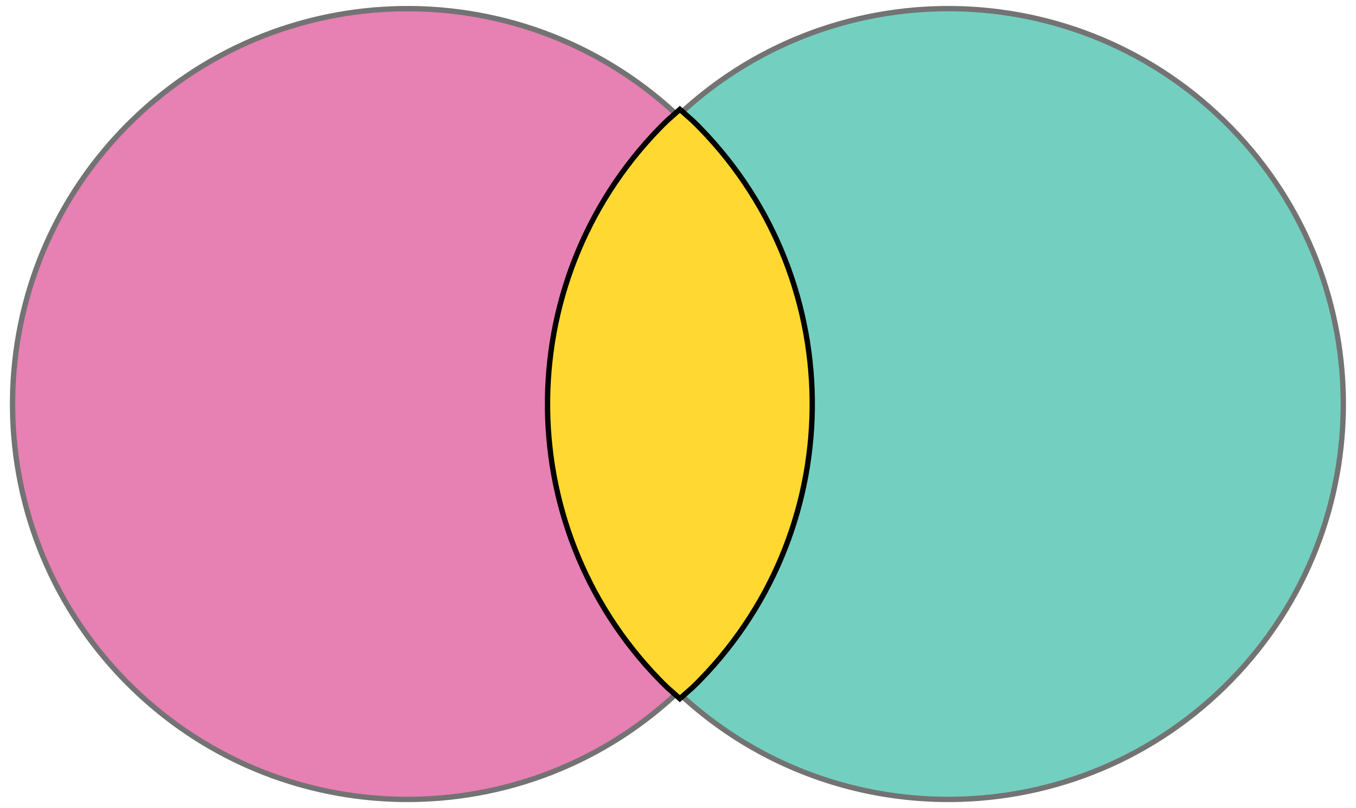
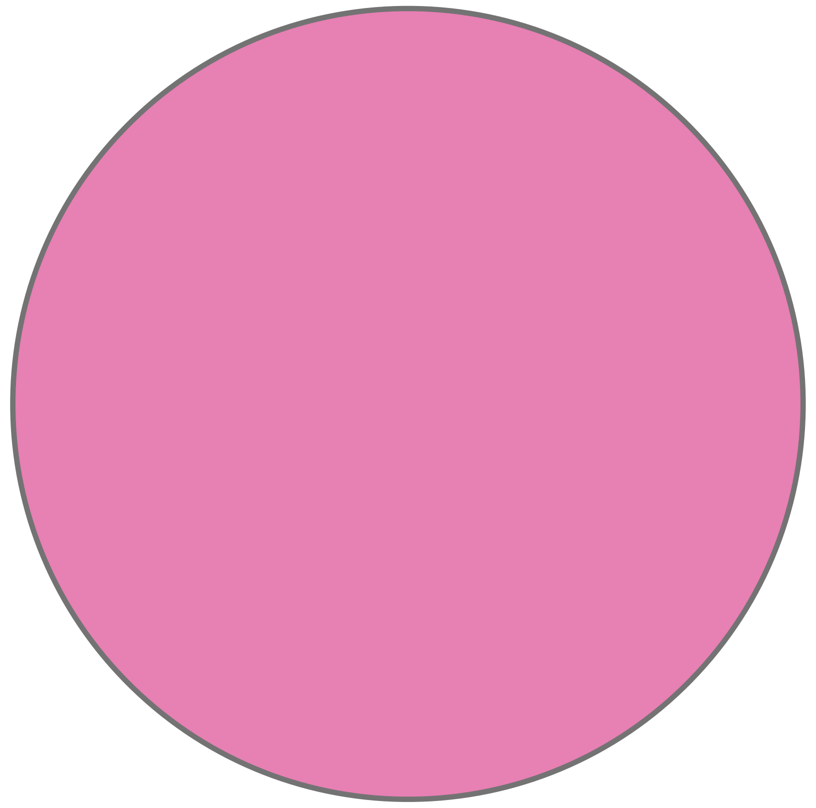
Probability

