Just Bayes

INFO 6105

6 February 2023

What does "*Bayesian inference*" even mean?

Inference = Educated guessing

Thomas Bayes = A nonconformist Presbyterian minister in London back when the United States were still The Colonies

He also wrote an essay on probability. His friend Richard Price edited and published it after he died

Bayesian inference = Guessing in the style of Bayes

An Example

Dilemma at the movies

This person dropped their ticket in the hallway, at the bathroom line.

Do you call out

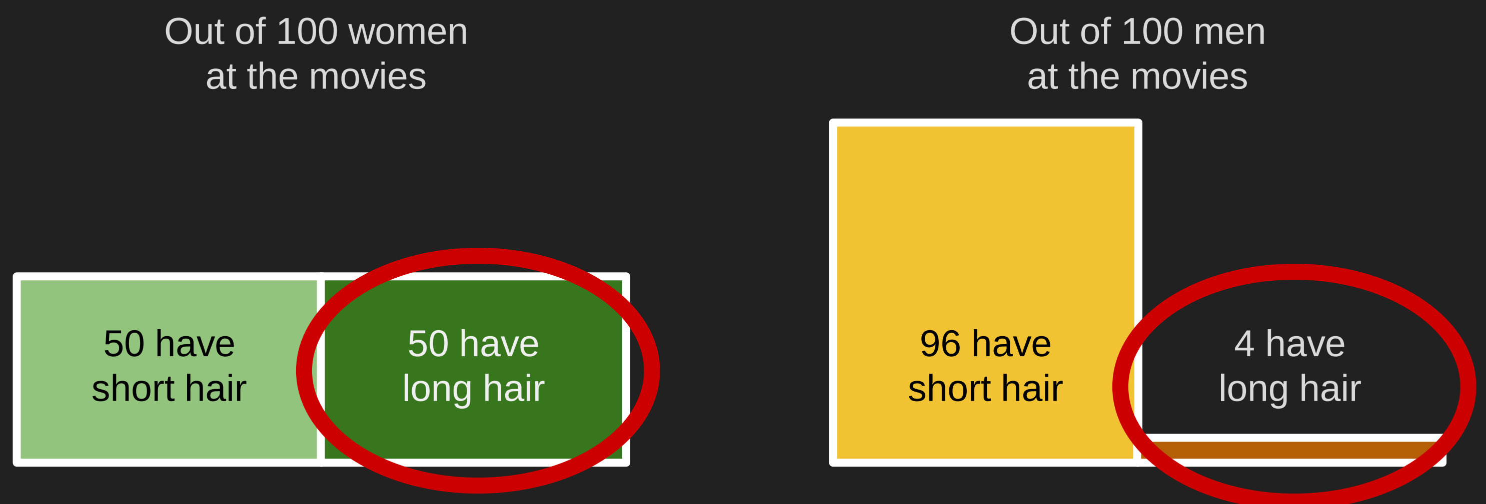
> “*Excuse me, ma’am!*”

or

> “*Excuse me, sir!*”

