Chapter 7: Probability

Figure 1: Roulette is a game whose outcomes are based entirely on the concept of probability.

Casinos are big business; according to the American Gaming Association, commercial casinos in the United States brought in over $43 billion in revenue in 2019. Casinos must walk a fine line in order to be profitable. Their customers must lose more money than they win, on average, in order to stay in business. But if the chances of a single customer winning more money than they lose is too small, people will stop coming in the door to play the games.

In this chapter, we'll study the techniques a casino must use to determine how likely it is that a customer will win a particular game, and then how the casino decides how much money a winner will rake in so that the customers are happy, but the casino also turns a profit in the long run. In order to figure out those likelihoods, we have to be able to somehow consider every possible outcome of these games. For example, in a game that involves players receiving 5 cards from a deck of 52, there are 2,598,960 possibilities for each player. We'll start off this chapter by learning how to count those possible outcomes.

7.1 The Multiplication Rule for Counting

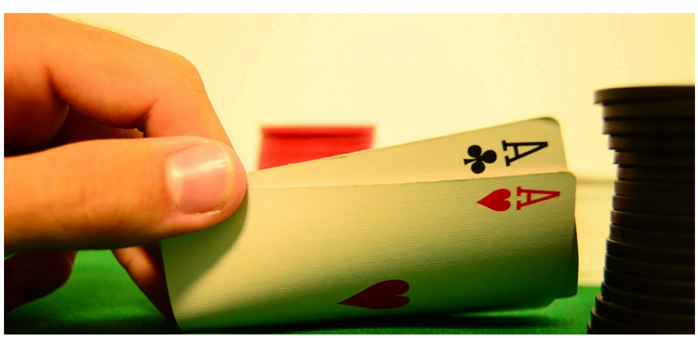


Figure 2: The Multiplication Rule for Counting allows us to compute more complicated probabilities, like drawing two aces from a deck.

Learning Objectives

After completing this section, you should be able to:

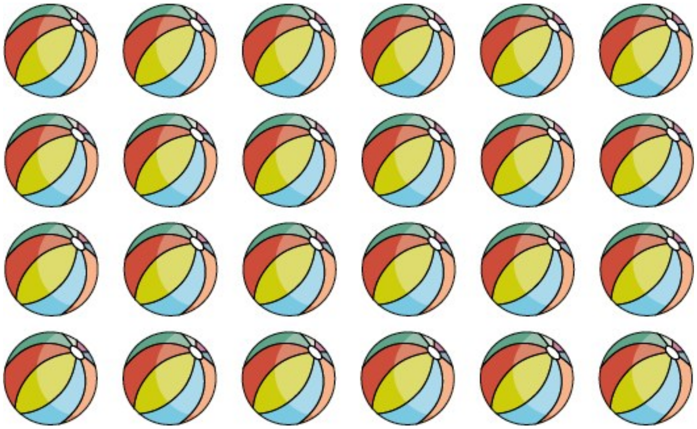
1. Apply the Multiplication Rule for Counting to solve problems.

One of the first bits of mathematical knowledge children learn is how to count objects by pointing to them in turn and saying: “one, two, three, ...” That’s a useful skill, but when the number of things that we need to count grows large, that method becomes onerous (or, for *very* large numbers, impossible for humans to accomplish in a typical human lifespan). So, mathematicians have developed short cuts to counting big numbers. These techniques fall under the mathematical discipline of **combinatorics**, which is devoted to counting.

Multiplication as a Combinatorial Short Cut

One of the first combinatorial short cuts to counting students learn in school has to do with areas of rectangles. If we have a set of objects to be counted that can be physically arranged into a rectangular shape, then we can use multiplication to do the counting for us. Consider this set of objects :

Certainly we can count them by pointing and running through the numbers, but it’s more efficient to group them .



If we group the balls by 4s, we see that we have 6 groups (or, we can see this arrangement as 4 groups of 6 balls). Since multiplication is repeated addition (i.e., $6 \times 4 = 4 + 4 + 4 + 4 + 4 + 4$), we can use this grouping to quickly see that there are 24 balls.

Let’s generalize this idea a little bit. Let’s say that we’re visiting a bakery that offers customized cupcakes. For the cake, we have three choices: vanilla, chocolate, and strawberry. Each cupcake can be topped with one of four types of frosting: vanilla, chocolate, lemon, and strawberry. How many different cupcake combinations are possible? We can think of laying out all the possibilities in a grid, with cake choices defining the rows and frosting choices defining the columns (Figure 7.5).

	frostings			
cakes	**vanilla**	**chocolate**	**lemon**	**strawberry**
vanilla	vanilla cake with vanilla frosting	vanilla cake with chocolate frosting	vanilla cake with lemon frosting	vanilla cake with strawberry frosting
chocolate	chocolate cake with vanilla frosting	chocolate cake with chocolate frosting	chocolate cake with lemon frosting	chocolate cake with strawberry frosting
strawberry	strawberry cake with vanilla frosting	strawberry cake with chocolate frosting	strawberry cake with lemon frosting	strawberry cake with strawberry frosting

Since there are 3 rows (cakes) and 4 columns (frostings), we have $3 \times 4 = 12$ possible combinations. This is the reasoning behind the **Multiplication Rule for Counting**, which is also known as the Fundamental Counting Principle. This rule says that if there are n ways to accomplish one task and m ways to accomplish a second task, then there are $n \times m$ ways to accomplish both tasks. We can tack on additional tasks by multiplying the number of ways to accomplish those tasks to our previous product.

Example 1 **Using the Multiplication Rule for Counting**

Every card in a standard deck of cards has two identifying characteristics: a suit (clubs, diamonds, hearts, or spades; these are indicated by these symbols, respectively: ♣, ♢, ♡, ♠) and a rank (ace, 2, 3, 4, 5, 6, 7, 8, 9, 10, jack, queen, and king; the letters A, J, Q, and K are used to represent the words). Each possible pair of suit and rank appears exactly once in the deck. How many cards are in the standard deck?

Solution

Since there are 4 suits and 13 ranks, the number of cards must be $4 \times 13 = 52$.

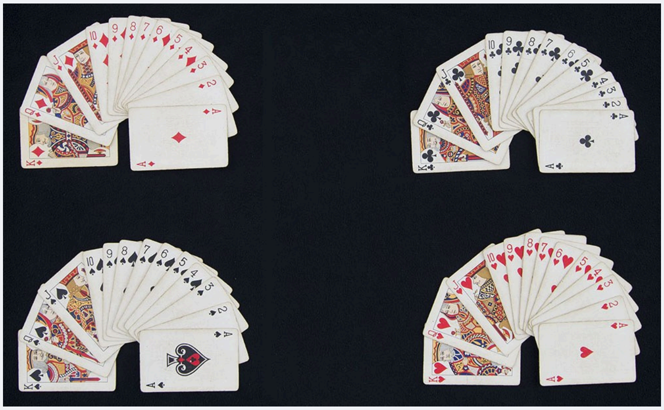


Figure 6: Standard Deck of Cards, Sorted by Rank and Suit

Example 2 **Using the Multiplication Rule for Counting for 4 Groups**

The University Combinatorics Club has 31 members: 8 seniors, 7 juniors, 5 sophomores, and 11 first-years. How many possible 4-person committees can be formed by selecting 1 member from each class?

Solution

Since we have 8 choices for the senior, 7 choices for the junior, 5 for the sophomore, and 11 for the first-year, there are $8 \times 7 \times 5 \times 11 = 3,080$ different ways to fill out the committee.

Example 3 **Using the Multiplication Rule for Counting for More Groups**

The standard license plates for vehicles in a certain state consist of 6 characters: 3 letters followed by 3 digits. There are 26 letters in the alphabet and 10 digits (0 through 9) to choose from. How many license plates can be made using this format?

Solution

Since there are 26 different letters and 10 different digits, the total number of possible license plates is $26 \times 26 \times 26 \times 10 \times 10 \times 10 = 17,576,000$.

Key Terms

- combinatorics
- Multiplication Rule for Counting (Fundamental Counting Principle)

Key Concepts

- The Multiplication Rule for Counting is used to count large sets.

7.2 Permutations

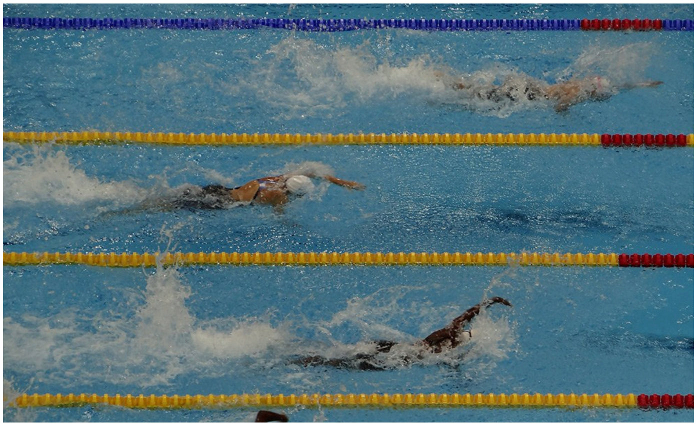


Figure 7: We can use permutations to calculate the number of different orders of finish in an Olympic swimming heat.

Learning Objectives

After completing this section, you should be able to:

1. Use the Multiplication Rule for Counting to determine the number of permutations.
2. Compute expressions containing factorials.
3. Compute permutations.
4. Apply permutations to solve problems.

Swimming events are some of the most popular events at the summer Olympic Games. In the finals of each event, 8 swimmers compete at the same time, making for some exciting finishes. How many different orders of finish are possible in these events? In this section, we'll extend the Multiplication Rule for Counting to help answer questions like this one, which relate to **permutations**. A permutation is an ordered list of objects taken from a given population. The length of the list is given, and the list cannot contain any repeated items.

Applying the Multiplication Rule for Counting to Permutations

In the case of the swimming finals, one possible permutation of length 3 would be the list of medal winners (first, second, and third place finishers). A permutation of length 8 would be the full order of finish (first place through eighth place). Let's use the Multiplication Rule for Counting to figure out how many of each of these permutations there are.

Example 1 **Using the Multiplication Rule for Counting to Find the Number of Permutations**

The final heat of Olympic swimming events features 8 swimmers (or teams of swimmers).

1. How many different podium placements (first place, second place, and third place) are possible?
2. How many different complete orders of finish (first place through eighth place) are possible?

Solution

1. Let's start with the first place finisher. How many options are there? Since 8 swimmers are competing, there are 8 possibilities. Once that first swimmer completes the race, there are 7 swimmers left competing for second place. After the second finisher is decided, there are 6 swimmers remaining who could possibly finish in third place. Thus, there are 8 possibilities for first place, 7 for second place, and 6 for third place. The Multiplication Rule for Counting then tells us there are $8 \times 7 \times 6 = 336$ different ways the winners' podium can be filled out.
2. To look at the complete order of finish, we can continue the pattern we can see in part 1 of this example: There are 5 possibilities for fourth place, 4 for fifth place, 3 for sixth place, 2 for seventh place, and then just 1 swimmer is left to finish in eighth place. Using the Multiplication Rule for Counting, we see that there are $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 40,320$ possible orders of finish.

Factorials

The pattern we see in Example 1 occurs commonly enough that we have a name for it: **factorial**. For any positive whole number n, we define the factorial of n (denoted $n!$ and read "n factorial") to be the product of every whole number less than or equal to n. We also define $0!$ to be equal to one. We will use factorials in a couple of different contexts, so let's get some practice doing computations with them.

Example 2 **Computing Factorials**

Compute the following:

1. $4!$

2. $\frac{8!}{6!}$
3. $\frac{9!}{3!4!}$

Solution

1. $4! = 4 \times 3 \times 2 \times 1 = 24$
2. There are two ways to approach this calculation. The first way is to compute the factorials first, then divide:

$$\frac{8!}{6!} = \frac{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = \frac{40{,}320}{720} = 56$$

However, there is an easier way! You may notice in the second step that there are several terms that can be canceled; that's always the case whenever we divide factorials. In this case, notice that we can rewrite the numerator like this:

$$8! = 8 \times 7 \times (6 \times 5 \times 4 \times 3 \times 2 \times 1) = 8 \times 7 \times 6!$$

With that in mind, we can proceed this way by canceling out the 6!:

$$\frac{8!}{6!} = \frac{8 \times 7 \times 6!}{6!} = 8 \times 7 = 56$$

That's much easier!

1. Let's approach this one using our canceling technique. When we see two factorials in either the numerator or denominator, we should focus on the larger one first. So:

$$\frac{9!}{3!4!} = \frac{9 \times 8 \times 7 \times 6 \times 5 \times 4!}{3!4!} = \frac{9 \times 8 \times 7 \times 6 \times 5}{3 \times 2 \times 1} = 9 \times 8 \times 7 \times 5 = 2,520$$

Permutations

As we've seen, factorials can pop up when we're computing permutations. In fact, there is a formula that we can use to make that connection explicit. Let's define some notation first. If we have a collection of n objects and we wish to create an ordered list of r of the objects (where $1 \leq r \leq n$), we'll call the number of those permutations nP_r (read "the number of permutations of n objects taken r at a time"). We formalize the formula we'll use to compute permutations below.

FORMULA

$$nP_r = \frac{n!}{(n-r)!}$$

If you wondered why we defined $0! = 1$ earlier, it was to make formulas like this one work; if we have n objects and want to order all of them (so, we want the number of permutations of n objects taken n at a time), we get $nP_n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = \frac{n!}{1} = n!$. Next, we'll get some practice computing these permutations.

Example 3 **Computing Permutations**

Find the following numbers:

1. The number of permutations of 12 objects taken 3 at a time
2. The number of permutations of 8 objects taken 5 at a time
3. The number of permutations of 32 objects taken 2 at a time

Solution

1. $12P_3 = \frac{12!}{(12-3)!} = \frac{12!}{9!} = \frac{12 \times 11 \times 10 \times 9!}{9!} = 12 \times 11 \times 10 = 1,320$
2. $8P_5 = \frac{8!}{(8-5)!} = \frac{8 \times 7 \times 6 \times 5 \times 4 \times 3!}{3!} = 8 \times 7 \times 6 \times 5 \times 4 = 6,720$
3. $32P_2 = \frac{32!}{(32-2)!} = \frac{32 \times 31 \times 30!}{30!} = 32 \times 31 = 992$

Example 4 **Applying Permutations**

1. A high school graduating class has 312 students. The top student is declared valedictorian, and the second-best is named salutatorian. How many possible outcomes are there for the valedictorian and salutatorian?
2. In the card game blackjack, the dealer's hand of 2 cards is dealt with 1 card faceup and 1 card facedown. If the game is being played with a single deck of (52) cards, how many possible hands could the dealer get?
3. The University Combinatorics Club has 3 officers: president, vice president, and treasurer. If there are 18 members of the club, how many ways are there to fill the officer positions?

Solution

1. This is the number of permutations of 312 students taken 2 at a time, and $312P_2 = 97,032$.
2. We want the number of permutations of 52 cards taken 2 at a time, and $52P_2 = 2,652$.
3. Here we're looking for the number of permutations of 18 members taken 3 at a time, and $18P_3 = 4,896$.

WHO KNEW?

Very Big Permutations

Permutations involving relatively small sets of objects can get very big, very quickly. A standard deck contains 52 cards. So, the number of different ways to shuffle the cards—in other words, the number of permutations of 52 objects taken 52 at a time—is $52! \approx 8 \times 10^{67}$ (written out, that's an 8 followed by 67 zeroes). The estimated age of the universe is *only* about 4×10^{17} seconds. So, if a very bored all-powerful being started shuffling cards at the instant the universe began, it would have to have averaged at least $\frac{8\times10^{67}}{4\times10^{17}} \approx 2 \times 10^{50}$ shuffles *per second since the beginning of time* to have covered every possible arrangement of a deck of cards. That means the next time you pick up a deck of cards and give it a good shuffle, it's almost certain that the particular arrangement you created has never been created before and likely never will be created again.

Key Terms

- permutation
- factorial

Key Concepts

- Using the Multiplication Rule for Counting to enumerate permutations.
- Simplifying and computing expressions involving factorials.
- Using factorials to count permutations.

Formulas

- ${}_nPr = \frac{n!}{(n-r)!}$

7.3 Combinations

Figure 8: Combinations help us count things like the number of possible card hands, when the order in which the cards were drawn doesn't matter.

Learning Objectives

After completing this section, you should be able to:

1. Distinguish between permutation and combination uses.
2. Compute combinations.
3. Apply combinations to solve applications.

In Permutations, we studied permutations, which we use to count the number of ways to generate an ordered list of a given length from a group of objects. An important property of permutations is that the order of the list matters: The results of a race and the selection of club officers are examples of lists where the order is important. In other situations, the order is not important. For example, in most card games where a player receives a hand of cards, the order in which the cards are received is irrelevant; in fact, players often rearrange the cards in a way that helps them keep the cards organized.

Combinations: When Order Doesn't Matter

In situations in which the order of a list of objects doesn't matter, the lists are no longer permutations. Instead, we call them **combinations**.

Example 1 **Distinguishing Between Permutations and Combinations**

For each of the following situations, decide whether the chosen subset is a permutation or a combination.

1. A social club selects 3 members to form a committee. Each of the members has an equal share of responsibility.
2. You are prompted to reset your email password; you select a password consisting of 10 characters without repeats.
3. At a dog show, the judge must choose first-, second-, and third-place finishers from a group of 16 dogs.
4. At a restaurant, the special of the day comes with the customer's choice of 3 sides taken from a list of 6 possibilities.

Solution

1. Since there is no distinction among the responsibilities of the 3 committee members, the order isn't important. So, this is a combination.
2. The order of the characters in a password matter, so this is a permutation.
3. The order of finish matters in a dog show, so this is a permutation.
4. A plate with mashed potatoes, peas, and broccoli is functionally the same as a plate with peas, broccoli, and mashed potatoes, so this is a combination.

Counting Combinations

Permutations and combinations are certainly related, because they both involve choosing a subset of a large group. Let's explore that connection, so that we can figure out how to use what we know about permutations to help us count combinations. We'll take a basic example. How many ways can we select 3 letters from the group A, B, C, D, and E? If order matters, that number is $_5P_3 = 60$. That's small enough that we can list them all out in the table below.