Pre-lecture 2 video

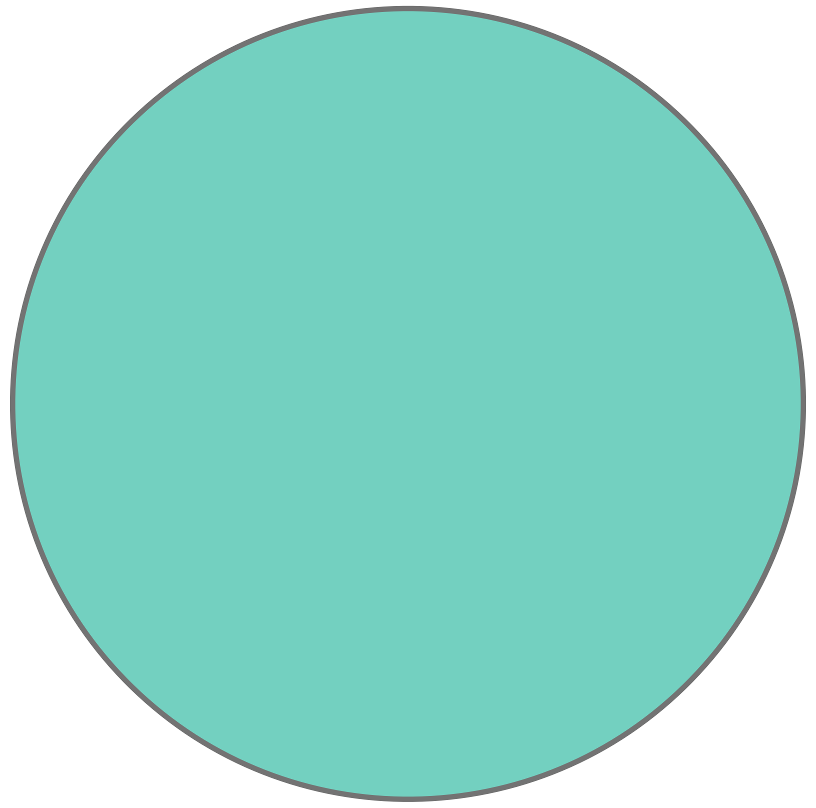
$$P(A, B) = P(A \cap B)$$

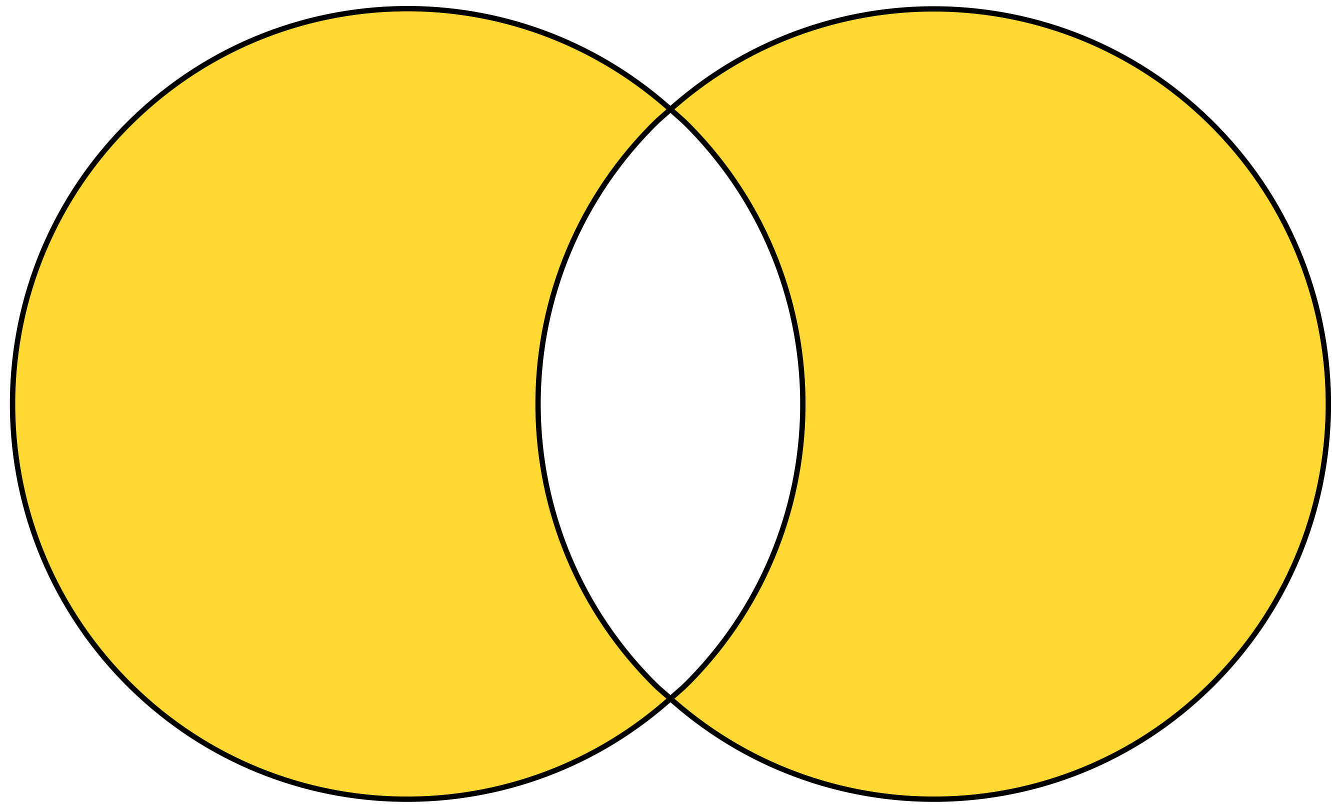


$P(A)$



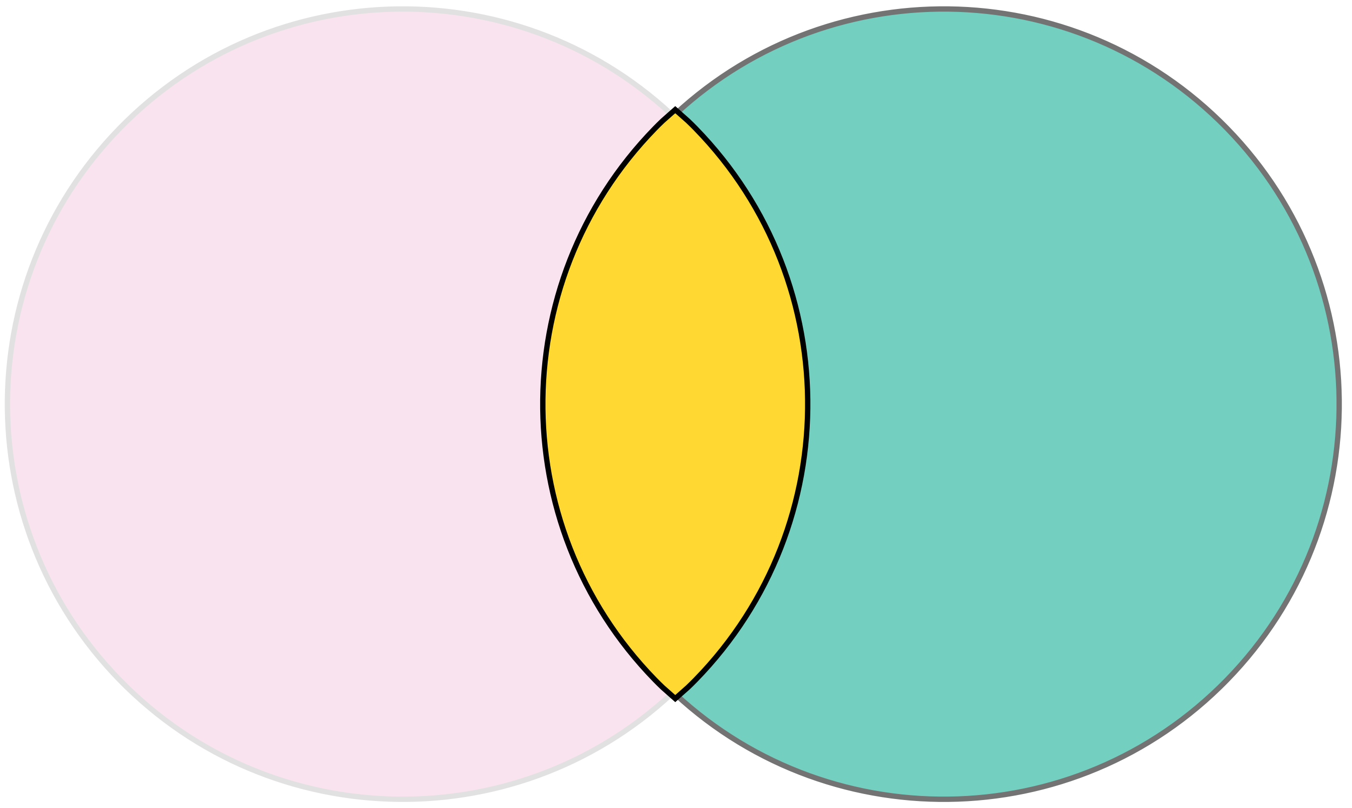
$P(B)$





$$P(A \cup B)$$

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



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

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

It is *conditional*!

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

Note the consequence of this viewpoint

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

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

Can be rearranged

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

disease X

symptoms Y	0	1
0	0.5	0.1
1	0.1	0.3

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

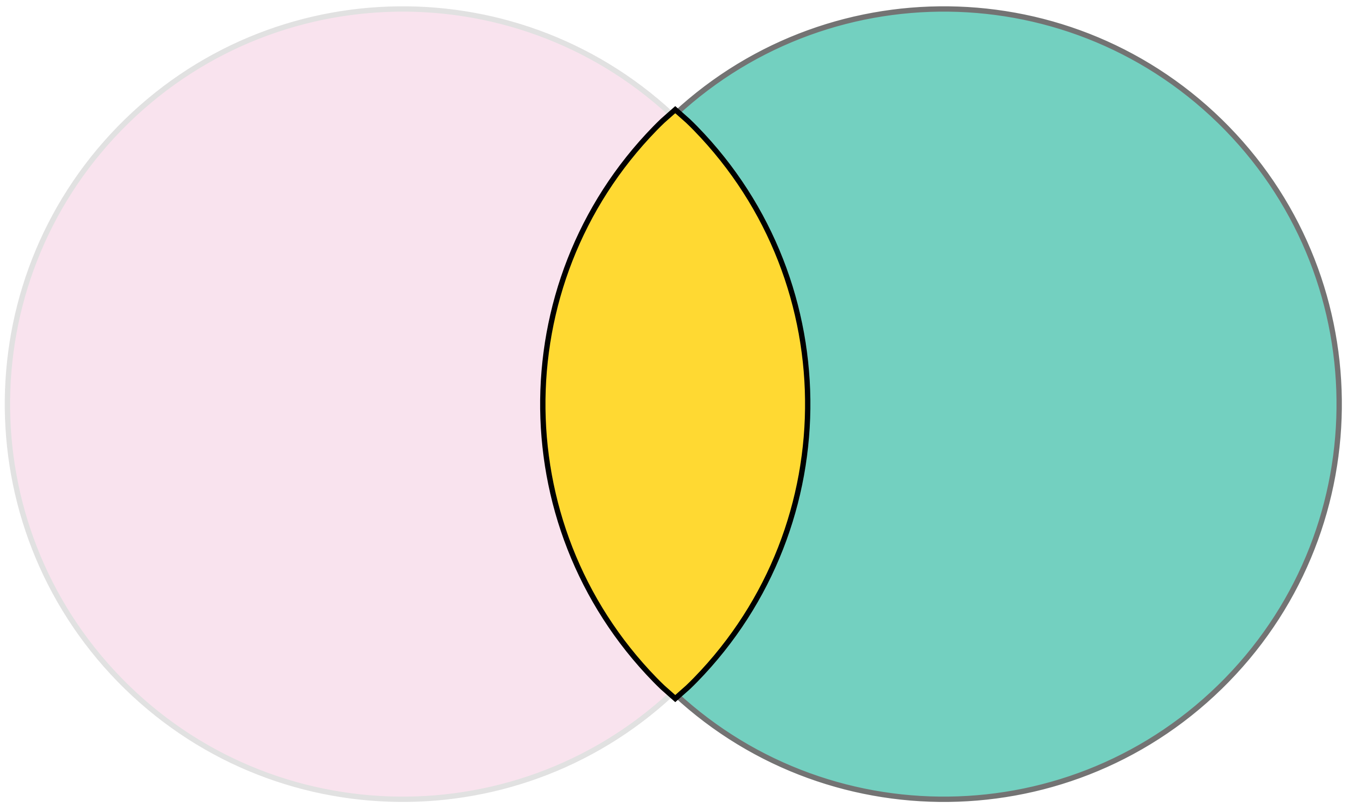
$$P(Y = 1) = 0.1 + 0.3 = 0.4$$

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

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

This is called the marginal probability,

it is NOT conditioned on anything



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

Most general case!

