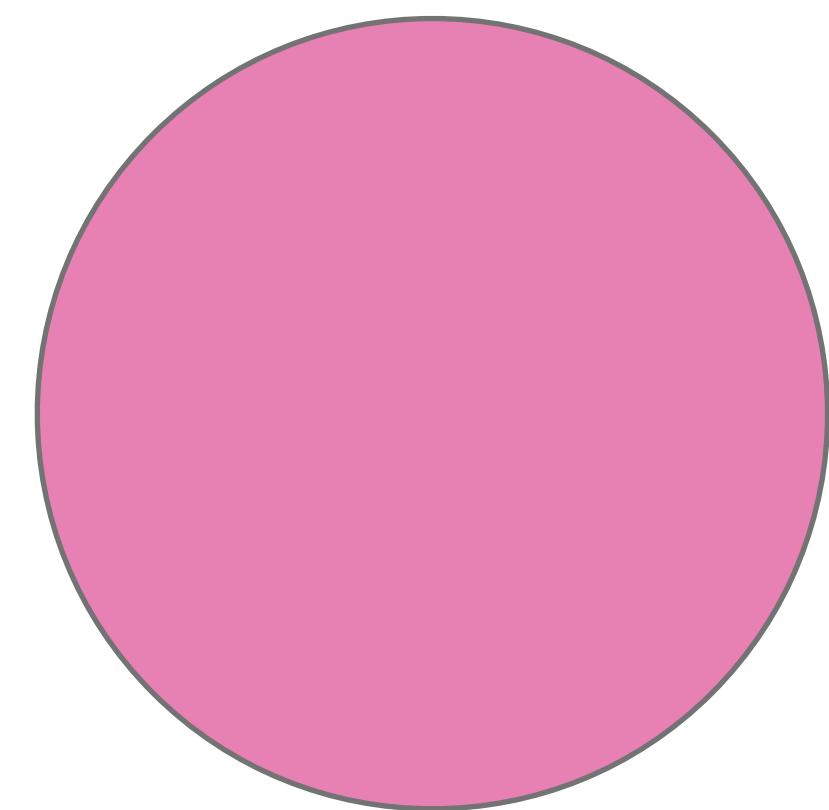
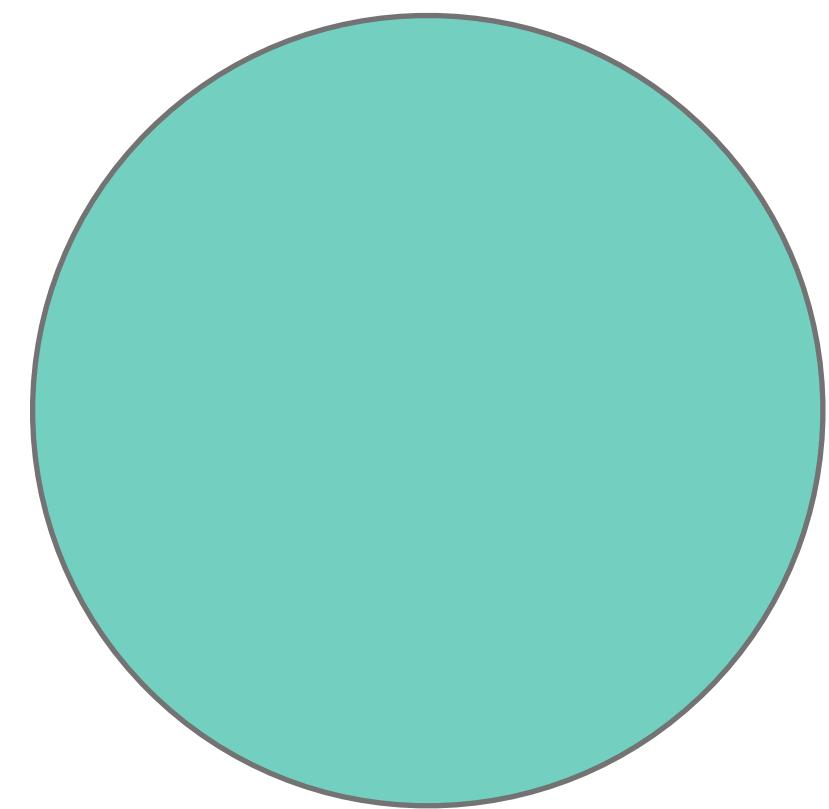
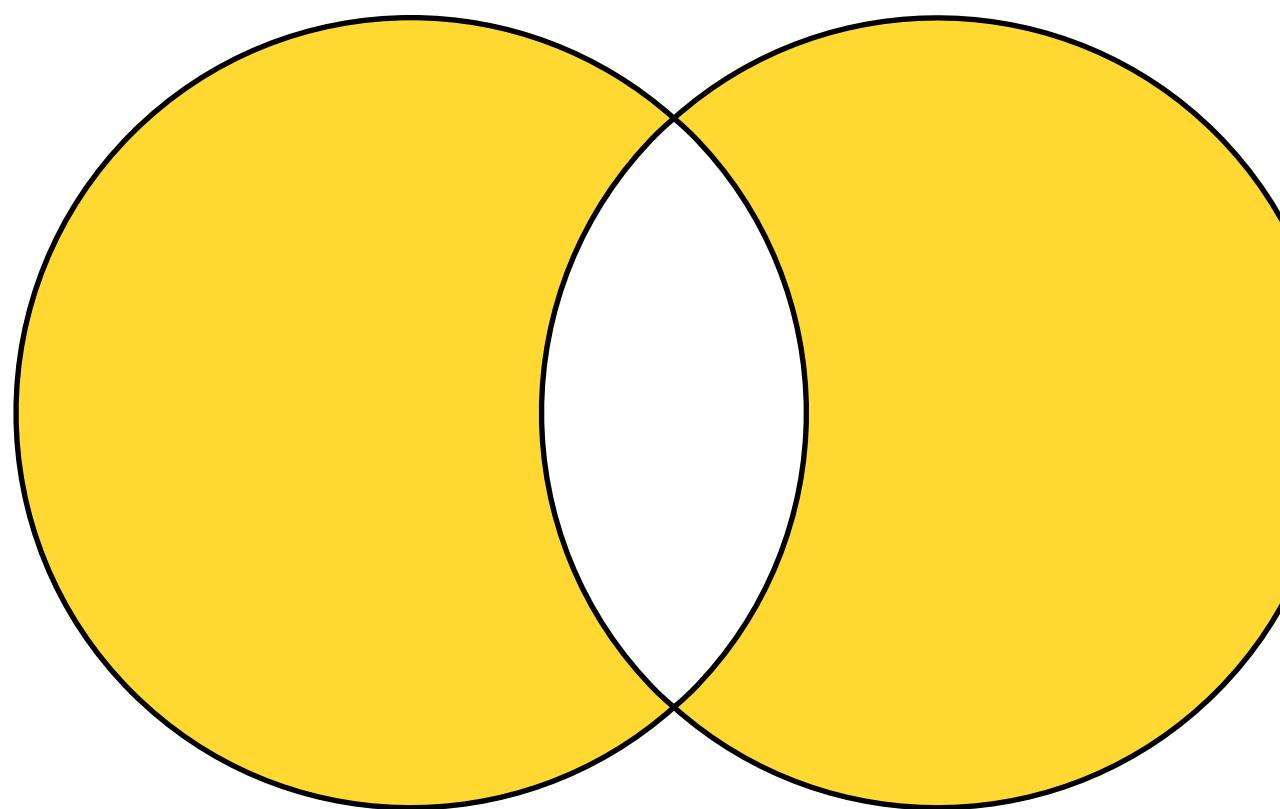
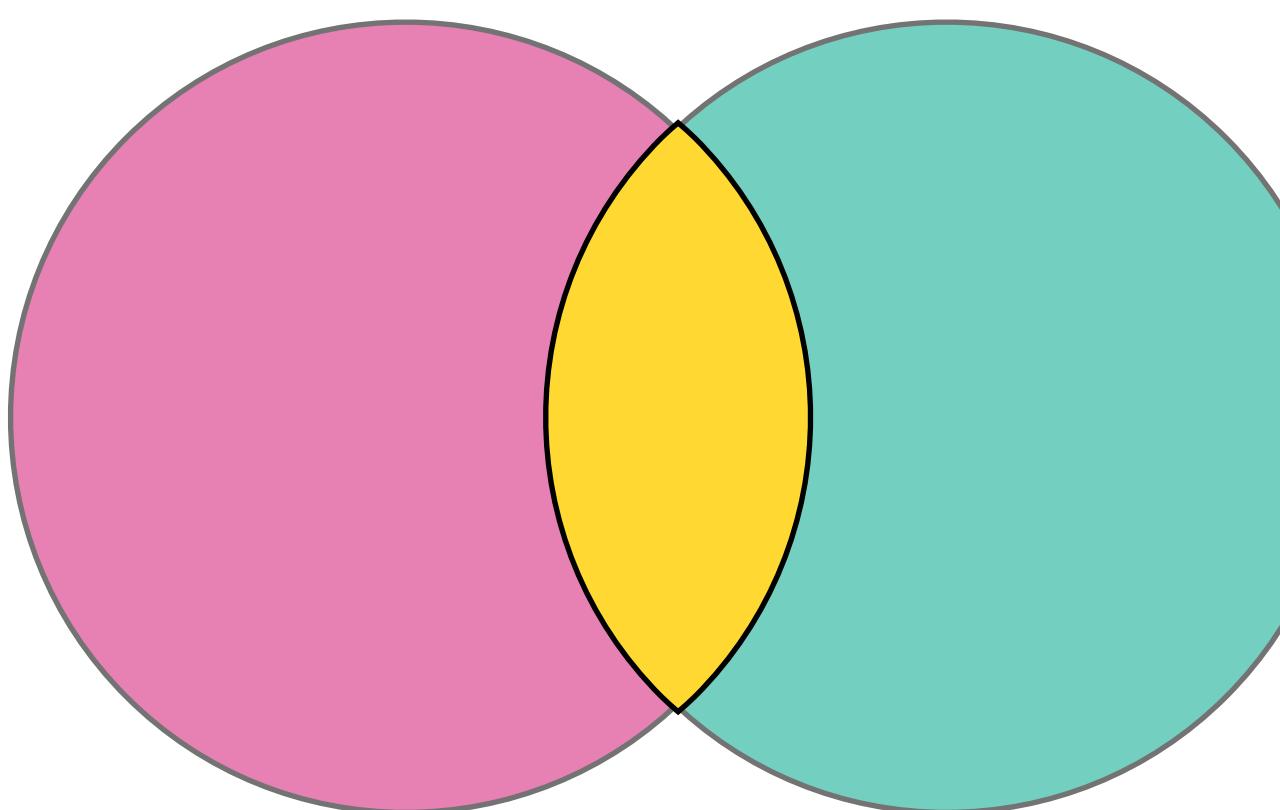


Probability

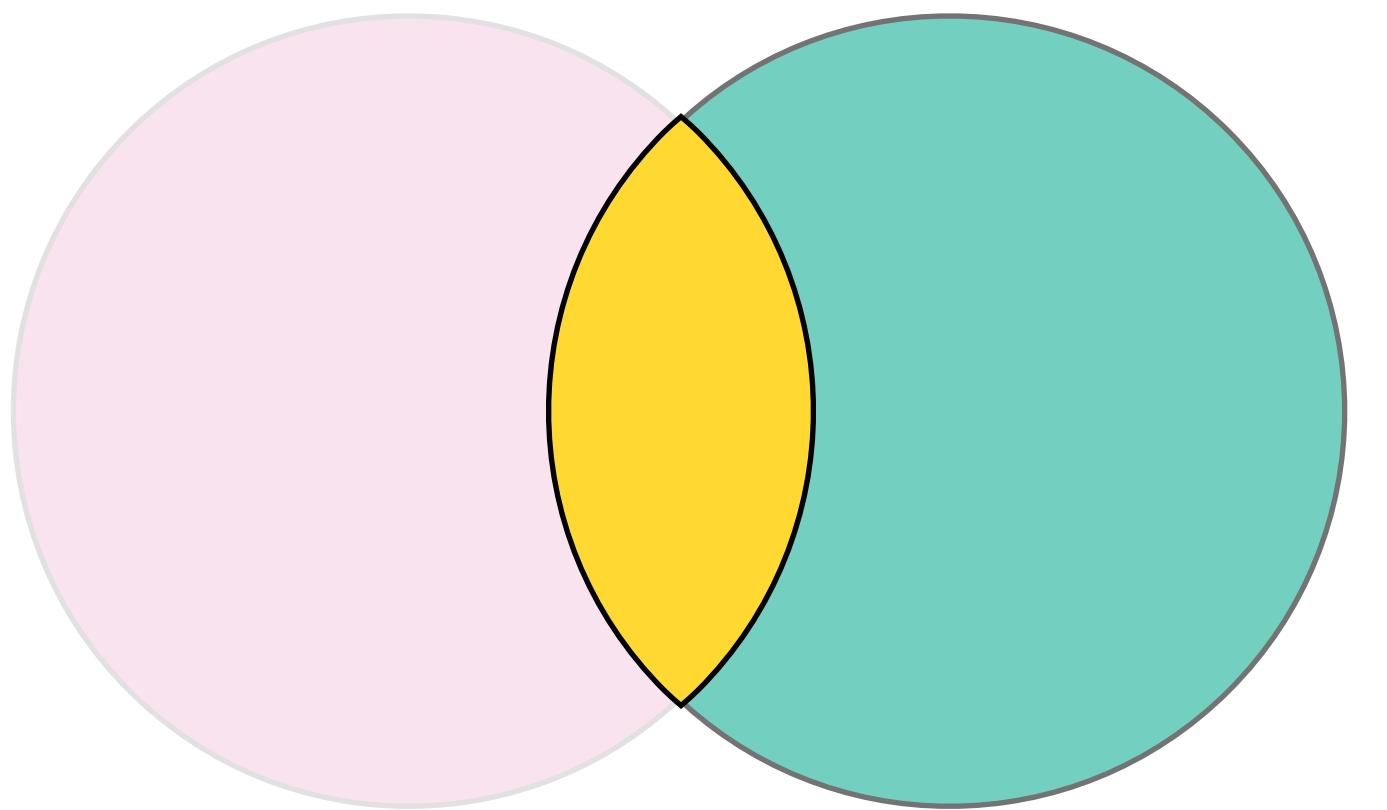
Pre-lecture 2 video

$P(A)$  $P(B)$ 

$$P(A, B) = P(A \cap B)$$

 $P(A \cup B)$

$$P(A \cup B) = P(A) + P(B) - P(A, B)$$



$$P(A | B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A | B) + P(\neg A | B) = 1$$

$$P(A \cap B) = P(A | B)P(B)$$

$$P(A, B) = P(A | B)P(B)$$

We talk about this as the probability of A *given* B;

