PROBABILITY and STOCHASTIC PROCESSES

A FRIENDLY INTRODUCTION FOR ELECTRICAL AND COMPUTER ENGINEERS

THIRD EDITION

ROY D. YATES • DAVID J. GOODMAN

WILEY

Probability and Stochastic Processes

Features of this Text

Who will benefit from using this text?

This text can be used in Junior or Senior level courses in probability and stochastic processes. The mathematical exposition will appeal to students and practitioners in many areas. The examples, quizzes, and problems are typical of those encountered by practicing electrical and computer engineers. Professionals in the telecommunications and wireless industry will find it particularly useful.

What's New?

This text has been expanded with new introductory material:

- Over 160 new homework problems
- New chapters on *Sequential Trials*, *Derived Random Variables* and *Conditional Probability Models*.
- MATLAB examples and problems give students hands-on access to theory and applications. Every chapter includes guidance on how to use MATLAB to perform calculations and simulations relevant to the subject of the chapter.
- Advanced material online in *Signal Processing* and *Markov Chains* supplements.

Notable Features

The Friendly Approach
The friendly and accessible writing style gives students an intuitive feeling for the formal mathematics.

Quizzes and Homework Problems
An extensive collection of in-chapter quizzes provides checkpoints for readers to gauge their understanding. Hundreds of end-of-chapter problems are clearly marked as to their degree of difficulty from beginner to expert.

Student Companion Website `www.wiley.com/college/yates`
Available for download: All MATLAB m-files in the text, the *Quiz Solutions Manual*, a *Student Solutions Manual*, the *Signal Processing Supplement*, and the *Markov Chains Supplement*.

Instructor Support

Instructors can register for the Instructor Companion Site at `www.wiley.com/college/yates`

Probability and Stochastic Processes

A Friendly Introduction for Electrical and Computer Engineers

Third Edition

Roy D. Yates
Rutgers, The State University of New Jersey

David J. Goodman
New York University

WILEY

V.P. & Executive Publisher	Don Fowley
Executive Editor	Dan Sayre
Sponsoring Editor	Mary O'Sullivan
Project Editor	Ellen Keohane
Production Editor	Eugenia Lee
Cover Designer	Samantha Low

This book was set in Computer Modern by the authors using LATEX and printed and bound by RRDonnelley. The cover was printed by RRDonnelley.

About the cover: The cover shows a circumhorizontal arc. As noted in Wikipedia, this is an ice-halo formed by plate-shaped ice crystals in high level cirrus clouds. The misleading term "fire rainbow" is sometimes used to describe this rare phenomenon, although it is neither a rainbow, nor related in any way to fire.

This book is printed on acid-free paper. ♾

ISBN 978-1-118-32456-1

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To Alissa, Brett, Daniel, Hannah, Leila, Milo, Theresa, Tony, and Zach

Preface

Welcome to the third edition

You are reading the third edition of our textbook. Although the fundamentals of probability and stochastic processes have not changed since we wrote the first edition, the world inside and outside universities is different now than it was in 1998. Outside of academia, applications of probability theory have expanded enormously in the past 16 years. Think of the 20 billion+ Web searches each month and the billions of daily computerized stock exchange transactions, each based on probability models, many of them devised by electrical and computer engineers.

Universities and secondary schools, recognizing the fundamental importance of probability theory to a wide range of subject areas, are offering courses in the subject to younger students than the ones who studied probability 16 years ago. At Rutgers, probability is now a required course for Electrical and Computer Engineering sophomores.

We have responded in several ways to these changes and to the suggestions of students and instructors who used the earlier editions. The first and second editions contain material found in postgraduate as well as advanced undergraduate courses. By contrast, the printed and e-book versions of this third edition focus on the needs of undergraduates studying probability for the first time. The more advanced material in the earlier editions, covering random signal processing and Markov chains, is available at the companion website (`www.wiley.com/college/yates`). To promote intuition into the practical applications of the mathematics, we have expanded the number of examples and quizzes and homework problems to about 600, an increase of about 35 percent compared to the second edition. Many of the examples are mathematical exercises. Others are questions that are simple versions of the ones encountered by professionals working on practical applications.

How the book is organized

Motivated by our teaching experience, we have rearranged the sequence in which we present the elementary material on probability models, counting methods, conditional probability models, and derived random variables. In this edition, the first chapter covers fundamentals, including axioms and probability of events, and the second chapter covers counting methods and sequential experiments. As before, we introduce discrete random variables and continuous random variables in separate chapters. The subject of Chapter 5 is multiple discrete and continuous random variables. The first and second editions present derived random variables and conditional random variables in the introductions to discrete and continuous random variables. In this third edition, derived random variables and conditional random variables appear in their own chapters, which cover both discrete and continuous random variables.

Chapter 8 introduces random vectors. It extends the material on multiple random variables in Chapter 5 and relies on principles of linear algebra to derive properties of random vectors that are useful in real-world data analysis and simulations. Chapter 12 on estimation relies on the properties of random vectors derived in Chapter 8. Chapters 9 through 12 cover subjects relevant to data analysis including Gaussian approximations based on the central limit theorem, estimates of model parameters, hypothesis testing, and estimation of random variables. Chapter 13 introduces stochastic processes in the context of the probability model that guides the entire book: an experiment consisting of a procedure and observations.

