Today's topics:

- **Descriptive statistics**
- **Inferential statistics**

Background preparation:

- *Kaplan (2016), Appendix*

0. Today's objectives

After today's class, you should be able to:

- Explain the different purposes of descriptive versus inferential statistics, and identify examples of each
- Explain the following key statistics terms and what they can tell us about a data set: *mean, standard deviation, correlation, null hypothesis,* p-*value*
- Discuss some points to watch out for when using statistics to describe and interpret a data set, especially *outliers*, *bimodal distributions,* and *correlation vs. causation*

1. Descriptive and inferential statistics

- How many Harry Potter movies have the members of this class seen?

  9 1 1 11 9 2 11 1 9 0 0 0 5 3 11 8 3 0 11
  7 9 9 2 9 9 8 3 10 9 8 11 8 8

- What are some ways we can **communicate** this data more effectively?

1. Descriptive and inferential statistics

- We can report the **mean**: 5.3

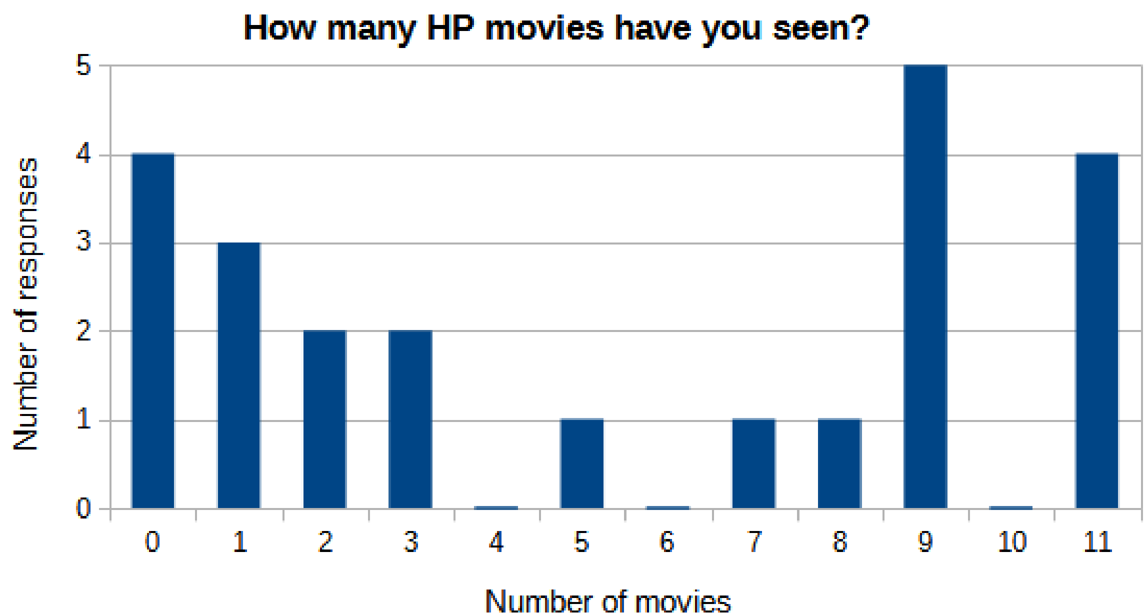
  the **median**: 5

  - These are ways of **summarizing** the data that we have collected in our experiment

- Are these **descriptive** or **inferential** statistics?

1. Descriptive and inferential statistics

- We can make a **data graphic**:



- How **representative** are the mean and the median of the individual values in this data set? Why?

1. Descriptive and inferential statistics

Descriptive statistics

- Purpose: To **summarize** the data we have collected
- Commonly encountered:
  - Mean
  - Standard deviation
  - Correlation

Inferential statistics

- Purpose: To determine whether we can make **predictions** or **generalizations** *beyond* the data we have collected

1. Descriptive and inferential statistics

Mean

- What is the mean of these values?

  4 4 2 4 16

  - How did you calculate this?

- What is the **concept** behind the mean?

1. Descriptive and inferential statistics

Mean

- What is the **concept** behind the mean?
  - The amount each item would contribute to the total if **all contributions were equal**

1. Descriptive and inferential statistics

Mean

- How is the mean (potentially) useful for:
  - **Describing** a data set?
  - Making **predictions**?