It is *conditional!*

Note the consequence of this viewpoint

Can be rearranged

$$\sum_{x,y} P(X = x, Y = y) = 1$$

		disease X	
		0	1
Y	0	0.5	0.1
	1	0.1	0.3

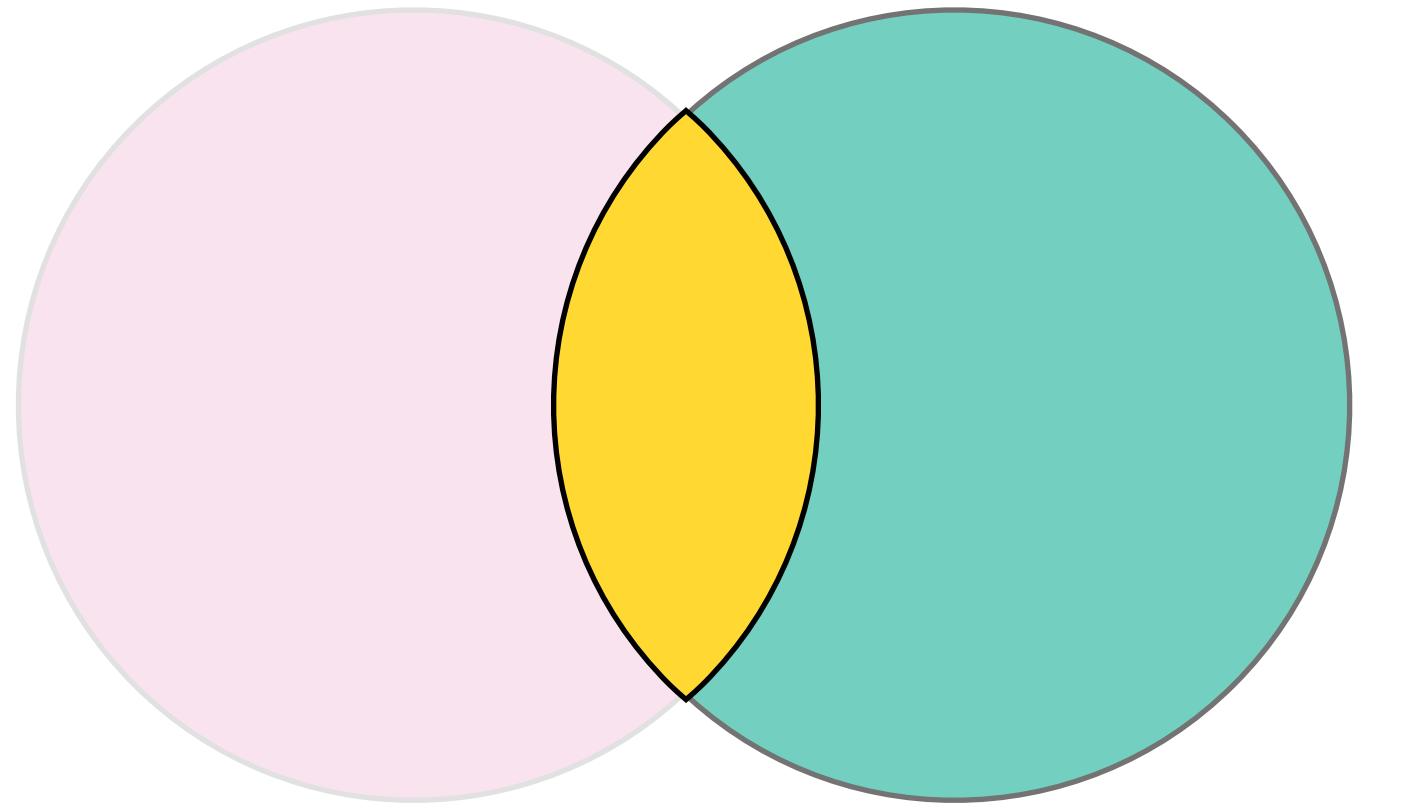
joint probability : $P(X = 0, Y = 1) = 0.1$

$$P(Y = 1) = 0.1 + 0.3 = 0.4$$

$$P(X = 0) = 0.1 + 0.5 = 0.6$$

$$P(X = x) = \sum_y P(X = x, Y = y)$$

This is called the marginal probability,
it is NOT conditioned on anything



$$P(A, B) = P(A | B)P(B)$$

$$P(A, B) = P(A)P(B)$$

Most general case!

But this is true if and only if A and B are *statistically independent*

$$P(A | B) = P(A)$$

$$P(B | A) = P(B)$$

The conditional probabilities are equivalent to the marginal probabilities

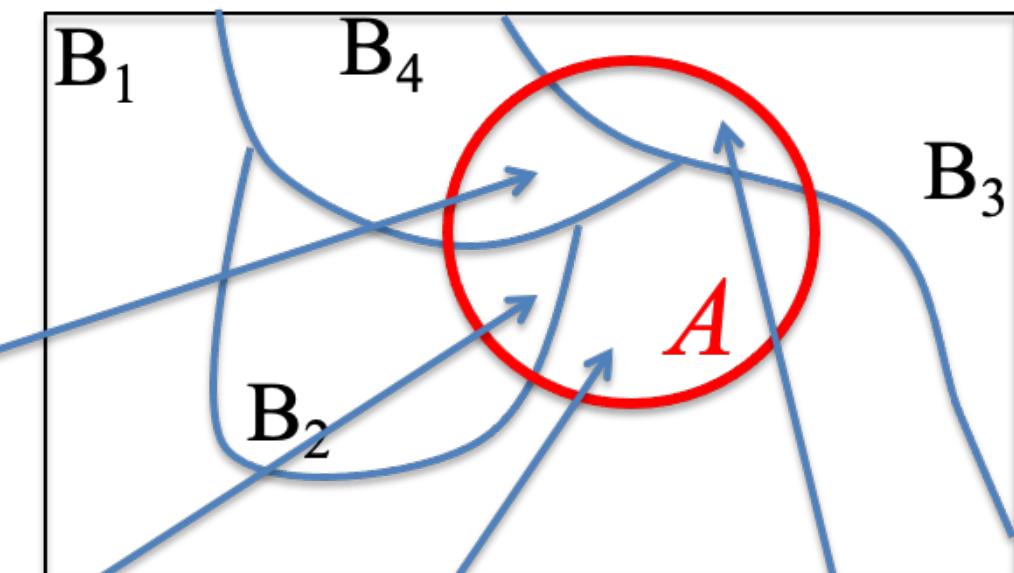
Suppose a sample space can be partitioned into disjoint sets B_i such that

$$B_1 \cup B_2 \cup \dots \cup B_m = \Omega$$

$$B_4 \cap A$$

$$B_2 \cap A$$

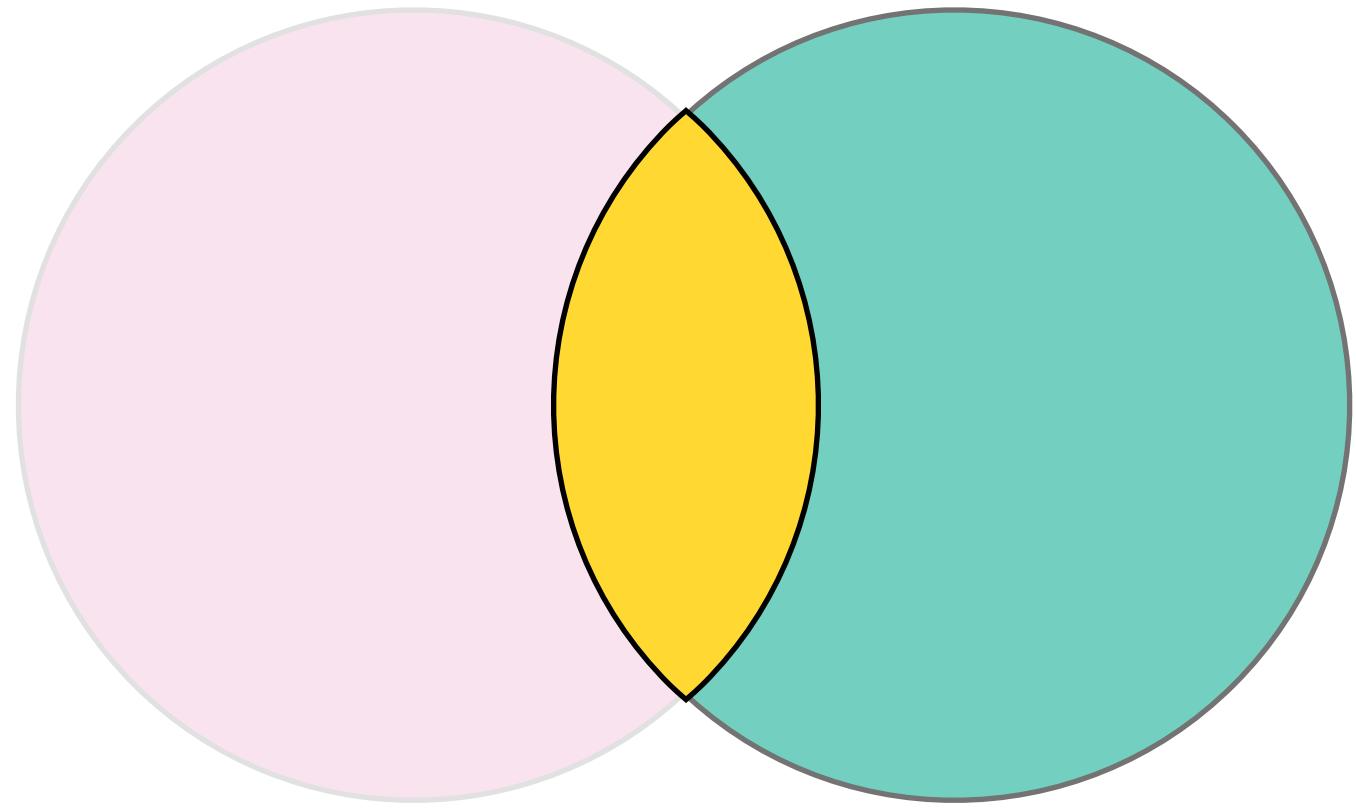
$$B_1 \cap A$$



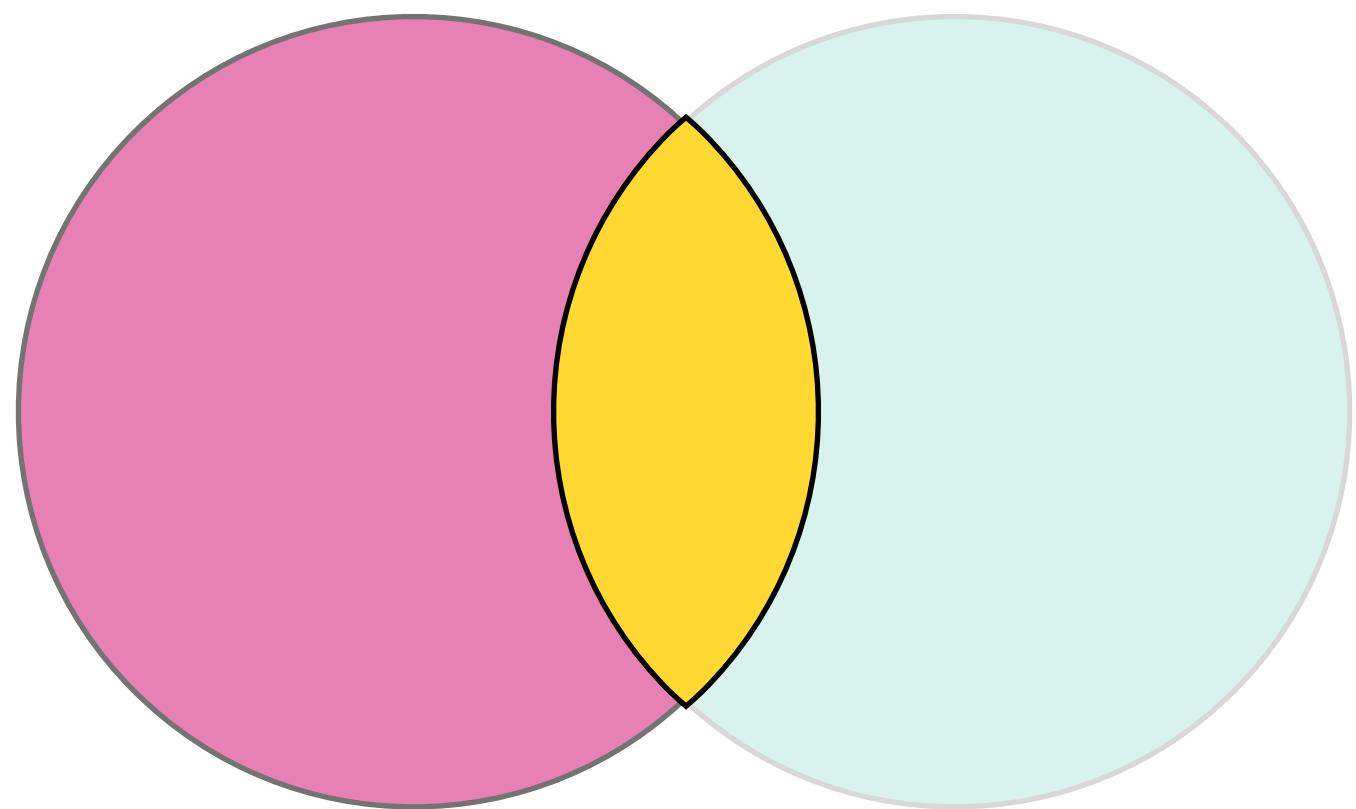
$$\begin{aligned} \Pr(A) &= \Pr(B_1 \cap A) + \Pr(B_2 \cap A) + \Pr(B_3 \cap A) + \dots + \Pr(B_m \cap A) \\ &= \Pr(A|B_1)\Pr(B_1) + \Pr(A|B_2)\Pr(B_2) + \Pr(A|B_3)\Pr(B_3) + \dots + \Pr(A|B_m)\Pr(B_m) \end{aligned}$$

Then the probability of an event A can be written by adding up the intersection of A with all the disjoint sets B_i

$$\Pr(A) = \sum_{i=1}^m \Pr(A|B_i)\Pr(B_i) \text{ The law of total probability}$$



$$P(A, B) = P(A | B)P(B)$$



Intersections are
commutative

$$P(B, A) = P(B | A)P(A)$$

So...

$$P(A | B)P(B) = P(B | A)P(A)$$

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

Bayes theorem

$$P(B | A) = \frac{P(A | B)P(B)}{P(A)}$$

Can be written more generally using the law of total probability

$$P(B_i | A) = \frac{P(A | B_i)P(B_i)}{\sum_{j=1}^m P(A | B_j)P(B_j)}$$

Suppose a sample space can be partitioned into disjoint sets B_i such that

$$B_1 \cup B_2 \cup \dots \cup B_m = \Omega$$

Probably, I'm Bayesian

COGS 118B Winter 2024

Jason G. Fleischer, PhD
Department of Cognitive Science
University of California San Diego

<https://jgfleischer.com>
[Book a slot in my office hours](#)

Playing with dice

$$P(X=1, Y=1) =$$

X:  #1

$$P(X=1) =$$

$$P(Y=4) =$$

$$P(X=3, Y>3) =$$

$$P(X=1 \mid Y=1) =$$

$$P(X=3 \mid Y>3) =$$

$$P(X=1 \mid X+Y \text{ is even}) =$$

	1	2	3	4	5	6
1						
2						
3						
4						
5						
6						

Y:  #2

Statistically independent pizza!



