivious, obscene, lying, obsequious, fond of the prattle of old men, changeable, irresolute, indecent, fond of women, quarrelsome, and because of the conflicts between my nature and soul I am not understood even by those with whom I associate most frequently.

Formally trained in medicine, Cardano's interest in probability derived from his addiction to gambling. His love of dice and cards was so all-consuming that he is said to have once sold all his wife's possessions just to get table stakes! Fortunately, something positive came out of Cardano's obsession. He began looking for a mathematical model that would describe, in some abstract way, the outcome of a random event. What he eventually formalized is now called the *classical definition of probability*: If the total number of possible outcomes, all equally likely, associated with some action is n and if m of those n result in the occurrence of some given event, then the probability of that event is m/n. If a fair die is rolled, there are $n = 6$ possible outcomes. If the event "outcome is greater than or equal to 5" is the one in which we are interested, then $m = 2$ (the outcomes 5 and 6) and the probability of the event is $2/6$, or $1/3$ (see Figure 1.1.3).

Cardano had tapped into the most basic principle in probability. The model he discovered may seem trivial in retrospect, but it represented a giant step forward: His was the first recorded instance of anyone computing a *theoretical*, as opposed to an empirical, probability. Still, the actual impact of Cardano's work was minimal. He

FIGURE 1.1.3



Possible outcomes

wrote a book in 1525, but its publication was delayed until 1663. By then, the focus of the Renaissance, as well as interest in probability, had shifted from Italy to France.

The date cited by many historians (those who are not Cardano supporters) as the "beginning" of probability is 1654. In Paris a well-to-do gambler, the Chevalier de Mere, asked several prominent mathematicians, including Blaise Pascal, a series of questions, the best-known of which was the *problem of points*:

> Two people, A and B, agree to play a series of fair games until one person has won six games. They each have wagered the same amount of money, the intention being that the winner will be awarded the entire pot. But suppose, for whatever reason, the series is prematurely terminated, at which point A has won five games and B three. How should the stakes be divided?

[The correct answer is that A should receive seven-eighths of the total amount wagered. (Hint: Suppose the contest were resumed. What scenarios would lead to A's being the first person to win six games?)]

Pascal was intrigued by de Mere's questions and shared his thoughts with Pierre Fermat, a Toulouse civil servant and probably the most brilliant mathematician in Europe. Fermat graciously replied, and from the now famous Pascal-Fermat correspondence came not only the solution to the problem of points but the foundation for more general results. More significantly, news of what Pascal and Fermat were working on spread quickly. Others got involved, of whom the best known was the Dutch scientist and mathematician Christiaan Huygens. The delays and the indifference that plagued Cardano a century earlier were not going to happen again.

Best remembered for his work in optics and astronomy, Huygens, early in his career, was intrigued by the problem of points. In 1657 he published *De Ratiociniis in Aleae Ludo* (Calculations in Games of Chance), a very significant work, far more comprehensive than anything Pascal and Fermat had done. For almost 50 years it was the standard "textbook" in the theory of probability. Huygens, of course, has supporters who feel that *he* should be credited as the founder of probability.

Almost all the mathematics of probability was still waiting to be discovered. What Huygens wrote was only the humblest of beginnings, a set of 14 propositions bearing little resemblance to the topics we teach today. But the foundation was there. The mathematics of probability was finally on firm footing.

Statistics: From Aristotle to Quetelet (Optional)

Historians generally agree that the subject of statistics began to take definite shape in the middle of the nineteenth century. What triggered its emergence was the union of three different "sciences," each of which had been developing along more or less independent lines (184).

The first of these sciences, what the Germans called *Staatenkunde*, involved the collection of comparative information on the history, resources, and military prowess of nations. Although efforts in this direction peaked in the seventeenth and eighteenth centuries, the concept was hardly new: Aristotle had done something similar in the fourth century B.C. Of the three movements, this one had the least influence on