applied mathematics practices
for the 21st century

MATHEMATICAL MODELING WITH PROBABILITY

Using Authentic Problem Contexts

Kenneth Chelst, Ph.D.
Thomas Edwards, Ph.D.

Applied Mathematics Practices for the 21st Century (AMP21) is an affiliation of professors and mathematics educators who share a common desire to bring relevance to K-12 mathematics by using authentic problem contexts in teaching and developing mathematics concepts and skills. The lead faculty are Kenneth Chelst (College of Engineering) and Thomas Edwards (College of Education) who are both professors at Wayne State University in Detroit, MI.

AMP21 is a Non-Profit developer and provider of curriculum that is aligned with the eight Standards for Mathematical Practice in the Common Core State Standards. Our team has published two textbooks for high school, 1) Mathematical Modeling with Algebra and 2) Mathematical Modeling with Probability. Our new curriculum development focuses on middle school topics central to proportional reasoning: percentages, rates, ratios and proportions. Everything is presented in authentic problem contexts that involve making decisions. AMP21 offers professional development workshops in conjunction with local universities to help teachers and schools develop the programs needed to enable students to succeed in the global economy. In addition, this curriculum can form a basis for project based learning in mathematics.

First Printing, 2018
ISBN -

Applied Mathematics Practices (AMP21)
In conjunction with Wayne State University
www.appliedmathpractices.com

Contact: Kenneth Chelst
4815 Fourth St. (Room 2017)
Detroit, MI 48201
kchelst@wayne.edu

Acknowledgements:

This work was developed initially with funding from the National Science Foundation (NSF Directorate for Education and Human Resources Project # DRL-0733137). The grant was a partnership between North Carolina State University, University of North Carolina – Charlotte, and Wayne State University. The following people contributed to the initial project, known as Project MINDSET:

Wayne State University

Dr. Thomas Edwards
Dr. Kenneth Chelst
Mr. Jay Johnson

North Carolina State University

Dr. Karen Keene
Dr. Karen Norwood
Dr. Robert Young
Dr. Molly Purser
Dr. Amy Craig-Reamer

University of North Carolina Charlotte

Dr. David Royster
Dr. David Pugalee
Ms. Sarah Johnson

Researchers

Dr. Saman Alaniazar
Dr. Hatice Ucar
Dr. Richelle Dietz
Dr. Krista Holstein
Mr. Will Hall
Ms. Zeyneb Yurtseven

Teachers

Derek Blackwelder, Jamie Blanchard James Bruckman, Kevin Busfield, Robert Butler, Molly Charles, Suzanne Christopher, Kristina Clayton, Krystle Corbin, Dean DiBasio (teacher writer), Robin Dixon, Caroline Geel, Michael Gumpp, Breanna Harrill, Jack Hunter, Amy Johnson, Jennifer Johnson, Jenise Jones (teacher writer), Anna Kamphaus, Chris Kennedy, George Lancaster (teacher writer), Craig Lazarski, Jeffrey Loewen, Charles Ludwick, Margaret Lumsden, Leon Martin, Cora McMillan, Kristen Meck (teacher writer), Karen Mullins, Erica Nelson, Bethany Peters, Joseph Price, Angela Principato (teacher writer), John Pritchett, Tyler Pulis, Paul Rallo, Tonya Raye, Kym Roberts, William Ruff, Amy Setzer, Rob Sharpe, Christy Simpson, Catherine Sitek, Cheryl Smith, Julia Smith, Wendy Carol Srinivasan, Susan Sutton, Greg Tucker, Laura Turner, Thad Wilhelm (teacher writer), Tracey Walden, Tracey Weigold, Lanette Wood

Cover Design: Susan Marie Dials

Significant contribution to the 2018 second edition revision: Thad Wilhelm

MATHEMATICAL MODELING WITH PROBABILITY
Using Authentic Problem Contexts

Kenneth Chelst and Thomas Edwards

Table of Contents

Chapter 1: Basic Probability and Randomness

We all face decisions in our jobs, in our communities, and in our personal lives that involve uncertainty. When making such decisions, there can be no guarantee that the outcome will be favorable. Probability decision models are designed to help an individual to assess the likelihood of various outcomes. The decision maker then uses that information to select a preferred choice that explicitly considers the relative likelihood of both positive and negative results.

In perceiving and responding to the world around us, it is critical to develop an understanding of the difference between patterns of random and non-random events. There is well documented fallacy in probabilistic thinking in which people see patterns when in fact all they have observed is random fluctuations. A decision maker faced with a series of negative outcomes must decide whether or not to act. Could this simply be a random fluctuation and no action needs to be taken? Or is there something causing these negative outcomes that he can and should address?

In this chapter we strive to develop a better understanding of random patterns through the use of simulated experiments. We begin with a physical simulation, flipping a coin. This can be used to model randomness with two equally likely events. Random number generators are introduced to represent randomness when two events are not equally likely. These random number generators are found in advanced calculators and Excel spreadsheets.

This chapter introduces the basic concepts and notation of probability. The multiplication rule is used to calculate the probability of occurrence of two independent random events. The principle of complementarity is used to indirectly compute the likelihood of an event. Missing from this chapter is any discussion of counting methods, combinations and permutations. Historically, these were developed primarily to apply probability to games of chance. In our experience we have never had to use counting rules to tackle non-gambling decisions involving uncertainty.

Chapter 2: Conditional Probability

Chapter 2 introduces the concept of conditional probability, P(B|A). Given that event A has occurred, what is the probability that event B will occur? Unlike most probability textbooks, we dedicate an entire chapter to explore conditional probability. Research has shown that many people have poor intuition regarding conditional probability concepts.

This chapter illustrates how to determine conditional probability from data tables as well as from a problem context. It illustrates the application of the multiplication rule to calculate joint probabilities of two events that are not independent. The chapter proceeds to develop the formula for a probabilistic partition. In this chapter we introduce the concept of a random variable. This is a function that translates the outcome of a random experiment into a unique numeric value. The measure of central tendency of this random variable is a probabilistic weighted sum called the expected value.

Chapter 3: Decision Trees

Decision trees provide a structure for determining the alternative that optimizes the expected value. Decisions are represented by rectangles and uncertain events by ovals. Branches emanating from a rectangle correspond to distinct decision alternatives. Branches from a random event correspond to different possible outcomes. The decision tree also provides the probability distribution for each of the alternatives.

The first example simply introduces the concept of tradeoff between more or less conservative decision alternatives and possible outcomes. Two subsequent decision contexts maximize the expected profit. The last example minimizes the expected cost of collision insurance.

Chapter 4: Binomial Distribution and Geometric Distribution

In Chapter 1 we used basic probability and simulation to study random events with two mutually exclusive outcomes: answer or do not answer phone, at work or absent from work, on time or late. In chapter 4 we introduce the concept of a probability distribution function. When repeated random events follow the same assumptions, there may be a formula that summarizes the probabilistic pattern. This formula will have parameters whose values vary from context to context.

The binomial distribution can be applied to a situation of identical independent repetitions of the same random experiment. The four common elements to the random experiment are:

1. only two possible outcomes: success and failure
2. p, the probability of success, is the same for each repetition
3. each repetition is independent of every other repetition
4. n identical repetitions of the random event

The parameters of the binomial distribution are n and p. The formula for the binomial distribution calculates the probability of X successes out of n trials. The random variable X has a range from zero to n. Statistical calculators and Excel spreadsheets include formulas for the binomial distribution and the cumulative binomial distribution. The binomial distribution examples in this chapter are all extensions of those that appeared in Chapter 1.

In the binomial distribution we count the number of success out of n trials. The geometric distribution includes the same first three elements as the binomial. However, the random variable that is tracked is the number of repetitions until the first success. This random variable can take on the values of one to infinity. One interesting application of the geometric distribution involves the number of shuttle flights until the first shuttle disaster.

Chapter 5: Poisson Distribution

The Poisson distribution is used to characterize a probabilistic environment in which random events occur totally independent of one another. Emergency calls to 911 or calls to a telephone helpline are examples of situations that have been modeled with the Poisson distribution. This distribution has one parameter, λ, the average number of incidents per unit of time. The distribution is used to estimate the probability that there will be X events in a unit time. Alternatively, it can be used to model an extended time period, t. In that case, the average or expected value of the random variable is λt. The random variable X is discrete and has a range of zero to infinity. Managers face the difficult task of scheduling the necessary resources to handle the fluctuations in workload caused by the unpredictability of these random events.

Chapter 6: Normal Distribution

The Normal distribution is used to describe a continuous random variable with a mean of μ and a standard deviation of σ. The sum of identically distributed independent random variables is known to approach the normal distribution. For that reason, the randomness of the sample average of a random variable can be approximated with the normal distribution. In the first example, the normal distribution is used to model the random error associated with cutting panels of material for use in constructing a parachute. The normal distribution can help determine the proportion of panels that are within specifications. The second example uses the normal distribution to help establish an appropriate warranty replacement policy. In the final example, a manager must decide how many units of a new toy to stock when the demand is normally distributed. The decision maker balances the risk of having too few items and losing out on sales against the risk of overstocking and having to steeply discount toys that are left over after the buying season passes.

Introduction to Managing and Deciding in an Uncertain World

The world around us is filled with uncertainty, risk, and variability that complicate day-to-day decisions and the development of long-term plans. This applies to personal decisions as well as decisions made by companies and government agencies. An applicant to college cannot be sure which schools will accept him. In a rush to get a meal between classes, a student faces uncertainty about the time it will take to be served at the school cafeteria. The school newspaper editor is concerned about how many members of the writing team will meet their deadlines. While reviewing alternative car insurance plans, the student driver struggles to decide on the size of the collision insurance deductible. Obviously, no one plans to have an accident, but the risk of an accident is always present. Companies that provide insurance look at the same problem. They come up with pricing strategies for insurance that pool the risk of an individual with large groups of similar people.

Variability is a characteristic of data that refers to the recognition that each data value is not the same. For example, there is variability in height of individuals or in their annual income. Two common statistics used to characterize a dataset's variability are variance and standard deviation.

Companies launching a new product or service must deal with uncertain demand. Police patrol supervisors must consider random fluctuations in the demand for service and patterns of crime. Plant managers and school officials must cope with workers who randomly do not show up for work due to illness.

Probability theory and sophisticated probability models enable us to identify individual occurrences that are more or less likely, as well as estimate long-term averages. This knowledge and understanding is essential for planning and making critical decisions in advance as well as adjustments on the spot. A plant or office manager must decide how many spare workers to have on duty to address the problem of absenteeism. The high school newspaper editor will also have to plan what to do if only eight of 10 writers have completed their assignments. Similarly, an automobile dealership needs to know how much inventory to carry. The police commander must schedule the number patrol officers on each shift and where to place them without knowing where and when the next crime will occur. Often, decision makers must balance the added costs of having reserves against the impact of running short.

This text is designed to provide decision-making guidance in the presence of uncertainty. One of the challenges in learning basic concepts of probability is that many of us do not have good intuition about randomness. We will address this problem while developing probabilistic decision-making skills. The text is therefore designed to develop your ability to recognize and understand patterns of random events. We will do so by having you simulate random experiments first with a coin flip. Then, you will simulate more complex situations with the random number function in your calculator. Lastly, you will use the random number generator in Excel to develop and analyze a large sample of randomly generated data.

In introducing basic concepts, we routinely use the concept of ***relative frequency*** as an estimate of probability. Thus, our introductory examples will use data rather than the counting methods that you may have seen in other probability courses.

Randomness is a characteristic of a process, experiment, or environment in which outcomes cannot be predicted with certainty. A random experiment, such as rolling a die, can yield different unpredictable outcomes. The gender of newborn is a result of a random process involved in becoming pregnant. The temperature on any day reflects randomness of our environment.

In later chapters we introduce mathematical formulas that can be used to describe different patterns of randomness. These are probability distributions. These formulas will be applied to two different types of variables: discrete and continuous. A ***discrete variable*** is countable. For example, the number of crimes in a day and the number of people absent from work are discrete variables. On the other hand, a ***continuous variable*** cannot be counted and is measured instead. For example, the time to complete an exam and the height of a randomly selected individual are continuous variables.

CHAPTER 1:

Basic Probability and Randomness

Section 1.0 Basic Probability and Randomness

In order to fully understand probabilistic modeling, you will need a good sense of the nature of ***random*** behavior. All of us have basic intuition about probability and randomness. This intuition develops over time from our experiences. These include what we watch on television, observe at sporting events, and read on the Internet. The problem is that research has shown that our intuition about randomness and probability is often flawed. For example, a manager being told that each and every one of his 10 suppliers is 95% reliable often believes the system he has in place is in good shape. In this chapter, you will learn why the manager needs to be concerned.

Because of the many misconceptions, the development of probability analysis skills requires demonstrating the flaws in our intuition. Only after this is accomplished, can you move on to learn formal approaches to probability. The first example we explore involves understanding what is often our preconceived notion of a random pattern. Specifically, we will ask you to imagine and write down a typical sequence of heads and tails when flipping a coin. You and your classmates will then be asked to actually flip a coin. By comparing the lists, we hope to dislodge a misconception you may have about random events.

One of the primary goals of this section and all future sections is to begin understanding how much **variability** there may be in a simple random process. This will be accomplished by comparing and contrasting the results of your coin flipping experiment with those of your neighbors and those of the class as a whole. For example, we will explore how much difference there is in the percentage of heads and tails in your list and those of each and every one of your classmates. We will then look at the average for the entire class.

One challenge managers and we personally face in a variety of situations involves understanding and interpreting fluctuations around a long-term average. Is a rise in absenteeism or lateness simply a reflection of a random fluctuation or an indication of a developing problem that needs to be addressed? Biostatisticians regularly are on the lookout for occurrences of disease that are out of the "normal" range. A report comes in that a small city has twice the national average rate of a specific type of cancer. There are tens of thousands of communities. Is it likely or unlikely purely by random fluctuation to see one or more cities with a rate twice the national average? If it is likely, no action should be taken, no detailed study initiated. If however, probability theory suggests this high rate for a small city is extremely rare, public health officials may commission an expensive study to gather more data to try to identify the contributing factors.

You will explore the pattern of which conference won the Super Bowl in an attempt to draw a conclusion about whether the data suggest that one conference is superior to the other. You will also look at data for a customer service telephone line to determine whether there is sufficient evidence that the service is not meeting the timeliness standard established by management. You will assist the faculty advisor to the school newspaper to better understand why they are having problems publishing the paper on time. Finally, you will assist a plant manager to decide how many spare workers to have on duty. These are needed to maintain productivity while coping with random fluctuations in the number of workers who are absent from work each day.

A key question you will always be addressing is, "What is the source of variability?" The fluctuation around the mean may be merely random variation or it could be the result of some external influence. Discerning whether the pattern of variation is random or due to some factor we might be able to control is a critical component of probabilistic and statistical analysis. We need these tools because, in trying to make this distinction, our own intuition often leads us to unfounded conclusions.

Section 1.1 Super Bowl – Conference Dominance

Every year since 1967, the American Football Conference (AFC) and National Football Conference (NFC) champions of the National Football League compete in a championship game known as the Super Bowl. Table 1.1.1 identifies by conference the winner of each of the 48 Super Bowls from I through XLVIII. The question is:

Super Bowl					
Year	**Winning Conference**	**Year**	**Winning Conference**	**Year**	**Winning Conference**
1967	NFC	1983	NFC	1999	AFC
1968	NFC	1984	AFC	2000	NFC
1969	AFC	1985	NFC	2001	AFC
1970	AFC	1986	NFC	2002	AFC
1971	AFC	1987	NFC	2003	NFC
1972	NFC	1988	NFC	2004	AFC
1973	AFC	1989	NFC	2005	AFC
1974	AFC	1990	NFC	2006	AFC
1975	AFC	1991	NFC	2007	AFC
1976	AFC	1992	NFC	2008	NFC
1977	AFC	1993	NFC	2009	AFC
1978	NFC	1994	NFC	2010	NFC
1979	AFC	1995	NFC	2011	NFC
1980	AFC	1996	NFC	2012	NFC
1981	AFC	1997	NFC	2013	AFC
1982	NFC	1998	AFC	2014	NFC

Table 1.1.1: Super Bowl winners by conference

What is it about the sequence of conference wins that has led most sports writers to believe that one conference or the other was superior at different extended periods in time?

