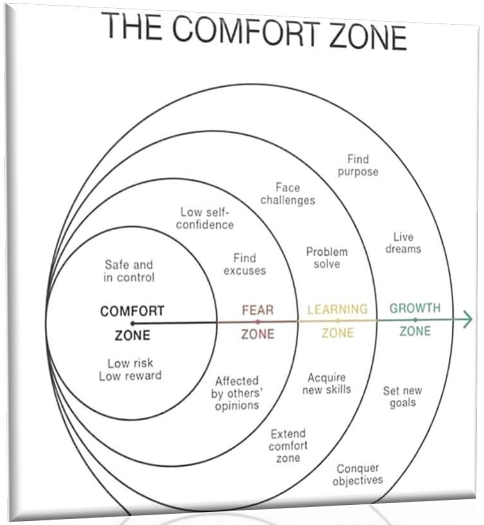
combra.cs.rutgers.edu



206
Discrete Structures II

Konstantinos P. Michmizos

Computational Brain Lab
Computer Science | Rutgers University | NJ, USA

Lecture 22 | Bayes Rule | Monday December 4th 2023

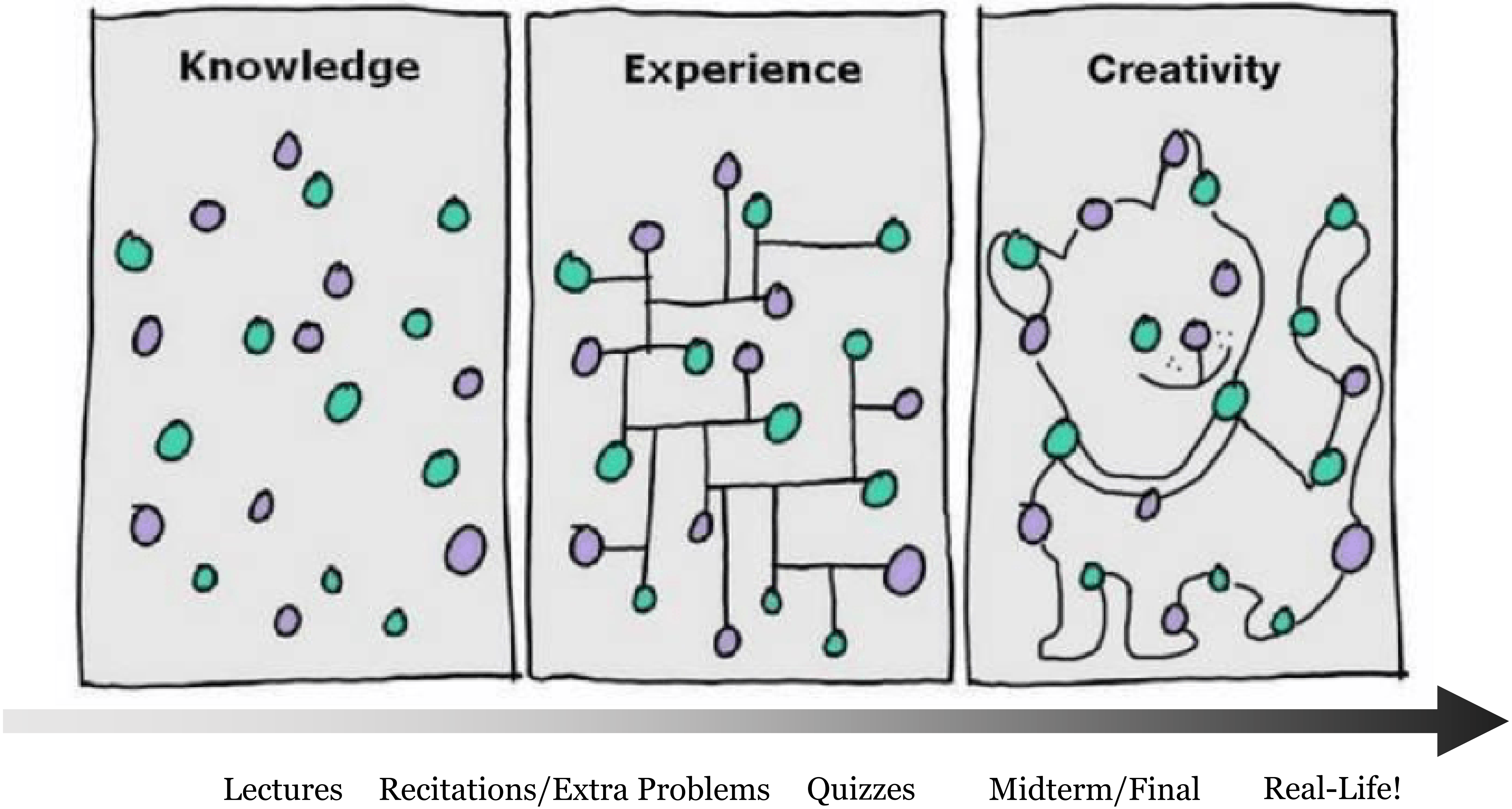
This is the... Last 4 Lectures

Part 1: Counting

- Proofs/induction
- Sum Rule
- Partition Method
- Difference Method
- Bijection Method
- Product Rule
- Generalized product rule
- Pigeonhole Principle
- Inclusion/Exclusion
- Combinatorial proofs, binomial coefficients

