Review of probability

Lecture 4

STA 371G

Who are these people?

Zack

Kelly
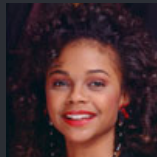

Screech

Slater

Jessie

Lisa

Some events

Suppose we pick a *Saved by the Bell* character at random. Each of these are events:

- $M =$ we select a male character

Some events

Suppose we pick a *Saved by the Bell* character at random. Each of these are events:

- M = we select a male character
- L = we select someone with lush hair, where "lush" means "long curly hair" (mullet included!)

Some events

Suppose we pick a *Saved by the Bell* character at random. Each of these are events:

- $M =$ we select a male character
- $L =$ we select someone with lush hair, where "lush" means "long curly hair" (mullet included!)
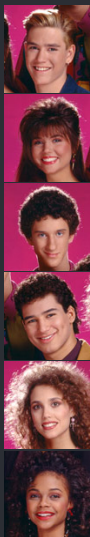- $M|L =$ we select a male *from among the lush-haired characters*

Some events

Suppose we pick a *Saved by the Bell* character at random. Each of these are events:

- M = we select a male character

- L = we select someone with lush hair, where "lush" means "long curly hair" (mullet included!)

- $M|L$ = we select a male *from among the lush-haired characters*

- $L|M$ = we select a lush-haired character *from among the male characters*

Some events

Suppose we pick a *Saved by the Bell* character at random. Each of these are events:

- M = we select a male character
- L = we select someone with lush hair, where "lush" means "long curly hair" (mullet included!)
- $M|L$ = we select a male *from among the lush-haired characters*
- $L|M$ = we select a lush-haired character *from among the male characters*
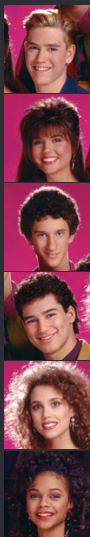
Some events

Suppose we pick a *Saved by the Bell* character at random. Each of these are events:

- M = we select a male character

- L = we select someone with lush hair, where "lush" means "long curly hair" (mullet included!)

- $M|L$ = we select a male *from among the lush-haired characters*

- $L|M$ = we select a lush-haired character *from among the male characters*
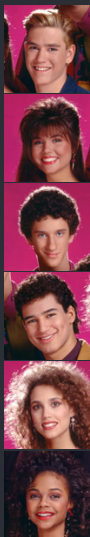
$P(M) =$

Some events

Suppose we pick a *Saved by the Bell* character at random. Each of these are events:

- M = we select a male character
- L = we select someone with lush hair, where "lush" means "long curly hair" (mullet included!)
- $M|L$ = we select a male *from among the lush-haired characters*
- $L|M$ = we select a lush-haired character *from among the male characters*
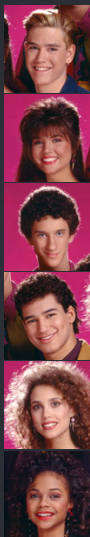
$P(M) = 3/6 \qquad P(L) =$

Some events

Suppose we pick a *Saved by the Bell* character at random. Each of these are events:

- M = we select a male character
- L = we select someone with lush hair, where "lush" means "long curly hair" (mullet included!)
- $M|L$ = we select a male *from among the lush-haired characters*
- $L|M$ = we select a lush-haired character *from among the male characters*

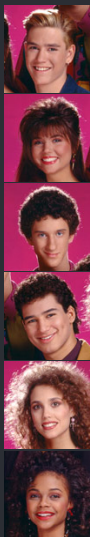$P(M) = 3/6$ $P(L) = 3/6$

$P(M|L) =$

Some events

Suppose we pick a *Saved by the Bell* character at random. Each of these are events:

- M = we select a male character
- L = we select someone with lush hair, where "lush" means "long curly hair" (mullet included!)
- $M|L$ = we select a male *from among the lush-haired characters*
- $L|M$ = we select a lush-haired character *from among the male characters*

$P(M) = 3/6 \qquad P(L) = 3/6$

$P(M|L) = 1/3 \qquad P(L|M) =$

Some events

Suppose we pick a *Saved by the Bell* character at random. Each of these are events:

- M = we select a male character
- L = we select someone with lush hair, where "lush" means "long curly hair" (mullet included!)
- $M|L$ = we select a male *from among the lush-haired characters*
- $L|M$ = we select a lush-haired character *from among the male characters*