1. What do you notice about these results?
2. Does one conference appear stronger than the other? Why or why not?
3. If neither conference is superior to the other, what number would it make sense to use as the probability that the winner will come from a given conference?
4. What was the overall percentage of wins by each conference? Each column represents a 16-year period. What was the winning percentage in each 16-year period?

In order to explore your conception of randomness, we will model the Super Bowl results since 1967. Assume that the two conferences are equally strong. If so, each conference winner has the same 50% chance of winning the Super Bowl in any year. We will use a coin flip to represent the 50% chance. First you will imagine and write down a typical sequence of heads and tails. Only afterwards, will you actually flip a coin.

5. On a sheet of paper, list the numbers from 1 to 48 to represent each of the Super Bowls from 1967 to 2014. Table 1.1.2 shows how to set up the list before entering the Hs and Ts. You will use an H to represent an AFC win and a T for an NFC win. Now imagine a random sequence of wins assuming the two conferences are equally strong. Record next to each number, an H or a T.

1		17		33	
2		18		34	
3		19		35	
4		20		36	
5		21		37	
6		22		38	
7		23		39	
8		24		40	
9		25		41	
10		26		42	
11		27		43	
12		28		44	
13		29		45	
14		30		46	
15		31		47	
16		32		48	
Number of Heads		Number of Heads		Number of Heads	
Percent of Heads		Percent of Heads		Percent of Heads	

Table 1.1.2: A format for recording simulated coin flips

After recording all 48 Hs and/or Ts, fill in the bottom rows of the table: the number of heads and percent of heads in each column. Also calculate the percent of Hs for all 48 imagined events.

6. Compare the percent of heads in your three columns and the overall percent. How much difference is there among the column percentages?

7. Compare the percent of heads in your list with those of two or three other students sitting close by. Are your results similar, or is there a lot of difference among them?

Next, make another list just like the one in Table 1.1.2 to record the results of your actual coin flips. After recording all 48 Hs and/or Ts, fill in the bottom rows of the table: the number of heads and percent of heads in each column. Also calculate the percent of Hs for all 48 flips.

8. Compare the percent of heads in your three columns and the overall percent. How much difference is there among the column percentages?

9. Compare the percent of heads in your list with those of two or three other students sitting close by. Are your results similar, or is there a lot of difference among them?

In the following questions, you will compare your imagined sequence with the sequence of the actual coin flips. Be sure to label your lists, so that you'll know which is which.

10. Was there more variability in the column percentages among the imagined lists or the coin flip lists?

11. What were the maximum and minimum percentages for you and your neighbors for each column for the imagined lists and coin flip lists? Are you surprised at how far these numbers are from 50% in the actual coin flip data?

12. What were the maximum and minimum percent of Hs and Ts for you and your neighbors for all 48 imagined events and coin flips? What differences do you notice about the ranges?

13. For the entire class, what was the minimum percent of Hs observed when flipping the coins 48 times? What was the maximum? Are you surprised at how far these numbers are from 50%?

14. Did anyone in the class have a percentage that low or that high in the imagined list?

15. For the entire class, what was the overall percent of Hs and Ts for the coin flip random experiment?

We will now return to the actual data for the Super Bowl in Table 1.2.1. Find the total the number of times in each of the three columns that the NFC won and calculate the percent.

16. Look at the percentage of NFC wins in the first column of the actual Super Bowl data. In your coin flip table, was there any percentage of Hs or Ts that low? Did anyone in the entire class have a percentage of Hs or Ts that low in a column of coin flips?

17. Look at the percentage of NFC wins in the second column of the actual Super Bowl data. In your coin flip table, was there any percentage of Hs or Ts that high? Did anyone in the entire class have a percentage of Hs or Ts that high in a column of coin flips?

18. Complete the Table 1.1.3 for the class as a whole. In this table we want the frequency of either a large number of Hs or Ts. In the summary of the 16 year periods, we would be surprised if either of the divisions won an unusually large proportion of times. Thus, in assessing the likelihood of one conference winning, for example, 11 or more times in 16 years, we need to consider both Hs and Ts.

	Large number of Hs		Large number of Ts		Hs and Ts Combined	
Number of Hs or Ts	Frequency	Relative Frequency	Frequency	Relative Frequency	Frequency	Relative Frequency
11						
12						
13						
14						
15						
16						

Table 1.1.3: Class frequency of high percentages

19. Based on Table 1.1.3, what is the likelihood that one conference would win 11 or more Super Bowls in a 16 year period? 12 or more? 13 or more? 14 or more?

20. Based on your comparisons between your actual coin flips and the Super Bowl wins, do the percentages strongly support the news writers' claims of dominance in any of the 16 year periods?

Up to this point we have explored different aspects of the percent of Hs and Ts. We have compared the imagined sequence with that generated by actual coin flips. We have also compared minimum and maximum percentages for columns and for the total. By including your classmates, you were able to consider larger and larger sets of data. For the class as a whole, you should have seen a wider range of percentages for the column percentages than you found in your data and your neighbor's. In addition, there should have been less variability in the total percent of conference wins over 48 years. The total percent for the entire class would likely be close to 50%.

Now we are going to look at the sequences of consecutive Hs or Ts. These sequences are called strings. There have been numerous class room experiments comparing made up lists and randomly generated lists. In general, the lists of imagined sequences of Hs and Ts tend to have shorter strings than randomly generated sequences. In addition, imagined lists rarely have any long strings of four more. We are going to see if your experience and those of your classmates replicate the research.

Record the lengths of every string of consecutive heads or tails in your imagined list. Then do the same for your coin flip list. For example, in Table 1.1.4, the first column records a sequence of Hs and Ts. The second column records the length of each string. In Table 1.1.4 the longest string is 3.

<table>
<tr><td>T</td><td rowspan="2">2</td></tr>
<tr><td>T</td></tr>
<tr><td>H</td><td>1</td></tr>
<tr><td>T</td><td>1</td></tr>
<tr><td>H</td><td rowspan="2">2</td></tr>
<tr><td>H</td></tr>
<tr><td>T</td><td>1</td></tr>
<tr><td>H</td><td rowspan="3">3</td></tr>
<tr><td>H</td></tr>
<tr><td>H</td></tr>
<tr><td>T</td><td rowspan="3">3</td></tr>
<tr><td>T</td></tr>
<tr><td>T</td></tr>
<tr><td>H</td><td rowspan="2">2</td></tr>
<tr><td>H</td></tr>
<tr><td>T</td><td>1</td></tr>
</table>

Table 1.1.4: Length of strings for imagined list

21. Compare the lengths of the strings in your two lists. Which list has the longest string? How long is that string? Compare your answers with those of two or three neighbors. In your group, what is the length of the longest string? Which list did it come from?

22. Compare your answers with the entire class. What is the length of the longest string? Which list did it come from?

23. In the Super Bowl results, what was the longest string of wins for one conference? Did any of the coin flip lists have a string this long?

24. Based on your comparisons do the data support the sports writers' claims of dominance?

25. Have you discovered any misconceptions in your own thinking about randomness?

1.1.1 Percentage, Long Strings, and Probabilistic Thinking

Every alternative sequence of 48 heads and tails has exactly the same probability of occurrence, one in 281.5 trillion. Thus the likelihood of seeing any particular sequence of alternating heads and tails is the same as the probability of seeing a sequence of all heads or all tails. However, the likelihood of producing exactly 24 heads out of the total of 48 is orders of magnitude higher than the chance of 48 heads in a row. That is due to the fact that there are many sequences of heads and tails that result in exactly 24 heads, but there is only one way to obtain 48 heads in a row. Thus, 48 heads in a row might lead you to believe that the coin is two-headed.

This same principle points to an interesting phenomenon. When random flips of a coin are carried out, the sequence often includes multiple examples of long runs of four or more heads or four or more tails. However, individuals recording their own made-up sequence have a tendency to believe that randomness means frequently alternating back and forth between heads and tails. Consequently, they tend to create lists that have few long strings of heads or tails. As a result, an experienced probabilistic thinker can often identify which of the two lists was generated by flipping a coin and which was generated by a human mind. The human mind tends to look at the list as it develops and work to balance the distribution of heads and tails.

> One of the primary goals of probabilistic analysis is to attempt to assess whether a pattern is simply random or if some factor is contributing to the pattern. It is never possible to resolve this issue with certainty, but an understanding of probability theory can help decide which driving force is more likely.

Section 1.2 Simulate Two-outcome Random Event

On the pages that follow, you will be introduced to a wide range of real-world problems that operations researchers routinely encounter. They approach these real-world problems using an array of techniques that apply principles of probability and statistics. You may already be familiar with many of these concepts, but some of them might be new. We believe that a major difference between our use of probability and statistics in this textbook and what you may have seen in the past is that we will use probability and statistics in the context of making decisions.

Along the way, you will also learn about the role of computers in exploring real-world problems by actually using computers. A computer allows us to simulate a situation rather than attempt to collect large amounts of data. It can also enable us to realistically model changes in the design of the system and determine the impact on system performance. This is much preferred to trying out each change in the real-world and observing the impact. When we use mathematics in this way, it is more as a modeling tool and less as a computational tool. The primary role of an operations research (OR) professional is to formulate the model and interpret the results that the model returns.

Studying mathematics in this way may require you to develop a new mindset about what mathematics is, how it can be used, and the best ways to learn it. Whatever mathematics is, it is certainly not a spectator sport. That is why a large portion of this book looks very similar to the Super Bowl simulation problem that you just read. Many questions were asked, but few were answered in the text. Where will the answers come from? You will provide those answers with the help of your classmates and teacher. In doing so, you will also discover that applying mathematics to solve real-world problems can be a relevant, creative, and exciting team sport. We hope that all of this will help you begin to form that new mindset.

1.2.1 Advance Planning

Imagine you and your extended family of 50 people have planned a major outdoor dinner barbecue. Do you have a plan for rain? What if a week in advance you are told there is a 5% chance of rain? What if it is 50% or 80%? What options do you have if you start planning a week in advance that you may not have the morning of the event? What communication challenges do you face? How can you plan for them to ensure as much as possible that everyone knows what to do? The situation a manager faces is similar. When he is trying to schedule work to be completed, he is not sure how many workers will be absent on any given day. He also cannot be sure how many of them will complete their assigned tasks on time. The more he understands about random phenomena, the better able he is to develop a plan to deal with these uncertainties.

1.2.2 A Simple Uncertainty Model – Two Possible Random Outcomes

The simplest kind of uncertainty to model is one that involves only two possible equally likely outcomes, such as a coin flip. There are many other uncertain situations with only two possible random outcomes: a worker or student is present or absent, an assignment is handed in on time or not, or a team wins or loses a game. However, in these cases the two possible outcomes are not necessarily equally likely. The critical number, often called a **parameter**, is the probability of the outcome occurring. For example, it would be 0.1 if there is a 10% chance a worker will be absent on any given day. It might be 0.95, if a student is extremely reliable in turning in his or her assignments on time. In both cases, the pairs of outcomes are both ***mutually exclusive*** and ***complementary***. They are mutually exclusive, because both outcomes cannot happen at the same time. A flip of a coin is either heads or tails but not both. They are complementary, because if one does not happen, the other must. If the coin does not show a head, then it must show a tail. If a worker shows up for work, then he is not absent and vice versa. Complementary events are two events that are also ***collectively exhaustive***. The two complementary events exhaust all of the possible outcomes. As a result, the sum of the two probabilities must equal one. The equation below uses mathematical notation to express this complementary relationship.

$$P(E) + P(E^c) = 1$$

Mathematical notation is used to concisely express ideas about quantities. However, it is difficult to interpret if you do now know what the symbols mean. In probability theory, a capital P followed by an event enclosed in parentheses represents the probability of that event occurring. So $P(E)$ represents the probability that event E will occur. The symbol E^c represents the complement of event E.

This leads to a useful formula for calculating the probability of the complement of an event.

$$P(E^c) = 1 - P(E)$$

This simple situation forms the basis for all of the examples in this chapter. We expand on this two-alternative context to study what happens when the same situation is repeated three, five, 10, 20, or more times. Here we are interested in the total number of times one outcome or the other occurs. In the worker situation, we are interested in how many workers are absent each day because management is concerned about the potential impact on productivity.

The starting point in modeling these uncertain situations is to determine the probability, p, of a specific occurrence. We arbitrarily call this a success. For a coin that is fair, we assume there it is a 0.5 probability of a head and a 0.5 probability of a tail. For the worker case, the manager would look at historical data and determine the proportion of days missed by the workers in his organization. The challenge is to use this value to understand what happens when the same situation is repeated multiple times. For example, how many heads will occur if we flip a coin 12 times? It is extremely unlikely if we repeat the 12 coin flips two, three, or four times that we will see exactly the same sequence of heads and tails. However, we will explore what types of predictions we can make, and with what degree of likelihood.

Anytime we gather data about an uncertain situation, we have information about how things were. What we do not know is how likely this situation is to be repeated. What happens if we change something, what will the random patterns look like then? For example, our small office of 10 workers may have extensive data on the random number of workers who are absent from work each day during flu season. How would we predict from this information what the situation is likely to be if we expanded to 14 workers or cut back to seven? A mathematical probability model can assist us in this analysis.

1.2.3 Simulations

The simplest and most flexible method of modeling probability is to ***simulate*** the situation. For example, assume there is a 50% probability of reaching a customer service representative right after opening time We could let "heads" represent reaching a customer service representative and "tails" represent not reaching a representative. We can simulate a work-week by flipping the coin five times and recording the results. Then we can count the number of times in the week that customer service was open on time. In this way, by flipping a coin, we can get a sense of what the results of actually calling might have been.

A 50-50 probability can be simulated by flipping a coin, but what about a random event in which the chance of success is 0.7 and the chance of failure is 0.3? In this chapter we will use the random number generator found on your calculator and also in EXCEL to simulate these types of situations. To simulate a 50-50 probability, we can generate any two random integers to represent the two equally likely outcomes. When the two events are not equally likely, we need to be more creative. To simulate a 0.7 chance of success, we will generate random integers between 1 and 10. If the number is 7 or less, we record a success and if the number is 8 or more, we record a failure. To simulate a 0.95 chance, we will generate random integers between 1 and 100. Any value less than or equal to 95 represents a success. If the probability is recorded to three decimal places, we will use the integers from 1 to 1000. With this simple modeling tool we will be able explore, study, and come to understand two-outcome random phenomena with different values of p, the probability of success.

Section 1.3 Customer Service at Koala Foods

Like most companies, Koala Foods, a wholesale distributor of foods to supermarkets, is concerned about its customer service. Koala Foods assures the supermarkets it serves that their needs can be addressed on weekdays beginning at 8:00 a.m. by calling a toll-free customer service hotline. The president of Koala Foods wants to increase customer satisfaction, so he hires a company that will monitor the customer service department. Specifically, the monitoring company will initially check daily that the customer service department is actually operational by 8:00 a.m. Later the monitoring company will track individual calls to assess the quality of service.

1.3.1 Forming a Strategy

AGB Company, a customer service consulting firm, has been contracted by Koala Foods to monitor its customer service department. There are various ways of measuring quality in this situation. Mr. Smith, the President of Koala Foods, proposes checking to see if the customer service department is operational by 8:02 a.m., at 8:05 a.m., and lastly at 8:10 a.m. He would be satisfied if the customer service department was ready and answering calls by 8:02 at least 50% of the time. He sets a higher standard of 70% for 8:05 a.m. He feels that by 8:10 a.m. the service should almost always be open and sets a standard of 95%. This is equivalent to the service being ready to answer calls at 8:10 a.m. 19 days out of 20.

On the first three days of the week, AGB Company calls the Koala Foods customer service department at 8:02 a.m. Each day they encounter an "hours of service" recording rather than an actual customer service representative.