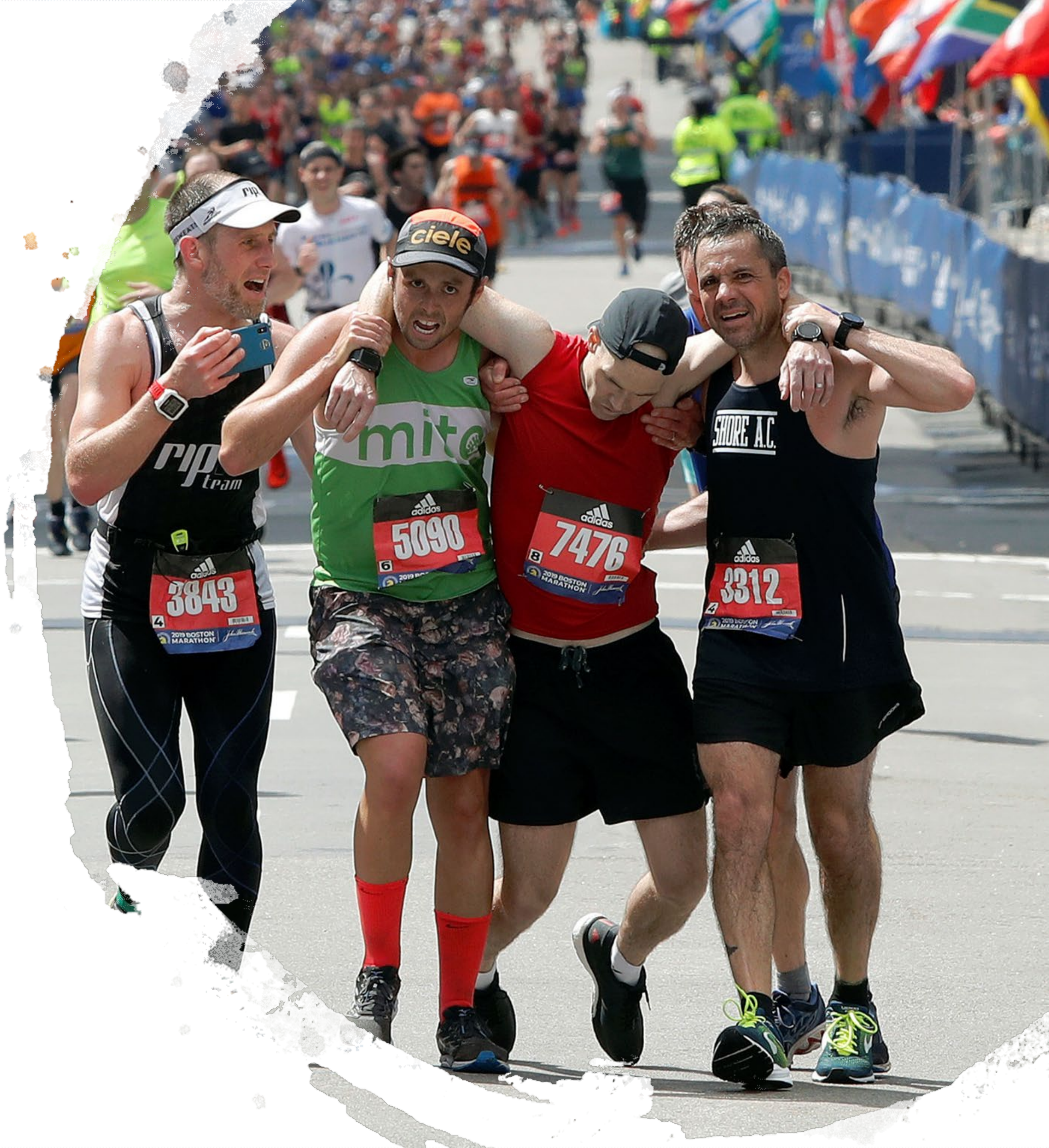
Part 2: Probability

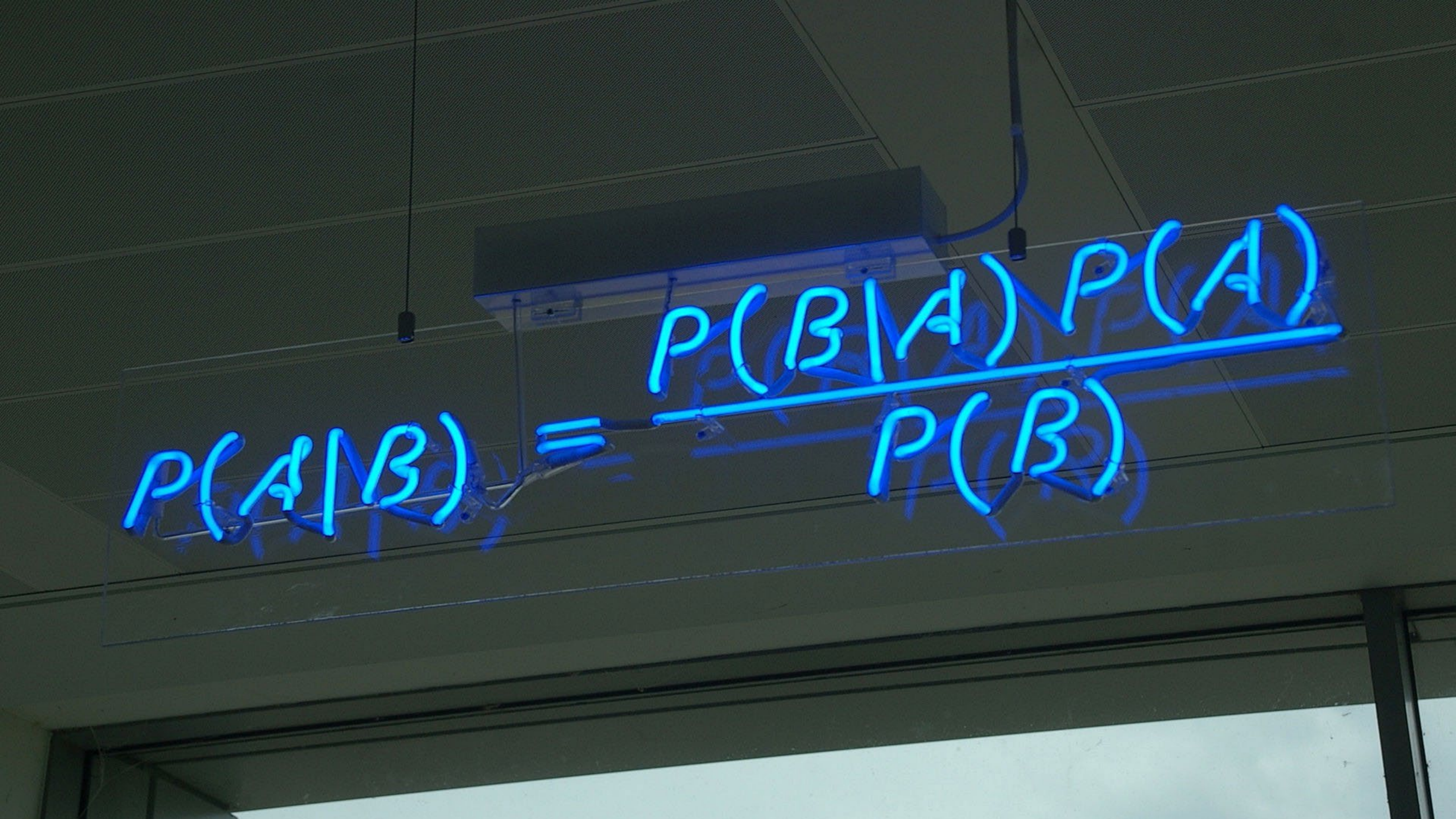
- Sample space
- Axioms of probability
- Conditional probability
- Independence
- Bayes rule
- Random Variables
- Expectation and Linearity
- Variance


Knowledge
Experience
Creativity
Lectures
Recitations/Extra Problems
Quizzes
Midterm/Final
Real-Life!

Announcements

- Quiz 5 is running
- Quiz 6 → This week
  - Conditional Probability
  - **Bayes rule**




P(A|B) = P(B|A) P(A) / P(B)

Conditional Probabilities

- P(B|A) means “Probability of event B **given** event A"

In other words, event A has already happened, now what is the chance of event B?

"Probability Of" *"Given"*

$$\mathrm{P}(\underbrace{A}_{\textit{Event A}} \text{ and } \underbrace{B}_{\textit{Event B}}) = \mathrm{P}(A) \times \mathrm{P}(B \mid A)$$

*"Probability of **event A and event B** equals the probability of **event A** times the probability of **event B given event A**"*

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

*"The probability of **event B given event A** equals the probability of **event A and event B** divided by the probability of **event A***

Independence

- A and B are independent events if $P(A|B) = P(A)$

$P(A|B) = P(A)$, A and B are independent

