Bayes Theorem is vital for inference in all realms - not just scientific data analysis.

For example, the "What is the probability it will rain" problem in the Homework.

Bayes Theorem can also be applied to cases of determining the probability that a patient has a disease given an outcome of a medical test, or the probability of guilt or innocence of a suspect in jurisprudence.
Bayesian Spam Filtering

You have developed a new spam filter. It is 100% accurate when it encounters spam! But it is a little over-cautious and 5% of the time it results in a false positive, identifying a valid email as spam.

Your client is internet-savvy and is careful about distributing her email address. As such, she gets spam, but only in 1 out of 1000 emails.

Given that your algorithm has identified an email as spam, what is the probability that it really is spam?

\[
P(S|+) = \frac{P(S|I) \cdot P(+|S,I)}{P(S|I) \cdot P(+|S,I) + P(\sim S|I) \cdot P(+|\sim S,I)}
\]

\[
P(S|I) = 0.001, \quad P(+|S,I) = 1.00, \quad P(+|\sim S,I) = 0.05
\]

\[
P(S|+) = \frac{0.001 \cdot 1.00}{0.001 \cdot 1.00 + 0.999 \cdot 0.05} = 0.0196
\]

\[
P(S|+) \approx 2\% \quad \text{only 2\% of the time is an email that has been identified as spam actually is spam!}
\]
Bayes in Medicine

You have just donated blood and the blood test has come back HIV positive.

One out of 1000 people have HIV. If you have HIV, the blood test will detect it. If you don't have HIV, the blood test will come back positive 5% of the time.

This is a made-up problem. Don't base a medical decision on these results.

Hypotheses | Data
--- | ---
HIV | + - test result
\neg HIV | - - test result

\[ P(\text{HIV} | +) = 0.001 \]
\[ P(+ | \text{HIV}, +) = 1.00 \]
\[ P(+ | \neg \text{HIV}, +) = 0.05 \]

\[ P(\text{HIV} | +) = \frac{P(\text{HIV} | +) \cdot P(+ | \text{HIV}, +)}{P(\text{HIV} | +) \cdot P(+ | \text{HIV}, +) + P(\neg \text{HIV} | +) \cdot P(+ | \neg \text{HIV}, +)} \]

\[ = \frac{0.001}{0.001 + 0.999 \cdot 0.05} = 0.0196 \]

\[ \approx 2% \] Working with probabilities can be non-intuitive.
Three Doors

You are on a famous game show with Monty Hall (a 1970's game show host).

There are three doors.

Behind one door is the prize... a brand new car!

Behind the other two doors are goats... not a prize.

You are asked to choose a door... call it door #1.

Monty Hall then shows you what is behind another door... it is a goat.

Monty Hall now asks if you would like to change doors.

WHAT SHOULD YOU DO?
Three Doors ... The Solution

Prior Probabilities

\[ P(1|I) = P(2|I) = P(3|I) = \frac{1}{3} \]

Where \( P(1|I) \) is the prior probability that the prize is behind door 1.

When Monty Hall shows you Door 2 ... that is data (information)

He could either have opened door 2 or door 3, but only if it had a goat. Let \( O_2 = \text{opened 2} \) and \( O_3 = \text{opened 3} \)

\[ P(O_2|2,I) = 0 \quad \text{He would not open door 2 if the prize was behind door 2} \]

\[ P(O_2|3,I) = 1 \quad \text{If the prize was behind door 3, he could have only opened door 2} \]

\[ P(O_2|1,I) = 0.5 \quad \text{If the prize was behind door 1, he could have opened either door 2 or 3.} \]

\[ P(1|O_2,I) = \frac{P(1,I)P(O_2|1,I)}{P(1,I)P(O_2|1,I) + P(3,I)P(O_2|3,I)} = \frac{\frac{1}{3} \cdot \frac{1}{2}}{\frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 1} = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3} \]

\[ P(3|O_2,I) = \frac{P(3,I)P(O_2|3,I)}{P(1,I)P(O_2|1,I) + P(3,I)P(O_2|3,I)} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3} \]

Switch doors! He didn't open #3 => INFORMATION
I have two children.
One is a boy.
What is the probability that both are boys?

Notation
B1 - Child #1 is a boy
B2 - Child #2 is a boy
S1 - Child #1
S2 - Child #2

B1 \& B2 - Both children are boys

(B1 \& S2) \& (S1 \& B2) - One child is a boy

Goal: Find \( P(B1 \& B2 | (B1 \& S2) \lor (S1 \& B2), I) \)

Bayes Theorem flips things around

\[
P(B1 \& B2 | (B1 \& S2) \lor (S1 \& B2), I) = \frac{P((B1 \& S2) \lor (S1 \& B2) | B1 \& B2, I) \times P(B1 \& B2 | I)}{P((B1 \& S2) \lor (S1 \& B2) | I)}
\]
Two Children I

First consider the denominator

\[ P((B_1 \land S_2) \lor (S_1 \land B_2)|I) = \]
\[ P(B_1 \land S_2|I) + P(S_1 \land B_2|I) - P(B_1 \land B_2|I) \]

\[ P(B_1 \land S_2|I) = P(B_1|S_2,I)P(S_2|I) \]
\[ = P(B_1|I)P(S_2|I) \]
\[ = \frac{1}{2} \cdot 1 \]
\[ = \frac{1}{2} \]

\[ P(S_1 \land B_2|I) = \frac{1}{2} \]

\[ P(B_1 \land B_2|I) = P(B_1|B_2,I)P(B_2|I) \]
\[ = P(B_1|I)P(B_2|I) \]
\[ = \frac{1}{2} \cdot \frac{1}{2} \]
\[ = \frac{1}{4} \]

\[ P((B_1 \land S_2) \lor (S_1 \land B_2)|I) = \frac{1}{2} + \frac{1}{2} - \frac{1}{4} = \frac{3}{4} \]

**NOTE:** THE \( \lor \) IS NOT DISJOINT

This is what makes the problem difficult.
Consider the likelihood

\[ P((B1 \land S2) \lor (S1 \land B2) \mid B1 \land B2, I) = \]

\[ P(B1 \land S2 \mid B1 \land B2, I) + P(S1 \land B2 \mid B1 \land B2, I) \]

\[ - P(B1 \land B2 \mid B1 \land B2, I) \]

\[ P(B1 \land S2 \mid B1 \land B2, I) = P(B1 \mid B1 \land B2 \land S2, I) P(S2 \mid B1 \land B2) \]

\[ = P(B1 \mid B1, I) P(S2 \mid B2, I) \]

\[ = 1 \]

\[ P(S1 \land B2 \mid B1 \land B2, I) = 1 \]

\[ P(B1 \land B2 \mid B1 \land B2, I) = 1 \]

Therefore

\[ P((B1 \land S2) \lor (S1 \land B2) \mid B1 \land B2, I) = 1 \]

Last

\[ P(B1 \land B2 \mid I) = P(B1 \mid B2, I) P(B2 \mid I) \]

\[ = P(B1 \mid I) P(B2 \mid I) \]

\[ = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} \]
Two Children II

\[ P(B_1 \& B_2 | (B_1 \& B_2) \lor (S_1 \& B_2), I) = \frac{1 - \frac{1}{4}}{\frac{3}{4}} = \frac{1}{3} \]

The probability that both children are boys is \( \frac{1}{3} \).