1. Does Koala Foods have a problem?

2. What are some possible reasons a call would not be answered at 8:02 a.m.?

3. Is this enough information to decide that there is a problem in meeting the 50% standard?

1.3.2 Simulating the Situation

We are interested in better understanding the significance of failing to reach customer service three days in a row. We need to decide whether this is sufficient evidence of a problem. If so, Koala Foods management will have to introduce greater supervision to ensure the standard is met. The alternative is to continue calling each morning to collect more data to see if there truly is a problem.

In order to gain the necessary understanding of this sequence of three consecutive failures, AGB Company decides to simulate three consecutive calls a large number of times. They want to know how likely it is to see three failures in a row if the 50% standard was being met overall. You and your classmates will assume the role of AGB and simulate this 50-50 random experiment by flipping a coin. Let the "heads" side of the coin represent a call that gets through to a customer service representative. The "tails" side of the coin represents a call that receives a recorded message.

Let H = answered call
T = unanswered call (recorded message)

4. Do you think that your first three consecutive flips will all be tails? If you repeat a set of three flips a total of 10 times, do you think you will ever see three tails in a row?

5. Do you think that anybody in the class will ever get all three tails?

6. Flip the coin three times and record the results. Repeat this process nine more times. How many times did you observe a series of three tails?

7. Did any of your classmates flip a series of three tails? Count the total number of times that you and your classmates observed a set of three tails in a row. What is the proportion of sets of three flips that were three tails in a row?

8. Think back to Koala Foods Customer Service Department and meeting the 50% standard. Does it seem unusual that AGB failed to reach customer service at 8:02 a.m. three days in a row if the 50% standard was being met?

Another method of simulation uses the random number generator in your graphing calculator. It is capable of simulating three repetitions of this random event with one single command. It is also very easy to repeat this process as many times as you want. Within the `MATH` menu, the `PRB` submenu contains the random integer command, `randInt(`. If you enter randInt(0,1,3), your calculator will return a set of three random integers that are either "0" or "1." Each time you push `ENTER`, your calculator will generate and display a new set of three values. Let "0" represent a call that receives a recorded message and "1" represent a call that gets through to a customer service representative.

9. Enter `randInt(0,1,3)` into your calculator ten times and count the number of times the string `{0 0 0}` occurs.

10. Compare your answer to #9 with those of your classmates.

11. Would you reconsider your answer to #4 and #5?

12. Should you reconsider your response to #1?

After seeing the results of the simulation, Mr. Smith decides the data presented by AGB Company was inconclusive. He now knows that he would need data for more than three days to conclude with confidence that the standard of 50% is not being met.

1.3.3 Probability and Relative Frequency

The relative frequency of an event in a large number of simulations can be used to estimate the underlying probability of the event happening. However, by making some basic assumptions, it is often possible to calculate the probability directly. This is true for the coin flipping example. Each time a fair coin is flipped, the chance of observing a head or tail is 0.5. It is independent of the previous outcome. The coin has no memory, so it does not "remember" the result of the previous flip. The probability that a set of ***independent*** events will occur can be found by multiplying the probabilities of each of the events.

> The ***relative frequency*** of an event is equal to the ratio of the number of times the event occurred to the total number of observations.

In this example, we are interested in the likelihood of observing three tails in a row. Using the multiplication rule, this probability is

$$(0.5)(0.5)(0.5) = 0.125$$

or

$$\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{8}.$$

The likelihood is one chance in eight. As a result, most of the students who carried out 10 repetitions of three coin flips would have observed at least one set of three tails. Earlier, the students were asked to accumulate the class's data to determine the overall relative frequency of three tails. The relative frequency should be close to 0.125, but not exactly that value because of randomness.

Thus, it is possible that Koala Foods is maintaining a 50% standard, but that these particular three days represent an unusual but not rare event. The standard that is often used to characterize an event as rare is less than one-in-twenty or 0.05. In some instances an even stricter standard of one-in-a-hundred, 0.01, is used to characterize something as a rare event.

> A ***random experiment*** is any activity under consideration in which the outcome is unpredictable, such as rolling a die or checking who is absent from class. Each possible observation in an experiment is called an ***outcome***. The set of possible outcomes of an experiment is called the ***sample space***. Any subset of a sample space is called an ***event***.
>
> A ***simple event*** is an event that cannot be decomposed. For example, 1, 2, 3, 4, 5, and 6 are all simple events associated with rolling a die once. The sample space in this experiment is {1,2,3,4,5,6}. Rolling an even number is a ***compound event*** that is made up of three simple events {2,4,6}.

In this example we used the flip of a coin to represent two equally likely outcomes. The sample space for the coin experiment consists of all possible outcomes of flipping a coin three times. The set of outcomes is listed below.

{H,H,H}
{H,H,T}
{H,T,H}
{T,H,H}
{H,T,T}
{T,H,T}
{T,T,H}
{T,T,T}

All of these eight outcomes are equally likely. As a result, the likelihood of any particular outcome, such as {T,T,T}, is 1/8.

In general, when all outcomes are equally likely, we can use the following formula to calculate the probability of an event E occurring.

$$P(E) = \frac{\text{Number of ways event } E \text{ can occur}}{\text{Number of possible outcomes}}$$

We can use this to determine the likelihood that on only one day out of three the phone is answered at 8:02 a.m.

Let E_1 = the event that only one call is answered and two calls are unanswered at 8:02 a.m.

This is a compound event corresponds to observing two tails and one head on three flips of a coin. The list of all possible ways of E_1 occurring consists of the following outcomes.

{H,T,T} {T,H,T} {T,T,H}

Thus, out of the eight total possible outcomes, there are three distinct, mutually exclusive, and equally likely outcomes that correspond to event E_1.

$$P(E_1) = \frac{3}{8} = 0.375$$

1.3.4 Dealing with a Higher Standard

Recall that the standard for reaching customer service at 8:05 a.m. was 70%.

13. Does the 70% standard for calls placed at 8:05 a.m. seem like a reasonable expectation if you are a
    a. manager at Koala Foods?
    b. customer service representative at Koala Foods?
    c. customer of Koala Foods who needs service?

AGB also called at 8:05 a.m. and three days in a row, no one answered. Mr. Smith wonders if there is any reason why these three days of data might be conclusive about whether the customer service department is meeting this higher standard. He decides to ask AGB to conduct another simulation. Once again, you and your classmates will play the role of AGB.

14. Will a coin be adequate to simulate this scenario? Why or why not?

To use the calculator, enter randInt(1,10,3). This will generate three random integers between 1 and 10, inclusive. Assign the numbers "1" through "7" to represent calls that are answered by customer service. The numbers "8" through "10"represent calls that receive machine recorded message.

15. Why does this method accurately model the 70% standard?

16. You and your classmates will perform this simulation 10 times each. Do you think there will be more, the same, or fewer simulated instances of receiving the recorded message three times in a row as compared to the first simulation?

Execute this simulation on your calculator 10 times. For each set of three, count and record the number of times a number between "8" and "10" occurs.

17. Did you have any lists of three numbers in which all three numbers were 8, 9, or 10?

18. Did any of your classmates observe a set of three numbers that were all 8, 9, or 10? Count the total number of times this occurred in the class's pooled data. In what proportion of total simulations did that occur?

19. Think back to Koala Foods Customer Service Department and meeting the 70% standard. Does it seem unusual that AGB failed to reach customer service at 8:05 a.m. three days in a row if the 70% standard was being met?

With the higher standard, the probability of answering a call is 0.7. Not answering the call is the complementary event, E^c. The probability of a complementary event is equal to one minus the probability of the event.

$$P\left(E^c\right) = 1 - P\left(E\right)$$

$$P\left(\text{Not answering call}\right) = 1 - P\left(\text{Answering call}\right) = 1 - 0.7 = 0.3$$

We can again use the multiplication rule to determine the likelihood of observing three days of non-answered calls at 8:05 if Koala Foods were achieving the 70% standard.

$$\begin{aligned} P(\text{No answer three days in a row}) &= P(\text{No answer day 1}) \cdot P(\text{No answer day 2}) \cdot P(\text{No answer day 3}) \\ &= (0.3)(0.3)(0.3) \\ &= 0.027 \end{aligned}$$

This is approximately one chance in 37. As a result, most of the students who carried out the 10 simulations would not have observed even one instance of a set of three numbers that were all 8, 9, or 10. Three days of unanswered calls is characterized as a rare event based on the 5% criterion mentioned earlier. Thus, it is reasonable to conclude that the customer service department of Koala Foods is not meeting the 70% standard for 8:05. Management will need to take action to correct the problem.

In calculating probabilities related to this event, we cannot apply the strategy of counting the number of possible outcomes that correspond to the event. All of the possible events ae listed below. No call answered three days in a row is just one of eight possible outcomes. However, its probability is not 1/8 or 0.125

Let A = answered call
$A^c = N$ = not answered call

{A,A,A}
{A,A,N}
{A,N,A}
{N,A,A}
{A,N,N}
{N,A,N}
{N,N,A}
{N,N,N}

These are all of the eight possible outcomes. However, they are not equally likely. Consider the likelihood that the 8:05 call was answered three days in a row. We can calculate the probability with the multiplication rule for independent events. If we assume the 70% standard, the probability is shown below.

$$P(A, A, A) = P(A) \cdot P(A) \cdot P(A) = 0.7^3 = 0.343$$

This is very different from the probability of no calls being answered.

$$P(N, N, N) = P(N) \cdot P(N) \cdot P(N) = 0.3^3 = 0.027$$

1.3.5 Dealing with a 95% Standard

Recall that the standard for reaching customer service at 8:10 a.m. was even higher. It was set at 95%. AGB also called on Monday, Tuesday, and Wednesday at 8:10 a.m. This time the results were different. One out of the three times, they reached customer service. The other two times, they received a recorded message. Once again Mr. Smith wonders if these results indicate conclusively that the standard of 95% is not being met. He decides to ask AGB to conduct another simulation. Once again, you and your classmates will play the role of AGB.

20. Determine the beginning and the end of an appropriate random integer interval that could be used to simulate this situation. Assume again a three-day monitoring period.

    randInt(_____,_____,_____)

21. What range of integers represents a call that is answered by a customer service representative?

22. What range of integers represents a call that receives a recording?

As you run your simulation and observe three numbers, you will need to record how many times a number representing an answered call appears. Run your simulation 10 times.

23. For your 10 simulations, how many times were 0, 1, 2, and 3 calls answered by customer service?

24. In your class, how many times were 0, 1, 2, and 3 calls answered by customer service?

Mr. Smith wonders whether only one answered call out of three calls is sufficient evidence to be concerned. If so zero out of three answered calls is even stronger evidence.

25. Did any member of the class observe either zero or one call being answered? How many times did this happen in the whole class? What proportion is this?

26. With regard to the 95% standard being met, what do you think Mr. Smith should conclude from AGB's three phone calls in which only one was answered at 8:10 a.m.?

In order apply probability to this context, we will introduce set notation and Venn diagrams. We introduce the event notation as follows:

M = call answered on Monday at 8:10 a.m.
T = call answered on Tuesday at 8:10 a.m.
W = call answered on Wednesday at 8:10 a.m.

M^c = call NOT answered on Monday at 8:10 a.m.
T^c = call NOT answered on Tuesday at 8:10 a.m.
W^c = call NOT answered on Wednesday at 8:10 a.m.

To develop the analysis, we begin by first considering just two days, Monday and Tuesday. Each circle corresponds to the event that an 8:10 a.m. call was answered on that day. The intersection does not mean that the days overlap. It refers to the event that the phone call was answered at 8:10 a.m. on both Monday and Tuesday.

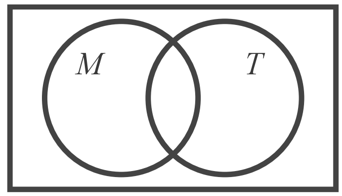


Figure 1.3.1: Venn diagram for calls 8:10 a.m. on Monday and Tuesday

The multiplication rule can be used to determine the probability of both events happening.

$$P(M \text{ and } T) = P(M \cap T) = P(M) \cdot P(T) = 0.95 \cdot 0.95 = 0.9025$$

Thus if the 8:10 a.m. standard were being met, there is more than 90% probability the phone would be answered on both days. The probability that the phone would be answered on at least one of the days corresponds to the ***union*** of these two events. The union of two sets, *A* and *B*, contains all the elements included either *A* or *B*. The formula for the number of elements in the union of two sets, *A* and *B,* is

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

In this formula $|A|$ represents the number of elements that belong to set *A*. The logic behind this formula is that adding the elements of the two sets together involves counting all elements in the intersection twice. The intersection is part of both sets *A* and *B*. This formula directly translates into a corresponding formula for the probability that the outcome will fall within the union of two sets.

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

In this context, the formula for the probability of union of *M* and *T* is

$$\begin{aligned} P(M \cup T) &= P(M) + P(T) - P(M \cap T) \\ &= 0.95 + 0.95 - (0.95 \cdot 0.95) \\ &= 0.9975 \end{aligned}.$$

Thus, we calculate the sum of the probabilities of falling within each individual set and subtract the probability of the intersection. The likelihood, that the call will be answered on at least one day (only Monday or only Tuesday or both Monday and Tuesday) is extremely high, almost 100%.

The probability of a call not being unanswered on a specific day is (1 – 0.95) or 0.05. Thus, if the 95% standard were generally being met, it is extremely unlikely that they did not reach an operator at 8:10 a.m.on both Monday and Tuesday.

$$\begin{aligned} P\left(M^c \cap T^c\right) &= P\left(M^c\right)\cdot P\left(T^c\right) \\ &= (0.05)(0.05) \qquad \text{or} \qquad \text{1 chance in 400} \\ &= 0.0025 \end{aligned}$$

This value could have also been determined by using the concept of complementary events. The complement to the event that the call was not answered on both Monday and Tuesday is the event that the call was answered on at least one of those days. That probability of that event was determined to be 0.9975. Thus

$$\begin{aligned} P\left(M^c \cap T^c\right) &= 1 - P\left(M \cup T\right) \\ &= 1 - 0.9975 \quad . \\ &= 0.0025 \end{aligned}$$

When tackling complex probability questions involving compound events, there is often more than one way to determine the probability.

Now let’s return to the three-day situation. First we will review the two extremes: calls are answered every day and calls are not answered each and every day.

$$\begin{aligned} P\left(M \cap T \cap W\right) &= P(M)\cdot P(T)\cdot P(W) \\ &= 0.95^3 \\ &\approx 0.857 \end{aligned}$$

Thus, if the standard were generally being met, the likelihood that calls were answered on three consecutive days is 0.857. However, there is still a significant probability, (1-0.857 = 0.143) that this will not happen, that at least one call will be missed. In contrast, it is extremely unlikely for no calls to be answered on three consecutive days if the 95% standard was generally being met.

$$\begin{aligned} P\left(M^c \cap T^c \cap W^c\right) &= P\left(M^c\right)\cdot P\left(T^c\right)\cdot P\left(W^c\right) \\ &= (1-0.95)^3 \qquad \text{or} \qquad \text{1 chance in 8,000} \\ &= 0.05^3 \\ &= 0.000125 \end{aligned}$$

The original question asked was, what can be concluded if a call was answered on only one day out of three. For now, the only way to calculate the probability involves detailed calculation. The event of interest can occur in any one of three mutually exclusive ways. The call was answered on Monday and not on Tuesday or Wednesday, $M \cap T^c \cap W^c$. Alternatively, the call was answered on Tuesday and not on Monday or Wednesday, $M^c \cap T \cap W^c$. Finally, the call was answered on

Wednesday but not on Monday or Tuesday, $M^c \cap T^c \cap W$. Because these patterns are mutually exclusive, the total probability equals the sum of the individual probabilities. The probability of each of these distinct outcomes is the same as calculated below:

$$P\left(M \cap T^c \cap W^c\right) = P\left(M\right) \cdot P\left(T^c\right) \cdot P\left(W^c\right) = 0.95\left(1-0.95\right)\left(1-0.95\right) = 0.002375$$

$$P\left(M^c \cap T \cap W^c\right) = P\left(M^c\right) \cdot P\left(T\right) \cdot P\left(W^c\right) = \left(1-0.95\right)0.95\left(1-0.95\right) = 0.002375$$

$$P\left(M^c \cap T^c \cap W\right) = P\left(M^c\right) \cdot P\left(T^c\right) \cdot P\left(W\right) = \left(1-0.95\right)\left(1-0.95\right)0.95 = 0.002375$$

Thus, the probability of calls being answered on only one day is three times 0.002375 or 0.0072. This probability is less than 1% and is considered highly unlikely. If this were the pattern observed, the conclusion would be that Koala Foods is generally not meeting the 95% standard for 8:10 a.m. The variability does not appear to simply be the result of random fluctuation. These observations suggest that the variability is due to a systemic problem that the management of Koala Foods needs to address.