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

$$\Rightarrow P(A \cap B) = P(A) \cdot P(B)$$

Conditional Probabilities – Example #1

Two fair coins are flipped.

- $A = \{first\ coin\ is\ H\}$,
- $B = \{second\ coin\ is\ H\}$.

Are A and B independent?

$\Omega = \{(H,H), (T,T), (H,T), (T,H)\}$

$P(A|B) = P(A)$, $P(A) = \frac{|A|}{|\Omega|} = \frac{2}{4} = \frac{1}{2}$

$P(A|B)$: new sample space $B = \{(H,H), (T,H)\}$

$P(A|B) = \frac{1}{2} = P(A)$

Conditional Probabilities – Example #2

Two fair coins are flipped.

- $A = \{first\ coin\ is\ H\}$
- B = two coins have different outcomes.

Are A and B independent?

$P(A) = \frac{1}{2}$

$\Omega = \{(H,H),(H,T),(T,H),(T,T)\}$

$B = \{(H,T),(T,H)\}$

$P(A|B) = \frac{1}{2} = P(A)$

Independence extends to 3 events

- A_1, A_2, A_3 are independent events if $P(A_1), P(A_2), P(A_3)$ does not change by knowing any subset of the other.

$$P(A_1|A_2) = P(A_1)$$
$$P(A_1|A_3) = P(A_1)$$
$$P(A_1|A_2, A_3) = P(A_1)$$
$$\Rightarrow P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$$
$$P(A_1 \cap A_3) = P(A_1)P(A_3)$$
$$P(A_1 \cap A_2) = P(A_1)P(A_2)$$
$$\vdots$$

Independence extends to n events

- $A_1, A_2, A_3, \ldots A_n$ are independent events if $P(A_i)$ does not change by knowing any subset of the other.

