PART III REVIEW

Some phenomena are random. That is, although their individual outcomes are unpredictable, there is a regular pattern in the long run. Gambling devices (rolling dice, spinning roulette wheels) and taking an SRS are examples of random phenomena. Probability and expected value give us a language to describe randomness. Random phenomena are not haphazard or chaotic any more than random sampling is haphazard. Randomness is instead a kind of order in the world, a long-run regularity as opposed to either chaos or a determinism that fixes events in advance. Chapter 17 discusses the idea of randomness, Chapter 18 presents basic facts about probability, and Chapter 20 discusses expected values.

When randomness is present, probability answers the question, "How often in the long run?" and expected value answers the question, "How much on the average in the long run?" The two answers are connected because the definition of "expected value" is in terms of probabilities. Much work with probability starts with a probability model that assigns probabilities to the basic outcomes. Any such model must obey the rules of probability. Another kind of probability model uses a density curve such as a Normal curve to assign probabilities as areas under the curve. Personal probabilities express an individual's judgment of how likely some event is. Personal probabilities for several possible outcomes must also follow the rules of probability if they are to be consistent with each other.

To calculate the probability of a complicated event without using complicated math, we can use random digits to simulate many repetitions. You can also find expected values by simulation. Chapter 19 shows how to do simulations. First give a probability model for the outcomes, then assign random digits to imitate the assignment of probabilities. The table of random digits now imitates repetitions. Keep track of the proportion of repetitions on which an event occurs to estimate its probability. Keep track of the mean outcome to estimate an expected value.

PART III SUMMARY

Here are the most important skills you should have acquired after reading Chapters 17 through 20.

A. RANDOMNESS AND PROBABILITY

1. Recognize that some phenomena are random. Probability describes the long-run regularity of random phenomena.

2. Understand the idea of the probability of an event as the proportion of times the event occurs in very many repetitions of a random phenomenon. Use the idea of probability as long-run proportion to think about probability.

3. Recognize that short runs of random phenomena do not display the regularity described by probability. Understand that randomness is unpredictable in the short run. Avoid seeking causal explanations for random occurrences.

B. PROBABILITY MODELS

1. Use basic probability facts to detect illegitimate assignments of probability: any probability must be a number between 0 and 1, and the total probability assigned to all possible outcomes must be 1.

2. Use basic probability facts to find the probabilities of events that are formed from other events: the probability that an event does not occur is 1 minus its probability. If two events cannot occur at the same time, the probability that one or the other occurs is the sum of their individual probabilities.

3. When probabilities are assigned to individual outcomes, find the probability of an event by adding the probabilities of the outcomes that make it up.

4. When probabilities are assigned by a Normal curve, find the probability of an event by finding an area under the curve.

C. EXPECTED VALUE

1. Understand the idea of expected value as the average of numerical outcomes in very many repetitions of a random phenomenon.

2. Find the expected value from a probability model that lists all outcomes and their probabilities (when the outcomes are numerical).

D. SIMULATION

1. Specify simple probability models that assign probabilities to each of several stages when the stages are independent of each other.

2. Assign random digits to simulate such models.

3. Estimate either a probability or an expected value by repeating a simulation many times.

PART III REVIEW EXERCISES

Review exercises are short and straightforward exercises that help you solidify the basic ideas and skills in each part of this book. We have provided "hints" that indicate where you can find the relevant material for the odd-numbered problems.

III.1. What's the probability? If you have access to a printed copy, open your local Yellow Pages telephone directory to any page in the Business White Pages listing. Look at the last four digits of each telephone number, the digits that specify an individual number within an exchange given by the first three digits. Note the first of these four digits in each of the first 100 telephone numbers on the page.

If you don't have access to a printed directory, you can use an online directory. Find an online directory with at least 100 different phone numbers. Look at the last four digits of each telephone number, the digits that specify an individual number within an exchange given by the first three digits. Note the first of these four digits in each of the first 100 telephone numbers in the directory.

One directory we found online was at http://education.ohio.gov/Contact/Phone-Directory. State government agencies often have directories with office phone numbers of employees.

(a) How many of the digits are 1, 2, or 3? What is the approximate probability that the first of the four "individual digits" in a telephone number is 1, 2, or 3? (*Hint*: See page 407.)

(b) If all 10 possible digits had the same probability, what would be the probability of getting a 1, 2, or 3? Based on your work in part (a), do you think the first of the four "individual digits" in telephone numbers is equally likely to be any of the 10 possible digits? (*Hint*: See page 429.)