You have to make a ***guess***

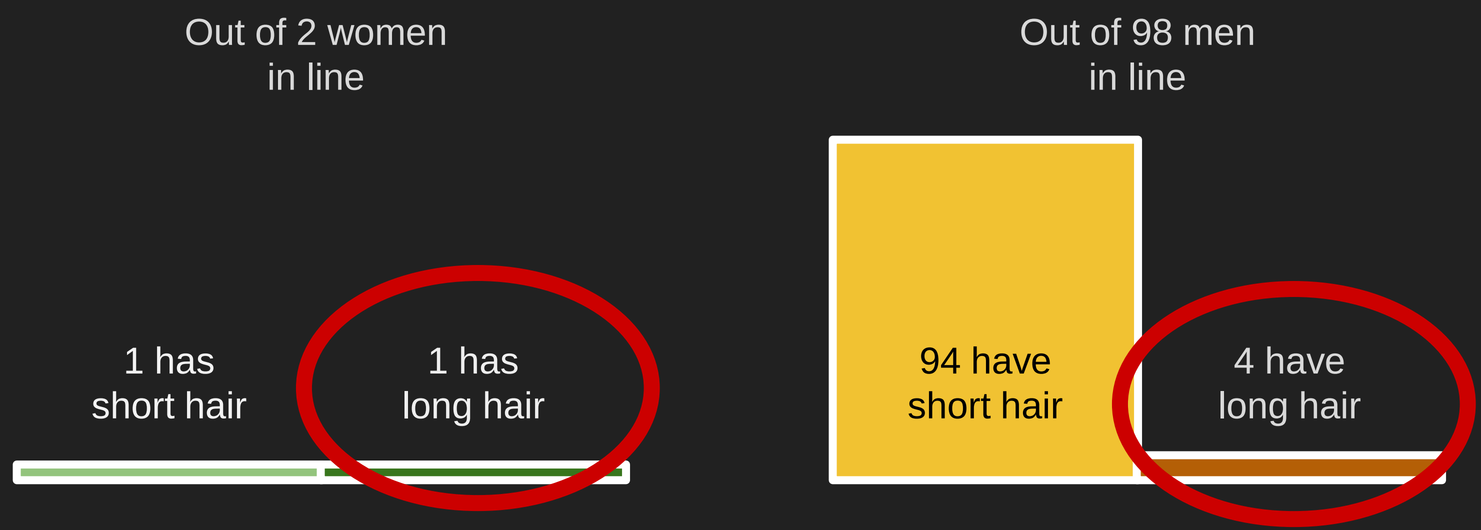
Put numbers to our dilemma: *At the movies*



About 12 times more women have long hair than men, so….

Not so fast!

Put numbers to our dilemma: *At the bathroom line*, *there are 98 men in line!*



Oh nooooo! The line in the bathroom is much much longer for men (that is because they drink a lot of beer…)

With perfect knowledge and perfect sight, we see that in line, 4 times more men have long hair than women!

Intelligence

- When you think about intelligence, what sort of things comes to mind? Intelligence is more than simply the accumulation of facts. It also encompasses the ability to learn new things
- Fluid intelligence refers to the ability to reason and think flexibly
- Crystallized intelligence refers to the accumulation of knowledge, facts, and skills that are acquired throughout life
- https://en.wikipedia.org/wiki/Fluid_and_crystallized_intelligence

Let's use some *crystallized* intelligence..

Conditional probabilities

P(long hair | woman)

If I know that a person is a **woman**, what is the probability that person has long hair?

P(long hair | woman)