ABC	ABD	ABE	ACB	ACD	ACE
ADB	ADC	ADE	AEB	AEC	AED
BAC	BAD	BAE	BCA	BCD	BCE
BDA	BDC	BDE	BEA	BEC	BED
CAB	CAD	CAE	CBA	CBD	CBE
CDA	CDB	CDE	CEA	CEB	CED
DCA	DAC	DAE	DBA	DBC	DBE
DCA	DCB	DCE	DEA	DEB	DEC
EAB	EAC	EAD	EBA	EBC	EBD
ECA	ECB	ECD	EDA	EDB	EDC

Now, let's look back at that list and color-code it so that groupings of the same 3 letters get the same color, as shown in :

Possible combinations

	Column					
Row	**1**	**2**	**3**	**4**	**5**	**6**
1	ABC	ABD	ABE	ACB	ACD	ACE
2	ADB	ADC	ADE	AEB	AEC	AED
3	BAC	BAD	BAE	BCA	BCD	BCE
4	BDA	BDC	BDE	BEA	BEC	BED
5	CAB	CAD	CAE	CBA	CBD	CBE
6	CDA	CDB	CDE	CEA	CEB	CED
7	DAB	DAC	DAE	DBA	DBC	DBE
8	DCA	DCB	DCE	DEA	DEB	DEC
9	EAB	EAC	EAD	EBA	EBC	EBD
10	ECA	ECB	ECD	EDA	EDB	EDC

After color-coding, we see that the 60 cells can be seen as 10 groups (colors) of 6. That's no coincidence! We've already seen how to compute the number of permutations using the formula To compute the number of combinations, let's count them another way using the Multiplication Rule for Counting. We'll do this in two steps:

Step 1: Choose 3 letters (paying no attention to order).

Step 2: Put those letters in order.

The number of ways to choose 3 letters from this group of 5 (A, B, C, D, E) is the number of combinations we're looking for; let's call that number $5C_3$ (read "the number of combinations of 5 objects taken 3 at a time"). We can see from our chart that this is ten (the number of colors used). We can generalize our findings this way: remember that the number of permutations of n things taken r at a time is $nP_r = \frac{n!}{(n-r)!}$. That number is *also* equal to $nC_r \times r!$, and so it must be the case that $\frac{n!}{(n-r)!} = nC_r \times r!$. Dividing both sides of that equation by $r!$ gives us the formula below.

FORMULA

$$nC_r = \frac{n!}{r!(n-r)!}$$

Example 2 **Using the Combination Formula**

Compute the following:

1. $8C_3$
2. $12C_5$
3. $15C_9$

Solution

1. $8C_3 = \frac{8!}{3!(8-3)!} = \frac{8\times7\times6\times5!}{3\times2\times1\times5!} = 8 \times 7 = 56$
2. $12C_5 = \frac{12!}{5!(12-5)!} = \frac{12\times11\times10\times9\times8\times7!}{5\times4\times3\times2\times1\times7!} = 11 \times 9 \times 8 = 792$
3. $15C_9 = \frac{15!}{9!(15-9)!} = \frac{15\times14\times13\times12\times11\times10\times9!}{9!\times6\times5\times4\times3\times2\times1} = \frac{14\times13\times11\times10}{4} = 5,005$

Example 3 **Applying the Combination Formula**

1. In the card game Texas Hold'em (a variation of poker), players are dealt 2 cards from a standard deck to form their hands. How many different hands are possible?
2. The board game *Clue* uses a deck of 21 cards. If 3 people are playing, each person gets 6 cards for their hand. How many different 6-card *Clue* hands are possible?
3. Palmetto Cash 5 is a game offered by the South Carolina Education Lottery. Players choose 5 numbers from the whole numbers between 1 and 38 (inclusive); the player wins the jackpot of $100,000 if the randomizer selects those numbers in any order. How many different sets of winning numbers are possible?

Solution

1. A standard deck has 52 cards, and a hand has 2 cards. Since the order doesn't matter, we use the formula for counting combinations:

$$52C_2 = \frac{52!}{2!(52-2)!} = \frac{52 \times 51 \times 50!}{2 \times 1 \times 50!} = \frac{52 \times 51}{2} = 1,326\ .$$

2. Again, the order doesn't matter, so the number of combinations is:

$$21C_6 = \frac{21!}{6!(21-6)!} = \frac{21 \times 20 \times 19 \times 18 \times 17 \times 16 \times 15!}{6 \times 5 \times 4 \times 3 \times 2 \times 1 \times 15!} = \frac{21 \times 19 \times 17 \times 16}{2} = 54,264\ .$$

3. There are 38 numbers to choose from, and we must pick 5. Since order doesn't matter, the number of combinations is:

$$38C_5 = \frac{38!}{5!(38-5)!} = \frac{38 \times 37 \times 36 \times 35 \times 34 \times 33!}{5 \times 4 \times 3 \times 2 \times 1 \times 33!} = 501,492\ .$$

CHECKPOINT

The notation and nomenclature used for the number of combinations is not standard across all sources. You'll sometimes see $\binom{n}{r}$ *instead of* nC_r *. Sometimes you'll hear that expression read as "n choose r" as shorthand for "the number of combinations of n objects taken r at a time."*

PEOPLE IN MATHEMATICS

Early Eastern Mathematicians

Although combinations weren't really studied in Europe until around the 13th century, mathematicians of the Middle and Far East had already been working on them for hundreds of years. The Indian mathematician known as Pingala had described them by the second century BCE; Varāhamihira (fl. sixth century) and Halayudha (fl. 10th century) extended Pingala's work. In the ninth century, a Jain mathematician named Mahāvīra gave the formula for combinations that we use today.

In 10th-century Baghdad, a mathematician named Al-Karaji also knew formulas for combinations; though his work is now lost, it was known to (and repeated by) Persian mathematician Omar Khayyam, whose work survives. Khayyam is probably best remembered as a poet, with his *Rubaiyat* being his most famous work.

Meanwhile, in 11th-century China, Jia Xian also was working with combinations, as was his 13th-century successor Yang Hui.

It is not known whether the discoveries of any of these men were known in the other regions, or if the Indians, Persians, and Chinese all came to their discoveries independently. We do know that mathematical knowledge and sometimes texts did get passed along trade routes, so it can't be ruled out.

Example 4 **Combining Combinations with the Multiplication Rule for Counting**

The student government at a university consists of 10 seniors, 8 juniors, 6 sophomores, and 4 first-years.

1. How many ways are there to choose a committee of 8 people from this group?

2. How many ways are to choose a committee of 8 people if the committee must consist of 2 people from each class?

Solution

1. There are 28 people to choose from, and we need 8. So, the number of possible committees is $28C_8 = 3,108,105$.
2. Break the selection of the committee members down into a 4-step process: Choose the seniors, then choose the juniors, then the sophomores, and then the first-years, as shown in the table below:

Class	Number of Ways to Choose Committee Representatives
senior	$10C_2 = 45$
junior	$8C_2 = 28$
sophomore	$6C_2 = 15$
first-year	$4C_2 = 6$

The Multiplication Rule for Counting tells us that we can get the total number of ways to complete this task by multiplying together the number of ways to do each of the four subtasks. So, there are $45 \times 28 \times 15 \times 6 = 113,400$ possible committees with these restrictions.

Key Terms

- combination

Key Concepts

- Permutations are used to count subsets when order matters; combinations work when order doesn't matter.
- Combinations can also be computed using factorials.

Formulas

- The formula for counting combinations is:
  $nC_r = \frac{n!}{r!(n-r)!}$

7.4 Tree Diagrams, Tables, and Outcomes

Figure 10: In genetics, the characteristics of an offspring organism depends on the characteristics of its parents.

Learning Objectives

After completing this section, you should be able to:

1. Determine the sample space of single stage experiment.
2. Use tables to list possible outcomes of a multistage experiment.
3. Use tree diagrams to list possible outcomes of a multistage experiment.

In the 19th century, an Augustinian friar and scientist named Gregor Mendel used his observations of pea plants to set out his theory of genetic propagation. In his work, he looked at the offspring that resulted from breeding plants with different characteristics together. For applications like this, it is often insufficient to only know in how many ways a process might end; we need to be able to list all of the possibilities. As we've seen, the number of possible outcomes can be very large! Thus, it's important to have a strategy that allows us to systematically list these possibilities to make sure we don't leave any out. In this section, we'll look at two of these strategies.

Single Stage Experiments

When we are talking about combinatorics or probability, the word "**experiment**" has a slightly different meaning than it does in the sciences. Experiments can range from very simple ("flip a coin") to very complex ("count the number of uranium atoms that undergo nuclear fission in a sample of a given size over the course of an hour"). Experiments have unknown outcomes that generally rely on something random, so that if the experiment is repeated (or **replicated**) the outcome might be different. No matter what the experiment, though, analysis of the experiment typically begins with identifying its sample space.

The **sample space** of an experiment is the set of all of the possible outcomes of the experiment, so it's often expressed as a set (i.e., as a list bound by braces; if the experiment is "randomly select a number between 1 and 4," the sample space would be written $\{1, 2, 3, 4\}$).

Example 1 **Finding the Sample Space of an Experiment**

For each of the following experiments, identify the sample space.

1. Flip a coin (which has 2 faces, typically called "heads" and "tails") and note which face is up.
2. Flip a coin 10 times and count the number of heads.
3. Roll a 6-sided die and note the number that is on top.
4. Roll two 6-sided dice and note the sum of the numbers on top.

Solution

1. If we use "H" to denote "heads is facing up" and "T" to denote "tails is facing up", then the sample space is {H, T}.
2. It's possible (though unlikely) that there will be no heads flipped; the outcome in that case would be "0." It's also possible (more likely, but still quite unlikely) that only one flip will result in heads. Any other whole number is possible, up to the maximum: We're flipping the coin 10 times, so we can't get any more than 10 heads. So, the sample space is {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10}.
3. There are 6 numbers on the die: 1, 2, 3, 4, 5, and 6. So, the sample space for a single roll of the die is {1, 2, 3, 4, 5, 6}.
4. If we roll 2 dice, the smallest possible sum we could get is $1 + 1 = 2$ and the biggest is $6 + 6 = 12$. Every other whole number between those two is possible. So, the sample space is {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}.

Multistage Experiments

Some experiments have more complicated sample spaces because they occur in stages. These stages can occur in succession (like drawing cards one at a time) or simultaneously (rolling 2 dice). Sample spaces get more complicated as the complexity of the experiment increases, so it's important to choose a systematic method for identifying all of the possible outcomes. The first method we'll discuss is the table.

Using Tables to Find Sample Spaces

Tables are useful for finding the sample space for experiments that meet two criteria: (1) The experiment must have only two stages, and (2) the outcomes of each stage must have no effect on the outcomes of the other. When the stages do not affect each other, we say the stages are **independent**. Otherwise, the stages are **dependent** and so we can't use tables; we'll look at a method for analyzing dependent stages soon.

Example 2 **Determining Independence**

Decide whether the two stages in these experiments are independent or dependent.

1. You flip a coin and note the result, and then flip the coin again and note the result.
2. You draw 2 cards from a standard deck (52 cards), one at a time.

Solution

1. No matter what happens on the first flip, the second flip has the same sample space: {H, T} (You'll sometimes hear the phrase "The coin has no memory"). So, these stages are independent.
2. Let's say that the first card you draw is A♠. The sample space for the second draw consists of all the cards *except* A♠ (since that card is no longer in the deck, you can't draw it again). If instead that first card was 2♡, the sample space for the second draw is different: it's every card except 2♡. Since the sample space for the second card changes based on the result of the first draw, these stages are dependent.

If you have a two-stage experiment with independent stages, a table is the most straightforward way to identify the sample space. To build a table, you list the outcomes of one stage of the experiment along the top of the table and the outcomes of the other stage down the side. The cells in the interior of the table are then filled using the outcomes associated with each cell's row and column. Let's look at an example.

Example 3 **Using Tables to Identify Sample Spaces**

Identify the sample spaces of these experiments using tables.

1. You roll two dice: one 4-sided and one 6-sided.
2. You're in an ice-cream shop, and you're going to get a single scoop of ice cream with a topping. The flavors of ice cream you're considering are vanilla, chocolate, and rocky road; the toppings are fudge, whipped cream, and sprinkles.
3. The pea plants you're breeding have two possible pod colors: green and yellow. These colors are decided by a particular gene, which comes in two types: "G" for green, and "g" for yellow (In genetics, capital letters usually denote *dominant* genes, while lower-case letters denote *recessive* genes). Each plant has two genes. If you breed a Gg pea plant with a gg plant, the offspring plant will get one gene from each parent. What are the possible outcomes?

Solution

1. **Step 1:** Make the outline of a table, with the results of the 4-sided roll on one side and the results of the 6-sided roll on the other. In practice, it doesn't matter which you choose; for this example, we'll put the 4-sided results on top (labeling the columns of the chart) and the 6-sided results on the side (labeling the rows of the chart) as shown in the following table:

		4-Sided Roll			
		1	2	3	4
6-Sided Roll	**1**				
2					
3					
4					
5					
6					

Step 2: Fill in the results in each cell of the table below. Use the notation of an ordered pair, with the row label first.

		4-Sided Roll			
		1	2	3	4
6-Sided Roll	**1**	(1,1)	(1,2)	(1,3)	(1,4)
2	(2,1)	(2,2)	(2,3)	(2,4)	
3	(3,1)	(3,2)	(3,3)	(3,4)	
4	(4,1)	(4,2)	(4,3)	(4,4)	
5	(5,1)	(5,2)	(5,3)	(5,4)	
6	(6,1)	(6,2)	(6,3)	(6,4)	

So, for example, (3,2) represents the outcome where the 6-sided roll results in a 3, and the 4-sided roll gives us a 2. Thus, the sample space of the experiment is {(1,1), (1,2), (1,3), (1,4), (2,1), (2,2), (2,3), (2,4), (3,1), (3,2), (3,3), (3,4), (4,1), (4,2), (4,3), (4,4), (5,1), (5,2), (5,3), (5,4), (6,1), (6,2), (6,3), (6,4)}.

1. **Step 1:** Let's put the flavors on the rows and the toppings on the columns of the following table:

		Toppings		
		fudge	whipped cream	sprinkles
Flavors	**vanilla**			
chocolate				
rocky road				

Step 2: We can fill in the cells of the table below with the resulting combinations.

		Toppings		
		fudge	whipped cream	sprinkles
Flavors	**vanilla**	vanilla with fudge	vanilla with whipped cream	vanilla with sprinkles
chocolate	chocolate with fudge	chocolate with whipped cream	chocolate with sprinkles	
rocky road	rocky road with fudge	rocky road with whipped cream	rocky road with sprinkles	

So, the sample space is {vanilla with fudge, vanilla with whipped cream, vanilla with sprinkles, chocolate with fudge, chocolate with whipped cream, chocolate with sprinkles, rocky road with fudge, rocky road with whipped cream, rocky road with sprinkles}.

1. **Step 1:** We'll put the parents' genes (P1 and P2) as labels on the rows and columns of the following table:

		P2	
		g	g
P1	**G**		
g			

Step 2: We'll fill in the offspring's gene composition, listing parent 1's gene first in the table below.

		P2	
		g	g
P1	**G**	Gg	Gg
g	gg	gg	

Thus, the sample space is {Gg, Gg, gg, gg}. (Diagrams like this, which allow us to identify the genotypes of offspring, are called *Punnett squares* in honor of Reginald Punnett (1875–1967), who first used them in the context of genetics.)

Using Tree Diagrams to Identify Sample Spaces

In experiments where there are more than two stages, or where the stages are dependent, a **tree diagram** is a helpful tool for systematically identifying the sample space. Tree diagrams are built by first drawing a single point (or *node*), then from that node we draw one branch (a short line segment) for each outcome of the first stage. Each branch gets its own node at the other end (which we typically label with the corresponding outcome for that branch); from each of these, we draw another branch for each outcome of the second stage, assuming that the outcome of the first stage matches the branch we were on. If there are other stages, we can continue from there by continuing to add branches and nodes. This sounds really complicated, but it's easier to understand through an example.

Example 4 **Using a Tree Diagram to Identify a Sample Space**

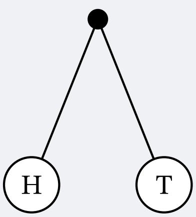
Use a tree diagram to find the sample spaces of each of the following experiments:

1. You flip a coin 3 times, noting the outcome of each flip in order.
2. You flip a coin. If the result is heads, you roll a 4-sided die. If it's tails, you roll a 6-sided die.
3. You are planning to go on a hike with a group of friends. There are 3 trails to consider: Abel Trail, Borel Trail, and Condorcet Trail. One of your friends, Jess, requires a wheelchair; if she joins you, the group couldn't handle the rocky Condorcet Trail.

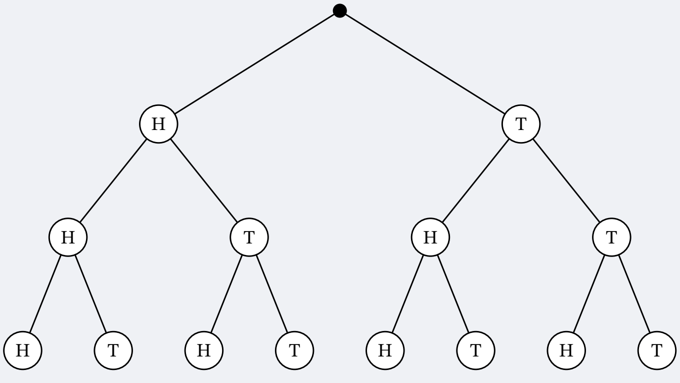
Solution

1. **Step 1:** Let's start by placing our first node .

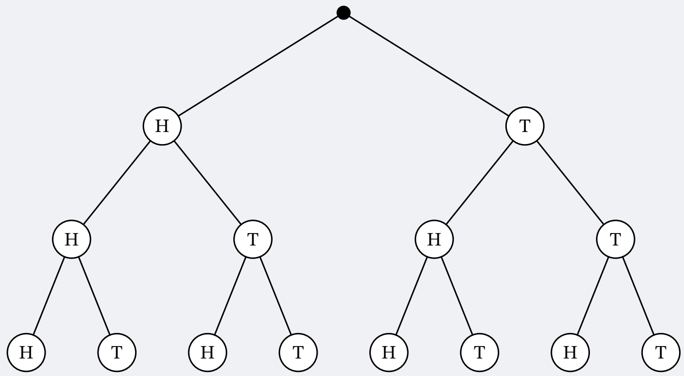
Step 2: We'll add two branches, one for each outcome of the first coin flip, and label them .