Section 1.4 Getting *The Lancer* to Press

Each month Al Mitchell, the faculty advisor for a school newspaper, oversees the production of the newspaper, *The Lancer*. He has 10 student writers, five of whom are editors and five are staff writers. In order to get to press on time, it is necessary that Al's students finish their articles by the required deadline. That is, he needs each and every student to meet his or her deadline.

1.4.1 The Problem

Over the years, Al has found that that student writers meet their deadlines only 95% of the time. On the surface, this does not seem bad, but *The Lancer* frequently goes to press late. He wants to change that. Al recalls something from his high school probability class called the ***multiplication principle***: The probability that a set of ***independent*** events will occur can be found by multiplying the probabilities of each event occurring. Formally stated, the multiplication principle says that:

> If X and Y are *independent* events, then the probability of both X and Y occurring is equal to the product of the probabilities of X occurring and Y occurring.
>
> $$P(X \cap Y) = P(X) \cdot P(Y)$$

Al wonders what independent events would mean in this context. One particular writer meeting a deadline does not seem to affect the likelihood of any other writer meeting the deadline. Thus, the events that writers X and Y meet the deadline are *independent*. Therefore, the multiplication principle can be used. If each of his 10 writers meets the deadline 95% of the time, the probability that all of them meet the deadline is given below.

$$\begin{aligned} P(\text{going to press on time}) &= (0.95)(0.95)(0.95)(0.95)(0.95)(0.95)(0.95)(0.95)(0.95)(0.95) \\ &= 0.95^{10} \\ &\approx 0.599 \end{aligned}$$

Now, for the paper to go to press on time, every one of the 10 writers must meet the deadline. If even one of the ten does not meet the deadline, the printing of the paper is delayed. This calculation shows that *The Lancer* would go to press on time approximately 60% of the time! No wonder it has been going to press late so often. Al wants to change this situation to ensure the paper goes to press on time 90% of the time.

Each writer meeting deadlines 95% of the time seems to be a reasonable expectation. However, having each of the10 students meet the deadlines 95% of the time results in a big problem for *The Lancer*. To improve this situation, Al believes he can get his editors to meet the deadlines 99% of the time. They have the most experience and are likely to work better under pressure compared to the staff writers who have just joined *The Lancer*'s staff.

Al's idea is to have his five editors meet deadlines 99% of the time, and the less experienced staff writers can continue to meet deadlines 95% of the time. The multiplication principle allows Al to calculate the probability of going to press on time.

$$\begin{aligned} P(\text{going to press on time}) &= (0.99)(0.99)(0.99)(0.99)(0.99)(0.95)(0.95)(0.95)(0.95)(0.95) \\ &= (0.99)^5 \cdot (0.95)^5 \\ &\approx 0.736 \end{aligned}$$

Going to press on time about 74% of the time is far better than the 60% it was previously. Nevertheless, he would still not achieve his 90% goal.

Al goes back to the drawing board and believes that he can push his editors to be on time 100% of the time. He feels the editors can handle that level of responsibility. If the five editors meet the deadlines 100% of the time and the staff writers meet the deadlines 95% of the time, the probability of going to press on time is shown below.

$$\begin{aligned} P(\text{going to press on time}) &= (1)(1)(1)(1)(1)(0.95)(0.95)(0.95)(0.95)(0.95) \\ &= (1)^5 \cdot (0.95)^5 \\ &\approx 0.775 \end{aligned}$$

Getting the editors to meet their deadlines 100% of the time barely made a difference!

Al sees that the only way to get *The Lancer* to press on time is to push also the staff writers to meet their deadlines at a higher rate. If the editors meet the deadlines 100% of the time, what must the probability be for the staff writers in order for *The Lancer* to not be delayed 90% of the time?

1. If each staff writer meets her/his deadline 96% of the time, would The Lancer go to press on time 90% of the time? If not, what if the staff writers meet their deadlines 97% of the time?

Al begins to wonder whether it is reasonable to have a goal of 90% for *The Lancer* to go to press on time. If the editors meet their deadlines 100% of the time, Al wonders what standard should be required of each staff writer. We can take an algebraic approach to determine this standard.

Let x represent the standard for each staff writer meeting the deadline. The probability of the paper going to press on time at least 90% of the time can be represented by the following inequality. Then we need to solve the inequality for x.

$$
\begin{aligned}
P(\text{going to press on time}) &\geq 0.90 \\
(1)(1)(1)(1)(1)(x)(x)(x)(x)(x) &\geq 0.90 \\
1^5 \cdot x^5 &\geq 0.90 \\
x^5 &\geq 0.90 \\
x &\geq \sqrt[5]{0.90}
\end{aligned}
$$

Now, take the fifth-root of a number, which is equivalent to raising the number to the $\frac{1}{5}$ power. This can be done with a calculator, so

$$
\begin{aligned}
&x \geq (0.90)^{\frac{1}{5}} \\
&x \geq 0.979
\end{aligned}
$$

The solution shows us that for *The Lancer* to go to press on time 90% of the time, each of the five staff writers would have to make deadline almost 98% of the time! Al wonders if this is a realistic goal. In the next chapter, Al will evaluate a change in policy. He will consider going to press on time if one or two writers are late. He will simply leave their columns out of the paper.

1.4.2 Complementary Events

Let's consider the timely publication of *The Lancer* from a slightly different perspective. Instead of focusing on how frequently *The Lancer* goes to press on time, let's consider how often it goes to press late.

Going back to the original case, the editors and staff writers originally each met the individual deadlines 95% of the time. This led to *The Lancer* going to press on time roughly 60% of the time.

2. That being the case, how often did *The Lancer* go to press late?

If they were on time 60% of the time, it makes sense to say that they were late 40% of the time. Mathematically, this happens because going to press on time and going to press late are ***complementary events***. Complementary events are both ***mutually exclusive*** and ***collectively exhaustive***. If two events are mutually exclusive, that means they cannot occur at the same time. For example, when rolling a die, the event of getting a three and the event of getting a five are mutually exclusive. However, getting a 3 and getting an odd number are not mutually exclusive. Two events are collectively exhaustive if together they cover all possible outcomes. This implies that one of them must occur. For example, when rolling a die, the event of rolling an odd number and the event of rolling an even number are collectively exhaustive. However, rolling an odd number and rolling a 2 are not collectively exhaustive.

3. Keeping with our die-rolling context, list three pairs of complementary events.

When we consider publishing a student newspaper, one pair of complementary events is publishing on time or not publishing on time (i.e., publishing late). Because an event and its complement include all the possible outcomes of an event, the probability that one or the other of them will occur is equal to one (i.e., one of them is certain to occur). Now, because they cannot both occur (why not?), you can find the probability of an event's complement by subtracting the probability of the original event from one. This is significant because calculating the probability of the complement of an event is sometimes easier than finding the probability of the event directly. Thus, in the case of *The Lancer*,

$$P(\text{publishing on time}) + P(\text{not publishing on time}) = 1$$
$$P(\text{not publishing on time}) = 1 - P(\text{publishing on time})$$

You might think that all of the writers on time and all of the writers not on time are complementary events. All of the writers on time corresponds directly to the event that all ten writers meet their individual deadlines. The complementary event that the paper is late is not equivalent to ***all*** writers being late. It only takes ***one*** writer being late for publication of *The Lancer* to be delayed. Thus the event the paper is late includes all of the following possibilities: 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 writers being late. Calculating directly the probability of the paper being late requires calculating the probability of each of these one through ten possibilities. The calculation of each of these is in itself complex and requires the advanced concepts discussed in the next chapter.

Section 1.5 Worker Absenteeism at BT Auto Industries

Unscheduled absenteeism is the bane of any manager of a manufacturing or service facility. When a worker is absent either some work does not get done, or other workers have to increase their productivity. Alternatively, substitute workers must be called in. A classroom is good example in the lost value when a teacher is absent and a substitute fills in.

1.5.1 Machine Repair Workers

According to The Centers for Disease Control and Prevention (CDC), the flu season usually ranges from November through March. Historically, the peaks usually occur sometime during a 100 day period from December to early March, with flu outbreaks cascading across the country. The CDC recommends that anyone who contracts the flu stay home to prevent spreading it to others. However, a recent study at the Children's Hospital of Philadelphia documented that most medical professionals at that facility do not follow this advice in their workplace. (JAMA Pediatr. 2015;169(9):815-821)

BT Auto Industries is a small manufacturer based in Detroit, Michigan. One particular piece of equipment is critical to its operation. When the equipment fails and cannot be repaired quickly, production is completely stopped. If the stoppage is long, workers are sent home in the middle of the day. Repair of the equipment requires special training and skills. Currently, two workers in the factory, Alejandro and Bernice have undergone extensive training. They are the only employees who can repair this equipment. David Plante, director of operations, is concerned about the upcoming flu season. There is an estimated 10% probability that a worker will be absent on a particular day due to the flu. His primary concern is the possibility that both Alejandro and Bernice would be absent on the same day.

Let A = the event that Alejandro is not sick and is at work
B = the event that Bernice is not sick and is at work

The complements of these events are
A^c = the event that Alejandro is out sick
B^c = the event that Bernice is out sick

The director of operations quickly estimated the likelihood that both would be absent on the same day. He assumes that the absences are independent events. Thus the multiplication rule can be applied.

$$\begin{aligned} P\left(A^c \cap B^c\right) &= P\left(A^c\right) \cdot P\left(B^c\right) \\ &= (0.1)(0.1) \\ &= 0.01 \end{aligned}$$

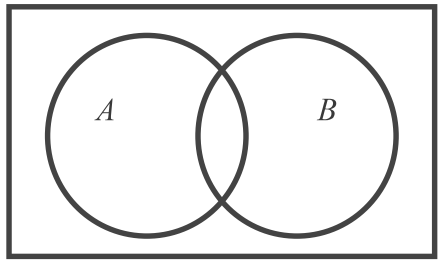


Figure 1.5.1: Venn diagram for Alejandro's and Bernice's presence at work

At first glance, his initial reaction is there is not much to worry about. The probability is only 0.01 or one in 100. However, his production supervisor, Amadeus Wolfe quickly points out that the flu season is 100 days long. Thus, on average there will be one day per season when both Alejandro and Bernice will be absent. In addition, Amadeus noted that some of the more difficult repairs were better handled by two people. In those instances the two working together could repair the equipment more than twice as fast as a single individual. The probability that both are at work is 0.81.

$$\begin{aligned} P(A \cap B) &= P(A) \cdot P(B) \\ &= (0.9)(0.9) \\ &= 0.81 \end{aligned}$$

On average, they will both be at work more than four days out of five. The likelihood that only one will be at work involves considering two mutually exclusive alternatives. Alejandro is at work and Bernice is out with the flu or Alejandro is out with the flu and Bernice is at work. Because these are mutually exclusive events their probabilities can be added.

$$\begin{aligned} P(\text{exactly one of them is at work}) &= P(A \cap B^c) + P(A^c \cap B) \\ &= (0.9)(0.1) + (0.1)(0.9) \\ &= 0.18 \end{aligned}$$

To double-check his calculation, he thinks about the three possible outcomes: both are at work, only one is at work and neither is at work. Since this is the set of all possible outcomes, their probabilities should sum to one.

1. Confirm that the probabilities of these three mutually exclusive events do sum to one.

David asks Amadeus to determine the value of investing \$1,000 in a week's worth of training for a third individual named Carmi. To do so, he first calculates the probability that all three would be absent from work on the same day. Again assuming that the three absences are independent events,

$$P(A^c \cap B^c \cap C^c) = P(A^c) \cdot P(B^c) \cdot P(C^c) = (0.1)^3 = 0.001$$, or one in a thousand.

Training a third individual reduces the risk of having no one available to fix the machine to an acceptable risk of once in a 1,000 days. Next, Amadeus tries to calculate the likelihood that at least two people are available to work together on repairing the equipment. There are two ways to approach this calculation. The more direct way is for Amadeus to calculate the probability that exactly two are on duty and add it to the probability that all three are at work.

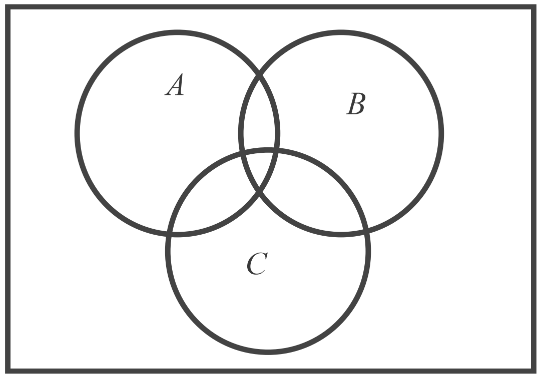


Figure 1.5.2: Venn diagram for Alejandro, Bernice, and Carmi

Let X = number of trained employees on duty.

$$P(X \geq 2) = P(X = 2) + P(X = 3)$$

$$\begin{aligned} P(X = 2) &= P(A \cap B \cap C^c) + P(A \cap B^c \cap C) + P(A^c \cap B \cap C) \\ &= (0.9)(0.9)(0.1) + (0.9)(0.1)(0.9) + (0.1)(0.9)(0.9) \\ &= 3\left[(0.9)^2(0.1)\right] \\ &= 3(0.081) \\ &= 0.243 \end{aligned}$$

$$\begin{aligned} P(X = 3) &= P(A \cap B \cap C) \\ &= (0.9)(0.9)(0.9) \\ &= (0.9)^3 \\ &= 0.729 \end{aligned}$$

$$\begin{aligned} P(X \geq 2) &= P(X = 2) + P(X = 3) \\ &= 0.243 + 0.729 \\ &= 0.972 \end{aligned}$$

With only two workers, the probability of two being at work was 0.81. However, with the added worker, the probability of at least two workers present increases to 0.972.

Amadeus wondered if it would have been easier to work with the complements to determine this probability.

2. Use the following equation to calculate the same probability.

$$P(X \geq 2) = 1 - P(X \leq 1)$$
$$= 1 - \left[P(X=0) + P(X=1)\right]$$

Dionne Trump had listened to the discussions between Amadeus and David. She believed they significantly overestimated the risk because they left out one important factor. They did not include the frequency of the equipment breaking down. If the equipment failed frequently, one or more times per day, then their concerns made sense. However, she was not sure how reliable the equipment was.

3. How would the above analysis change if the equipment broke down on average only once every five days?

4. Would it make sense to increase the maintenance budget by $1,000 so that the frequency of breakdowns was reduced to one in 20 days?

1.5.2 Simulating Day-to-Day Operations

David and Amadeus decide to turn their attention to day-to-day operations. The production floor has 12 workers. In particular, an average of one out of every 10 workers is absent every day during the flu season! In order to maintain their level of productivity, BT Auto Industries hires spare workers to cover the shifts of the employees who have called in sick. Due to tight budget constraints this year, the management wants to hire just enough spare workers to achieve the probability that 95% of the time the absent workers are covered. Anything less and productivity is impacted.

During December, they experienced a number of worker shortages but they were not overly concerned. The demand for their products generally drops during the pre-holiday season. However, they are concerned about the first 10 weeks of the new year when demand picks up. Derek Hall, the plant manager, is interested in determining the number of spare workers to hire on standby. He asked Donald Champ, an industrial engineer, to run a computer simulation of the number of workers who show up each day. The simulation will use random numbers to replicate each worker's attendance during the 10-week peak flu season. They will need to simulate the attendance of each of the 12 workers on each of 50 work days.