Independence for n events

- $A_1, A_2, A_3, \ldots A_n$ are independent events if for all $k = 2,3, \ldots n$, and for all indices $i_1, i_2, \ldots i_k$
- $P\left(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\right) = P\left(A_{i_1}\right) P(A_{i_2}) \ldots P(A_{i_k})$

Conditional Probability

- Direct enumeration of sample space
- Tree based enumeration
- Direct use of formula
- Independence
- **Bayes Rule**

$P(A|B)$

Suppose computing $P(B|A)$ is much easier

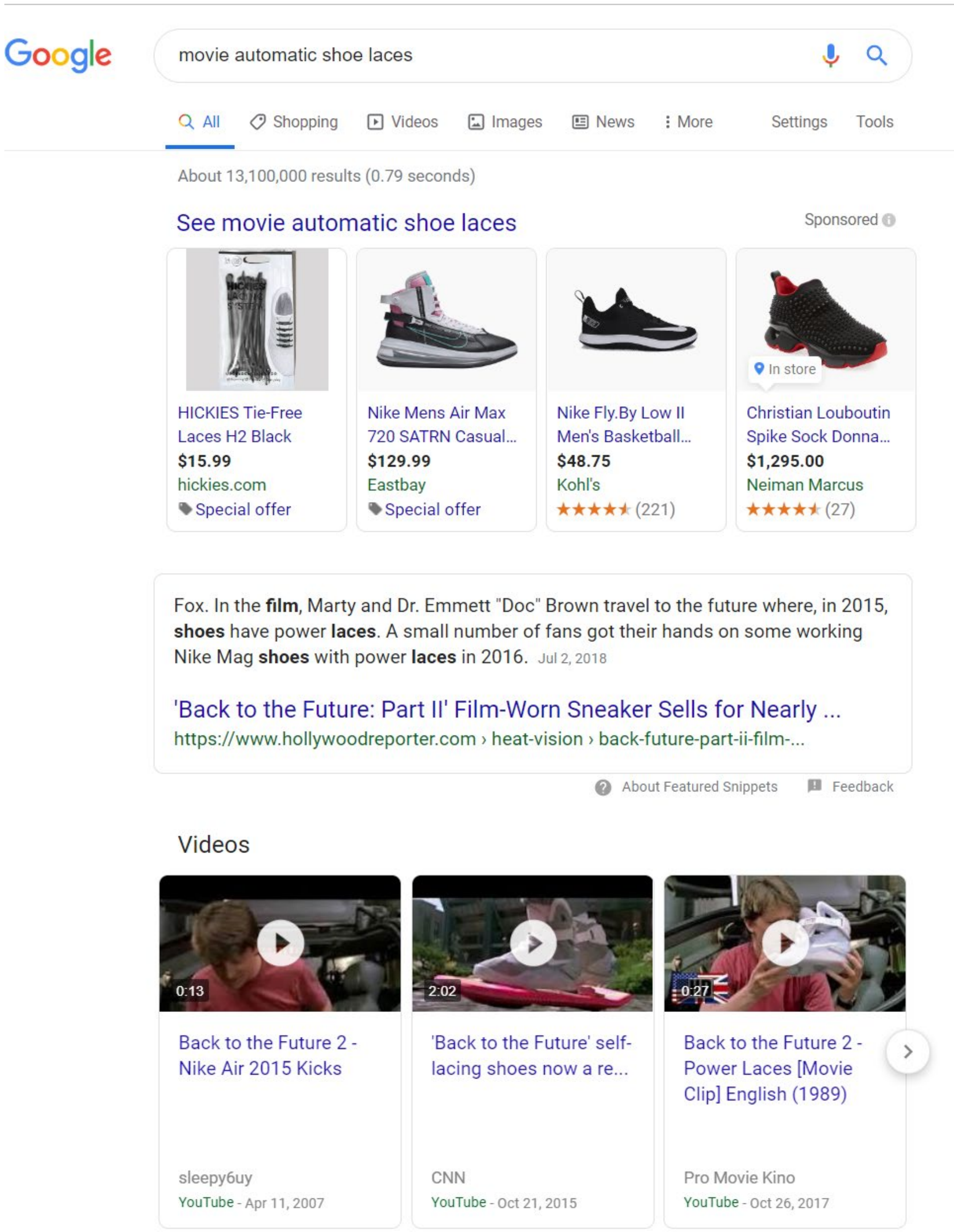
Bayes Rule

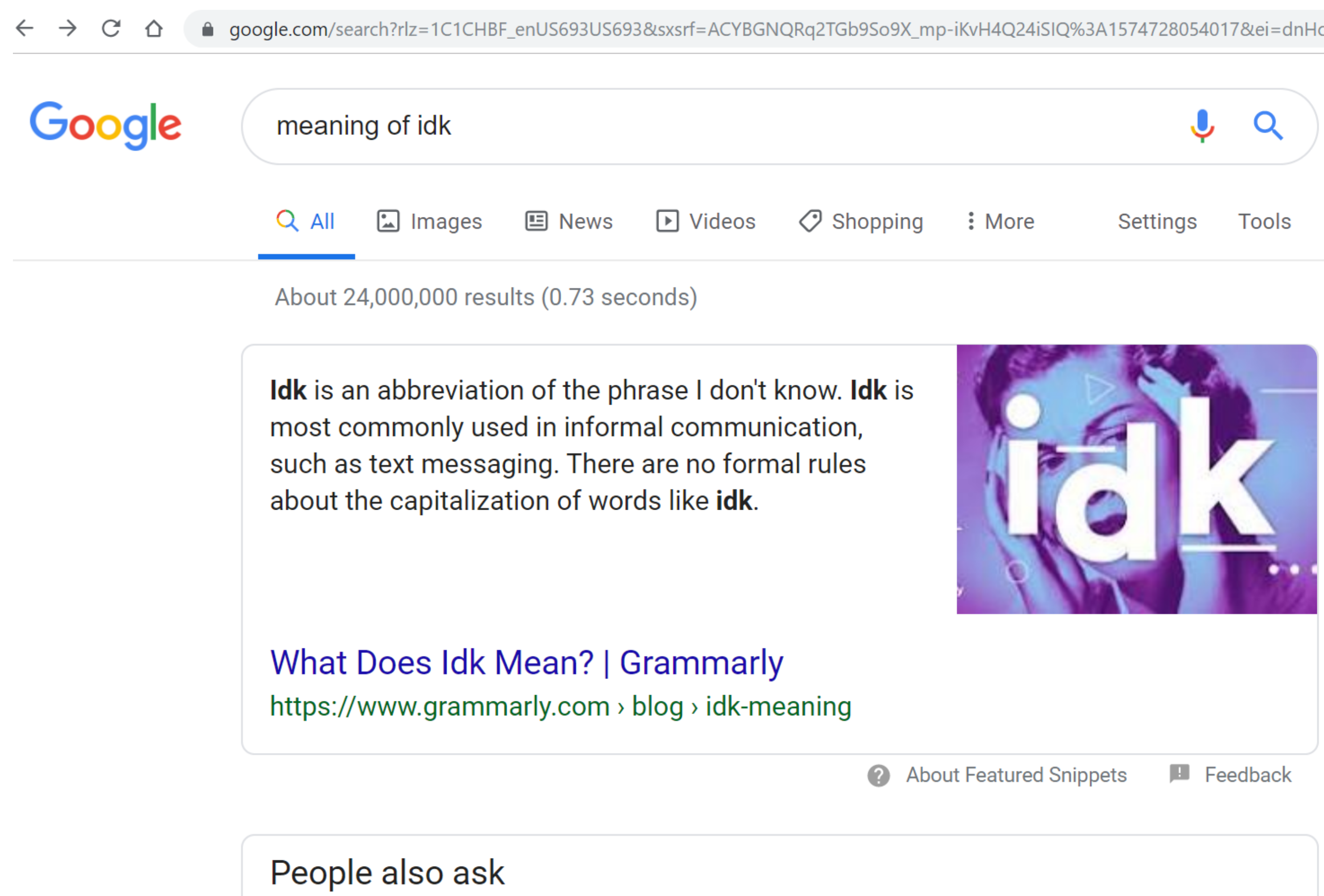
An internet search for "movie automatic shoe laces" brings up "Back to the future"

- Has the search engine watched the movie?

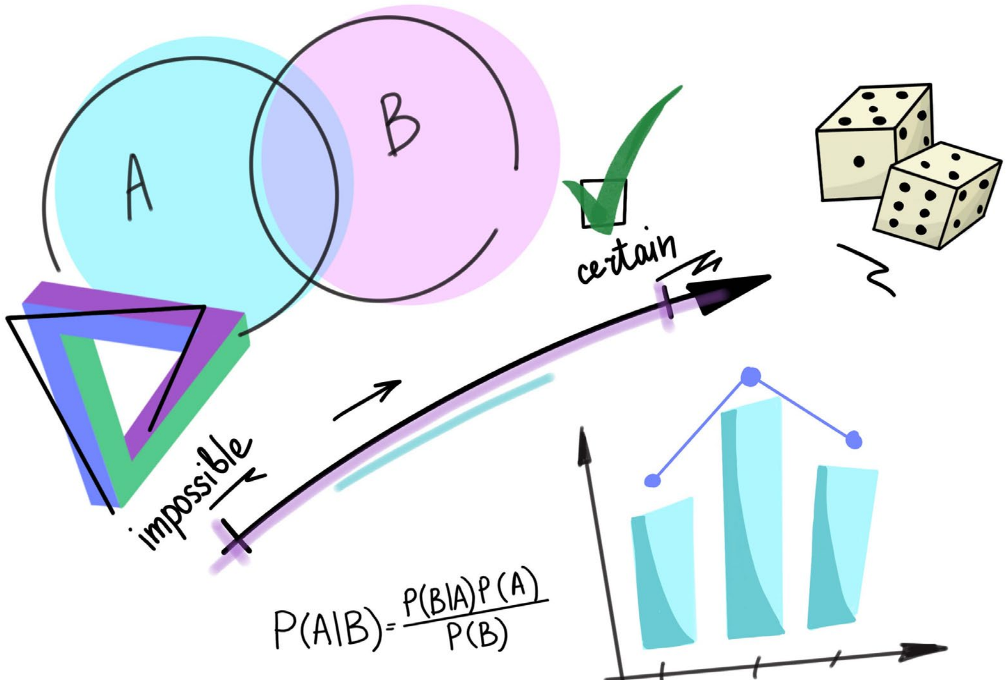