1. Descriptive and inferential statistics

Mean

- How is the mean (potentially) useful for:
  - **Describing** a data set?
    - Describes a **central tendency** of the data set
  - Making **predictions**?
    - *Might* give "**expected value**" for future cases
    - *Whether* this is a legitimate prediction to make can be **tested** with **inferential statistics**

1. Descriptive and inferential statistics

Mean

- What are some potential **pitfalls** with using the mean in these ways?

1. Descriptive and inferential statistics

Mean

- What are some potential **pitfalls** with using the mean in these ways?

- *Descriptive and predictive:* The mean might not resemble any actual value in the data set
  - Extreme **outliers** can skew the mean
    - Geography majors at UNC—the highest average salary after graduation?
  - The data set might be **bimodal**
    - Age: parents and toddlers

1. Descriptive and inferential statistics

Mean

- What are some potential **pitfalls** with using the mean in these ways?
- *Predictive:* Were the items measured actually **representative** of their category?
  → *inferential statistics*

1. Descriptive and inferential statistics

Standard deviation

- What does the standard deviation of a set of numbers indicate?

1. Descriptive and inferential statistics

Standard deviation | Kaplan (2016: 266; my emphasis):

- The standard deviation reflects the **amount of 'spread'** in the data: a small SD means that the numbers in the set are clustered tightly around the mean, while a larger SD means that the numbers are more spread out.
- As a rule of thumb, more than half of the numbers in the set will fall within one standard deviation of the mean, and the vast majority will fall within two standard deviations—
- but all this depends on the specific properties of the set, and there are no guarantees.

1. Descriptive and inferential statistics

Standard deviation

- How to calculate standard deviation (FYI only)
  - *Deviance* (for each data point): Find the difference from the mean
  - *Sum of squared deviances*: Square each deviance value and add them up
  - Variance: Divide the sum of squared deviances by (the number of data points minus 1)
  - ***Standard deviation:*** Square root of the variance
- A good source for basic statistics: *Concepts & Applications of Inferential Statistics*, by Richard Lowry—free online textbook at *http://vassarstats.net/textbook/*

1. Descriptive and inferential statistics

Standard deviation

- Why is it important?

  Are words *meaningfully* longer in intoxicated speech?

  (Kaplan 2016: 266)

Table A.1 *Average pitch (fundamental frequency), duration, and loudness (intensity) of consonant-vowel-consonant words in sober and intoxicated speech.*

	Sober		Intoxicated	
Measure	*M*	(*SD*)	*M*	(*SD*)
Pitch (Hz)	154.0	(51.7)	156.4	(52.3)
Duration (ms)	442	(120)	467	(127)
Loudness (dB)	78.2	(5.8)	75.1	(5.3)

- Many inferential statistics methods **use standard deviation** in their calculations

1. Descriptive and inferential statistics

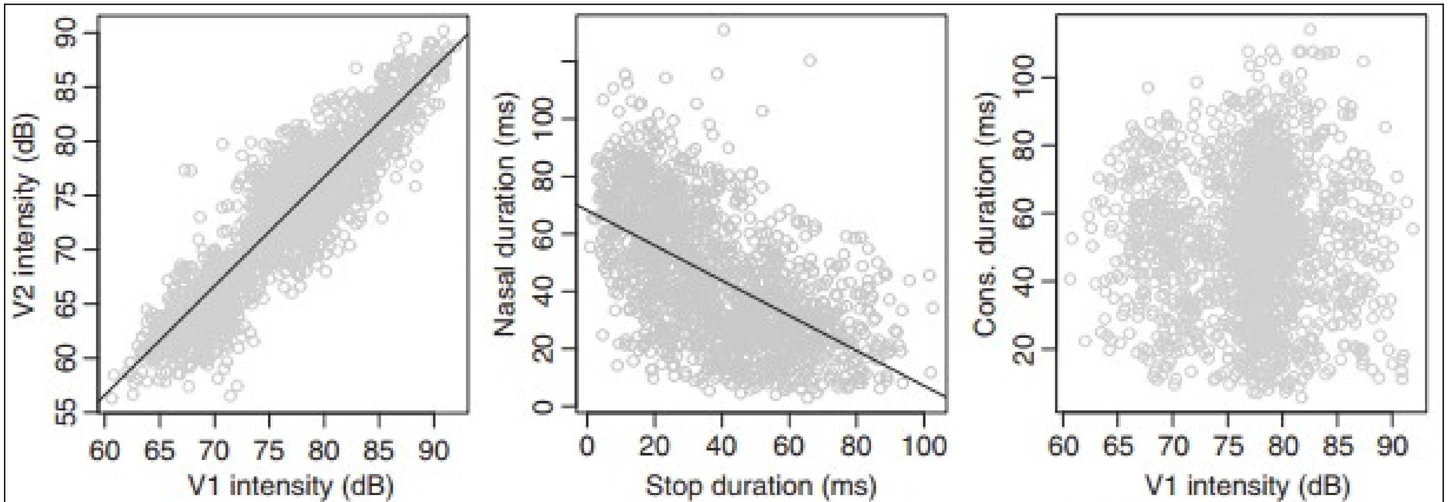
Correlation

- What is correlation (***r***)? What does positive / zero / negative correlation mean?

