Chapter 9

Understanding Randomness

D. Raffle
5/28/2015

What is randomness?

Most people have some idea what a random event is.

· Flipping a coin
· Drawing a card from a deck
· Rolling a die

The concept of randomness isn't just restricted to games. In statistics, we rely on the concept of randomness.

· To describe uncertainty
· To collect **representative** samples
· To come up with methods to draw conclusions and make estimations

2/9

Defining Randomness

What does it mean for something to be random? First, we need to establish definitions

· Let's call something which has a random outcome an **trial**.
· We'll call the possible outcome of a trial **events**.

So what makes a trial random?

· For the trial, we don't know what event will occur ahead of time

Example: Flipping a coin

· The actual flipping of the coin is a trial
· The possible events are getting a heads or getting a tails

3/9

Random ≠ Unpredictable

Random trials only have a certain number of outcomes

- If we flip a coin, it can only land heads up or tails up
- If the coin is fair, both events are equally likely

They are well-behaved in the longterm

- If we flip a coin once, we don't which side will land up
- In the longterm, however, we should see about half heads and half tails

4/9

In Real Life

We can often view real-world events as random trials. Think about a morning commute.

· When we leave in the morning, we're can't be sure exactly how long it will take to get to work (or school)

· Traffic or construction can hold us up

· It's not completely unpredictable, however

· We might know how long an average (or expected) commute is, and can plan around it

Using Randomness

Why do statisticians care about randomness? Mostly to eliminate **bias** in our studies.

· We usually need to estimate values we don't know, like the percentage of people who will vote for a presidential candidate

· We can try to estimate this by asking a fraction of the people as they leave their polls

· This only works if we get a representative cross section

· If we poll too many people from one party, we might predict that they'll win when they didn't

· We ensure that our samples are representative by making sure they are random

· Unfortunately, people are very poor at operating randomly

6/9

Which Plot is More Random?
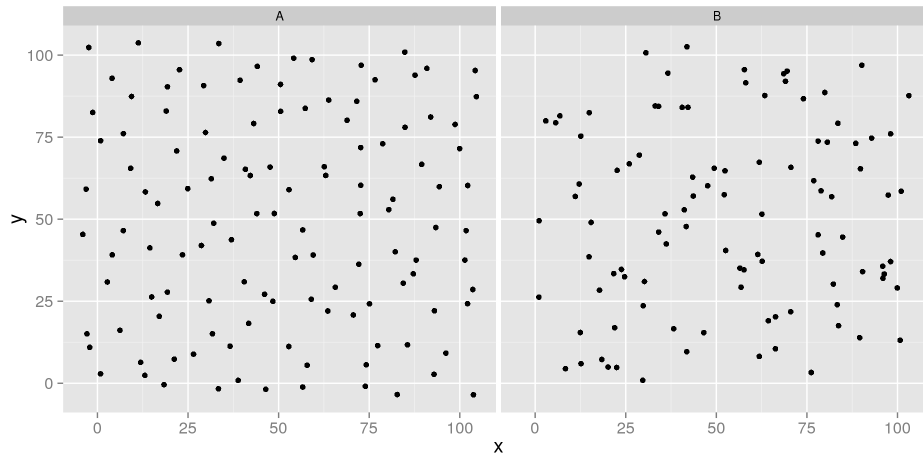
Pick a Number at Random

$$1 \quad 2 \quad 3 \quad 4$$

Did we get random results?

· Almost 75% of people choose 3
· 20% chooose 3 or 4
· Only 5% of people choose 1

8/9

So what do we do?

How can statisticians and researchers get random samples if we can't trust
ourselves?

· Historically, we used what are called random number tables
· Modern researchers used software

```
sample(1:20, 5)
```

```
## [1]  2 19  7 10  9
```

```
rnorm(n = 5, mean = 100, sd = 5)
```

```
## [1]  95.85033  96.60495  99.02627  95.46636 101.26399
```