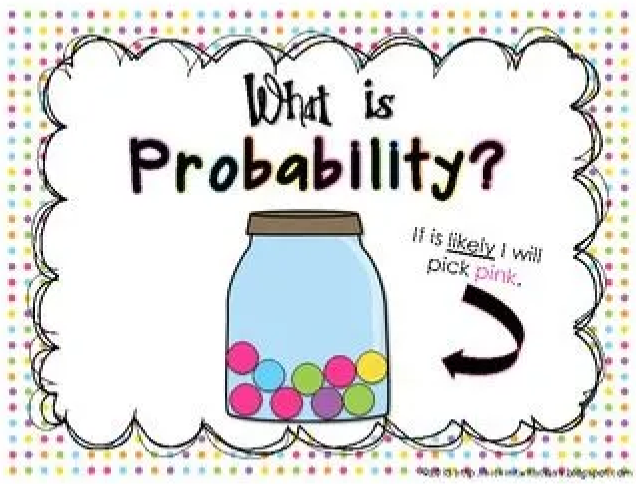
Probability Theory: Bayes Theorem

Ashish Arora · Follow

9 min read · Aug 2

Probability distributions are the backbone of statistical inferences and to understand these distributions we should at least have some basic understanding of Probability Theory.



The objective of this post is to build a foundation in probability and the topics that we will cover in this article are:

- Understanding of Probability
- Basic Terminologies such as sample space, event, experiment, & outcome.
- Probability Axioms
- Types of Events: Independent, Dependent Events, etc.
- Probability Rules: Addition and Multiplication rules.
- Concept of conditional probability

- Bayes’ theorem

Understanding Probability Theory

Probability theory deals with the study of uncertainty and randomness, providing tools to analyze events and quantify their likelihood of occurrence.

Basic Terminologies

Probability

Probability is the likelihood of an event happening.

Experiment and Outcome

An experiment *refers to a process or activity that generates a set of possible results or outcomes.* It is the fundamental action that we are interested in studying from a probabilistic perspective.

An outcome is a *single result that occurs as a consequence of an experiment.* It represents one of the possible individual results that can happen in the given experiment. **For example, when rolling a fair six-sided dice, the possible outcomes are the numbers 1, 2, 3, 4, 5, or 6.**

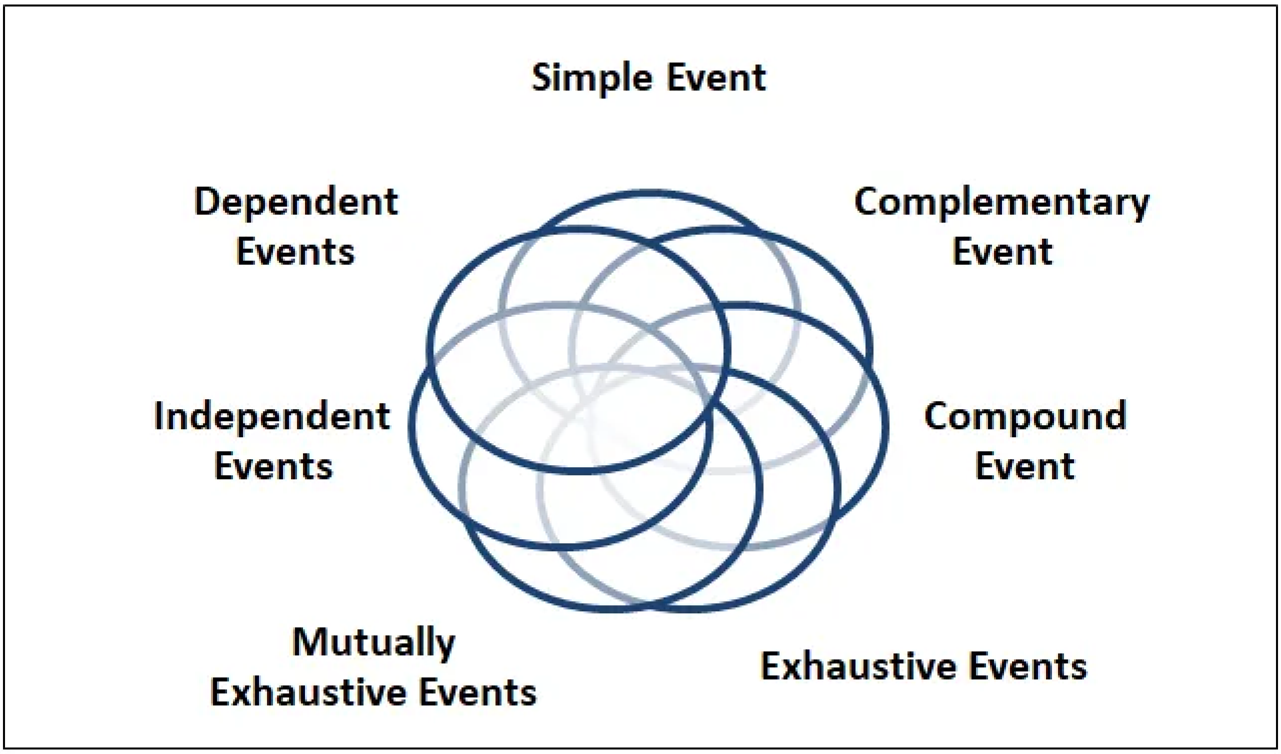
Equally Likely Outcomes: When all the outcomes in a sample space have the same chance of occurring, they are called equally likely outcomes. For example, when flipping a fair coin, the outcomes (heads or tails) are equally likely.