$P(M) = 3/6$ $P(L) = 3/6$
$P(M|L) = 1/3$ $P(L|M) = 1/3$

Joint probability

$P(M \text{ and } L) =$ probability of selecting a lush-haired male

Joint probability

$P(M$ and $L) =$ probability of selecting a lush-haired male

$\qquad = 1/6$

Joint probability

$P(M \text{ and } L)$ = probability of selecting a lush-haired male

$= 1/6$

$= P(M)P(L|M) = \dfrac{3}{6} \cdot \dfrac{1}{3}$

$= P(L)P(M|L) = \dfrac{3}{6} \cdot \dfrac{1}{3}$

Joint probability

$P(M \text{ and } L) =$ probability of selecting a lush-haired male

$= 1/6$

$= P(M)P(L|M) = \dfrac{3}{6} \cdot \dfrac{1}{3}$

$= P(L)P(M|L) = \dfrac{3}{6} \cdot \dfrac{1}{3}$

$\neq P(M)P(L)$

One or the other

$P(M \text{ or } L) =$ probability of selecting a lush-haired character, a male character, or both

One or the other

$P(M \text{ or } L) =$ probability of selecting a lush-haired character, a male character, or both

$= 5/6$

We have to subtract off $P(M \text{ and } L)$ because otherwise we are double-counting Slater!

One or the other

$P(M \text{ or } L) =$ probability of selecting a lush-haired
character, a male character, or both
$= 5/6$
$= P(M) + P(L) - P(M \text{ and } L)$
$= 3/6 + 3/6 - 1/6$

We have to subtract off $P(M \text{ and } L)$ because otherwise
we are double-counting Slater!

Conditional probability

When we say $P(A|B)$, what we mean is: "In a world where we know

B has already happened, how likely is it that A also happened?"

Now that you've found a strange pair of underwear in your dresser, what is the probability that your partner is cheating on you?

$P(\text{partner is cheating} \mid \text{underwear found})$

Now that you've found a strange pair of underwear in your dresser, what is the probability that your partner is cheating on you?

$P($partner is cheating | underwear found$)$

It's hard to know how to estimate this directly!

Let's come up with estimates for the following:

- $P($underwear found | partner is cheating$)$

Let's come up with estimates for the following:

- P(underwear found | partner is cheating)
- P(underwear found | partner is not cheating)

Let's come up with estimates for the following:

- $P($underwear found | partner is cheating$)$
- $P($underwear found | partner is not cheating$)$
- $P($partner is cheating$)$ (before we found the underwear!)

Bayes' Rule

Bayes' Rule lets us reverse the conditional probabilities!

$U =$ underwear found, $C =$ partner is cheating

$$P(C|U) = \frac{P(U|C)P(C)}{P(U|C)P(C) + P(U|\overline{C})P(\overline{C})}$$

Think of Bayes' Rule as a way to update our thinking based on new information:

$P(C)$ ⟵ Prior probability

$P(C|U)$ ⟵ Posterior probability (includes new information)

HIV Testing

- HIV testing is important for public health, but HIV tests are not perfect

HIV Testing

- HIV testing is important for public health, but HIV tests are not perfect
- The OraQuick ADVANCE Rapid HIV-1/2 Antibody Test has the following properties:

HIV Testing

- HIV testing is important for public health, but HIV tests are not perfect
- The OraQuick ADVANCE Rapid HIV-1/2 Antibody Test has the following properties:
    - If you **have HIV**, there is a 99.3% chance the test will show a **positive result**

HIV Testing

- HIV testing is important for public health, but HIV tests are not perfect
- The OraQuick ADVANCE Rapid HIV-1/2 Antibody Test has the following properties:
  - If you **have HIV**, there is a 99.3% chance the test will show a **positive result**
  - If you **do not have HIV**, there is a 99.8% chance the test will show a **negative result**

HIV Testing

- HIV testing is important for public health, but HIV tests are not perfect
- The OraQuick ADVANCE Rapid HIV-1/2 Antibody Test has the following properties:
  - If you **have HIV**, there is a 99.3% chance the test will show a **positive result**
  - If you **do not have HIV**, there is a 99.8% chance the test will show a **negative result**
- 0.4% of people in the US are HIV-positive

HIV Testing

We know $P(TP|HP) = 0.993$, but we really want to know $P(HP|TP)$!

HIV Testing

- If you have HIV, there is a 99.3% chance the test will show a positive result

HIV Testing

- If you have HIV, there is a 99.3% chance the test will show a positive result

- If you do not have HIV, there is a 99.8% chance the test will show a negative result

HIV Testing

- If you have HIV, there is a 99.3% chance the test will show a positive result

- If you do not have HIV, there is a 99.8% chance the test will show a negative result

- But if you test positive there is about a 1/3 chance you do NOT have HIV!

HIV Testing

- If you have HIV, there is a 99.3% chance the test will show a positive result
- If you do not have HIV, there is a 99.8% chance the test will show a negative result
- But if you test positive there is about a 1/3 chance you do NOT have HIV!
- This is counterintuitive — it's because of the way we are wired (it even has a name: "base rate fallacy")