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

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

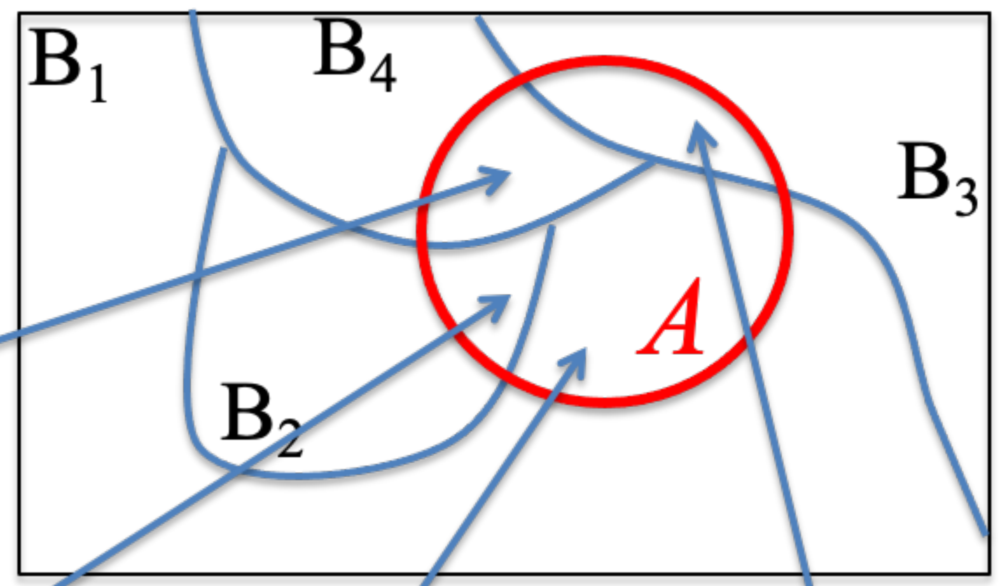
$$P(A \mid B) = P(A)$$

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

The conditional probabilities are equivalent to the marginal probabilities

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

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

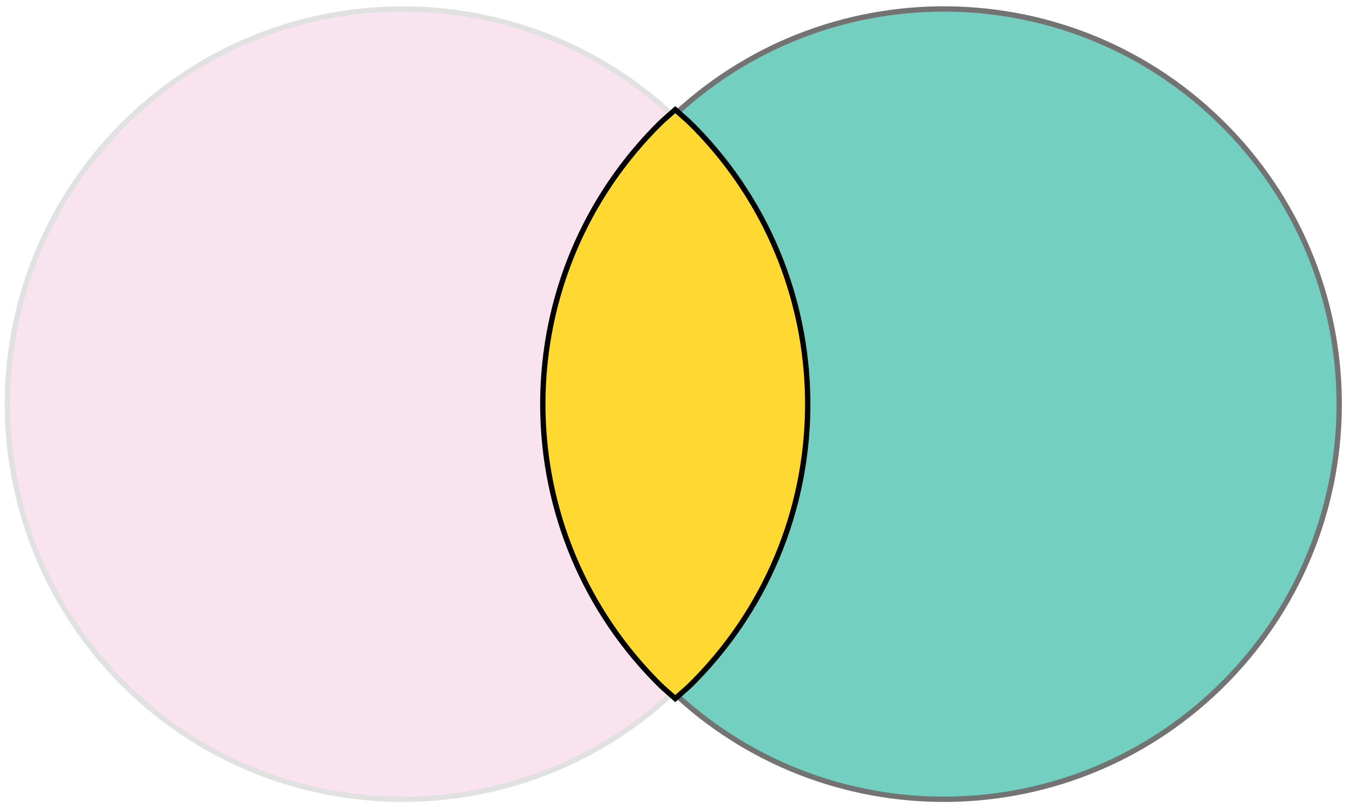


Then the probability of an event A can be written by adding up the intersection of A with all the disjoint sets B_i

$$\Pr(A) = \Pr(B_1 \cap A) + \Pr(B_2 \cap A) + \Pr(B_3 \cap A) + \cdots + \Pr(B_m \cap A)$$

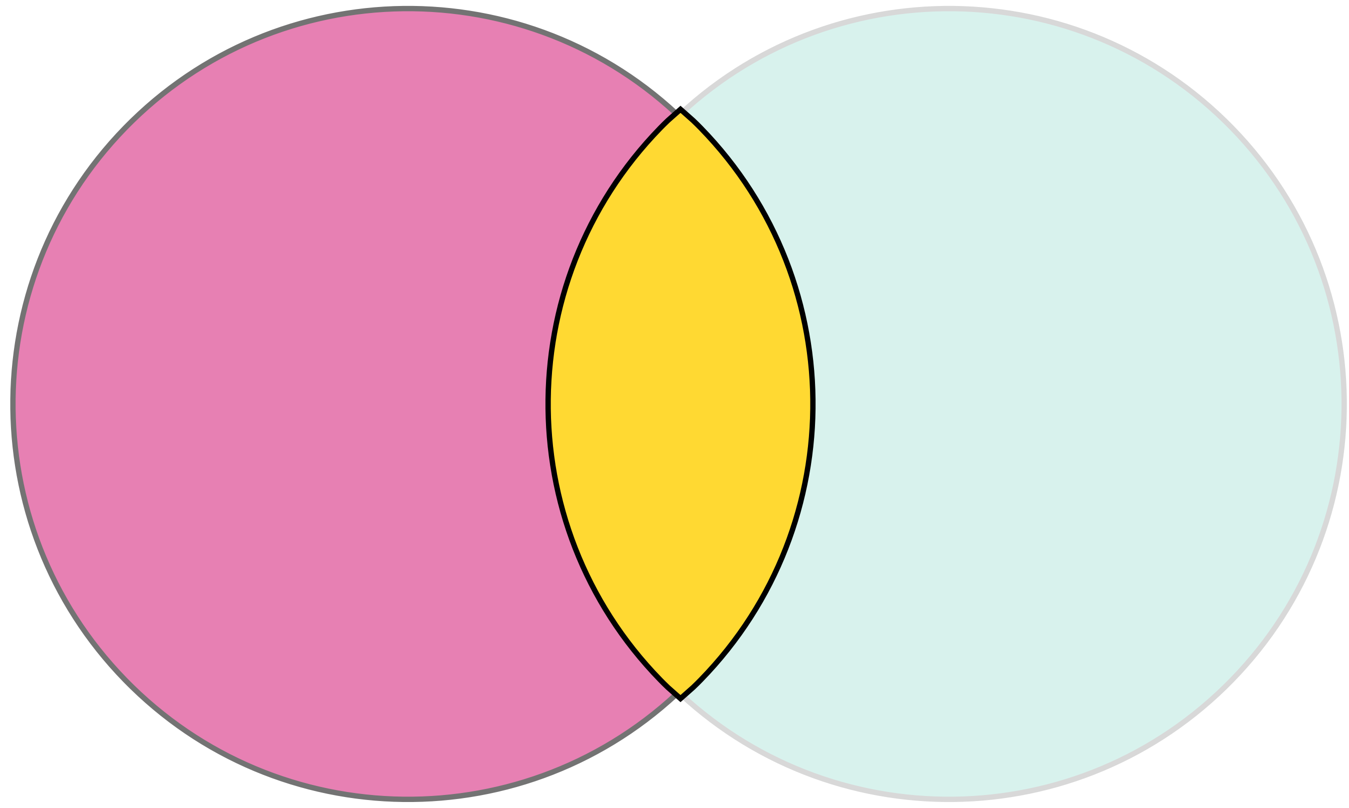
$$= \Pr(A|B_1)\Pr(B_1) + \Pr(A|B_2)\Pr(B_2) + \Pr(A|B_3)\Pr(B_3) + \cdots + \Pr(A|B_m)\Pr(B_m)$$

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



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

Intersections are commutative



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

So...

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

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

$$\mathrm{B_1 \cup B_2 \cup \ldots \cup B_m = \Omega}$$

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

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

$P(X=1) =$

$P(Y=4) =$

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

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

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

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

X: 🎲#1

Y: 🎲#2

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

Statistically independent pizza!