- What does correlation look like in a **scatterplot**?

1. Descriptive and inferential statistics

Correlation | In a **scatterplot** (Kaplan 2016: 268)



(a) Relationship between intensity of the first and second vowels in words like *rapid*, $r = .894$.

(b) Relationship between duration of the stop (*p*) and nasal (*m*) in words like *trumpet*, $r = -.565$.

(c) Relationship between intensity of the first vowel and duration of the middle consonant in words like *rapid*, $r = -.001$.

Figure A.1 Examples of strong (a), moderate (b), and no (c) correlation.

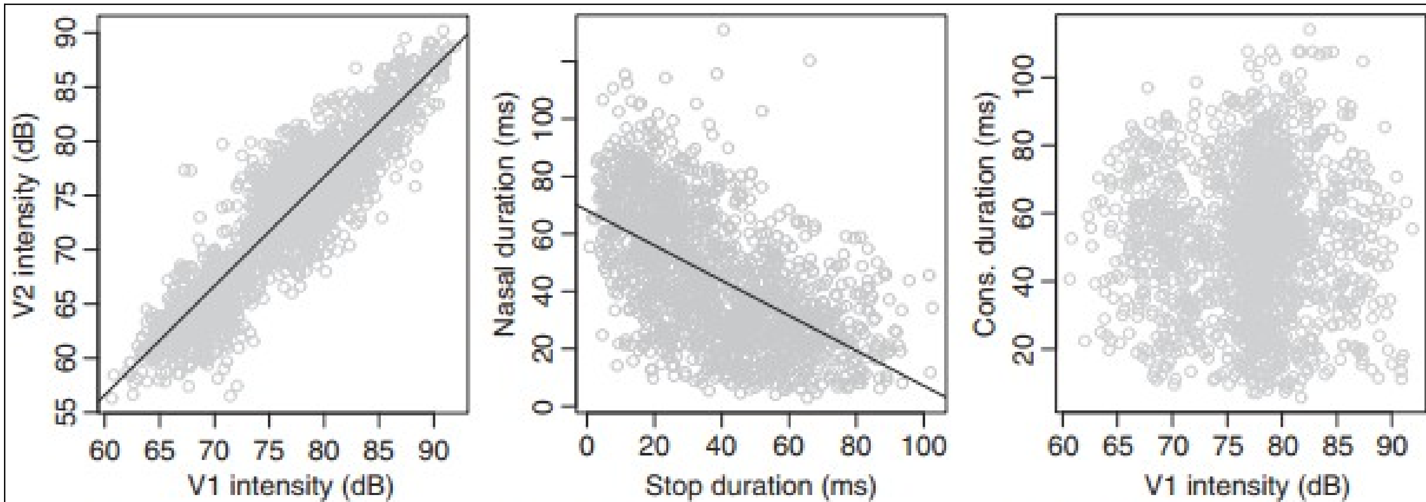
- Is correlation an example of **descriptive** or **inferential** statistics?

1. Descriptive and inferential statistics

Correlation (*descriptive* statistics! → current data only)

- Correlation (***r***): Measures to what extent the value of variable *y* is **predicted by** the value of variable *x*
  - If we know how long each student studied for an exam, can we predict how well they did?
- Correlation can be **positive**, **zero**, or **negative**
  - Positive ($0 < r \leq 1$): when *x* increases, *y* increases
  - Zero correlation: no relationship between *x*, *y*
  - Negative ($-1 \leq r < 0$): when *x* increases, *y* decreases
- $\boldsymbol{r^2}$ shows what **% of variation** in *y* is explained by *x*

Correlation | In a **scatterplot** (Kaplan 2016:268)



(a) Relationship between intensity of the first and second vowels in words like *rapid*, $r = .894$.

(b) Relationship between duration of the stop (*p*) and nasal (*m*) in words like *trumpet*, $r = -.565$.