III.2. Blood types. Choose a person at random and record his or her blood type. Here are the probabilities for each blood type:

Blood type:	Type O	Type A	Type B	Type AB
Probability:	0.4	0.3	0.2	?

(a) What must be the probability that a randomly chosen person has Type AB blood?

(b) To simulate the blood types of randomly chosen people, how would you assign digits to represent the four types?

III.3. Grades in an economics course. Indiana University posts the grade distributions for its courses online. Students in Economics 201 in the fall 2013 semester received 18% A's, 8% A−'s, 7% B+'s, 16% B's, 11% B−'s, 6% C+'s, 12% C's, 4% C−'s, 3% D+'s, 6% D's, 4% D−'s, and 6% F's. Choose an Economics 201 student at random. The probabilities for the student's grade are:

Grade	A	A−	B+	B
Probability	0.18	0.08	0.07	0.16

Grade	B−	C+	C	C−
Probability	0.11	0.06	0.12	0.04
Grade	D+	D	D−	F
Probability	0.03	0.06	0.04	?

(a) What must be the probability of getting an F? (*Hint*: See page 429.)

(b) To simulate the grades of randomly chosen students, how would you assign digits to represent the five possible outcomes listed? (*Hint*: See page 447.)

III.4. Blood types. People with Type B blood can receive blood donations from other people with either Type B or Type O blood. Tyra has Type B blood. What is the probability that two or more of Tyra's six close friends can donate blood to her? Using your work in Exercise III.2, simulate 10 repetitions and estimate this probability. (Your estimate from just 10 repetitions isn't reliable, but you have shown in principle how to find the probability.)

III.5. Grades in an economics course. If you choose four students at random from all those who have taken the course described in Exercise III.3, what is the probability that all the students chosen got a B or better? Simulate 10 repetitions of this random choosing and use your results to estimate the probability. (Your estimate from only 10 repetitions isn't reliable, but if you can do 10, you could do 10,000.) (*Hint*: See page 447.)

III.6. Grades in an economics course. Choose a student at random from the course described in Exercise III.3 and observe what grade that student earns (with A = 4, A− = 3.7, B+ = 3.3, B = 3.0, B− = 2.7, C+ = 2.3, C = 2.0, C− =1.7, D+ = 1.3, D = 1.0, D− = 0.7, and F = 0.0).

(a) What is the expected grade of a randomly chosen student?

(b) The expected grade is not one of the 12 grades possible for one student. Explain why your result nevertheless makes sense as an expected value.

III.7. Dice. What is the expected number of spots observed in rolling a carefully balanced die once? (*Hint*: See page 467.)

III.8. Profit from a risky investment. Rotter Partners is planning a major investment. The amount of profit X is uncertain, but a probabilistic estimate gives the following distribution (in millions of dollars):

Profit:	−1	0	1	2
Probability:	0.1	0.1	0.2	0.2
Profit:	3	5	20	
Probability:	0.2	0.1	0.1	

What is the expected value of the profit?

III.9. Poker. Deal a five-card poker hand from a shuffled deck. The probabilities of several types of hand are approximately as follows:

Hand:	Worthless	One pair
Probability:	0.50	0.42
Hand:	Two pairs	Better hands
Probability:	0.05	?

(a) What must be the probability of getting a hand better than two pairs? (*Hint*: See page 429.)

(b) What is the expected number of hands a player is dealt before the first hand better than one pair appears? Explain how you would use simulation to answer this question, then simulate just two repetitions. (*Hint*: See page 472.)

III.10. How much education? The Census Bureau gives this distribution of education for a randomly chosen American over 25 years old in 2014:

Education:	Less than high school	High school graduate	College, no bachelor's
Probability:	0.117	0.297	0.167

Education:	Associate's degree	Bachelor's degree	Advanced degree
Probability:	0.099	0.202	0.118

(a) How do you know that this is a legitimate probability model?

(b) What is the probability that a randomly chosen person over age 25 has at least a high school education?

(c) What is the probability that a randomly chosen person over age 25 has at least a bachelor's degree?

III.11. Language study. Choose a student in grades 9 to 12 at random and ask if he or she is studying a language other than English. Here is the distribution of results:

Language:	Spanish	French	German
Probability:	0.300	0.080	0.021

Language:	Latin	All others	None
Probability:	0.013	0.022	0.564

(a) Explain why this is a legitimate probability model. (*Hint*: See page 429.)

(b) What is the probability that a randomly chosen student is studying a language other than English? (*Hint*: See page 429.)

(c) What is the probability that a randomly chosen student is studying French, German, or Spanish? (*Hint*: See page 429.)

