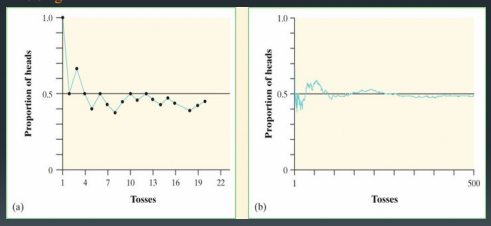
Chapter 5: Probability: What are the Chances?

Section 5.1
Randomness, Probability, and Simulation

HAPPY HALLOWEEN!!!!!!

Example 1: When you toss a coin, there are only two possible outcomes, heads or tails. The figure below on the left shows the results of tossing a coin 20 times. For each number of tosses from 1 to 20, we have plotted the proportion of those tosses that gave a head. You can see that the proportion of heads starts at 1 on the first toss, falls to 0.5 when the second toss gives a tail, then rises to 0.67, and then falls to 0.5, and 0.4 as we get two more tails. After that, the proportion of heads continues to fluctuate but never exceeds 0.5 again.



Suppose we keep tossing the coin until we have made 500 tosses. The figure above on the right shows the results. The proportion of tosses that produce heads is quite variable at first. As we make more and more tosses, however, the proportion of heads gets close to 0.5 and stays there.

The Idea of Probability

Chance behavior is unpredictable in the short run, but has a regular and predictable pattern in the long run.

The **law of large numbers** says that if we observe more and more repetitions of any chance process, the proportion of times that a specific outcome occurs approaches a single value.

Definition:
The **probability** of any outcome of a chance process is a number between 0 and 1 that describes the proportion of times the outcome would occur in a very long series of repetitions.

Example 2: How much should a company charge for an extended warranty for a specific type of cell phone? Suppose that 5% of these cell phones under warranty will be returned, and the cost to replace the phone is $150. If the company knew which phones would go bad, it could charge $150 for these phones and $0 for the rest. However, the company can't know which phones will be returned. Because about 1 in every 20 will be returned, they should charge at least 150/20 = $7.50 for the extended warranty.

Example 3: According to the "Book of Odds," the probability that a randomly selected U.S. adult usually eats breakfast is 0.61.

a) Explain what probability 0.61 means in this setting.

This probably means that if you asked a large sample of U.S. adults whether they usually eat breakfast, about 61% of them will answer "Yes."

b) Why doesn't this probability say that if 100 U.S. adults are chosen at random, exactly 61 of them usually eat breakfast?

In a random sample of 100 adults, we would expect that around 61 of them will usually eat breakfast. However, the exact number will vary from sample to sample.

Myths about Randomness

The idea of probability seems straightforward. However, there are several myths of chance behavior we must address.

The myth of short-run regularity:

The idea of probability is that randomness is predictable in the long run. Our intuition tries to tell us random phenomena should also be predictable in the short run. However, probability does not allow us to make short-run predictions.

The myth of the "law of averages":

Probability tells us random behavior evens out in the long run. Future outcomes are not affected by past behavior. That is, past outcomes do not influence the likelihood of individual outcomes occurring in the future.

"Random" doesn't mean haphazard. In statistics, random means "by chance."

Don't confuse the law of large numbers, which describes the big idea of probability, with the "law of averages" described here.

Example 4: Toss a coin six times and record heads (H) or tails (T) on each toss. Which of the following outcomes is more probable?

HTHTTH TTTHHH

Almost everyone says that HTHTTH is more probable, because TTTHHH does not "look random." In fact, both are equally likely. That heads and tails are equally probable says only that about half of a very long sequence of tosses will be heads. It doesn't say that heads and tails must come close to alternating in the short run. The coin has no memory. It doesn't know what past outcomes were, and it can't try to create a balanced sequence.

Example 5: Is there such a thing as a "hot hand" in basketball? Belief that runs must result from something other than "just chance" influences behavior. If a basketball player makes several consecutive shots, both the fans and her teammates believe that she has a "hot hand" and is more likely to make the next shot. Several studies have shown that runs of baskets made or missed are no more frequent in basketball than would be expected if the result of each shot is unrelated to the outcomes of the player's previous shots. If a player makes half her shots in the long run, her made shots and misses behave just like tosses of a coin—and that means that runs of makes and misses are more common than our intuition expects.