Step 3: We're ready for stage two of the experiment: another coin flip. At each node, we add in branches that represent those outcomes .

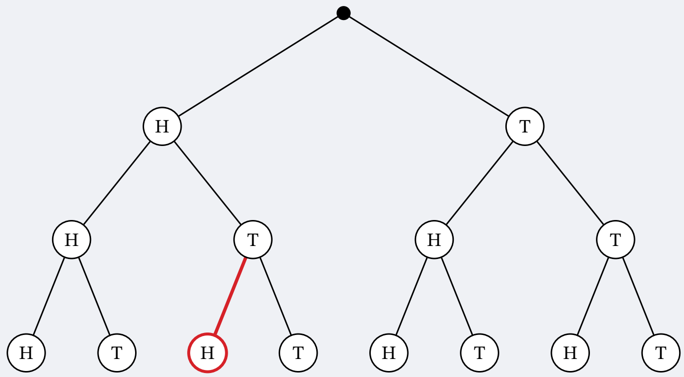


Finally, we can add another set of branches for the outcomes of the third stage .

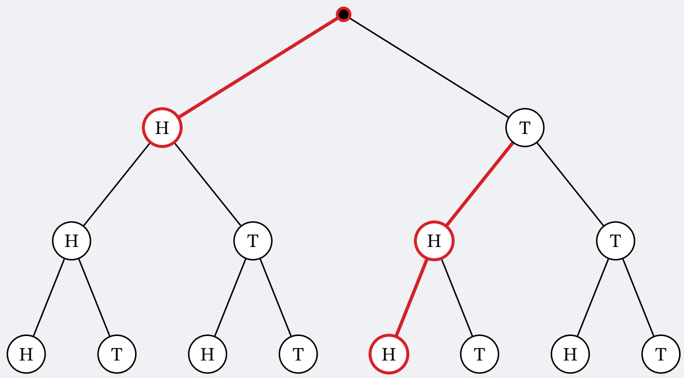


(These final nodes are called *leaves.*)

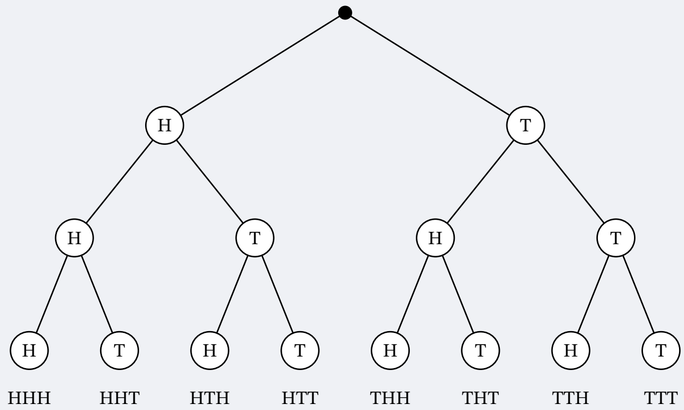
Step 4: We can write down the outcomes in the sample space by tracing the path out to each leaf, writing down the outcome at each node we pass through. For example, this leaf :



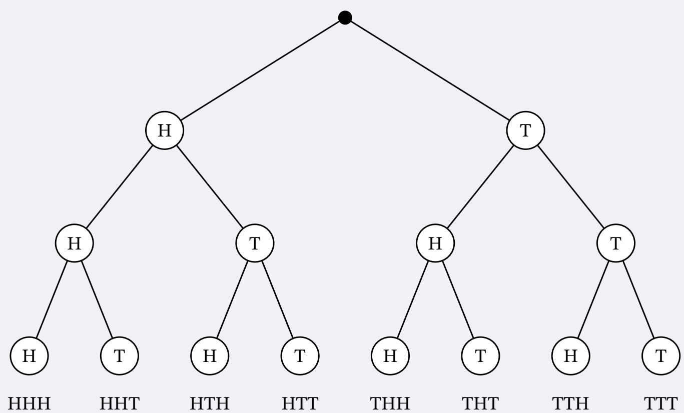
is reached via this path :



Step 5: We'll label that leaf as "HTH", since the path passes through nodes labeled H, T, and H on its way out to our leaf.

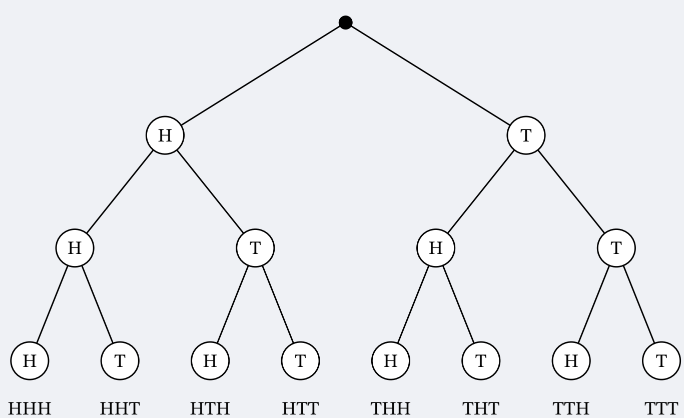


Step 6: We can label the remaining leaves using the same method .

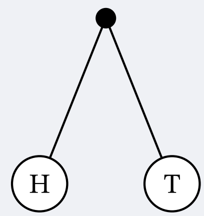


The sample space is the labels on the leaves: {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}.

1. **Step 1:** We'll start with our initial node .

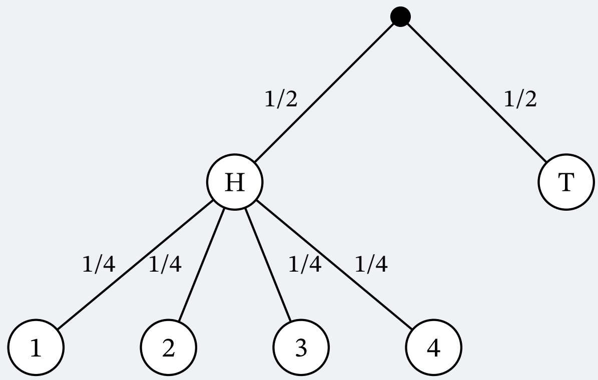


Step 2: We'll add in branches for the outcomes of the first stage , which is the coin flip.

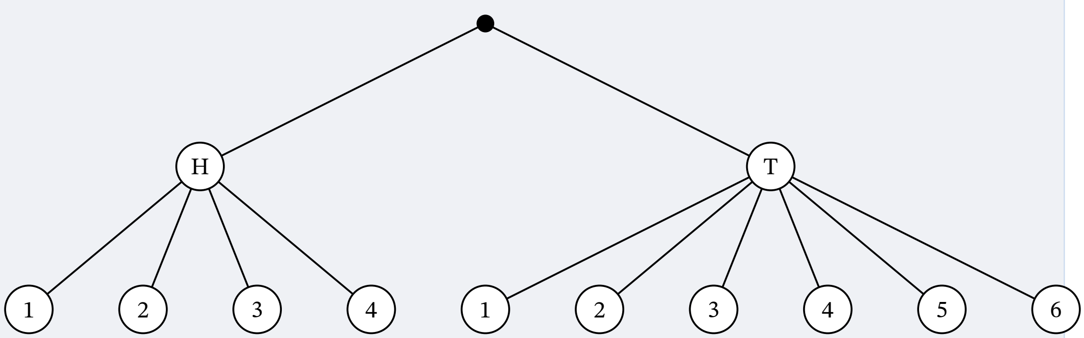


Step 3: The second stage of the experiment depends on the outcome of the first stage.

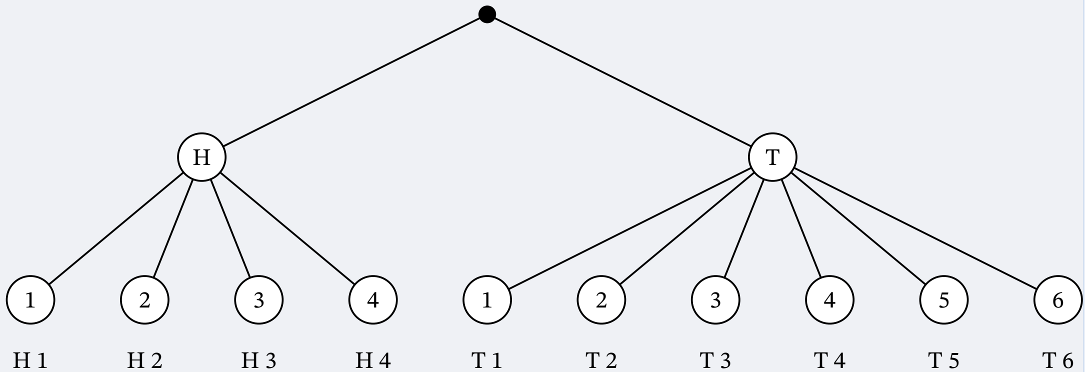
1. If the outcome of the first stage was H, then we roll a 4-sided die. So, *only* on the node for H, we'll add in the outcomes of a 4-sided die roll .



1. If the outcome of the first stage was T, then we roll a 6-sided die. So, we'll add those branches to the node for T .

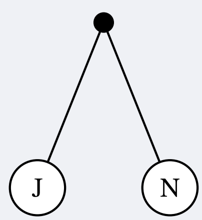


Step 4: We can label the leaves to get the sample space .

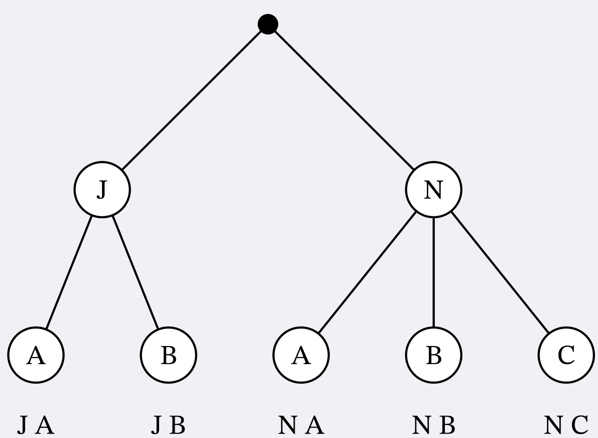


The sample space is: {H1, H2, H3, H4, T1, T2, T3, T4, T5, T6}.

1. **Step 1:** Let's label the trails A, B, and C for ease of labeling. Even though the trails are listed first in the exercise, we can't use the trail choice as our first stage: the trails available to us depend on whether Jess is able to join the trip. So, the first stage is whether Jess joins us (J) or not (N) .



Step 2: We list the appropriate trails on each branch .



So, the sample space is {JA, JB, NA, NB, NC}.

Key Terms

- experiment
- replication

- sample space
- independent/dependent

Key Concepts

- We identify the sample space of an experiment by identifying all of its possible outcomes.
- Tables can help us find a sample space by keeping the possible outcomes organized.
- Tree diagrams provide a visualization of the sample space of an experiment that involves multiple stages.

7.5 Basic Concepts of Probability

Figure 26: When you roll two dice, some outcomes (like rolling a sum of seven) are more likely than others (rolling a sum of twelve).

Learning Objectives

After completing this section, you should be able to:

1. Define probability including impossible and certain events.
2. Calculate basic theoretical probabilities.
3. Calculate basic empirical probabilities.
4. Distinguish among theoretical, empirical, and subjective probability.
5. Calculate the probability of the complement of an event.

It all comes down to this. The game of *Monopoly* that started hours ago is in the home stretch. Your sister has the dice, and if she rolls a 4, 5, or 7 she'll land on one of your best spaces and the game will be over. How likely is it that the game will end on the next turn? Is it more likely than not? How can we measure that likelihood? This section addresses this question by introducing a way to measure uncertainty.

Introducing Probability

Uncertainty is, almost by definition, a nebulous concept. In order to put enough constraints on it that we can mathematically study it, we will focus on uncertainty strictly in the context of experiments. Recall that experiments are processes whose outcomes are unknown; the sample space for the experiment is the collection of all those possible outcomes. When we want to talk about the likelihood of particular outcomes, we sometimes group outcomes together; for example, in the *Monopoly* example at the beginning of this section, we were interested in the roll of 2 dice that might fall as a 4, 5, or 7. A grouping of outcomes that we're interested in is called an **event**. In other words, an event is a subset of the sample space of an experiment; it often consists of the outcomes of interest to the experimenter.

Once we have defined the event that interests us, we can try to assess the likelihood of that event. We do that by assigning a number to each event (E) called the **probability** of that event ($P(E)$). The probability of an event is a number between 0 and 1 (inclusive). If the probability of an event is 0, then the event is impossible. On the other hand, an event with probability 1 is certain to occur. In general, the higher the probability of an event, the more likely it is that the event will occur.

Example 1 **Determining Certain and Impossible Events**

Consider an experiment that consists of rolling a single standard 6-sided die (with faces numbered 1-6). Decide if these probabilities are equal to zero, equal to one, or somewhere in between.

1. $P(\text{roll a } 4)$
2. $P(\text{roll a } 7)$
3. $P(\text{roll a positive number})$
4. $P\left(\text{roll a } \frac{1}{3}\right)$
5. $P(\text{roll an even number})$
6. $P(\text{roll a single-digit number})$

Solution

Let's start by identifying the sample space. For one roll of this die, the possible outcomes are $\{1, 2, 3, 4, 5{,}6\}$. We can use that to assess these probabilities:

1. We see that 4 is in the sample space, so it's possible that it will be the outcome. It's not certain to be the outcome, though. So, $0 < P(\text{roll a } 4) < 1$.
2. Notice that 7 is not in the sample space. So, $P(\text{roll a } 7) = 0$.
3. Every outcome in the sample space is a positive number, so this event is certain. Thus, $P(\text{roll a positive number}) = 1$.

4. Since $\frac{1}{3}$ is not in the sample space, $P\left(\text{roll a } \frac{1}{3}\right) = 0$.
5. Some outcomes in the sample space are even numbers (2, 4, and 6), but the others aren't. So, $0 < P(\text{roll an even number}) < 1$.
6. Every outcome in the sample space is a single-digit number, so $P(\text{roll a single-digit number}) = 1$.

Three Ways to Assign Probabilities

The probabilities of events that are certain or impossible are easy to assign; they're just 1 or 0, respectively. What do we do about those in-between cases, for events that might or might not occur? There are three methods to assign probabilities that we can choose from. We'll discuss them here, in order of reliability.

Method 1: Theoretical Probability

The theoretical method gives the most reliable results, but it cannot always be used. If the sample space of an experiment consists of equally likely outcomes, then the **theoretical probability** of an event is defined to be the ratio of the number of outcomes in the event to the number of outcomes in the sample space.

FORMULA

For an experiment whose sample space S consists of equally likely outcomes, the **theoretical probability** of the event E is the ratio

$$P(E) = \frac{n(E)}{n(S)},$$

where $n(E)$ and $n(S)$ denote the number of outcomes in the event and in the sample space, respectively.

Example 2 **Computing Theoretical Probabilities**

Recall that a standard deck of cards consists of 52 unique cards which are labeled with a rank (the whole numbers from 2 to 10, plus J, Q, K, and A) and a suit (♣, ♢, ♡, or ♠). A standard deck is thoroughly shuffled, and you draw one card at random (so every card has an equal chance of being drawn). Find the theoretical probability of each of these events:

1. The card is 10♠.
2. The card is a ♡.
3. The card is a king (K).

Solution

There are 52 cards in the deck, so the sample space for each of these experiments has 52 elements. That will be the denominator for each of our probabilities.

1. There is only one 10♠ in the deck, so this event only has one outcome in it. Thus, $P(10♠) = \frac{1}{52}$.
2. There are 13 ♡ s in the deck, so $P(♡) = \frac{13}{52} = \frac{1}{4}$.
3. There are 4 cards of each rank in the deck, so $P(\mathrm{K}) = \frac{4}{52} = \frac{1}{13}$.

CHECKPOINT

It is critical that you make sure that every outcome in a sample space is equally likely before you compute theoretical probabilities!

Example 3 Using Tables to Find Theoretical Probabilities

In the Basic Concepts of Probability, we were considering a *Monopoly* game where, if your sister rolled a sum of 4, 5, or 7 with 2 standard dice, you would win the game. What is the probability of this event? Use tables to determine your answer.

Solution

We should think of this experiment as occurring in two stages: (1) one die roll, then (2) another die roll. Even though these two stages will usually occur simultaneously in practice, since they're independent, it's okay to treat them separately.

Step 1: Since we have two independent stages, let's create a table , which is probably the most efficient method for determining the sample space.

	first die					
second die	**1**	**2**	**3**	**4**	**5**	**6**
1	(1, 1)	(2, 1)	(3, 1)	(4, 1)	(5, 1)	(6, 1)
2	(1, 2)	(2, 2)	(3, 2)	(4, 2)	(5, 2)	(6, 2)
3	(1, 3)	(2, 3)	(3, 3)	(4, 3)	(5, 3)	(6, 3)
4	(1, 4)	(2, 4)	(3, 4)	(4, 4)	(5, 4)	(6, 4)
5	(1, 5)	(2, 5)	(3, 5)	(4, 5)	(5, 5)	(6, 5)
6	(1, 6)	(2, 6)	(3, 6)	(4, 6)	(5, 6)	(6, 6)

Now, each of the 36 ordered pairs in the table represent an equally likely outcome.

Step 2: To make our analysis easier, let's replace each ordered pair with the sum .

	first die					
second die	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Step 3: Since the event we're interested in is the one consisting of rolls of 4, 5, or 7. Let's shade those in .

	first die					
second die	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Our event contains 13 outcomes, so the probability that your sister rolls a losing number is $\frac{13}{36}$.