P(pepperoni) =

P(mushroom) =

P(pep, mush) =

P(pep | mush) =

P(mush | 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*

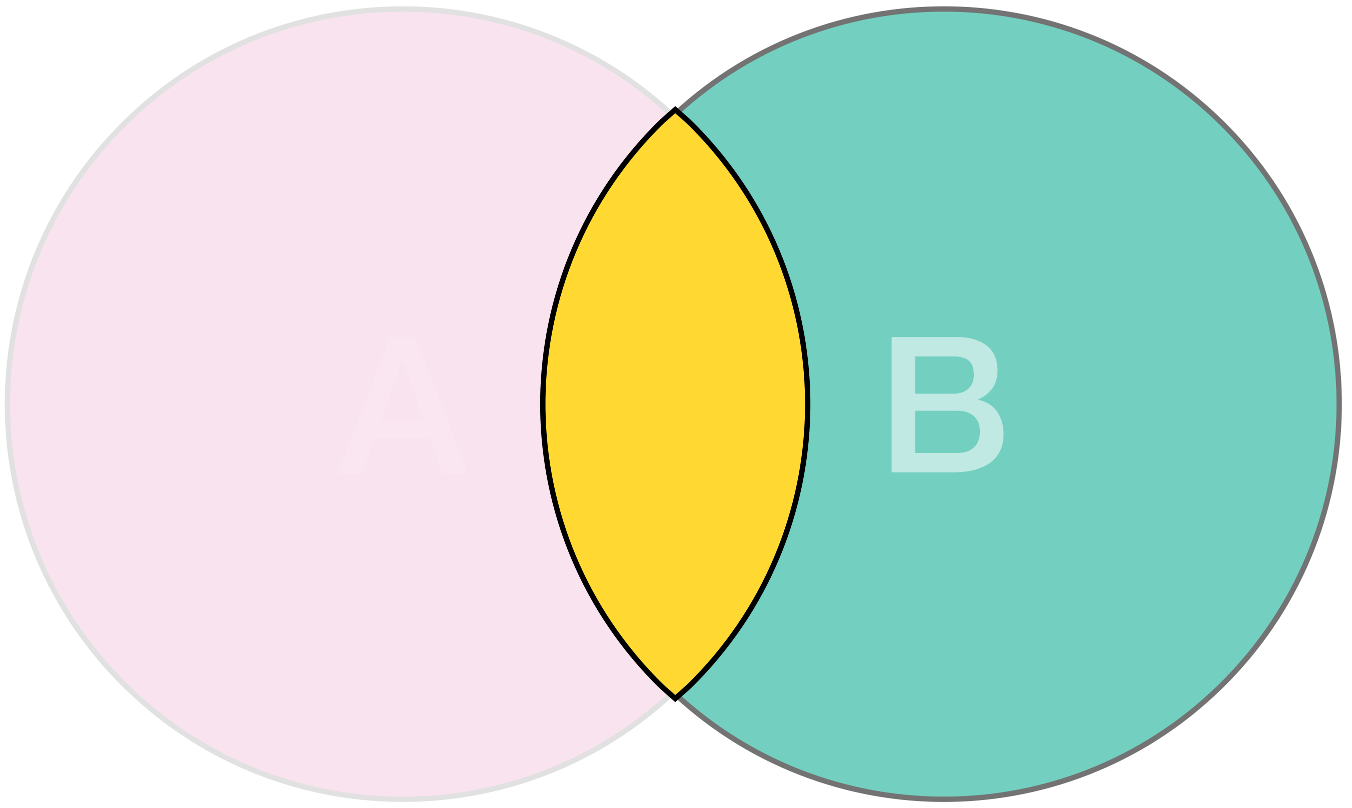
P(pepperoni) =

P(mushroom) =

P(pineapple) =

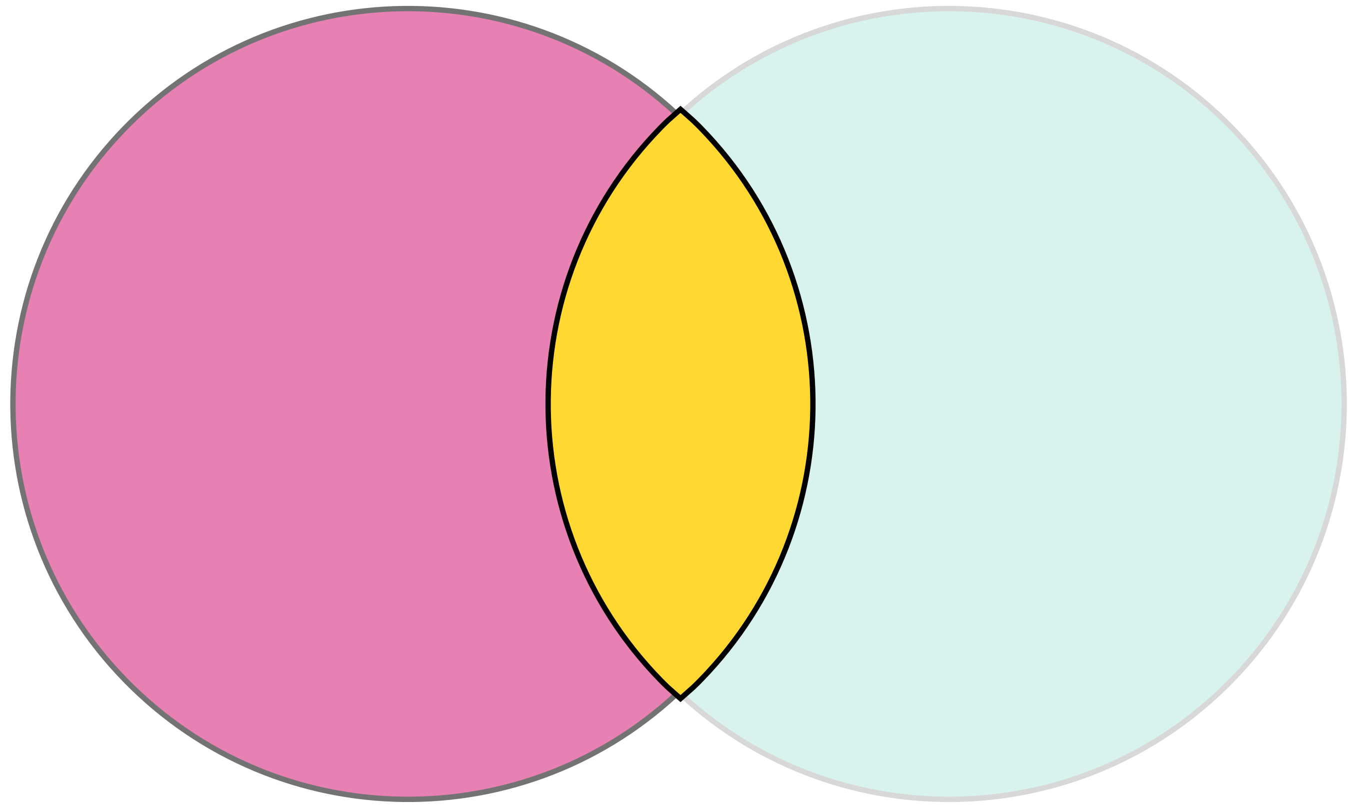
P(pine,mush | pep) =

P(pep | pine,mush) =



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

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

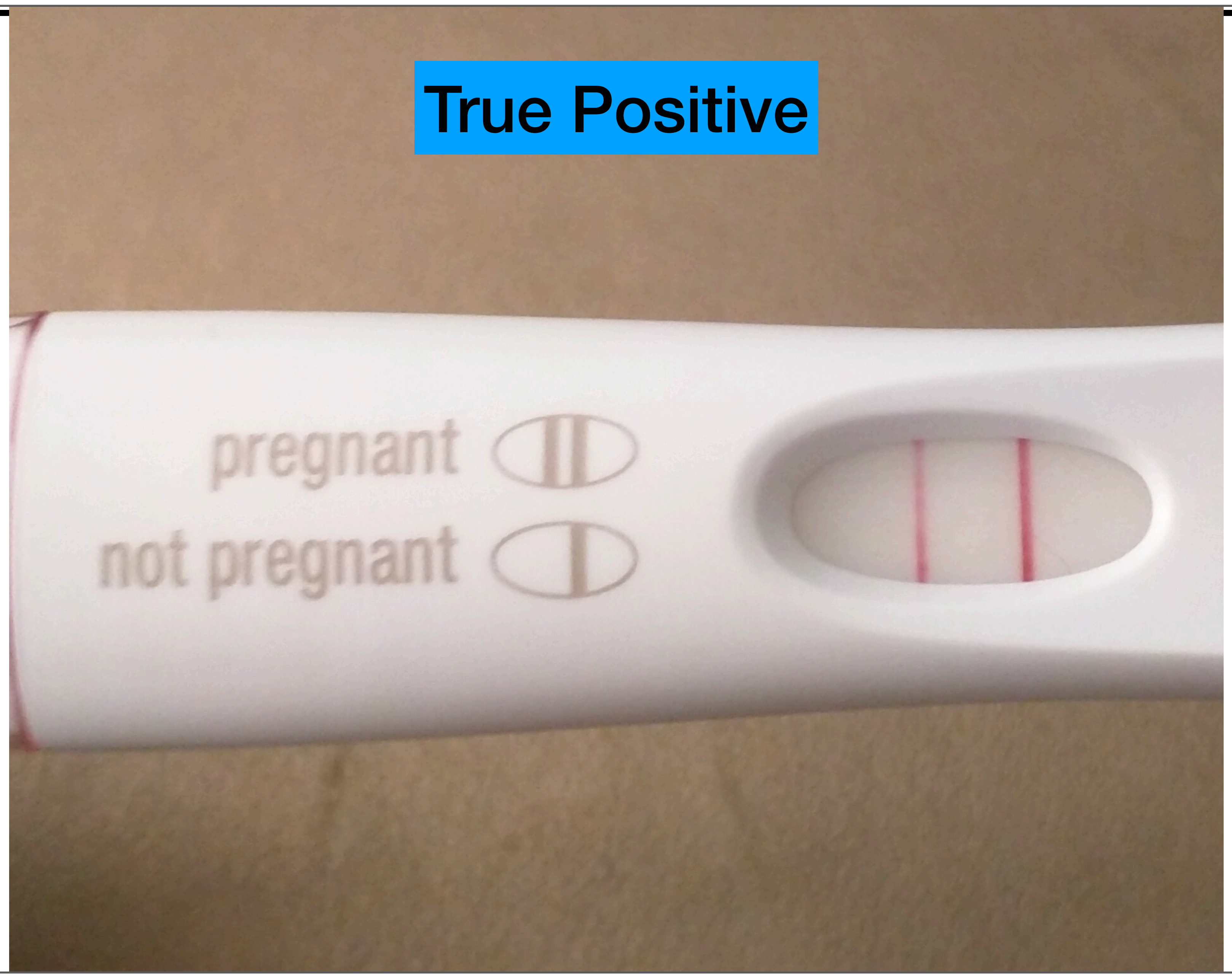
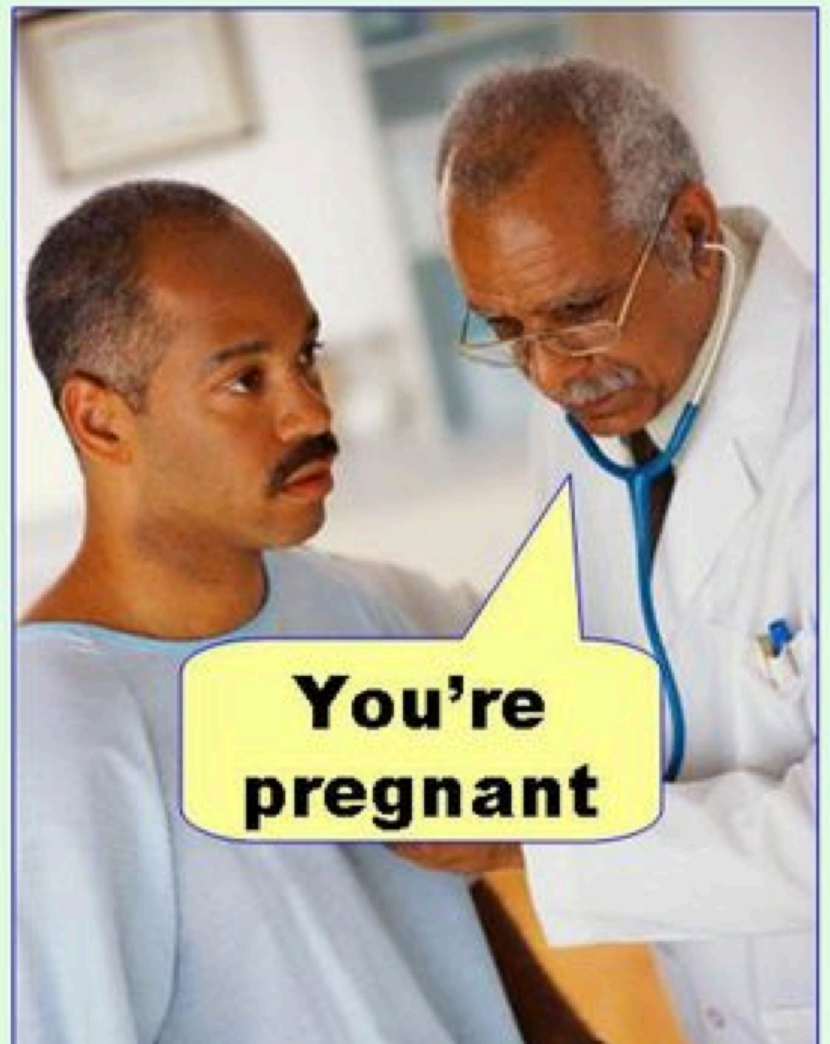
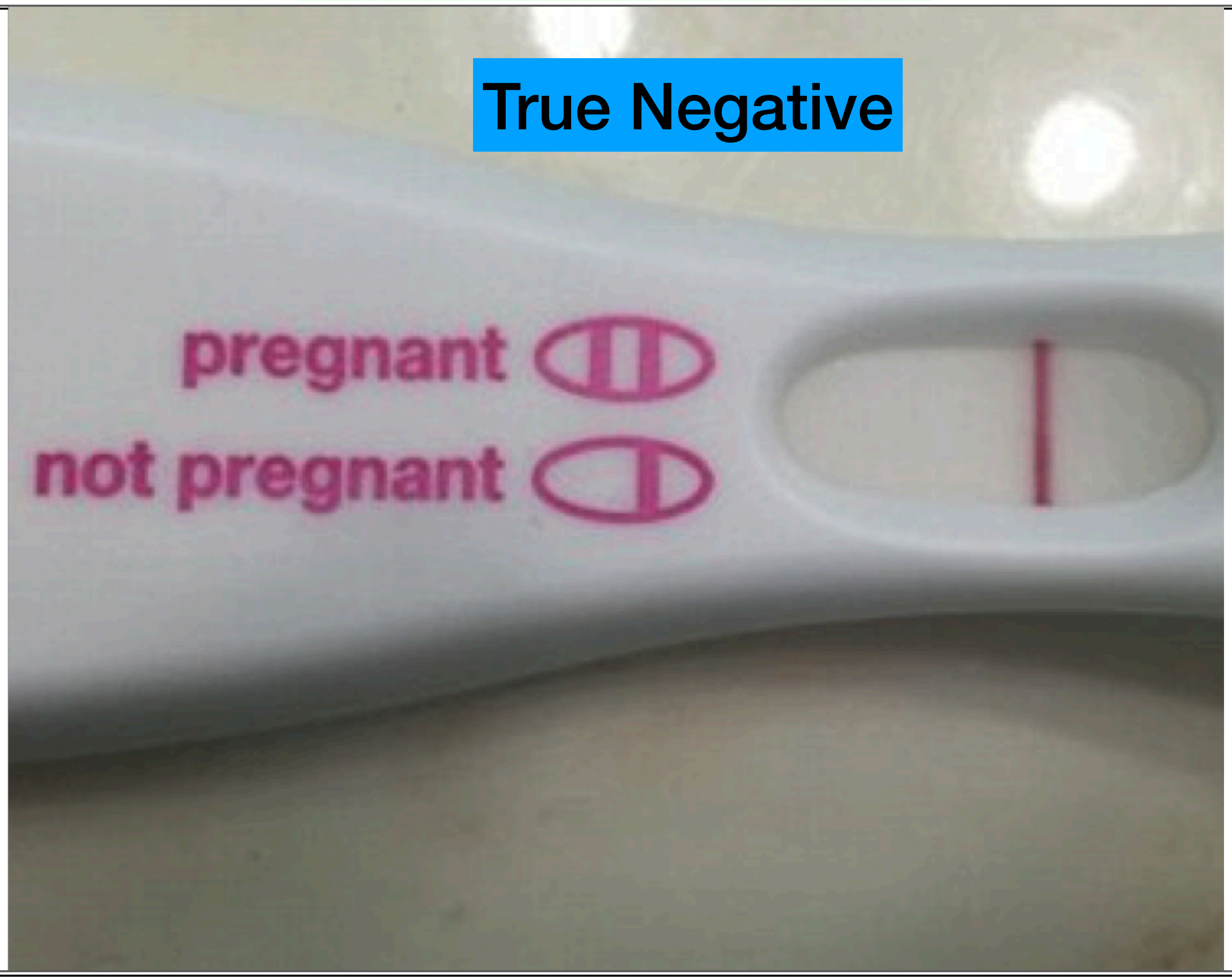