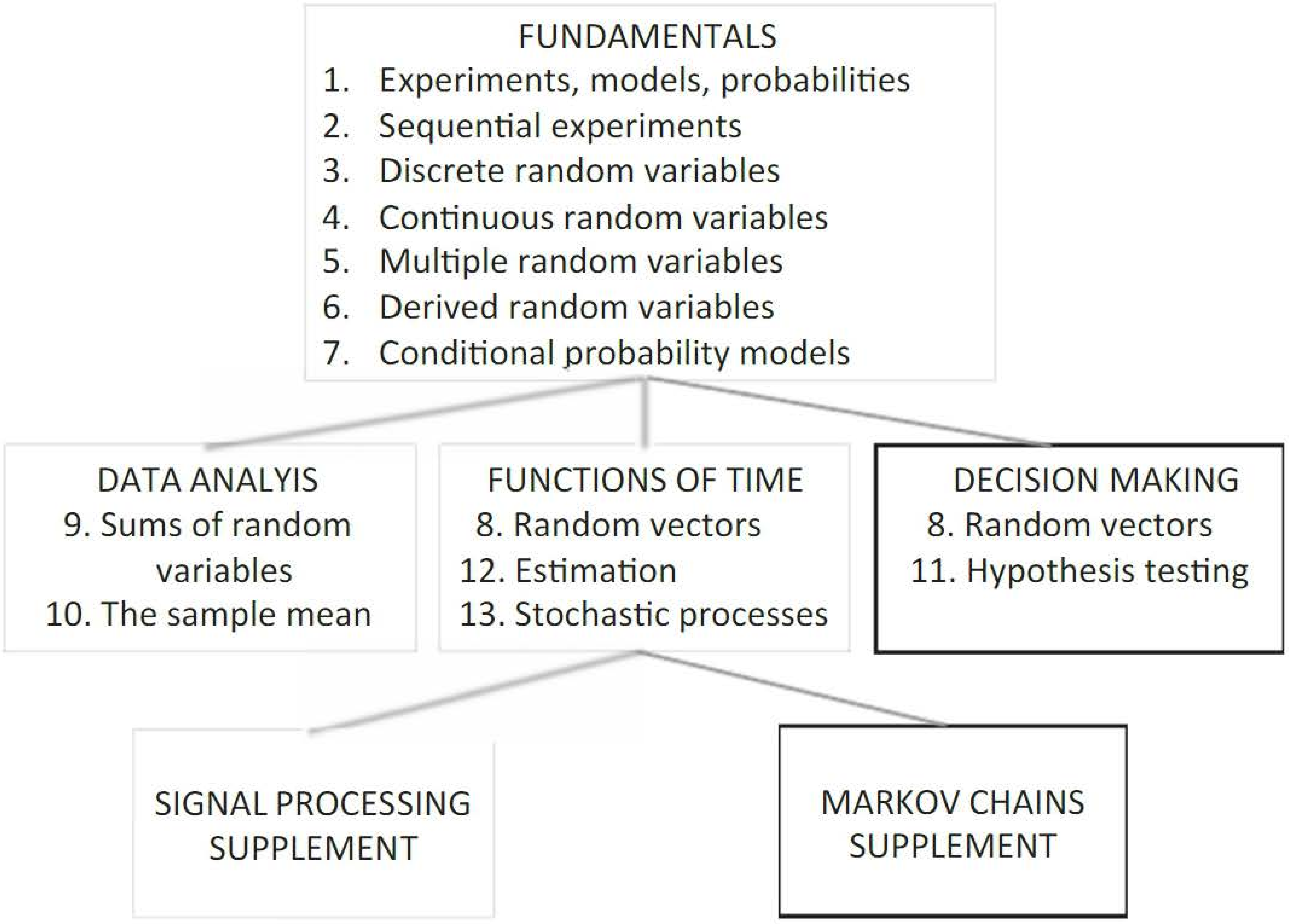
Each of the 92 sections of the 13 chapters ends with a quiz. By working on the quiz and checking the solution at the book's website, students will get quick feedback on how well they have grasped the material in each section.

We think that 60- 80% (7 to 10 chapters) of the book would fit into a one semester undergraduate course for beginning students in probability. We anticipate that all courses will cover the first five chapters, and that instructors will select the remaining course content based on the needs of their students. The "roadmap" on page ix displays the thirteen chapter titles and suggests a few possible undergraduate syllabi.

The Signal Processing Supplement (SPS) and Markov Chains Supplement (MCS) are the final chapters of the third edition. They are now available at the book's website. They contain postgraduate-level material. We, and colleagues at other universities, have used these two chapters in graduate courses that move very quickly through the early chapters to review material already familiar to students and to fill in gaps in learning of diverse postgraduate populations.

What is distinctive about this book?

- The entire text adheres to a single model that begins with an experiment consisting of a procedure and observations.
- The mathematical logic is apparent to readers. Every fact is identified clearly as a definition, an axiom, or a theorem. There is an explanation, in simple English, of the intuition behind every concept when it first appears in the text.
- The mathematics of discrete random variables is introduced separately from the mathematics of continuous random variables.
- Stochastic processes and statistical inference fit comfortably within the unifying model of the text.
- An abundance of exercises puts the theory to use. New ideas are augmented with detailed solutions of numerical examples.
- Each section begins with a brief statement of the important concepts introduced in the section and concludes with a simple quiz to help students gauge their grasp of the new material.



A road map for the text.

- Each problem at the end of a chapter is labeled with a reference to a section in the chapter and a degree of difficulty ranging from "easy" to "experts only." For example Problem 3.4.5 requires material from Section 3.4 but not from later sections. Each problem also has a label that reflects our estimate of degree of difficulty. Skiers will recognize the following symbols:

  ● Easy ■ Moderate ♦ Difficult ♦♦ Experts Only

  Every ski area emphasizes that these designations are relative to the trails at that area. Similarly, the difficulty of our problems is relative to the other problems in this text.