No, but it knows from lots of other searches what people are **probably** looking for.

- And it calculates that probability using Bayes' Theorem.





How does Google do this?



Previously: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

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

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

Bayes Rule

$P(A|B) \rightarrow$ want

$P(B|A) \rightarrow$ easily compute

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

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

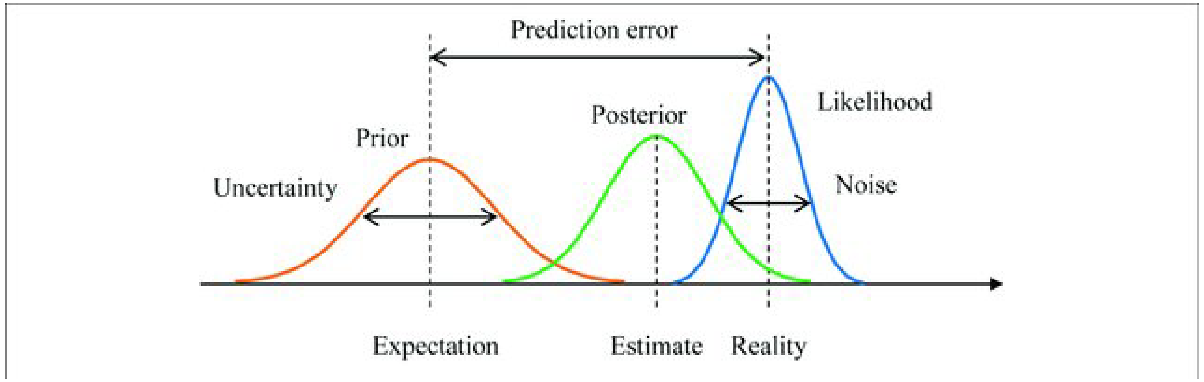
$$\Rightarrow \quad P(B)\,P(A|B) = P(A)\,P(B|A)$$

$$\Rightarrow \quad \boxed{P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)}} \rightarrow \text{Bayes Rule}$$