$$P(B,A) = P(B|A)P(A)$$

So…

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

$$\boxed{P(A|B) = \frac{P(B|A)P(A)}{P(B)} \quad \text{Baye's theorem}}$$

	Predicted +	Predicted -
Truly +		**Type II error** (false negative) 
Truly -	**Type I error** (false positive) 	

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

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

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

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

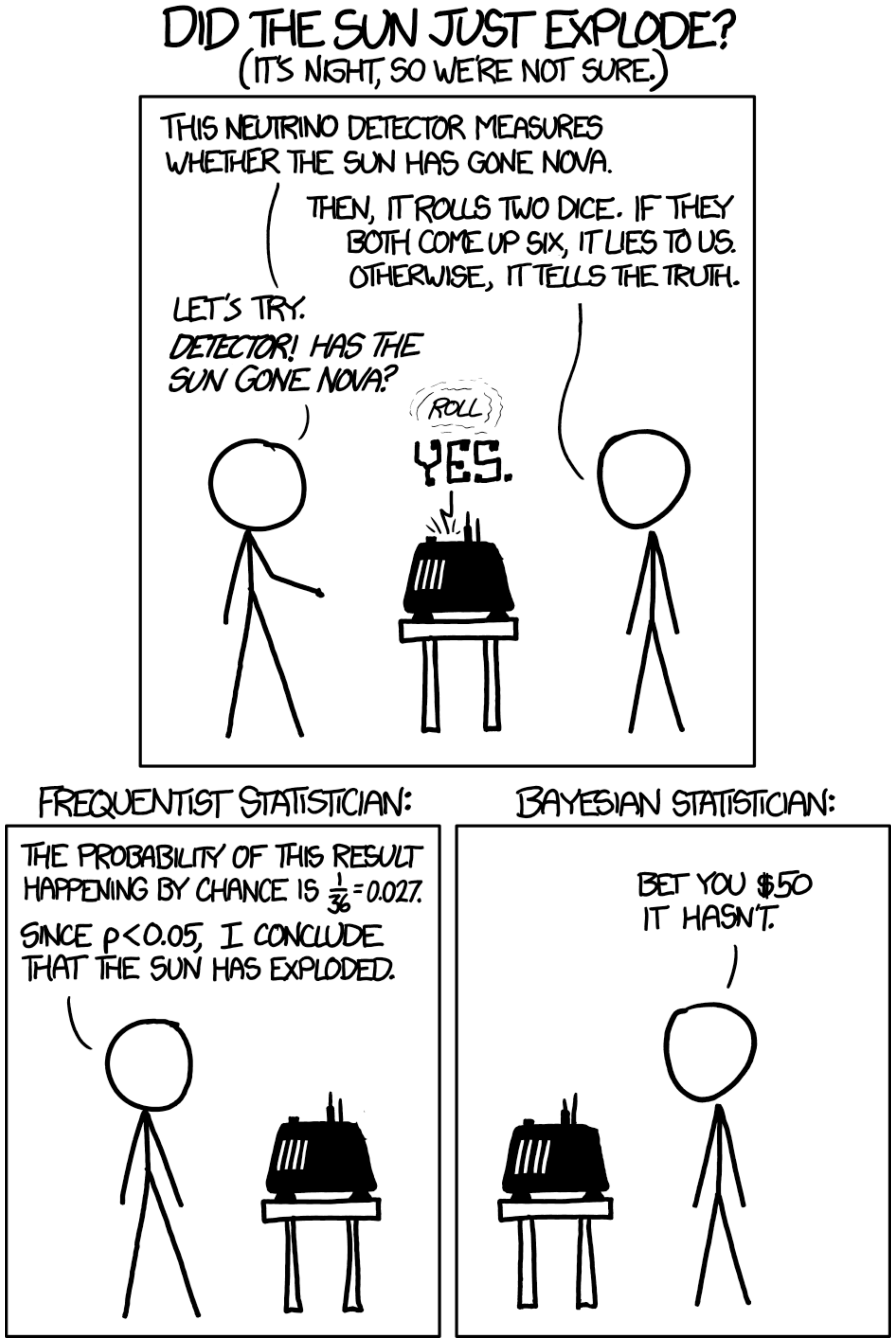
$$P(H_i \mid D) = \frac{P(D \mid H_i)P(H_i)}{\sum_{j=1}^{m} P(D \mid H_j)P(H_j)}$$

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
$P(D \mid H_i)$	Probability of observing data D under H_i
$P(H_i \mid D)$	Probability of H_i given the data



https://xkcd.com/1132

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

$$p(B|F) = \frac{p(F|B)p(B)}{p(F|B)p(B) + p(F|\overline{B})p(\overline{B})}$$

$$= \frac{\binom{50}{33}(0.6)^{33}(0.4)^{17}(1/100)}{\binom{50}{33}(0.6)^{33}(0.4)^{17}(1/100) + \binom{50}{33}(0.5)^{50}(99/100)}$$

$$= \frac{(0.6)^{33}(0.4)^{17}(1/100)}{(0.6)^{33}(0.4)^{17}(1/100) + (0.5)^{50}(99/100)}$$

$$= \frac{1.2^{33}0.8^{17}(0.01)}{0.99 + 1.2^{33}0.8^{17}(0.01)} \qquad \approx \frac{.09237}{.99 + .09237}$$

$$\approx 0.0853$$

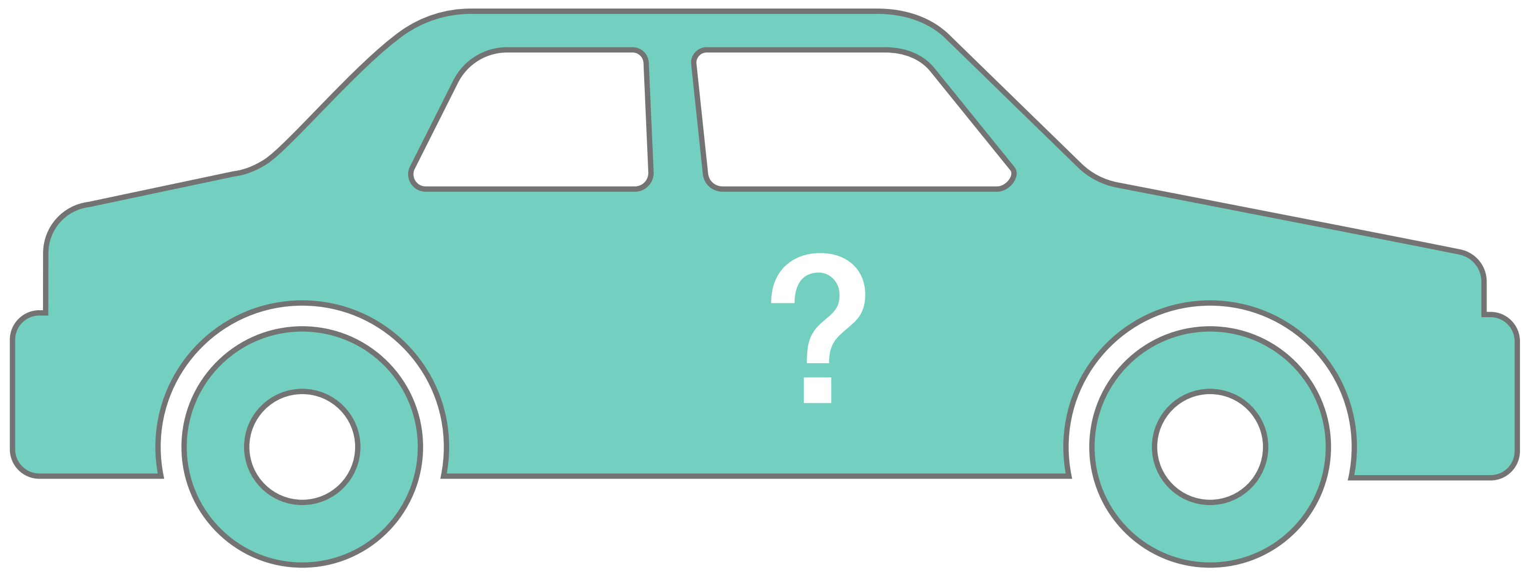
You decided that this coin is suspicious and flip it 50 more times, getting another 29 heads. If your priors are taken from the answer to the previous question, what are the new odds that the coin is biased?