- There is considerable support on the World Wide Web for students and instructors, including MATLAB programs and solutions to the quizzes and problems.

Further Reading

Libraries and bookstores contain an endless collection of textbooks at all levels covering the topics presented in this textbook. We know of two in comic book format [GS93, Pos01]. The reference list on page 489 is a brief sampling of books that can add breadth or depth to the material in this text. Most books on probability, statistics, stochastic processes, and random signal processing contain expositions of the basic principles of probability and random variables, covered in Chapters 1–5. In advanced texts, these expositions serve mainly to establish notation for more specialized topics. [LG11] and [Gub06] share our focus on electrical and computer engineering applications. [BT08], [Ros12] and [Dra67] and introduce the fundamentals of probability and random variables to a general audience of students with a calculus background. [KMT12] is a comprehensive graduate level textbook with a thorough presentation of fundamentals of probability, stochastic processes, and data analysis. It uses the basic theory to develop techniques including hidden Markov models, queuing theory, and machine learning used in many practical applications. [Bil12] is more advanced mathematically; it presents probability as a branch of measure theory. [MR10] and [SMM10] introduce probability theory in the context of data analysis. [Dav10] and [HL11] are beginners' introductions to MATLAB. [Ber98] is in a class by itself. It presents the concepts of probability from a historical perspective, focusing on the lives and contributions of mathematicians and others who stimulated major advances in probability and statistics and their application in various fields including psychology, economics, government policy, and risk management.

Acknowledgments

We are grateful for assistance and suggestions from many sources including our students at Rutgers and New York Universities, instructors who adopted the previous editions, reviewers, and the Wiley team.

At Wiley, we are pleased to acknowledge the encouragement and enthusiasm of our executive editor Daniel Sayre and the support of sponsoring editor Mary O'Sullivan, project editor Ellen Keohane, production editor Eugenia Lee, and cover designer Samantha Low.

We also convey special thanks to Ivan Seskar of WINLAB at Rutgers University for exercising his magic to make the WINLAB computers particularly hospitable to the electronic versions of the book and to the supporting material on the World Wide Web.

The organization and content of the second edition has benefited considerably from the input of many faculty colleagues including Alhussein Abouzeid at Rensselaer Polytechnic Institute, Krishna Arora at Florida State University, Frank Candocia at Florida International University, Robin Carr at Drexel University, Keith Chugg at USC, Charles Doering at University of Michigan, Roger Green at North Dakota State University, Witold Krzymien at University of Alberta, Edl Schamiloglu at University of New Mexico, Arthur David Snider at University of South Florida, Junshan Zhang at Arizona State University, and colleagues Narayan Mandayam, Leo Razumov, Christopher Rose, Predrag Spasojević, and Wade Trappe at Rutgers.

Unique among our teaching assistants, Dave Famolari took the course as an undergraduate. Later as a teaching assistant, he did an excellent job writing homework solutions with a tutorial flavor. Other graduate students who provided valuable feedback and suggestions on the first edition include Ricki Abboudi, Zheng

Cai, Pi-Chun Chen, Sorabh Gupta, Vahe Hagopian, Amar Mahboob, Ivana Maric, David Pandian, Mohammad Saquib, Sennur Ulukus, and Aylin Yener.

The first edition also benefited from reviews and suggestions conveyed to the publisher by D.L. Clark at California State Polytechnic University at Pomona, Mark Clements at Georgia Tech, Gustavo de Veciana at the University of Texas at Austin, Fred Fontaine at Cooper Union, Rob Frohne at Walla Walla College, Chris Genovese at Carnegie Mellon, Simon Haykin at McMaster, and Ratnesh Kumar at the University of Kentucky.

Finally, we acknowledge with respect and gratitude the inspiration and guidance of our teachers and mentors who conveyed to us when we were students the importance and elegance of probability theory. We cite in particular Robert Gallager and the late Alvin Drake of MIT and the late Colin Cherry of Imperial College of Science and Technology.

A Message to Students from the Authors

A lot of students find it hard to do well in this course. We think there are a few reasons for this difficulty. One reason is that some people find the concepts hard to use and understand. Many of them are successful in other courses but find the ideas of probability difficult to grasp. Usually these students recognize that learning probability theory is a struggle, and most of them work hard enough to do well. However, they find themselves putting in more effort than in other courses to achieve similar results.

Other people have the opposite problem. The work looks easy to them, and they understand everything they hear in class and read in the book. There are good reasons for assuming this is easy material. There are very few basic concepts to absorb. The terminology (like the word *probability*), in most cases, contains familiar words. With a few exceptions, the mathematical manipulations are not complex. You can go a long way solving problems with a four-function calculator.

For many people, this apparent simplicity is dangerously misleading because it is very tricky to apply the math to specific problems. A few of you will see things clearly enough to do everything right the first time. However, most people who do well in probability need to practice with a lot of examples to get comfortable with the work and to really understand what the subject is about. Students in this course end up like elementary school children who do well with multiplication tables and long division but bomb out on word problems. The hard part is figuring out what to do with the numbers, not actually doing it. Most of the work in this course is that way, and the only way to do well is to practice a lot. Taking the midterm and final are similar to running in a five-mile race. Most people can do it in a respectable time, provided they train for it. Some people look at the runners who do it and say, “I’m as strong as they are. I’ll just go out there and join in.” Without the training, most of them are exhausted and walking after a mile or two.

So, our advice to students is, if this looks really weird to you, keep working at it. You will probably catch on. If it looks really simple, don’t get too complacent. It may be harder than you think. Get into the habit of doing the quizzes and problems, and if you don't answer all the quiz questions correctly, go over them until you understand each one.

