

Theorem Of Wisdom: Bayes' Theorem As The Most Possible Rational Way Of Making Decisions

Sandra(Yijia) Cai

July 2, 2025

Abstract

Often time, people are struggling with making decisions. According to [Harvard Business Review](#), average adults make 33,000 to 35,000 decisions each day, including easy and hard ones. However, how many of them are actually close to the best absolute rationality for the best possible outcome is yet a myth. Let alone the decision making process should be driven by good reasoning and rationale, which normally consume a certain amount of brain power to do so. Bayes' Theorem, a theorem in probability theory that describes the probability of an event occurring under certain known conditions, created by Mathematician Thomas Bayes in 1763, can be a useful tool for making it happen given its comprehension, causality, and factorization of objective facts. Coming from a scientific approach of observing and evaluating things we are experiencing, it provides a quantitative method of weighing out different options choices to believe. In other words, the leanest way of touching absolute rationality.

1 Introduction

Bayes' Theorem, categorized under Conditional Probability, which is a way of predicting the probability of an event when there was another event already occurred. Considering the first event as event A and the second corresponding event as event B, this probability follows the format of being written as

$$P(B|A)$$

, notation for the probability of B given A. If event A has 0 influence over the probability of event B, then the conditional probability of event B is simply

$$P(B)$$

. In this case, event A and event B are also viewed as independent events. If the events are not independent, the probability of both events A and B occurring simultaneously, which is also known as the intersection of A and B, is defined by

$$P(A \text{ and } B) = P(A) * P(B|A)$$



Figure 1: Mathematician Thomas Bayes (1702 - 1761)

, considering the influence on event B from event A. In this case, the conditional probability

$$P(B|A)$$

can be calculated by

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

, which is considered valid only when

$$P(A) > 0$$

. As an useful means to reason information, for example, when FBI is backtracking what has happened at the crime scene to identify the criminal; considering there will be a massive pouring rain given the density of drips dropping from the sky in the next minutes; knowing other people are going to cry when they frown their forehead and curve their mouth corners. Information reasoning exists in everywhere of our daily lives. However, the reality has complex variables and grey zones, for example, a homeless person suddenly wins the lottery ticket, people may think this person cheated, which presents an only-right-or-only-wrong mindset without factorizing other possible reasons contributing to the event altogether, which includes

No.1 This person cheated

and

No.2 This person has been observing which number wins lottery the most

and

No.3 This person used Bayes' Theorem to calculate the chance of winning at that time

. Therefore, it is important to maintain a mindset of probability, displaying all of the possible reasons of the event and then choose the one that has the biggest probability. Intuition is the most used tool for reasoning in people's lives, which is to make decisions based on the most intuitively possible reason relevant to the event but not Bayes' Theorem. Sometimes, it can be dangerously misleading.

2 Practicality

2.1 Scenario 1: Who is going to be more drunk

Ben and Jerry are drinking together. Based on their alcohol tolerance, Ben will get drunk when drinking 9 shots out of 10 total shots whereas Jerry will get drunk when drinking only 1 shot out of total 10 shots. This means Ben has higher alcohol tolerance than Jerry.

Now, suppose we observe that someone is drunk, and we want to determine whether it's more likely to be Ben or Jerry. To apply Bayes' Theorem properly, we need to consider both the likelihood and prior probabilities.

The likelihood represents how probable it is to observe someone being drunk given that it's a specific person:

$$P(\text{drunk}|\text{Ben}) = \frac{9}{10} = 90\%$$

$$P(\text{drunk}|\text{Jerry}) = \frac{1}{10} = 10\%$$

However, to determine who is more likely to be the drunk person (i.e., $P(\text{Ben}|\text{drunk})$ vs. $P(\text{Jerry}|\text{drunk})$), we also need to consider the prior probabilities. If we assume both Ben and Jerry are equally likely to be the person we're observing initially:

$$P(\text{Ben}) = P(\text{Jerry}) = 50\%$$

Using Bayes' Theorem:

$$P(\text{Ben}|\text{drunk}) = \frac{P(\text{drunk}|\text{Ben}) \times P(\text{Ben})}{P(\text{drunk})}$$

$$P(Jerry|drunk) = \frac{P(drunk|Jerry) \times P(Jerry)}{P(drunk)}$$

Where $P(drunk) = P(drunk|Ben) \times P(Ben) + P(drunk|Jerry) \times P(Jerry) = 0.9 \times 0.5 + 0.1 \times 0.5 = 0.5$
Therefore:

$$P(Ben|drunk) = \frac{0.9 \times 0.5}{0.5} = 90\%$$

$$P(Jerry|drunk) = \frac{0.1 \times 0.5}{0.5} = 10\%$$

This shows that if we observe someone drunk, it's more likely to be Ben (90%) than Jerry (10%), because Ben gets drunk more frequently despite having higher tolerance. This demonstrates how Bayes' Theorem properly combines likelihood with prior probability to make rational inferences.