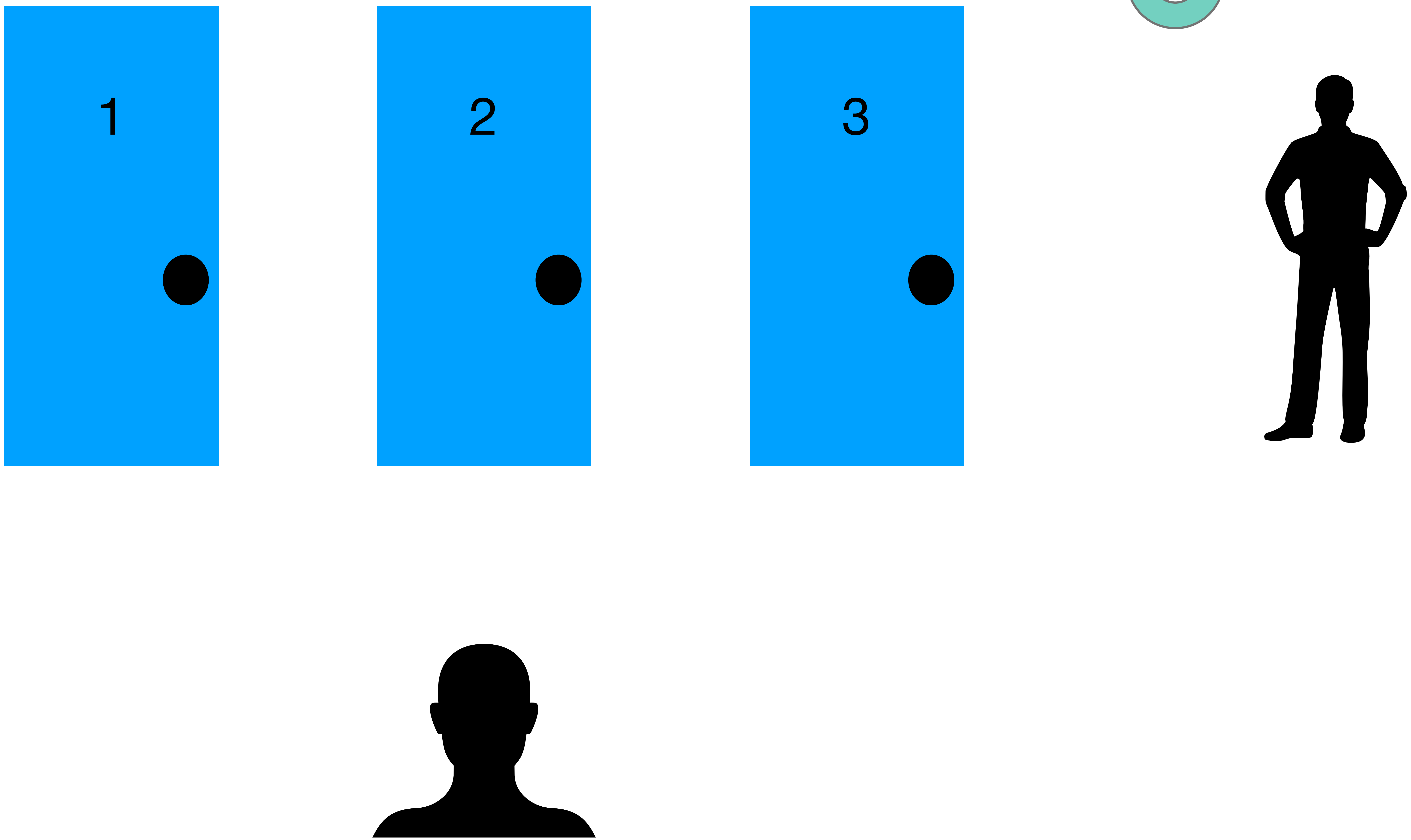
$$
\begin{aligned}
p(B|F) &= \frac{p(F|B)p(B)}{p(F|B)p(B) + p(F|\overline{B})p(\overline{B})} \\
&= \frac{1.2^{33}0.8^{17}(0.0853)}{0.915 + 1.2^{33}0.8^{17}(0.0853)} \\
&\approx \frac{.788}{.915 + .788} \\
&\approx .463
\end{aligned}
$$

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

The Monty Hall Problem

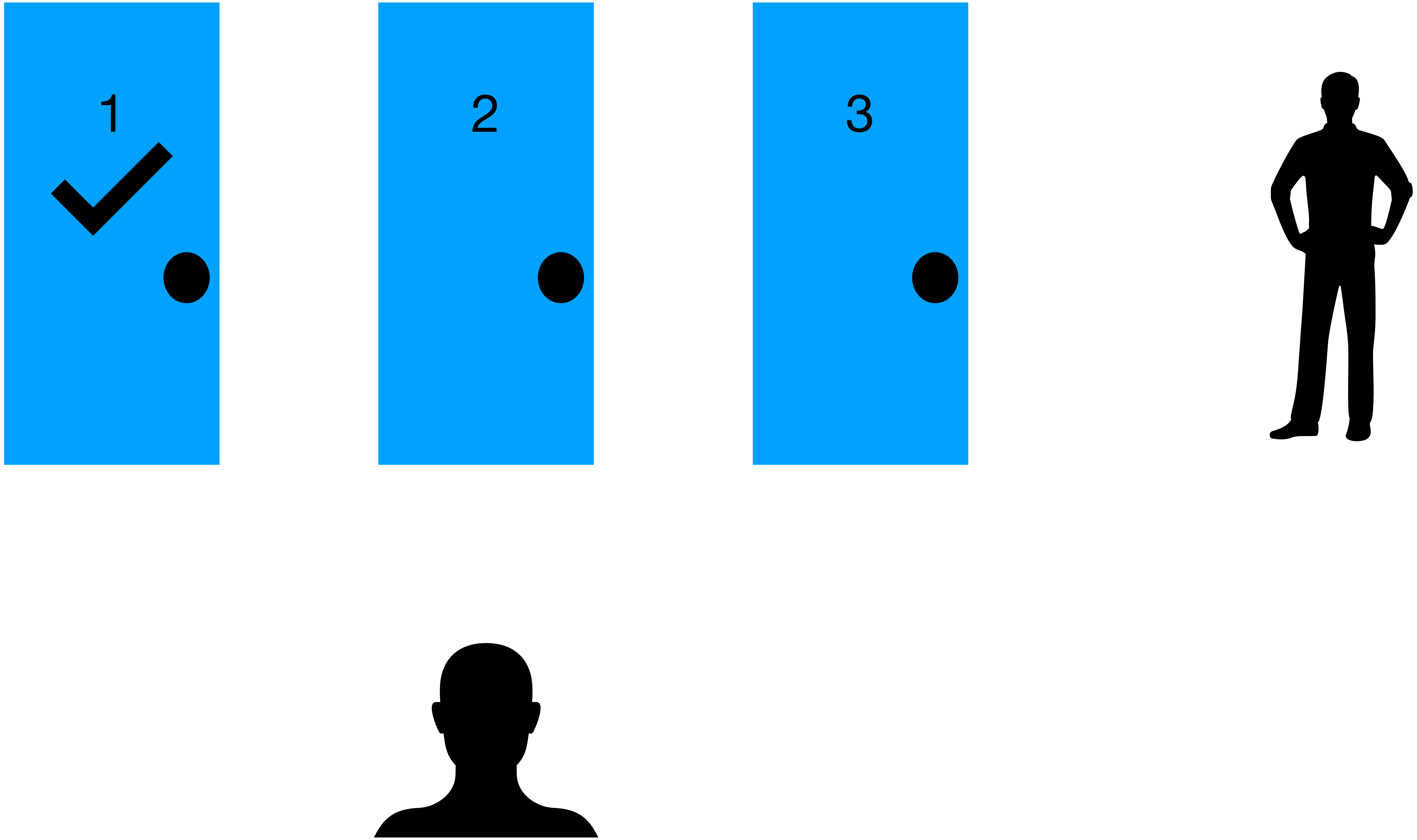
Let’s make a deal!





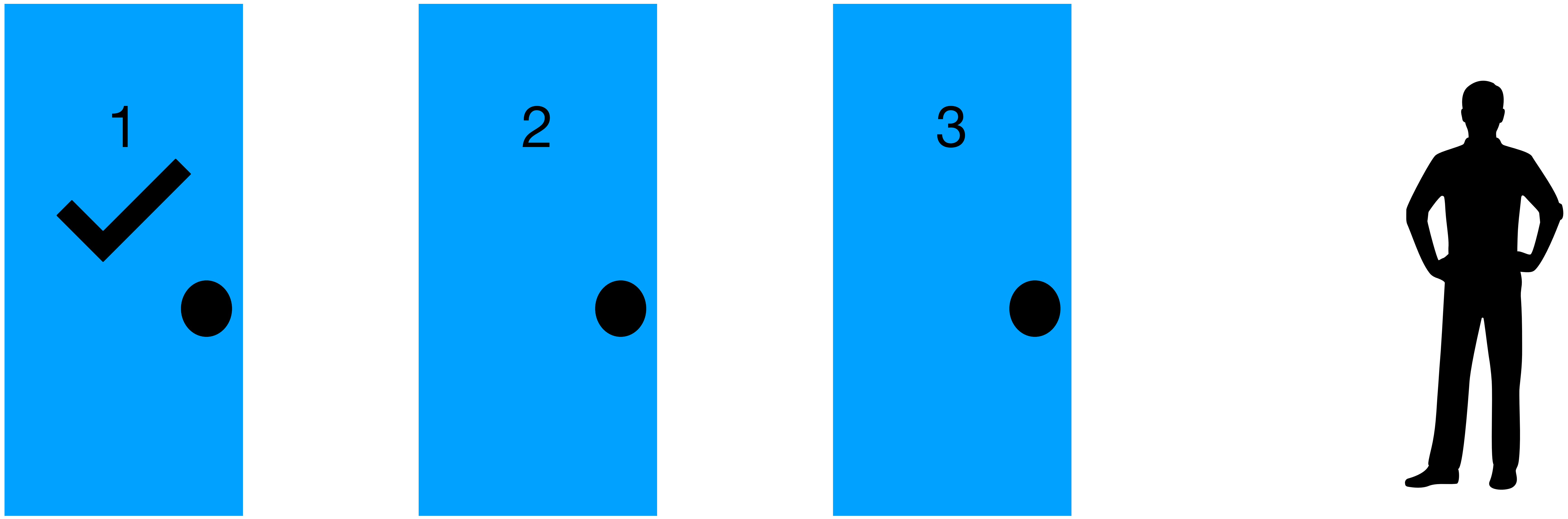
The Monty Hall Problem

Let’s make a deal!



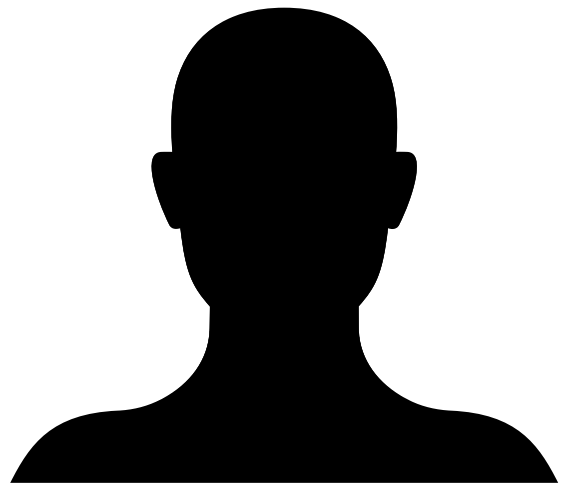
The Monty Hall Problem

Let's make a deal!



