MATH 180A: INTRO TO PROBABILITY (FOR DATA SCIENCE)

www.math.ucsd.edu/~tkemp/180A

Today: § 2.2 – 2.3

Next: § 1.5, 3.1

Lab 1 due TONIGHT!

HW2 due Friday, 10/11.

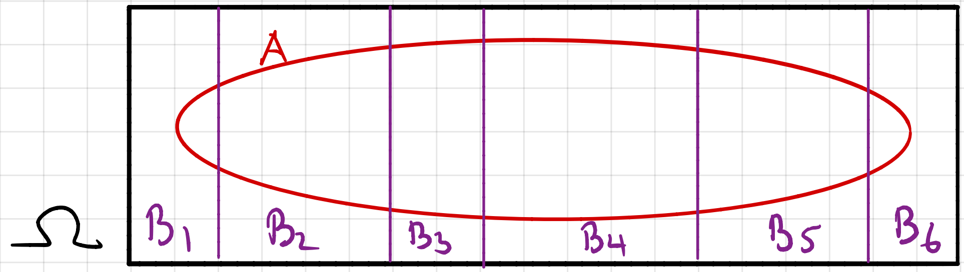
Screencast & video available after each lecture @ podcast.ucsd.edu

Before/After slides now available on course webpage.

Lots of active discussion on Piazza.

Law of Total Probability

If $B_1, B_2, \ldots, B_n$ partition Ω (disjoint, $B_1 \cup \cdots \cup B_n = \Omega$, $\mathbb{P}(B_j) > 0$)



then for any event A:

$$\mathbb{P}(A) = \mathbb{P}(AB_1 \cup AB_2 \cup \cdots \cup AB_n) = \sum_{j=1}^{n} \mathbb{P}(B_j)\,\mathbb{P}(A \mid B_j).$$

Eg. 90% of coins are fair, 9% are biased to come up heads 60%.
1% are biased to come up heads 80%.

You find a coin on the street. How likely is it to come up heads?

$B_1 = \{\text{fair coins}\}$ $\mathbb{P}(B_1) = 90\%$ $\mathbb{P}(A \mid B_1) = 50\%$

$B_2 = \{60\% \text{ heads}\}$ $\mathbb{P}(B_2) = 9\%$ $\mathbb{P}(A \mid B_2) = 60\%$

$B_3 = \{80\% \text{ heads}\}$ $\mathbb{P}(B_3) = 1\%$ $\mathbb{P}(A \mid B_3) = 80\%$

$A = \{\text{heads}\}$

$\mathbb{P}(A) = 51.2\%$

2.2

Question:

90% of coins are fair. 9% are biased to come up heads 60%.

1% are biased to come up heads 80%.

You find a coin on the street. You toss it, and it comes up heads.

How likely is it that this coin is heavily biased?

Eg. According to Forbes Magazine, as of April 10, 2019, there are 2208 billionaires in the world.

1964 of them are men.

Bayes' Rule

(A relationship between $\mathbb{P}(A|B)$ and $\mathbb{P}(B|A)$)

Let $B_1, B_2, \ldots, B_n$ partition the sample space. Then for any event A with $\mathbb{P}(A) > 0$,

$$\mathbb{P}(B_k | A) = \mathbb{P}(B_k A) / \mathbb{P}(A)$$

$$= \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_k)\mathbb{P}(B_k)}{\sum_{j=1}^{n} \mathbb{P}(A|B_j)\mathbb{P}(B_j)}$$

Eg. (Coins)

Epidemiological Confusion

An HIV test is 99% accurate. (1% false positives, 1% false negatives)
0.33% of US residents have HIV.

If you test positive, what is the probability you have HIV?