= # women with long hair / # women

= 25 / 50 = **.5**

Out of 100 people at the movies

50 are women

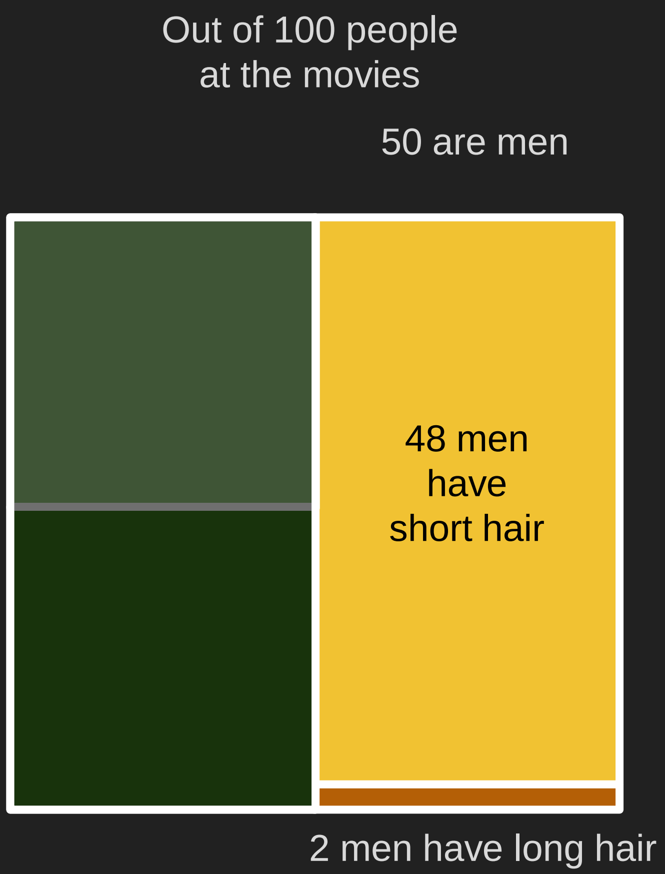
Conditional probabilities

If I know that a person is a **man**, what is the probability that person has long hair?

P(long hair | man)

= # men with long hair / # men

= 2 / 50 = **.04**



Joint probabilities (∩)

P(A and B) is the probability that both A and B are the case.

Also written P(A, B) or P(A ∩ B)

P(A + B) = P(A) * P(B)

P(A and B) is the same as P(B and A)

The probability that I am having a jelly donut with my milk is the same as the probability that I am having milk with my jelly donut.

P(donut and milk) = P(milk and donut)

Joint (∩) probabilities

What is the probability that a person is both a **woman** and has **short hair**?

P(woman with short hair)

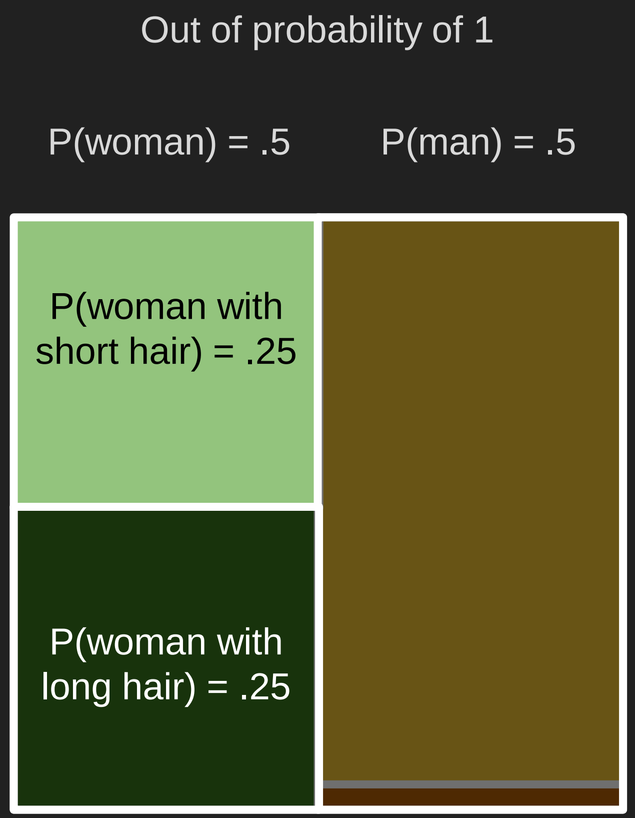
= P(woman) * P(short hair | woman)

= .5 * .5 = **.25**

P(woman with long hair)

= P(woman) * P(long hair | woman)

= .5 * .5 = **.25**



Joint (∩) probabilities

What is the probability that a person is both a **man** and has **short hair**?

P(man with short hair)