$$P(\text{pepperoni}) =$$

$$P(\text{mushroom}) =$$

$$P(\text{pep, mush}) =$$

$$P(\text{pep} \mid \text{mush}) =$$

$$P(\text{mush} \mid \text{pep}) =$$

$$P(A, B) = P(A)P(B)$$

$$P(A \mid B) = P(A)$$

$$P(B \mid A) = P(B)$$

iff A and B are
statistically independent



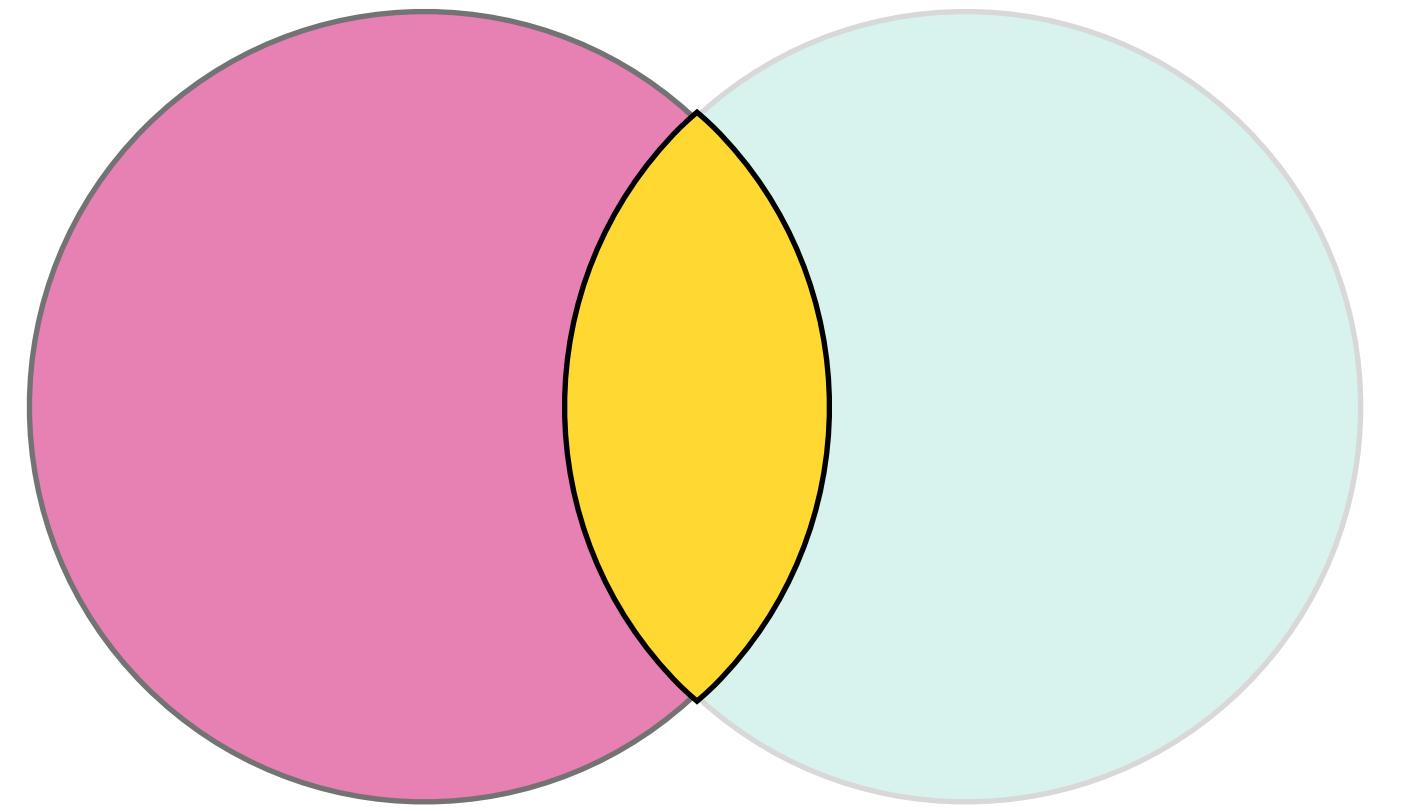
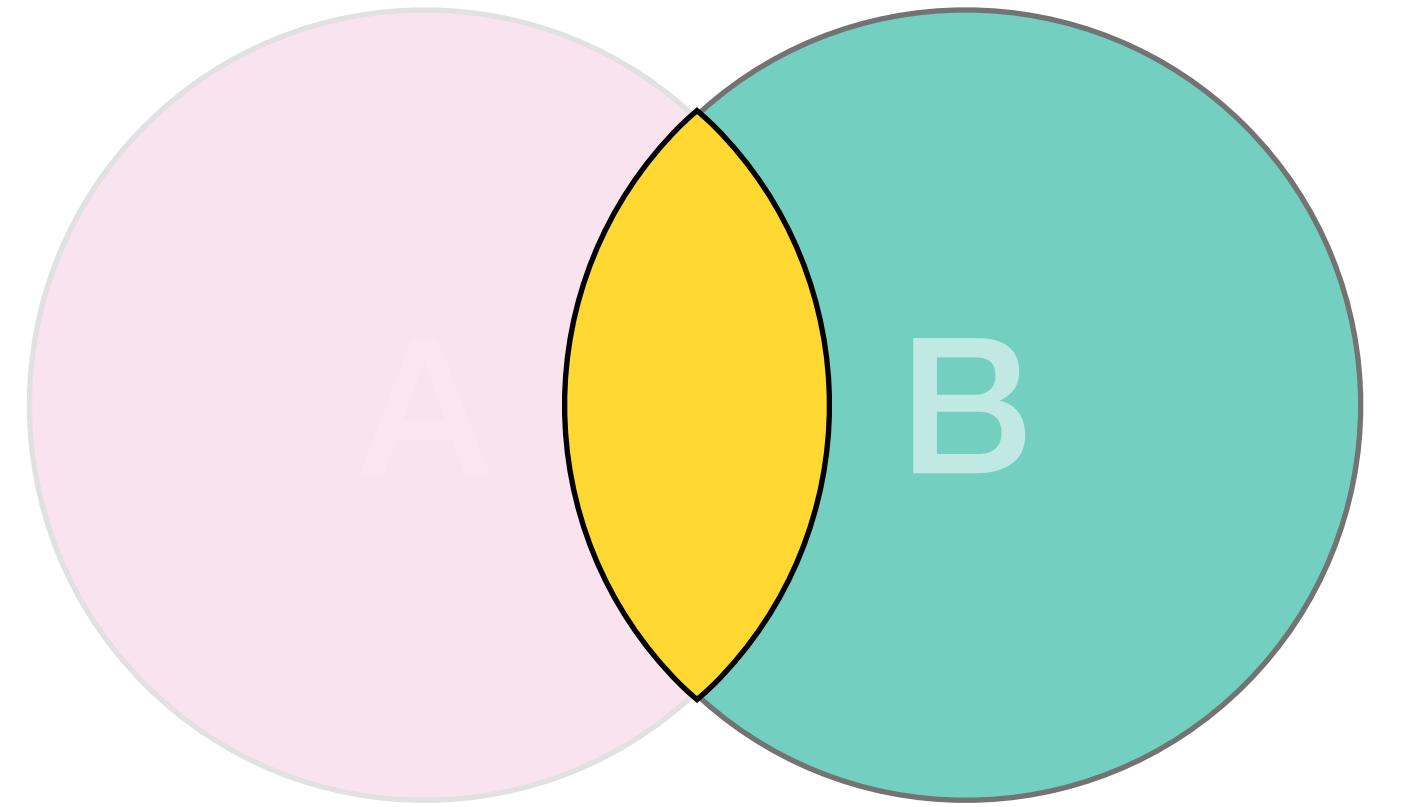
$P(\text{pepperoni}) =$

$P(\text{mushroom}) =$

$P(\text{pineapple}) =$

$P(\text{pine,mush} \mid \text{pep}) =$

$P(\text{pep} \mid \text{pine,mush}) =$



$$P(A | B) = \frac{P(A, B)}{P(B)}$$

$$P(A, B) = P(A | B)P(B)$$

$$P(B, A) = P(B | A)P(A)$$

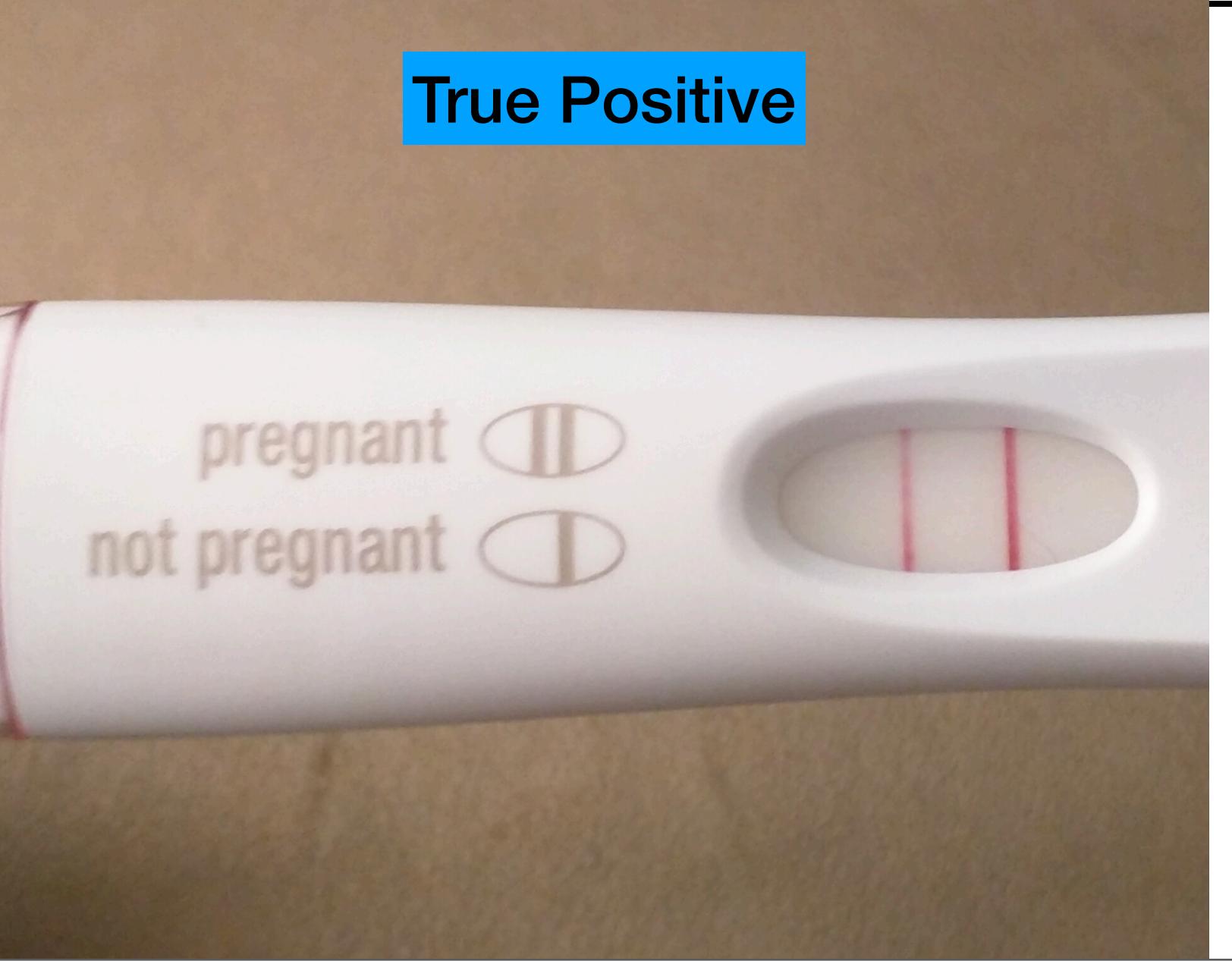
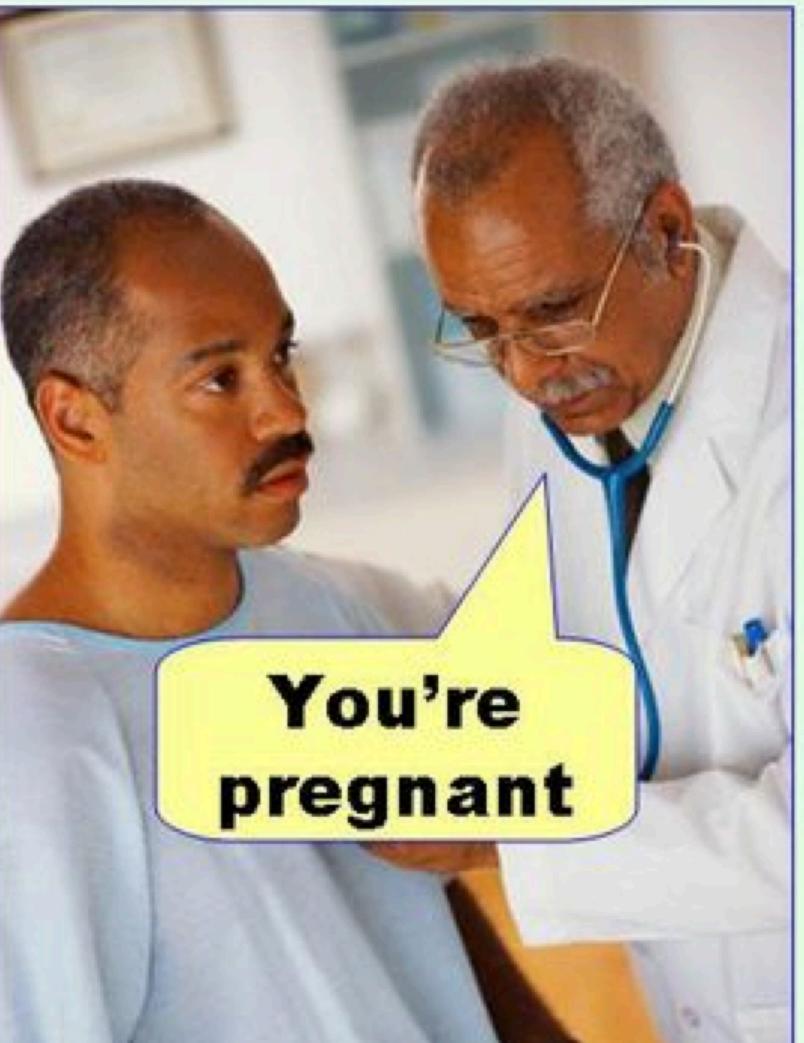
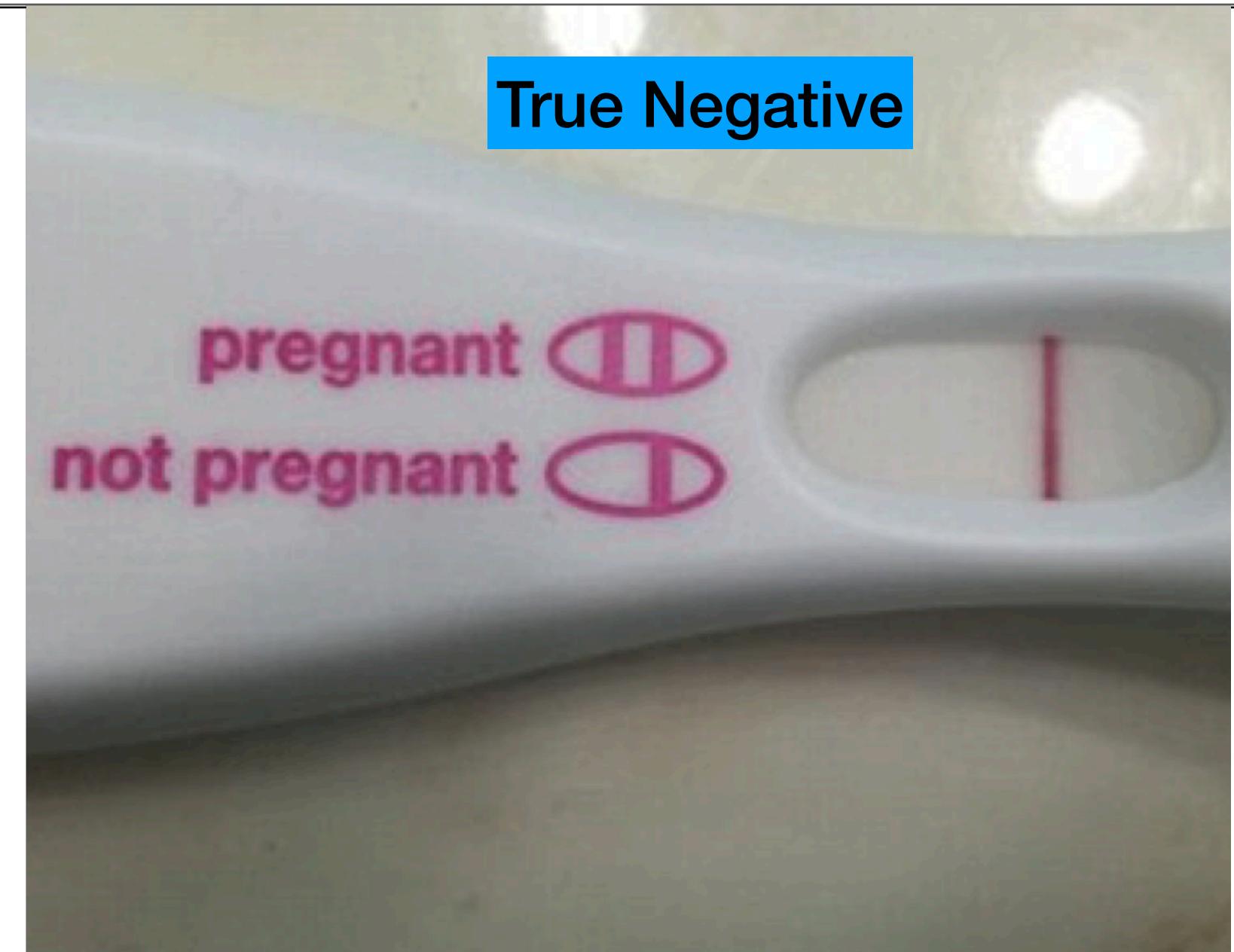
So...