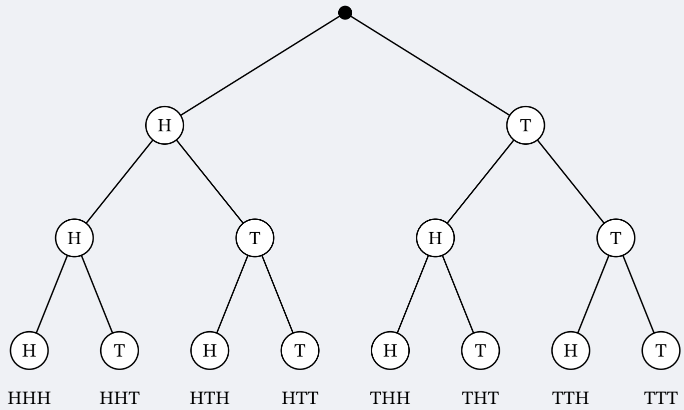
Example 4 **Using Tree Diagrams to Compute Theoretical Probability**

If you flip a fair coin 3 times, what is the probability of each event? Use a tree diagram to determine your answer

1. You flip exactly 2 heads.
2. You flip 2 consecutive heads at some point in the 3 flips.
3. All 3 flips show the same result.

Solution

Let's build a tree to identify the sample space .



The sample space is {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}, which has 8 elements.

1. Flipping exactly 2 heads occurs three times (HHT, HTH, THH), so the probability is $\frac{3}{8}$.
2. Flipping 2 consecutive heads at some point in the experiment happens 3 times: HHH, HHT, THH. So, the probability is $\frac{3}{8}$.
3. There are 2 outcomes that all show the same result: HHH and TTT. So, the probability is $\frac{2}{8} = \frac{1}{4}$.

PEOPLE IN MATHEMATICS

Gerolamo Cardano

The first known text that provided a systematic approach to probabilities was written in 1564 by Gerolamo Cardano (1501–1576). Cardano was a physician whose illegitimate birth closed many doors that would have otherwise been open to someone with a medical degree in 16th-century Italy. As a result, Cardano often turned to gambling to help ends meet. He was a remarkable mathematician, and he used his knowledge to gain an edge when playing at cards or dice. His 1564 work, titled *Liber de ludo aleae* (which translates as *Book on Games of Chance*), summarized everything he knew about probability. Of course, if that book fell into the hands of those he played against, his advantage would disappear. That's why he never allowed it to be published in his lifetime (it was eventually published in 1663). Cardano made other contributions to mathematics; he was the first person to publish the third degree analogue of the Quadratic Formula (though he didn't discover it himself), and he popularized the use of negative numbers.

Method 2: Empirical Probability

Theoretical probabilities are precise, but they can't be found in every situation. If the outcomes in the sample space are not equally likely, then we're out of luck. Suppose you're watching a baseball game, and your favorite player is about to step up to the plate. What is the probability that he will get a hit?

In this case, the sample space is {hit, not a hit}. That doesn't mean that the probability of a hit is $\frac{1}{2}$, since those outcomes aren't equally likely. The theoretical method simply can't be used in this situation. Instead, we might look at the player's statistics up to this point in the season, and see that he has 122 hits in 531 opportunities. So, we might think that the probability of a hit in the next plate appearance would be about $\frac{122}{531} \approx 0.23$. When we use the outcomes of previous replications of an experiment to assign a probability to the next replication, we're defining an **empirical probability**. Empirical probability is assigned using the outcomes of previous replications of an experiment by finding the ratio of the number of times in the previous replications the event occurred to the total number of previous replications.

Empirical probabilities aren't exact, but when the number of previous replications is large, we expect them to be close. Also, if the previous runs of the experiment are not conducted under the exact set of circumstances as the one we're interested in, the empirical probability is less reliable. For instance, in the case of our favorite baseball player, we might try to get a better estimate of the probability of a hit by looking only at his history against left- or right-handed pitchers (depending on the handedness of the pitcher he's about to face).

WHO KNEW?

Probability and Statistics

One of the broad uses of statistics is called statistical inference, where statisticians use collected data to make a guess (or inference) about the population the data were collected from. Nearly every tool that statisticians use for inference is based on probability. Not only is the method we just described for finding empirical probabilities one type of statistical inference, but some more advanced techniques in the field will give us an idea of how close that empirical probability might be to the actual probability!

Example 5 **Finding Empirical Probabilities**

Assign an empirical probability to the following events:

1. Jose is on the basketball court practicing his shots from the free throw line. He made 47 out of his last 80 attempts. What is the probability he makes his next shot?
2. Amy is about to begin her morning commute. Over her last 60 commutes, she arrived at work 12 times in under half an hour. What is the probability that she arrives at work in 30 minutes or less?
3. Felix is playing *Yahtzee* with his sister. Felix won 14 of the last 20 games he played against her. How likely is he to win this game?

Solution

1. Since Jose made 47 out of his last 80 attempts, assign this event an empirical probability of $\frac{47}{80} \approx 59\%$.

2. Amy completed the commute in under 30 minutes in 12 of the last 60 commutes, so we can estimate her probability of making it in under 30 minutes this time at $\frac{12}{60} = 20\%$.
3. Since Felix has won 14 of the last 20 games, assign a probability for a win this time of $\frac{14}{20} = 70\%$.

WORK IT OUT

Buffon's Needle

A famous early question about probability (posed by Georges-Louis Leclerc, Comte de Buffon in the 18th century) had to do with the probability that a needle dropped on a floor finished with wooden slats would lay across one of the seams. If the distance between the slats is exactly the same length as the needle, then it can be shown using calculus that the probability that the needle crosses a seam is $\frac{2}{\pi}$. Using toothpicks or matchsticks (or other uniformly long and narrow objects), assign an empirical probability to this experiment by drawing parallel lines on a large sheet of paper where the distance between the lines is equal to the length of your dropping object, then repeatedly dropping the objects and noting whether the object touches one of the lines. Once you have your empirical probability, take its reciprocal and multiply by 2. Is the result close to π?

Method 3: Subjective Probability

In cases where theoretical probability can't be used and we don't have prior experience to inform an empirical probability, we're left with one option: using our instincts to guess at a **subjective probability**. A subjective probability is an assignment of a probability to an event using only one's instincts.

Subjective probabilities are used in cases where an experiment can only be run once, or it hasn't been run before. Because subjective probabilities may vary widely from person to person and they're not based on any mathematical theory, we won't give any examples. However, it's important that we be able to identify a subjective probability when we see it; they will in general be far less accurate than empirical or theoretical probabilities.

Example 6 **Distinguishing among Theoretical, Empirical, and Subjective Probabilities**

Classify each of the following probabilities as theoretical, empirical, or subjective.

1. An eccentric billionaire is testing a brand new rocket system. He says there is a 15% chance of failure.
2. With 4 seconds to go in a close basketball playoff game, the home team need 3 points to tie up the game and send it to overtime. A TV commentator says that team captain should take the final 3-point shot, because he has a 38% chance of making it (greater than every other player on the team).

3. Felix is losing his *Yahtzee* game against his sister. He has one more chance to roll 2 dice; he'll win the game if they both come up 4. The probability of this is about 2.8%.

Solution

1. This experiment has never been run before, so the given probability is subjective.
2. Presumably, the commentator has access to each player's performance statistics over the entire season. So, the given probability is likely empirical.
3. Rolling 2 dice results in a sample space with equally likely outcomes. This probability is theoretical. (We'll learn how to calculate that probability later in this chapter.)

WHO KNEW?

Benford's Law

In 1938, Frank Benford published a paper ("The law of anomalous numbers," in *Proceedings of the American Philosophical Society*) with a surprising result about probabilities. If you have a list of numbers that spans at least a couple of orders of magnitude (meaning that if you divide the largest by the smallest, the result is at least 100), then the digits 1–9 are not equally likely to appear as the first digit of those numbers, as you might expect. Benford arrived at this conclusion using empirical probabilities; he found that 1 was about 6 times as likely to be the initial digit as 9 was!

New Probabilities from Old: Complements

One of the goals of the rest of this chapter is learning how to break down complicated probability calculations into easier probability calculations. We'll look at the first of the tools we can use to accomplish this goal in this section; the rest will come later.

Given an event E, the **complement** of E (denoted E') is the collection of all of the outcomes that are *not* in E. (This is language that is taken from set theory, which you can learn more about elsewhere in this text.) Since every outcome in the sample space either is or is not in E, it follows that $n(E) + n(E') = n(S)$. So, if the outcomes in S are equally likely, we can compute theoretical probabilities $P(E) = \frac{n(E)}{n(S)}$ and $P(E') = \frac{n(E')}{n(S)}$. Then, adding these last two equations, we get

$$\begin{aligned} P(E) + P(E') &= \frac{n(E)}{n(S)} + \frac{n(E')}{n(S)} \\ &= \frac{n(E) + n(E')}{n(S)} \\ &= \frac{n(S)}{n(S)} \\ &= 1 \end{aligned}$$

Thus, if we subtract $P(E')$ from both sides, we can conclude that $P(E) = 1 - P(E')$. Though we performed this calculation under the assumption that the outcomes in S are all equally likely, the last equation is true in every situation.

FORMULA

$$P(E) = 1 - P(E')$$

How is this helpful? Sometimes it is easier to compute the probability that an event *won't* happen than it is to compute the probability that it *will*. To apply this principle, it's helpful to review some tricks for dealing with inequalities. If an event is defined in terms of an inequality, the complement will be defined in terms of the opposite inequality: Both the direction and the inclusivity will be reversed, as shown in the table below.

If E is defined with:	then E' is defined with:
$<$	$\geq$
$\leq$	$>$
$>$	$\leq$
$\geq$	$<$

Example 7 **Using the Formula for Complements to Compute Probabilities**

1. If you roll a standard 6-sided die, what is the probability that the result will be a number greater than one?
2. If you roll two standard 6-sided dice, what is the probability that the sum will be 10 or less?
3. If you flip a fair coin 3 times, what is the probability that at least one flip will come up tails?

Solution

1. Here, the sample space is $\{1, 2, 3, 4, 5, 6\}$. It's easy enough to see that the probability in question is $\frac{5}{6}$, because there are 5 outcomes that fall into the event "roll a number greater than 1." Let's also apply our new formula to find that probability. Since E is defined using the inequality roll > 1, then E' is defined using roll ≤ 1. Since there's only one outcome (1) in E', we have $P(E') = \frac{1}{6}$. Thus, $P(E) = 1 - P(E') = \frac{5}{6}$.
2. In Example 3, we found the following table of equally likely outcomes for rolling 2 dice :

	first die					
second die	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Here, the event E is defined by the inequality sum ≤ 10. Thus, E' is defined by sum $>$ 10. There are three outcomes in E': two 11s and one 12. Thus, $P(E) = 1 - P(E') = 1 - \frac{3}{36} = \frac{11}{12}$.

1. In Example 4 in Section 7.4, we found the sample space for this experiment consisted of these equally likely outcomes: {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}. Our event E is defined by $\text{T} \geq 1$, so E' is defined by $\text{T} < 1$. The only outcome in E' is the first one on the list, where zero tails are flipped. So, $P(E) = 1 - P(E') = 1 - \frac{1}{8} = \frac{7}{8}$.

Key Terms

- event
- probability
- theoretical probability
- empirical probability
- subjective probability

Key Concepts

- The theoretical probability of an event is the ratio of the number of equally likely outcomes in the event to the number of equally likely outcomes in the sample space.
- Empirical probabilities are computed by repeating the experiment many times, and then dividing the number of replications that result in the event of interest by the total number of replications.
- Subjective probabilities are assigned based on subjective criteria, usually because the experiment can't be repeated and the outcomes in the sample space are not equally likely.
- The probability of the complement of an event is found by subtracting the probability of the event from one.

Formulas

- For an experiment whose sample space S consists of equally likely outcomes, the *theoretical probability* of the event E is the ratio $P(E) = \frac{n(E)}{n(S)}$ where $n(E)$ and $n(S)$ denote the number of outcomes in the event and in the sample space, respectively.
- $P(E) = 1 - P(E')$

7.6 Probability with Permutations and Combinations

Figure 32: Bingo and many lottery games depend on selecting one or more numbers at random from a list; often this is done by drawing numbered balls from a bin.

Learning Objectives

After completing this section, you should be able to:

1. Calculate probabilities with permutations.

2. Calculate probabilities with combinations.

   In our earlier discussion of theoretical probabilities, the first step we took was to write out the sample space for the experiment in question. For many experiments, that method just isn't practical. For example, we might want to find the probability of drawing a particular 5-card poker hand. Since there are 52 cards in a deck and the order of cards doesn't matter, the sample space for this experiment has $52C_5 = 2,598,960$ possible 5-card hands. Even if we had the patience and space to write them all out, sorting through the results to find the outcomes that fall in our event would be just as tedious.

   Luckily, the formula for theoretical probabilities doesn't require us to know every outcome in the sample space; we just need to know how many outcomes there are. In this section, we'll apply the techniques we learned earlier in the chapter (The Multiplication Rule for Counting, permutations, and combinations) to compute probabilities.

Using Permutations to Compute Probabilities

Recall that we can use permutations to count how many ways there are to put a number of items from a list in order. If we're looking at an experiment whose sample space looks like an ordered list, then permutations can help us to find the right probabilities.

Example 1 **Using Permutations to Compute Probabilities**

1. In horse racing, an exacta bet is one where the player tries to predict the top two finishers in particular race in order. If there are 9 horses in a race, and a player decided to make an exacta bet at random, what is the probability that they win?
2. You are in a club with 10 people, 3 of whom are close friends of yours. If the officers of this club are chosen at random, what is the probability that you are named president and one of your friends is named vice president?
3. A bag contains slips of paper with letters written on them as follows: A, A, B, B, B, C, C, D, D, D, D, E. If you draw 3 slips, what is the probability that the letters will spell out (in order) the word BAD?

Solution

1. Since order matters for this situation, we'll use permutations. How many different exacta bets can be made? Since there are 9 horses and we must select 2 in order, we know there are $9P_2 = 72$ possible outcomes. That's the size of our sample space, so it will go in the denominator of the probability. Since only one of those outcomes is a winner, the numerator of the probability is 1. So, the probability of randomly selecting the winning exacta bet is $\frac{1}{72}$.
2. There are 10 people in the club, and 2 will be chosen to be officers. Since the order matters, there are $10P_2 = 90$ different ways to select officers. Next, we must figure out how many outcomes are in our event. We'll use the Multiplication Rule for Counting to find that number. There is only 1 choice for president in our event, and there are 3 choices for vice president. So, there are $1 \times 3 = 3$ outcomes in the event. Thus, the

probability that you will serve as president with one of your friends as vice president is $\frac{3}{90} = \frac{1}{30}$.

3. There are 12 slips of paper in the bag, and 3 will be drawn. So, there are $12P_3 = 1320$ possible outcomes. Now, we'll compute the number of outcomes in our event. The first letter drawn must be a B, and there are 3 of those. Next must come an A (2 of those) and then a D (4 of those). Thus, there are $3 \times 2 \times 4 = 24$ outcomes in our event. So, the probability that the letters drawn spell out the word BAD is $\frac{24}{1320} = \frac{1}{55}$.

Combinations to Computer Probabilities

If the sample space of our experiment is one in which order doesn't matter, then we can use combinations to find the number of outcomes in that sample space.

Example 2 **Using Combinations to Compute Probabilities**

1. Palmetto Cash 5 is a game offered by the South Carolina Education Lottery. Players choose 5 numbers from the whole numbers between 1 and 38 (inclusive); the player wins the jackpot of $100,000 if the randomizer selects those numbers in any order. If you buy one ticket for this game, what is the probability that you win the top prize by choosing all 5 winning numbers?
2. There's a second prize in the Palmetto Cash 5 game that a player wins if 4 of the player's 5 numbers are among the 5 winning numbers. What's the probability of winning the second prize?
3. *Scrabble* is a word-building board game. Players make hands of 7 letters by selecting tiles with single letters printed on them blindly from a bag (2 tiles have nothing printed on them; these blanks can stand for any letter). Players use the letters in their hands to spell out words on the board. Initially, there are 100 tiles in the bag. Of those, 44 are (or could be) vowels (9 As, 12 Es, 9 Is, 8 Os, 4 Us, and 2 blanks; we'll treat Y as a consonant). What is the probability that your initial hand has no vowels?

Solution

1. There are 38 numbers to choose from, and the order of the 5 we pick doesn't matter. So, there are $38C_5 = 501,492$ outcomes in the sample space. Only one outcome is in our winning event, so the probability of winning is $\frac{1}{501,492}$.
2. As in part 1 of this example,, there are 501,492 outcomes in the sample space. The tricky part here is figuring out how many outcomes are in our event. To qualify, the outcome must contain 4 of the 5 winning numbers, plus one losing number. There are $5C_4 = 5$ ways to choose the 4 winning numbers, and there are $38 - 5 = 33$ losing numbers. So, using the Multiplication Rule for Counting, there are $5 \times 33 = 165$ outcomes in our event. Thus, the probability of winning the second prize is $\frac{165}{501,492} = \frac{55}{167,164}$, which is about 0.00033.

3. The number of possible starting hands is $100C_7 = 16,007,560,800$. There are $100 - 44 = 56$ consonants in the bag, so the number of all-consonant hands is $56C_7 = 231,917,400$. Thus, the probability of drawing all consonants is $\frac{231,917,40}{16,007,560,800} = \frac{32,139}{2,425,388} \approx 0.0145$.

Key Concepts

- We use permutations and combinations to count the number of equally likely outcomes in an event and in a sample space, which allows us to compute theoretical probabilities.

7.7 What Are the Odds?

Figure 33: Scratch-off lottery tickets, as well as many other games, represent the likelihood of winning using odds.

Learning Objectives

After completing this section, you should be able to:

1. Compute odds.
2. Determine odds from probabilities.
3. Determine probabilities from odds.

A particular lottery instant-win game has 2 million tickets available. Of those, 500,000 win a prize. If there are 500,000 winners, then it follows that there are 1,500,000 losing tickets. When we evaluate the risk associated with a game like this, it can be useful to compare the number of ways to win the game to the number of ways to lose. In the case of this game, we would compare the 500,000 wins to the 1,500,000 losses. In other words, there are 3 losing tickets for every winning ticket. Comparisons of this type are the focus of this section.

Computing Odds

The ratio of the number of equally likely outcomes in an event E to the number of equally likely outcomes *not* in the event E' is called the **odds for** (or **odds in favor** of) the event. The opposite ratio (the number of outcomes not in the event to the number in the event E' to the number in the event E is called the **odds against** the event.