An analysis of attendance records for the past five years indicates that an average of one in 10 workers is absent every day during these 10 weeks. They will, therefore, use a 10% absenteeism rate for each of the 12 workers. They generated random integer values between 1 and 10. They assigned a value of 1 to mean that the worker was absent on a given day. An integer between 2 and 10 represented a worker who was present that day.

Table 1.5.1 illustrates the simulation of three days. Each of the random numbers takes on a value between 1 and 10. Table 1.5.2 illustrates how these numbers are then translated into whether or not a worker is absent or present. On Day 1, the number 1 appears only under worker A. There was a total of one absence that day. On Day 2, the number 1 appears only under worker C; thus, there was one worker absent that day. On Day 3, the simulation generated the number 1 under workers E, I, and L. There were three workers absent on Day 3.

	Workers												Total Absent each day
Day	A	B	C	D	E	F	G	H	I	J	K	L	
1	1	6	6	7	2	8	8	4	9	6	9	9	1
2	3	8	1	4	7	9	9	8	8	10	3	10	1
3	6	10	7	6	1	4	2	5	1	3	7	1	3

Table 1.5.1: Simulation of absenteeism on first three days

	Workers												Total Absent Each day
Day	A	B	C	D	E	F	G	H	I	J	K	L	
1	A	P	P	P	P	P	P	P	P	P	P	P	1
2	P	P	A	P	P	P	P	P	P	P	P	P	1
3	P	P	P	P	A	P	P	P	A	P	P	A	3

Table 1.5.2: Translate numbers into Absent (A) or Present (P)

Table 1.5.3 provides the detailed simulated attendance data for each worker on each day for 50 days. It also includes daily totals and individual worker totals for the simulated 10-week period of 50 work days.

If you were to use a calculator to generate data as in Table 1.5.1, you would have to generate the random numbers in stages. The screen of most calculators cannot display more than five numbers at a time. You would then have to record the information in a table. Afterwards, you would identify which values correspond to absent and present. This would be a tedious task. In contrast, Excel can easily simulate 50 days for 12 workers and produce Table 1.5.2 directly. This is demonstrated in section 1.5.3.

	Workers												
Day	A	B	C	D	E	F	G	H	I	J	K	L	**Total absent each day**
1	1	6	6	7	2	8	8	4	9	6	9	9	**1**
2	3	8	1	4	7	9	9	8	8	10	3	10	**1**
3	6	10	7	6	1	4	2	5	1	3	7	1	**3**
4	10	5	3	7	9	8	9	6	9	9	7	1	**1**
5	9	10	8	2	5	4	6	5	9	10	3	5	**0**
6	10	7	5	1	2	2	1	4	6	6	10	3	**2**
7	8	4	9	4	8	5	3	7	10	10	5	9	**0**
8	9	5	6	8	2	8	4	10	9	10	7	7	**0**
9	1	8	2	9	2	3	2	4	6	10	8	7	**1**
10	6	4	7	4	10	9	2	4	7	10	6	5	**0**
11	6	4	2	6	7	2	3	5	9	10	4	3	**0**
12	6	1	8	4	9	6	10	7	9	7	9	8	**1**
13	9	7	3	9	8	5	9	5	6	6	3	10	**0**
14	3	8	4	6	7	8	10	9	1	5	1	8	**2**
15	4	9	5	2	7	5	5	5	8	5	1	5	**1**
16	5	8	2	10	9	3	7	3	6	2	6	1	**1**
17	7	10	10	4	4	3	6	3	7	8	7	9	**0**
18	6	7	9	2	2	6	8	8	5	9	7	2	**0**
19	9	9	8	3	10	7	7	10	9	1	3	4	**1**
20	3	9	10	6	7	4	10	8	9	3	9	8	**0**
21	8	5	8	7	6	1	4	4	8	2	9	7	**1**
22	9	4	7	6	8	9	2	10	10	7	8	7	**0**
23	10	2	7	1	4	9	8	1	5	4	9	4	**2**
24	2	1	5	7	3	1	5	3	4	1	8	4	**3**
25	5	8	7	9	1	10	9	6	7	5	7	2	**1**
26	3	2	5	5	2	6	8	5	3	9	9	3	**0**
27	10	2	8	3	1	10	5	5	9	3	1	2	**2**
28	4	2	6	9	7	3	7	2	6	7	2	1	**1**
29	1	2	3	1	10	10	10	4	2	9	2	7	**2**
30	9	3	5	8	8	1	4	1	10	3	4	5	**2**
31	3	2	8	1	3	1	5	7	8	1	1	7	**4**
32	10	3	5	10	6	3	6	2	7	5	1	6	**1**
33	5	1	9	10	6	10	6	1	10	6	2	10	**2**
34	3	2	10	7	9	2	8	9	4	4	1	5	**1**
35	4	9	5	7	1	8	2	4	5	6	4	8	**1**
36	10	3	9	2	9	3	6	3	1	3	6	6	**1**
37	5	6	4	4	6	4	7	3	4	2	3	6	**0**
38	1	6	9	5	7	4	8	10	8	4	7	6	**1**
39	7	4	1	8	4	6	3	9	4	9	2	6	**1**
40	8	9	5	10	1	5	1	1	6	8	10	4	**3**
41	6	9	8	7	7	5	10	1	3	7	5	1	**2**
42	5	2	3	9	9	6	7	5	3	9	6	9	**0**
43	9	6	3	10	5	7	2	8	9	1	6	8	**1**
44	6	5	1	7	4	4	1	7	10	2	9	3	**2**
45	8	7	4	1	6	7	8	5	1	2	7	9	**2**
46	8	7	7	3	1	7	5	5	8	3	6	6	**1**
47	10	1	5	3	3	5	5	7	7	2	4	1	**2**
48	1	6	6	5	8	7	3	4	8	10	4	6	**1**
49	2	4	5	2	1	3	10	4	5	4	8	8	**1**
50	5	1	6	10	10	6	4	6	3	6	9	9	**1**
Absent	**5**	**5**	**3**	**5**	**7**	**4**	**3**	**5**	**4**	**4**	**6**	**6**	

1 = absent; 2 through 10 = present

Table 1.5.3: Fifty daily results of computer simulation of worker absenteeism

Table 1.5.4 summarizes the frequency distribution of the number of absent workers for the 50-day simulation. The absolute frequency is converted into the relative frequency by dividing each value by 50, the number of simulated workdays. For example, on 13 out of 50 days, there were no workers absent for a relative frequency of 26%. The cumulative relative frequency sums the percentages up to the corresponding value. For example, the cumulative relative frequency of having 3 or fewer absent workers was 98%. This is the sum of the percentages for 0, 1, 2, or 3 absent workers. Use the information in the Table 1.5.3 and 1.5.4 to answer the following questions.

Absent Workers	**Absolute Frequency**	**Relative Frequency**	**Cumulative Relative Frequency**
0	13	26%	26%
1	22	44%	70%
2	11	22%	92%
3	3	6%	98%
4	1	2%	100%
5	0	0%	100%

Table 1.5.4 Frequency distribution of the simulated number of absent workers

5. For what percentage of the 50 simulated days would one spare worker be enough?

6. For what percentage of the 50 simulated days would two spare workers be enough?

7. If the goal is to have enough spare workers 95% of the time, based on this simulation how many spare workers are needed?

8. If the goal is to have enough spare workers 99% of the time, based on this simulation how many spare workers are needed?

9. Can you explain why there are gaps in the actual percentages in questions 5-8?

10. Based on the Table 1.5.3, what is each worker's actual rate of absenteeism? What is the minimum percentage rate and what is the maximum?

11. Why is each worker's actual rate not equal to 10%?

12. What is the overall average absenteeism rate for the 12 workers? Why is this rate closer to the one-in-ten average as compared to many of the individual workers?

13. In setting up the simulation, what assumption have we made about each worker's rate of absenteeism? Do you think the assumption is valid?

1.5.3 Creating Your Own Simulation

Because this is a random process, the results from repeated simulations can be different. To be confident about their spare worker policy, BT Auto's management decides to replicate this 50-day period many times over. They also would like to understand how much variability there might be from year to year in achieving their 95% goal.

Each member of the class will create a spreadsheet to simulate the scenario of a 12-person workforce with a 10% absenteeism rate. The model will show a 50-day period. (If you are reading the text on your own, we have provided an Excel file worksheet with simulated data for an imaginary group of classmates.)

Open Excel and create a table with a column labeled for each worker and a row for each day. Use row "1" to assign an identifying number to each workers. Use column "A" to number the days. Within the table, in the cell for your first worker on the first day, you need to enter a formula whose output will indicate whether that worker was absent on that day. The formula is similar to the "randInt(" command on the graphing calculator. In Excel the formula is "RANDBETWEEN." As you might imagine, this formula will return an integer value between any two integers you input.

For the situation where each worker is absent 10% of the time, we need to think of the absentee rate as a probability. A 10% absenteeism rate means that on any given day a worker may be absent with a probability of one out of 10. Thinking about it this way, we can choose a random number between 1 and 10. Assign an output of "1" to represent a worker being absent. The outputs "2" through "10" represent a worker being present.

To do this, enter an equal sign (=) and the following formula into the appropriate cell.

=RANDBETWEEN(1,10)

Once this formula is entered, an integer between 1 and 10 should appear in the cell. To fill the rest of your spreadsheet, first drag the formula across the row and then down the column to copy it into the remaining cells.

Your spreadsheet should now be populated with integers from 1 to 10 for each of your 12 workers (columns labeled B through M) for 50 days (rows). You will need to prevent these randomly generated numbers from constantly changing. To do this, copy all of the random values using a *copy*, *paste special*, *value* command. You should paste these values into the same cell locations the random numbers originally appear in.

Now we need to determine how many workers were absent each day and how many times each worker was absent during the 50 days. Excel has a formula that enables us to determine this statistic.

The "COUNTIF" formula will search a range of cells and count how many times a particular value occurs. We are concerned with absences, which are indicated by a "1" in a cell. We can give Excel a command to search each row and each column for ones. It will report back how many times the number “1” occurred. You need to include in the formula the range of cells to search and the value it should count.

For example, assume the record of each of the 12 workers on Day 1 is in cells B6 through M6. The appropriate command to count all “1”s is

=COUNTIF(B6:M6, 1)

The resultant count will represent the number of workers absent on day “1”. Drag the formula down the column to obtain a value for each of the 50 days.

The resultant count will represent the number of times this worker was absent during the 50 day period. Drag the formula across the row to obtain a value for each worker.

We apply the same type of formula for each column to determine the number of times a specific worker was absent during the 50 days. For example, if the 50 days of data for worker A are in cells B6 through B55, the appropriate command to count all “1”s is

=COUNTIF(B6:B55, 1)

Drag the formula across the row to obtain a value for each worker.

14. Compare your row and column totals with those of your neighbors. What do you notice?

15. What could explain the similarities and differences you noted?

16. Can you find anyone else with a simulation spreadsheet that is exactly like yours?

Managers at BT Auto with different responsibilities will look at the data presented in Table 1.5.3 from diverse perspectives. The plant manager is primarily concerned with getting the work done. This official is most interested in how many workers showed up for work. This assumes that all of the workers can do all of the jobs and there is no specialization. The manager will, therefore, focus on the row totals. In a later section, we take the plant manager’s perspective and explore the randomness in the number of workers who were absent. This analysis is critical to determining the need and cost of having spare workers to cover for absentees.

The payroll manager is concerned with the details as to who showed up for work each day. He has to submit that information to ensure each individual is paid appropriately. In addition, the worker’s absences are recorded against his sick leave allowance. The Human Resource (HR) manager may be interested in the pattern of absences for individual workers. Is a specific worker absent an unusual amount of time? Is another worker always there and perhaps eligible for a bonus? The HR manager will be interested in the column totals for each worker. In the next section we look at the variability in the simulated data from the perspective of each worker’s record.

Before proceeding let's explore your gut instinct. Since there is 0.1 probability a worker will be absent on any given day, workers will be absent on average five days out of 50.

17. However, what do you think the chances are that a worker will not miss a single day of the 50 days? Is it one in 10? One in 100? One in 1,000? Or less than one in 1,000?

18. Also, what do you think the chances are that a specific worker will be absent 10 or more days during the flu season?

19. Similarly, what do you think the chances are that all 12 workers will show up for work on a given day?

1.5.4 Individual Worker Absences – Your Individual and Class Pooled Data

The Assistant Vice President of Human Resources, Charlene Fine, is interested in identifying the most reliable workers. These are workers who show up every day no matter what. She is also interested in monitoring those workers who are frequently out sick.

20. From the column totals, identify the minimum and maximum number of days a worker was absent. How many workers were not absent even once during a 50 day period in the spreadsheet in Table 1.5.3? What about the workers in your spreadsheet?

21. Which worker in your simulation had the best attendance record? How much better than the average of five absences was that worker?

22. Which worker had the worst attendance record? How much worse was that worker than the average?

23. Compare your answers with your neighbors. Were the workers with the best and worst attendance records the same for you and your neighbors?

24. What was the least number of days a worker was absent among in your small group? What was the most number of days a worker was absent?

25. Justify whether or not the worker with best attendance in the simulated dataset should be rewarded for good attendance.

You are going to use the class values to develop a frequency distribution table of the number of days a worker was absent during the 50-day time period. You will first do this just for your own data. Then the class will pool all the data to obtain a larger sample. This will be a better estimate of the relative frequency of different values. Then the percentages will be accumulated in order to create a cumulative relative frequency.

Identify the worker(s) with lowest number of absences for the class. This is the class minimum. Identify the worker(s) with highest number of absences for the class. This is the class maximum.

Create a sequence of integer values beginning at the class minimum and ending at the class maximum. Next to each integer value, record the number of times in ***your*** spreadsheet that each value appeared in the column totals.

26. How many workers in your spreadsheet were absent the class minimum and class maximum number of days?

27. Now you will pool the frequency distribution results for the whole class and then determine the *relative frequency* for the whole class. Find the observed relative frequency for each value for each student and record it in a class pooled table. To start, count the total number of times the minimum value appeared in the pooled data. To obtain the relative frequency, divide this by 12 times the number of students in the class. The 12 represents the number of workers in each spreadsheet. Repeat this for every value in the table to determine its relative frequency.

28. Create another column for cumulative relative frequency. To determine, for example the cumulative relative frequency of five or fewer absences, sum the relative frequency values for zero through five.

29. For class pooled data, what is the relative frequency of a worker being present each and every day during a 50-day period? How does this compare to your gut feel estimate noted in question 17?

30. In the class pooled data, how many times was a worker absent exactly once during a 50-day period? What is its relative frequency?

31. In the class pooled data, how many times was a worker absent one or fewer times during a 50-day period? What is its cumulative relative frequency?

32. In the class pooled data, how many times was a worker absent five or fewer times during a 50-day period? What is its cumulative relative frequency?

33. In the class pooled data, how many times was a worker absent nine or fewer times during a 50-day period? What is its cumulative relative frequency?

34. In the class pooled data, how many times was a worker absent 10 or more times during a 50-day period? What is its relative frequency? How does this relate to the answer to question 33? How does this compare to your gut feel estimate noted in question 18?

1.5.5 Total Number of Workers Absent Each Day

In this section, you will take the plant manager's perspective. The focus will be on the frequency distribution of the number of workers absent each day.

35. In your spreadsheet look at the row totals. What were the minimum and maximum number workers that were absent?

36. How many days experienced no absences? How does this compare to your gut feel estimate noted in question 19?

37. What are the minimum and maximum values for the whole class?

You are going to use the class minimum and maximum values to develop a frequency distribution table of the number of workers absent during each day of the 50-day time period. You will first do this just for your own data. Then the class will pool all the data to obtain larger sample. This is a better estimate of the relative frequency of different values. The relative frequency distribution will then be used to create a cumulative relative frequency distribution.