$$P(A | B)P(B) = P(B | A)P(A)$$



$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

Baye's theorem

	Predicted +	Predicted -
Truly +	<p>True Positive</p>  <p>A close-up photograph of a white digital pregnancy test. The screen displays the text "pregnant" above two vertical red lines, and "not pregnant" below two empty oval icons. A blue rectangular box at the top left of the image contains the text "True Positive".</p>	<p>Type II error (false negative)</p>  <p>A photograph of a female doctor in a white coat and stethoscope around her neck, examining the belly of a pregnant woman. A yellow speech bubble originates from the doctor and contains the text "You're not pregnant". The background is a green vertical bar.</p>
Truly -	<p>Type I error (false positive)</p>  <p>A photograph of a male patient in a light blue shirt looking up at a male doctor wearing a white coat and glasses. A yellow speech bubble originates from the doctor and contains the text "You're pregnant". The background is a green vertical bar.</p>	<p>True Negative</p>  <p>A close-up photograph of a white digital pregnancy test. The screen displays the text "not pregnant" above one vertical red line, and "pregnant" below two empty oval icons. A blue rectangular box at the top right of the image contains the text "True Negative".</p>

There is a 0.1% (0.001) chance that a random person on the street has cancer

A test to find cancer will give a false positive 2% (0.02) of the time.

The same test will give a false negative 10% (0.1) of the time.

Person is	Test +	Test -	Totals
Cancerous			
Healthy			
Totals			10,000

There is a 0.1% (0.001) chance that a random person on the street has cancer

A test to find cancer will give a false positive 2% (0.02) of the time.

The same test will give a false negative 10% (0.1) of the time.

Person is	Test +	Test -	Totals
Cancerous			10
Healthy			9990
Totals			10,000

There is a 0.1% (0.001) chance that a random person on the street has cancer

A test to find cancer will give a false positive 2% (0.02) of the time.

The same test will give a false negative 10% (0.0003) of the time.

Person is	Test +	Test -	Totals
Cancerous	9	1	10
Healthy	200	9789	9990
Totals			10,000

There is a 0.1% (0.001) chance that a random person on the street has cancer

A test to find cancer will give a false positive 2% (0.02) of the time.

The same test will give a false negative 10% (0.0003) of the time.

Person is	Test +	Test -	Totals
Cancerous	9	1	10
Healthy	200	9789	9990
Totals	209	9790	10,000

$$P(\text{cancer} | \text{test} +) = \frac{P(\text{test} + | \text{cancer})P(\text{cancer})}{P(\text{test} +)}$$

$$P(\text{cancer}) = 0.001$$

Person is	Test +	Test -	Totals
Cancerous	9	1	10
Healthy	200	9789	9990
Totals	209	9790	10,000

$$P(\text{cancer} | \text{test} +) = \frac{P(\text{test} + | \text{cancer})P(\text{cancer})}{P(\text{test} +)}$$

$$P(\text{cancer}) = 0.001$$

$$P(\text{test} + | \text{cancer}) = 9/10 = 0.9$$

Person is	Test +	Test -	Totals
Cancerous	9	1	10
Healthy	200	9789	9990
Totals	209	9790	10,000

$$P(\text{cancer} | \text{test} +) = \frac{P(\text{test} + | \text{cancer})P(\text{cancer})}{P(\text{test} +)}$$

$$P(\text{cancer}) = 0.001$$

$$P(\text{test} + | \text{cancer}) = 9/10 = 0.9$$

$$P(\text{test} +) = 209/10000 = 0.020$$

Person is	Test +	Test -	Totals
Cancerous	9	1	10
Healthy	200	9789	9990
Totals	209	9790	10,000

$$P(\text{cancer} | \text{test} +) = \frac{P(\text{test} + | \text{cancer})P(\text{cancer})}{P(\text{test} +)}$$

$$P(\text{cancer}) = 0.001$$

$$P(\text{test} + | \text{cancer}) = 9/10 = 0.9$$

$$P(\text{test} +) = 209/10000 = 0.020$$

$$P(\text{cancer} | \text{test} +) = \frac{0.9 * 0.001}{0.020}$$

$$= 0.045$$

Person is	Test +	Test -	Totals
Cancerous	9	1	10
Healthy	200	9789	9990
Totals	209	9790	10,000

$$P(H_i | D) = \frac{P(D | H_i)P(H_i)}{\sum_{j=1}^m P(D | H_j)P(H_j)}$$

A very useful form of Bayes rule

“Initial belief plus new evidence equals a new improved belief”

In particular, we are accumulating evidence about multiple hypotheses and comparing them to each other

$$P(H_i)$$

Prior probability of hypothesis H_i

$$P(D | H_i)$$

Probability of observing data D under H_i

$$P(H_i | D)$$

Probability of H_i given the data

DID THE SUN JUST EXPLODE? (IT'S NIGHT, SO WE'RE NOT SURE.)

THIS NEUTRINO DETECTOR MEASURES WHETHER THE SUN HAS GONE NOVA.

THEN, IT ROLLS TWO DICE. IF THEY BOTH COME UP SIX, IT LIES TO US. OTHERWISE, IT TELLS THE TRUTH.

LET'S TRY.

DETECTOR! HAS THE SUN GONE NOVA?

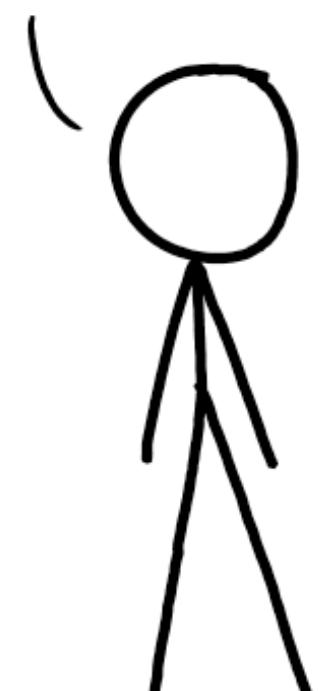
ROLL

YES.



FREQUENTIST STATISTICIAN:

THE PROBABILITY OF THIS RESULT HAPPENING BY CHANCE IS $\frac{1}{36} = 0.027$. SINCE $p < 0.05$, I CONCLUDE THAT THE SUN HAS EXPLODED.



BAYESIAN STATISTICIAN:

BET YOU \$50 IT HASN'T.



You are a professional coin flipper looking for biased coins at the mint. You know that 99 out of every 100 coins are perfectly fair and that 1 out of 100 lands on heads 60% of the time. You flip a coin 50 times and get 33 heads. What are the odds that this coin is biased?

Let B = biased, F = the chance of the flip. Then

$$\begin{aligned}
 p(B|F) &= \frac{p(F|B)p(B)}{p(F|B)p(B) + p(F|\bar{B})p(\bar{B})} \\
 &= \frac{\binom{50}{33}(0.6)^{33}(0.4)^{17}(1/100)}{\binom{50}{33}(0.6)^{33}(0.4)^{17}(1/100) + \binom{50}{33}(0.5)^{50}(99/100)} \\
 &= \frac{(0.6)^{33}(0.4)^{17}(1/100)}{(0.6)^{33}(0.4)^{17}(1/100) + (0.5)^{50}(99/100)} \\
 &= \frac{1.2^{33}0.8^{17}(0.01)}{0.99 + 1.2^{33}0.8^{17}(0.01)} \\
 &\approx 0.0853 \qquad \approx \frac{.09237}{.99 + .09237}
 \end{aligned}$$