CHECKPOINT
Both odds and probabilities are calculated as ratios. To avoid confusion, we will always use fractions, decimals, or percents for probabilities, and we'll use colons to indicate odds. The rules for simplifying fractions apply to odds, too. Thus, the odds for winning a prize in the game described in the section opener are $500,000 : 1,500,000 = 1 : 3$ *and the odds against winning a prize are* $3 : 1$. *These would often be described in words as "the odds of winning are one to three in favor" or "the odds of winning are three to one against.".*

CHECKPOINT
Notice that, while probabilities must always be between zero and one inclusive, odds can be any (non-negative) number, as we'll see in the next example.

Example 1 **Computing Odds**

1. If you roll a fair 6-sided die, what are the odds for rolling a 5 or higher?
2. If you roll two fair 6-sided dice, what are the odds against rolling a sum of 7?
3. If you draw a card at random from a standard deck, what are the odds for drawing a ♡?
4. If you draw 2 cards at random from a standard deck, what are the odds against them both being ♠?

Solution

1. The sample space for this experiment is $\{1, 2, 3, 4, 5, 6\}$. Two of those outcomes are in the event "roll a five or higher," while four are not. So, the odds for rolling a five or higher are $2 : 4 = 1 : 2$.

2. In Example 3 in Section 7.5, we found the sample space for this experiment using the following table :

	first die					
second die	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

There are 6 outcomes in the event "roll a sum of 7," and there are 30 outcomes not in the event. So, the odds against rolling a 7 are $30 : 6 = 5 : 1$.

1. There are 13 $\heartsuit$ in a standard deck, and $52 - 13 = 39$ others. So, the odds in favor of drawing a $\heartsuit$ are $13 : 39 = 1 : 3$.
2. There are $13C_2 = 78$ ways to draw 2 $\spadesuit$, and $52C_2 - 78 = 1,248$ ways to draw 2 cards that are *not* both $\spadesuit$. So, the odds against drawing 2 $\spadesuit$ are $1,248 : 78 = 16 : 1$.

Odds as a Ratio of Probabilities

We can also think of odds as a ratio of probabilities. Consider again the instant-win game from the section opener, with 500,000 winning tickets out of 2,000,000 total tickets. If a player buys one ticket, the probability of winning is $\frac{500,000}{2,000,000} = \frac{1}{4}$, and the probability of losing is $1 - \frac{1}{4} = \frac{3}{4}$. Notice that the ratio of the probability of winning to the probability of losing is $\frac{1}{4} : \frac{3}{4} = 1 : 3$, which matches the odds in favor of winning.

FORMULA

For an event E,

$$\text{odds for } E = n(E) : n(E') = P(E) : P(E') = P(E) : (1 - P(E))$$
$$\text{odds against } E = n(E') : n(E) = P(E') : P(E) = (1 - P(E)) : P(E)$$

We can use these formulas to convert probabilities to odds, and vice versa.

Example 2 **Converting Probabilities to Odds**

Given the following probabilities of an event, find the corresponding odds for and odds against that event.

1. $P(E) = \frac{3}{5}$
2. $P(E) = 17\%$

Solution

1. Using the formula, we have:

$$\begin{aligned} \text{odds for } E = P(E) : (1 - P(E)) \\ = \frac{3}{5} : \left(1 - \frac{3}{5}\right) \\ = \frac{3}{5} : \frac{2}{5} \\ = 3 : 2. \end{aligned}$$

(Note that in the last step, we simplified by multiplying both terms in the ratio by 5.)
Since the odds for E are 3 : 2, the odds against E must be 2 : 3.

1. Again, we'll use the formula:

$$\begin{aligned} \text{odds for } E = P(E) : (1 - P(E)) \\ = 0.17 : (1 - 0.17) \\ = 0.17 : 0.83 \\ \approx 1 : 4.88. \end{aligned}$$

(In the last step, we simplified by dividing both terms in the ratio by 0.17.)
It follows that the odds against E are approximately 4.88 : 1.

Now, let's convert odds to probabilities. Let's say the odds for an event are $A : B$. Then, using the formula above, we have $A : B = P(E) : (1 - P(E))$. Converting to fractions and solving for $P(E)$, we get:

$$\begin{aligned} \frac{A}{B} &= \frac{P(E)}{1 - P(E)} \\ A(1 - P(E)) &= B \times P(E) \\ A - A \times P(E) &= B \times P(E) \\ A &= A \times P(E) + B \times P(E) \\ A &= (A + B) \times P(E) \\ \frac{A}{A + B} &= P(E). \end{aligned}$$

Let's put this result in a formula we can use.

FORMULA

If the odds in favor of E are $A : B$, then

$$P(E) = \frac{A}{A + B}$$

.

Example 3 **Converting Odds to Probabilities**

Find $P(E)$ if E:

1. The odds of E are $2:1$ in favor
2. The odds of E are $6:1$ against

Solution

1. Using the formula we just found, we have $P(E) = \frac{2}{2+1} = \frac{2}{3}$.
2. If the odds against are $6:1$, then the odds for are $1:6$. Thus, using the formula, $P(E) = \frac{1}{1+6} = \frac{1}{7}$.

CHECKPOINT

Some places, particularly state lottery websites, will use the words "odds" and "probability" interchangeably. Never assume that the word "odds" is being used correctly! Compute one of the odds/probabilities yourself to make sure you know how the word is being used!

Key Terms

- odds (for/against)

Key Concepts

- Odds are computed as the ratio of the probability of an event to the probability of its compliment.

Formulas

- For an event E,

$$\text{odds for } E = n(E) : n(E') = P(E) : P(E') = P(E) : (1 - P(E))$$
$$\text{odds against } E = n(E') : n(E) = P(E') : P(E) = (1 - P(E)) : P(E)$$

- If the odds in favor of E are $A : B$, then $P(E) = \frac{A}{A+B}$.

7.8 The Addition Rule for Probability

Figure 35: Students can be sorted using a variety of possible categories like class year, major, whether they are a varsity athlete, and so forth.

Learning Objectives

After completing this section, you should be able to:

1. Identify mutually exclusive events.
2. Apply the Addition Rule to compute probability.
3. Use the Inclusion/Exclusion Principle to compute probability.

   Up to this point, we have looked at the probabilities of *simple* events. Simple events are those with a single, simple characterization. Sometimes, though, we want to investigate more complicated situations. For example, if we are choosing a college student at random, we might want to find the probability that the chosen student is a varsity athlete *or* in a Greek organization. This is a **compound event**: there are two possible criteria that might be met. We might instead try to identify the probability that the chosen student is *both* a varsity athlete *and* in a Greek organization. In this section and the next, we'll cover probabilities of two types of compound events: those build using "or" and those built using "and." We'll deal with the former first.

Mutual Exclusivity

Before we get to the key techniques of this section, we must first introduce some new terminology. Let's say you're drawing a card from a standard deck. We'll consider 3 events: H is the event "the card is a ♡," T is the event "the card is a 10," and S is the event "the card is a ♠." If the card drawn is J ♠, then H and T didn't occur, but S did. If the card drawn is instead 10♠, then H didn't occur, but both T and S did.

We can see from these examples that, if we are interested in several possible events, more than one of them can occur simultaneously (both T and S, for example). But, if you think about all the possible outcomes, you can see that H and S can *never* occur simultaneously; there are no cards in the deck that are both ♡ and ♠. Pairs of events that cannot both occur simultaneously are called **mutually exclusive**. Let's go through an example to help us better understand this concept.

Example 1 **Identifying Mutually Exclusive Events**

Decide whether the following events are mutually exclusive. If they are not mutually exclusive, identify an outcome that would result in both events occurring.

1. You are about to roll a standard 6-sided die. E is the event "the die shows an even number" and F is the event "the die shows an odd number."
2. You are about to roll a standard 6-sided die. E is the event "the die shows an even number" and S is the event "the die shows a number less than 4."
3. You are about to flip a coin 4 times. J is the event "at least 2 heads are flipped" and K is the event "fewer than 3 tails are flipped."

Solution

1. Let's look at the outcomes for each event: $E = \{2, 4, 6\}$ and $F = \{1, 3, 5\}$. There are no outcomes in common, so E and F are mutually exclusive.
2. Again, consider the outcomes in each event: $E = \{2, 4, 6\}$ and $S = \{1, 2, 3\}$. Since the outcome 2 belongs to both events, these are *not* mutually exclusive.
3. Suppose the results of the 4 flips are HTTH. Then at least 2 heads are flipped, and fewer than 3 tails are flipped. That means that both J and K occurred, and so these events are *not* mutually exclusive.

The Addition Rule for Mutually Exclusive Events

If two events are mutually exclusive, then we can use addition to find the probability that one or the other event occurs.

FORMULA

If E and F are mutually exclusive events, then

$$P(E \text{ or } F) = P(E) + P(F)$$

Why does this formula work? Let's consider a basic example. Suppose we're about to draw a *Scrabble* tile from a bag containing A, A, B, E, E, E, R, S, S, U. What is the probability of drawing an E or an S? Since 3 of the tiles are marked with E and 2 are marked with S, there are 5 tiles that satisfy the criteria. There are ten tiles in the bag, so the probability is $\frac{5}{10} = \frac{1}{2}$. Notice that the probability of drawing an E is $\frac{3}{10}$ and the probability of drawing an S is $\frac{2}{10}$ adding those together, we get $\frac{3}{10} + \frac{2}{10} = \frac{5}{10}$. Look at the numerators in the fractions involved in the sum: the 3 represents the number of E tiles and the 2 is the number of S tiles. This is why the Addition Rule works: The total number of outcomes in one event or the other is the sum of the numbers of outcomes in each of the individual events.

Example 2 **Using the Addition Rule**

For each of the given pairs of events, decide if the Addition Rule applies. If it does, use the Addition Rule to find the probability that one or the other occurs.

1. You are rolling a standard 6-sided die. Event A is "roll an even number" and event B is "roll a 3."
2. You are drawing a card at random from a standard 52-card deck. Event R is "draw a ♡" and event S is "draw a king."
3. You are rolling a pair of standard 6-sided dice. Event E is "roll an odd sum" and event F is "roll a sum of 10." The table we constructed in Example 3 in Section 7.5 might help.

	first die					
second die	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Solution

1. Since 3 is not an even number, these events are mutually exclusive. So, we can use the Addition Rule: since $P(A) = \frac{3}{6}$ and $P(B) = \frac{1}{6}$, we get $P(A \text{ or } B) = \frac{3}{6} + \frac{1}{6} = \frac{2}{3}$.
2. If the card drawn is K ♡, then both R and S occur. So, they aren't mutually exclusive, and the Addition Rule doesn't apply.
3. Since 10 is not odd, these events are mutually exclusive. Since $P(E) = \frac{18}{36}$ and $P(F) = \frac{3}{36}$, the Addition Rule gives us $P(E \text{ or } F) = \frac{18}{36} + \frac{3}{36} = \frac{7}{12}$.

Finding Probabilities When Events Aren't Mutually Exclusive

Let's return to the example we used to explore the Addition Rule: We're about to draw a *Scrabble* tile from a bag containing A, A, B, E, E, E, R, S, S, U. Consider these events: J is "draw a vowel" and K is "draw a letter that comes after L in the alphabet." Since there are 6 vowels, $P(J) = \frac{6}{10}$. There are 4 tiles with letters that come after L alphabetically, so $P(K) = \frac{4}{10}$. What is $P(J \text{ or } K)$? If we blindly apply the Addition Rule, we get $\frac{6}{10} + \frac{4}{10} = 1$, which would mean that the compound event J or K is certain. However, it's possible to draw a B, in which case neither J nor K happens. Where's the error?

The events are not mutually exclusive: the outcome U belongs to both events, and so the Addition Rule doesn't apply. However, there's a way to extend the Addition Rule to allow us to find this probability anyway; it's called the **Inclusion/Exclusion Principle**. In this example, if we just add the two probabilities together, the outcome U is included in the sum twice: It's one of the 6 outcomes represented in the numerator of $\frac{6}{10}$, and it's one of the 4 outcomes represented in the numerator of $\frac{4}{10}$. So, that particular outcome has been "double counted." Since it has been *included* twice, we can get a true accounting by *excluding* it once: $\frac{6}{10} + \frac{4}{10} - \frac{1}{10} = \frac{9}{10}$. We can generalize this idea to a formula that we can apply to find the probability of any compound event built using "or."

FORMULA

Inclusion/Exclusion Principle: If E and F are events that contain outcomes of a single experiment, then

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

It's worth noting that this formula is truly an extension of the Addition Rule. Remember that the Addition Rule requires that the events E and F are mutually exclusive. In that case, the compound event (E and F) is impossible, and so $P(E \text{ and } F) = 0$. So, in cases where the events in question are mutually exclusive, the Inclusion/Exclusion Principle reduces to the Addition Rule.

Example 3 **Using the Inclusion/Exclusion Principle**

Suppose we have events E, F, and G, associated with these probabilities:

$$\begin{aligned} P(E) &= 0.45 \\ P(F) &= 0.6 \\ P(G) &= 0.55 \\ P(E \text{ and } F) &= 0.2 \\ P(E \text{ and } G) &= 0.2 \\ P(F \text{ and } G) &= 0.25 \end{aligned}$$

Compute the following:

1. $P(E \text{ or } F)$
2. $P(E \text{ or } G)$
3. $P(F \text{ or } G)$

Solution

1. Using the Inclusion/Exclusion Principle, we get:

$$\begin{aligned} P(E \text{ or } F) &= P(E) + P(F) - P(E \text{ and } F) \\ &= 0.45 + 0.6 - 0.2 \\ &= 0.85. \end{aligned}$$

2. Again, we'll apply the Inclusion/Exclusion Principle:

$$\begin{aligned} P(E \text{ or } G) &= P(E) + P(G) - P(E \text{ and } G) \\ &= 0.45 + 0.55 - 0.2 \\ &= 0.8. \end{aligned}$$

3. Applying the Inclusion/Exclusion Principle one more time:

$$\begin{aligned} P(F \text{ or } G) &= P(F) + P(G) - P(F \text{ and } G) \\ &= 0.6 + 0.55 - 0.25 \\ &= 0.9. \end{aligned}$$

Key Terms

- mutually exclusive

Key Concepts

- The Addition Rule is used to find the probability that one event or another will occur when those events are mutually exclusive.
- The Inclusion/Exclusion Principle is used to find probabilities when events are not mutually exclusive.

Formulas

- If E and F are mutually exclusive events, then

$$P(E \text{ or } F) = P(E) + P(F)$$

- If E and F are events that contain outcomes of a single experiment, then

$$P(E \text{ or } F) = P(E) + P(F) - P(E \text{ and } F)$$

7.9 Conditional Probability and the Multiplication Rule

Figure 37: If you roll two dice by throwing them one at a time, the face showing on the first die will affect the possible outcomes for the sum of the two dice.

Learning Objectives

After completing this section, you should be able to:

1. Calculate conditional probabilities.
2. Apply the Multiplication Rule for Probability to compute probabilities.

Back in Example 3 in Section 7.5, we constructed the following table to help us find the probabilities associated with rolling two standard 6-sided dice:

	first die					
second die	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

For example, 3 of these 36 equally likely outcomes correspond to rolling a sum of 10, so the probability of rolling a 10 is $\frac{3}{36} = \frac{1}{12}$. However, if you choose to roll the dice one at a time, the probability of rolling a 10 will change after the first die comes to rest. For example, if the first die shows a 5, then the probability of rolling a sum of 10 has jumped to $\frac{1}{6}$—the event will occur if the second die also shows a 5, which is 1 of 6 equally likely outcomes for the second die. If instead the first die shows a 3, then the probability of rolling a sum of 10 drops to 0—there are no outcomes for the second die that will give us a sum of 10.

Understanding how probabilities can shift as we learn new information is critical in the analysis of our second type of compound events: those built with "and." This section will explain how to compute probabilities of those compound events.

Conditional Probabilities

When we analyze experiments with multiple stages, we often update the probabilities of the possible final outcomes or the later stages of the experiment based on the results of one or more of the initial stages. These updated probabilities are called **conditional probabilities**.

In other words, if O is a possible outcome of the first stage in a multistage experiment, then the probability of an event E conditional on O (denoted $P(E \mid O)$, read "the probability of E given O") is the updated probability of E under the assumption that O occurred.

In the example that opened this section, we might consider rolling two dice as a multistage experiment: rolling one, then the other. If we define E to be the event "roll a sum of 10," O to be the event "first die shows 5," and Q to be the event "first die shows 3," then we computed $P(E) = \frac{1}{12}$, $P(E \mid O) = \frac{1}{6}$, and $P(E \mid Q) = 0$.

Example 1 **Computing Conditional Probabilities**

1. April is playing a coin-flipping game with Ben. She will flip a coin 3 times. If the coin lands on heads more than tails, April wins; if it lands on tails more than heads, Ben wins. Let A be the event "April wins," H be "first flip is heads," and T be "first flip is tails." Compute $P(A)$, $P(A \mid H)$, and $P(A \mid T)$.
2. You are about to draw 2 cards without replacement from a deck containing only these 10 cards: A ♡, A ♠, A ♣, A ◇, K ♠, K ♣, Q ♡, Q ♠, J ♡, J ♠. We'll define the following events: F is "both cards are the same rank," A is "first card is an ace," and K is "first card is a king." Compute $P(F \mid A)$ and $P(F \mid K)$.
3. Jim's sock drawer contains 5 black socks and 3 blue socks. To avoid waking his partner, Jim doesn't want to turn the lights on, so he puts on 2 socks at random. Let M be the event "Jim's 2 socks match," let K be the event "the sock on Jim's left foot is black," and let L be the event "the sock on Jim's left foot is blue." Compute $P(M)$, $P(M \mid K)$, and $P(M \mid L)$.

Solution

1. **Step 1.** The sample space is {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}. The event A consists of the first 4 of those outcomes: HHH, HHT, HTH, and THH. Thus, $P(A) = \frac{4}{8} = \frac{1}{2}$.

Step 2. Now, let's compute $P(A \mid H)$. We are assuming the result of the first flip is heads. That leaves us with 4 possible outcomes: HHH, HHT, HTH, and HTT. Of those, April wins 3 (HHH, HHT, HTH) and loses one (HTT). So, $P(A \mid H) = \frac{3}{4}$.

Step 3. If the result of the first flip is instead tails, the 4 possible outcomes are THH, THT, TTH, and TTT. Of those, April wins 1 (THH) and loses 3 (THT, TTH, TTT). So, $P(A \mid T) = \frac{1}{4}$.