Create a sequence of integer values beginning at the class minimum and ending at the class maximum. Next to each integer value, record the number of times in your spreadsheet that each value appeared in the row totals. Calculate the relative frequency by dividing by 50.

38. In your row totals, how many days were exactly zero workers absent during the 50-day period?

39. Now tabulate the results for the whole class. How many days was exactly zero workers absent during the 50-day period? For the class as a whole, the total number of simulated days is 50 times the number of students who participated in the activity. If there are 20 students in the class, 1,000 days have been simulated. Divide this total of days with zero absentees by the class total number of simulated days to obtain the relative frequency of observing zero absentees.

40. Repeat this calculation of relative frequency for each value ranging from the class minimum to the class maximum.

41. Also calculate the cumulative relative frequency from the minimum to the maximum observed. The cumulative relative frequency for any value below the minimum is zero. The cumulative relative frequency for any value equal to the maximum or higher is one.

42. How do these relative frequencies compare to the data in Table 1.5.4?

1.5.6 Cost of Spare Workers

The company is considering hiring three steady spare workers to show up each day for work. They are on standby to replace absent workers. To encourage them to show up for work, the company established a compensation policy. These standby workers are paid $55 just for coming to work even if they do no work. However, if the individual actually works, the worker is paid an additional $70, for a total of $125.

Assume three spare workers showed up for work.

43. What is the cost to the company on a day that only one worker was absent. Just one spare worker was needed to work and the other two were sent home?

44. What is the cost to the company on a day that two workers were needed and only one was sent home?

45. On any day when the number of absent workers is three or more, what is the cost for standby workers that day?

46. In general let N be the number of standby workers who actually work. With three workers on regular standby, write an equation for the total cost as a function of N.

In Table 1.5.5 we added a column that includes the cost for standby workers for each day in our simulation reported in Table 1.5.3. It is the next to last column on the right.

47. What is the simulated total 50-day cost for the current policy of three spares on standby? What is the average daily cost?

1.5.7 New Policy: Pay Guarantee

Management is considering paying one worker $100 per day every day no matter what because most days there is at least one worker absent. The other two workers are paid as before, either $55 or $125. If the first worker is not needed for the production line, he is given a special task such as cleaning up some set of files or counting inventory. It is not a high-value job but it keeps him busy. We have added another column to our original table and recorded each day's cost for this new policy.

48. On days with zero absences, what does this new policy cost? Does it cost more or less than before?

49. On days with one absence, how much money does this new policy cost? Is it more or less than before?

50. On days with two absences, how much money does this new policy cost? Three absences?

51. When does the policy cost more money and when does it save money?

52. What is the most common dollar amount under the new policy? Explain why it occurs so often.

53. Based on Table 1.5.5, how many days out of the 50 did the new policy save money?

54. Would you recommend the new policy or the current policy? Explain your answer.

Day	Workers												Absent each day	Cost	
	A	B	C	D	E	F	G	H	I	J	K	L		Current	New
1	1	6	6	7	2	8	8	4	9	6	9	9	**1**	$235	$210
2	3	8	1	4	7	9	9	8	8	10	3	10	**1**	$235	$210
3	6	10	7	6	1	4	2	5	1	3	7	1	**3**	$375	$350
4	10	5	3	7	9	8	9	6	9	9	7	1	**1**	$235	$210
5	9	10	8	2	5	4	6	5	9	10	3	5	**0**	$165	$210
6	10	7	5	1	2	2	1	4	6	6	10	3	**2**	$305	$280
7	8	4	9	4	8	5	3	7	10	10	5	9	**0**	$165	$210
8	9	5	6	8	2	8	4	10	9	10	7	7	**0**	$165	$210
9	1	8	2	9	2	3	2	4	6	10	8	7	**1**	$235	$210
10	6	4	7	4	10	9	2	4	7	10	6	5	**0**	$165	$210
11	6	4	2	6	7	2	3	5	9	10	4	3	**0**	$165	$210
12	6	1	8	4	9	6	10	7	9	7	9	8	**1**	$235	$210
13	9	7	3	9	8	5	9	5	6	6	3	10	**0**	$165	$210
14	3	8	4	6	7	8	10	9	1	5	1	8	**2**	$305	$280
15	4	9	5	2	7	5	5	5	8	5	1	5	**1**	$235	$210
16	5	8	2	10	9	3	7	3	6	2	6	1	**1**	$235	$210
17	7	10	10	4	4	3	6	3	7	8	7	9	**0**	$165	$210
18	6	7	9	2	2	6	8	8	5	9	7	2	**0**	$165	$210
19	9	9	8	3	10	7	7	10	9	1	3	4	**1**	$235	$210
20	3	9	10	6	7	4	10	8	9	3	9	8	**0**	$165	$210
21	8	5	8	7	6	1	4	4	8	2	9	7	**1**	$235	$210
22	9	4	7	6	8	9	2	10	10	7	8	7	**0**	$165	$210
23	10	2	7	1	4	9	8	1	5	4	9	4	**2**	$305	$280
24	2	1	5	7	3	1	5	3	4	1	8	4	**3**	$375	$350
25	5	8	7	9	1	10	9	6	7	5	7	2	**1**	$235	$210
26	3	2	5	5	2	6	8	5	3	9	9	3	**0**	$165	$210
27	10	2	8	3	1	10	5	5	9	3	1	2	**2**	$305	$280
28	4	2	6	9	7	3	7	2	6	7	2	1	**1**	$235	$210
29	1	2	3	1	10	10	10	4	2	9	2	7	**2**	$305	$280
30	9	3	5	8	8	1	4	1	10	3	4	5	**2**	$305	$280
31	3	2	8	1	3	1	5	7	8	1	1	7	**4**	$445	$420
32	10	3	5	10	6	3	6	2	7	5	1	6	**1**	$235	$210
33	5	1	9	10	6	10	6	1	10	6	2	10	**2**	$305	$280
34	3	2	10	7	9	2	8	9	4	4	1	5	**1**	$235	$210
35	4	9	5	7	1	8	2	4	5	6	4	8	**1**	$235	$210
36	10	3	9	2	9	3	6	3	1	3	6	6	**1**	$235	$210
37	5	6	4	4	6	4	7	3	4	2	3	6	**0**	$165	$210
38	1	6	9	5	7	4	8	10	8	4	7	6	**1**	$235	$210
39	7	4	1	8	4	6	3	9	4	9	2	6	**1**	$235	$210
40	8	9	5	10	1	5	1	1	6	8	10	4	**3**	$375	$350
41	6	9	8	7	7	5	10	1	3	7	5	1	**2**	$305	$280
42	5	2	3	9	9	6	7	5	3	9	6	9	**0**	$165	$210
43	9	6	3	10	5	7	2	8	9	1	6	8	**1**	$235	$210
44	6	5	1	7	4	4	1	7	10	2	9	3	**2**	$305	$280
45	8	7	4	1	6	7	8	5	1	2	7	9	**2**	$305	$280
46	8	7	7	3	1	7	5	5	8	3	6	6	**1**	$235	$210
47	10	1	5	3	3	5	5	7	7	2	4	1	**2**	$305	$280
48	1	6	6	5	8	7	3	4	8	10	4	6	**1**	$235	$210
49	2	4	5	2	1	3	10	4	5	4	8	8	**1**	$235	$210
50	5	1	6	10	10	6	4	6	3	6	9	9	**1**	$235	$210
Absent	**5**	**5**	**3**	**5**	**7**	**4**	**3**	**5**	**4**	**4**	**6**	**6**		**$12,240**	**$11,900**

Table 1.5.5: Workers' absences and standby worker costs

Chapter 1 (Basic Probability) Homework Questions

Complement

1. Two Complementary events
   a. Describe complementary events with regard to the outcome of a baseball game.
   b. Be careful to describe complementary events with regard to the outcome of a football or soccer game. How is this answer different than the answer to a?
   c. Describe complementary events with regard to performance on a test.
   d. Describe complementary events with regard to the launch of the Space Shuttle.

2. Provide examples of events that are mutually exclusive but are not complementary in the context of:
   a. The outcome of a soccer game or college football game.
   b. Performance on a test
   c. Timeliness of completing a project

Simulation

3. Simulate by hand flip coin – Two coins
   a. Flip two coins. List all of the possible outcomes of the experiment.
   b. In the coin flip experiment if you only record the number of heads, list all of the possible outcomes of the experiment?
   c. Provide an example of complementary events with regard to the outcome of flipping two coins
   d. Flip two coins 10 times. Record your results into the Table 1 below. (Number of Heads equals 0 means experiment resulted in all Tails.)

Number of Heads	Experiment										Absolute Frequency	Relative Frequency	Cumulative Relative Frequency
	1	2	3	4	5	6	7	8	9	10			
0													
1													
2													

Table 1: Record 2 coin flip experiment

   e. How many of the times did you get one Head and one Tail in 10 experiments? Write your answer into the corresponding row under the absolute frequency column.
   f. Write the frequency of other possibilities into the table.
   g. Calculate the relative frequency of each possibility and compare relative frequency of each outcome with two of your friends.
   h. Calculate the cumulative relative frequency of each possibility and compare the cumulative relative frequency of each outcome with two of your friends.

4. Simulate by calculator flip coin – Two coins
    a. Go to the APPS menu of your calculator, press 0 for probability simulation, and then press 1 to toss coins. Press on the SET button: change the number of trial set to 50, change the number of coins to 2, and press the OK button. Now, press on the TOSS button. You will see a graph with bars showing the frequency of each possible outcome. What do you observe in the graph?

    b. Now, go to the DATA menu and count the frequency of each outcome.
    c. What are the relative frequencies of each outcome? Compare your results with your friends.

5. Calculate Probabilities – Use Multiplication Rule (Two coins)
    a. Calculate the probability of two heads when two coins are flipped
    b. Calculate the probability of zero heads when two coins are flipped
    c. Use complement to calculate the probability of exactly one head when two coins are flipped
    d. Compare the theoretical probabilities with the observed relative frequencies you found in questions 3 and 4.

6. Simulate by hand flip coin – Three coins
    a. Flip three coins. List all of the possible outcomes of the experiment.
    b. If you only record the number of heads, list all of the possible outcomes of the experiment?
    c. In the three coin flip experiment, you can have either all (three) Heads or less than three (zero, one, or two) Heads. The probability of having three Heads or less than three Heads add up to one. What are the other complementary events with regard to the outcome of flipping three coins?
    d. In the three coin flip experiment, one Head and two Tails and three Heads and zero Tails are mutually exclusive events and they are not complementary. Give an example of two mutually exclusive events that are not complementary.
    e. Flip three coins 10 times. After each flip, record the number of heads. Next, in Table 2 below record a 1corresponding to the number of heads observed on that experiment.

Number of Heads	**Experiment**										**Absolute Frequency**	**Relative Frequency**	**Cumulative Relative Frequency**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**			
0													
1													
2													
3													

Table 2: Record 3 coin flip experiment

f. Count and record the number of times the outcome is all Heads and its complement. Calculate the relative frequency. Do relative frequencies add to one?
g. What is the complementary event to observing all Heads? Calculate its relative frequency. How does the answer relate to the value obtained in the previous question?
h. Determine the total frequency of for each value: 0, 1, 2, and 3.
i. Calculate relative frequency of each value.
j. Calculate the cumulative relative frequency. What is the relative frequency of obtaining 2 or fewer heads?
k. From the table above look at the relative frequencies of each outcome. Do they add up to 1?
l. In part d above, you created two mutually exclusive events. What are their relative frequencies? Do these relative frequencies add up to one?
m. The probability of getting only one Head or only one Tail in flipping three fair coins must be equal. Is your result a representation of this statement?

7. Simulate by calculator flip coin – Three coins
   a. Go to the APPS menu of your calculator again. Press 0 for probability simulation, and then press 1 to toss coins. Press on the SET button: change the number of trial set to 50, change the number of coins to 3, and press the OK button. Now, press on the TOSS button. You will see a graph with bars showing the frequency of each possible outcome. What do you observe in the graph?
   b. Now, go to the DATA menu and count the frequency of each outcome.
   c. What are the relative frequencies of each outcome? Compare your results with your friends.
   d. What are the cumulative relative frequencies of each outcome? Compare your results with your friends.

8. Calculate Probabilities – Use Multiplication Rule (Three coins)
   a. Calculate the probability of three heads when three coins are flipped
   b. Calculate the probability of zero heads when three coins are flipped
   c. Explain why the probability of getting one head and two tails should be the same as the probability of getting two heads and one tail.
   d. Use Complement and the answer to part c to calculate the probability of exactly one head. Exactly two heads.
   e. Compare the theoretical probabilities with the observed relative frequencies you found in questions 6 and 7.

9. Physical simulation – Pick one with replacement
   a. Take two red and four blue M&M'S® and put them into a bag. Then, blindly withdraw an M&M®. Record your outcome as R or B. Put the M&M® that you picked back into the bag. Perform this experiment six times. How many times did you pick a red M&M® in six trials?
   b. What is the relative frequency of picking R? What is the relative frequency of picking B?
   c. Do the relative frequencies you found in part (b) make sense? Why or why not?

d. Compare the relative frequency of picking an R to 1/3.
e. Record your results from part a in the Table 3 below. Repeat part (a) nine more times. Namely you will blindly withdraw an M&M® and put it back into the bag. Record in Table 3 the result of each withdrawal. Each set of six selections with replacement represents a single trial. Repeat this set of six withdrawals with replacement a total of 10 times.

	Withdrawal							
Trial	**1**	**2**	**3**	**4**	**5**	**6**	**# of Reds**	**Relative Frequency of Reds**
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

Table 3: Simulate selection of M&M'S®

f. How many trials recorded a relative frequency of exactly 1/3?
g. Pool your results from all 10 trials. There were a total of 60 selections. Is the relative frequency of picking Reds close to 1/3? Explain why the pooled data should be closer to 1/3 than the individual row recorded relative frequency.

10. Physical simulation – Pick two with replacement and record the values.
   a. You have two red and four blue M&M'S® in the bag. Blindly withdraw an M&M®, look at its color. Put it back into the bag. Withdraw a second M&M®, look at its color and put it back into the bag. Record in Table 4 what you observed for the pair of withdrawals.
   b. Repeat this experiment 20 times. Record your results in Table 4.

	Trial																					
Outcomes	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**Frequency**	**Relative Frequency**
RR																						
RB																						
BR																						
BB																						

Table 4: Simulate selection of M&M'S® with replacement – record actual

c. Which was the most frequent outcome? Which was the least frequent outcome?
d. Now, record in Table 5 your results as the "number" of reds. Both RB and BR equal one red.

	Trial																					
Number of Reds	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**Frequency**	**Relative Frequency**
0																						
1																						
2																						

Table 5: Simulate selection of M&M'S® with replacement – record number of reds

e. Calculate theoretical probabilities of having zero, one, or two reds in the two M&M® withdrawal experiment.
f. Compare the relative frequencies you observed in part (d) with the theoretical probabilities.

11. Using MS Excel Random Generator
    a. Repeat the experiment "pick two with replacement" 50 times using the Randbetween function of Excel. Set up the Excel sheet as in the Figure 1. Write the formula in cell B2 as in the Formula Bar and copy it to other cells.

The formula, IF(RANDBETWEEN(1,6)<=2,"R","B") generates a random number between 1 and 6. If the number is 2 or less, it records an R in the cell. Otherwise it records a B. Copy this function into cell C2. Then copy the values in cells B2 and C2 all the way down to cells B51 and C51. To stabilize the random numbers, use the copy and past values commands. (If you do not do this, the random numbers will constantly change.) In column D you can use the CountIf command to count the number of Rs.

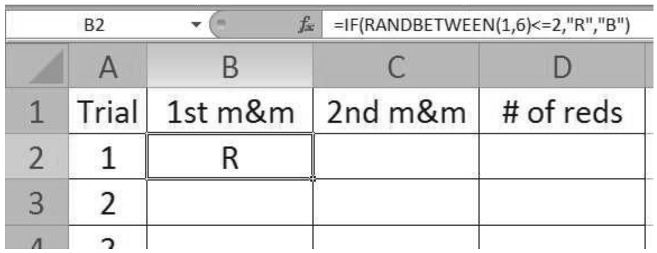


Figure 1: Randbetween function in Excel

b. Record in Table 6 the number of times you observed zero, one, or two Rs.