Sample Space and Event:

Sample space encompasses all possible outcomes of an experiment or observation.

Events are subsets of the sample space, representing specific outcomes of interest.

Different Types of Events:



1. **Simple Event and Compound Event:** If an event consists of a single result from the sample space, it is termed a simple event. For example, in rolling a fair six-sided dice we have a sample space of 1,2,3,4,5,6, but getting a "3" is a simple event. However, if want to have a number less than 3, then the event has more than two outcomes that is 1 and 2. Such events are known as **compound events.**

2. **Exhaustive Events:** Exhaustive events are a set of events that together cover the entire sample space. In other words, the union of these events is equal to the entire sample space. For example, when rolling a fair six-sided dice, getting a number less than 7 and more than 0 will give us the entire space "1," "2," "3," "4," "5," or "6". Such events are exhaustive events as they cover all possible outcomes.

3. **Mutually Exclusive Events:** Mutually exclusive events are events that cannot occur simultaneously. If two events are mutually exclusive, it means that they have no outcomes in common. Mathematically, if A and B are mutually exclusive events, then their intersection ($A \cap B$) is an empty set ($\emptyset$). For example, when flipping a coin, the events "Heads" and "Tails" are mutually exclusive.

4. **Independent Events:** Two events A and B are independent if the occurrence of one event does not affect the probability of the other event. In other words, the probability of one event happening does not depend on whether the other event occurs. Mathematically, for independent events A and B, $P(A \cap B) = P(A) * P(B)$.

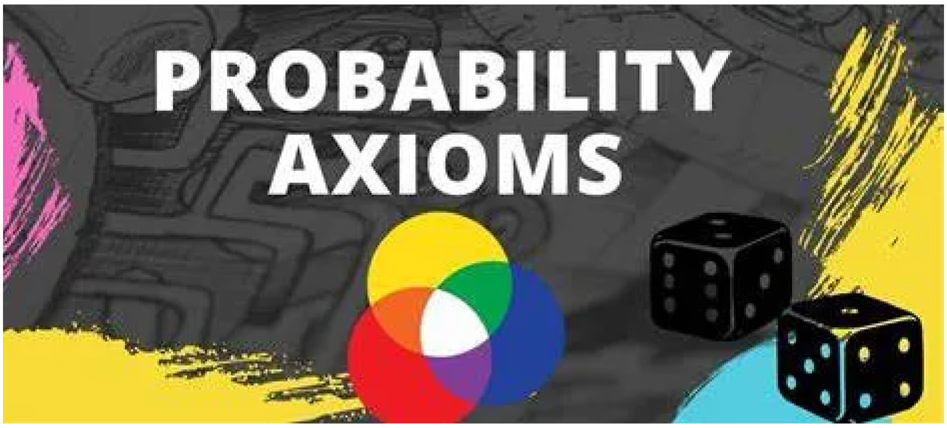
5. **Dependent Events:** Two events A and B are dependent if the occurrence of one event affects the probability of the other event. In this case, the probability of one event happening depends on whether the other event occurs. Mathematically, for dependent events A and B, $P(A \cap B) \neq P(A) * P(B)$.