1. **Step 1.** If the event A happens, then 1 of the 4 aces is drawn first; the remaining cards in the deck are 3 aces, 2 kings, 2 queens, and 2 jacks. In order for the event F to occur, the second card drawn has to be an ace. Since there are 3 aces among the remaining 9 cards, $P(F \mid A) = \frac{3}{9} = \frac{1}{3}$.

Step 2. If the event K happens instead, then the first card drawn is a king. That leaves 4 aces, 1 king, 2 queens, and 2 jacks in the deck. Under the assumption that the first card is a king, the event F will occur only if the second card is also a king. Since only one of the remaining 9 cards is a king, we have $P(F \mid K) = \frac{1}{9}$.

1. **Step 1.** We can view the event M as a compound event using "or": both socks are blue or both socks are black. Let's compute the probability that both socks are blue using combinations. We're choosing 2 socks from a group of 8; 3 of the 8 are blue. So, $P(\text{both socks blue}) = \frac{3C_2}{8C_2} = \frac{3}{28}$. Similarly, $P(\text{both socks black}) = \frac{5C_2}{8C_2} = \frac{10}{28}$. Therefore, since these events are mutually exclusive, we can use the Addition Rule: $P(M) = P(\text{both socks blue}) + P(\text{both socks black}) = \frac{3}{28} + \frac{10}{28} = \frac{13}{28}$.

Step 2. If the sock on Jim's left foot is black (i.e., K occurred), then there are 4 remaining black socks of the 7 in the drawer. So, $P(M \mid K) = \frac{4}{7}$.

Step 3. If the sock on Jim's left foot is blue (L occurred), then there are 2 blue socks among the 7 remaining in the drawer. So, $P(M \mid L) = \frac{2}{7}$.

CHECKPOINT

In Tree Diagrams, Tables, and Outcomes, we introduced the concept of dependence between stages of a multistage experiment. We stated at the time that two stages were dependent if the result of one stage affects the other stage. We explained that dependence in terms of the sample space, but sometimes that dependence can be a little more subtle; it's more properly understood in terms of conditional probabilities. Two stages of an experiment are dependent if $P(E \mid F) \neq P(E \mid F')$ for some outcome of the second stage E and outcome of the first stage F.

WHO KNEW?

Protecting Bombers in World War II

In his book *How Not to Be Wrong*, Jordan Ellenberg recounts this anecdote: During World War II, the American military wanted to add additional armor plating to bomber aircraft, in order to reduce the chances that they get shot down. So, they collected data on planes after returning from missions. The data showed that the fuselage, wings, and fuel system had many more bullet holes (per unit area) than the engine compartments, so the military brass wanted to add additional armor to the parts of the plane that were hit most often. Luckily, before they added the armor to the planes, they asked for a second opinion. Abraham Wald, a Jewish mathematician who had fled the rising Nazi regime, pointed out that it was far more important that the armor plating be added to areas where there were *fewer* bullet holes. Why? The planes they were studying had already completed their missions, so the military was essentially looking at conditional probabilities: the probability of suffering a bullet strike, *given that* the plane made it back safely. More bullet holes in an area on the plane indicated that was a region that wasn't as important for the plane's survival!

Compound Events Using "And" and the Multiplication Rule

For multistage experiments, the outcomes of the experiment as a whole are often stated in terms of the outcomes of the individual stages. Commonly, those statements are joined with "and." For example, in the sock drawer example just above, one outcome might be "the left sock is black and the right sock is blue." As with "or" compound events, these probabilities can be computed with basic arithmetic.

FORMULA

Multiplication Rule for Probability: If E and F are events associated with the first and second stages of an experiment, then $P(E \text{ and } F) = P(E) \times P(F \mid E)$.

CHECKPOINT

In The Addition Rule for Probability, we considered probabilities of events connected with *"and" in the statement of the Inclusion/Exclusion Principle. These two scenarios are different; in the statement of the Inclusion/Exclusion Principle, the events connected with "and" are both events associated with the **same** single-stage experiment (or the same stage of a multistage experiment). In the Multiplication Rule, we're looking at events associated with **different** stages of a multistage experiment.*

Example 2 **Using the Multiplication Rule for Probability**

You are president of a club with 10 members: 4 seniors, 3 juniors, 2 sophomores, and 1 first-year. You need to choose 2 members to represent the club on 2 college committees. The first person selected will be on the Club Awards Committee and the second will be on the New Club Orientation Committee. The same person cannot be selected for both. You decide to select these representatives at random.

1. What is the probability that a senior is chosen for both positions?
2. What is the probability that a junior is chosen first and a sophomore is chosen second?
3. What is the probability that a sophomore is chosen first and a senior is chosen second?

Solution

1. We need the probability that a senior is chosen first and a senior is chosen second. These are two stages of a multistage experiment, so we'll apply the Multiplication Rule for Probability: $P(\text{ senior chosen first and senior chosen second}) = P(\text{senior chosen first}) \times P(\text{senior chosen second} \mid \text{senior chosen first})$. Since there are 4 seniors among the 10 members, $P(\text{senior chosen first}) = \frac{4}{10} = \frac{2}{5}$. Next, assuming a senior is chosen first, there are 3 seniors among the 9 remaining members. So, $P(\text{senior chosen second} \mid \text{senior chosen first}) = \frac{3}{9} = \frac{1}{3}$. Putting this all together, we get $P(\text{senior chosen first and senior chosen second}) = \frac{2}{5} \times \frac{1}{3} = \frac{2}{15}$.
2. There are 3 juniors among the 10 members, so $P(\text{junior chosen first}) = \frac{3}{10}$. Assuming a junior is chosen first, there are 2 sophomores among the remaining 9 members, so $P(\text{sophomore chosen second} \mid \text{junior chosen first}) = \frac{2}{9}$. Thus, using the Multiplication Rule for Probability, we have $P(\text{junior chosen first and sophomore chosen second}) = \frac{3}{10} \times \frac{2}{9} = \frac{1}{15}$.
3. The probability that a sophomore is chosen first is $\frac{2}{10} = \frac{1}{5}$, and the probability that a senior is chosen second given that a sophomore was chosen first is $\frac{4}{9}$. Thus, using the Multiplication Rule for Probability, we have: $P(\text{sophomore chosen first and senior chosen second}) = \frac{1}{5} \times \frac{4}{9} = \frac{4}{45}$.

WORK IT OUT

The Birthday Problem

One of the most famous problems in probability theory is the Birthday Problem, which has to do with shared birthdays in a large group. To make the analysis easier, we'll ignore leap days, and assume that the probability of being born on any given date is $\frac{1}{365}$. Now, if you have 366 people in a room, we're guaranteed to have at least one pair of people who share a single birthday. Imagine filling the room by first admitting someone born on January 1, then someone born on January 2, and so on... The 365th person admitted would be born on December 31. If you add one more person to the room, that person's birthday would *have* to match someone else's.

Let's look at the other end of the spectrum. If you choose two people at random, what is the probability that they share a birthday? As with many probability questions, this is best addressed by find out the probability that they do *not* share a birthday. The first person's birthday can be anything (probability 1), and the second person's birthday can be anything other than the first person's birthday (probability $\frac{364}{365}$). The probability that they have *different* birthdays is $1 \times \frac{364}{365} = \frac{364}{365}$. So, the probability that they share a birthday is $1 - \frac{364}{365} = \frac{1}{365}$. What if we have three people? The probability that they all have different birthdays can be obtained by extending our previous calculation: The probability that two people have different birthdays is $\frac{364}{365}$, so if we add a third to the mix, the probability that they have a different birthday from the other two is $\frac{363}{365}$. So, the probability that all three have different birthdays is $\frac{364}{365} \times \frac{363}{365} \approx 0.9918$, and thus the probability that there's a shared birthday in the group is $1 - 0.9918 \approx 0.0082$.

The big question is this: How many people do we need in the room to have the probability of a shared birthday greater than $\frac{1}{2}$? Make a guess, then with a partner keep adding hypothetical people to the group and computing probabilities until you get there!

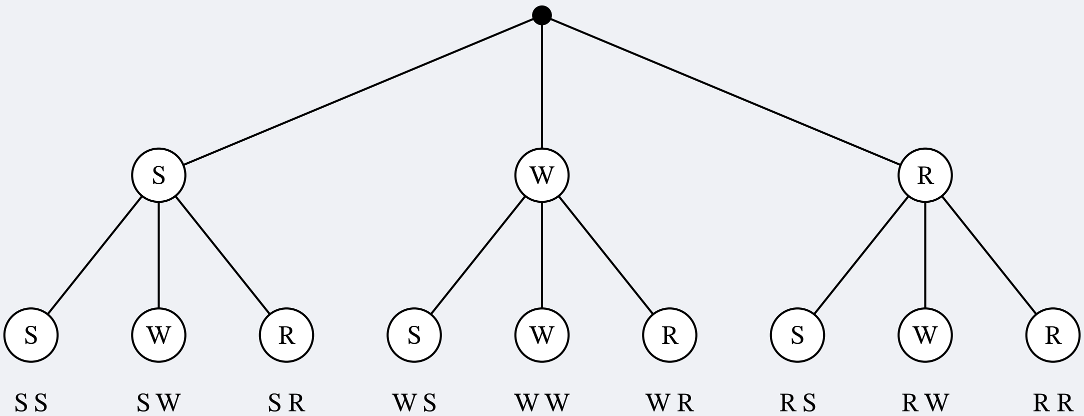
It is often useful to combine the rules we've seen so far with the techniques we used for finding sample spaces. In particular, trees can be helpful when we want to identify the probabilities of every possible outcome in a multistage experiment. The next example will illustrate this.

Example 3 **Using Tree Diagrams to Help Find Probabilities**

The board game *Clue* uses a deck of 21 cards: 6 suspects, 6 weapons, and 9 rooms. Suppose you are about to draw 2 cards from this deck. There are 6 possible outcomes for the draw: 2 suspects, 2 weapons, 2 rooms, 1 suspect and 1 weapon, 1 suspect and 1 room, or 1 weapon and 1 room. What are the probabilities for each of these outcomes?

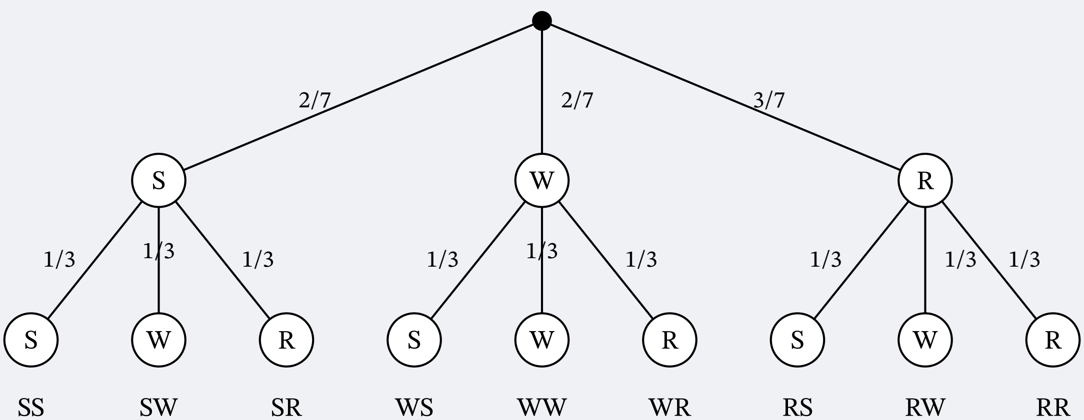
Solution

Step 1: Let's start by building a tree diagram that illustrates both stages of this experiment. Let's use S, W, and R to indicate drawing a suspect, weapon, and room, respectively .



Step 2: We want to start computing probabilities, starting with the first stage. The probability that the first card is a suspect is $\frac{6}{21} = \frac{2}{7}$. The probability that the first card is a weapon is the same: $\frac{2}{7}$. Finally, the probability that the first card is a room is $\frac{9}{21} = \frac{3}{7}$.

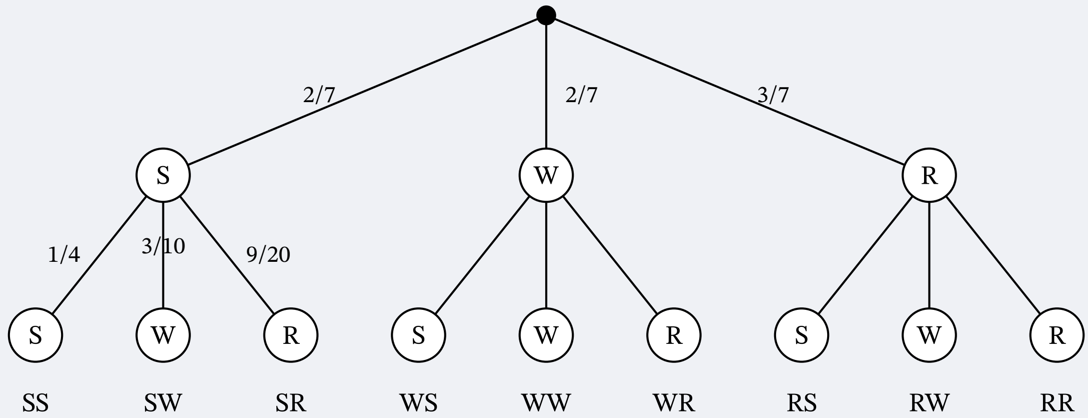
Step 3: Let's incorporate those probabilities into our tree: label the edges going into each of the nodes representing the first-stage outcomes with the corresponding probabilities .



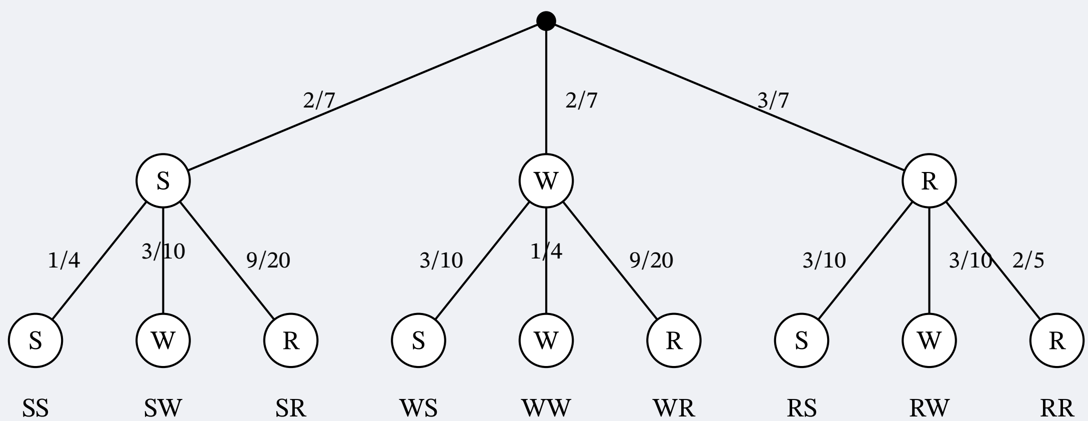
Note that the sum of the probabilities coming out of the initial node is 1; this should always be the case for the probabilities coming out of any node!

Step 4: Let's look at the case where the first card is a suspect. There are 3 edges emanating from that node (leading to the outcomes SS, SW, and SR). We'll label those edges with the appropriate *conditional* probabilities, under the assumption that the first card is a suspect. First, there are 5 remaining suspect cards among the 20 left in the deck, so $P(\text{second is suspect} \mid \text{first is suspect}) = \frac{5}{20} = \frac{1}{4}$. Using similar reasoning, we can compute $P(\text{second is weapon} \mid \text{first is suspect}) = \frac{6}{20} = \frac{3}{10}$ and $P(\text{second is room} \mid \text{first is suspect}) = \frac{9}{20}$.

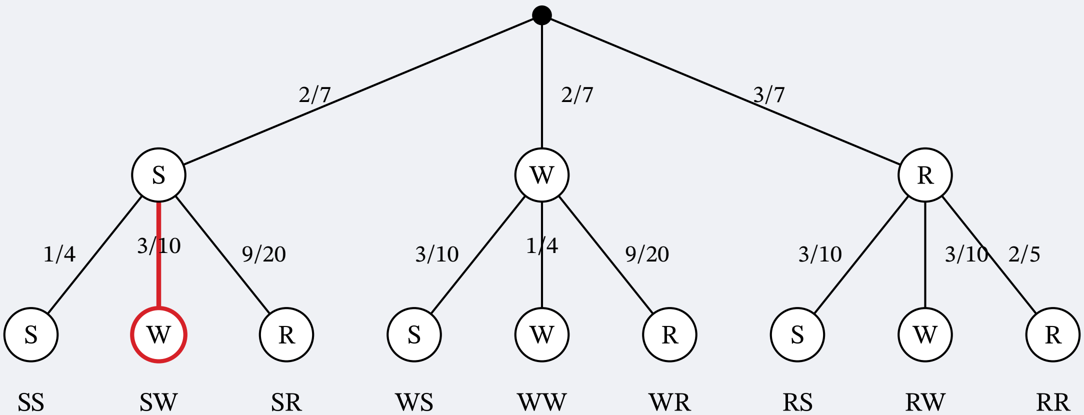
Step 5: Checking our work, we see that the sum of these 3 probabilities is again equal to 1. Let's add those to our tree .



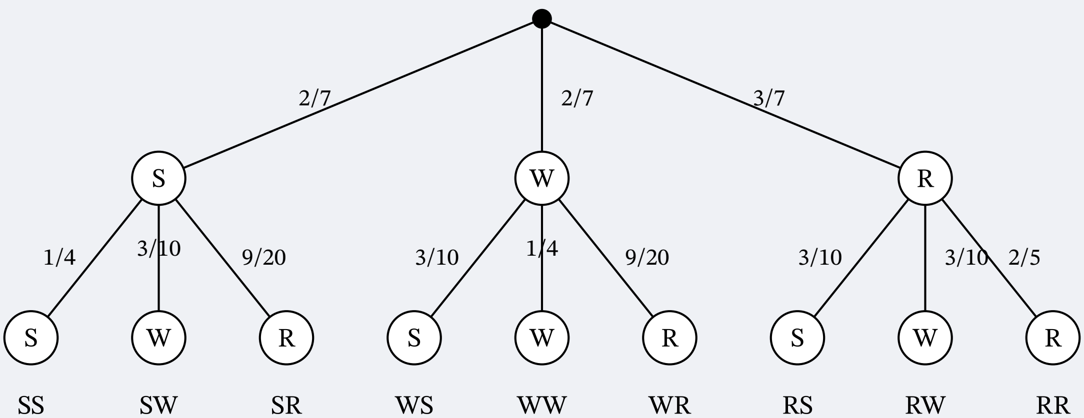
Step 6: Let's continue filling in the conditional probabilities at the other nodes, always checking to make sure the sum of the probabilities coming out of any node is equal to 1 .