2.2 Scenario 2: Is the elevator going to crash

Some people are afraid of taking elevators. When the elevator suddenly slows down from one floor to another, people start to think about it is going to crash down, and here is what people are reasoning:

Severe turbulence in the speed of the elevator were observed

and then there are following possible reasons:

a. elevator crashes down

and

b. elevator is normal but it is overused and a little bit clunky

However, if an elevator crashes down, the likelihood of the people in this elevator feeling obvious differences in the speed of the elevator is

$$P(severe\ turbulence|elevator\ crashes\ down) = 100\%$$

. When the elevator is overused but is functioning normally, assuming it is 1 time out of 10 times of the total usage, then the likelihood of crashing down is

$$P(severe\ turbulence|elevator\ is\ overused) = 10\%$$

. In this case, elevator crashing down has stronger ability of explaining people feeling severe turbulence than elevator being overused, which means the likelihood is higher. Hence people will reason the phenomenon with elevator is crashing down and they will feel great panicked.

However, elevators being overused is way more common to happen in reality, not crashing down. So we notice there is the problem of utilizing the maximum likelihood estimation — when we infer the cause from the phenomenon, we have to not only consider the ability of explaining but also the the basic probability of this cause itself occurring, which is prior probability, which is independent to the phenomenon itself. According to [Elevator Injury Lawyer](#), the chance of dying in one elevator crash-down is about 1 in 10.5 million, which means if a person takes an elevator once per day, it takes 28767 years for this person to experience this event. Therefore, the prior probability of elevator crashing down is

$$P(elevator\ crashing\ down) = \frac{1}{28767 * 365} = 9.52 * 10^{-8}$$

. However, it is common to experience elevator being overused and it does not have an averaged speed while going up or down per floor consistently, which means we can estimate the prior probability

$$P(elevator\ overused) = 0.9$$

. Then we multiply two reasons' prior probability and likelihood to make a comparison, we have

$$elevator\ crashing\ down : 100\% * 9.52 * 10^{-8} = 9.52 * 10^{-8}$$

, and

$$elevator\ overused : 10\% * 0.9 = 0.09$$

. It is noticed that when elevator has severe turbulence, elevator being overused is 94537.8 times elevator getting crashed down, which means there is almost no need to worry about the crash down.

2.3 Interpretation

Now that we can come to an interpretation of Bayes' Theorem. We have

$$P(\text{reason}|\text{phenomenon}) = \frac{P(\text{reason}) * P(\text{phenomenon}|\text{reason})}{P(\text{phenomenon})}$$

, which means the left side of the equation is information reasoning calculates the probability of inferring the cause after observing the phenomenon, which is also called the posterior probability; the first numerator from left to right is prior probability, the second numerator is the likelihood which is the probability of observing a phenomenon given the cause, also known as the ability of the cause to explain phenomenon; the denominator is the basic probability of the phenomenon itself given the listed various reasons. However, regardless of reasons, the denominator remains the same. Therefore, we can compare numerators, which is, the result of prior probability multiplying likelihood, to calculate the best possible reason. It is also noticed that the denominator is to scale the likelihood of the cause by a fixed ratio, which provides a role of standardization. Hence we have

$$\text{posterior probability} = \text{prior probability} * \text{standardized likelihood}$$

3 Summary

We can now draw a conclusion that the process of Bayes' Theorem is ...

1. Observe the phenomenon
2. List all of the possible causes
3. Calculate different prior probabilities of different causes
4. Calculate different likelihood of different causes
5. Compare the multiplication of prior probability and likelihood
6. Choose the highest multiplication to be the most possible cause to the phenomenon

It also allows us to observe new phenomena, and to continuously update and iterate the posterior probabilities of causes, which is also considered [The Generalization of Bayesian Theory](#). The process from the Bayes' Theorem to The Generalization of Bayesian Theory is a good interpretation for science reasoning too. In science reasoning, we have ...

- Propose theoretical hypotheses based on observations
- Continuously update confidence in the hypothetical theory based on new observations

Bayes' Theorem is currently considered to be the most powerful scientific view on earth in science and philosophy. According to [Cornell University](#), it can be discussed with the existence of God. This is not only useful to scientific world, but also our daily lives, with the examples provided above, because there are many causes to a certain phenomenon and people are limited to their own experiences coming up with different causes, hence endowing different prior probability to different causes. Different prior probability combined with different existing experience will lead to different basic confidence on a certain phenomenon, let alone the causes that lack of data as the backup to be a valid proof. Hence, calculate the likelihood of the cause for new evidence based on reliable data, thus updating the posterior probability of each cause. When there is more and more evidence, people turn to have similar or same basic confidence on a certain phenomenon.

Discussing this further, Bayes' Theorem is a good evidence against conspiracy theory because the likelihood is either 0 or almost 100%, so we have

$$P(\text{observed phenomenon}|\text{conspiracy}) \approx 100\%$$

. However, the prior probability is extremely low,

$$P(\text{conspiracy}) \approx \text{extremely low}$$

, and

$$P(\text{observed phenomenon}|\text{conspiracy}) \approx \text{extremely low}$$

. Therefore, in the current environment of political history in the making and media [infodemic](#) out bursting, people who place trust on conspiracy is using maximum likelihood estimation to think about problems, not Bayes' mindset, and given the ability of conspiracy to explain certain phenomena, it is highly difficult for people to change their mind on this because the posterior probability keeps increasing when iterating with new data. Therefore, if people are not willing to change their belief on prior probability, there is almost no way to reach a common ground, which is not within the scope of rational thinking and discussion.

Bayes' Theorem is worthy to be learned by everyone given the interpretation and embodiment of scientific view and validation. It is probably the most possible rational way for human beings to reach an iteration-friendly best-optimal decision making process while doing the calculations.