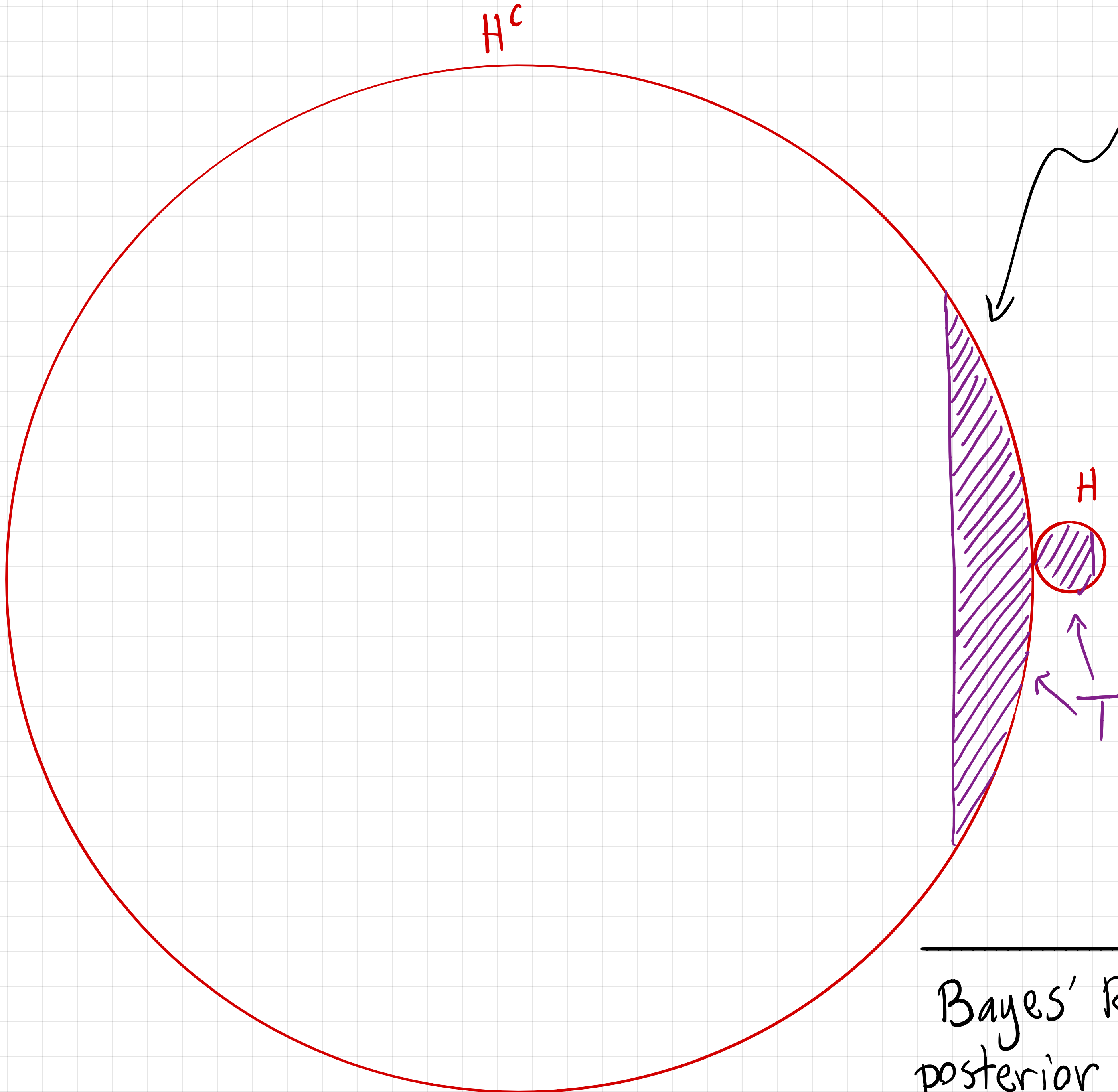
(a) 99%

(b) 1%

(c) 25%

(d) 0.33%

(e) There is not enough information to answer.

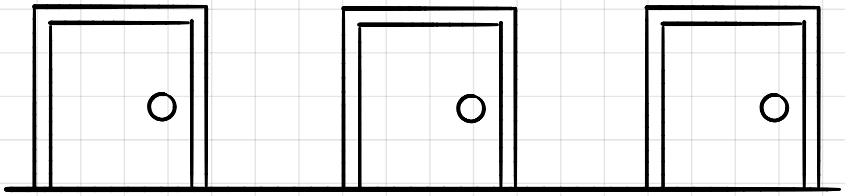


even though this part of T is only 1% of H^c, it is **3 times as big** as the part of T in H (which takes up 99% of H).