Step 7: We can compute the probability of landing on any final node by multiplying the probabilities along the path we would take to get there. For example, the probability of drawing a suspect first and a weapon second (i.e., ending up on the node labeled "SW") is $\frac{2}{7} \times \frac{3}{10} = \frac{3}{35}$, as illustrated in .



Step 8: Let's fill in the rest of the probabilities .



Step 9: A helpful feature of tree diagrams is that the final outcomes are always mutually exclusive, so the Addition Rule can be directly applied. For example, the probability of drawing one suspect and one room (in any order) would be $P(SR) + P(RS) = \frac{9}{70} + \frac{9}{70} = \frac{9}{35}$. We can find the probabilities of the other outcomes in a similar fashion, as shown in the following table:

Outcome	Probability
2 suspects	$\frac{1}{14}$
2 weapons	$\frac{1}{14}$
2 rooms	$\frac{6}{35}$
1 suspect and 1 weapon	$\frac{3}{35} + \frac{3}{35} = \frac{6}{35}$
1 suspect and 1 room	$\frac{9}{70} + \frac{9}{70} = \frac{9}{35}$
1 weapon and 1 room	$\frac{9}{70} + \frac{9}{70} = \frac{9}{35}$

Checking once again, the sum of these 6 probabilities is 1, as expected.

WORK IT OUT

The Monty Hall Problem

On the original version of the game show *Let's Make a Deal*, originally hosted by Monty Hall and now hosted by Wayne Brady, one contestant was chosen to play a game for the grand prize of the day (often a car). Here's how it worked: On the stage were three areas concealed by numbered curtains. The car was hidden behind one of the curtains; the other two curtains hid worthless prizes (called "Zonks" on the show). The contestant would guess which curtain concealed the car. To build tension, Monty would then reveal what was behind one of the *other* curtains, which was always one of the Zonks (Since Monty knew where the car was hidden, he always had at least one Zonk curtain that hadn't been chosen that he could reveal). Monty then turned to the contestant and asked: "Do you want to stick with your original choice, or do you want to switch your choice to the other curtain?" What should the contestant do? Does it matter?

With a partner or in a small group, simulate this game. You can do that with a small candy (the prize) hidden under one of three cups, or with three playing cards (just decide ahead of time which card represents the "Grand Prize"). One person plays the host, who knows where the prize is hidden. Another person plays the contestant and tries to guess where the prize is hidden. After the guess is made, the host should reveal a losing option that wasn't chosen by the contestant. The contestant then has the option to stick with the original choice or switch to the other, unrevealed option. Play about 20 rounds, taking turns in each role and making sure that both contestant strategies (stick or switch) are used equally often. After each round, make a note of whether the contestant chose "stick" or "switch" and whether the contestant won or lost. Find the empirical probability of winning under each strategy. Then, see if you can use tree diagrams to verify your findings.

Key Terms

- conditional probability

Key Concepts

- Conditional probabilities are computed under the assumption that the condition has already occurred.
- The Multiplication Rule for Probability is used to find the probability that two events occur in sequence.

Formulas

- If E and F are events associated with the first and second stages of an experiment, then $P(E \text{ and } F) = P(E) \times P(F \mid E)$.

7.10 The Binomial Distribution

Figure 45: If one baseball team has a 65% chance of beating another in any single game, what's the likelihood that they win a best-of-seven series?

Learning Objectives

After completing this section, you should be able to:

1. Identify binomial experiments.
2. Use the binomial distribution to analyze binomial experiments.

It's time for the World Series, which determines the champion for this season in Major League Baseball. The scrappy Los Angeles Angels are facing the powerhouse Cincinnati Reds. Computer models put the chances of the Reds winning any single game against the Angels at about 65%. The World Series, as its name implies, isn't just one game, though: it's what's known as a "best-of-seven" contest: the teams play each other repeatedly until one team wins 4 games (which could take up to 7 games total, if each team wins three of the first 6 matchups). If the Reds truly have a 65% chance of winning a single game, then the probability that they win the series should be greater than 65%. Exactly how much bigger?

If you have the patience for it, you could use a tree diagram like we used in Example 3 in Section 7.9 to trace out all of the possible outcomes, find all the related probabilities, and add up the ones that result in the Reds winning the series. Such a tree diagram would have $2^7 = 128$ final nodes, though, so the calculations would be very tedious. Fortunately, we have tools at our disposal that allow us to find these probabilities very quickly. This section will introduce those tools and explain their use.

Binomial Experiments

The tools of this section apply to multistage experiments that satisfy some pretty specific criteria. Before we move on to the analysis, we need to introduce and explain those criteria so that we can recognize experiments that fall into this category. Experiments that satisfy each of these criteria are called **binomial experiments**. A binomial experiment is an experiment with a fixed number of repeated independent binomial trials, where each trial has the same probability of success.

Repeated Binomial Trials

The first criterion involves the structure of the stages. Each stage of the experiment should be a replication of every other stage; we call these replications **trials**. An example of this is flipping a coin 10 times; each of the ten flips is a trial, and they all occur under the same conditions as every other. Further, each trial must have only two possible outcomes. These two outcomes are typically labeled “success” and “failure,” even if there is not a positive or negative connotation associated with those outcomes. Experiments with more than two outcomes in their sample spaces are sometimes reconsidered in a way that forces just two outcomes; all we need to do is completely divide the sample space into two parts that we can label “success” and “failure.” For example, your grade on an exam might be recorded as A, B, C, D, or F, but we could instead think of the grades A, B, C, and D as “success” and a grade of F as “failure.” Trials with only two outcomes are called **binomial trials** (the word *binomial* derives from Latin and Greek roots that mean “two parts”).

Independent Trials

The next criterion that we’ll be looking for is independence of trials. Back in Tree Diagrams, Tables, and Outcomes, we said that two stages of an experiment are independent if the outcome of one stage doesn’t affect the other stage. Independence is necessary for the experiments we want to analyze in this section.

Fixed Number of Trials

Next, we require that the number of trials in the experiment be decided before the experiment begins. For example, we might say “flip a coin 10 times.” The number of trials there is fixed at 10. However, if we say “flip a coin until you get 5 heads,” then the number of trials could be as low as 5, but theoretically it could be 50 or a 100 (or more)! We can’t apply the tools from this section in cases where the number of trials is indeterminate.

Constant Probability

The next criterion needed for binomial experiments is related to the independence of the trials. We must make sure that the probability of success in each trial is the same as the probability of success in every other trial.

Example 1 **Identifying Binomial Experiments**

Decide whether each of the following is a binomial experiment. For those that aren’t, identify which criterion or criteria are violated.

1. You roll a standard 6-sided die 10 times and write down the number that appears each time.
2. You roll a standard 6-sided die 10 times and write down whether the die shows a 6 or not.
3. You roll a standard 6-sided die until you get a 6.
4. You roll a standard 6-sided die 10 times. On the first roll, we define “success” as rolling a 4 or greater. After the first roll, we define “success” as rolling a number greater than the result of the previous roll.

Solution

1. Since we're noting 1 of 6 possible outcomes, the trials are not binomial. So, this isn't a binomial experiment.
2. We have 2 possible outcomes ("6" and "not 6"), the trials are independent, the probability of success is the same every time, and the number of trials is fixed. This is a binomial experiment.
3. Since the number of trials isn't fixed (we don't know if we'll get our first 6 after 1 roll or 20 rolls or somewhere in between), this isn't a binomial experiment.
4. Here, the probability of success might change with every roll (on the first roll, that probability is $\frac{1}{2}$ if the first roll is a 6, the probability of success on the next roll is zero). So, this is not a binomial experiment.

The Binomial Formula

If we flip a coin 100 times, you might expect the number of heads to be around 50, but you probably wouldn't be surprised if the actual number of heads was 47 or 52. What is the probability that the number of heads is exactly 50? Or falls between 45 and 55? It seems unlikely that we would get more than 70 heads. Exactly how unlikely is that?

Each of these questions is a question about the number of successes in a binomial experiment (flip a coin 100 times, "success" is flipping heads). We could theoretically use the techniques we've seen in earlier sections to answer each of these, but the number of calculations we'd have to do is astronomical; just building the tree diagram that represents this situation is more than we could complete in a lifetime; it would have $2^{100} \approx 1.3 \times 10^{30}$ final nodes! To put that number in perspective, if we could draw 1,000 dots every second, and we started at the moment of the Big Bang, we'd currently be about 0.00000003% of the way to drawing out those final nodes. Luckily, there's a shortcut called the Binomial Formula that allows us to get around doing all those calculations!

FORMULA

Binomial Formula: Suppose we have a binomial experiment with n trials and the probability of success in each trial is p. Then:

$$P(\text{number of successes is } a) = {}_nCa \times p^a \times (1-p)^{n-a}$$

We can use this formula to answer one of our questions about 100 coin flips. What is the probability of flipping exactly 50 heads? In this case, $n = 100$, $p = \frac{1}{2}$, and $a = 50$, so $P(\text{flip 50 heads}) = {}_{100}C50 \times \left(\frac{1}{2}\right)^{50} \times \left(1-\frac{1}{2}\right)^{100-50}$. Unfortunately, many calculators will balk at this calculation; that first factor ($100C_{50}$) is an enormous number, and the other two factors are very close to zero. Even if your calculator can handle numbers that large or small, the arithmetic can create serious errors in rounding off.

TECH CHECK

Luckily, spreadsheet programs have alternate methods for doing this calculation. In Google Sheets, we can use the BINOMDIST function to do this calculation for us. Open up a new sheet, click in any empty cell, and type "=BINOMDIST(50,100,0.5,FALSE)" followed by the Enter key. The cell will display the probability we seek; it's about 8%. Let's break down the syntax of that function in Google Sheets: enter "=BINOMDIST(a, n, p, FALSE)" to find the probability of a successes in n trials with probability of success p.

Example 2 **Using the Binomial Formula**

1. Find the probability of rolling a standard 6-sided die 4 times and getting exactly one 6 *without using technology*.
2. Find the probability of rolling a standard 6-sided die 60 times and getting exactly ten 6s using technology.
3. Find the probability of rolling a standard 6-sided die 60 times and getting exactly eight 6s using technology.

Solution

1. We'll apply the Binomial Formula, where $n = 4$, $a = 1$, and $p = \frac{1}{6}$:

$$\begin{aligned} P(\text{rolling one 6}) &= {}_4C_1 \times \left(\frac{1}{6}\right)^1 \times \left(\frac{5}{6}\right)^{4-1} \\ &= \frac{4!}{1!(4-1)!} \times \frac{1}{6} \times \left(\frac{5}{6}\right)^3 \\ &= 4 \times \frac{1}{6} \times \frac{5^3}{6^3} \\ &= \frac{4 \times 5^3}{6^4} \\ &= \frac{500}{1{,}296}. \end{aligned}$$

2. Here, $n = 60$, $a = 10$, and $p = \frac{1}{6}$. In Google Sheets, we'll enter "=BINOMDIST(10, 60, 1/6, FALSE)" to get our result: 0.137.
3. This experiment is the same as in Exercise 2 of this example; we're simply changing the number of successes from 10 to 8. Making that change in the formula in Google Sheets, we get the probability 0.116.

The Binomial Distribution

If we are interested in the probability of more than just a single outcome in a binomial experiment, it's helpful to think of the Binomial Formula as a function, whose input is the number of successes and whose output is the probability of observing that many successes. Generally, for a small number of trials, we'll give that function in table form, with a complete list of the possible outcomes in one column and the probability in the other.

For example, suppose Kristen is practicing her basketball free throws. Assume Kristen always makes 82% of those shots. If she attempts 5 free throws, then the Binomial Formula gives us these probabilities:

Shots Made	Probability
0	0.000189
1	0.004304
2	0.0392144
3	0.1786432
4	0.4069096
5	0.3707398

A table that lists all possible outcomes of an experiment along with the probabilities of those outcomes is an example of a **probability density function** (PDF). A PDF may also be a formula that you can use to find the probability of any outcome of an experiment.

CHECKPOINT

Because they refer to the same thing, some sources will refer to the Binomial Formula as the Binomial PDF.

If we want to know the probability of a range of outcomes, we could add up the corresponding probabilities. Going back to Kristen's free throws, we can find the probability that she makes 3 or fewer of her 5 attempts by adding up the probabilities associated with the corresponding outcomes (in this case: 0, 1, 2, or 3):

$$\begin{aligned} P(\text{makes 3 or fewer}) = & \quad P(a=0)+P(a=1)+P(a=2)+P(a=3) \\ & = 0.000189+0.004304+0.0392144+0.1786432 \\ & = \quad 0.2223506 \end{aligned}$$

The probability that the outcome of an experiment is less than or equal to a given number is called a *cumulative probability*. A table of the cumulative probabilities of all possible outcomes of an experiment is an example of a **cumulative distribution function** (CDF). A CDF may also be a formula that you can use to find those cumulative probabilities.

CHECKPOINT

Cumulative probabilities are always associated with events that are defined using $\leq$. If other inequalities are used to define the event, we must restate the definition so that it uses the correct inequality.

Here are the PDF and CDF for Kristen's free throws:

Shots Made	Probability	Cumulative
0	0.000189	0.000189
1	0.004304	0.004493
2	0.0392144	0.0437073
3	0.1786432	0.2223506
4	0.4069096	0.6292602
5	0.3707398	1

TECH CHECK

Google Sheets can also compute cumulative probabilities for us; all we need to do is change the "FALSE" in the formulas we used before to "TRUE."

Example 3 **Using the Binomial CDF**

Suppose we are about to flip a fair coin 50 times. Let H represent the number of heads that result from those flips. Use technology to find the following:

1. $P(H \leq 22)$
2. $P(H < 26)$
3. $P(H > 28)$
4. $P(H \geq 20)$
5. $P(20 < H < 25)$

Solution

1. The event here is defined by $H \leq 22$, which is the inequality we need to have if we want to use the Binomial CDF. In Google Sheets, we'll enter "=BINOMDIST(22, 50, 0.5, TRUE)" to get our answer: 0.2399.
2. This event uses the wrong inequality, so we need to do some preliminary work. If $H < 26$, that means $H \leq 25$ (because H has to be a whole number). So, we'll enter "=BINOMDIST(25, 50, 0.5, TRUE)" to find $P(H < 26) = P(H \leq 25) = 0.5561$.
3. The inequality associated with this event is pointing in the wrong direction. If E is the event $H > 28$, that means that E contains the outcomes $\{29, 30, 31, 32, 33, \ldots\}$. Thus, E' must contain the outcomes $\{\ldots, 25, 26, 27, 28\}$. In other words, E' is defined by $H \leq 28$. Since it uses $\leq$, we can find $P(E')$ using "=BINOMDIST(28, 50, 0.5, TRUE)": 0.8389 So, using the formula for probabilities of complements, we have

$$\begin{aligned} P(E) &= 1 - P(E') \\ &= 1 - 0.8389 \\ &= 0.1611. \end{aligned}$$

4. As in part 3, this inequality is pointing in the wrong direction. If F is the event $H \geq 20$, then F contains the outcomes $\{20, 21, 22, 23, \ldots\}$. That means F' contains the outcomes $\{\ldots, 16, 17, 18, 19\}$, and so F' is defined by $H \leq 19$. So, we can find $P(F')$ using "=BINOMDIST(19, 50, 0.5, TRUE)": 0.0595. Finally, using the formula for probabilities of complements, we get:
$$\begin{aligned} P(F) &= 1 - P(F') \\ &= 1 - 0.0595 \\ &= 0.9405. \end{aligned}$$
5. If $20 < H < 25$, that means we are interested in the outcomes $\{21, 22, 23, 24\}$. This doesn't look like any of the previous situations, but there is a way to find this probability using the Binomial CDF. We need to put everything in terms of "less than or equal to," so we'll first note that all of our outcomes are less than or equal to 24. But we don't want to include values that are less than or equal to 20. So, we have three events: let I be the event defined by $20 < H < 25$ (note that we're trying to find $P(I)$). Let J be defined by $H \leq 24$, and let K be defined by $H \leq 20$. Of these three events, J contains the most outcomes. If J occurs, then either K or I must have occurred. Moreover, K and I are mutually exclusive. Thus, $P(J) = P(K) + P(I)$, by the Addition Rule. Solving for the probability that we want, we get
$$\begin{aligned} P(I) &= P(J) - P(K) \\ &= P(H \leq 24) - P(H \leq 20) \\ &= 0.44386 - 0.10132 \\ &= 0.34254. \end{aligned}$$

Finally, we can answer the question posed at the beginning of this section. Remember that the Reds are facing the Angels in the World Series, which is won by the team who is first to win 4 games. The Reds have a 65% chance to win any game against the Angels. So, what is the probability that the Reds win the World Series? At first glance, this is not a binomial experiment: The number of games played is not fixed, since the series ends as soon as one team wins 4 games. However, we can extend this situation to a binomial experiment: Let's assume that 7 games are always played in the World Series, and the winner is the team who wins more games. In a way, this is what happens in reality; it's as though the first team to lose 4 games (and thus cannot win more than the other team) forfeits the rest of their games. So, we can treat the actual World Series as a binomial experiment with seven trials. If W is the number of games won by the Reds, the probability that the Reds win the World Series is $P(W \geq 4)$. Using the techniques from the last example, we get $P(\text{Reds win the series}) = 0.8002$.

PEOPLE IN MATHEMATICS

Abraham de Moivre

Abraham de Moivre was born in 1667 in France to a Protestant family. Though he was educated in Catholic schools, he remained true to his faith; in 1687, he fled with his brother to London to escape persecution under the reign of King Louis XIV. Once he arrived in England, he supported himself as a freelance math tutor while he conducted his own research. Among his interests was probability; in 1711, he published the first edition of *The Doctrine of Chances: A Method of Calculating the Probabilities of Events in Play*. This book was the second textbook on probability (after Cardano's *Liber de ludo aleae*). De Moivre discovered an important connection between the binomial distribution and the*normal distribution* (an important concept in statistics; we'll explore that distribution and its connection to the binomial distribution in Chapter 8). De Moivre also discovered some properties of a new probability distribution that later became known as the Poisson distribution.