How to Find the Probability of an Event?

To find the likelihood of the occurrence of events in probability the steps are as follows:

1. Determine the sample space or the total number of possible outcomes of the experiment.
2. Determine the number of favorable outcomes of the event.
3. Divide the value from step 2 by the value obtained in step 1 to get the required probability.

Probability Axioms



1. Non-Negativity: The probability of any event is a non-negative value. That is, for any event A, $\mathbf{P(A) \geq 0}$.
2. Normalized: The probability of an event is a number between 0 and 1, where 0 indicates impossibility, 1 signifies certainty, and values in between represent the likelihood of occurrence.
3. Additive: For any mutually exclusive events (events that cannot occur simultaneously), the probability of their union is equal to the sum of their individual probabilities. That is, for mutually exclusive events A and B, where $A \cap B = \emptyset$ (empty set), and $P(A \cup B) = P(A) + P(B)$.

Probability Rules

Probability

Test for Independence

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

Addition Rule

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

Multiplication Rule of All Events

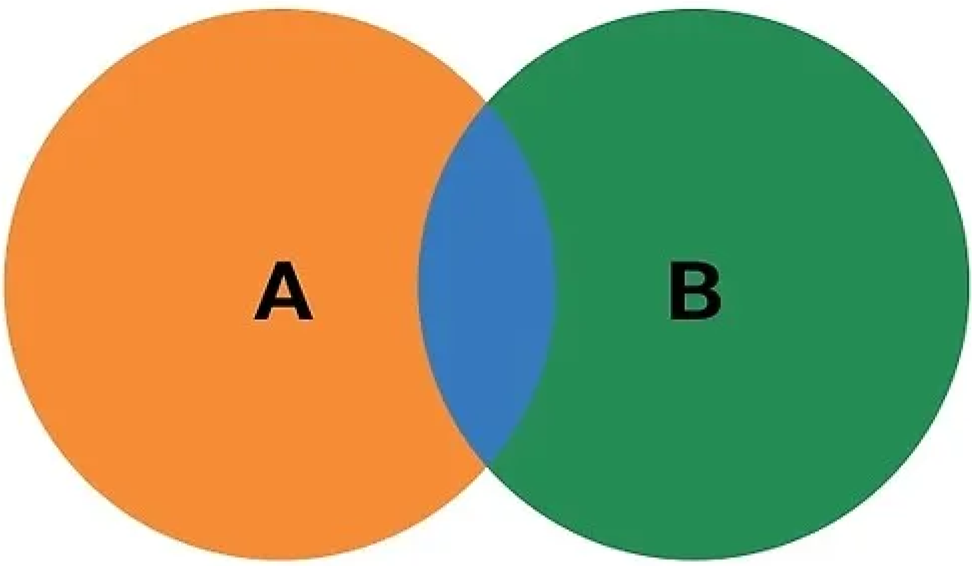
$$P(A \text{ and } B) = P(A) \cdot P(B \mid A)$$

Multiplication Rule for Independent Events

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

Conditional Probability

$$P(A \mid B) = P(A \text{ and } B) \,/\, P(B)$$



The Addition Rule and Multiplication Rule are fundamental principles in probability theory that help calculate the probabilities of compound events.

Multiplication Rule

The Multiplication Rule states that the probability of the intersection of two independent events (A ∩ B) is equal to the product of their individual probabilities:

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

If there are more than two independent events, the rule extends to:

$P(A \cap B \cap C \cap \ldots) = P(A) * P(B) * P(C) * \ldots$

Addition Rule

The Addition Rule states that the probability of the union of two mutually exclusive events (A ∪ B) is equal to the sum of their individual probabilities:

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

If there are more than two mutually exclusive events, the rule extends to:

$P(A \cup B \cup C \cup \ldots) = P(A) + P(B) + P(C) + \ldots$

Complementary Rule

Whenever an event is the complement of another event, then the sum of them is equal to 1.