This is possible because H^c **dwarfs** H in this example.

Bayes' Rule shows that posterior probabilities are **highly sensitive** to prior inputs.

The Monty Hall Problem



At the climax of a gameshow, you are shown 3 doors. The host tells you that, behind one of them is a valuable prize (a car), while the other two hide nothing of value (goats).

You choose one. The host then opens one of the two doors you did not choose, revealing a goat. He then asks you if you want to stick with your original choice, or switch to the other closed door.

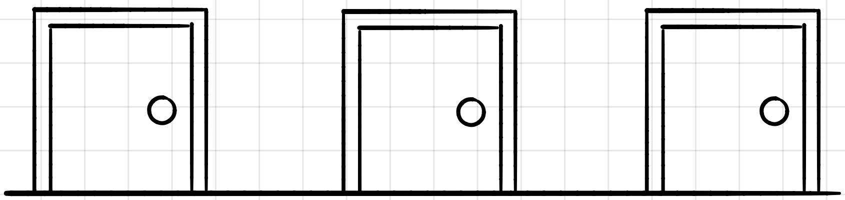
Should you switch??

(a) Yes.

(b) No.

(c) Doesn't matter.

The Monty Hall Problem



Let's decide to call the door you chose originlly #1.

$\therefore$ Monty will open #2 or #3. We'll focus our analysis on #2.

$B_i = \{$ the car is behind door #i $\}$.

$A = \{$ Monty opens door #2 $\}$

We want to know $\mathbb{P}(B_3 \mid A)$.

$\mathbb{P}(B_1) = \mathbb{P}(B_2) = \mathbb{P}(B_3) = \frac{1}{3}$

$\mathbb{P}(A \mid B_2) =$

$\mathbb{P}(A \mid B_3) =$

$\mathbb{P}(A \mid B_1) =$

2.3

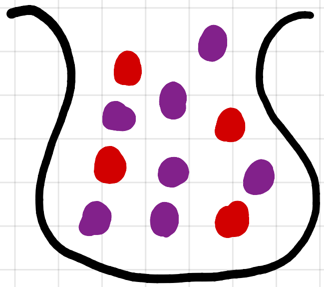
Suppose two events A and B really have nothing to do with each other. That doesn't mean they're disjoint; it means they have no influence on each other.

Eg. Flip a coin 3 times. $A = \{$the first toss is heads$\}$
$B = \{$the second toss is tails$\}$.

$A = \{HHH, HHT, HTH, HTT\}$
$B = \{HTH, HTT, TTH, TTT\}$

Def: Two events A, B are (statistically) **independent** if

Eg.



An urn has 4 red and 7 blue balls.
Two balls are sampled

$A = \{1^{st} \text{ ball is red}\}$

$B = \{2^{nd} \text{ ball is blue}\}$

Are A and B independent?

(a) Yes.

(b) No.

(c) Can't tell from the question.

How about 3 events?

Eg. A coin is tossed 3 times.

$A = \{$there is exactly 1 tails in the first two$\}$

$B = \{$there is exactly 1 tail in the second two$\}$

$C = \{$there is exactly 1 tail in the first & third$\}$

$A = \{TH*, HT*\}$ $B = \{*TH, *HT\}$ $C = \{T*H, H*T\}$

$AB = \{THT, HTH\}$ $\rightarrow$ but $ABC = \emptyset$.

$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) =$

$\mathbb{P}(AB) =$

A, B

B, C

A, C

Def: A collection $A_1, A_2, \ldots, A_n$ of events are **independent** if: for every subcollection $A_{i_1}, A_{i_2}, \ldots, A_{i_m}$ $(1 \le i_1 < i_2 < \cdots < i_m \le n)$

$$\mathbb{P}(A_{i_1} A_{i_2} \cdots A_{i_m}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2})\cdots\mathbb{P}(A_{i_m}).$$

Eg. When $n=3$, this means we must have

$$\mathbb{P}(A_1A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$$
$$\mathbb{P}(A_1A_3) = \mathbb{P}(A_1)\mathbb{P}(A_3)$$
$$\mathbb{P}(A_2A_3) = \mathbb{P}(A_2)\mathbb{P}(A_3)$$

Eg. A fair coin is tossed n times.