(c) Relationship between intensity of the first vowel and duration of the middle consonant in words like *rapid*, $r = -.001$.

Figure A.1 Examples of strong (a), moderate (b), and no (c) correlation.

- Can you get a sense here of what it means to say that knowing *x* does/does not help us predict *y*?

2. Correlation and causation

Discussion

- You've likely heard: *"Correlation is not causation!"*
  - What does this actually mean?
  - How should we use it in interpreting experiment results?

2. Correlation and causation

- You’ve likely heard: *“Correlation is not causation!”*
  - What does this actually mean?

- What this does NOT mean:
  - Correlation is not “real”
  - Correlation tells us nothing
  - Correlation means sample size wasn’t large enough

2. Correlation and causation

- You've likely heard: *"Correlation is not causation!"*
  - What does this actually mean?

- What this DOES mean:
  - Finding that *x* and *y* are **correlated** is **not enough** for us to conclude that *x* **causes** *y*

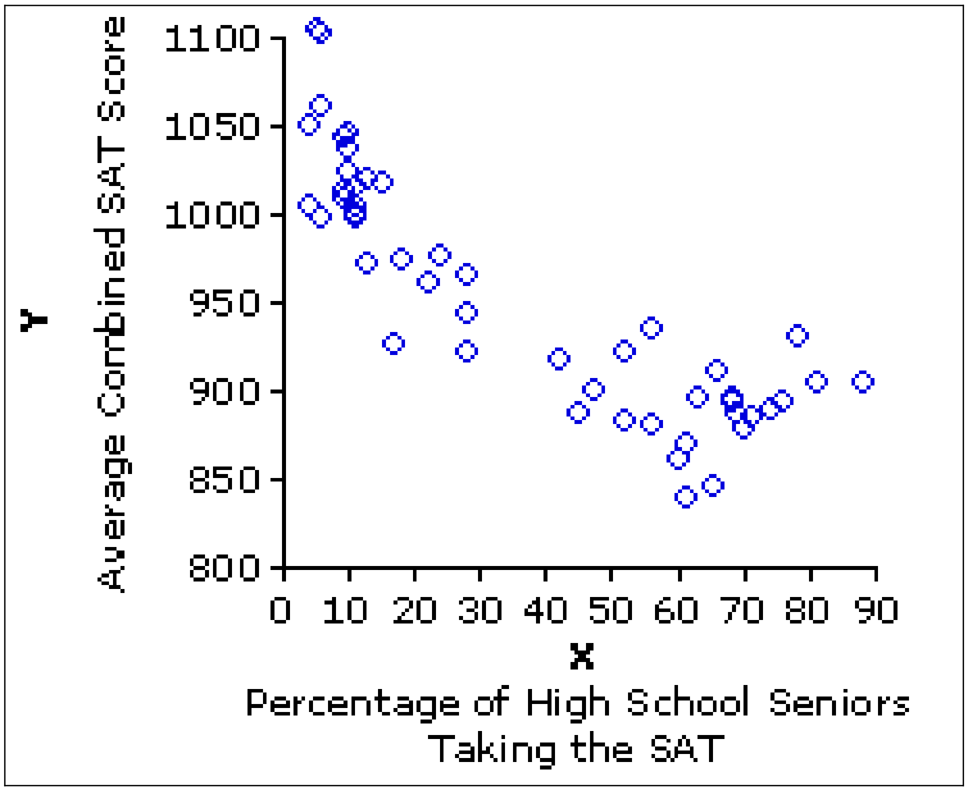
- Suppose we find that *x* and *y* are **correlated**. What are the logical possibilities for **causation**?

2. Correlation and causation

- Suppose we find that *x* and *y* are **correlated**. What are the logical possibilities for **causation**?
  - Maybe *x* causes *y*
  - Maybe *y* causes *x*
  - Maybe *z* causes both *x* and *y*:
    in this case, *z* is a **confounding factor**

2. Correlation and causation

- Example: Are *x* and *y* correlated? Can we **prove** causation? Is there a **plausible** causal relationship?