$P(A') + P(A) = 1$

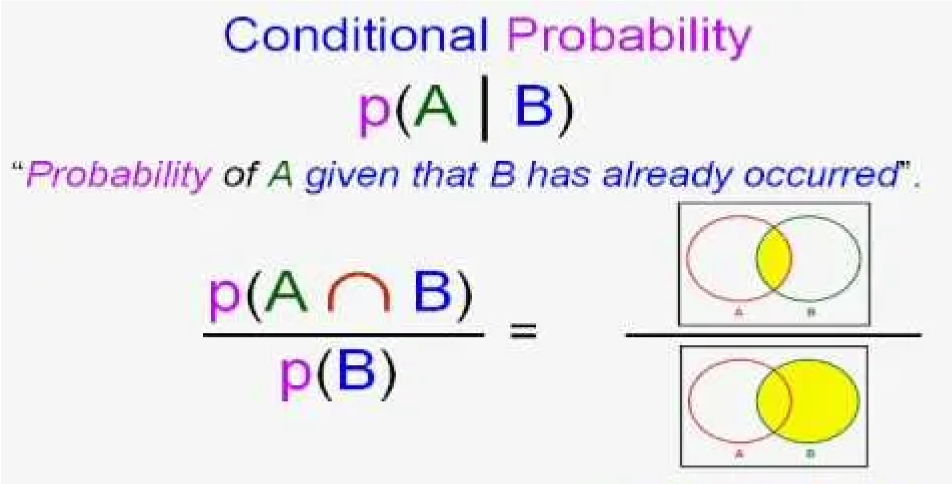
or

$P(A') = 1 — P(A)$

Conditional Probability

Conditional probability refers to the probability of an event occurring given that another event has already occurred.

The formula for conditional probability is given by the Multiplication Rule for Conditional Probability:



Where:

- P(A | B) is the conditional probability of event A given event B.
- P(A ∩ B) is the probability of the intersection of events A and B (both events occurring).
- P(B) is the probability of event B occurring.

Example:

A jar contains 4 green marbles and 6 yellow marbles. Two marbles have been drawn from the jar. The second marble has been drawn without replacement. What is the probability that both the drawn marbles will be yellow?

Solution

Let A = the event that the first marble is yellow; and let B = the event that the second marble is yellow. We know the following:

