CHAPTER 5
Probability: What Are the Chances?

5.1

Randomness, Probability, and Simulation

The Practice of Statistics, 5th Edition
Starnes, Tabor, Yates, Moore

Bedford Freeman Worth Publishers

Randomness, Probability, and Simulation

After this section, you should be able to:

- INTERPRET probability as a long-run relative frequency.
- USE simulation to MODEL chance behavior.

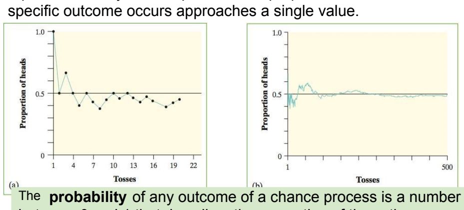
We make decisions based on uncertainty every day.

- Most days it takes me 20 minutes to drive to work, but how often will it take longer?
- Should I take my umbrella with me today, even though it's not raining now?
- Should I buy an extended warranty for my new phone?
- Should I buy a lottery ticket?

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.



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.

How is probability calculated?

Long-run relative frequency (experimental): The probability of any outcome in a chance process is a number between 0 and 1, and it describes the proportion of times that outcome would occur in a very long series of repetitions.

Subjective approach: individual judgment based on facts and including a personal evaluation of available information.

Theoretical: when outcomes are equally likely, we can state the probability of an outcome with a high degree of certainty

Myths About Randomness

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.

If you flip a coin ten times a get 10 heads,does this affect the probability of getting a head on the next flip?

Simulation

Simulation is a powerful tool for modeling chance behavior.

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.

STEP 4

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

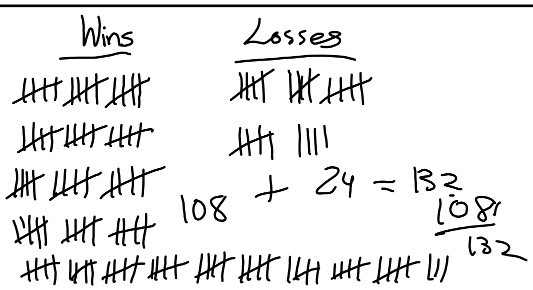
The dice game of 21 - each player rolls a die repeatedly trying to get the highest score without going over 21.

Play!

Suppose your opponent has 18. How many rolls should you expect to make and what are the chances of you winning.

Simulate!

- What are the components and how will you model them?
- How can we combine components to model a trial?



The baseball World Series consists of up to 7 games, with the first team to win 4 games the winner. The first 2 games are played at Team A's home field, then 3 games at Team B's field, then if necessary, back to Team A's field. Records over the past century show that there is a home field advantage, with the home team winning about 55% of the time. Does the system of alternating fields even out the home field advantage? How often will Team A win the series?

1) What component will be repeated?

2) How can we model this using random digits?

3) How can we model a trial by combining components?

4) How will we analyze the reponse variable?

Estimating probability:

Tossing the pig:

Toss the pig many times, and record the number of each of the following landings:

On feet

On snout

On left side

On right side

On back

Calculate the probability for each type of landing.

Example: Simulations with technology

In an attempt to increase sales, a breakfast cereal company decides to offer a promotion. Each box of cereal will contain a collectible card featuring one of these rappers: Migos, Lil Uzi Vert, Chance the Rapper, Kyle and Drake.

The company says that each of the 5 cards is equally likely to appear in any box of cereal.

A fan decides to keep buying boxes of the cereal until she has all 5 cards. She is surprised when it takes her 23 boxes to get the full set of cards. Should she be surprised?

Problem What is the probability that it will take 23 or more boxes to get a full set of all 5 collectible cards?

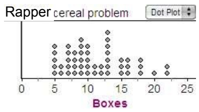
Example: Simulations with technology

Plan: We need five numbers to represent the five possible cards.
Let's let 1 = Migos,
2 = Lil Uzi Vert,
3 = Chance the Rapper,
4 = Kyle, and
5 = Drake.

We'll use randInt(1,5) to simulate buying one box of cereal and looking at which card is inside.

Because we want a full set of cards, we'll keep pressing Enter until we get all five of the labels from 1 to 5. We'll record the number of boxes that we had to open.

Example: Simulations with technology



Here is some output from a Fathom simulation.

Conclude: We never had to buy more than 22 boxes to get the full set of rapper cards in 50 repetitions of our simulation. So our estimate of the probability that it takes 23 or more boxes to get a full set is roughly 0. The fan should be surprised about how many boxes she had to buy.

Randomness, Probability, and Simulation

Section Summary

In this section, we learned how to…

- INTERPRET probability as a long-run relative frequency.
- USE simulation to MODEL chance behavior.