We can't resist commenting on the role of probability and stochastic processes in our careers. The theoretical material covered in this book has helped both of us devise new communication techniques and improve the operation of practical systems. We hope you find the subject intrinsically interesting. If you master the basic ideas, you will have many opportunities to apply them in other courses and throughout your career.

We have worked hard to produce a text that will be useful to a large population of students and instructors. We welcome comments, criticism, and suggestions. Feel free to send us e-mail at *ryates@winlab.rutgers.edu* or *dgoodman@poly.edu*. In addition, the website `www.wiley.com/college/yates` provides a variety of supplemental materials, including the MATLAB code used to produce the examples in the text.

Roy D. Yates
Rutgers, The State University of New Jersey

David J. Goodman
New York University

September 27, 2013

Contents

1

Experiments, Models, and Probabilities

Getting Started with Probability

The title of this book is *Probability and Stochastic Processes*. We say and hear and read the word *probability* and its relatives (*possible, probable, probably*) in many contexts. Within the realm of applied mathematics, the meaning of *probability* is a question that has occupied mathematicians, philosophers, scientists, and social scientists for hundreds of years.

Everyone accepts that the probability of an event is a number between 0 and 1. Some people interpret probability as a physical property (like mass or volume or temperature) that can be measured. This is tempting when we talk about the probability that a coin flip will come up heads. This probability is closely related to the nature of the coin. Fiddling around with the coin can alter the probability of heads.

Another interpretation of probability relates to the knowledge that we have about something. We might assign a low probability to the truth of the statement, *It is raining now in Phoenix, Arizona*, because we know that Phoenix is in the desert. However, our knowledge changes if we learn that it was raining an hour ago in Phoenix. This knowledge would cause us to assign a higher probability to the truth of the statement, *It is raining now in Phoenix*.

Both views are useful when we apply probability theory to practical problems. Whichever view we take, we will rely on the abstract mathematics of probability, which consists of definitions, axioms, and inferences (theorems) that follow from the axioms. While the structure of the subject conforms to principles of pure logic, the terminology is not entirely abstract. Instead, it reflects the practical origins of probability theory, which was developed to describe phenomena that cannot be predicted with certainty. The point of view is different from the one we took when we started studying physics. There we said that if we do the same thing in the same way over and over again — send a space shuttle into orbit, for example —

the result will always be the same. To predict the result, we have to take account of all relevant facts.

The mathematics of probability begins when the situation is so complex that we just can't replicate everything important exactly, like when we fabricate and test an integrated circuit. In this case, repetitions of the same procedure yield different results. The situation is not totally chaotic, however. While each outcome may be unpredictable, there are consistent patterns to be observed when we repeat the procedure a large number of times. Understanding these patterns helps engineers establish test procedures to ensure that a factory meets quality objectives. In this repeatable procedure (making and testing a chip) with unpredictable outcomes (the quality of individual chips), the *probability* is a number between 0 and 1 that states the proportion of times we expect a certain thing to happen, such as the proportion of chips that pass a test.

As an introduction to probability and stochastic processes, this book serves three purposes:

- It introduces students to the logic of probability theory.
- It helps students develop intuition into how the theory relates to practical situations.
- It teaches students how to apply probability theory to solving engineering problems.

To exhibit the logic of the subject, we show clearly in the text three categories of theoretical material: definitions, axioms, and theorems. Definitions establish the logic of probability theory, and axioms are facts that we accept without proof. Theorems are consequences that follow logically from definitions and axioms. Each theorem has a proof that refers to definitions, axioms, and other theorems. Although there are dozens of definitions and theorems, there are only three axioms of probability theory. These three axioms are the foundation on which the entire subject rests. To meet our goal of presenting the logic of the subject, we could set out the material as dozens of definitions followed by three axioms followed by dozens of theorems. Each theorem would be accompanied by a complete proof.

While rigorous, this approach would completely fail to meet our second aim of conveying the intuition necessary to work on practical problems. To address this goal, we augment the purely mathematical material with a large number of examples of practical phenomena that can be analyzed by means of probability theory. We also interleave definitions and theorems, presenting some theorems with complete proofs, presenting others with partial proofs, and omitting some proofs altogether. We find that most engineering students study probability with the aim of using it to solve practical problems, and we cater mostly to this goal. We also encourage students to take an interest in the logic of the subject — it is very elegant — and we feel that the material presented is sufficient to enable these students to fill in the gaps we have left in the proofs.

Therefore, as you read this book you will find a progression of definitions, axioms, theorems, more definitions, and more theorems, all interleaved with examples and comments designed to contribute to your understanding of the theory. We also include brief quizzes that you should try to solve as you read the book. Each one

This notation tells us to form a set by performing the operation to the left of the vertical bar, $|$, on the numbers to the right of the bar. Therefore,

$$C = \{1, 4, 9, 16, 25\}. \tag{1.4}$$

Some sets have an infinite number of elements. For example

$$D = \{x^2 | x = 1, 2, 3, \ldots\}. \tag{1.5}$$

The dots tell us to continue the sequence to the left of the dots. Since there is no number to the right of the dots, we continue the sequence indefinitely, forming an infinite set containing all perfect squares except 0. The definition of D implies that $144 \in D$ and $10 \notin D$.

In addition to set inclusion, we also have the notion of a *subset*, which describes a relationship between two sets. By definition, A is a subset of B if every member of A is also a member of B. We use the symbol $\subset$ to denote subset. Thus $A \subset B$ is mathematical notation for the statement "the set A is a subset of the set B." Using the definitions of sets C and D in Equations (1.3) and (1.5), we observe that $C \subset D$. If

$$I = \{\text{all positive integers, negative integers, and } 0\}, \tag{1.6}$$

it follows that $C \subset I$, and $D \subset I$.