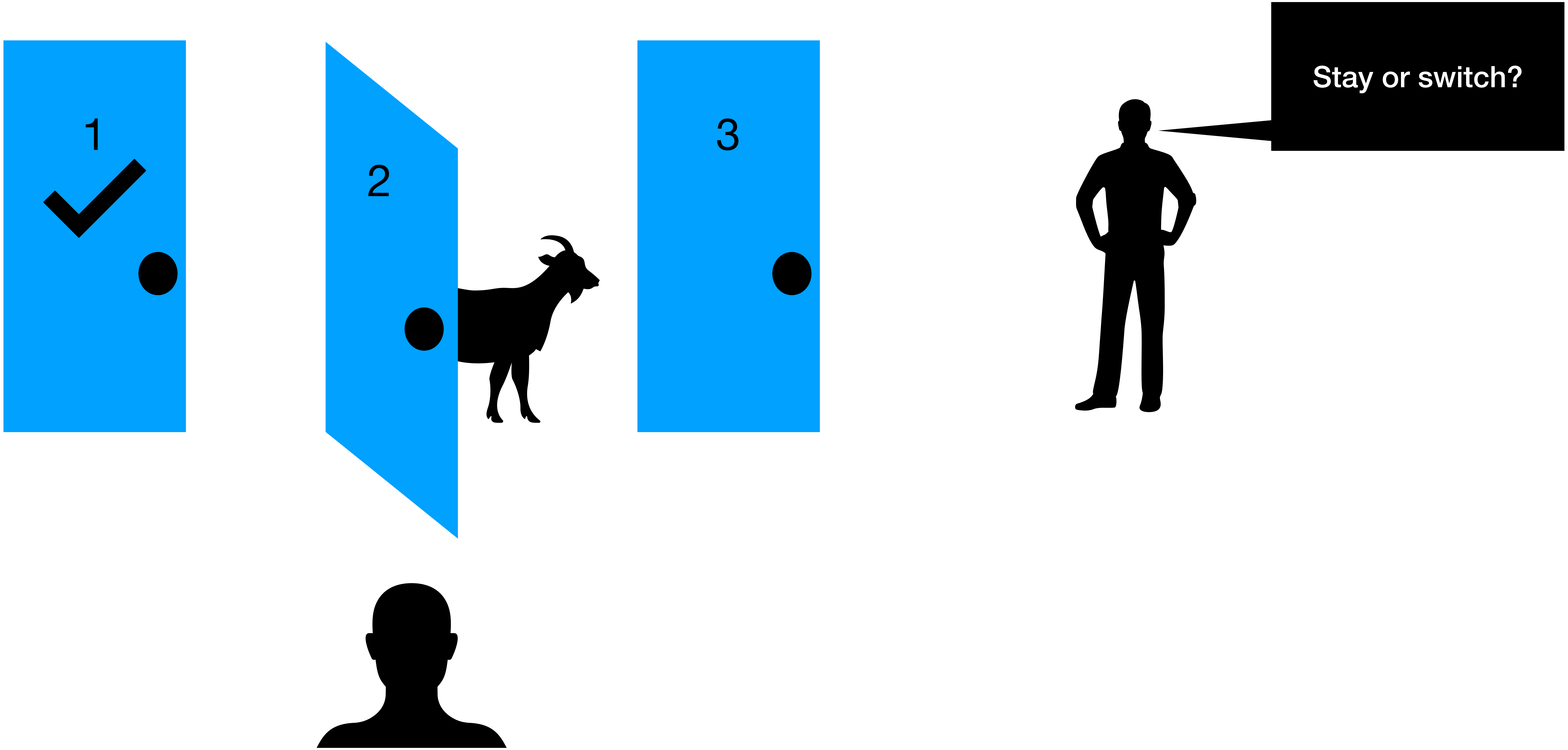
P(car behind D1)?

P(car behind D2 and D3)?



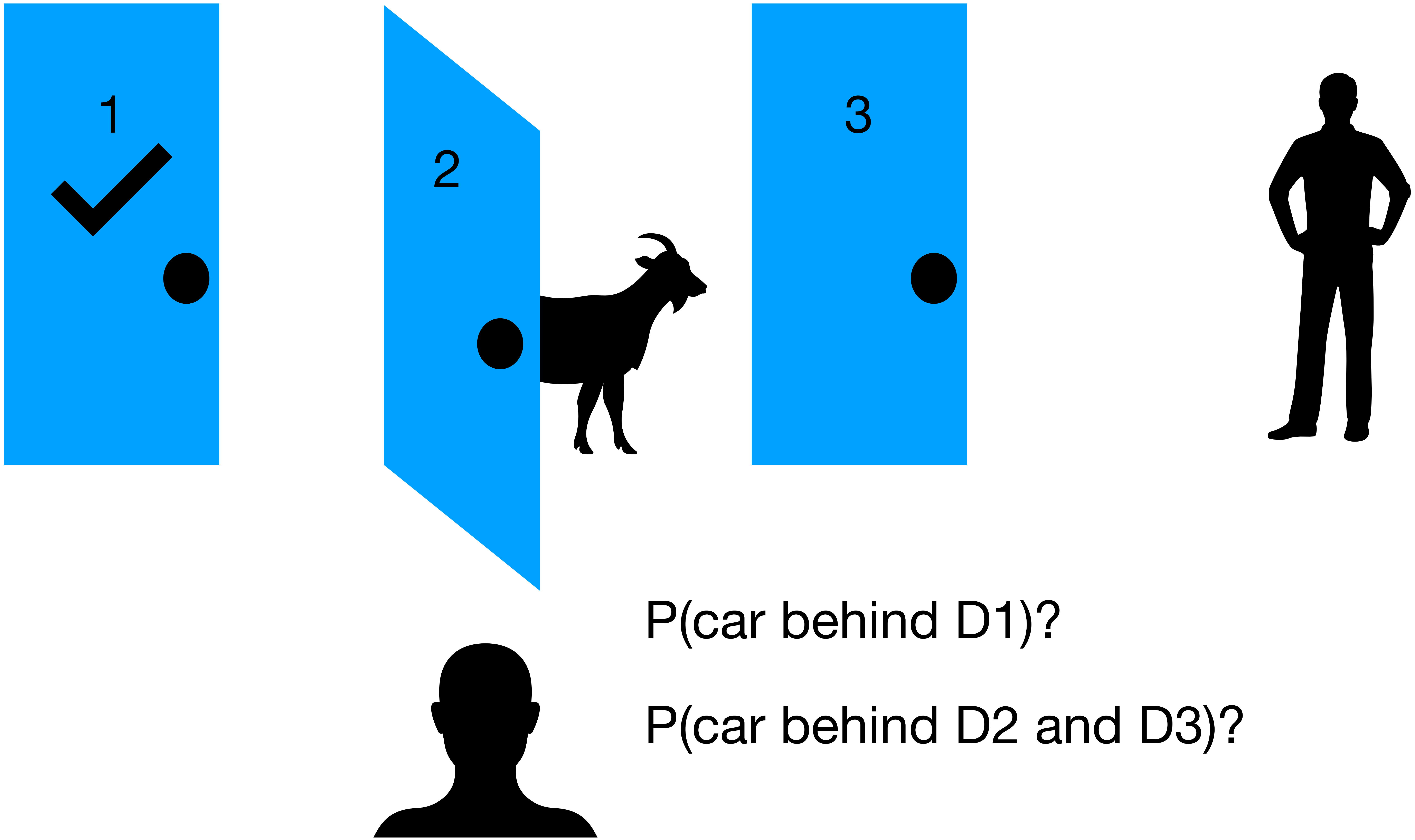
The Monty Hall Problem

Let's make a deal!



The Monty Hall Problem

Let's make a deal!

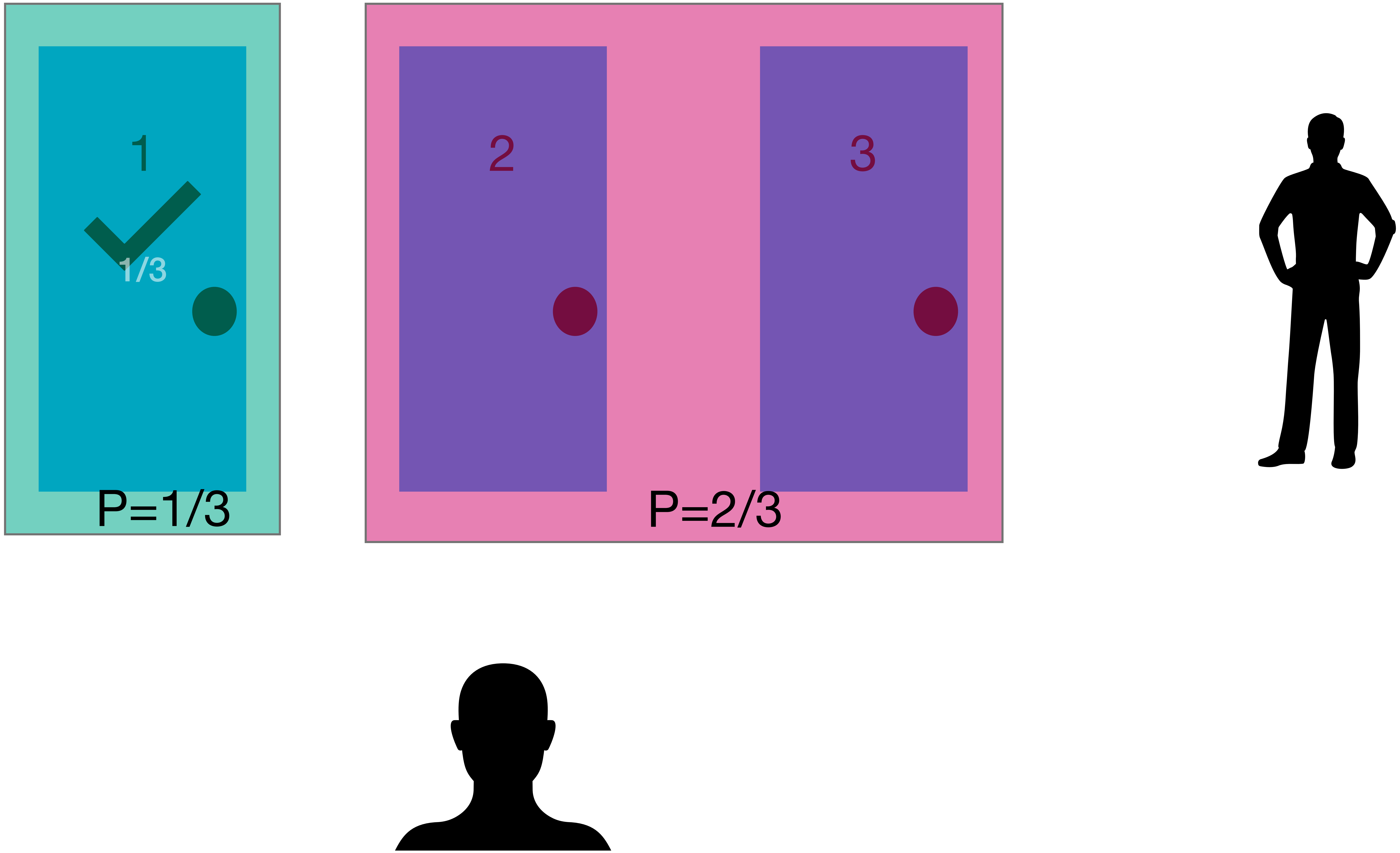


P(car behind D1)?