Bayesian Inference*

$$\overbrace{P(A|B)}^{\text{Posterior}} = \frac{\overbrace{P(B|A)}^{\text{Likelihood}} * \overbrace{P(A)}^{\text{Prior}}}{\underbrace{P(B)}_{\text{Evidence}}}$$

* *Inference* = *Educated Guess*



- Bayesian inference with a prior distribution, a posterior distribution, and a likelihood function.
- The prediction error is the difference between the prior expectation and the peak of the likelihood function (i.e., reality).
- Uncertainty is the variance of the prior. Noise is the variance of the likelihood function.

Example #1 – Bayes rule

- What is the probability of a couple having two girls given that they have at least one girl?

- P(2G|at least 1G) = {GG,GB,BB,BG}

$$= \frac{P(at\ least\ 1G \parallel 2G)\, P(2G)}{P(at\ least\ 1G)} =$$

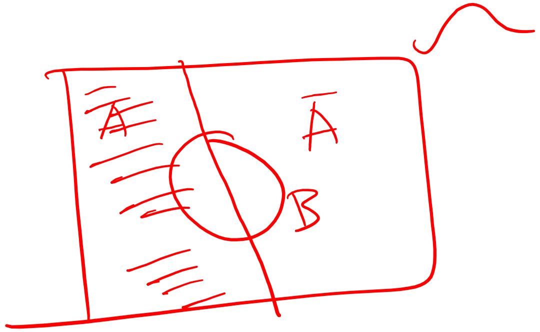
$$= \frac{1 \bullet 1/4}{3/4} = 1/3$$

Law of Total Probability

- Sometimes conditional probability are used to easily compute (un)conditional probabilities.

$P(B)$

Let A be an event

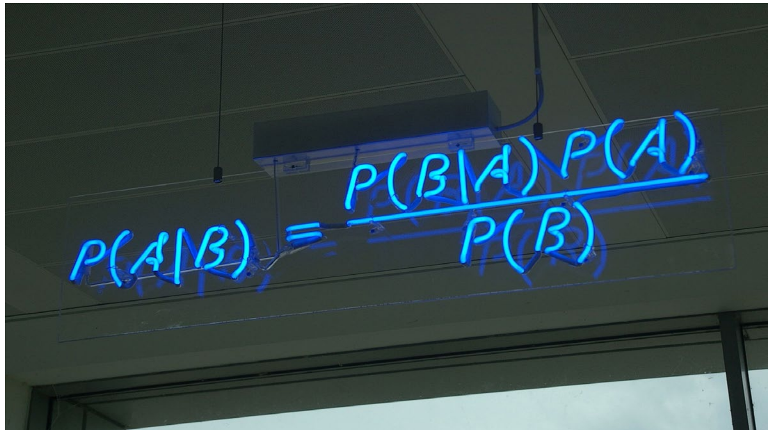


Suppose $P(B|A)$ and $P(B|\bar{A})$ is easy

$$P(B) = P(A) \cdot P(B|A) + P(\bar{A}) \cdot P(B|\bar{A})$$

$$= P(A) \cdot \frac{P(B \cap A)}{P(A)} + P(\bar{A}) \cdot \frac{P(B \cap \bar{A})}{P(\bar{A})}$$

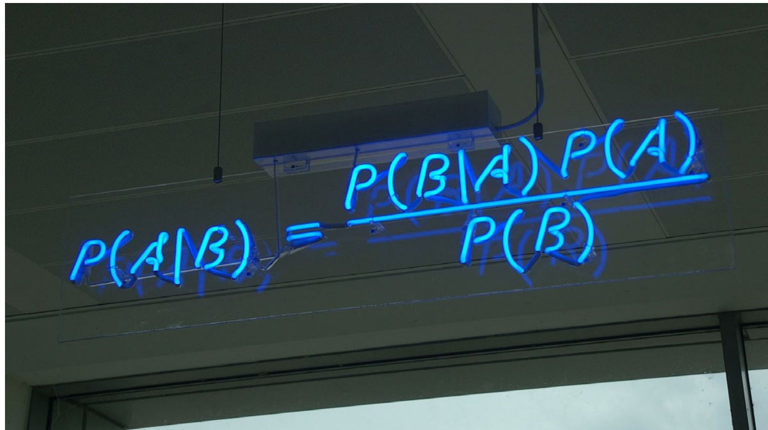
$$= P(B \cap A) + P(B \cap \bar{A})$$



Example #2 – Bayes rule

- If I randomly draw a cyan ball from two buckets that I can randomly choose from, B1 and B2, what is the probability of getting the cyan ball from the 1st bucket?

- Bayes Rule: $P(B1|A) = \dfrac{P(A|B1)\,P(B1)}{P(A)}$ = ???