- Average combined SAT scores by state in 1993

From Richard Lowry's *Concepts & Applications of Inferential Statistics*, http://vassarstats.net/textbook/ch3pt1.html#top

2. Correlation and causation

Some take-home points:

- Although correlation does NOT imply causation, finding a correlation *can* still be **informative**
  - Might suggest a *possible* causation relationship (for further investigation)
  - Might be a way to test a theory's prediction
- We should try to minimize **confounding** factors in experiments
  - Helps increase the chances that relationships between factors in the study are *meaningful*

3. Inferential statistics and probability

- **Descriptive statistics** give us a **summary** of the information in a particular data set
  - We can describe phenomena we have observed
- But usually, we do experiments to understand **general questions**, not specific cases
  - Do the phenomena we have observed allow us to **make broader predictions** about the world?

    Some examples:

    - Will different people behave similarly?
    - Will different stimuli produce similar results?

3. Inferential statistics and probability

- **Inferential statistics** — How likely are the patterns in the data to have **arisen by coincidence**?
  - **Lower** **probability** of **coincidence** means patterns in the data are **more** **likely** to represent **facts about the world**

3. Inferential statistics and probability

- In its most basic form, an experiment **compares** two **conditions** to see if they are **different**

Group discussion

- What are the **two conditions** in these studies discussed in Kaplan (2016), sec 2.3?
  - Österberg (1961)
  - Yiakoumetti (2006)
- For each study, what is the **null hypothesis**?

3. Inferential statistics and probability

- In its most basic form, an experiment **compares** two **conditions** to see if they are **different**
- What is the **null hypothesis** for such experiments?
    - The **null hypothesis** is that there is **no actual difference** between the conditions

3. Inferential statistics and probability

- The **null hypothesis** is that there is **no actual difference** between conditions in an experiment
  - Any *apparent* difference between conditions *in our data* would therefore be due to **coincidence**
- Inferential statistics helps us ask:
  - How **likely** are the differences we observed to have **occurred** (i.e., by coincidence)...
  - ...**if the null hypothesis is correct**?
- If **very unlikely**, we **reject** the null hypothesis
  - We conclude: the differences are **meaningful**

3. Inferential statistics and probability

- We can ask: What is the **probability (*p*)** that a difference of this size would be observed if the null hypothesis is actually correct?
  - **Low** probability → **unlikely** to have arisen **by chance → statistically significant** (we **reject** the null hypothesis)

- "Low" probability — how low is low enough?

  Thresholds (α levels) often seen in research articles:

$p<0.001$	very highly significant
$p<0.01$	highly significant
$p<0.05$	significant
$p<0.1$	'marginally significant' (sometimes noted)

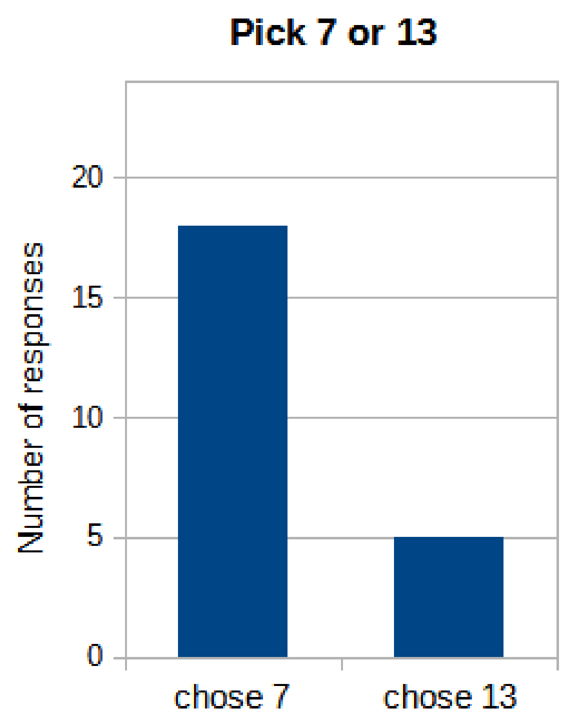
3. Inferential statistics and probability

- LING 60, Sp 2024 | Pick a number: 7 or 13?

- Results:

picked 7	*picked 13*
18	5

- Data graphic:



- Is this pattern of results **evidence** that people prefer 7, or just a **coincidence**?

3. Inferential statistics and probability

- Is this pattern of results **evidence** that people preferred 7, or just a **coincidence**?
- Probability of 18+ / 23 participants choosing 7 if everyone was equally likely to pick either: **0.0053**
  (exact binomial test)
  - Highly unlikely to be a coincidence!
- Do these results *tell* us why 7 was preferred?