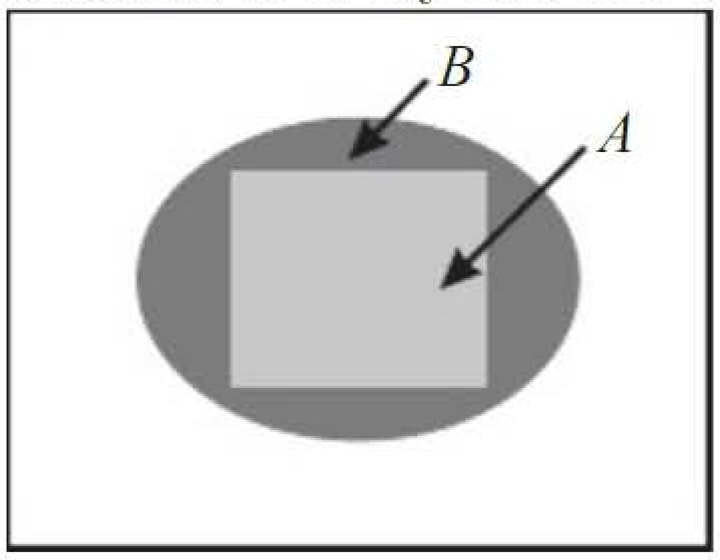
The definition of set equality, $A = B$, is

$$A = B \text{ if and only if } B \subset A \text{ and } A \subset B.$$

This is the mathematical way of stating that A and B are identical if and only if every element of A is an element of B and every element of B is an element of A. This definition implies that a set is unaffected by the order of the elements in a definition. For example, $\{0, 17, 46\} = \{17, 0, 46\} = \{46, 0, 17\}$ are all the same set.

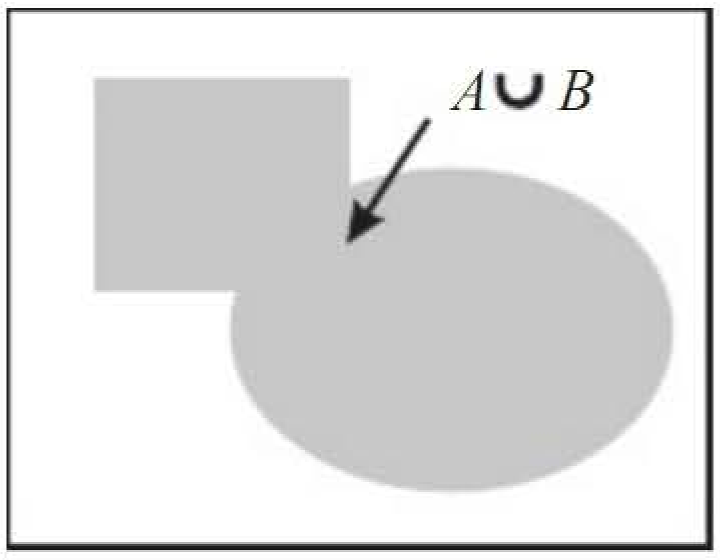
To work with sets mathematically it is necessary to define a *universal set*. This is the set of all things that we could possibly consider in a given context. In any study, all set operations relate to the universal set for that study. The members of the universal set include all of the elements of all of the sets in the study. We will use the letter S to denote the universal set. For example, the universal set for A could be $S = \{\text{all universities in the United States, all planets}\}$. The universal set for C could be $S = I = \{0, 1, 2, \ldots\}$. By definition, every set is a subset of the universal set. That is, for any set X, $X \subset S$.

The *null set*, which is also important, may seem like it is not a set at all. By definition it has no elements. The notation for the null set is $\varnothing$. By definition $\varnothing$ is a subset of every set. For any set A, $\varnothing \subset A$.



It is customary to refer to Venn diagrams to display relationships among sets. By convention, the region enclosed by the large rectangle is the universal set S. Closed surfaces within this rectangle denote sets. A Venn diagram depicting the relationship $A \subset B$ is shown on the left.

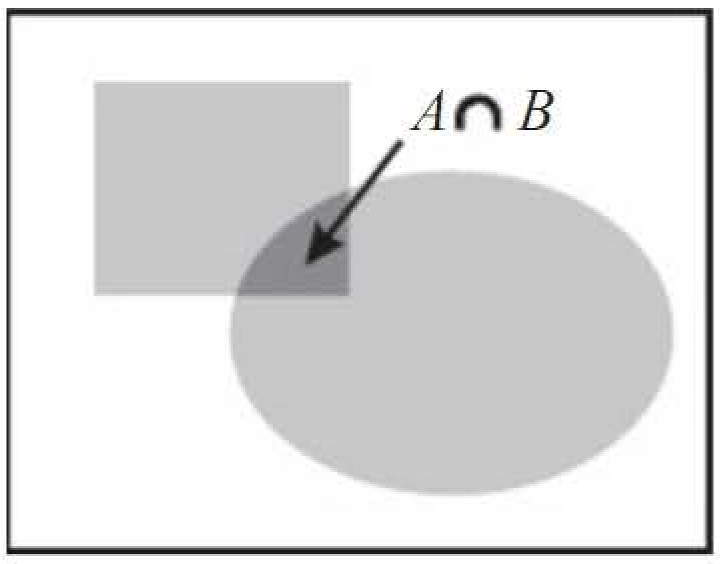
When we do set algebra, we form new sets from existing sets. There are three operations for doing this: *union*, *intersection*, and *complement*. Union and intersection combine two existing sets to produce a third set. The complement operation forms a new set from one existing set. The notation and definitions follow.