Example #2 – Bayes rule

- If I randomly draw a cyan ball from two buckets that I can randomly choose from, B1 and B2, what is the probability of getting the cyan ball from the 1st bucket?

- $P(B1) = P(B2) = 0.5$
- A: select a cyan ball
- $P(A \mid B1) = 0.5$
- $P(A \mid B2) = 0.333$
- $P(A)=?$ → Remember: <u>B1 and B2 are disjoint! ! !</u>
  [**SUM RULE** → Disjoint sets →] $P(A) = P(A \cap B1) + P(A \cap B2) =$

$= P(A \mid B1)\, P(B) + P(A \mid B2)\, P(B2) = 0.5 \times 0.5 + 0.333 \times 0.5 = 5/12$

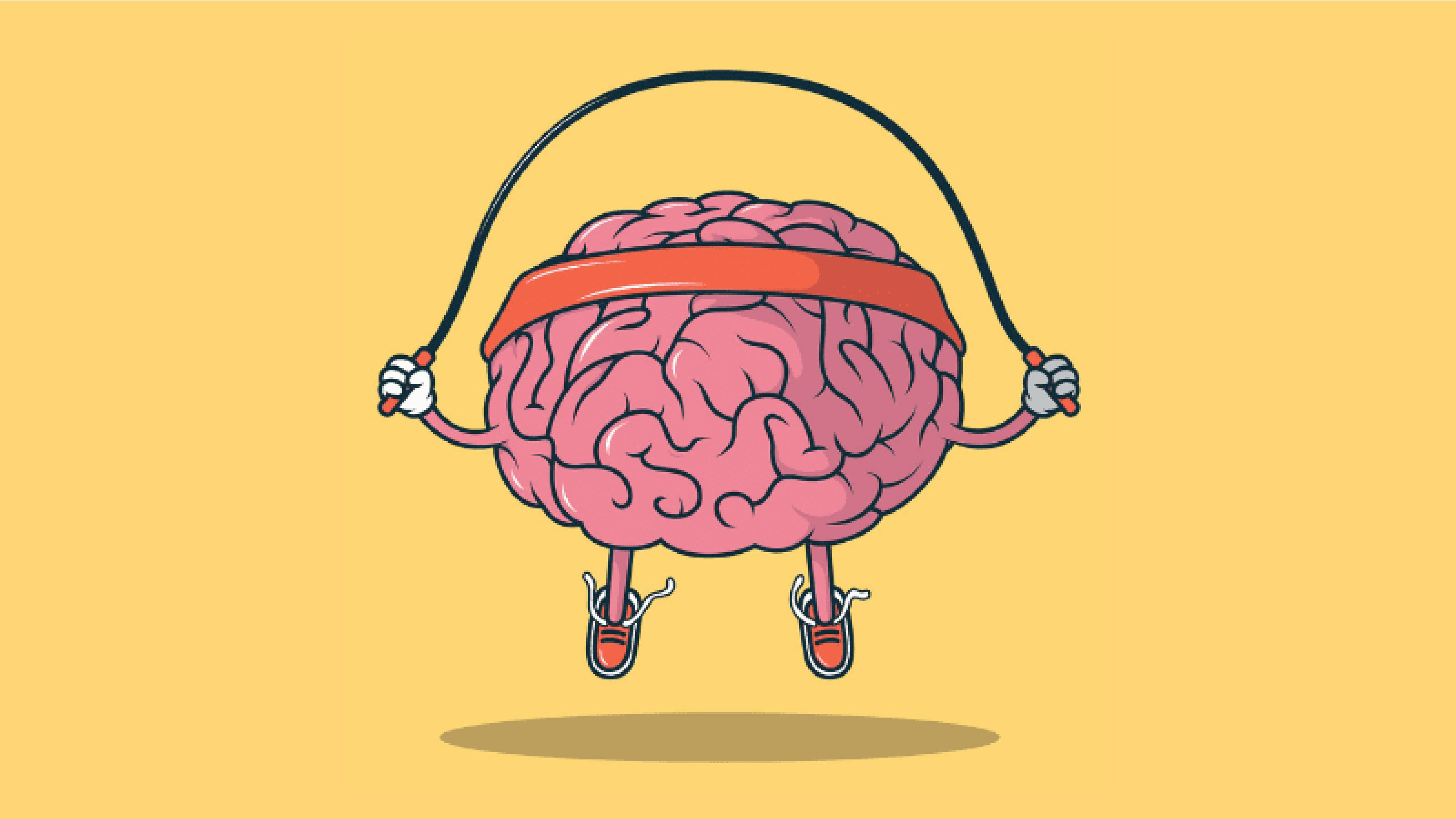
Example #2 – Bayes rule

- If I randomly draw a cyan ball from two buckets that I can randomly choose from, B1 and B2, what is the probability of getting the cyan ball from the 1st bucket?

- Bayes Rule: $P(B1|A) = \dfrac{P(A|B1)\,P(B1)}{P(A)} = \dfrac{0.5 \times 0.5}{5/12} = 3/5.$

- *Does the result make sense???*

Brain Break – 1 min



CTAAR survey

https://sirs.rutgers.edu/blue

CTAAR Survey

Section 1

Section 2

Section 3

Example #3 – Bayes rule

- 4 cards are drawn from a randomly shuffled deck of 52 cards. What is the probability that at least 2 Aces are drawn, given that at least one card is an Ace?

$A \to$ at least 2 Aces

$B \to$ at least 1 Ace

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

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\binom{52}{4} - \binom{48}{4} - 4\binom{48}{3}}{\binom{52}{4}}$$

$$P(B) = \frac{|B|}{|\Omega|} = \frac{\binom{52}{4} - \binom{48}{4}}{\binom{52}{4}}$$

Example #4 - Bayes Rule

- James has two coins in his hand. One is a real coin and the second one is a faulty one with Tales on both sides. He blind folds himself choose a random coin and tosses it in the air. The coin falls down with Tale facing upwards.