III.12. Choosing at random. Abby, Deborah, Mei-Ling, Sam, and Roberto work in a firm's public relations office. Their employer must choose two of them to attend a conference in Paris. To avoid unfairness, the choice will be made by drawing two names from a hat. (This is an SRS of size 2.)

(a) Write down all possible choices of two of the five names. These are the possible outcomes.

(b) The random drawing makes all outcomes equally likely. What is the probability of each outcome?

(c) What is the probability that Mei-Ling is chosen?

(d) What is the probability that neither of the two men (Sam and Roberto) is chosen?

III.13. Languages in Canada. Canada has two official languages: English and French. Choose a resident of Quebec at random and ask, "What is your mother tongue?" Here is the distribution of responses, combining many separate languages from the province of Quebec:

Language:	English	French	Other
Probability:	0.083	0.789	?

(a) What is the probability that a randomly chosen resident of

Quebec's mother tongue is either English of French? (*Hint*: See page 429.)

(b) What is the probability that a randomly chosen resident of Quebec's mother tongue is "Other"? (*Hint*: See page 429.)

III.14. Is college worth the cost? A September 29, 2015, Gallup poll asked recent college graduates (those who obtained their bachelor's degree beginning in 2006) whether their education was worth the cost. Assume that the results of the poll accurately reflect the opinions of all recent college graduates. Here is the distribution of responses:

Response:	Agree strongly	Agree somewhat	Neither agree nor disagree
Probability:	0.38	0.27	0.17
Response:	Disagree somewhat	Disagree strongly	
Probability:	0.09	?	

(a) What is the probability that a randomly chosen recent college graduate disagrees strongly?
(b) What is the probability that a randomly chosen recent college graduate agrees strongly or agrees somewhat that his or her education was worth the cost?

III.15. An IQ test. The Wechsler Adult Intelligence Scale (WAIS) is a common IQ test for adults. The distribution of WAIS scores for persons over 16 years of age is approximately Normal with mean 100 and standard deviation 15. Use the 68–95–99.7 rule to answer these questions.
(a) What is the probability that a randomly chosen individual has a WAIS score of 115 or higher? (*Hint*: See pages 432–435.)

(b) In what range do the scores of the middle 95% of the adult population lie? (*Hint*: See pages 434–437.)

III.16. Worrying about crime. How much do Americans worry about crime and violence? Suppose that 40% of all adults worry a great deal about crime and violence. (According to sample surveys that ask this question, 40% is about right.) A polling firm chooses an SRS of 2400 people. If they do this many times, the percentage of the sample who say they worry a great deal will vary from sample to sample following a Normal distribution with mean 40% and standard deviation 1.0%. Use the 68–95–99.7 rule to answer these questions.

(a) What is the probability that one such sample gives a result within $\pm 1.0\%$ of the truth about the population?

(b) What is the probability that one such sample gives a result within $\pm 2\%$ of the truth about the population?

III.17. An IQ test (optional). Use the information in Exercise III.15 and Table B to find the probability that a randomly chosen person has a WAIS score of 112 or higher. (*Hint*: See pages 434–435.)

III.18. Worrying about crime (optional). Use the information in Exercise III.16 and Table B to find the probability that one sample misses the truth about the population by 2.5% or more. (This is the probability that the sample result is either less than 37.5% or greater than 42.5%.)

III.19. An IQ test (optional). How high must a person score on the WAIS test to be in the top 10% of all scores? Use the information in Exercise III.15 and Table B to answer this question. (*Hint*: See pages 434–435.)

III.20. Models, legitimate and not. A bridge deck contains 52 cards, four of each of the 13 face values ace, king, queen, jack, ten, nine, . . . , two. You deal a single card from such a deck and record the face value of the card dealt. Give an assignment of probabilities to the possible outcomes that should be correct if the deck is thoroughly shuffled. Give a second assignment of probabilities that is legitimate (that is, obeys the rules of probability) but differs from your first choice. Then give a third assignment of probabilities that is *not* legitimate, and explain what is wrong with this choice.

III.21. Mendel's peas. Gregor Mendel used garden peas in some of the experiments that revealed that inheritance operates randomly. The seed color of Mendel's peas can be either green or yellow. Suppose we produce seeds by "crossing" two plants, both of which carry the G (green) and Y (yellow) genes. Each parent has probability 1-in-2 of passing each of its genes to a seed, independently of the other parent. A seed will be yellow unless both parents contribute the G gene. Seeds that get two G genes are green.

What is the probability that a seed from this cross will be green? Set up a simulation to answer this question, and estimate the probability from 25 repetitions. (*Hint*: See page 447.)

