Lesson 16: What's randomness all about?

What's your idea

of random...?

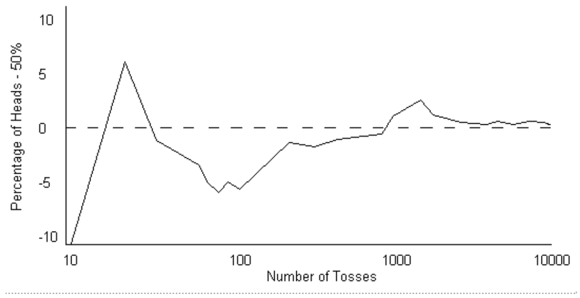
What is it about chance outcomes being random that makes random selection seem fair?

Two things:

1. Nobody can guess the outcome before it happens.

2. When we want things to be fair, usually some underlying set of outcomes will be equally likely (although in many games some combinations of outcomes are more likely than others).

Random, think again?? Structure in chaos.

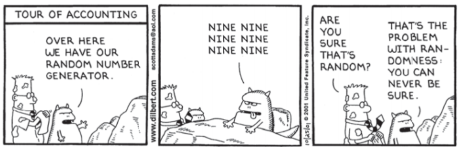


Random outcomes have structure when viewed in the long run.

Look at the numbers quickly, pick a number at random, and write it down........quickly!!

1 2 3 4

75 %	pick 3
20%	pick 2 or 4
5%	pick 1



It's not easy being random.
Computers generate pseudorandom numbers

produced by mathematical procedure...but passes some statistical test for randomness.

Aren't you done shuffling yet? Even something as common as card shuffling may not be as random as you might think. If you shuffle cards by the usual method in which you split the deck in half and try to let cards fall roughly alternately from each half, you're doing a "riffle shuffle."

How many times should you shuffle cards to make the deck random? A surprising fact was discovered by statisticians Persi Diaconis, Ronald Graham, and W. M. Kantor. It takes seven riffle shuffles. Fewer than seven leaves order in the deck, but after that, more shuffling does little good. Most people, though, don't shuffle that many times.

When computers were first used to generate hands in bridge tournaments, some professional bridge players complained that the computer was making too many "weird" hands—hands with 10 cards of one suit, for example. Suddenly these hands were appearing more often than players were used to when cards were shuffled by hand. The players assumed that the computer was doing something wrong. But it turns out that it's humans who hadn't been shuffling enough to make the decks really random and have those "weird" hands appear as often as they should.

There is no one right answer to many important questions.

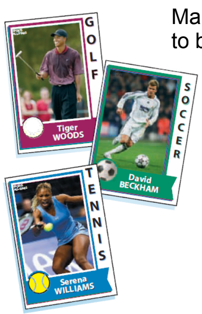
the course changes character from a discussion of ways to describe the world by looking at, displaying, and summarizing data, to a lab course in which we generate results experimentally.

The best ways we know to generate data that give a fair and accurate picture of the world rely on randomness, and the ways in which we draw conclusions from those data depend on randomness, too.

Simulation

"random" is the same as "equally likely." ????

Just because there are two outcomes (win or lose), your chance of winning the lottery is not 50%.



Manufacturer puts cards in cereal boxes to boost sales.

"You gotta catch 'em all"

- 20% of boxes contain picture of Tiger Woods
- 30% of boxes contain picture of David Becham
- 50% of boxes contain picture of Serena Williams

How many cereal boxes do you expect to buy in order to get a complete set?

Simulation

it's cheaper than buying boxes

imitation of the real process that we can manipulate and control

Can't be less than three boxes ?

Building a Simulation

1. **Identify the component to be repeated.** In this case, our component is the opening of a box of cereal.
2. **Explain how you will model the component's outcome.** The digits from 0 to 9 are equally likely to occur. Because 20% of the boxes contain Tiger's picture, we'll use 2 of the 10 digits to represent that outcome. Three of the 10 digits can model the 30% of boxes with David Beckham cards, and the remaining 5 digits can represent the 50% of boxes with Serena. One possible assignment of the digits, then, is

   0, 1 Tiger   2, 3, 4 Beckham   5, 6, 7, 8, 9 Serena.

Specify how to simulate trials:

3. **Explain how you will combine the components to model a trial.** We pretend to open boxes (repeat components) until our collection is complete. We do this by looking at each random digit and indicating what picture it represents. We continue until we've found all three.
4. **State clearly what the response variable is.** What are we interested in? We want to find out the number of boxes it might take to get all three pictures.

Put it all together to run the simulation:

5. **Run several trials.** For example, consider the third line of random digits shown earlier (p. 257):

   89064273086456814121982266538858732858016990278431103804200676 64.

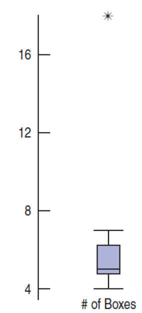
   Let's see what happened.

Trial Number	Component Outcomes	Trial Outcomes: y = Number of boxes
1	89064 = Serena, Serena, Tiger, Serena, Beckham	5

Analyze the response variable:

outcome of trial

6. **Collect and summarize the results of all the trials.** You know how to summarize and display a response variable. You'll certainly want to report the shape, center, and spread, and depending on the question asked, you may want to include more.
7. **State your conclusion,** as always, in the context of the question you wanted to answer. Based on this simulation, we estimate that customers hoping to complete their card collection will need to open a median of 5 boxes, but it could take a lot more.



Model vs. Real Outcomes

the mean number of cereal boxes the simulation suggests you might have to buy to get all three pictures is not what will actually happen when you walk into the grocery store.

Example: The game of 21 can be played with an ordinary 6-sided die. Competitors each roll the die repeatedly, trying to get the highest total less than or equal to 21. If your total exceeds 21, you lose.

Suppose your opponent has rolled an 18. Your task is to try to beat him by getting more than 18 points without going over 21. How many rolls do you expect to make, and what are your chances of winning?

1. Identify the component to be repeated.

2. Explain how you will model the component's outcome.

3. Explain how you will combine the components to model a trial.

4. State clearly what the response variable is.

5. Run several trials.

6. Collect and summarize the results of all the trials.

7. State your conclusion.

Check with page 259

Using the calculator for random digits