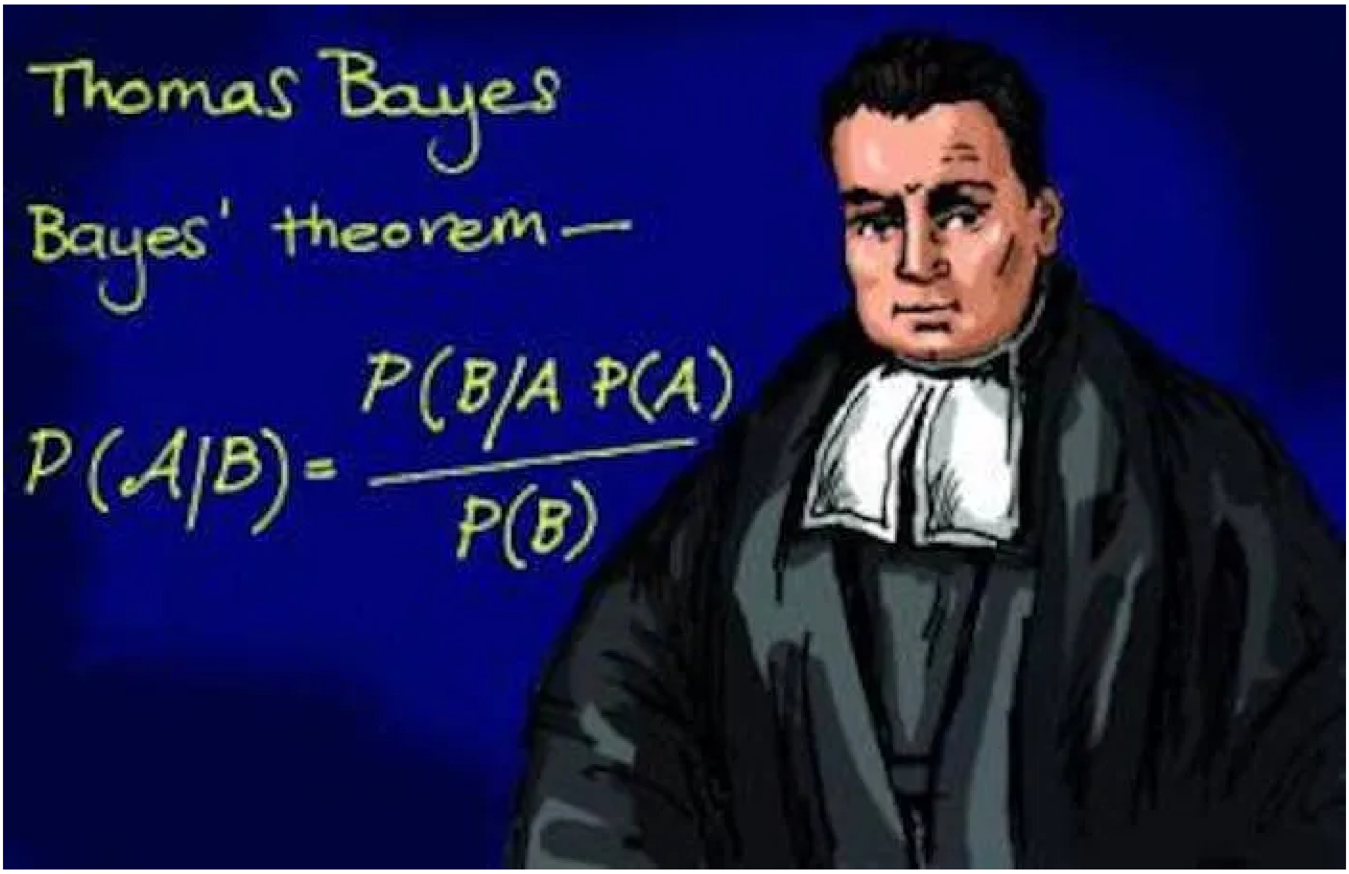
- In the beginning, there are 10 marbles in the box, 6 of which are yellow. Therefore, P(A) = 6/10

- After the first selection, there are 9 marbles in the jar, 5 of which are yellow. Therefore, P(B|A) = 5/9

Therefore, based on the rule of multiplication:

$P(A \cap B) = P(A).P(B|A)$
$P(A \cap B) = (6/10)*(5/9) = 30/90 = 1/3 = 0.33$

Bayes Theorem



Bayes' theorem describes the probability of an event based on prior knowledge of the conditions that might be relevant to the event.