Number of Rs	Frequency	Relative Frequency
0		
1		
2		

Table 6: Simulate 50 times the selection of M&M'S® with replacement – record number of reds

c. Compare the relative frequencies you observed in part (b) with the theoretical probabilities determined in question 10e. Are the relative frequencies closer to the theoretical probabilities when compared to physical simulation that was repeated only 20 times?

12. Multiple Choice Exam
A student has to take an exam, but he has not studied at all. Therefore, he has no idea about the answers. The exam is a multiple choice exam with 10 questions; each has 4 choices with only one correct answer. The passing grade for this exam is 6 correct answers. The student has decided to take his chances. Table 7 contains simulated data for 30 repetitions of a 10-question exam in which a student randomly guesses the correct answer. (C= Correct and I = Incorrect) Answer the questions below according to the data in Table 7.

Trial #	Questions										No. of Cs
	1	2	3	4	5	6	7	8	9	10	
1	C	C	I	I	I	C	I	C	C	C	
2	I	C	I	C	I	I	C	I	I	I	
3	I	I	C	I	I	I	I	I	C	C	
4	C	I	I	I	C	I	I	I	I	I	
5	I	I	C	C	I	I	C	I	C	I	
6	I	I	I	I	I	C	I	I	I	I	
7	I	I	C	C	C	C	I	C	C	I	
8	I	I	C	I	I	I	I	I	I	I	
9	I	I	C	I	I	I	I	I	I	I	
10	I	I	I	I	I	I	I	I	I	I	
11	C	C	C	I	C	I	I	I	C	C	
12	I	C	C	I	I	I	I	C	I	I	
13	I	I	I	C	I	I	I	C	C	I	
14	I	I	I	I	I	I	I	I	I	I	
15	I	C	I	C	I	I	C	I	I	I	
16	I	I	C	I	I	I	I	I	C	I	
17	I	I	I	C	C	I	C	I	I	I	
18	I	C	I	C	I	I	C	I	I	I	
19	C	I	I	C	C	I	I	I	C	I	
20	C	I	I	I	I	I	I	I	C	I	
21	C	I	C	I	I	C	I	C	I	I	
22	C	C	I	I	I	I	I	I	C	I	
23	I	I	I	I	I	I	C	I	C	I	
24	I	I	I	I	I	C	I	I	I	I	
25	I	I	I	I	I	I	C	I	C	I	
26	I	I	I	I	C	C	I	I	I	C	
27	I	I	C	I	C	C	I	I	I	I	
28	I	I	C	C	I	I	C	I	I	C	
29	I	C	I	C	I	I	I	C	I	I	
30	I	I	C	I	C	I	C	I	C	I	
No. of Cs											
Proportion											

Table 7: Multiple choice exam with random guesses

a. Record the number of correct answers in each trial in the last column. How many of the trials did the student fail to score a 6 or higher?
b. Record the number of correct answers for each question in the last row. Calculate each proportion. Look at the proportion of correct answers to each of the questions. Is it exactly 0.25?

c. Using the last column of Table 7, record in Table 8 the frequency of the number of correct answers on each 10 question exam.

No. of Cs	Frequency	Relative Frequency	Cumulative Relative Frequency
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Table 8: Frequency distribution of the number of correct answers

d. What is the most frequent score on the exam? Is this number surprising?
e. What is the observed relative frequency that the student has six correct answers?
f. What is the relative frequency of scoring five or fewer correct answers?
g. What is the relative frequency of scoring six or more correct answers?
h. What is the observed relative frequency that the student has zero correct answers?
i. Calculate the theoretical probability of getting zero correct answers using the Multiplication Rule. (In Chapter 3, you will learn the formula used to calculate all of the theoretical probabilities for each of the possible outcomes.)

13. Simulate a Multiple Choice Exam
    a. Simulate his performance in MS Excel; make 30 simulation runs. Set the Excel sheet as in Figure 2. Write the formula in cell B3 as it appears in the Formula Bar. Copy it to other cells. The IF command in Excel states that if the random number is equal to 1, Excel will record a value of "C" for correct in the cell. Otherwise it records and "I" for incorrect.

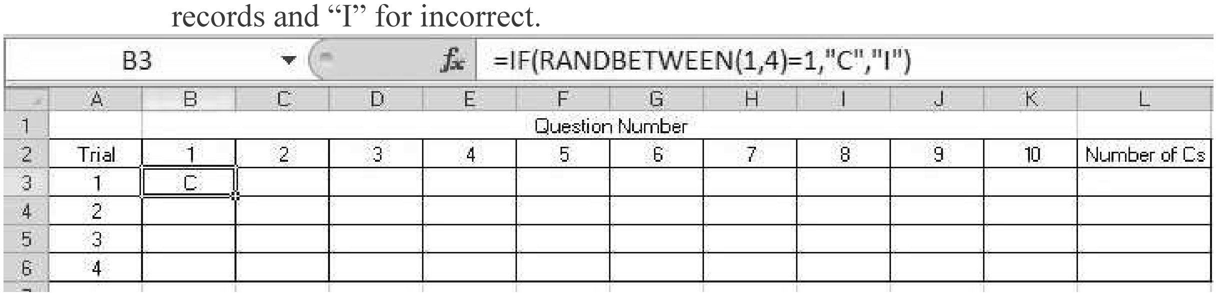


Figure 2: Simulate multiple choice exam in Excel

    b. Why is there an "= 1" in the statement?
    c. He will pass the exam if he answers at least six of the questions correctly. According to spreadsheet experiment, how many trials resulted in failure?
    d. In any of the trials, did he answer all the questions correctly?

e. Record in Table 9 the number of correct answers and identify which is the most frequent score.

No. of Cs	Frequency	Relative Frequency	Cumulative Relative Frequency
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

Table 9: Simulated data - frequency distribution of the number of correct answers

14. Refer to Section 1.5.1 which discusses absenteeism of machine repair workers. David Plante, director of operations at BT Auto Industries, wondered what would happen if there were an unusually serious flu outbreak. In that case absenteeism could be as high as 15%. When the rate was 10%, David felt that it was necessary to train a 3rd worker and that would address his concerns.
   a. Assume absenteeism is 15% and there are three workers trained to repair machines. What proportion of days would three trained workers be enough to ensure that there was always at least one machine repair worker on duty? (Use basic concepts of probability to answer this question.)
   b. Assume you want to simulate absenteeism in a serious flu outbreak. What is the range of values you will use in your random number generator?
   c. What values correspond to a worker being absent with 0.15 probability?
   d. Now simulate 50 days of three worker attendance using a 15% absentee probability. Record your results in a table similar to Table 10.

	Worker			
Day	A	B	C	**Total at Work**
1				
2				
3				
.				
.				
.				
50				

Table 10: Trained worker attendance during a serious flu outbreak

e. How many times was there at least one trained worker at work? What is the relative frequency? Compare this result to the theoretical probability calculated in part a.
f. Recall that there was significant value in having two workers to jointly repair a machine. How many times were there at least two trained workers at work? What is the relative frequency?
g. Would three trained workers adequately meet the plant's needs even during a serious outbreak of flu? Explain your answer.

Expanded Homework Questions

I. X-Press School Newspaper

X-Press is a monthly school newspaper which started in September 2008. It has three writers and three editors. According to the records from last year, it is estimated that the probability that each student will meet the deadline is 90%. Even though, this year has started out well for all the writers, the paper was delayed twice in six months due to lateness by an editor. Table 11 shows the data for the six months of 2009.

	Editors			Writers			
Month	**Tom**	**Bob**	**May**	**Ann**	**Lee**	**Eric**	**Paper Late?**
1	on time	on time	on time	on time	on time	on time	No
2	on time	on time	on time	on time	on time	on time	No
3	on time	on time	**late**	on time	on time	on time	Yes
4	on time	on time	on time	on time	on time	on time	No
5	on time	on time	on time	on time	on time	on time	No
6	on time	**late**	**late**	on time	on time	on time	Yes
Number of late times per student	0	1	2	0	0	0	

Table 11: September 2008-February 2009 performance data

1. The data in Table 11 shows that the student writers' on time performance was 100% in the last six months.
   a. Is this strong evidence of an improvement of the timeliness of the student writers?
   b. The two times late in the last six months was due to two of the editors, Bob, and May. Looking at the data, can we say that the editors have some problems?
   c. May was late twice over the six months period. Do you think May has some problems?

Actually, the data in Table 11 was randomly simulated. It is formed by generating random numbers for all the editors and writers as if all of them having a likelihood of 0.9 of meeting each deadline. It is just pure randomness that all the writers in this simulation met the deadline 100% of the time. Therefore, these results provide no evidence that they will show the same performance in the next month.

   d. If you assume that each writer and editor is 90% reliable, what is the probability of all of them meeting the deadline in any month?
   e. What is the probability of the paper being late in any month?

2. Simulate the performance of the newspaper over the 6 months yourself. Create a table as Table 11 above in Excel. Write **=IF(RANDBETWEEN(1,10)=1,"late","on time")** in B3 and copy it to other cells until G8. In cell H3, write **=IF(COUNTIF(B3:G3,"late")>0,"Yes","No")** and copy it until H8. To count the number of lateness of an individual student write **=COUNTIF(B3:B8,"late")** in cell B9 and copy it until G9.

B3 f_x =IF(RANDBETWEEN(1,10)=1,"late","on time")

	A	B	C	D	E	F	G	H
1		Editors			Writers			
2	Months	Tom	Bob	May	Ann	Lee	Eric	Late?
3	1	on time						
4	2							
5	3							
6	4							
7	5							
8	6							
9	Number of late times/student							

Figure 3: Simulation of X-Press Newspaper

a. Is there any editor or writer, whose performance is similar to May's in your experiment? Namely was any editor or writer late twice over the six-month period?
b. Make nine copies of rows one through nine in other rows of the spreadsheet. In other words, simulate the performance of the newspaper over the six months a total of 10 times.
c. In how many of the 10 runs, was a specific editor or writer late twice over the six month period?
d. How many times out of the 10 replications was the newspaper published on time for each and every one of the six months?
e. The ten runs represent simulating 60 months. What was the overall percentage of months the paper was late?
f. Record in Table 12 the results of your 10 simulation runs. Calculate the relative frequencies.

	Number of times late in six months	Relative Frequency	Cumulative Relative Frequency
0			
1			
2			
3			
4			
5			
6			

Table 12: Frequency distribution of number of times late in six months

g. Calculate the cumulative relative frequencies of the number of times late.
h. What is the relative frequency of the student paper being late two or more times over the six-month period?

3. The newspaper has changed its policy. Now, it will publish the paper on time if just one of the students has not submitted his work on time. The newspaper will still be late if two or more are late. Use the previously simulated data to evaluate this new policy.
   a. Make a frequency table similar to Table 12 which shows the number of times the paper is late in 10 simulation runs. Calculate the relative frequencies and cumulative relative frequencies.
   b. How much of a change was there in the relative frequency of never being late during a six-month period when compared to the policy used in question 2?

II. High School Hockey – Five Game Series

1. The Red Run School District has two hockey teams, the Quacks and the Tops. They are going to play a best-of-five series to determine the champion this year. The series ends when one team wins three games. We believe that both teams are equally likely to win in each game.
   a. Take a piece of paper, and cut it into two equal pieces. Write “Quacks” on one piece, and “Tops” on the other piece. Fold the papers, shuffle them, and pick one piece. Record what you have observed. Refold the piece, reshuffle and select again. Repeat this until you record three of the same team’s name. The total number of repetitions must not exceed five. It could be as few as three.
   b. How many times did you fold-shuffle-pick until you collected three of the same team’s name?
   c. Repeat the experiment in part (a) a total of 10 times. Record the results in a table similar to Table 13.

Trials	1	2	3	4	5	Number of Games
1	Q	T	T	T		4
2	Q	Q	Q			3
3	T	T	Q	Q	T	5
4						

Table 13: Physical simulation of a five game series

   d. How many trials resulted in the order of “Quacks – Quacks – Tops – Tops – Tops”?
   e. What is the probability of this specific sequence, “Quacks – Quacks – Tops – Tops – Tops”? Compare this probability with the relative frequency in your simulation. Are you surprised by the result?
   f. Tabulate the data you created and complete Table 14. This summarizes the relative frequency of the number of games played until a winner was determined.

Games Played	Number of Trials	Relative Frequency	Cumulative Relative Frequency
3			
4			
5			

Table 14: Physical simulation of a five game series - relative frequency

g. What proportion of times did the series last only three games?
h. What is the theoretical probability the series would last only three games? (Remember that there are two distinct ways this can happen.) Compare this probability to the observed relative frequency.

2. Now, let's simulate the series by flipping a coin. Let heads be the Quacks and tails be the Tops. Repeat flipping the coin until you get either three heads or three tails. The total number of flips must not exceed five. This will simulate one five-game series.
   a. Do the same thing 20 times to complete the simulation runs and record the results in a table similar to Table 13.
   b. How many times did Tops win the series in the 20 simulated experiments? Compare your results with your friends.
   c. What proportion of times did the series last only three games?
   d. What other physical experiments can you use to simulate the series? How?

3. Suppose 3 games have already been played, and the Quacks have won two out of three games.
   a. Assume that both teams are equally likely to win in each additional game. What is the probability that the Tops will come back and win the series?
   b. Historical data suggests that a team winning two out of the first three games will win the series 85% of the time. There is only a 15% chance for the other team to win. This percentage is smaller than the value found in part a. According to this information, the Quacks and the Tops are not equally likely to win in each game. Assume the probability that the Tops will win two games in a row is only 0.15. What is the probability of Tops winning an individual game?

4. You will simulate the fourth game using the RAND (Math → Prb → 1: Rand) function of your calculator. If the random number is less than or equal to the probability of the Tops winning that was found in 3b, then the Tops win the game and the series continues. Otherwise the Quacks win the game and the series ends.
   a. Simulate the fourth game. Who won the game?
   b. Simulate the fourth and the fifth games 20 times. What percentage of time did the series end with the fourth game?
   c. Explain why you cannot flip a coin to simulate the fourth game?

5. Simulate the fourth and the fifth games 100 times using MS Excel. Write the formulation in cell B2 and drag it to the bottom and to the right until you cover all the games and the simulation runs. The letter x corresponds the probability of the Tops winning that you have found earlier.

	A	B	C	D
1	Run Number	4th game	5th game	
2	1	=IF(RAND()<=x,"Tops","Quacks")		
3	2			
4	3			

Figure 4: Simulate fourth and fifth games of hockey series

a. How many times did Quacks win the series in four games?
b. How many times did Quacks win the series in five games?
c. How many times did Tops win the series?

III. 2010 Bradley Cup – Seven Game Series

1. The Finals of 2010 Bradley Cup Basketball Playoffs have begun. The Raleigh Elks and Lansing Moose are playing a best-of-seven series for the championship. Four games have already been played. The Elks won three out of the first four games and the Moose won only one game. Assuming that the Moose and the Elks are equally likely to win in each game, answer the three questions below.
   a. What is the probability that the Moose will win the series?
   b. If the Moose wins the fifth game, what is the probability that the Moose will win the series?
   c. What is the probability that the Elks will win the series if the Moose wins the fifth game?

   Historical data indicates that a team winning three out of the first four games won the series 95% of the time. This data suggests that the two teams are not equally likely to win each game.
   d. Explain why the team that is ahead three to one may have a higher than 50% chance of winning each game.
   e. Perhaps the chance of winning in each game is only 40% for the Moose. What would be the probability that the Moose will win the series?
   f. Based on historical data, the chances of a team like Moose winning the series is only 0.05. What value raised to third power yields 0.05? What does this number represent?

2. Simulate the fifth, sixth and seventh games 100 times using MS Excel. Use the value you found in part (f) above rounded to three decimal places to run your simulation.
   a. How many times did Elks win the series in five games?
   b. How many times did Elks win the series in six games?
   c. How many times did Moose win the series?