The *union* of sets A and B is the set of all elements that are either in A or in B, or in both. The union of A and B is denoted by $A \cup B$. In this Venn diagram, $A \cup B$ is the complete shaded area. Formally,

$$x \in A \cup B \text{ if and only if } x \in A \text{ or } x \in B.$$

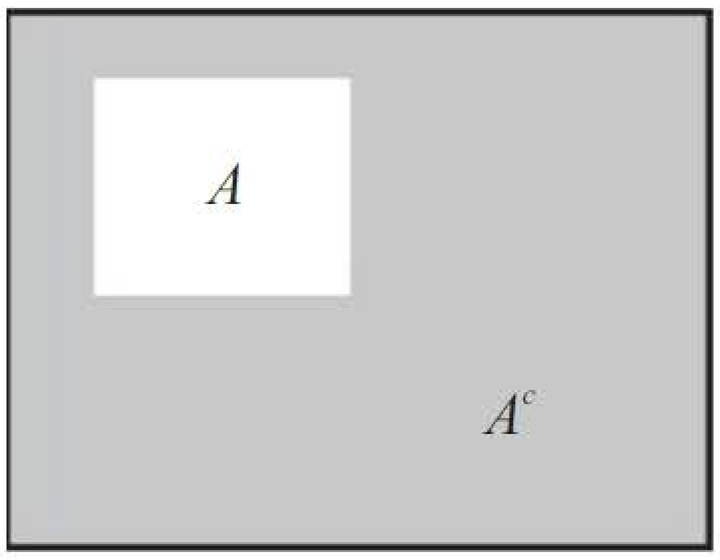
The set operation union corresponds to the logical "or" operation.



The *intersection* of two sets A and B is the set of all elements that are contained both in A and B. The intersection is denoted by $A \cap B$. Another notation for intersection is AB. Formally, the definition is

$$x \in A \cap B \text{ if and only if } x \in A \text{ and } x \in B.$$

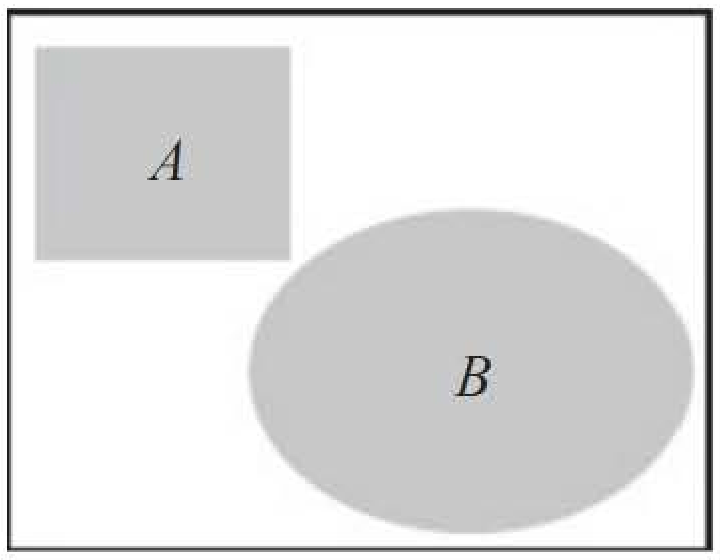
The set operation intersection corresponds to the logical "and" function.



The *complement* of a set A, denoted by A^c, is the set of all elements in S that are not in A. The complement of S is the null set $\varnothing$. Formally,

$$x \in A^c \text{ if and only if } x \notin A.$$

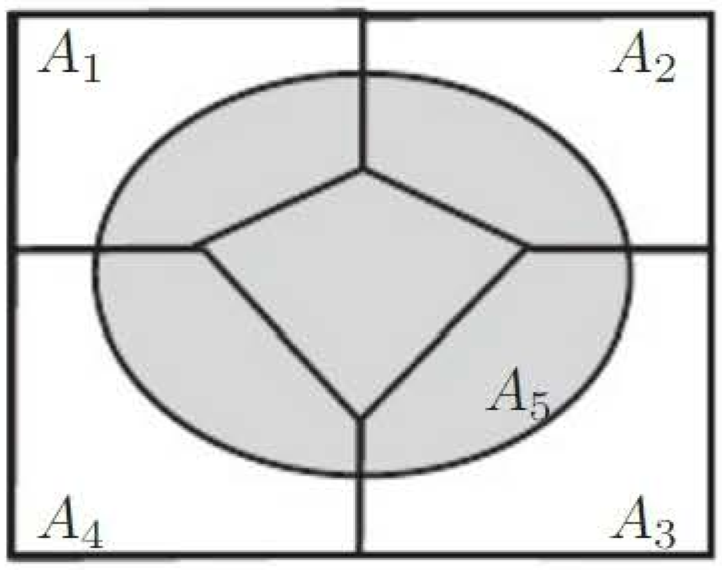
In working with probability we will often refer to two important properties of collections of sets. Here are the definitions.



A collection of sets $A_1, \ldots, A_n$ is *mutually exclusive* if and only if

$$A_i \cap A_j = \varnothing, \qquad i \neq j. \tag{1.7}$$

The word *disjoint* is sometimes used as a synonym for mutually exclusive.