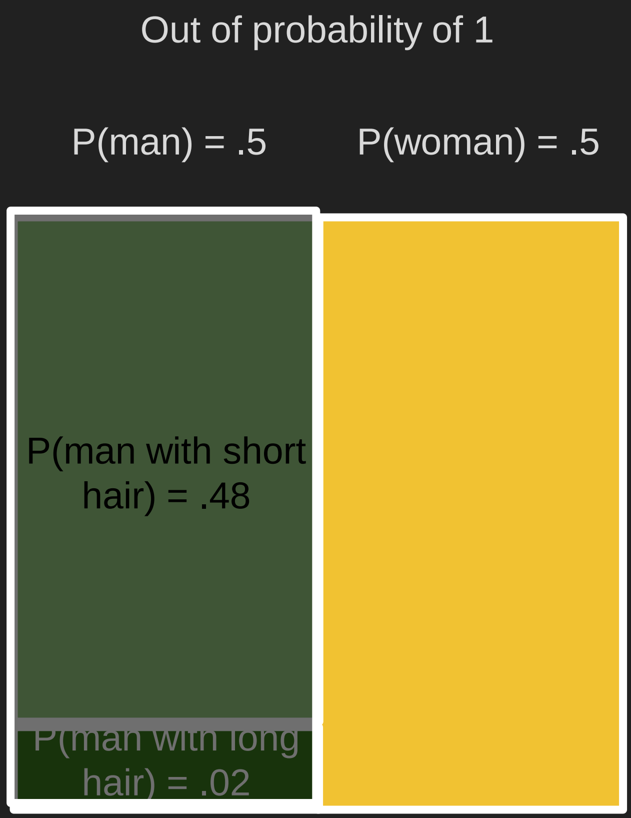
= P(man) * P(short hair | man)

= .5 * .96 = **.48**

P(man with long hair)

= P(man) * P(long hair | man)

= .5 * .04 = **.02**



At the bathroom line (@b) though…

Joint (∩) probabilities

P(man@bathroom with long hair)

= P(man@b) * P(long hair | man)

= .98 * .04 = **.04**

P(woman@bathroom with long hair)

= P(woman@b) * P(long hair | woman)

= .02 * .5 = **.01**

Marginal probabilities (∪)

P(A or B) is the probability that either A or B is the case

Also written P(A or B) or $P(A \cup B)$

$P(A \text{ or } B) = P(A) + P(B) - P(A \cap B)$

Marginal probabilities

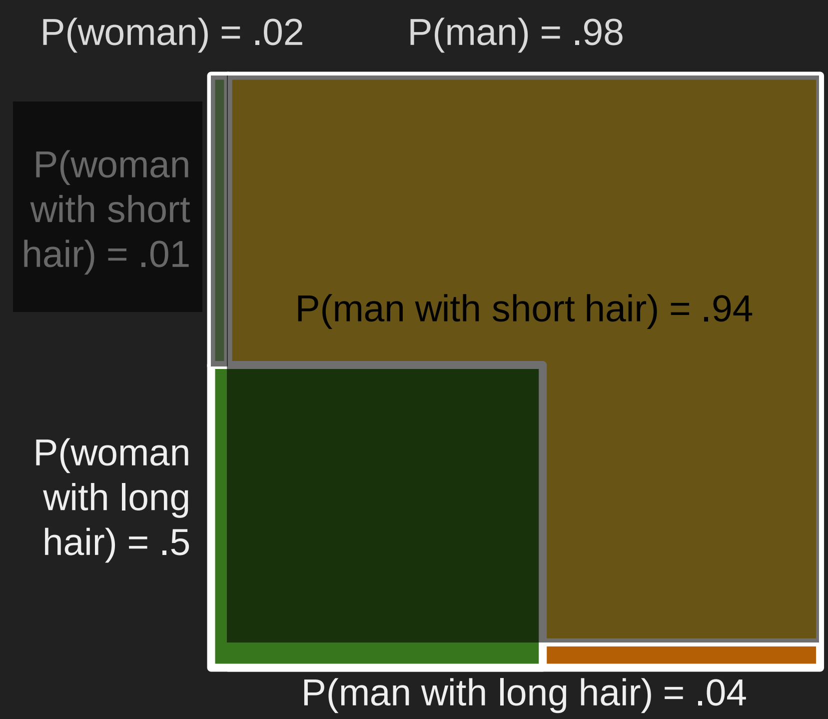
Either **man** or **woman** with **long hair**:

P(long hair) = P(woman with long hair) +

P(man with long hair)