P(car behind D2 and D3)?

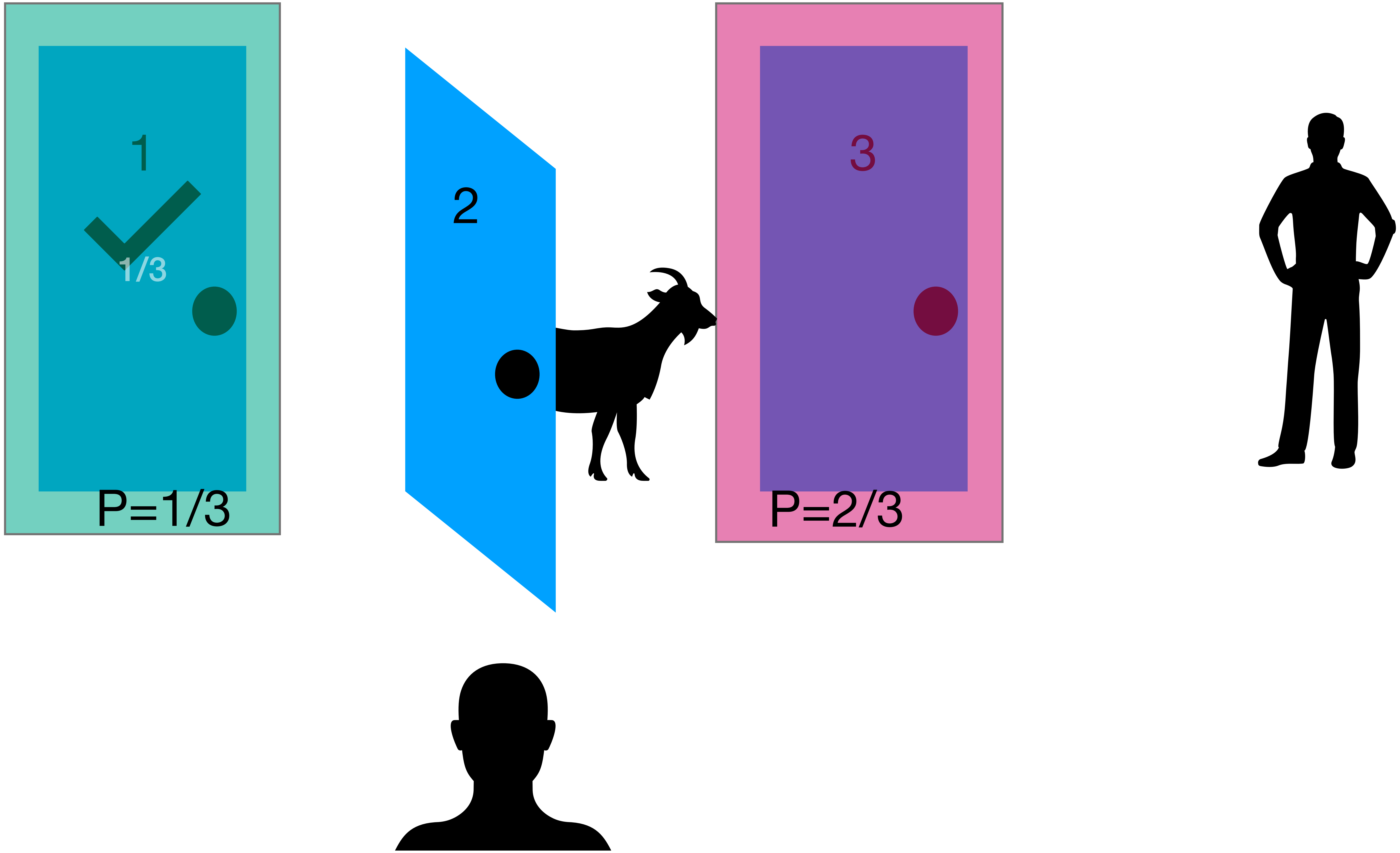
The Monty Hall Problem - Intuition

Let's make a deal!



The Monty Hall Problem - Intuition

Let's make a deal!



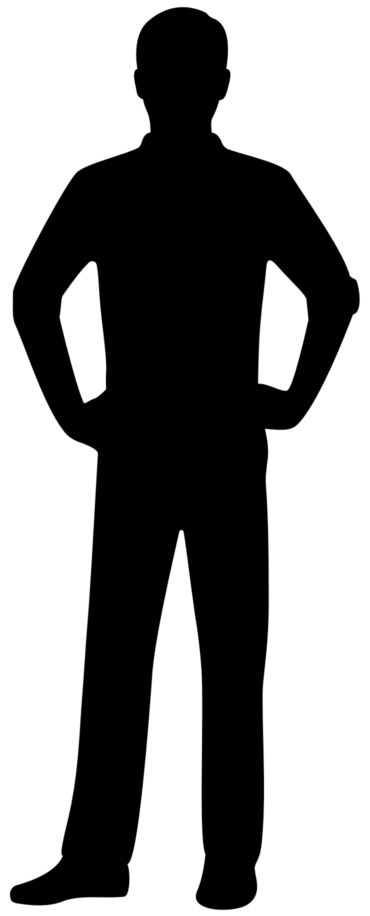
The Monty Hall Problem - Bayes math

Let's make a deal!

Want to calculate:

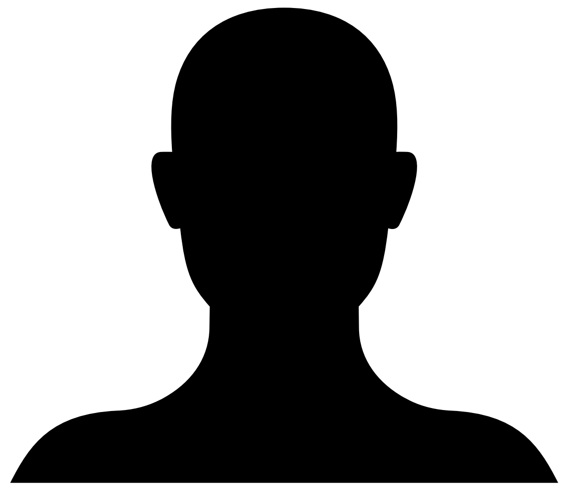
$$P(\text{car} = 3 \,|\, \text{opens} = 2, \text{choose} = 1)$$

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



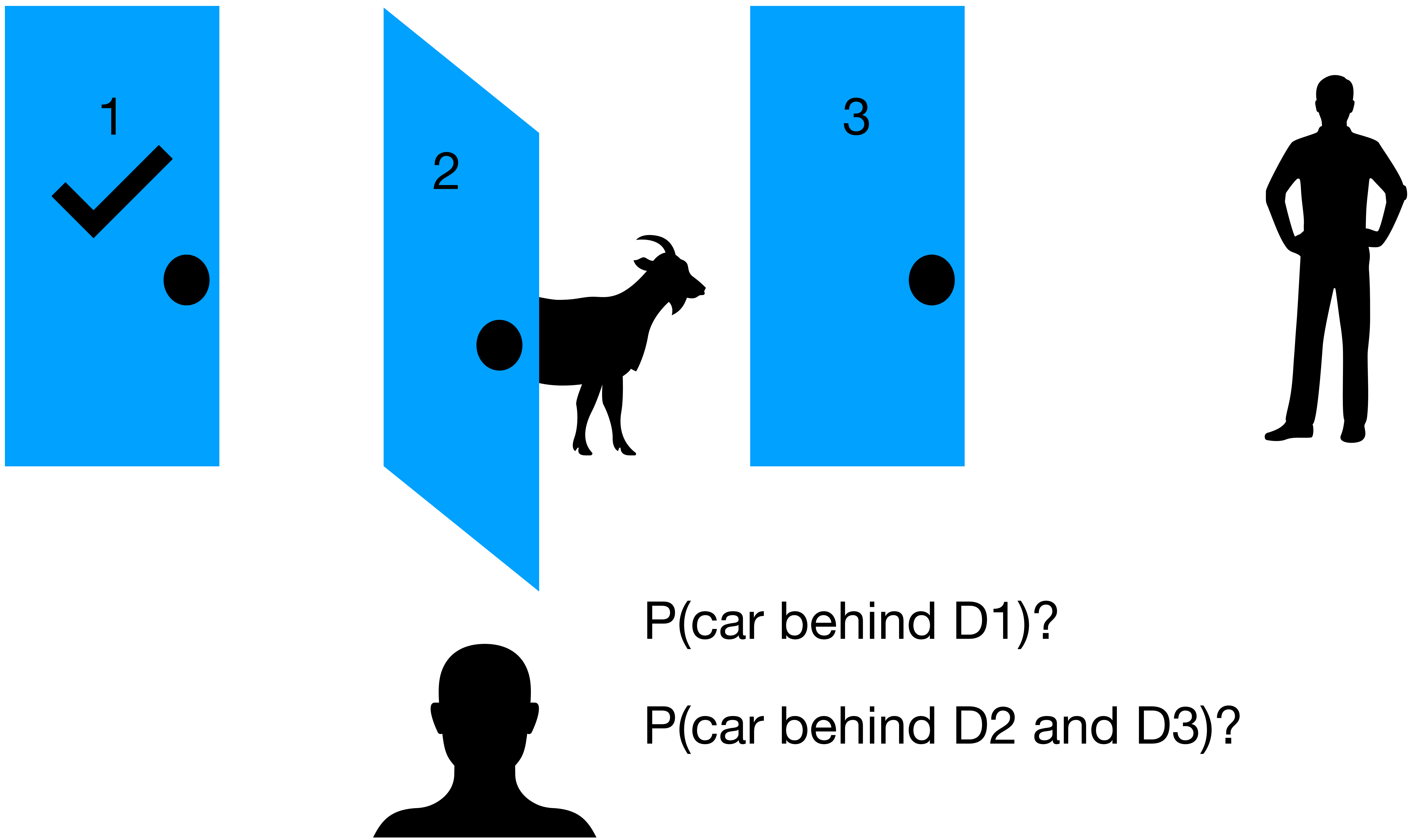
Left to you as an exercise…

Literally. On A1.



The Monty Hall Problem

Let’s make a deal!



P(car behind D1)?

P(car behind D2 and D3)?