A collection of sets $A_1, \ldots, A_n$ is *collectively exhaustive* if and only if

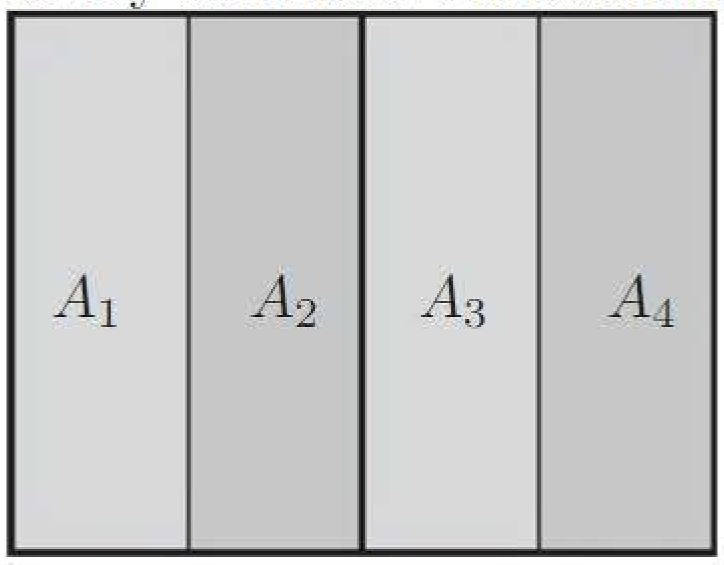
$$A_1 \cup A_2 \cup \cdots \cup A_n = S. \tag{1.8}$$

In the definition of *collectively exhaustive*, we used the somewhat cumbersome notation $A_1 \cup A_2 \cup \cdots \cup A_n$ for the union of N sets. Just as $\sum_{i=1}^{n} x_i$ is a shorthand for $x_1 + x_2 + \cdots + x_n$, we will use a shorthand for unions and intersections of n sets:

$$\bigcup_{i=1}^{n} A_i = A_1 \cup A_2 \cup \cdots \cup A_n, \tag{1.9}$$

$$\bigcap_{i=1}^{n} A_i = A_1 \cap A_2 \cap \cdots \cap A_n. \tag{1.10}$$

We will see that collections of sets that are both mutually exclusive and collectively exhaustive are sufficiently useful to merit a definition.



A collection of sets $A_1, \ldots, A_n$ is a *partition* if it is both mutually exclusive and collectively exhaustive.

From the definition of set operations, we can derive many important relationships between sets and other sets derived from them. One example is

$$A \cap B \subset A. \tag{1.11}$$

To prove that this is true, it is necessary to show that if $x \in A \cap B$, then it is also true that $x \in A$. A proof that two sets are equal, for example, $X = Y$, requires two separate proofs: $X \subset Y$ and $Y \subset X$. As we see in the following theorem, this can be complicated to show.

Theorem 1.1

De Morgan's law relates all three basic operations:

$$(A \cup B)^c = A^c \cap B^c.$$

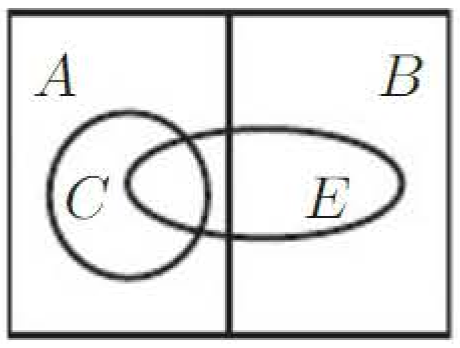
Proof There are two parts to the proof:

- To show $(A \cup B)^c \subset A^c \cap B^c$, suppose $x \in (A \cup B)^c$. That implies $x \notin A \cup B$. Hence, $x \notin A$ and $x \notin B$, which together imply $x \in A^c$ and $x \in B^c$. That is, $x \in A^c \cap B^c$.

- To show $A^c \cap B^c \subset (A \cup B)^c$, suppose $x \in A^c \cap B^c$. In this case, $x \in A^c$ and $x \in B^c$. Equivalently, $x \notin A$ and $x \notin B$ so that $x \notin A \cup B$. Hence, $x \in (A \cup B)^c$.

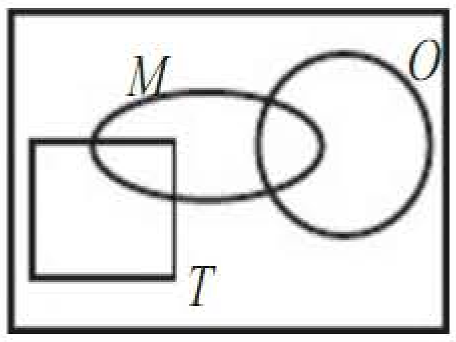
Example 1.1

Phonesmart offers customers two kinds of smart phones, Apricot (A) and Banana (B). It is possible to buy a Banana phone with an optional external battery E. Apricot customers can buy a phone with an external battery (E) or an extra memory card (C) or both. Draw a Venn diagram that shows the relationship among the items A,B,C and E available to Phonesmart customers.



Since each phone is either Apricot or Banana, A and B form a partition. Since the external battery E is available for both kinds of phones, E intersects both A and B. However, since the memory card C is available only to Apricot customers, $C \subset A$. A Venn diagram representing these facts is shown on the right.

Quiz 1.1



Gerlandas offers customers two kinds of pizza crust, Tuscan (T) and Neapolitan (N). In addition, each pizza may have mushrooms (M) or onions (O) as described by the Venn diagram at right. For the sets specified below, shade the corresponding region of the Venn diagram.

(a) N (b) $N \cup M$

(c) $N \cap M$ (d) $T^c \cap M^c$

1.2 Applying Set Theory to Probability

> Probability is based on a repeatable experiment that consists of a procedure and observations. An *outcome* is an observation. An *event* is a set of outcomes.

The mathematics we study is a branch of measure theory. Probability is a number that describes a set. The higher the number, the more probability there is. In this sense probability is like a quantity that measures a physical phenomenon; for example, a weight or a temperature. However, it is not necessary to think about probability in physical terms. We can do all the math abstractly, just as we defined sets and set operations in the previous paragraphs without any reference to physical phenomena.