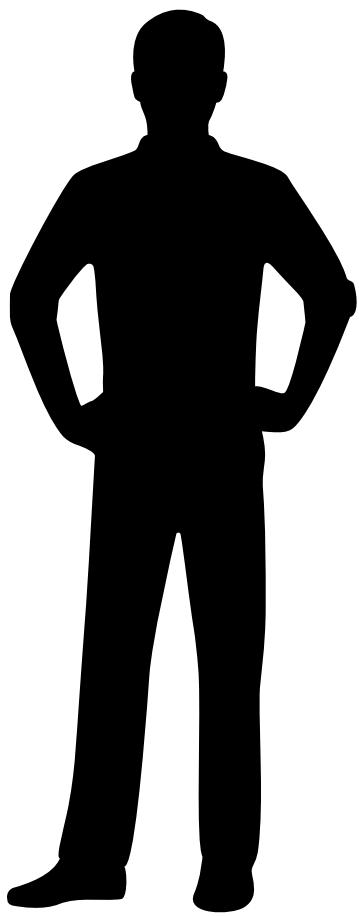
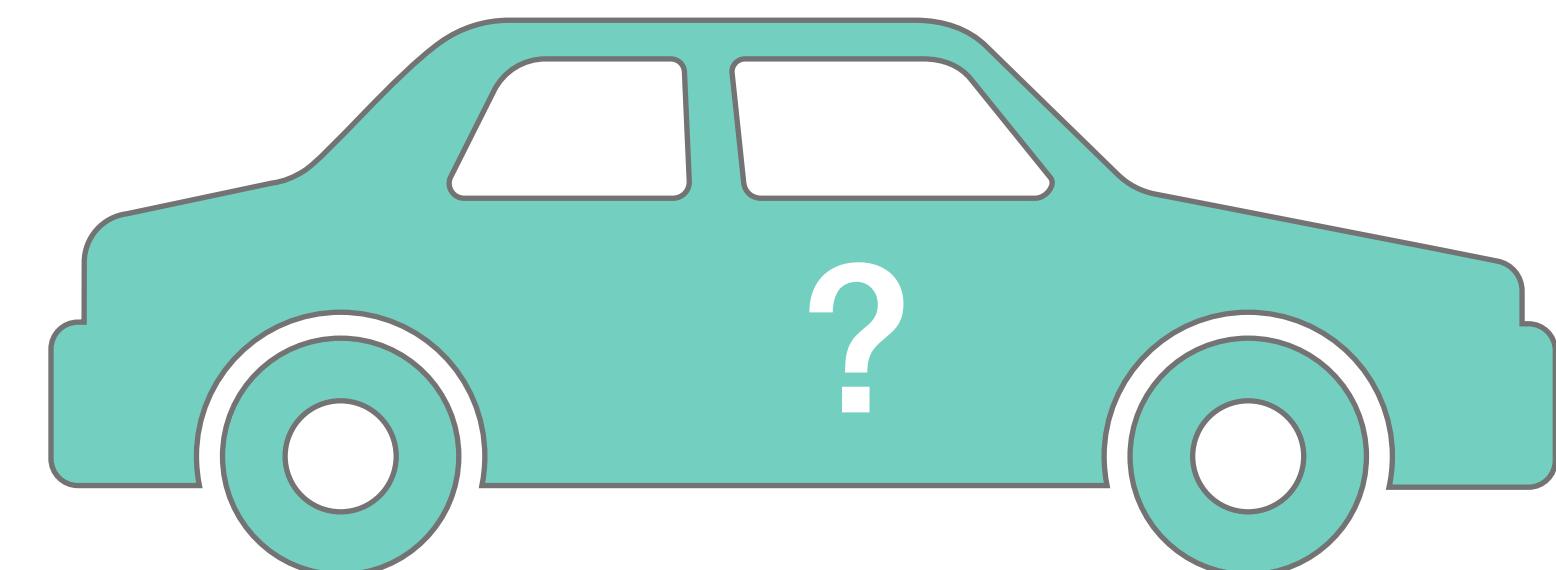
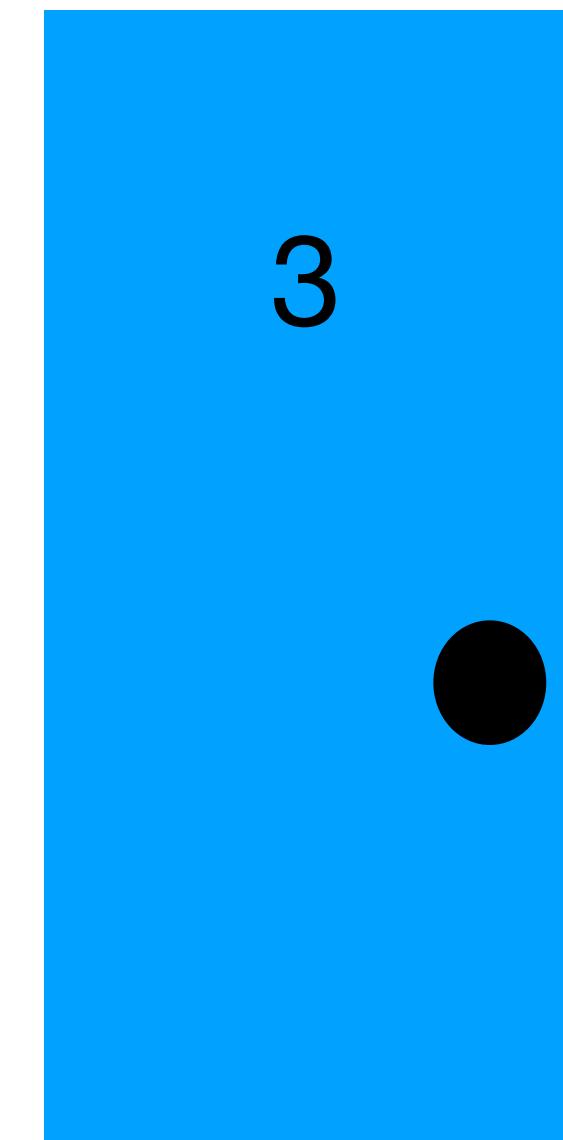
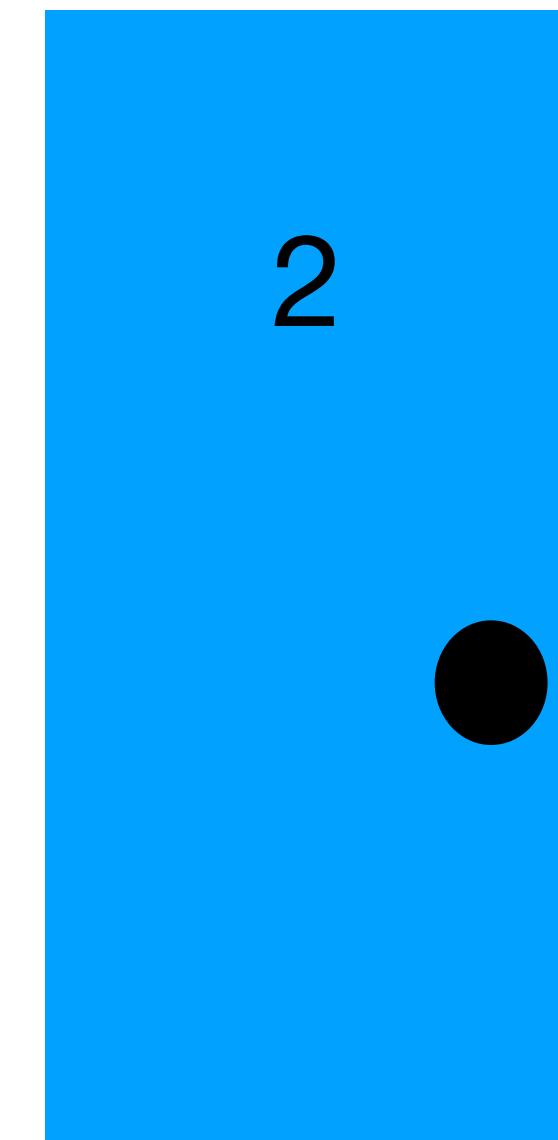
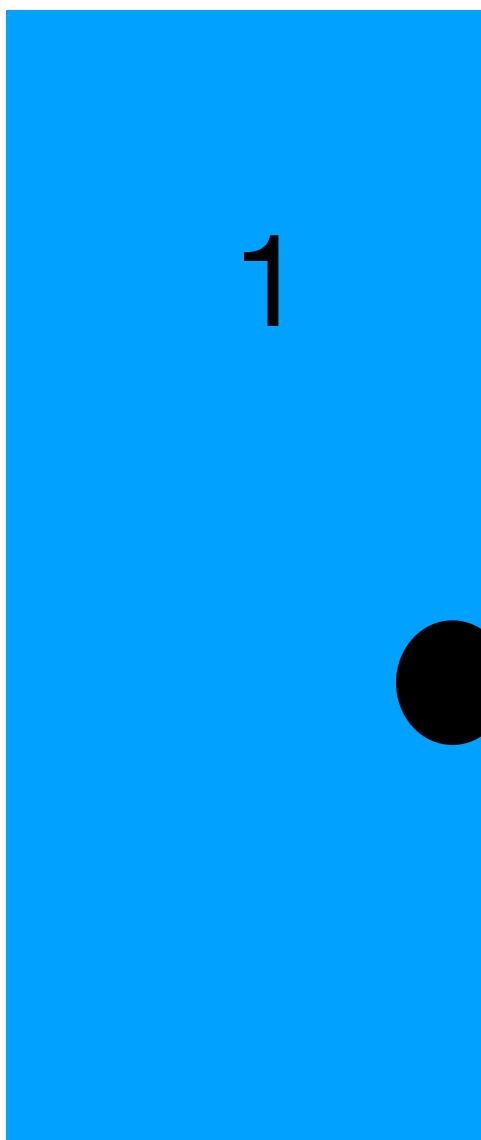
You decided that this coin is suspicious and flip it 50 more times, getting another 29 heads. If your priors are taken from the answer to the previous question, what are the new odds that the coin is biased?

$$\begin{aligned} p(B|F) &= \frac{p(F|B)p(B)}{p(F|B)p(B) + p(F|\bar{B})p(\bar{B})} \\ &= \frac{1.2^{33}0.8^{17}(0.0853)}{0.915 + 1.2^{33}0.8^{17}(0.0853)} \\ &\approx \frac{.788}{.915 + .788} \\ &\approx .463 \end{aligned}$$

The chance that the coin is biased is now a little less than 50%.

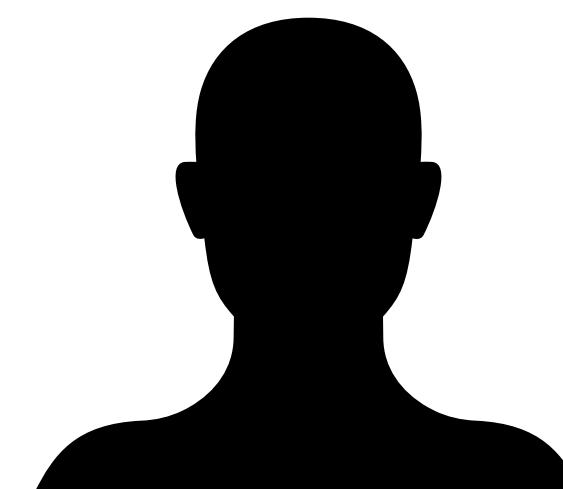
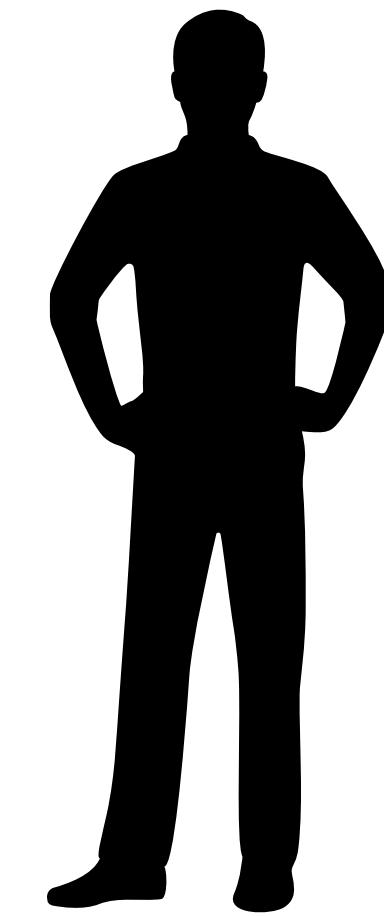
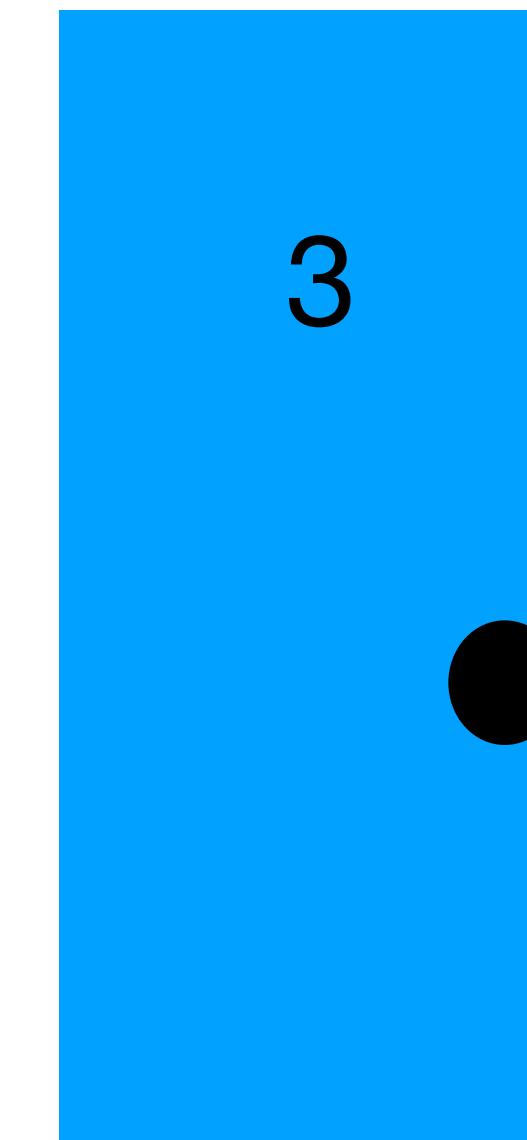
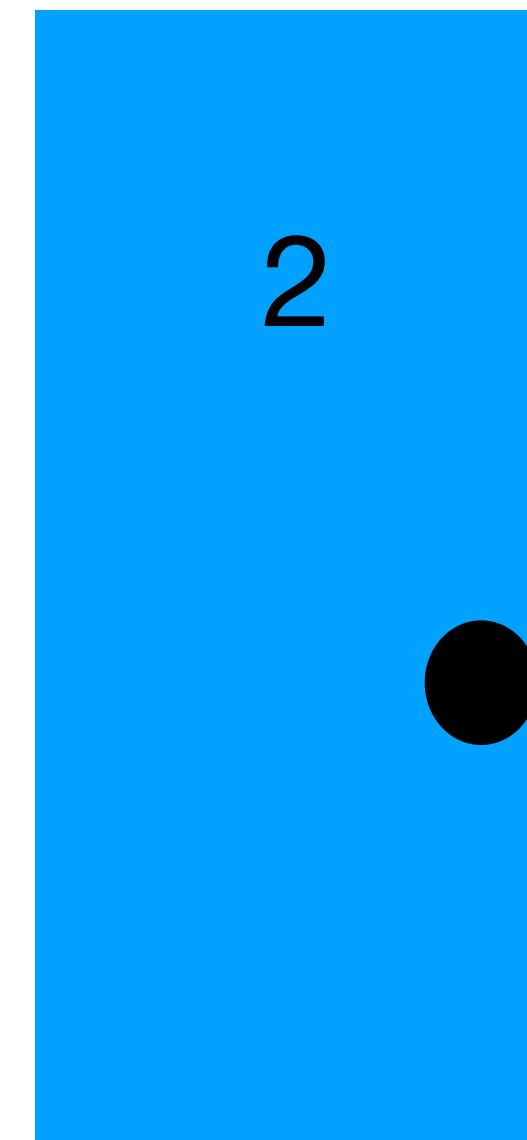
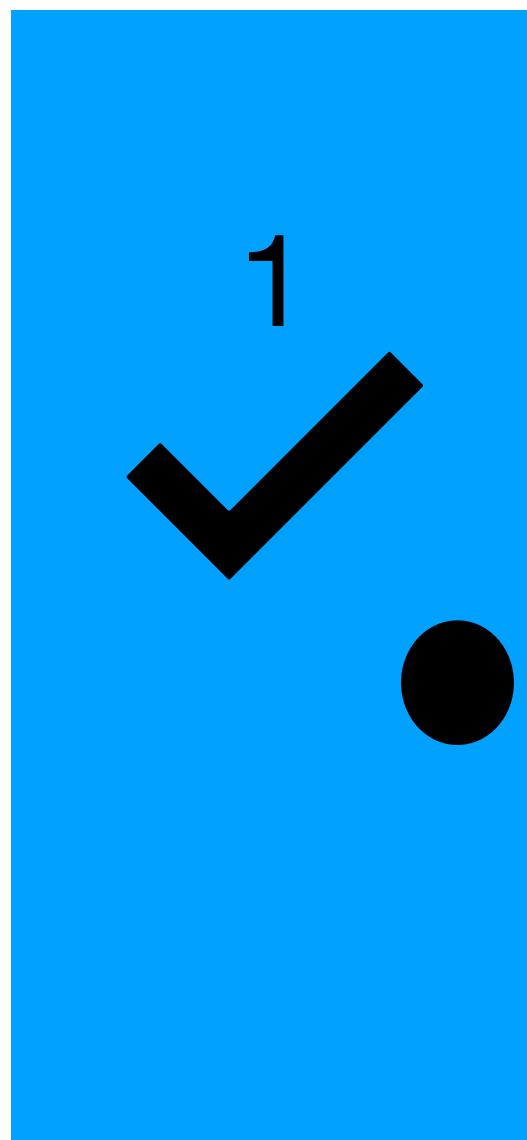
The Monty Hall Problem

Let's make a deal!