Key Terms

- binomial experiment
- probability density function (PDF)
- cumulative distribution function (CDF)

Key Concepts

- Binomial experiments result when we count the number of successful outcomes in a fixed number of repeated, independent trials with a constant probability of success.
- The binomial distribution is used to find probabilities associated with binomial experiments.
- Probability density functions (PDFs) describe the probabilities of individual outcomes in an experiment; cumulative distribution functions (CDFs) give the probabilities of ranges of outcomes.

Formulas

- Suppose we have a binomial experiment with n trials and the probability of success in each trial is p. Then:
  $P(\text{number of successes is } a) = {}_nCa \times p^a \times (1-p)^{n-a}$

7.11 Expected Value

Figure 46: The concept of expected value allows us to analyze games that involve randomness, like Roulette.

Learning Objectives

After completing this section, you should be able to:

1. Calculate the expected value of an experiment.
2. Interpret the expected value of an experiment.
3. Use expected value to analyze applications.

The casino game roulette has dozens of different bets that can be made. These bets have different probabilities of winning but also have different payouts. In general, the lower the probability of winning a bet is, the more money a player wins for that bet. With so many options, is there one bet that's "smarter" than the rest? What's the best play to make at a roulette table? In this section, we'll develop the tools we need to answer these questions.

Expected Value

Many experiments have numbers associated with their outcomes. Some are easy to define; if you roll 2 dice, the sum of the numbers showing is a good example. In some card games, cards have different point values associated with them; for example, in some forms of the game rummy, aces are worth 15 points; 10s, jacks, queens, and kings are worth 10; and all other cards are worth 5. The outcomes of casino and lottery games are all associated with an amount of money won or lost. These outcome values are used to find the **expected value** of an experiment: the mean of the values associated with the outcomes that we would observe over a large number of repetitions of the experiment. (See Conditional Probability and the Multiplication Rule for more on means.)

That definition is a little vague; How many is "a large number?" In practice, it depends on the experiment; the number has to be large enough that every outcome would be expected to appear at least a few times. For example, if we're talking about rolling a standard 6-sided die and we note the number showing, a few dozen replications should be enough that the mean would be representative. Since the probability of each outcome is $\frac{1}{6}$, we would expect to see each outcome about 8 times over the course of 48 replications. However, if we're talking about the Powerball lottery, where the probability of winning the jackpot is about $\frac{1}{292,000,000}$, we would need several *billion* replications to ensure that every outcome appears a few times. Luckily, we can find the theoretical expected value before we even run the experiment the first time.

FORMULA

Expected Value: If O represents an outcome of an experiment and $n(O)$ represents the value of that outcome, then the expected value of the experiment is:

$$\sum n(O) \times P(O)$$

where Σ is the "sum," meaning we add up the results of the formula that follows over all possible outcomes.

Example 1 **Finding Expected Values**

Find the expected values of the following experiments.

1. Roll a standard 6-sided die and note the number showing.
2. Roll two standard 6-sided dice and note the sum of the numbers showing.
3. Draw a card from a well-shuffled standard deck of cards and note its rummy value (15 for aces; 10 for tens, jacks, queens, and kings; 5 for everything else).

Solution

1. **Step 1:** Let's start by writing out the PDF table for this experiment.

Value	Probability
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

Step 2: To find the expected value, we need to find $n(O) \times P(O)$ for each possible outcome in the table below.

Value	Probability	$n(O) \times P(O)$
1	$\frac{1}{6}$	$1 \times \frac{1}{6} = \frac{1}{6}$
2	$\frac{1}{6}$	$2 \times \frac{1}{6} = \frac{1}{3}$
3	$\frac{1}{6}$	$3 \times \frac{1}{6} = \frac{1}{2}$
4	$\frac{1}{6}$	$4 \times \frac{1}{6} = \frac{2}{3}$
5	$\frac{1}{6}$	$5 \times \frac{1}{6} = \frac{5}{6}$
6	$\frac{1}{6}$	$6 \times \frac{1}{6} = 1$

Step 3: We add all of the values in that last column: $\frac{1}{6} + \frac{1}{3} + \frac{1}{2} + \frac{2}{3} + \frac{5}{6} + 1 = \frac{7}{2} = 3.5$. So, the expected value of a single roll of a die is 3.5.

1. Back in Example 3 in Section 7.5, we made this table of all of the equally likely outcomes :

	first die					
second die	**1**	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Step 1: Let's use to create the PDF for this experiment, as shown in the following table:

Value	Probability
2	$\frac{1}{36}$
3	$\frac{1}{18}$
4	$\frac{1}{12}$
5	$\frac{1}{9}$
6	$\frac{5}{36}$
7	$\frac{1}{6}$
8	$\frac{5}{36}$
9	$\frac{1}{9}$
10	$\frac{1}{12}$
11	$\frac{1}{18}$
12	$\frac{1}{36}$

Step 2: We can multiply each row to find $n(O) \times P(O)$as shown in the following table:

Value	Probability	$n(O) \times P(O)$
2	$\frac{1}{36}$	$\frac{1}{18}$
3	$\frac{1}{18}$	$\frac{1}{6}$
4	$\frac{1}{12}$	$\frac{1}{3}$
5	$\frac{1}{9}$	$\frac{5}{9}$
6	$\frac{5}{36}$	$\frac{5}{6}$
7	$\frac{1}{6}$	$\frac{7}{6}$
8	$\frac{5}{36}$	$\frac{10}{9}$
9	$\frac{1}{9}$	1
10	$\frac{1}{12}$	$\frac{5}{6}$
11	$\frac{1}{18}$	$\frac{11}{18}$
12	$\frac{1}{36}$	$\frac{1}{3}$

Step 3: We can add the last column to get the expected value:

$$\frac{1}{18}+\frac{1}{6}+\frac{1}{3}+\frac{5}{9}+\frac{5}{6}+\frac{7}{6}+\frac{10}{9}+1+\frac{5}{6}+\frac{11}{18}+\frac{1}{3}=7$$

So, the expected value is 7.

1. **Step 1:** Let's make a PDF table for this experiment. There are 3 events that we care about, so let's use those events in the table below:

Event	Probability
{A}	$\frac{1}{13}$
{10, J, Q, K}	$\frac{4}{13}$
{2, 3, 4, 5, 6, 7, 8, 9}	$\frac{8}{13}$

Step 2: Let's add a column to the following table for the values of each event:

Event	Probability	Value
{A}	$\frac{1}{13}$	15
{10, J, Q, K}	$\frac{4}{13}$	10
{2, 3, 4, 5, 6, 7, 8, 9}	$\frac{8}{13}$	5

Step 3: We'll add the column for the product of the values and probabilities to the table below:

Event	Probability	Value	$n(O) \times P(O)$
{A}	$\frac{1}{13}$	15	$\frac{15}{13}$
{10, J, Q, K}	$\frac{4}{13}$	10	$\frac{40}{13}$
{2, 3, 4, 5, 6, 7, 8, 9}	$\frac{8}{13}$	5	$\frac{40}{13}$

Step 4: We'll find the sum of the last column: $\frac{15}{13} + \frac{40}{13} + \frac{40}{13} = \frac{95}{13} \approx 7.3$. Thus, the expected *Rummy* value of a randomly selected card is about 7.3.

Let's make note of some things we can learn from Example 1. First, as Exercises 1 and 3 demonstrate, the expected value of an experiment might not be a value that could come up in the experiment. Remember that the expected value is interpreted as a mean, and the mean of a collection of numbers doesn't have to actually be one of those numbers.

Second, looking at Exercise 1, the expected value (3.5) was just the mean of the numbers on the faces of the die: $\frac{1+2+3+4+5+6}{6} = 3.5$. This is no accident! If we break that fraction up using the addition in the numerator, we get $\frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6}$, which we can rewrite as $1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6}$. That's exactly the computation we did to find the expected value! In fact, expected values can always be treated as a special kind of mean called a **weighted mean**, where the weights are the probabilities associated with each value. When the probabilities are all equal, the weighted mean is just the regular mean.

Interpreting Expected Values

As we noted, the expected value of an experiment is the mean of the values we would observe if we repeated the experiment a large number of times. (This interpretation is due to an important theorem in the theory of probability called the Law of Large Numbers.) Let's use that to interpret the results of the previous example.

Example 2 **Interpreting Expected Values**

Interpret the expected values of the following experiments.

1. Roll a standard 6-sided die and note the number showing.
2. Roll 2 standard 6-sided dice and note the sum of the numbers showing.
3. Draw a card from a well-shuffled standard deck of cards and note its *Rummy* value (15 for aces; 10 for tens, jacks, queens, and kings; 5 for everything else).

Solution

1. If you roll a standard 6-sided die many times, the mean of the numbers you roll will be around 3.5.

2. If you roll a pair of standard 6-sided dice many times, the mean of the sums of the numbers you roll will be about 7.
3. If you draw a card from a well-shuffled deck many times, the mean of the *Rummy* values of the cards would be around 7.3.

WHO KNEW?

Pascal's Wager

The French scholar Blaise Pascal (1623–1662) was among the earliest mathematicians to study probabilities, and was the first to accurately describe and compute expected values. In his book *Pensées* (*Thoughts*), he turned the analysis of expected values to his belief in the Christian God. He said that there is no way for people to establish the probability that God exists, but since the "winnings" on a bet that God exists (and that you then lead your life accordingly) are essentially infinite, the expected value of taking that bet is always positive, no matter how unlikely it is that God exists.

Using Expected Value

Now that we know how to find and interpret expected values, we can turn our attention to using them. Suppose someone offers to play a game with you. If you roll a die and get a 6, you get \$10. However, if you get a 5 or below, you lose \$1. Is this a game you'd want to play? Let's look at the expected value: The probability of winning is $\frac{1}{6}$ and the probability of losing is $\frac{5}{6}$, so the expected value is $\$10 \times \frac{1}{6} + (-\$1) \times \frac{5}{6} = \frac{5}{6} \approx \$0.83$. That means, on average, you'll come out ahead by about 83 cents every time you play this game. It's a great deal! On the other hand, if the winnings for rolling a 6 drop to \$3, the expected value becomes $\$3 \times \frac{1}{6} + (-\$1) \times \frac{5}{6} = -\frac{1}{3} \approx -\$0.33$, meaning you should expect to lose about 33 cents on average for every time you play. Playing that game is not a good idea! In general, this is how casinos and lottery corporations make money: Every game has a negative expected value for the player.

WHO KNEW?

Expected Values in Football

In the 21st century, data analytics tools have revolutionized the way sports are coached and played. One tool in particular is used in football at crucial moments in the game. When a team faces a fourth down (the last possession in a series of four possessions, a fairly common occurrence), the coach faces a decision: Run one play to try to gain a certain number of yards, or kick the ball away to the other team. Here's the interesting part of the decision: If the team "goes for it" and runs the play and they are successful, then they keep possession of the ball and can continue in their quest to score more points. If they are unsuccessful, then they lose possession of the ball, giving the other team an opportunity to score points. If, instead, the team punts, or kicks the ball away, then the other team gets possession of the ball, but in a worse position for them than if the original team goes for it and fails. To analyze this situation, data analysts have generated empirical probabilities for every fourth down situation, and computed the expected value (in terms of points) for each decision. Coaches frequently use those calculations when they decide which option to take!

PEOPLE IN MATHEMATICS

Pierre de Fermat and Blaise Pascal

In 1654, a French writer and amateur mathematician named Antoine Gombaud (who called himself the Chevalier du Mére) reached out to his gambling buddy Blaise Pascal to answer a question that he'd read about called the "problem of points." The question goes like this: Suppose you're playing a game that is scored using points, and the first person to earn 5 points is the winner. The game is interrupted with the score 4 points to 2. If the winner stood to win $100, how should the prize money be divided between the players? Certainly the person who is 1 point away from victory should get more, but *how much* more?

We have developed tools in this section to answer this question. At its heart, it's a question about conditional probabilities and expected value. At the time that Pascal first started thinking about it, though, those ideas hadn't yet been invented. Pascal reached out to a colleague named Pierre de Fermat, and over the course of a couple of months, their correspondence with each other would eventually solve the problem. In the process, they first described conditional probabilities and expected values!

Apart from their work in probability, these men are famous for other work in mathematics (and, in Pascal's case, philosophy and physics). Fermat is remembered for his work in geometry and in number theory. After his death, the statement of what came to be called "Fermat's Last Theorem" was discovered scribbled in the margin of a book, with the note that Fermat had discovered a "marvelous proof that this margin is too small to contain." The theorem says that any equation of the form $a^n + b^n = c^n$ has no positive integer solutions if $n \geq 3$. No proof of that theorem was discovered until 1994, when Andrew Wiles used computers and new branches of geometry to finally prove the theorem!

Pascal is remembered for the "arithmetical triangle" that is named for him (though he wasn't the first person to discover it; see the section on the binomial distribution for more), as well as work in geometry. In physics, Pascal worked on hydrodynamics and air pressure (the SI unit for pressure is named for him), and in philosophy, Pascal advocated for a mathematical approach to philosophical problems.

Example 3 **Using Expected Values**

In the casino game keno, a machine chooses at random 20 numbers between 1 and 80 (inclusive) without replacement. Players try to predict which numbers will be chosen. Players don't try to guess all 20, though; generally, they'll try to predict between 1 and 10 of the chosen numbers. The amount won depends on the number of guesses they made and the number of guesses that were correct.

1. At one casino, a player can try to guess just 1 number. If that number is among the 20 selected, the player wins \$2; otherwise, the player loses \$1. What is the expected value?
2. At the same casino, if a player makes 2 guesses and they're both correct, the player wins \$14; otherwise, the player loses \$1. What is the expected value?
3. Players can also make 3 guesses. If 2 of the 3 guesses are correct, the player wins \$1. If all 3 guesses are correct, the player wins \$42. Otherwise, the player loses \$1. What is the expected value?
4. Which of these games is the best for the player? Which is the best for the casino?

Solution

1. There are 20 winning numbers out of 80, so if we try to guess one of them, the probability of guessing correctly is $\frac{20}{80} = \frac{1}{4}$. The probability of losing is then $\frac{3}{4}$, and so the expected value is $\$2 \times \frac{1}{4} + (-\$1) \times \frac{3}{4} = -\$0.25$.
2. There are $20C_2 = 190$ winning choices out of $80C_2 = 3,160$ total ways to choose 2 numbers. So, the probability of winning is $\frac{190}{3,160}$ and the probability of losing is $\frac{3,160-190}{3,160} = \frac{2,970}{3,160}$. So, the expected value of the game is $\$14 \times \frac{190}{3,160} + (-\$1) \times \frac{2970}{3160} \approx -\$0.10$.
3. **Step 1:** Let's start with the big prize. There are $20C_3 = 1,140$ ways to correctly guess 3 winning numbers out of $80C_3 = 82,160$ ways to guess three numbers total. That means the probability of winning the big prize is $\frac{1,140}{82,160} \approx 0.01388$. **Step 2:** Let's find the probability of the second prize. The denominator is the same: 82,160. Let's figure out the numerator. To win the second prize, the player must pick 2 of the 20 winning numbers and one of the 60 losing numbers. The number of ways to do that can be found using the Multiplication Rule for Counting: there are $20C_2 = 190$ ways to pick 2 winning numbers and 60 ways to pick 1 losing number, so there are $190 \times 60 = 11,400$ ways to win the second prize. So, the probability of winning that second prize is $\frac{11,400}{82,160} \approx 0.13875$.

Step 3: Since the overall probability of winning is $\frac{1,140}{82,160} + \frac{11,400}{82,160} = \frac{12,540}{82,160} \approx 0.15263$, the probability of losing must be $1 - 0.15263 = 0.84737$. So, the expected value is $\$42 \times 0.01388 + \$1 \times 0.13875 + (-\$1) \times 0.84737 \approx -\$0.13$.

1. The bet that's the best for the player is the one with the highest expected value for the player, which is guessing two numbers. The best one for the casino is the one with the lowest expected value for the player, which is guessing one number.

WORK IT OUT

Make Your Own Lottery

By yourself or with a partner, devise your own lottery scheme. Assume you would have access to one or more machines that choose numbers randomly. What will a lottery draw look like? How many numbers are players choosing from? How many will be drawn? Will they be drawn with replacement or without replacement? What conditions must be met for a player to win first or second (or more!) prize? Once you've decided that, decide the payoff structure for winners, and how much the game will cost to play. Try to make the game enticing enough that people will want to play it, but with enough negative expected value that the lottery will make money. Aim for the expected value to be about -0.25 times the cost of playing the game.

Key Terms

- expected value

Key Concepts

- The expected value of an experiment is the sum of the products of the numerical outcomes of an experiment with their corresponding probabilities.
- The expected value of an experiment is the most likely value of the average of a large number of replications of the experiment.

Formulas

- If O represents an outcome of an experiment and $n(O)$ represents the value of that outcome, then the expected value of the experiment is:

$$\sum n(O) \times P(O)$$

where Σ stands for the sum, meaning we add up the results of the formula that follows over all possible outcomes.

Projects

1. The Binomial Distribution is one of many examples of a *discrete probability distribution*. Other examples include the Geometric, Hypergeometric, Multinomial, Poisson, and Negative Binomial Distributions. Choose one of these distributions, and find out what makes it different from the Binomial Distribution. In what situations can it be applied? How is it used? Once you have an idea of how it's used, write a series of five questions like the ones in this chapter that can be answered with that distribution, and find the answers.
2. *Binomial* is a word that also comes up in algebra; the word describes polynomials with two terms. At first glance, there isn't much to indicate that these two uses of the word are related, but it turns out there is a connection. Explore the connection between the Binomial Distribution and the algebraic concept of binomial expansion, (the process of multiplying out expressions like $(x + y)^n$ for a positive whole number n). Search for a connection with the mathematical object known as Pascal's Triangle.
3. Hazard is a dice game that was mentioned in Chaucer's *Canterbury Tales*. It was a popular game of chance played in taverns and coffee houses well into the 18th century; its popularity at the time of the foundation of probability theory means that it was a common example in

early texts on finding expected values and probabilities. Find the rules of the game, and get some practice playing it. Then, analyze the choices that the caster gets to make, and decide which is most advantageous, using the language of expected values.