IV. Airline Overbooking

1. Profits and Break-Even Analysis

Great Lakes Airlines (GLA) operates out of Lansing, MI (LAN) and serves 20 cities in seven states. They want to conduct a profit break-even analysis on Flight 425 which flies from Lansing (LAN) to in Milwaukee, WI (MKE). The average one-way airfare price is $300. The basic operating cost of the plane is $6,000 for each flight with an additional cost of $20 per customer. Flight 425 has 30-passenger seat capacity. Customers who cancel their flights or don't show up at the gate will pay GLA $ 75 but get back the rest of the ticket price.

a. What is the revenue if the number of reserved passengers is 20 and all passengers show up?
b. What is the cost if the number of reservations is 20 and all passengers show up? What is the profit or loss with 20 passengers?
c. The number of passengers where the cost equals revenue is called the break-even point. According to this, what is the break-even point for this case in which all reservations ow up?

   According to GLA flight records, 15% of the people who make reservation either cancel or just don't show up.
d. In Table 15 below, five customers out of 30 did not show up, so what is the profitability? (Remember to include the penalty cost for not showing up.)

Customer Number	Show-up?	Customer Number	Show-up?
1	Yes	**17**	Yes
2	No	**18**	No
3	Yes	**19**	Yes
4	Yes	**20**	Yes
5	Yes	**21**	Yes
6	Yes	**22**	No
7	Yes	**23**	Yes
8	Yes	**24**	Yes
9	Yes	**25**	Yes
10	Yes	**26**	Yes
11	Yes	**27**	Yes
12	Yes	**28**	No
13	Yes	**29**	No
14	Yes	**30**	Yes
15	Yes	**Number of No-Shows**	5
16	Yes		

Table 15: Customers showing up out of total of 30 reservations

e. Write a mathematical expression to calculate GLA's net profit as a function of the number of reservations and the number of no shows. Be careful to define all of your variables.
f. Assume that 30 customers have reserved tickets for a flight. Simulate the showing up behavior of each customer in Excel Worksheet. Use Figure 5 as your reference. Then calculate the net profit GLA earned on the flight. Use RANDBETWEEN(1,100) function to generate numbers between 1 and 100. If this random number is less than or equal to 15, it means the person did not show up. Count the number of no-shows using the function COUNTIF(B2:B31,"No").

	A	B	C	D
1	customer #	Show-up?		
2	1	Yes		
3	2	Yes		
4	3	Yes		
5	=IF(RANDBETWEEN(1,100)<=15,"No",			
6	"Yes")			
7	6	Yes		
8	7	Yes		
9	8	Yes		
10	9	Yes		
11	10	No		
12	11	Yes		

Figure 5: Simulation of customers showing up at the gate

g. Is it possible to fill all the seats without booking a maximum of 30 passengers? What is the probability of that happening?
h. If GLA has 250 flights per year, on average how many times will all 30 people show up?

2. Overbooking and Bumping: 31 Reservations Allowed

Overbooking means that more tickets are sold (or booked) than there are seats on the plane. Overbooking is a common practice in the airline industry and it is legal. Sometimes people cancel their flights or just don't show up at the gate. This results in empty seats on the planes and loss opportunity for more revenue. By overbooking flights, the likelihood of empty seats because of those cancellations and no-shows is reduced.

If the number of ticketholders who arrive at the gate is more than the flight capacity, then some of the passengers will be bumped to other flights. They will receive some compensation for the disruption. They generally receive another ticket for the next available flight and compensation in the form of cash or a voucher for future travel. If the delay involves an overnight stay, they will be placed in a hotel free of charge. Compensation will vary depending upon whether the passenger volunteers to be bumped or is involuntarily bumped.
GLA wants to determine a good overbooking strategy to reduce the lost revenue due to cancellations and no shows on fully booked flights. They will apply the initial overbooking strategy on Flight 425. If a customer is denied boarding in case of overbooking, GLA will bump the customer to another flight. GLA will $250 to compensate that customer for the inconvenience. GLA starts its overbooking strategy with 31 reservations on each flight.

a. Determine GLA's net profit on this flight if all 31 customers show up. (Do not count the 31st ticket as revenue for this flight. The bumped customer will use his ticket on another flight.)

b. Determine GLA's net profit if only 30 customers show up. (Remember a no-show is charged only a $75 penalty and not the full $300 price of the ticket.)
c. Determine GLA's net profit if only 29 customers show up.
d. Write a mathematical expression to calculate GLA's net profit for the case of overbooking as a function of: the number of reservations, the number of no shows and the number of people bumped. Be careful to define all of your variables.
e. Table 16 shows the frequency of the number of customers who did not show up over 100 days when GLA allows 31 reservations and overbooks one customer consistently. Fill in the blank cells under the number bumped and the profit/day column. Multiply the number of days by profit/day. Write the solution in the fifth column. Take the average of the fifth column to find the average profit per day.

Number of No-Shows	Number Bumped	Profit/Day	Number of Days	Profit/Day × Number of Days
0			1	
1			1	
2			8	
3			17	
4			23	
5			15	
6			15	
7			7	
8			6	
9			5	
10			2	
	Average Profit/Day			___

Table 16: Frequency of no shows for 100 days with 31 bookings allowed

f. What number of no-shows was the most profitable? Why?

3. Overbooking and Bumping: 32 Reservations Allowed

a. Table 17 shows the frequency of number of no-shows over 100 days when GLA allows 32 reservations. Fill in the blank cells under the number bumped column. Calculate the relative frequency for each number of no-shows and write it to the corresponding row in the fourth column. Write the cumulative relative frequencies in the fifth column.

Number of No-Shows	Number Bumped	Number of Days	Relative Frequency	Cumulative Relative Frequency
0		1		
1		1		
2		8		
3		17		
4		23		
5		15		
6		15		
7		7		
8		6		
9		5		
10		2		

Table 17: Frequency table for 100 days

Now, answer the questions below using the fifth column of Table 17.

b. What percent of the time did all 32 customers appear at the gate?
c. What percent of the time was at least one customer bumped?

d. What percent of the time did more than four customers not show up for their flight?
e. Complete Table 18 to determine the average profit per day..

Number of No-Shows	Number Bumped	Profit/Day	Number of Days	Profit/Day × Number of Days
0			1	
1			1	
2			8	
3			17	
4			23	
5			15	
6			15	
7			7	
8			6	
9			5	
10			2	
	Average Profit/Day			___

Table 18: Frequency of no shows for 100 days with 32 bookings allowed

4. Optimal Overbooking Policy

Table 19 shows the 100-day simulation results of Flight 425 with different overbooking policies (35, 36, and 37 reservations). Each policy was simulated independently. For example, according to this table, when GLA allowed 35 customers to reserve for Flight 425, three customers did not show up eight times out of 100 days. In this case, two customers of the 32 customers who did show up have to be bumped to other flights. Use the data in Table 19 to answer the following questions.

Number of No-Shows	**Frequency of No-Shows for Different Policies**		
	35	**36**	**37**
0	2	1	0
1	2	0	3
2	7	5	2
3	8	2	9
4	16	17	14
5	21	21	14
6	16	21	17
7	12	16	16
8	8	9	14
9	6	3	7
10	0	3	2
11	2	0	2
12	0	2	0
13	0	0	0
14	0	0	0
15	0	0	0

Table 19: No-Shows history of Flight 425 for different overbooking policies

a. Fill in the blank cells in Table 20 using the data from Table 19.

	Overbooking Policy		
	35	36	37
Probability of no bumping			
Probability of bumping 2 or more			
Average number of no-shows			
Average number of bumped passengers			
Average number of flying passengers			

Table 20: Evaluation of overbooking policies – 35, 36, and 37

b. Create a table similar to Table 18 for each of these three overbooking policies.
c. Determine the average profit for each policy? Which policy is the best?
d. How many random numbers would you need to generate in order to simulate 100 days in which 35 seats were booked?

Chapter 1 Summary

What have we learned?

The purpose of this chapter is to gain a better understanding of the nature of random behavior. Random behavior means that we do not know what will happen in the future. Often we know one of a number of possibilities will happen. However, we do not know which of them will occur. One way to study situations that include multiple random events is to simulate them. We started by simulating two equally likely outcomes by flipping coins. When we have outcomes that are not equally likely we use a random number generator on a calculator or computer. The purpose of the simulation is to try to discover what would happen if a situation occurred a large number of times. This can give us an idea of how likely a particular outcome is. We can use this information to make decisions.

We have also learned how to combine the probabilities of two or more events if we know the probability of each of them occurring. For independent events, we can multiply the probabilities together to find the probability that they both occur.

Finally, we learned how to use an Excel spreadsheet for larger scale simulations. Excel can be used for generating random numbers and can also do calculations that can be used to evaluate the results. When we flip coins or use a calculator to generate the results, we have to write the results down. Excel can keep and easily display all of the simulated results. It is important to remember to use the paste special “values” command to save the numbers generated. This is to prevent random numbers that were generated from changing the next time we enter a command.

Our intuition when it comes to random behavior is often wrong. People have a hard time getting rid of the belief that the future will correct for a surprising run of random behavior in the past. There is no “Law of Averages” that states that the outcomes of a random event will "even out" within a small sample. However, the average of the results of a large number of trials will get closer to the expected value as an event is repeated. This is not because future results somehow make up for the past. It is simply that a small number of results have a small impact on the average of a large number of results.

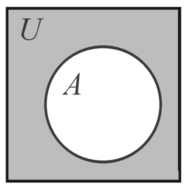
Terms

Chance event Events whose outcomes are random.

Collectively exhaustive events A set of events is collectively exhaustive if at least one of them must occur (i.e., a collectively exhaustive set of events includes all possible outcomes).

Collectively exhaustive sets If sets A and B are collectively exhaustive subsets of the universal set U, then $A \bigcup B = U$.

Complement of a set The complement of set A, A', contains all of the elements that are not included in set A. The shaded region of the Venn diagram below shows the complement of A. The complement of set A is often written as A^c or A' or $\overline{A}$.

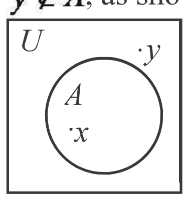


Complementary events and probability Two events A and B are complementary if they are mutually exclusive and collectively exhaustive.

$$P(A) + P(B) = 1$$
$$P(B) = 1 - P(A)$$
$$P(A) = 1 - P(B)$$

Compound events A compound event consists of two or more simple events.

Element of a set The objects that make up a set are referred to as elements of the set. If x is an element of set A, then $x \in A$, and if y is not an element of set A, $y \notin A$, as shown in the Venn diagram below.



Empty (or null) set The empty set is the set containing no elements, denoted $\varnothing$ or { }.

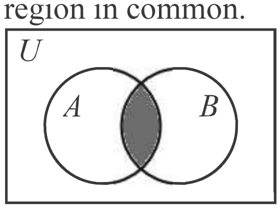
Event Any subset of the sample space is called an event.

Experiment An experiment is an activity under consideration whose outcome is left to chance.

Independent events and probability Two events A and B are independent if the outcome of one event does not affect the outcome of the other. If A and B are independent, then

$$P(A \text{ and } B) = P(A \cap B) = P(A) \cdot P(B)$$

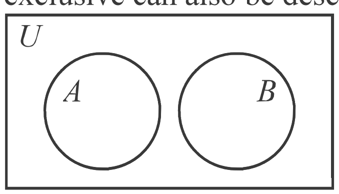
Intersection of sets The intersection of sets A and B contains all the elements that belong to set A and also to set B. The shaded region of the Venn diagram below shows the intersection of sets A and B. Notice that A and B have this region in common.



Mutually exclusive events and probability Two events are mutually exclusive if the occurrence of one event precludes the possibility of the other event occurring. If events A and B are mutually exclusive, then the probability of event A or event B occurring is the sum of the probabilities of events A and B.

$$P(A \text{ or } B) = P(A \cap B) = P(A) + P(B)$$

Mutually exclusive sets If sets A and B are mutually exclusive subsets of the universal set U, then they have no elements in common. Two sets that are mutually exclusive can also be described as disjointed sets.



Outcome Each possible observation or occurrence in an experiment is called an outcome.

Probability The probability of an event is a measure of the likelihood of that particular event occurring. The probability of an event is a fraction or decimal between zero and one. Events that are unlikely have probabilities closer to zero. Events that are likely to occur have probabilities closer to one.
If all possible outcomes are equally likely, the probability of an event occurring equals the ratio of the number of outcomes that result in a particular event to the total number of possible outcomes in the sample space. The probability P of event E occurring can be written using the following notation:

$$P(E) = \frac{\text{number of favorable outcomes}}{\text{total number of possible outcomes}}.$$

Random process A random process models changes in a system over time, where the changes lack a definite plan, purpose, or pattern.

Random variable A random variable is an expression whose value is subject to chance variations lacking any definite plan, purpose, or pattern. Like any other variable in mathematics, a random variable can take on different values.

Randomness Randomness implies that the outcome of an event does not follow any prescribed pattern and is thus unpredictable. Another way of saying this is that the outcome of the event occurs strictly by chance.

Observed relative frequency The ratio of the number of times an event is observed to occur to the total number of observations. The observed relative frequency of an event can be used as an estimate of the probability of the event. The observed relative frequency of an event is also referred to as the empirical probability of the event.

Outcome The result of an experiment or simulation involving uncertainty.

Sample space The sample space is the set consisting of all the possible outcomes of an experiment. This set is often denoted S.

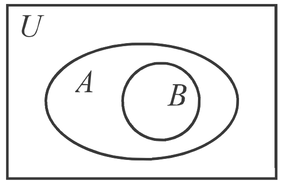
Sequence A list of outcomes when an event is repeated.

Set The term set is undefined, but may be thought of as a collection of objects. In set-builder notation, the elements of a set are enclosed in braces; e.g., $S = \{0, 1, 2\}$.

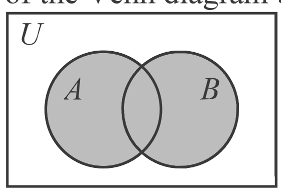
Simulation A simulation is a process used to model random events.

String A sequence of consecutive results with the same outcome

Subset A subset of a set contains elements, all of which are contained in another set. If B is a subset of *A,* then every element of *B* is also an element of *A*, denoted by $B \subseteq A$. The Venn diagram below shows the subset relationship where *B* is completely contained within *A*; i.e., *B* is a subset of *A*.



Union of sets The union of set *A* and set *B* contains all of the elements that belong to set *A* or set *B* or *both* set *A* and set *B*. In this sense, we are using the word *or* inclusively, which means either or both. The shaded region of the Venn diagram below shows the union of sets *A* and *B*.



Universal set The set containing all possible elements is denoted as *U*.

Variability Variability is a characteristic of data that refers to the fact that each data value is not necessarily the same.

Chapter 1 (Basic Probability and Randomness) Objectives

You should be able to:

- Identify the possible outcomes for an event.
- Determine the probability of each of the possible outcomes for an event.
- Determine whether or not two events are complementary.
- Use the concept of complementarity to calculate the probability of an event.
- Use the multiplication principle to determine the probability of two independent events both occurring.
- Use “randInt” on the TI calculator and “randbetween” in Excel to generate a random integer between two values.
- Use “countif” function in Excel to find out how often a particular value appears in a list.
- Evaluate the results of a simulation to find the likelihood of a particular outcome.

Chapter 1 Study Guide

1. What is the difference between independent events and mutually exclusive events? Give an example of each.

2. What does it mean if two events are complementary? Give an example.

3. Give an example of two events that are mutually exclusive but not complementary?

4. If two events are independent and both have a 90% chance of occurring.

    a. What is the probability both will occur?

    b. What is the probability one will occur and one will not?

    c. What is the probability neither will occur?

    d. If three independent events, what is the probability of them all occurring?

5. What does the notation $P(A)$ mean? What about the notation $P(A^c)$

6. If $P(A) = 0.75$ what is $P(A^c)$?

7. Are the results of a simulation always reliable? Explain.