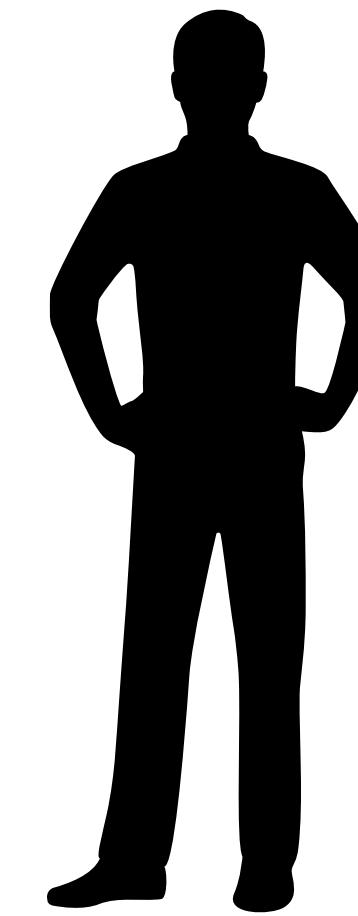
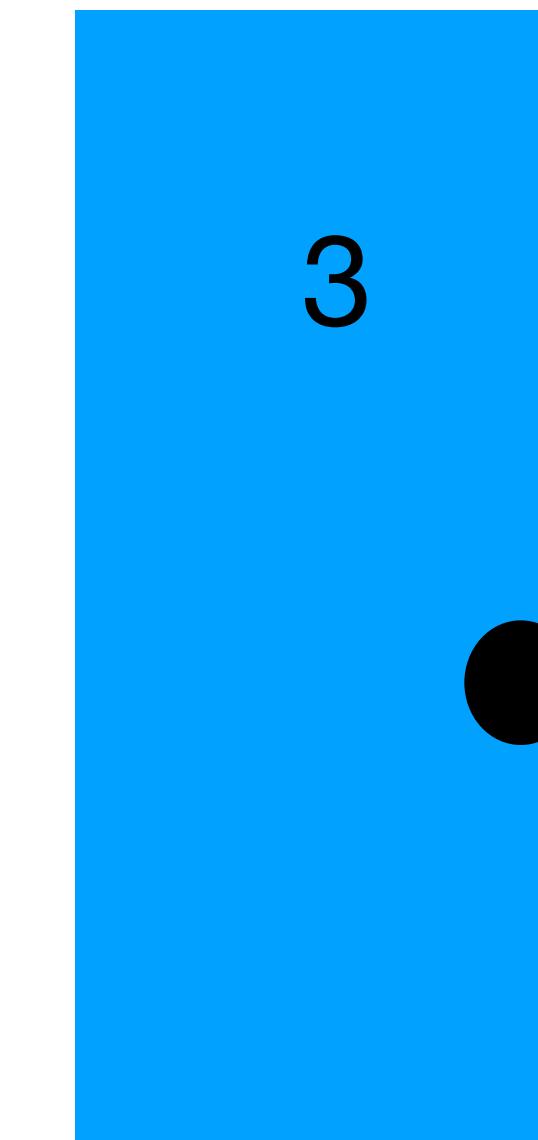
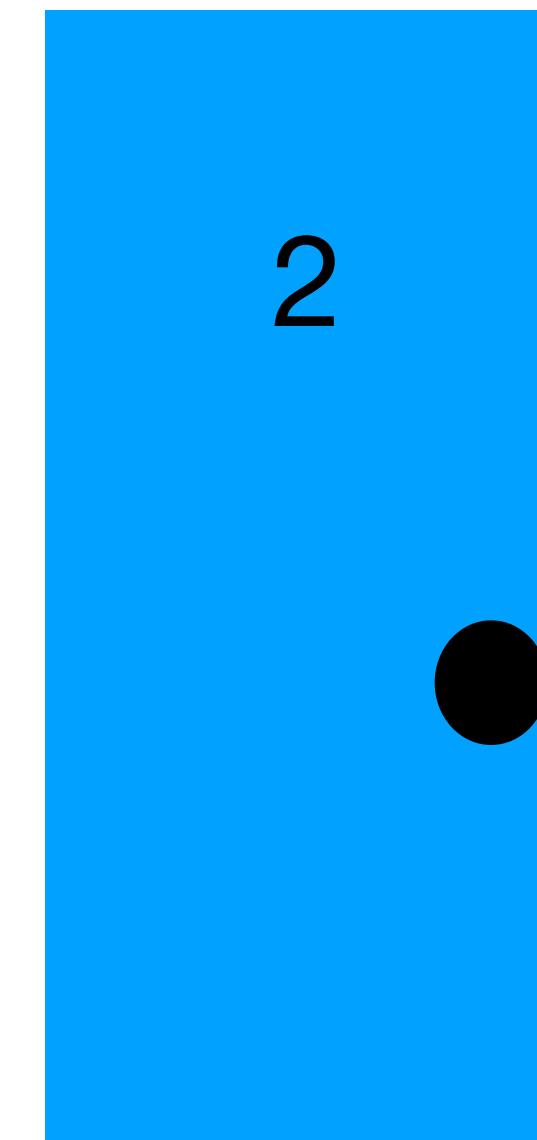
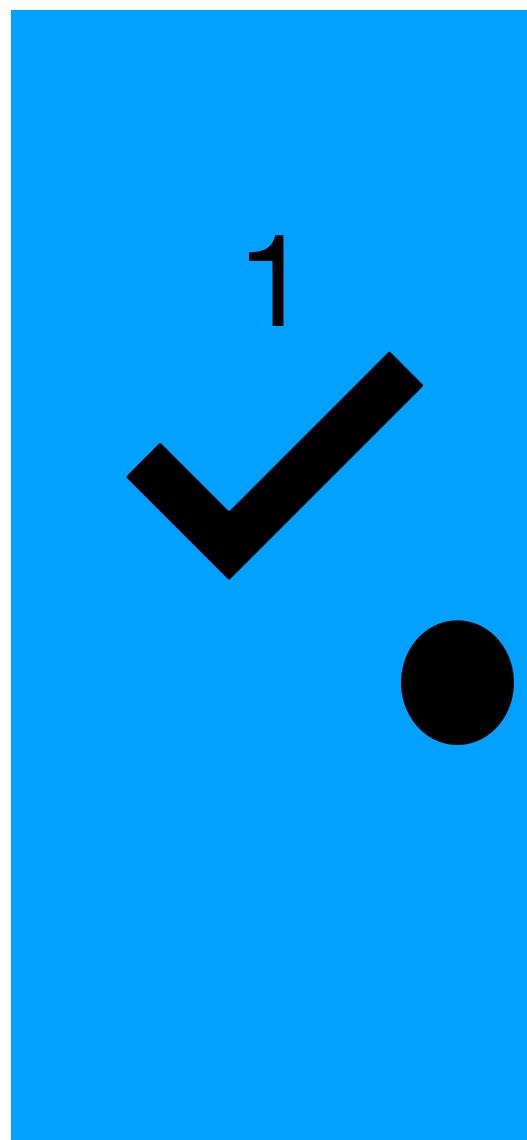
The Monty Hall Problem

Let's make a deal!



The Monty Hall Problem

Let's make a deal!

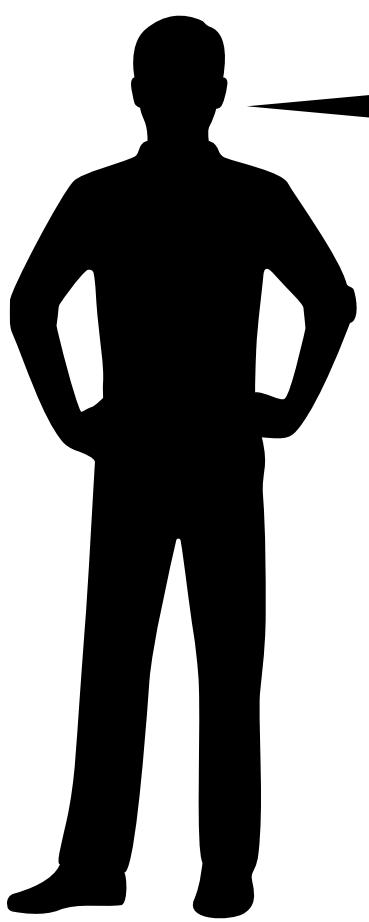
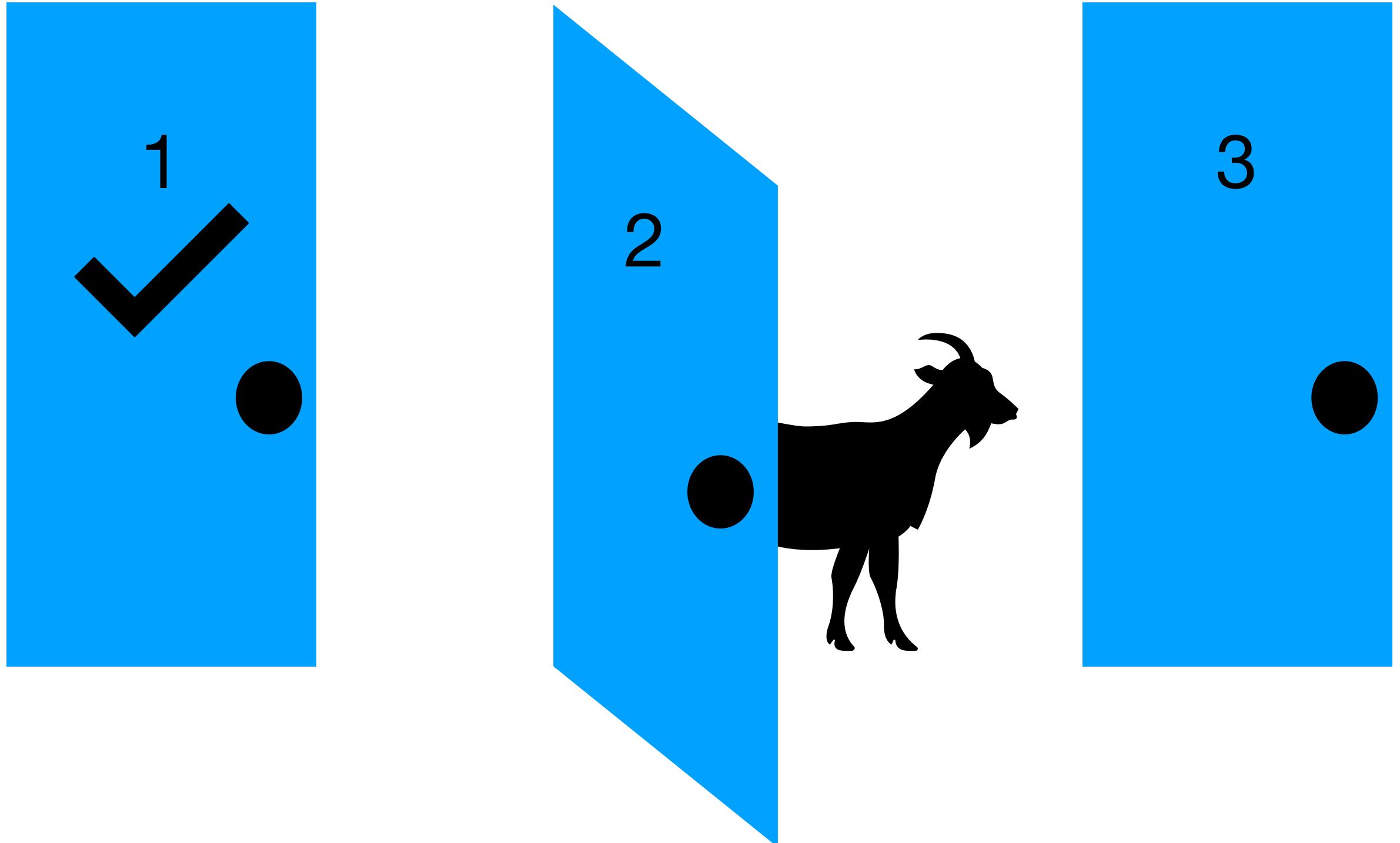


$P(\text{car behind D1})?$

$P(\text{car behind D2 and D3})?$

The Monty Hall Problem

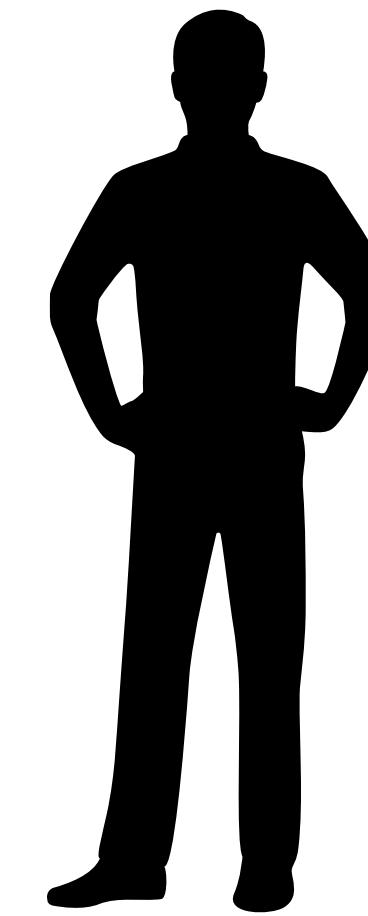
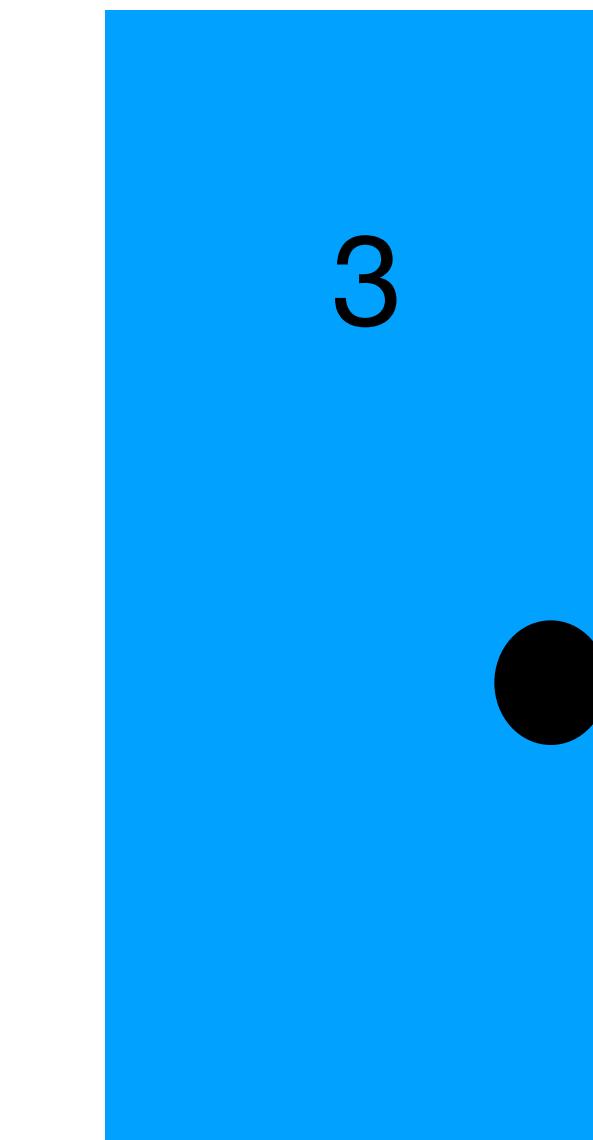
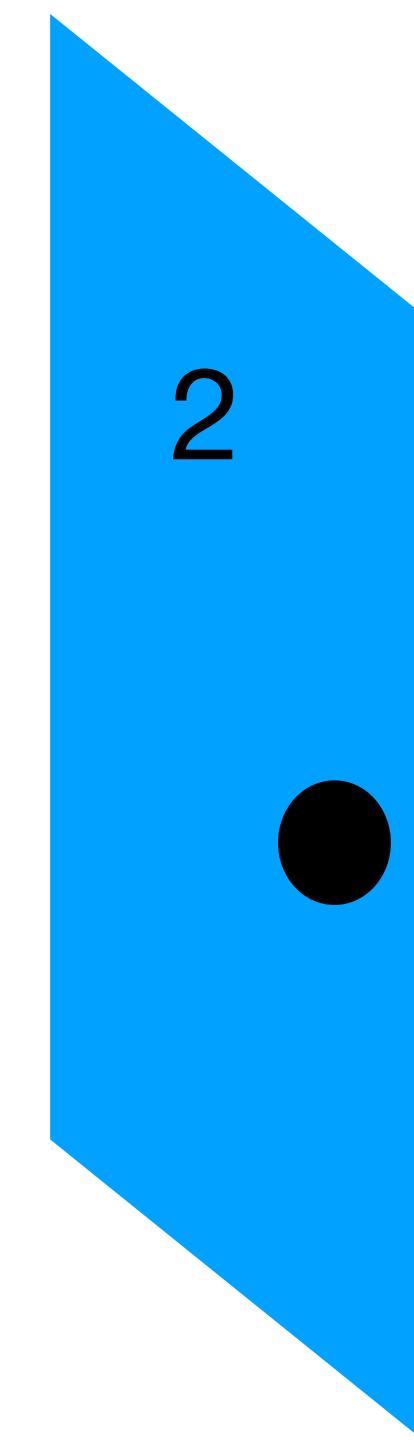
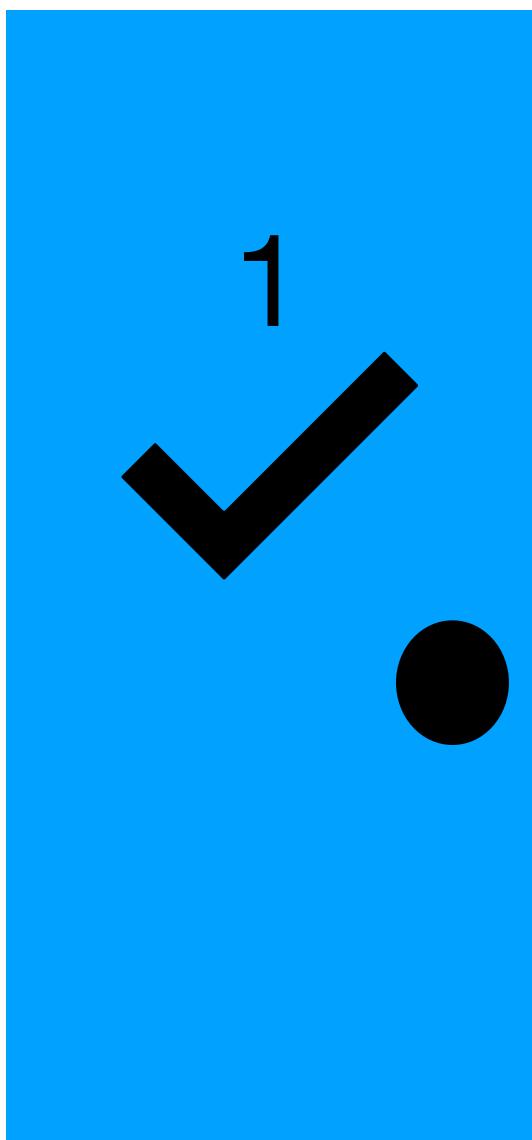
Let's make a deal!