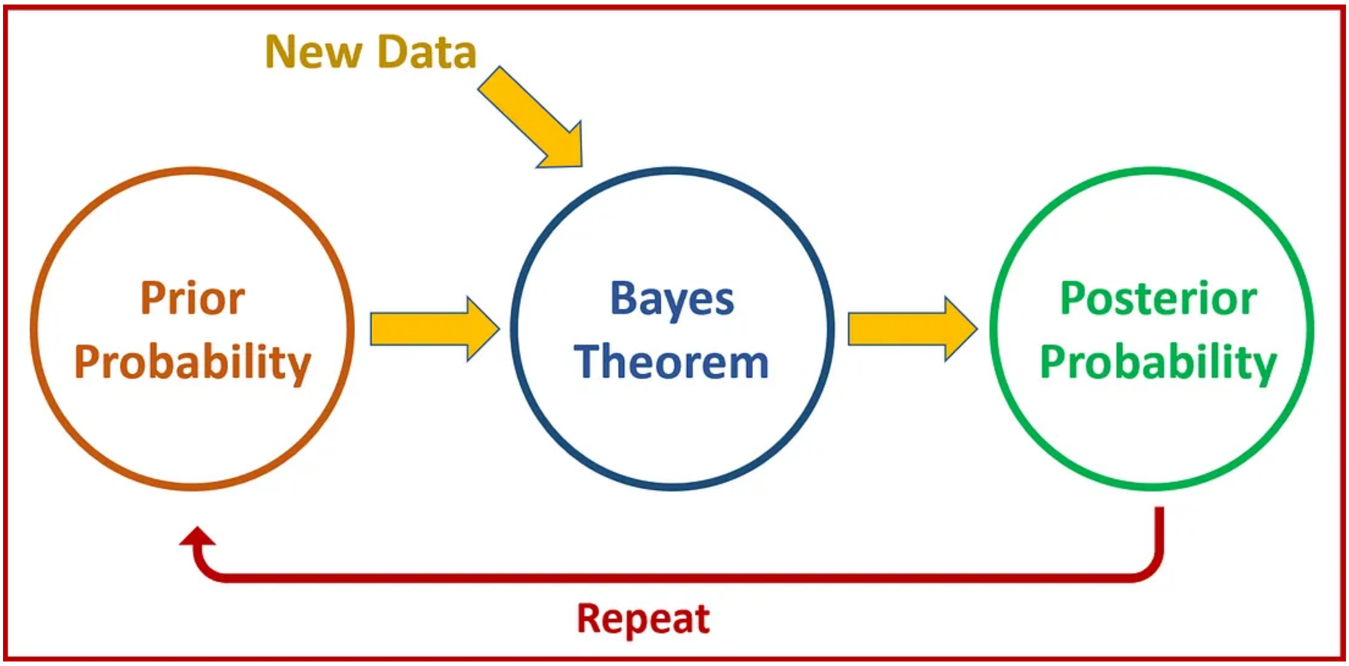
Seems Confusing? Don't worry! I will make it simple for you.

Bayes' Theorem serves as a logical framework to incorporate new evidence into our existing beliefs, allowing us to make more informed decisions. It provides a systematic way to update probabilities when we encounter new data, observations, or test results.

It is one of the most remarkable developments in probability theory by incorporating human cognition into its application. Just as humans adjust their beliefs with new experiences, Bayes' Theorem enhances probabilities after incorporating a piece of new evidence.

Let's take one example:

Imagine you are a doctor faced with diagnosing a patient's illness. You have a prior belief (***prior probability***) about the likelihood of different diseases based on your experience and knowledge. Now, the patient has undergone a diagnostic test (***new evidence***) that provides certain test results (***conditional probability***). You need to update your initial belief (***prior probability***) to arrive at a more accurate diagnosis (***posterior probability***).



Mathematical Journey: From Prior Probability to Posterior Probability

But how does it update the probability mathematically? Let's understand this.

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

Here, P(A) = The probability of a hypothesis (Prior knowledge) being true before any evidence is present.

P(B) = The probability of event B by considering all possible values of the evidence variable

P(B|A) = The probability of seeing the evidence given the hypothesis is true.