III.22. Predicting the winner. There are 14 teams in the Big Ten athletic conference. Here's one set of personal probabilities for next year's basketball champion: Michigan State has probability 0.3 of winning. Illinois, Iowa, Minnesota, Nebraska, Northwestern, Penn State, and Rutgers have no chance. That leaves six teams. Indiana, Michigan, Purdue, Ohio State, and Wisconsin also all have the same probability of winning, but that probability is one-half that of Maryland. What probability does each of the 14 teams have?

III.23. Selling cars. Bill sells new cars in a small town for a living. On a weekday afternoon, he will deal with one customer with probability 0.6, two customers with probability 0.3, and three customers with probability 0.1. Each customer has probability 0.2 of buying a car. Customers buy independently of each other.

Describe how you would simulate the number of cars Bill sells in an afternoon. You must first simulate the number of customers, then simulate the buying decisions of one, two, or three customers. Simulate one afternoon to demonstrate your procedure. (*Hint*: See page 452.)

PART III PROJECTS

Projects are longer exercises that require gathering information or producing data and that emphasize writing a short essay to describe your work. Many are suitable for teams of students.

Project 1. A bit of history. On page 409, we said, "The systematic study of randomness . . . began when seventeenth-century French gamblers asked French mathematicians for help in figuring out the 'fair value' of bets on games of chance." Pierre de Fermat and Blaise Pascal were two of the mathematicians who responded. Both are interesting characters. Choose one of these men. Write a brief essay giving his dates, some anecdotes you find noteworthy from his life, and at least one example of a probability problem he studied. (A Web search on the name will produce abundant information. Remember to use your own words in writing your essay.)

Project 2. Reacting to risks. On page 418, we quoted a writer as saying, "Few of us would leave a baby sleeping alone in a house while we drove off on a 10-minute errand, even though car-crash risks are much greater than home risks." Take it as a fact that the probability that the baby will be injured in the car is very much higher than the probability of any harm occurring at home in the same time period. Would you leave the baby alone? Explain your reasons in a short essay. If you would not leave the baby alone, be sure to explain why you choose to ignore the probabilities.

Project 3. First digits. Here is a remarkable fact: the first digits of the numbers in long tables are usually *not* equally likely to have any of the 10 possible values 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9. The digit 1 tends to occur with probability roughly 0.3, the digit 2 with probability about 0.17, and so on. You can find more information about this fact, called "Benford's law," on the Web or in two articles by Theodore P. Hill, "The Difficulty of Faking Data," *Chance,* 12, No. 3 (1999), pp. 27–31; and "The First Digit Phenomenon," *American Scientist,* 86 (1998), pp. 358–363. You don't have to read these articles for this project.

Locate at least two long tables whose entries could plausibly begin with any digit. You may choose data tables, such as populations of many cities, the number of shares traded on the New York Stock Exchange on many days, or mathematical tables such as logarithms or square roots. We hope it's clear that you can't use the table of random digits. Let's require that your examples each contain at least 300 numbers. Tally the first digits of all entries in each table. Report the distributions (in percentages) and compare them with each other, with Benford's law, and with the "equally likely" distribution.

Project 4. Personal probability. Personal probabilities are personal, so we expect them to vary from person to person. Choose an event that most students at your school should have an opinion about, such as rain next Friday or a victory in your team's next game. Ask many students (at least 50) to tell you what probability they would assign to rain or a victory. Then analyze the data with a graph and numbers—shape, center, spread, and all that. What do your data show about personal probabilities for this future event?

Project 5. Making decisions. Exercise 20.12 (page 476) reported the results of a study by the psychologist Amos Tversky on the effect of wording on people's decisions about chance outcomes. His subjects were college students. Repeat Tversky's study at your school. Prepare two typed cards. One says:

You are responsible for treating 600 people who have been exposed to a fatal virus. Treatment A has probability 1-in-2 of saving all 600 and probability 1-in-2 that all 600 will die. Treatment B is guaranteed to save exactly 400 of the 600 people. Which treatment will you give?

The second card says:

You are responsible for treating 600 people who have been exposed to a fatal virus. Treatment A has probability 1-in-2 of saving all 600 and probability 1-in-2 that all 600 will die. Treatment B will definitely lose exactly 200 of the lives. Which treatment will you give?

Show each card to at least 25 people (25 different people for each, chosen as randomly as you can conveniently manage and chosen from people who have not studied probability). Record the choices. Tversky claims that people shown the first card tend to choose B, while those shown the second card tend to choose A. Do your results agree with this claim? Write a brief summary of your findings: Do people use expected values in their decisions? Does the frame in which a decision is presented (the wording, for example) influence choices?