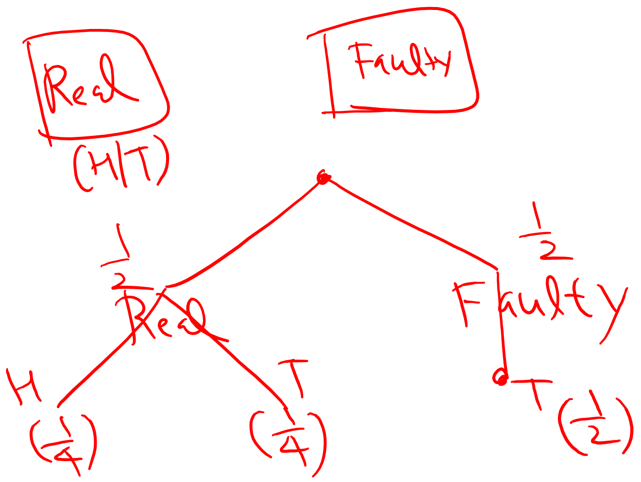
  What is the probability that this is the faulty coin?

B → outcome is Tale

A → coin chosen was faulty

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

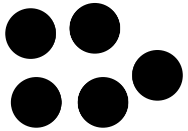
$P(A) = \frac{1}{2}$, $P(B) = \frac{3}{4}$



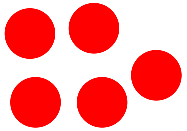
Law of Total Probability – Example #2

- You’ve been captured by pirates on an island.
- Need to play the following game to survive

100 black rocks



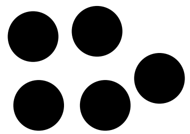
100 red rocks



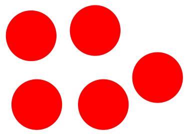
- Divide the rocks among two bags as you wish.
- Toss a fair coin and depending on the outcome draw a rock at random from the corresponding bag.
- If rock is black you win!

Law of Total Probability – Example #2

100 black rocks



100 red rocks



X Black rocks
Y Red rocks

100−X Black rocks
100−Y Red rocks

P(black rock is selected)

$$P(\text{black rock} \mid \text{Coin}=H) = \frac{X}{X+Y}$$

$$P(\text{black rock} \mid \text{coin}=T) = \frac{100-X}{200-X-Y}$$

$$P(\text{black rock is selected}) = P(\text{Coin}=H)\,P(\text{Black rock} \mid H) + P(\text{Coin}=T)\,P(\text{Black} \mid T)$$

$$= \left(\frac{1}{2}\right)\left(\frac{X}{X+Y}\right) + \left(\frac{1}{2}\right)\left(\frac{100-X}{200-X-Y}\right)$$

Bayes Rule – One more example

Find the probability for "when there is smoke, there is fire"

Example: If dangerous fires are rare (1%) but smoke is fairly common (10%) due to barbecues, and 90% of dangerous fires make smoke then:

$$P(\text{Fire}|\text{Smoke}) = \frac{P(\text{Fire})\, P(\text{Smoke}|\text{Fire})}{P(\text{Smoke})}$$

$$= \frac{1\% \times 90\%}{10\%}$$

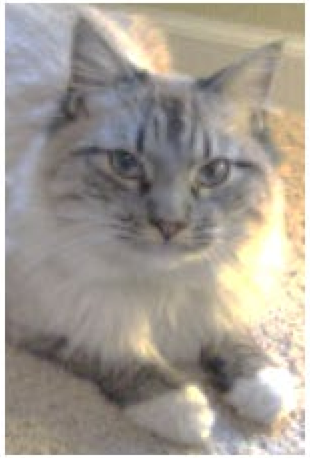
$$= 9\%$$

So the "Probability of dangerous Fire when there is Smoke" is 9%

Bayes Rule – Yet another example

Example: Allergy or Not?

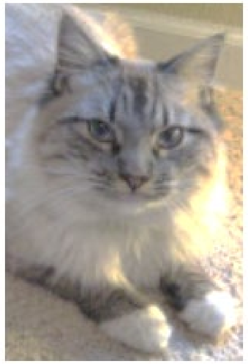
Hunter says she is itchy. There is a test for Allergy to Cats, but this test is not always right:



- For people that **really do** have the allergy, the test says "Yes" **80%** of the time
- For people that **do not** have the allergy, the test says "Yes" **10%** of the time ("false positive")

If 1% of the population have the allergy, and **Hunter's test says "Yes"**, what are the chances that Hunter really has the allergy?

Bayes Rule – Yet another example



$$P(\text{Allergy}|\text{Yes}) = \frac{P(\text{Allergy})\ P(\text{Yes}|\text{Allergy})}{P(\text{Yes})}$$

P(Allergy) is Prob of Allergy = 1%

P(Yes|Allergy) is Prob of test saying "Yes" for people with allergy = 80%

P(Yes) is Prob of test saying "Yes" (to anyone) = ??%

- We **don't know** what the **general** chance of the test saying "Yes" is but we can calculate it by adding up those **with**, and those **without** the allergy:
- 1% have the allergy, and the test says "Yes" to 80% of them
- 99% do **not** have the allergy and the test says "Yes" to 10% of them

P(Yes) = 1% × 80% + 99% × 10% = 10.7% of the population.

P(Allergy|Yes) = about **7%**