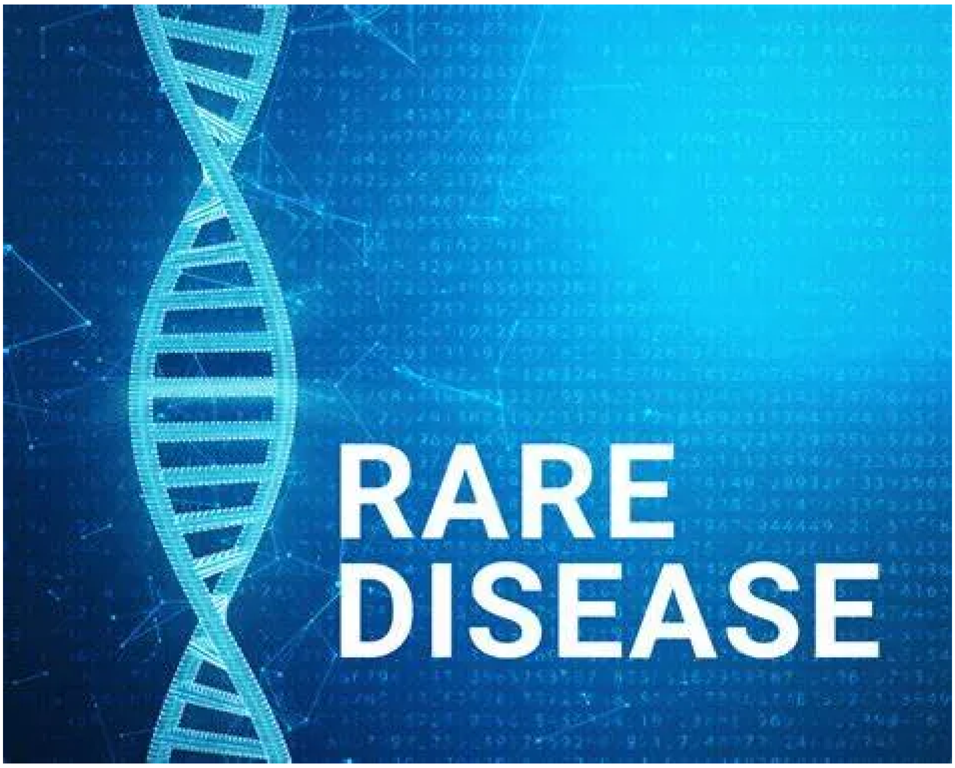
P(A|B) = (Posterior)-The probability of a hypothesis is true given the evidence.

Note: Since there are prior events or beliefs, events A and B are independent events.

Note: The denominator (P(B)) acts as a normalization factor, ensuring that the probabilities sum to 1.

Let's go through a numerical example to illustrate Bayes' Theorem step-by-step.

Consider a medical scenario where a patient is being tested for a rare disease.



Suppose we have the following information:

1. **Prior Probability: P(Disease) = 0.01** (The prior probability of the patient having the disease is 1% because it's a rare disease in the general population.)
2. **Sensitivity: P(Positive Test Result | Disease) = 0.95** (The sensitivity of the diagnostic test is 95%, meaning the probability of getting a positive test result given that the patient actually has the disease.)

3. **Specificity: P(Negative Test Result | No Disease) = 0.90** (The specificity of the diagnostic test is 90%, meaning the probability of getting a negative test result given that the patient does not have the disease.)

Now, the patient undergoes the diagnostic test, and the test result is positive (Event B). We want to calculate the probability of the patient actually having the disease (Event A | B).

Step 1: Start with the Prior Probability: P(Disease) = 0.01

Step 2: Calculate the Likelihood: P(Positive Test Result | Disease) = 0.95

Step 3: Compute the Marginal Probability: P(Positive Test Result) = P(Positive Test Result | Disease) * P(Disease) + P(Positive Test Result | No Disease) * P(No Disease) = 0.95 * 0.01 + (1–0.90) * (1–0.01) = 0.0495 + 0.009 = 0.0585

Step 4: Apply Bayes' Theorem: P(Disease | Positive Test Result) = (P(Positive Test Result | Disease) * P(Disease)) / P(Positive Test Result) = (0.95 * 0.01) / 0.0585 = 0.0095 / 0.0585 ≈ 0.1624

Step 5: Obtain the Posterior Probability: P(Disease | Positive Test Result) ≈ 0.1624

In this example, despite getting a positive test result, the probability of the patient actually having the disease is only around 16.24%. However, the prior information of having the disease was just 1%, though after given the evidence it has increased.

Though, it is still low for a person diagnosed with some disease. The low prior probability of the disease and the imperfect test accuracy contribute to the relatively low posterior probability. So, this gives us the idea that there could be a chance that the testing lab's probability is less accurate.

— — — — — — — — — — — — —

So, this is all from my end of this post covering the basics of probability theory.

If you don't know, how to solve probabilities using permutation and combination then do learn the basic of it. Though it is not widely used. Knowing the Counting principle makes life easy.

You can take a reference from this.