Stay or switch?

The Monty Hall Problem

Let's make a deal!

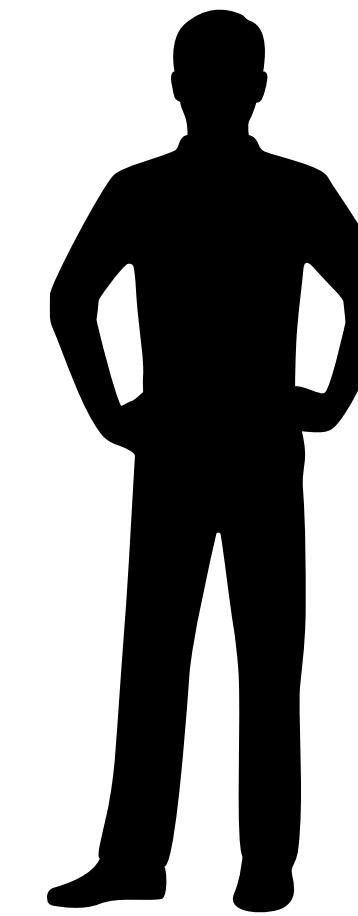
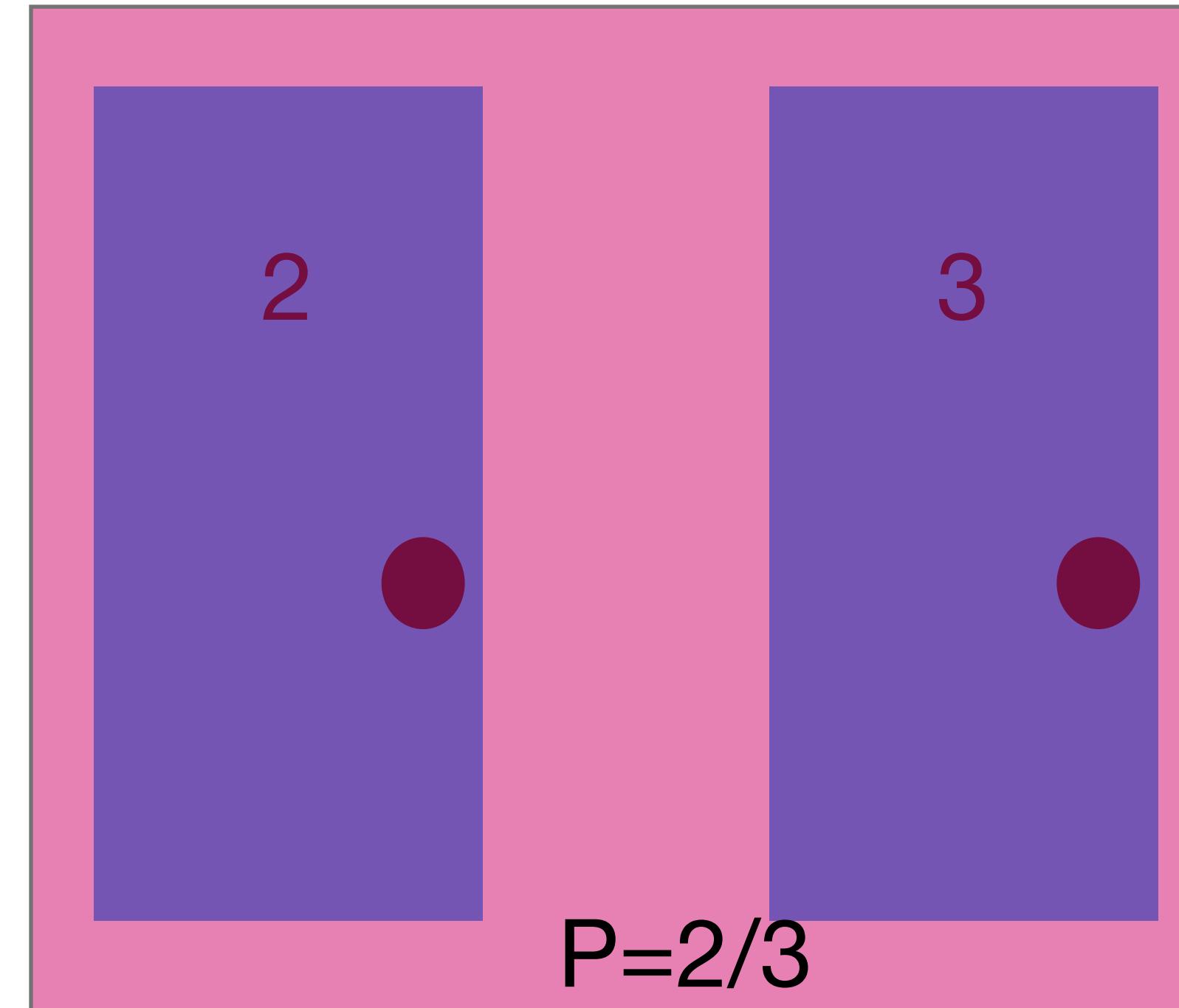
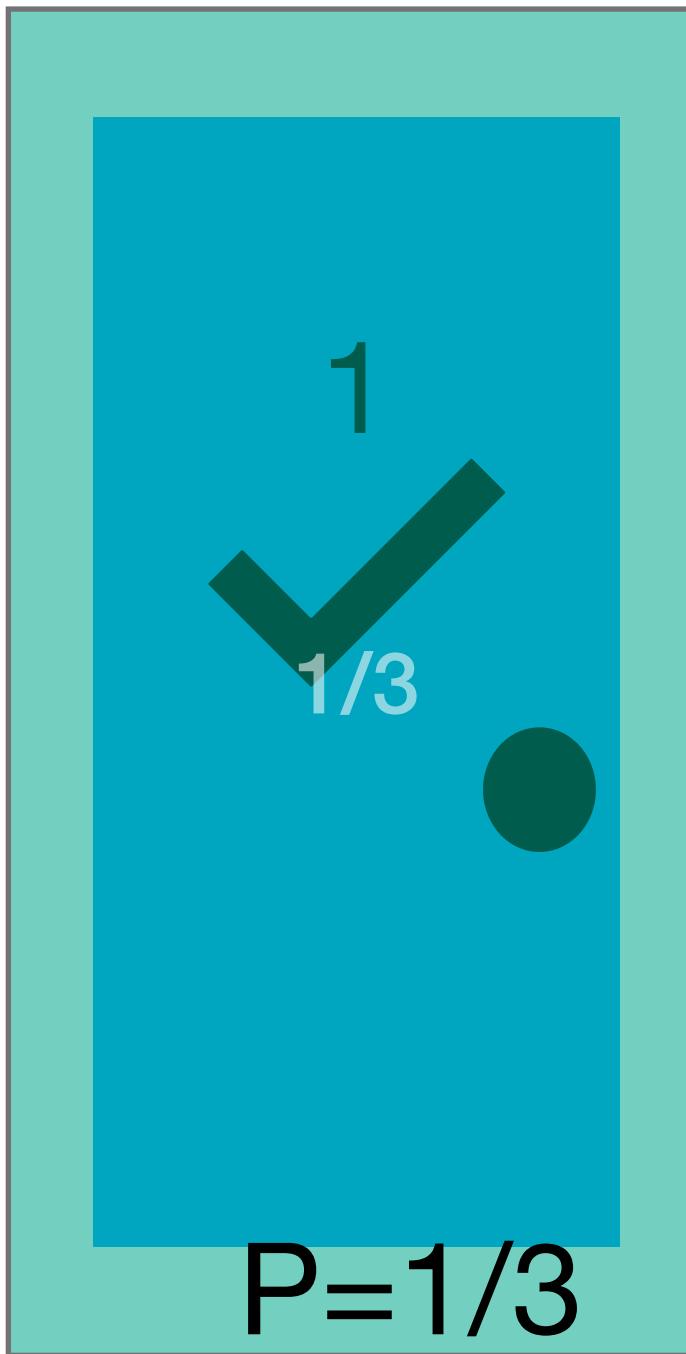


$P(\text{car behind D1})?$

$P(\text{car behind D2 and D3})?$

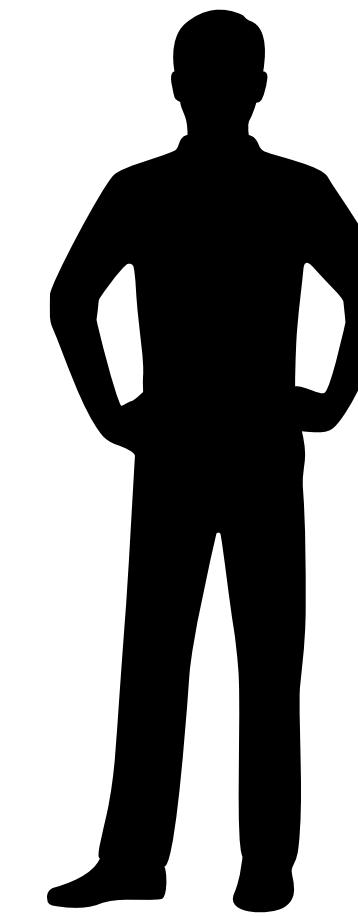
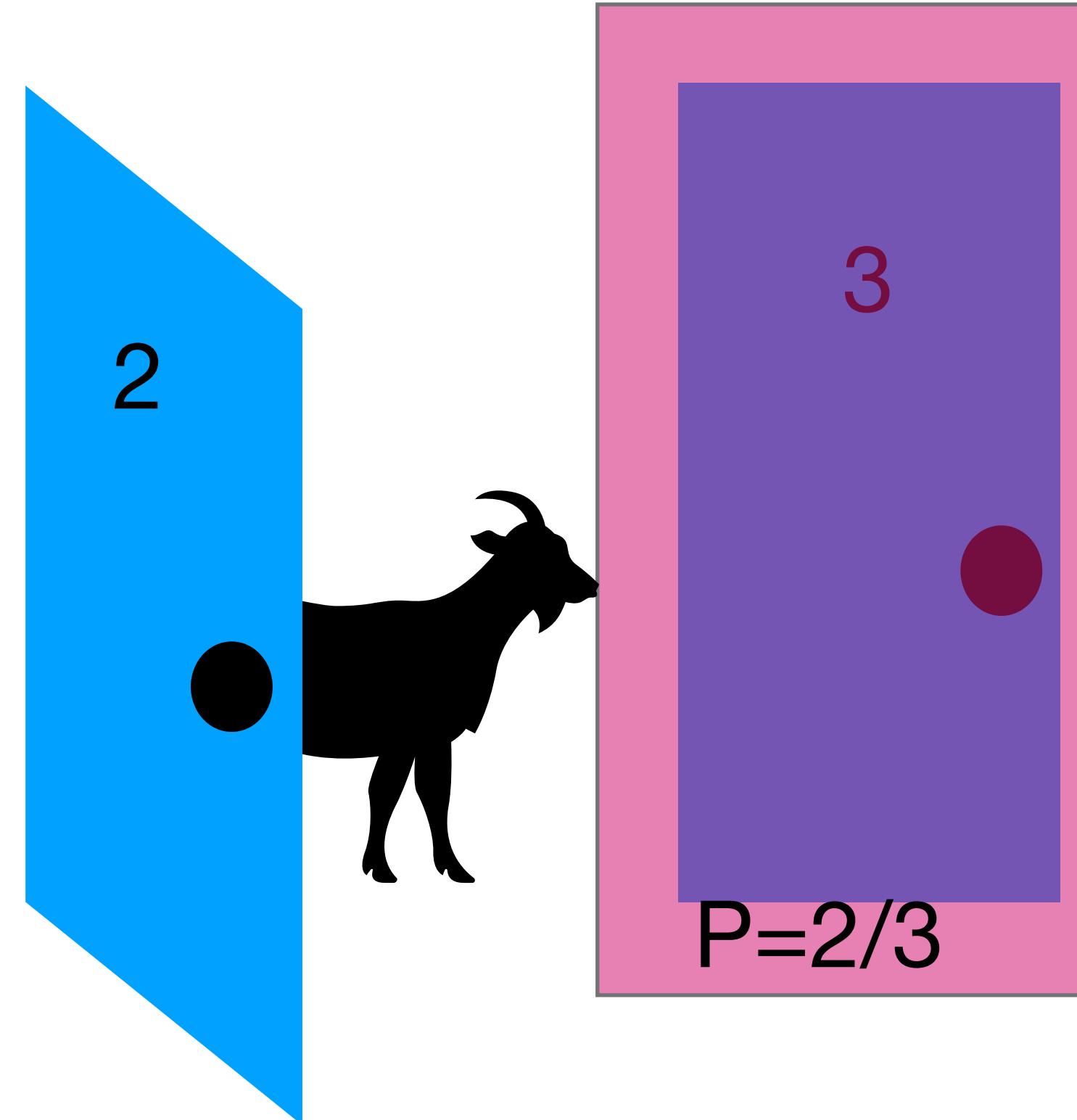
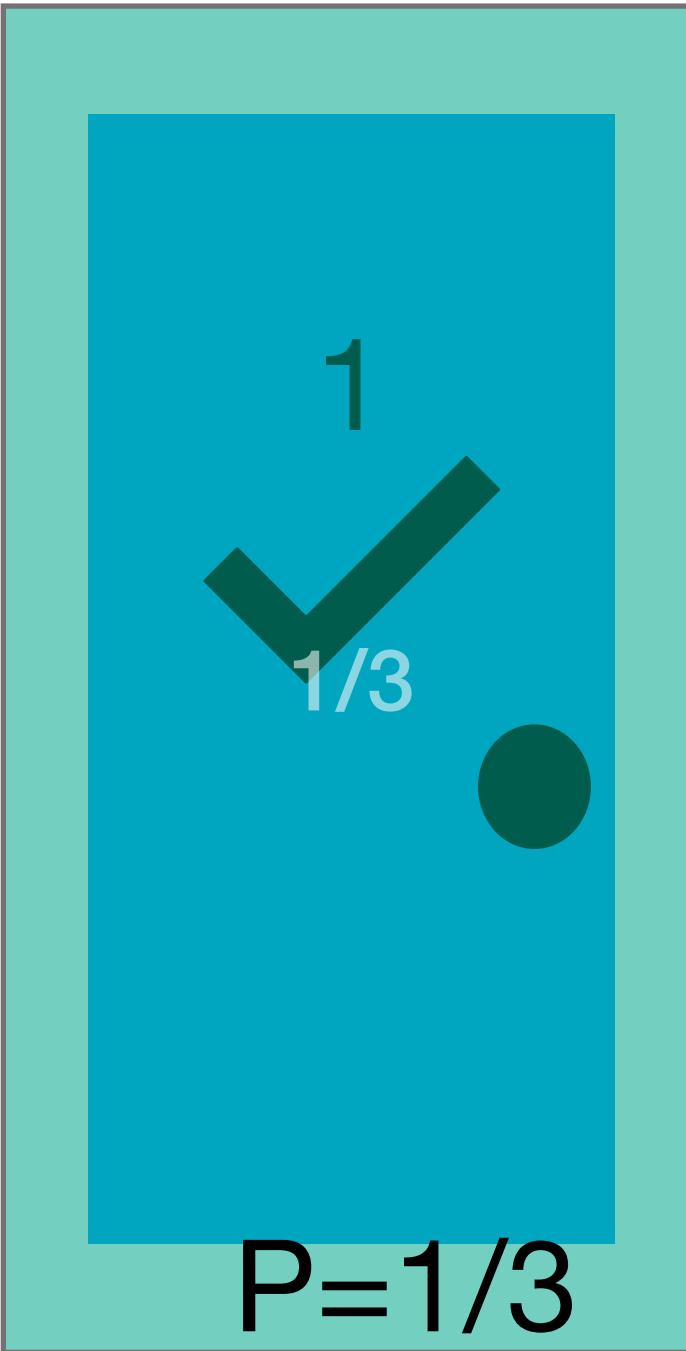
The Monty Hall Problem - Intuition