= .25 + .02 = **.27**

Out of probability of 1



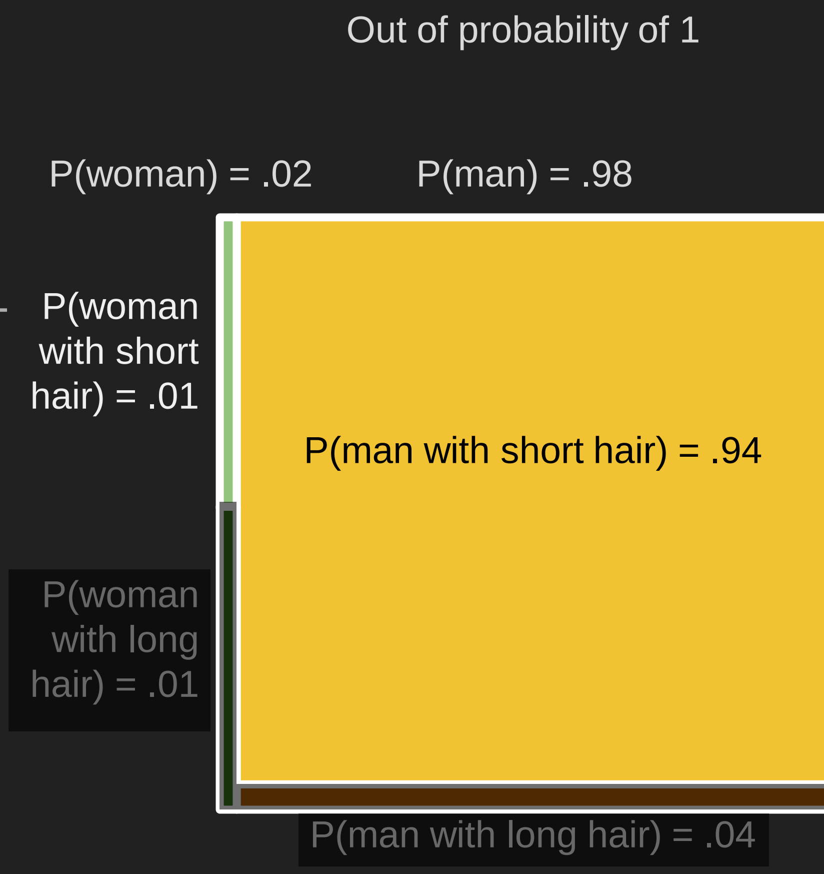
Marginal probabilities

Either **man** or **woman** with **short hair**:

P(short hair) = P(woman with short hair) +

P(man with short hair)

= .25 + .48 = **.73**



What we really care about

We know the person has long hair.
Are they a man or a woman?

P(man | long hair)

We don’t know this answer yet!

Thomas Bayes noticed something cool

$P(\text{man with long hair}) = \boxed{P(\text{long hair}) * P(\text{man} \mid \text{long hair})}$

$P(\text{long hair and man}) = \boxed{P(\text{man}) * P(\text{long hair} \mid \text{man})}$

Because $P(\text{man and long hair}) = P(\text{long hair and man})$

$\boxed{P(\text{long hair}) * P(\text{man} \mid \text{long hair})} = \boxed{P(\text{man}) * P(\text{long hair} \mid \text{man})}$

Thomas Bayes noticed something cool

P(man with long hair) = P(long hair) * P(man | long hair)

P(long hair and man) = P(man) * P(long hair | man)

Because P(man and long hair) = P(long hair and man)

P(long hair) * P(man | long hair) = P(man) * P(long hair | man)

P(man | long hair) = P(man) * P(long hair | man) / P(long hair)

Thomas Bayes noticed something cool

P(man with long hair) = P(long hair) * P(man | long hair)

P(long hair and man) = P(man) * P(long hair | man)

Because P(man and long hair) = P(long hair and man)

P(long hair) * P(man | long hair) = P(man) * P(long hair | man)

P(man | long hair) = P(man) * P(long hair | man) / P(long hair)