Free throws may be a different story. A recent study suggests that players who shoot two free throws are *slightly* more likely to make the second shot if they make the first one.

Example 6: In casinos, there is often a large display next to every roulette table showing the outcomes of the previous spins of the wheel. The results of previous spins reveal nothing about the results of future spins, so why do the casinos pay for these displays? The casinos know that many people will be more willing to place a bet if they observe a pattern in the previous outcomes. And as long as people are placing bets, the casino is making money.

Simulation

The imitation of chance behavior, based on a model that accurately reflects the situation, is called a **simulation**.

Performing a Simulation

State: Ask a question of interest about some chance process.

Plan: Describe how to use a chance device to imitate one repetition of the process. Tell what you will record at the end of each repetition.

Do: Perform many repetitions of the simulation.

Conclude: Use the results of your simulation to answer the question of interest.

We can use physical devices, random numbers (e.g. Table D), and technology to perform simulations.

Example 7: At a local high school, 95 students have permission to park on campus. Each month, the student council holds a "golden ticket parking lottery" at a school assembly. The two lucky winners are given reserved parking spots next to the school's main entrance. Last month, the winning tickets were drawn by a student council member from the AP Statistics class. When both golden tickets went to members of that same class, some people thought the lottery had been rigged. There are 28 students in the AP Statistics class, all of whom are eligible to park on campus. Design and carry out a simulation to decide whether it's plausible that the lottery was carried out fairly.

Students	Labels
AP Statistics Class	01-28
Other	29-95
Skip numbers from 96-00	

Reading across row 139 in Table D, look at pairs of digits until you see two different labels from 01-95. Record whether or not both winners are members of the AP Statistics Class.

55 | 58	89 | 94	04 | 70	70 | 84	10|98|43	56 | 35	69 | 34	48 | 39	45 | 17
X | X	X | X	✓ | X	X | X	✓|Sk|X	X | X	X | X	X | X	X | ✓
No	No	No	No	No	No	No	No	No

19 | 12	97|51|32	58 | 13	04 | 84	51 | 44	72 | 32	18 | 19	40|00|36	00|24|28
✓ | ✓	Sk|X|X	X | ✓	✓ | X	X | X	X | X	✓ | ✓	X|Sk|X	Sk|✓|✓
Yes	No	No	No	No	No	**Yes**	No	**Yes**

Based on 18 repetitions of our simulation, both winners came from the AP Statistics class 3 times, so the probability is estimated as 16.67%.

Example 8: In an attempt to increase sales, a breakfast cereal company decides to offer a NASCAR promotion. Each box of cereal will contain a collectible card featuring one of these NASCAR drivers: Jeff Gordon, Dale Earnhardt, Jr., Tony Stewart, Danica Patrick, or Jimmie Johnson. The company says that each of the 5 cards is equally likely to appear in any box of cereal. A NASCAR fan decides to keep buying boxes of the cereal until she has all 5 drivers' cards. She is surprised when it takes her 23 boxes to get the full set of cards. Should she be surprised? Design and carry out a simulation to help answer this question.

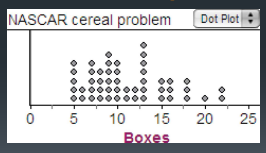
Driver	Label
Jeff Gordon	1
Dale Earnhardt, Jr.	2
Tony Stewart	3
Danica Patrick	4
Jimmie Johnson	5

Use randInt(1,5) to simulate buying one box of cereal and looking at which card is inside. Keep pressing Enter until we get all five of the labels from 1 to 5. Record the number of boxes we had to open.

3 5 2 1 5 2 3 5 4 **9 boxes**

4 3 5 3 5 1 1 1 5 3 1 5 4 5 2 **15 boxes**

5 5 5 2 4 1 2 1 5 3 **10 boxes**



We never had to buy more than 22 boxes to get the full set of cards in 50 repetitions of our simulation. Our estimate of the probability that it takes 23 or more boxes to get a full set is roughly 0.

AP EXAM TIP: On the AP exam, you may be asked to describe how you will perform a simulation using rows of random digits. If so, provide a clear enough description of your simulation process for the reader to get the same results you did from only your written explanation.