3. Inferential statistics and probability

- Is this pattern of results **evidence** that people preferred 7, or just a **coincidence**?
- Probability of 18+ / 23 participants choosing 7 if everyone was equally likely to pick either: **0.0053**

  (exact binomial test)

  - Highly unlikely to be a coincidence!
- Do these results *tell* us why 7 was preferred? | No!
  - Lucky/unlucky numbers?
  - Smaller/bigger number?
  - Appeared first in the list of answers?

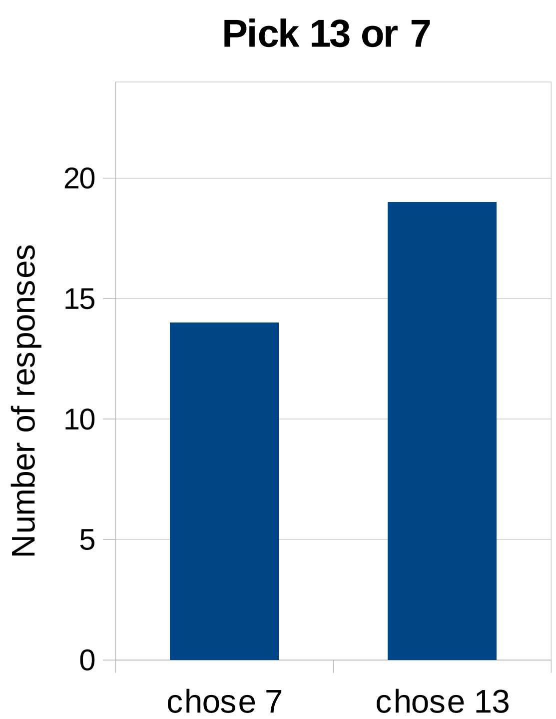
3. Inferential statistics and probability

- THIS TIME | Pick a number: 13 or 7?

- Results:

  picked 7 *picked 13*
  14 19

- Data graphic:



- Is this pattern of results **evidence** that people prefer 13, or just a **coincidence**?

3. Inferential statistics and probability

- Is this pattern of results **evidence** that people preferred 13, or just a **coincidence**?
- Probability of 19+ / 33 participants choosing 13 if everyone was equally likely to pick either: **0.24**
  (exact binomial test)
  - Highly likely to be a coincidence!

- *Combined results:*
  Probability of 32+ / 56 participants choosing 7 if everyone was equally likely to pick either: **0.175**
  (exact binomial test)

3. Inferential statistics and probability

- Trade-off: There is no magically "right" *p*-value
  - Threshold (α) too strict? Might reject results too often
  - But α of $p<0.05$ is sometimes too lax (xkcd #882)

- Recent trend in research: Focus on measures such as effect size and confidence intervals instead
  - But you will encounter *p*-values in many articles

3. Inferential statistics and probability

- What is the **probability (*p*)** that a difference of this size would be observed if null hyp. is correct?

  [More info: VassarStats Binomial Distributions, Binomial Probabilities]

Table A.2 p-*values for various outcomes of a coin-tossing experiment, testing the null hypothesis that heads and tails are equally likely.*

Tosses	Heads	*p*
10	6	.754
20	12	.503
50	30	.203
100	60	.0569
200	120	.00569
500	300	.00000894

(Kaplan 2016: 272)

- Which of these coins do you think are **unfair**?

3. Inferential statistics and probability

- Reading about experiment results: What to look for
  - What was the **null hypothesis**?
    (might be assumed rather than stated explicitly!)
  - What **statistical test** was performed?
  - Were any comparisons **statistically significant**?
  - Do the results show
    - a **main effect** (factor matters in the same way across all experiment conditions)?
    - an **interaction** (factor matters differently in different conditions)?

3. Inferential statistics and probability

- Some things to watch out for…

  Kaplan (2016: 274)

  - It's tempting to use the *p*-value of a statistical test as a binary decision-making tool: if $p < 0.05$, the result is real; otherwise, it's not.

  - If null hypothesis can't be rejected (null result): Really no difference, or sample size too small?
  - A statistically significant difference can still be too small to matter in practical terms
  - Correlation does not prove causation

Null results and experimental power

- If the sample size in an experiment is too small, it may not produce a low enough *p*-value, even if the effect is real
  - A ‘null result’ doesn’t **prove** there is **no** effect
  - But we can trust a null result more confidently if the experiment was large, or many experiments have found a null result
  - Compare the coin-toss example above...
    If we get **60% heads**, is the coin unfair?

4. Next time

- Data graphics
- Experiment design
- More generally:
  Language myths and research questions