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

Bayes' Theorem

$$P(A \mid B) = \left[\frac{P(B \mid A)\ P(A)}{P(B)} \right]$$

Back to the movie theater, this time with Bayes

$$P(\text{man} \mid \text{long hair}) = \left\{ \frac{P(\text{man}) * P(\text{long hair} \mid \text{man})}{P(\text{long hair})} \right\}$$

$$= \frac{P(\text{man}) * P(\text{long hair} \mid \text{man})}{P(\text{woman with long hair}) + P(\text{man with long hair})}$$

$$P(\text{man} \mid \text{long hair}) = \left\{ \frac{.5 * .04}{.25 + .02} \right\} = .02 / .27 = \boxed{.07}$$

Back to the bathroom line (@b), this time with Bayes

$$P(\text{man} \mid \text{long hair}) = \frac{P(\text{man}) * P(\text{long hair} \mid \text{man})}{P(\text{long hair})}$$

$$= \frac{P(\text{man}) * P(\text{long hair} \mid \text{man})}{\underbrace{P(\text{w@b with long hair})}_{.02 * .5} + \underbrace{P(\text{m@b with long hair})}_{.98 * .04}}$$

$$P(\text{man} \mid \text{long hair}) = \frac{.98 * .04}{.01 + .04} = .04 / .05 = \boxed{.80}$$

Conclusion

This person dropped their ticket in the hallway

> *"Excuse me, sir!"*

Conclusion

- You update your belief based on ***evidence***

Bayes' Theorem

$$P(w \mid m) = \left[\frac{P(m \mid w)\ P(w)}{P(m)} \right]$$

Bayes' Theorem

$$P(w \mid m) = \left[\frac{P(m \mid w)\ \overset{\textbf{prior}}{\boxed{P(w)}}}{P(m)} \right]$$

Bayes' Theorem

likelihood

$$P(w \mid m) = \left[\frac{\boxed{P(m \mid w)}\ P(w)}{P(m)} \right]$$

Bayes' Theorem

posterior

$$\boxed{P(w \mid m)} = \left[\frac{P(m \mid w)\ P(w)}{P(m)} \right]$$

Bayes' Theorem

$$P(w \mid m) = \left[\frac{P(m \mid w)\; P(w)}{\boxed{P(m)}} \right]$$

marginal likelihood

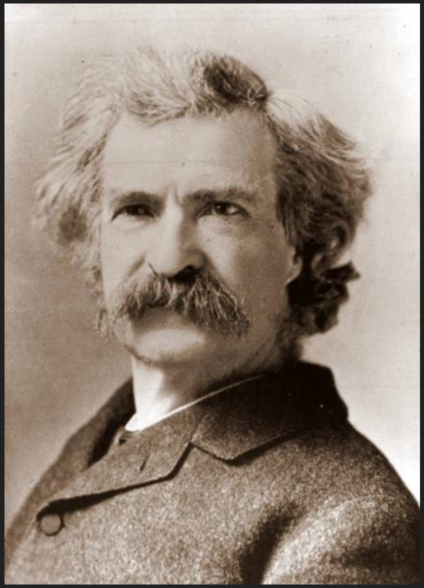
Why Bayesian inference makes us nervous

We’re not always aware of what we believe.

Putting what we believe into a distribution correctly is tricky.

We want to be able to be surprised by our data.
Inaccurate beliefs can make it hard or impossible to learn.

> “*It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so*.”
>
> \- Mark Twain



Believe the impossible, at least a little bit

Leave room for believing the unlikely. Leave a non-zero probability unless you are absolutely certain.

> “When you have excluded the impossible, whatever remains, however improbable, must be the truth”
>
> - Sherlock Holmes (Sir Arthur Conan Doyle)

Believe the *impossible*, at least a little bit

"Alice laughed: "There's no use trying," she said; "one can't believe impossible things."

"I dare say you haven't had much practice," said the Queen. "When I was younger, I always did it for half an hour a day. Why, sometimes I've believed as many as six impossible things before breakfast."

- Lewis Carroll (Alice’s Adventures in Wonderland)


CLASS
DISMISSED
iStockphoto