Fortunately for engineers, the language of probability (including the word *probability* itself) makes us think of things that we experience. The basic model is a repeatable *experiment*. An experiment consists of a *procedure* and *observations*. There is uncertainty in what will be observed; otherwise, performing the experiment would be unnecessary. Some examples of experiments include

1. Flip a coin. Did it land with heads or tails facing up?
2. Walk to a bus stop. How long do you wait for the arrival of a bus?
3. Give a lecture. How many students are seated in the fourth row?
4. Transmit one of a collection of waveforms over a channel. What waveform arrives at the receiver?
5. Transmit one of a collection of waveforms over a channel. Which waveform does the receiver identify as the transmitted waveform?

For the most part, we will analyze *models* of actual physical experiments. We create models because real experiments generally are too complicated to analyze. For example, to describe *all* of the factors affecting your waiting time at a bus stop, you may consider

- The time of day. (Is it rush hour?)
- The speed of each car that passed by while you waited.
- The weight, horsepower, and gear ratios of each kind of bus used by the bus company.
- The psychological profile and work schedule of each bus driver. (Some drivers drive faster than others.)
- The status of all road construction within 100 miles of the bus stop.

It should be apparent that it would be difficult to analyze the effect of each of these factors on the likelihood that you will wait less than five minutes for a bus. Consequently, it is necessary to study a *model* of the experiment that captures the important part of the actual physical experiment. Since we will focus on the model of the experiment almost exclusively, we often will use the word *experiment* to refer to the model of an experiment.

Example 1.2

An experiment consists of the following procedure, observation, and model:

- Procedure: Monitor activity at a Phonesmart store.
- Observation: Observe which type of phone (Apricot or Banana) the next customer purchases.
- Model: Apricots and Bananas are equally likely. The result of each purchase is unrelated to the results of previous purchases.

As we have said, an experiment consists of both a procedure and observations. It is important to understand that two experiments with the same procedure but with different observations are different experiments. For example, consider these two experiments:

Example 1.3

Monitor the Phonesmart store until three customers purchase phones. Observe the sequence of Apricots and Bananas.

Example 1.4

Monitor the Phonesmart store until three customers purchase phones. Observe the number of Apricots.

These two experiments have the same procedure: monitor the Phonesmart store until three customers purchase phones. They are different experiments because they require different observations. We will describe models of experiments in terms of a set of possible experimental outcomes. In the context of probability, we give precise meaning to the word *outcome*.

Definition 1.1 Outcome

An ***outcome*** *of an experiment is any possible observation of that experiment.*

Implicit in the definition of an outcome is the notion that each outcome is distinguishable from every other outcome. As a result, we define the universal set of all possible outcomes. In probability terms, we call this universal set the *sample space*.

Definition 1.2 Sample Space

The ***sample space*** *of an experiment is the finest-grain, mutually exclusive, collectively exhaustive set of all possible outcomes.*

The *finest-grain* property simply means that all possible distinguishable outcomes are identified separately. The requirement that outcomes be mutually exclusive says that if one outcome occurs, then no other outcome also occurs. For the set of outcomes to be collectively exhaustive, every outcome of the experiment must be in the sample space.

Example 1.5

- The sample space in Example 1.2 is $S = \{a, b\}$ where a is the outcome "Apricot sold," and b is the outcome "Banana sold."
- The sample space in Example 1.3 is

$$S = \{aaa, aab, aba, abb, baa, bab, bba, bbb\} \tag{1.12}$$

- The sample space in Example 1.4 is $S = \{0, 1, 2, 3\}$.

Example 1.6

Manufacture an integrated circuit and test it to determine whether it meets quality objectives. The possible outcomes are "accepted" (a) and "rejected" (r). The sample space is $S = \{a, r\}$.

Set Algebra	Probability
Set	Event
Universal set	Sample space
Element	Outcome

Table 1.1 The terminology of set theory and probability.

In common speech, an event is something that occurs. In an experiment, we may say that an event occurs when a certain phenomenon is observed. To define an event mathematically, we must identify *all* outcomes for which the phenomenon is observed. That is, for each outcome, either the particular event occurs or it does not. In probability terms, we define an event in terms of the outcomes in the sample space.

Definition 1.3 Event

An ***event*** *is a set of outcomes of an experiment.*

Table 1.1 relates the terminology of probability to set theory. All of this may seem so simple that it is boring. While this is true of the definitions themselves, applying them is a different matter. Defining the sample space and its outcomes are key elements of the solution of any probability problem. A probability problem arises from some practical situation that can be modeled as an experiment. To work on the problem, it is necessary to define the experiment carefully and then derive the sample space. Getting this right is a big step toward solving the problem.

Example 1.7

Suppose we roll a six-sided die and observe the number of dots on the side facing upwards. We can label these outcomes $i = 1, \ldots, 6$ where i denotes the outcome that i dots appear on the up face. The sample space is $S = \{1, 2, \ldots, 6\}$. Each subset of S is an event. Examples of events are

- The event $E_1 = \{\text{Roll 4 or higher}\} = \{4, 5, 6\}$.
- The event $E_2 = \{\text{The roll is even}\} = \{2, 4, 6\}$.
- $E_3 = \{\text{The roll is the square of an integer}\} = \{1, 4\}$.

Example 1.8

Observe the number of minutes a customer spends in the Phonesmart store. An outcome T is a nonnegative real number. The sample space is $S = \{T|T \geq 0\}$. The event "the customer stays longer than five minutes is $\{T|T > 5\}$.