Let's make a deal!



The Monty Hall Problem - Intuition

Let's make a deal!



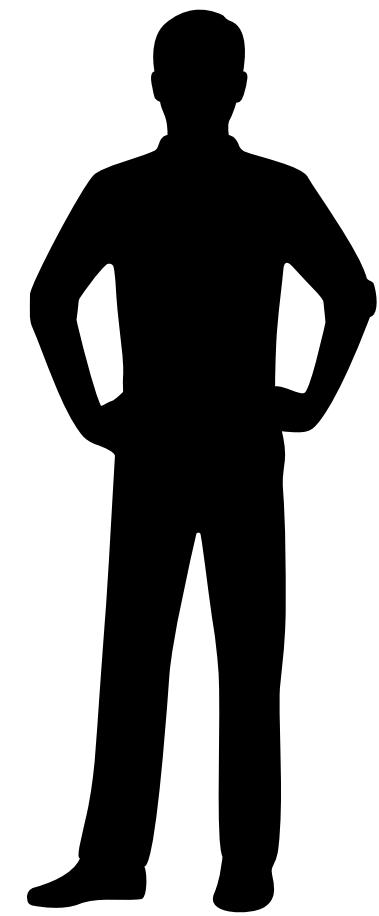
The Monty Hall Problem - Bayes math

Let's make a deal!

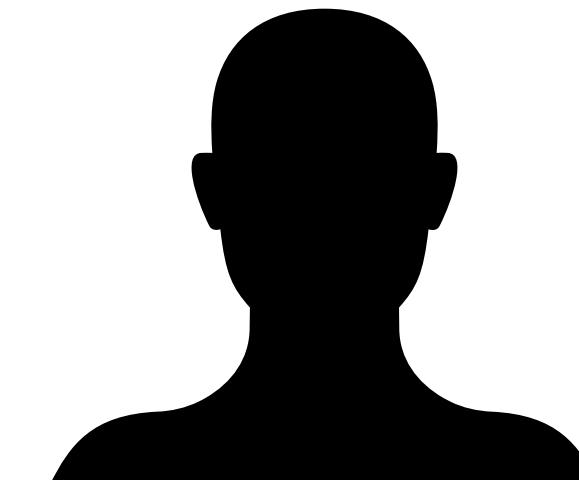
Want to calculate:

$$P(\text{car} = 3 \mid \text{opens} = 2, \text{choose} = 1)$$

$$= \frac{P(\text{opens} = 2, \text{choose} = 1 \mid \text{car} = 3)P(\text{car} = 3)}{P(\text{opens} = 2, \text{choose} = 1)}$$



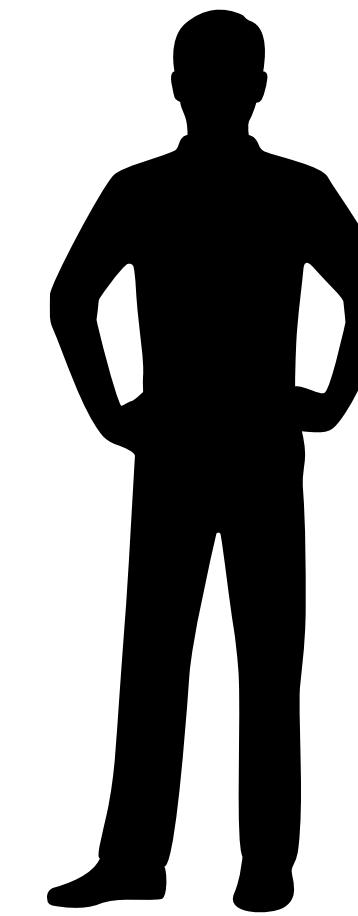
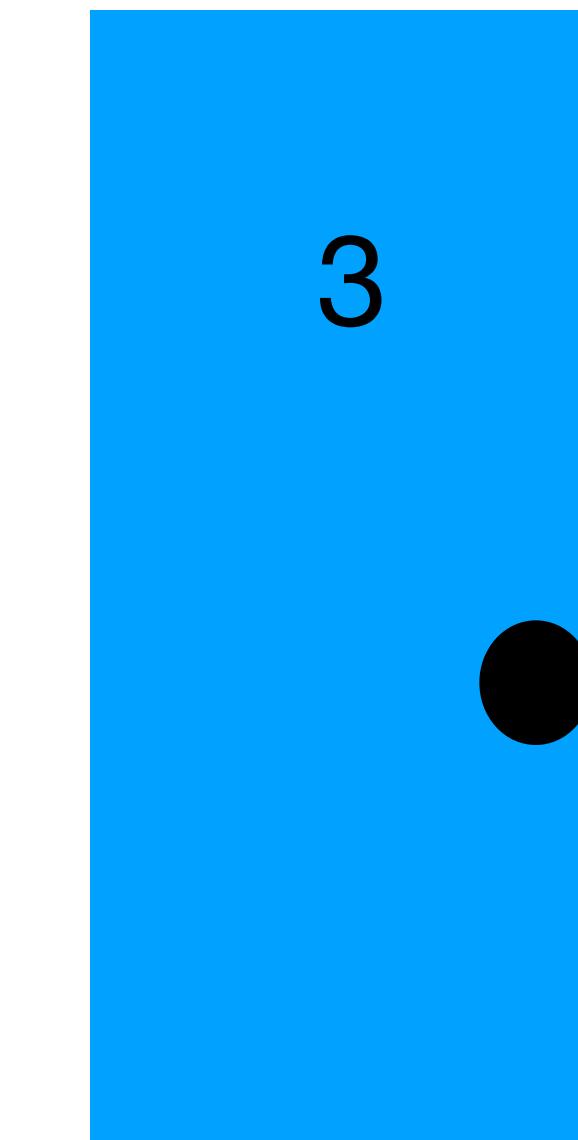
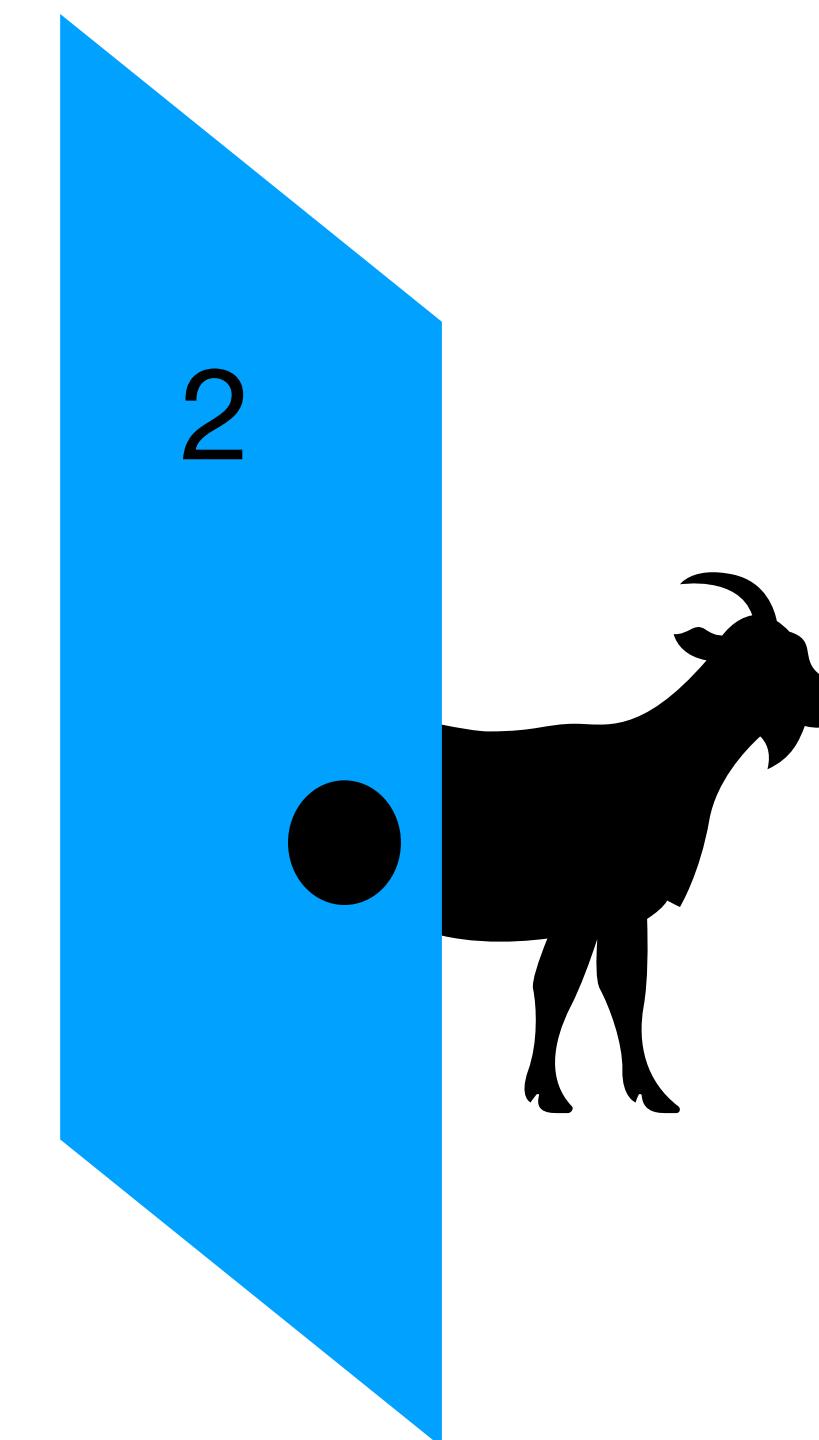
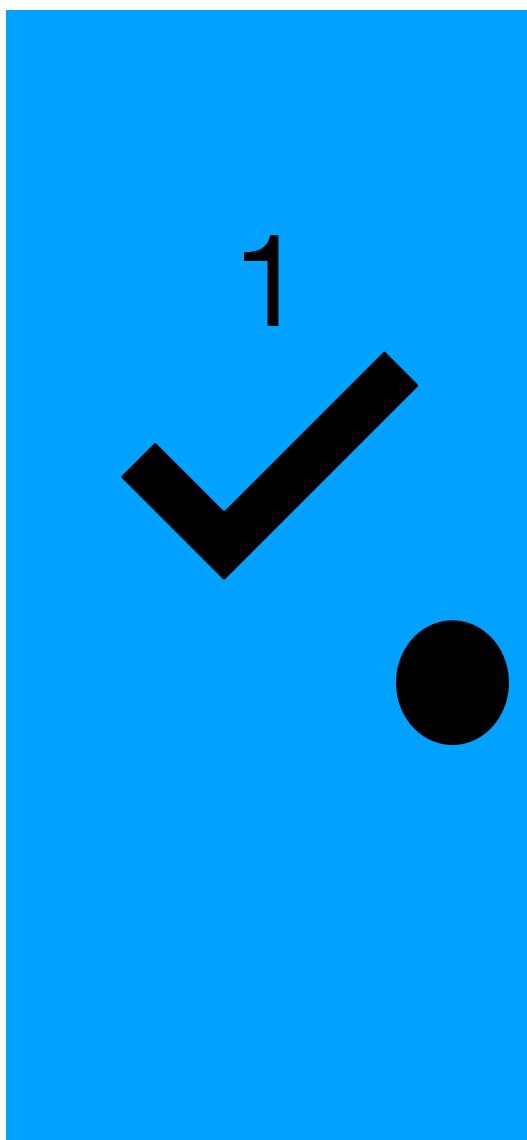
Left to you as an exercise...



Literally. On A1.

The Monty Hall Problem

Let's make a deal!



$P(\text{car behind D1})?$

$P(\text{car behind D2 and D3})?$